Constraining Effective Field Theories with Machine Learning

Johann Brehmer

NYU

HET seminar, BNL
June 27, 2018

NYU

MSDSE
Moore - Sloan
Data Science Environments

Kyle Cranmer

Gilles Louppe

Juan Pavez

Constraining Effective Field Theories with Machine Learning
arXiv:1805.00013

A Guide to Constraining Effective Field Theories with Machine Learning
arXiv:1805.00020

Mining gold from implicit models to improve likelihood-free inference
arXiv:1805.12244

Thanks to Kyle and Gilles for ~~letting me steal~~ inspiring many slides!

Long Island Rail Road

Key

Full Time rail station

Accessible station

Part Time rail station

Major Transit Hub

Long Island Sound

Port Jefferson Steamship Co.

Port Jeffers

Stony Brook

PORT JEFFERSON BRANCH

Kings Park

St. James

SUFFOL

Constraining EFTs with ML

Likelihood-free inference

Locust Valley

Oyster Bay

Glen Cove

Glen Street

OYSTER BAY BRANCH

Sea Cliff

Glen Head

Greenvale

Port Washington

PORT WASHINGTON BRANCH

Plandome

Manhasset

Great Neck

Little Neck

Douglaston

Bayside

Auburndale

Broadway

Murray Hill

Flushing-Main St

Mets-Willets Point

NASSAU

Roslyn

Albertson

East Williston

Hicksville

Westbury

Carle Place

Bethpage

Huntington

Cold Spring Harbor

Syosset

Greenlawn

Deer Park

Wyandanch

Pinelawn

Farmingdale

Central Islip

Brentwood

Ronkonkoma

MA BRANCH

Great River

Oakdale

Sayville

Islip

Bay Shore

Sayville Ferry Service

Fire Island Ferries, Inc.

Mineola

Merillon Av

New Hyde Pk

Floral Park

Bellerose

Queens Village

Hollis

QUEENS

Forest Hills

Kew Gardens

Woodside

Hunterspoint Av

Long Island City

Penn Station

HEMPSTEAD BRANCH

Hempstead

West Hempstead

WEST HEMPSTEAD BRANCH

Hempstead Gardens

Garden City

Nassau Blvd

Life Press

Established methods

Mining gold from the simulator

CITY TERMINAL ZONE

Atlantic Terminal Brooklyn

Nostrand Av

East New York

AIRTRAIN JFK

Locust Manor

Laurelton

BROOKLYN

Jamaica Bay

Gibson

Hewlett

Woodmere

Cedarhurst

Lawrence

Inwood

Far Rockaway

FAR ROCKAWAY BRANCH

Lynbrook

Centre Av

Rockville Centre

E. Rockaway

Oceanside

Island Park

Long Beach

LONG BEACH BRANCH

Baldwin

Freeport

Merrick

Bellmore

Wantagh

Seaford

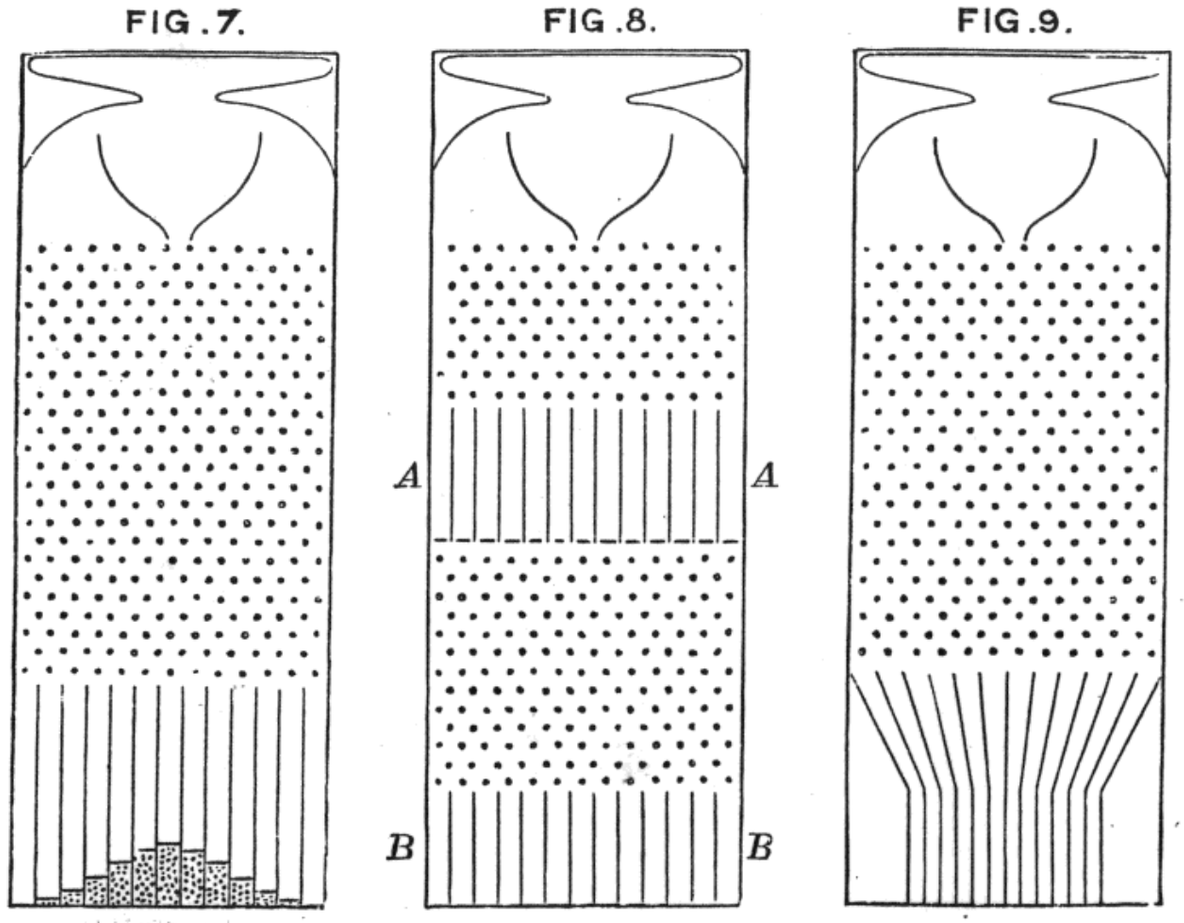
Massapequa

Massapequa Park

Atlantic Ocean

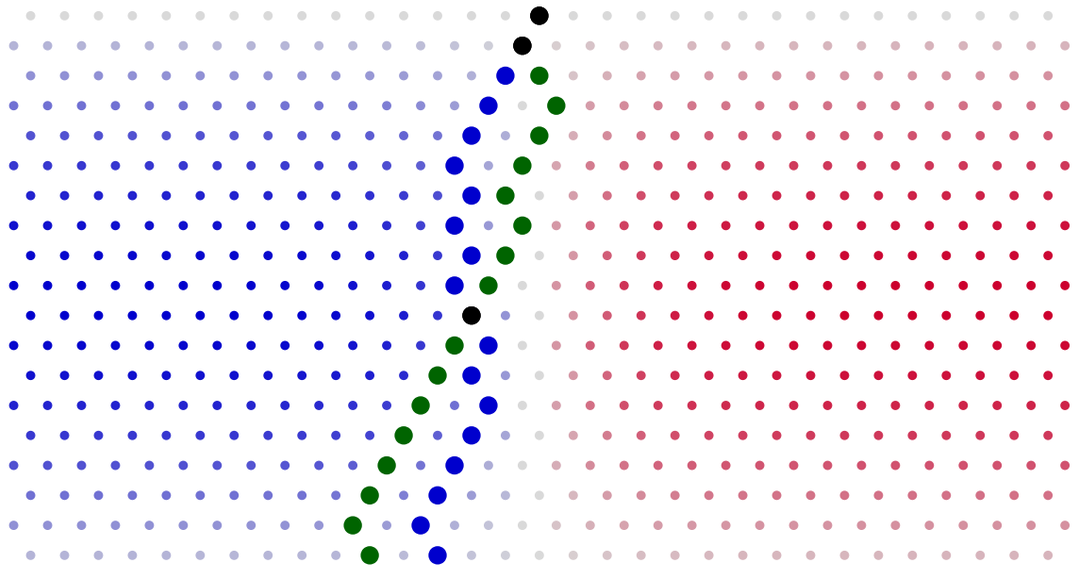
Likelihood-free inference

The Galton board



[Source: F. Galton 1889]

Probabilities



Probability of ending in bin x:

$$p(x) = \int \mathrm{d}z \; p(x, z)$$

Sum over all trajectories ("latent variables")

Probability of each path z from start to x

The generalized Galton board

- What if probability to go left at a nail is not always 0.5, but some (known) function of some parameters θ ?
- Prediction: given θ, generate observations $\{x_i\}$ … just drop balls!
- Inference: given observations $\{x_i\}$, what are the most likely values for θ ?

The generalized Galton board

- What if probability to go left at a nail is not always 0.5, but some (known) function of some parameters θ ?
- Prediction: given θ, generate observations $\{x_i\}$ … just drop balls!
- Inference: given observations $\{x_i\}$, what are the most likely values for θ ?

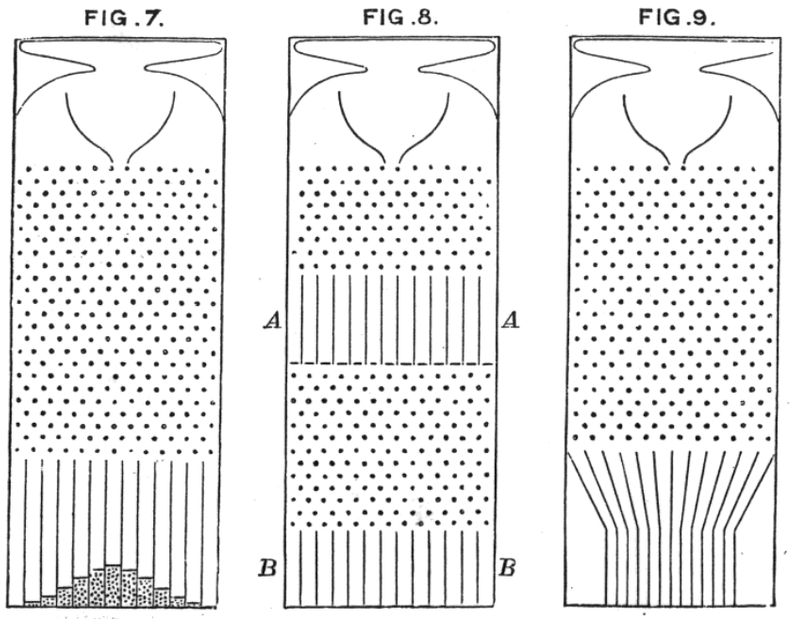
"Easy" problem if we can evaluate likelihood $p(x_i|\theta)$. But

$$p(x|\theta) = \int \mathrm{d}z \; p(x, z|\theta)$$

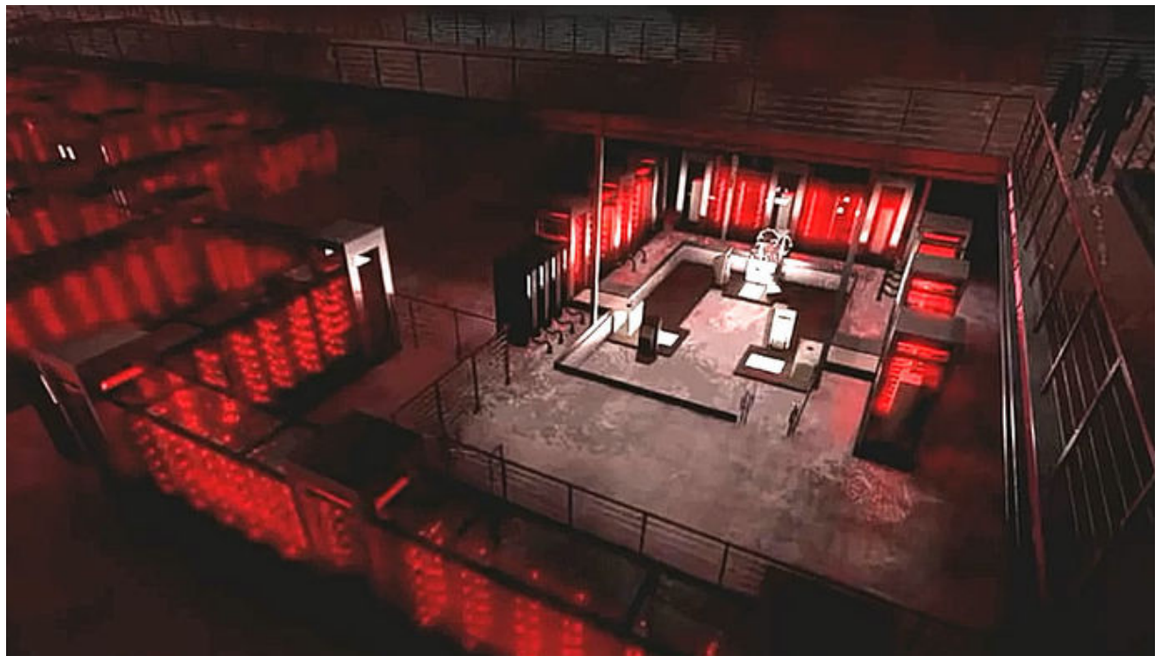
The number of possible paths z can be huge, and it becomes impossible to calculate the integral

⇒ Likelihood not tractable, only implicitly defined through "simulator"

Galton board: metaphor for simulator-based science



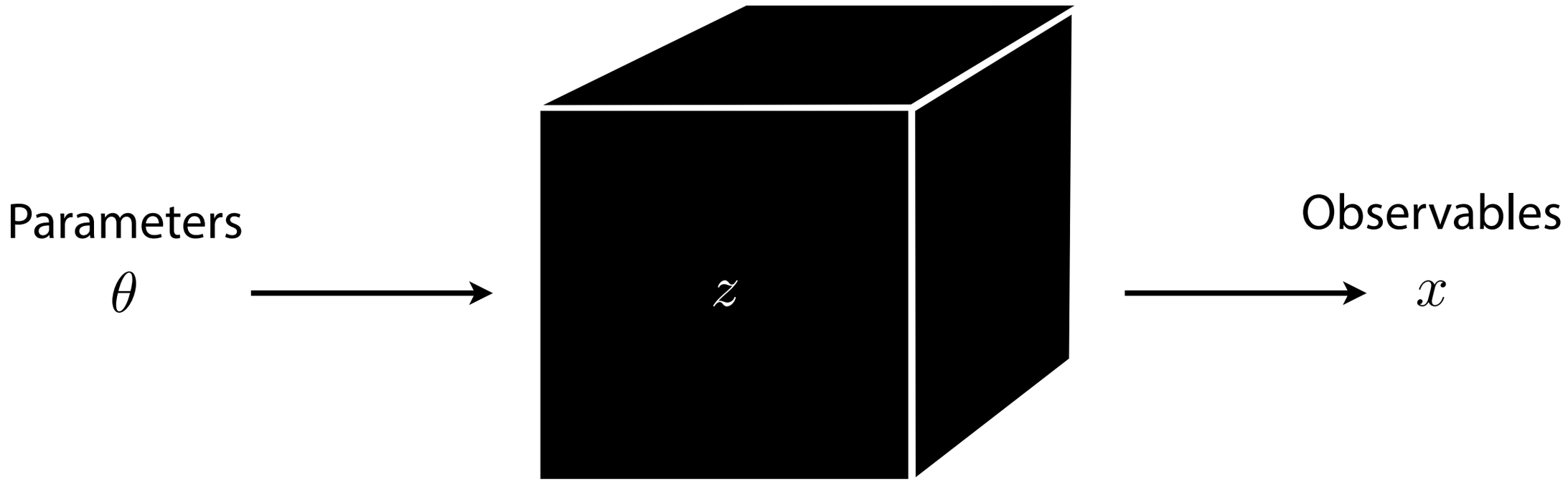
[Source: F. Galton 1889]



[Source: HBO 2018]

Galton board device	$\longrightarrow$	Computer simulation
Parameters θ	$\longrightarrow$	Parameters of interest θ
Bins x	$\longrightarrow$	Observables x
Path z	$\longrightarrow$	Latent variables z (stochastic execution trace through simulator)

Likelihood-free inference



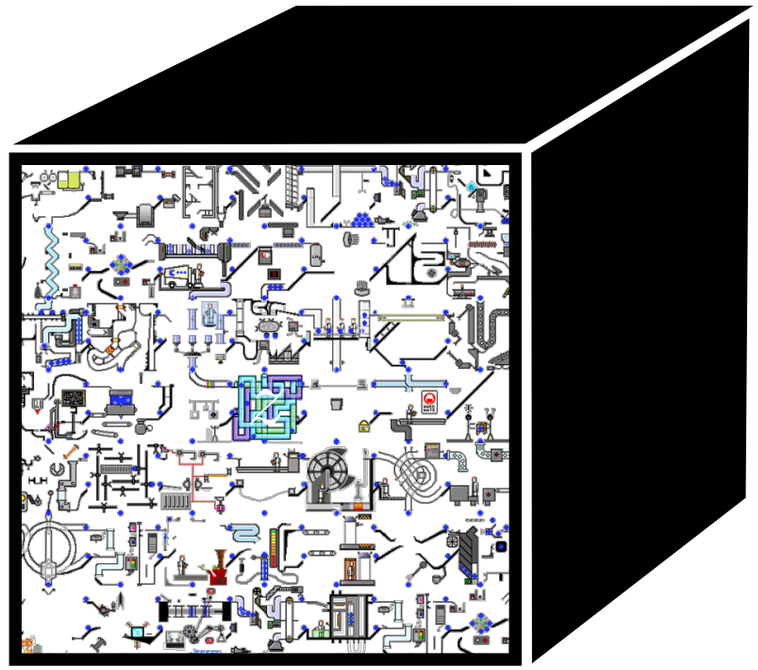
Prediction (simulation):

- Well-understood mechanistic model
- Simulator can generate samples

Likelihood-free inference

Parameters

θ



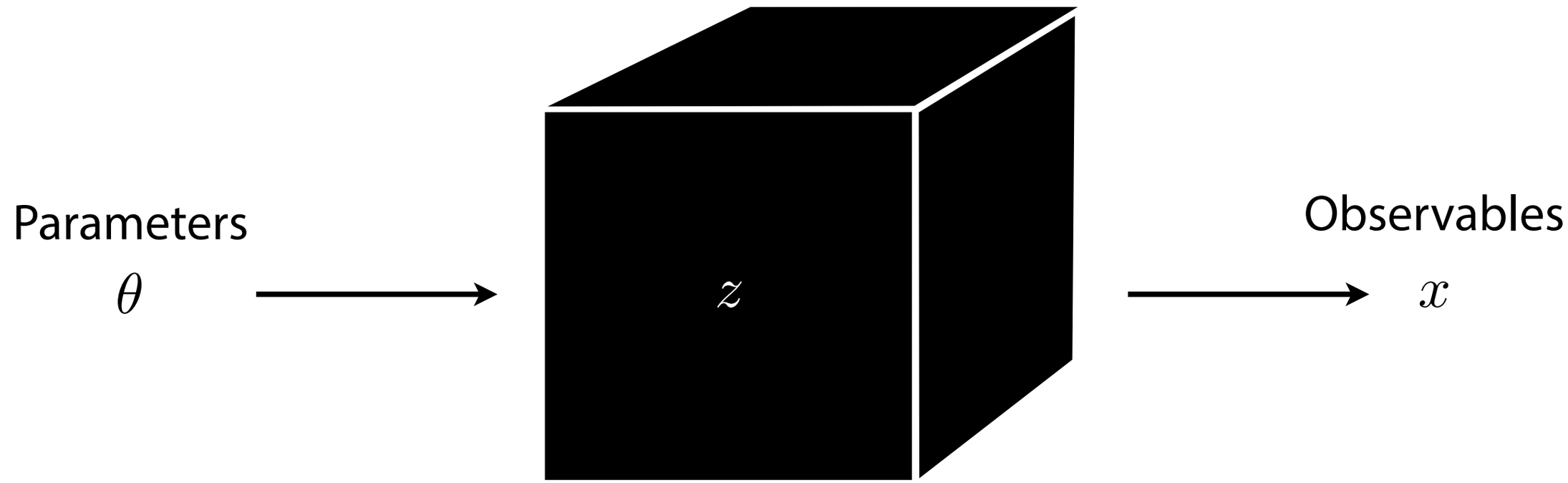
Observables

x

Prediction (simulation):

