

ПРОВЕРКА ПАКЕТА «PROFESSOR2» ДЛЯ НАСТРОЙКИ ПАРАМЕТРОВ ПРОЦЕССА ФРАГМЕНТАЦИИ В РЕАКЦИИ e^+e^- АННИГИЛЯЦИИ

В. А. МУРАДЯН*, Н. З. АКОПОВ, А. М. ГУМАРЯН, Г. А. КАРЯН

Национальная научная лаборатория им. А.И. Алиханяна, Ереван, Армения

*e-mail: v.muradyan@yerphi.am

(Поступила в редакцию 3 марта 2021 г.)

Описана схема верификации пакета «Professor2» для настройки параметров одного из наиболее используемых физических генераторов PYTHIA8.2 для симулирования процессов в физике высоких энергий. Цель состоит в настройке списка параметров PYTHIA8.2 для процесса фрагментации ниже порога рождения $b\bar{b}$. В качестве исходного списка параметров мы используем настройки параметров для фрагментации в Belle (предшественник Belle II) по умолчанию. Для псевдо-данных Монте Карло (МК) используется программный пакет Belle II, то есть basf2 и программа генерация, разработанная основным автором этой статьи на языке Python. Чтобы проверить всю процедуру, используется импульсное распределение заряженных пионов и каонов для настройки двух свободных параметров модели фрагментации в Pythia8, то есть модели фрагментации LUND. Показано, что пакет «Professor2» эффективен для одновременного выполнения мультипараметрической настройки.

1. Введение

Настройка МК генераторов является важнейшей проблемой для современной физики высоких энергий [1,2]. Эксперимент Belle II на ускорителе КЕК (Япония) [3] планирует набрать $50ab^{-1}$ данных для исследований физики выходящей за рамки Стандартной Модели [4]. Подобные исследования требуют не только наличия большой статистики, но также точного описания данных МК генераторами, для адекватного вычитания фоновых событий. В этом контексте описание континуум фона играет важнейшую роль для изучения редких распадов В мезонов.

До создания пакета «Professor2» [5] в традиционных методах ручной (также грубой) настройки всегда были проблемы с поиском свободных параметров в моделях, особенно с увеличением числа параметров модели. Метод реализованный в пакете «Professor2» называется настройкой на основе параметризации.

Подход, реализованный в пакете «Professor2» – это параметризация поведения генератора с помощью полиномов различной степени. Затем полином фитируется по откликам генератора в каждом бине по наблюдаемой переменной из массива данных за счет изменений компонент P -мерного вектора параметров $p = (p_1, \dots, p_P)$. Далее метод использует критерий согласия для фитированной функции и минимизирует его. В итоге получается набор наилучших параметров $p_{(\text{tune})}$, описывающий экспериментальные данные. Одним из основных преимуществ пакета «Professor2» является его способность обрабатывать взаимные корреляции между различными параметрами, выполняя многомерную минимизацию в пространстве параметров RUTHIA [6] (см. Рис.1), что не относится к вышеупомянутым методам «ручной» и «грубой» настройки.

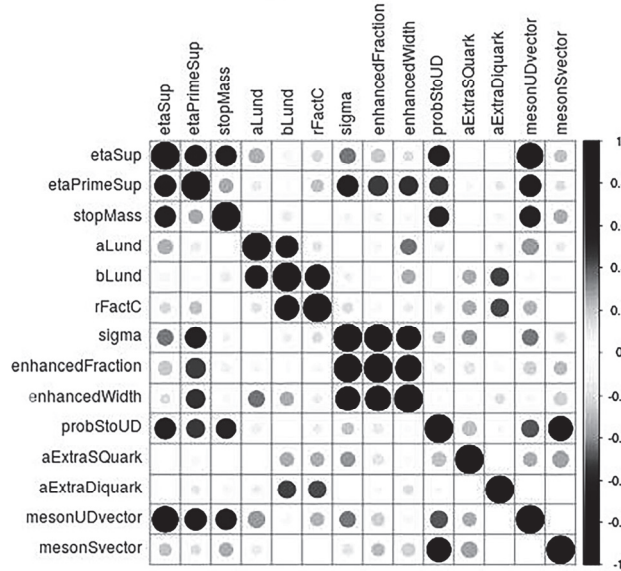


Рис.1. Корреляция между некоторыми параметрами Pythia8.

