

Байесовская деконволюция для эксперимента СФЕРА-2

Игорь Вайман

13 марта 2021 г.

1 Описание эксперимента

Эксперимент СФЕРА-2 основан на сборе и регистрации черенковского света ШАЛ, отражённого от «экрана» – ровной заснеженной поверхности льда озера Байкал. Установка представляет собой сферическое зеркало и мозаику ФЭУ, установленную вблизи фокальной поверхности. Черенковский свет ШАЛ, рассеиваясь на снегу, проходит через диафрагму установки, отражается от зеркала, и попадает на фотокатод. При попадании фотона с длиной волны в области чувствительности фотокатода с него выбиваются 1 – 10 (уточнить) вторичных электронов, которые проходят систему динодов с разностью потенциалов между каждой парой, создавая в результате избыточный на аноде. Этот заряд стекает с анода через цепь усиления, создавая в результате напряжение на входе АЦП. Это напряжение считывается электроникой, в результате и составляя (вместе с некоторой телеметрией) экспериментальные данные.

Этот набор физических процессов мы описываем в следующих предположениях:

1. В системе нет электронных шумов – единственным источником шума являются фоновые фотоны (от звёздного и зодиакального света, так как эксперимент проводится в ясные безлунные ночи). Фотоны шума попадают в установку аналогично «сигнальным» фотонам – после рассеяния на поверхности. Их поток можно приблизительно оценить по каталогам светимости звёзд и из моделей зодиакального света. (Ссылка на работу Энтиной)

2. Работу ФЭУ можно эффективно описать одной «рандомизированной» импульсной характеристикой (РИХ). Понятие и применение РИХ будет раскрыто далее, в общем виде это случайная функция времени, описывающая отклик системы на один фотон, падающий на фотокатод. Главный источник случайности в установке СФЕРА-2 – неопределённость числа вторичных фотонов, вылетевших с фотокатода, приводящая к неопределённости амплитуды импульса тока. Стоит отметить, что случайный характер ИХ проявляется именно на границе между режимом счёта фотонов и режимом измерения потока, поскольку поток фотонов уже слишком велик, чтобы разрешать отдельные импульсы, но недостаточно – чтобы случайные характеристики эффективно усреднялись. Мы также предполагаем, что случайность импульсной характеристики проявлялась независимо между отдельными фотонами и между разными ФЭУ.
3. Используя экспериментальные данные (данные, записанные АЦП, постоянную компоненту анодного тока и абсолютную калибровку ФЭУ) мы можем восстановить показания анодного тока (ссылка на работу по электронике). Неопределённость, вносимая анодной цепью, усилителем и (главным образом) дискретизацией сигнала на входе АЦП, хорошо известна и учитывается во вторую очередь.

2 Постановка задачи

Рассмотрим N последовательных равновеликих временных бинов. Для простоты будем считать единицей времени длительность одного бина. Тогда мы можем задать временные бины интервалами $[i-1, i]$, $i = 1 \dots N$. Обратим внимание, что при такой нумерации бинов эффект от фотонов, попавших в i -тый временной бин проявляется впервые в момент времени i . Это приводит к симметричной нумерации бинов и временных отсчётов. Обозначим число фотонов в каждом бине как n_i , $i = 1, \dots, N$.