- Well-understood mechanistic model
- Simulator can generate samples

Likelihood-free inference



Prediction (simulation):

- Well-understood mechanistic model
- Simulator can generate samples

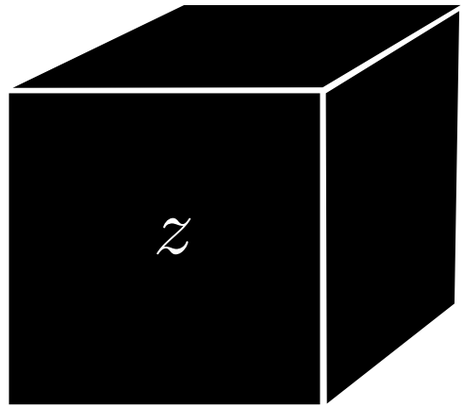
Inference:

- Likelihood function $p(x|\theta)$ is intractable
- Goal: estimator $\hat{p}(x|\theta)$

Cosmological N-body simulations

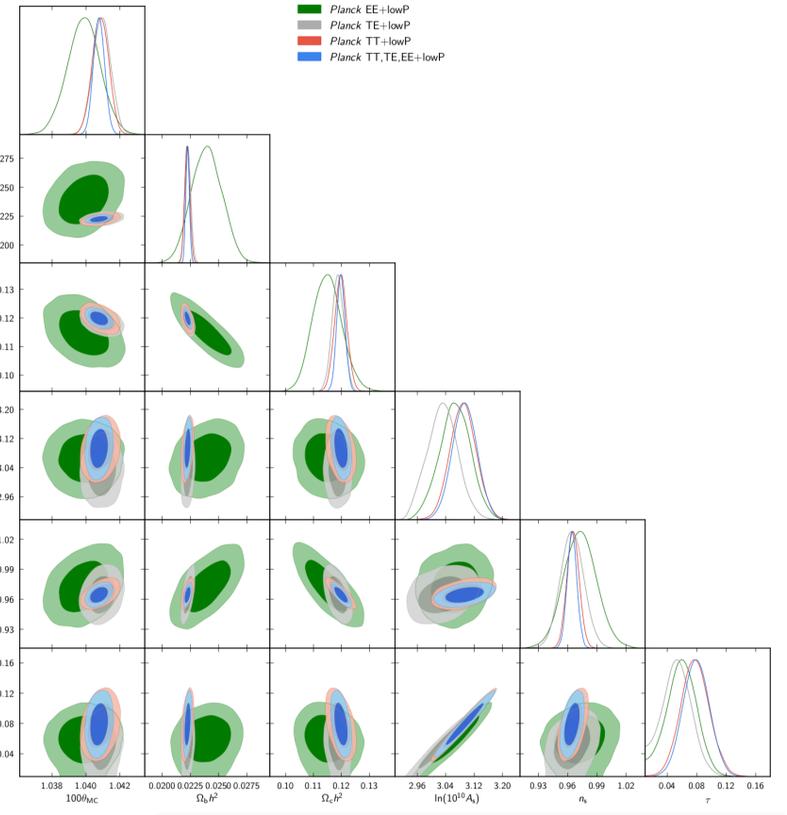
Parameters

θ

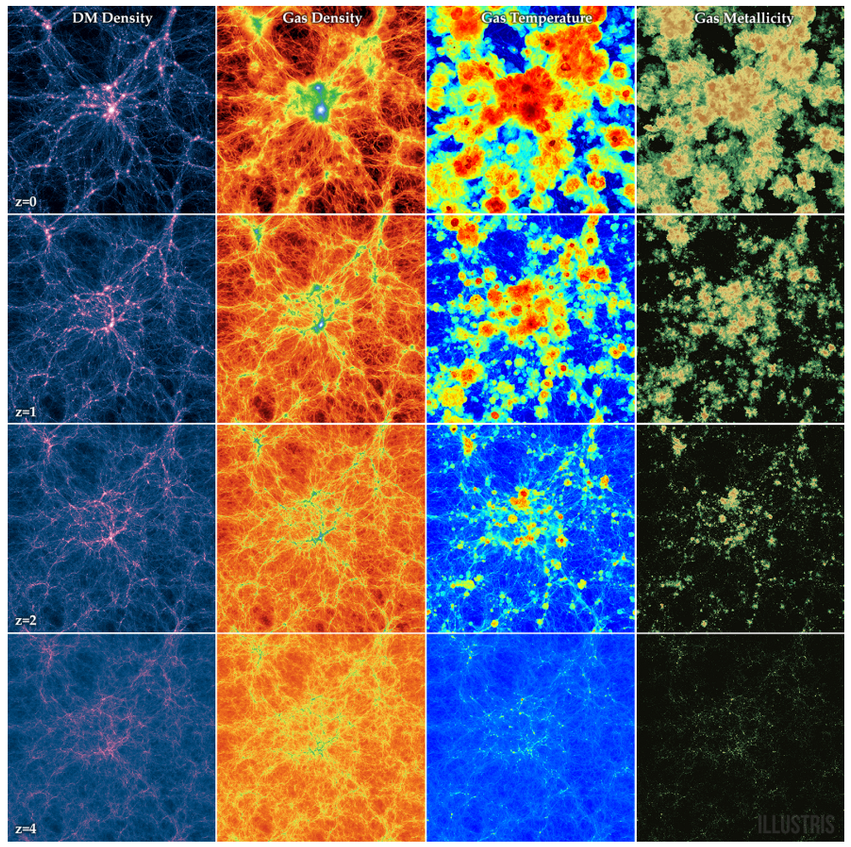


Observables

x



[Source: Planck 1502.01589]

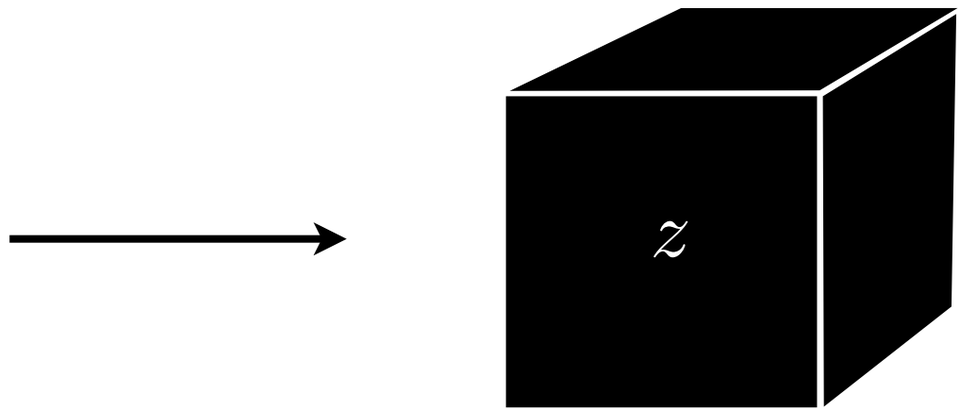


[Source: Illustris 1405.2921]

Computational topography

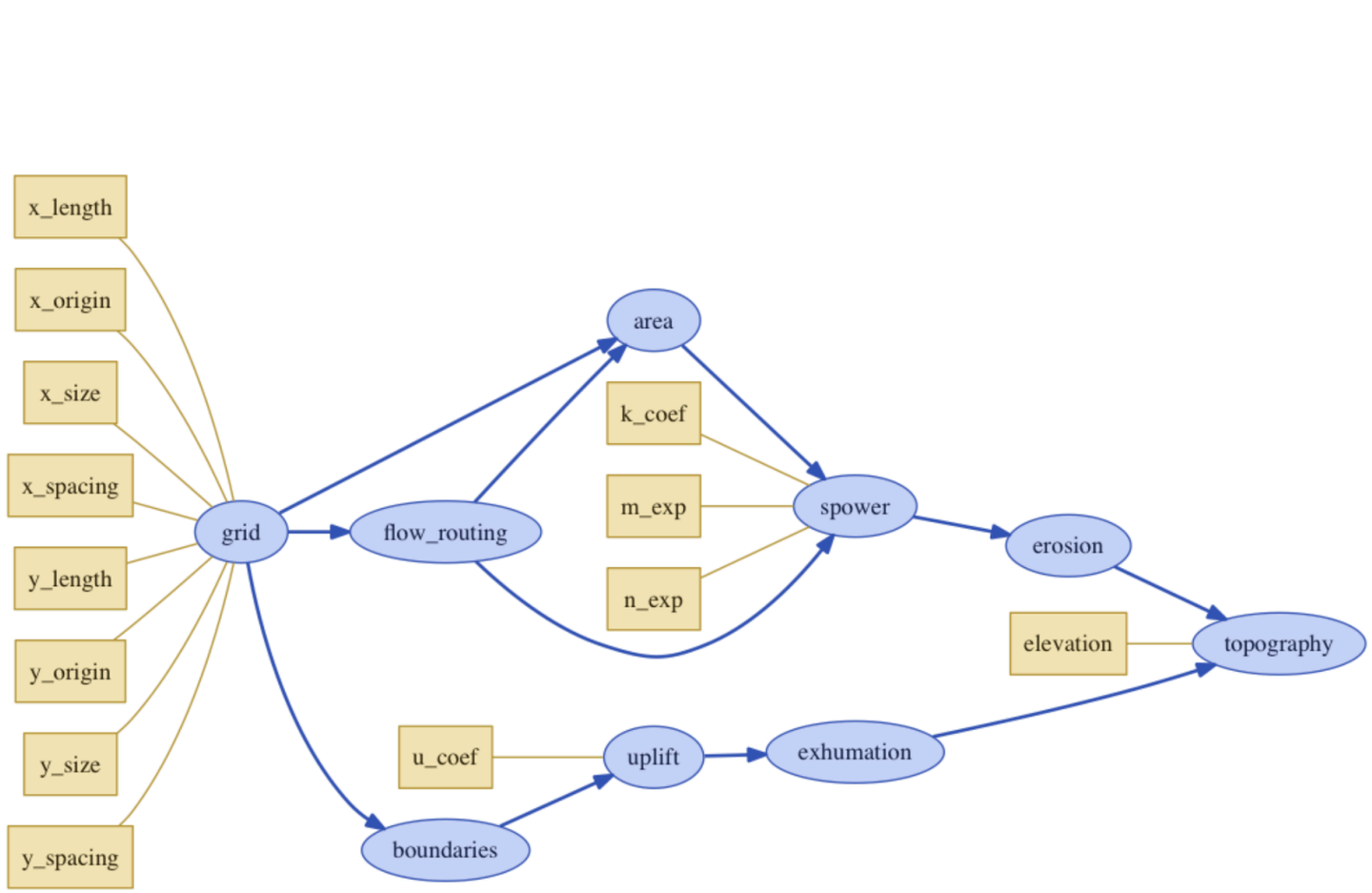
Parameters

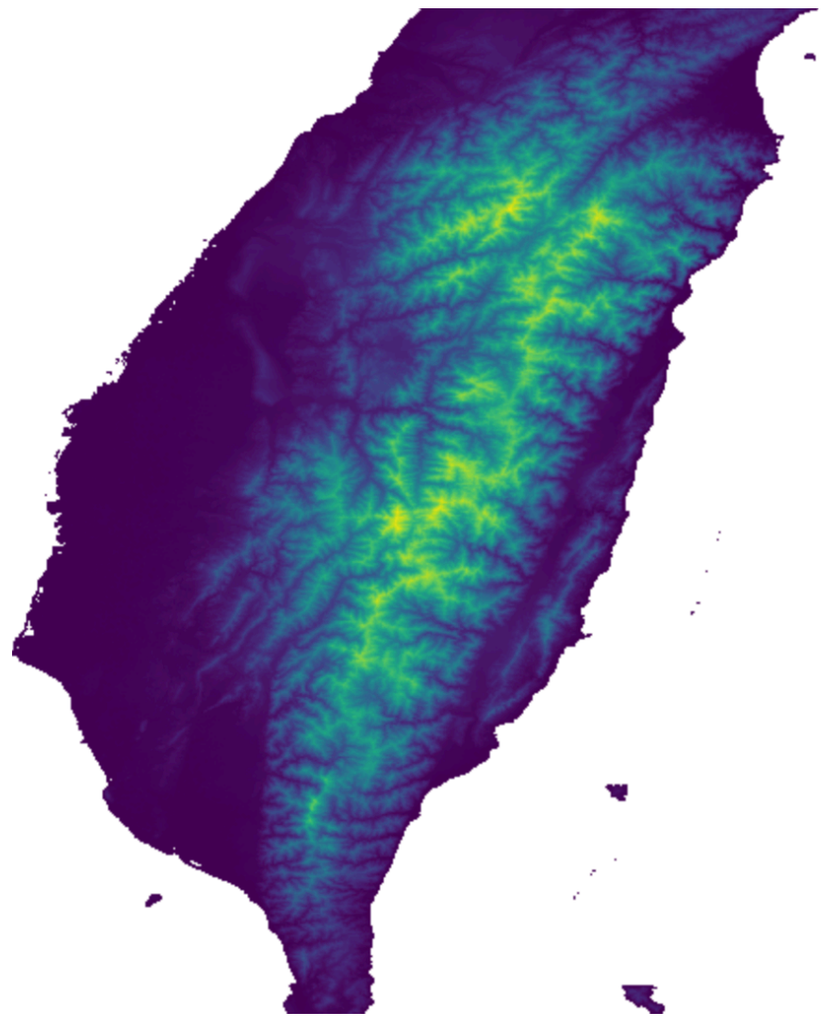
θ



Observables

x



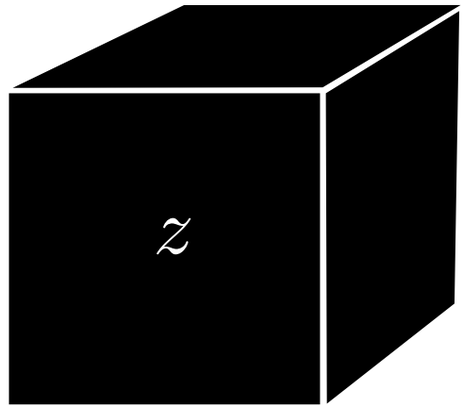


[Source: Benoit Bovy]

Epidemiology

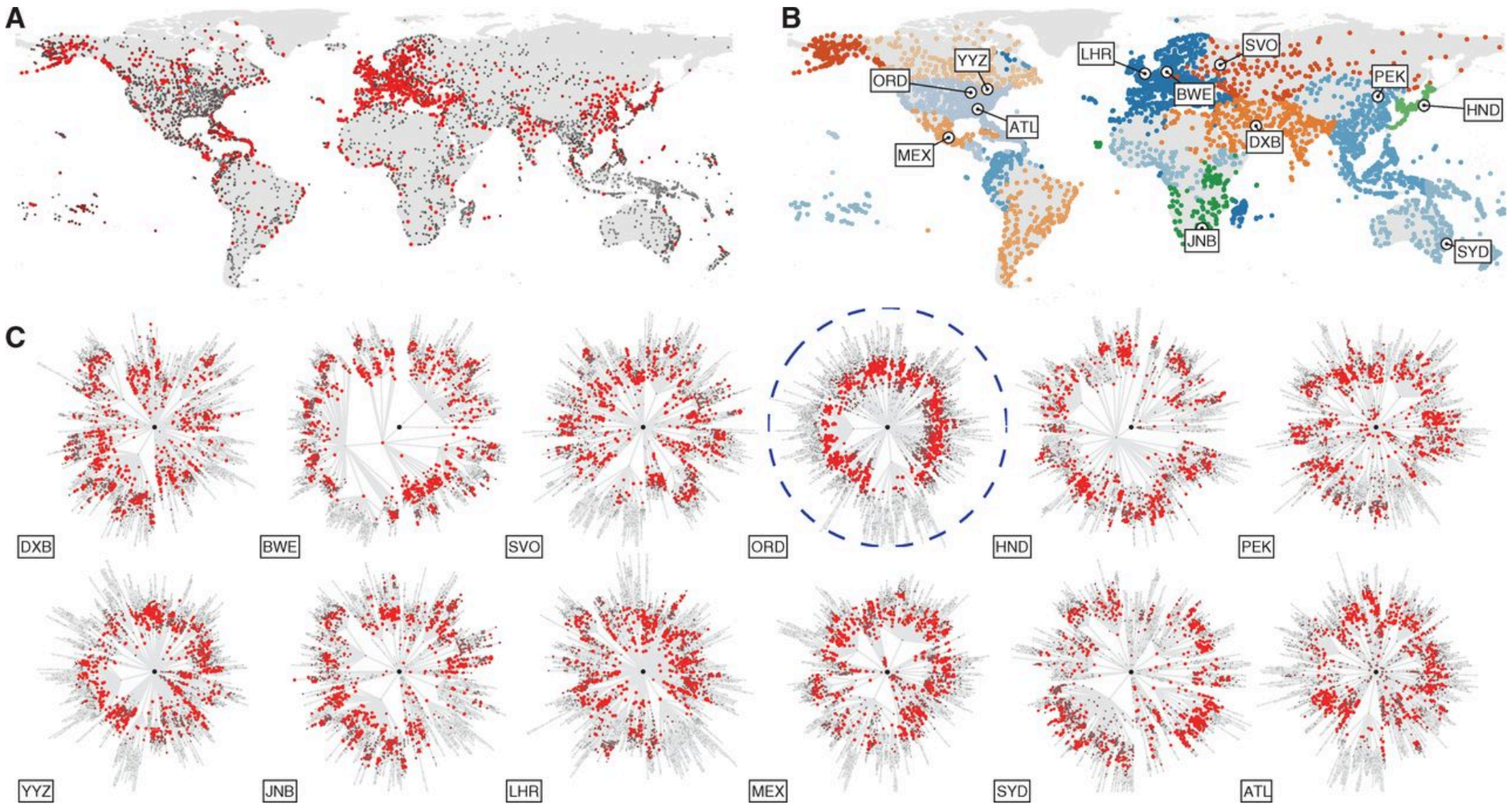
Parameters

θ



Observables

x



[Source: D. Brockmann, D. Helbing 2013]

