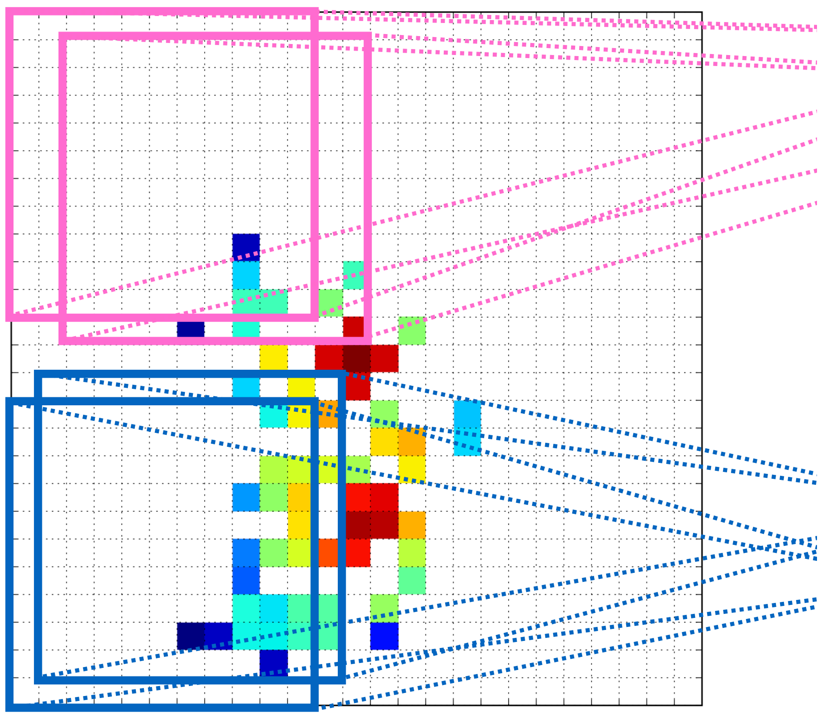
New Physics from Precision at High Energy

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com
bpnachman@lbl.gov
@bpnachman
bnachman



BERKELEY LAB

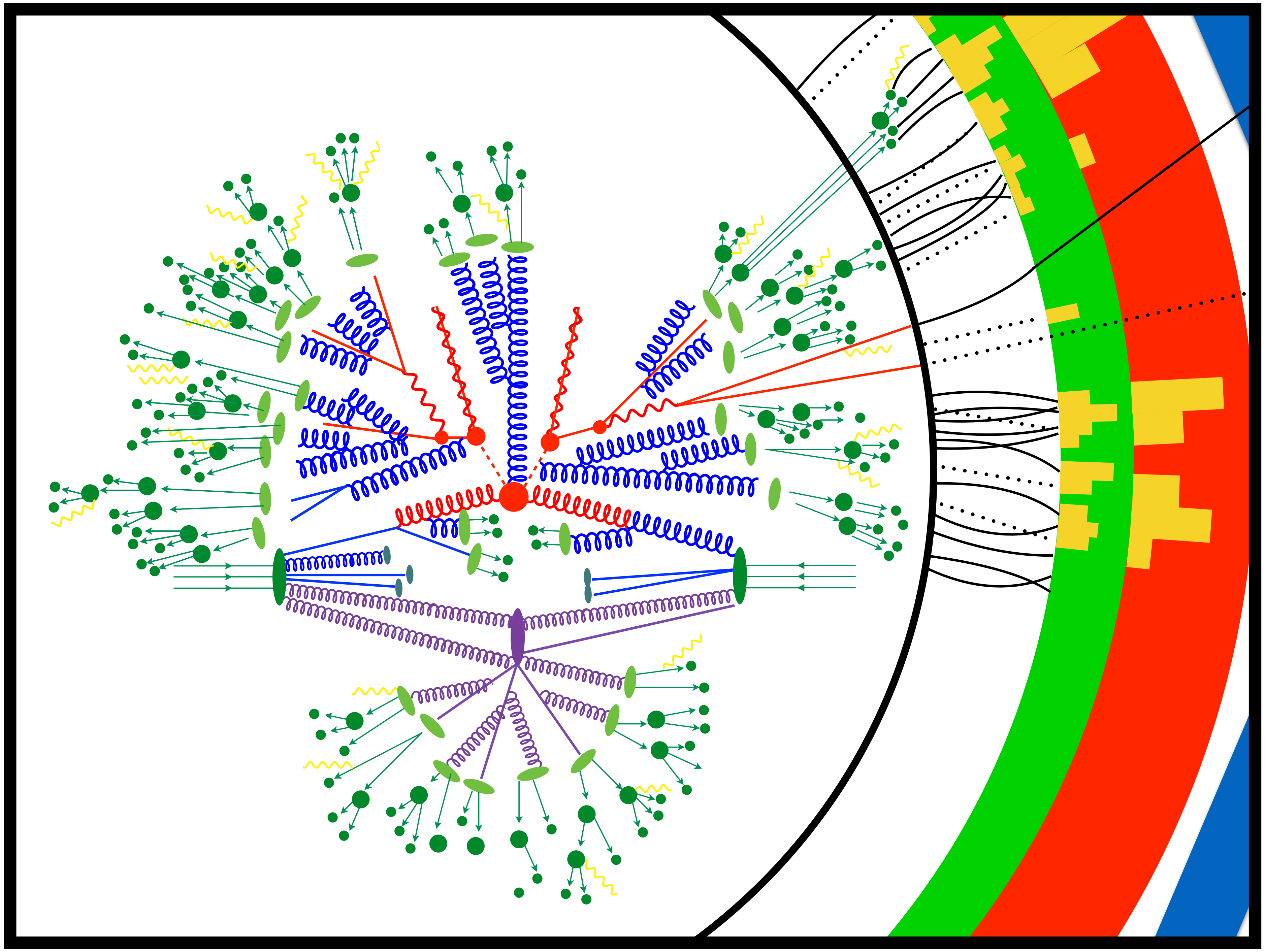
BIDS
BERKELEY INSTITUTE FOR DATA SCIENCE

LISHEP
Session C
July 8, 2021

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM

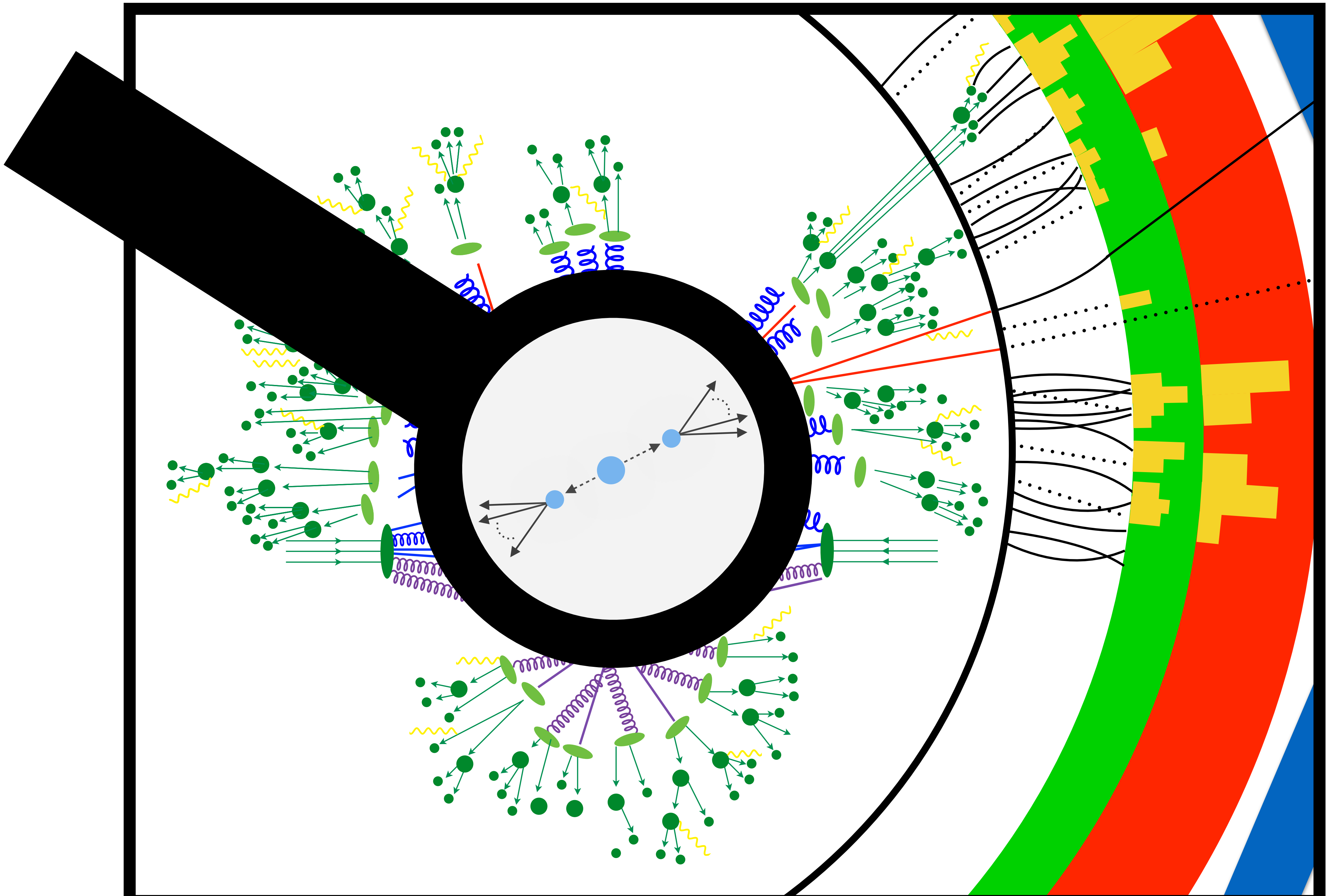


→ KITP workshop: New Physics from Precision at High Energies

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM

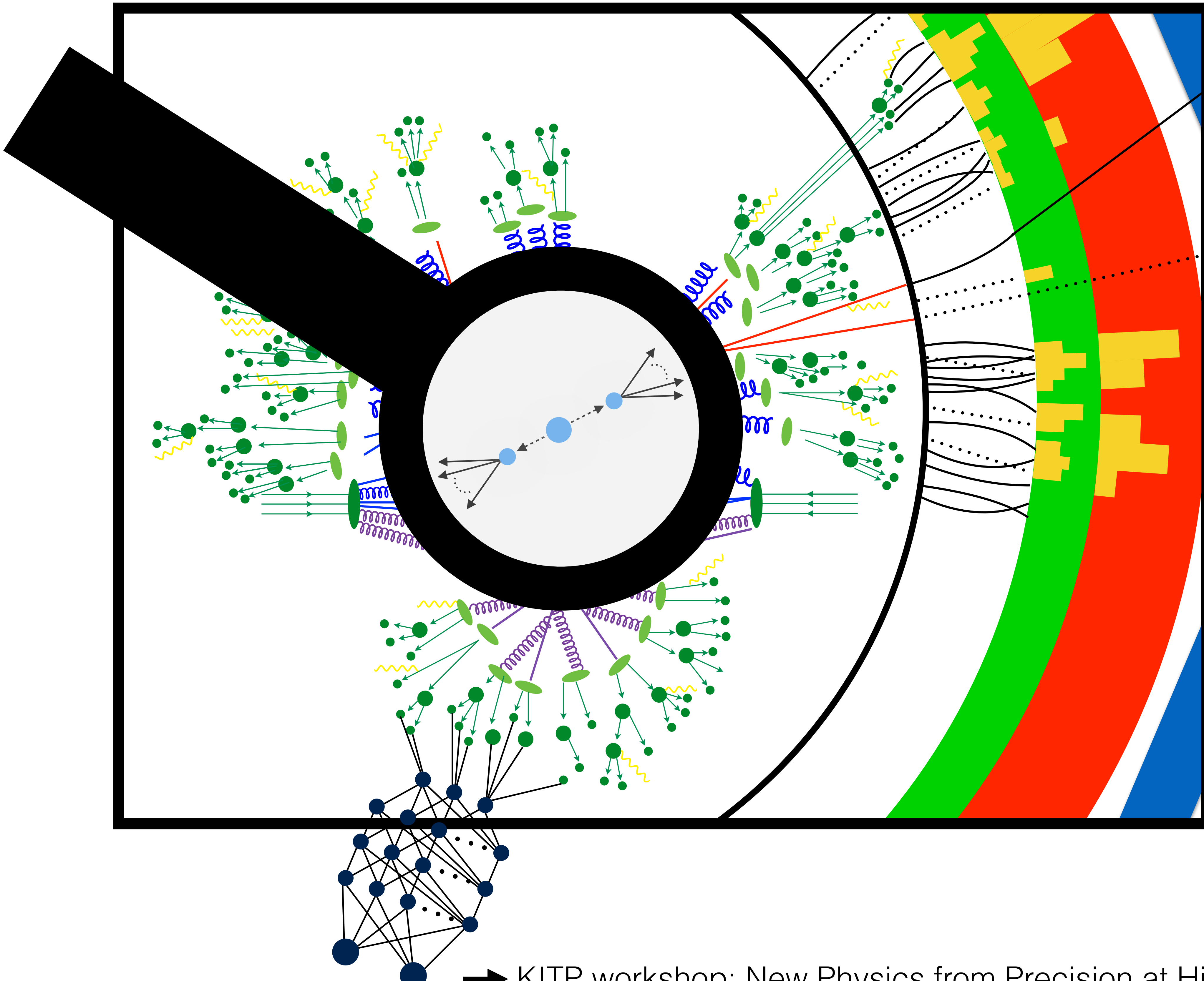


→ KITP workshop: New Physics from Precision at High Energies

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM

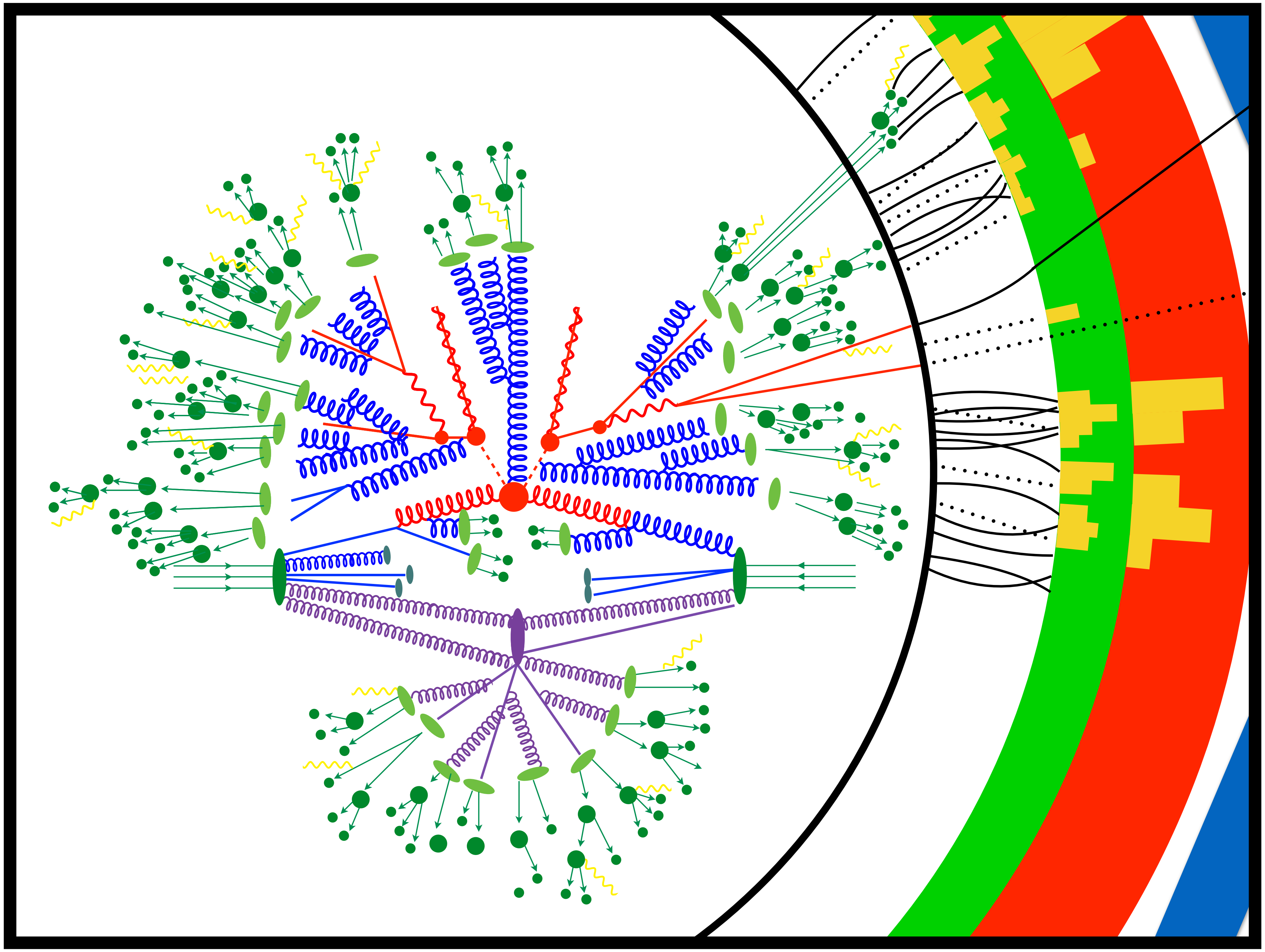


→ KITP workshop: New Physics from Precision at High Energies

(1) Precision measurements at high energy

(2) New physics from precision measurements

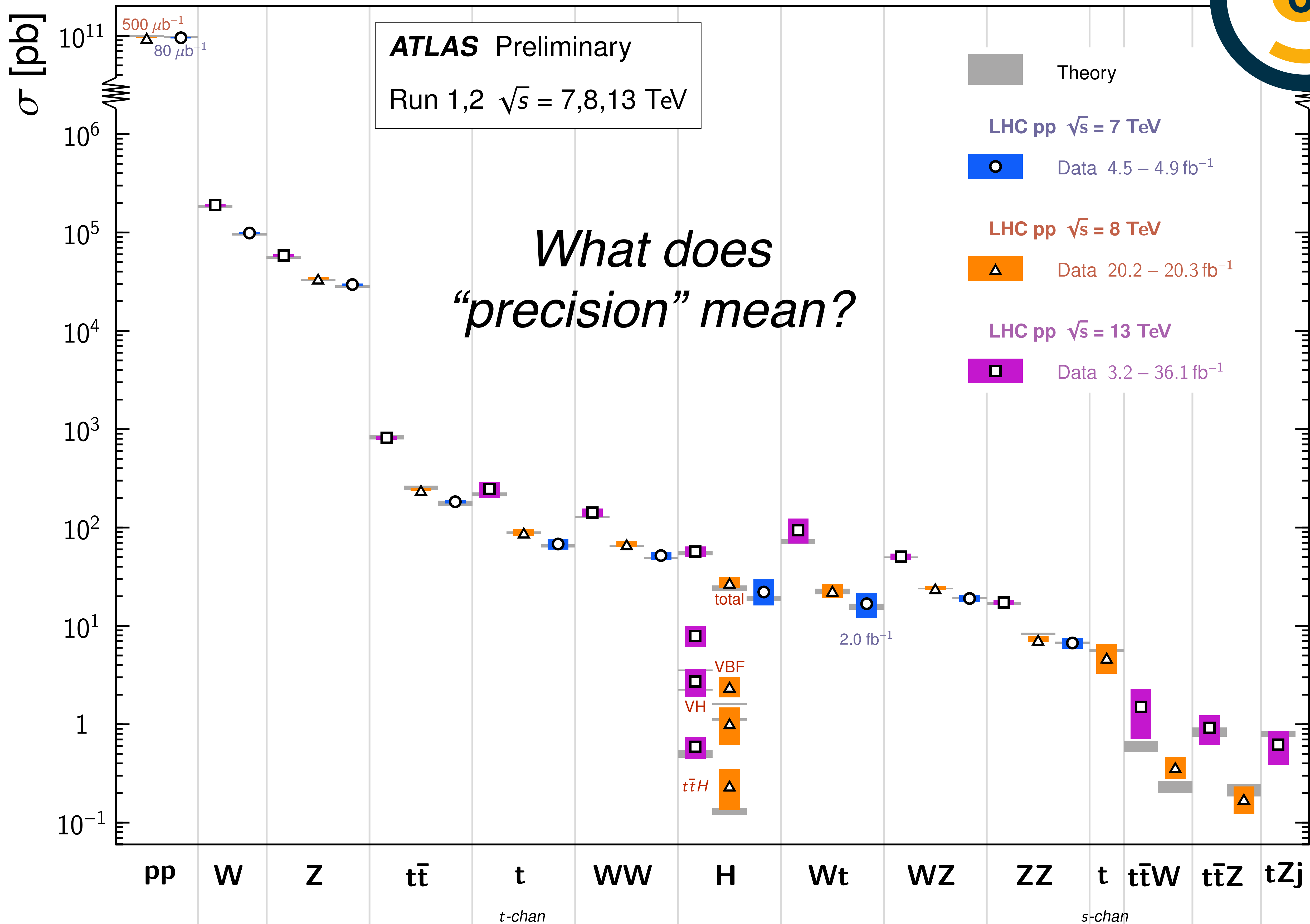
(3) Machine learning + measurements + BSM

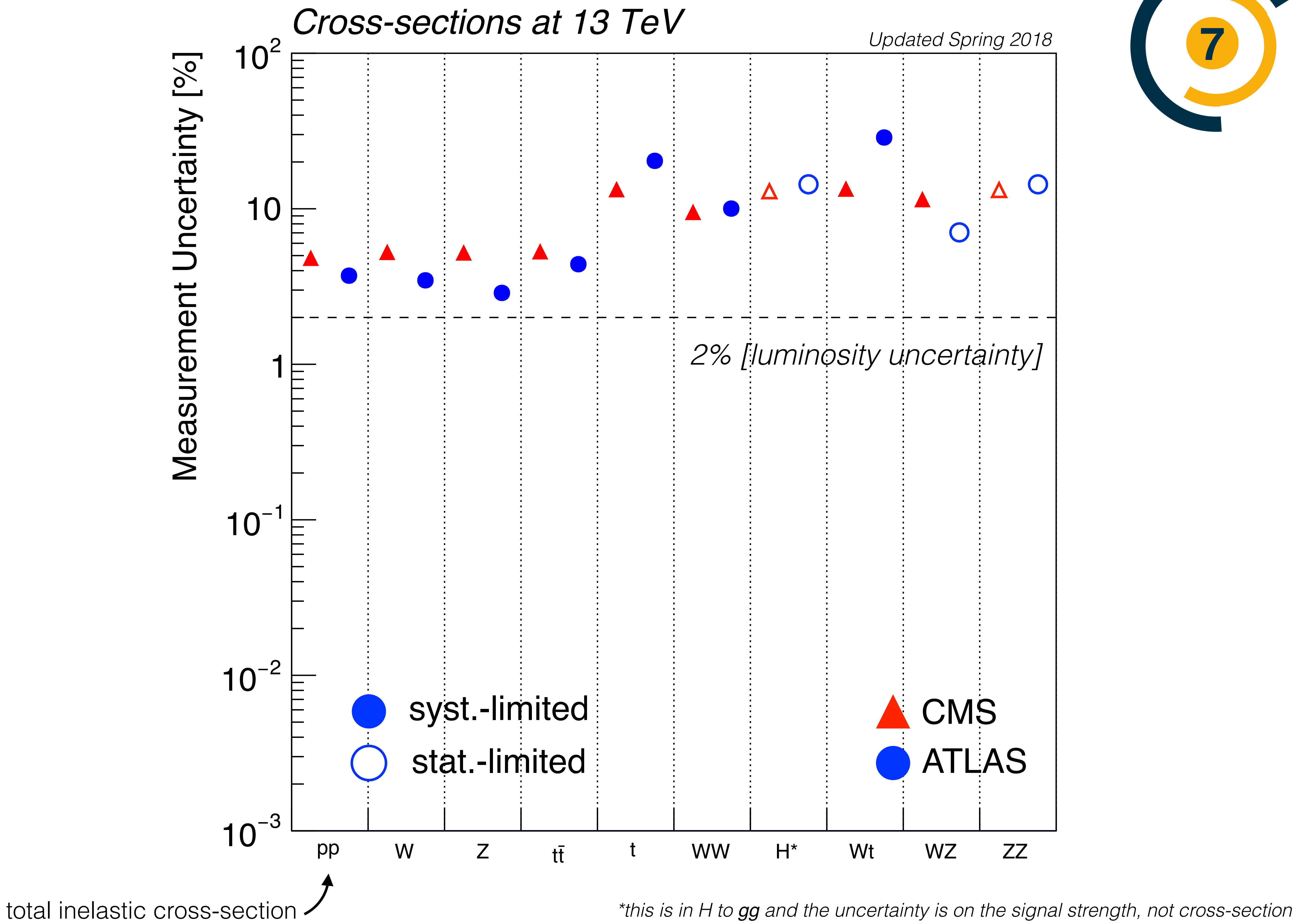


→ KITP workshop: New Physics from Precision at High Energies

Standard Model Total Production Cross Section Measurements

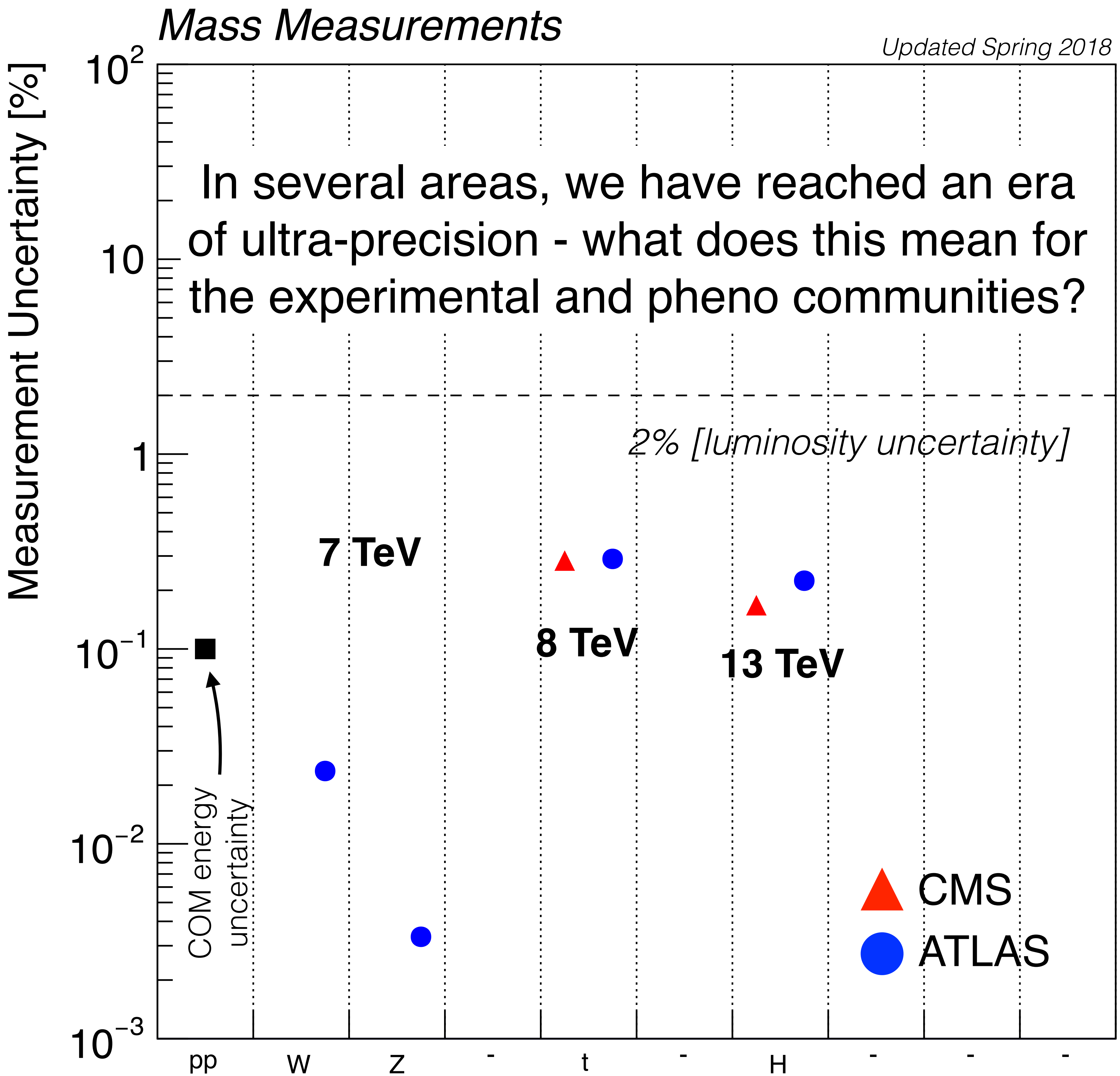
Status: March 2018



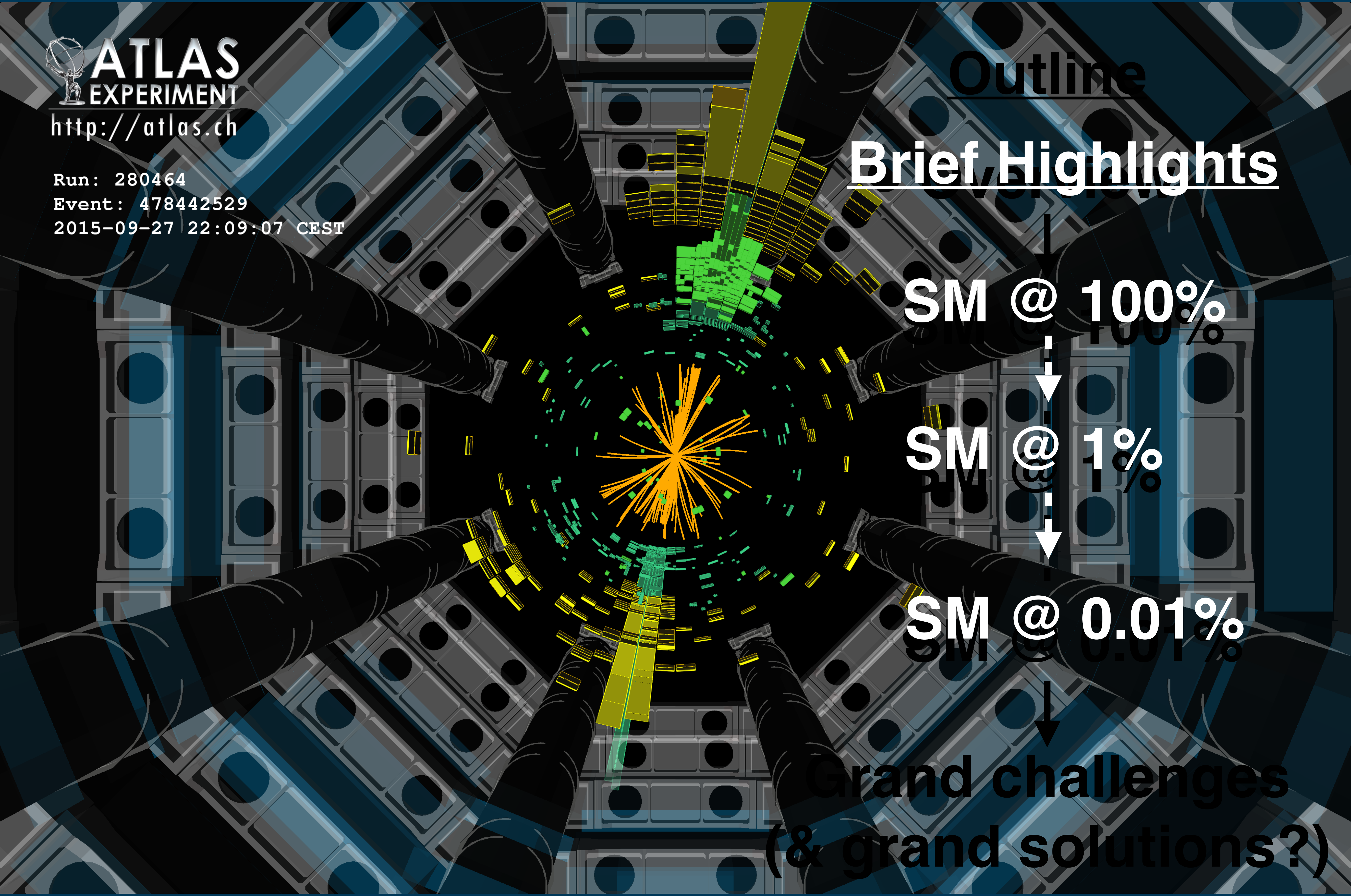


Cross-sections at 13 TeV
Updated Spring 2018
Measurement Uncertainty [%]
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
2% [luminosity uncertainty]
syst.-limited
stat.-limited
CMS
ATLAS
pp
W
Z
$t\bar{t}$
t
WW
H*
Wt
WZ
ZZ
total inelastic cross-section
*this is in H to **gg** and the uncertainty is on the signal strength, not cross-section

In several areas, we have reached an era of ultra-precision - what does this mean for the experimental and pheno communities?



Fun fact: beam energy uncertainty (■) recently improved - significant impact on ttbar cross-section uncertainty!

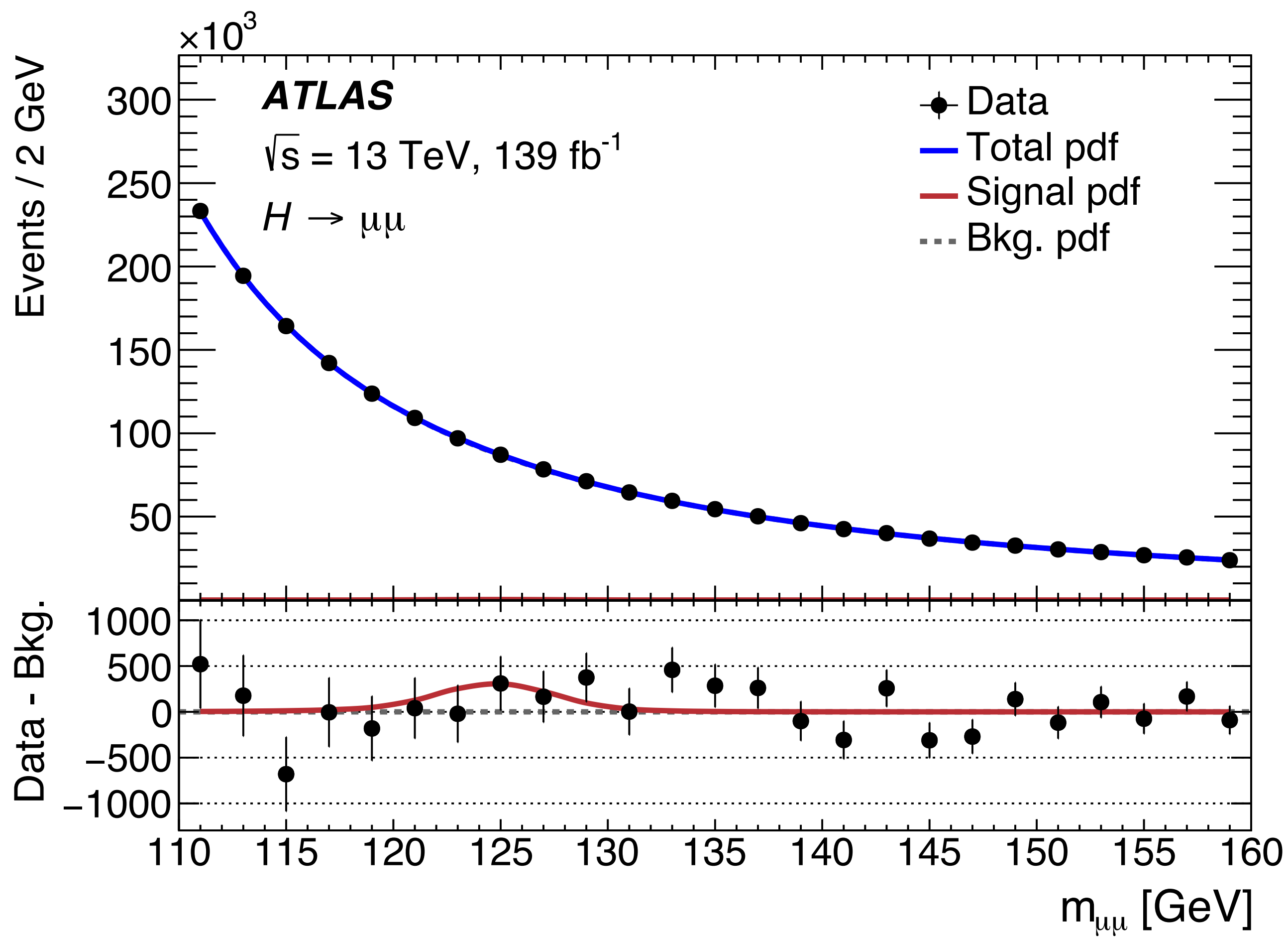

ATLAS
EXPERIMENT
http://atlas.ch
Run: 280464
Event: 478442529
2015-09-27 22:09:07 CEST
Outline
Brief Highlights
SM @ 100%
SM @ 1%
SM @ 0.01%
Grand challenges
(& grand solutions?)

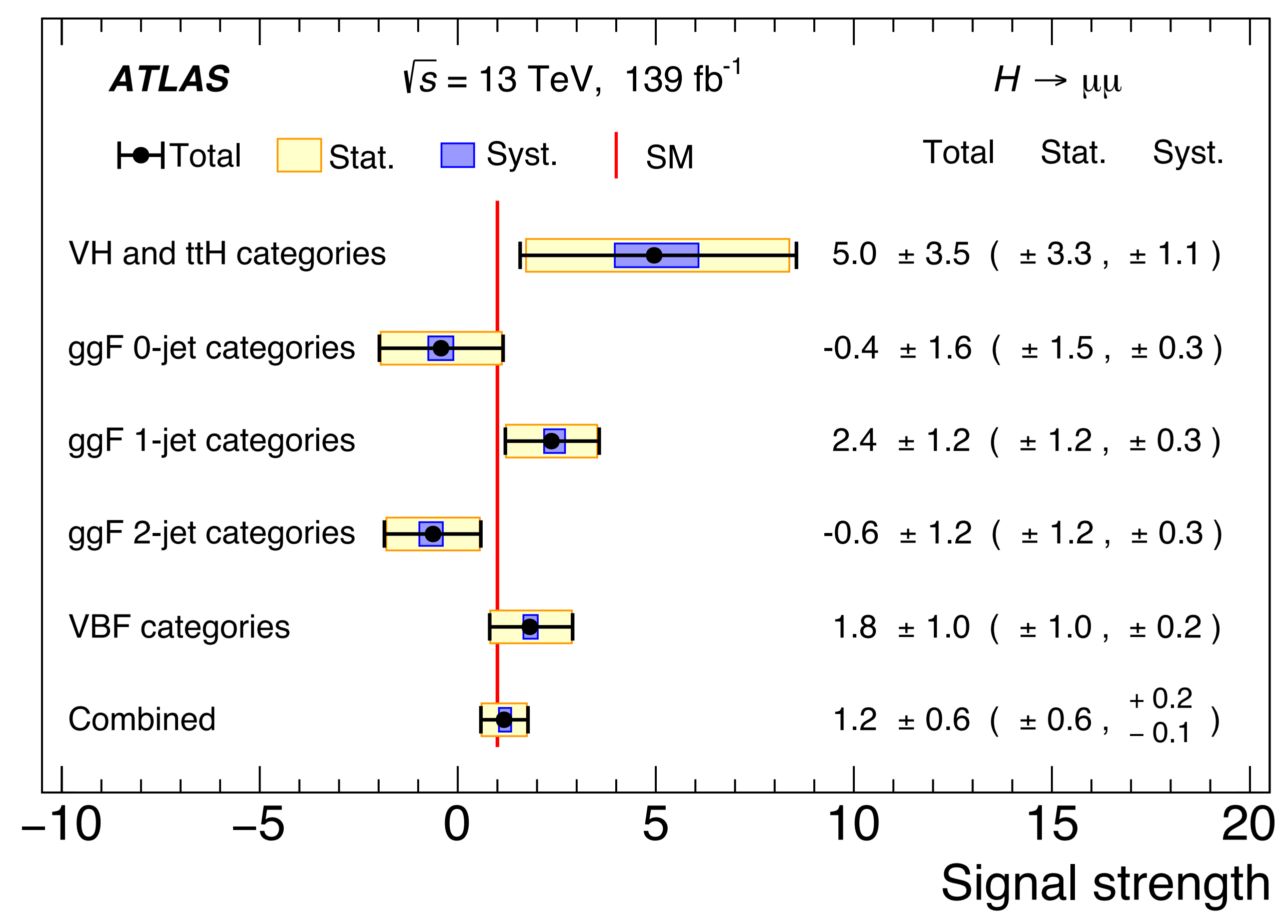
SM @ 100%

"Measurements" with ~100% uncertainty
are still in the "search" mode.

SM @ 100%

"Measurements" with ~100% uncertainty are still in the "search" mode.

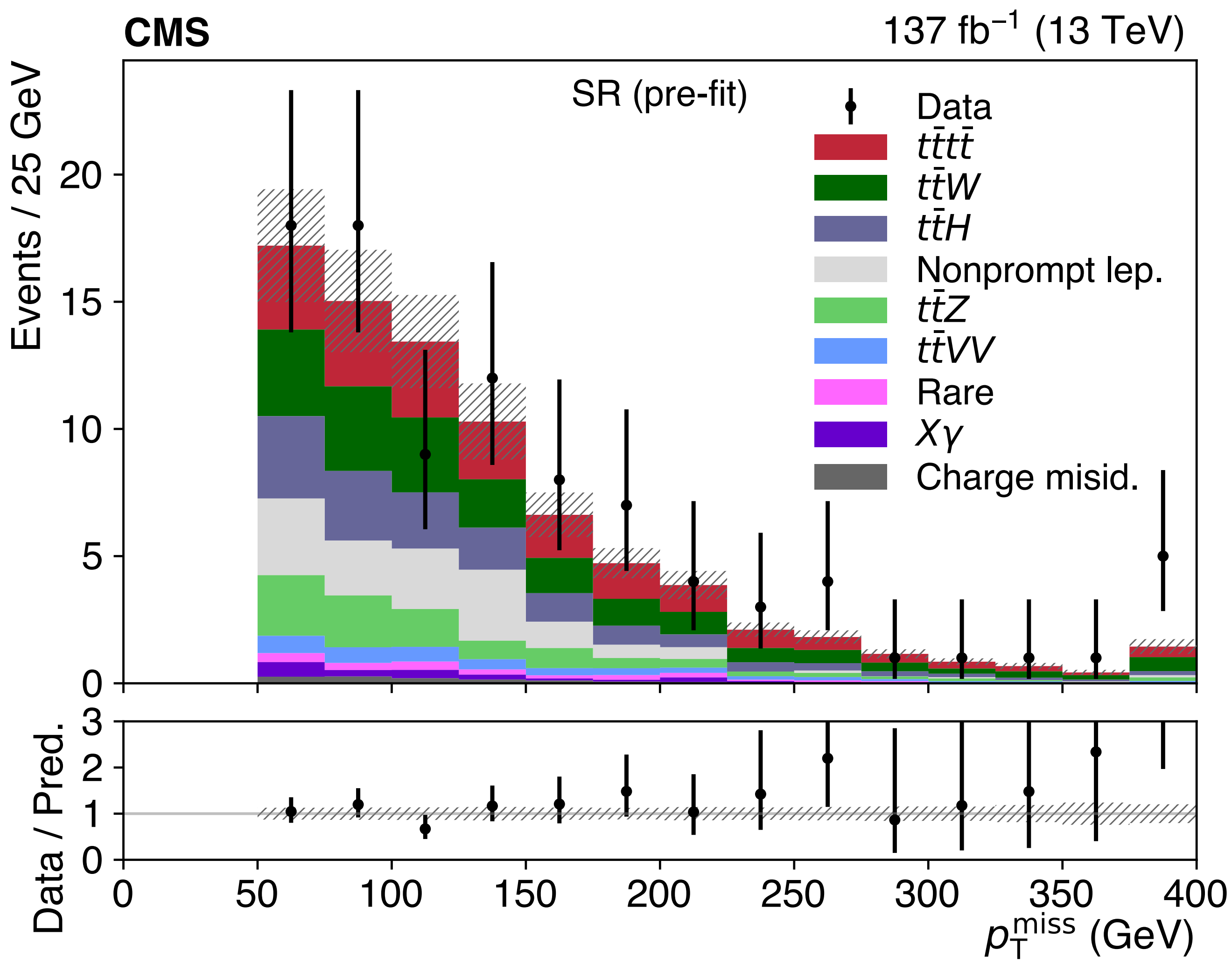


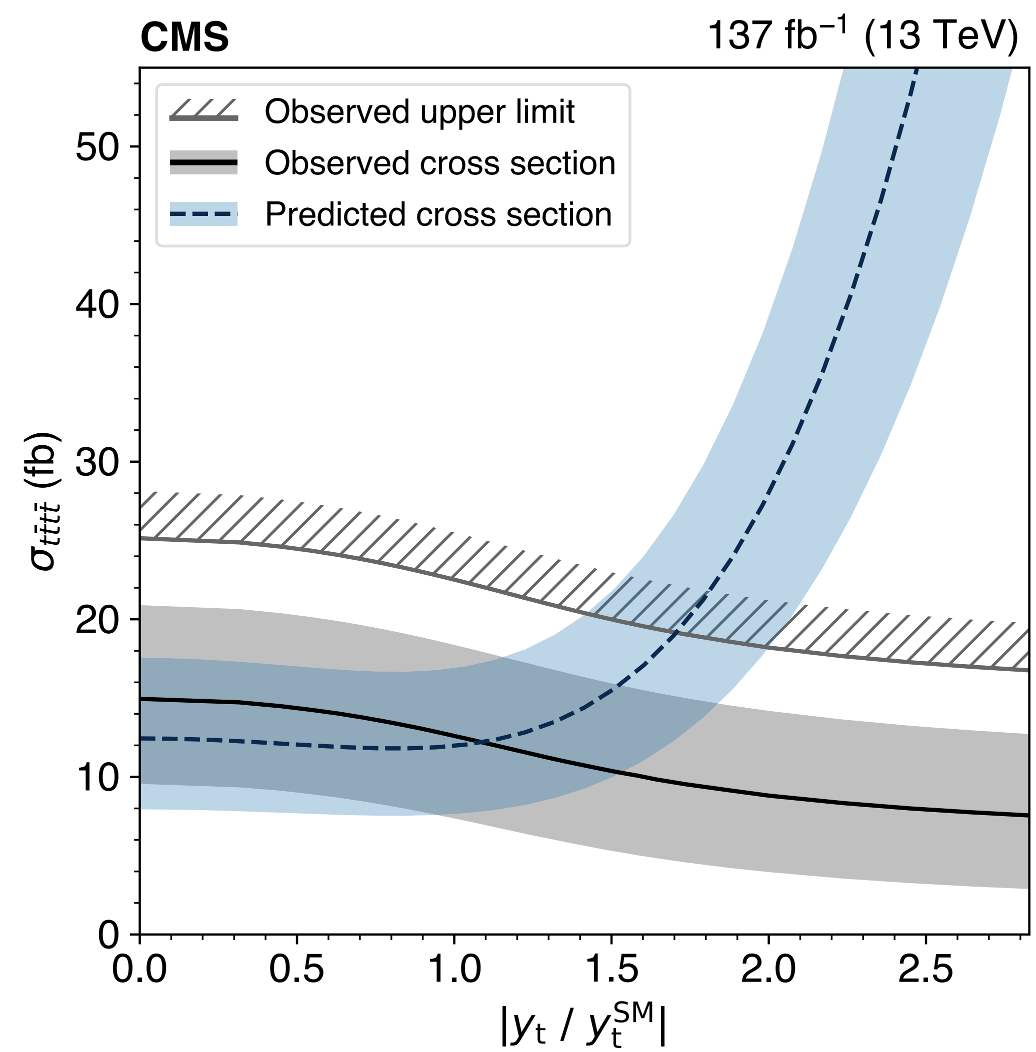


Example: Higgs to muons

SM @ 100%

“Measurements” with ~100% uncertainty are still in the “search” mode.

