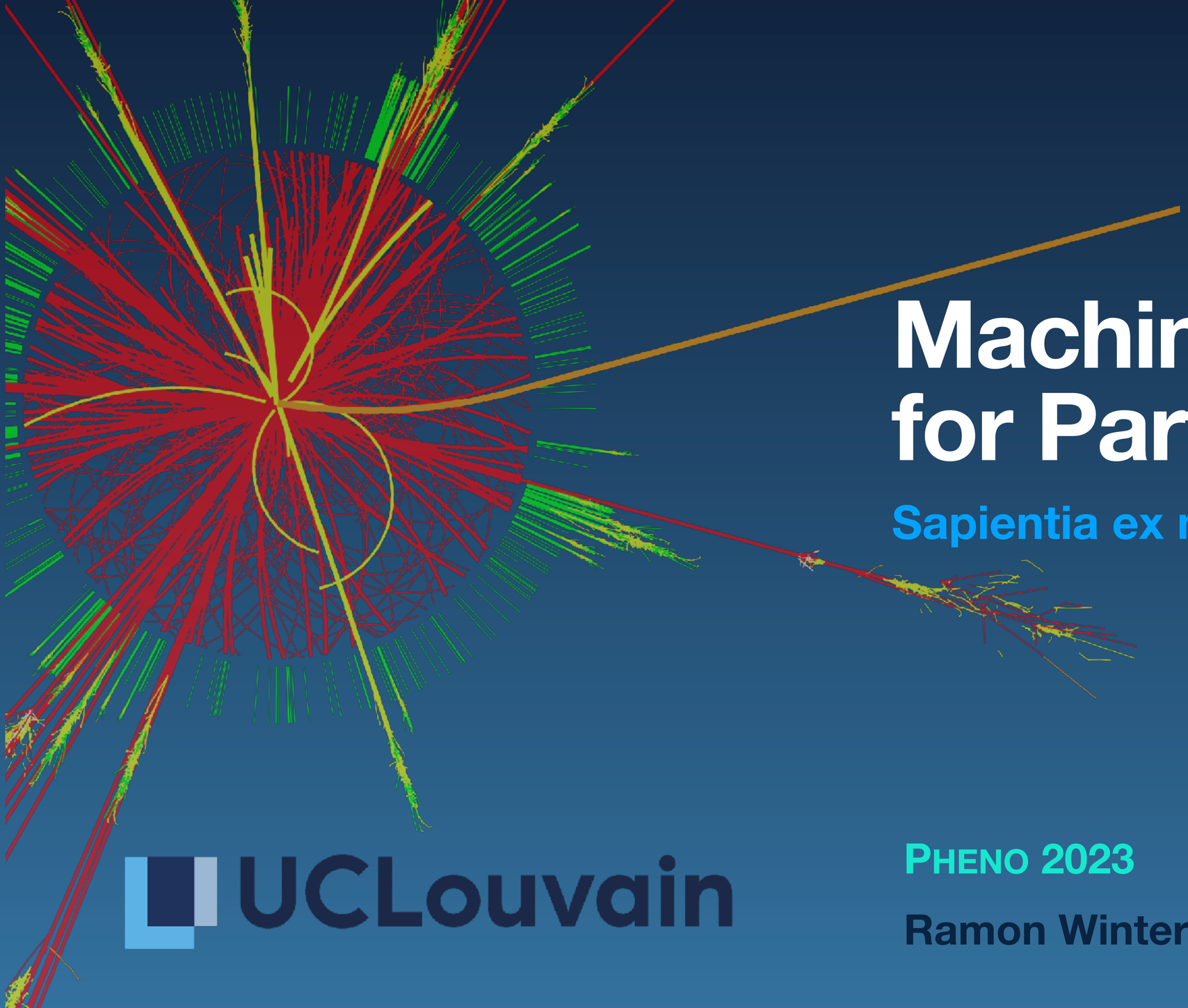


Machine Learning for Particle Physics

Sapientia ex machina?

UCLouvain

PHENO 2023

Ramon Winterhalder — UC Louvain

Why does Pheno care about ML?

Why does Pheno care about ML?

> Pheno cares about machine learning because it can improve data analysis, simulation and modeling, lead to new discoveries, and foster cross-disciplinary collaboration

ChatGPT

Why ML in HEP?

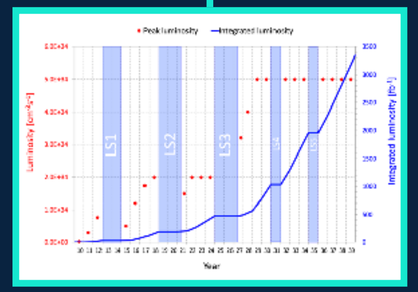
Data volume

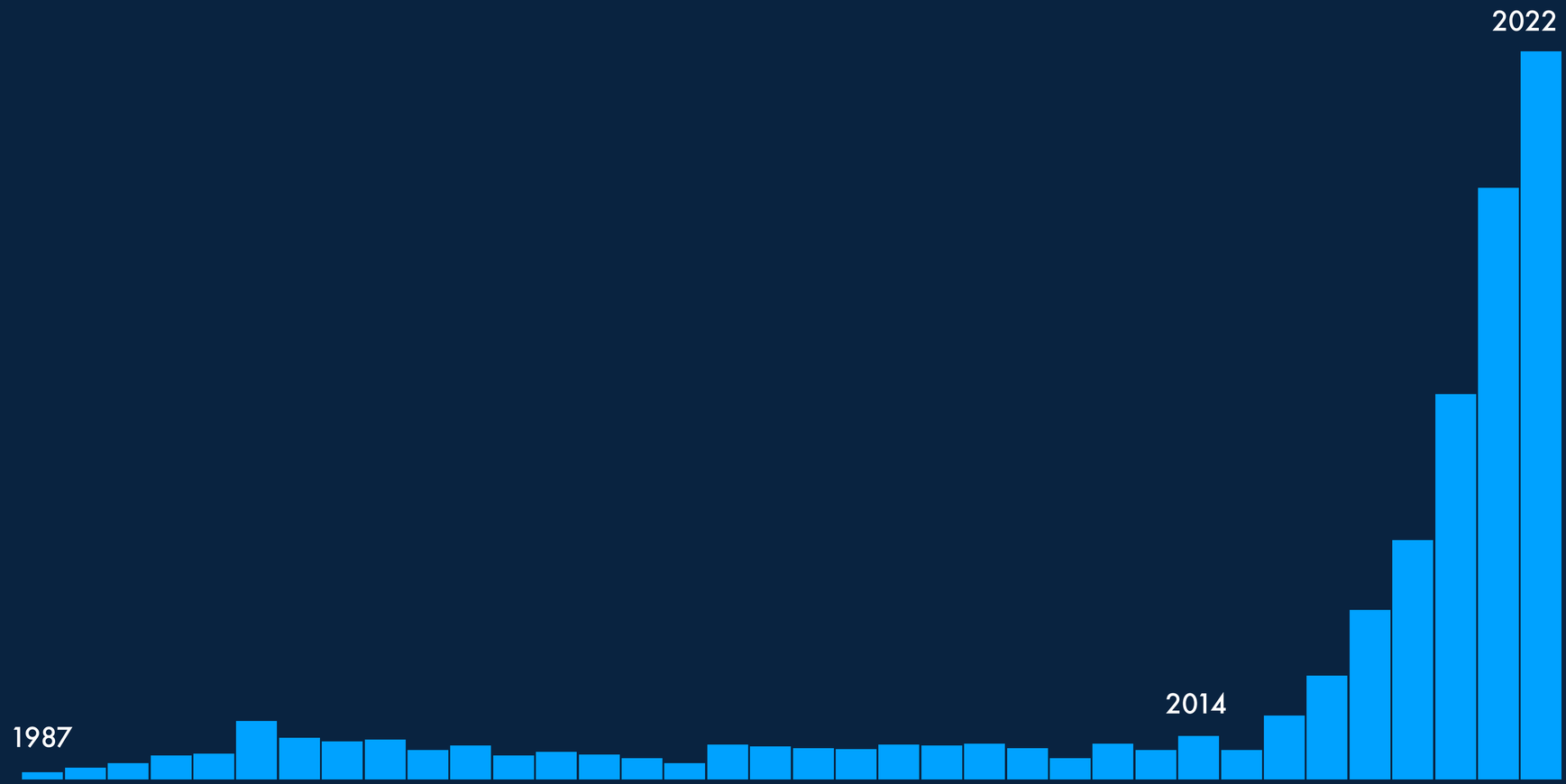
Large amounts of data

1. labeled (Simulation)
2. unlabeled (Detector)

ML wants lots of data

1





Why ML in HEP?

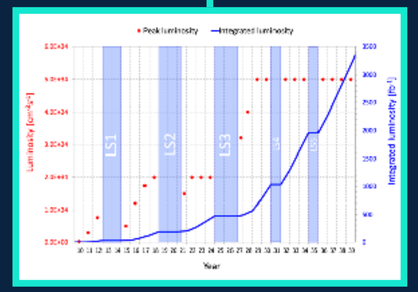
Data volume

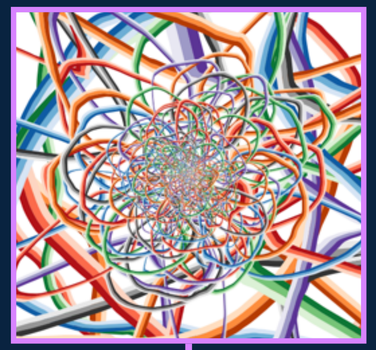
Large amounts of data

1. labeled (Simulation)
2. unlabeled (Detector)

ML wants lots of data

1



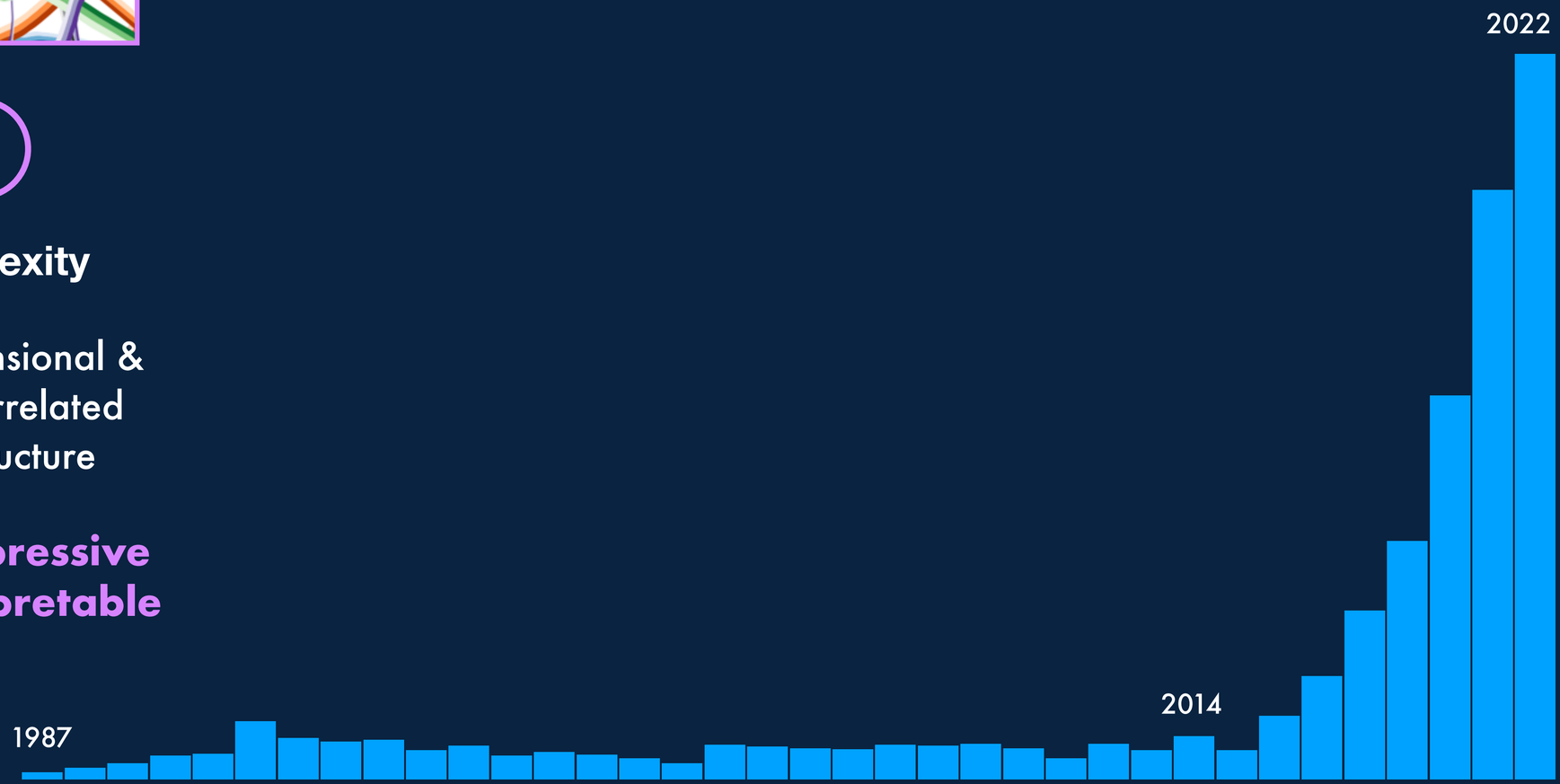


2

Complexity

High-dimensional & highly correlated data structure

ML is expressive and interpretable



Why ML in HEP?

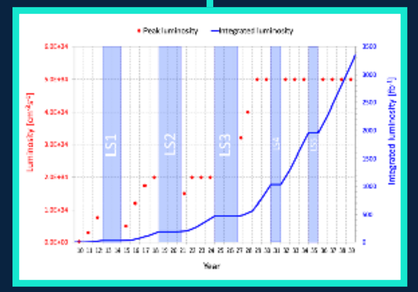
Data volume

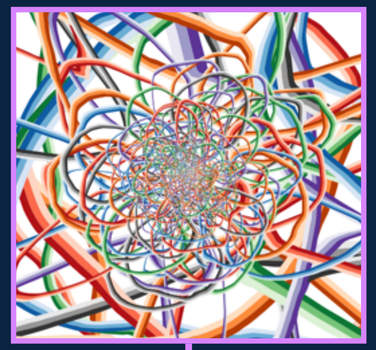
Large amounts of data

1. labeled (Simulation)
2. unlabeled (Detector)

ML wants lots of data

1





2

Complexity

High-dimensional & highly correlated data structure

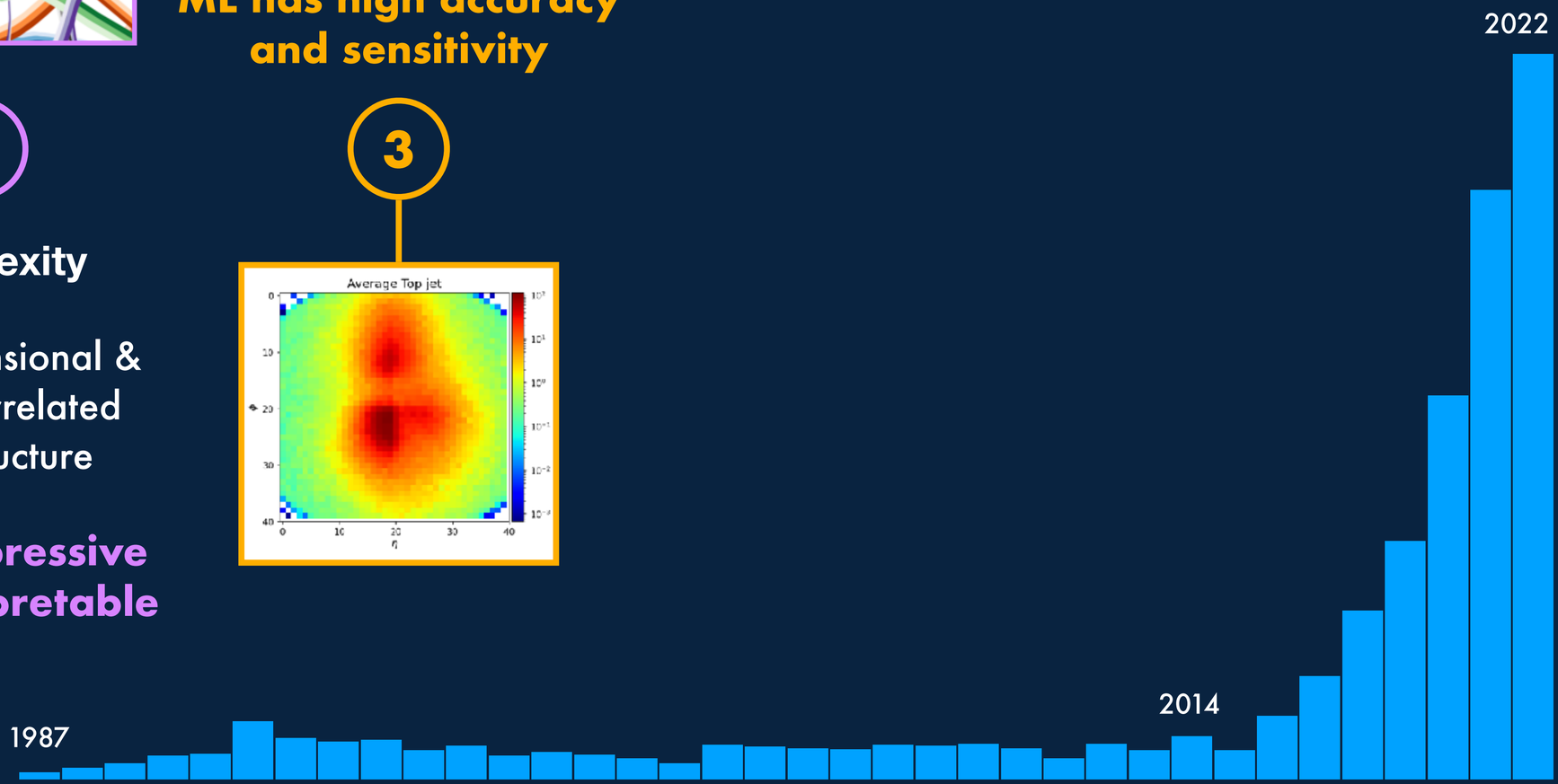
ML is expressive and interpretable

Signal detection

Rare and elusive signals among large backgrounds

ML has high accuracy and sensitivity

3



Why ML in HEP?

