# SBI and FASER

# Motivation

- I got a DFG grant to look at FASER neutrino measurements with SBI.
- Postdoc will be hired to (partially) work on this.
- We should at least vaguely do what I promised, and avoid doing things twice

- collider neutrino measurements are sensitive to various phenomena affecting neutrino production, propagation and interaction
- neutrino event rates and kinematic distributions depend on all of these effects
- flux and interaction phenomena must be constrained together

- requires a coherent framework that simultaneously describes both neutrino production and interactions, and inference tools to constrain them

# Likelihood Function

- CC neutrino event characterized by an interaction vertex position **x** and the observed momenta and flavors $\mathbf{p_{obs}}$ for the final state lepton and hadrons
- The distribution of events in a theory, characterized by parameters **θ**, is then described by a likelihood function $p(x,p_{obs}|\theta)$

$$p(x, p_{obs}|\theta_F, \theta_f, \theta_\sigma) \sim \sum_{\nu,i} F_\nu(x, p_\nu|\theta_F) \otimes f_i(p_i|\theta_f) \otimes d\sigma(p_{parton}|\theta_\sigma) \otimes R(p_{obs}|p_{parton}) \ .$$

| Flux | PDF | Hard Scattering | Hadronization |
|---|---|---|---|
| Generated flux | Your favorite PDF set | POWHEG | Pythia |

- Example questions to address:
  - How well can we constrain PDFs in presence of flux uncertainties?

# Review: Simulation and Measurements

$$\theta \xrightarrow{\quad f(x|\theta) \quad} x \xrightarrow{\quad f(x|\theta) \quad} \hat{\theta} \pm \Delta\theta$$

parameters — simulation — observables — inference — estimator/uncertainty

all physics is encoded in the likelihood function f(x|θ):
*likelihood of an observation x as a function in a theory θ*

If we know f(x|θ), we can calculate likelihood ratio and do likelihood ratio test.

# Review: Simulation and Measurements

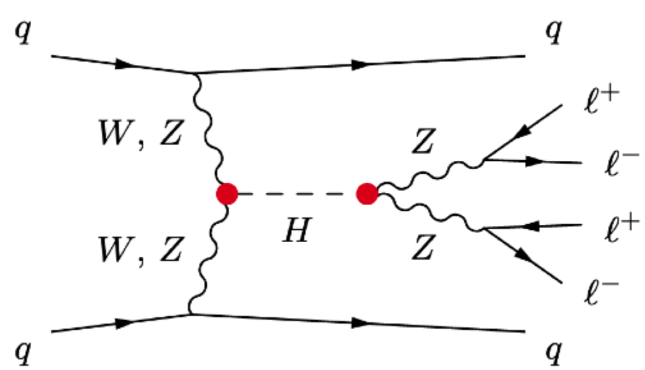
**How can we compute the likelihood function?**

