

第9章 关于粒子表

对于粒子物理研究者,《粒子物理综论》(*Review of Particle Physics*)及其缩略本《粒子物理手册》(*Particle Physics Booklet*)^[1]几乎是须臾不能离开的工具书. 它们的纸质本每两年更新一次,电子版则每年更新一次. 它是由百名以上来自世界各个著名研究所和大学的、高水准的粒子物理学家组成的**粒子数据组**(Particle Data Group, PDG), 收集在相关学术刊物上正式发表的粒子物理数据后编撰的出版物.

《粒子物理综论》的主体部分是描述粒子性质的**汇总表** (summary table) 和**粒子表** (particle listings); 《粒子物理手册》的主体则是汇总表. 汇总表的内容包括已经确认的粒子性质的当前最优值, 假设的 (尚未证实的) 粒子的搜寻上限值的汇总, 以及守恒定律实验检验的汇总. 粒子表的内容包含推定汇总表给出的当前最优值所依据的全部数据, 并紧跟着给出近期或足够重要的其他测量的数据, 它 (们) 只是由于某种理由没有用来作为最优值的计算; 粒子表还给出了尚未证实的粒子的信息和搜寻新粒子的信息, 以及关于特别有意义或有争议的课题的简短综述.

《粒子物理综论》(2014 年版) 是一本超过 1500 页、质量逾 3kg 的鸿篇巨著, 而《粒子物理手册》则是尺寸 9cm×15cm×1cm 的袖珍小册子, 便于粒子物理学家随身携带, 甚至在旅途中查阅, 所以汇总表的数据更为常用. 本章的内容根据《粒子物理综论》(2014 年版) 引言的部分内容编写, 目的是帮助读者了解汇总表给出的粒子性质的数据是怎样根据粒子表列出的所有测量值加以确定的. 了解 PDG 在处理这些数据中采用的原则和方法, 对于理解汇总表给出的数据的真实含义十分重要.

9.1 数据的选择和处理

粒子表包含了我们已经知道的、发表在刊物上的关于粒子性质的全部相关数据. 20 种与粒子物理相关的重要刊物在截止日期前的相关数据被包含在内, 作者寄给 PDG 的即将发表的论文的相关数据也被包含在内. 除极少数例外, 没有列入预印本或会议报告的结果, 也不列入只具有历史重要性的数据.

在粒子表中, 对用来计算或估计汇总表中给出的数值的测量, 与不用来计算这些数值的测量, 进行了明确的区分. 一个测量被排除在外的可能原因如下:

- (1) 它已被后来的测量取代或包含于后来的测量之中.
- (2) 没给出误差.
- (3) 包含有疑问的假设.

- (4) 信噪比差, 统计显著性低, 或者数据质量比其他已有的数据质量差.
- (5) 与其他更可靠的多个结果明显不一致.
- (6) 与别的结果不相独立.
- (7) 不是最优的限值 (limit)(见后). 限值指上限或下限.
- (8) 引自预印本或会议报告.

某些情形下, 测量不是完全可靠的, 因此测量值不用来进行求平均的计算. 例如, 许多重子共振态的质量得自于分波分析, 粒子表引用的不是带误差的均值, 而是一个估计的范围, 在此范围内可能包含质量的真值. 这一点在粒子表的重子部分加以讨论.

对于上限, 在汇总表中通常引用最强的限值. 对上限不求平均或合并值, 除非在极少数情形下它们被重新表示为高斯型误差的测量值.

习惯上假定, 粒子和反粒子有相同的自旋、质量和平均寿命. 汇总表后面的守恒律的检验列出了 CPT 和其他守恒律的实验检验.

粒子表中 PDG 利用以下标识来说明, 怎样从粒子表所列的测量值来求得汇总表中所给出的值:

- (1) OUR AVERAGE——由选定的数据的加权平均求得.
- (2) OUR FIT——由选定的数据的约束拟合或超定的多参数拟合求得.
- (3) OUR EVALUATION——不由直接测量值求得, 而由相关的物理量的测量值求得.
- (4) OUR ESTIMATE——根据数据的观测值范围求得, 不由正规的统计方法求得.
- (5) OUR LIMIT——不是根据直接测量获得的限值, 而是 PDG 根据测量得到的比值或别的数据计算得到的限值.