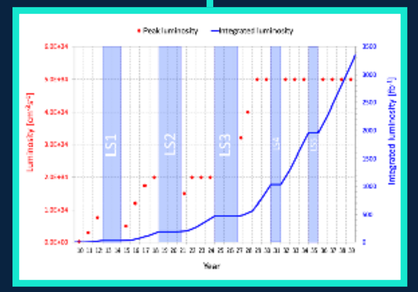
Data volume

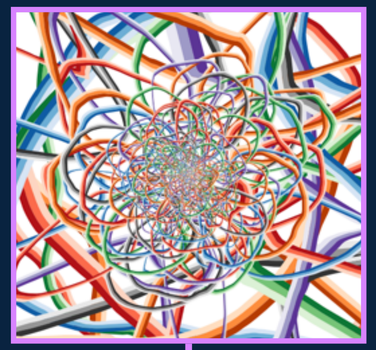
Large amounts of data

1. labeled (Simulation)
2. unlabeled (Detector)

ML wants lots of data

1





2

Complexity

High-dimensional & highly correlated data structure

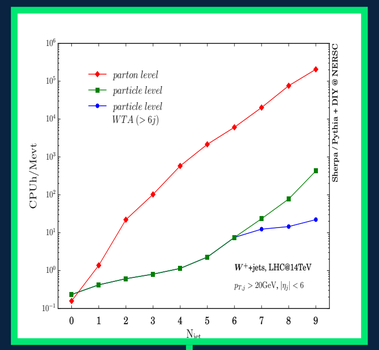
ML is expressive and interpretable

Signal detection

Rare and elusive signals among large backgrounds

ML has high accuracy and sensitivity

3

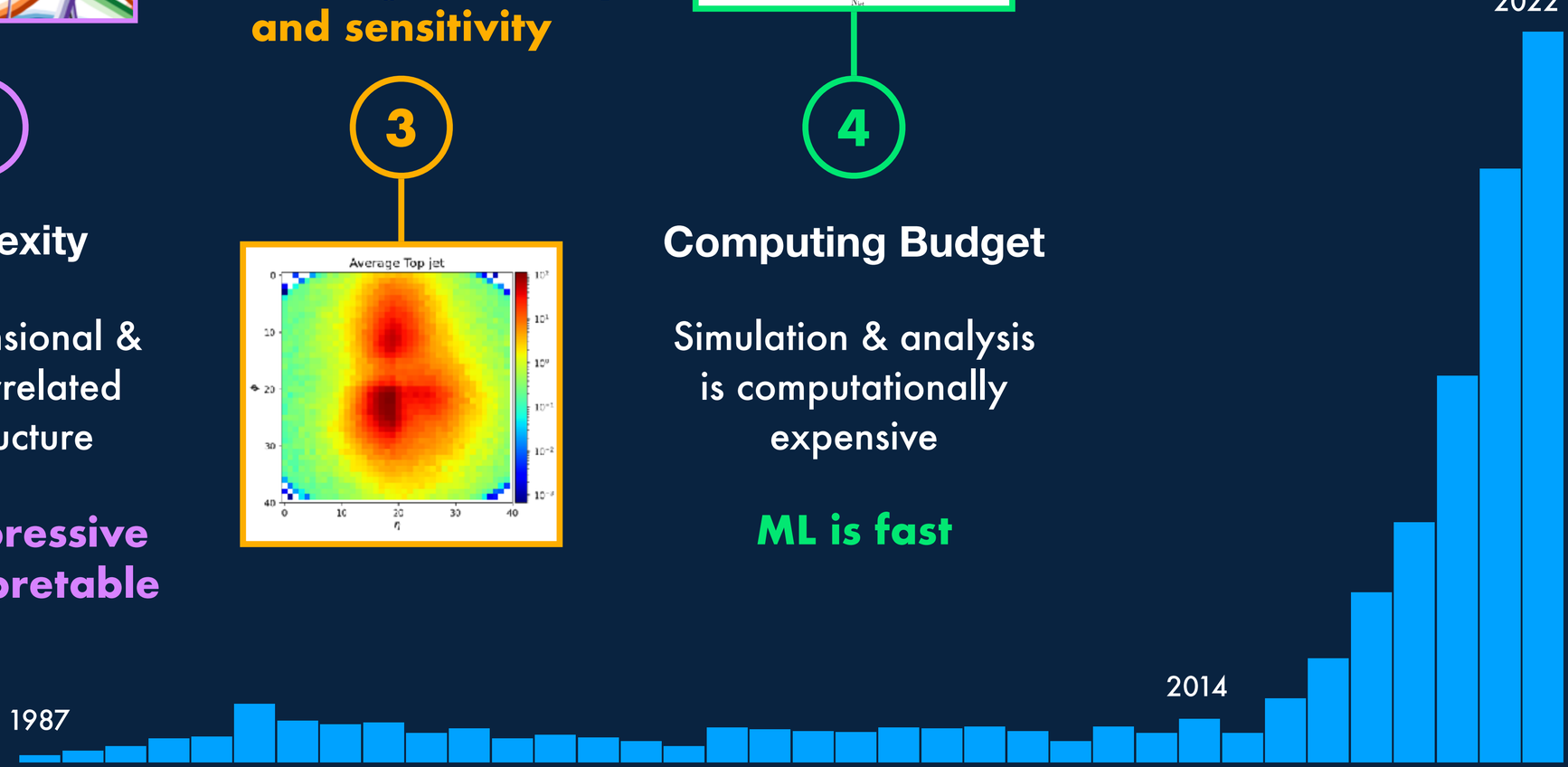


4

Computing Budget

Simulation & analysis is computationally expensive

ML is fast



Why ML in HEP?

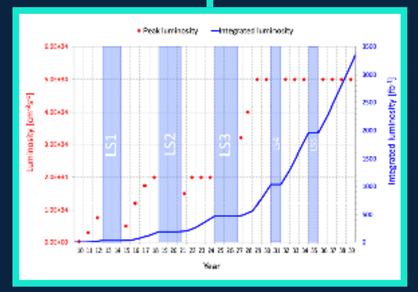
Data volume

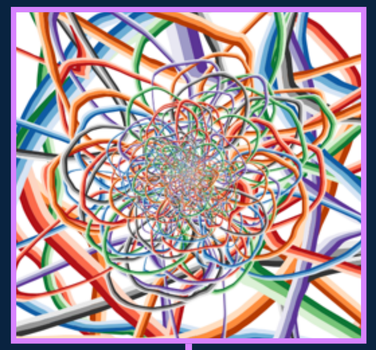
Large amounts of data

1. labeled (Simulation)
2. unlabeled (Detector)

ML wants lots of data

1





2

Complexity

High-dimensional & highly correlated data structure

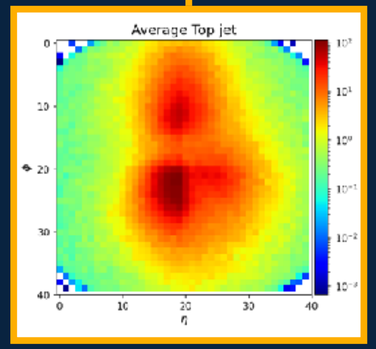
ML is expressive and interpretable

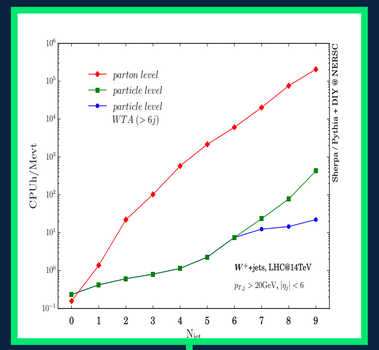
Signal detection

Rare and elusive signals among large backgrounds

ML has high accuracy and sensitivity

3





4

Computing Budget

Simulation & analysis is computationally expensive

ML is fast

Increasing interest

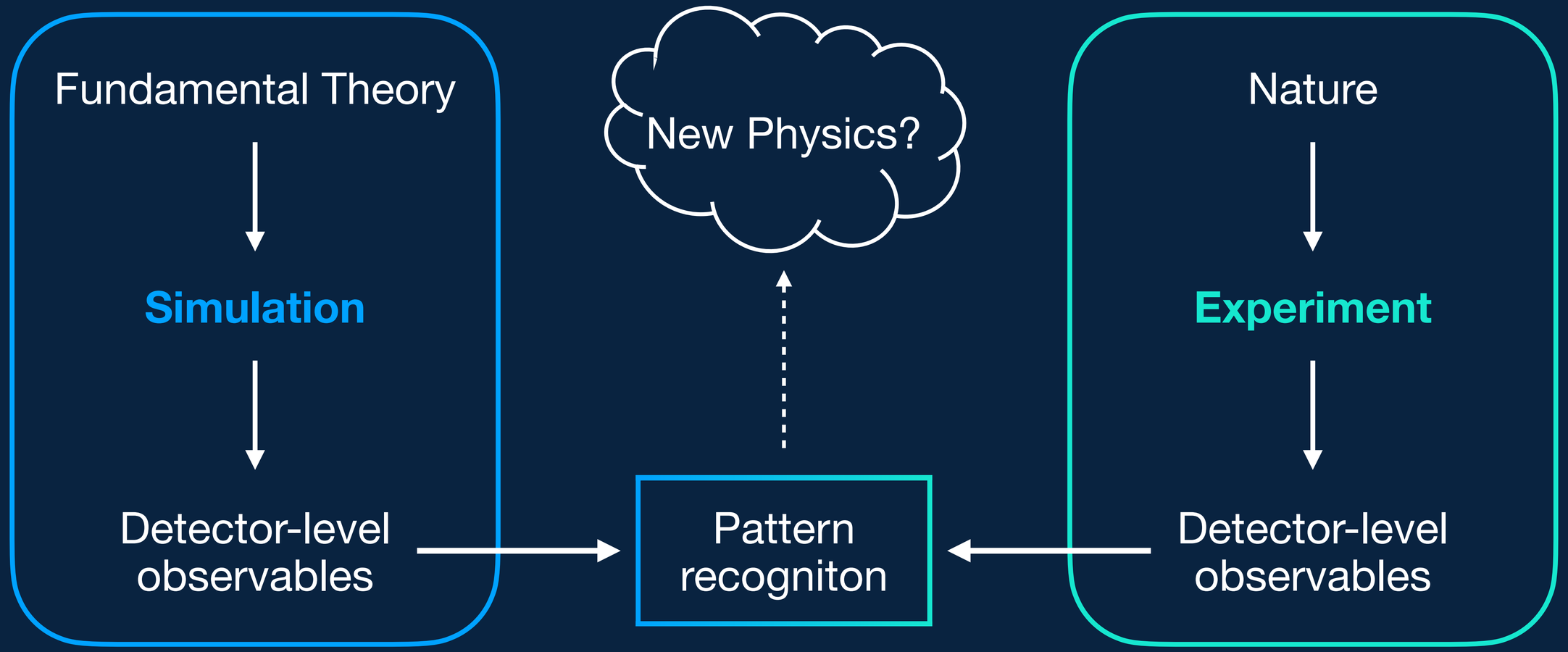
> 150 paper/year

Future of HEP?

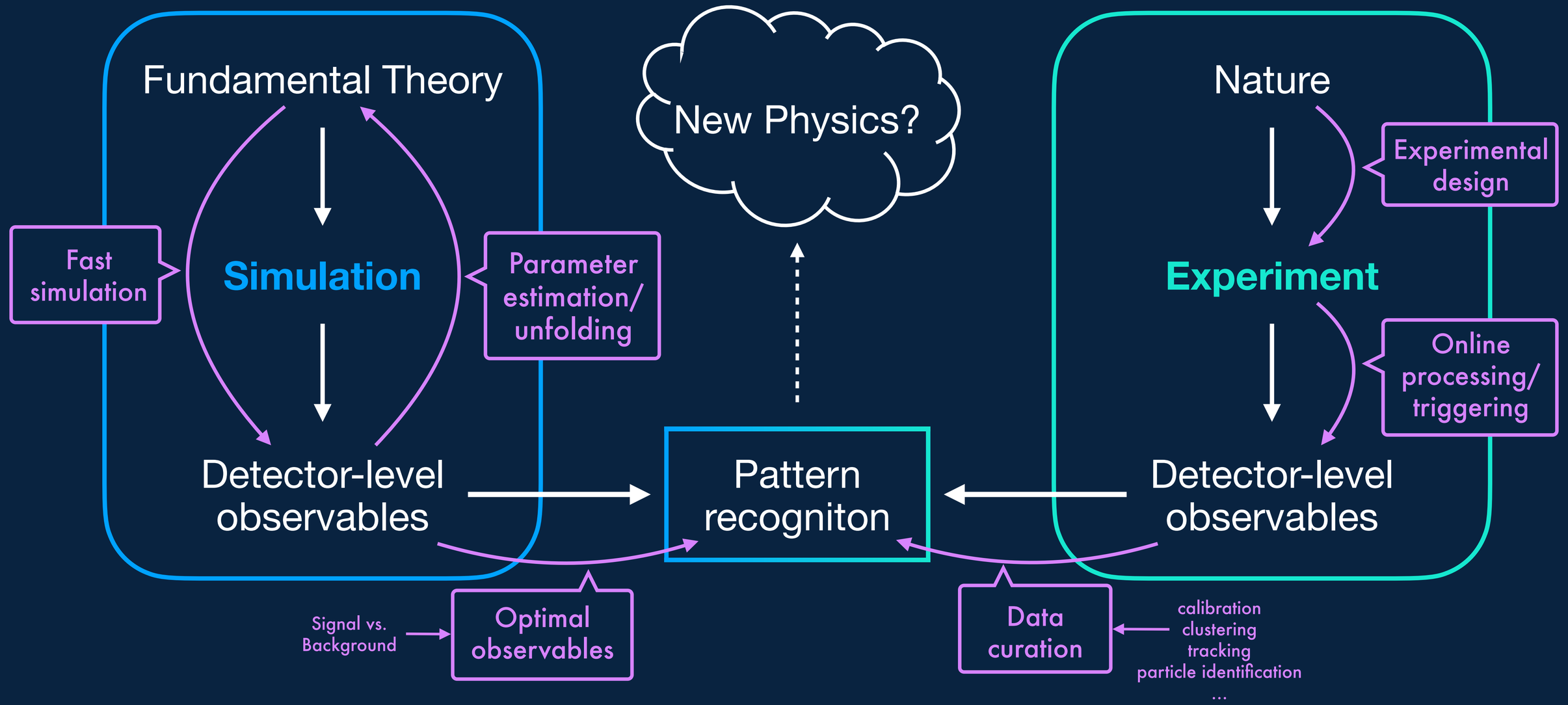
ML is fun!

