

不同的多个数据样本联合分析方法中 共振态显著性的差异

张敬庆

(南京师范大学物理科学与技术学院, 江苏 南京 210023)

[摘要] 粒子物理实验中数据统计量影响着物理量的测量精度和显著性的大小. 为了尽可能的扩大统计量, 减小测量中的统计误差, 提高显著性, 同时采用多个数据样本进行同一分析一直是一个有效的方法. 在采用多个不同样本时, 通常有两种处理方式. 第一种是直接将不同的数据样本累加, 然后作为一个总样本进行分析. 第二种是不进行样本累加, 而是在分析中采用联合拟合的方式同时提取出各个样本中信号的相关信息和显著性. 本文对两种方法中共振态显著性的差异进行了研究和比较, 研究结果可供使用多个数据样本开展的相关分析参考.

[关键词] 多个数据样本, 联合拟合, 显著性

[中图分类号] N56 [文献标志码] A [文章编号] 1001-4616(2022)04-0021-05

Difference of Resonance Significance in Different Approach of Analysis Using Multiple Datasets

Zhang Jingqing

(Nanjing Normal University, School of Physics and Technology, Nanjing 210023, China)

Abstract: Statistics is important in experiment of particle physics. It will affect not only statistical uncertainty but also significance. To improve sensitive and increase significance, a common way is to combine multiple datasets to enlarge the statistics in the analysis. There are two frequently used approaches. The first one is that the datasets are added together as one whole dataset, then the analysis is performed on this unique dataset. The second approach doesn't need to construct a whole dataset, instead, a simultaneous fit will be applied to all the datasets, and information of the signal will be extracted in this simultaneous fit. In this paper, the difference of resonance significance in these two different approaches are studied. The research and conclusion can be used as a reference for future analysis using multiple datasets.

Key words: multiple datasets, simultaneous fit, significance

量子色动力学 (Quantum Chromodynamics, QCD) 是粒子物理标准模型中用来描述强相互作用的理论. QCD 理论中, 强子可以分为两类, 一种是由夸克反夸克组成的介子, 自旋为整数, 另一种是由 3 个夸克 (反夸克) 组成的重子 (反重子), 自旋为半整数^[1-2]. QCD 理论的有效性已经得到了粒子物理实验方面的大量证明.

在 QCD 理论中, 除介子和重子这两类普通强子以外, 也可能存在由 3 个以上的夸克和/或反夸克组成的奇异强子, 比如由两个夸克和两个反夸克组成的四夸克态, 这些奇异强子为研究强相互作用和禁闭机制提供了独特的条件^[3]. 在实验上对这些由多个夸克反夸克组成的奇异强子进行寻找, 是对 QCD 理论有效性的重要验证, 因此也一直是粒子物理实验中的一个前沿领域.

在新粒子的寻找中, 统计显著性大于 5 倍标准差通常是实验物理学家认为观测到一个新粒子的标准. 为了尽可能提高实验观测中信号的显著性, 需要在寻找新粒子的实验中, 在现有的实验条件下, 尽量扩大数据样本的统计量.

以 $X(3872)$ 粒子的发现作为例子^[4], 在该文章中, $X(3872)$ 粒子在 B 介子衰变末态中的 $\pi^+\pi^-J/\psi$ 不变质

收稿日期: 2022-05-16.

基金项目: 国家自然科学基金面上项目 (12075123).