Particle physics

Parameters of interest

Theory parameters

θ

Evolution

Particle physics

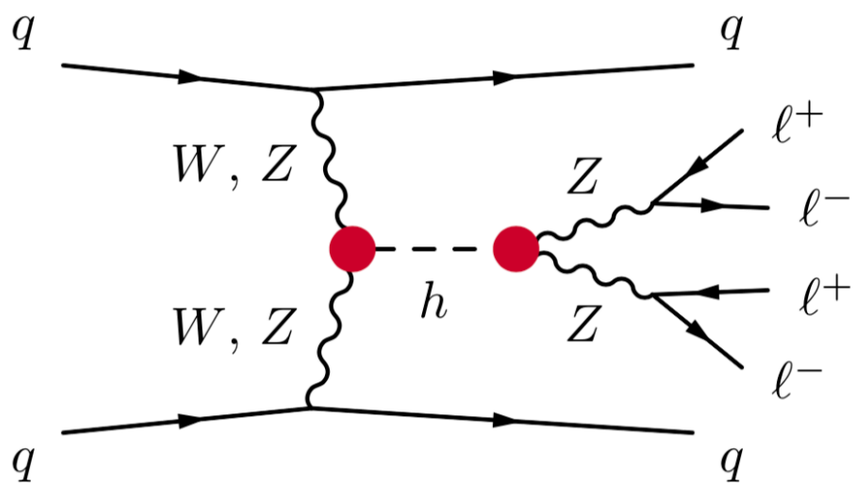
Latent variables

Parameters of interest

Parton-level momenta

Theory parameters

$z_p \longleftarrow \theta$



Evolution

Particle physics

Latent variables

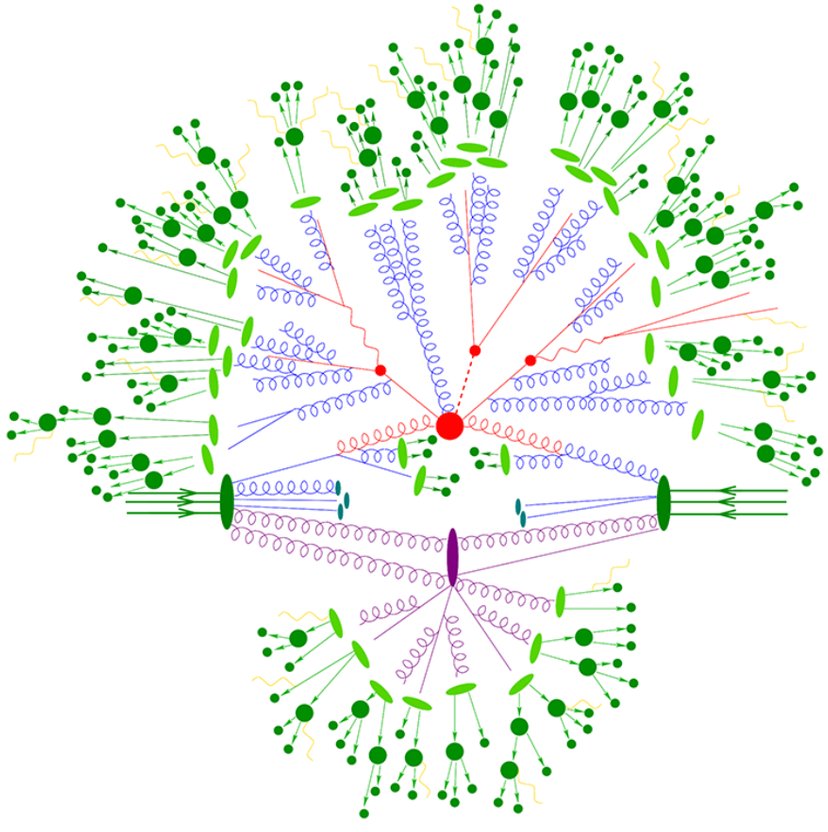
Parameters of interest

Shower splittings

Parton-level momenta

Theory parameters

$z_s \longleftarrow z_p \longleftarrow \theta$



[Source: F. Krauss]

Evolution

Particle physics

Latent variables

Parameters of interest

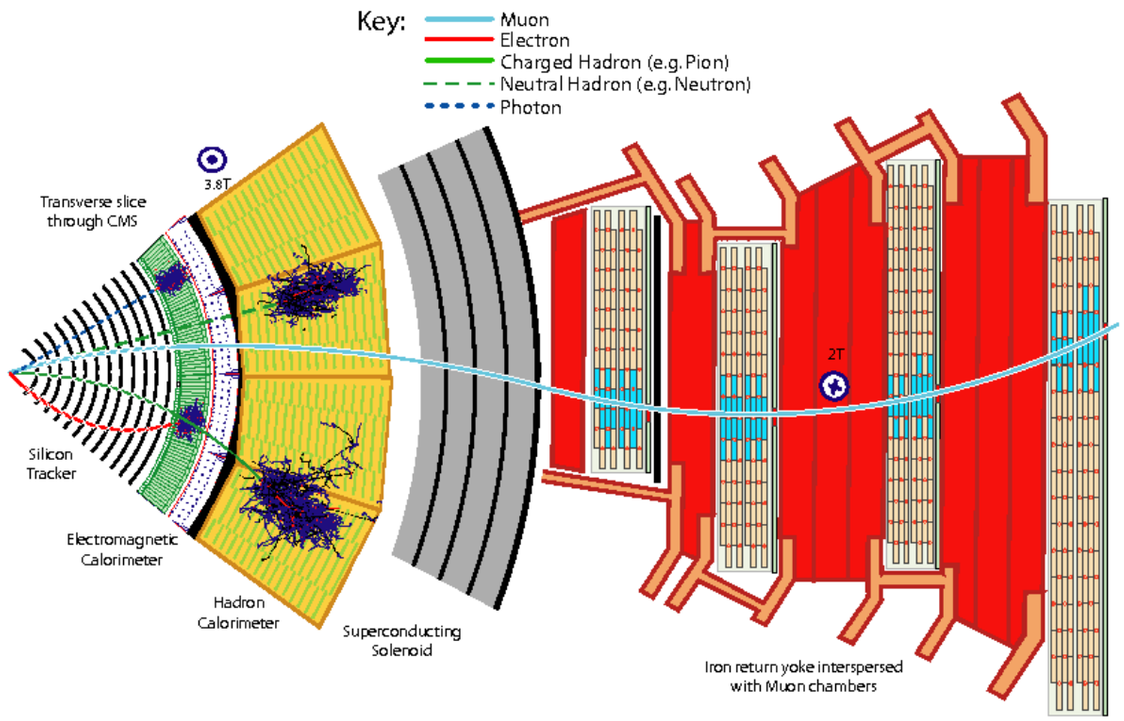
Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

$z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$



[Source: CMS]

Evolution

Particle physics

Features

Latent variables

Parameters of interest

Observables

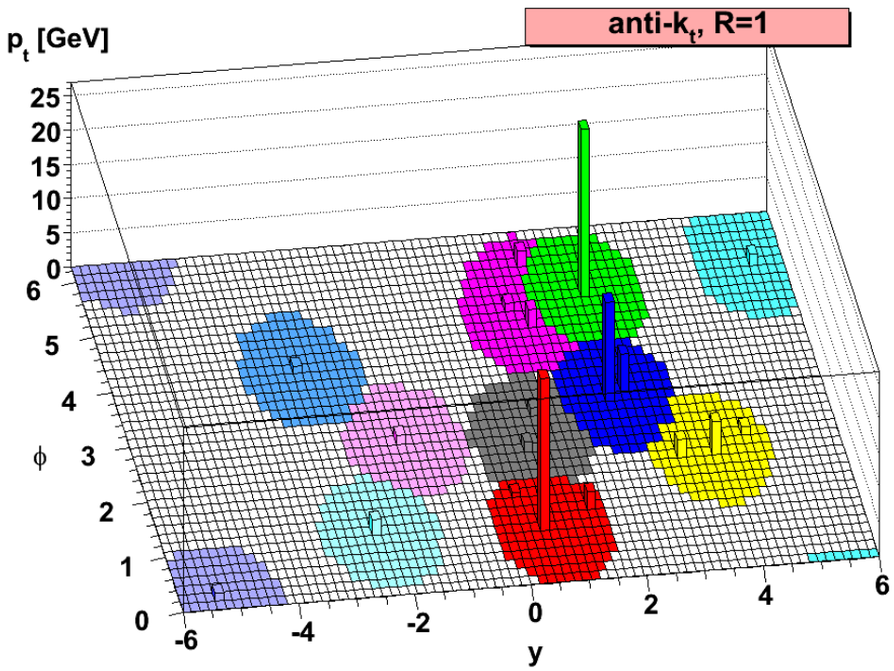
Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$



[Source: M. Cacciari, G. Salam, G. Soyez 0802.1189]

Evolution

Particle physics

Features

Latent variables

Parameters of interest

Observables

Detector interactions

Shower splittings

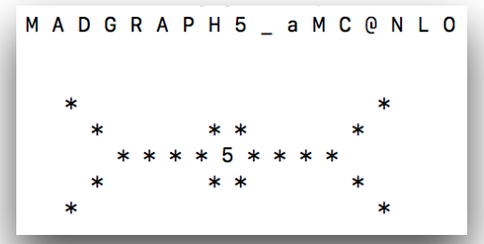
Parton-level momenta

Theory parameters

$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$

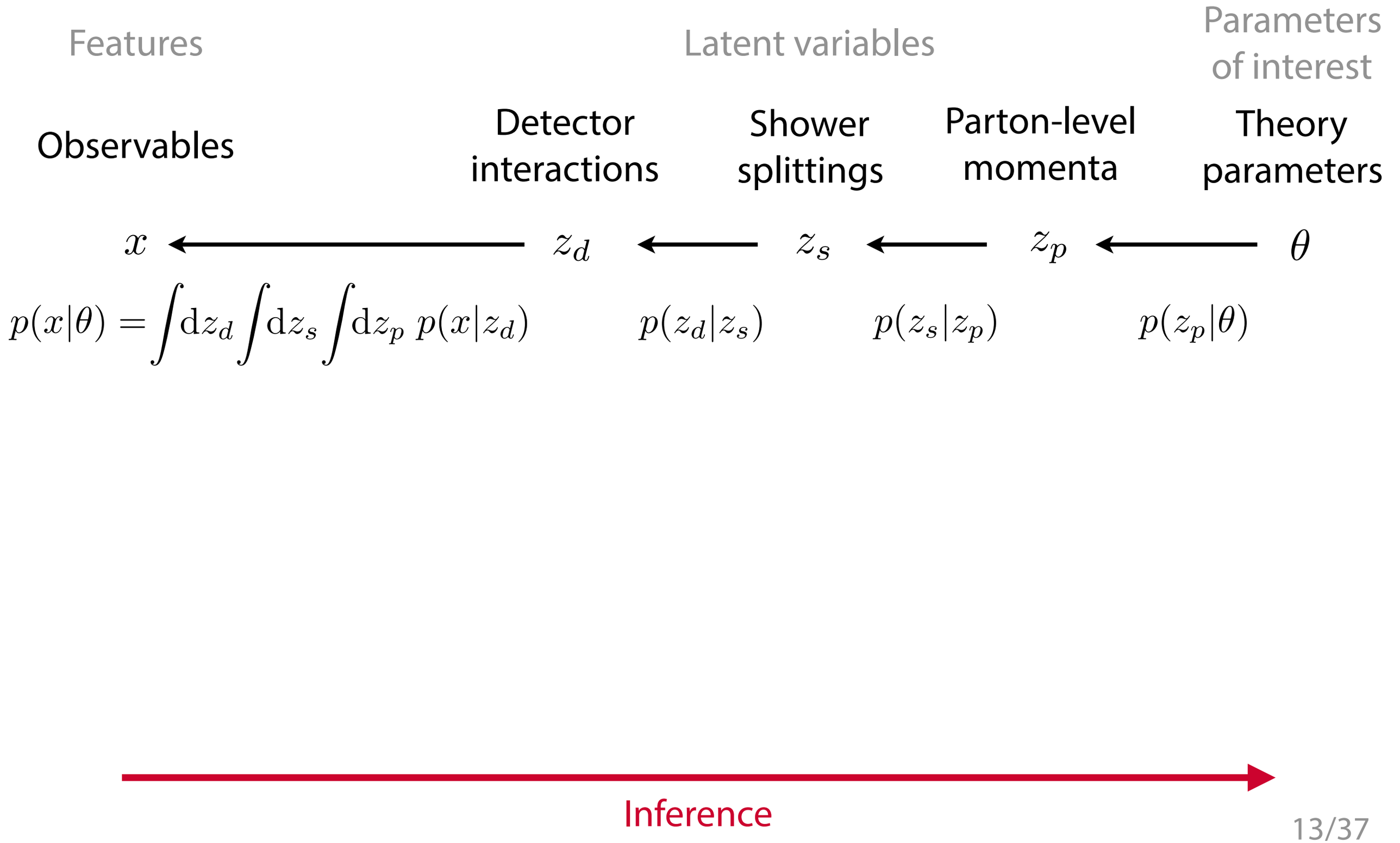
Sample from

$p(x|z_d)$ $p(z_d|z_s)$ $p(z_s|z_p)$ $p(z_p|\theta)$



Prediction (simulation)

Particle physics



Particle physics

Features

Latent variables

Parameters of interest

Observables

Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

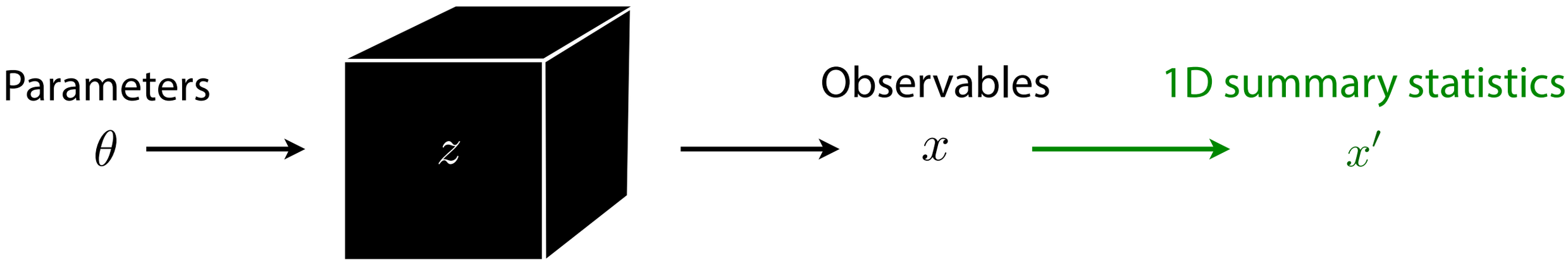
$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \, p(x|z_d) \quad p(z_d|z_s) \quad p(z_s|z_p) \quad p(z_p|\theta)$$

Infeasible to calculate the integral over this enormous space

Inference

Why has that not stopped us so far?

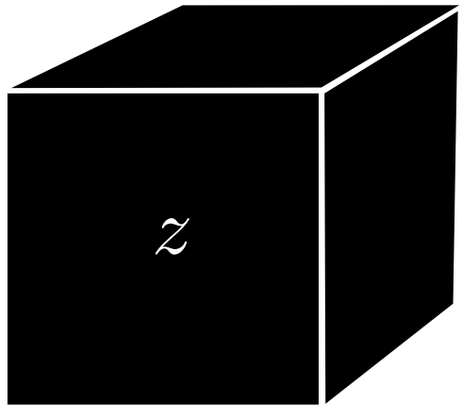
The physicist's way



The physicist's way

Parameters

θ

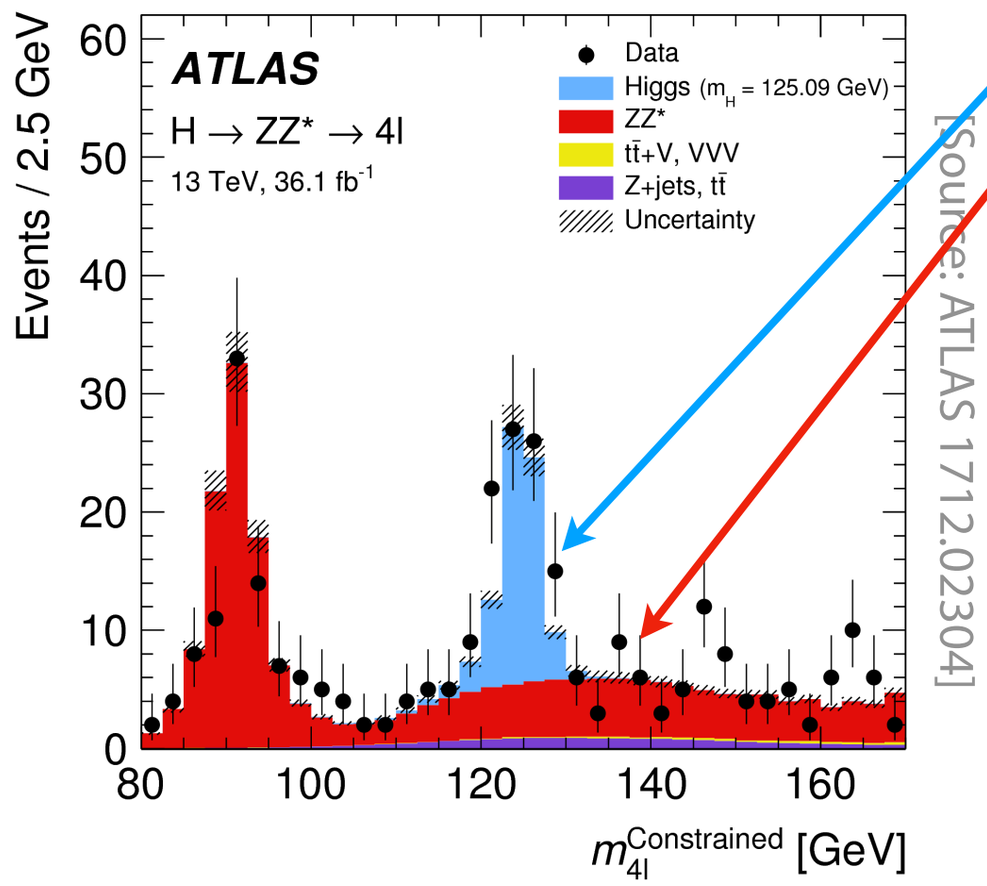


Observables

x

1D summary statistics

x'



[Source: ATLAS 1712.02304]

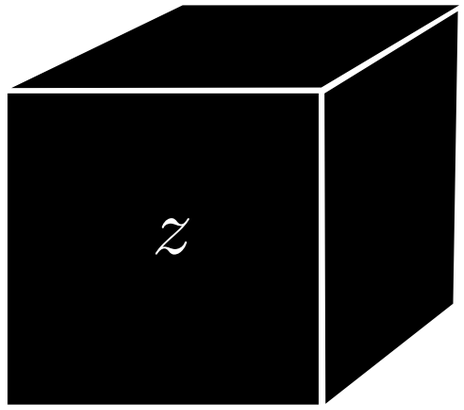
Run simulator for different θ, fill histograms

x'