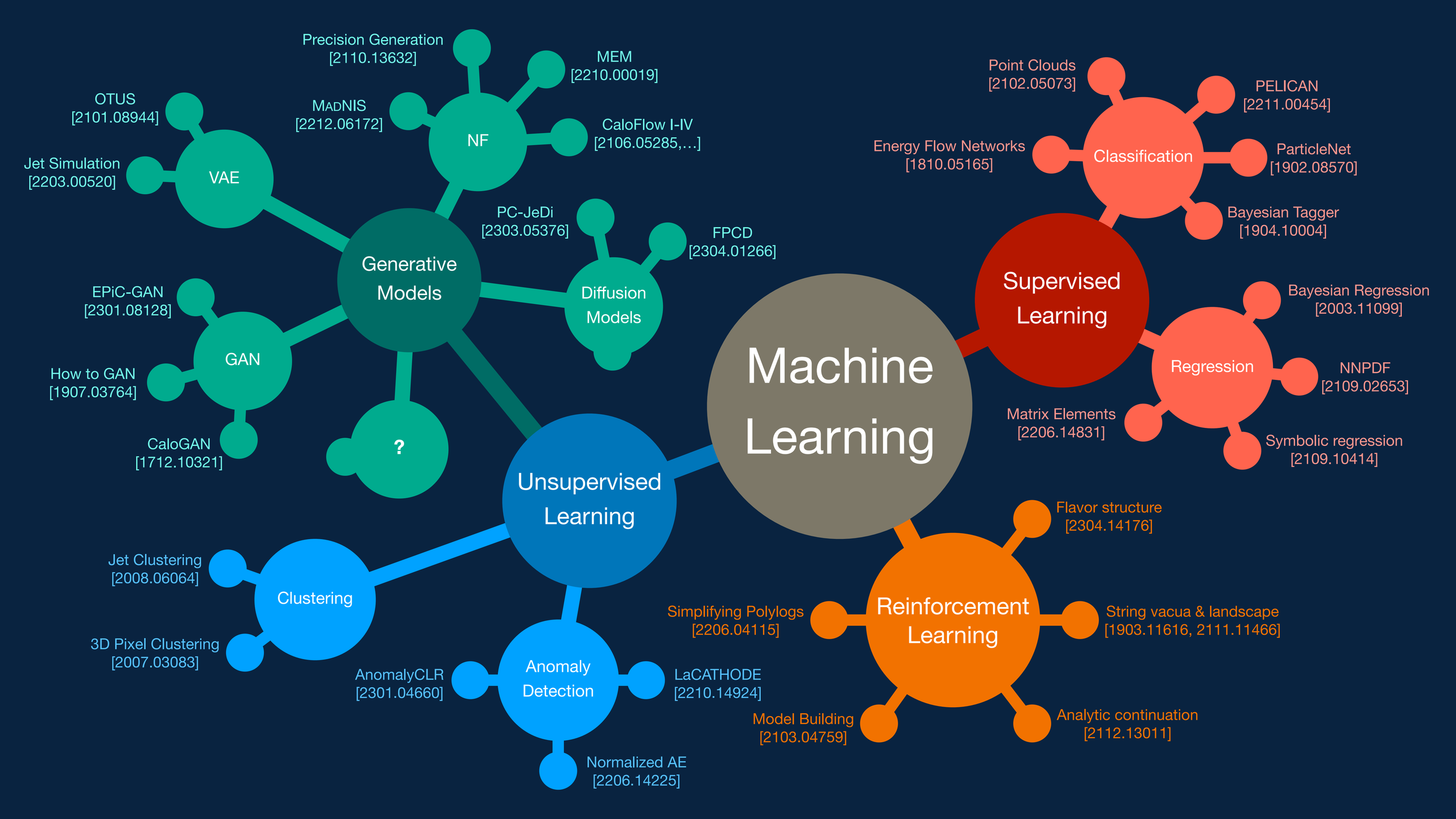
5

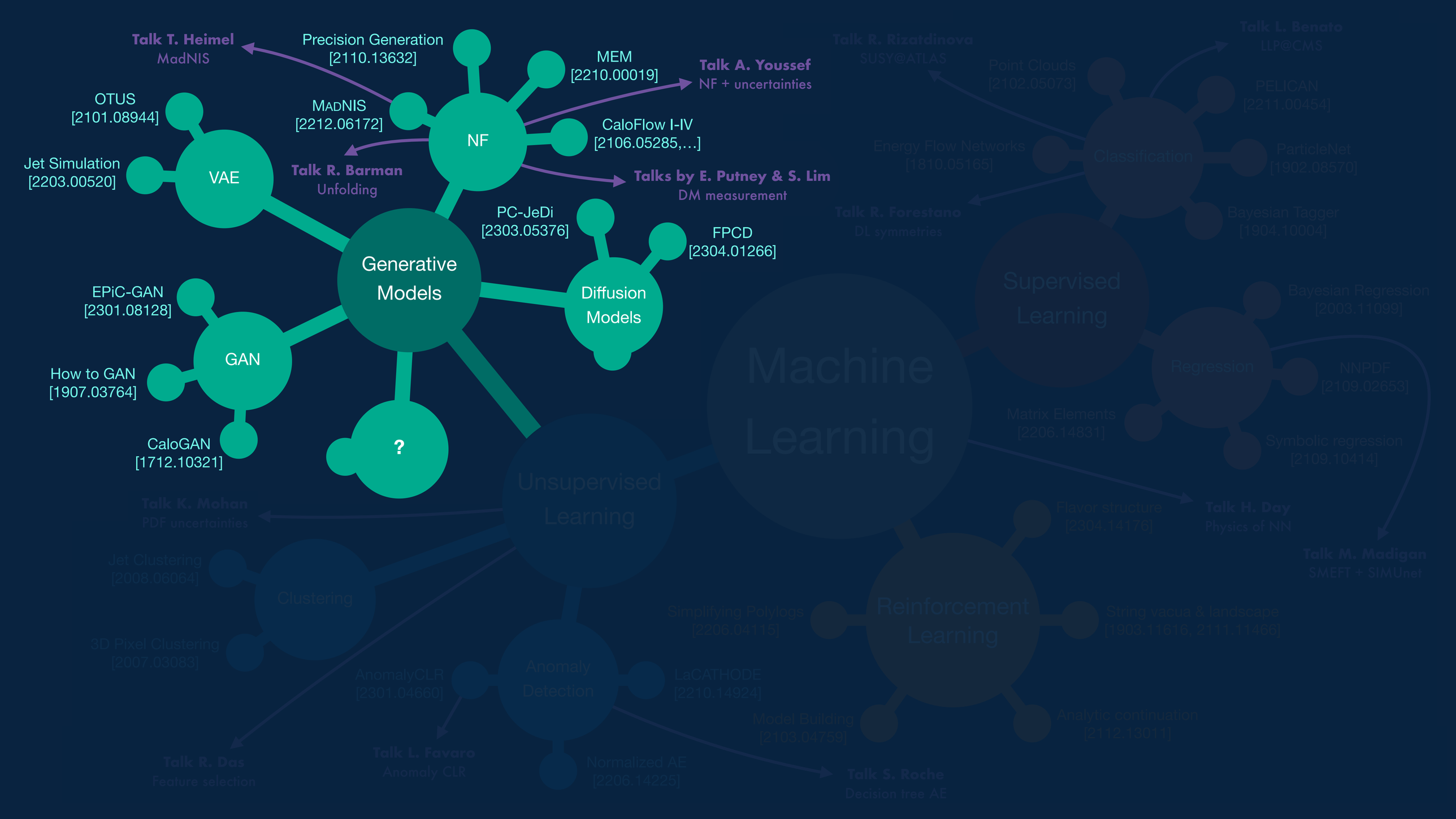
LHC analysis (oversimplified)

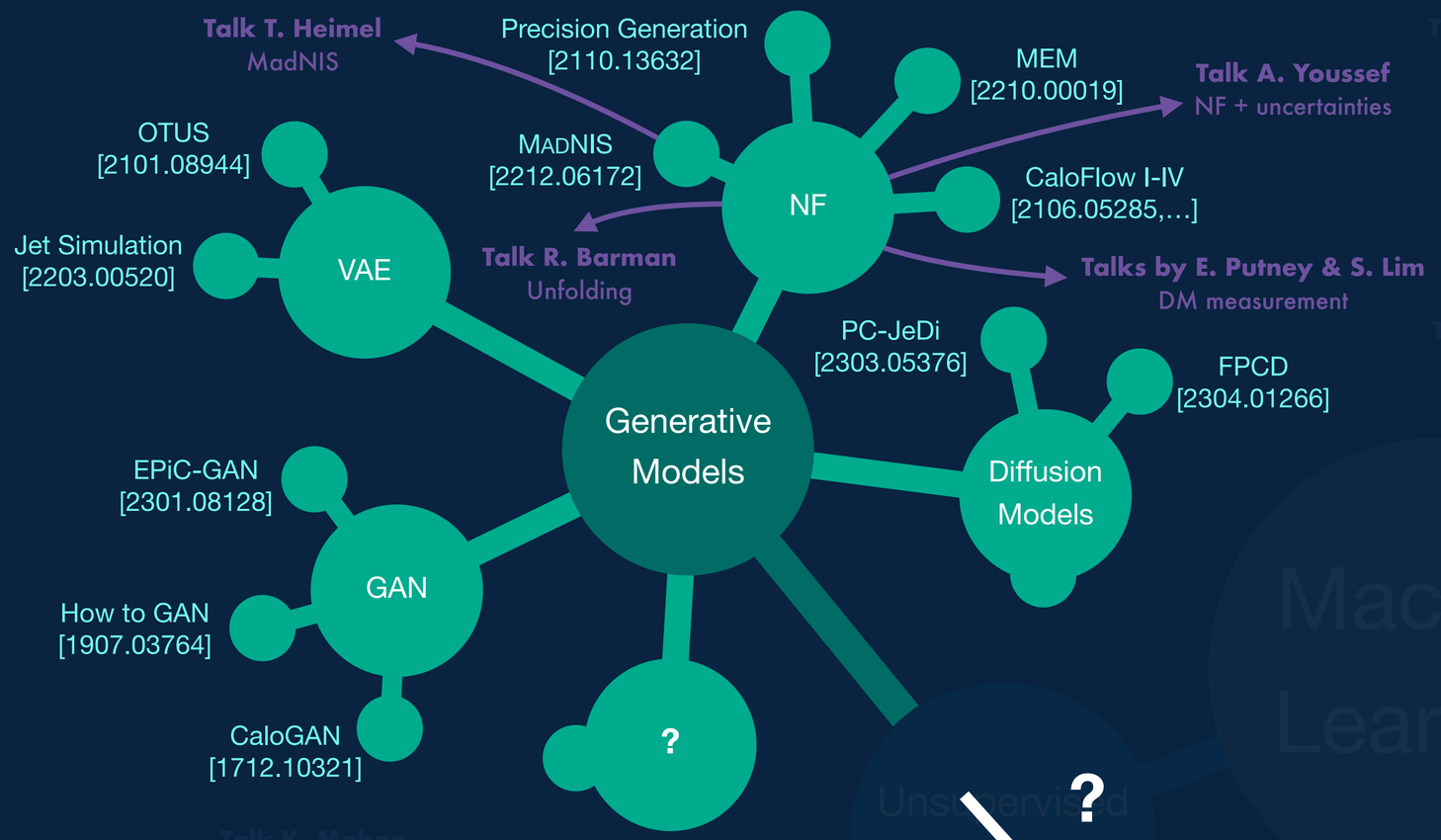


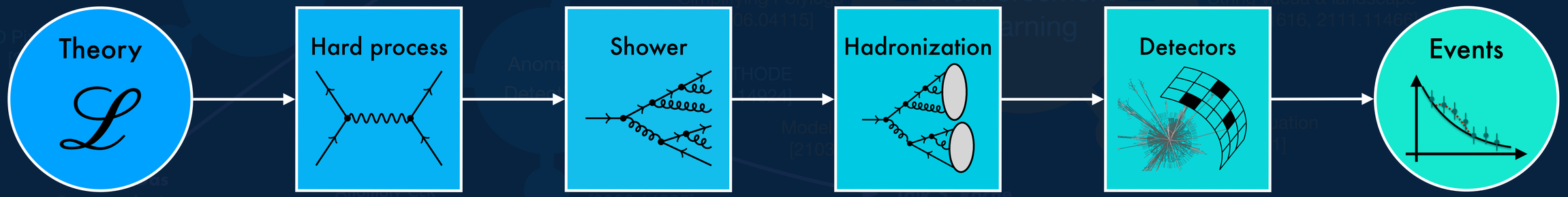
LHC analysis + ML




Machine Learning
Generative Models
VAE
OTUS [2101.08944]
Jet Simulation [2203.00520]
NF
Precision Generation [2110.13632]
MADNIS [2212.06172]
MEM [2210.00019]
CaloFlow I-IV [2106.05285,…]
GAN
EPiC-GAN [2301.08128]
How to GAN [1907.03764]
CaloGAN [1712.10321]
?
Diffusion Models
PC-JeDi [2303.05376]
FPCD [2304.01266]
Supervised Learning
Classification
Point Clouds [2102.05073]
PELICAN [2211.00454]
Energy Flow Networks [1810.05165]
ParticleNet [1902.08570]
Bayesian Tagger [1904.10004]
Regression
Bayesian Regression [2003.11099]
NNPDF [2109.02653]
Matrix Elements [2206.14831]
Symbolic regression [2109.10414]
Unsupervised Learning
Clustering
Jet Clustering [2008.06064]
3D Pixel Clustering [2007.03083]
Anomaly Detection
AnomalyCLR [2301.04660]
LaCATHODE [2210.14924]
Normalized AE [2206.14225]
Reinforcement Learning
Flavor structure [2304.14176]
Simplifying Polylogs [2206.04115]
String vacua & landscape [1903.11616, 2111.11466]
Model Building [2103.04759]
Analytic continuation [2112.13011]


Talk T. Heimel
MadNIS
Precision Generation
[2110.13632]
MEM
[2210.00019]
Talk A. Youssef
NF + uncertainties
OTUS
[2101.08944]
MADNIS
[2212.06172]
NF
CaloFlow I-IV
[2106.05285,…]
Jet Simulation
[2203.00520]
VAE
Talk R. Barman
Unfolding
Talks by E. Putney & S. Lim
DM measurement
PC-JeDi
[2303.05376]
FPCD
[2304.01266]
Generative
Models
Diffusion
Models
EPiC-GAN
[2301.08128]
GAN
How to GAN
[1907.03764]
CaloGAN
[1712.10321]
?
Talk R. Rizatdinova
SUSY@ATLAS
Point Clouds
[2102.05073]
Talk L. Benato
LLP@CMS
PELICAN
[2211.00454]
Energy Flow Networks
[1810.05165]
Classification
ParticleNet
[1902.08570]
Talk R. Forestano
DL symmetries
Bayesian Tagger
[1904.10004]
Supervised
Learning
Bayesian Regression
[2003.11099]
Machine
Learning
Regression
NNPDF
[2109.02653]
Matrix Elements
[2206.14831]
Symbolic regression
[2109.10414]
Unsupervised
Learning
Talk K. Mohan
PDF uncertainties
Flavor structure
[2304.14176]
Talk H. Day
Physics of NN
Talk M. Madigan
SMEFT + SIMUnet
Jet Clustering
[2008.06064]
Clustering
Simplifying Polylogs
[2206.04115]
Reinforcement
Learning
String vacua & landscape
[1903.11616, 2111.11466]
3D Pixel Clustering
[2007.03083]
AnomalyCLR
[2301.04660]
Anomaly
Detection
LaCATHODE
[2210.14924]
Model Building
[2103.04759]
Analytic continuation
[2112.13011]
Talk R. Das
Feature selection
Talk L. Favaro
Anomaly CLR
Normalized AE
[2206.14225]
Talk S. Roche
Decision tree AE


Talk T. Heimel
MadNIS
Precision Generation
[2110.13632]
MEM
[2210.00019]
Talk A. Youssef
NF + uncertainties
OTUS
[2101.08944]
MADNIS
[2212.06172]
NF
CaloFlow I-IV
[2106.05285,…]
Jet Simulation
[2203.00520]
VAE
Talk R. Barman
Unfolding
Talks by E. Putney & S. Lim
DM measurement
PC-JeDi
[2303.05376]
FPCD
[2304.01266]
Generative
Models
Diffusion
Models
EPiC-GAN
[2301.08128]
GAN
How to GAN
[1907.03764]
CaloGAN
[1712.10321]
?
?

