

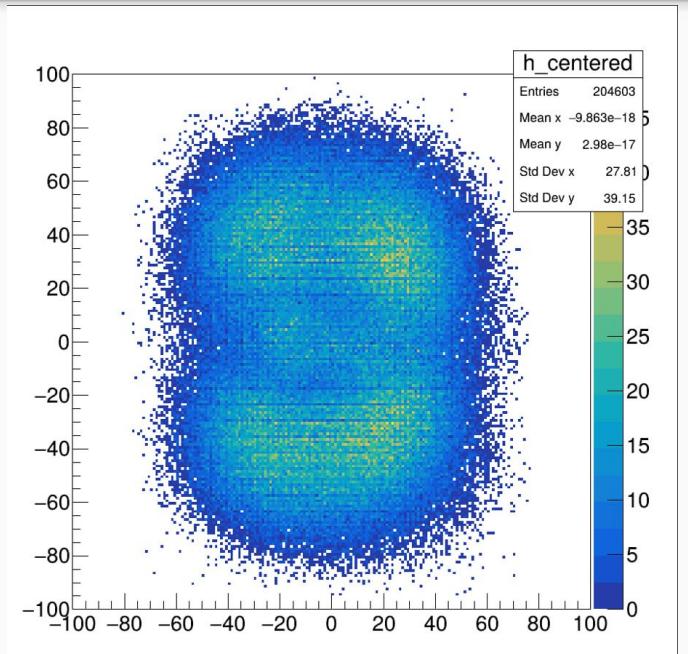
Доклад|17.11

Агафонов Глеб

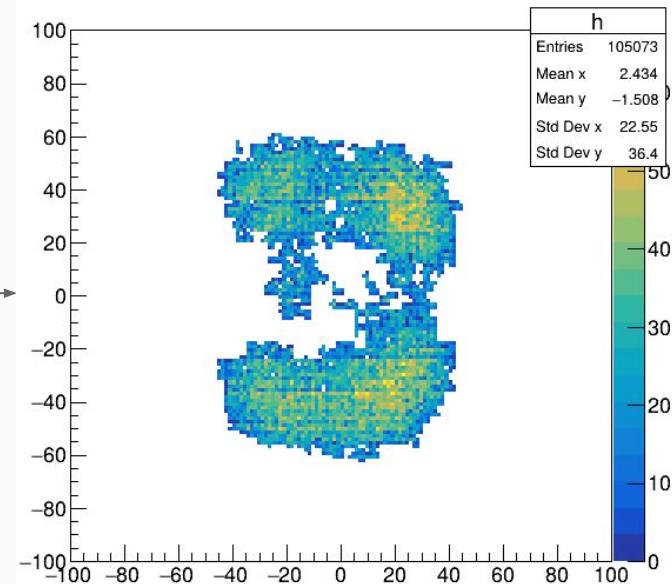
Что было сделано

- Реализован алгоритм кластеризации на основе DBSCAN
- Реализован метод ММ-оценок
- Реализован Гибридный метод
- Реализован алгоритм очистки от фона
- Сделан промежуточный анализ методов против друг друга
- Проведен семинар
- Анализировать эффективность алгоритма очистки от фона

DBSCAN: результаты



НЕЭФФЕКТИВЕН



ММ-оценки

- Получаем первичные оценки параметров кривой используя обычный МНК
- Далее следует S-оценка: минимизируем дисперсию оценок полученных в начале.
- Далее следует M-оценка: используя параметры, которые дала нам S-оценка реализуем минимизацию особой функции от алгебраической ошибки. Это непрерывная, строго положительная, интегрируемая, симметричная и монотонно-возрастающая функция с $p(0) = 0$. Вычисляются весы оценок.
- Это повторяется пока оценки не будут сходиться.

Гибридный метод

Скрестили метод Таубина и Метода ММ-оценок.

