

用MATLAB 探索基本粒子

计算机工程大作业汇报



汇报人：

尹琪钦 211870080 物理学



LHC 和其上的若干探测器实验



Outline

作为计算机工程的课程报告，本次是 MATLAB 作为物理分析工具的介绍讲座

报告内容主要包括：

- 粒子物理的基本内容介绍：

粒子的动力学、基本粒子标准模型

- 粒子物理实验数据处理、分析技术：

物理量重建、质量谱峰值分析、实验信号拟合、统计分析

既有物理知识也有分析演示，是一次专业性和趣味性兼备的报告

- 希望可以激发大家对 MATLAB 和粒子物理的兴趣，让大家有所收获



目录

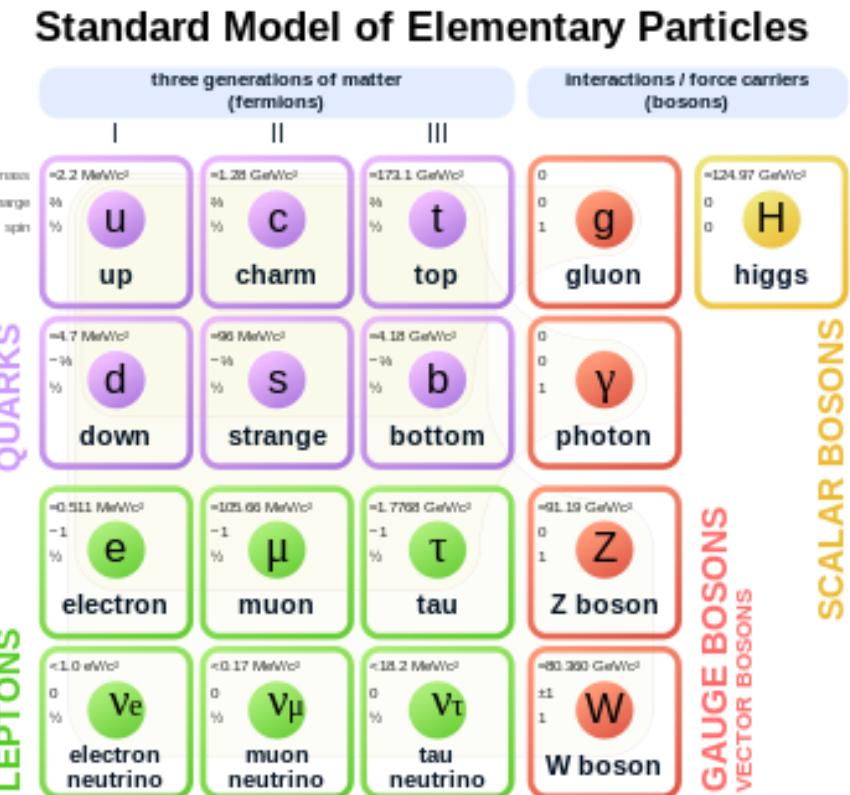
一、粒子物理概要

二、粒子质量谱与统计分析

三、寻找Higgs 粒子

结语：总结与感悟

第一节：粒子物理概要



1. 粒子的动力学简介
2. 粒子物理标准模型

1、粒子的动力学基础



(1)、狭义相对论动力学 (2)、粒子的寿命和质量

- 通修课-大学物理中我们曾初探狭义相对论，

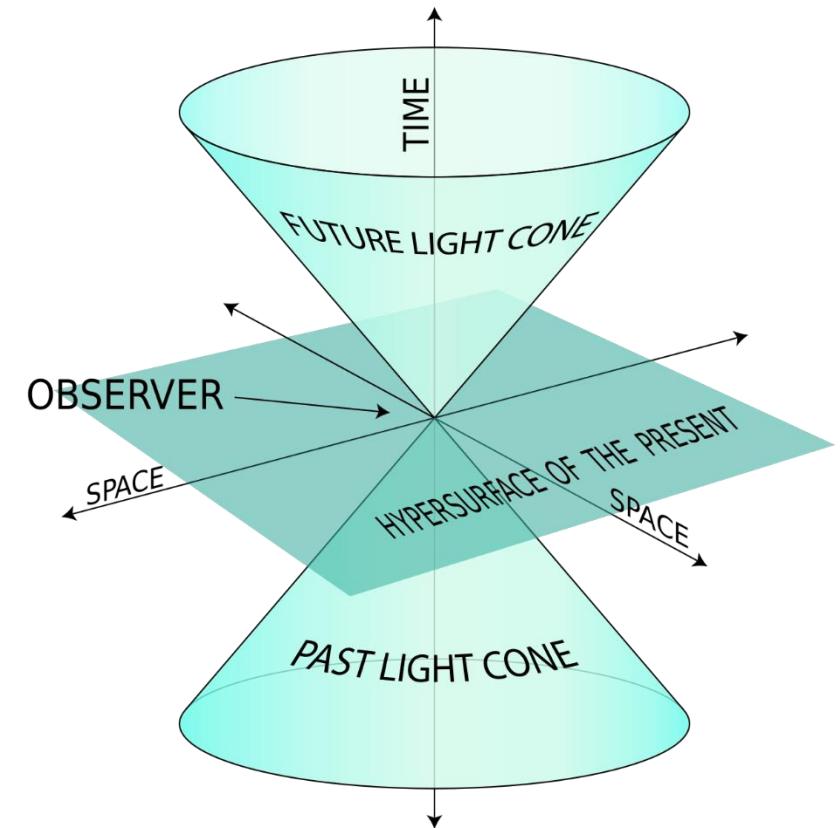
印象最深的莫过于Lorentz 变换：

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma ct - \gamma\beta x \\ \gamma x - \beta\gamma ct \\ y \\ z \end{pmatrix} \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

低速极限下($v \ll c$)退化到牛顿力学中我们熟知的 Galilean 变换：

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ x - vt \\ y \\ z \end{pmatrix}$$

- 实质上反映了不同的时空几何结构



Minkowski 时空示意图

1、粒子的动力学基础



南京大學
NANJING UNIVERSITY

(1)、狭义相对论动力学 (2)、粒子的寿命和质量

- 采用四动量 \mathbf{P}^μ 刻画粒子在时空中的运动

$$\mathbf{P}^\mu = m\mathbf{U}^\mu = \gamma_u(mc, m\vec{u}) = \left(\frac{E}{c}, \vec{p}\right)^T$$

- 为了方便，往往采用自然单位制，
取 $c = \hbar = 1$

- 四动量 \mathbf{P}^μ 的内积给出Lorentz 不变量

1. 由内积的定义:

$$\mathbf{P}^2 = \eta^{\mu\nu} P_\mu P_\nu = -\left(\frac{E}{c}\right)^2 + p^2$$

2. 在粒子静止系计算:

$$\mathbf{P}^2 = -\left(\frac{E_{\text{rest}}}{c}\right)^2 = -(mc)^2.$$

$$\Rightarrow E^2 = \vec{p}^2 + m^2$$

相对论能-动量关系

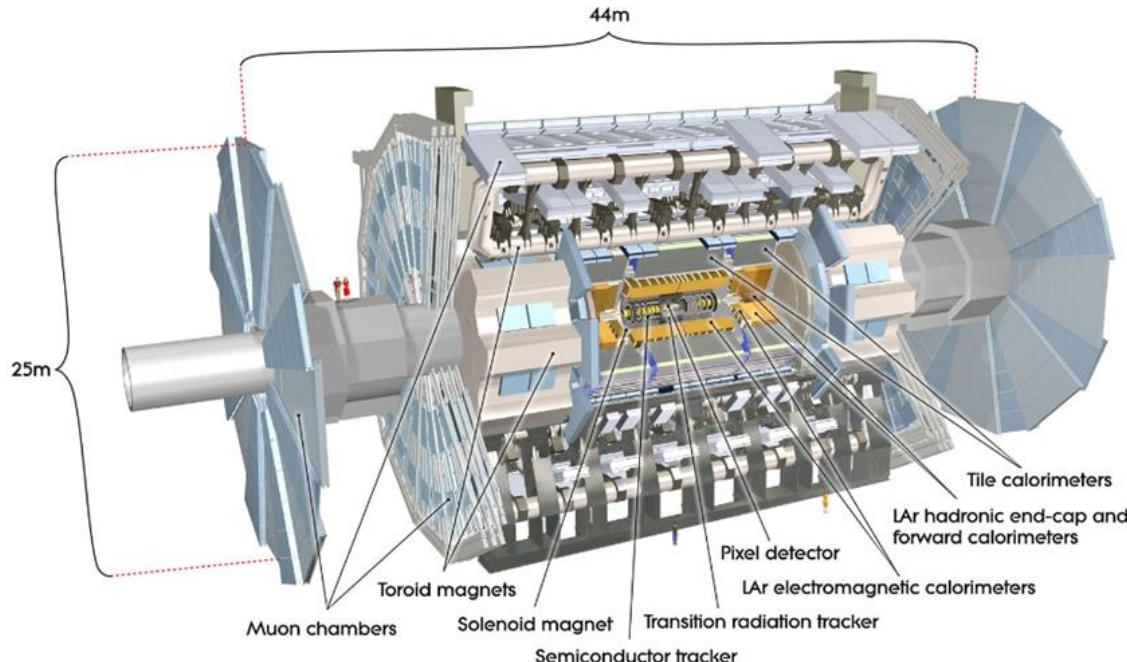
vs. 牛顿力学能-动量关系: $E = \frac{\vec{p}^2}{2m}$