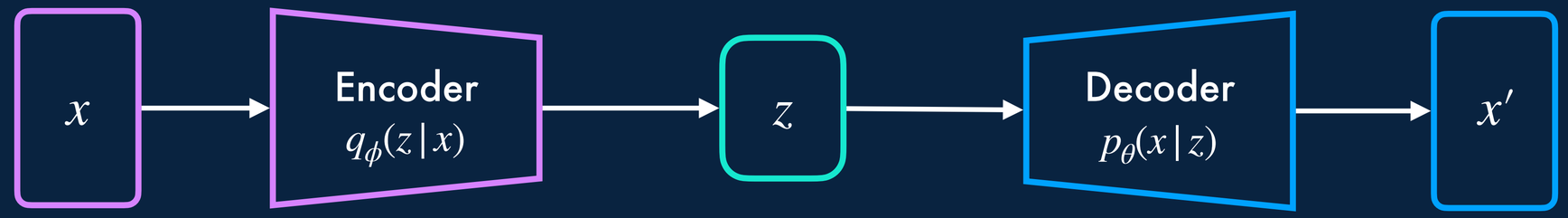

Theory
$\mathcal{L}$
Hard process
Shower
Hadronization
Detectors
Events

Deep generative models

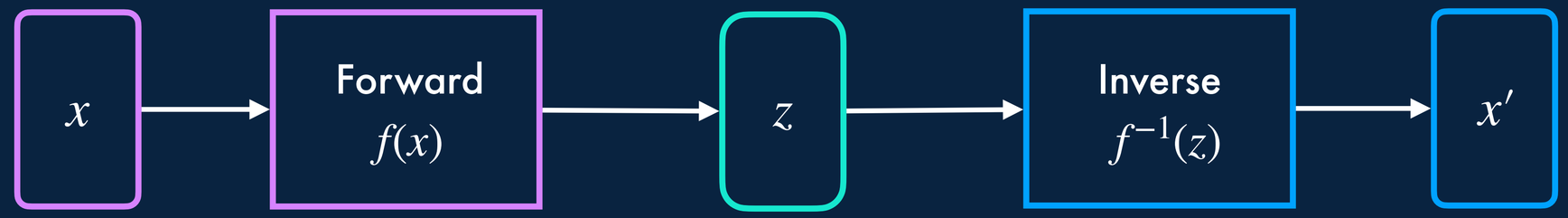
Deep generative models

Generative
Adversarial
Network

Variational
Auto-
Encoder

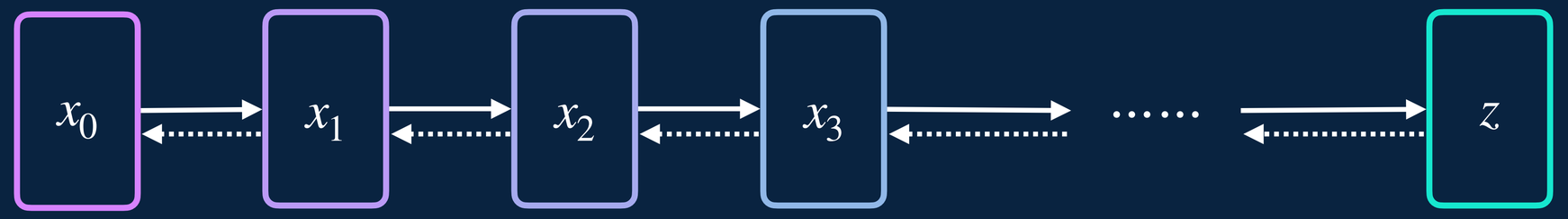


Normalizing
Flow



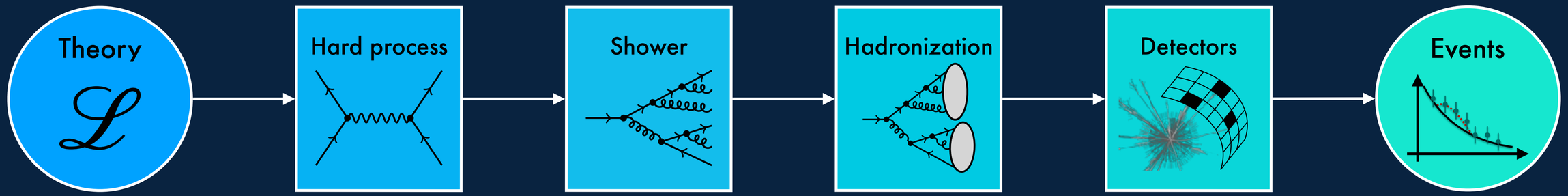
New kids in town

Diffusion
Probabilistic
Model

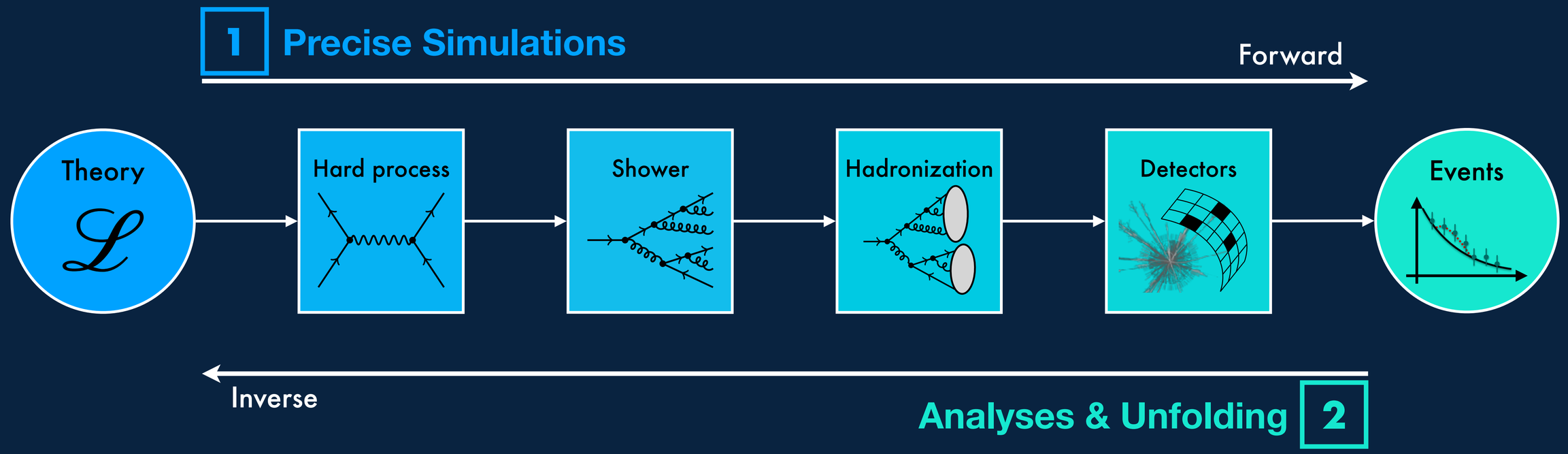


Generative
Pretrained
Transformer

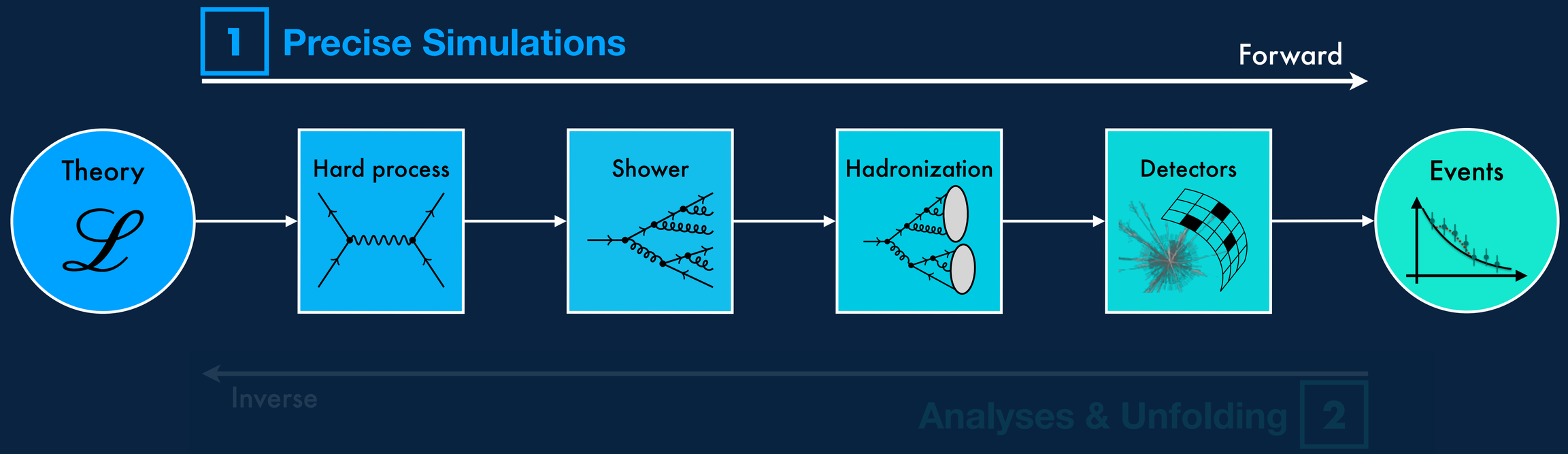
ML aided simulation chain



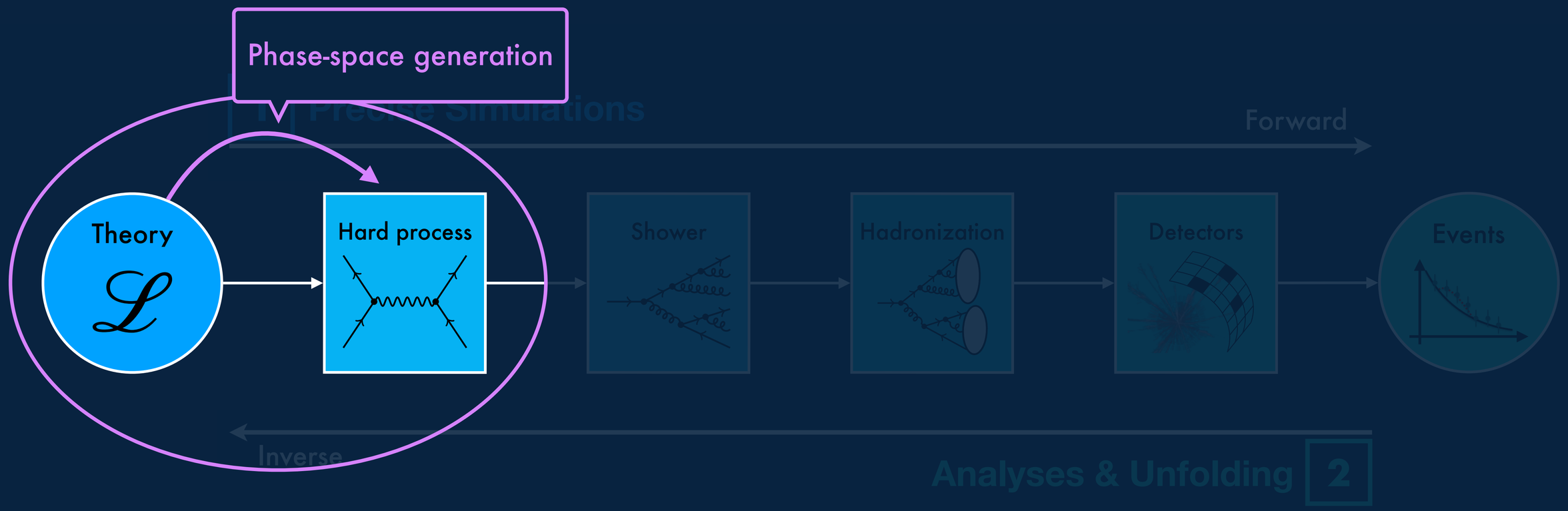
ML aided simulation chain



ML improved simulations

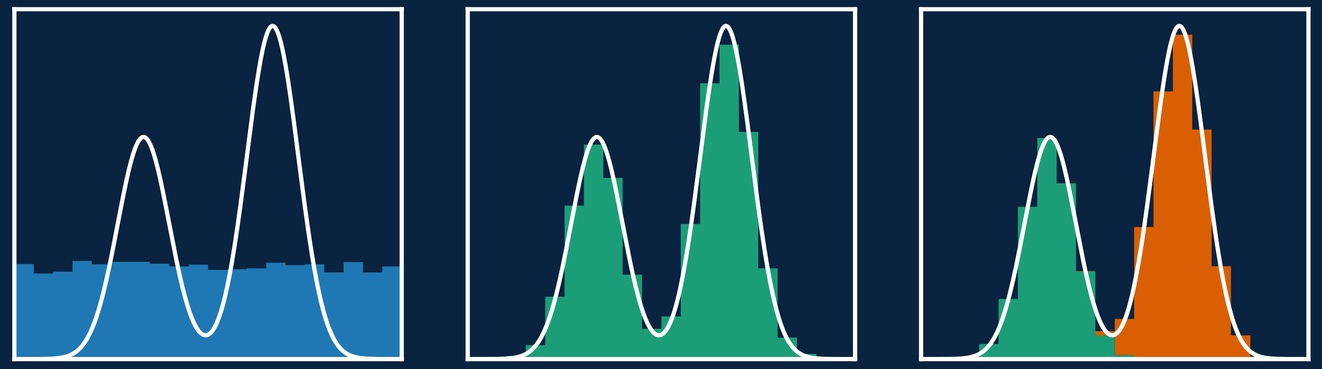


ML improved simulations



BDT [1707.00028, …], NN [1810.11509, 2009.07819, …]
NF [2001.05486, 2001.05478, 2001.10028, 2005.12719, 2112.09145, 2212.06172, …]

MadNIS — Neural importance sampling

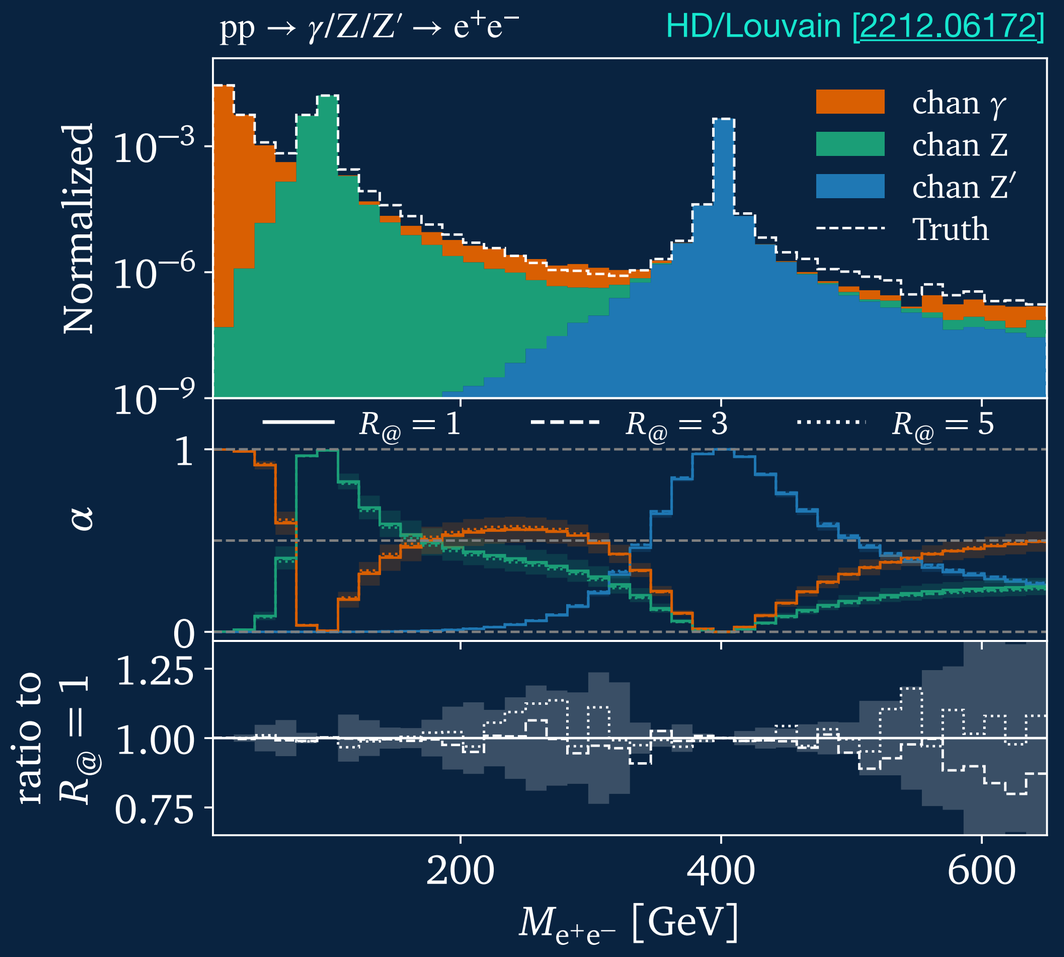


Flat sampling | Importance Sampling | Multi-channel

$$I = \sum_i \left\langle \alpha_i(x) \frac{f(x)}{g_i(x)} \right\rangle_{x \sim g_i(x)}$$

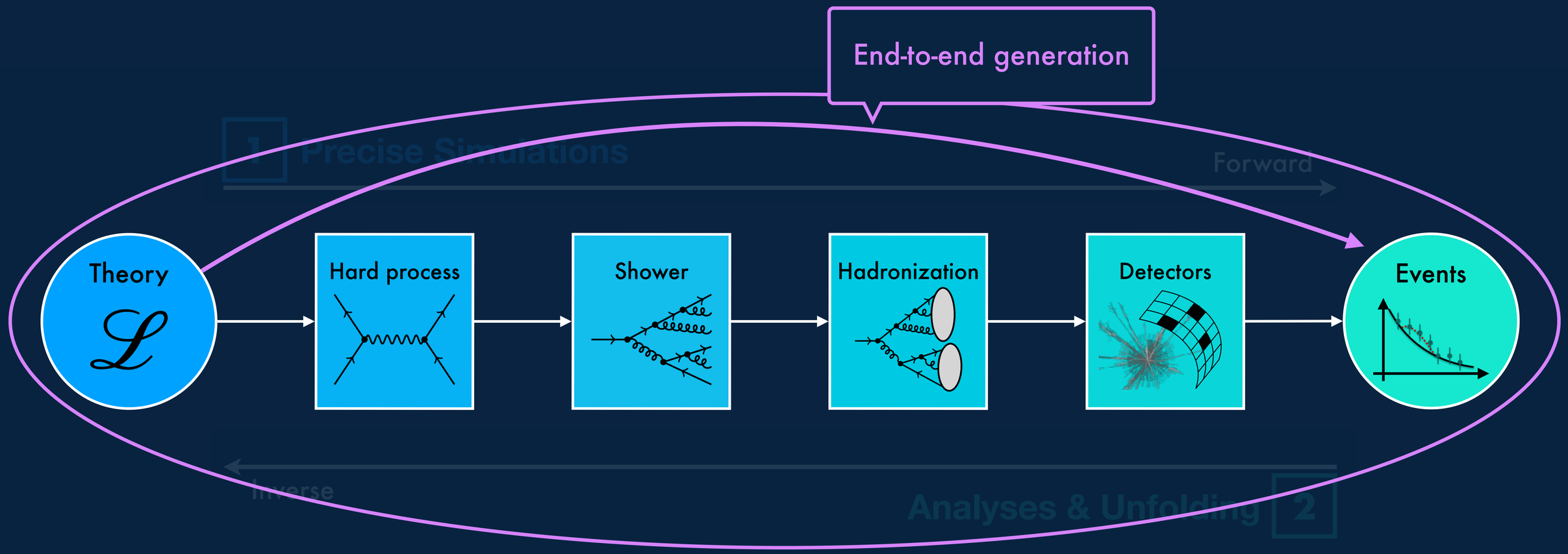
Parametrize with **NN**

Parametrize with **NF**



→ **Details in talk by T. Heimel**

End-to-end generation

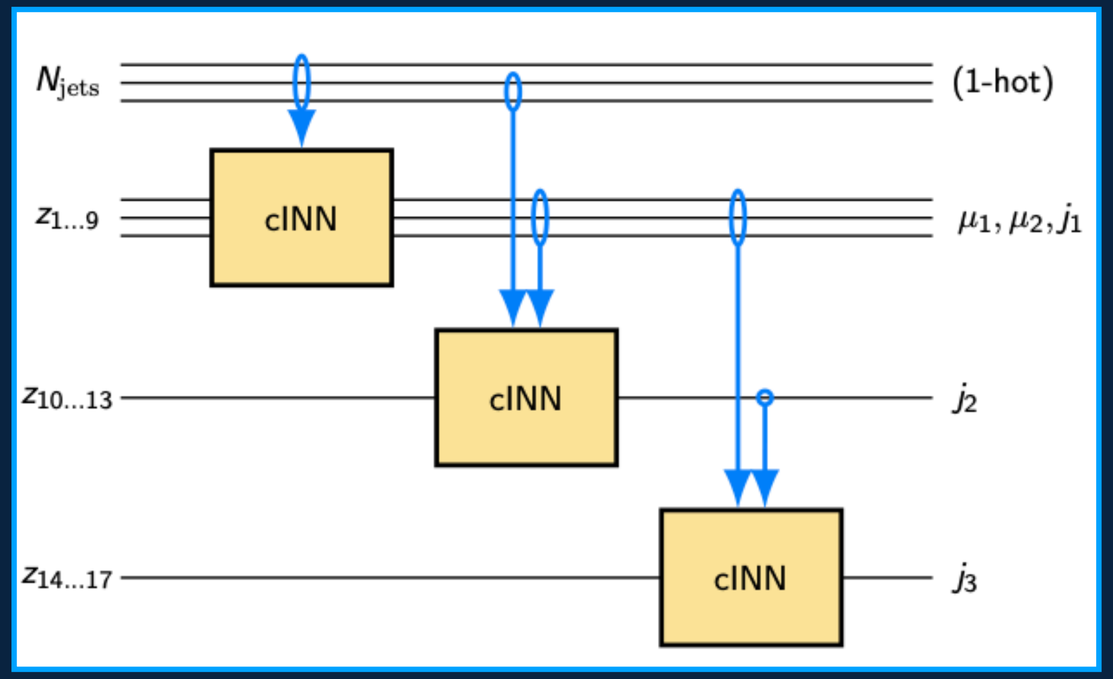


VAE [2101.08944, …], NF [2011.13445, 2110.13632, …]
GAN [1901.00875, 1901.05282, 1903.02433,1907.03764, 1912.02748, 2001.11103, …]

Precision generation

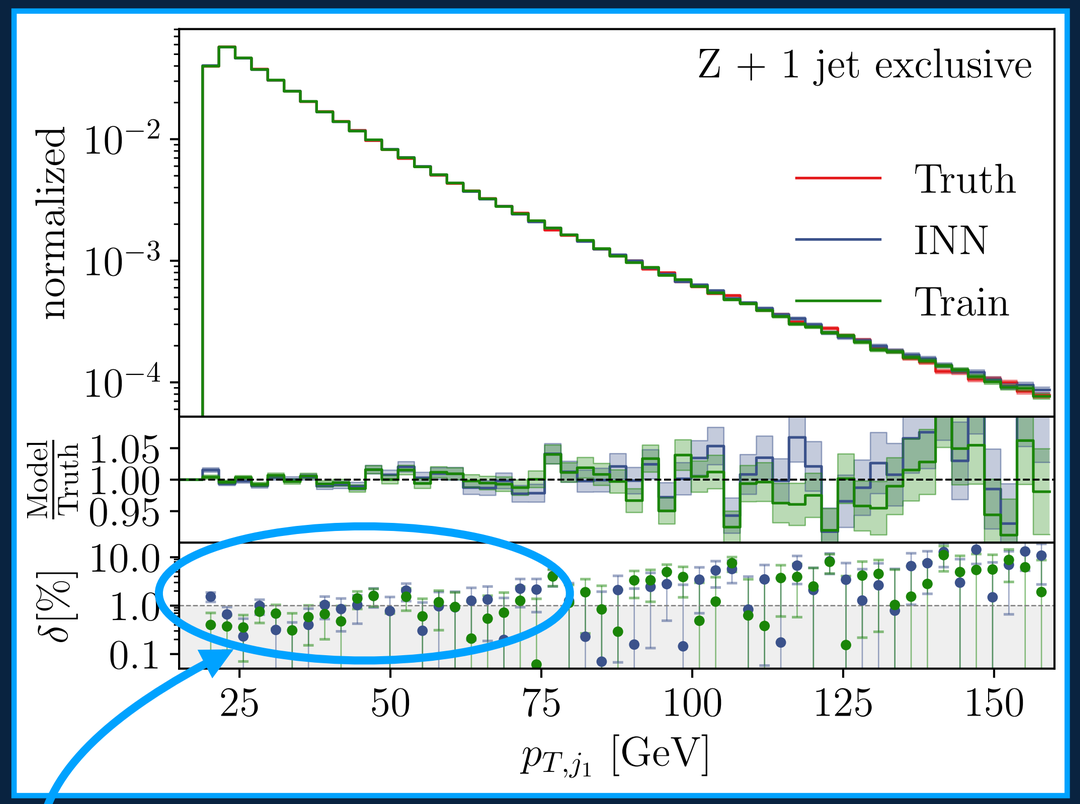
Z + jets production:

$$\mathrm{pp} \to \mathrm{Z} + \{1,2,3\}\,\mathrm{jets}$$
$$\to \mu^+\mu^- + \{1,2,3\}\,\mathrm{jets}$$