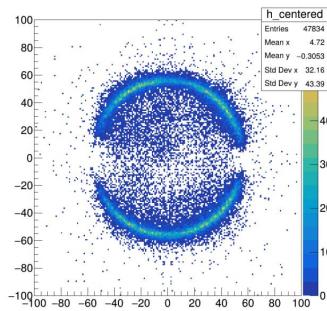
Таубин дает старт этому методы, что увеличивает его скорость и дает нужную точность

Сравнительный анализ

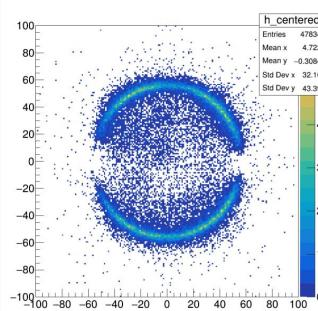
- Количество удачных фитов
- Количество неудачных фитов
- Среднее значение выбросов, минимальное значение выбросов и максимальное значение выбросов.
- Время фитирования одного события
- Насколько близко точки лежат к идеальной окружности

Анализ на разных выборках

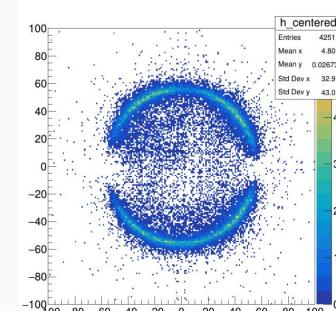
Каса



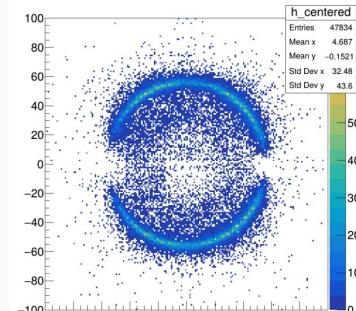
Таубин



ММ



Гибрид

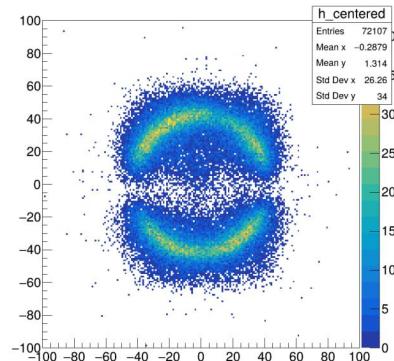


Анализ на разных выборках

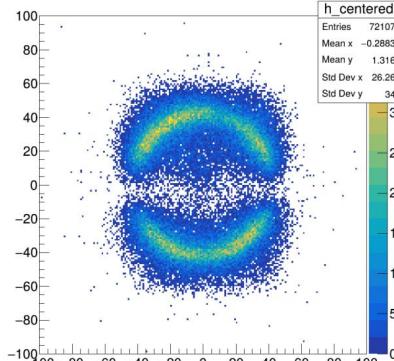
Параметр	Каса	Таубин	ММ	ММ+Таубин	
Успешные фиты	99.7907%	99.814%	99.7156%	99.8023%	Таубин
Неуспешные фиты	0.209302%	0.186047%	0.284372%	0.197674%	Таубин
Ср. значение выбросов	2.81022	2.76363	2.55361	2.45042	ММ + Таубин
Мин. значение выбросов	0	0	0	0	-
Макс. значение выбросов	45.6176	48.7655	45.6176	65.8389	ММ
Ср. время на событие	0.000977406 ms	0.000243233 ms	0.0254954 ms	0.0139876 ms	Таубин
95% выбросов <=	12.177	11.5509	10.9936	11.0238	ММ
Значение выбросов < 1.0 мм	47.6056%	47.9638%	50.6091%	55.9406%	ММ + Таубин