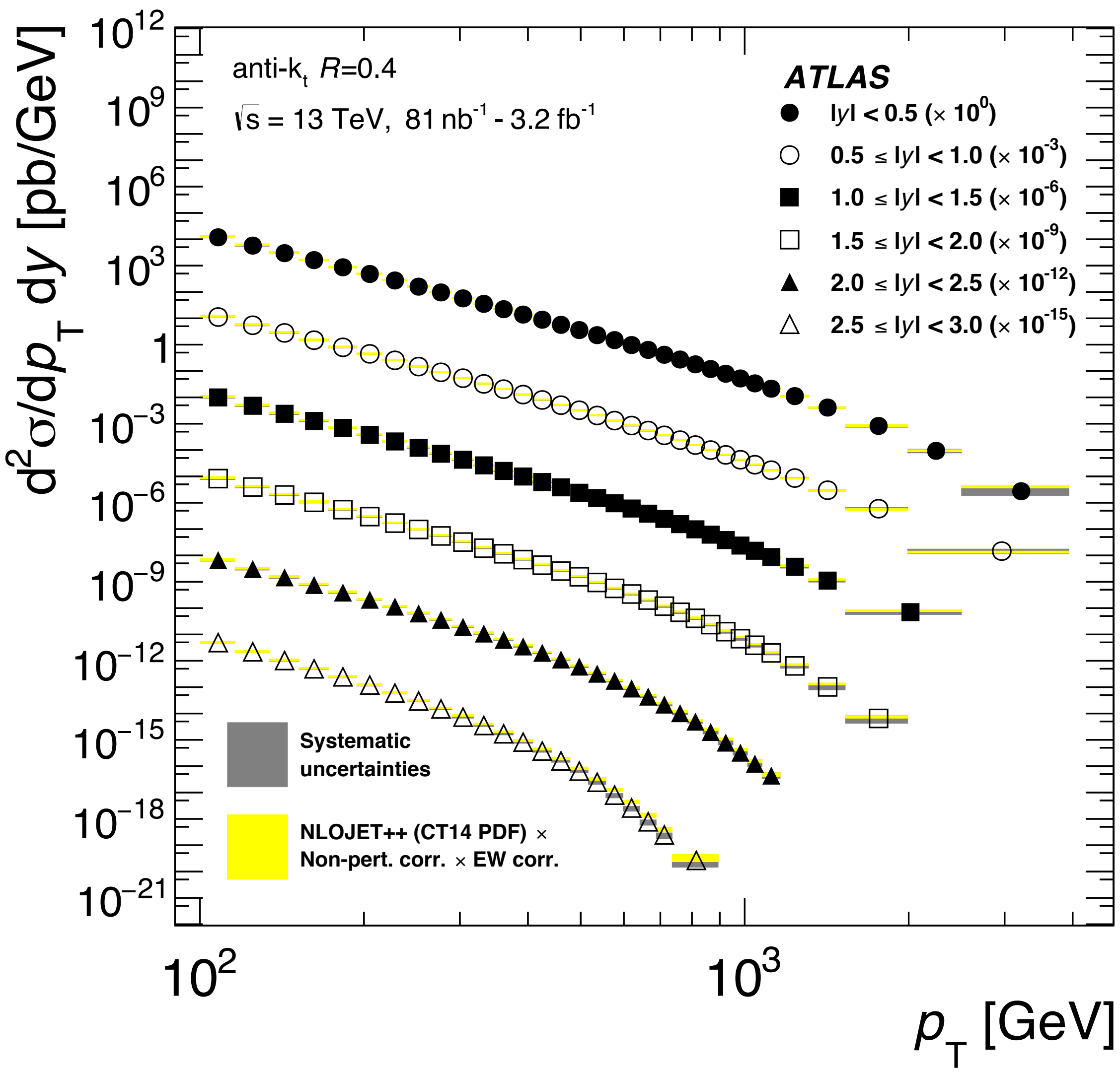


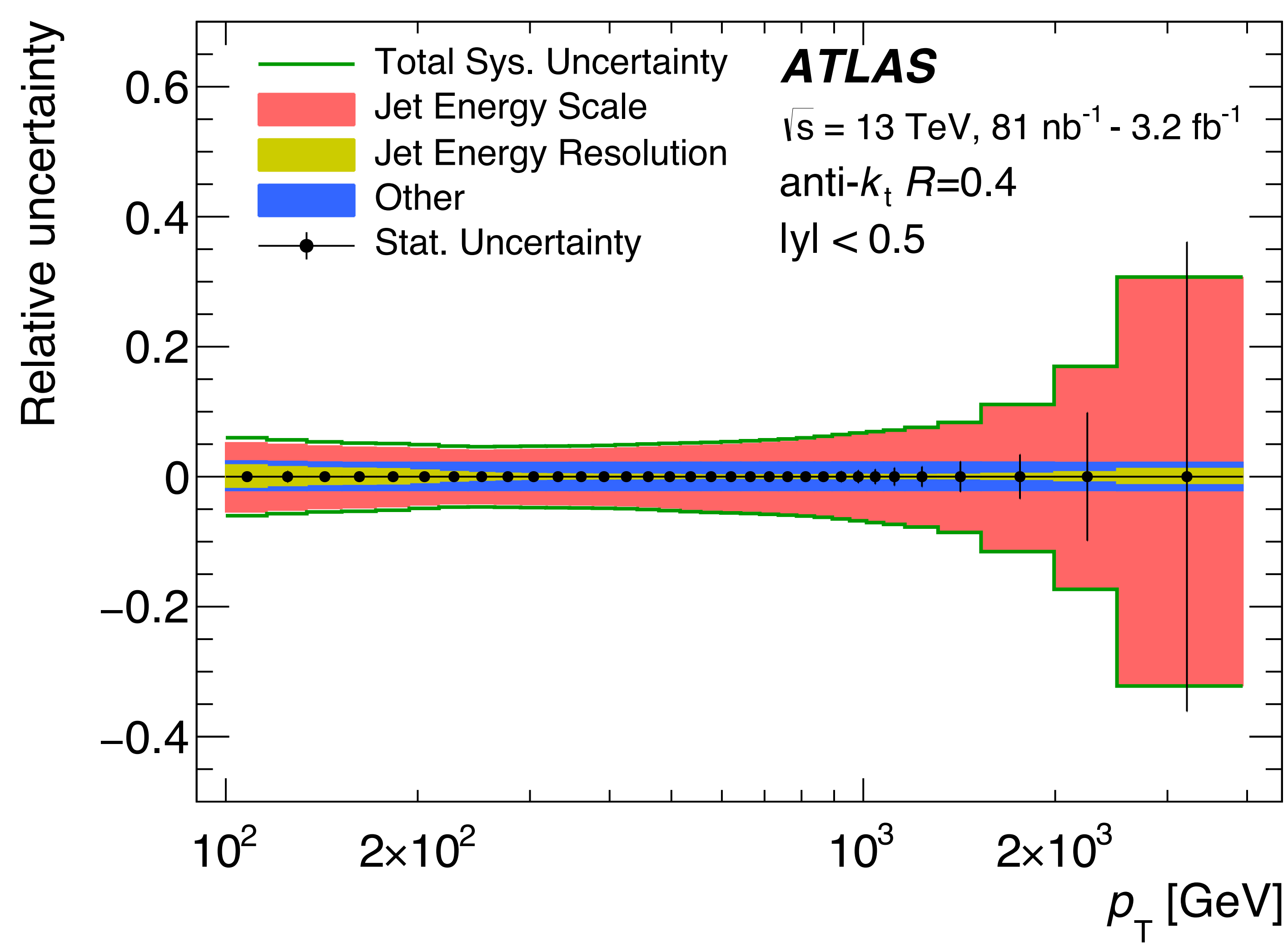


Example: Four top quarks

SM @ 10%

“Measurements” with 10% uncertainty can begin to probe differential cross sections.

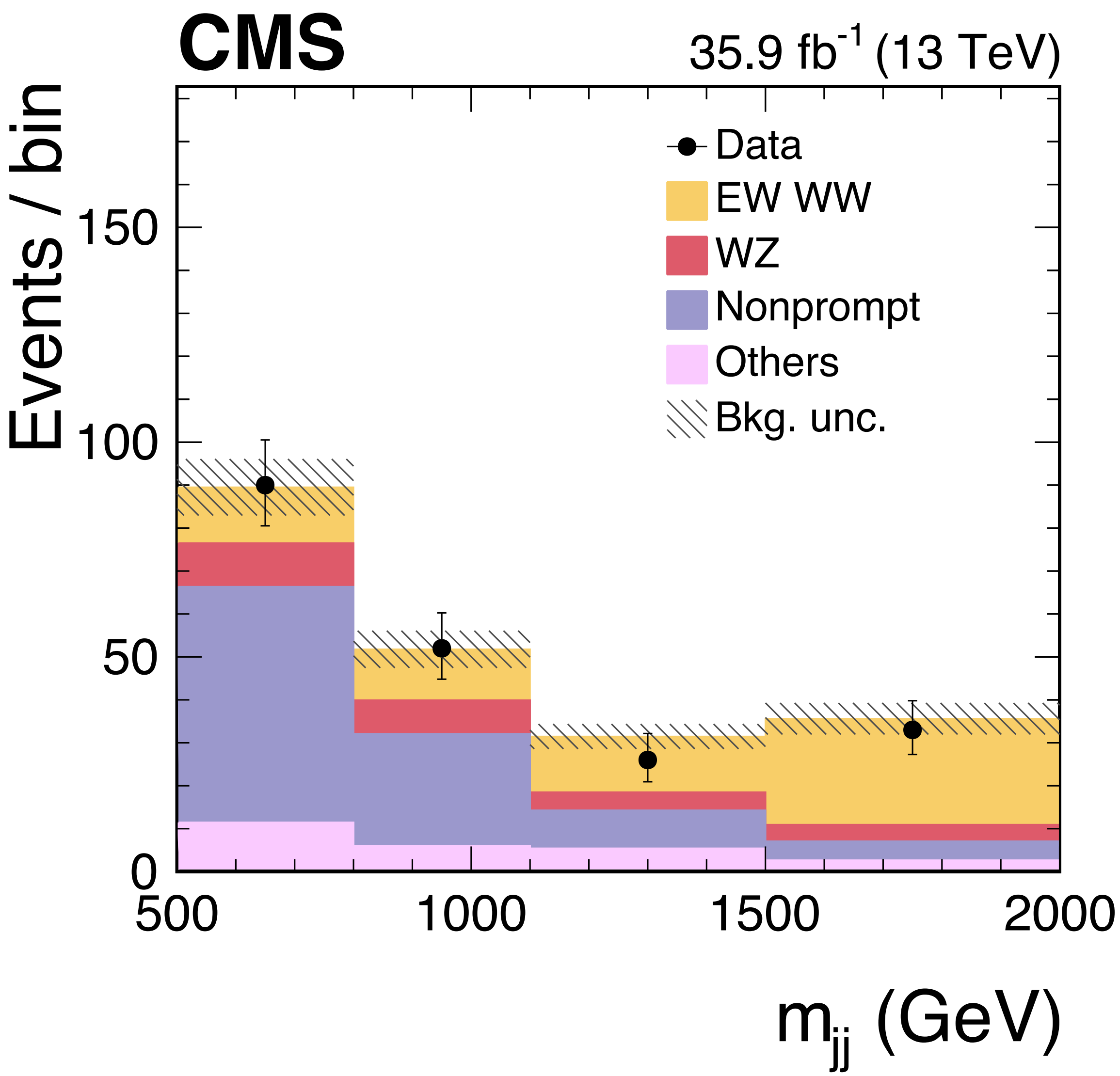


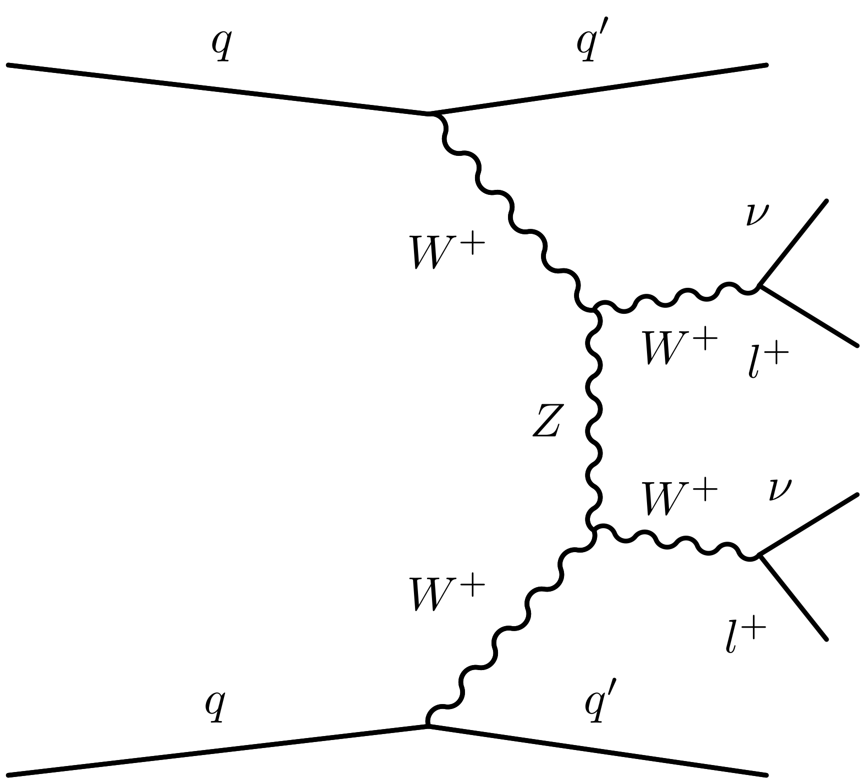


Example: Dijets

SM @ 10%

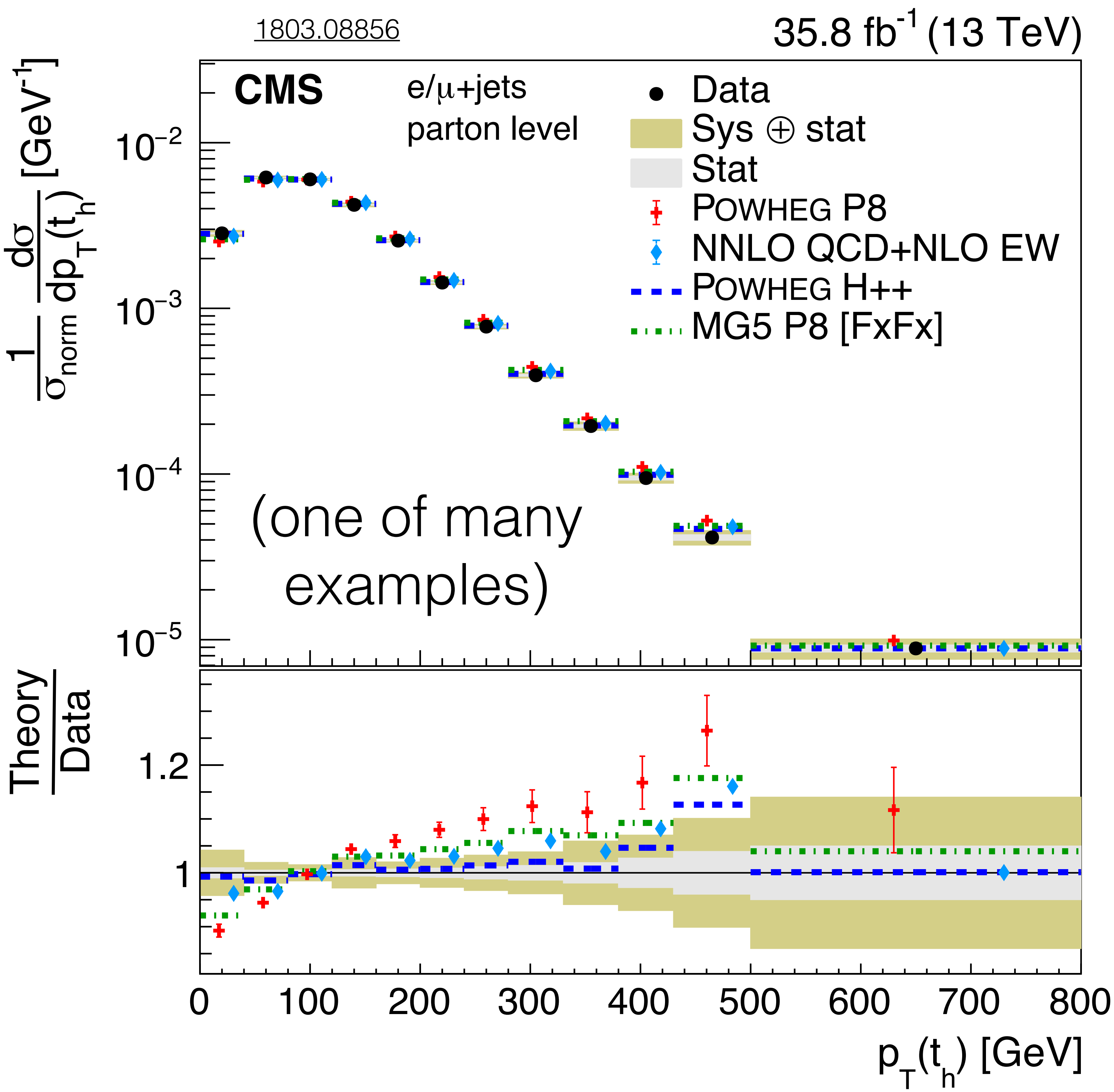
"Measurements" with 10% uncertainty can begin to probe differential cross sections.





Example: vector boson scattering

SM @ 1%



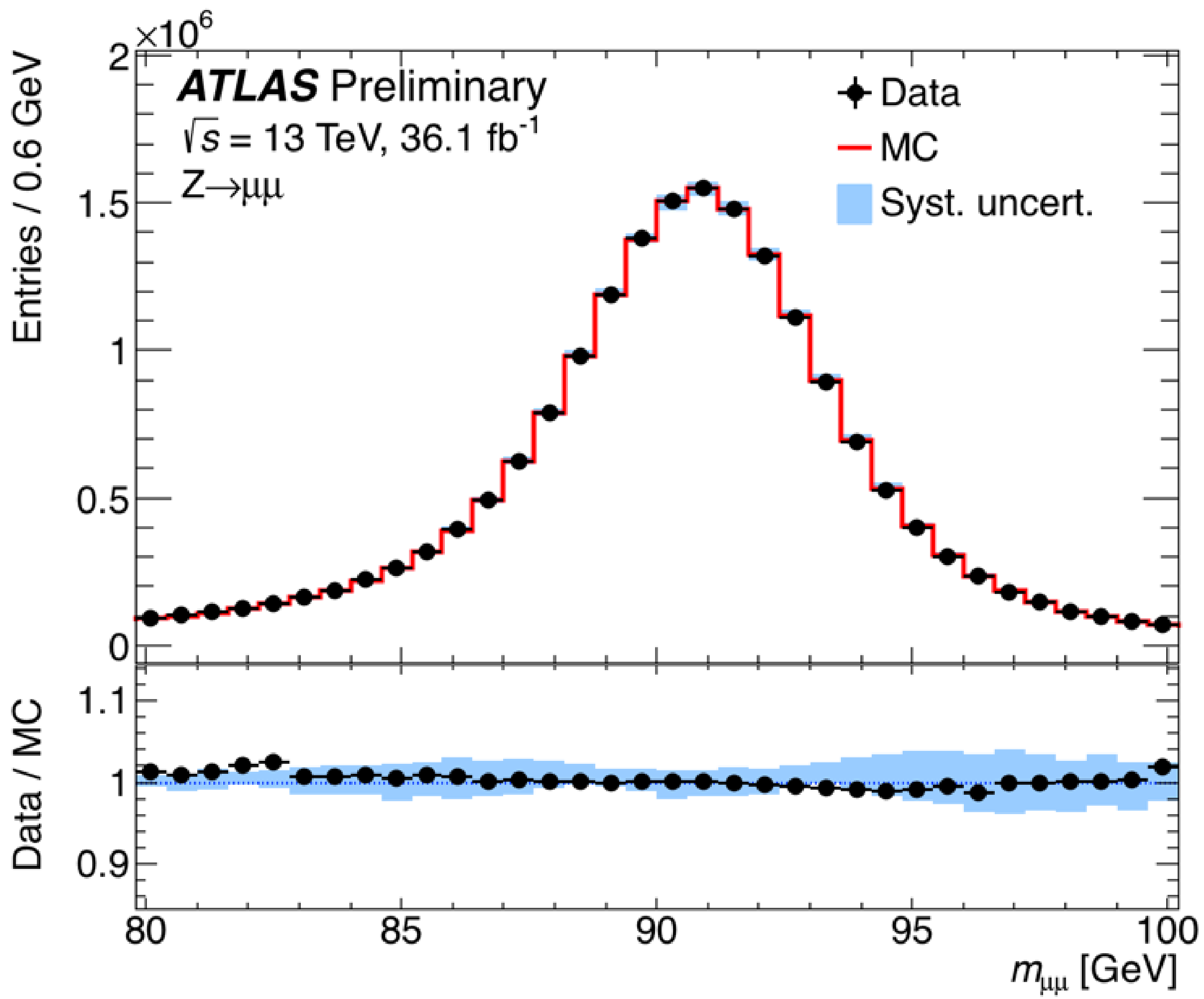
Differential cross-sections of W, Z, top are reaching 1%

Matching or exceeding precision of corresponding calculations

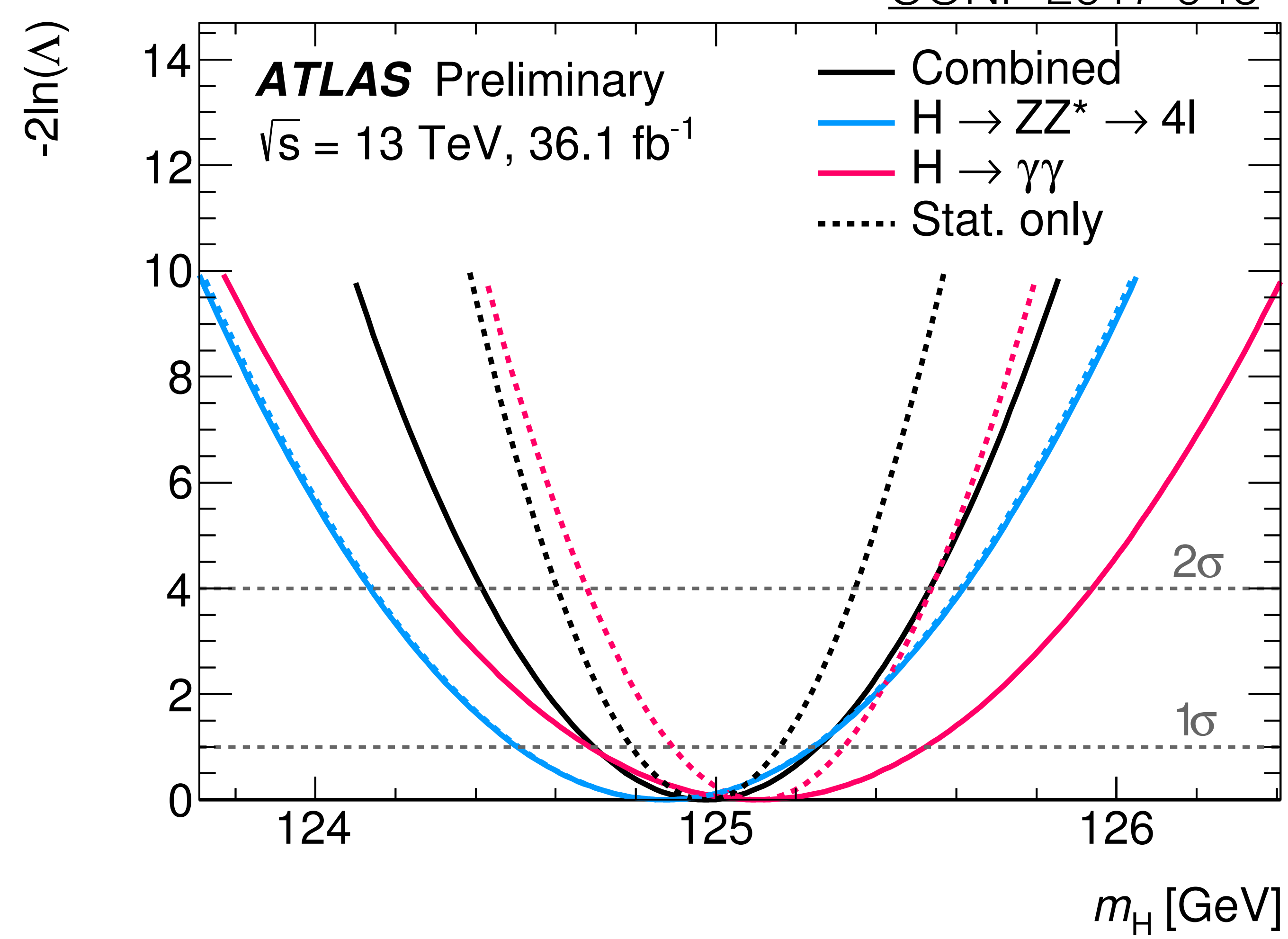
At this level of precision need to be careful what is measured!
(e.g. parton != particle)

SM @ 0.1%

Even though newest fundamental particle, m_H very well known

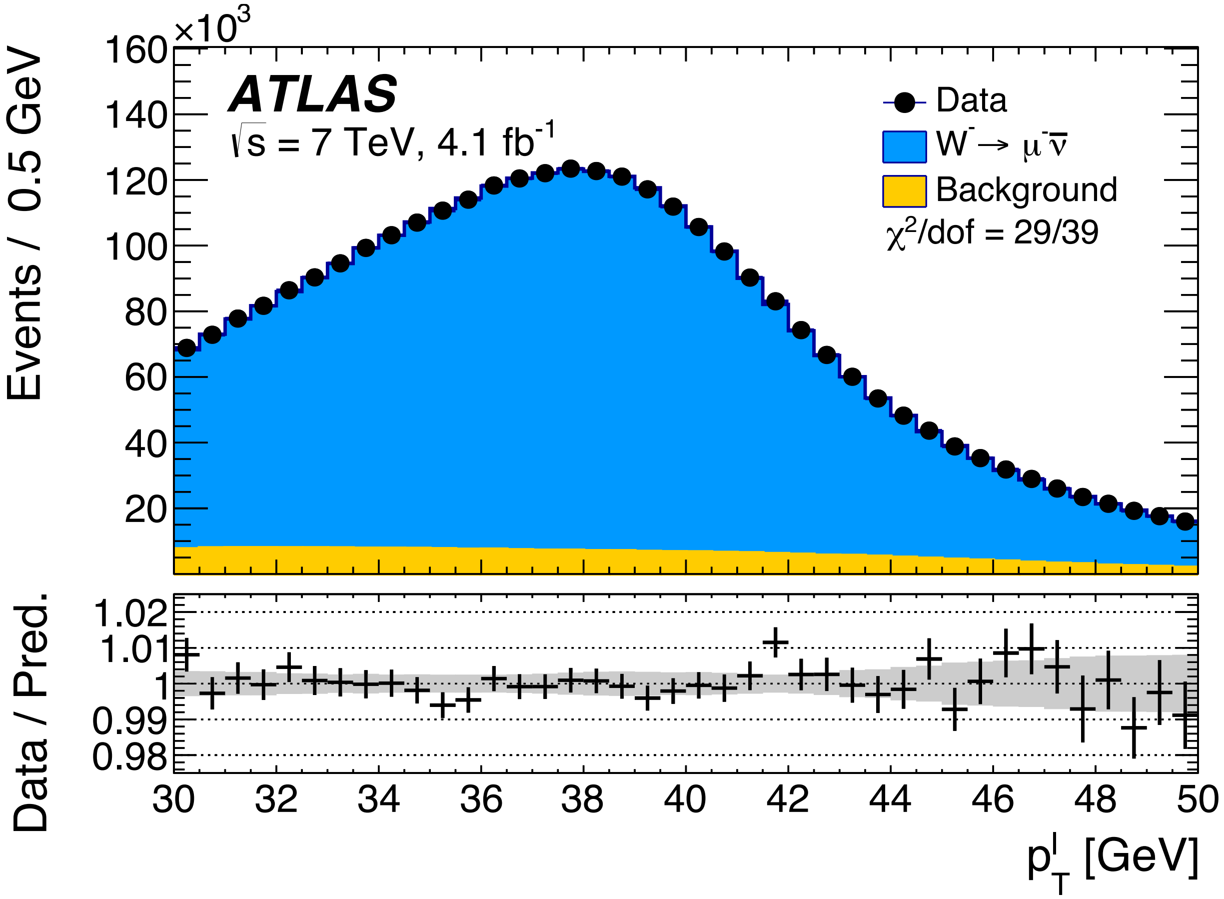


CONF-2017-046



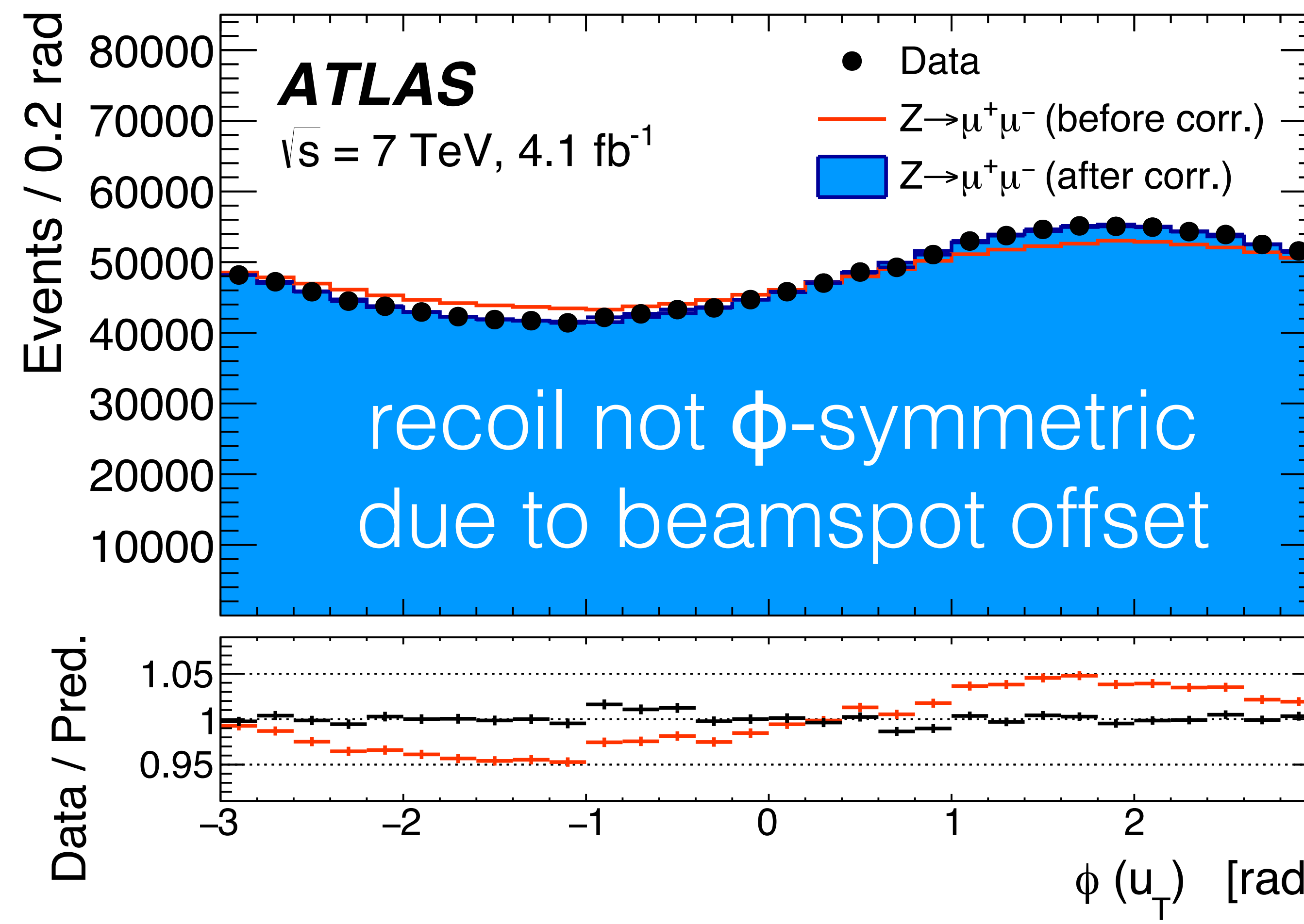
Requires superb understanding of muon momentum scale and e/γ energy scale and resolution

SM @ 0.01% - Ultra Precision: W mass



Even though it is 7 TeV, the W mass measurement was published in 2018

Eur. Phys. J. C 78 (2018) 110

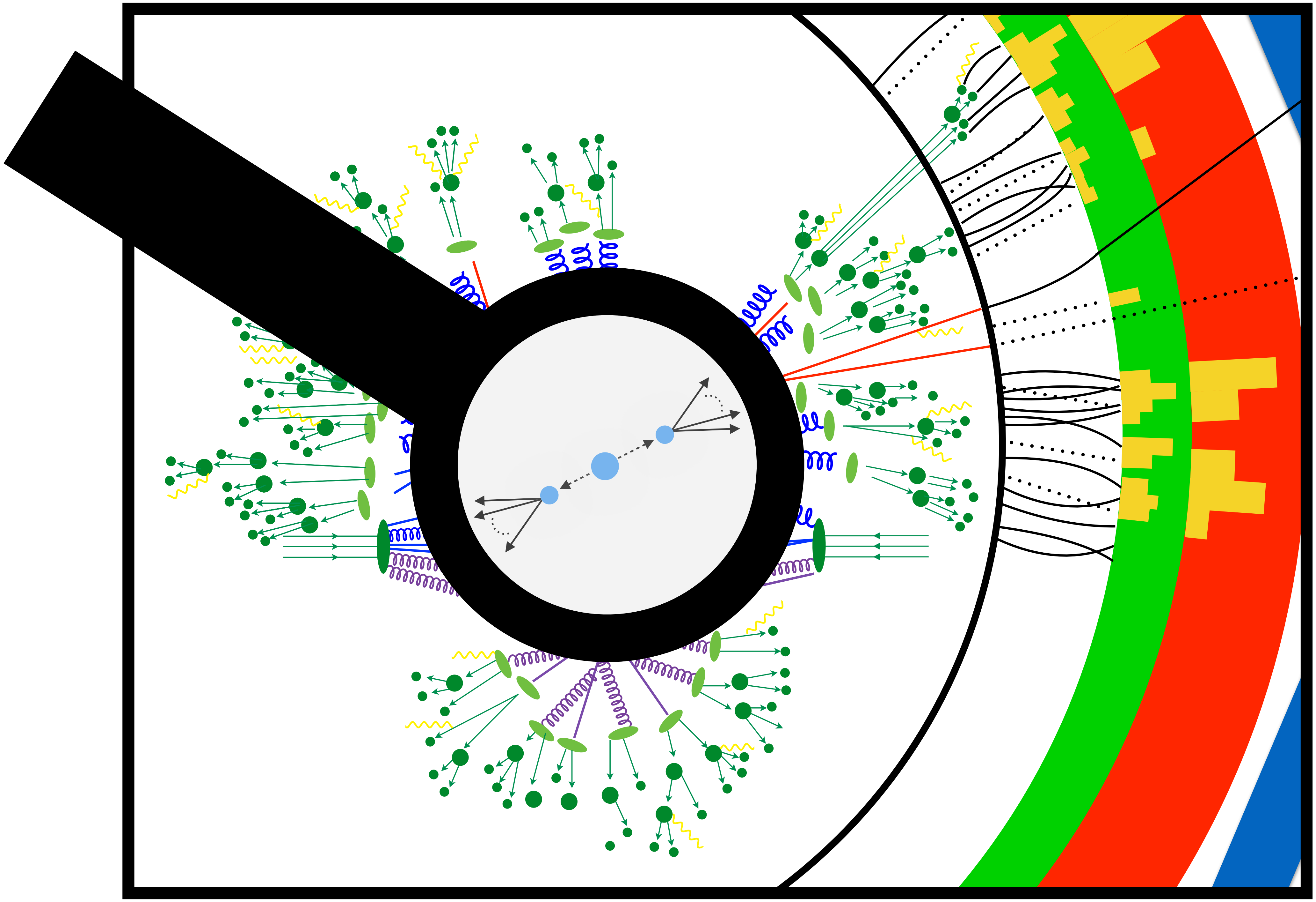


To achieve a 20 MeV uncertainty, need not only excellent uncertainties but dedicated calibrations (e.g. with Z p_T)

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM



→ KITP workshop: New Physics from Precision at High Energies

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM

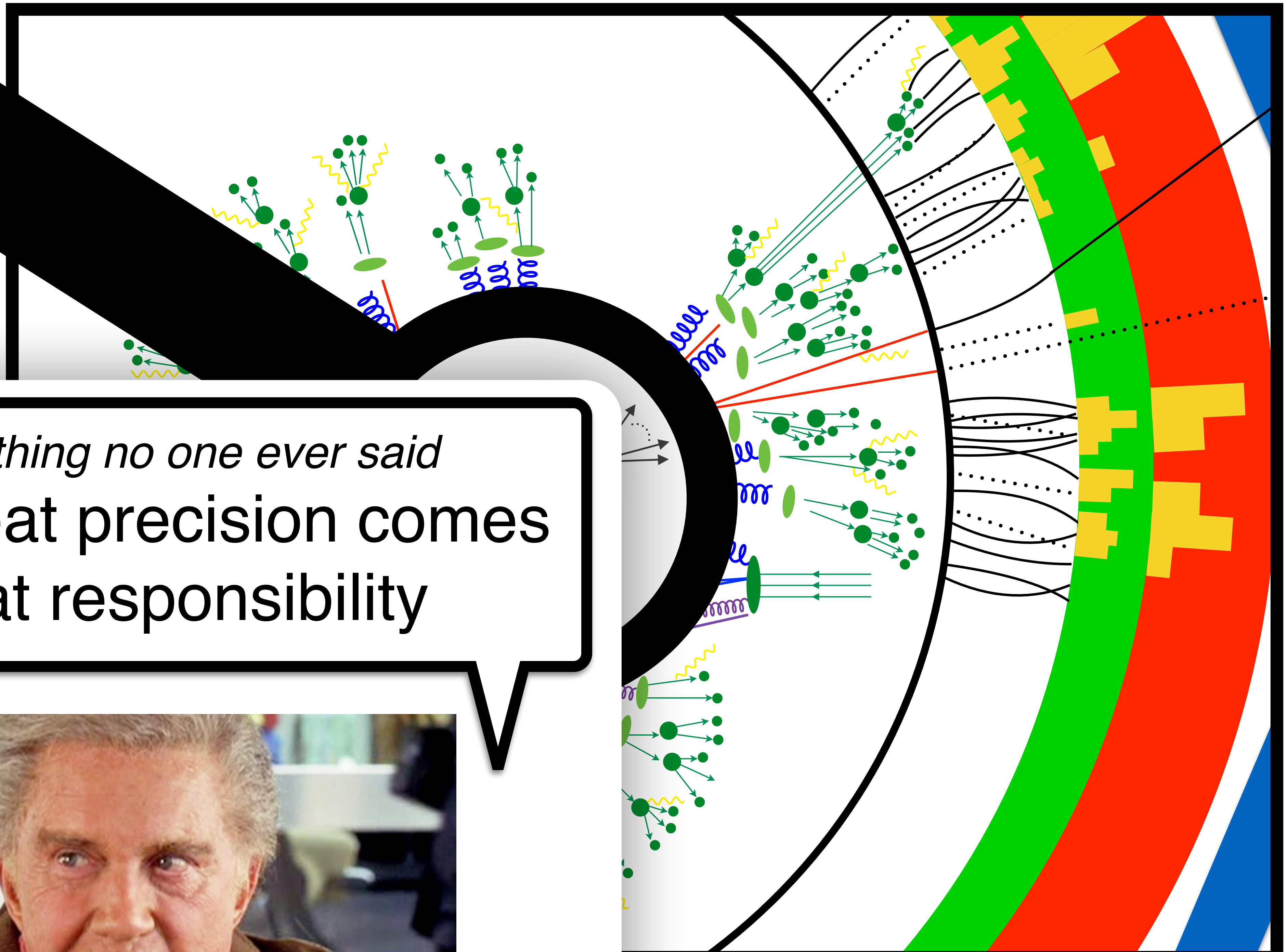
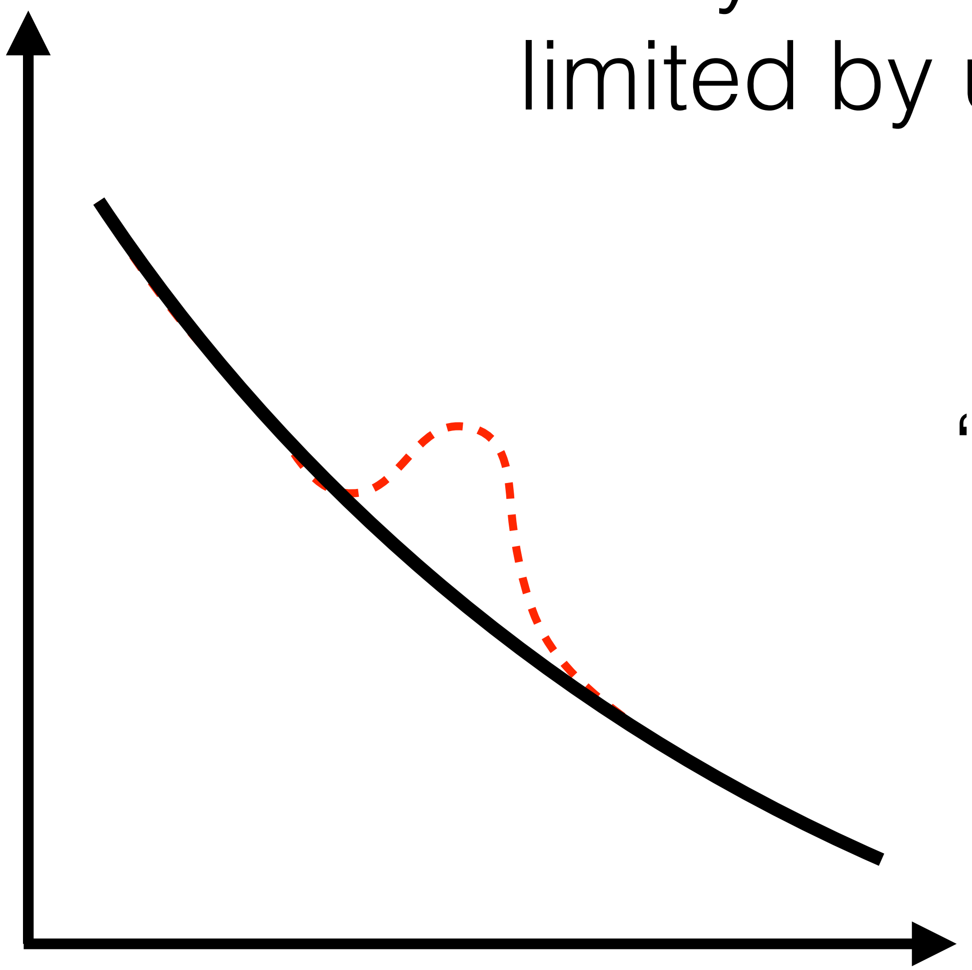