Для реализации этого метода необходим достаточно мощный компьютер. При этом набор необходимой статистики для большого числа параметров может потребовать нескольких дней, и всего нескольких минут, чтобы найти лучший параметр. Затем следует запустить генератор с набором наилучших параметров и сравнить результаты симуляции с экспериментальными данными. В следующих разделах мы более подробно опишем, как использовать пакет «Professor2».

2. Описание метода «Professor2»

Основная цель метода – определить функцию соответствия между сгенерированными и экспериментальными данными, а затем минимизировать ее. Далее описываются шаги в целом и реализация метода для процесса $e^+ + e^-$ аннигиляции в кварки, где затем каждый из кварков с некоторыми вероятностями превращается в адроны (процесс адронизации или фрагментации).

2.1. Генератор отклика

Как уже упоминалось, имеется генератор, который надо запускать для многих наборов параметров. Для выбора параметров настройки вначале производится проверка чувствительности параметров. Для этого параметр сдвигается от его номинального (по умолчанию) значения и рассматривается влияние сдвига на распределения образованных частиц по импульсам.

В качестве примера проверки чувствительности на Рис.2 и Рис.3 продемонстрирован двух-образцовый статистический тест, выполненный для параметров a (StringZ:aLund) и b (StringZ:bLund), показывающий статистически значимое (Рис.2), не значимое (Рис.3) влияние на спектры различных адронов в конечном состоянии.

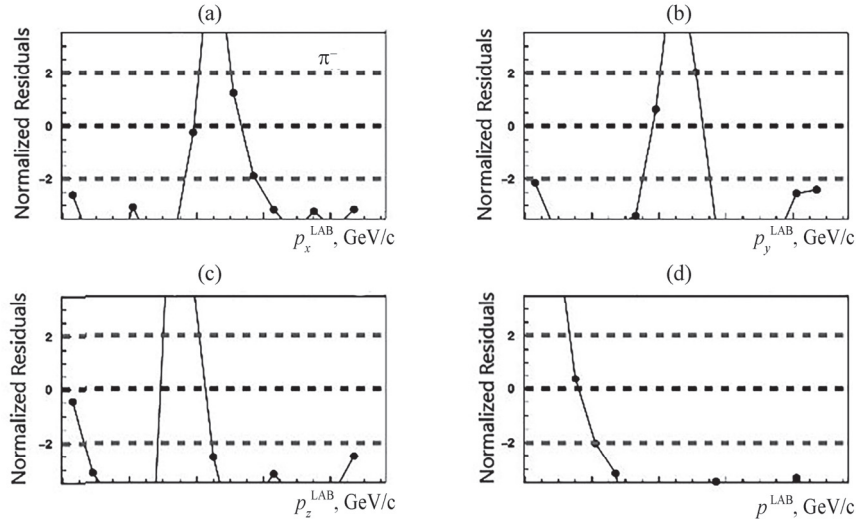


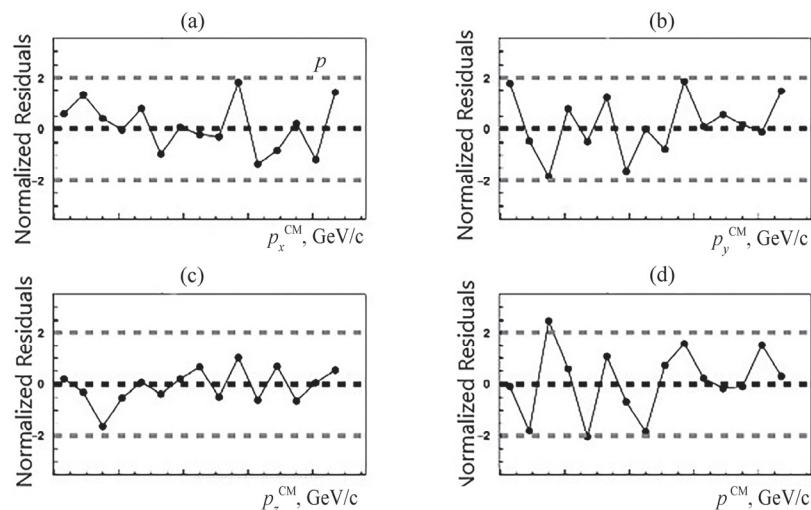
Рис.2. Тест на чувствительность к параметрам a и b для компонент импульса отрицательно заряженных пионов (p_x , p_y , p_z , p , соответственно, (a), (b), (c) и (d)).