Анализ на разных выборках

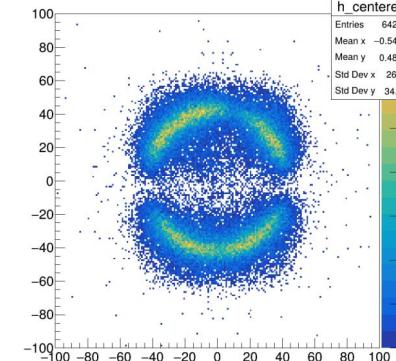
Каса



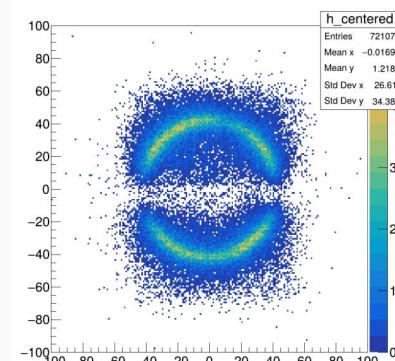
Таубин



ММ



Гибрид

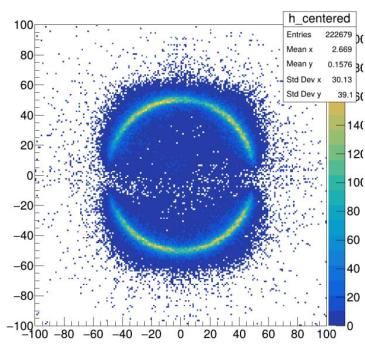


Анализ на разных выборках

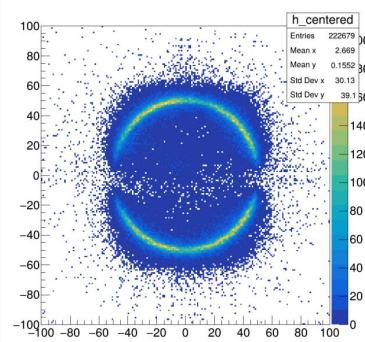
Параметр	Каса	Таубин	ММ	ММ+Таубин	
Успешные фиты	99.9631%	99.9754%	99.9503%	99.9754%	ММ + Таубин/Таубин
Неуспешные фиты	0.036914%	0.0246093%	0.0497451%	0.0246093%	ММ + Таубин/Таубин
Ср. значение выбросов	5.44884	5.39956	5.05502	4.88903	ММ + Таубин
Мин. значение выбросов	0	0	0	0	-
Макс. значение выбросов	40.7719	42.7548	39.3026	54.5749	ММ
Ср. время на событие	0.00151225 ms	0.000260786 ms	0.0424438 ms	0.0241408 ms	Таубин
95% выбросов <=	16.2876	16.3832	15.1739	18.1738	ММ
Значение выбросов < 1.0 мм	15.7744%	16.3437%	17.8417%	25.0052%	ММ + Таубин/Таубин