Image credit: Mavel (Spiderman)

→ KITP workshop: New Physics from Precision at High Energies

Searches

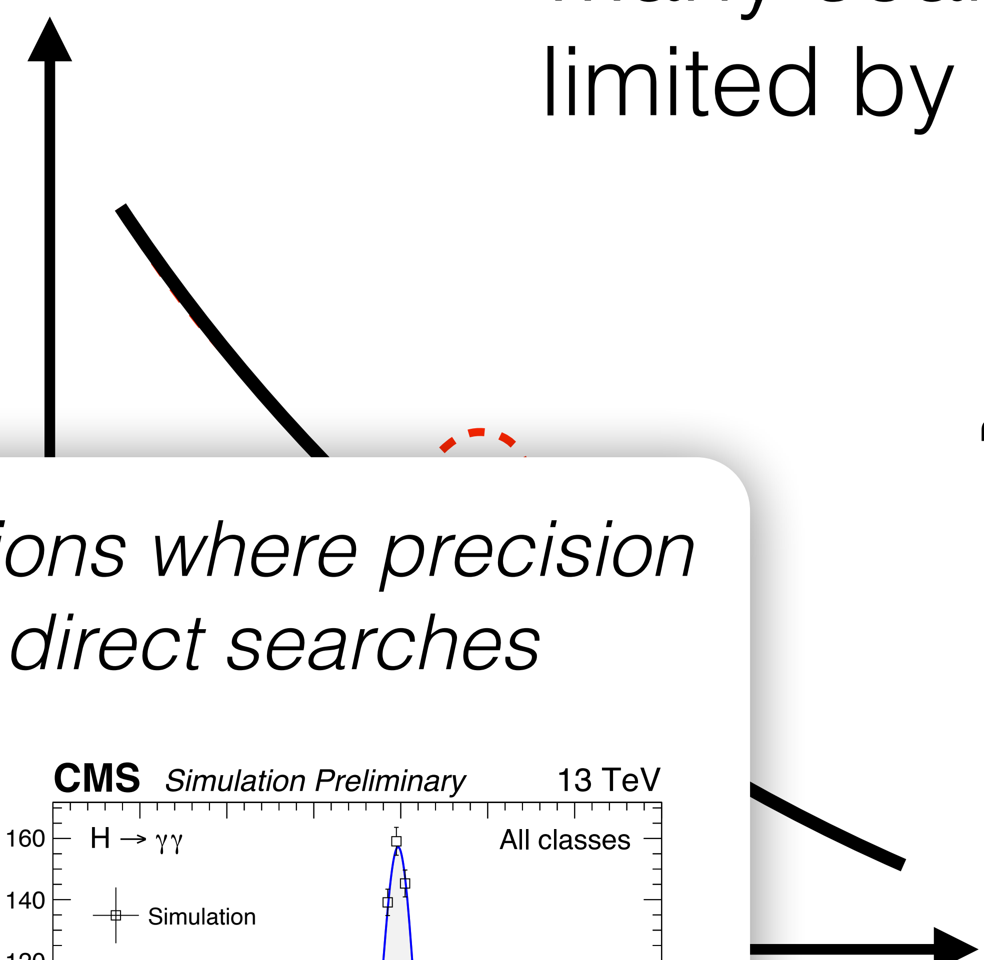


Many searches are not limited by uncertainties

Looking for "**big**" effects

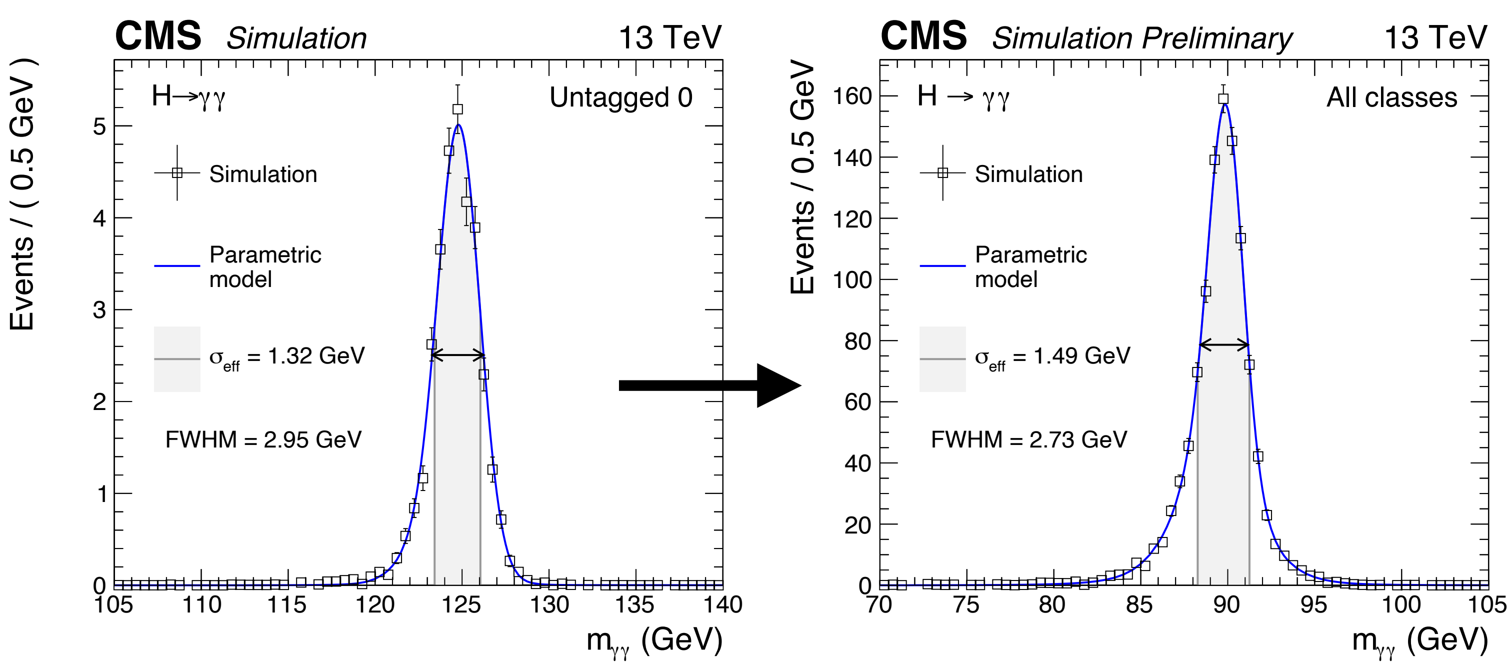
Searches

Many searches are not limited by uncertainties

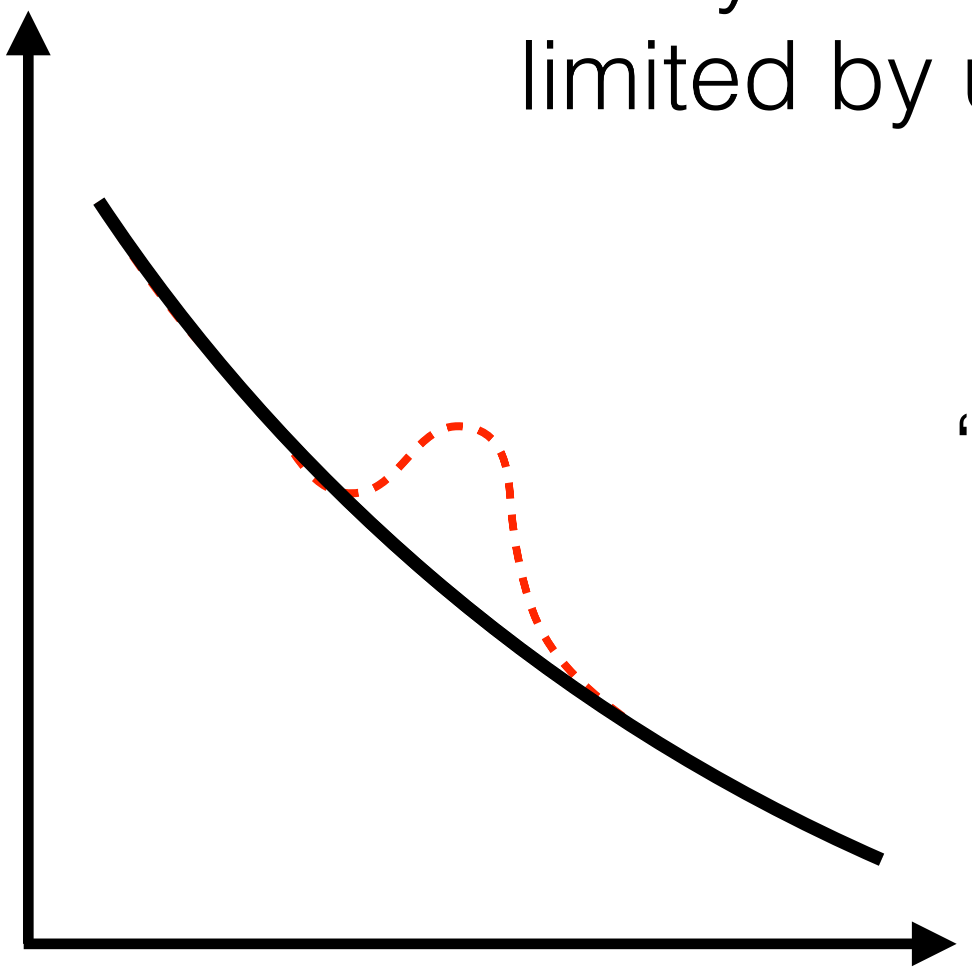


Looking for "**big**" effects

There are some exceptions where precision is also required for direct searches



Searches

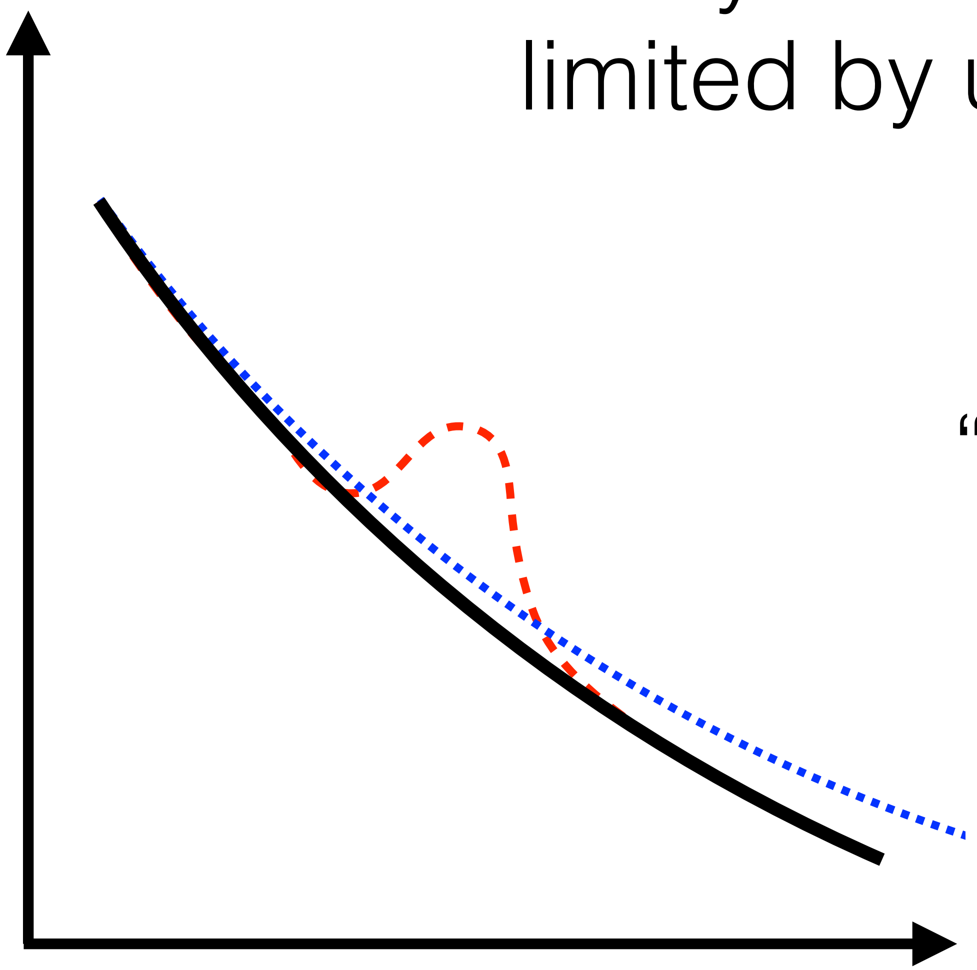


Many searches are not limited by uncertainties

Looking for "**big**" effects

Searches

Many searches are not limited by uncertainties

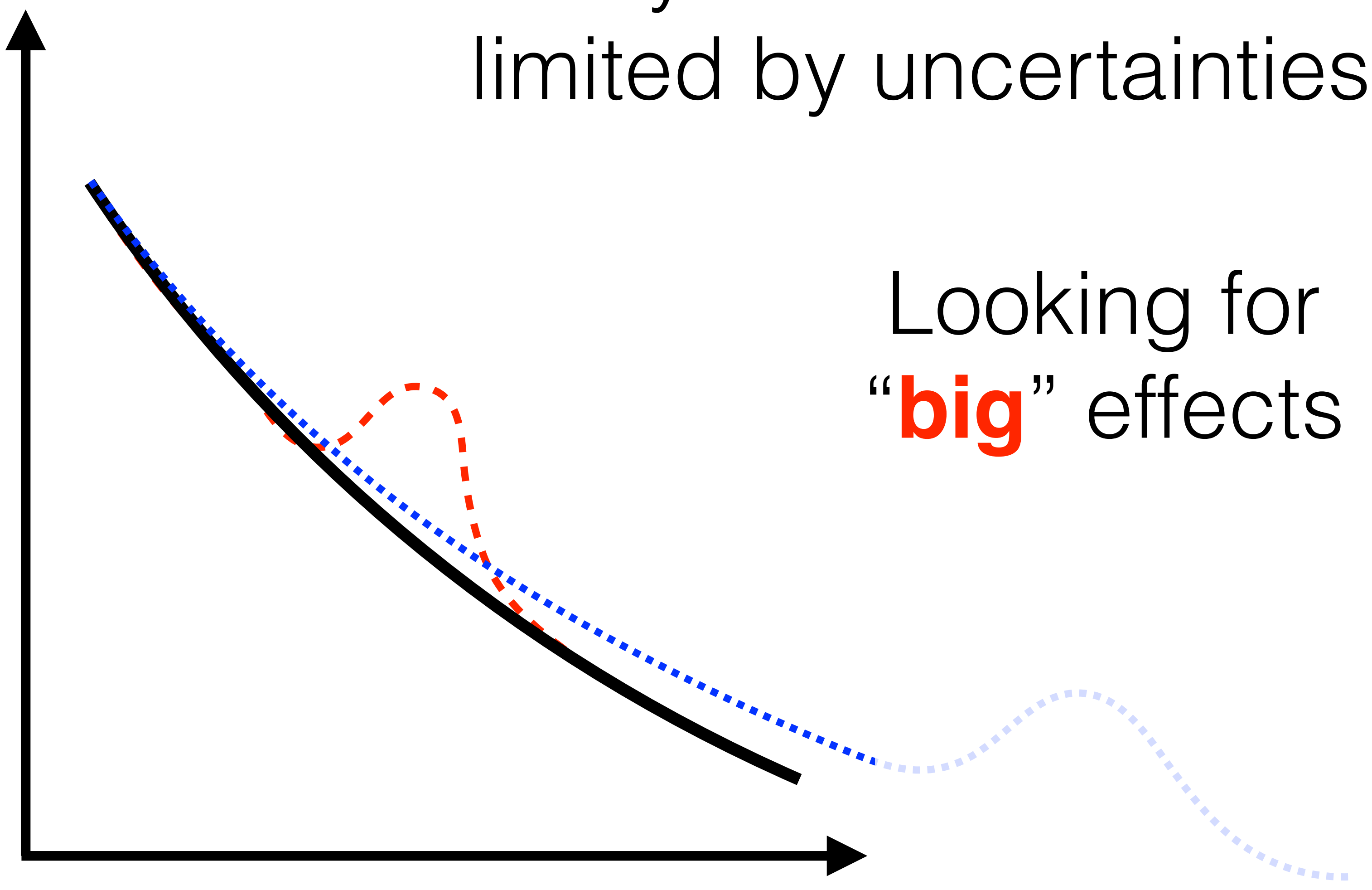


Looking for "**big**" effects

Indirect searches with precision measurements instead look for "**small**" deviations

Searches

Many searches are not limited by uncertainties

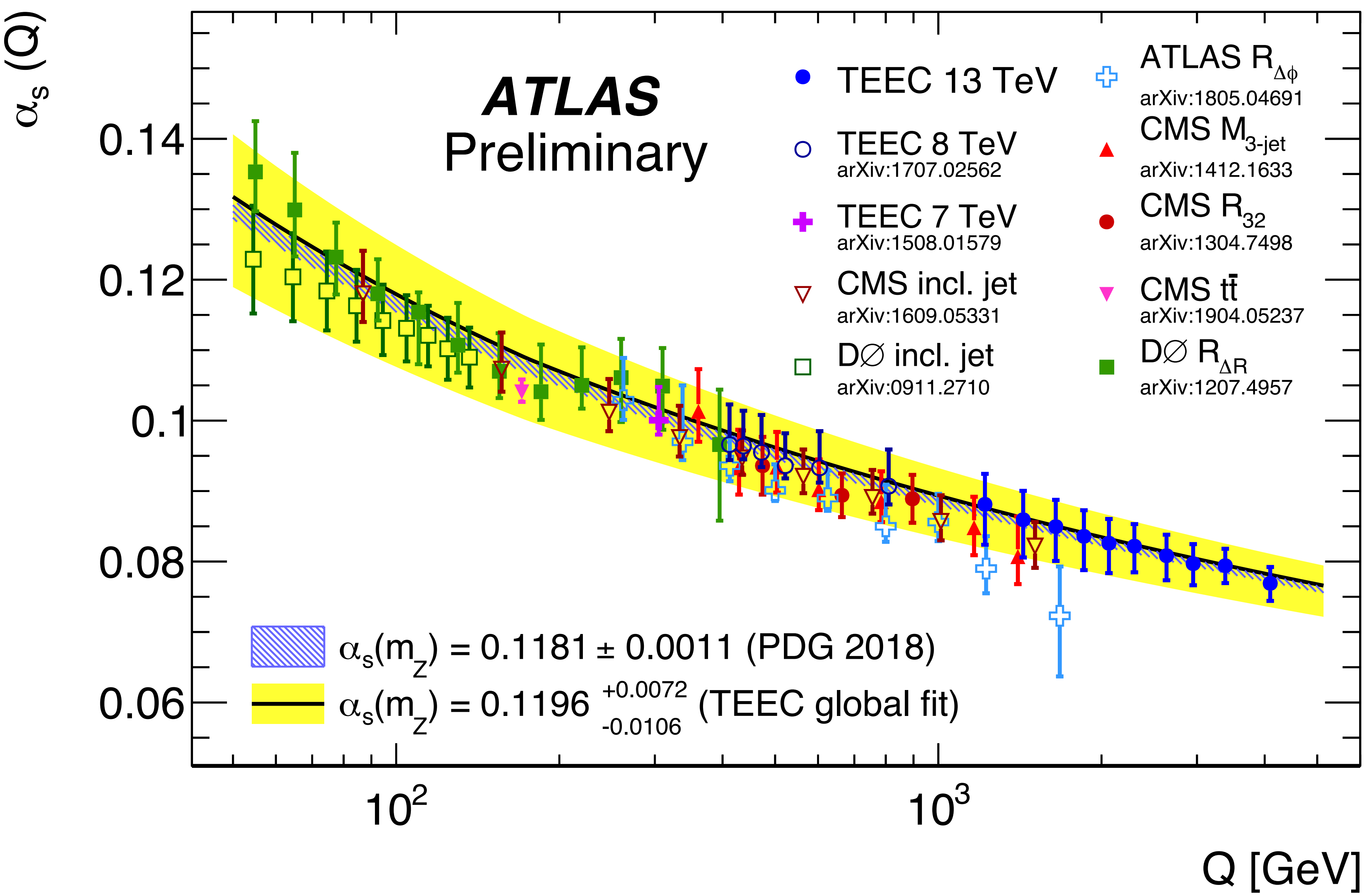


Looking for "**big**" effects

Indirect searches with precision measurements instead look for "**small**" deviations

Often the result of new particles beyond the kinematic each

Example

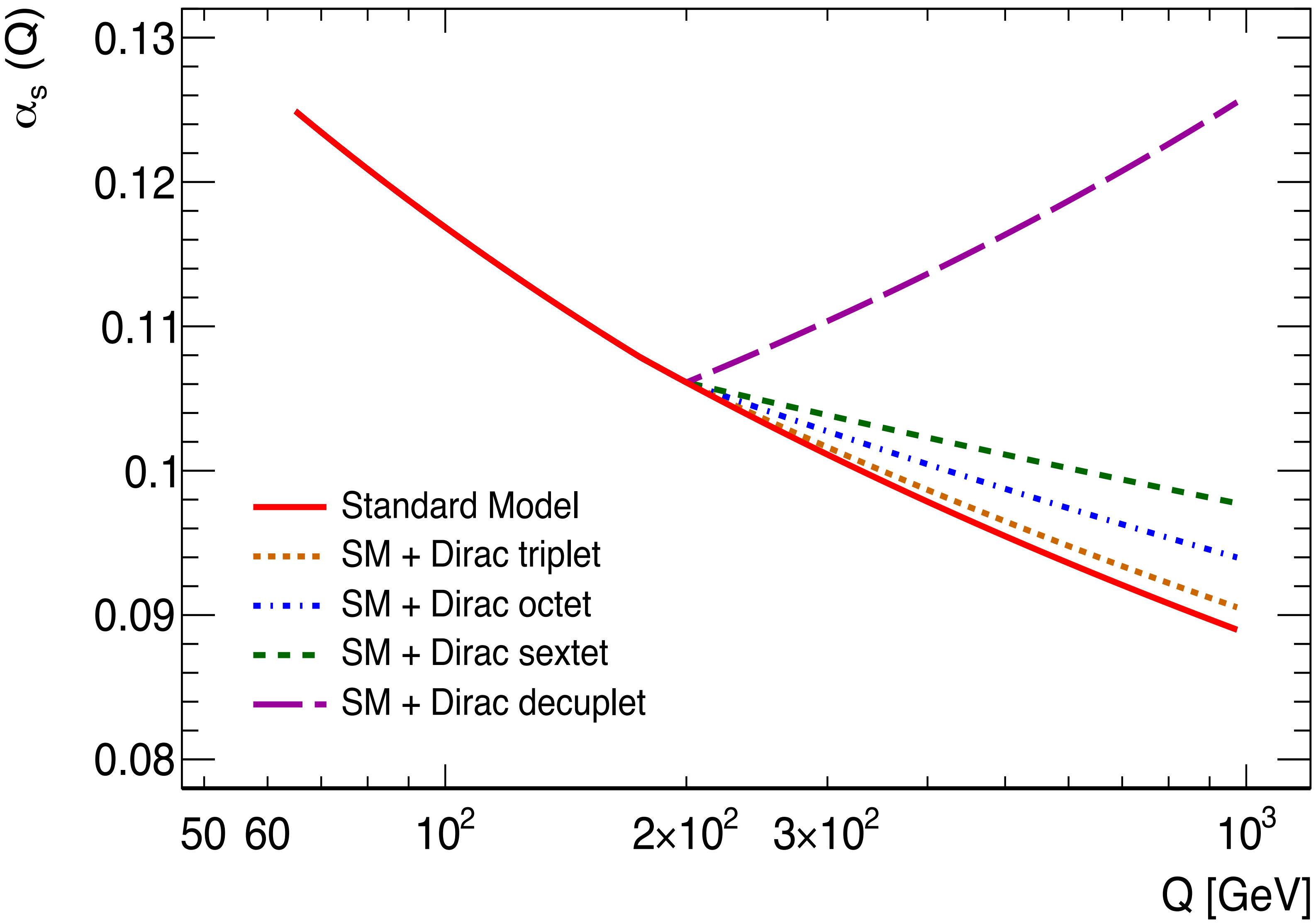


Running of the strong coupling constant

What if there are new colored particles?

Example

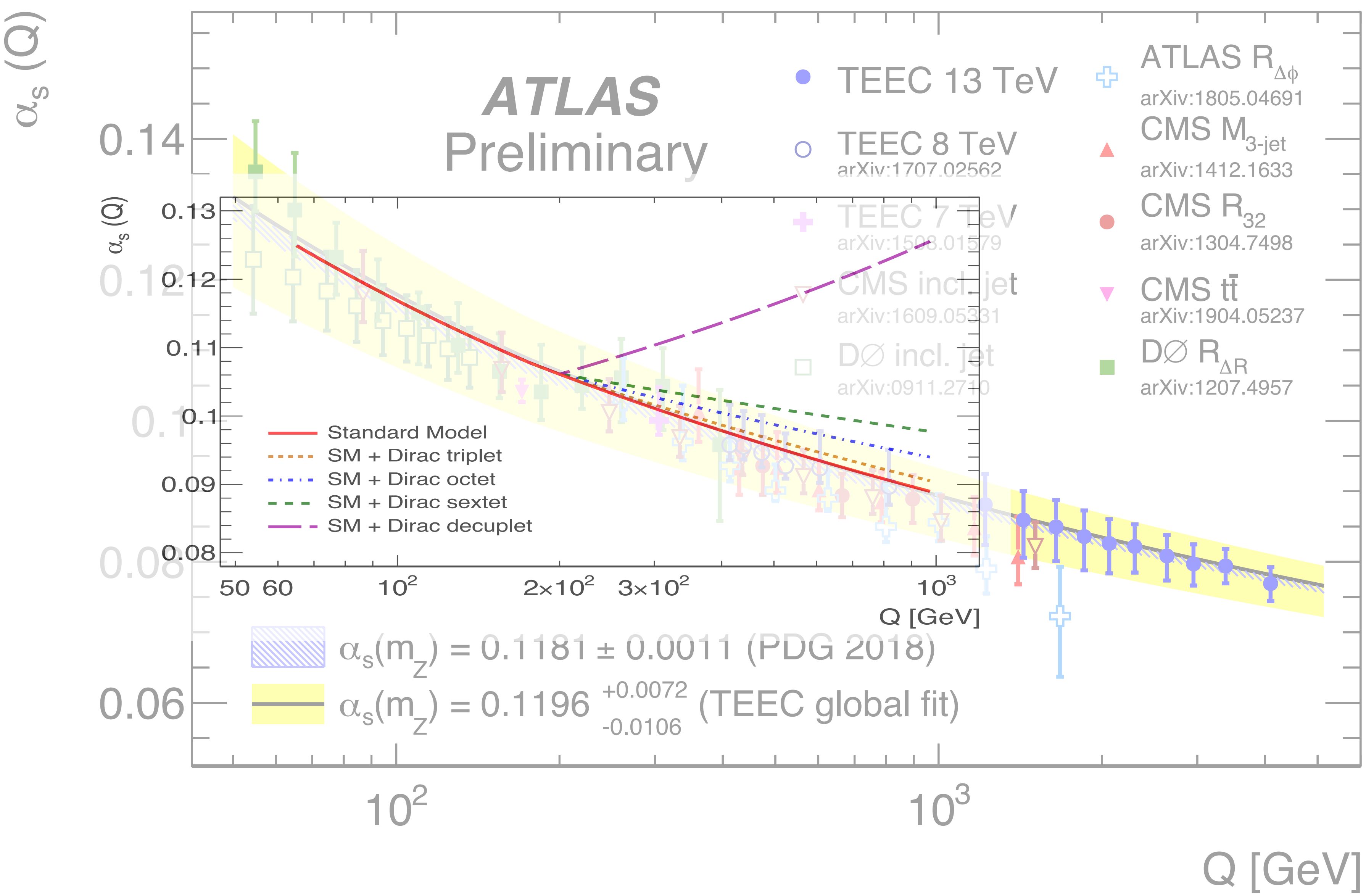
J. Llorente and BN, Nucl. Phys. B 936 (2018) 106



Running of the strong coupling constant

What if there are new colored particles?

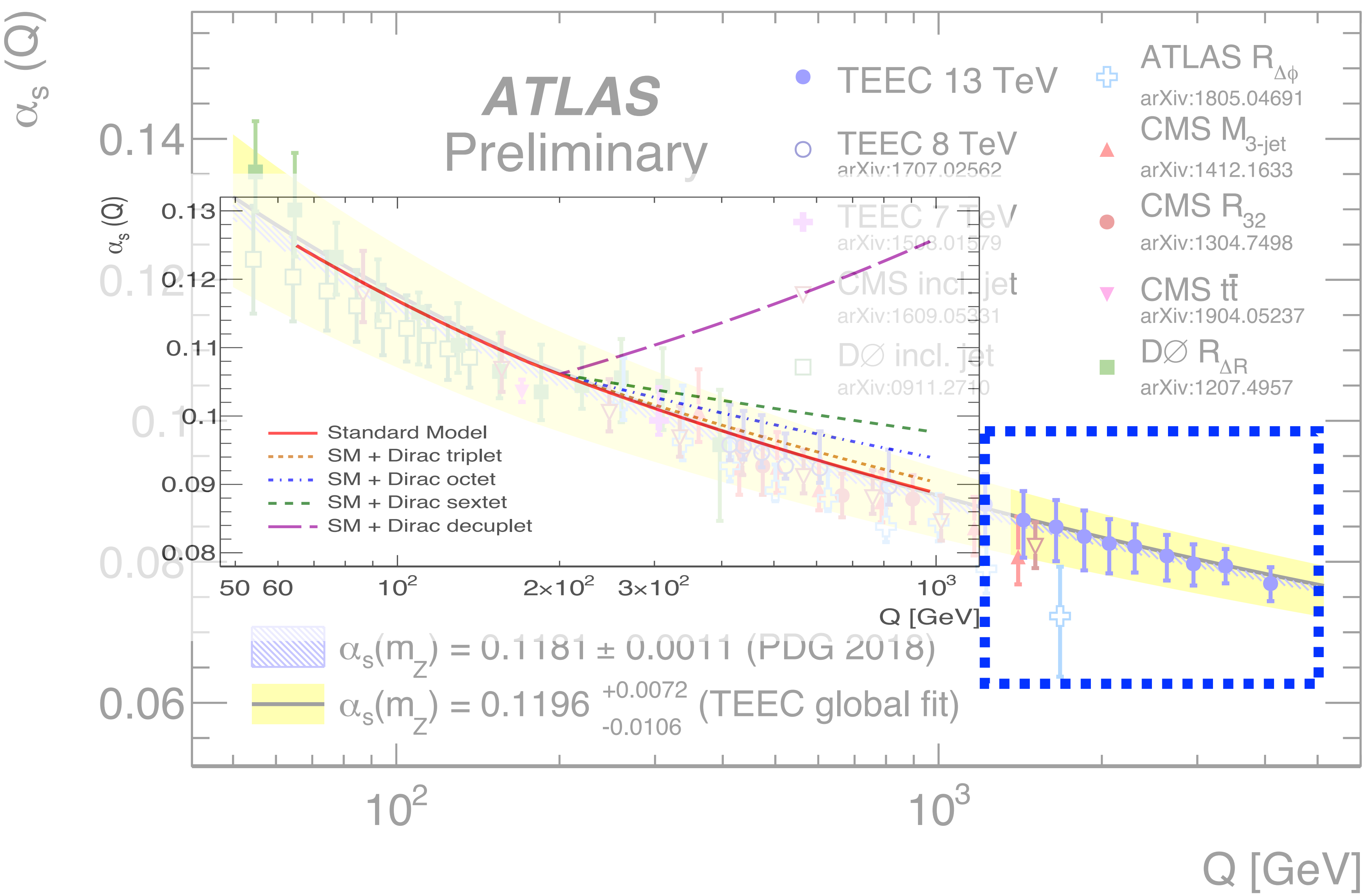
Example



Running of the strong coupling constant

What if there are new colored particles?

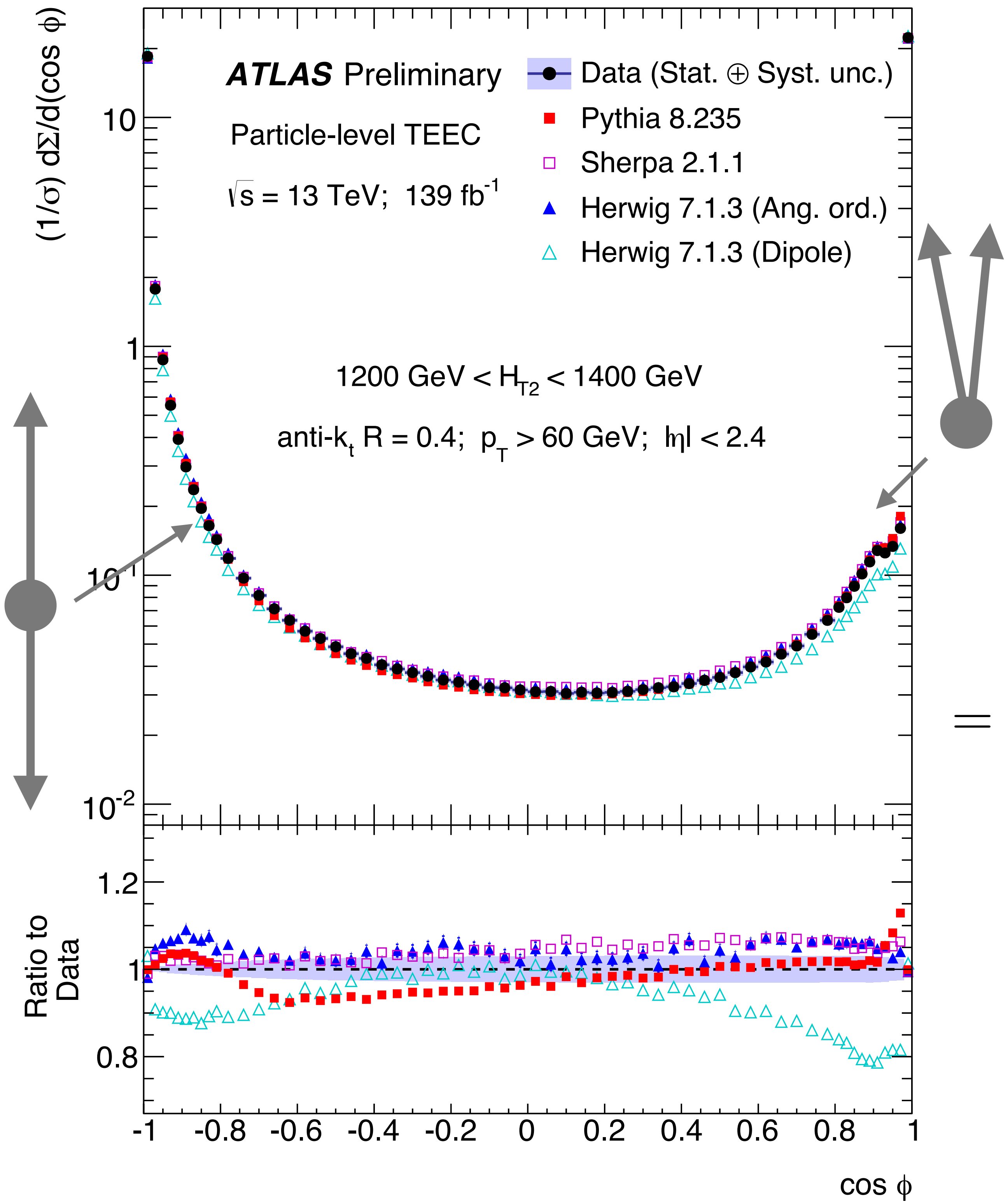
Example



Running of the strong coupling constant

What if there are new colored particles?

Strong coupling at the highest scales



Powerful probe: event shapes

Example: energy-energy correlation function (TEEC)

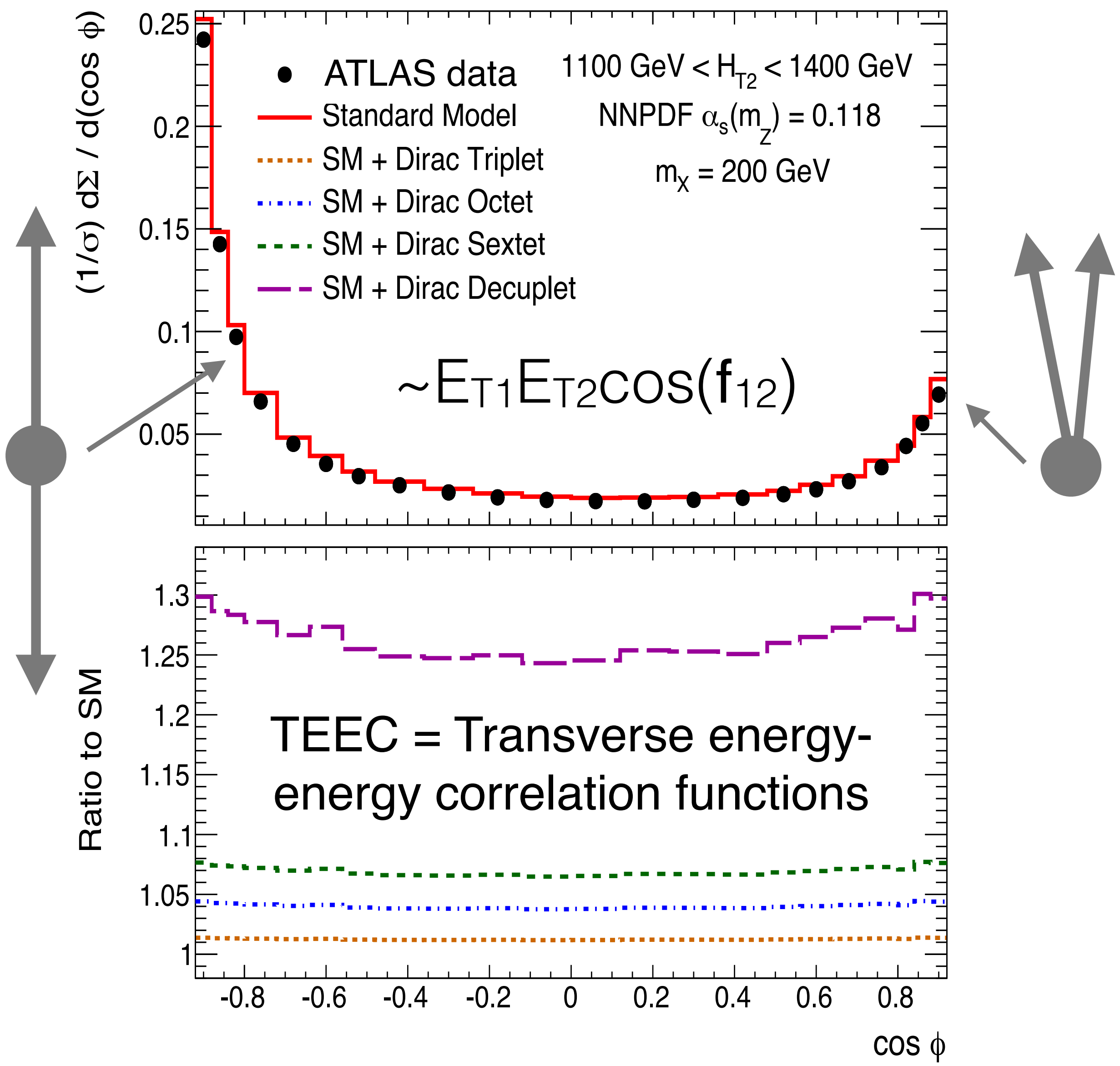
$$\frac{1}{\sigma}\frac{\mathrm{d}\Sigma}{\mathrm{d}\cos\phi}(\cos\phi)$$

$$= \sum_{\mathrm{events}} \sum_{i,j=1}^{n_{\mathrm{jets}}} \frac{E_{T,i}E_{T,j}}{\left(\sum_{k=1}^{n_{\mathrm{jet}}} E_{T,k}\right)^2}\delta(\cos\phi - \cos\phi_{ij})$$

Measure in bins of ~p_T and compare to theory predictions. Uncertainties are %-level.

(2) BSM

J. Llorente and BN, Nucl. Phys. B 936 (2018) 106



Jets provide the largest lever arm to study deviations from the SM QCD running of the strong coupling.