Parametrize with **NF**

Heidelberg (HD) [2110.13632]

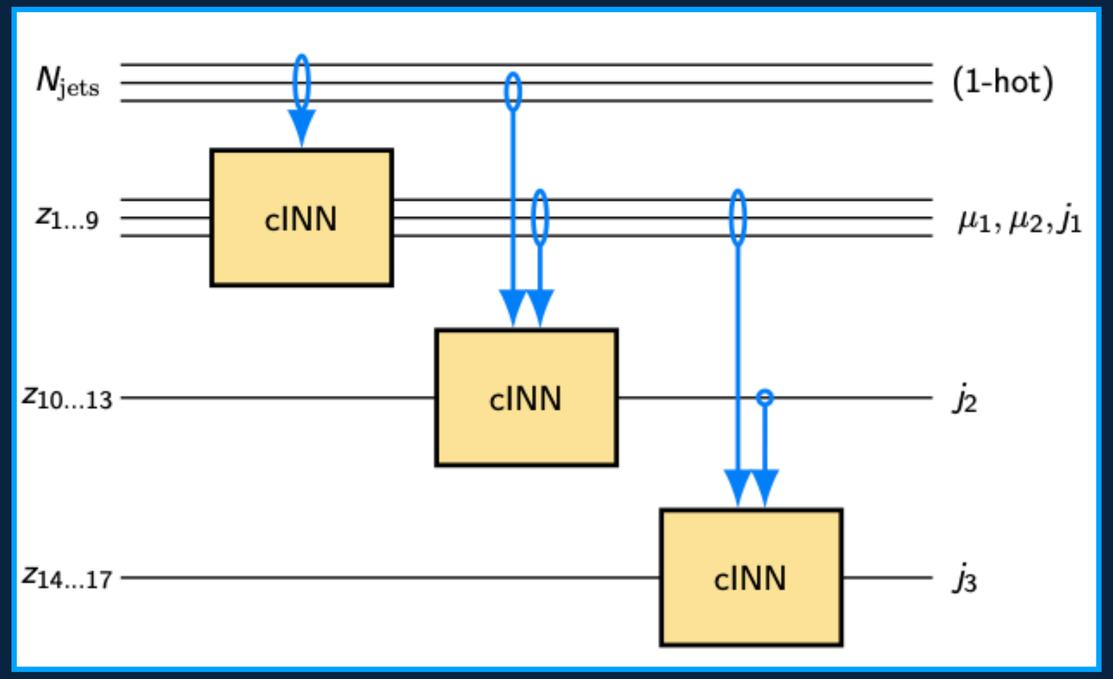


⊕⊕ Sub-percent precision in bulk region

Precision generation

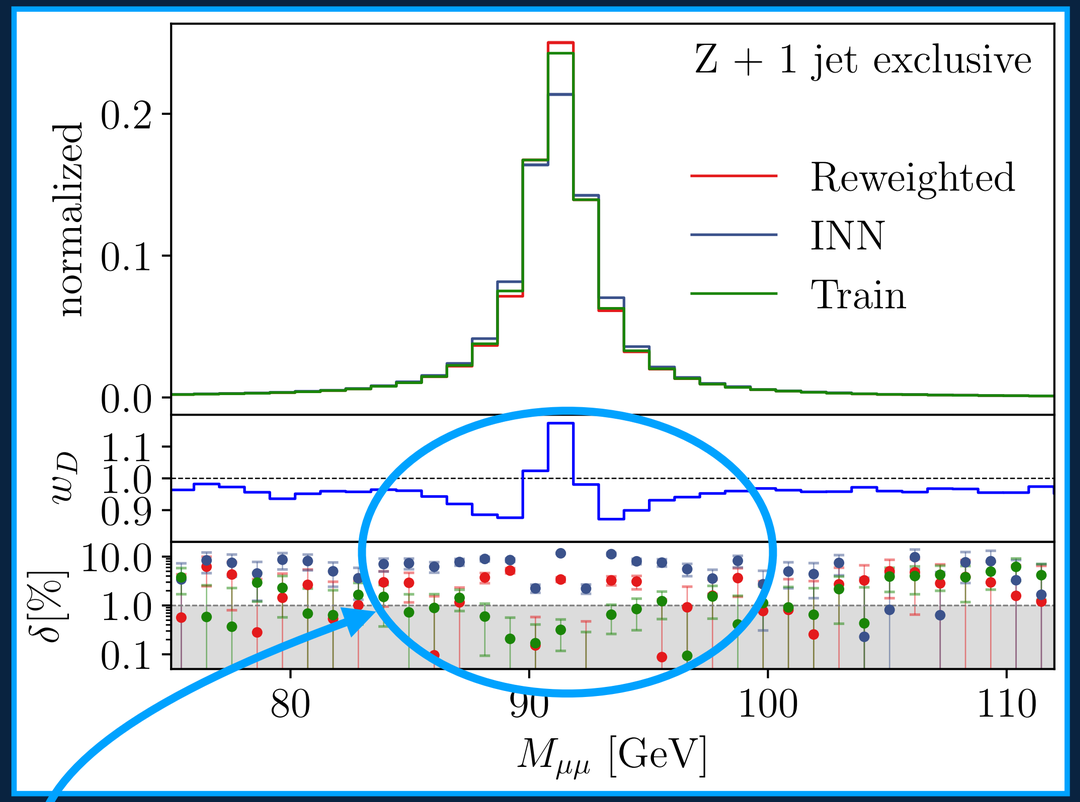
Z + jets production:

$$pp \to Z + \{1,2,3\}\text{ jets}$$
$$\to \mu^+\mu^- + \{1,2,3\}\text{ jets}$$



Parametrize with **NF**

Heidelberg (HD) [2110.13632]



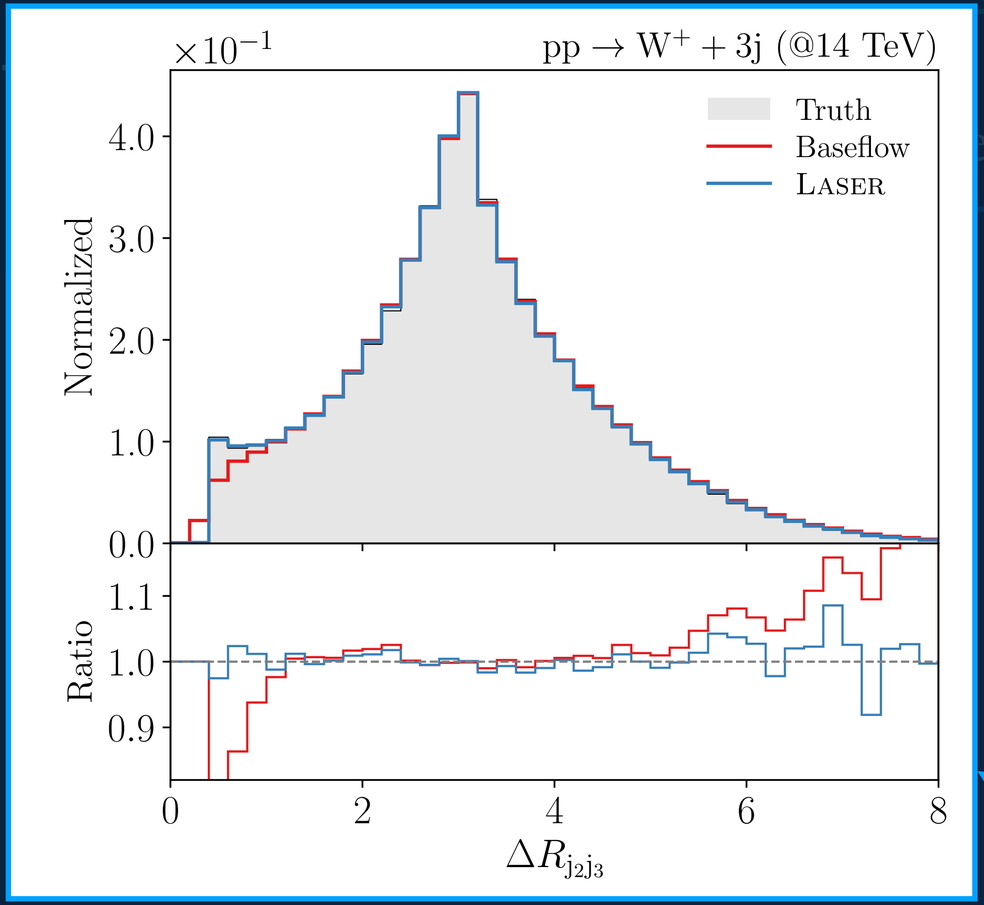
⊕ Additional **classifier** improves precision

DCTRGAN [1907.08209, 2009.03796], LASER [2106.00792]

⊕ Berkeley/Louvain + HD/Rutgers → **Soon!**

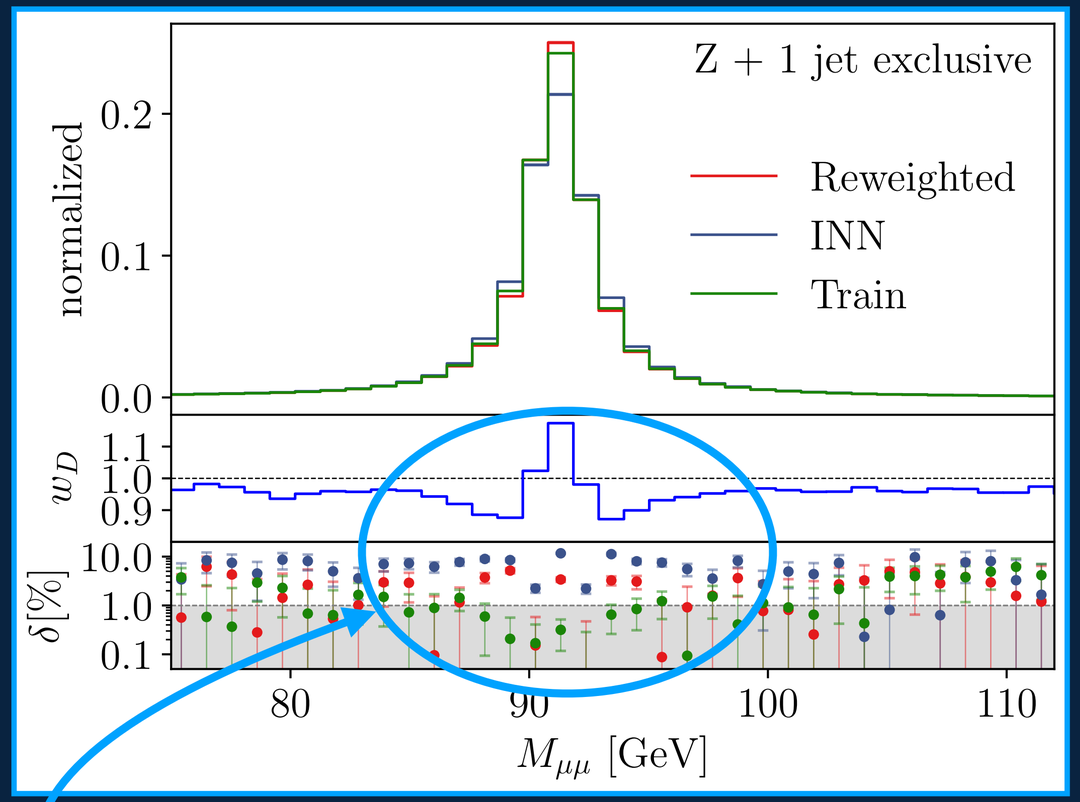
Precision generation

Berkeley/Louvain [2305.xxxxx]



⊕ LASER gives **unweighted events**

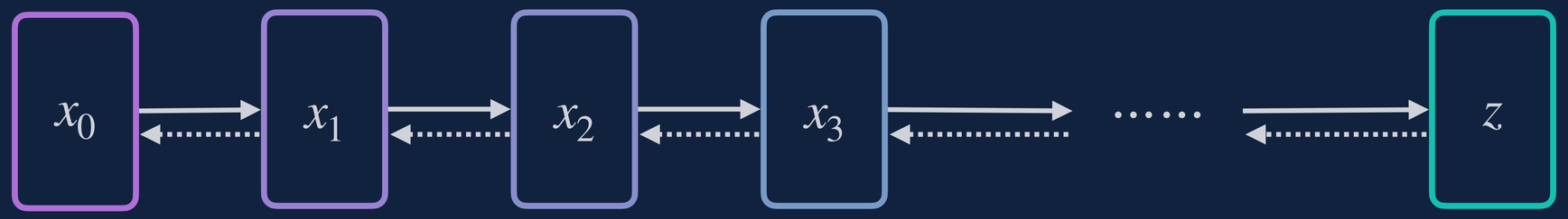
Heidelberg (HD) [2110.13632]



⊕ Additional **classifier** improves precision

DCTRGAN [1907.08209, 2009.03796], LASER [2106.00792]

⊕ Berkeley/Louvain + HD/Rutgers → **Soon!**



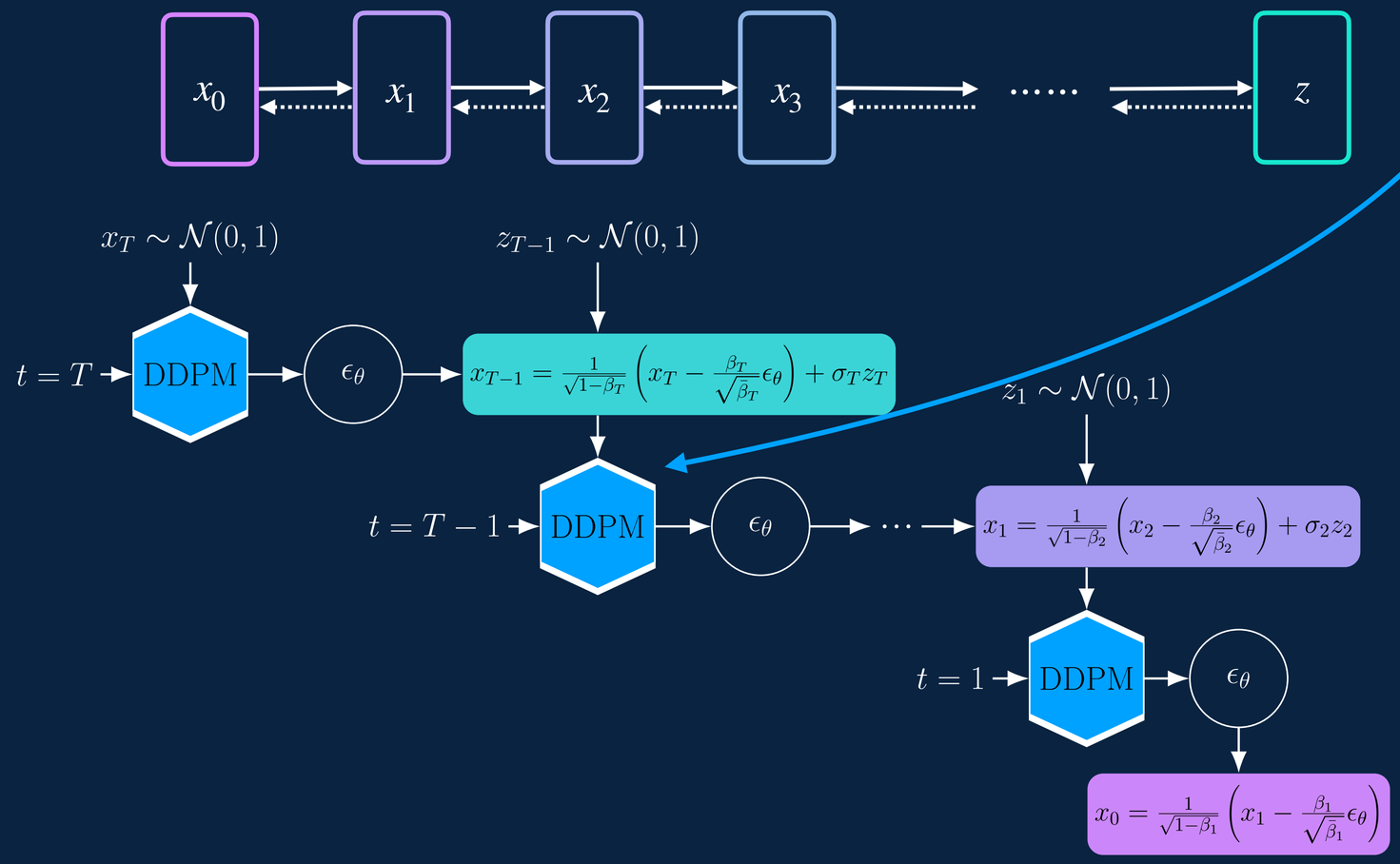
New kids in town

Pushing limits in the precision era?