1、粒子的动力学基础



(1)、狭义相对论动力学 (2)、粒子的寿命和质量

• 对撞机结构与动力学量的测量：



ATLAS 探测器实验布局

- 由对撞机中的径迹探测器、量能器等结构，我们测量得到相关的动力学量，如能量、三维动量.....
- 探测器轴向为 z 轴，粒子与 z 轴夹角 θ ，定义赝快度 (pseudorapidity):

$$\eta = -\ln \tan \frac{\theta}{2}$$

于是轴向和横向($x - y$ 平面内)动量关系：

$$\begin{aligned} p_z &= p_t \sinh(\eta) \\ &= p_t \cot(\theta) \end{aligned}$$

1、粒子的动力学基础

(1)、狭义相对论动力学 (2)、粒子的寿命和质量

- 不稳定粒子-指数衰变规律: $dN = -\Gamma_{fi} N dt$ (不难由费米黄金规则得到)

- 积分后得到: $N(t) = N(0)e^{-\Gamma_{fi} t} = N(0)e^{-\frac{t}{\tau}}$

得到得到粒子平均寿命的定义: $\tau \equiv \frac{1}{\Gamma_{fi}}$

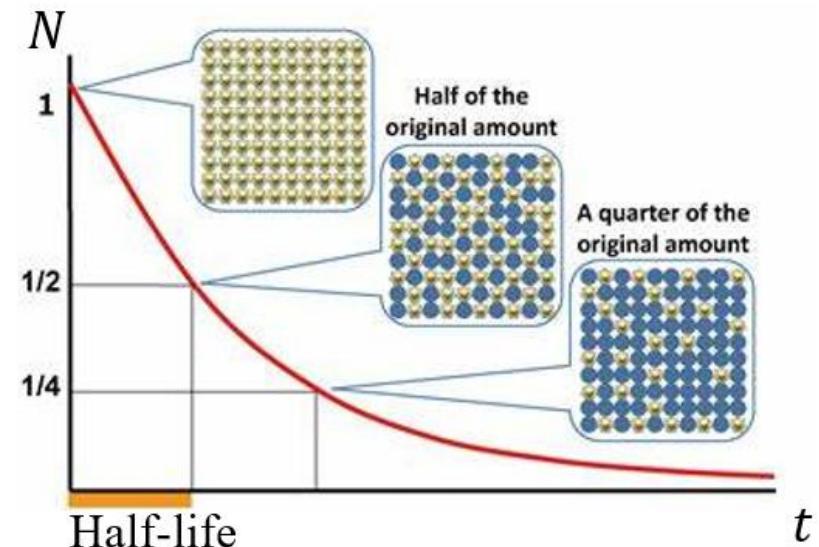
- 微观粒子的演化规律-Schrödinger 方程:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \hat{H} \psi(\mathbf{r}, t)$$

自由粒子静止系有: $\psi(\mathbf{r}, t) = e^{-imt} \psi(\mathbf{r})$

$$\int_{-\infty}^{\infty} |\psi(\mathbf{r}, t)|^2 d^3x = 1$$

- 归一化条件暗示: 非相对论框架下
粒子不会衰变。如何描述不稳定性?



指数衰变与半衰期

1、粒子的动力学基础



(1)、狭义相对论动力学 (2)、粒子的寿命和质量

- 不稳定粒子的唯象描述： $\psi(\mathbf{r}, t) = e^{-imt} e^{-\frac{t}{2\tau}} \psi(\mathbf{r})$ 原有波函数上引入衰减因子
- 的确可以描述粒子概率的衰减： $|\psi(\mathbf{r}, t)|^2 \propto e^{-\frac{t}{\tau}}$

以 t 为变量做 Fourier 变换，共轭变量设为 M ，
得到表示不稳定粒子质量分布概率的波函数：

$$\begin{aligned}\phi(\mathbf{r}, M) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(\mathbf{r}, t) e^{iMt} dt = \frac{\psi(\mathbf{r})}{\sqrt{2\pi}} \int_0^{\infty} e^{i(M-m)t} e^{-\frac{t}{2\tau}} dt \\ &= \frac{\psi(\mathbf{r})}{\sqrt{2\pi}} \frac{i}{M - m + \frac{i}{2\tau}}\end{aligned}$$

不稳定粒子质量分布的概率密度： $|\phi(\mathbf{r}, M)|^2$

$$W(M) = N \phi^*(\mathbf{r}, M) \phi(\mathbf{r}, M) = N \frac{|\psi(\mathbf{r})|^2}{2\pi} \frac{1}{(M - m)^2 + \frac{1}{4\tau^2}} \quad N \text{ 为归一化系数}$$

1、粒子的动力学基础

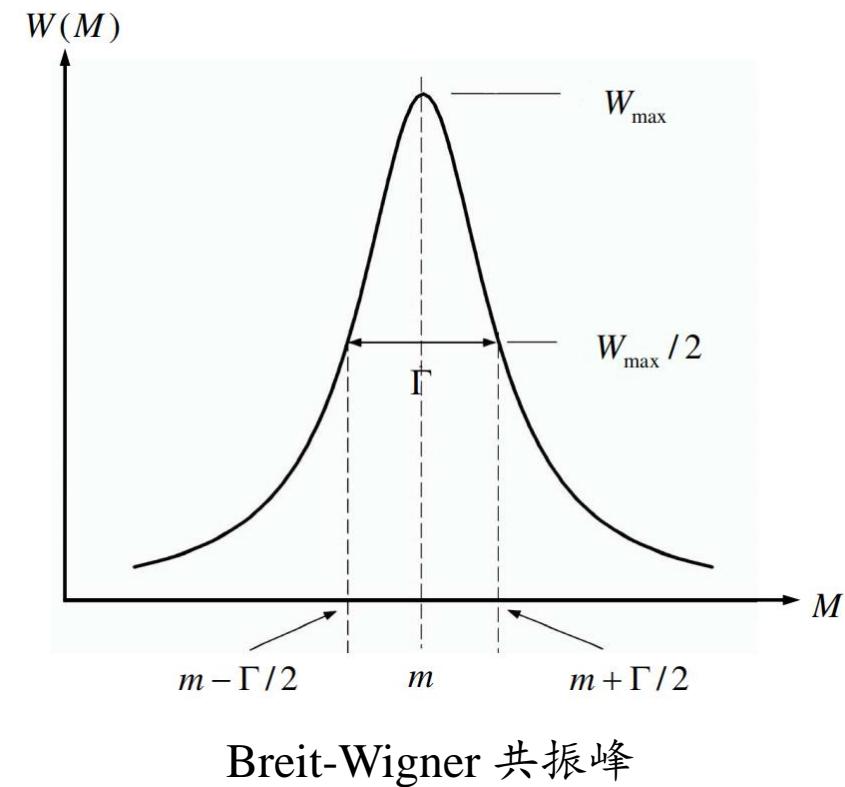
(1)、狭义相对论动力学 (2)、粒子的寿命和质量

- 不稳定粒子的唯象描述: $\psi(\mathbf{r}, t) = e^{-imt} e^{-\frac{t}{2\tau}} \psi(\mathbf{r})$

归一化后得到不稳定粒子质量概率分布, 也就是 Breit-Wigner 公式, 实验上对应质量分布谱:

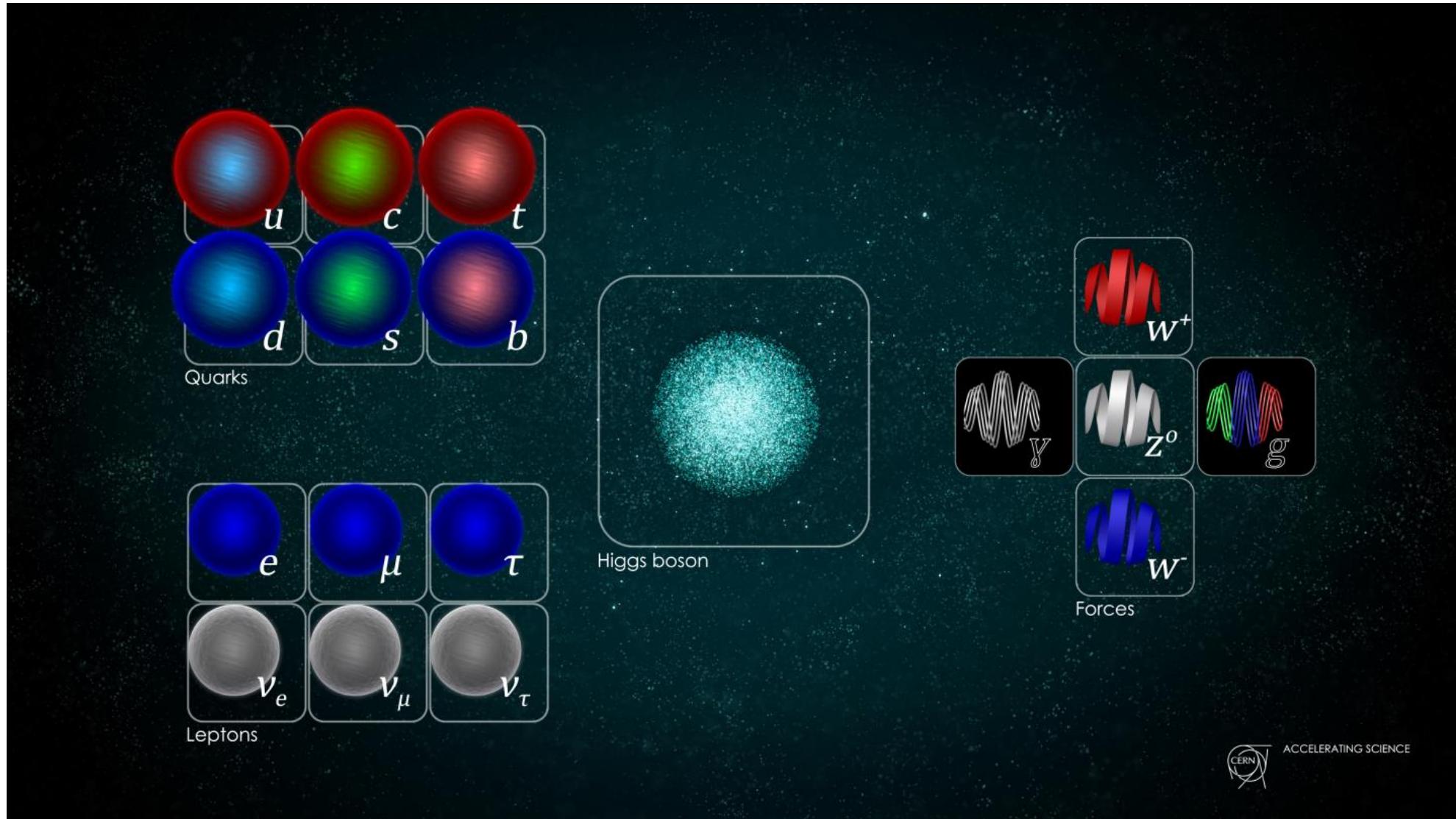
$$W(M) = \frac{1}{2\pi} \frac{\Gamma}{(M - m)^2 + \left(\frac{\Gamma}{2}\right)^2}$$