通讯作者: 张敬庆, 博士, 研究方向: 粒子物理实验. E-mail: jingqing.zhang@njnu.edu.cn

量谱上被观测到,具体的衰变过程是 $B^+ \rightarrow K^+ X(3872)$, $X(3872) \rightarrow \pi^+ \pi^- J/\psi$ 和电荷共轭过程 $B^- \rightarrow K^- X(3872)$, $X(3872) \rightarrow \pi^+ \pi^- J/\psi$, 在实验中 J/ψ 粒子通过电子或缪子末态重建,即 $J/\psi \rightarrow e^+ e^-$ 或 $J/\psi \rightarrow \mu^+ \mu^-$. 在该文章中,为了扩大数据样本的统计量,进而提高 $X(3872)$ 粒子的显著性,作者同时采用了 B^+ 和 B^- 的衰变事例,使得数据量比单独使用 B^+ 或者 B^- 的衰变事例提高了一倍. 另外, J/ψ 介子通过电子对或者缪子对重建,数据量又比单独使用一种末态重建时增加了约一倍,即该文章的作者在研究中联合使用了 4 个数据样本($B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi \rightarrow K^+ \pi^+ \pi^- e^+ e^-$, $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ 和 $B^- \rightarrow K^- \pi^+ \pi^- J/\psi \rightarrow K^- \pi^+ \pi^- e^+ e^-$, $B^- \rightarrow K^- \pi^+ \pi^- J/\psi \rightarrow K^- \pi^+ \pi^- \mu^+ \mu^-$). 通过联合使用这 4 个数据样本扩大统计量,该文章的作者最终在研究中确定 $X(3872)$ 的统计显著性高达 10 倍标准差,远高于粒子物理实验中的发现标准(5 倍标准差).

1 粒子物理实验中联合使用多个数据样本的两种方法

然而,我们注意到在不同的文章中,在联合使用多个数据样本时有两种不同的方法. 方法一是文献[4]中所采用的方法,将多个数据样本简单的累加到一起,作为一个样本来使用,即将 4 个数据样本的不变质量谱 $M(\pi^+ \pi^- J/\psi)$ 叠加到一起,然后对总的 $M(\pi^+ \pi^- J/\psi)$ 谱进行拟合,得到 $X(3872)$ 的质量、宽度等共振态参数,并根据似然值和拟合参数个数相对于使用纯本底拟合时的变化估计 $X(3872)$ 的显著性.

方法二是在联合使用多个数据样本时采用联合拟合(simultaneous fit)的方法对各个数据样本所含的信号信息进行提取,并根据似然值和拟合参数个数相对于使用纯本底拟合时的变化估计信号的显著性. 方法二不需要将多个数据样本叠加到一起. 文献[5]是使用方法二的一个例子. 在文献[5]中,为了对 $X(1835)$ 的性质进行研究,文章作者使用了 $J/\psi \rightarrow \gamma \pi^+ \pi^- \eta'$, $\eta' \rightarrow \pi^+ \pi^- \eta$ 和 $J/\psi \rightarrow \gamma \pi^+ \pi^- \eta'$, $\eta' \rightarrow \gamma \pi^+ \pi^-$ 两个数据样本,在研究中使用联合拟合对两个样本的 $\pi^+ \pi^- \eta'$ 不变质量谱进行同时拟合.

方法一将多个数据样本叠加到一起构成一个总样本,在分析这个总数据样本时,共振态信号质量谱的分辨率和效率随质量谱的变化(效率曲线)以及本底形状反应了各个样本叠加后的综合效应,在拟合中也只能考虑总样本的共振态信号的分辨率、效率曲线和本底形状.

方法二对多个数据样本进行联合拟合,在联合拟合中,各个数据样本共享共振态信号的参数(质量、宽度等),同时采用不同的共振态信号的分辨率、效率曲线以及本底形状.

2 用蒙特卡洛样本研究两种方法在共振态显著性估计上的差异

考虑到方法一和方法二在描述数据样本中的共振态信号和本底时的差别,我们认为这可能会对共振态信号的统计显著性产生影响,并使用蒙特卡洛方法^[6]对二者在共振态显著性上的差别进行了研究,具体研究方法和结论如下所述.

我们以 $X(3872)$ 粒子的衰变为例. $X(3872)$ 衰变到 $\pi^+ \pi^- J/\psi$, 而 J/ψ 又有多种衰变末态存在. 我们用蒙特卡洛方法产生了两种含有 $X(3872)$ 信号和本底的 $M(\pi^+ \pi^- J/\psi)$ 不变质量分布,一种代表 $J/\psi \rightarrow \mu^+ \mu^-$ 的数据样本,另一种代表 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 的数据样本. 这两种样本中 $X(3872)$ 的质量一样,但是宽度、 $X(3872)$ 事例数、本底形状和本底事例数不一样,用来表示两种数据样本中不同的分辨率、效率、本底形状和本底水平.