当一个实验者在一个实验中观测到一种粒子的存在迹象时, 当然希望知道过去在这个区域中曾经看到过些什么. 因此 PDG 在粒子表中列出了认为具有足够统计意义, 并且没有被更可靠的数据证明是错误的所有的报道过的粒子态. 但是在汇总表中只推荐 PDG 认为是确认得很好的态. 当然这种判断带有某种主观性, 无法给出精确的判据. 详细的讨论参见粒子表中的短评 (minireviews).

9.2 均值与拟合

关于获得均值和误差的讨论分为三节: ① 误差处理; ② 无约束求平均; ③ 约束拟合.

9.2.1 误差处理

下面, 误差 δx 表示 $x \pm \delta x$ 这一范围是关于中心值 x 的 68.3% 置信区间. PDG 将它处理为高斯误差. 因此当 (实际) 误差确实是高斯误差时, δx 就是通常的一倍标准差 (1σ). 现在, 许多实验者分别给出统计误差和系统误差, 这种情形下, PDG 通常引用这两种误差, 统计误差在前, 系统误差在后. 为了求平均和进行拟合, 将两者平方相加开根作为合并后的总误差 δx .

当实验者给出测量值 x 的不对称误差 $(\delta x)^+$ 和 $(\delta x)^-$, 且与其他测量结果进行求平均或拟合运算时, 测量值 x 的误差处理为 x 、 $(\delta x)^+$ 和 $(\delta x)^-$ 这三个量的一个连续函数. 当所得的均值或拟合值小于 $x - (\delta x)^-$ 时, 利用 $(\delta x)^-$; 当所得的均值或拟合值大于 $x + (\delta x)^+$ 时, 利用 $(\delta x)^+$. 当所得的均值或拟合值在 $x - (\delta x)^-$ 和 $x + (\delta x)^+$ 之间时, 利用的误差值是 x 的线性函数. 由于 PDG 利用的误差依赖于求均值或拟合值的结果, 所以需要进行迭代运算来获得最终结果. 不对称的输出误差根据输入量和输出量间存在线性关系由输入误差所确定.

在求平均或拟合运算中, 通常不考虑不同测量之间的关联, 而是试图以减小关联的方式去选择数据. 但当若干个结果以 $A_i \pm \sigma_i \pm \Delta$ 的形式给出 (其中 Δ 是公共的系统误差) 时, 则关联误差是以确定的方式进行处理. 这种情形下, 首先对 $A_i \pm \sigma_i$ 求平均, 然后所得的统计误差与 Δ 求总误差. 不过, 如果对 $A_i \pm (\sigma_i^2 \pm \Delta_i^2)^{1/2}$ 求平均可得到相同的结果, 这里 $\Delta_i = \sigma_i \Delta \left| \sum (1/\sigma_j^2) \right|^{1/2}$. 利用后一方法的优点是, 当用了修改过的系统误差 Δ_i 时, 各个测量可视为相互独立, 可利用常规的方法与别的数据求平均. 因此, PDG 采用这种方法. 在求平均之前, 将 Δ 值列表, 借助于自动化的计算程序来计算 Δ_i , 并给出注解, 说明存在公共的系统误差.

出现关联误差的另一常见的情形是, 实验者测量两个量, 然后引用这两个量以及两者之差, 如 m_1, m_2 , 以及 $\Delta = m_2 - m_1$. 不能将这三个量都引入约束拟合, 因为三者不独立. 某些情形下, 略去误差最大的那个量, 而将另两个量引入约束拟合是一种好的近似方法. 但是在某些情形下, m_1, m_2 和 Δ 的误差相当, 这三个值没有一个值可以忽略. 这种情形下, 将三个值都输入拟合过程中, 在进行拟合运算之前, 借助于自动化的计算程序来增大误差, 使得这三个量在约束拟合中可以作为独立测量来处理. 这种情形下, PDG 给出注解, 说明采用了这种处理方法.

9.2.2 无约束求平均

为了对数据求平均, PDG 利用标准的最小二乘法. 假定某个物理量的若干个测量是不相关联的, 加权平均及误差可用下式计算:

$$\bar{x} \pm \delta \bar{x} = \frac{\sum_i w_i x_i}{\sum_i w_i} \pm \left(\sum_i w_i \right)^{-1/2}, \quad (9.2.1)$$

式中

$$w_i = 1/(\delta x_i)^2,$$

其中, x_i 和 δx_i 是实验 i 报道的值和误差, 求和遍及 N 个实验. 然后计算 $\chi^2 = \sum w_i (\bar{x} - x_i)^2$ 并与 χ^2 的期望值 $N - 1$ (如果测量值服从高斯分布) 进行比较.