The physicist's way

Parameters

θ

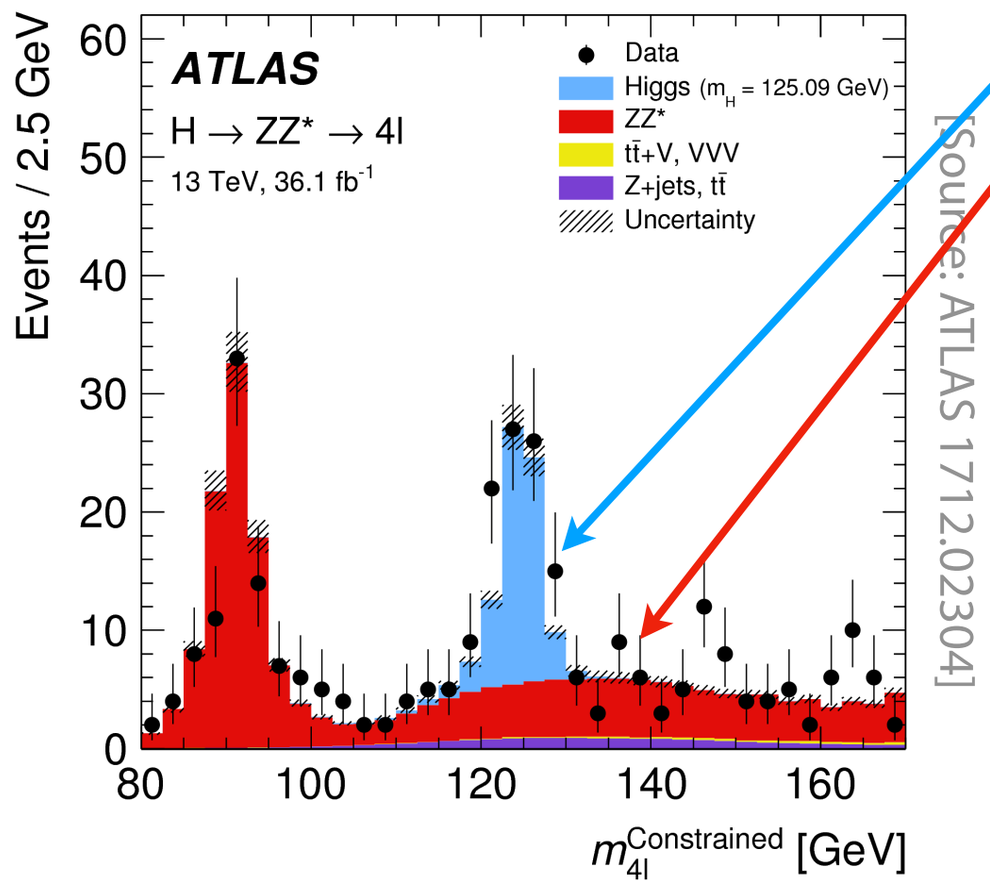


Observables

x

1D summary statistics

x'



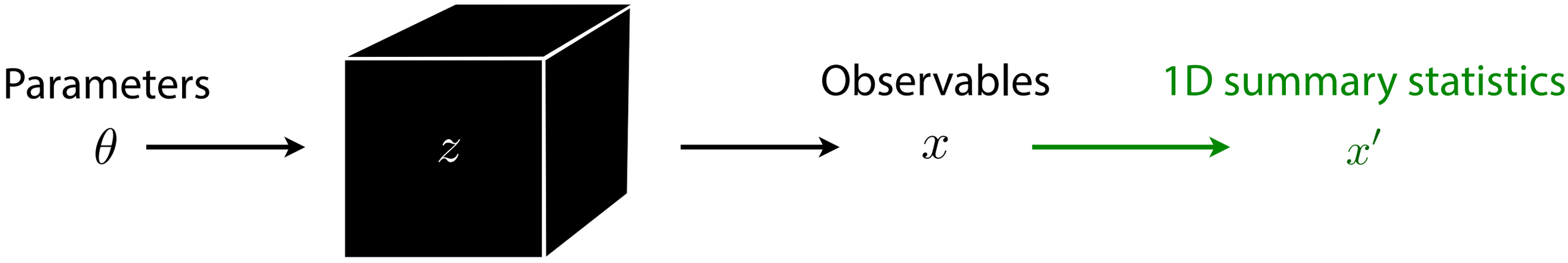
[Source: ATLAS 1712.02304]

Run simulator for different θ, fill histograms

$$\hat{p}(x|\theta) = p(x'|\theta) =$$

x'

Does the histogram method scale?



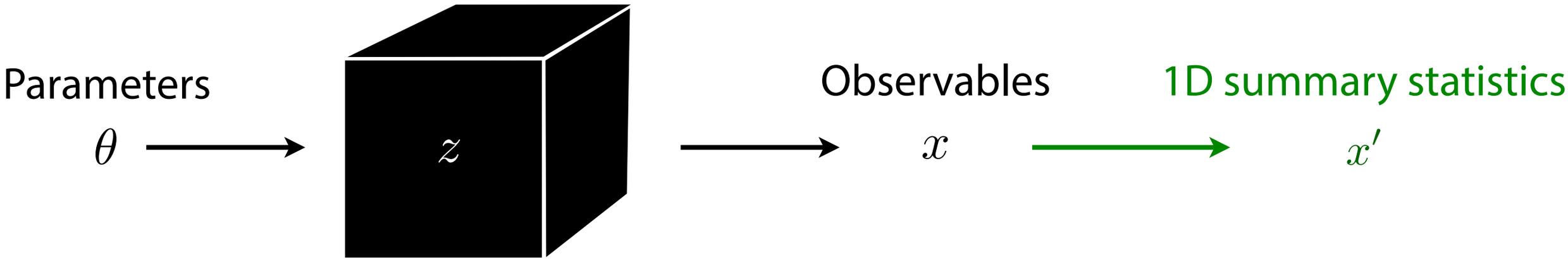
- Choosing x' is difficult and problem-dependent

- Often there is no single good variable — compressing to any x' loses information!

  [JB, K. Cranmer, F. Kling, T. Plehn 1612.05261;
  JB, F. Kling, T. Plehn , T. Tait 1712.02350]

- Ideally: analyze high-dimensional x including all correlations

Does the histogram method scale?

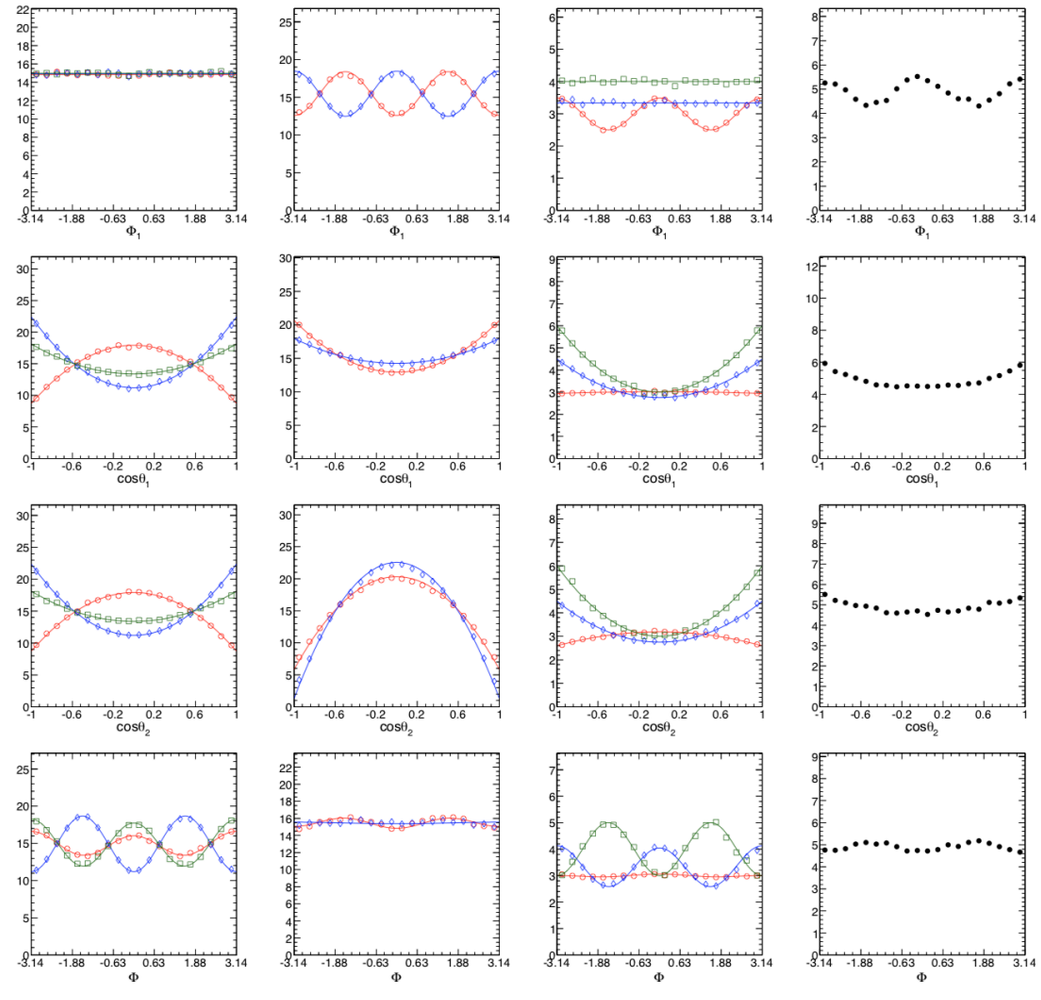


- Choosing x' is difficult and problem-dependent

- Often there is no single good variable — compressing to any x' loses information!

  [JB, K. Cranmer, F. Kling, T. Plehn 1612.05261; JB, F. Kling, T. Plehn , T. Tait 1712.02350]

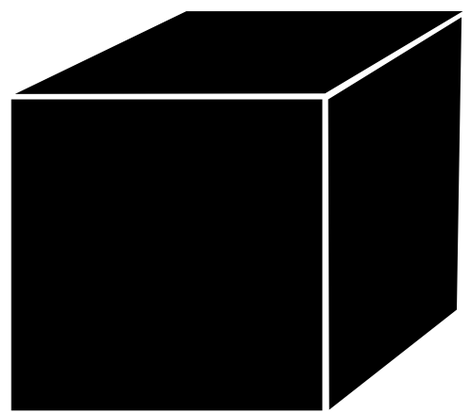
- Ideally: analyze high-dimensional x including all correlations



[Source: Bolognesi et al. 1208.4018]

An incomplete list of established methods

- Histograms of observables
  Summary statistics
- Approximate Bayesian Computation
  Summary statistics
- Machine Learning techniques
  Density networks, CARL, autoregressive models,
  normalizing flows, …

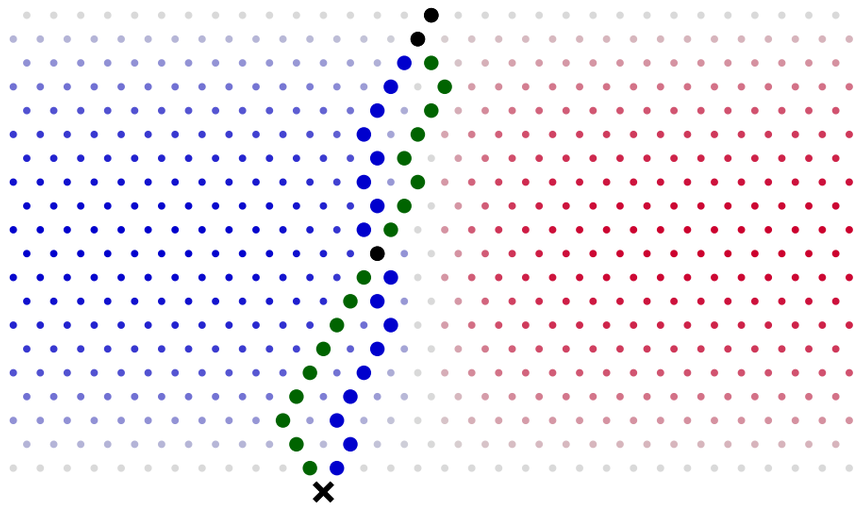


- Matrix Element Method / Optimal Observables
  Neglect or approximate shower + detector,
  explicitly calculate integral

$$\hat{p}(x|\theta) = \int \mathrm{d}z_p \, \tilde{p}(x|z_p) \, p(z_p|\theta)$$

Mining gold from the simulator

Back to the Galton board



- Likelihood

$$p(x|\theta) = \int \mathrm{d}z \; p(x, z|\theta)$$

is usually intractable because of the integral over all possible paths z

- But: we can calculate the probability of each individual path

$$p(x, z|\theta) = \prod_{\text{nails } i} p_i(x, z|\theta)$$

Turns out that that can be useful…

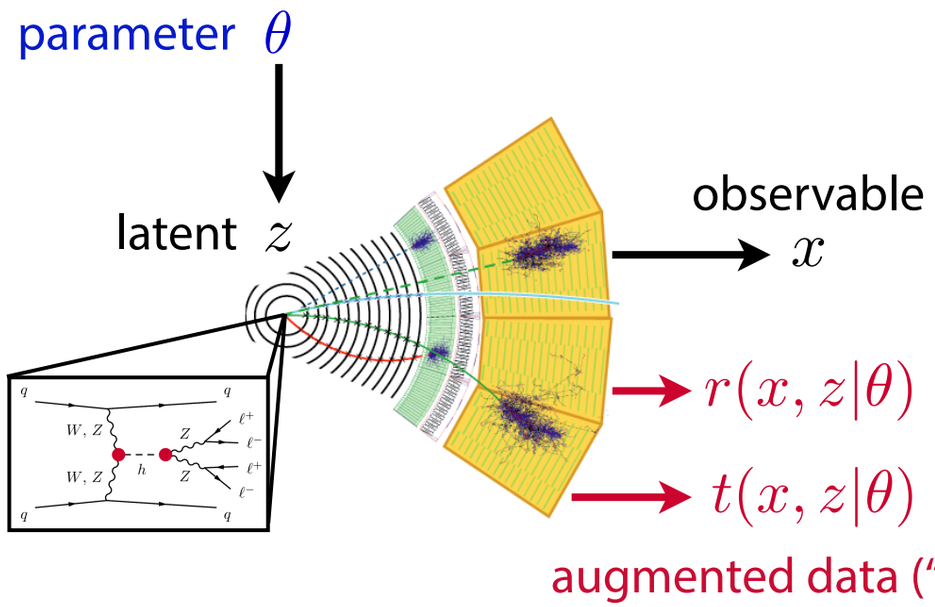
A new approach

- New algorithms for simulator-based inference
  - Key idea: Extract more information from simulator, use it with ML
  - Applicable to many (but not all) likelihood-free inference problems

- Sales pitch:
  - Higher-fidelity measurements with less training data
  - Scales well to many observables (no need to pick summary statistics) and parameters
  - No approximations in simulator needed (LHC: supports NLO, parton shower, full detector simulation)
  - Evaluation in microseconds

- Motivated by LHC legacy measurements, but applications far beyond physics

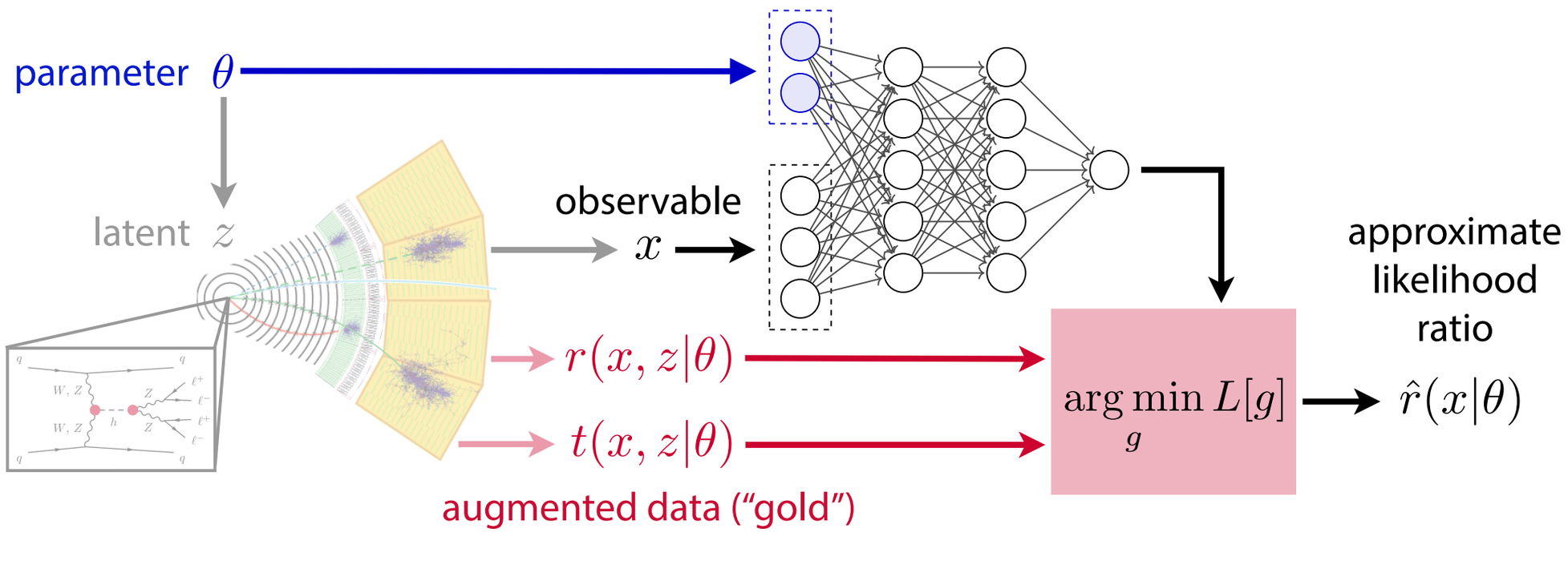
Bird's-eye view



1. Simulation

"Mining gold": Extract additional information from simulator

Bird's-eye view



1. Simulation

"Mining gold": Extract additional information from simulator

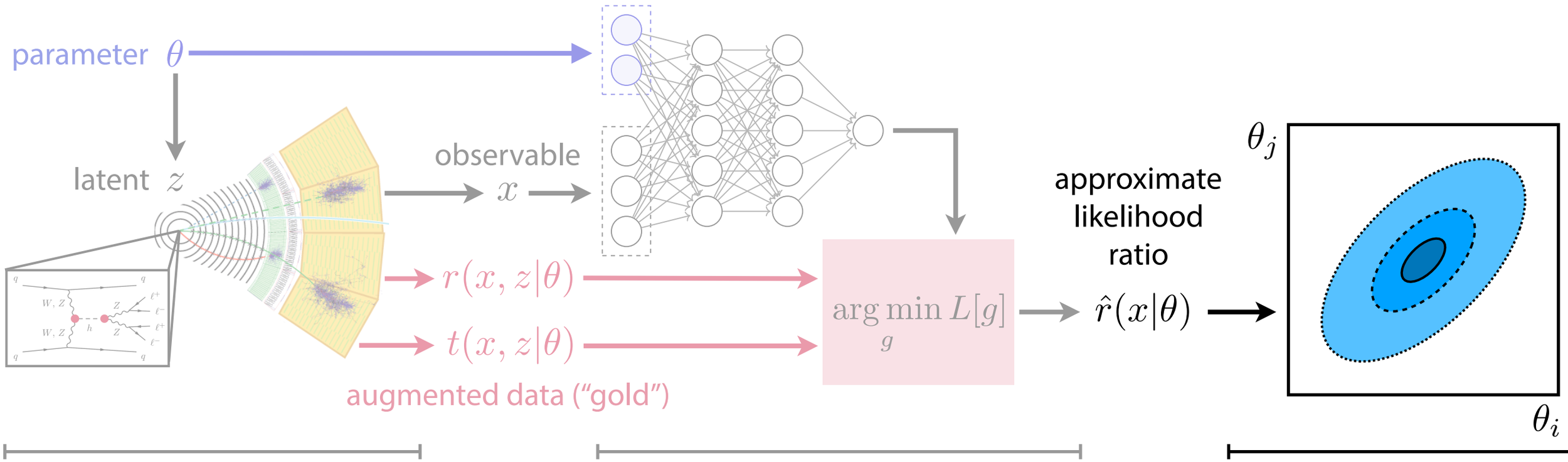
2. Machine Learning

Use this information to train precise estimator for likelihood ratio

$$r(x|\theta_0, \theta_1) \equiv \frac{p(x|\theta_0)}{p(x|\theta_1)}$$

observables model parameters

Bird's-eye view



1. Simulation

"Mining gold": Extract additional information from simulator

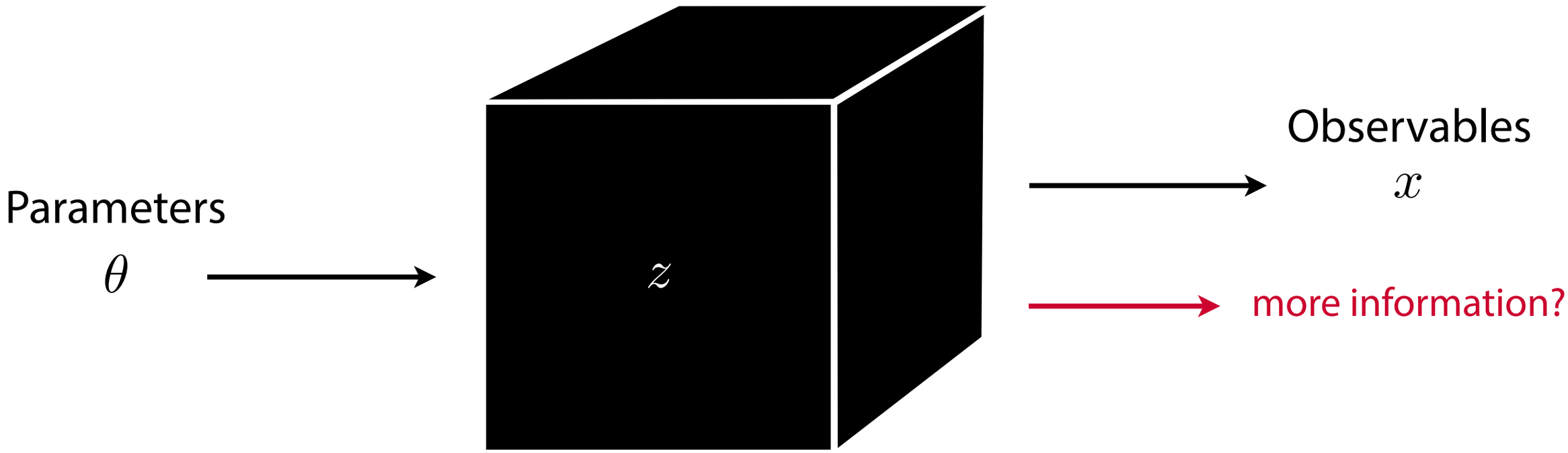
2. Machine Learning

Use this information to train precise estimator for likelihood ratio

3. Inference

Limit setting with standard hypothesis tests

Mining gold from particle physics simulators



Mining gold from particle physics simulators

Features

Latent variables

Parameters of interest

Observables

Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$

$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \; p(x|z_d) \qquad p(z_d|z_s) \qquad p(z_s|z_p) \qquad p(z_p|\theta)$$