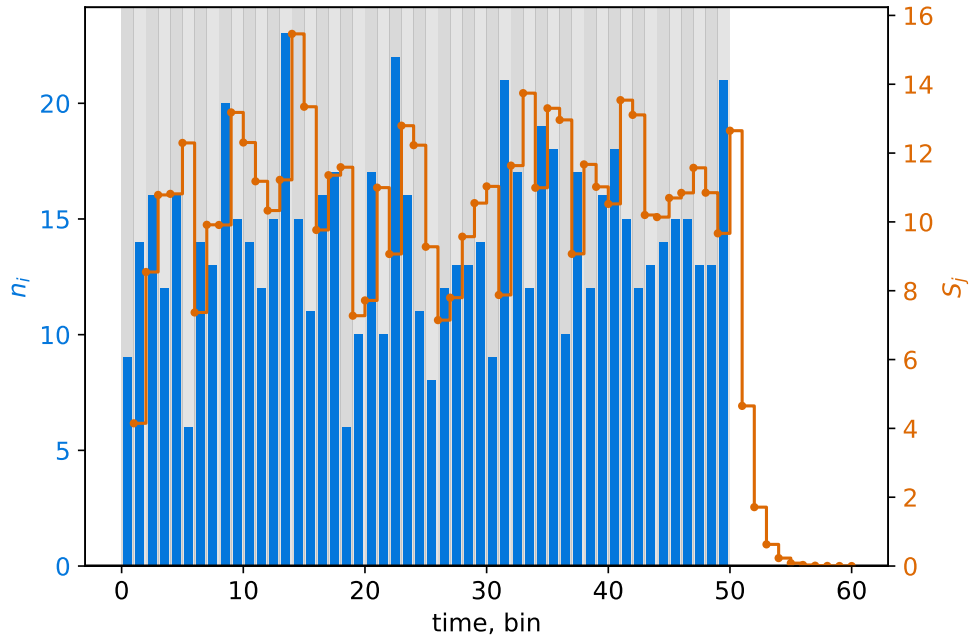


Рис. 1: Пример данных для задачи байесовской деконволюции. Здесь $N = 50$, количество фотонов в каждом бине определено из пуассоновского распределения с $\lambda = \mathbb{E}(n_i) = 15$; РИХ для простоты положена детерминистической функцией $\exp(-t)$ с $L = 10$. Задача состоит в том, чтобы, зная значения выходного сигнала (оранжевый) и РИХ, оценить апостериорные распределения количества фотонов в каждом бине (синих столбцов).

Время прихода фотона внутри бина – также случайная величина, обозначим её t_{inbin} . Она может быть, вообще говоря, распределена произвольным образом в интервале $[0, 1]$, однако мы в простейшем случае будем считать $t_{inbin} \sim U(0, 1)$. Это оправдано для независимых друг от друга фоновых фотонов, и может служить приближением для фотонов ШАЛ в случае, если дисперсия времён прихода фотонов внутри «пакета» сильно превышает длительность временного бина. По данным модельных ливней это не всегда так, поэтому влияние неравномерности распределения времён прихода фотонов будет исследовано отдельно.

Предположим, что импульсная характеристика системы – случайная

функция $\tilde{h}(t)$, любая реализация которой удовлетворяет условиям:

1. Каузальности, то есть $\forall t < 0 \quad \tilde{h}(t) = 0$
2. Конечности во времени, то есть $\exists \tilde{L}$ такое, что $\forall t > \tilde{L} \quad \tilde{h}(t) = 0$.

Заметим, что эффект от фотонов в i -том бине проявляется в отсчётах с i по $i + \lfloor \tilde{L} \rfloor$, где $\lfloor \tilde{L} \rfloor$ – целая часть или округление вниз \tilde{L} . Обозначим $L \equiv \lfloor \tilde{L} \rfloor$. Тогда полный сигнал от фотонов гарантированно содержится в отсчётах с 1 по $N + L$.

Будем считать, что АЦП записывает значения сигнала S_j в точках $j = 1, \dots, N + L$.

Поставим задачу «байесовской деконволюции» следующим образом, используя байесовскую терминологию:

Пусть дана рандомизированная импульсная характеристика системы $\tilde{h}(t)$ и значения S_j , $j = 1, \dots, N + L$. Найти апостериорные функции плотности вероятности для значений n_i , $i = 1, \dots, N$. В качестве априорного распределения n_i используем наивное неограниченное однородное распределение ($p_{n_i}(x) = \text{Const} \quad \forall x \in [0, \infty)$).

Заметим, что, в отличие от обычной деконволюции, мы не ставим задачу оценить исходный сигнал – в данном случае представляющий собой сумму δ -функций с соответствующими сдвигами – но только его агрегированную характеристику. Информация об отдельных фотонах не является необходимой для обработки экспериментальных событий, и её получение представляет собой, по-видимому, более трудную задачу.

Легко заметить также, что в реальном эксперименте мы имеем дело с неограниченным во времени потоком фотонов, а не с изолированными N бинами – эффекты на краях области регистрации сигнала нужно будет учитывать отдельно.

Иллюстрация постановки задачи приведена на рис. 1.

3 Решение

3.1 Выходной сигнал как реализация случайного процесса

Дадим также несколько определений для промежуточных величин, описывающих $\{S_j\}$ как реализацию случайного процесса.

Очевидно, что $\{S_j\}$ – случайные величины, как в силу того, что РИХ в общем случае представляет собой случайную функцию, так и в силу принципиально случайного распределения фотонов внутри временного бина.

Запишем S_j как сумму вкладов от фотонов разных бинов

$$S_j = \sum_{l=0}^L C(n_{j-l}, l) \quad (1)$$

Здесь $C(n, l)$ – случайная величина, описывающая вклад в сигнал на j -том временном отсчёте от n фотонов в бине $j-l$, иначе говоря, вклад с задержкой l бинов (от lag). В силу индексации бинов с 1 и определения L задержка изменяется в пределах от 0 до L .

Случайную величину $C(n, l)$, очевидно, можно выразить через РИХ. Проще всего сделать это через Монте-Карло-сэмплирование распределения. Получим с произвольной точностью эмпирическую функцию плотности распределения для величины $C(1, l)$. Для этого сгенерируем значения $t_i \sim t_{inbin}$; и функции $h_i(t) \sim \hat{h}(t)$ для $i = 1 \dots N_{sample}$. Выборка для $C(1, l)$ тогда будет состоять из значений $h_i(l + 1 - t_i)$. Выборка для $C(n, l)$ легко получить, проделав описанную процедуру n раз и сложив все n реализаций N_{sample} -мерных векторов выборок.