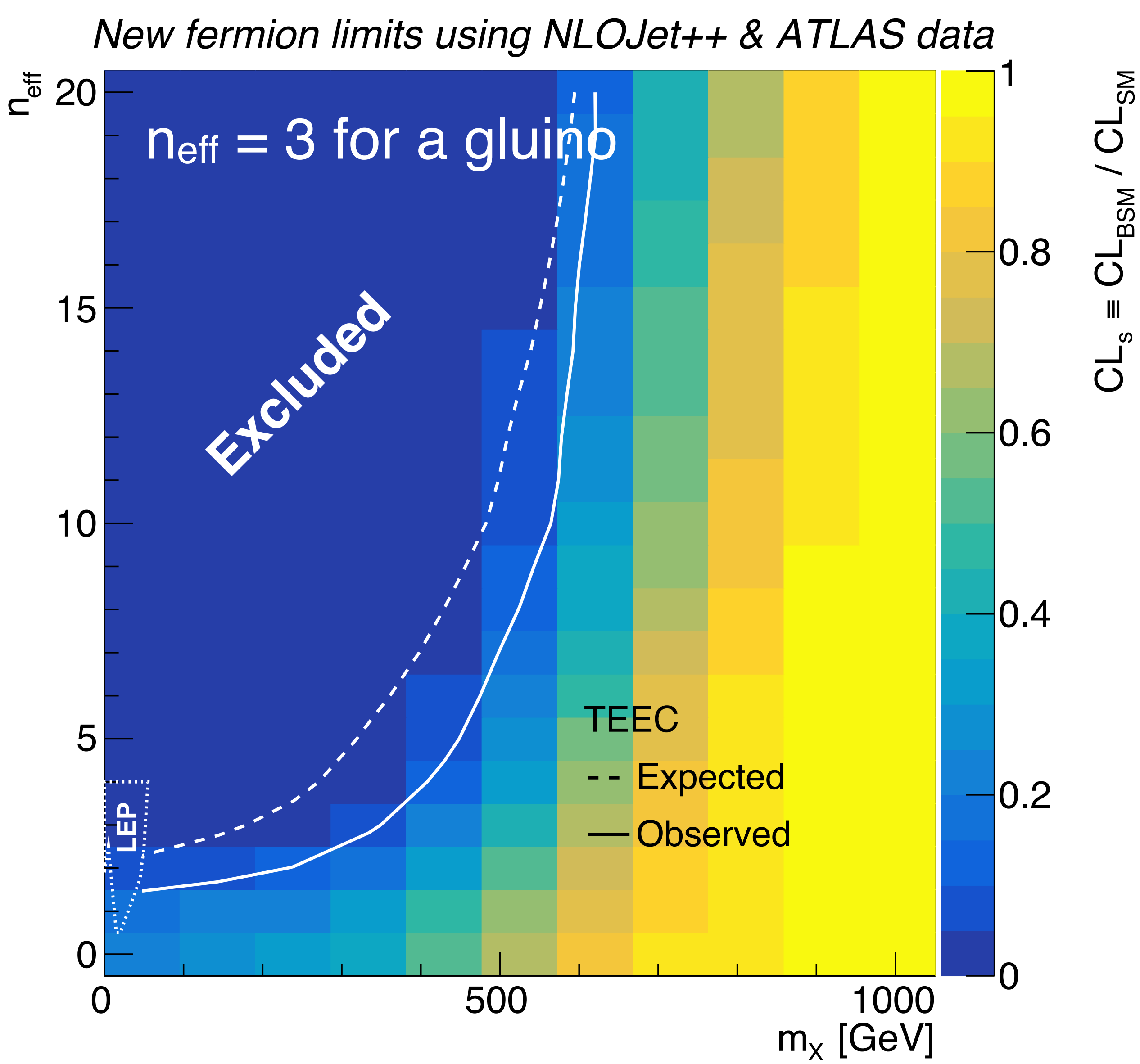
Such an approach is complementary to direct searches as this is ~agnostic to the decay.

$$\frac{1}{\sigma}\frac{d\Sigma}{d\cos\phi}(\cos\phi) = \sum_{\text{events}} \sum_{i,j=1}^{n_{\text{jets}}} \frac{E_{T,i}E_{T,j}}{\left(\sum_{k=1}^{n_{\text{jet}}} E_{T,k}\right)^2}\delta(\cos\phi - \cos\phi_{ij})$$

(defined as dimensionless cross-section)

Jets as a precision probe for BSM

Using the TEEC, we have set stringent exclusion limits on colored BSM (except when n_{eff} ~ 1 (e.g. single squark)



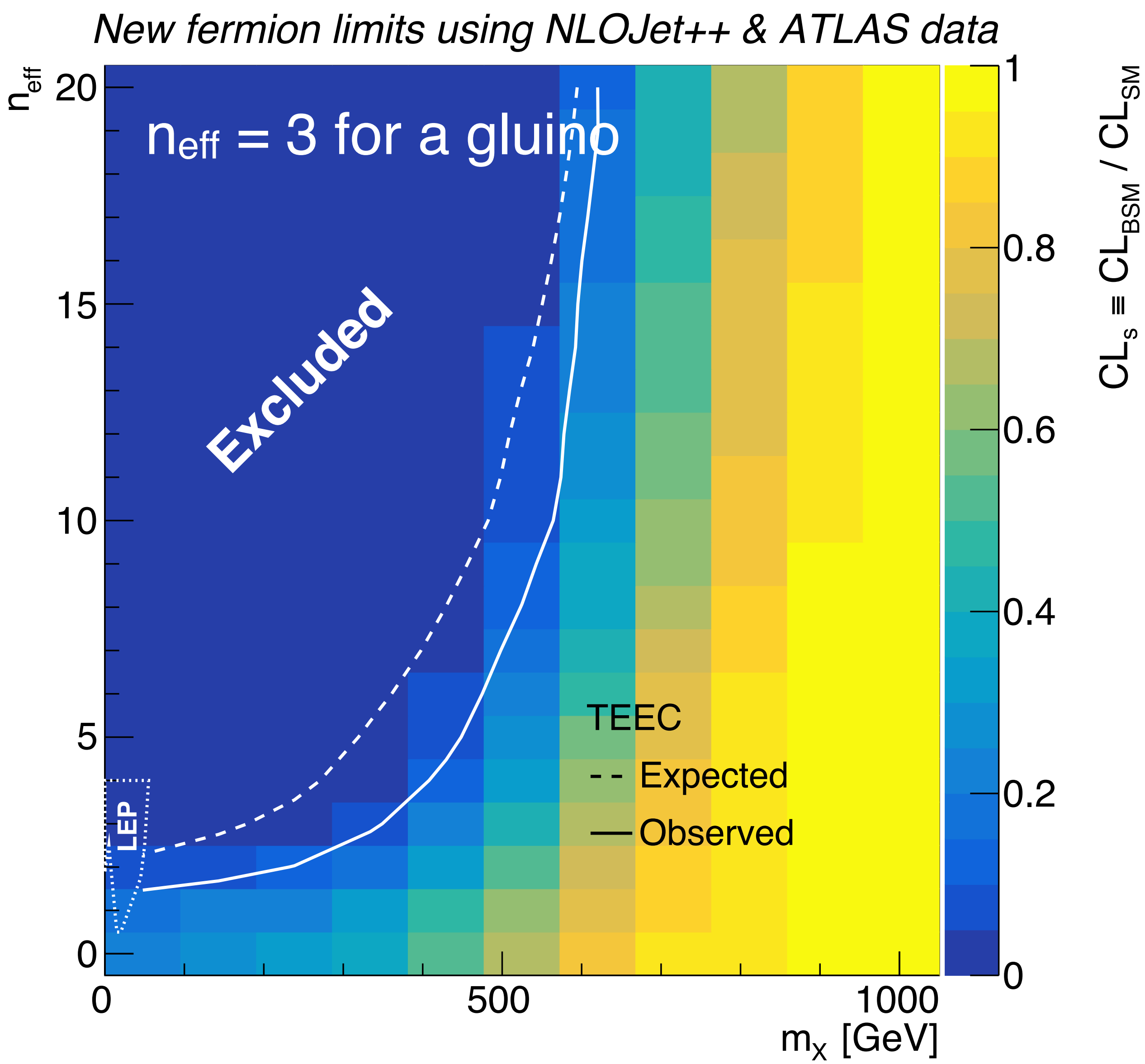
J. Llorente and BN, Nucl. Phys. B 936 (2018) 106

Jets as a precision probe for BSM

Using the TEEC, we have set stringent exclusion limits on colored BSM (except when n_{eff} ~ 1 (e.g. single squark)

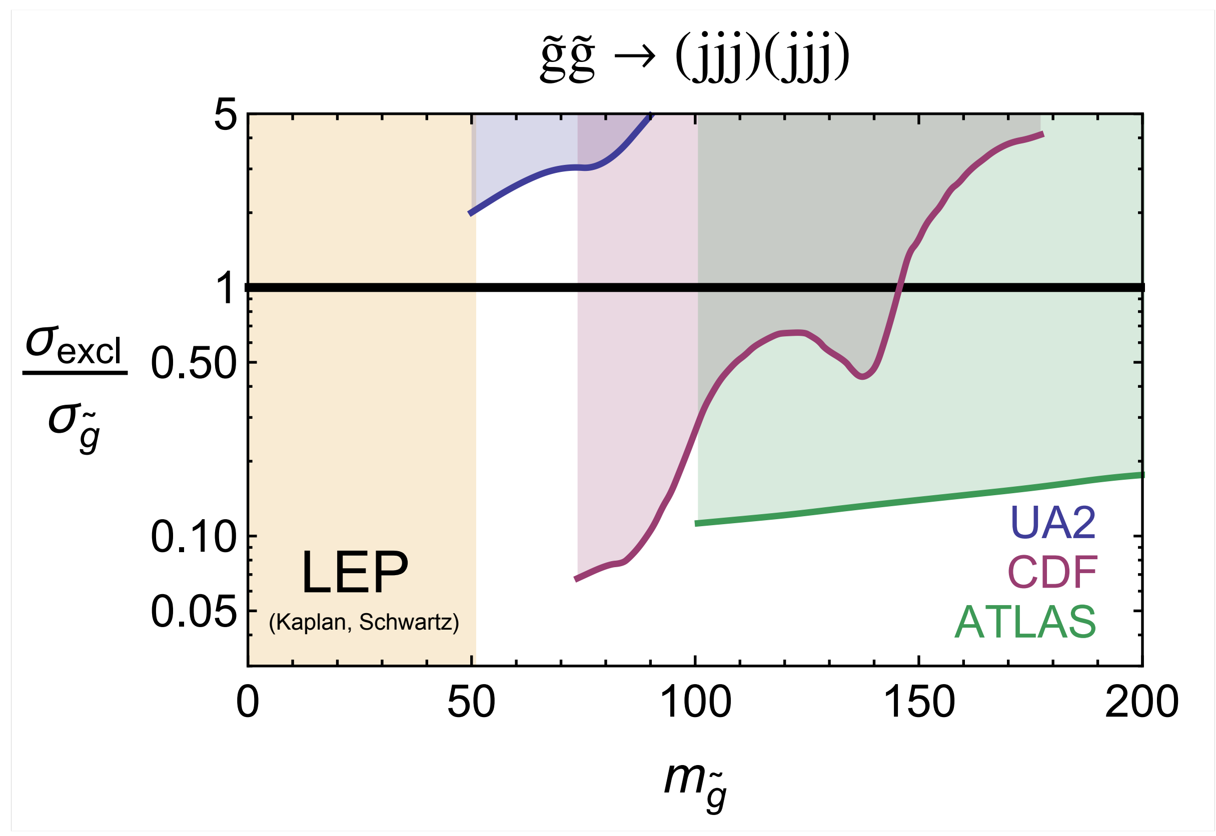
You may ask:

But I thought limits on gluinos were >(>) 1 TeV??

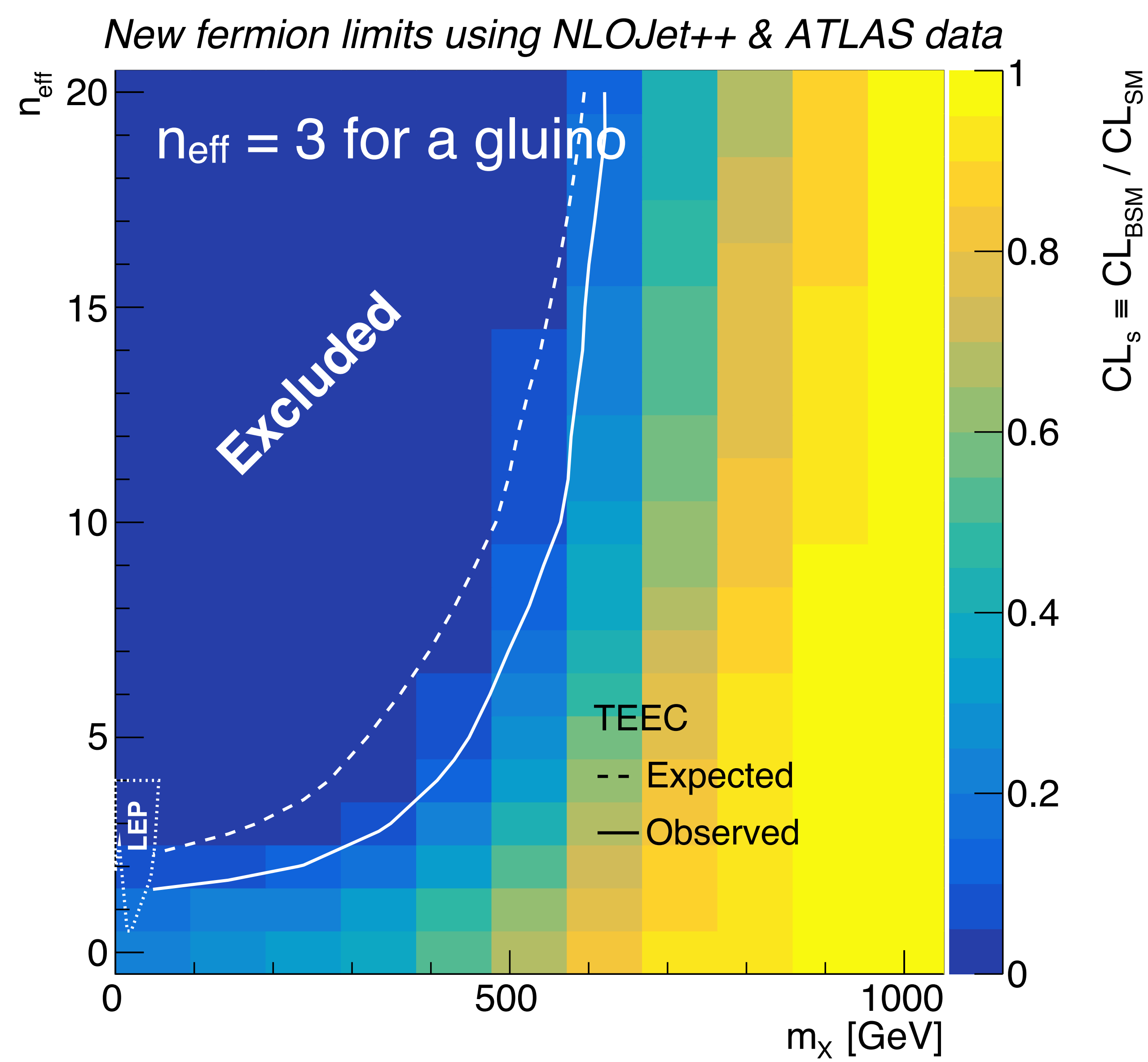


J. Llorente and BN, Nucl. Phys. B 936 (2018) 106

Jets as a precision probe for BSM



J. Evans and D. McKeen, 1803.01880



J. Llorente and BN, Nucl. Phys. B 936 (2018) 106

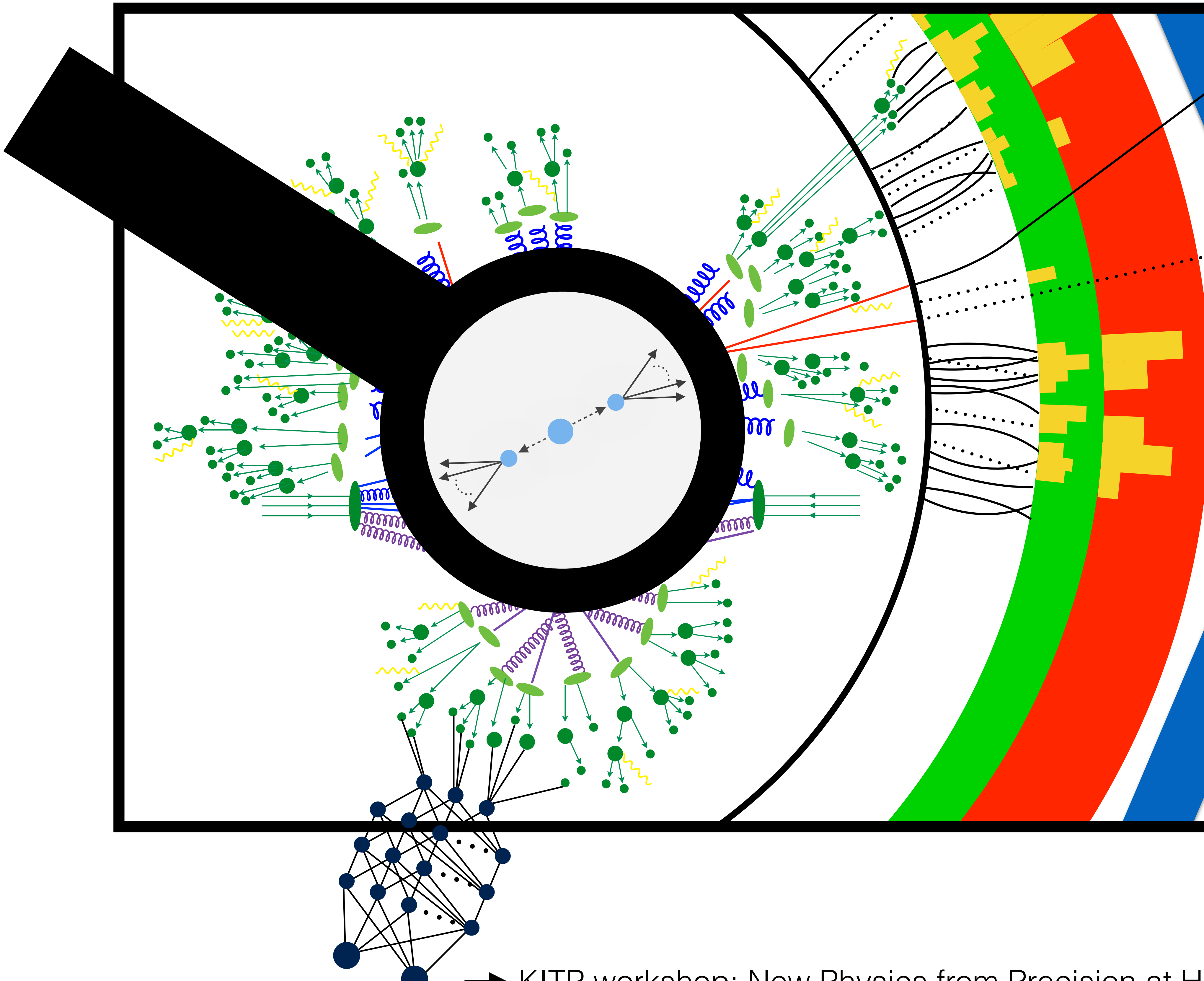
There may be gaps!

Indirect searches tend to have less/ different assumptions than direct searches and are thus essential.

(1) Precision measurements at high energy

(2) New physics from precision measurements

(3) Machine learning + measurements + BSM



→ KITP workshop: New Physics from Precision at High Energies

Data analysis in HEP

Theory of everything

↓

Physics simulators

↓

Detector-level observables

↓

Pattern recognition

↔

Nature

↓

Experiment

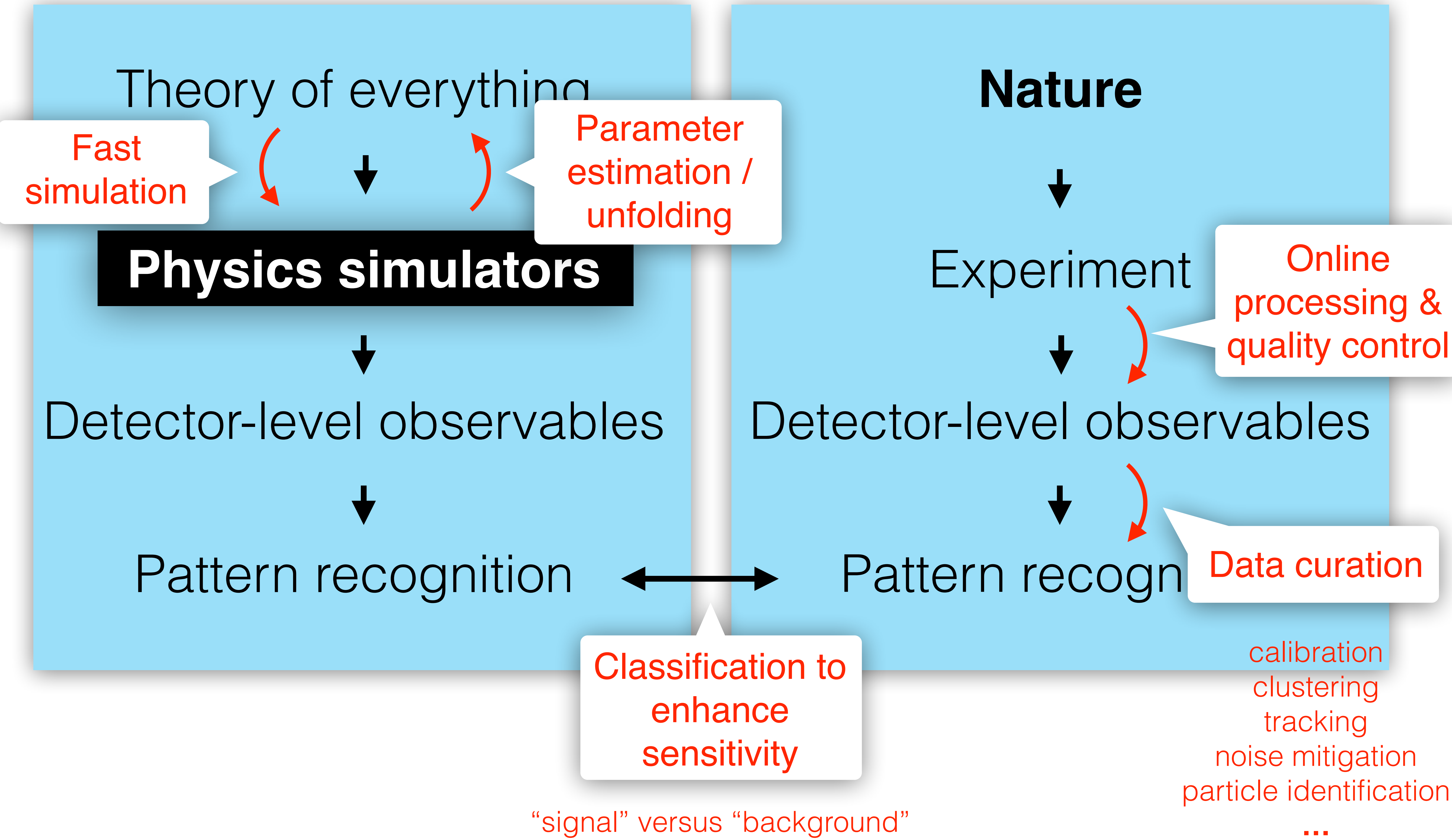
↓

Detector-level observables

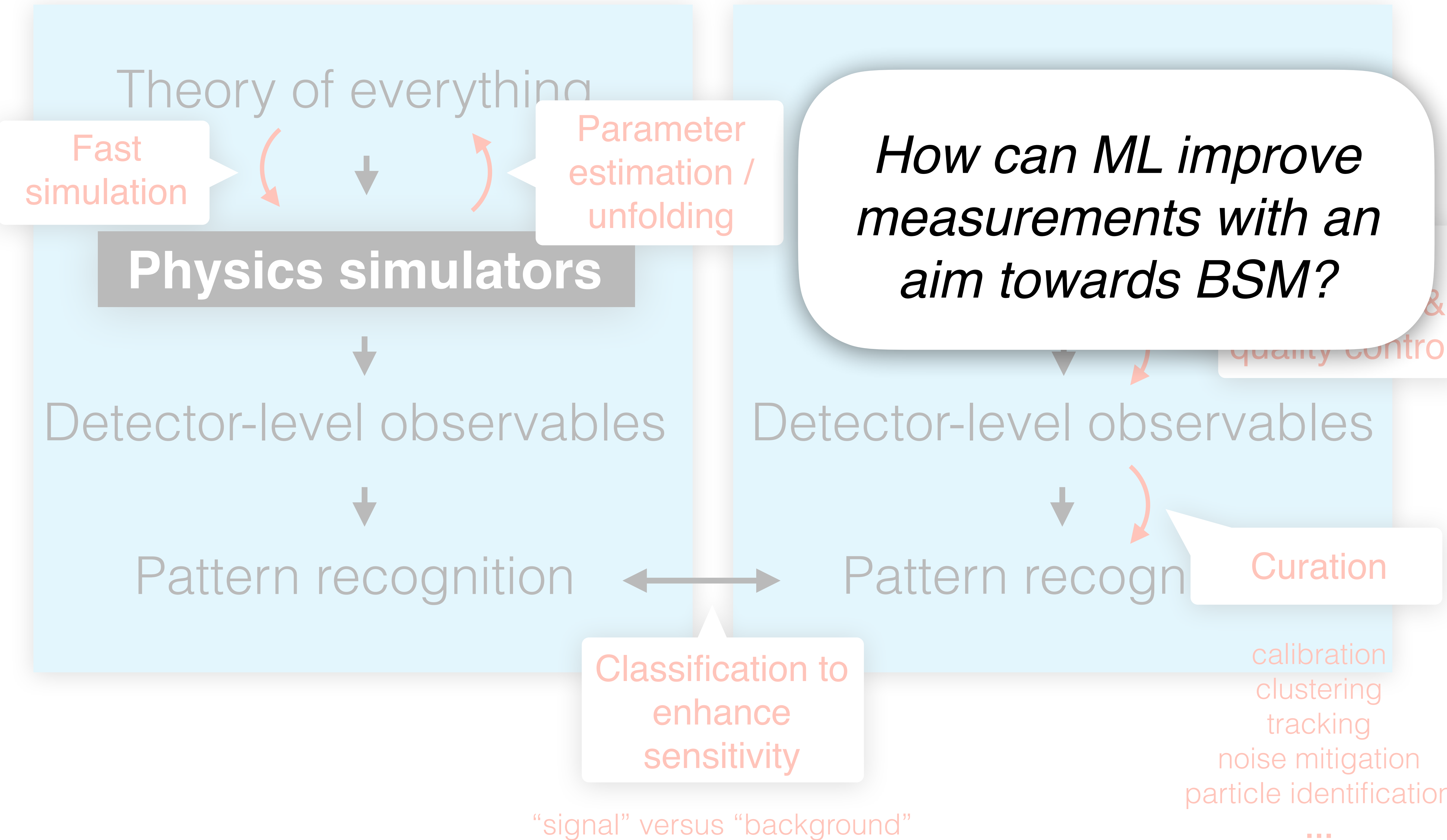
↓

Pattern recognition

Data analysis in HEP + Deep Learning

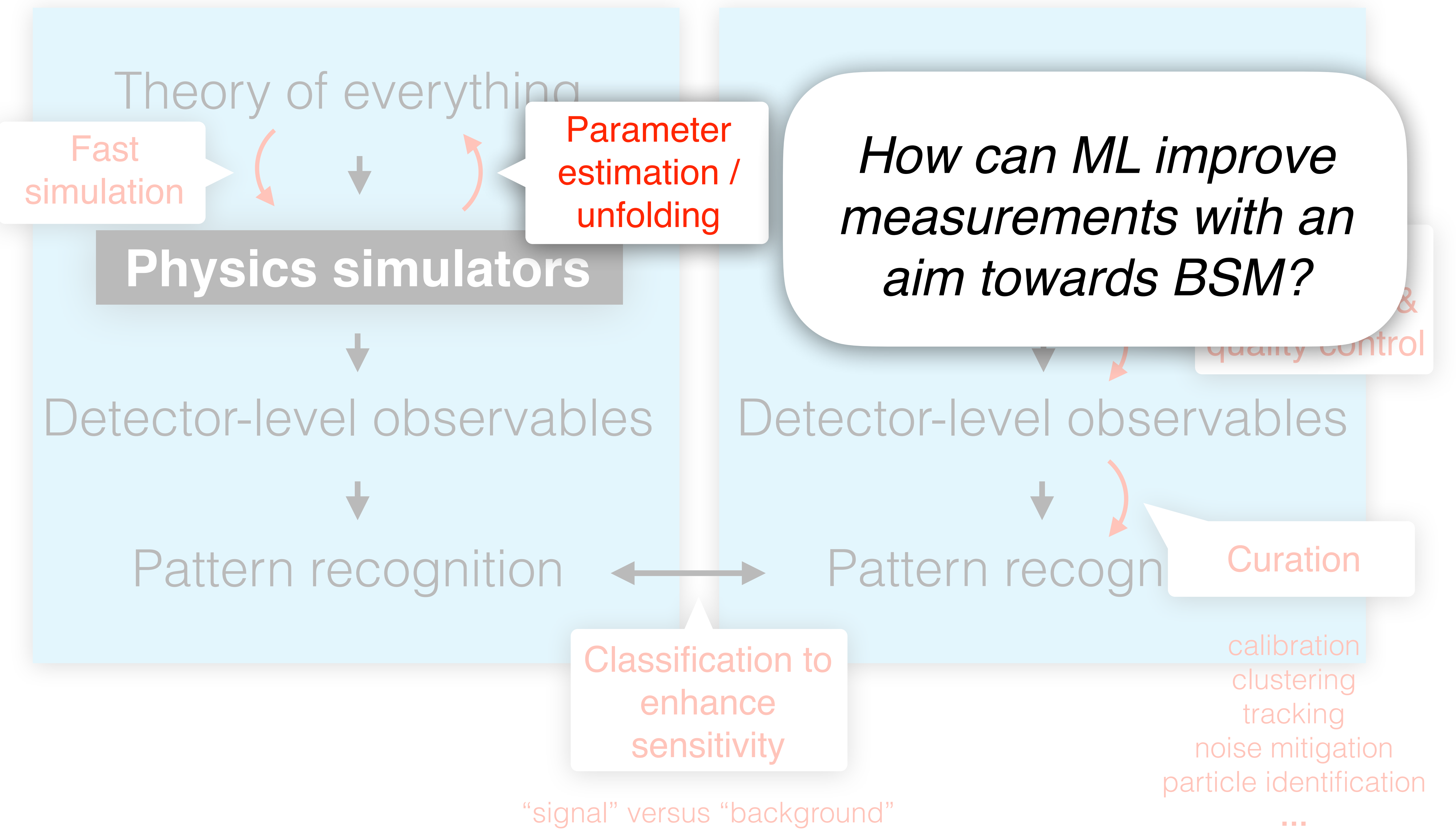


Data analysis in HEP + Deep Learning



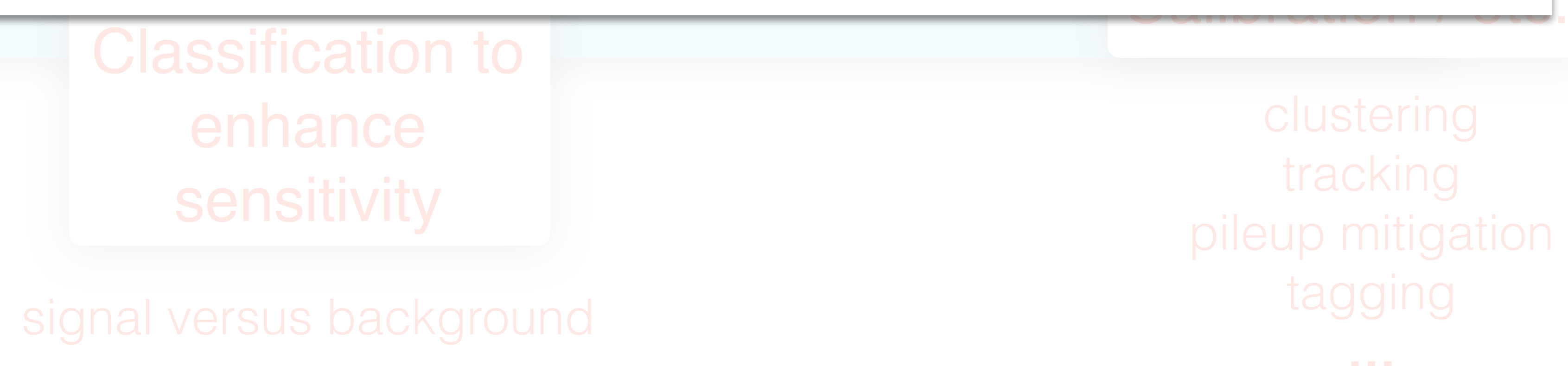
How can ML improve measurements with an aim towards BSM?

Data analysis in HEP + Deep Learning



Data analysis in HEP + Deep Learning

In low-dimensions, can try to explicitly model p(x|params).
To make the most of our data, want many dimensions.
In high-dimensions, p is intractable. Need likelihood-free!



Data analysis in HEP + Deep Learning

Theory of everything

Fast simulation

Parameter estimation / unfolding

Physics simulators

Nature

Experiment

In low-dimensions, can try to explicitly model *p*(x|params).
To make the most of our data, want many dimensions.
In high-dimensions, *p* is intractable. Need likelihood-free!

Classification to enhance sensitivity

signal versus background

clustering
tracking
pileup mitigation
tagging
...

Data analysis in HEP + Deep Learning

In low-dimensions, can try to explicitly model *p*(x|params).
To make the most of our data, want many dimensions.
In high-dimensions, *p* is intractable. Need likelihood-free!

Classification to enhance sensitivity

signal versus background

clustering
tracking
pileup mitigation
tagging
...

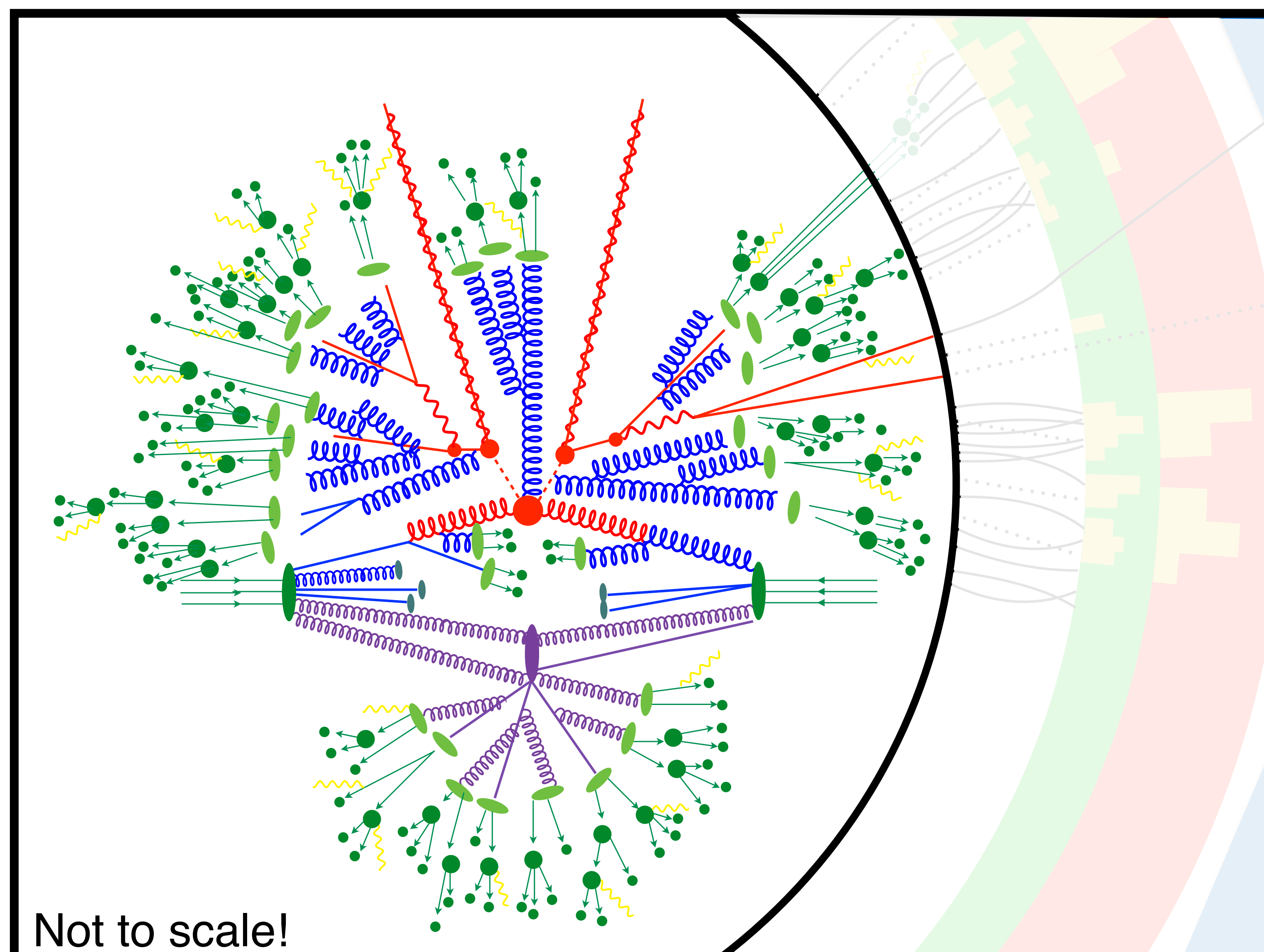
A *hyper* challenge

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



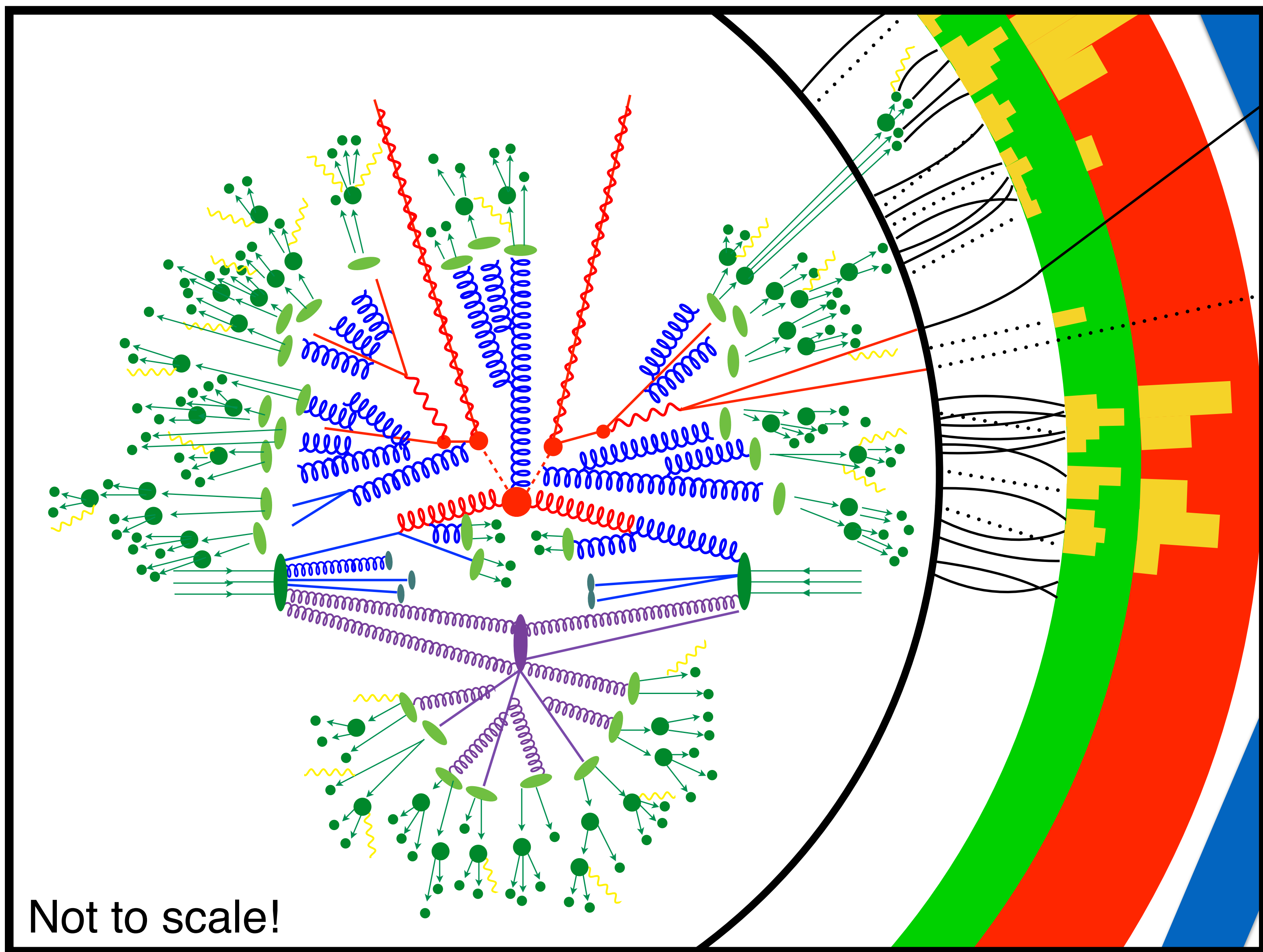
A *hyper* challenge

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

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

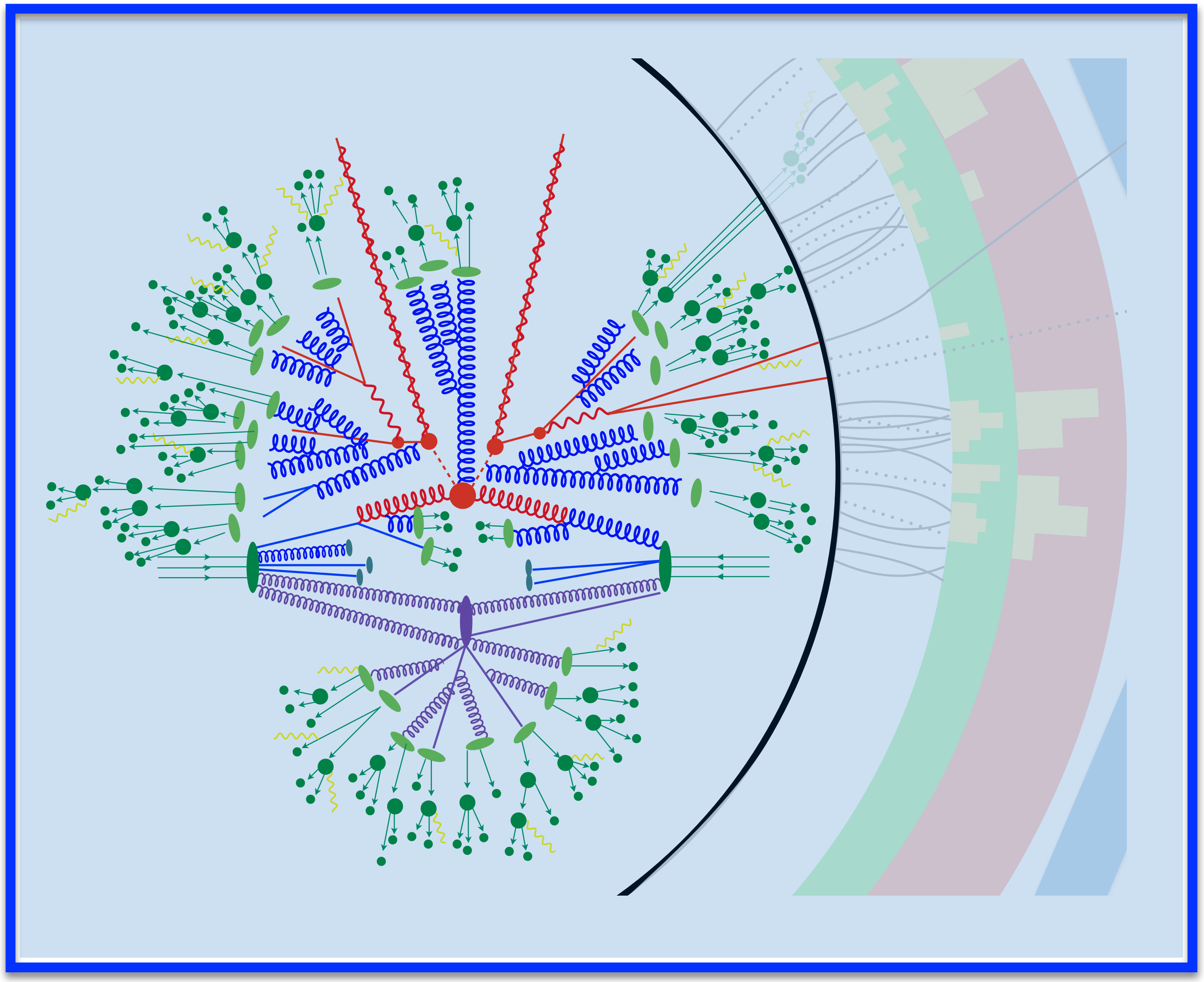
We detect these particles with **O(100 M)** readout channels

Image inspired by JHEP 02 (2009) 007

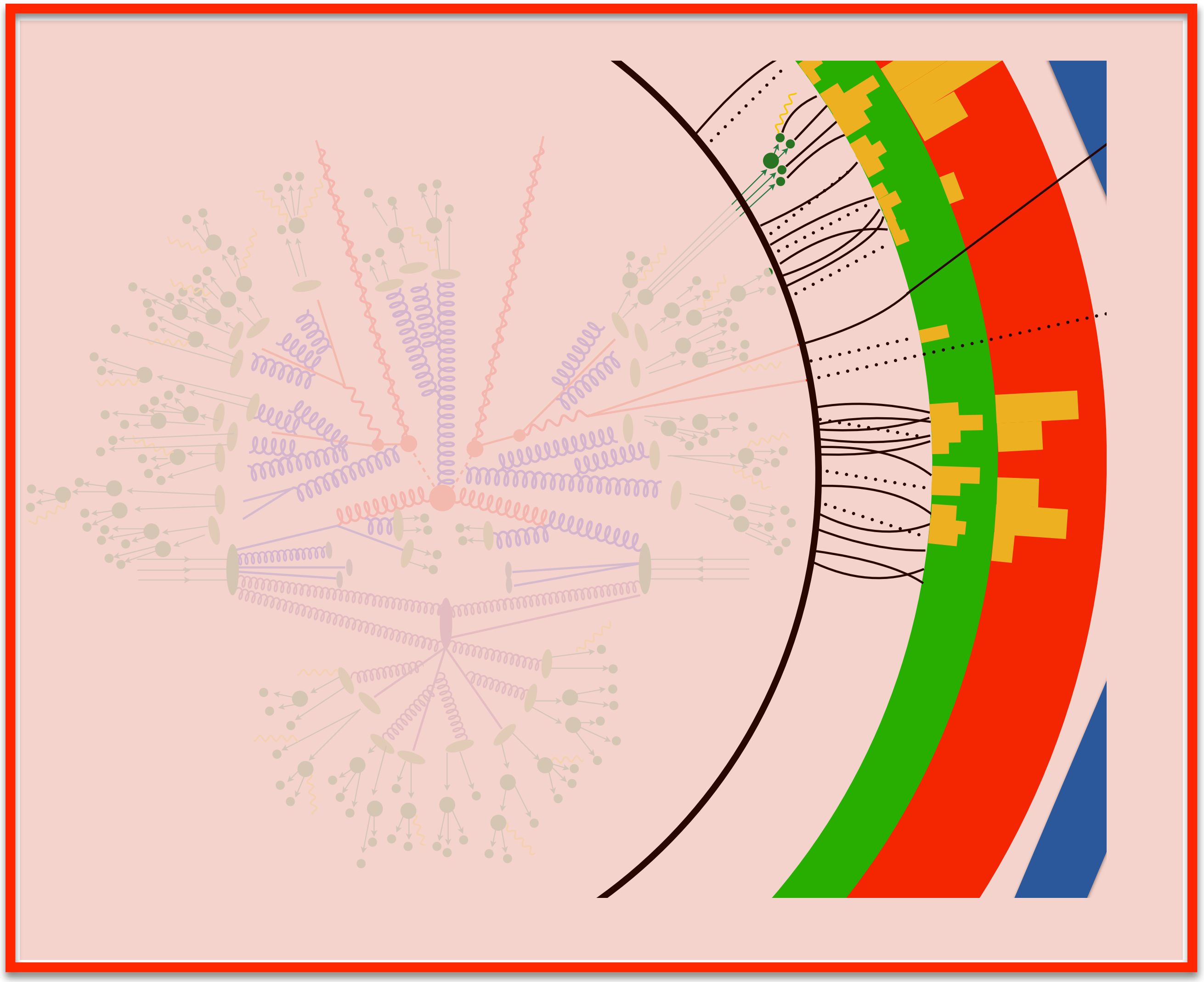


Example: Unfolding (Deconvolution)