Credits to Nathan Huetsch, Sofia Schweitzer, Peter Sorrenson and Jonas Spinner

Denoising diffusion probabilistic model

Diffusion probabilistic model



Parametrize with **DDPM**

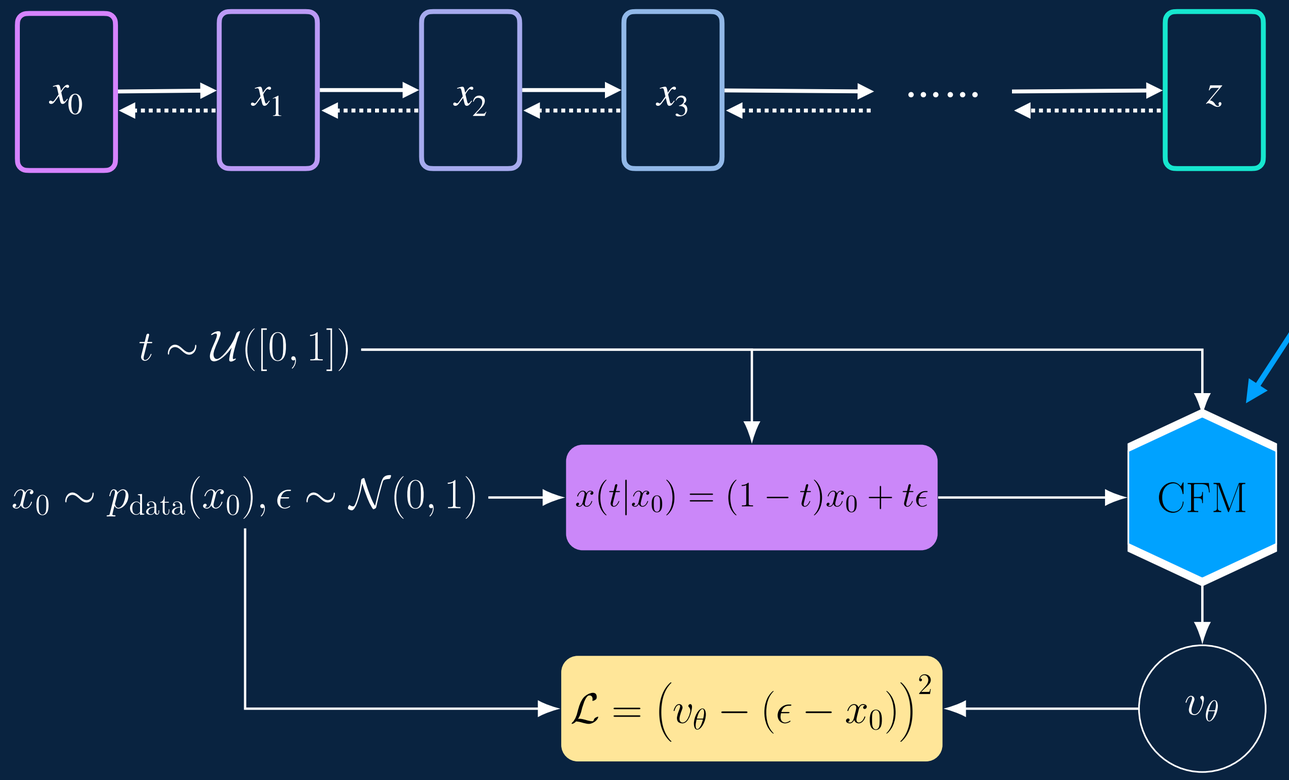
- **Discrete** diffusion time steps **T**
- Invert each step iteratively

⊖ Requires **T** network evaluations

⊕ State-of-the-art image generation

Conditional flow matching

Diffusion probabilistic model



Parametrize with **CFM**

- **Continuous** time evolution
- Solve ODE numerically

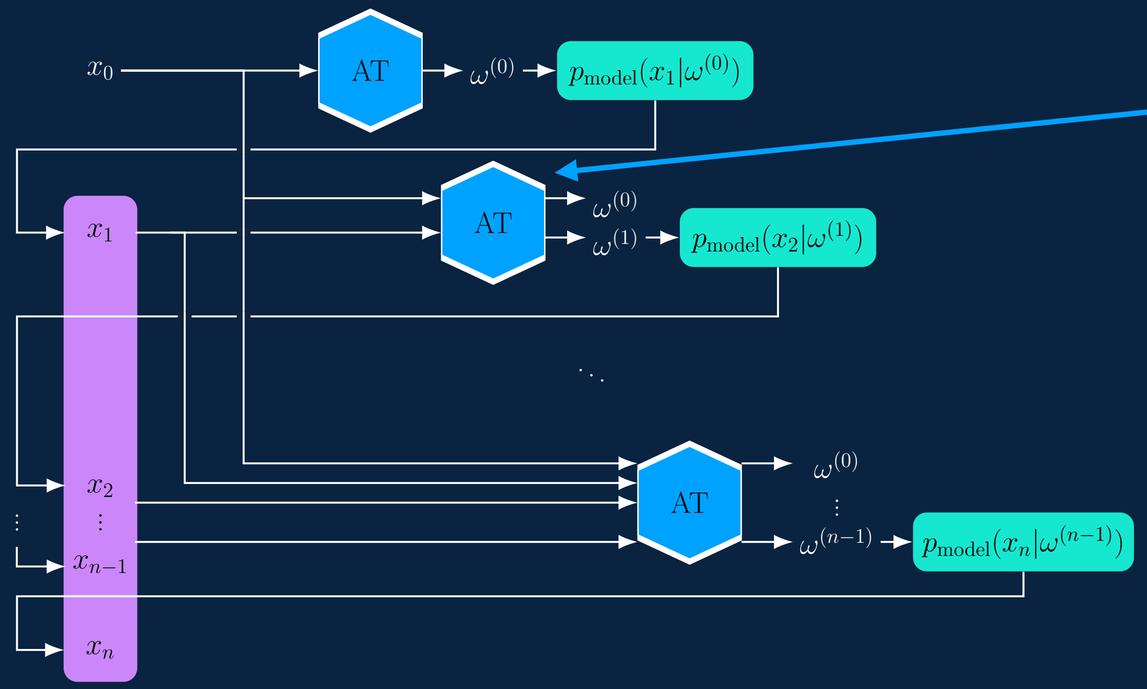
$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = v_\theta(x(t), t)$$

⊕ No fine tuning of finite time steps

⊖ Solving ODE numerically is slow and requires **multiple** network evals

Autoregressive transformer

Generative (pre-trained) transformer



Parametrize with **AT**

- **Autoregressive** sampling
- **Self-attention** for complicated correlations and combinatorics

⊕ Arbitrary input length **L**

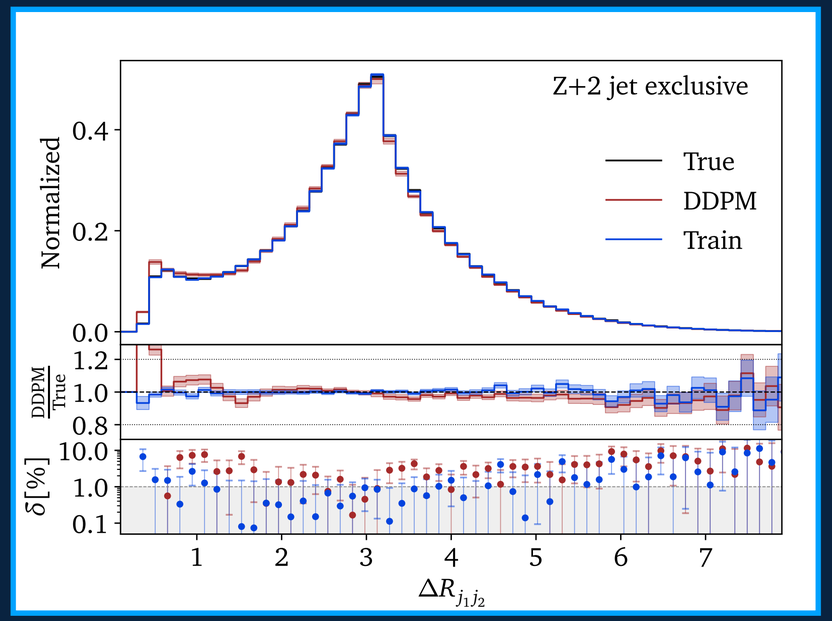
⊖ Requires **L** network evaluations

⊕ State-of-the-art text generation

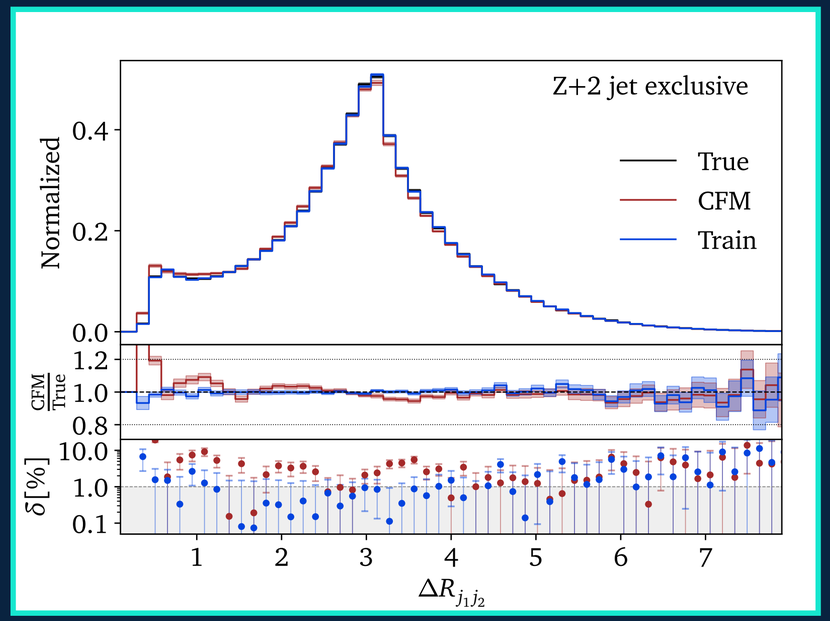
New kids in town — Future?

— Jet separation —

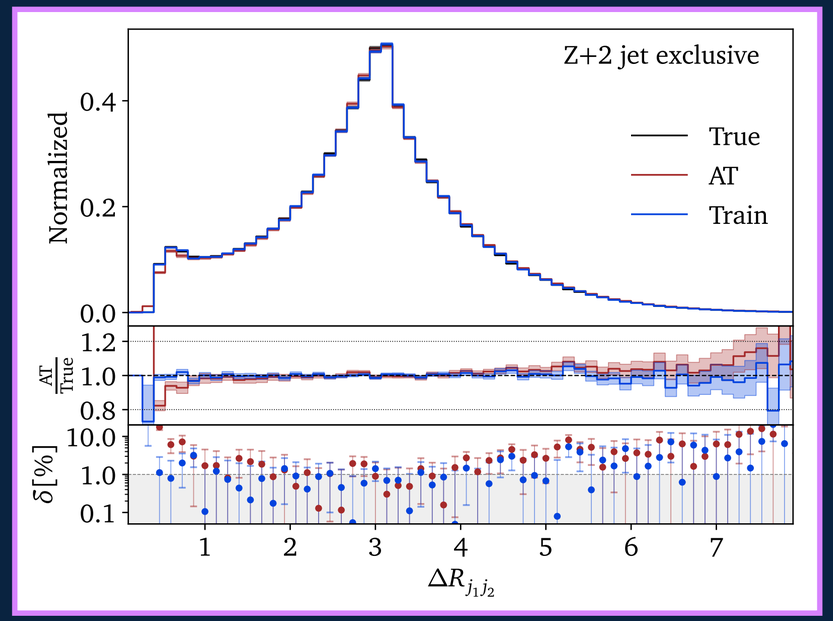
DDPM



CFM



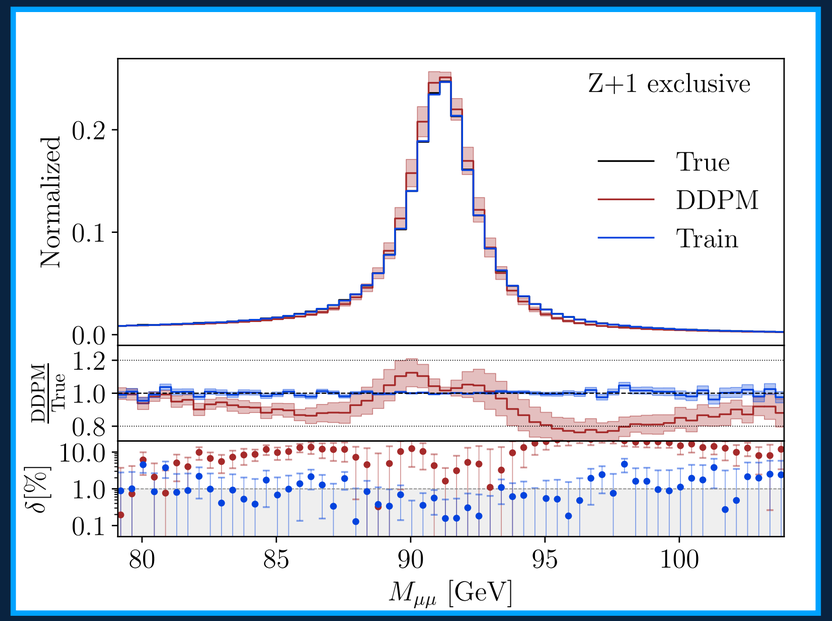
AT



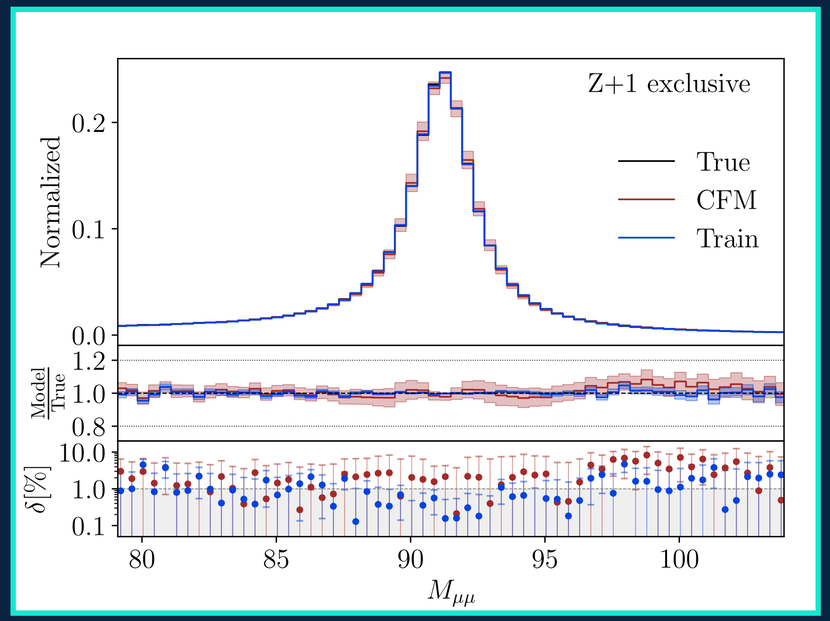
New kids in town — Future?

— Invariant lepton mass —

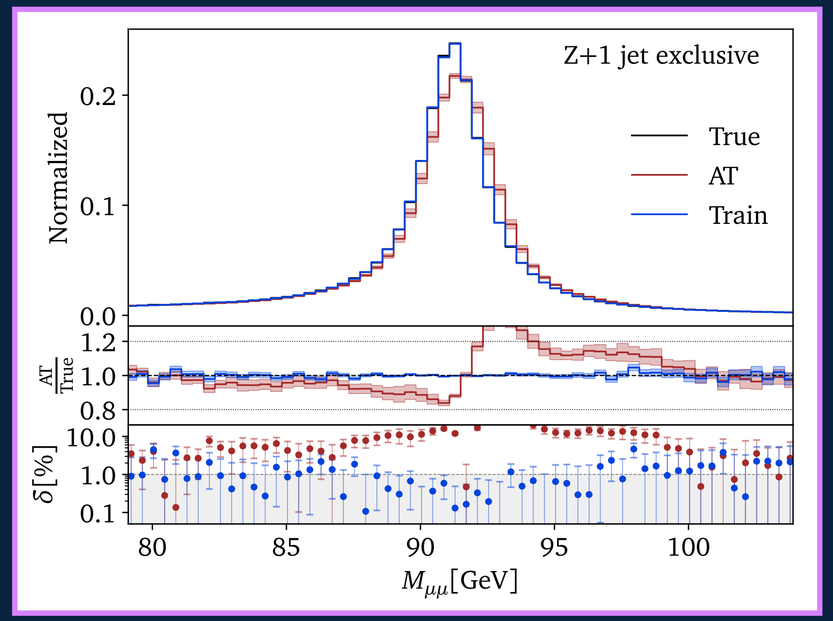
DDPM



CFM



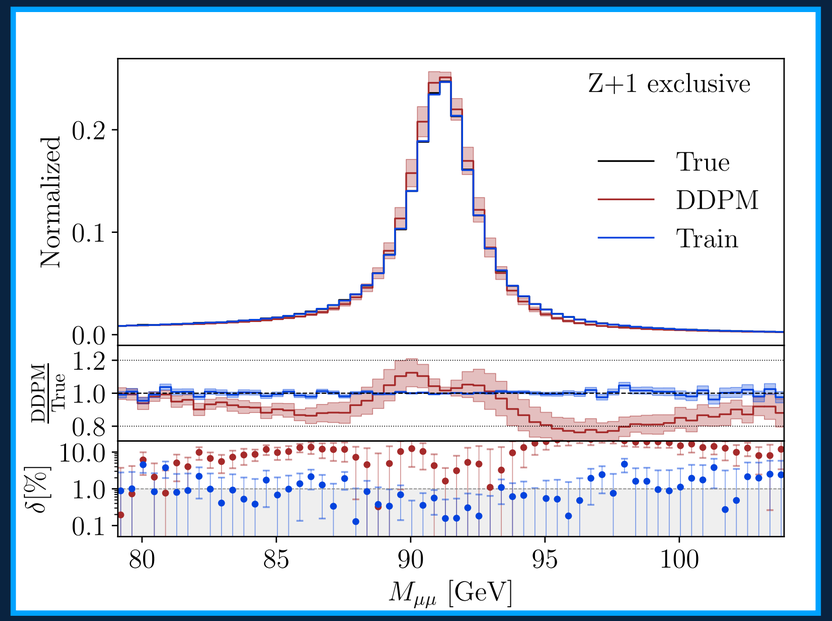
AT



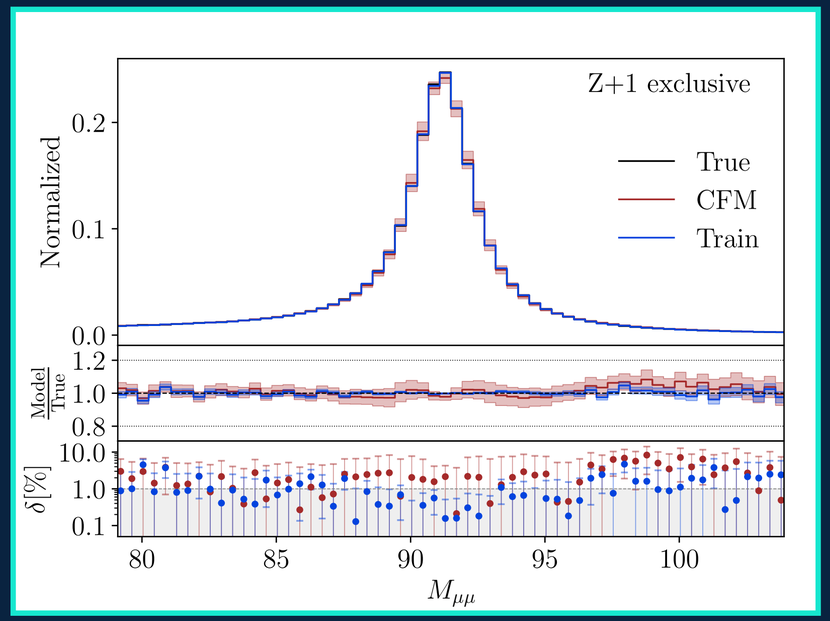
New kids in town — Future?