两种数据样本的蒙特卡洛样本中, $X(3872)$ 信号形状和本底形状分别由高斯函数(函数(1))和一阶车比雪夫多项式描述(函数(2)).

$$f(x, \mu, \sigma) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (1)$$

$$f(x, p_0, p_1) = p_0 + p_1 \cdot (x - 0.5 \cdot (3.822 + 3.922)) / (0.5 \cdot (3.922 - 3.822)). \quad (2)$$

蒙特卡洛样本中,两种样本的信号、本底参数以及信号事例数(N_s)、本底事例数(N_b)设置如表 1 所示.

在设置两种样本的信号事例数 N_s 时,我们假设 $J/\psi \rightarrow \mu^+ \mu^-$ 的分支比是 0.06^[7], 该衰变模式下 $X(3872)$ 的探测效率是 0.3, $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 的分支比是 0.02^[7], 该模式下的探测效率是 0.3.

我们使用舍选法产生了 1 000 组蒙特卡洛样本,每组样本包含一个代表 $J/\psi \rightarrow \mu^+ \mu^-$ 模式的 $M(\pi^+ \pi^- J/\psi)$ 分布和一个代表 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 模式的 $M(\pi^+ \pi^- J/\psi)$.

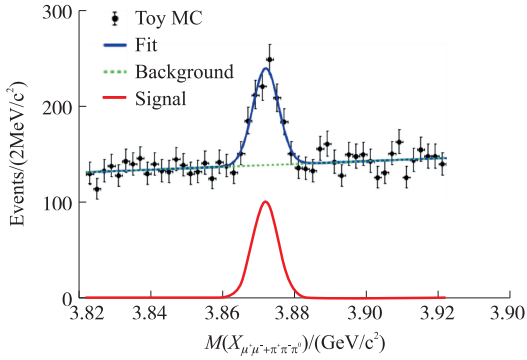
表 1 蒙特卡洛样本的产生参数

Table 1 Parameters for Monte Carlo sample generation

样本类型	N_s	N_b	μ (GeV/c ²)	σ (GeV/c ²)	p_0	p_1
$J/\psi \rightarrow \mu^+ \mu^-$	300	4 000	3.872	0.003	1.0	-0.3
$J/\psi \rightarrow \pi^+ \pi^- \pi^0$	100	3 000	3.872	0.003 2	1.0	0.5

为了比较方法一和方法二中 $X(3872)$ 显著性的差别,我们对这 1 000 组蒙特卡洛样本分别使用方法一和方法二对 $J/\psi \rightarrow \mu^+ \mu^-$ 样本和 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 进行联合分析,在分析中使用最大似然法对数据样本进行拟合,并估计 $X(3872)$ 的统计显著性. 拟合中,所有拟合参数都是自由变化的,初始值设为蒙特卡洛样本产生时的设定值.

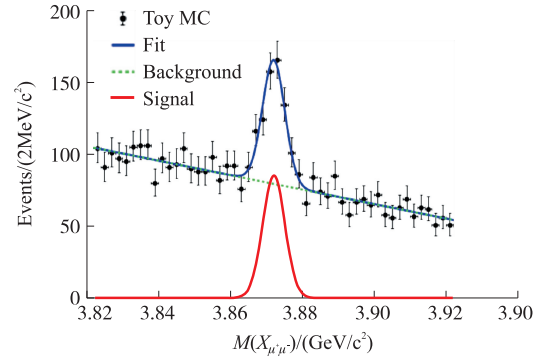
在估计 $X(3872)$ 显著性时,我们先用信号加本底的函数(高斯函数加一阶车比雪夫多项式)拟合数据样本,得到对应的似然值的对数($\ln L_{s+b}$). 图 1,图 2 和图 3 分别是采用方法一和方法二对第一组蒙特卡洛样本的拟合结果,总体拟合曲线和蒙特卡洛样本符合得很好.