- 形式上和经典力学共振十分相似
- 不难验证 Γ 就是共振峰的半高全宽(FWHM)
- 峰的位置 m 和宽度 Γ 是实验上感兴趣的物理量



Breit-Wigner 共振峰

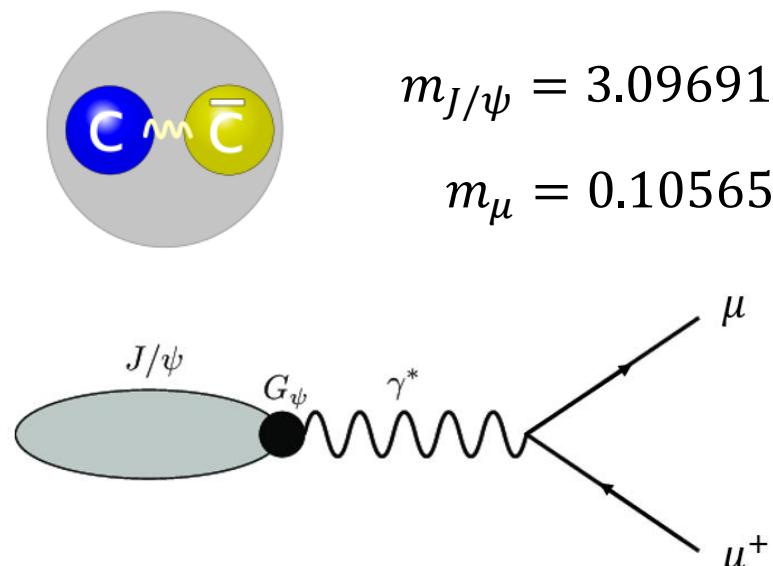
2、粒子物理标准模型



2、粒子物理标准模型

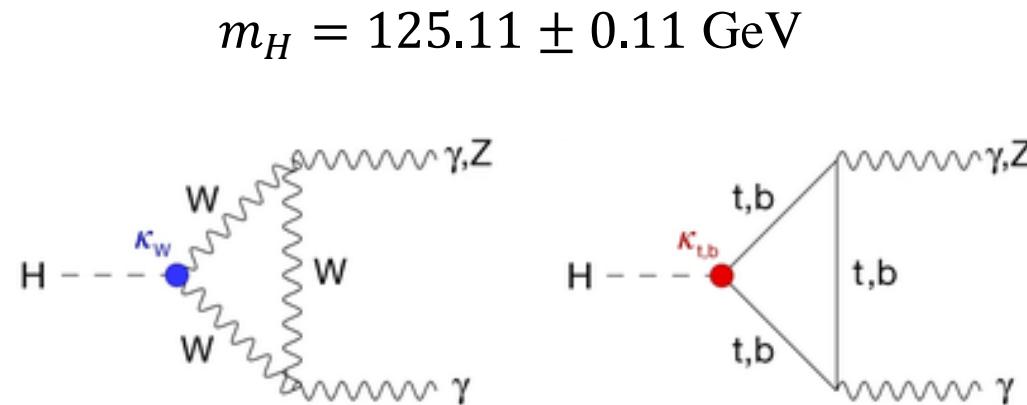
- 标准模型是目前人类理解微观世界最成功的理论，
本次报告将聚焦两种粒子：

1. J/ ψ Meson:



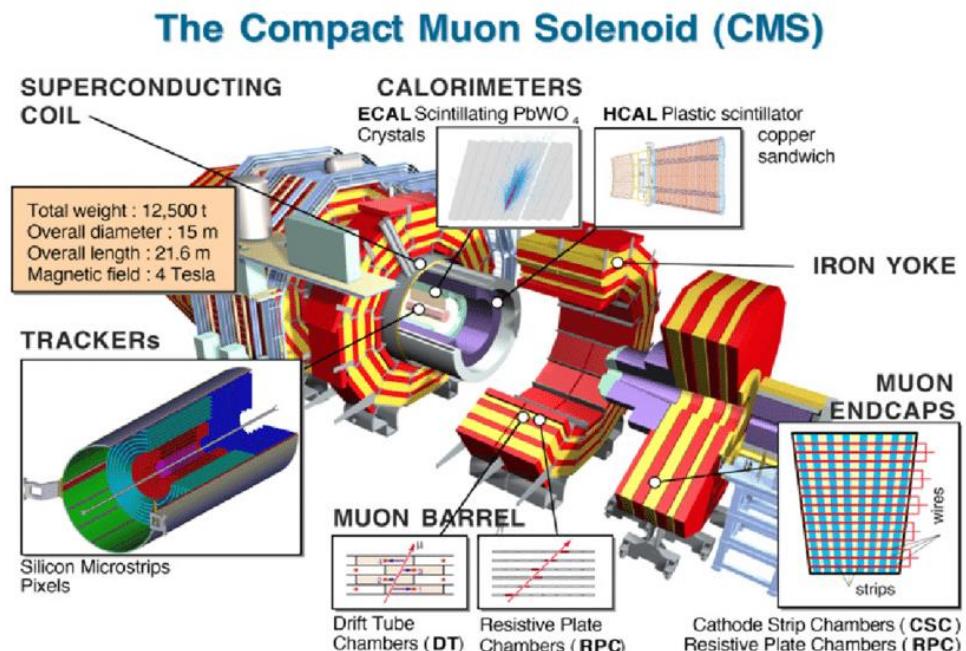
$J/\psi \rightarrow \mu^+ \mu^-$ 过程费曼图

2. Higgs Boson:

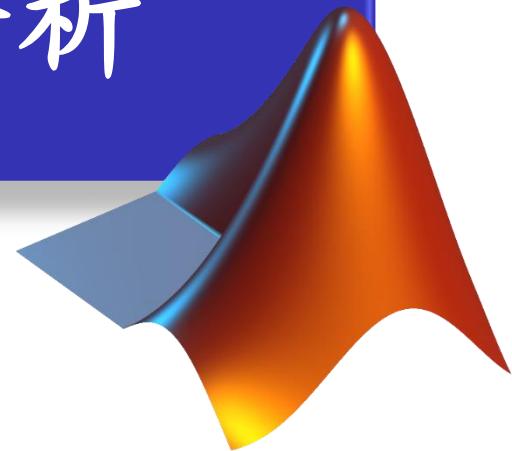


Higgs 粒子衰变为光子的可能模式

第二节：粒子质量谱与统计分析



CMS 探测器实验



1. 探索实验数据
2. 质量共振峰参数重建

1、数据介绍与探索

(1)、数据来源、基本结构 (2)、动力学量分布

- 数据来源： Datasets derived from the Run2011A

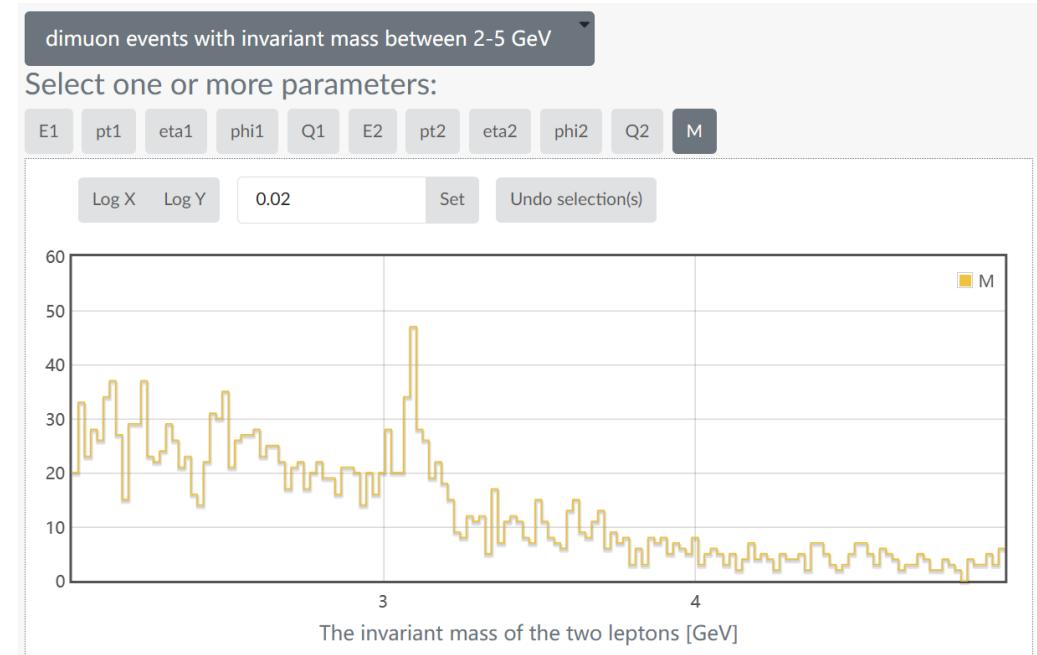
<https://opendata.cern.ch/record/545>

- CMS 探测器采集的数据经过初步筛选得到对应粒子反应的多个数据文件：

- Wmumu.csv
- Zmumu.csv
- Jpsimumu.csv

(每个事件有两个异号muon, $|\eta| < 2.4$,
系统不变质量 $2 \text{ GeV} < m_{ll} < 5 \text{ GeV}$)