Intractable integrals

Mining gold from particle physics simulators

Features

Latent variables

Parameters of interest

Observables

Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$

$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \; p(x|z_d) \qquad p(z_d|z_s) \qquad p(z_s|z_p) \qquad \boxed{p(z_p|\theta)}$$

This parton-level likelihood can be evaluated!

Mining gold from particle physics simulators

Features | Latent variables | Parameters of interest

Observables | Detector interactions | Shower splittings | Parton-level momenta | Theory parameters

$x \leftarrow z_d \leftarrow z_s \leftarrow z_p \leftarrow \theta$

$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \; p(x|z_d) \qquad p(z_d|z_s) \qquad p(z_s|z_p) \qquad \boxed{p(z_p|\theta)}$$

This parton-level likelihood can be evaluated!

⇒ We can calculate the "joint" likelihood ratio conditional on a specific evolution:

$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z_d, z_s, z_p|\theta_0)}{p(x, z_d, z_s, z_p|\theta_1)} = \frac{p(x|z_d)}{p(x|z_d)} \frac{p(z_d|z_s)}{p(z_d|z_s)} \frac{p(z_s|z_p)}{p(z_s|z_p)} \qquad \boxed{\frac{p(z_p|\theta_0)}{p(z_p|\theta_1)}}$$

The value of gold

We have joint likelihood ratio

$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z_d, z_s, z_p|\theta_0)}{p(x, z_d, z_s, z_p|\theta_1)}$$

We want likelihood ratio

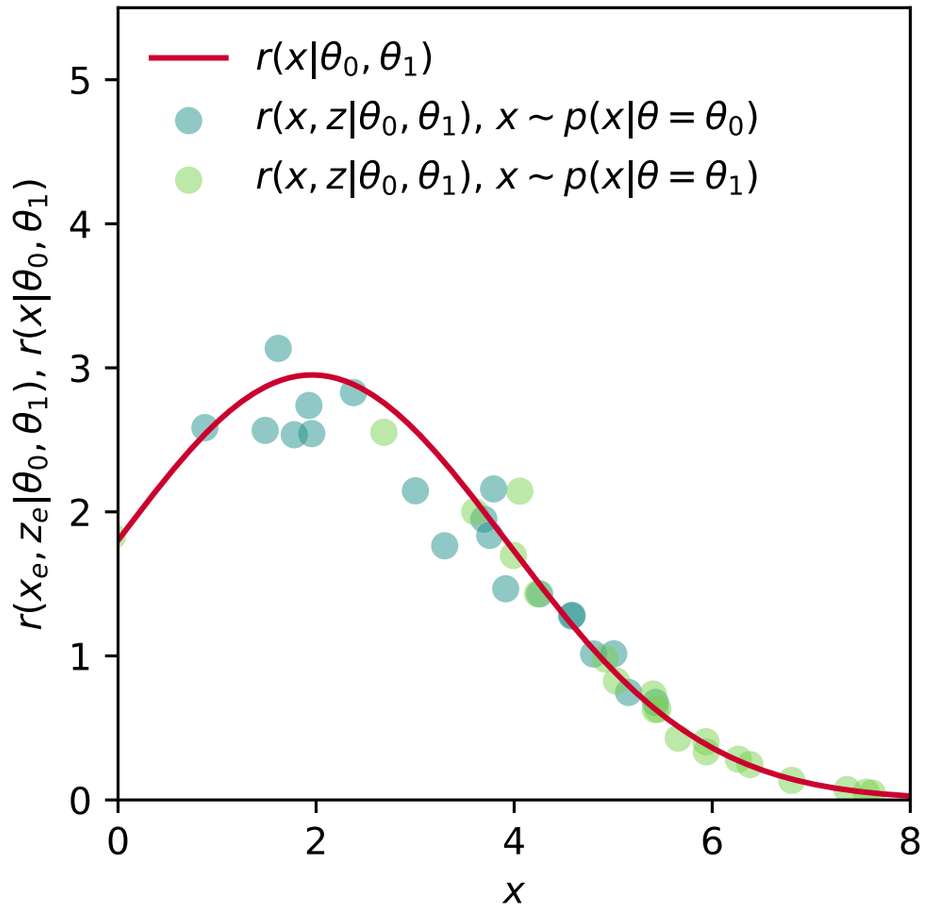
$$r(x|\theta_0, \theta_1) \equiv \frac{p(x|\theta_0)}{p(x|\theta_1)}$$

The value of gold

We have joint likelihood ratio

$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z_d, z_s, z_p|\theta_0)}{p(x, z_d, z_s, z_p|\theta_1)}$$

$r(x, z|\theta_0, \theta_1)$ are scattered around $r(x|\theta_0, \theta_1)$



We want likelihood ratio

$$r(x|\theta_0, \theta_1) \equiv \frac{p(x|\theta_0)}{p(x|\theta_1)}$$

The value of gold

We have joint likelihood ratio

$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z_d, z_s, z_p|\theta_0)}{p(x, z_d, z_s, z_p|\theta_1)}$$

With $r(x, z|\theta_0, \theta_1)$, we define the functional

$$L_r[\hat{r}(x)] = \mathbb{E}_{p(x,z|\theta_1)}\left[\left(\hat{r}(x) - r(x, z|\theta_0, \theta_1)\right)^2\right].$$

One can show it is minimized by

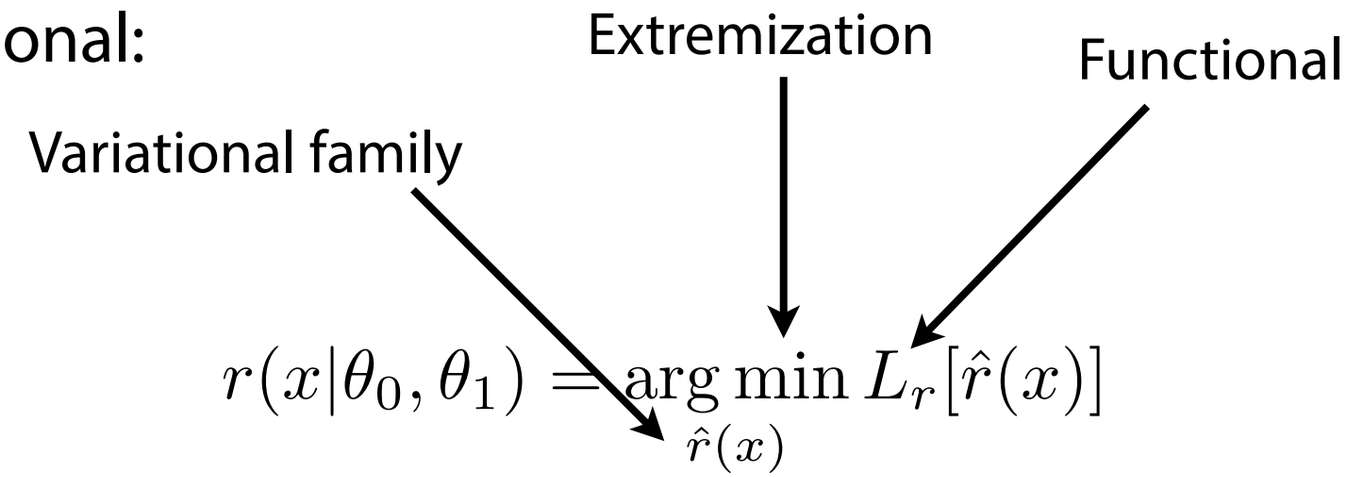
$$\arg\min_{\hat{r}(x)} L_r[\hat{r}(x)] = r(x|\theta_0, \theta_1) \text{ !}$$

We want likelihood ratio

$$r(x|\theta_0, \theta_1) \equiv \frac{p(x|\theta_0)}{p(x|\theta_1)}$$

How to do this in practice?

Need to minimize a functional:



How to do this in practice?

Need to minimize a functional:

Extremization

Functional

Variational family

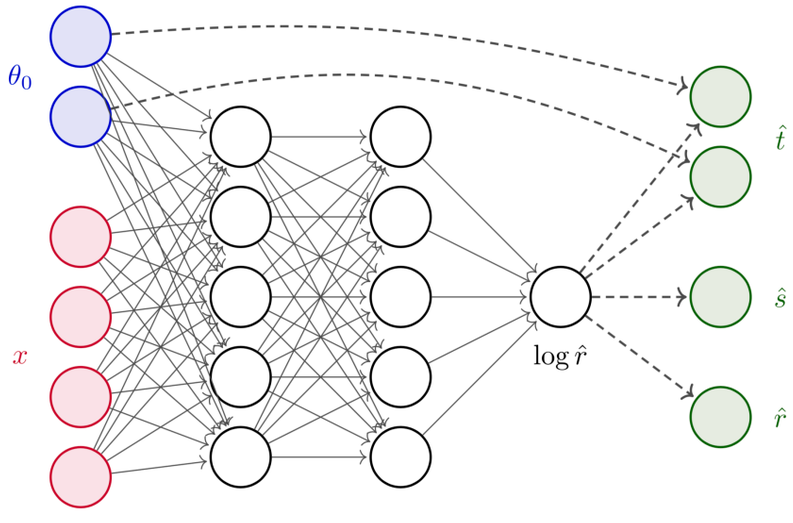
$$r(x|\theta_0,\theta_1) = \underset{\hat{r}(x)}{\arg\min}\, L_r[\hat{r}(x)]$$

This is exactly what machine learning does!

Loss function

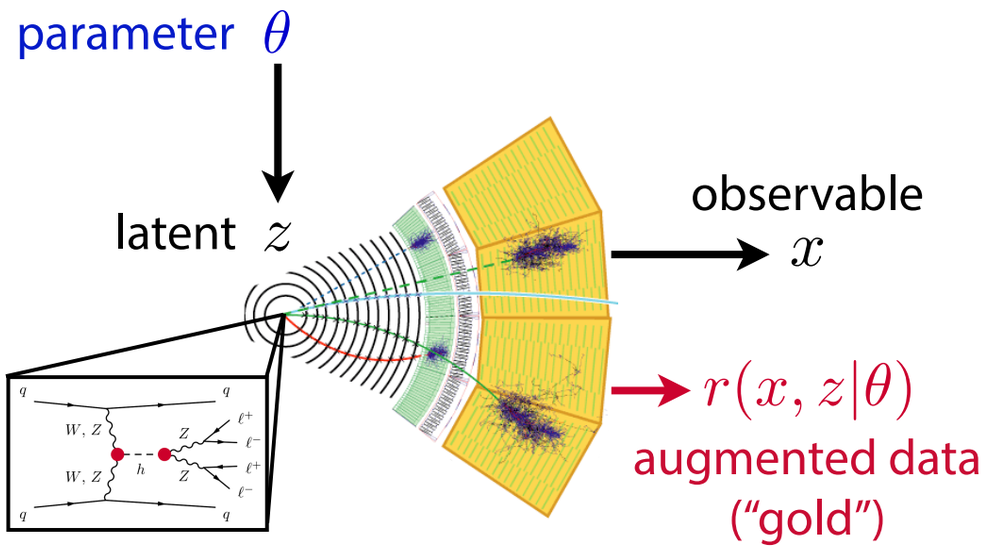
Stochastic gradient descent

Neural network

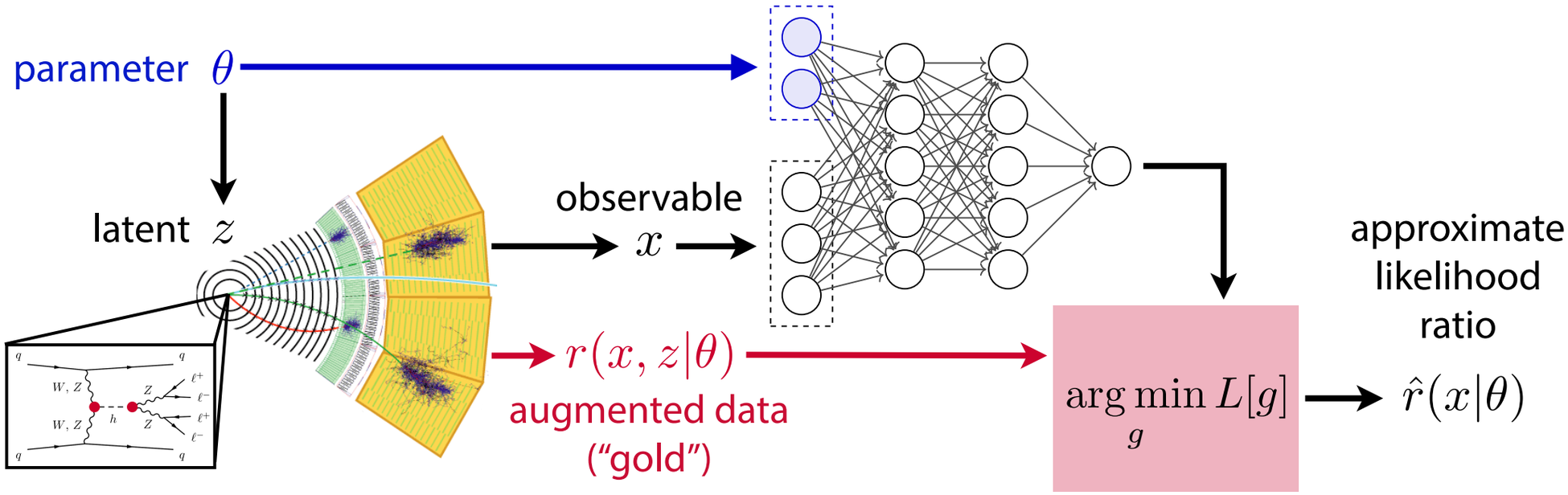


$\Rightarrow$ We implement $\hat{r}(x|\theta_0,\theta_1)$ as a neural network trained on the data available from the simulator

What we have so far



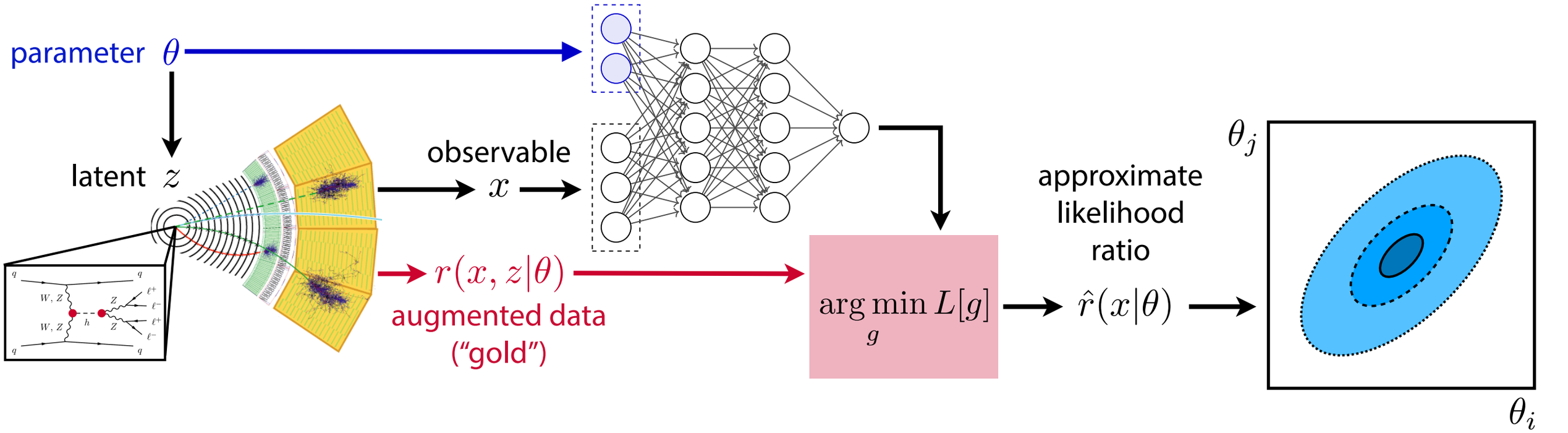
What we have so far



1. Simulation

2. Machine Learning

What we have so far



The score

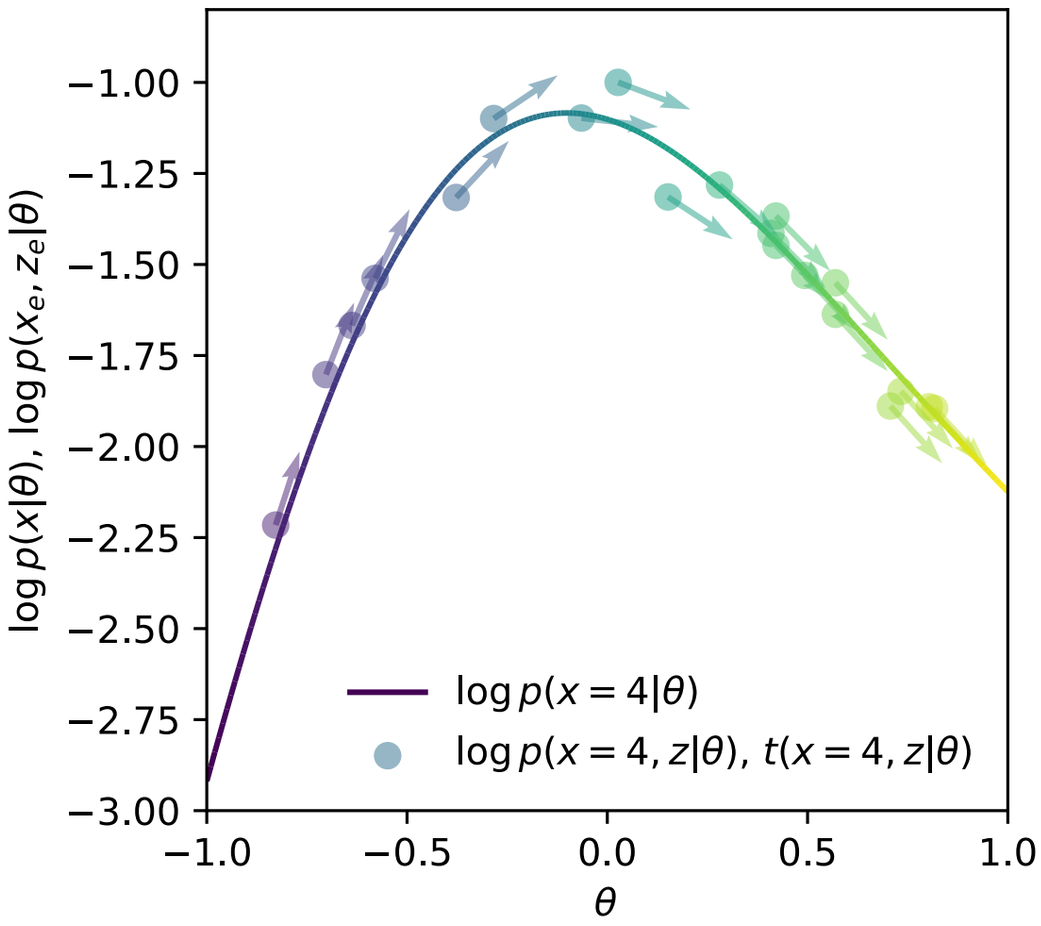
- Inference just based on the joint likelihood ratio works well, but there is another powerful piece of information