Анализ на разных выборках

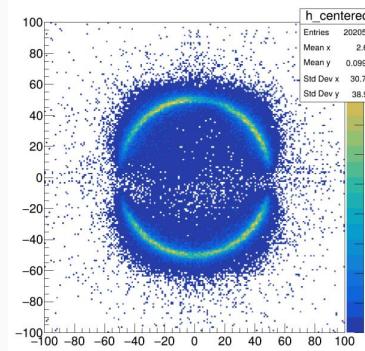
Каса



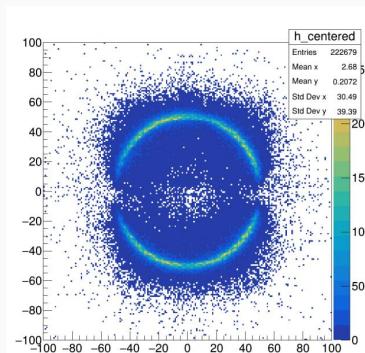
Таубин



ММ



Гибрид

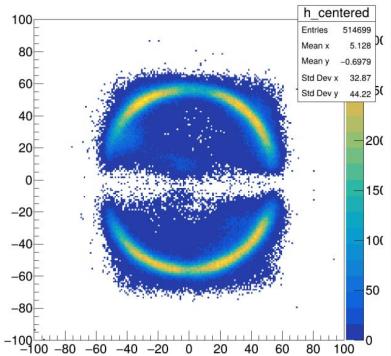


Анализ на разных выборках

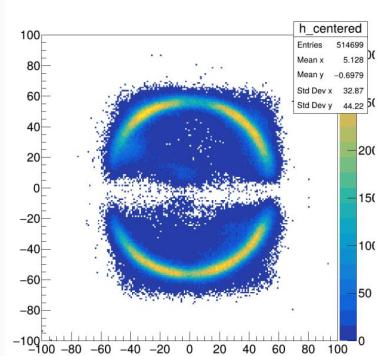
Параметр	Каса	Таубин	ММ	ММ+Таубин	
Успешные фиты	99.8618%	99.8895%	99.8648%	99.8839%	Таубин
Неуспешные фиты	0.138179%	0.110543%	0.13519%	0.11607%	Таубин
Ср. значение выбросов	3.32774	3.26323	2.98292	2.89078	ММ + Таубин
Мин. значение выбросов	0	0	0	0	-
Макс. значение выбросов	43.5402	45.8663	42.883	60.7314	ММ
Ср. время на событие	0.000799304 ms	0.000184915 ms	0.0264444 ms	0.0149915 ms	Таубин
95% выбросов <=	13.5874	13.1548	12.1648	12.8972	ММ
Значение выбросов < 1.0 мм	41.6093%	42.0018%	44.6757%	50.889%	ММ + Таубин

Анализ на разных выборках

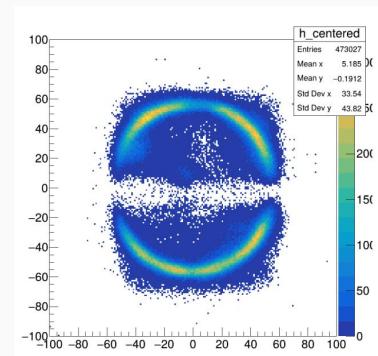
Каса



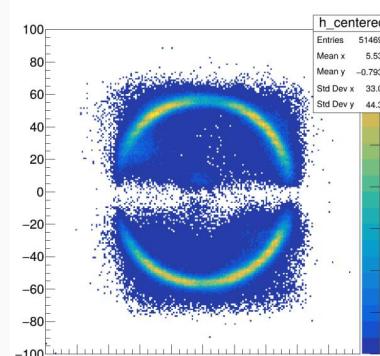
Таубин



ММ



Гибрид



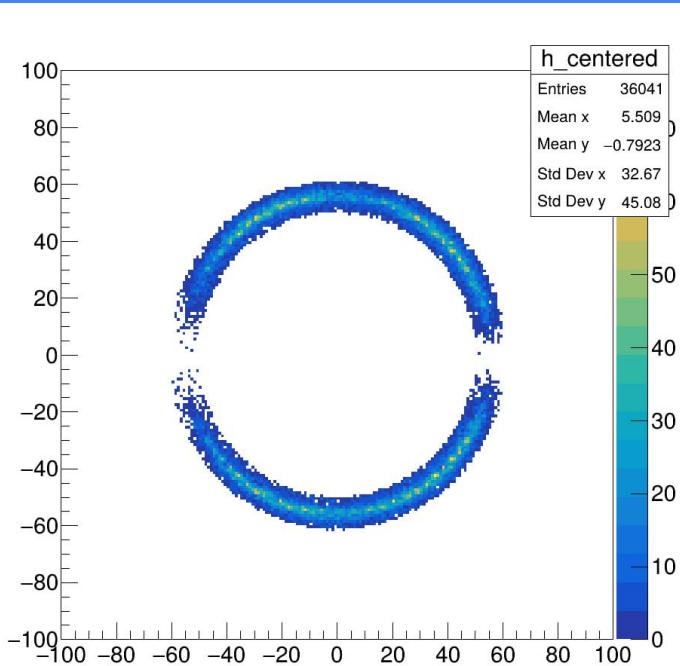
Анализ на разных выборках

Параметр	Каса	Таубин	ММ	ММ+Таубин	
Успешные фиты	99.993%	99.993%	99.993%	99.993%	-
Неуспешные фиты	0.00698861%	0.00698861%	0.00698861%	0.00698861%	Таубин
Ср. значение выбросов	4.00844	3.90836	3.84219	3.68488	ММ + Таубин
Мин. значение выбросов	0	0	0	0	-
Макс. значение выбросов	50.5544	52.0061	51.4038	514693	Каса
Ср. время на событие	0.00202881 ms	0.000419305 ms	0.0709095 ms	0.0393689 ms	Таубин
95% выбросов <=	13.0733	12.7597	12.2231	13.8728	ММ
Значение выбросов < 1.0 мм	22.1054%	22.7951%	22.4223%	28.6526%	ММ + Таубин