- Dimuon_DoubleMu.csv
-



Open data 网站提供的交互式直方图可视化
(以dimuon 事件为例，质量谱存在一个峰)

1、数据介绍与探索



(1)、数据来源、基本结构 (2)、动力学量分布

- 数据结构([Jpsimumu.csv](#)): 共20000行，每行分别记录两个muon的相关信息

Run	Event	type1	E1	px1	py1	pz1	pt1	eta1	phi1	Q1	type2	E2	px2	py2	pz2	pt2	eta2	phi2	Q2
165617	75206813	G	10.1623	0.4763	-8.5164	5.5231	8.5297	0.6091	-1.5149	-1	G	19.7754	2.7508	-13.9871	13.7059	14.255	0.8539	-1.3766	1
165617	75678475	G	15.8799	15.0618	-1.6658	-4.7464	15.1536	-0.3083	-0.1102	1	G	11.9338	11.9084	-0.767	-0.0851	11.933	-0.0071	-0.0643	-1
165617	74428554	G	21.8279	-6.2214	11.0845	17.7447	12.7111	1.1357	2.0823	-1	G	36.6513	-10.6874	15.0692	31.6544	18.4743	1.3076	2.1877	1
165617	75193169	G	19.4923	2.7612	-5.5769	-18.4719	6.223	-1.8084	-1.1111	1	G	14.9334	1.5525	-6.1254	-13.5302	6.319	-1.505	-1.3226	-1
165617	74832715	G	8.0972	4.6127	-1.8389	6.3949	4.9657	1.071	-0.3793	-1	T	3.5131	-0.5305	0.688	3.4024	0.8688	2.0742	2.2276	1
165617	74981507	G	30.5862	15.5218	5.1293	-25.8509	16.3474	-1.2391	0.3192	1	G	12.4868	7.764	1.0425	-9.7233	7.8337	-1.0421	0.1335	-1
165617	75612982	G	7.5544	1.2913	-5.3588	-5.1646	5.5122	-0.8361	-1.3343	-1	T	3.8416	-0.9896	-3.4548	-1.3534	3.5938	-0.3682	-1.8498	1
165617	74760204	G	24.6376	9.9836	19.1426	-11.8697	21.5896	-0.5253	1.0901	1	G	11.0213	3.3785	7.8272	-6.9841	8.5252	-0.7476	1.1633	-1
165617	75017052	G	22.5057	5.6296	-9.9497	19.3857	11.4319	1.2987	-1.0559	1	G	30.5709	4.9468	-13.636	26.9102	14.5055	1.3769	-1.2228	-1
165617	75204668	G	10.7051	-8.4664	-2.2374	-6.1564	8.7571	-0.6551	-2.8832	1	G	55.1896	-45.0429	-16.6152	-27.2205	48.0096	-0.5403	-2.7882	-1
165617	75682819	G	22.6067	5.0267	-0.3155	22.0383	5.0366	2.182	-0.0627	1	G	17.8797	2.7987	2.0567	17.5388	3.4732	2.3222	0.6338	-1
165617	75891163	G	11.6778	10.7636	2.2556	-3.9266	10.9974	-0.3499	0.2066	1	G	11.1458	9.0065	3.4698	-5.5732	9.6517	-0.5494	0.3677	-1
165617	74949022	G	22.4059	1.3759	7.2096	-21.1694	7.3397	-1.7812	1.3822	-1	T	4.7813	-0.5663	0.7255	-4.6907	0.9204	-2.3312	2.2335	1
165617	75818801	G	25.2596	23.5556	3.1909	8.5432	23.7708	0.3521	0.1346	1	G	11.7166	10.4432	1.1288	5.1896	10.504	0.4759	0.1077	-1
165617	74875638	G	23.0765	4.4733	7.9045	21.2138	9.0825	1.5844	1.0558	-1	T	25.5421	1.9891	7.7265	24.2637	7.9785	1.8314	1.3188	1
165617	76187234	G	9.7908	-1.6247	2.0609	9.432	2.6243	1.9912	2.2384	1	G	11.708	1.3215	1.7092	11.5064	2.1605	2.3744	0.9126	-1
165617	76650121	G	14.5875	-6.1673	12.5515	-4.1482	13.9849	-0.2924	2.0275	1	G	10.6562	-3.3095	8.6271	-5.3071	9.2401	-0.5467	1.9371	-1
165617	76624220	G	8.8197	-0.0263	2.5635	-8.4383	2.5636	-1.9068	1.581	-1	T	7.4905	-1.2713	-0.6276	-7.3543	1.4178	-2.3485	-2.683	1
165617	76796604	G	13.5186	5.4299	7.9615	-9.4801	9.6369	-0.8698	0.9722	1	G	12.9792	7.0747	5.3189	-9.4925	8.8511	-0.9317	0.6447	-1
165617	76691627	G	4.886	-4.2066	-2.0383	1.4184	4.6744	0.299	-2.6904	1	T	5.2678	-2.2166	-2.7244	3.9246	3.5122	0.962	-2.2538	-1

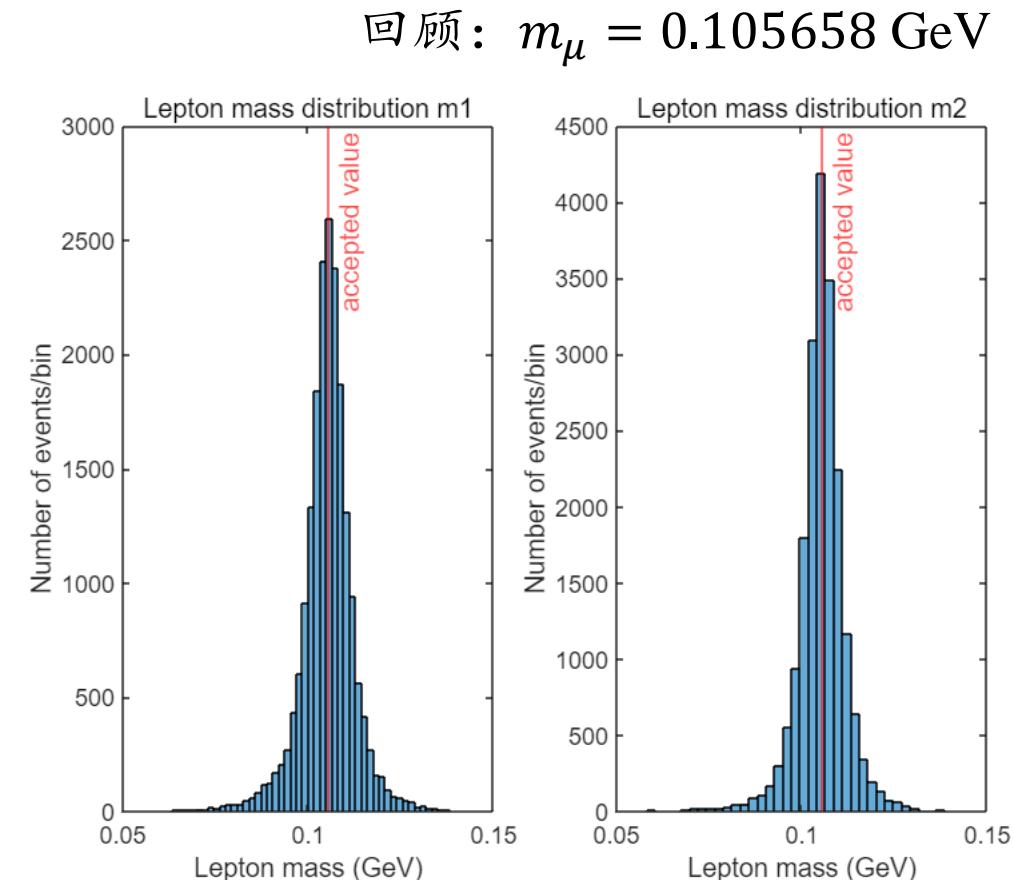
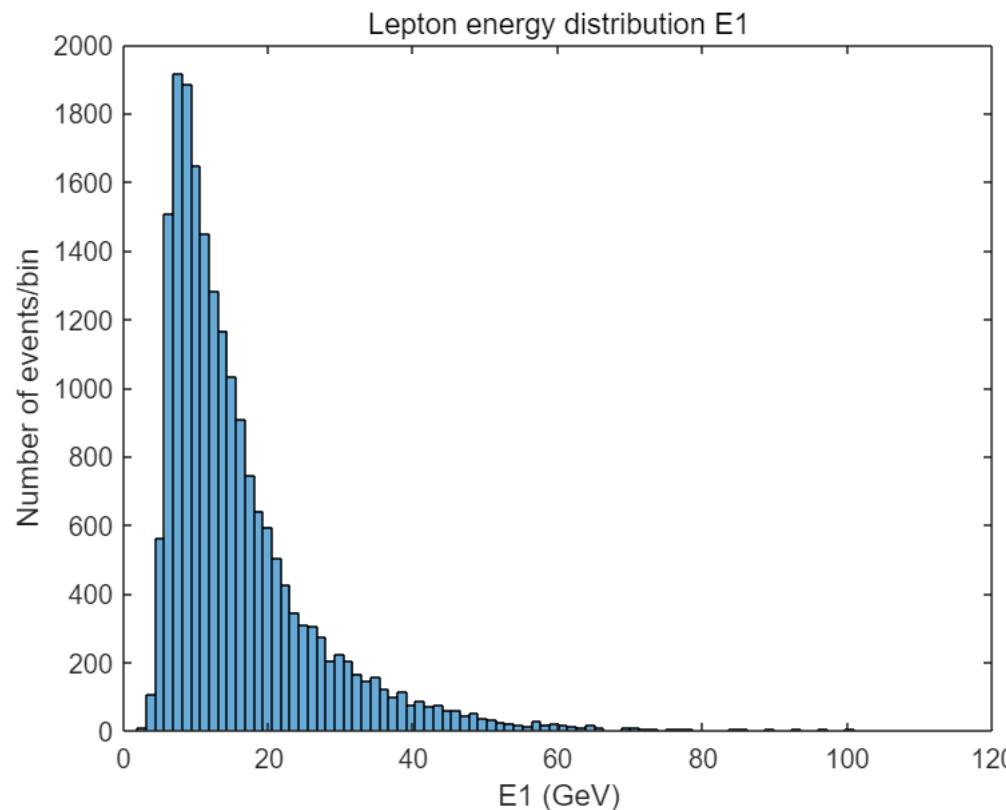
表格每列名称与部分数据展示