sophisticated tools model parts of process

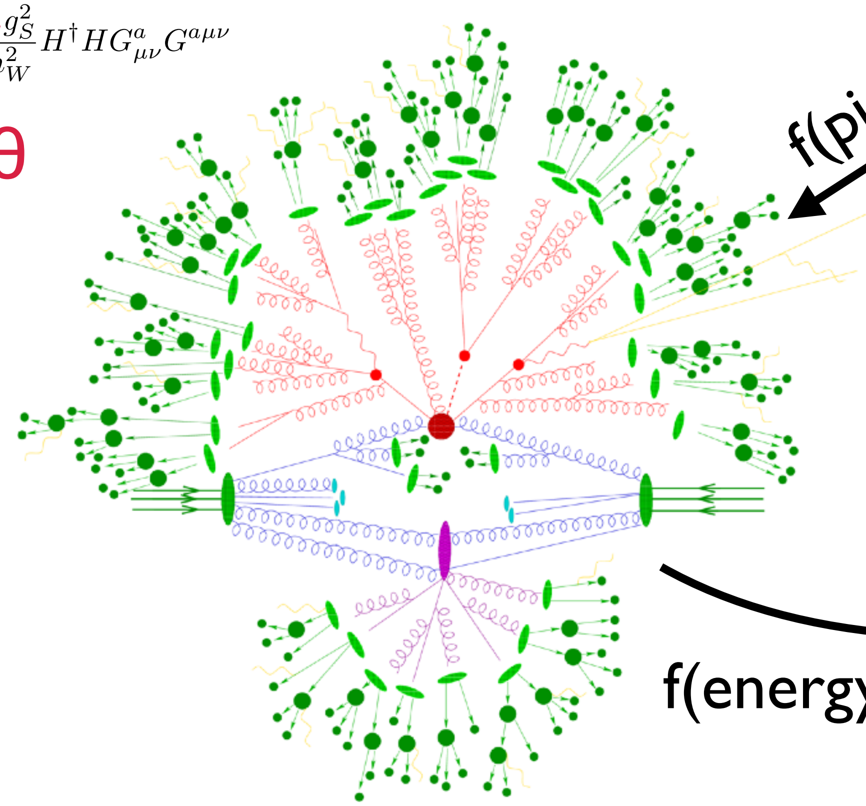
$$
\begin{aligned}
\Delta\mathcal{L}_{SILH} = & \frac{\bar{c}_H}{2v^2}\partial^\mu\left(H^\dagger H\right)\partial_\mu\left(H^\dagger H\right) + \frac{\bar{c}_T}{2v^2}\left(H^\dagger \overleftrightarrow{D}^\mu H\right)\left(H^\dagger \overleftrightarrow{D}_\mu H\right) - \frac{\bar{c}_6\lambda}{v^2}\left(H^\dagger H\right)^3 \\
& + \left(\left(\frac{\bar{c}_u}{v^2} y_u H^\dagger H \bar{q}_L H^c u_R + \frac{\bar{c}_d}{v^2} y_d H^\dagger H \bar{q}_L H d_R + \frac{\bar{c}_l}{v^2} y_l H^\dagger H \bar{L}_L H l_R\right) + \text{h.c.}\right) \\
& + \frac{i\bar{c}_W g}{2m_W^2}\left(H^\dagger \sigma^i \overleftrightarrow{D}^\mu H\right)\left(D^\nu W_{\mu\nu}\right)^i + \frac{i\bar{c}_B g'}{2m_W^2}\left(H^\dagger \overleftrightarrow{D}^\mu H\right)\left(\partial^\nu B_{\mu\nu}\right) \\
& + \frac{i\bar{c}_{HW} g}{m_W^2}\left(D^\mu H\right)^\dagger \sigma^i \left(D^\nu H\right) W^i_{\mu\nu} + \frac{i\bar{c}_{HB} g'}{m_W^2}\left(D^\mu H\right)^\dagger \left(D^\nu H\right) B_{\mu\nu} \\
& + \frac{\bar{c}_\gamma g'^2}{m_W^2} H^\dagger H B_{\mu\nu} B^{\mu\nu} + \frac{\bar{c}_g g_S^2}{m_W^2} H^\dagger H G^a_{\mu\nu} G^{a\mu\nu}
\end{aligned}
$$