Want this ↓



Measure this ↓

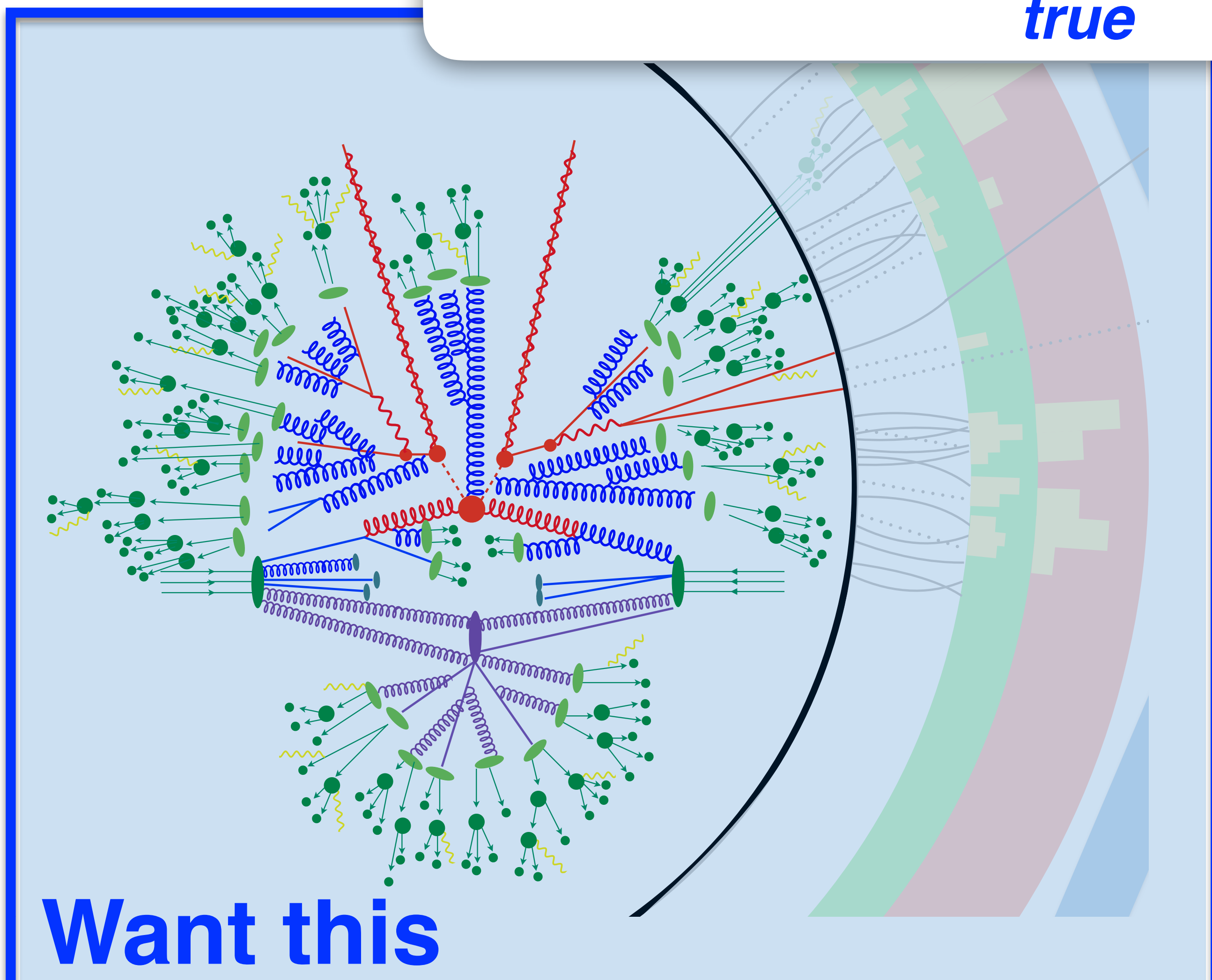


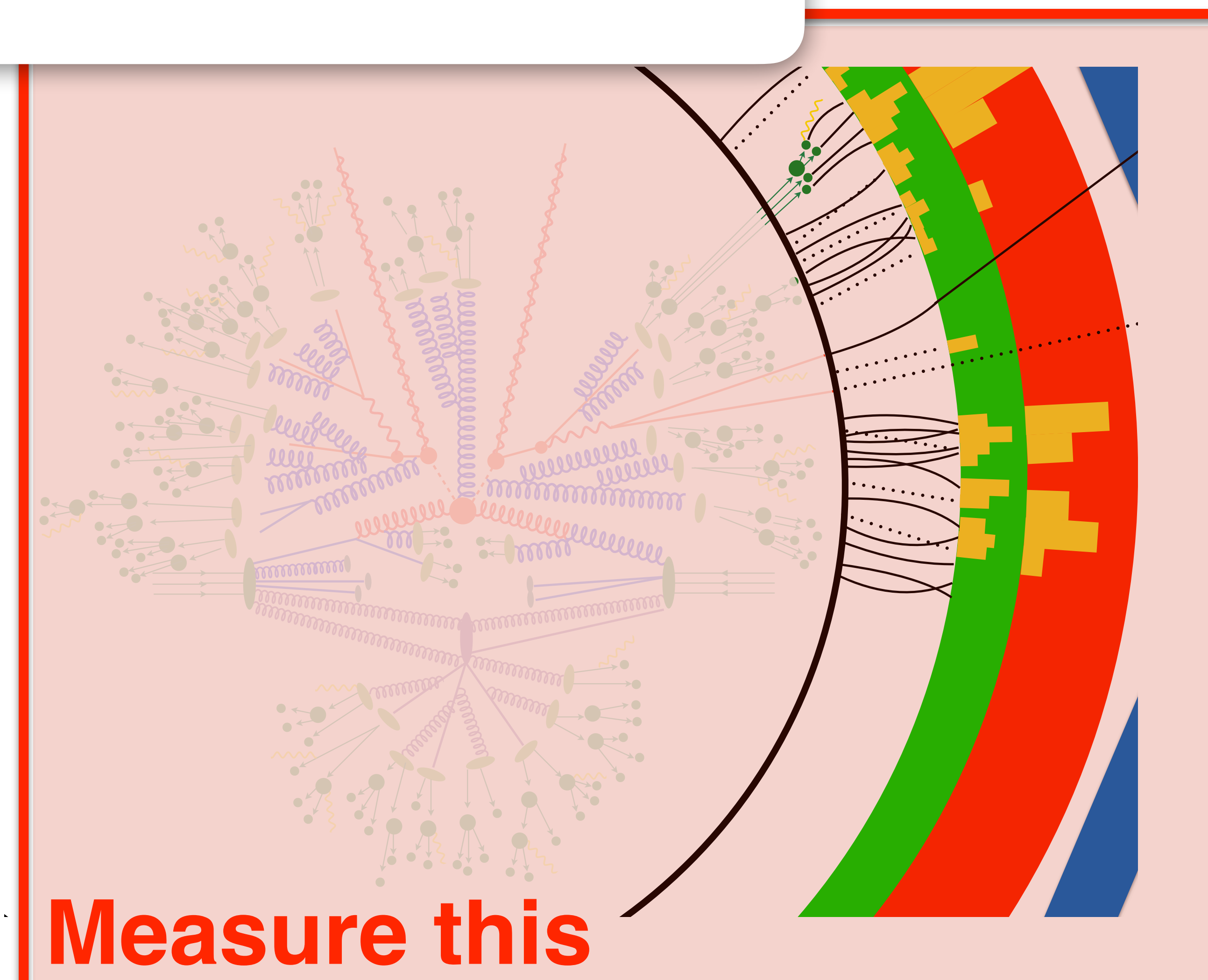
i.e. remove detector distortions

Example: Unfolding (Deconvolution)

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

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





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

Example: Unfolding (Deconvolution)

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

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

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

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

Example: 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

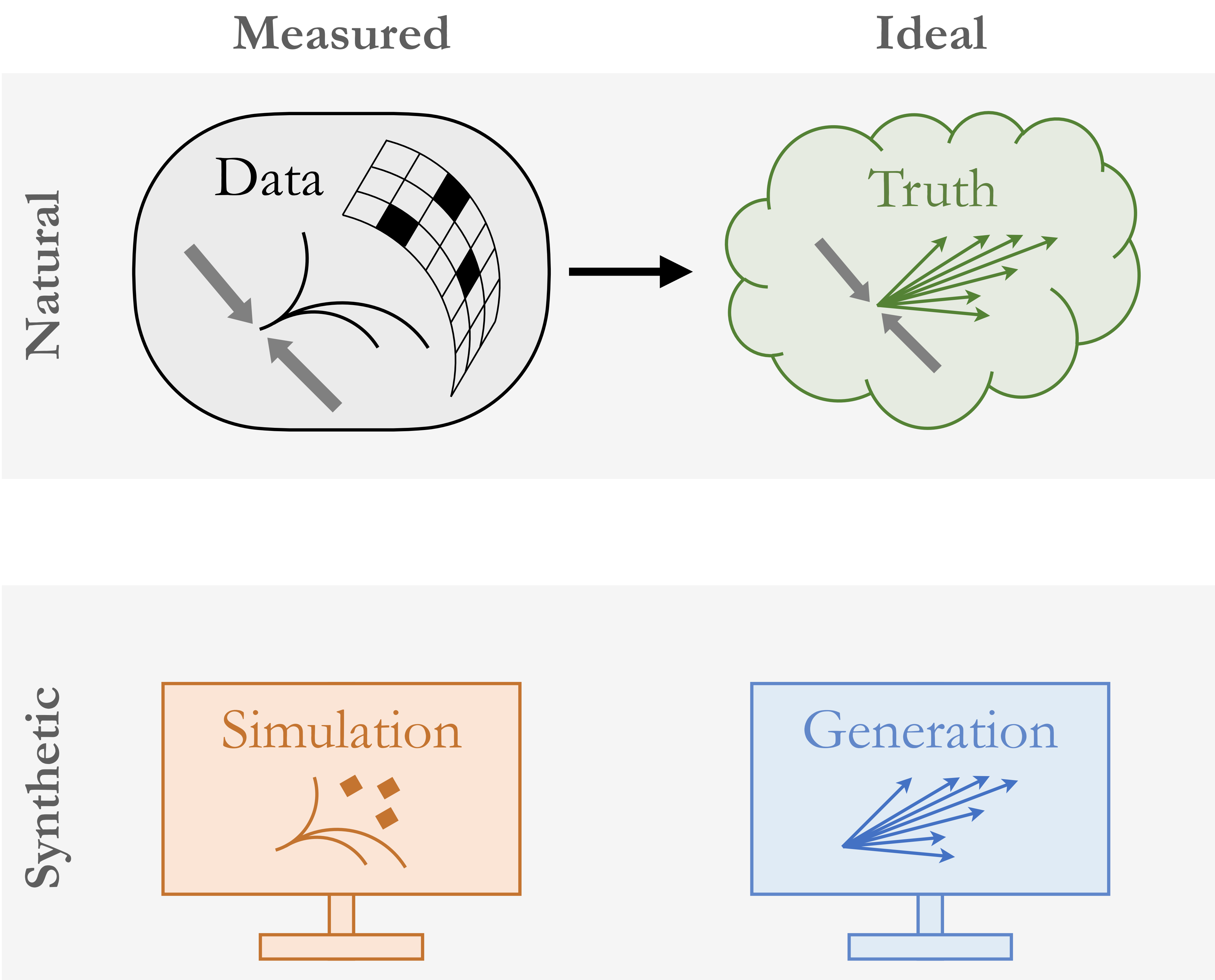
Fact: Neutral networks learn to approximate the likelihood ratio = $q(x)/p(x)$

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

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

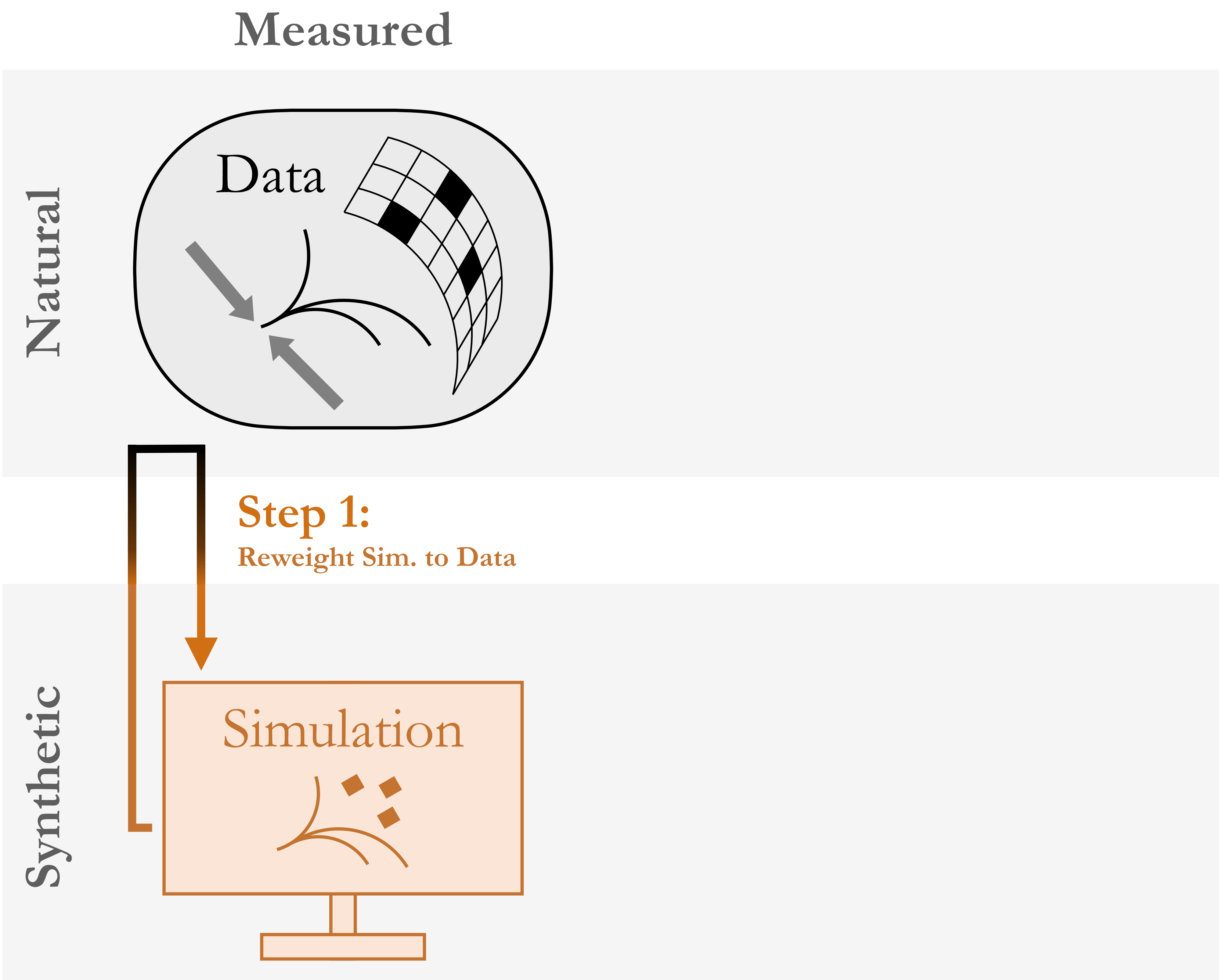
This turns the problem of **density estimation** (**hard**) into a problem of **classification** (**easy**)

Unfold by iterating: OmniFold



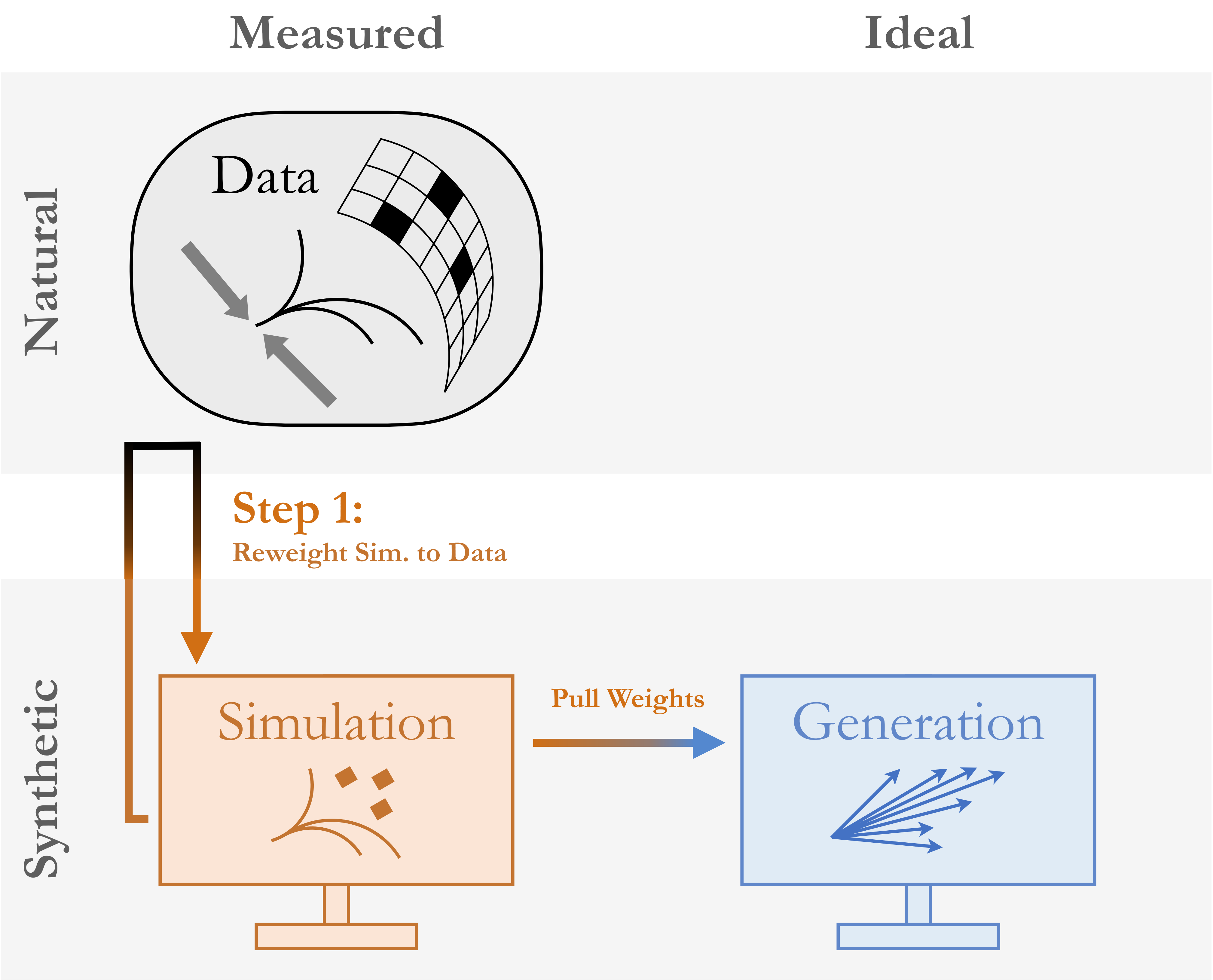
A. Andreassen, P. Komiske, E. Metodiev, BN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



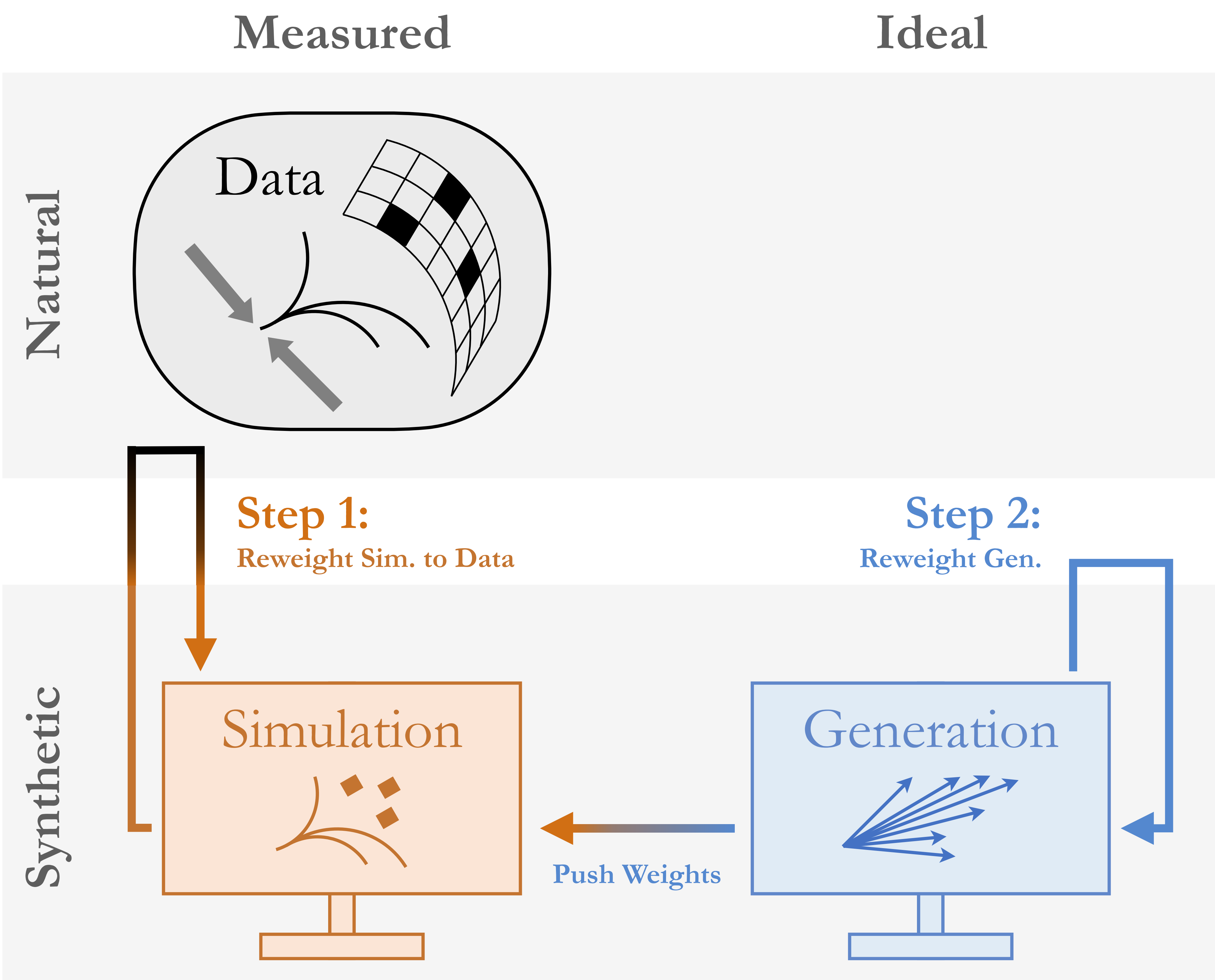
A. Andreassen, P. Komiske, E. Metodiev, BN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



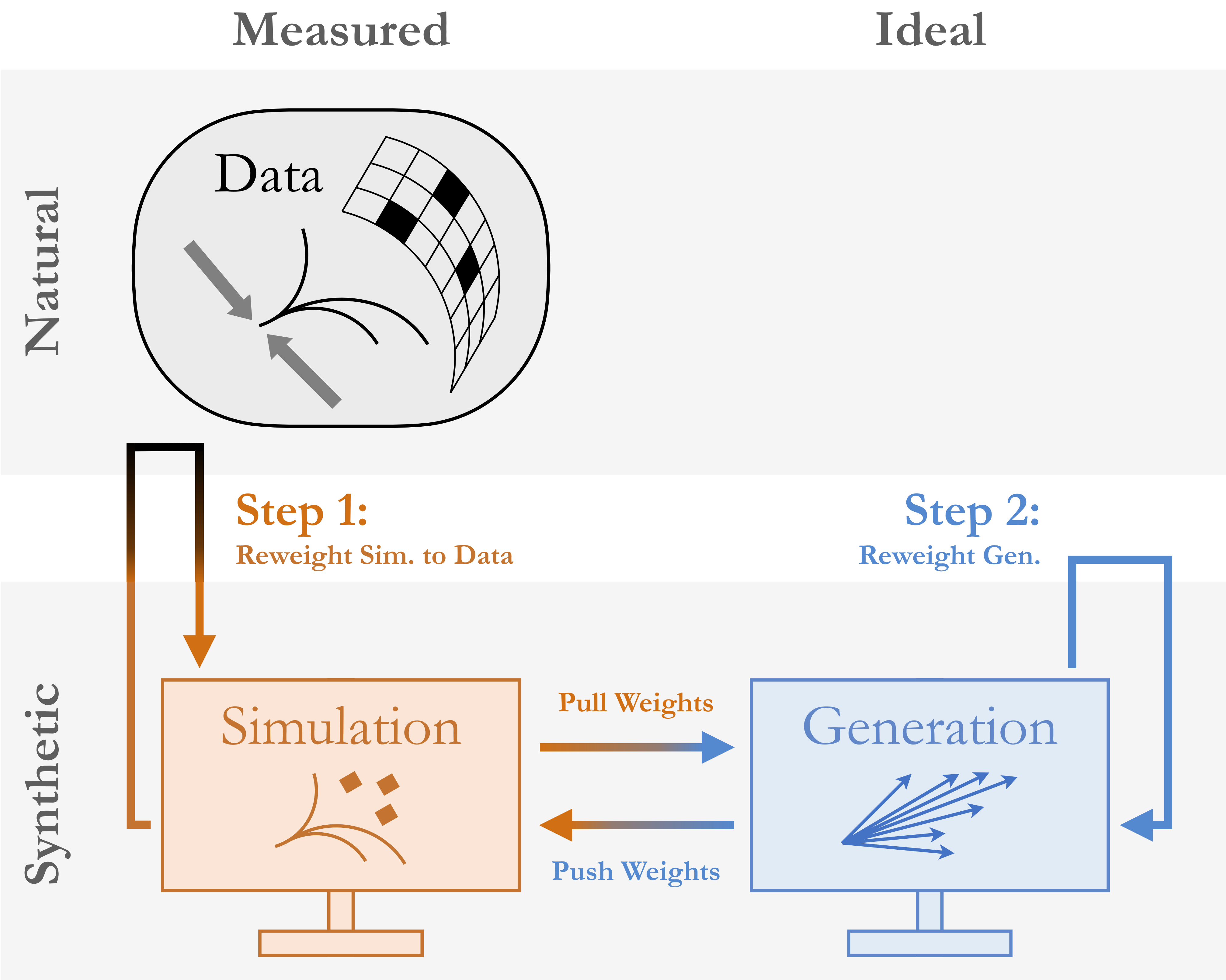
A. Andreassen, P. Komiske, E. Metodiev, BN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



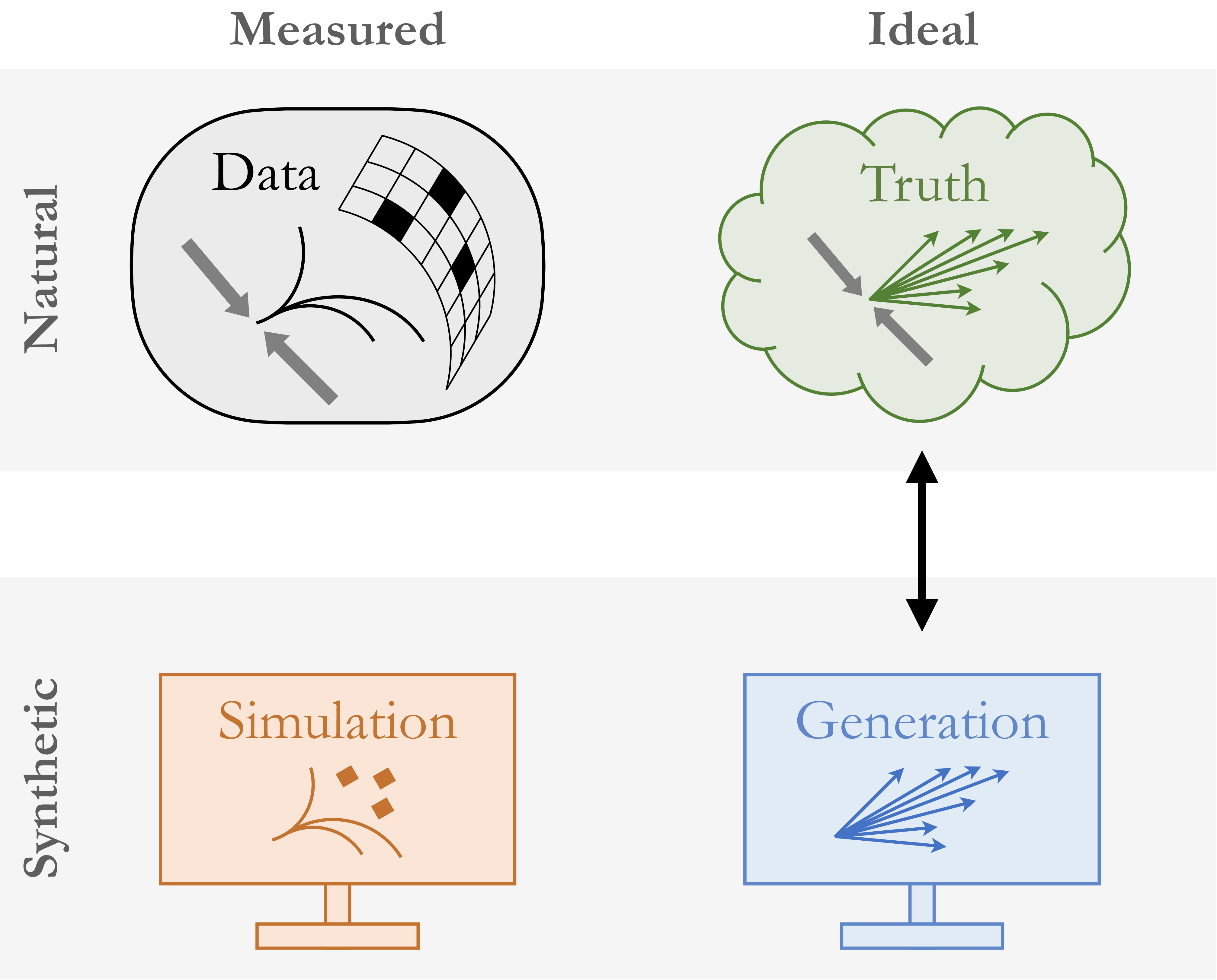
A. Andreassen, P. Komiske, E. Metodiev, BN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



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

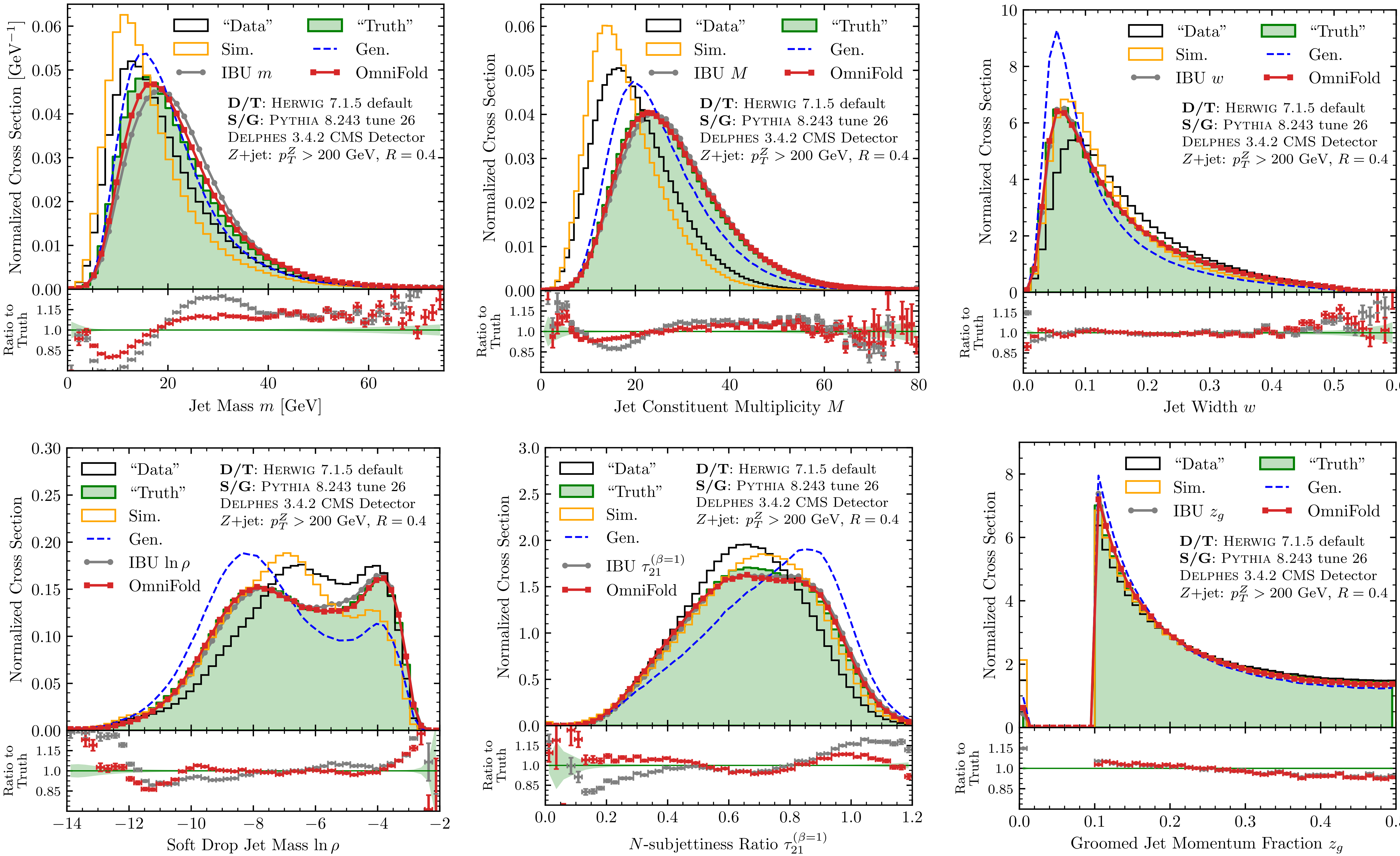
Unfold by iterating: OmniFold



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

Example: unfold all particles in Z+jets

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

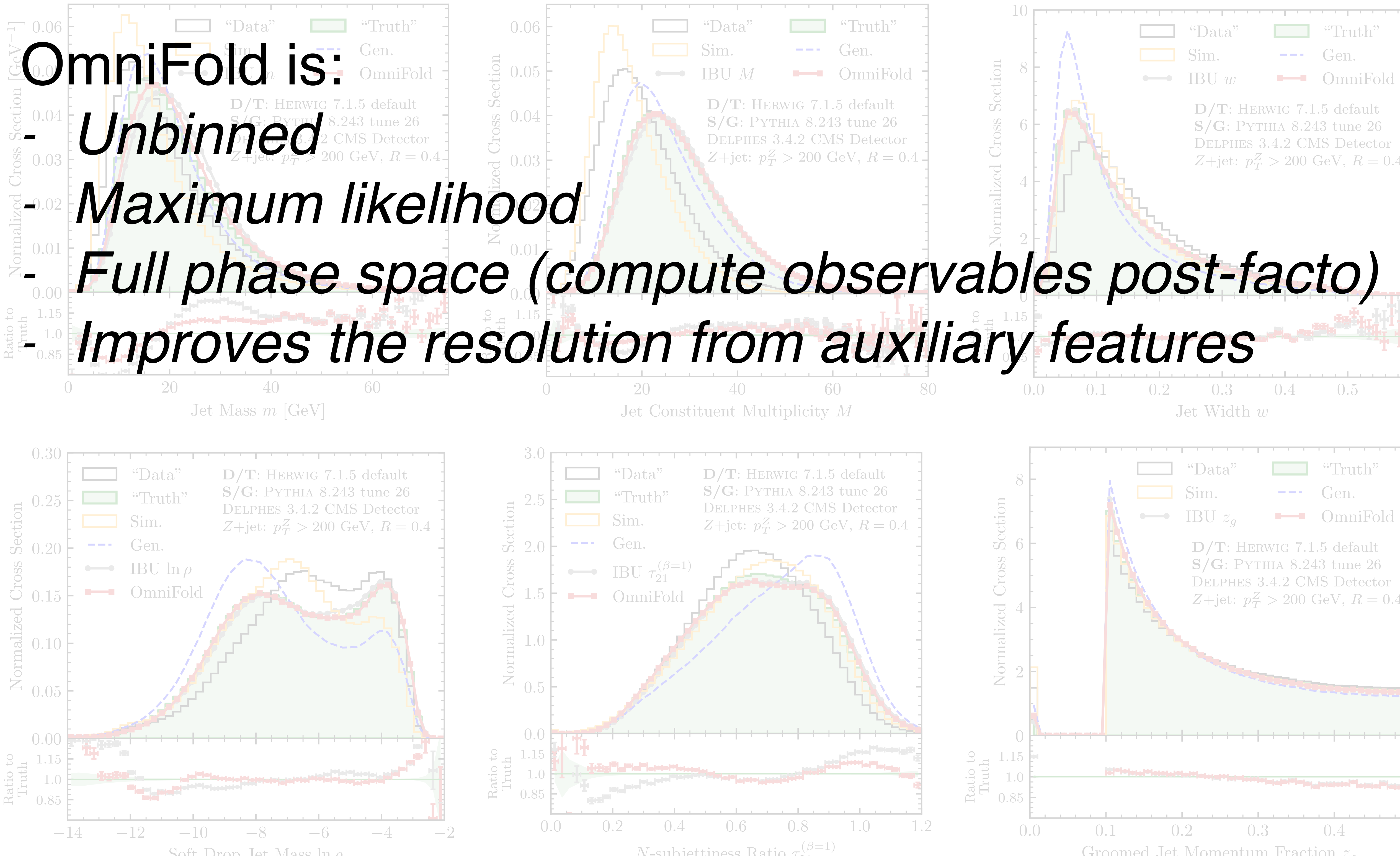


Example: unfold all particles in Z+jets

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

OmniFold is:

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

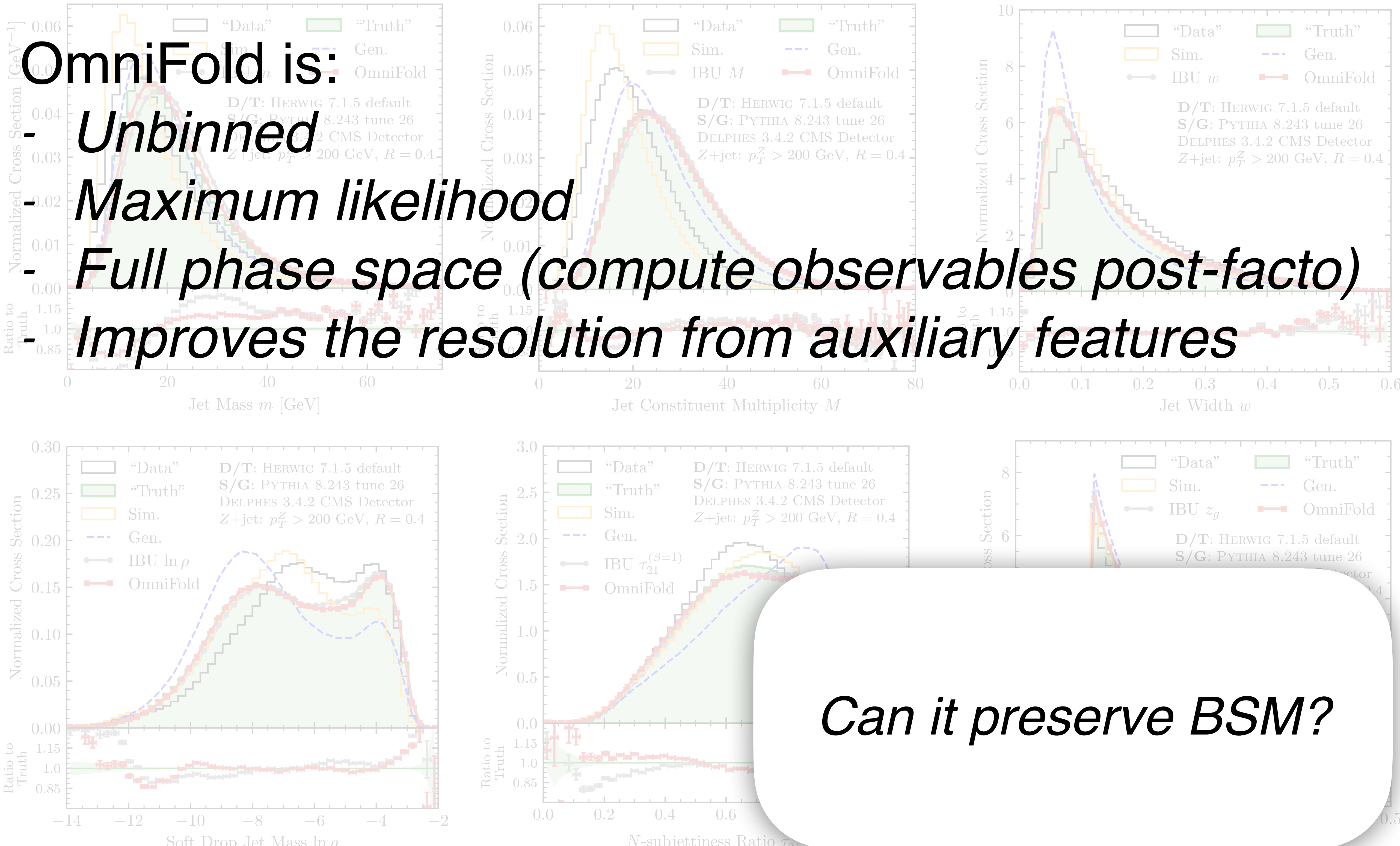


Example: unfold all particles in Z+jets

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

OmniFold is:

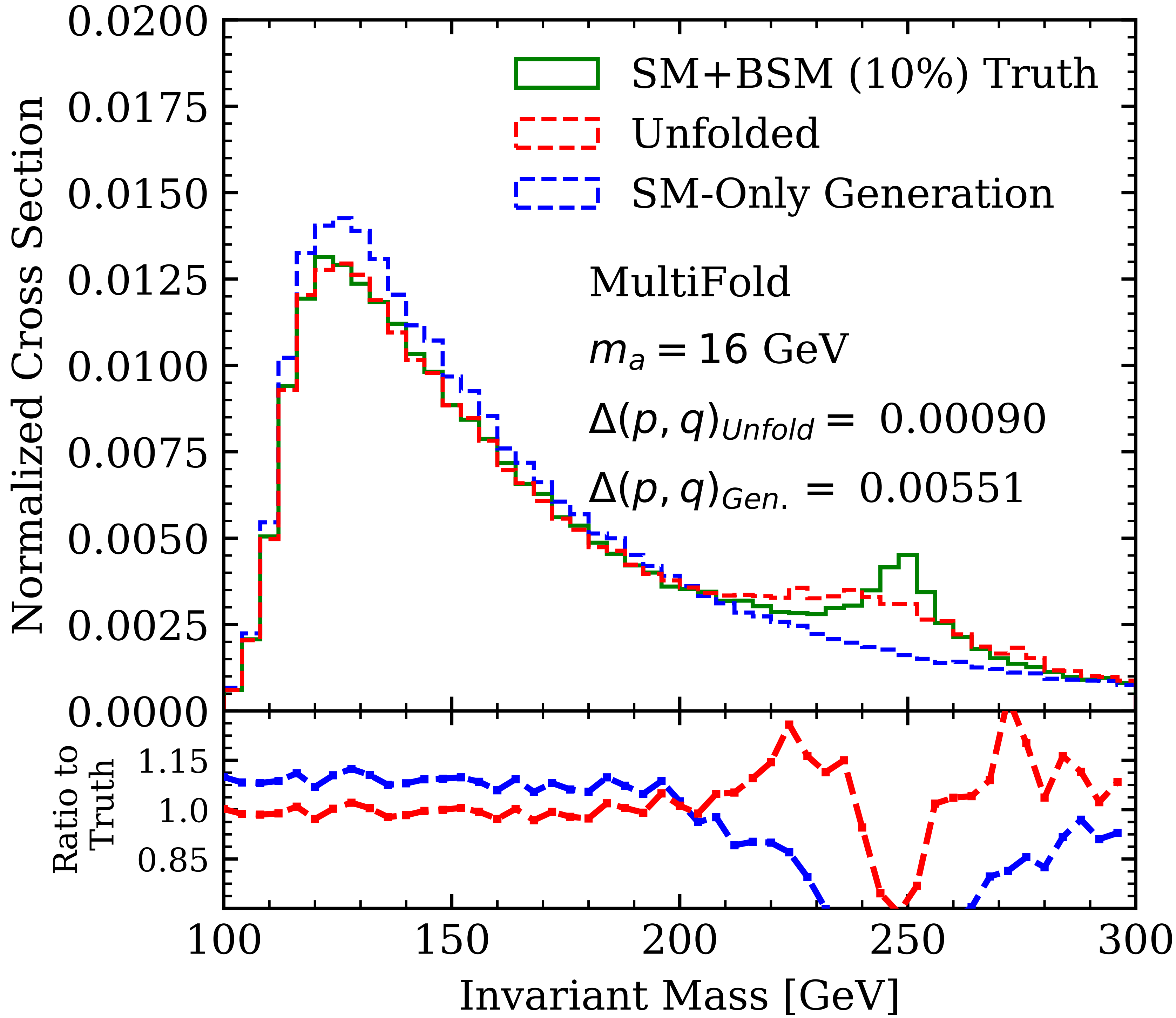
- *Unbinned*
- *Maximum likelihood*
- *Full phase space (compute observables post-facto)*
- *Improves the resolution from auxiliary features*



Can it preserve BSM?