Формула проверки чувствительности (Normalized Residuals) выражается как:

$$r_i = \frac{n_i - N\hat{p}_i}{\sqrt{N\hat{p}_i \sqrt{(1 - N/(N + M))(1 - (n_i + m_i)/(N + M))}}}, \quad (1)$$

где $\hat{p}_i = (n_i + m_i)/(N + M)$, n_i / m_i – числа событий i -ого бина в эталонной/модифицированной выборках, N/M – общие числа событий в эталонной/модифицированной выборках.

Метод обеспечивает полиномиальное соответствие содержания многих сгенерированных бинов для каждого набора параметров. Особенность использованного в данной работе подхода состоит в том, что содержимое бинов для набора параметров p_{tune} сравниваются не с реальными данными от ускорителей, а с откликами генератора, сгенерированными со значениями параметров по



умолчанию. Для выбора генератора, как было выше упомянуто, нас интересует процесс фрагментации в событиях $e^+ + e^-$, который изучается в эксперименте Belle2 на ускорителе KEK, поэтому очевидно использовать среду Belle2: basf2 [7] (Рис.4). Основой для генерации является генератор событий Pythia8. Фрагментация в Pythia8 реализована на основе струнной модели LUND [8] и представляется в виде:

$$f(z) = (1/z)(1-z)^a \exp(-bm_T^2/z), \quad (2)$$

где a и b – свободные параметры, изменяющиеся в интервалах: $[0, 2]$ и $[0.2, 2]$



(со значениями по умолчанию в Belle: 0.32, 0.62), m_T – поперечная масса, z – относительная энергия адрона. Программа для генерации написана на языке Python и адаптирована для среды Belle2: basf2.

На первом шаге процесса настройки следует определить набор параметров со значениями которых будут сгенерированы спектры различных наблюдаемых. Для этого в пакете «Professor2» есть функция, которая позволяет нам генерировать случайные значения параметров в заданном диапазоне – prof2-sample (Рис.5.). Эта функция генерирует 100 (по умолчанию) наборов параметров с разными значениями.

Минимально необходимое число наборов параметров, реализуется командой prof2-ncoeffs (Рис. 6). В нашем случае при 2-х параметрах, и, скажем, полиноме второго порядка – это минимальное число равно 6. (на самом деле на практике нужно намного больше, чем этот минимум), в данной работе мы взяли 100.

```
prof2-sample -o output params.dat
```

Рис.5. Это команда будет генерировать 100 наборов и будет хранить их в «output» файле. А в «params.dat» файле даны имена параметров и их максимальные и минимальные значения.

Следующим шагом является запуск генератора со всеми полученными параметрами. В этой части мы будем использовать не «Professor2», а наш код, который позволяет нам генерировать события: $e^+ + e^- \rightarrow q\bar{q}$, для кварковых комбинаций $u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}$ (континуальные данные). Псевдо-данные, которые будут использованы сохраняются в ROOT [9] файлах. После генерации событий MC у нас будет 4 ROOT файла (Рис. 7) для каждой комбинации кварков, и все это повторится 100 раз (по числу наборов параметров).

```
prof2-ncoeffs 2
```



Polynomial	order	Minimum	samples
	0		1
	1		3
	2		6
	3		10
	4		15
	5		21
	6		28
	7		36
	8		45
	9		55
	10		66

Рис.6. Даны минимальные необходимое число наборов параметров с помощью команды prof2-ncoeffs.

```
ccbar.root ddbar.root ssbar.root uubar.root
```

Рис.7. ROOT файлы.