如果 $\chi^2/(N - 1) \leq 1$, 并且数据没有问题, 则接受所得的结果.

如果 $\chi^2/(N - 1)$ 非常大, 可以选择不求平均值. 替代的方法是: 引述计算得到的均值, 但给出对于误差的有根据的推测值, 即一个保守估计用来考虑有问题的数据导致的效应.

最后, 如果 $\chi^2/(N - 1) > 1$ 但大得不多, 仍可以对数据求平均, 但需要进行以下的处理.

(1) 将式 (9.2.1) 中的误差 $\delta \bar{x}$ 增大到原来的 S 倍, S 称为尺度因子, 定义为

$$S = [\chi^2/(N - 1)]^{1/2}. \quad (9.2.2)$$

这样做的理由是: 大的 χ^2 值看来可能是由于至少在一个实验中误差被过低估计. 在不知道哪个实验误差估计过低的情形下, 我们假定所有实验都过低估计误差 S 倍. 如果所有的输入误差增大到原来的 S 倍, 则 χ^2 变成 $N - 1$, 因此输出误差 $\delta \bar{x}$ 增大到原来的 S 倍 [130].

当对于误差相差很大的多个实验数据进行合并时, 上述做法需要稍加修改. PDG 只用误差较小的几个实验数据计算 S 值. δx_i 截断值 (上限) 是带有任意性地选择为

$$\delta_0 = 3N^{1/2}\delta \bar{x},$$

其中, $\delta \bar{x}$ 是所有实验的均值的误差 (没有乘尺度因子 S). 这样做的理由是, 尽管低精度实验对于值 \bar{x} 和 $\delta \bar{x}$ 作用较小, 但对 χ^2 却可能有显著的贡献, 使得高精度实验的贡献被掩盖. 应当指出, 如果每个实验误差 δx_i 都相等, 则 $\delta \bar{x}$ 等于 $\delta x_i/N^{1/2}$, 那么每个 δx_i 都远低于截断值. (但更为常见的是, PDG 只是简单地排除离均值和拟合值较远的相对误差大的测量值, 即新的精确数据“赶走”老的、不精确数据.)

尺度因子方法具有以下性质: 如果两个误差相近的值之差远大于它们的误差 (不论是否有别的低精度测量值), 则考虑尺度因子后的误差 $\delta \bar{x}$ 近似等于该差值的 $1/2$.

需要强调, 误差的尺度因子方法对于中心值没有影响. 如果希望回复到非尺度因子误差 $\delta \bar{x}$, 只需要将引用的误差除以 S 就可以了.

(2) 如果误差小于 δ_0 的实验个数 M 不小于 3, 并且 $\chi^2/(M - 1) > 1.25$, 则 PDG 在粒子表中给出数据的表意图 (ideogram of the data). 图 9.1 是一个具体例子. 有时, 一个或两个数据点位于多数数据点集中的区域之外较远的地方; 有时, 数据点

劈分为两组或多组. PDG 并不从这些表意图中提取任何数值; 表意图仅仅是为了对相关的数数据有直观的整体性了解, 读者可以用他们认为恰当的方式利用它们.