3.2 Грубая оценка методом наименьших квадратов

Перед тем, как полностью решать саму задачу байесовской деконволюции, сделаем грубую оценку \vec{n} , зная выходной сигнал и РИХ. Для этого заменим все случайные величины S_j на их соответствующие математические ожидания. Это аналогично замене РИХ на её среднее значение, а также замене всех времён прихода фотонов на средние.

При этом мы используем независимость фотонов в пределах одного бина, из чего следует $\mathbb{E} C(n, l) = n \mathbb{E} C(1, l)$. Нетрудно заранее вычис-

лечь для данной РИХ выборки значений $C(1, l)$ для $l = 0 \dots L$. Тогда получим, обозначая $c_l \equiv \mathbb{E} C(1, l)$,

$$\hat{S}_j = \mathbb{E} S_j = \sum_{l=0}^L \mathbb{E} C(n_{j-l}, l) = \sum_{l=0}^L n_{j-l} \mathbb{E} C(1, l) = \sum_{l=0}^L n_{j-l} c_l \quad (2)$$

Суммирование можно записать в матричном виде:

$$\begin{pmatrix} c_0 & 0 & 0 & \dots & 0 \\ c_1 & c_0 & 0 & \dots & 0 \\ c_2 & c_1 & c_0 & & \vdots \\ c_3 & c_2 & c_1 & \ddots & 0 \\ \vdots & c_3 & c_2 & \ddots & c_0 \\ c_L & \vdots & c_3 & \ddots & c_1 \\ 0 & c_L & \vdots & \ddots & c_2 \\ \vdots & & c_L & & c_3 \\ 0 & \dots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & 0 & c_L \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_N \end{pmatrix} = \begin{pmatrix} \hat{S}_1 \\ \hat{S}_2 \\ \vdots \\ \hat{S}_N \\ \hat{S}_{N+1} \\ \vdots \\ \hat{S}_{N+L} \end{pmatrix} \quad (3)$$

Это переопределённая система линейных уравнений, для которой можно найти решение в смысле наименьших квадратов с помощью псевдообратной матрицы Мура-Пенроуза (ссылка). Псевдообратная матрица C^+ для C определяется следующими условиями: (1) $CC^+C = C$, (2) $C^+CC^+ = C^+$, (3) CC^+ и C^+C – эрмитовы матрицы. Псевдообратная матрица всегда существует, и для системы $C\vec{n} = \vec{S}$ вектор $C^+\vec{S}$ даёт искомое МНК-решение системы.

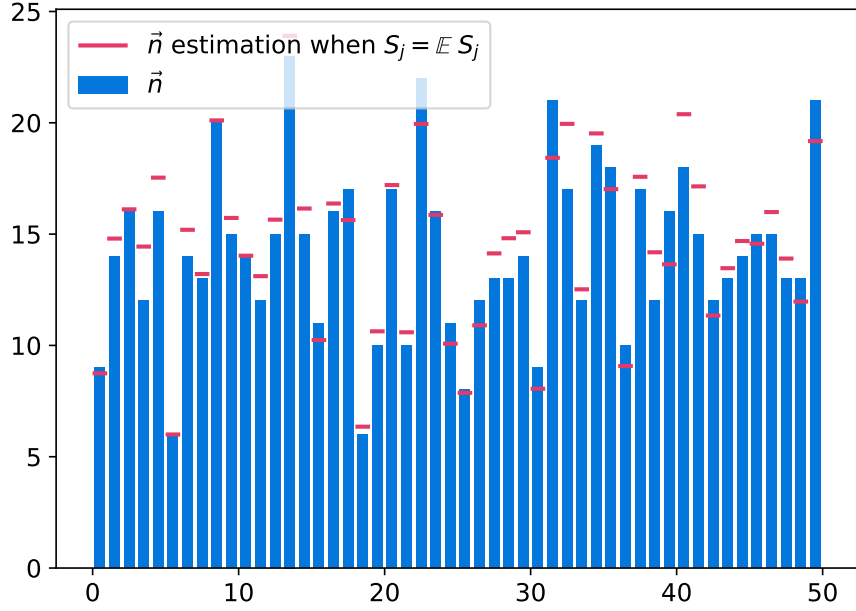


Рис. 2: Оценка \vec{n} в предположении, что выходной сигнал \vec{S} равен своему математическому ожиданию.

Приближённый численный расчёт такой матрицы можно проводить, например, с помощью функции `pinv` модуля `numpy.linalg` в Python (ссылка).

Результат этой процедуры приведён на рис. 2. Видно, что такая процедура даёт по крайней мере корректное первое приближение – именно так эти величины и будут использованы далее.

3.3 Метод Монте-Карло с марковскими цепями (Markov Chain Monte Carlo)

TBD