2.2. «Вычисления»

После определения содержания всех бинов для каждого отдельного набора параметров, обозначим значение содержания бина как $MC_b(p)$ (очевидно, это будет функция от параметров), метод использует полином второго порядка в качестве основы для параметризации $MC_b(p)$:

$$MC_b(p) \approx \alpha_0 + \sum_i \beta_i^{(b)} p_i' + \sum_{i < j} \gamma_{ij}^{(b)} p_i' p_j', \quad (3)$$

где сдвинутый вектор параметров: $p' \equiv p - p^0$. Для данной полиномиальной формы, нужно определить коэффициенты α, β, γ , чтобы наилучшим образом описать поведение генератора в каждом бине.

Для проверки критерии согласия «Professor2» вычисляет функционал χ^2 и минимизирует его. Функционал χ^2 определяется как:

$$\chi^2(p) = \sum_O \omega_O \sum_{b \in O} \frac{f^{(b)}(p) - R_b}{\Delta_b^2},$$

где R_b – эталонное значение для бина b , а ошибка Δ_b – это общая неопределенность эталонного значения для бина b , также включены веса ω_O для каждого наблюдаемого O («observable», физические величины: проекции импульсов, энергия, ...). Таким образом, можно присвоить статистические веса, для увеличения (уменьшения) значимости различных распределений.

3. Процесс настройки

Имея 100 различных наборов параметров, то есть 100 запусков генератора для каждой комбинации кварков, необходимо объединить выборки полученные в результате генерации 4 комбинаций кварков для формирования континуум выборки (Рис. 8).

Следующим шагом является получение распределений компонентов импульса ($p_x p_y p_z$) для заряженных пионов и каонов (π^+, π^-, K^+, K^-) из континуум выборки. Эти распределения также сохраняются в ROOT файлах. В каждом ROOT-файле содержится 12 разных распределений в виде гистограмм. В итоге имеется 1200 различных гистограмм с разными параметрами.

Важное замечание состоит в том, что в гистограммах не должно быть бинов с нулевым содержанием или их веса должны быть равны нулю. Прежде чем найти лучшие комбинации параметров, которые будут описывать эталонные данные, необходимо выполнить команду `prof2-ipl`.

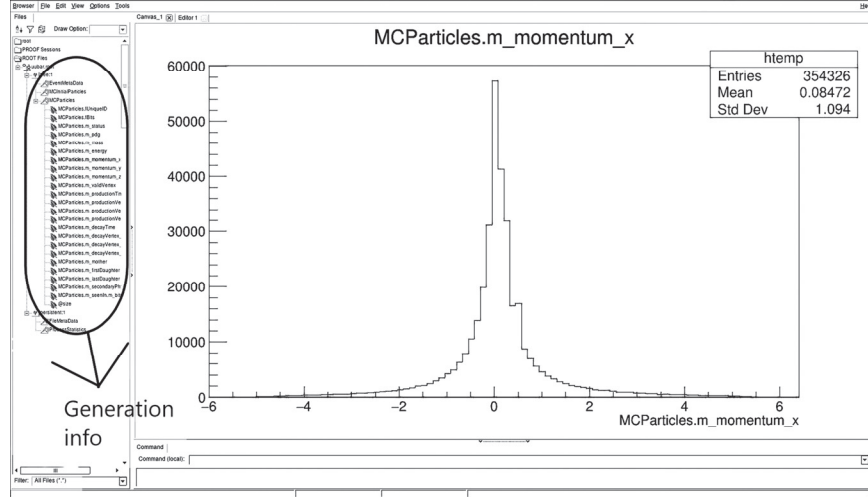


Рис.8. Распределение по x -компоненте импульса полученное от континуума выборки.

Как было упомянуто выше, в качестве эталонных данных, используются запуски генератора со значениями параметров: $a = 0.32$ и $b = 0.62$. На завершающем шаге, запускается команда `prof2-tune` для получения конкретных a и b , описывающие эталонные данные. Затем сгенерированные выборки с настроенными параметрами сравниваются с эталонными выборками. Сравнение показывает очень хорошее соответствие между эталонными и настроенными образцами (Рис. 9), что подтверждает правильность процедуры, описанной в этой статье.