model of nature θ

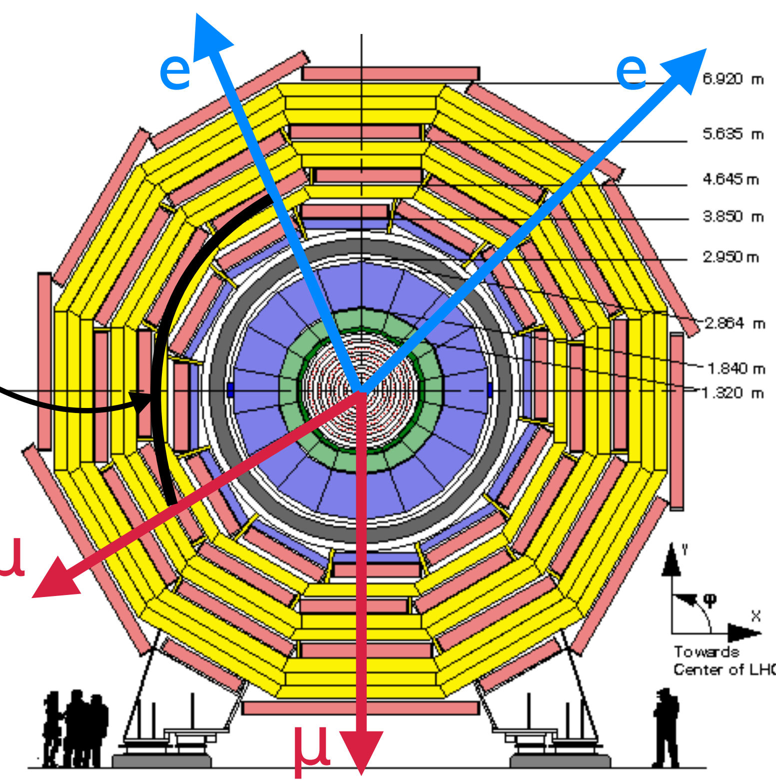
f(Higgs|model)




f(pion|quark)



f(energy deposit|pion)




interesting observables x

we obtain events distributed according to f(x|θ)

we cannot calculate f(x|θ)

learn likelihood from simulated data - likelihood free inference

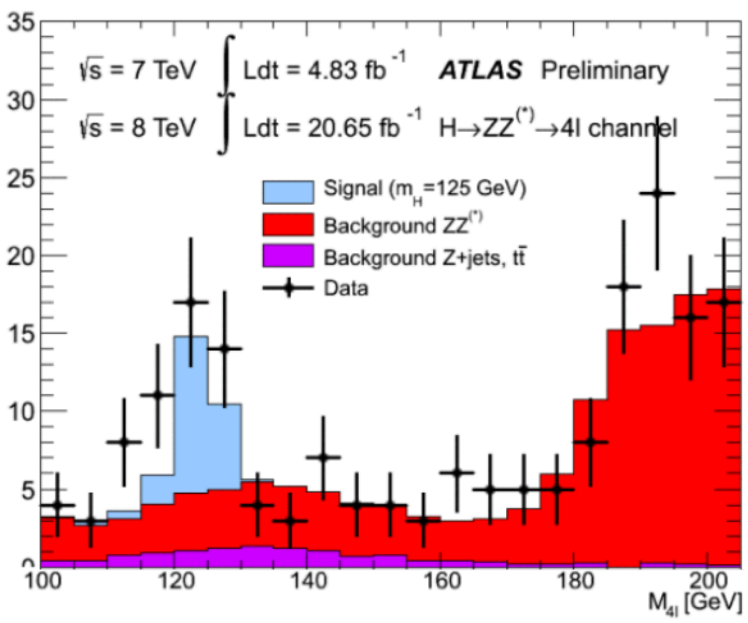
# Review: Simulation and Measurements

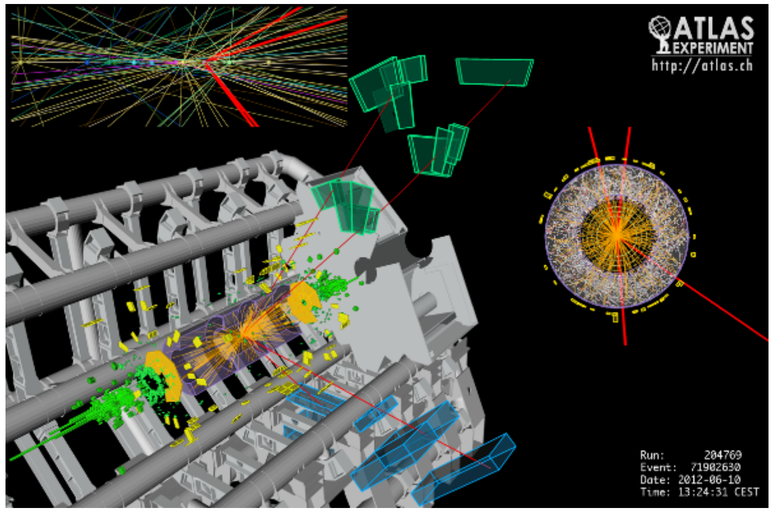
### Traditional Method: Histograms

analyze one carefully chosen variable

histogram ↔ p(x|θ)