图中黑色圆点是产生的蒙特卡洛样本(既包含 $J/\psi \rightarrow \mu^+ \mu^-$ 的样本,也包含 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 的样本),红色曲线是 $X(3872)$ 的信号形状,蓝色部分是总体拟合曲线,绿色虚线是本底贡献.

图 1 采用方法一时对蒙特卡洛样本的拟合结果

Fig. 1 Fit result using approach 1



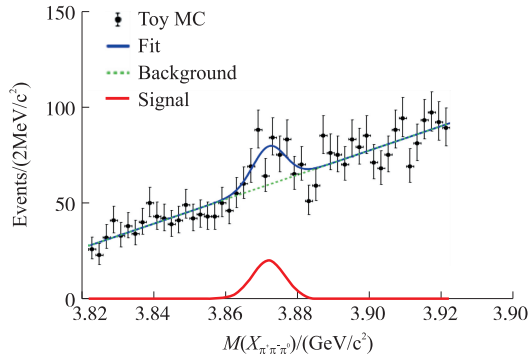
图中的黑色圆点是产生的蒙特卡洛样本($J/\psi \rightarrow \mu^+ \mu^-$),红色曲线是 $X(3872)$ 的信号形状,蓝色部分是总体拟合曲线,绿色虚线是本底贡献.

图 2 采用方法二时对蒙特卡洛样本联合拟合结果在 $M(X_{\mu^+\mu^-})$ 的投影Fig. 2 Fit projection in $M(X_{\mu^+\mu^-})$ using approach 2

图 1 展示了对于第一组的 $J/\psi \rightarrow \mu^+ \mu^-$ 和 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 样本采用方法一的拟合结果.

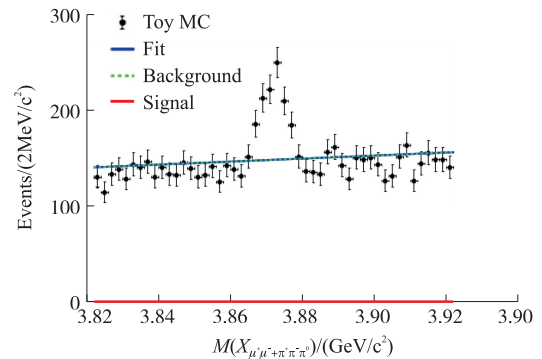
图 2 和图 3 展示了对于第一组 $J/\psi \rightarrow \mu^+ \mu^-$ 和 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 样本采用方法二的联合拟合的结果.

为估计 $X(3872)$ 的显著性,我们还要对相应的蒙特卡洛样本只用代表本底形状的一阶车比雪夫多项式进行拟合,得到对应的似然值的对数($\ln L_b$). 图 4、图 5 和图 6 分别是采用方法一和方法二对第一组蒙特卡洛样本不含信号的拟合结果.



图中的黑色圆点是产生的蒙特卡洛样本($J/\psi \rightarrow \pi^+ \pi^- \pi^0$),红色曲线是 $X(3872)$ 的信号形状,蓝色部分是总体拟合曲线,绿色虚线是本底贡献.

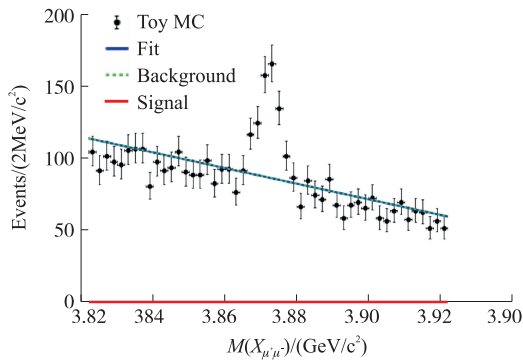
图 3 采用方法二时对蒙特卡洛样本联合拟合结果

在 $M(X_{\pi^+\pi^-\pi^0})$ 的投影Fig. 3 Fit projection in $M(X_{\pi^+\pi^-\pi^0})$ using approach 2

图中黑色圆点是产生的蒙特卡洛样本(既包含 $J/\psi \rightarrow \mu^+ \mu^-$ 的样本,也包含 $J/\psi \rightarrow \pi^+ \pi^- \pi^0$ 的样本),蓝色部分是总体拟合曲线,绿色虚线是本底贡献.