1、数据介绍与探索



(1)、数据来源、基本结构 (2)、动力学量分布

- 先在MATLAB中下载网站数据到本地
- 接下来，初步探索物理量的分布：



2、质量共振峰分析

(1)、共振峰的位置与宽度 (2)、估计质量置信区间

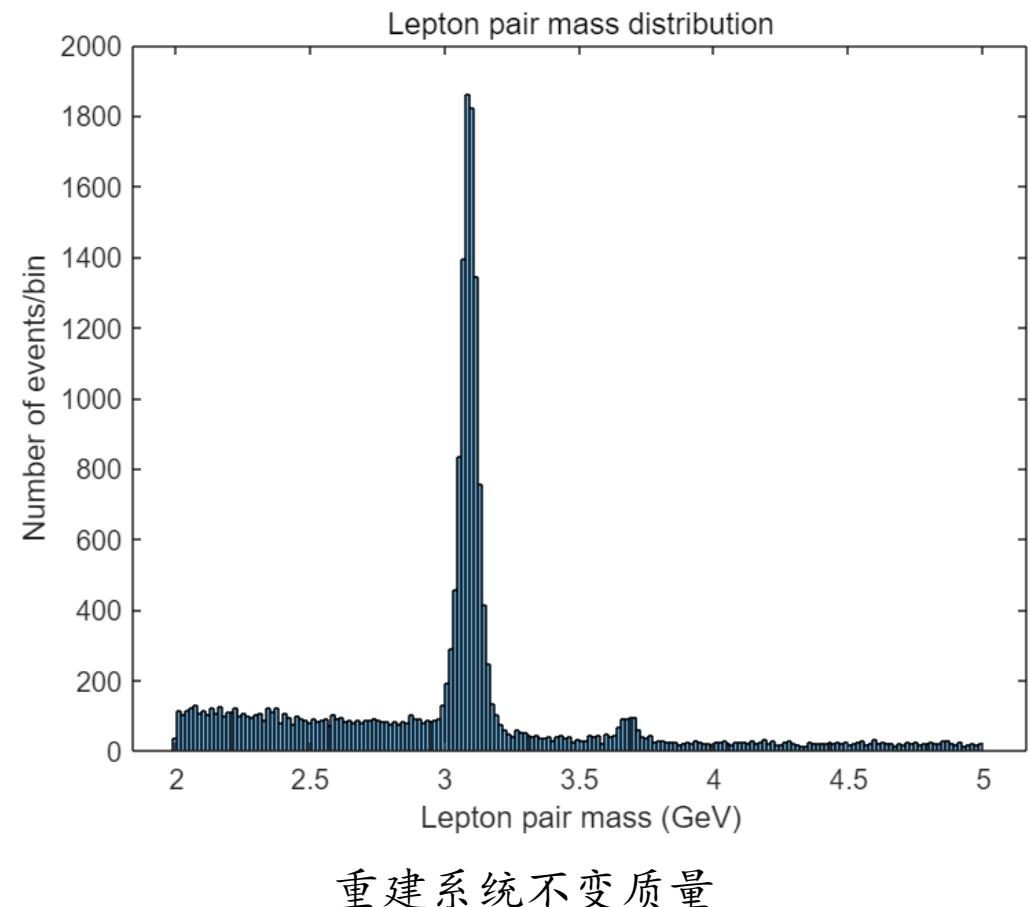
- 通过能-动量关系重建dimuon 末态系统不变质量:

$$m_{ll}^2 = \left(\sum_i P_i^\mu \right)^2 = \left(\sum_i E_i \right)^2 - \left(\sum_i \vec{p}_i \right)^2$$

- 得到 m_{ll} 统计直方图(右图):

系统不变质量谱存在两个**共振峰**, 其中左侧的峰十分尖锐, 大致对应 J/ ψ 粒子质量 3.096916 GeV

- 需要更详细**分析共振峰**, 从而估计粒子质量和**衰变宽度**



2、质量共振峰分析

(1)、共振峰的位置与宽度 (2)、估计质量置信区间

- 借助findpeaks 函数确定质量共振峰参数:

峰值高度: pks = 1×2

1862 97

峰位置: locs = 1×2

3.0862 3.7061

峰宽度: widths = 1×2

0.0675 0.0825

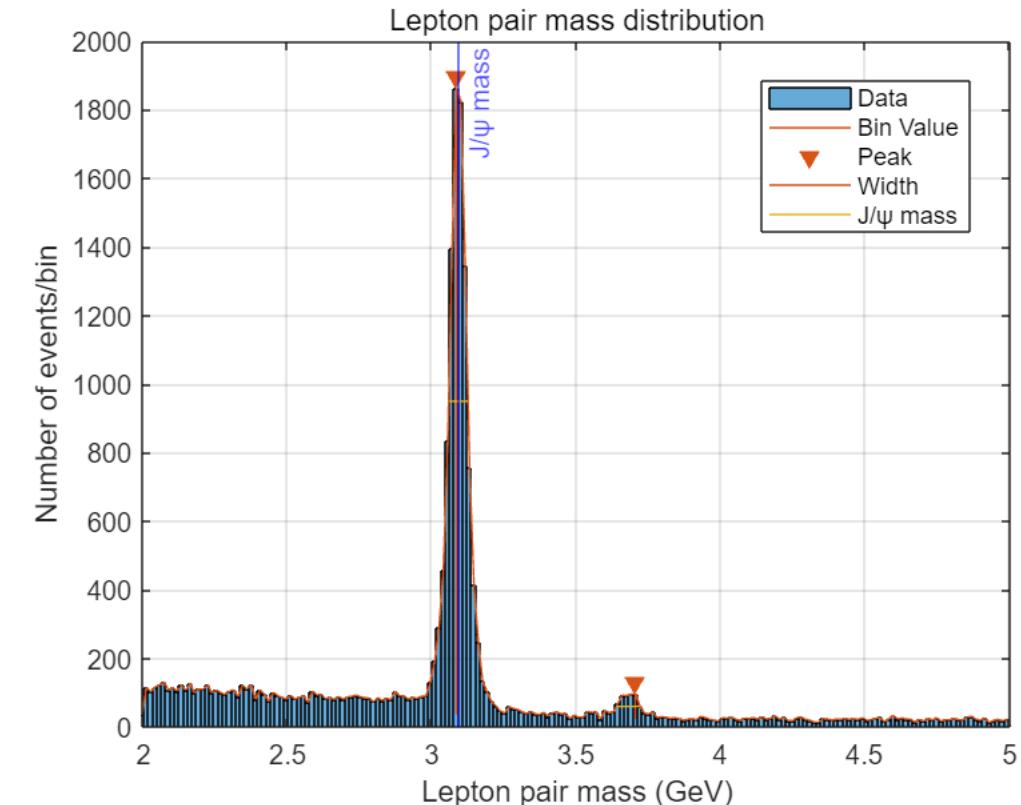
相对突出度: proms = 1×2

1825 74

- 计算共振峰信号显著度:

$$\frac{N_{\text{sig}}}{\sqrt{N_{\text{bkg}}}} \longrightarrow N_{\text{tot}}(\text{峰区域内总事例数}) - N_{\text{bkg}}$$