OmniFold + BSM

P. Komiske, W. P. McCormack, BN, 2105.09923



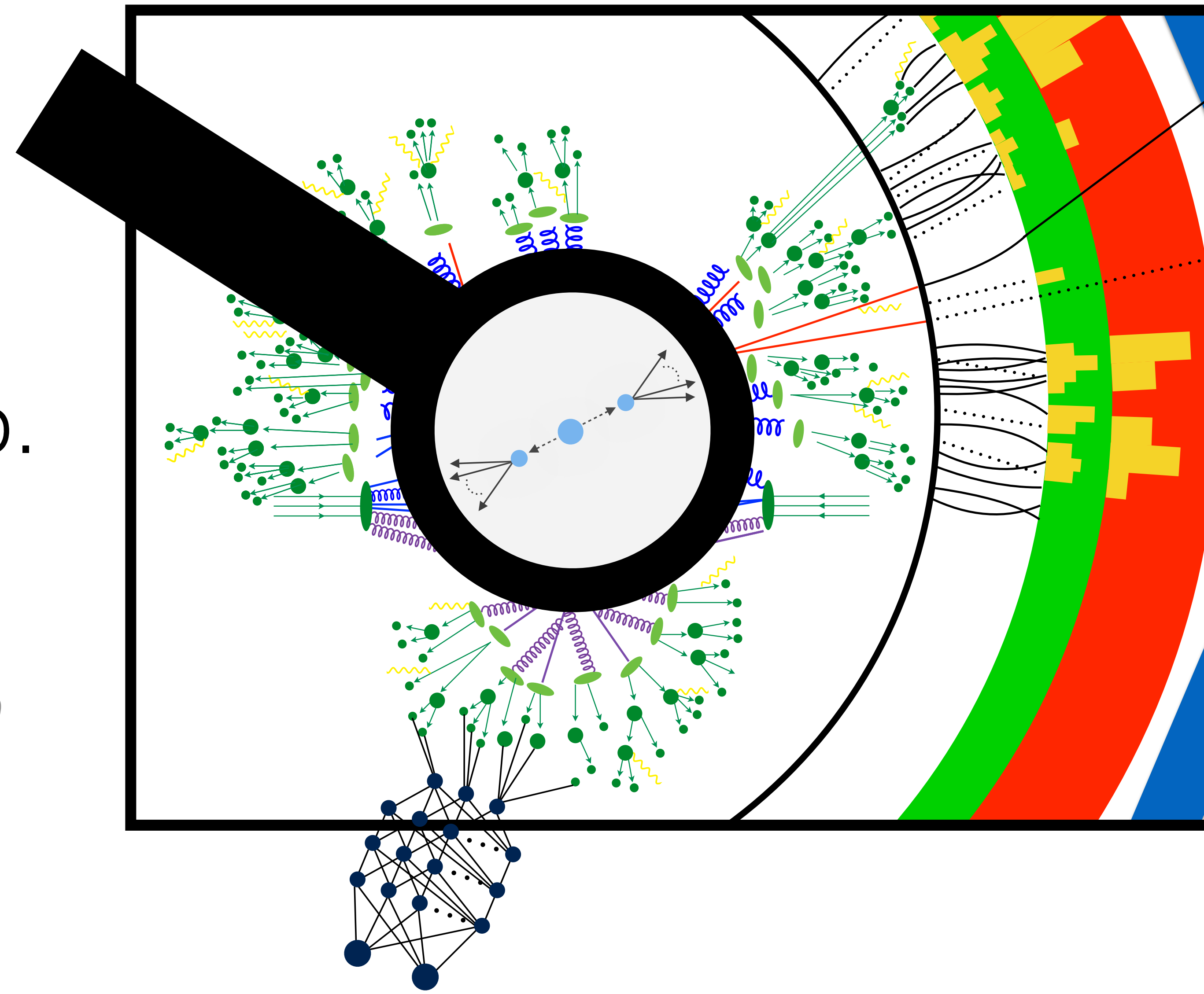
Z+jets with BSM in data, but not simulation

Non-local BSM is well-preserved; local BSM is preserved if (a) it is big enough and (b) it is in a region with enough phase space overlap with the background

Conclusions and outlook

Today I have focused on indirect searches for new physics with precision measurements. I also discussed how ML may help.

This is only a taste of both the physics and methodology; there is also a rich program in direct searches (with and without ML)

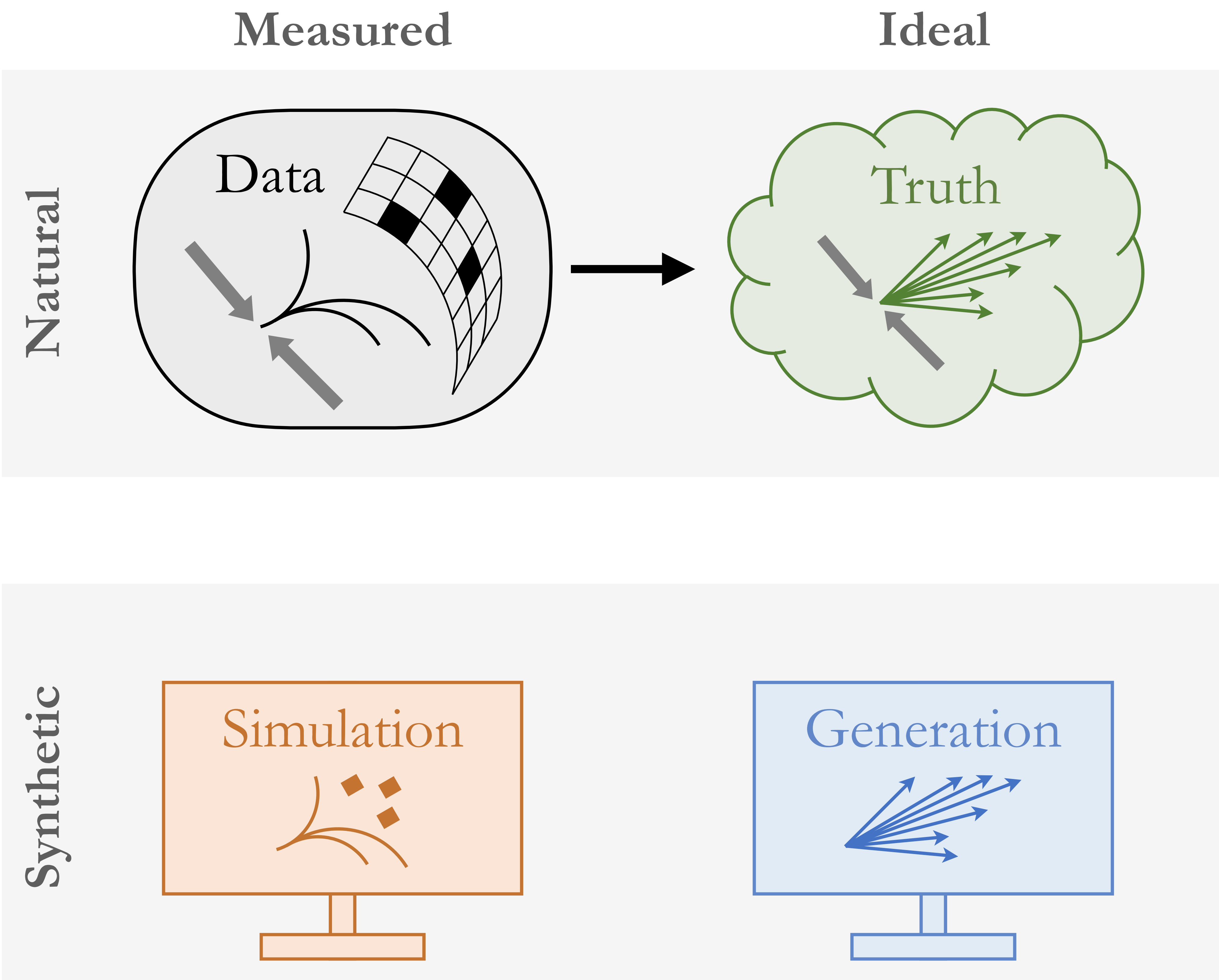


The **full phase space** of our experiments is now explorable, and with new measurements combined with new theory insight, we will be able to be maximally sensitive to BSM!

Backup

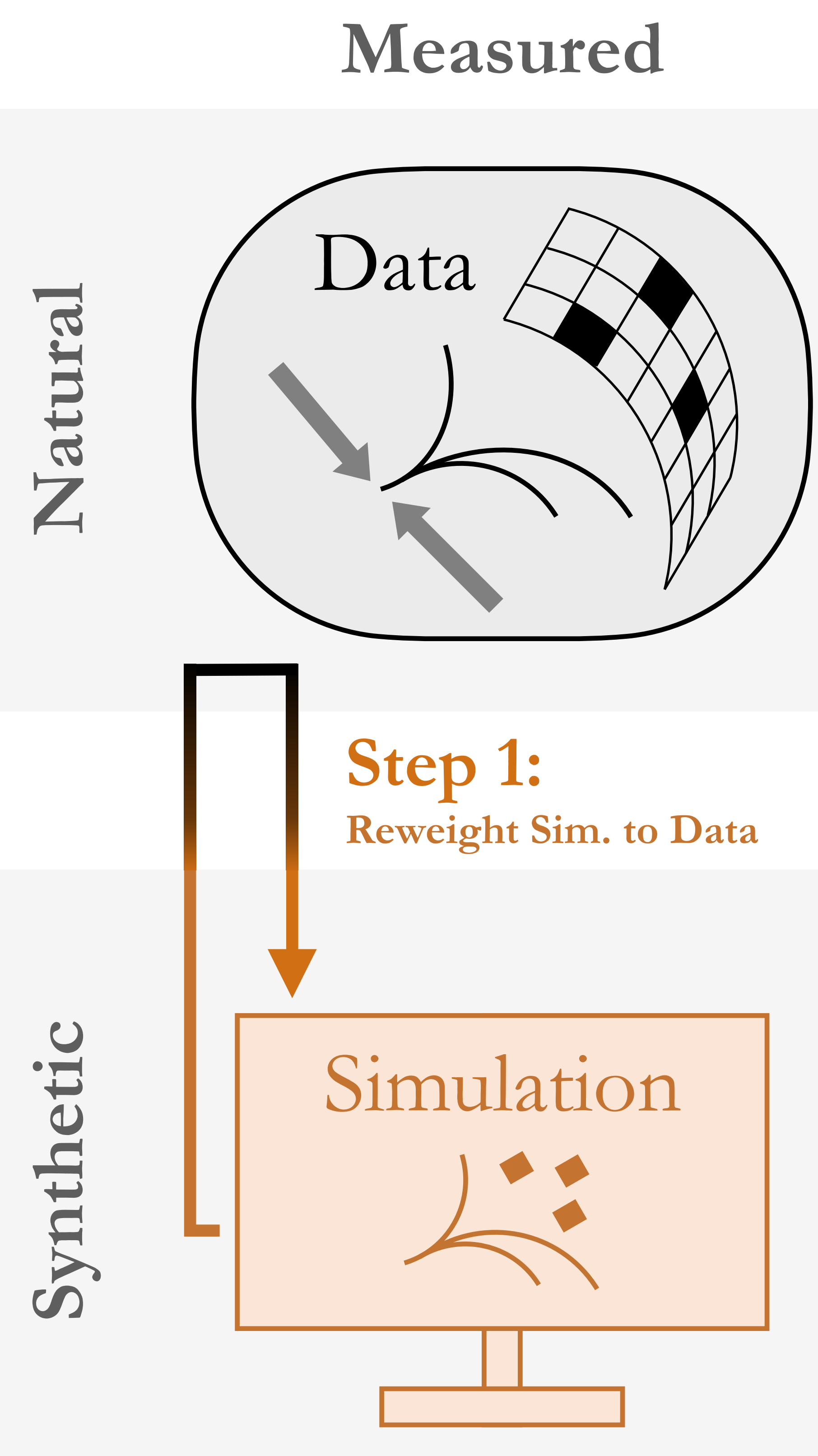
63

Unfold by iterating: OmniFold



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

Unfold by iterating: OmniFold

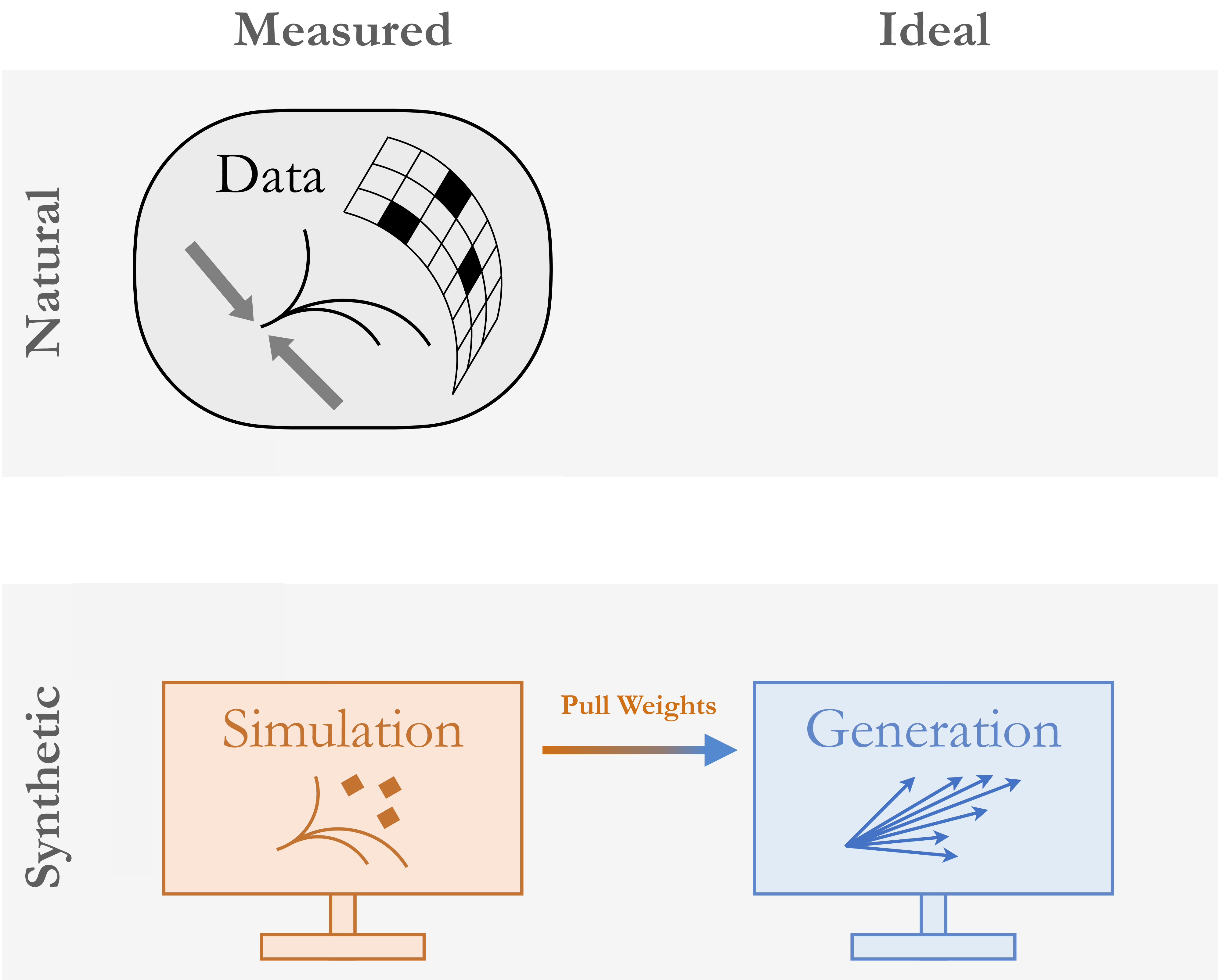


Ideal

Detector effects are often simulated with programs like Geant4

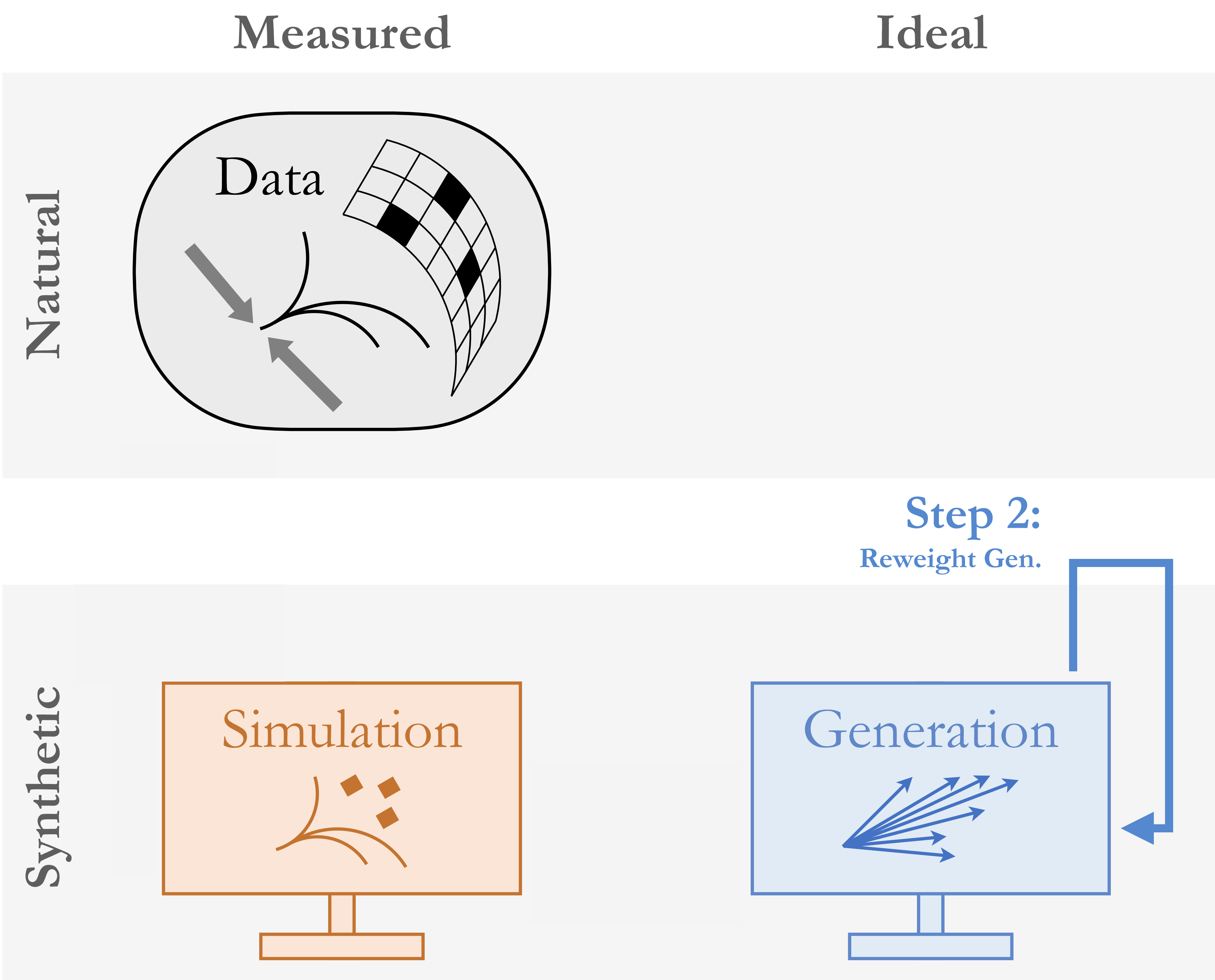
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



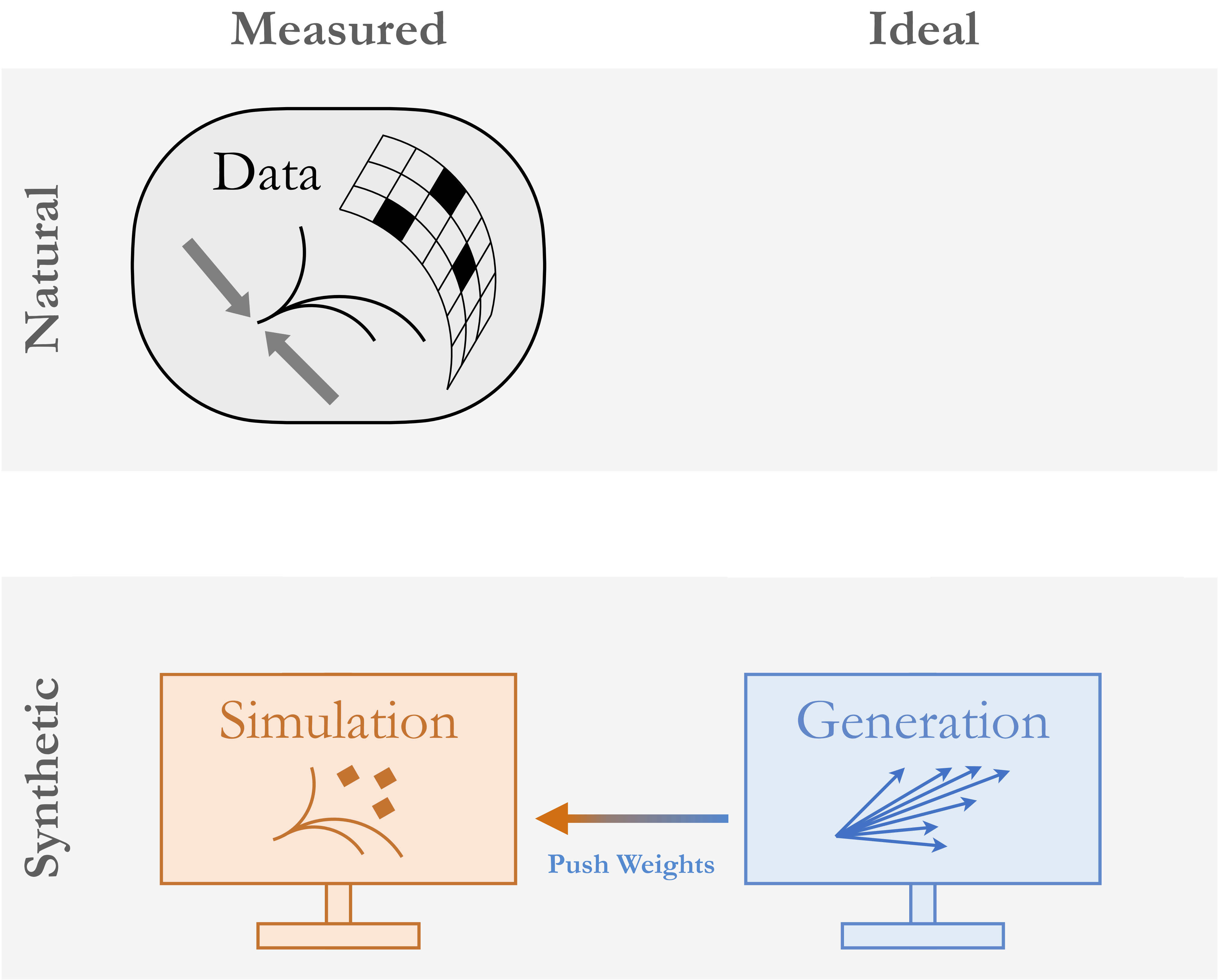
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



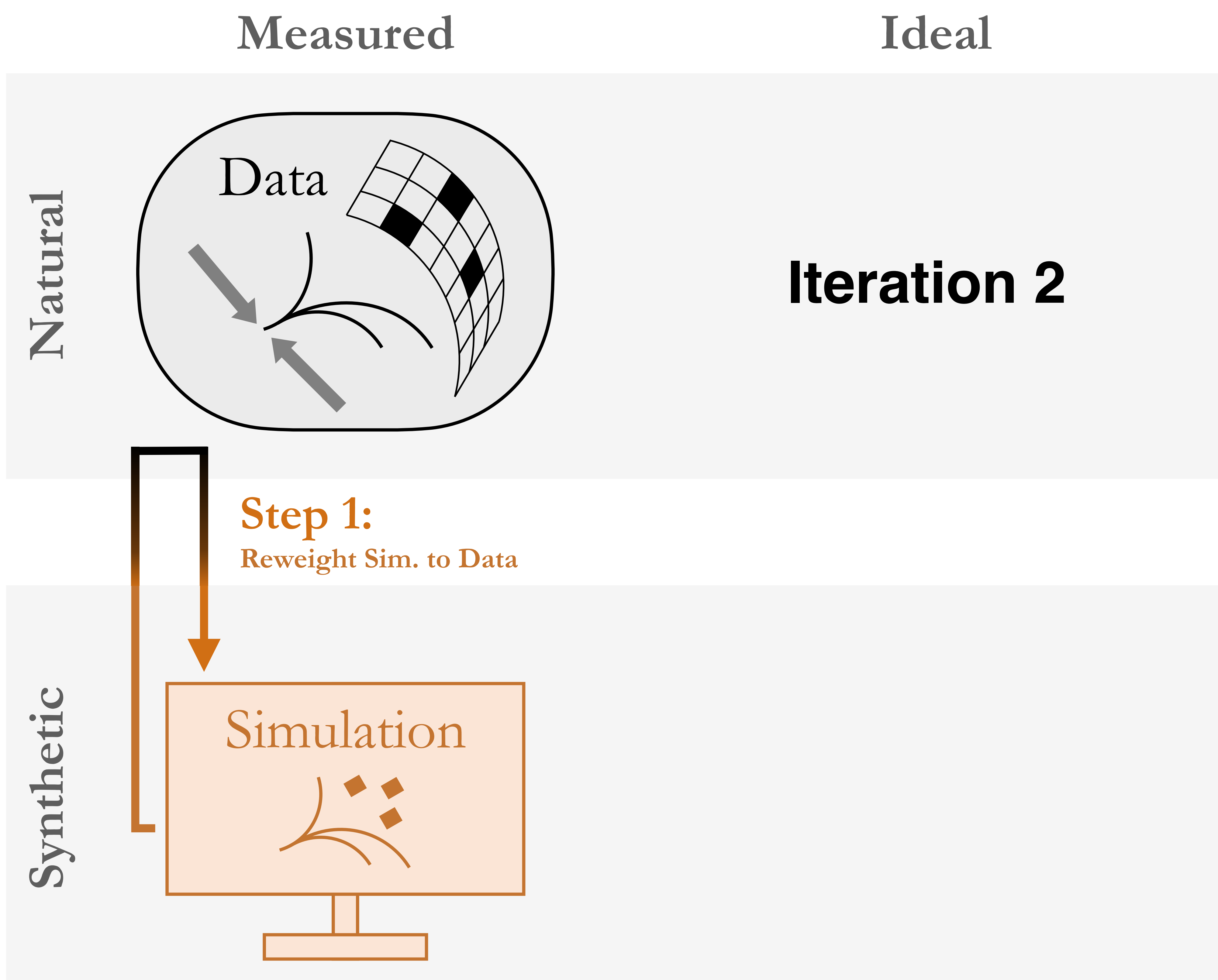
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



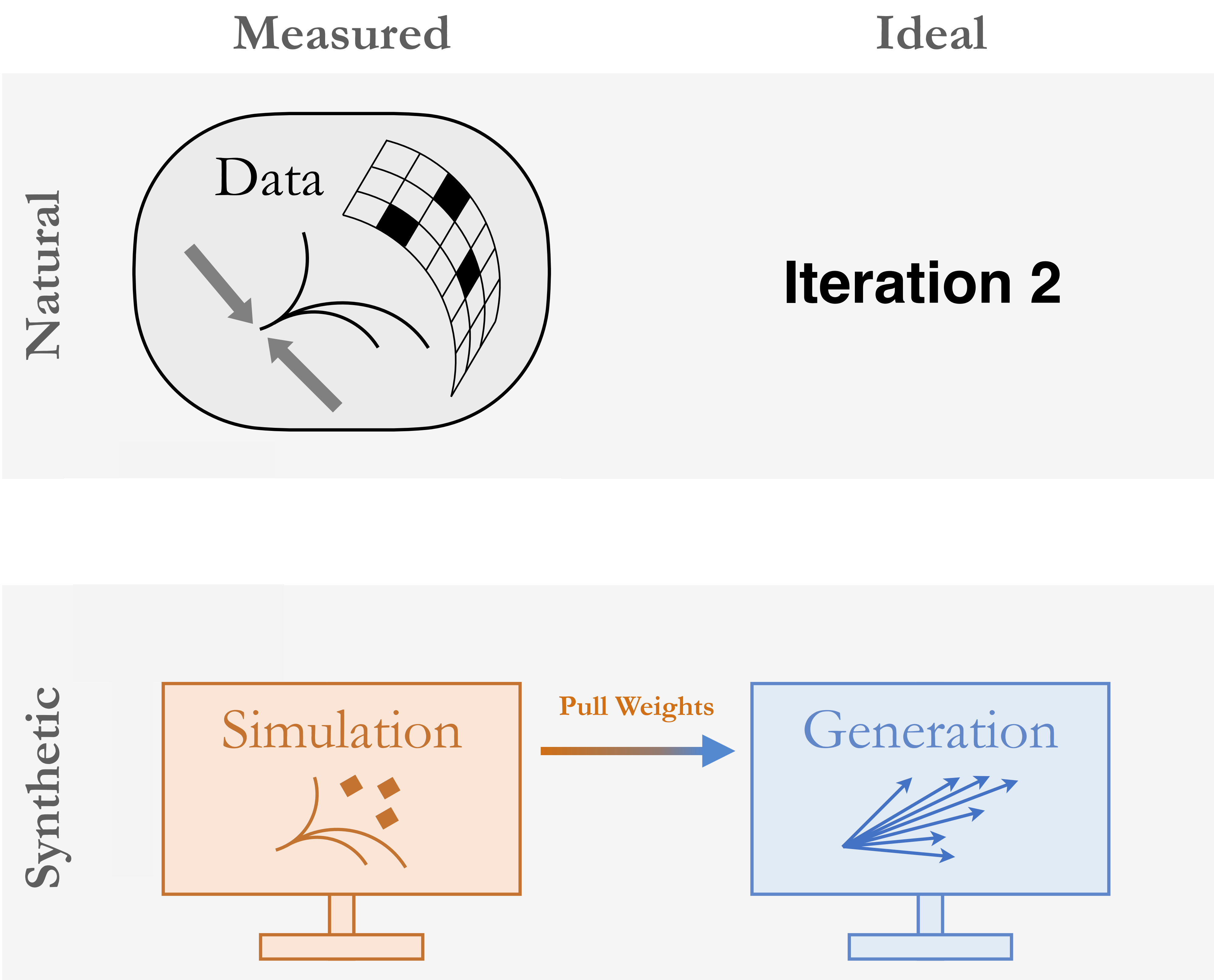
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



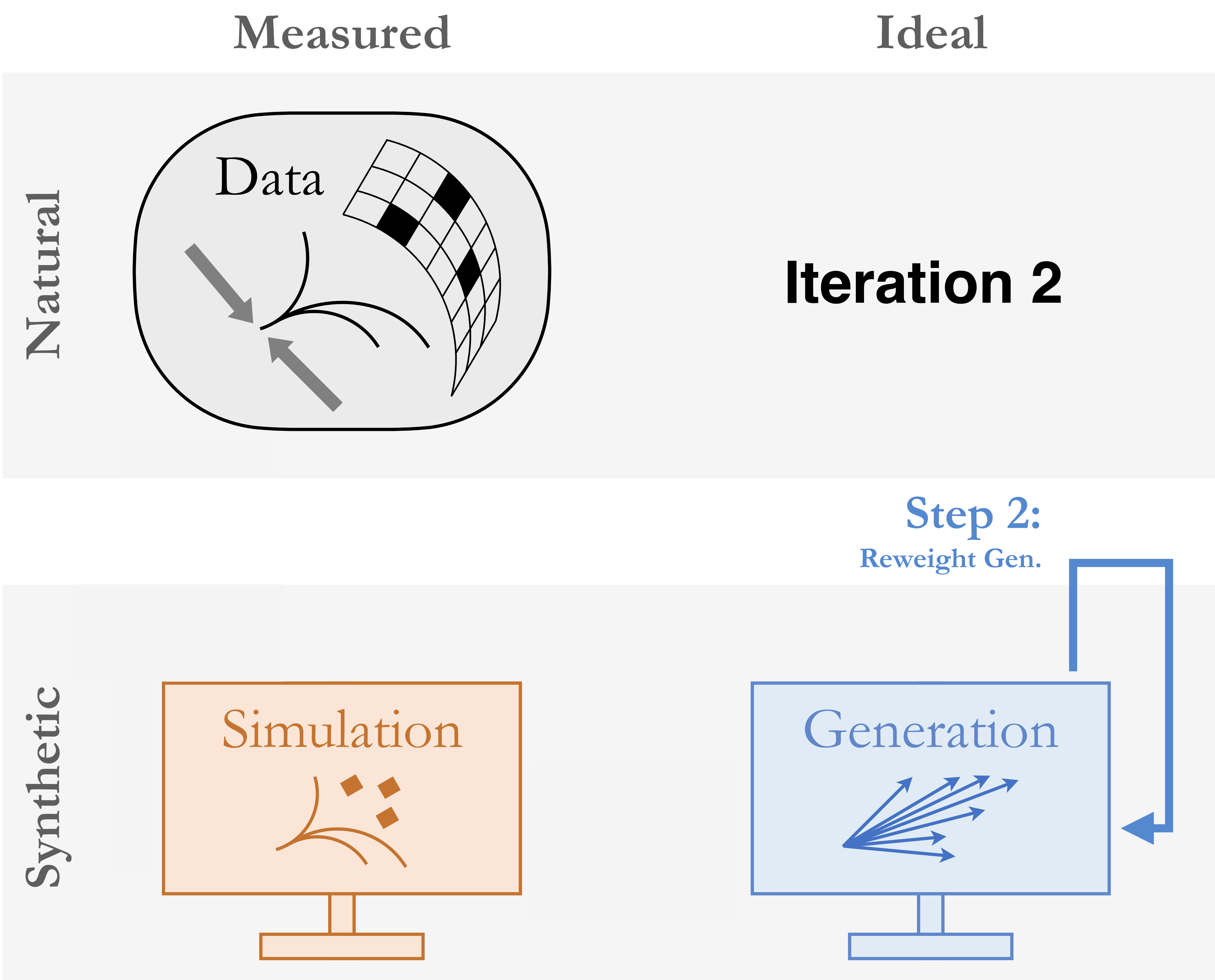
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



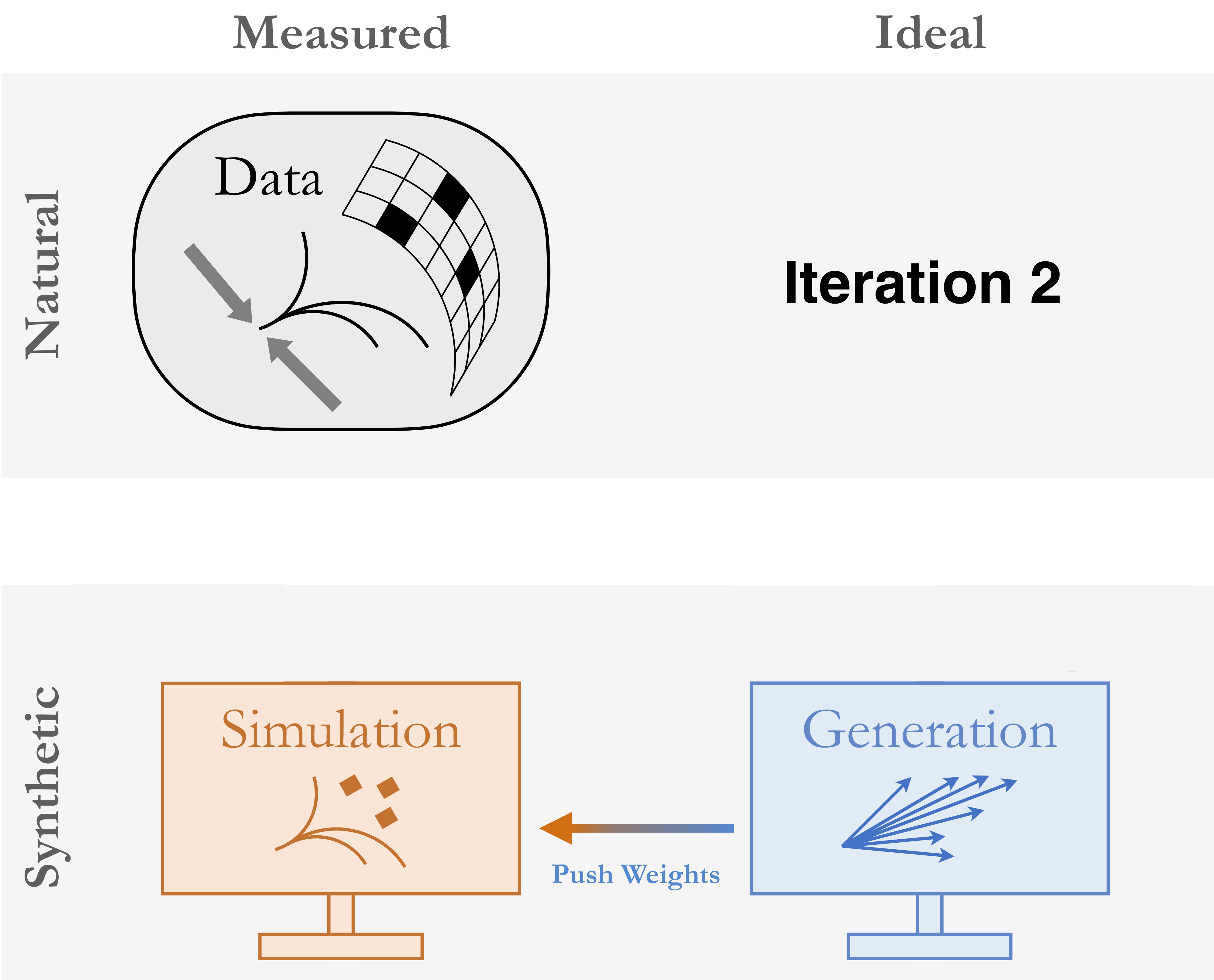
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



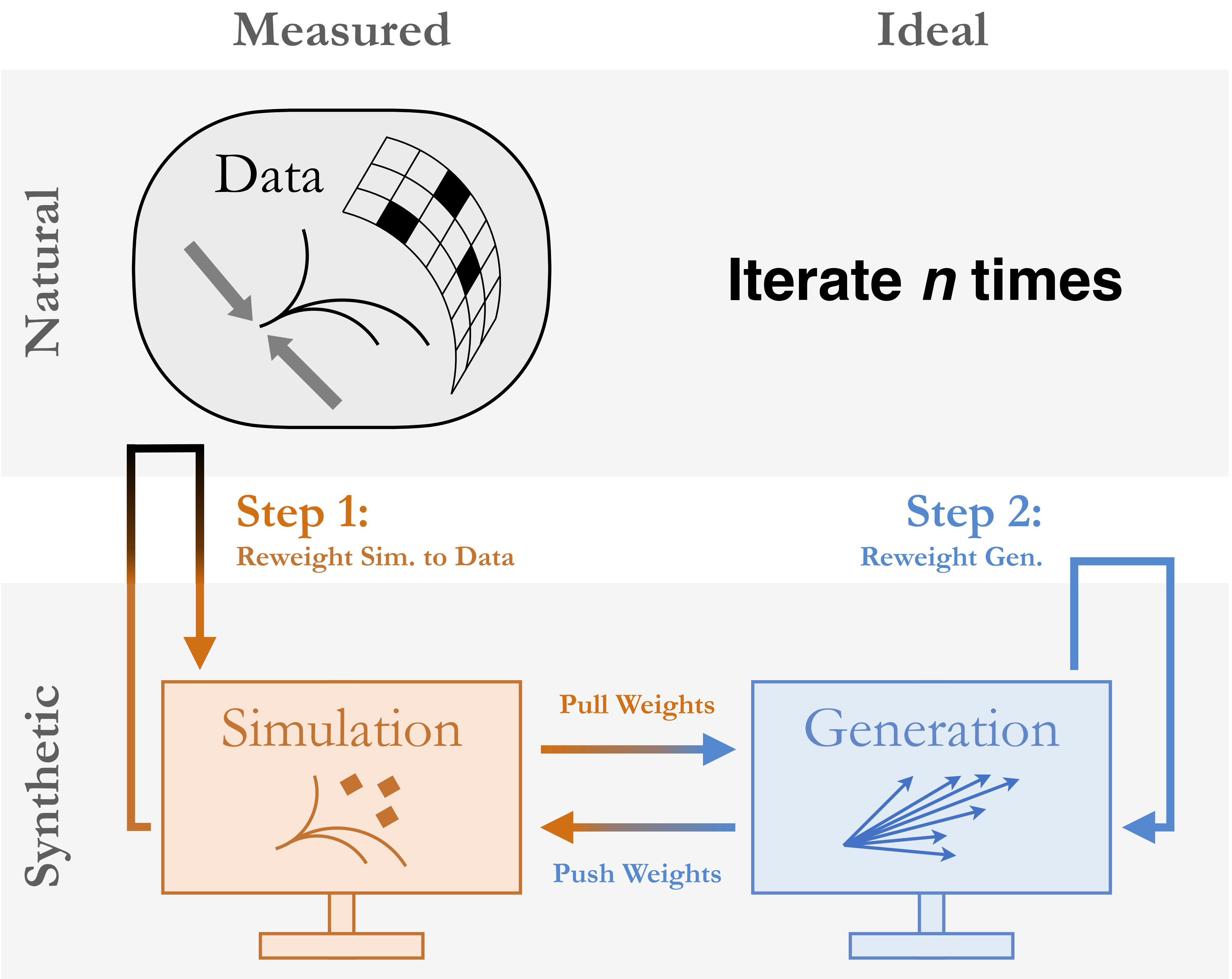
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



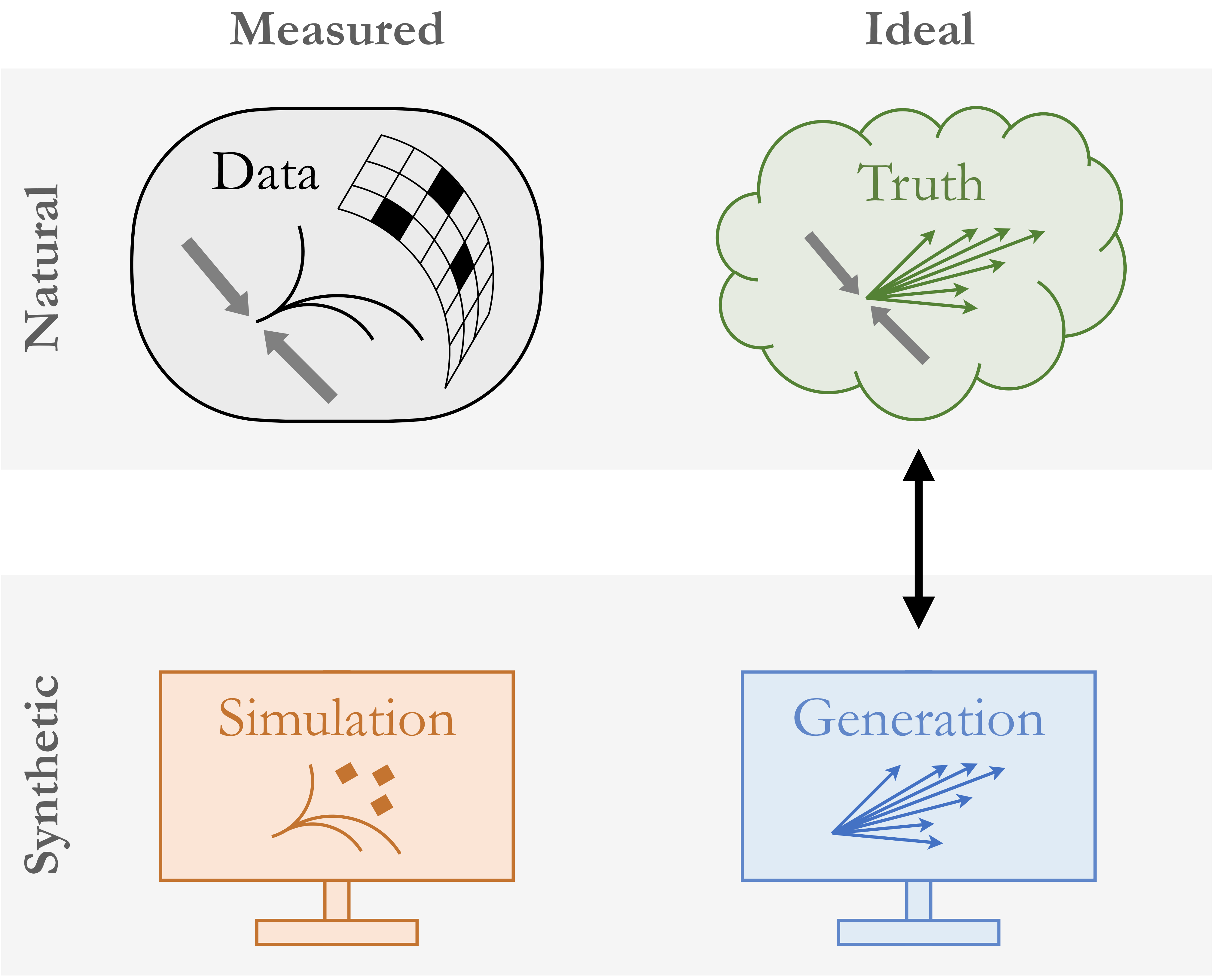
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



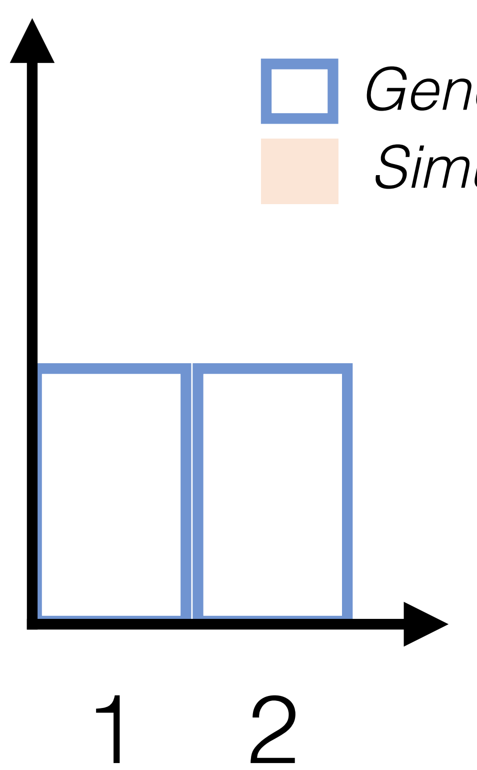
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