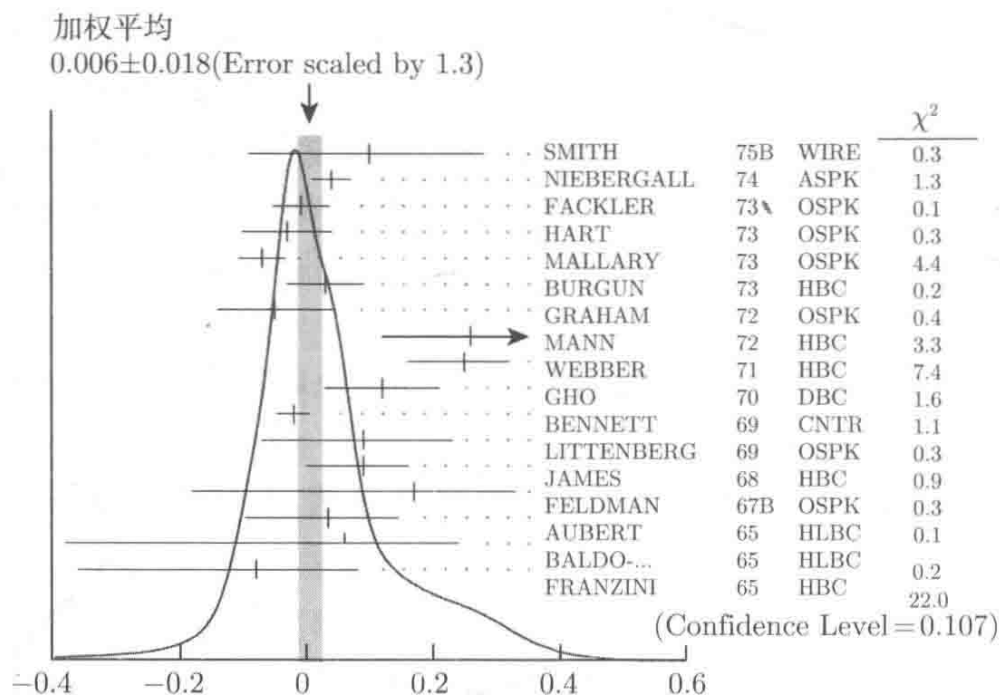


图 9.1 一个典型的表意图

顶端的箭头指示加权平均的位置, 阴影区的宽度表示考虑尺度因子 S 之后的均值的误差. 右端的一系列数字给出每个实验对 χ^2 的贡献. 注意倒数第二个实验的误差杆没有画全, 在计算 S 时没有利用该实验的数据 (参见正文)

表意图中的每个测量值表示为中心值 x_i 、误差 δx_i 、面积正比于 $1/\delta x_i$ 的高斯函数. $1/\delta x_i$ 的选择带有任意性. 这样的选择下表意图的重心对应于各测量值权重 $1/\delta x_i$ 的求平均, 而不是实际使用的各测量值权重 $(1/\delta x_i)^2$ 求平均. 这种做法对于某些实验严重低估系统误差的情形可能是适当的. 但是, 这种选择方法中, 每个测量的高斯函数的高度正比于 $(1/\delta x_i)^2$, 表意图的峰值位置将通常更多地偏向于高精度的测量值. 关于利用表意图的详细讨论参见 1986 年版《粒子物理综论》^[131].

9.2.3 约束拟合

某些情形下, 如粒子的衰变分支比、粒子质量和粒子间的质量差, 需要用约束拟合来获得一组参数的最优值. 例如, 大多数的分支比和事例率的测量是通过对全部数据的最小二乘拟合来进行分析的, 并提取衰变分支比 P_i 、分宽度 Γ_i 、全宽度 Γ (或平均寿命) 及其相关的误差矩阵.

作为一个例子, 假定一个态有 m 种衰变末态, 末态 i 的衰变分支比为 P_i , 且有 $\sum P_i = 1$. 实际测量的是 N_r 个不同的比值 R_r , 如 $R_1 = P_1/P_2$, $R_2 = P_1/P_3$ 等. (我们可以处理任何形式为 $\sum \alpha_i P_i / \sum \beta_i P_i$ 的比值 R , 这里 α_i 和 β_i 是常数, 通常等于 1 或 0. R 的形式也可以是 $R = P_i P_j$ 和 $R = (P_i P_j)^{1/2}$.) 进一步假定每一个比值

R 由 N_k 个实验进行了测量 (第 k 个实验的测量值用下标 k 标记, 如 R_{1k}). 然后通过 $m-1$ 个独立参数的函数 χ^2 的极小化求得 P_i 的最优值:

$$\chi^2 = \sum_{r=1}^{N_r} \sum_{k=1}^{N_k} \left(\frac{R_{rk} - R_r}{\delta R_{rk}} \right)^2, \tag{9.2.3}$$

其中, R_{rk} 是测量值; R_r 是拟合值.

除了拟合值 \bar{P}_i 之外, 还需计算误差矩阵 $\langle \delta \bar{P}_i \delta \bar{P}_j \rangle$. 列出对角元素 $\delta \bar{P}_i = \langle \delta \bar{P}_i \delta \bar{P}_i \rangle^{1/2}$ 的表 (某些误差是乘了尺度因子的情形除外, 这一点将在下面讨论). 在粒子表中, PDG 给出完整的关联矩阵, 同时计算每一比值的拟合值以与输入数据对比; 并且拟合值列在相关输入值的上方, 同时给出这些输入量的简单无约束平均.