ИТОГИ

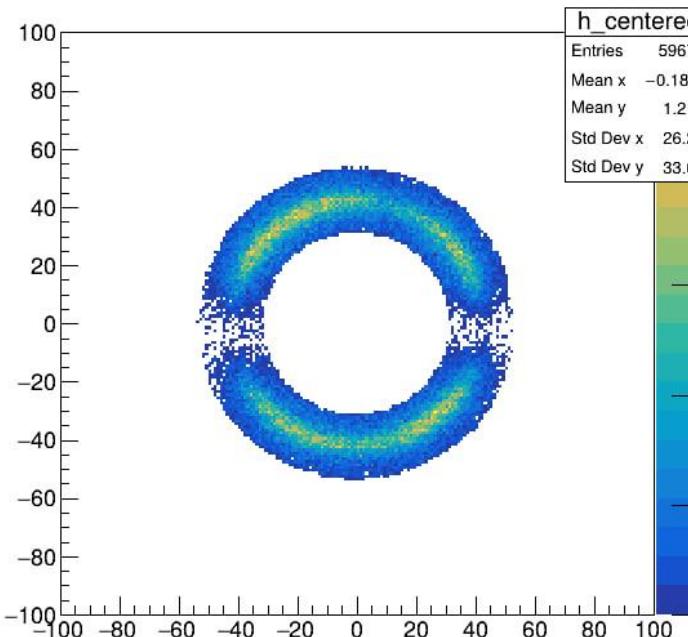
- Каса - самый неточный, неуспешный метод. Терпимо устойчив к выбросам. Скорость невысокая, но и не низкая.
- Таубин - самый быстрый и один из самых успешных методов. Достаточная точность, высокая восприимчивость к выбросам
- ММ-оценки - самый устойчивый к выбросам, самый медленный. Один из самых точных методов, достаточно успешный
- Гибрид - самый точный и успешный метод. Скорость ближе к низкой. Терпимая устойчивость к выбросам в среднем, но допускает сильные выбросы.