图 4 采用方法一时对蒙特卡洛样本的拟合结果(不含信号)

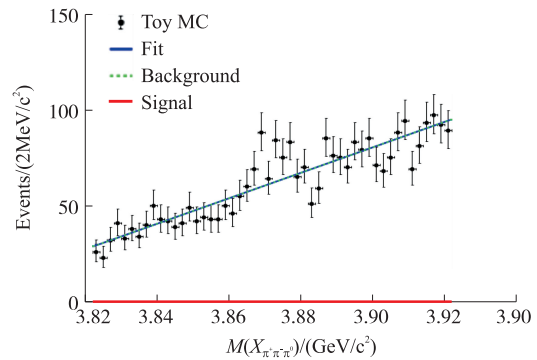
Fig. 4 Fit result using approach 1 (background only fit)



图中黑色圆点是产生的蒙特卡洛样本 ($J/\psi \rightarrow \mu^+ \mu^-$), 蓝色部分是总体拟合曲线, 绿色虚线是本底贡献。

图 5 采用方法二时对蒙特卡洛样本联合拟合结果在 $M(X_{\mu^+\mu^-})$ 的投影 (不含信号)

Fig. 5 Fit projection in $M(X_{\mu^+\mu^-})$ using approach 2 (background only fit)



图中黑色圆点是产生的蒙特卡洛样本 ($J/\psi \rightarrow \pi^+ \pi^- \pi^0$), 蓝色部分是总体拟合曲线, 绿色虚线是本底贡献。

图 6 采用方法二时对蒙特卡洛样本联合拟合结果在 $M(X_{\pi^+\pi^-\pi^0})$ 的投影 (不含信号)

Fig. 6 Fit projection in $M(X_{\pi^+\pi^-\pi^0})$ using approach 2 (background only fit)

在考虑了似然值变化 ($\Delta \ln L = \ln L_{s+b} - \ln L_b$) 以及自由度个数的变化 (Δndf , 即拟合中自由参数数目的变化) 后, 我们可以利用高能物理实验领域通用的数据分析软件 ROOT^[8-9] 提供的函数计算出 $X(3872)$ 的显著性, 显著性为 $m = \text{RooStats}::\text{PValueToSignificance}(\text{TMath}::\text{Prob}(\Delta \ln L, \Delta ndf)/2.0)$ 倍标准差, 即显著性为 $m\sigma$ 。

图 7 展示了 1 000 组蒙特卡洛样本分别采用方法一和方法二得到的 $X(3872)$ 的显著性, 以及它们之间的差别。从图中可以看出, 多数情况下, 方法二得到的显著性更高。这是因为方法二在联合拟合中能够更精细的考虑不同样本中信号的分辨率、效率曲线和本底形状, 因此对信号更灵敏。

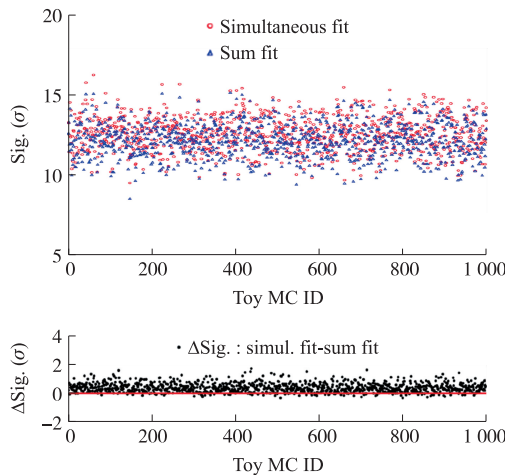


图 7 采用方法一 (Sum fit, 蓝色三角形) 和方法二 (Simultaneous fit, 红色圆圈) 得到的 $X(3872)$ 的显著性, 以及方法二的显著性与方法一的显著性的差别 (方法二显著性减方法一显著性, 黑色实心点)。

Fig. 7 Significance of the $X(3872)$ in approach 1 (Sum fit, blue triangles) and approach 2 (Simultaneous fit, red circles), and the significance difference in two approaches (significance in approach 2 minus significance in approach 1, black dots).