对于上面的例子有 3 条评述意见:

(1) 假定全宽度和分支比的测量值之间不存在任何关联性. 但通常我们会有关于分宽度 Γ_i 和全宽度 Γ 的信息. 这种情形下 PDG 将 Γ 与 P_i 都作为拟合参数, 并在粒子表中给出各分宽度的关联矩阵.

(2) PDG 试图选取相互独立且尽可能与原始数据接近的那些比值和宽度值. 当一个实验测量了所有的分支比并将它们之和约束为 1 时, PDG 将其中的 (通常是确定性最差的) 一个分支比不作为拟合参数, 使得输入数据更接近于相互独立. 然后再容许输入数据之间存在关联.

(3) 当任一 R 的测量值给出高于预期的 χ^2 贡献时, PDG 同时对 R_r 和 P_i 计算尺度因子. 根据式 (9.2.3), 对于 χ^2 的二重求和的第一重求和是从实验 $k=1$ 到实验 $k=N_k$, 然后是求和 $\chi^2 = \sum \chi_r^2$. 看起来比值 r 的尺度因子为 $\chi_r^2 / \langle \chi_r^2 \rangle$, 但由于 $\langle \chi_r^2 \rangle$ 并不是一个确定值 (位于 N_k 和 N_{k-1} 之间), 我们不知道如何给出它的表式达. 作为替代, 定义

$$S_r^2 = \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{(R_{rk} - \bar{R}_r)^2}{\langle (R_{rk} - \bar{R}_r)^2 \rangle}, \tag{9.2.4}$$

利用该定义, S_r^2 的期望值为 1. 可以证明这时有

$$\langle (R_{rk} - \bar{R}_r)^2 \rangle = \langle (\delta R_{rk})^2 \rangle - (\delta \bar{R}_r)^2, \tag{9.2.5}$$

其中, $\delta \bar{R}_r$ 是比值 r 的拟合误差.

利用乘以尺度因子 S_r (若 $S_r < 1$ 则乘以 1) 之后的分支比误差重新进行拟合, 通常会得到新的较大的误差 $\delta \bar{P}'_i$. 这种情形下 PDG 最后给出的尺度因子定义为 $S_i = \delta \bar{P}'_i / \delta \bar{P}_i$. 但是, 按照 PDG 不让 S 影响中心值的原则, PDG 给出由原拟合 (即误差不乘以尺度因子) 得到的 \bar{P}_i 值.

存在一种特殊的情形, 前面所述的方法求得的误差可能需要改变. 若一个拟合的分支比 (或事例率) \bar{P}_i 与 0 的差别小于 3 倍标准偏差 ($\delta \bar{P}'_i$), 那么需要计算一个新

的、较小的低端误差 $(\delta\bar{P}_i'')^-$, 它使得 $\bar{P}_i - (\delta\bar{P}_i'')^-$ 和 \bar{P}_i 之间高斯函数下的面积为 0 到 \bar{P}_i 之间面积的 68.3%. 对于分支比小于 3 倍标准偏差的情形需要作同样的修正. 这种方法使得引用的误差不至于超出物理边界.

9.3 舍 入

粒子表中列出的结果通常是科学刊物中报道的实验测量数值, 而在汇总表中列出的数值 (均值和限值) 是经过一组舍入规则后的数值.

基本规则可陈述如下: 如果误差值 3 个最高位的数值在 100~354, 只舍入为 2 位有效数字; 如果在 355~949, 只舍入为 1 位有效数字; 如果在 950~999, 则舍入为 1000 且保留两位有效数字. 所有这些情形中给出的中心值的精度与误差的有效数字位匹配. 例如, 求均值的结果为 0.827 ± 0.119 , 应该显示为 0.83 ± 0.12 ; 而 0.827 ± 0.367 应该显示为 0.8 ± 0.4 .

如果汇总表中的结果来源于单次测量而非求平均, 则无需作舍入处理. 这种情形下保留原始论文的发表数据的位数, 除非认为数据不恰当. 需要指出, 即便对于单次测量, 当合并统计和系统误差时, 舍入规则也适用于合并后的总误差. 还需指出, 汇总表中的大多数限值来自单个实验 (最好的限值), 因此不进行舍入.