→ ignore all the information in rest of the data



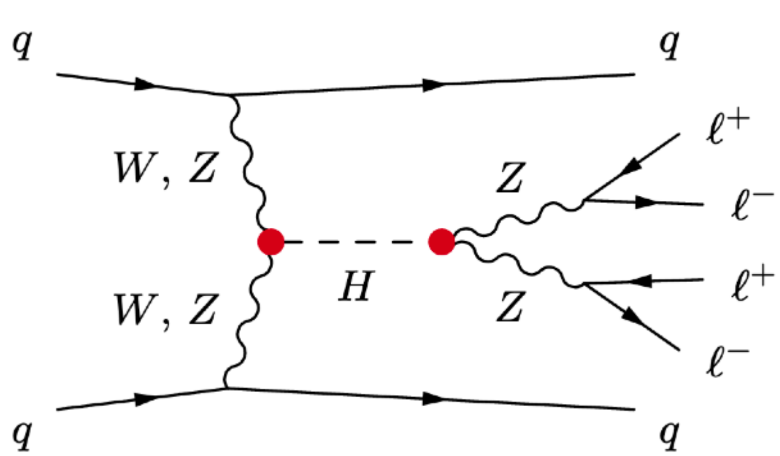




### Matrix-Element Method:

use matrix element information, approximate detector with transferfunction p(x|z) and calculate the integral directly.

$$p(x|\theta) \sim \int dz\ p(x|z)\ |M(z|\theta)|^2$$

→ big integral with thousands of dimensions



# Method

$$p(x, p_{obs}|\theta_F, \theta_f, \theta_\sigma) \sim \sum_{\nu,i} F_\nu(x, p_\nu|\theta_F) \otimes f_i(p_i|\theta_f) \otimes d\sigma(p_{parton}|\theta_\sigma) \otimes R(p_{obs}|p_{parton}) \; .$$

- Idea: If $p(x,p_{obs}|\theta)$ were known, we could calculate likelihood ratio (LR) and do a likelihood ratio test

- Problem: $p(x,p_{obs}|\theta)$ not calculable.

- Solutions:
  - binned histogram based analysis: loses information loss
  - matrix-element method: approximates approximates parton shower, hadronization and detector response in a way that is calculable
  - SBI: infer $p(x,p_{obs}|\theta)$ using ML

$$r(x, p_{obs}; p_{parton}|\theta_1, \theta_0) = \frac{F_\nu(x, p_\nu|\theta_{F,1})}{F_\nu(x, p_\nu|\theta_{F,0})} \frac{f_i(p_i|\theta_{f,1})}{f_i(p_i|\theta_{f,0})} \frac{d\sigma(p_{parton}|\theta_{\sigma,1})}{d\sigma(p_{parton}|\theta_{\sigma,0})}$$

this is calculable

$$r(x, p_{obs}|\theta_1, \theta_0)$$

can be used to infer LR

# MadMiner Method

**Particle Physics**

$$p(x|\theta) = \int dz\; p(x,z|\theta) = \int dz\; p(x|z)p(z|\theta)$$

factorizes

$$p(z|\theta) \sim |\mathcal{M}(z|\theta)|^2$$

**Statistics**

we want: $r(x|\theta) = p(x|\theta)/p(x|0)$

we know: $r(x,z|\theta) = \frac{p(x,z|\theta)}{p(x,z|0)} = \frac{p(z|\theta)}{p(z|0)}$