— Invariant lepton mass —

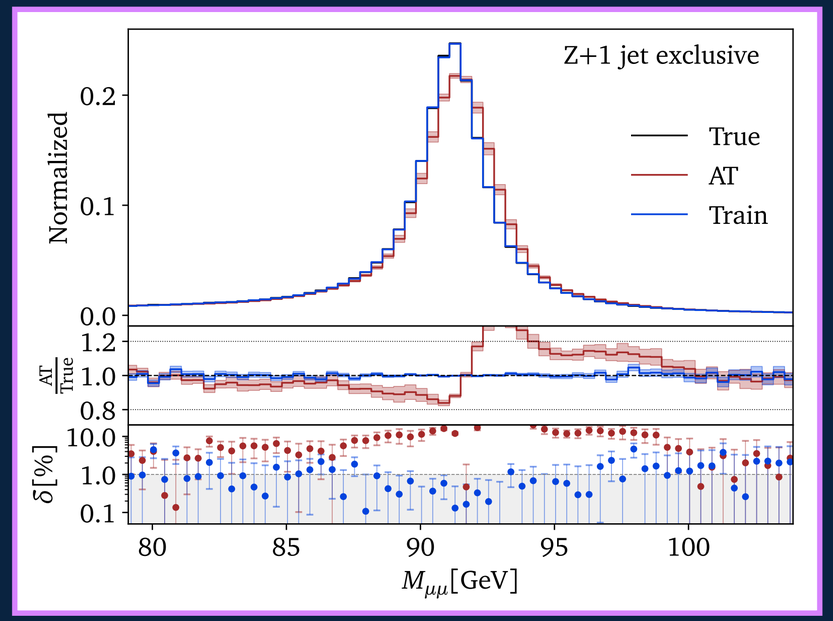
DDPM



CFM

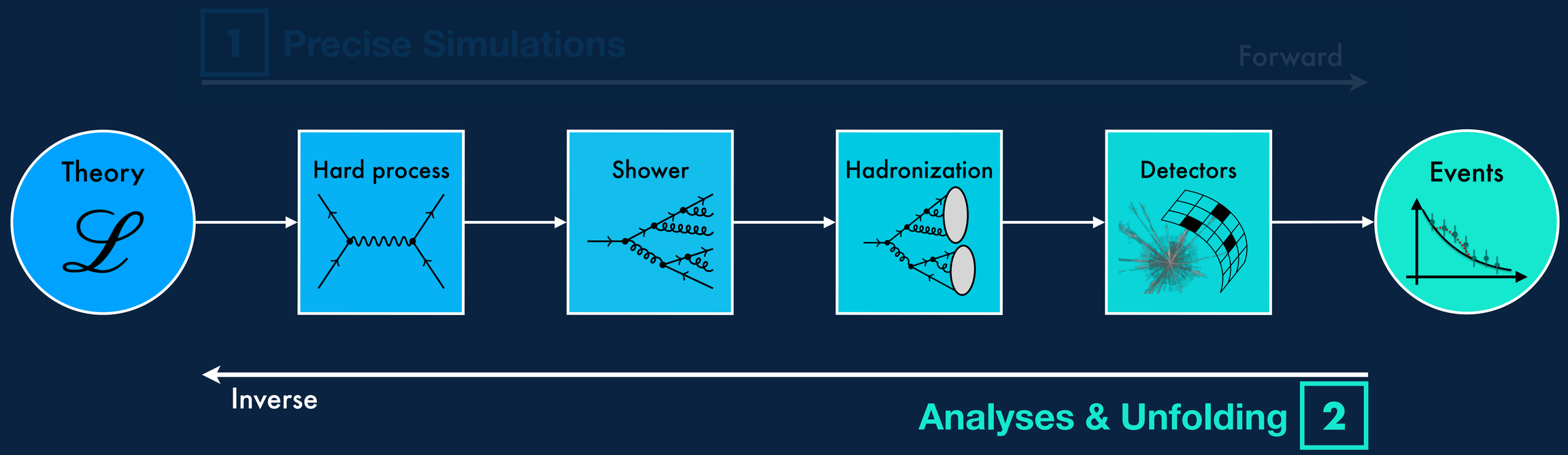


AT

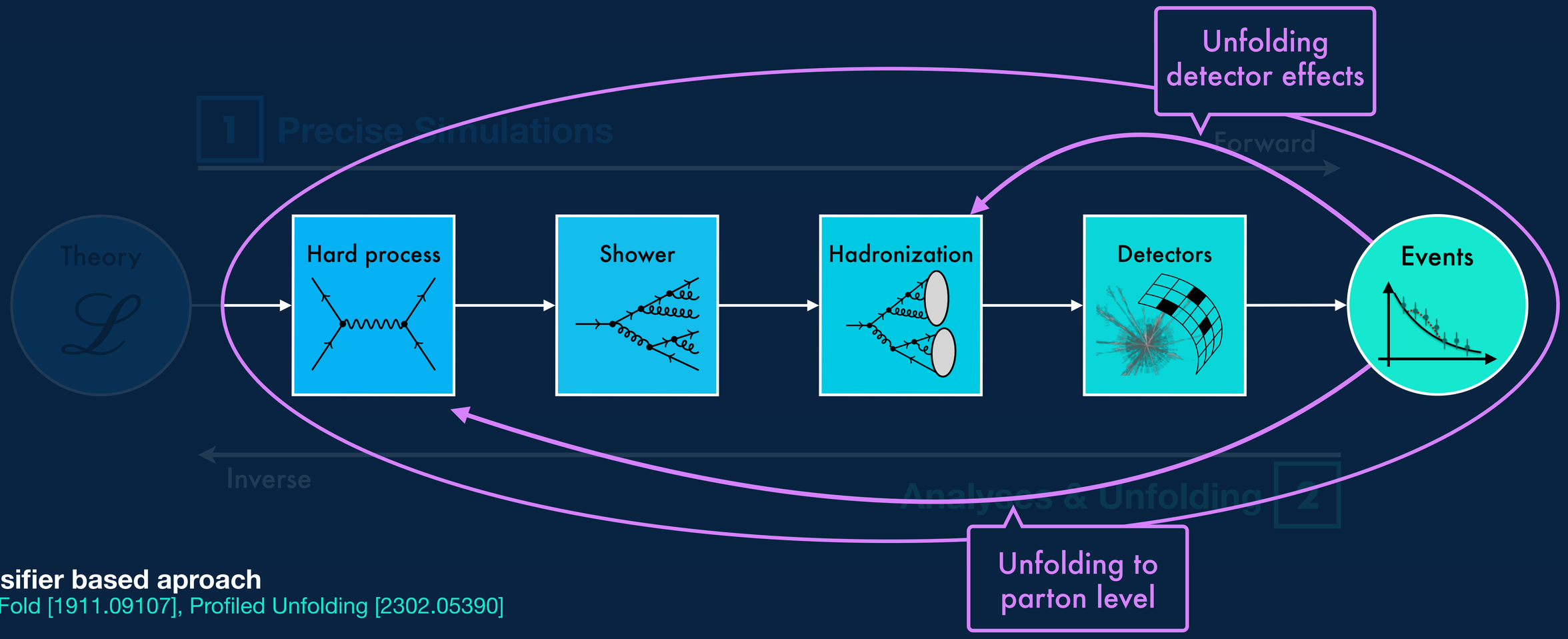


→ Soon to be published [2305.xxxx]

Inverting the simulation chain



Inverting the simulation chain

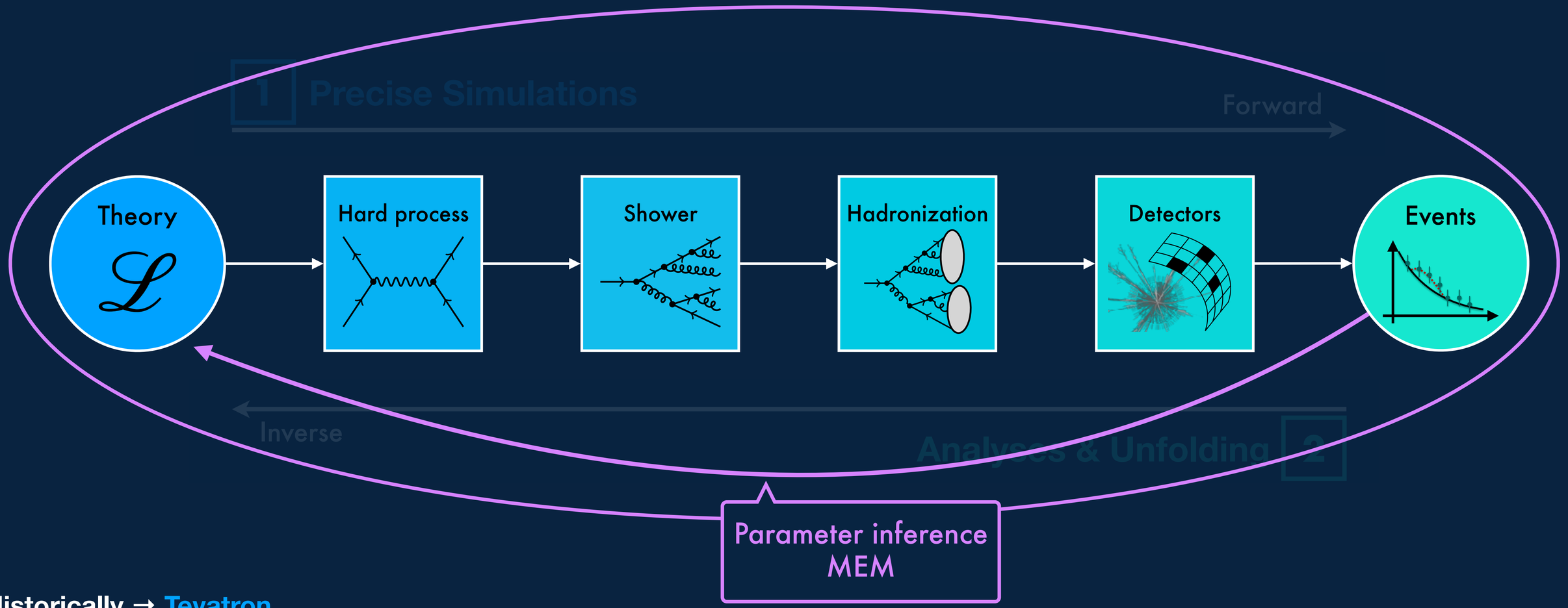


Classifier based aproach
OmniFold [1911.09107], Profiled Unfolding [2302.05390]

Density based approach
FCGAN [1912.00477], cINN [2006.06685],
IcINN [2212.08674], OTUS [2101.08944]

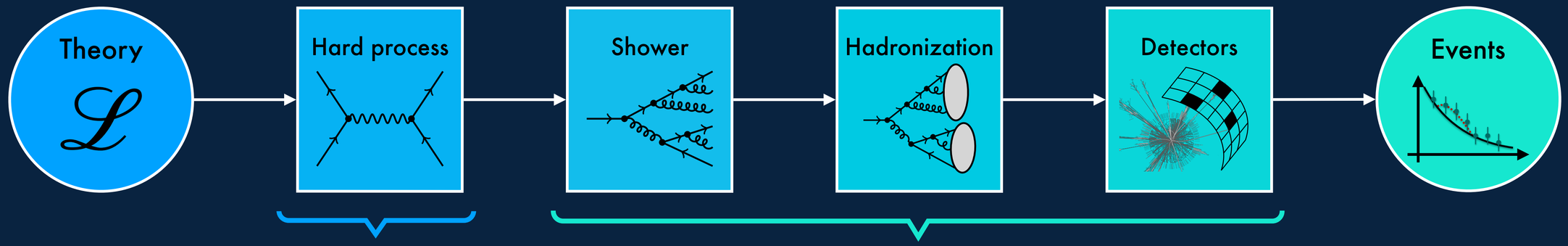
→ **More in talk by R. Barman**

Inverting the simulation chain



Historically → Tevatron
Top mass: D0 (98', 04'), CDF 06', Fiedler et al. [1003.1316]
Single-top: Review [1710.10699]

Inference with normalizing flows



Theory parameter α

Known from theory

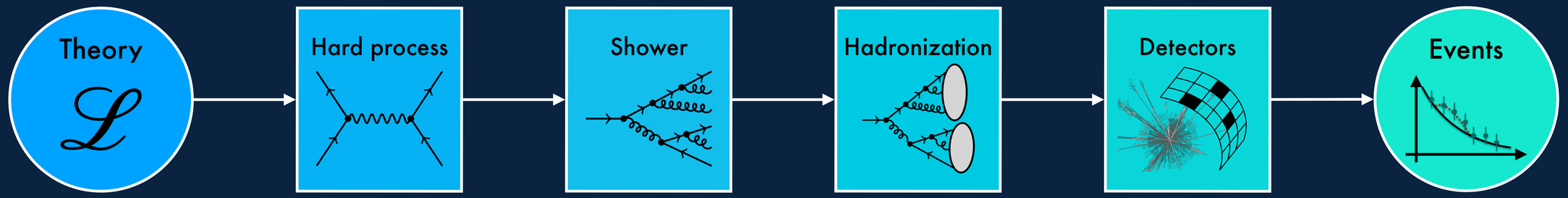
Likelihood intractable → parametrize with **NF**

Reconstructed momenta x_{reco}

MEM master formula:

$$p(x_{\text{reco}} \,|\, \alpha) = \int \mathrm{d}x_{\text{hard}}\, p(x_{\text{hard}} \,|\, \alpha)\, p(x_{\text{reco}} \,|\, x_{\text{hard}})$$

Inference with normalizing flows



Theory parameter α

Known from theory

Likelihood intractable → parametrize with **NF**

Reconstructed momenta x_{reco}

MEM master formula:

$$p(x_{\text{reco}} \,|\, \alpha) = \int \mathrm{d}x_{\text{hard}}\, p(x_{\text{hard}} \,|\, \alpha)\, p(x_{\text{reco}} \,|\, x_{\text{hard}})$$

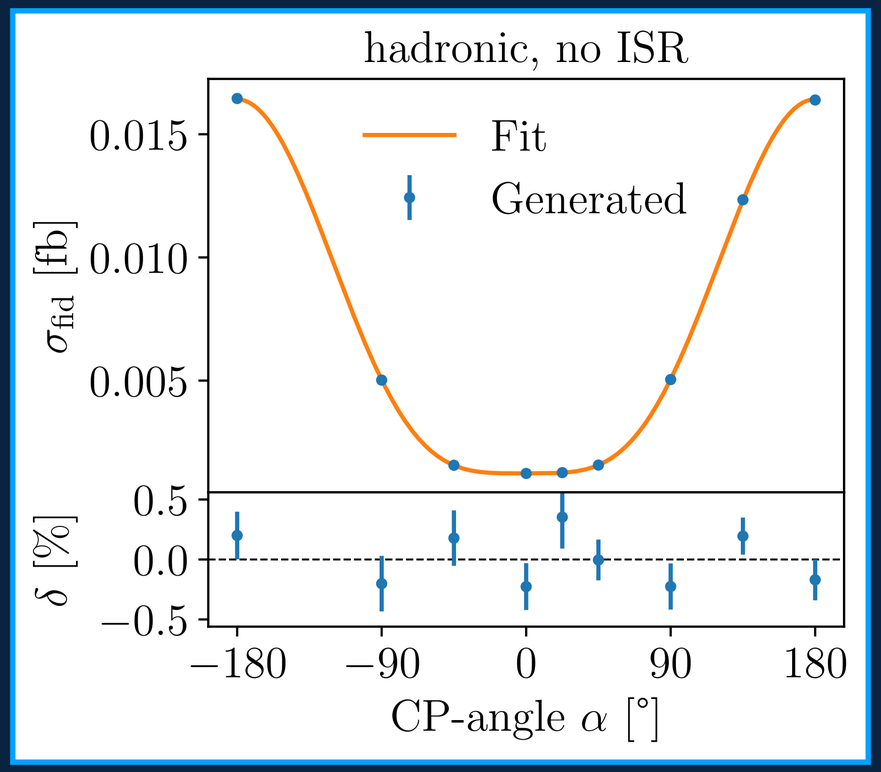
In practice → perform integral **numerically**

parametrize with additional **NF**

Talk by T. Heimel

Matrix element method

Heidelberg/Louvain [2210.00019, 23XX.xxxxx]



tHj production:

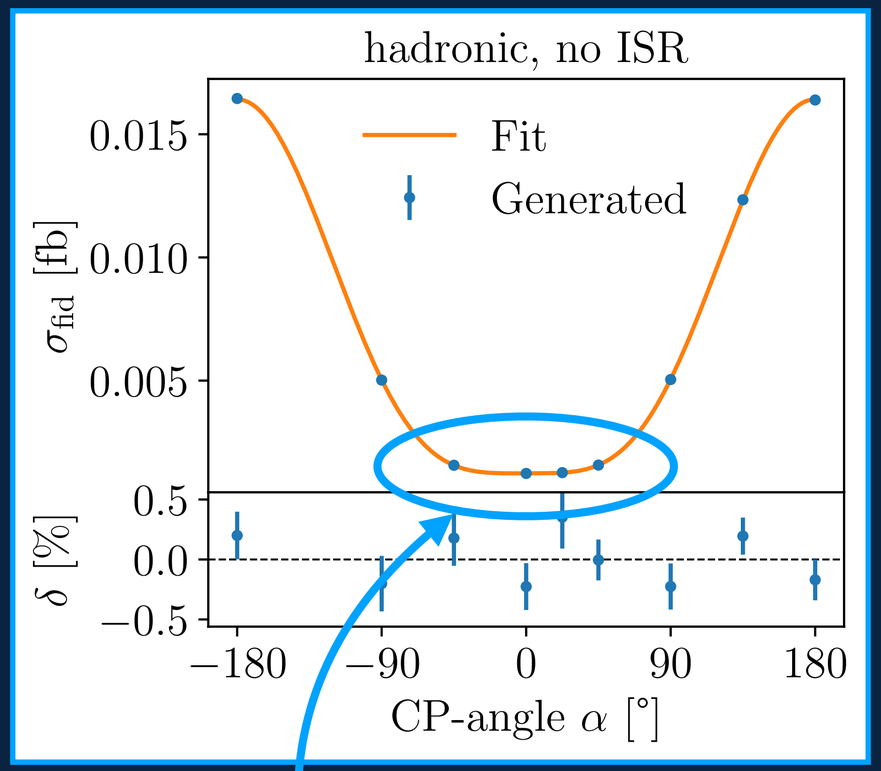
$$\begin{aligned} pp &\to tHj \\ &\to (bW)\,(\gamma\gamma)\,j \end{aligned}$$

$$\mathscr{L}_{t\bar{t}H} = -\frac{y_t}{\sqrt{2}}\left[\cos\alpha\ \bar{t}t + \frac{2}{3}i\sin\alpha\ \bar{t}\gamma_5 t\right]H$$

