

Вимірювання перерізів. Розділення сигналу і фону

Олександр Зенаєв

Переріз народження (визначення)

- Переріз народження, або переріз реакції (production cross section, or just cross section, x-section, σ) – фізична величина, яка характеризує ймовірність утворення певного процесу або реакції у зіткненнях частинок
- Переріз народження визначає частоту подій певного типу (наприклад, народження нових частинок). Його можна уявити як ефективну площу, яку “бачать” частинки, що взаємодіють певним чином
- Переріз вимірюється в одиницях площі барн (barn, b): $1\text{b} = 10^{-24}\text{cm}^2$. У HEP зазвичай використовують $1\text{mb} = 10^{-27}\text{cm}^2$, $1\mu\text{b} = 10^{-30}\text{cm}^2 \dots 1\text{fb} = 10^{-39}\text{cm}^2$
- Переріз визначається за формулою:

$$\sigma = \frac{N}{L},$$

де N – кількість подій, L – світимість (luminosity, іноді позначають \mathcal{L})

- L характеризує кількість зіткнень частинок у точці взаємодії та визначається параметрами пучків частинок (інтенсивністю, частотою зіткнень тощо), вимірюється в b^{-1} (inversed barn) або cm^{-2}
- Розрізняють миттєву світимість (instantaneous luminosity) за одиницю часу (наприклад, для LHC Run-3 це приблизно $10^{34}\text{cm}^{-2}\text{s}^{-1}$), та інтегральну світимість (integrated luminosity) за певний час, яка визначає, скільки всього було накопичено даних в експерименті (наприклад, з початку 2024 р. і до вересня ATLAS і CMS накопичили близько 100fb^{-1})

- Для експериментального вимірювання перерізу використовується та сама формула:

$$\sigma = \frac{N}{L}$$

- Втім, N може містити не тільки сигнал, але і фон:

$$N_s = N - N_{bg}$$

Фон або віднімається за передбаченням, або визначається через template fitting

- Крім того, не всі події можуть бути зареєстровані

$$N_s = N_s^{rec} / \epsilon$$

Ефективність реєстрації ϵ визначається з Монте-Карло (МК) симуляції

- Іноді розрізняють аксептанс (A) як частину подій, яка геометрично потрапляє в детектор, і ефективність реєстрації (ϵ) подій детектором (сюди також можуть входити обмеження, що накладаються для зменшення фону)
- Більш повна формула для обрахунку перерізу:

$$\sigma = \frac{N - N_{Bg}}{\epsilon \cdot A \cdot L},$$

Задача: є N подій. Треба визначити частки сигналу (N_s) і фону (N_{bg}).

(1) відняти очікувану кількість фонових подій

- ▶ потребує знання перерізу фонових подій
- ▶ симулюють фонові події, використовуючи певну модель, далі симулюють взаємодію частинок з детектором, проводять реконструкцію подій, визначають N_{Bg} і потім $N_s = N - N_{bg}$
- ▶ результат залежить від моделі (model dependence), яку використовували для симуляції фону (кінематичні розподіли)
- ▶ рідко використовується в чистому вигляді, оскільки фон у НЕР експериментах занадто складний і містить багато компонент

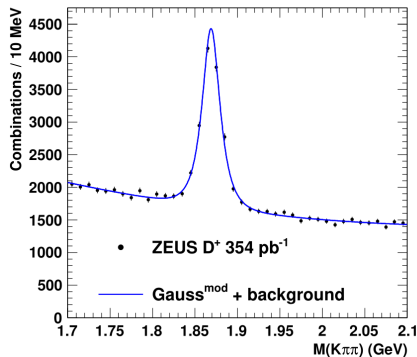
(2) шаблонна підгонка (template fitting)