$$\longrightarrow \text{用峰两侧区域事件数估计}$$



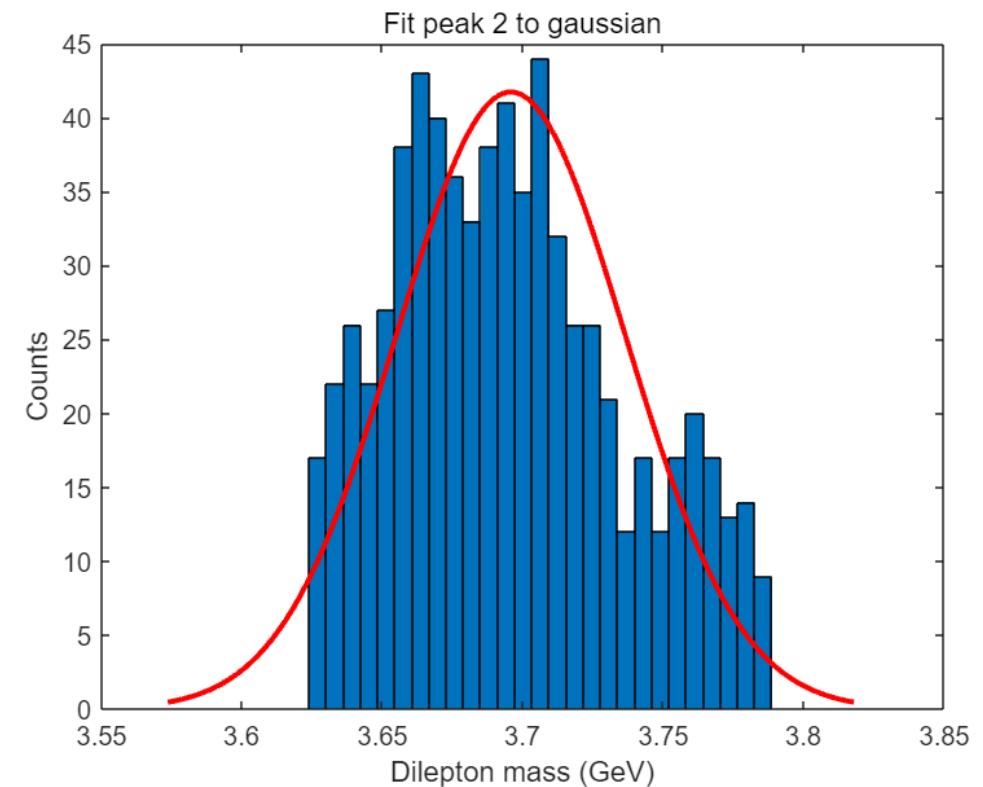
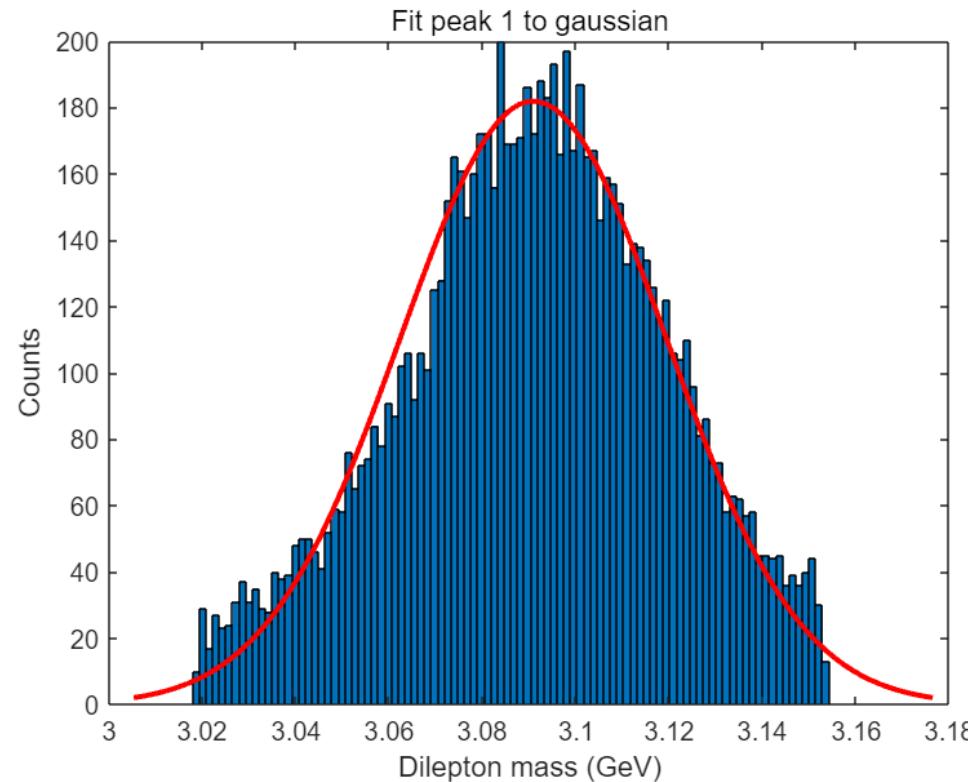
$$\Rightarrow$$

PeakSig = 2×1
360.6501
38.2485

2、质量共振峰分析

(1)、共振峰的位置与宽度 (2)、估计质量置信区间

- 用histfit 函数以正态分布拟合共振峰，得到拟合参数 μ, σ :

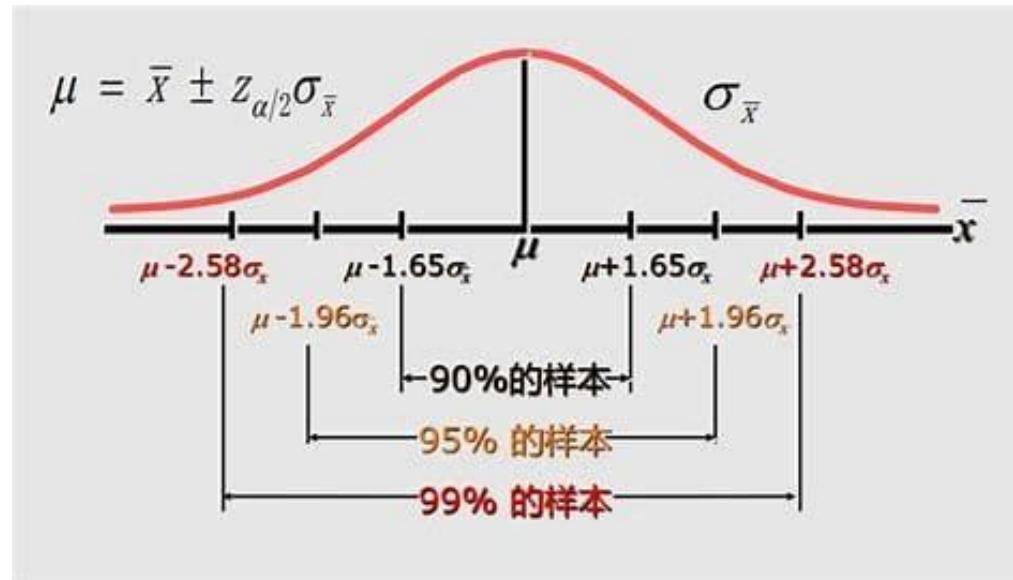




2、质量共振峰分析

(1)、共振峰的位置与宽度 (2)、估计质量置信区间

- 用借助正态分布性质和参数估计 m_{ll} 置信区间:

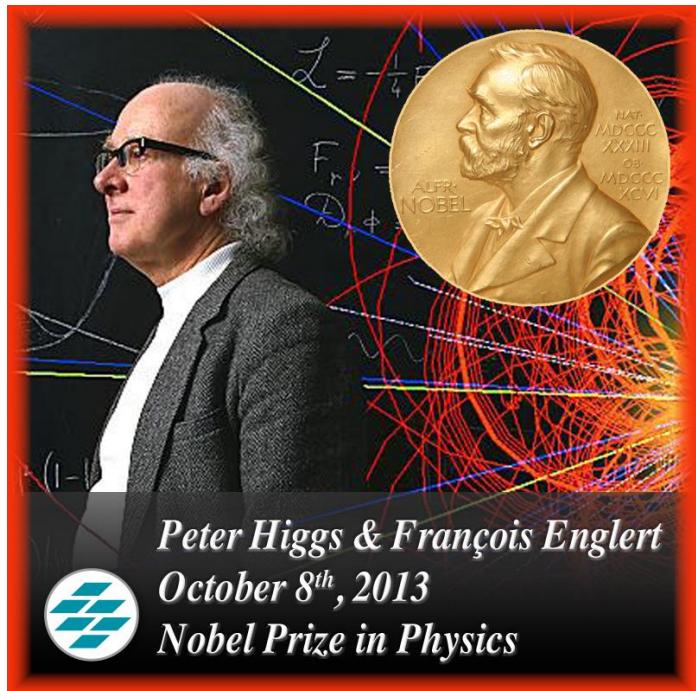


正态 分布
Peak 1: mu = 3.09087 [3.09029, 3.09146]
 sigma = 0.0285251 [0.0281179, 0.0289443]

正态 分布
Peak 2: mu = 3.69568 [3.69266, 3.69871]
 sigma = 0.0407036 [0.0386746, 0.042959]

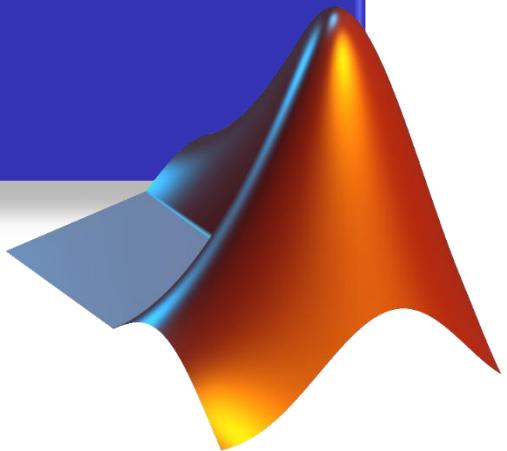
- 用信号区域内样本方差开根估计正态分布标准差
- $\mu \pm 1.96 \text{ std}$ 给出质量分布 95% 置信区间:
 $m_1 - (3.0903, 3.0915)$
 $m_2 - (3.6935, 3.6978)$
- 结果和 J/ψ 粒子质量接近, 误差在 0.2%
若将所有系统误差全部考虑、采用更好的统计方法将提高质量估计精度

第三节：寻找Higgs 玻色子



2013年诺奖得主-Peter Higgs

1. 数据结构处理和筛选
2. 质量谱拟合寻找信号事件



1、数据介绍与处理

(1)、数据来源、基本结构 (2)、数据筛选

- 数据来源：



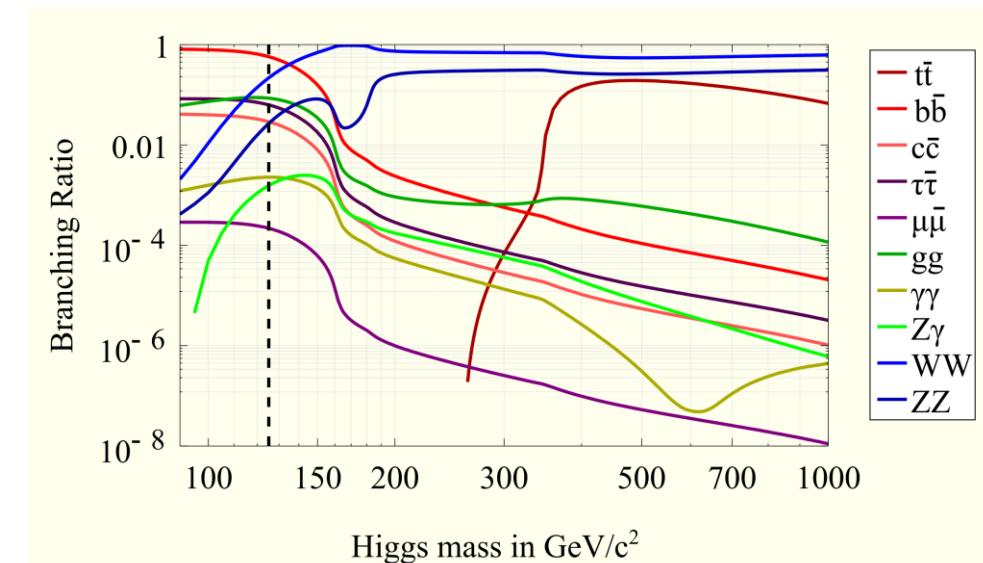
ATLAS 13 TeV Open Data

[Datasets in ROOT format | ATLAS Open Data](#)