Anomalous coupling with **CP-angle** α

Matrix element method

Heidelberg/Louvain [2210.00019, 23XX.xxxxx]



tHj production:

$$\begin{aligned} pp &\to tHj \\ &\to (bW)\,(\gamma\gamma)\,j \end{aligned}$$

$$\mathscr{L}_{t\bar{t}H} = -\frac{y_t}{\sqrt{2}}\left[\cos\alpha\ \bar{t}t + \frac{2}{3}i\sin\alpha\ \bar{t}\gamma_5 t\right]H$$

Anomalous coupling with **CP-angle** α

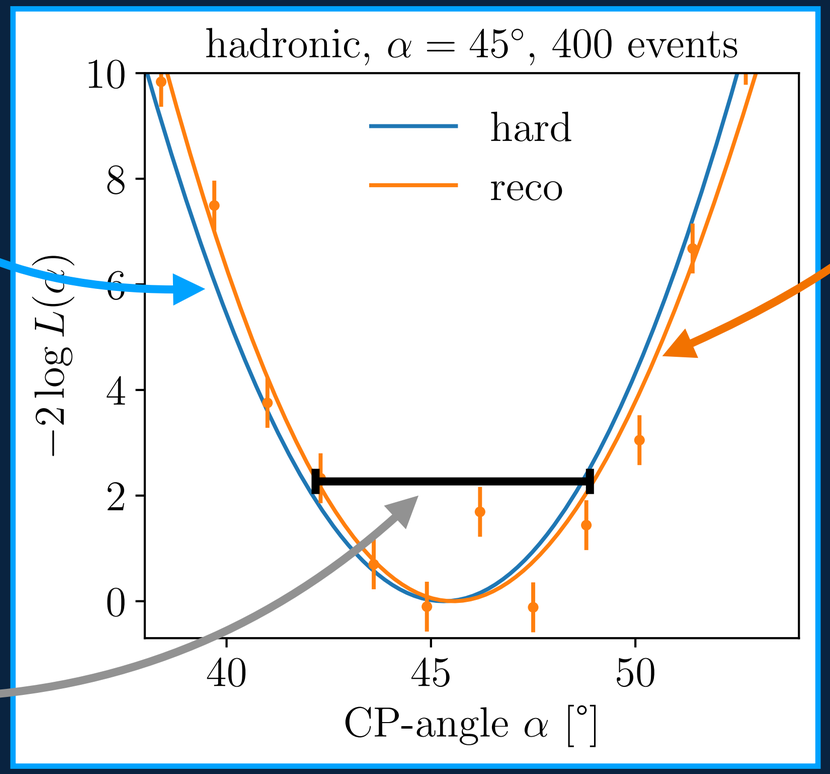
Around SM ($\alpha = 0$):

- ⊖ low total cross section (few events)
- ⊖ low variation of rate
- ⊕ kinematics sensitive → ideal use case for **MEM**

Matrix element method

Hard-scattering truth for comparison

Heidelberg/Louvain [2210.00019, 23XX.xxxxx]



Result from MEM

Statistical uncertainty

tHj production:

$$\mathrm{pp} \to \mathrm{tH\,j}$$
$$\to (\mathrm{bW})\,(\gamma\gamma)\,\mathrm{j}$$

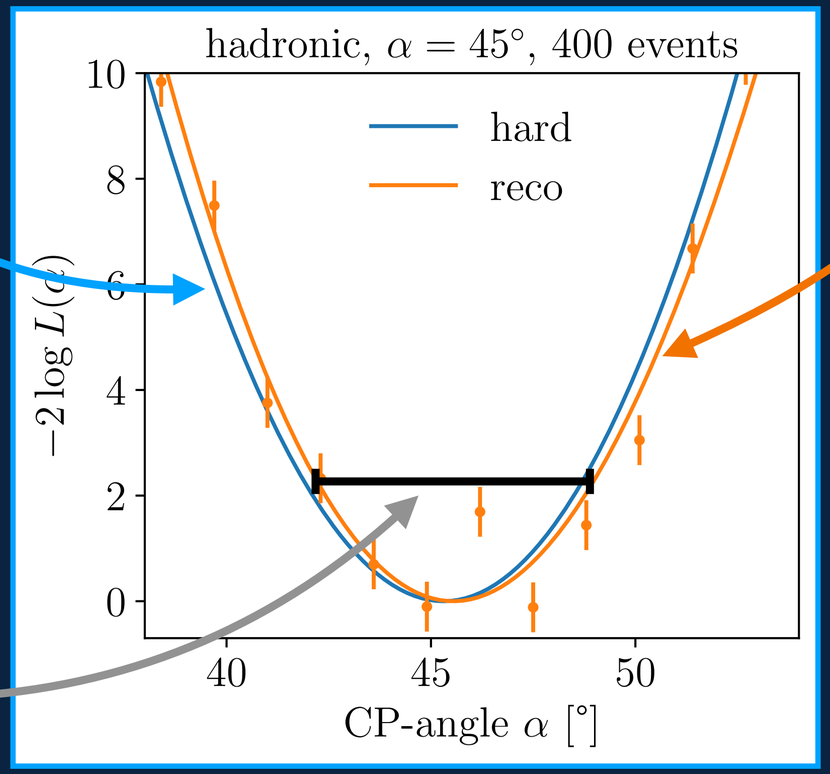
$$\mathscr{L}_{\mathrm{t\bar{t}H}} = -\frac{y_t}{\sqrt{2}}\left[\cos\alpha\ \bar{\mathrm{t}}\mathrm{t} + \frac{2}{3}\mathrm{i}\sin\alpha\ \bar{\mathrm{t}}\gamma_5\mathrm{t}\right]\mathrm{H}$$

Anomalous coupling with **CP-angle** α

Matrix element method

Hard scattering truth for comparison

Heidelberg/Louvain [2210.00019, 23XX.xxxxx]



Result from MEM

Statistical uncertainty

tHj production:

$$pp \to tH\,j$$
$$\to (bW)\,(\gamma\gamma)\,j$$

$$\mathscr{L}_{t\bar{t}H} = -\frac{y_t}{\sqrt{2}}\left[\cos\alpha\ \bar{t}t + \frac{2}{3}i\sin\alpha\ \bar{t}\gamma_5 t\right]H$$

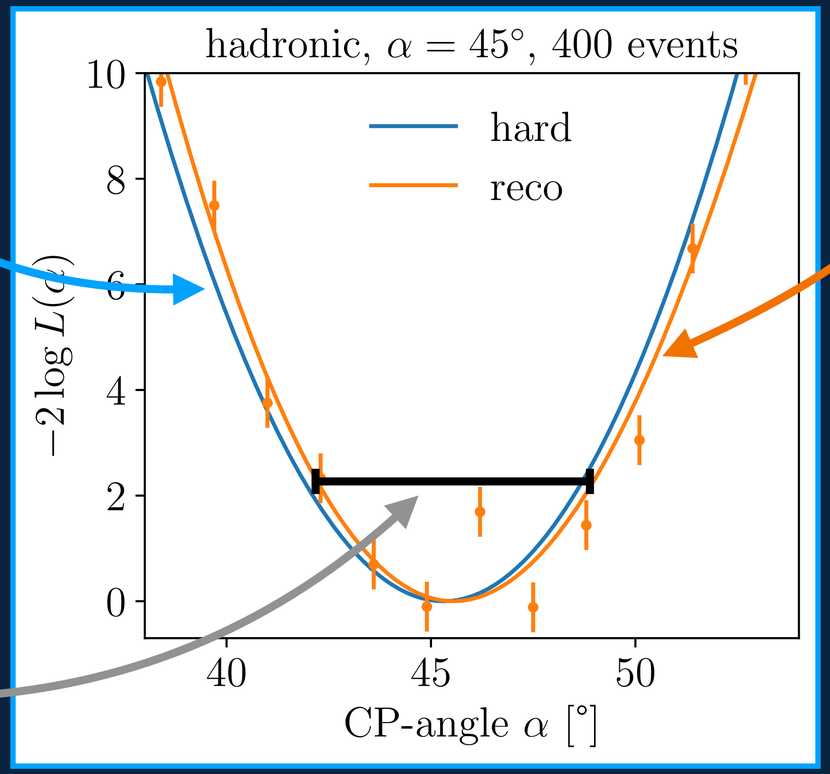
Anomalous coupling with **CP-angle α**

Uncertainties from training of neural network?
→ Bayesian neural networks

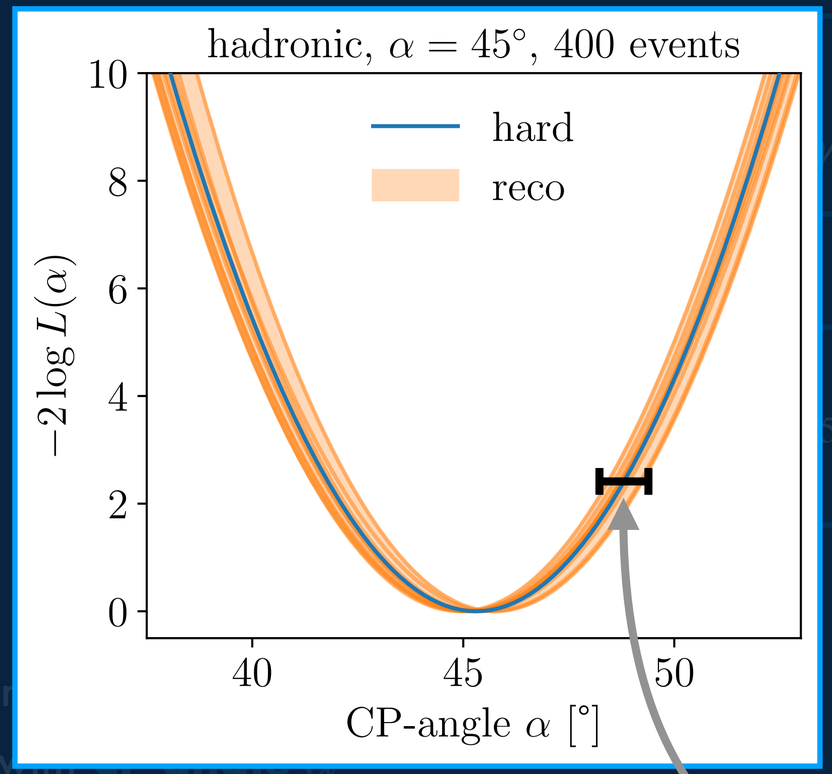
Matrix element method

Hard scattering truth for comparison

Heidelberg/Louvain [2210.00019, 23XX.xxxxx]



Result from MEM



Statistical uncertainty

Systematic uncertainty from training

Uncertainties from training of neural network?
→ Bayesian neural networks

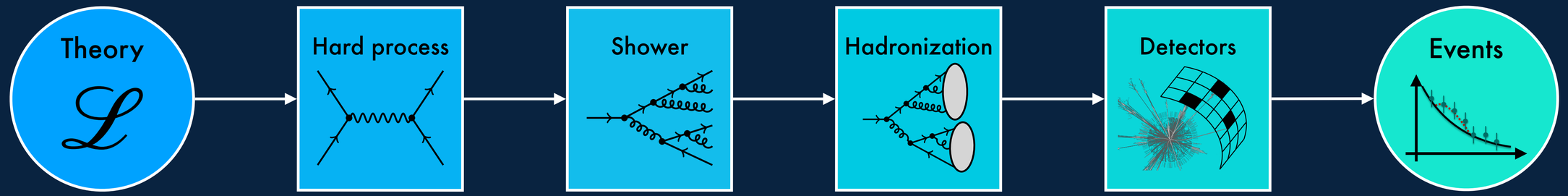
Summary and Outlook

Summary

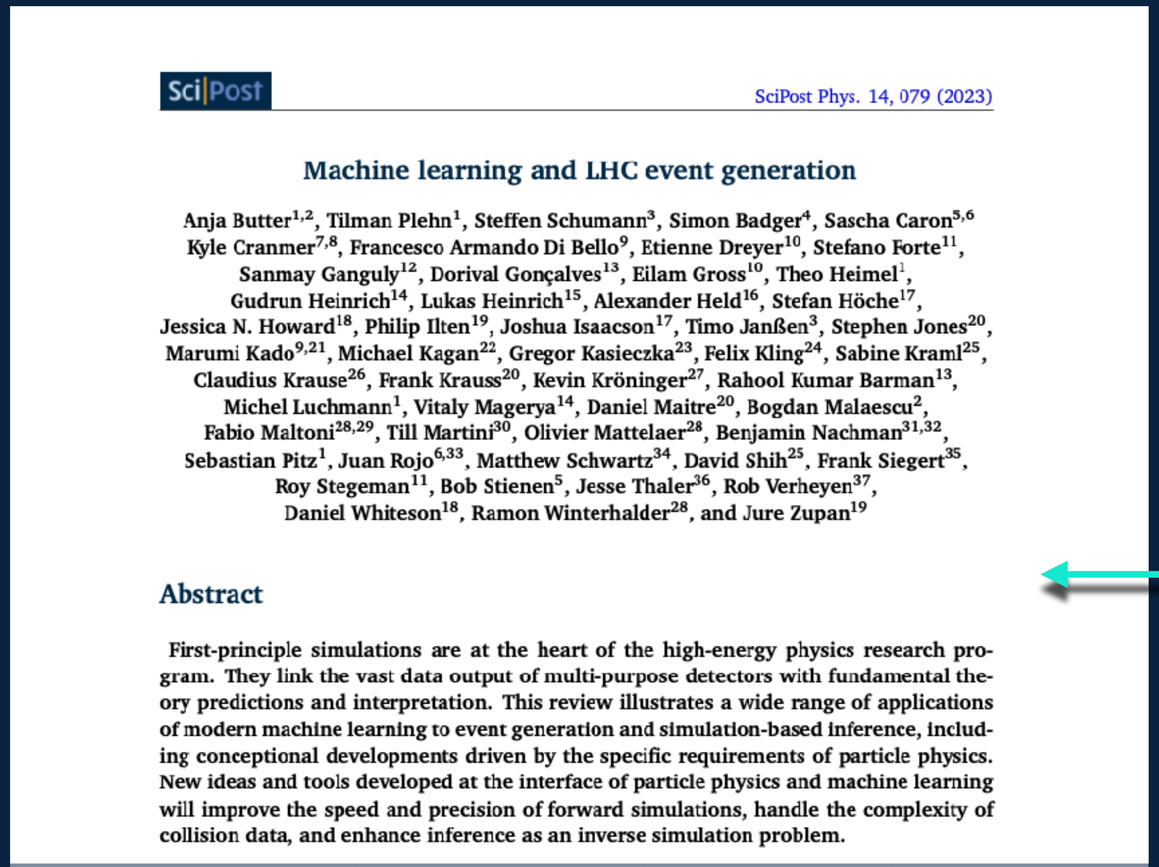
- DGMs provide **fast** and **precise** simulations
- Flows (+ Transformers) provide statistically **well-defined likelihoods** for inference
- Account for **uncertainties** with **Bayesian neural networks**

Outlook

- Full integration of DGMs into **standard tools**
- Make everything run on the **GPU and differentiable** (MadJax - Heinrich et al. [2203.00057])



Summary and Outlook



SciPost Phys. 14, 079 (2023)

Machine learning and LHC event generation

Anja Butter[1,2], Tilman Plehn[1], Steffen Schumann[3], Simon Badger[4], Sascha Caron[5,6], Kyle Cranmer[7,8], Francesco Armando Di Bello[9], Etienne Dreyer[10], Stefano Forte[11], Sanmay Ganguly[12], Dorival Gonçalves[13], Eilam Gross[10], Theo Heimel[1], Gudrun Heinrich[14], Lukas Heinrich[15], Alexander Held[16], Stefan Höche[17], Jessica N. Howard[18], Philip Ilten[19], Joshua Isaacson[17], Timo Janßen[3], Stephen Jones[20], Marumi Kado[9,21], Michael Kagan[22], Gregor Kasieczka[23], Felix Kling[24], Sabine Kraml[25], Claudius Krause[26], Frank Krauss[20], Kevin Kröninger[27], Rahool Kumar Barman[13], Michel Luchmann[1], Vitaly Magerya[14], Daniel Maitre[20], Bogdan Malaescu[2], Fabio Maltoni[28,29], Till Martini[30], Olivier Mattelaer[28], Benjamin Nachman[31,32], Sebastian Pitz[1], Juan Rojo[6,33], Matthew Schwartz[34], David Shih[25], Frank Siegert[35], Roy Stegeman[11], Bob Stienen[5], Jesse Thaler[36], Rob Verheyen[37], Daniel Whiteson[18], Ramon Winterhalder[28], and Jure Zupan[19]

Abstract

First-principle simulations are at the heart of the high-energy physics research program. They link the vast data output of multi-purpose detectors with fundamental theory predictions and interpretation. This review illustrates a wide range of applications of modern machine learning to event generation and simulation-based inference, including conceptional developments driven by the specific requirements of particle physics. New ideas and tools developed at the interface of particle physics and machine learning will improve the speed and precision of forward simulations, handle the complexity of collision data, and enhance inference as an inverse simulation problem.

Outlook

- Full integration of DGMs into **standard tools**
- Make everything run on the **GPU and differentiable** (MadJax - Heinrich et al. [2203.00057])
- More details in our **Snowmass report**
- Stay tuned for many other **ML4HEP applications**

HEPML