最后应当指出, 在若干情形下, 当一组结果出自一组数据的单次拟合时, 所有结果都选择保留 2 位有效数字. 例如, W 和 Z 玻色子以及 τ 轻子的某些性质就是这种情形.

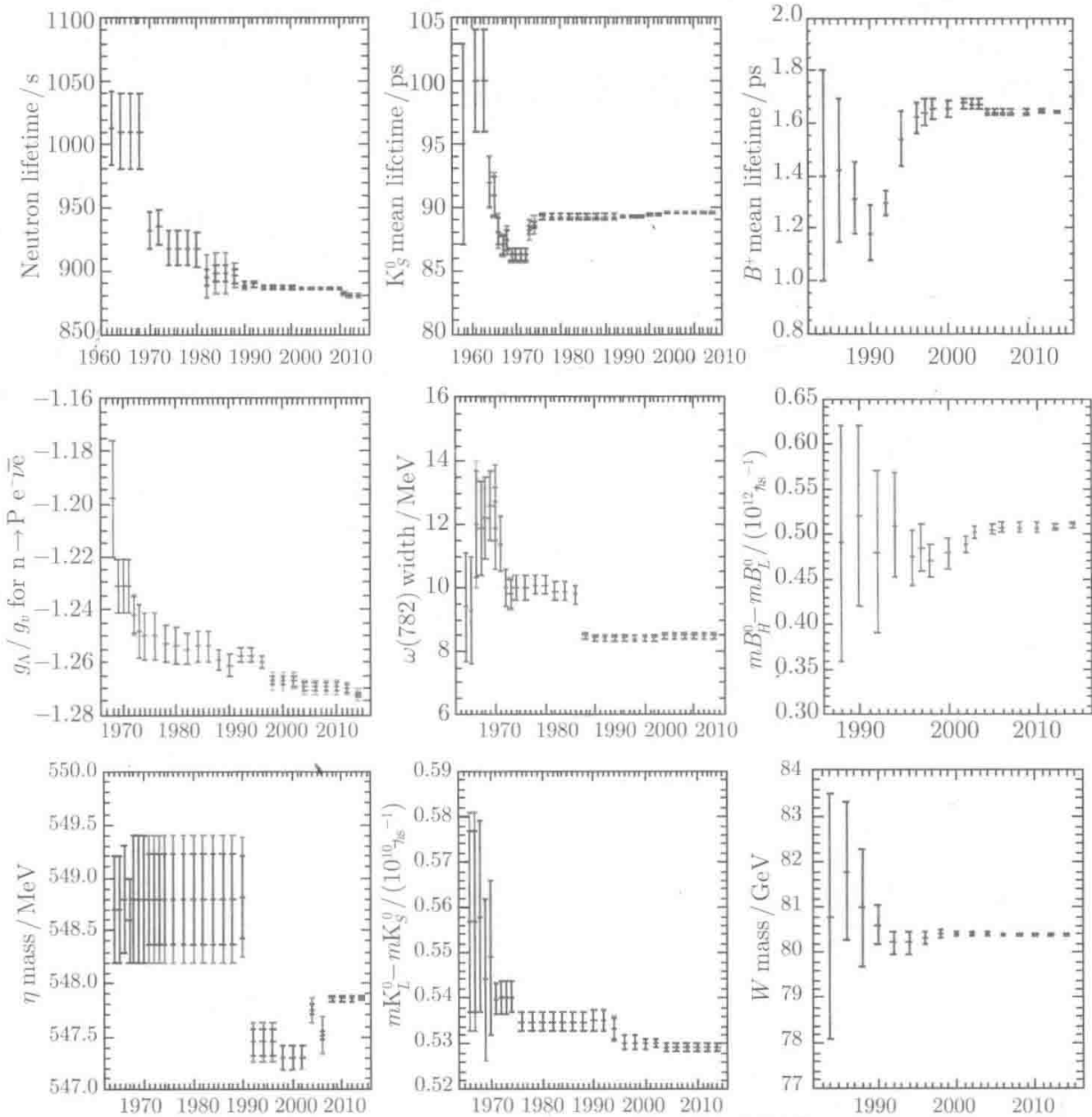
9.4 讨 论

Taylor 在文献 [132] 中讨论了包含相互矛盾的结果的数据求平均的问题. 他考虑了若干种算法来尝试将不一致的数据求得合理的平均值. 但是很难研发出一种方法, 能够以一种合理的方式同时来处理两类基本的情况: ① 远离数据主体的数据是不正确的并且没有给出误差; ② 与上述情况相反, 即数据主体是不正确的. 令人遗憾的是, Taylor 发现情况② 并不少见. 他的结论是: 比较而言, 数据的留用或丢弃的准则的选择比较重要, 而求平均方法的选择则不那么重要.