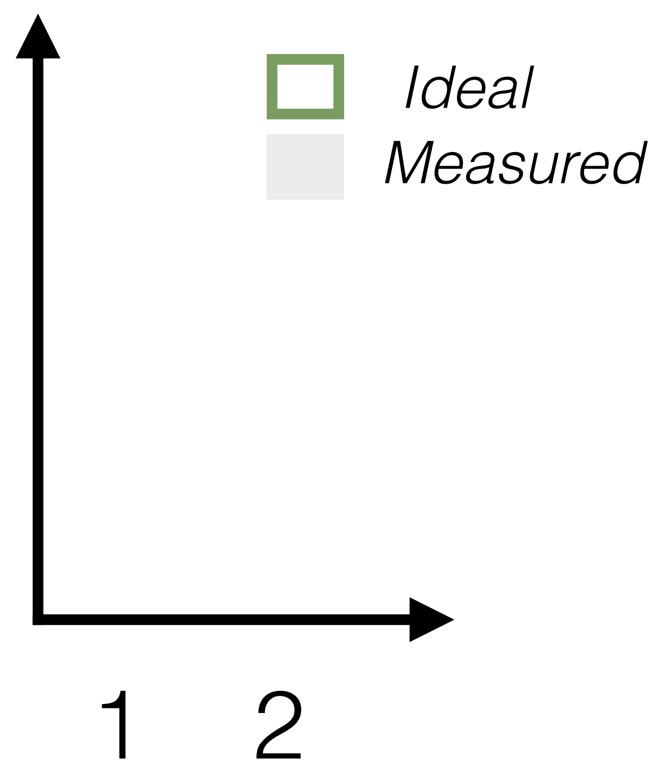
Unfold by iterating: OmniFold

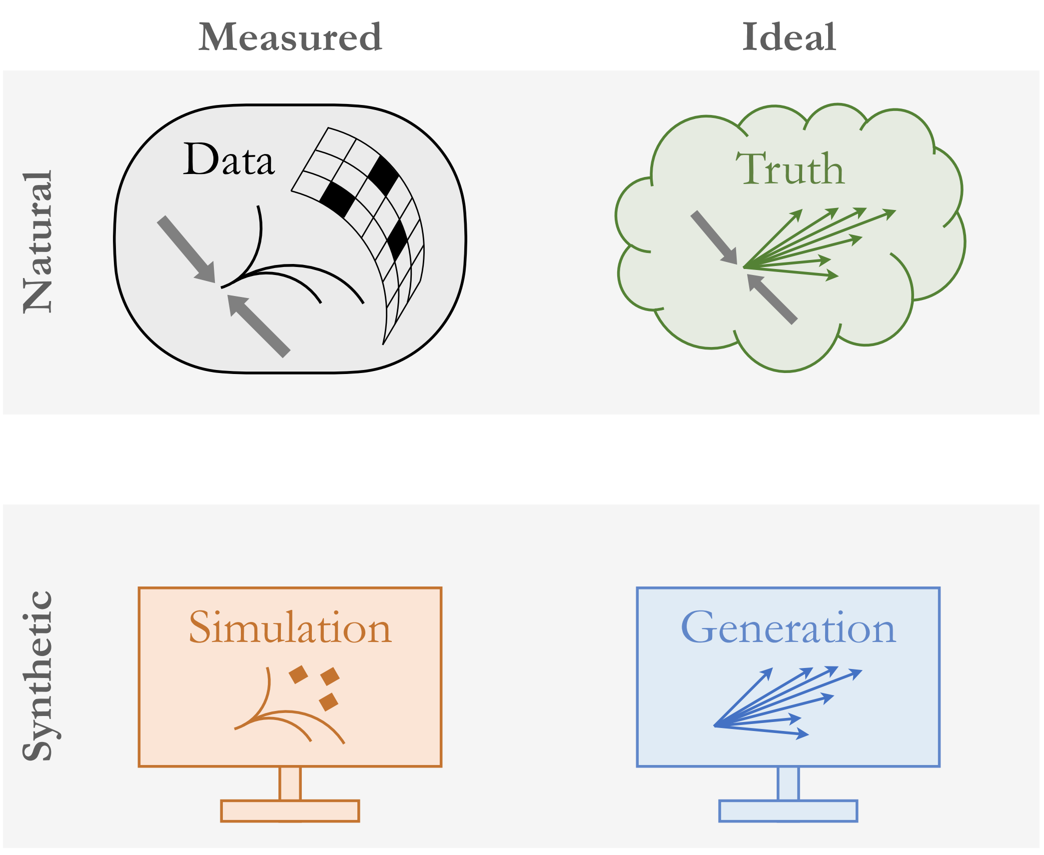


A. Andreassen, P. Komiske, E. Metodiev,
BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold

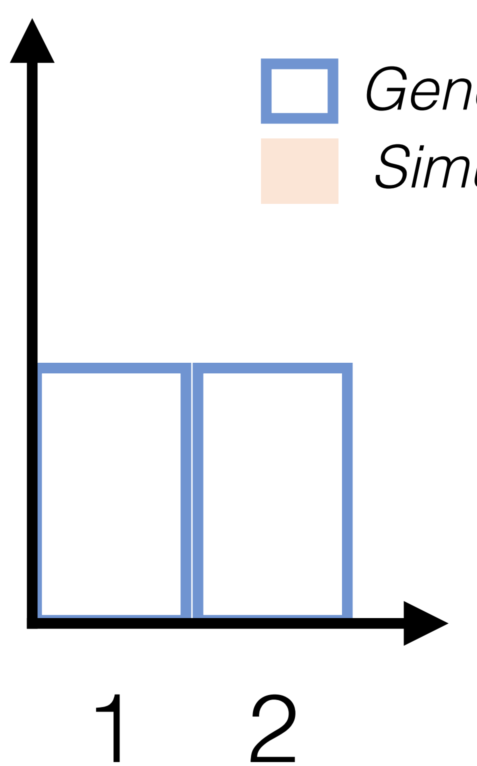


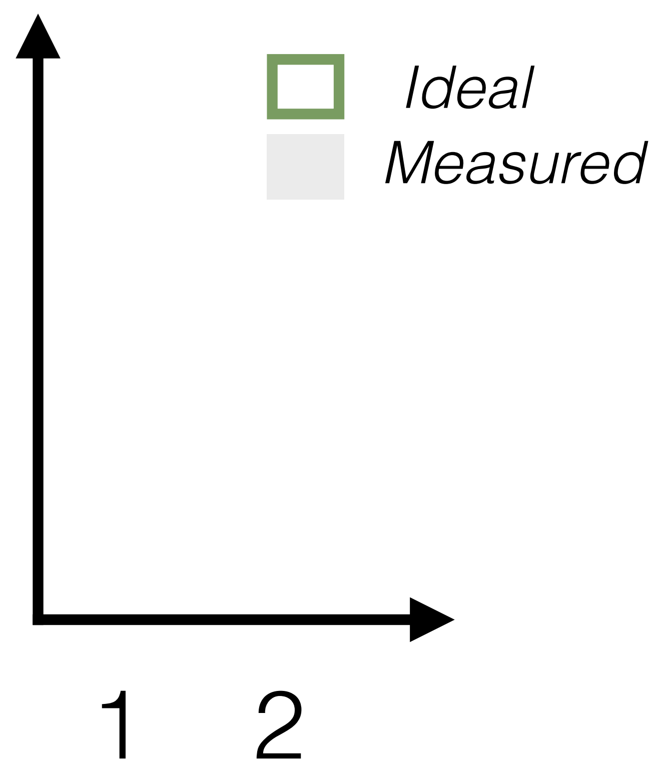




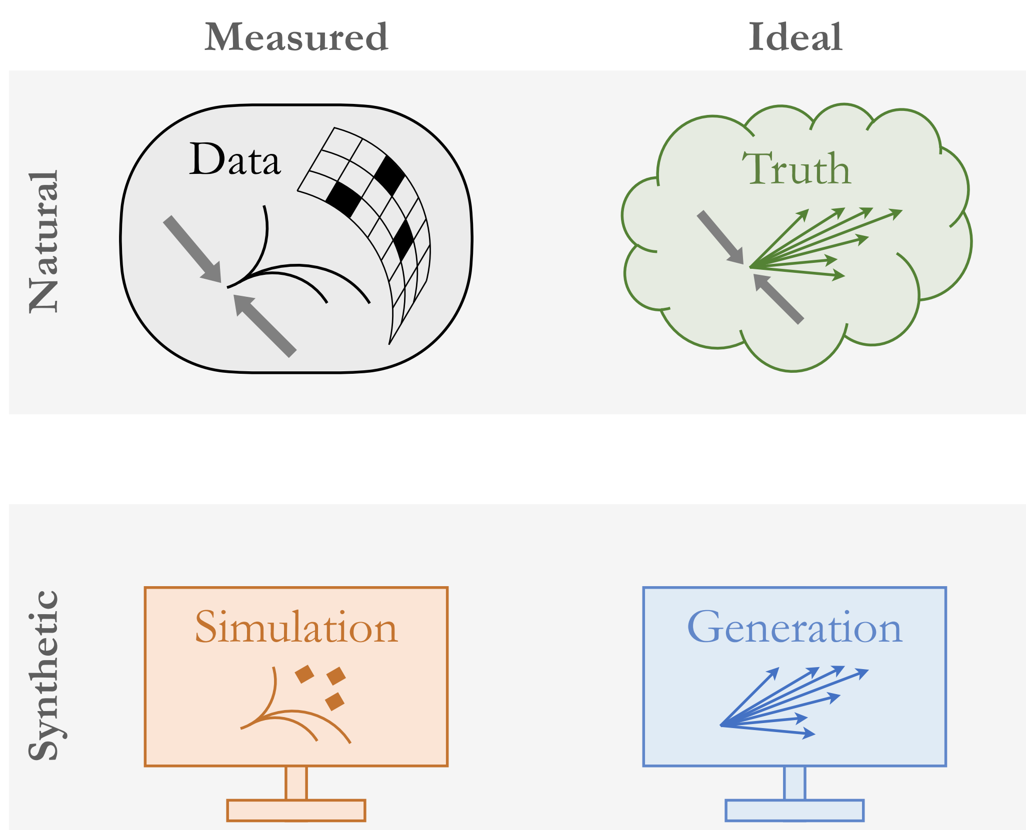
A. Andreassen, P. Komiske, E. Metodiev,
BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



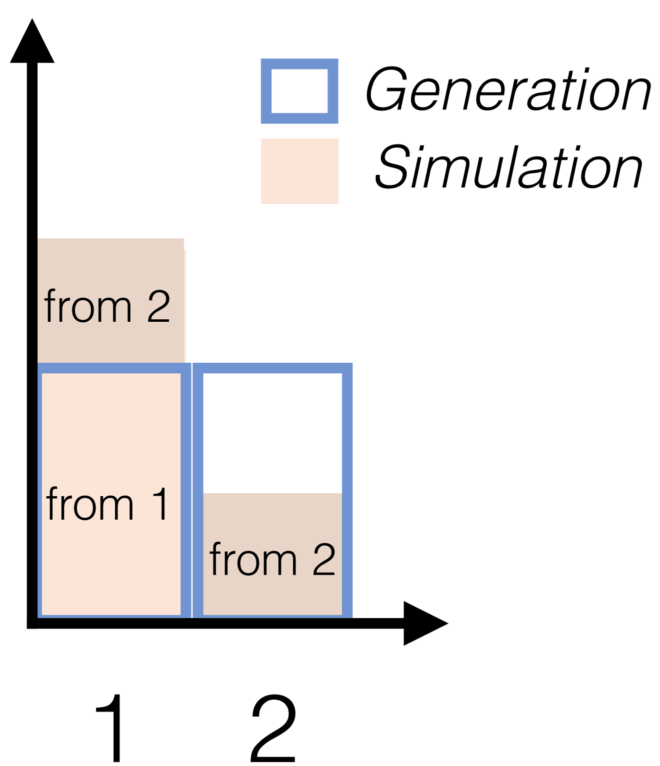


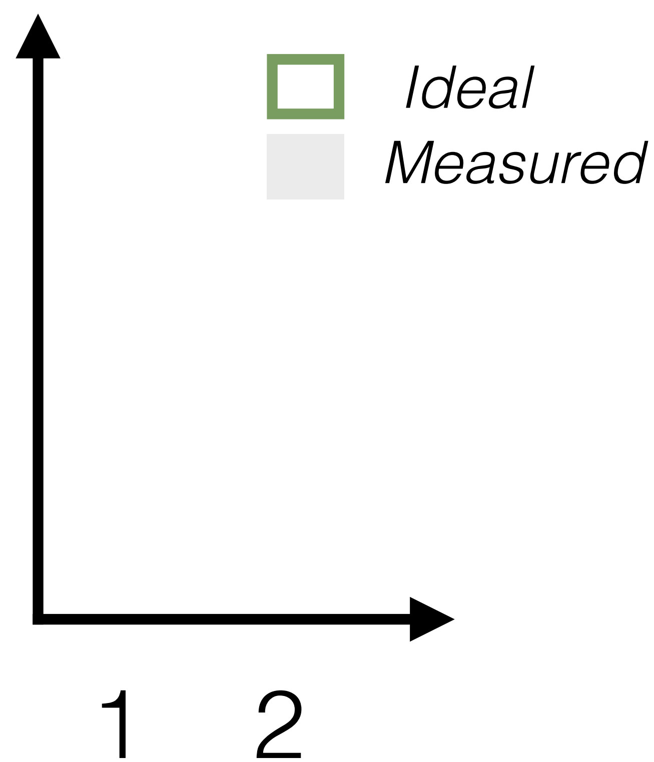
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%



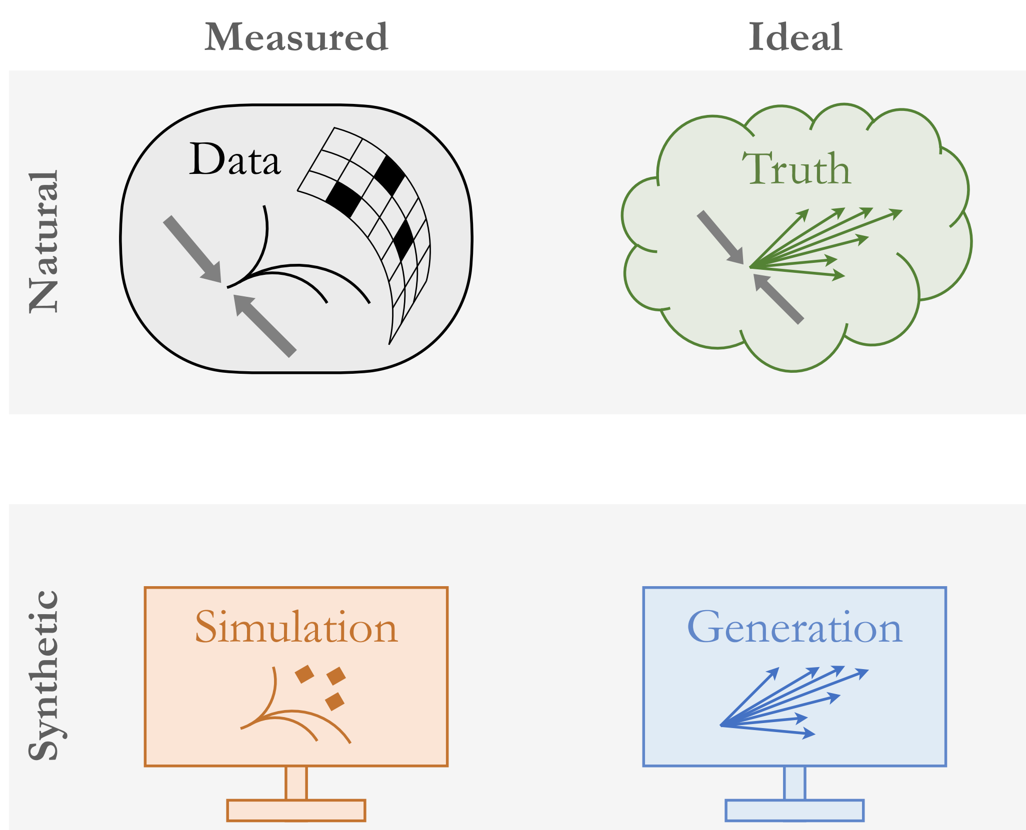
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



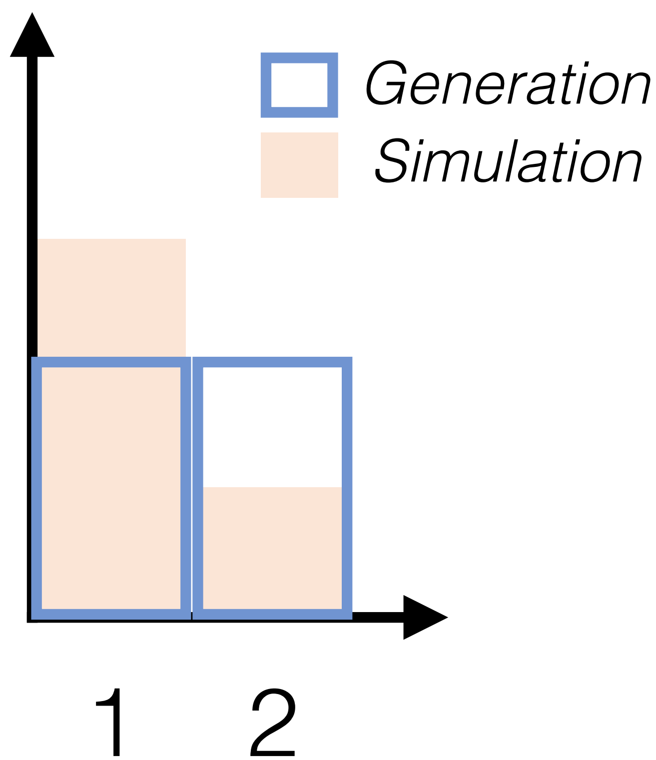


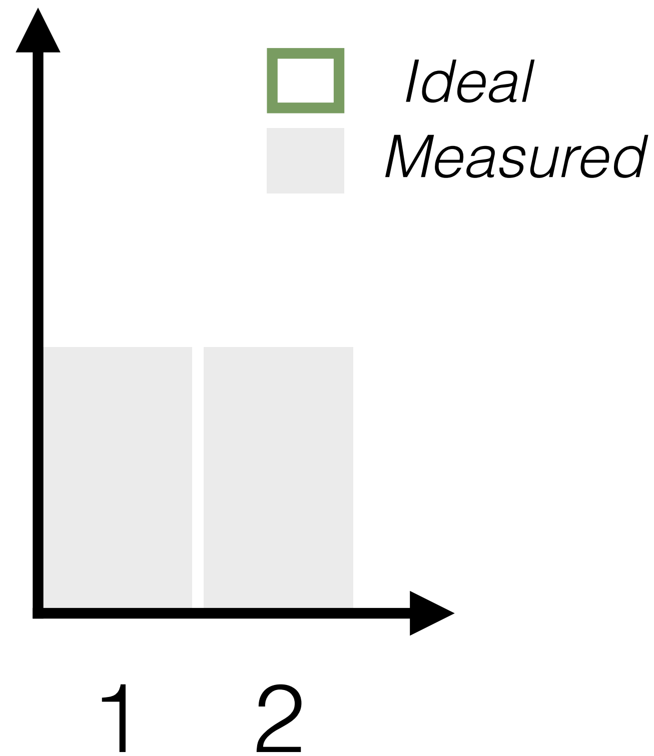
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%



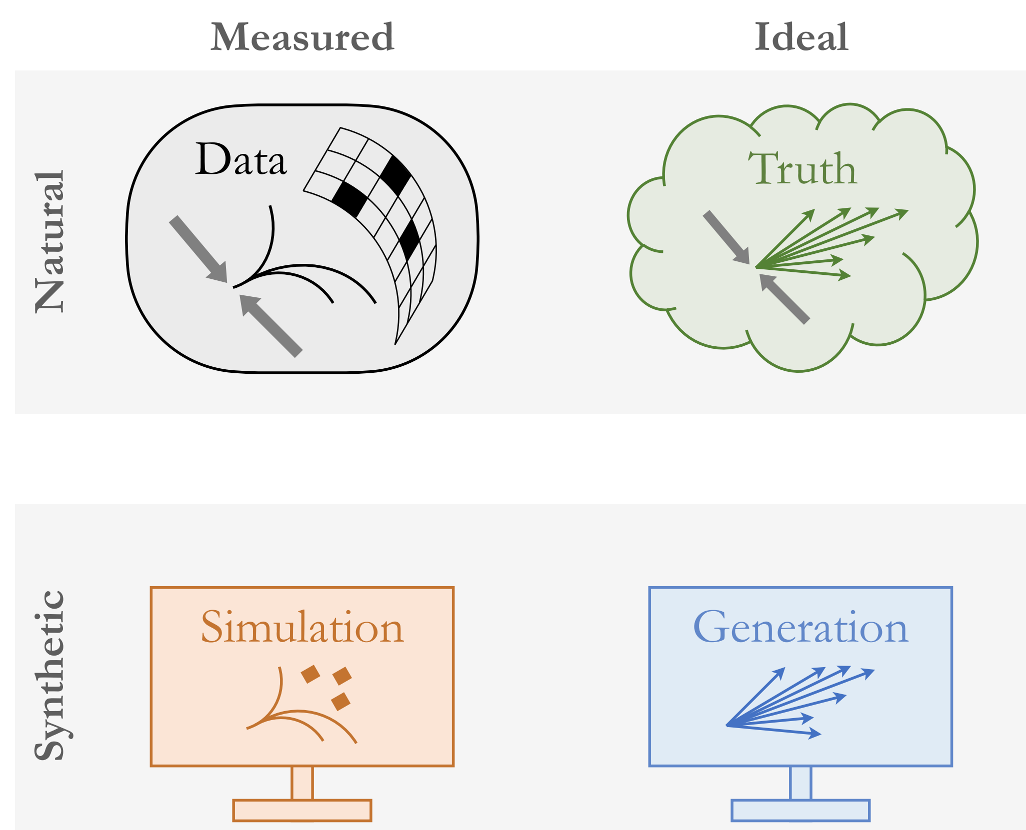
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



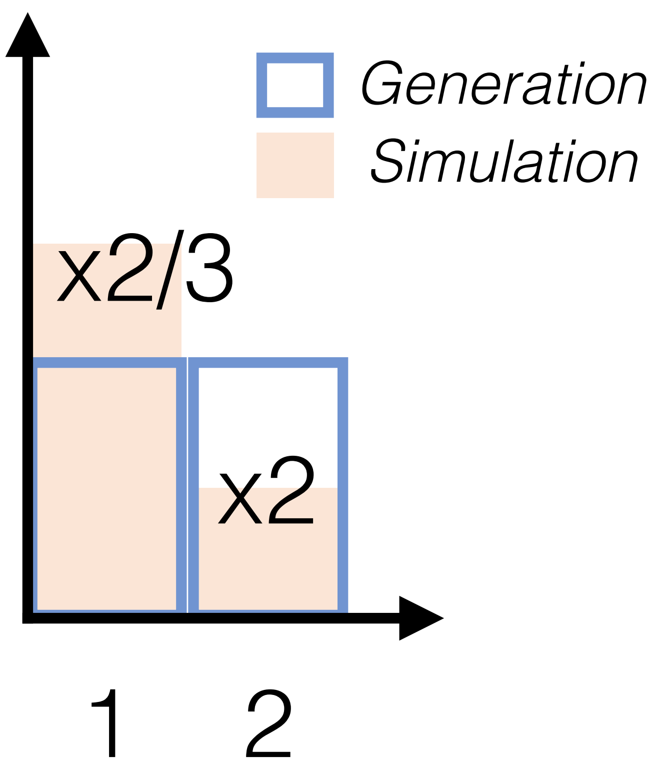


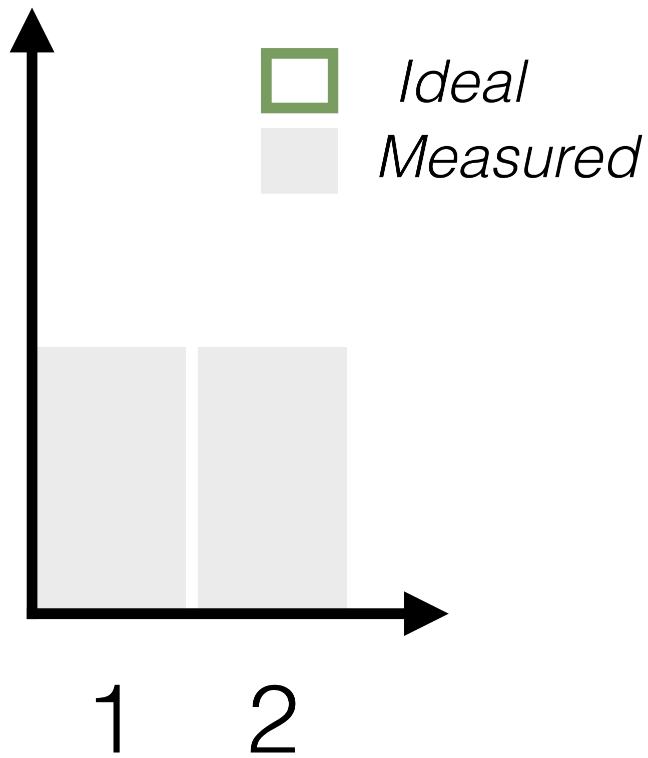
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%



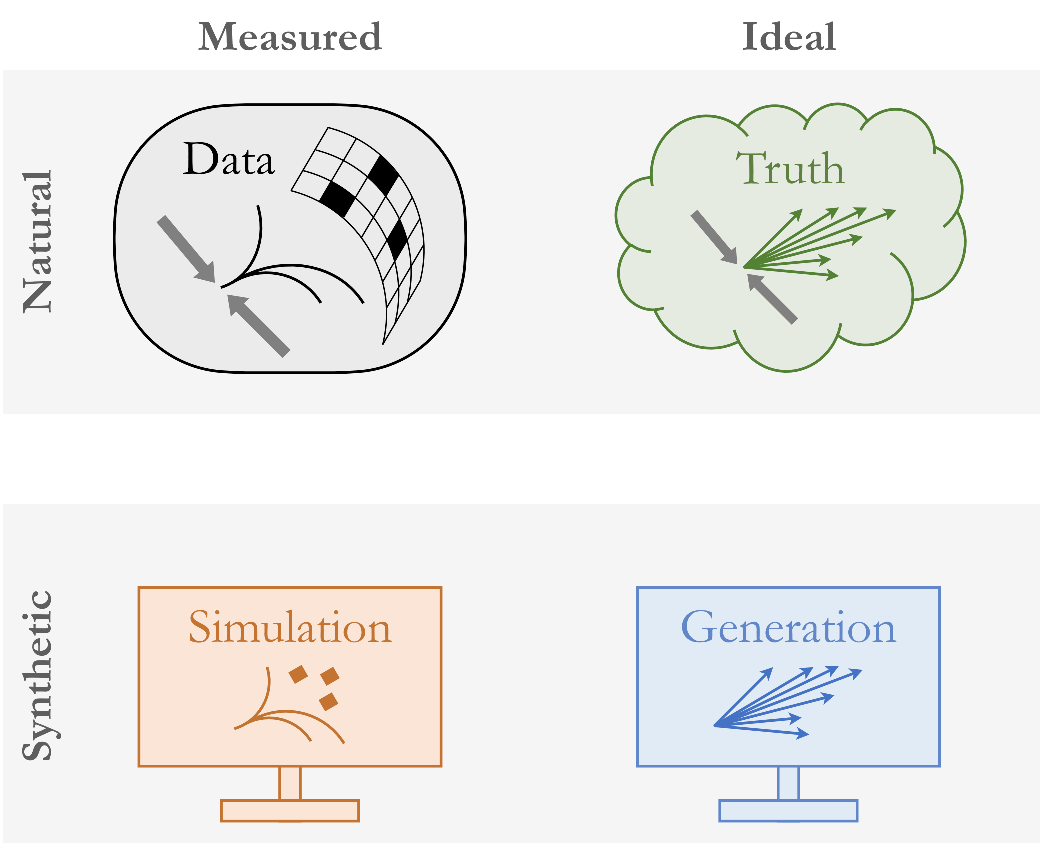
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



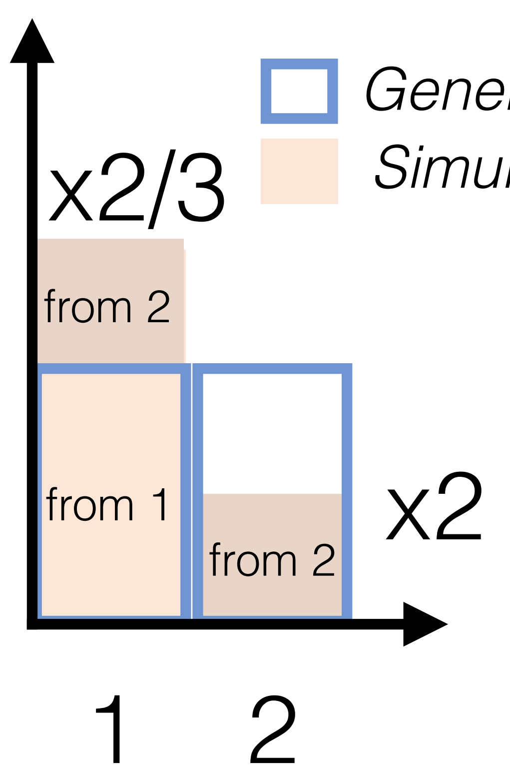


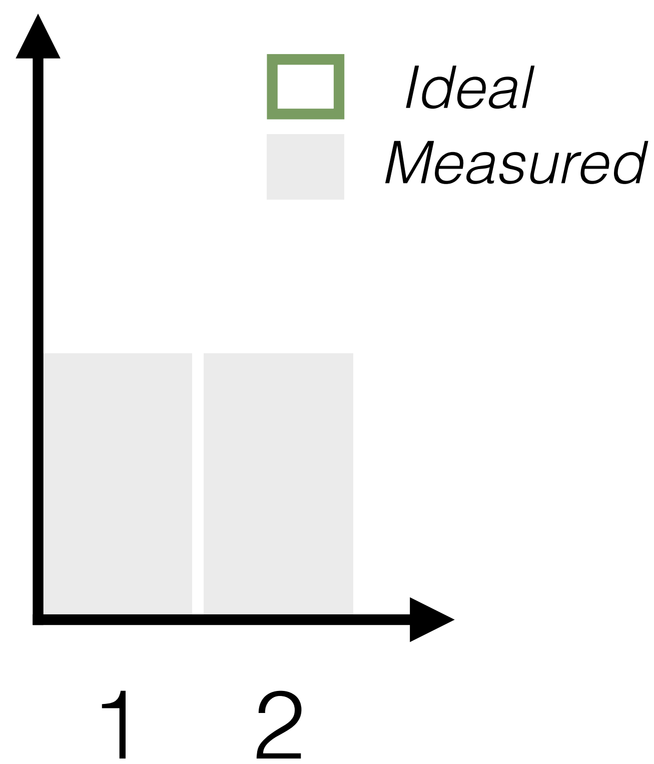
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%



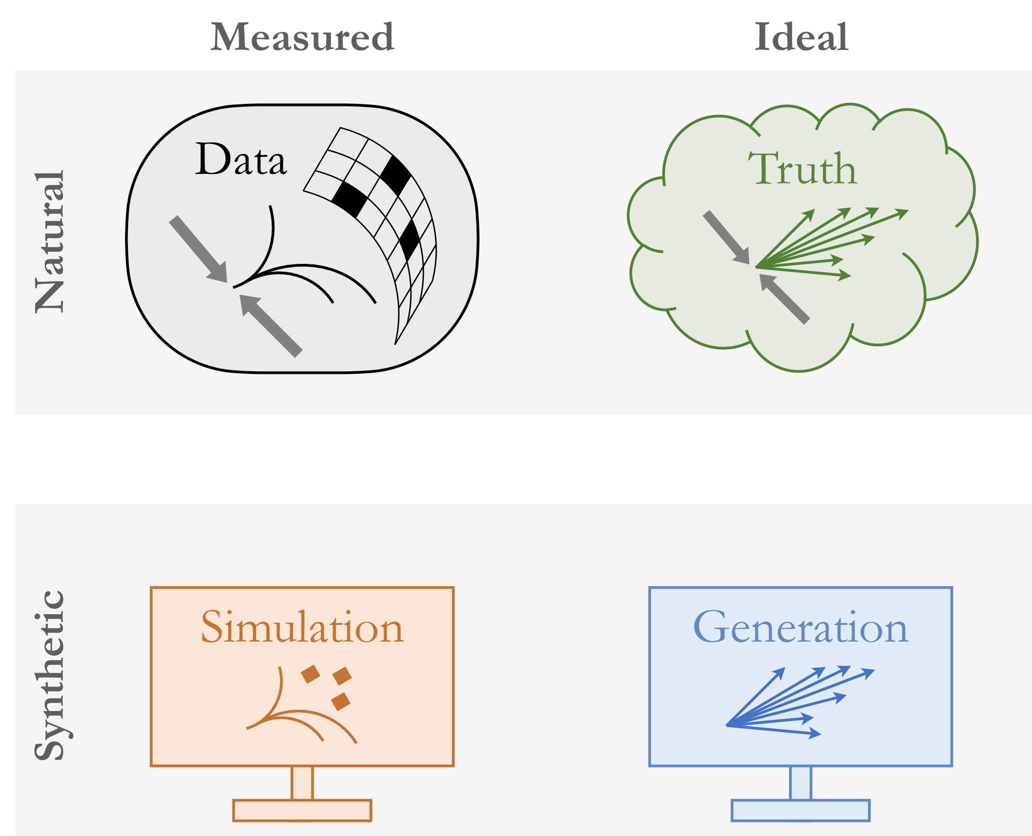
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold





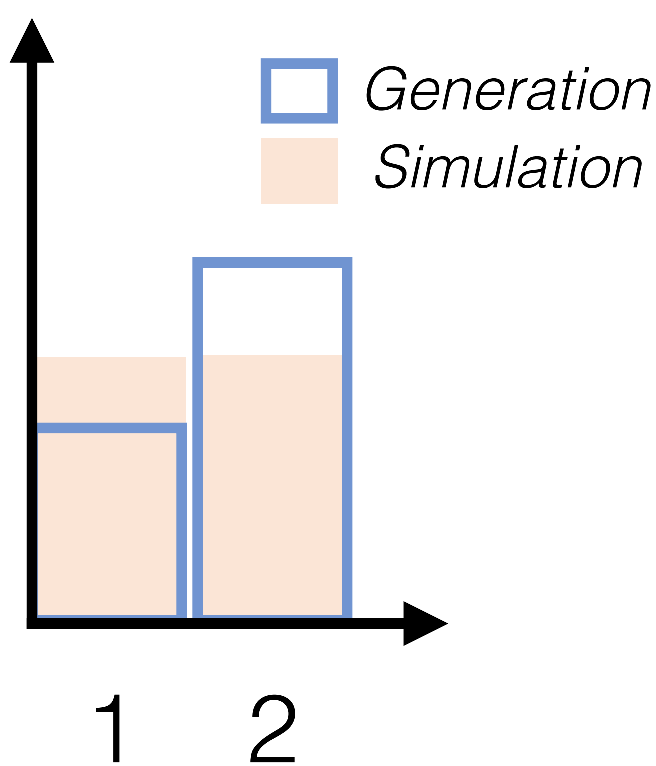
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

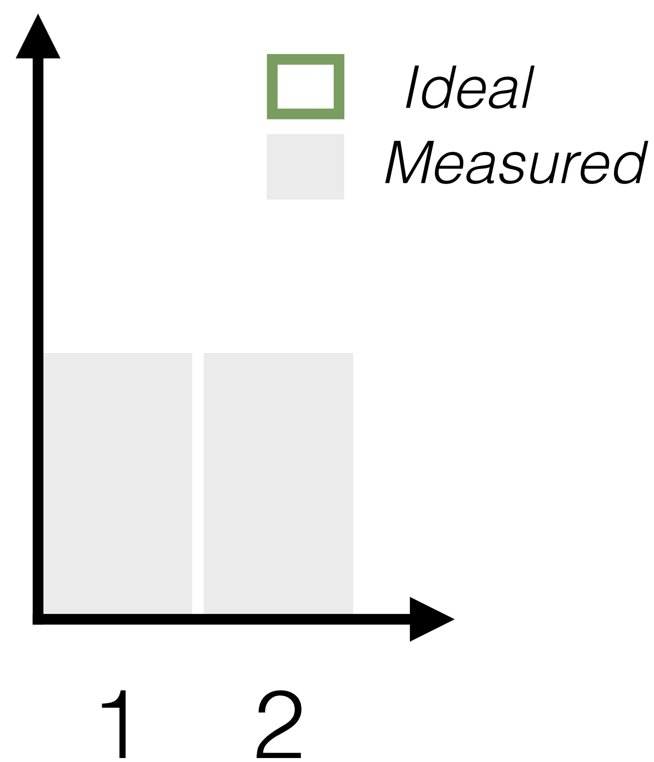


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

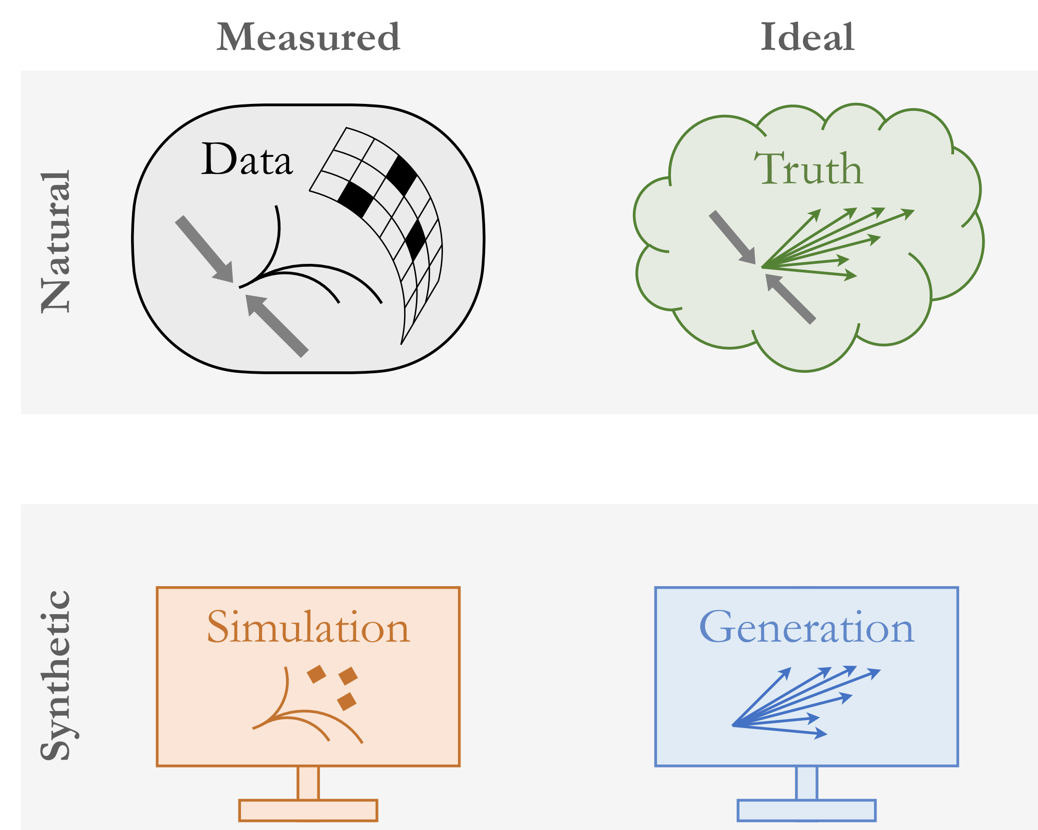
Unfold by iterating: OmniFold

After iteration 1





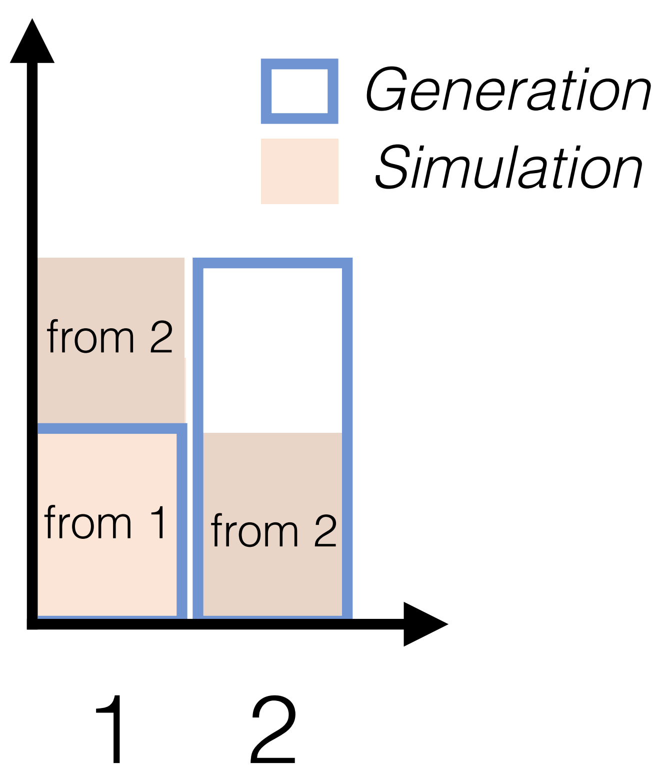
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

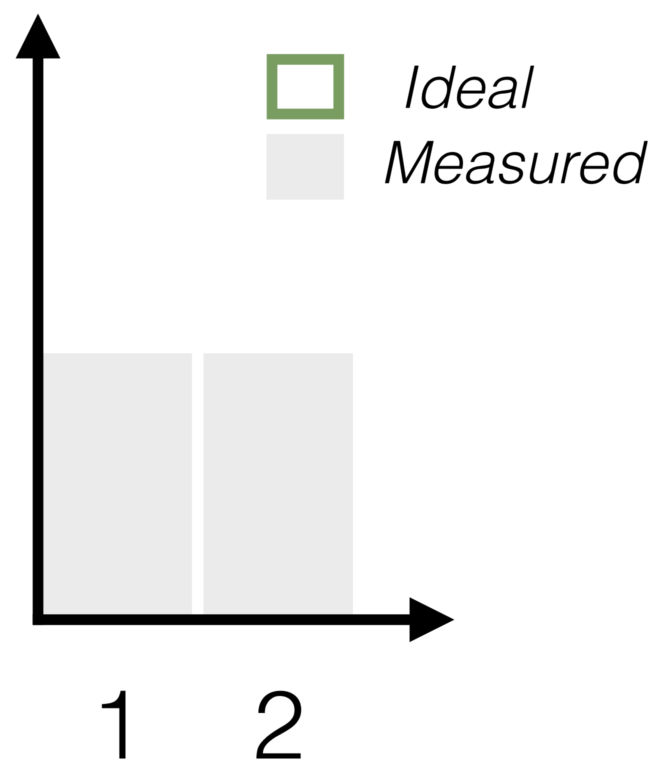


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

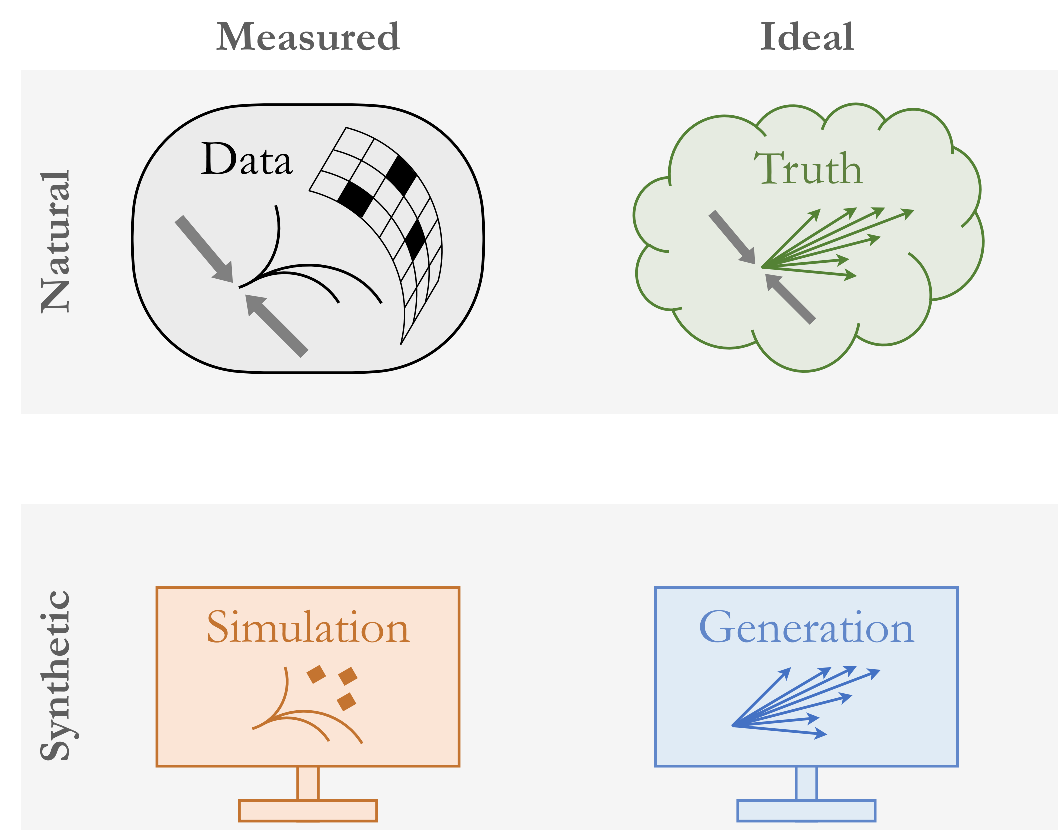
Unfold by iterating: OmniFold

After iteration 1





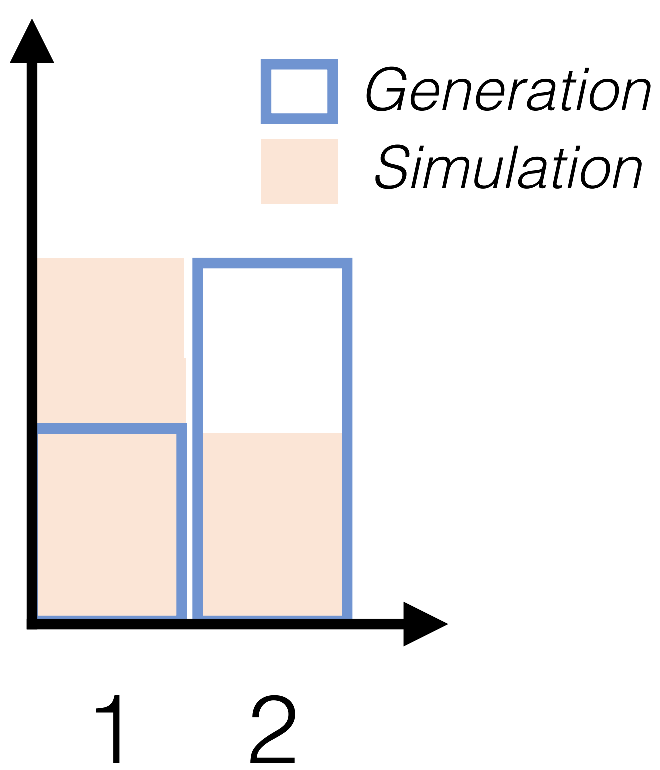
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