<https://opendata.atlas.cern/docs/datasets/files>

- 选择其中 $H \rightarrow \gamma\gamma$ 衰变道的数据进行分析

尽管该衰变道分支比略低，但是光子事件的重建、确认效率高，使得信号峰较明显、尖锐



Higgs boson 衰变分支比

1、数据介绍与处理

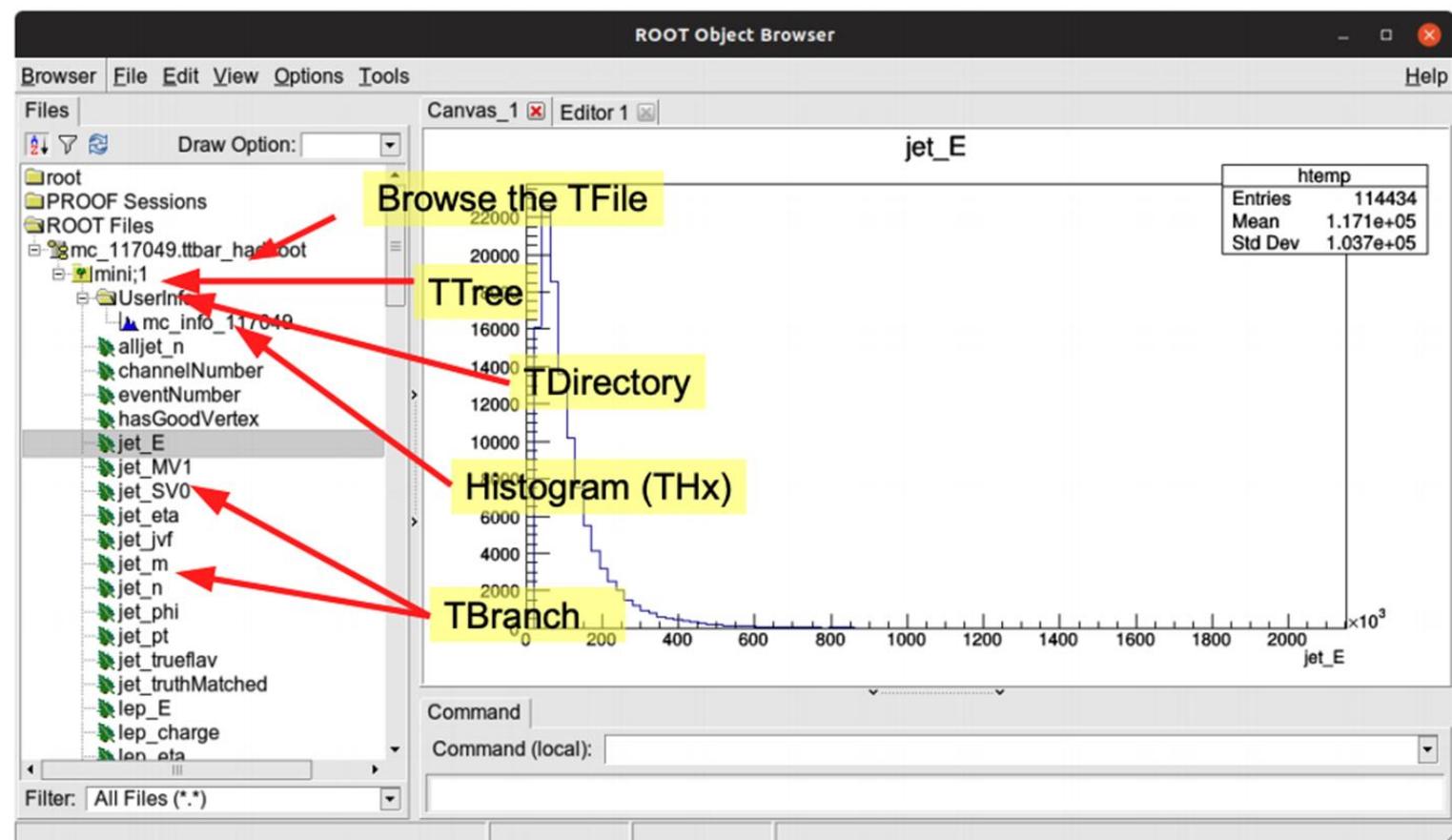


(1)、数据来源、基本结构 (2)、数据筛选

- 原始数据以.root 格式提供

- 不同于常见的 .csv 格式以表格存储信息
- .root 文件以 TTree 存储数据，**结构层次性更明显**

需要借助 **python uproot3** 框架转换为 MATLAB 方便处理的格式



CERN ROOT GUI

1、数据介绍与处理



(1)、数据来源、基本结构 (2)、数据筛选

- 原始数据转换为.csv格式后两个一个entry 双光子的物理信息对应两行
- 保持了.root 文件的层次信息，共约 860000 行数据

entry	subentry	photon_pt	photon_eta	photon_phi	photon_E	photon_isTightID	photon_etcone20
0	0	46690.223	-1.779089	-2.186474	142247.6	-1	1209.1973
0	1	29442.607	1.1680628	1.1273633	51918	-1	575.0436
1	0	60888.74	0.1741194	1.8595004	61814.07	-1	5396.2334
1	1	37795.01	-0.968964	-0.888251	56970.11	-1	2005.2532
2	0	39186.21	0.1393366	1.0165337	39567.22	-1	3261.8389
2	1	25327.361	-1.309055	-1.787857	50309.61	-1	21145.145
3	0	35529.688	1.0483403	1.5874366	56908.46	1	-23.80545
3	1	29579.598	1.2933013	-1.533586	57963.63	1	-614.49365

- 在MATLAB 中将层次数据扁平化，由0, 1标记不同的subentry

photon_pt0	photon_pt1	photon_eta0	photon_eta1	photon_isTightID0	photon_isTightID1	photon_etcone200	photon_etcone201
46690	29443	-1.7791	1.1681	-1	1209.2	575.04
60889	37795	0.17412	-0.96896	-1	5396.2	2005.3
39186	25327	0.13934	-1.3091	-1	3261.8	21145

原本一个双光子事件的物理量由一个(2,7)矩阵编码

subentry 层次扁平化之后，一个双光子事件由(1,6)向量编码

1、数据介绍与处理



(1)、数据来源、基本结构 (2)、数据筛选

- 筛选 (“CUTS”) 的目的：最大化信号（需要研究的物理事件）与背景（不相关或噪声事件）的区分度

```
condition1 = dataN.photon_isTightID0 == 1 & dataN.photon_isTightID1 == 1;
condition2 = dataN.photon_pt0 > 40000 & dataN.photon_pt1 > 30000;
condition3 = dataN.photon_etcone200 < 4000 & dataN.photon_etcone201 < 4000;
condition4 = ((abs(dataN.photon_eta0) > 1.52 | abs(dataN.photon_eta0) < 1.37) & ...
              (abs(dataN.photon_eta1)>1.52 | abs(dataN.photon_eta1)<1.37));
```

Cuts 是一个技术活，一定程度上基于反应的物理图像，
近来也有通过机器学习寻找更好cuts 的研究。

- 数十万行数据筛选后只留下 22350 行信号事件占比更大的数据

2、拟合质量谱寻找Higgs 粒子



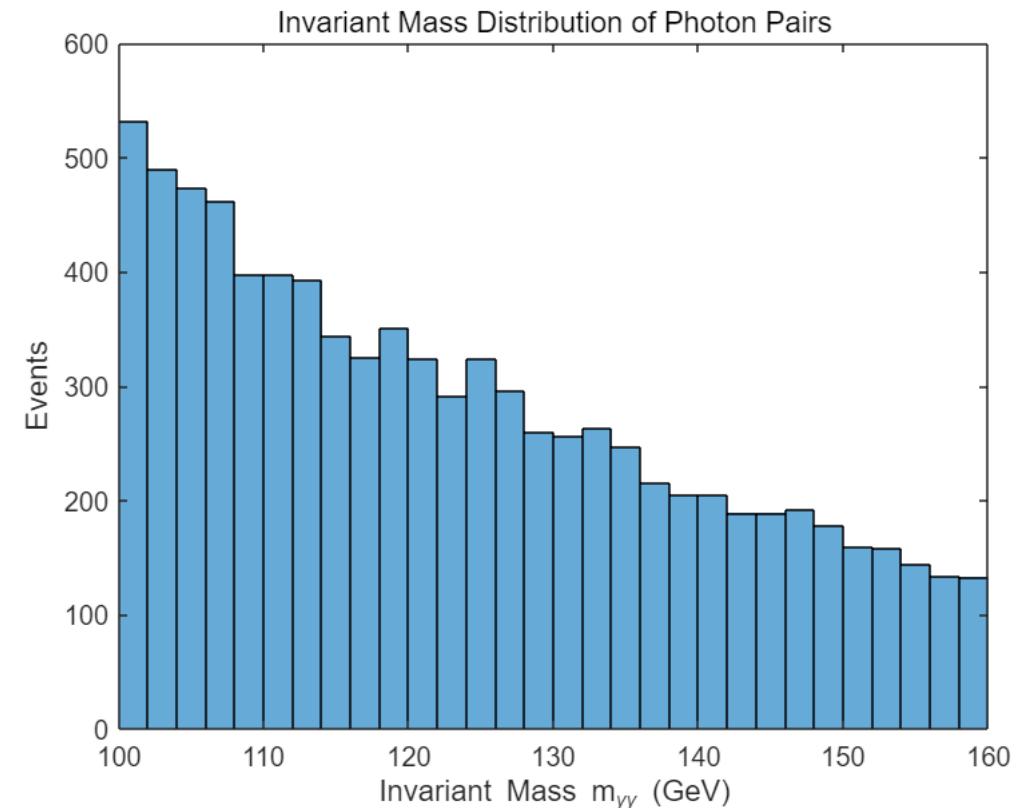
(1)、初步统计质量谱 (2)、曲线拟合寻找信号

- 通过能-动量关系重建系统不变质量:

$$m_{\gamma\gamma}^2 = \left(\sum_i P_i^\mu \right)^2 = \left(\sum_i E_i \right)^2 - \left(\sum_i \vec{p}_i \right)^2$$