因此 PDG 更强调数据的选择. 通常 PDG 会征求外部专家 (顾问) 的帮助, 但有时候不可能确定一组矛盾的数据中哪一个是正确的. 尺度因子方法是通过增大误差来处理这种“无知”的一种尝试. 事实上这种情况相当于告知大家, 当前的实验结果无法对该物理量作出精确的测量, 因为存在不可解决的矛盾, 故需要等待进一步的测量值. 读者可以从尺度因子的大小获知已有数据相互矛盾的程度. 如果需要, 可以通过汇总表中的原文献的数据, 利用数据的不同选择重新求平均.

PDG 中遇到的情况比 Taylor 考虑的大多数情况 (基本常数如 \hbar 的估计) 要好一些. Taylor 考虑的情况中大多数是系统误差起主导作用. 对于 PDG 数据, 通常统计误差至少与系统误差相当, 而统计误差一般是容易估计的. 值得注意的一个例外出现于分波分析中, 在那里不同的方法应用于同样的数据会产生不同的结果. 这种情形下, 如前所述, PDG 通常不用该结果求平均, 而只是引述数值的范围.

文献 [1] 给出了早先的 PDG 平均值的简要的历史. 图 9.2 显示了若干个粒子性质的数值的历史演变, 有的发生了很大的变化. 这通常反映了下述现象: 增添了重要的新数据或丢弃了比较老的数据. 当新数据系统误差较小, 或对系统误差有比较多的检验, 或进行了修正 (这种修正在老实验进行时尚不知道), 或者新数据的误差比老数据小得多时, 则倾向于丢弃老数据. 有时, 数据出现大的跳跃的地方尺度因子变得很大, 这反映了新数据的加入或不相一致的数据的加入导致的不确定



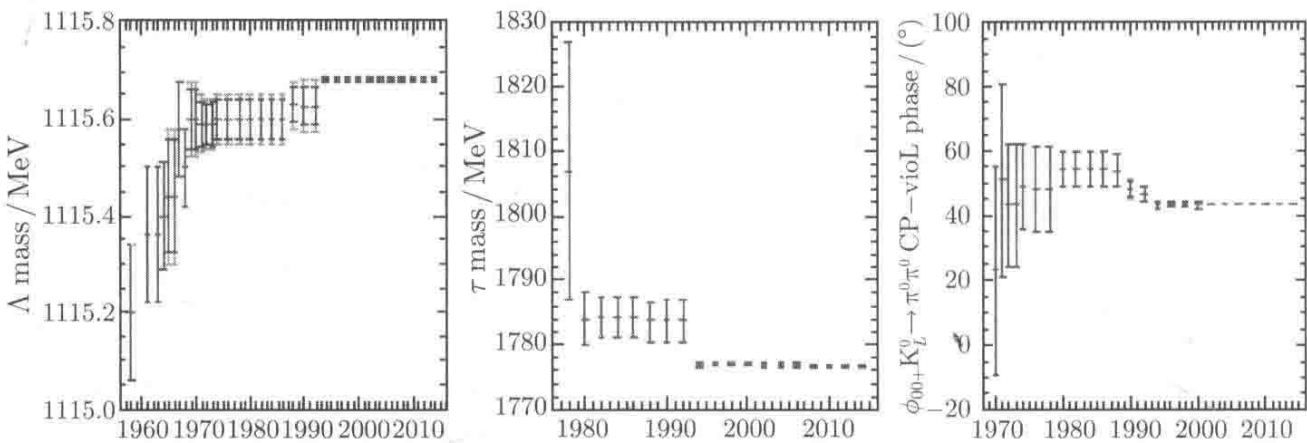


图 9.2 粒子表中一些粒子性质参数数值的年代演变
横坐标是粒子表的出版年份

性. 但大体说来, 数据的历史标绘的全景图显示了一种向更高精度的中心值缓慢逼近的过程, 而中心值与第一批数据点是相当一致的.

可以得出结论, 实验数据的合并和 PDG 求平均方法的可靠性一般是好的, 但需要提醒, 涨落超出所引述误差的现象可能并且确实存在.