- The score

$$t(x|\theta_0) \equiv \nabla_\theta \log p(x|\theta)\Big|_{\theta_0}$$

fully characterizes the likelihood function in the neighborhood of θ_0

- The score itself is intractable. But…



Learning the score

Similar to the joint likelihood ratio,
we can calculate the joint score

$$t(x, z|\theta_0) \equiv \nabla_\theta \log p(x, z_d, z_s, z|\theta)\Big|_{\theta_0}$$

We want score

$$t(x|\theta_0) \equiv \nabla_\theta \log p(x|\theta)\Big|_{\theta_0}$$

Learning the score

Similar to the joint likelihood ratio,
we can calculate the joint score

$$t(x, z|\theta_0) \equiv \nabla_\theta \log p(x, z_d, z_s, z|\theta)\Big|_{\theta_0}$$

We want score

$$t(x|\theta_0) \equiv \nabla_\theta \log p(x|\theta)\Big|_{\theta_0}$$

Given $t(x, z|\theta_0)$,
we define the functional

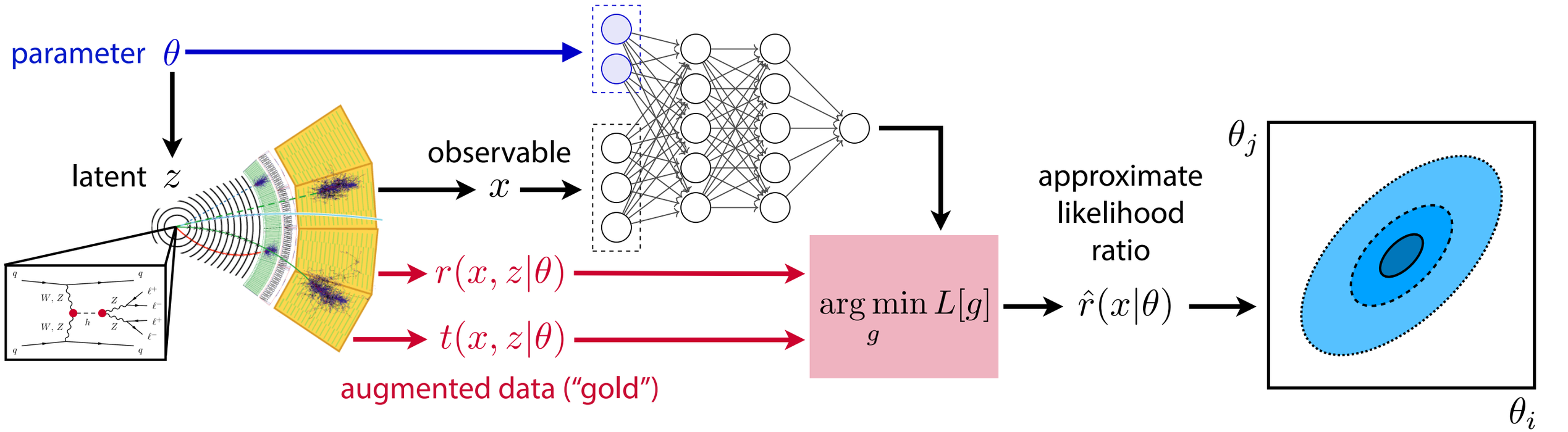
$$L_t[\hat{t}(x)] = \mathbb{E}_{p(x,z|\theta_0)}\left[\left(\hat{t}(x) - t(x, z|\theta_0)\right)^2\right].$$

One can show it is minimized by

$$\arg\min_{\hat{t}(x)} L_t[\hat{t}(x)] = t(x|\theta_0).$$

Again, we implement this
with machine learning

Putting the pieces together



A family of likelihood-free inference strategies

Different strategies to combine the different pieces of information:

Method	L_{XE}	L_{MLE}	L_r	L_t	θ sampling
ABC (**A**pproximate **B**ayesian **C**omputation)					$\theta \sim \pi(\theta)$
NDE (**N**eural **d**ensity **e**stimation)		✓			$\theta \sim \pi(\theta)$
LRT / CARL (**L**ikelihood **r**atio **t**rick / **c**alibrated **a**pproximate **r**atios of **l**ikelihoods)	✓				$\theta \sim \pi(\theta)$
ROLR (**R**egression **o**n **l**ikelihood **r**atio)			✓		$\theta \sim \pi(\theta)$
SCANDAL (**Sc**ore **a**ugmented **n**eural **d**ensity **a**pproximates **l**ikelihood)		✓		✓	$\theta \sim \pi(\theta)$
CASCAL (**C**ARL **a**nd **sc**ore **a**pproximate **l**ikelihood ratio).	✓			✓	$\theta \sim \pi(\theta)$
RASCAL (**R**atio **a**nd **sc**ore **a**pproximate **l**ikelihood ratio)			✓	✓	$\theta \sim \pi(\theta)$
SALLY (**Sc**ore **a**pproximates **l**ikelihood **l**ocal**ly**)				✓	$\theta = \theta_0$
SALLINO (**Sc**ore **a**pproximates **l**ikelihood **l**ocally **in** **o**ne dimension)				✓	$\theta = \theta_0$

A family of likelihood-free inference strategies

Different strategies to combine the different pieces of information:

Method	L_{XE}	L_{MLE}	L_r	L_t	θ sampling
ABC (**A**pproximate **B**ayesian **C**omputation)					$\theta \sim \pi(\theta)$
NDE (**N**eural **d**ensity **e**stimation)		✓			$\theta \sim \pi(\theta)$
LRT / CARL (**L**ikelihood **r**atio **t**rick / **c**alibrated **a**pproximate **r**atios of **l**ikelihoods)	✓				$\theta \sim \pi(\theta)$
ROLR (**R**egression **o**n **l**ikelihood **r**atio)			✓		$\theta \sim \pi(\theta)$
SCANDAL (**Sc**ore **a**ugmented **n**eural **d**ensity **a**pproximates **l**ikelihood)		✓		✓	$\theta \sim \pi(\theta)$
CASCAL (CARL **and** **sc**ore **a**pproximate **l**ikelihood ratio).	✓			✓	$\theta \sim \pi(\theta)$
RASCAL (**R**atio **a**nd **sc**ore **a**pproximate **l**ikelihood ratio)			✓	✓	$\theta \sim \pi(\theta)$
SALLY (**Sc**ore **a**pproximates **l**ikelihood **l**ocally**y**)				✓	$\theta = \theta_0$
SALLINO (**Sc**ore **a**pproximates **l**ikelihood **l**ocally **in** **o**ne dimension)				✓	$\theta = \theta_0$

RASCAL loss function:

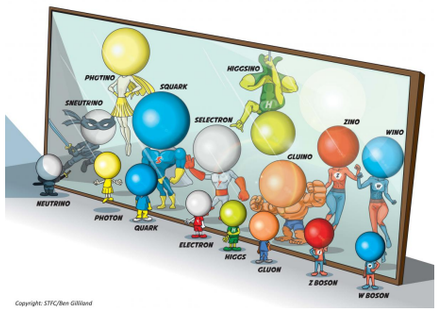
$$L_{\text{RASCAL}}[\hat{r}(x|\theta_0,\theta_1)] = L_r[\hat{r}(x|\theta_0,\theta_1)] + \alpha\ L_t[\nabla_{\theta_0} \log \hat{r}(x|\theta_0,\theta_1)]$$

Proof of concept:
Constraining EFTs at the LHC

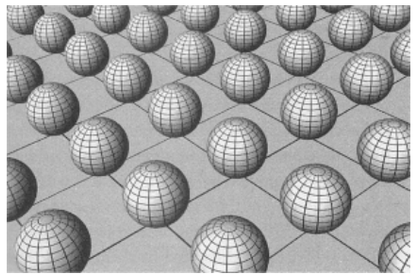
[Image sources: STFC / Ben Gililand, Sean Carroll, Friedrich Justin Bertuch 1806, symmetry]

Effective field theory

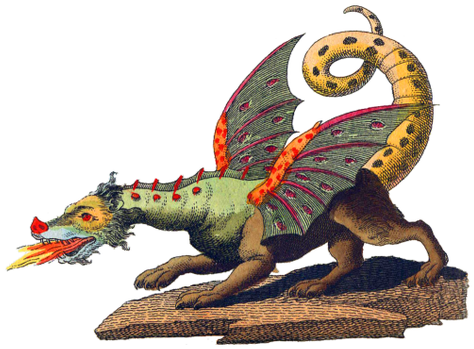
Energy



Supersymmetry?



Extra dimensions?



Dragons?

Composite Higgs?

…

Λ

$E \ll \Lambda$

LHC Higgs

Other LHC measurements

LEP

Flavour

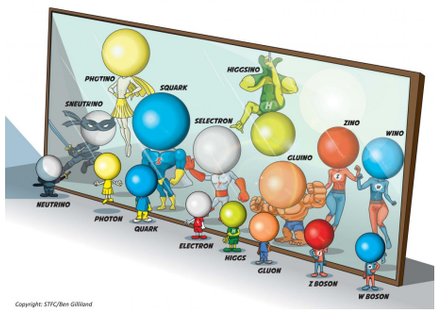
…

[Image sources: STFC / Ben Gililand, Sean Carroll, Friedrich Justin Bertuch 1806, symmetry]

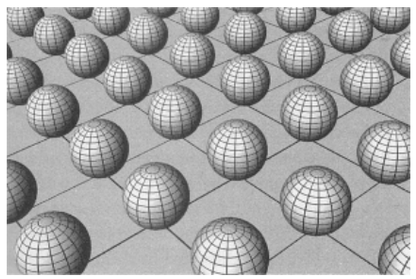
Effective field theory

Energy

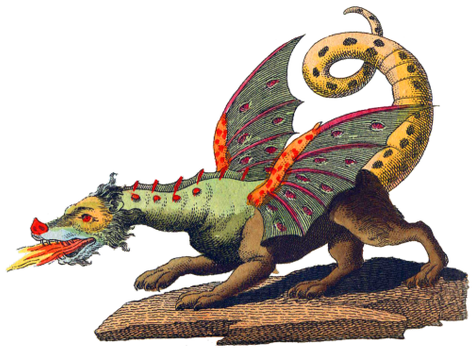
Λ



Supersymmetry?



Extra dimensions?



Dragons?

Composite Higgs?

…

Matching

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{f_i}{\Lambda^2}\,\mathcal{O}_i + \dots$$

Effective field theory

Measurements

$E \ll \Lambda$

LHC Higgs

Other LHC measurements

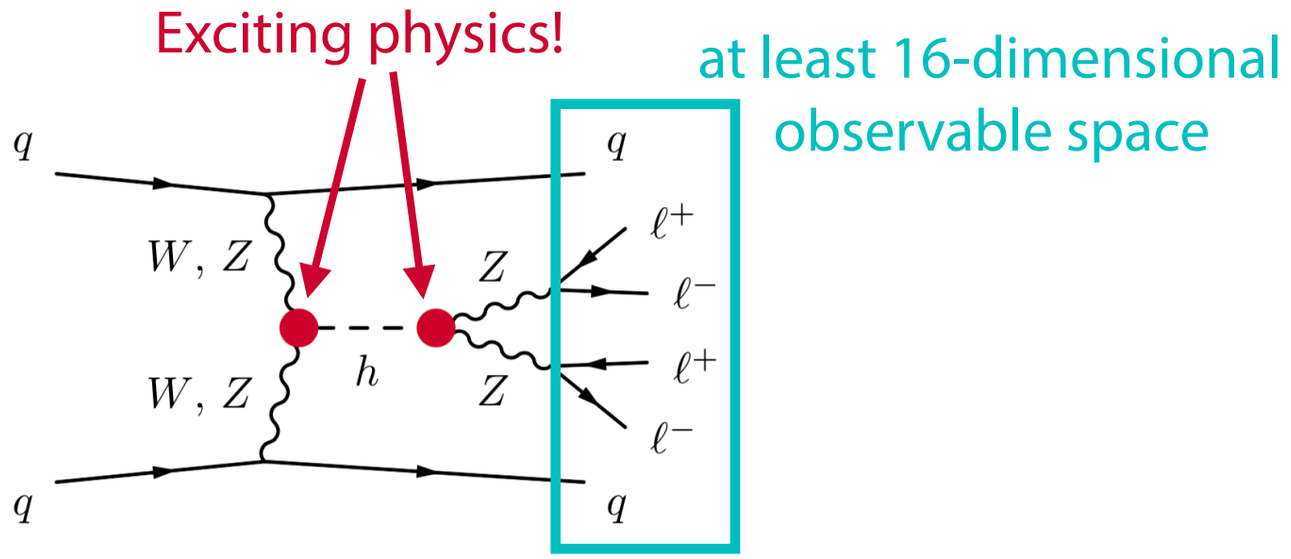
LEP

Flavour

…

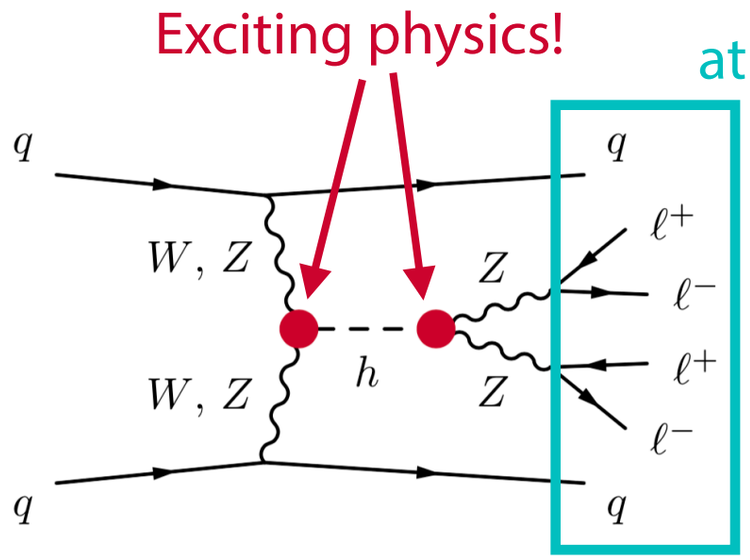
Proof of concept

- Higgs production in weak boson fusion:



Proof of concept

- Higgs production in weak boson fusion:

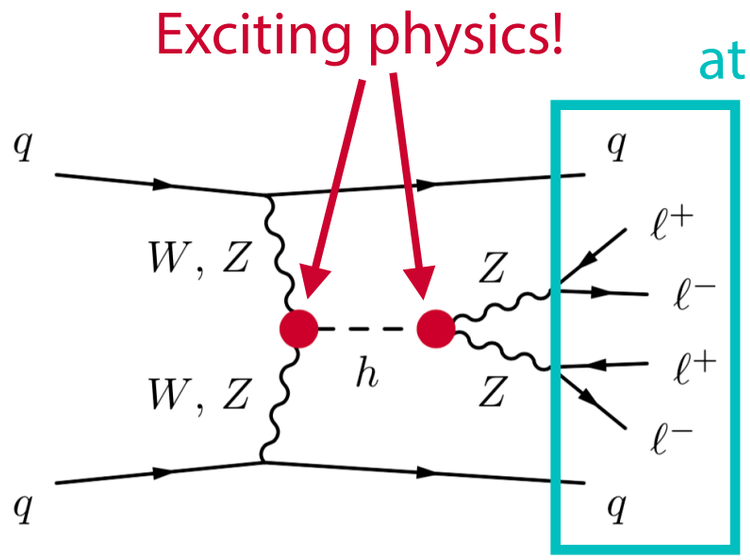


- Goal: constraints on two EFT parameters

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \boxed{\frac{f_W}{\Lambda^2}} \underbrace{\frac{\mathrm{i}g}{2} (D^\mu \phi)^\dagger \, \sigma^a \, D^\nu \phi \, W^a_{\mu\nu}}_{\mathcal{O}_W} - \boxed{\frac{f_{WW}}{\Lambda^2}} \underbrace{\frac{g^2}{4} (\phi^\dagger \phi) \, W^a_{\mu\nu} \, W^{\mu\nu \, a}}_{\mathcal{O}_{WW}}$$

Proof of concept

- Higgs production in weak boson fusion:



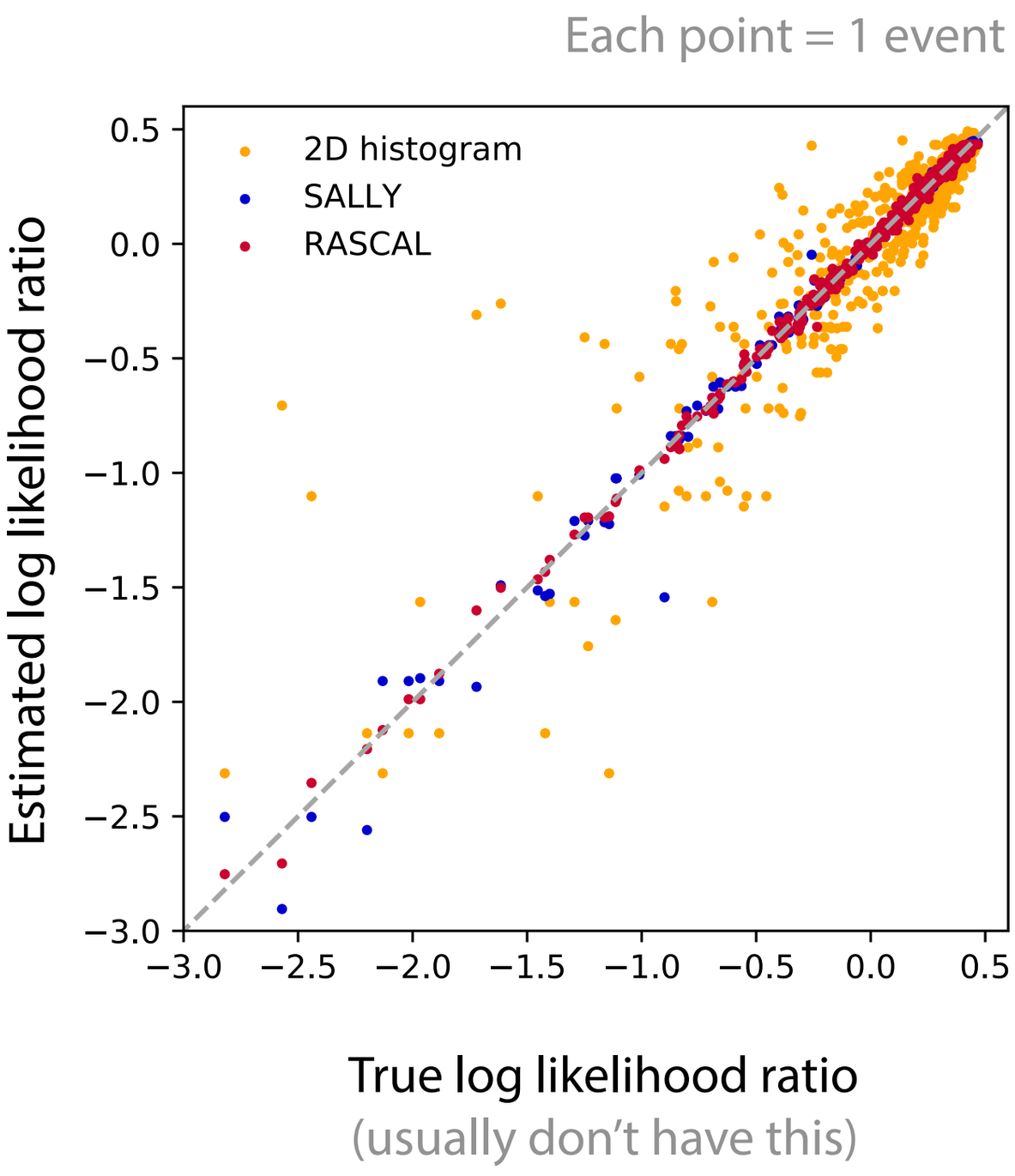
- Goal: constraints on two EFT parameters

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \boxed{\frac{f_W}{\Lambda^2}} \underbrace{\frac{\mathrm{i}g}{2} \left(D^\mu \phi\right)^\dagger \sigma^a D^\nu \phi \, W^a_{\mu\nu}}_{\mathcal{O}_W} - \boxed{\frac{f_{WW}}{\Lambda^2}} \underbrace{\frac{g^2}{4} \left(\phi^\dagger \phi\right) W^a_{\mu\nu} W^{\mu\nu\, a}}_{\mathcal{O}_{WW}}$$

- Two setups:
  - Simplified setup in which we can compare to true likelihood
  - "Realistic" simulation with approximate detector effects

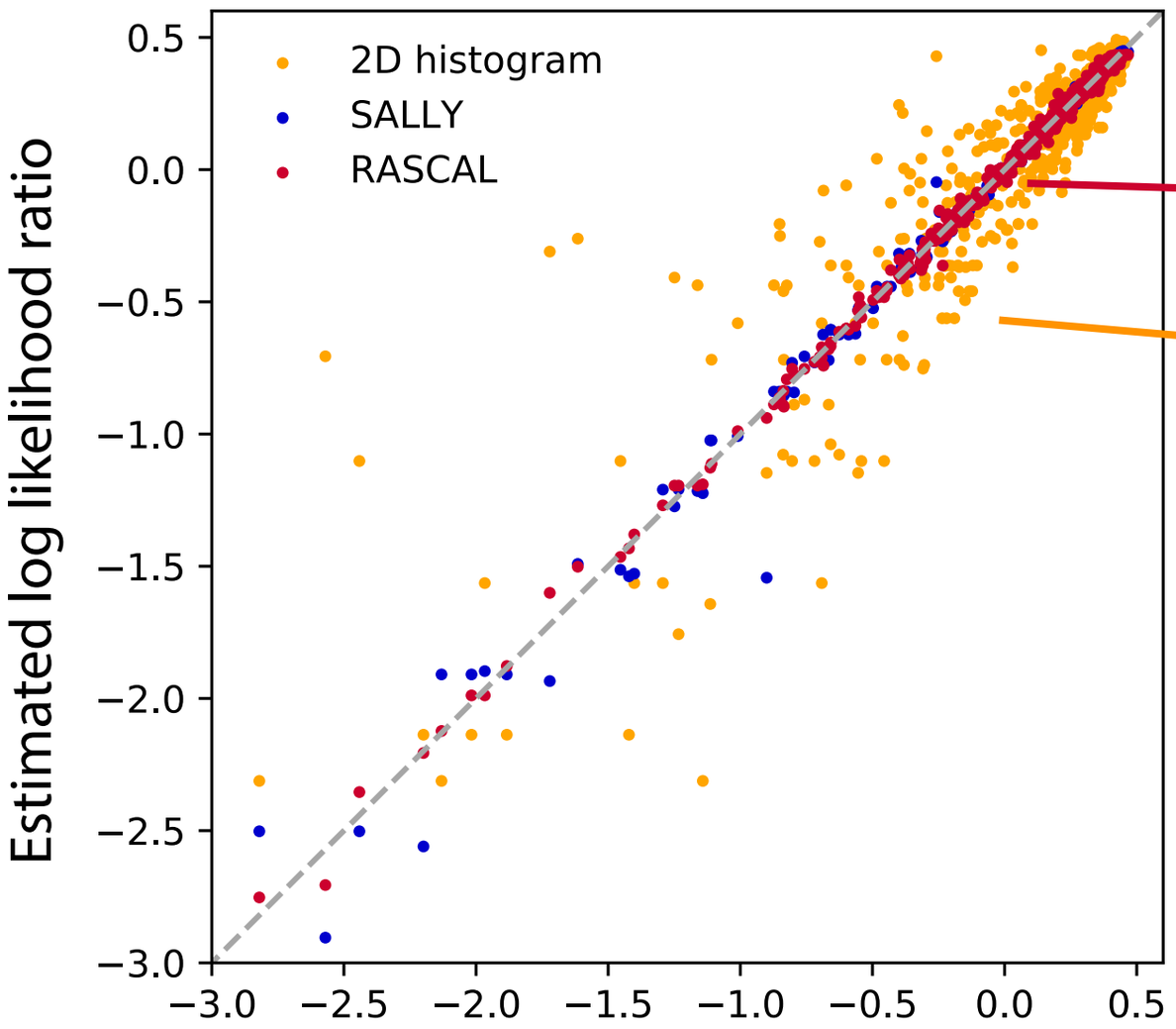
- Simulation: MadGraph [J. Alwall et al. 1405.0301] + MadMax
  [K. Cranmer, T. Plehn hep-ph/0605268; T. Plehn, P. Schichtel, D. Wiegand 1311.2591]

Precise likelihood ratio estimates



Precise likelihood ratio estimates

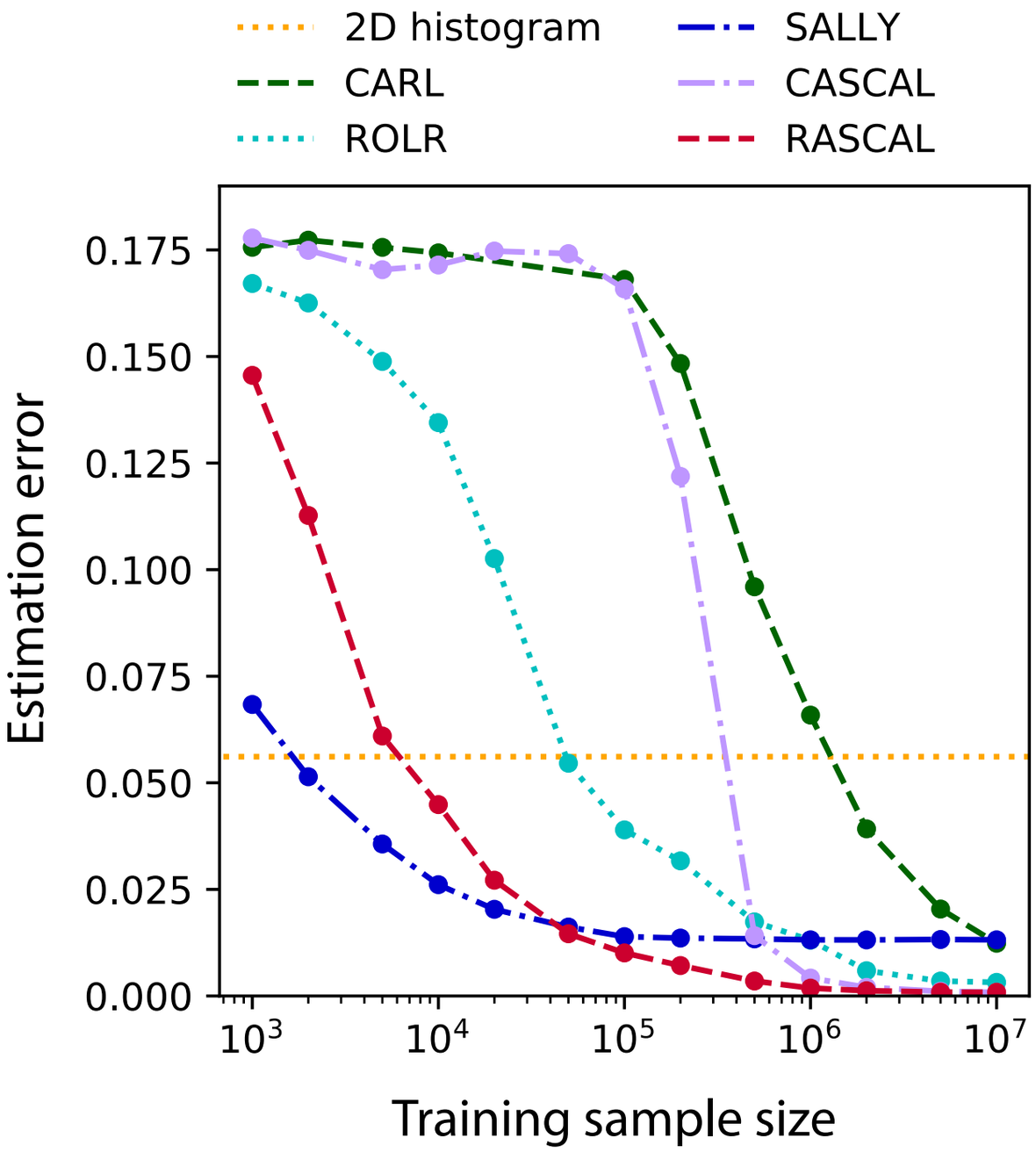
Each point = 1 event



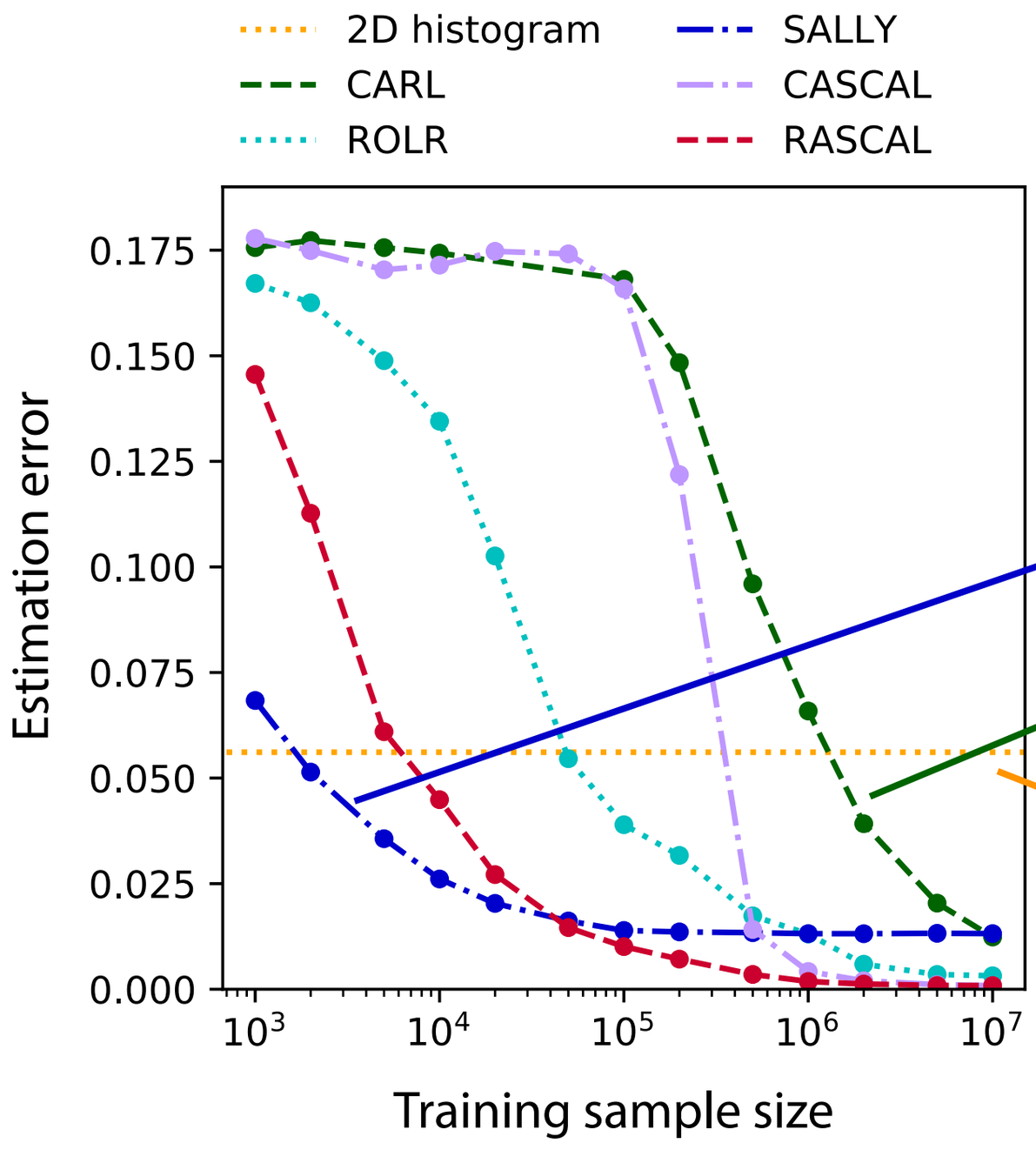
(usually don't have this)

RASCAL
much more precise than
traditional histogram

Less training data needed



Less training data needed

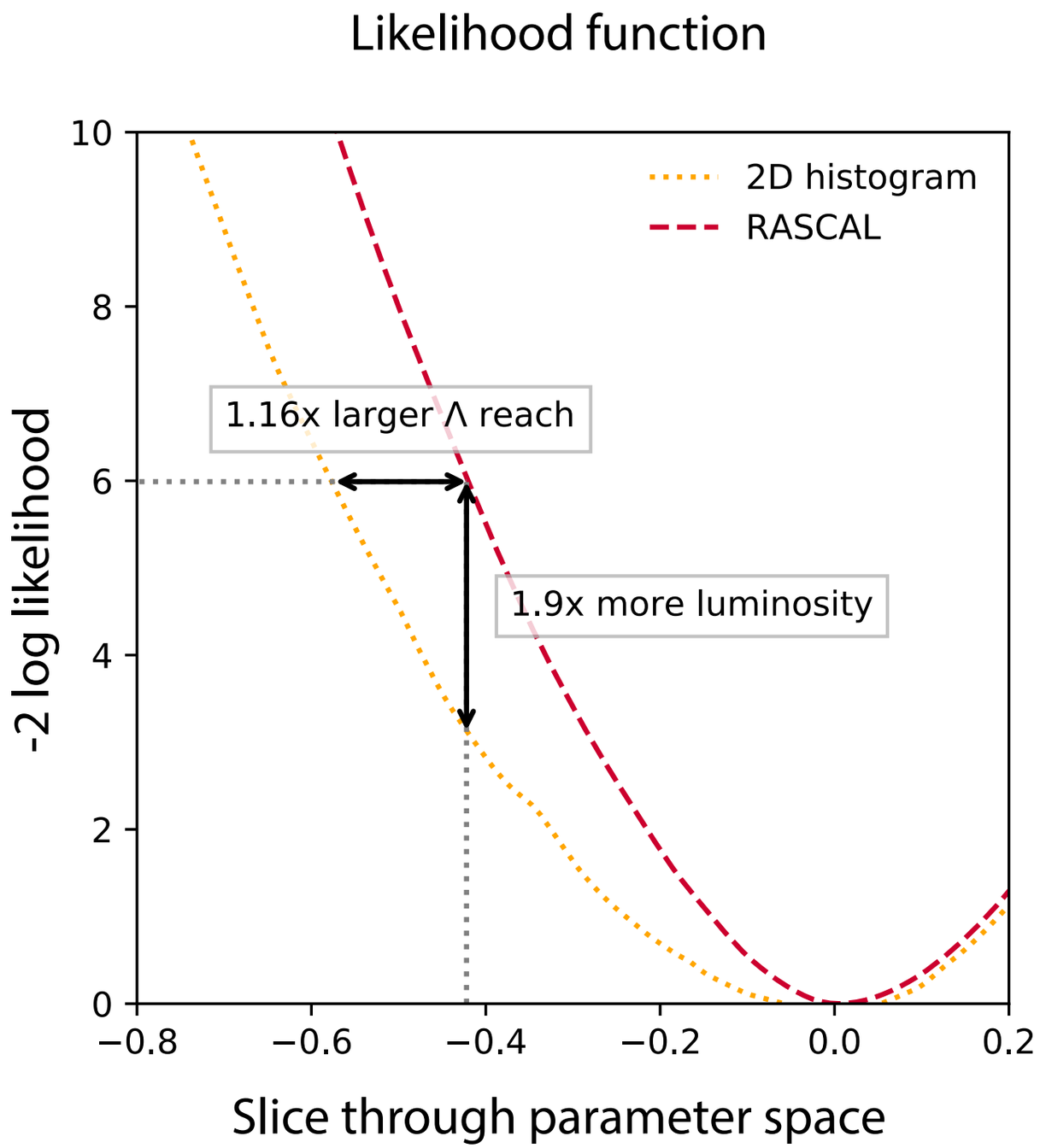


New techniques require less data than generic ML method

[CARL, see K. Cranmer, G. Louppe, J. Pavez 1506.02169]