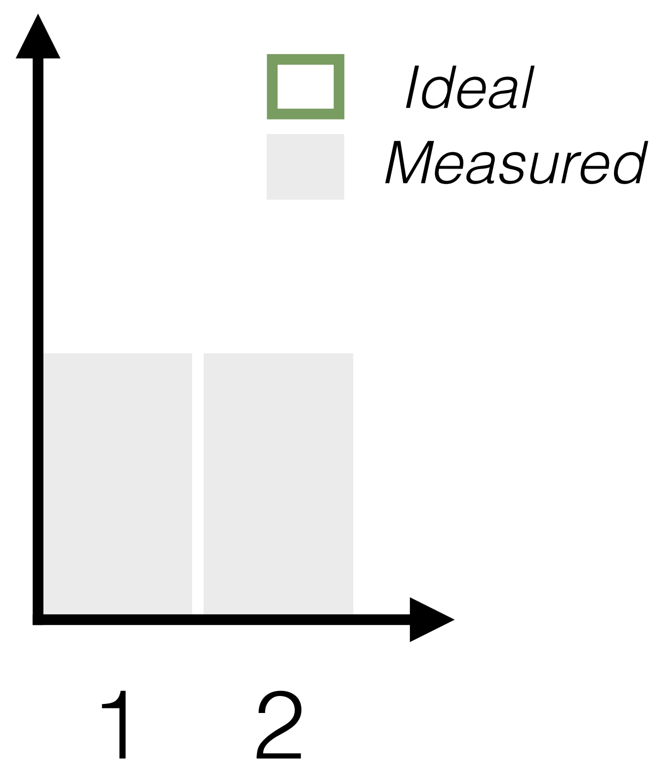


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

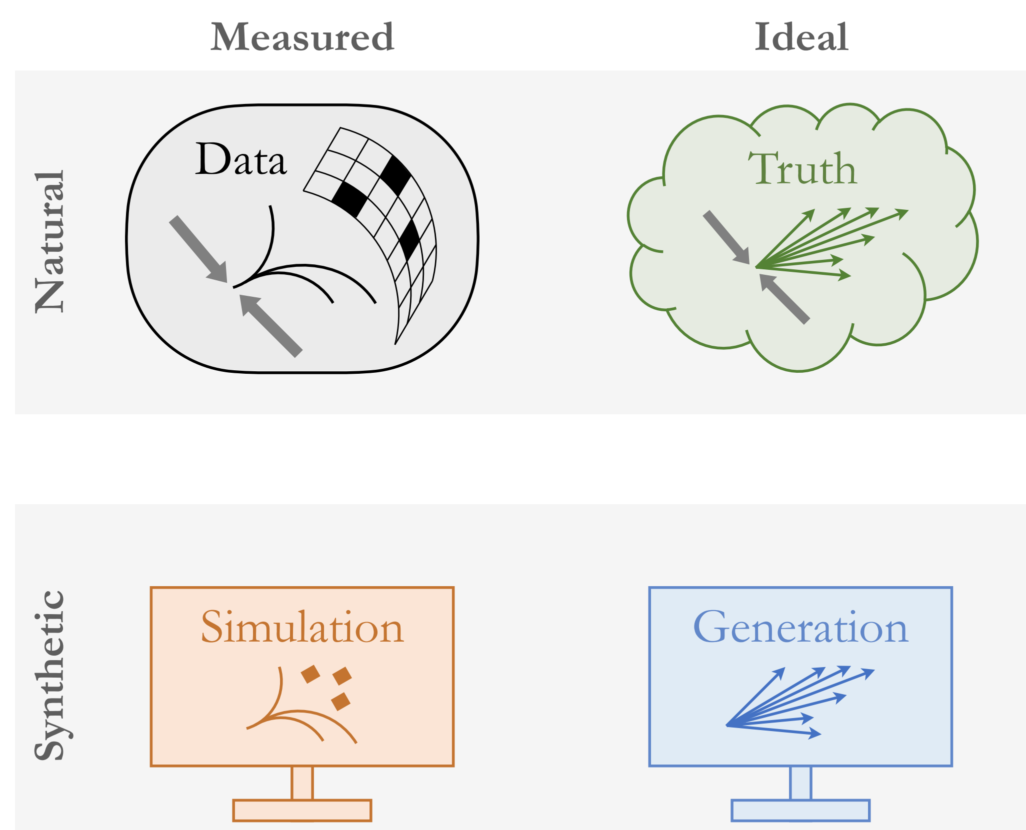
Unfold by iterating: OmniFold

After iteration 1





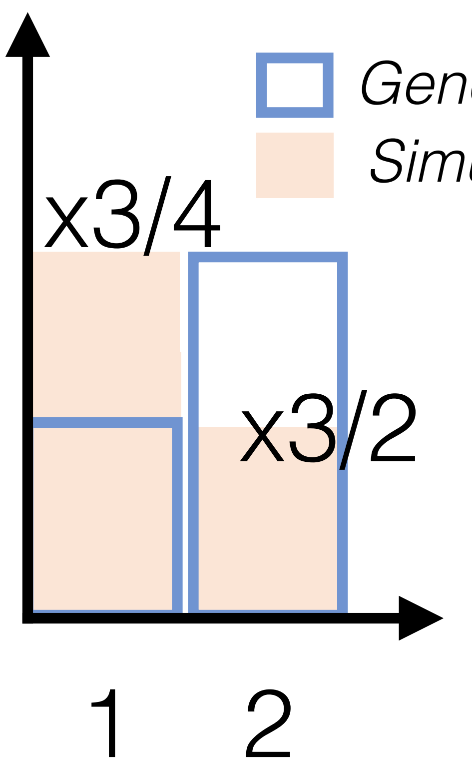
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

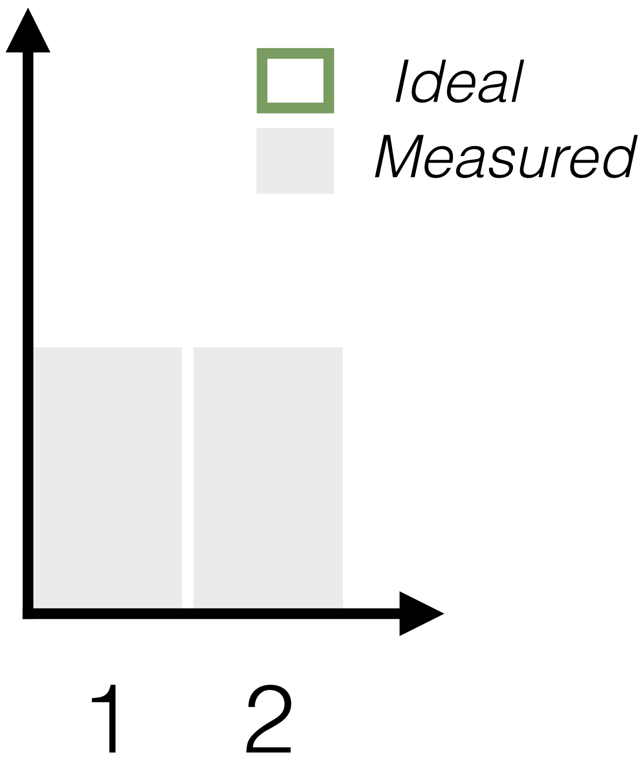


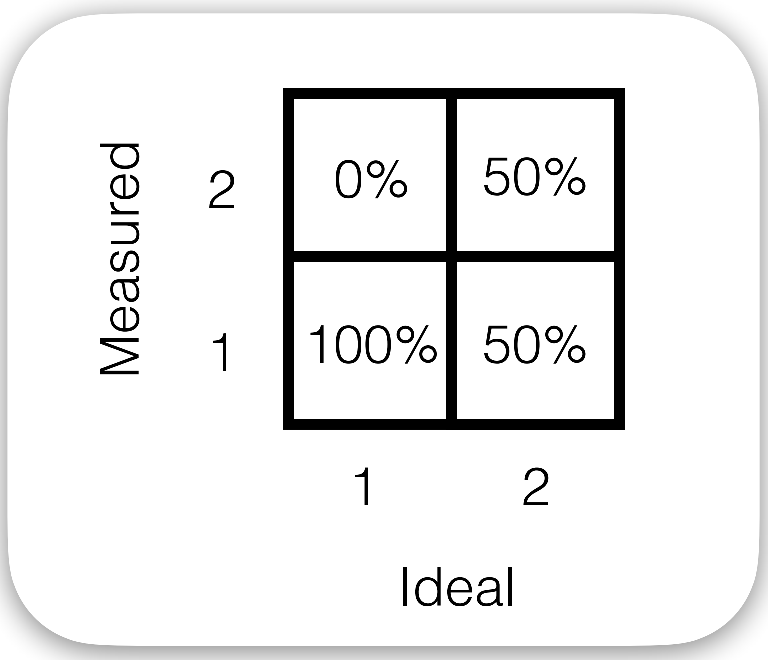
A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

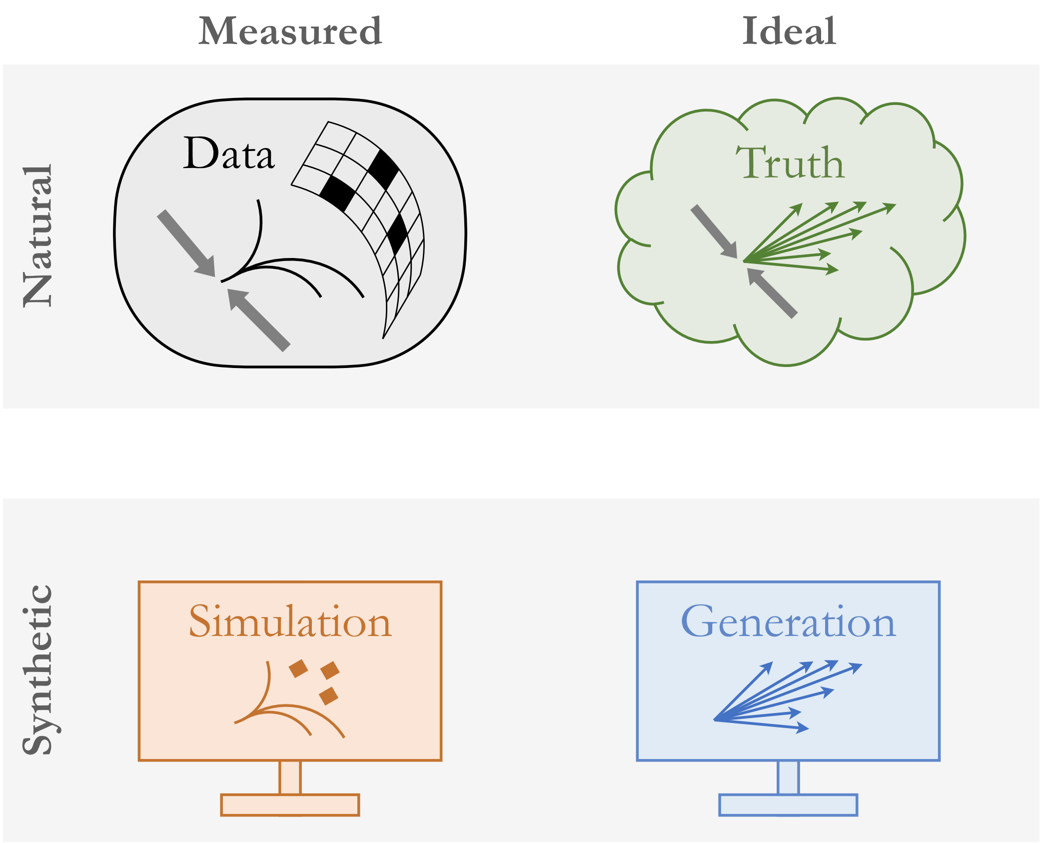
Unfold by iterating: OmniFold

After iteration 1





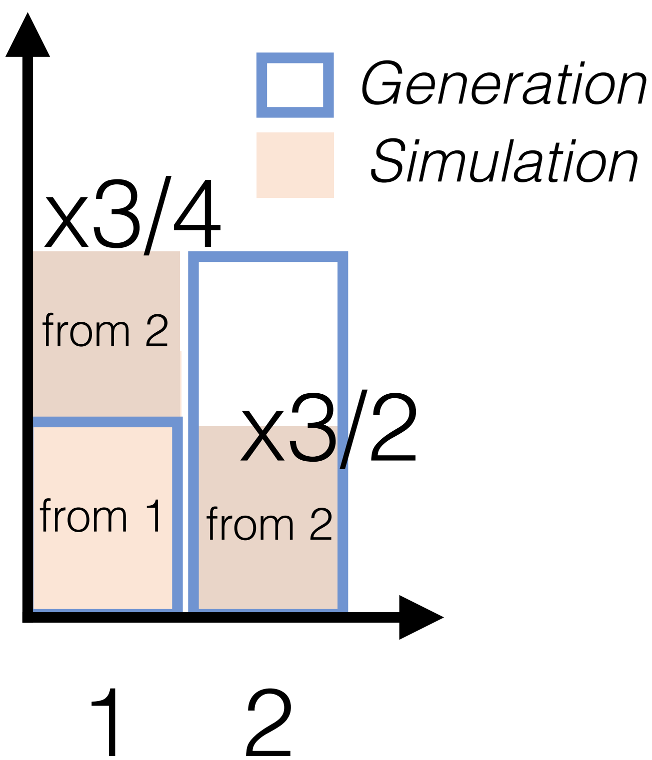


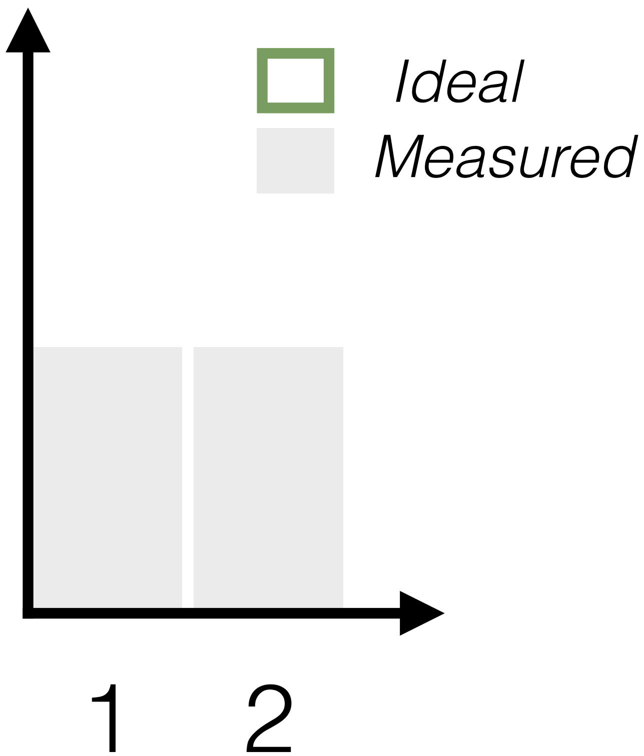


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

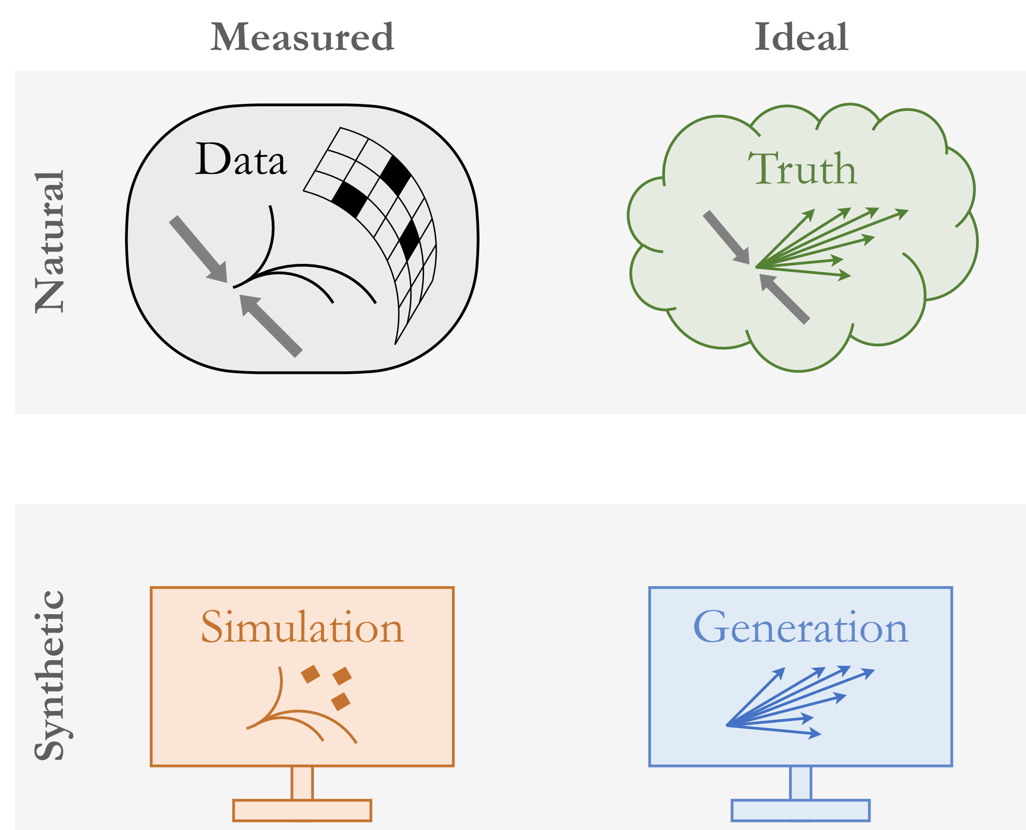
Unfold by iterating: OmniFold

After iteration 1





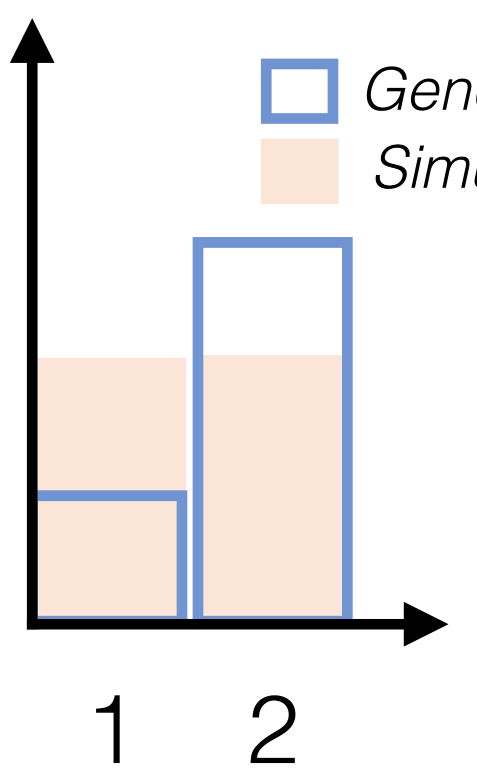
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

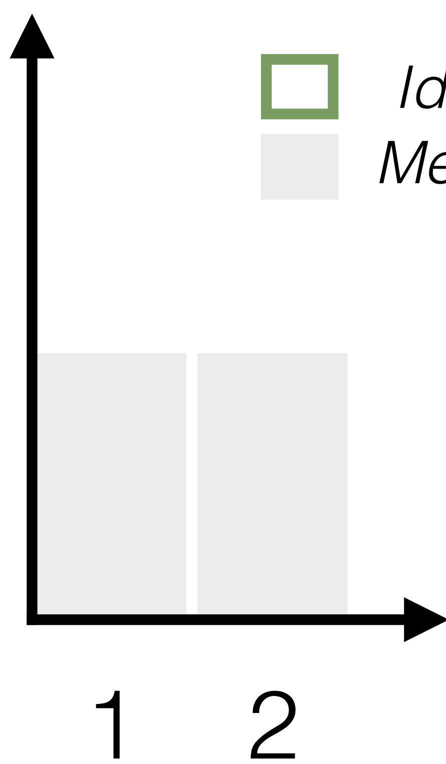


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

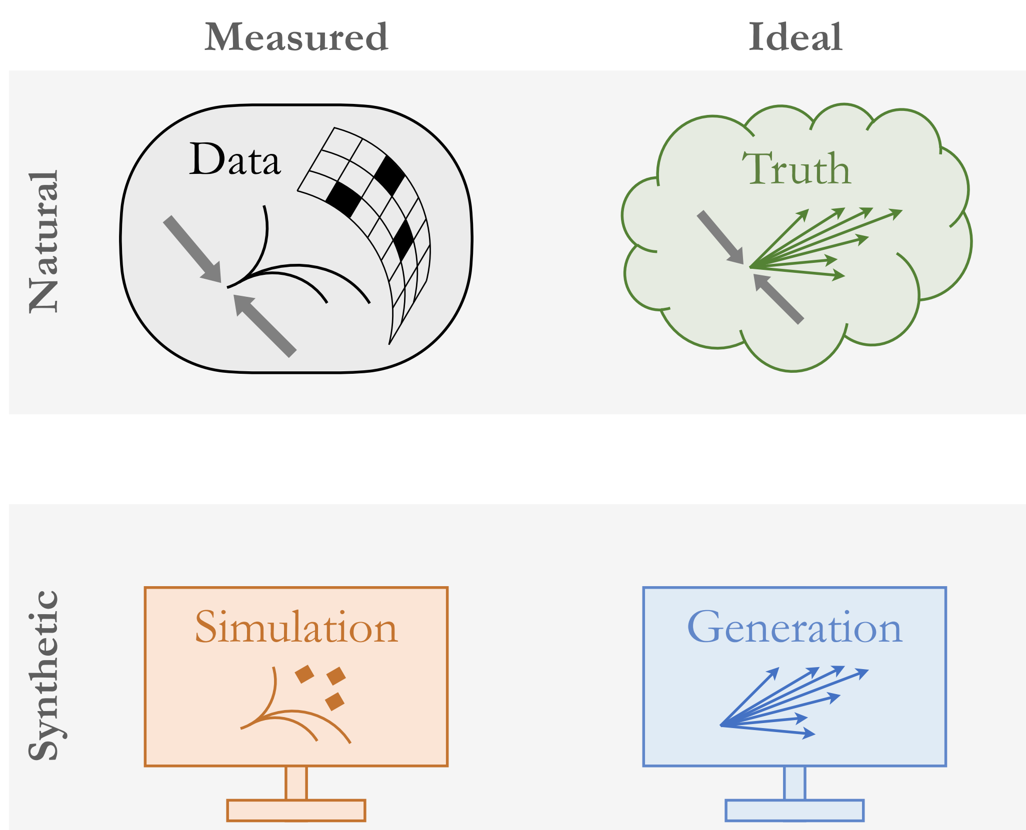
Unfold by iterating: OmniFold

After iteration 2





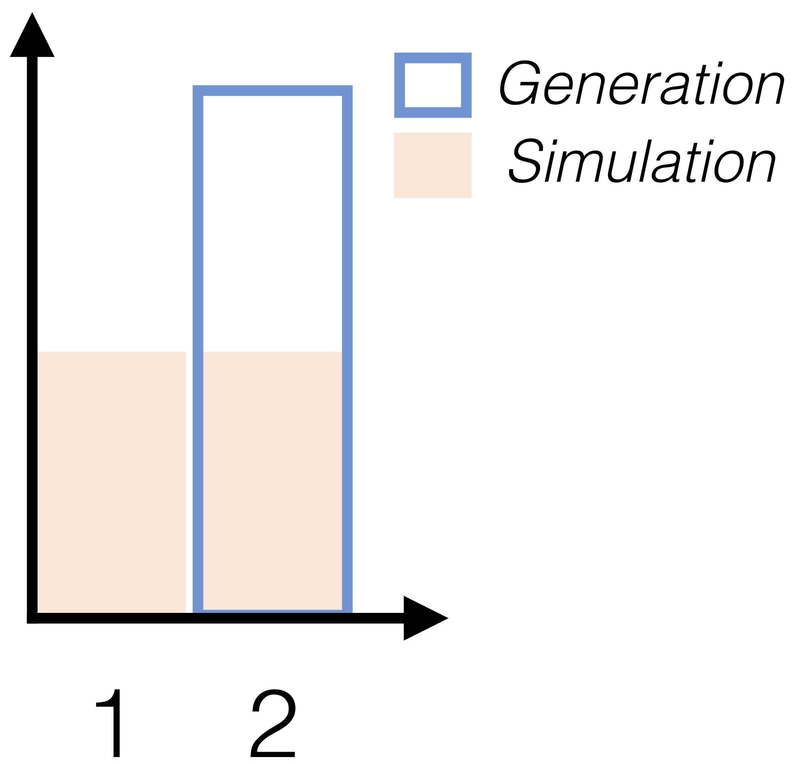
Measured \ Ideal	1	2
2	0%	50%
1	100%	50%

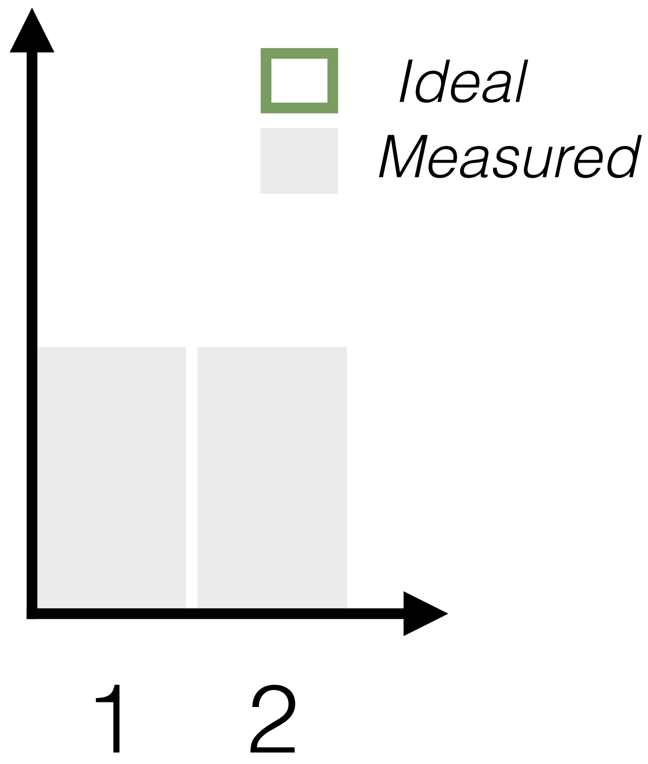


A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

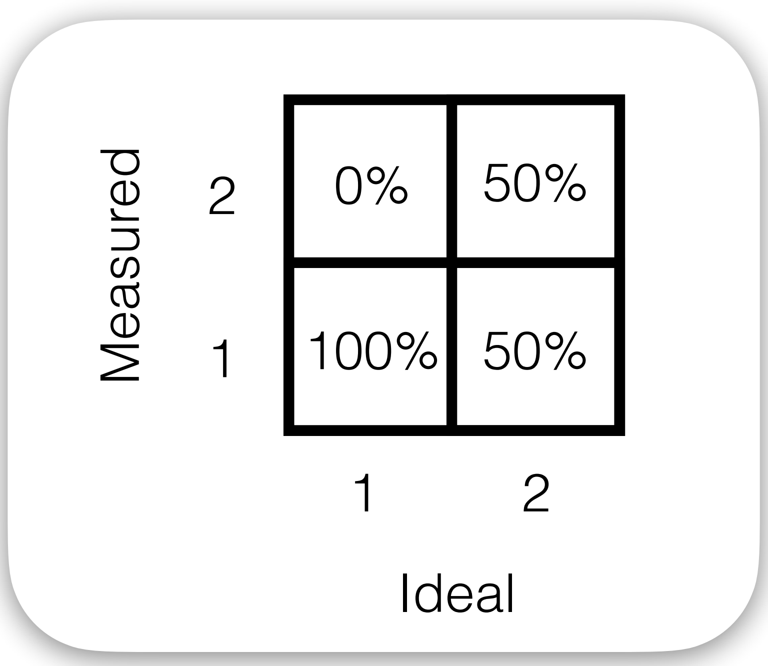
Unfold by iterating: OmniFold

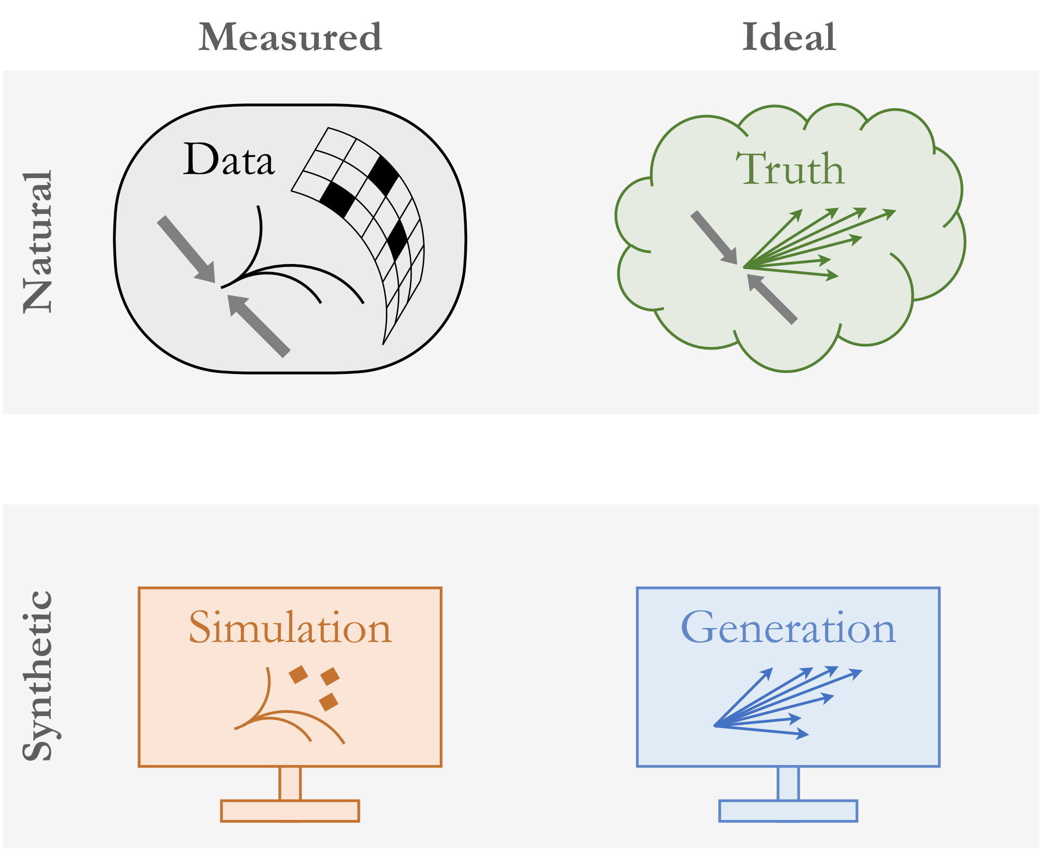
After iteration ∞





N.B. if you just apply p(ideal | measured), *you would have gotten the wrong answer!*

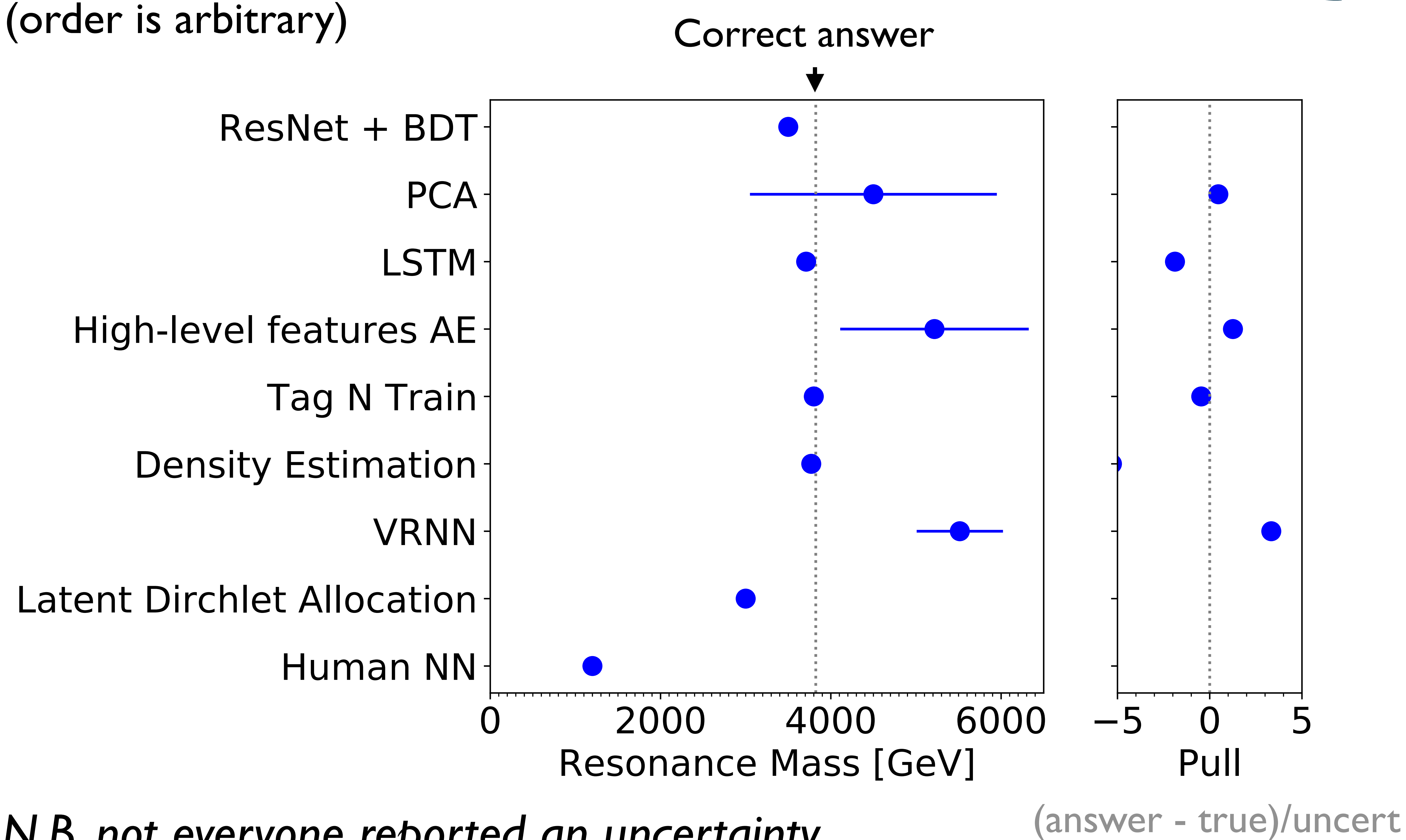




A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, PRL 124 (2020) 182001

Results - resonance mass

(order is arbitrary)



N.B. not everyone reported an uncertainty

Full phase space unfolding: OmniFold

Emily Dickinson, #975

The Mountain sat upon the Plain
In his tremendous Chair –
His observation omnifold,
His inquest, everywhere –

The Seasons played around his knees
Like Children round a sire –
Grandfather of the Days is He
Of Dawn, the Ancestor –

A. Andreassen, E. Metodiev, P. Komiske, **BPN**, J. Thaler, 1911.09107

Parameter estimation

Mean and standard deviation over 20 runs:

	Parameter	Target value	Fit value
Val.	TimeShower:alphaSvalue	0.1200	0.1195 ± 0.0022
	StringZ:aLund	0.6000	0.6276 ± 0.0373
	StringFlav:probStoUD	0.1200	0.1203 ± 0.0071
Blinded	TimeShower:alphaSvalue	0.1700	0.1707 ± 0.0022
	StringZ:aLund	0.7500	0.7425 ± 0.0453
	StringFlav:probStoUD	0.1400	0.1422 ± 0.0065

Similar spread

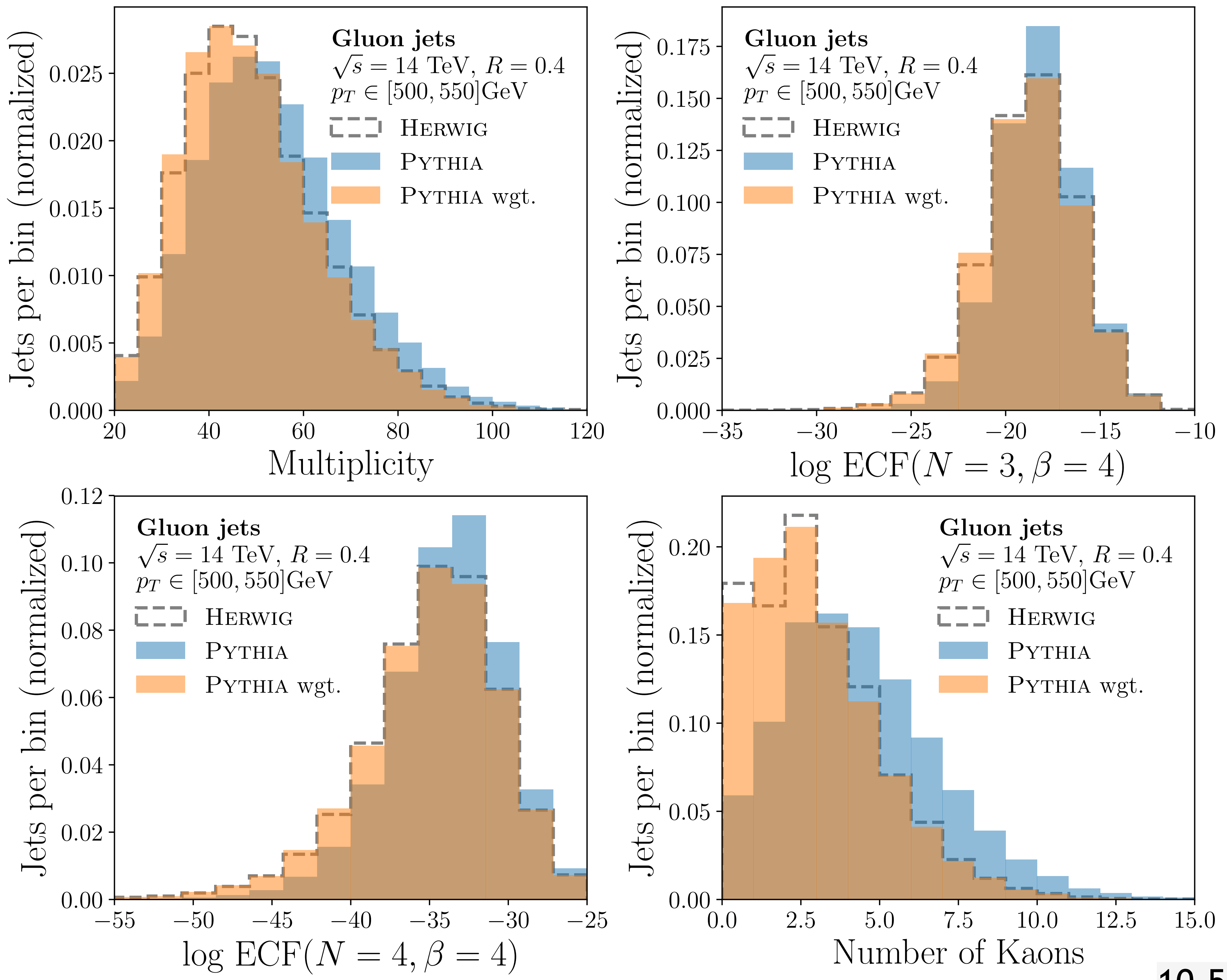
1D:

Parameter	Target value	Fit value
TimeShower:alphaSvalue	0.1600	0.1601 ± 0.0018
StringZ:aLund	0.8000	0.7980 ± 0.0257
StringFlav:probStoUD	0.2750	0.2754 ± 0.0065

The meaning of this “uncertainty” is discussed later.

Pythia versus Herwig

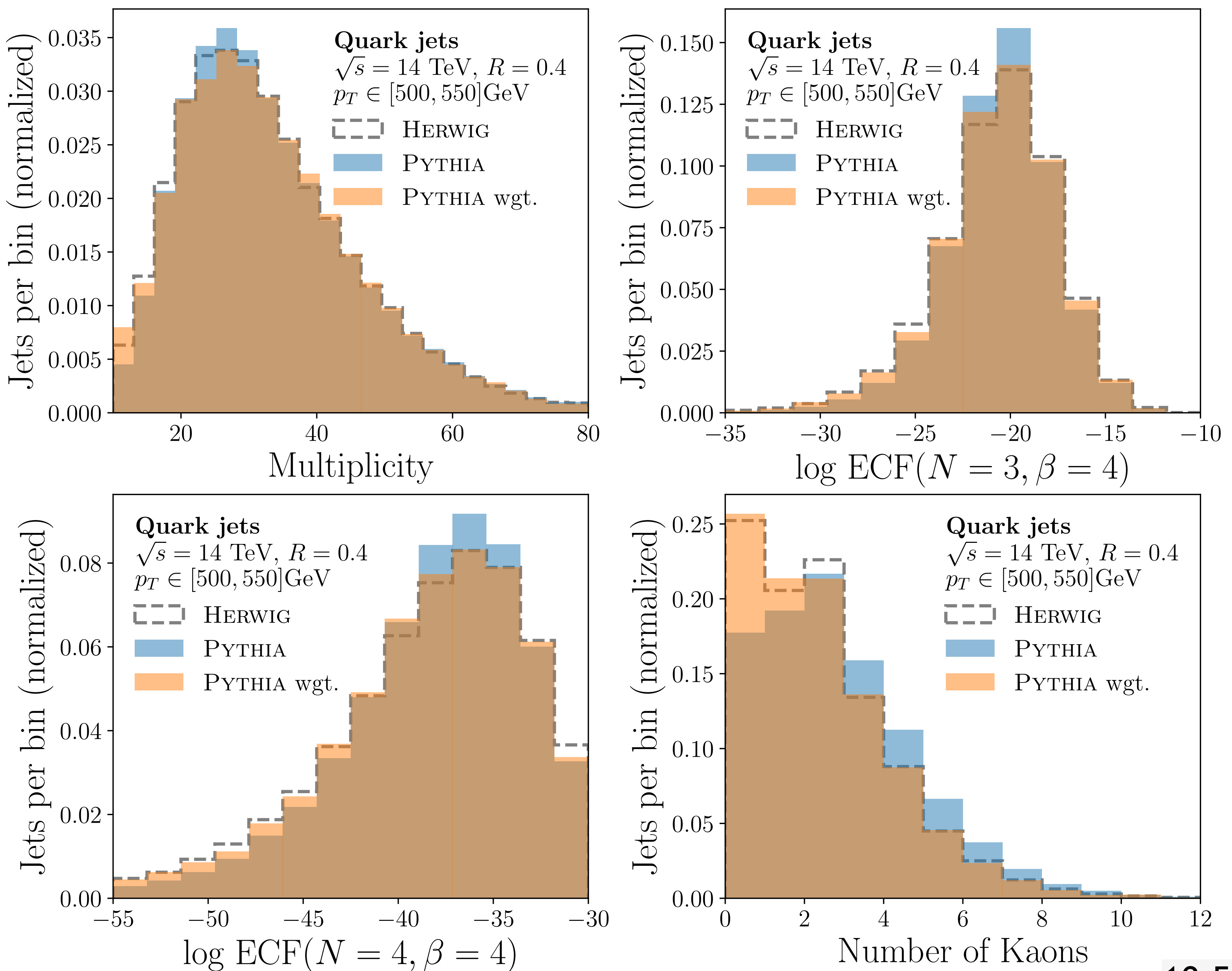
No hyper-parameter tuning - out of the box!



Samples from
10.5281/zenodo.2658764
10.5281/zenodo.3164691

Pythia versus Herwig

No hyper-parameter tuning - out of the box!



Samples from
10.5281/zenodo.2658764
10.5281/zenodo.3164691