- ▶ одним із варіантів є реконструкція і фітування розподілу інваріантної маси
- ▶ в цьому випадку N_s визначається як площа піку
- ▶ функція для опису комбінаторного фону зазвичай підбирається емпірично (data driven) і може варіюватися для оцінки систематичної похибки
- ▶ чим менший фон, тим менша буде статистична похибка ΔN_s

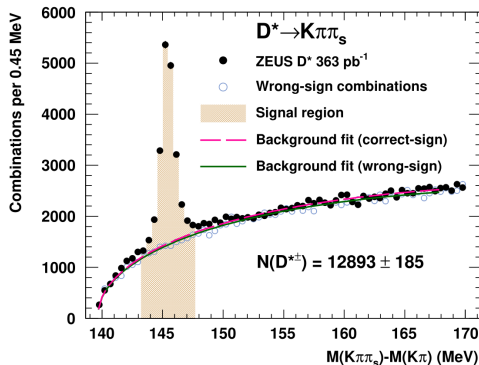
Method	Statistical uncertainty	Model dependence
Background subtraction	$\Delta N_s = \sqrt{N_s}$	+
Template fitting	$\Delta N_s \sim \sqrt{N_s + N_{Bg}}$	-

Н.В. Можуть бути інші специфічні методи, такі як віднімання фону через використання подій з неправильним електричним зарядом та ін.

ZEUS



ZEUS



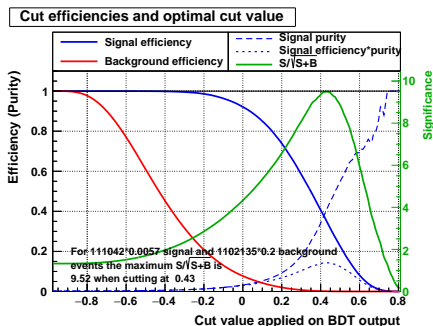
- left: “Measurement of D^\pm production in deep inelastic ep scattering with the ZEUS detector at HERA”: fitting of invariant mass distribution $D^+ \rightarrow K^- \pi^+ \pi^+$
[JHEP 05 (2013) 023, arXiv:1302.5058]
- right: ZEUS Coll., “Measurement of $D^{*\pm}$ production in deep inelastic scattering at HERA”: wrong sign $D^{*+} \rightarrow D^0(\rightarrow K^+ \pi^-) \pi_s^+$ subtraction from correct sign $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi_s^+$
[JHEP 05 (2013) 097, arXiv:1303.6578]

Задача: є N подій. Відсіяти якомога більше фонових подій і залишити якомога більше сигнальних подій, щоб мінімізувати $\Delta N_s / N_s$ (задача класифікації)

- Найважливіша складова аналізу даних: дозволяє “побачити” кілька (десятків, сотень, тисяч) сигнальних подій серед мільярдів усіх зареєстрованих подій
- Найпростіший підхід: накладати обмеження (cuts) на кінематичні властивості подій та/або окремих реконструйованих частинок (імпульси, кути)
- Більш сучасний метод: алгоритми машинного навчання (дерева рішень та ін.)
- Маємо N подій, що містять сигнал і фон ($N = N_s + N_b$). Із них відбираємо N^{sel} подій, які також будуть містити фон і сигнал ($N^{sel} = N_s^{sel} + N_b^{sel}$). Результат характеризується ефективністю (efficiency, ϵ) та чистотою (purity, p):

$$\epsilon = \frac{N_s^{sel}}{N_s}$$

$$p = \frac{N_s^{sel}}{N_s^{sel} + N_b^{sel}}$$



- Використовуючи генератор подій, згенерувати додатковий комбінаторний фон (декілька π^{\pm})
- Накладаючи обмеження на імпульси частинок, зменшити фон
- Визначити ефективність та чистоту
- Визначити переріз народження, скоректований на ефективність відбору подій

Github:

https://github.com/zenaiev/hep/blob/main/invmass/invmass_adv.py

Google Colab:

https://colab.research.google.com/github/zenaiev/hep/blob/main/invmass/invmass_adv.ipynb