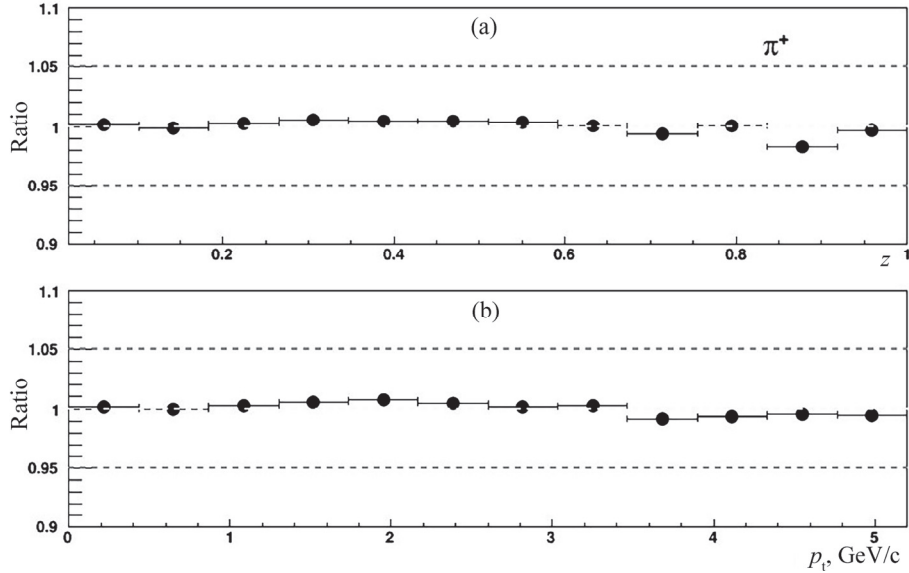


Рис.9. Отношения содержания бинов настроенной и эталонной выборок для положительно заряженного пиона, как функций от относительной энергии (вверху) и поперечного импульса (внизу).

4. Заключение

Основной целью этой статьи была проверка корректности пакета «Professor2», используемого для процесса фрагментации, кварков, которые были произведены в столкновения $e^+ + e^-$. Достижение этой цели продемонстрировано в данной работе. Подобная верификация пакета Professor до сих пор не была реализована для реакции $e^+ + e^-$. Следующим шагом будет проверка модели фрагментации с использованием экспериментальных данных.

ЛИТЕРАТУРА

1. **A.M. Sirunyan, A.R. Tumasyan, V.A. Khachatryan, et al.** J. Contemp. Phys., **52**, 311 (2017).
2. **A.M. Sirunyan, A.R. Tumasyan, V.A. Khachatryan, A.G. Petrosyan.** J. Contemp. Phys., **51**, 229 (2016).
3. **T. Abe et al.** Belle II Technical Design Report, e-Print: 1011.0352 [physics.ins-det] (2010).
4. **E. Kou et al.** PTEP, **2019**, **12** 123C01 (2019), PTEP **2020**, **2** 029201 (2020), (erratum).
5. **A. Buckley, H. Hoeth, H. Lacker, H. Schulz, J.E. von Seggern.** The European Physical Journal C, **65**, 331 (2010).
6. **T. Sjöstrand et al.** Comput. Phys. Commun., **191**, 159 (2015).
7. **T. Kuhr, C. Pulvermacher, M. Ritter, et al.** Comput Softw Big Sci, **3**, 1 (2019).
8. **B.-G. Andersson.** Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol., **7**, 1 (1997).
9. **R. Brun, F. Rademakers.** Nucl. Instrum. Meth. A, **389** 81 (1997).

PROFESSOR2 PACKAGE FOR TUNING THE FRAGMENTATION PROCESS IN e^+e^- ANNIHILATION

V.A. MURADYAN, N.Z. AKOPOV, H.M. GHUMARYAN, G.A. KARYAN

In this paper the hadronization process implemented in the most used physics generator PYTHIA8.2 for simulating the processes in high energy physics has been studied. The goal is to tune PYTHIA8.2 parameter list for the fragmentation process below $b\bar{b}$ production threshold. As a default parameter list we use Belle (predecessor of the Belle II experiment) default parameter settings for fragmentation. For MC pseudo-data we use Belle II software package i.e. basf2 and generator script developed by the main author of this paper. To validate the whole procedure we use momentum distribution of charged pions and kaons in order to tune two free parameters of the fragmentation model in Pythia8 i.e. LUND fragmentation model. We demonstrate that the Professor2 package is handful to perform multi-parameter tune simultaneously.