End-to-end Toy Analysis (iSTEP2023)

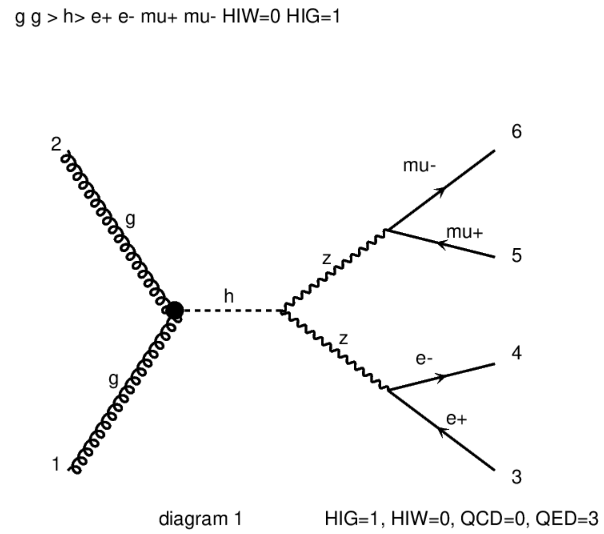
Tongguang Cheng

2023 年 8 月 26 日

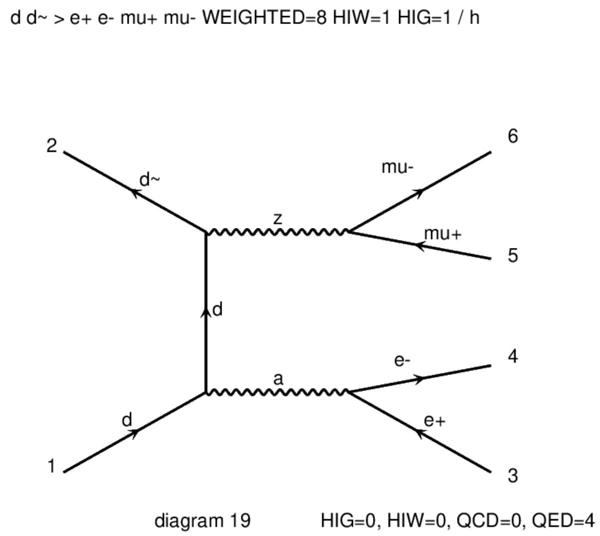
目录

1 定义你的目标

- 信号是什么，有哪些参数需要从数据中获得（粒子的质量/粒子产生过程的散射截面）？
- 背景是什么，背景是指可以给出和信号相似特征（比如模态粒子）的非信号过程
- *Rank background by their importance, and see to what extend you can predict the background yield (as a function of observables) with a reasonable precision.



(a) process a



(b) process b

2 制定事例选择条件

- 用 Madgraph 生成信号和背景过程的样本
- 分析 Delphes 层级（经过了 parton shower , hadronization , detector simulation/reconstruction）

的数据

  - 通过柱状图找到可以区分信号和背景过程的可观测量，通过这些变量选择合理的信号区间（合理是指有好的信号选择效率和合理的信号背景比）

- 计算在信号区间的信号和背景的期望的事例数 N_s, N_b

$$
\begin{aligned}
N^s_{expected} &= Luminosity \times \sigma^s \times AccEff^s(Delphes) \times Eff^{signal}(selection), \\
& AccEff^s(Delphes) = N^s_{Delphes}/N^s_{LHE}, \\
& Eff^s(Selection) = N^s_{Delphes\ final}/N^s_{Delphes} \\
& \sigma^s = \mu_s \times \sigma^s_{LHE}
\end{aligned}
$$

$$
\begin{aligned}
N^b_{expected} &= \sum_{i=1}^{N^b_{Delphes\ final}} weight_i \\
& weight_i = \frac{Luminosity \times \sigma^b_{LHE}}{N^b_{LHE}}
\end{aligned}
$$

3 建立统计模型，拟合信号参数

- 选择对信号参数敏感的可观测量 (observable), 比如 $pp > H > 2e2mu$ 中的 $m(e^+e^-\mu^+\mu^-)$
- 建立信号和背景的 $ExtendedPdf$ 模型:
  - signal $(N^s_{exp},\ PDF^s(m(e^+e^-\mu^+\mu^-)|m_H)$
  - backgroud,$(N^b_{exp},\ PDF^b(m(e^+e^-\mu^+\mu^-)))$
  - model = RooAddPdf(”sum”,”sum”,RooArgList(signal,background),RooArgList(nsig,nbkg))
- 生成数据
- 模型拟合数据，展示结果

模型直接产生数据

$$
\begin{aligned}
dataset &= model.generate(massReco)\ (N_{data} = N_{exp}) \\
dataset &= model.generate(massReco, Extend = True)(N_{data}\ followsPoissondistributionofN_{exp})
\end{aligned}
$$

产生 Weighted events

$$
\begin{aligned}
dataset &= ROOT.RooDataSet(\text{"}dataset\text{"}, \text{"}dataset\text{"}, massReco, eventWeight, Import = tree)commen \\
weight &= ROOT.RooFormulaVar(\text{"}weight\text{"}, \text{"}1000 * lumi * eventWeight\text{"}, ROOT.RooArgSet(lumi, \\
& dataset.addColumn(weight) \\
dataset_weighted &= ROOT.RooDataSet(\text{"}dataset_w\text{"}, \text{"}dataset_w\text{"}, dataset, dataset.get(), \text{""}, \text{"}weight\text{"})
\end{aligned}
$$

拟合数据

$results = p2.fitTo(wdata, Save = True, SumW2Error = True(weighted/events), Extend = True, PrintLevel =$

展示结果