Фильтрация



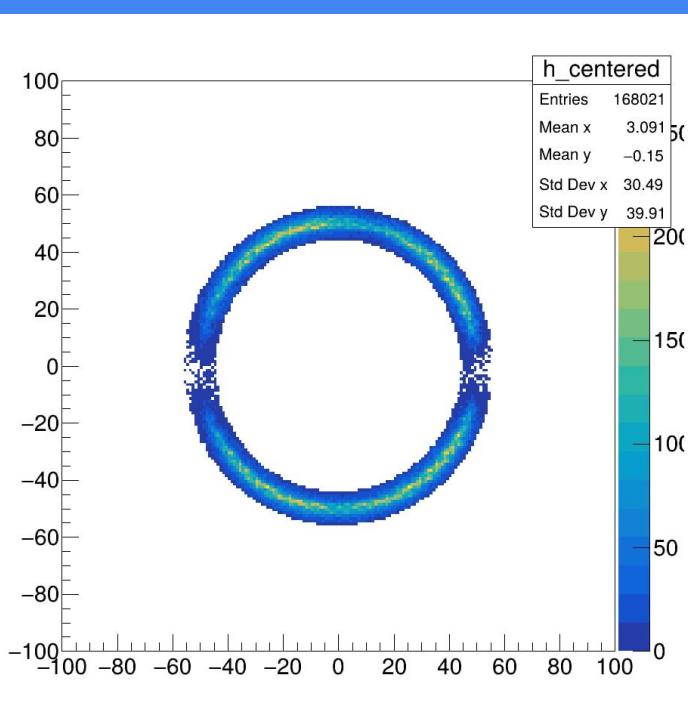
Потеря 32% событий

Фильтрация

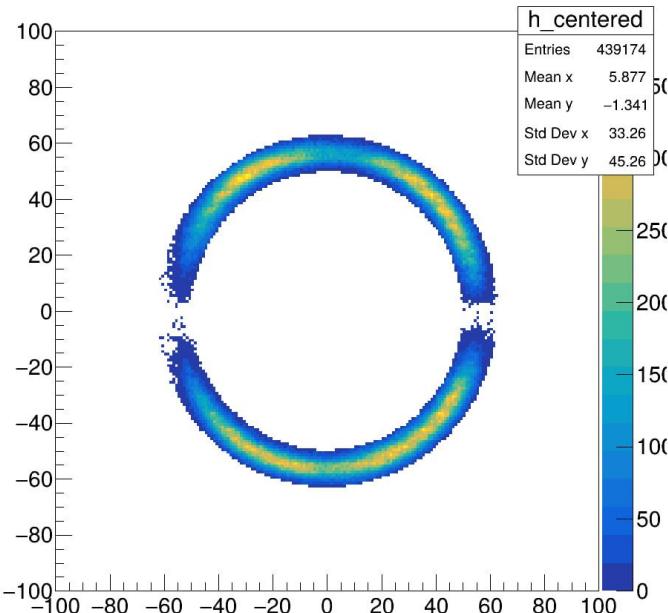


Потеря 16% событий

Фильтрация



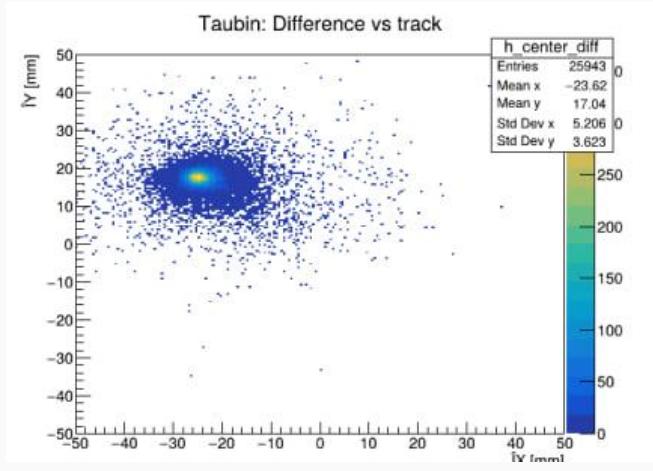
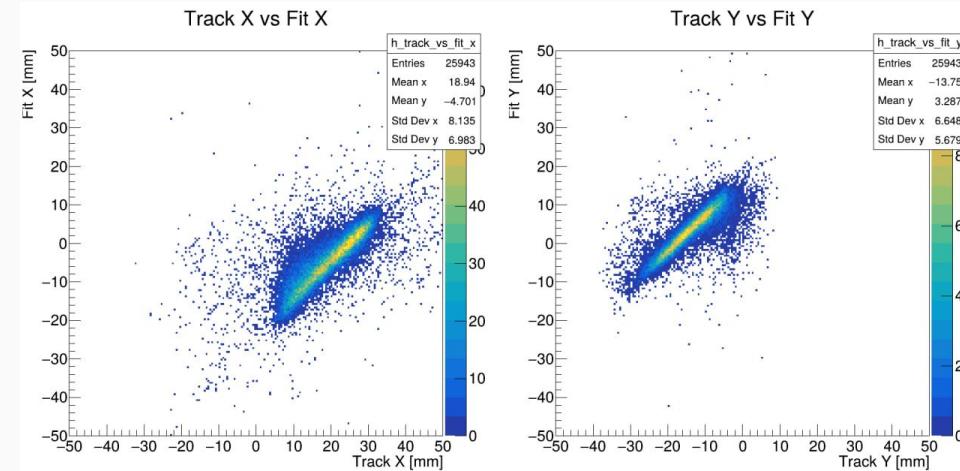
Фильтрация



Потеря 3% событий

Анализ корректности методов

СЛОЖНОСТИ!



Планы

- Завершить анализ аналитических методов
- Усовершенствовать алгоритм фильтрации
- Провести митинг в ИЯФ
- Начать внедрять машинного обучения. Работа с модельными данными