3 两种方法的时间成本

两种方法在分析样本数据时, 差异主要体现在对效率、分辨率的具体处理方式上, 因此所耗费的时间略有差异, 但是这点用时差异相对于粒子物理实验中一个完整的数据分析来说并不显著。

在粒子物理实验中信号的探测效率和分辨率一般用产生的蒙特卡洛样本进行估计。虽然两种方法对效率和分辨率的具体处理方式有差异, 但都需要对各个子样本产生对应的信号蒙特卡洛样本, 如果保持相同的精度, 那么两种方式产生的蒙特卡洛样本的数目和样本大小是一致的, 耗费的时间也是相同的。另一方面, 在一般的数据分析中, 蒙特卡洛样本产生所需要的时间只占整个分析中的很小一部分, 由此导致

的两种方法时间成本的差异可以忽略.

在具体处理效率和分辨率时,在第一种方法中,只需要用按比例混合后的蒙特卡洛样本提取总样本的效率和分辨率,在第二种方法中,则需要对各个蒙特卡洛样本分别提取效率和分辨率,因此所需时间也略多于第一种方法.但是,相比于一个完整的数据分析的时间尺度,两种方法在这一步上时间成本的差异并不显著.

4 结论

本文讨论了粒子物理实验中联合使用多个数据样本的两种方法在共振态显著性估计上的差异.方法一是将多个数据样本直接叠加,操作简单,但是在对共振态信号的拟合中只能包含叠加后的总样本的分辨率、效率曲线和本底形状.方法二不需要将多个数据样本直接累加,而是对多个数据样本进行联合拟合(simultaneous fit),在联合拟合中可针对不同的数据样本采用不同的分辨率、效率曲线和本底形状.方法二能够更精细地考虑不同样本中的分辨率、效率曲线和本底形状,因此对共振态信号更灵敏.我们的蒙特卡洛研究结果显示,使用方法二时得到的信号显著性要比使用方法一时得到的信号显著性更高.同时,两种方法由于对效率和分辨率采用不同的处理方式所造成的时间成本的差别相对于粒子物理实验中一个完整的数据分析来说并不显著.因此,在两种方法都可以使用的实验中,我们推荐优先使用方法二(联合拟合)对实验数据进行分析.

[参考文献]

- [1] GELL-MANN M. A schematic model of baryons and mesons[J]. Physics letters, 1964, 8: 214–215.
- [2] ZWEIG G. An SU3 model for strong interaction symmetry and its breaking[R]. Version 1, Technical Report. CERN-TH-401, CERN, Geneva; 1964.
- [3] BRAMBILLA N, EIDELMAN S, FOKA P, et al. QCD and strongly coupled gauge theories: challenges and perspectives[J]. The European physical journal C, 2004, 74: 2981.
- [4] BELLE COLLABORATION, CHOI S K. Observation of a narrow charmonium-like state in exclusive $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi$ decays[J]. Physical review letters, 2003, 91: 262001.
- [5] BELLE COLLABORATION, ABLIKIM M. Observation of an anomalous line shape of the $\eta' \pi^+ \pi^-$ mass spectrum near the $p\bar{p}$ mass threshold in $J/\psi \rightarrow \gamma \eta' \pi^+ \pi^-$ [J]. Physical review letters, 2016, 117: 042002.
- [6] 朱永生. 实验物理中的概率和统计[M]. 北京: 科学出版社, 2020.
- [7] PARTICLE DATA GROUP, ZYLA P Z. The review of particle physics[J]. Progress of theoretical and experimental physics, 2020, 2020: 083C01.
- [8] BRUN R, RADEMAKERS F. ROOT: An object oriented data analysis framework[J]. Nuclear instruments and methods A, 1997, 389: 81–86.
- [9] VERKERKE W, KIRKBY D. The RooFit toolkit for data modeling[C]//Computing in High energy and Nuclear Physics, California, USA, 2003. <https://arxiv.org/abs/physics/0306116>.

[责任编辑:顾晓天]