- 得到 $m_{\gamma\gamma}$ 统计直方图(右图):

系统不变质量谱信号十分微弱，但是因为理论上基本可以推测标准模型Higgs 粒子质量约125 GeV，我们仍可感受到在此处存在微弱的信号峰。



重建系统不变质量

2、拟合质量谱寻找Higgs 粒子



(1)、初步统计质量谱 (2)、曲线拟合寻找信号

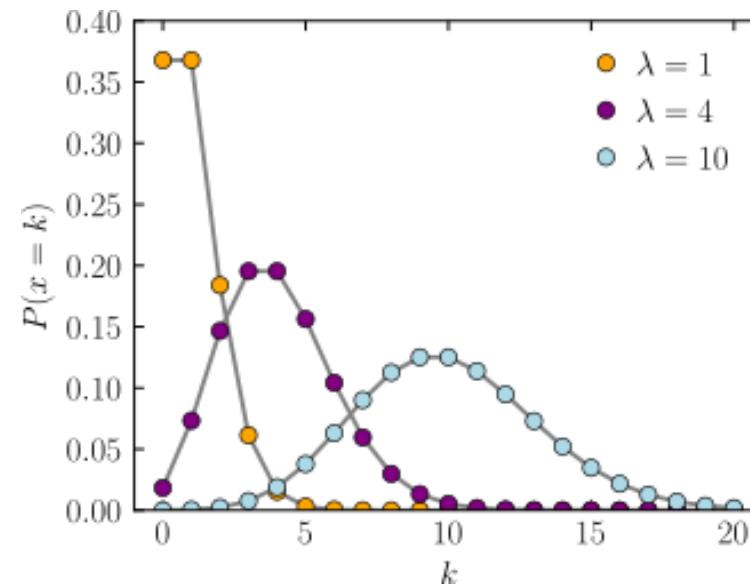
- 计算直方图每个bin 的统计误差:

- 假设每个bin 的事例数服从Poisson 分布:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

- 借助Poisson 分布数学期望和方差相等的性质，
并用每个bin 事例数估计这个bin 的数学期望

$$E(X) = V(X) = \lambda$$



Poisson 分布图像

```
data_x = h.Values;
data_x_errors = sqrt(data_x);
data_x_errors(data_x_errors == 0) = 1; % 避免除以零
```

由此得到的统计误差将
在图像中表现为误差棒

2、拟合质量谱寻找Higgs 粒子



(1)、初步统计质量谱 (2)、曲线拟合寻找信号

- 用光滑函数对直方图进行拟合:

```
% 自定义组合模型, 包括多项式和高斯函数
% x - 自变量, p0-p4 - 4次多项式系数, A - 振幅, mu - 均值, sigma - 标准差
gausspoly = fittype('p0 + p1*x + p2*x^2 + p3*x^3 + p4*x^4 + A*exp(-((x-mu)^2)/(2*sigma^2))',...
    'independent', 'x', 'coefficients', ...
    {'p0', 'p1', 'p2', 'p3', 'p4', 'A', 'mu', 'sigma'});

% 对数据进行拟合
% 给各系数提供一个初始估计值
startPoints = [1000, 0, 0, 0, 0, 90, 125.09, 2]; % 根据数据调整初始值
[coeff, gof] = fit(bin_centres', data_x', gausspoly, 'Start', startPoints, ...
    'Weights', 1./(data_x_errors'.^2));

% 输出总拟合曲线、信号曲线、本底曲线
out=feval(coeff, bin_centres);

sigfit=coeff.A*exp(-((bin_centres'-coeff.mu).^2)/(2*coeff.sigma^2));

background = coeff.p0 + coeff.p1*bin_centres' + coeff.p2*(bin_centres').^2 + ...
    coeff.p3*(bin_centres').^3 + coeff.p4*(bin_centres').^4;

% 用于绘制拟合残差
signal=data_x'-background;
```

- 正态分布函数拟合信号(1)
- 4次多项式函数拟合本底(2)
- 联合 1, 2 函数由MATLAB 中 fit 函数给出合适的拟合参数

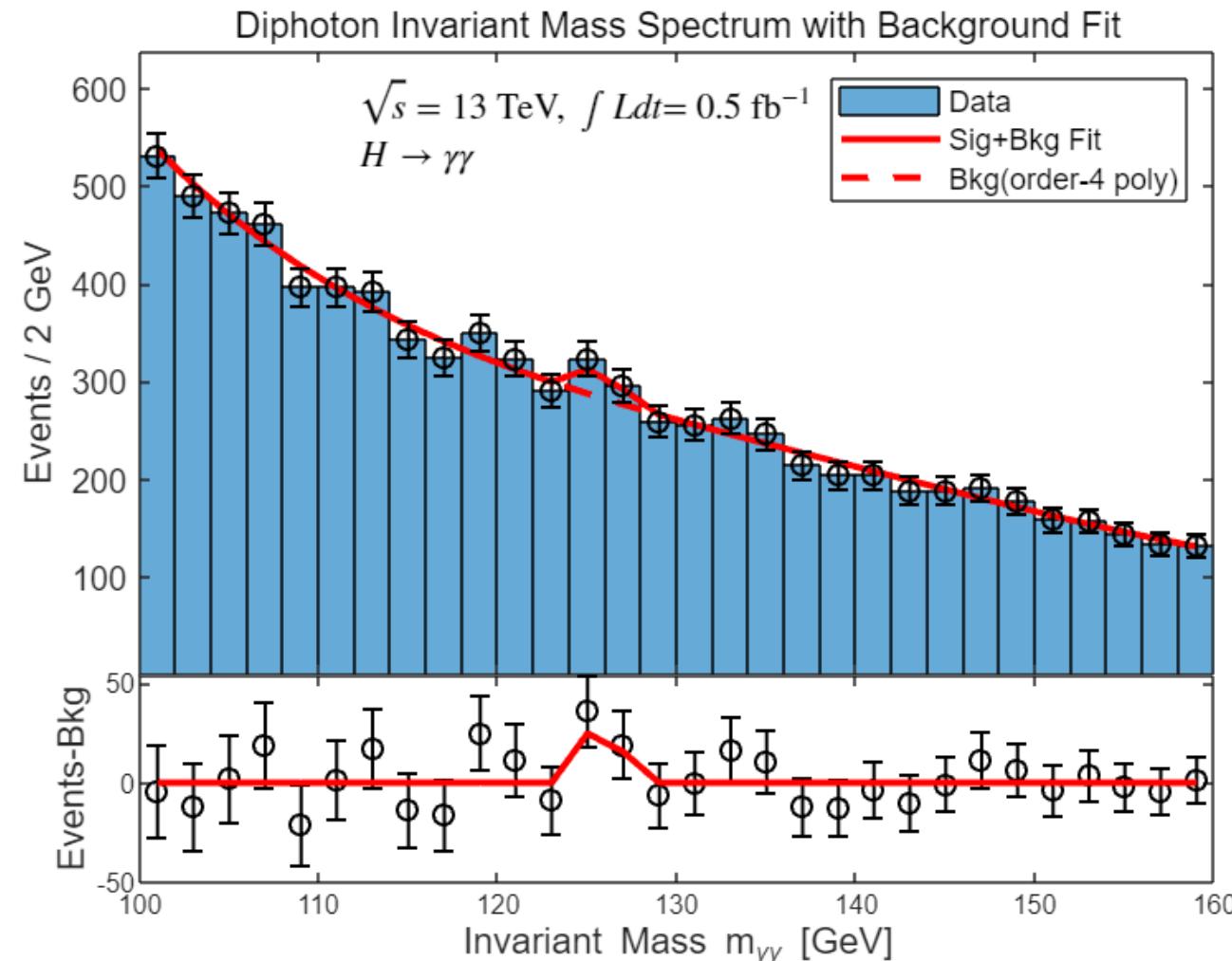
2、拟合质量谱寻找Higgs 粒子



(1)、初步统计质量谱 (2)、曲线拟合寻找信号

- 最终，得到质量谱的本底+信号拟合结果：
 - 曲线拟合帮助发现微弱信号，虚线给出拟合的本底，**125 GeV** 处的突起对应 Higgs 粒子的质量共振峰。
 - 并绘制了**残差图**检验拟合效果

恭喜你！发现了Higgs 粒子！



总结 Conclusion

在本次报告中，我们：

- 介绍了粒子物理的基本内容：

粒子的动力学性质、基本粒子标准模型

- 演示了粒子物理实验数据处理、分析技术：

物理量重建、质量谱峰值分析、实验信号拟合、统计分析

既有物理知识也有分析演示，是一次专业性和趣味性兼备的报告



南京大學
NANJING UNIVERSITY

感谢老师同学的垂听！
Thanks for Listening!

尹琪钦 (物理学 211870080)
211870080@smail.nju.edu.cn

2024年春