**Calculus of Variations**

$$\mathcal{L}[f(x)] = \int dx\; dz\, p(x,z|\theta)\,|f(x,z) - f(x)|^2$$

$\downarrow$ minimized by

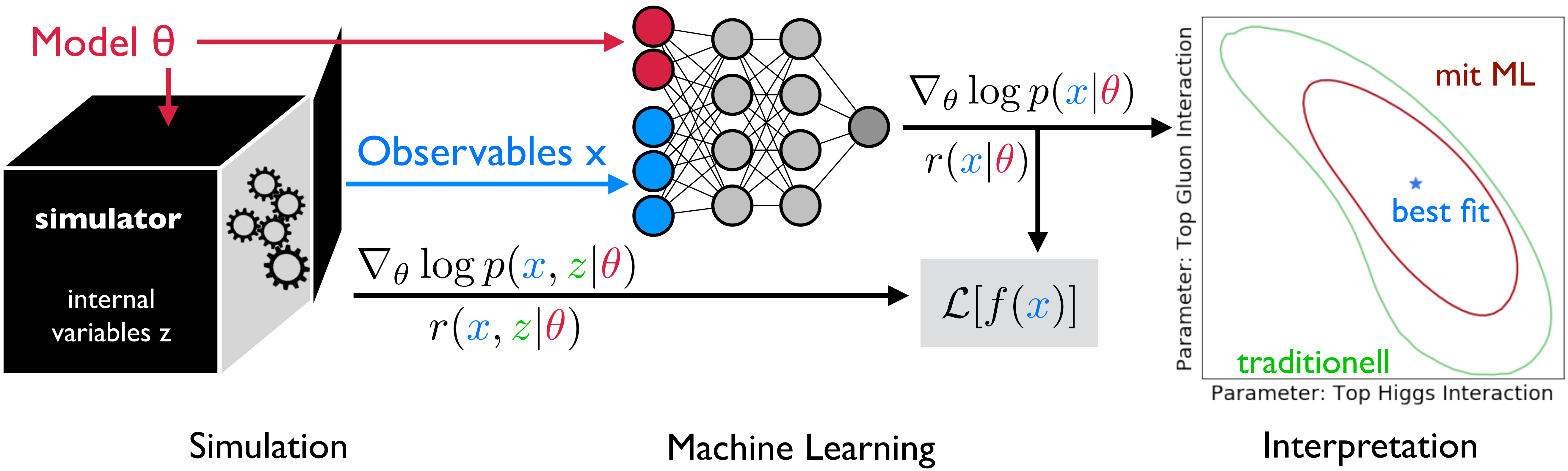
$$f(x) = \int dz\, p(x,z|\theta) f(x,z)/p(x|\theta)$$

$$f(x,z) = r(x,z|\theta)$$

$\downarrow$ $\mathcal{L}[f(x)]$ minimized by

$$f(x) = r(x|\theta)$$

# MadMiner Method




**MadMiner automatizes this method**

**MadMiner: Machine learning-based inference for particle physics**

[Brehmer, Cranmer, Espejo, FK 1907.10621], [Brehmer, Cranmer, Espejo, FK, Louppe, Paves 1906.01578]
[Brehmer, Dawson, Homiller, FK, Plehn 1908.06980]

github.com/diana-hep/madminer

pip install madminer

# Method

- <u>Work to be done:</u>
  - interface MadMiner method/code with POWHEG (instead with MadGraph)
  - deal with negative wights due to NLO: implement unweighting
  - fast detector simulation (partially done)

- Sensitivity estimates trivial due to Cramer Rao Bound

$$C_{ij}(\theta) \geq (I_{ij})^{-1}, \qquad I_{ij}(\theta) = \mathbb{E}\left[\frac{\partial \log p_{\mathrm{full}}(\{x\}|\theta)}{\theta_i} \frac{\partial \log p_{\mathrm{full}}(\{x\}|\theta)}{\theta_j}\middle|\theta\right]$$

  - how information is distributed over phase space,
  - identify the most useful observables,
  - compare the sensitivity of multivariate and traditional analysis strategies

# Applications

- How well can we constrain PDFs in presence of flux uncertainties?
- Can we see IC/low-x effects in presence of other uncertainties
- Sterile neutrino oscillations?
- New Tau-philic BSM effects? NSI?

Framework can be used for various things. So let's build it with this in mind.