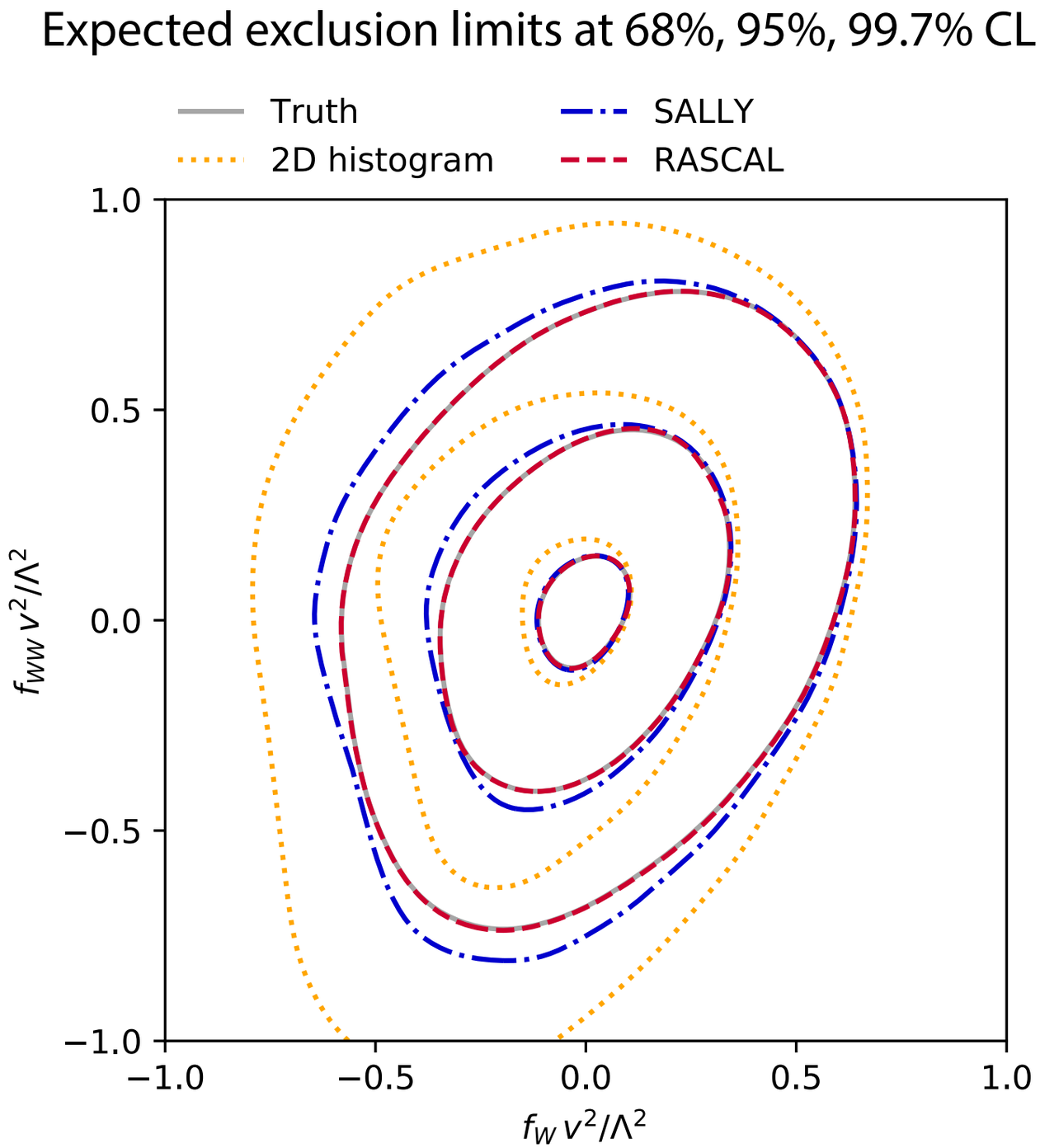
Traditional histogram

Better sensitivity



36 events, assuming SM

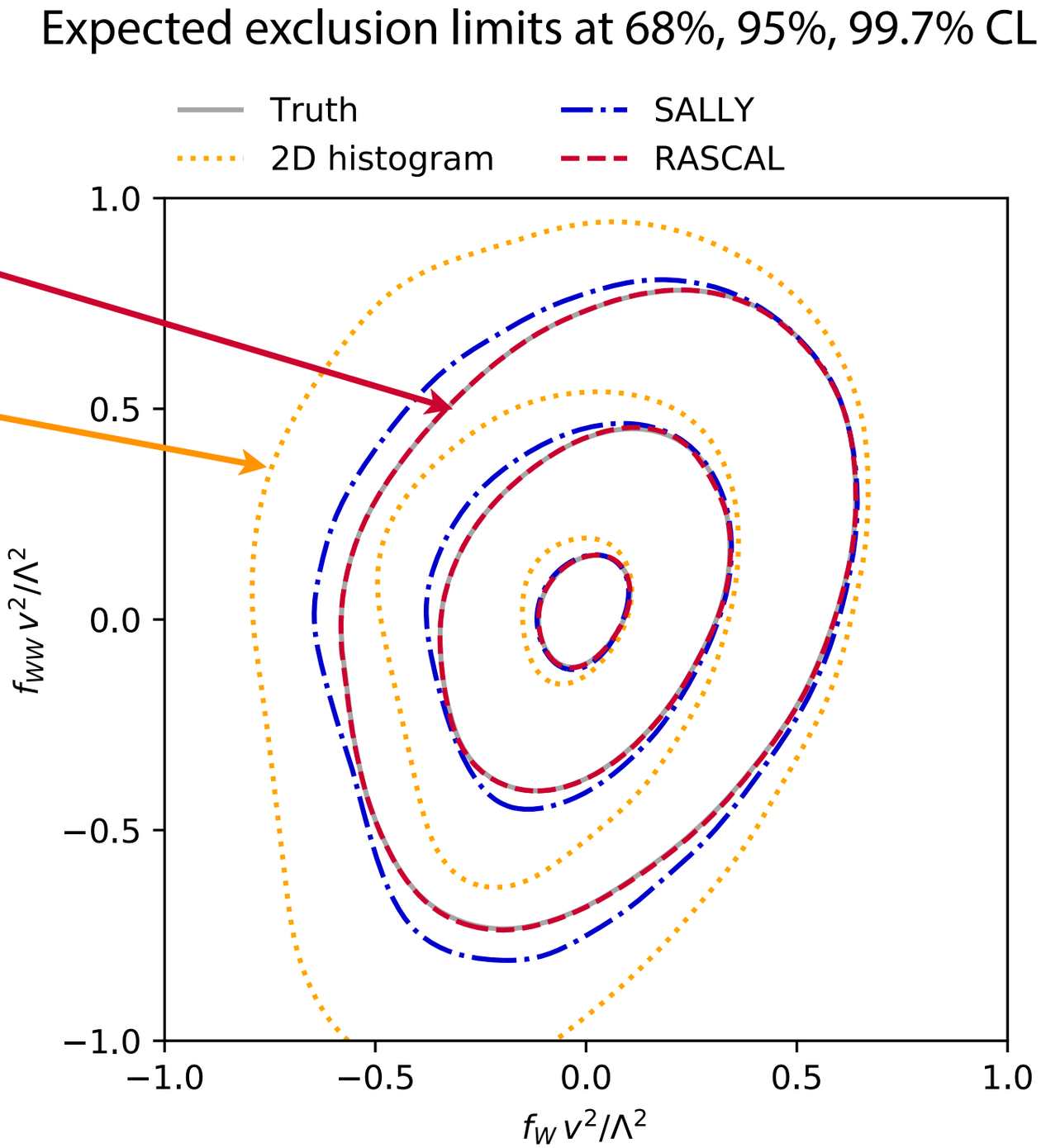
Stronger bounds



36 events, assuming SM

Stronger bounds

RASCAL
enables stronger
limits than
traditional histogram

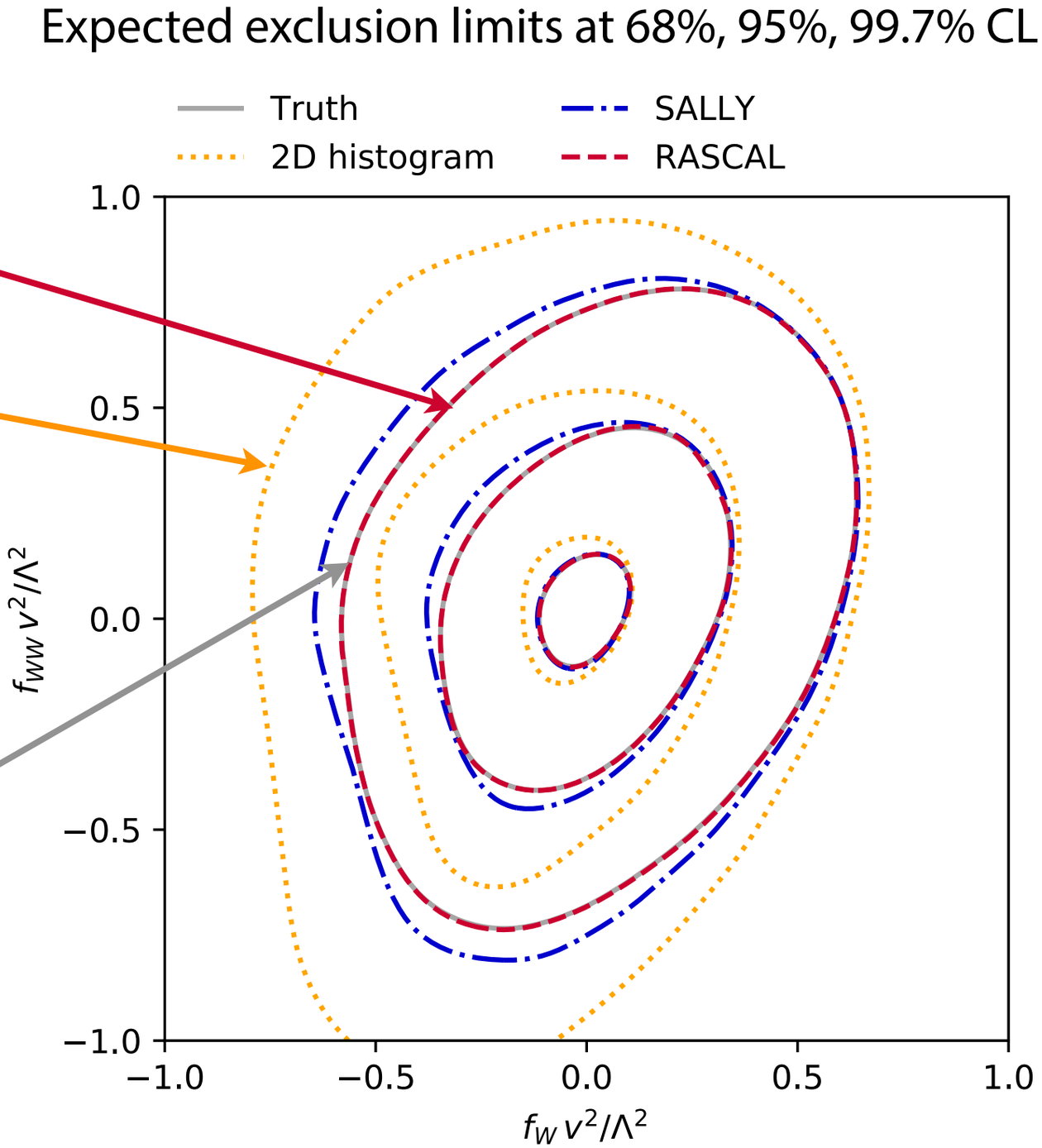


36 events, assuming SM

Stronger bounds

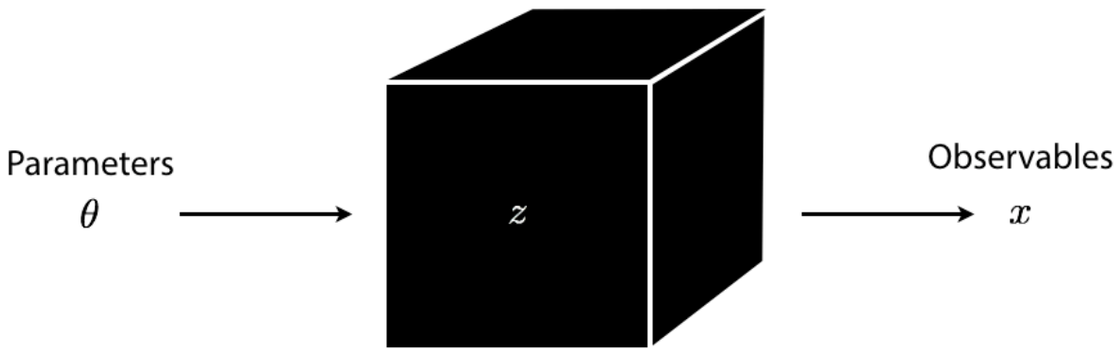
RASCAL
enables stronger
limits than
traditional histogram

Limits from RASCAL
virtually indistinguishable
from true likelihood
(usually we don't have that)



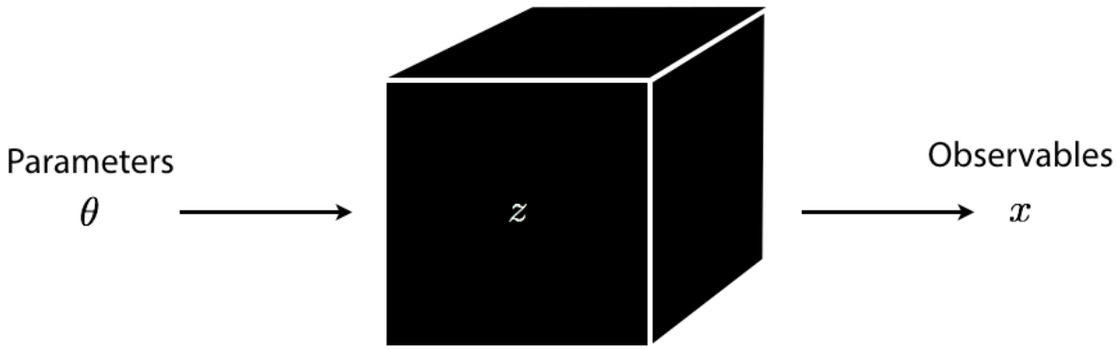
36 events, assuming SM

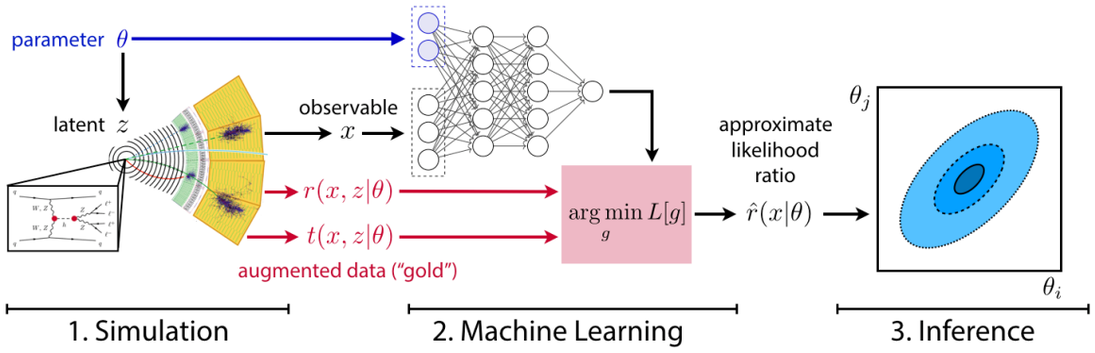
A new approach to simulator-based inference



- Much of modern science is based on simulations, “likelihood-free”
- Established inference methods treat simulator as black box

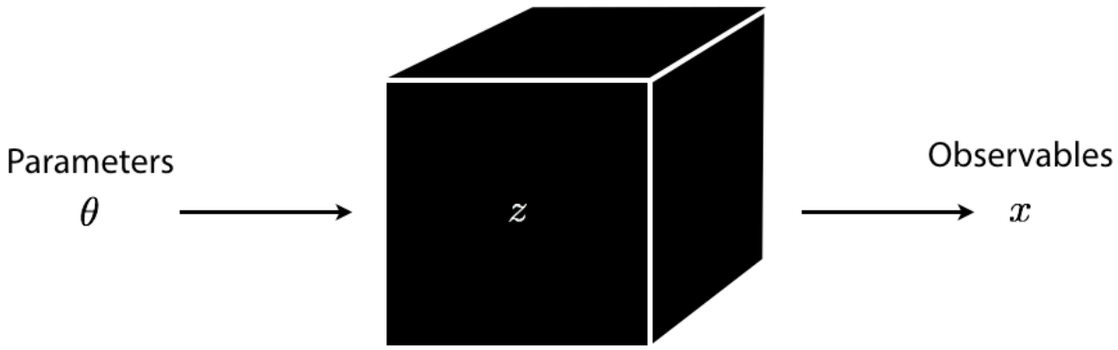
A new approach to simulator-based inference

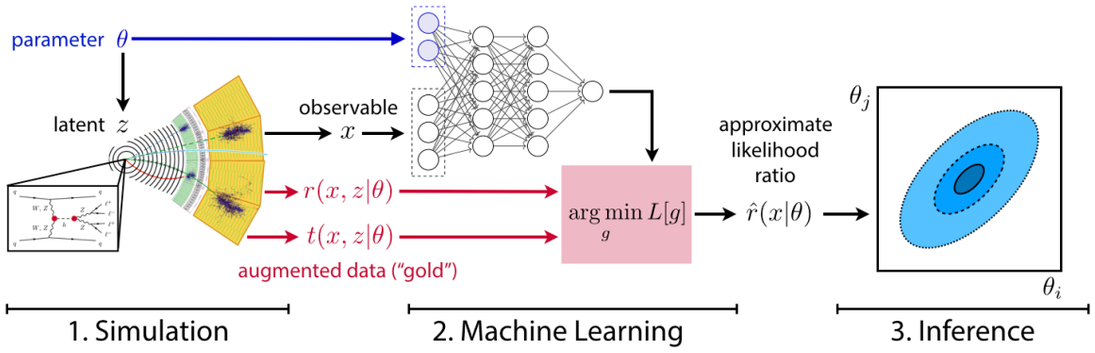


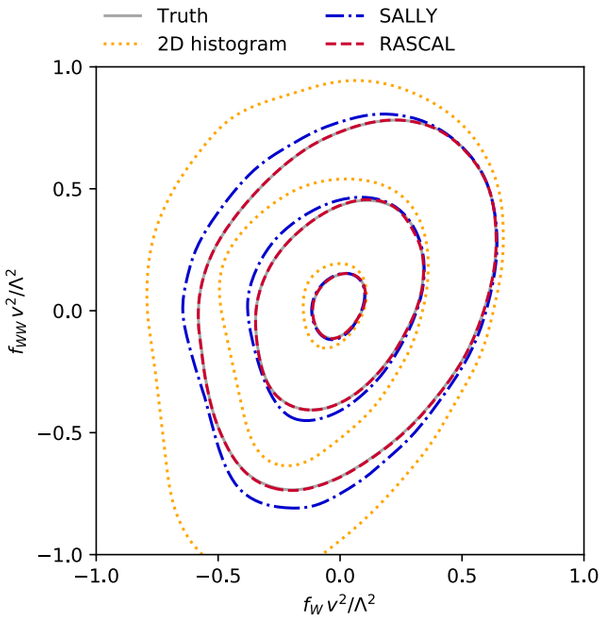


- Much of modern science is based on simulations, "likelihood-free"
- Established inference methods treat simulator as black box
- New inference techniques: Leverage more information from simulator + power of machine learning

A new approach to simulator-based inference







- Much of modern science is based on simulations, "likelihood-free"
- Established inference methods treat simulator as black box
- New inference techniques: Leverage more information from simulator + power of machine learning
- First application to LHC physics: New methods allow for stronger EFT constraints with less data