

Байесовская деконволюция для эксперимента СФЕРА-2

Игорь Вайман

17 мая 2021 г.

Глава 1

Введение

1.1 Цели и задачи

1.2 Актуальность

Глава 2

Методы регистрации и восстановления параметров ШАЛ

В этой главе приводится краткий обзор современных методов регистрации ШАЛ на примере крупнейших установок: **TBD список**; обобщаются методы восстановления параметров первичной частицы, применяемые в этих экспериментах, и ассоциированные с ними неопределённости. Обсуждается важность оценки систематических погрешностей (см. [9]) для корректной интерпретации результатов эксперимента, кратко описываются источники этой погрешности.

6 ГЛАВА 2. МЕТОДЫ РЕГИСТРАЦИИ И ВОССТАНОВЛЕНИЯ ПАРАМЕТРОВ ШАЛ

Глава 3

Эксперимент СФЕРА-2

В этой главе приводится идея, мотивация и краткая история развития метода регистрации отражённого черенковского света, который развивается в эксперименте СФЕРА-2; обзревается проведённые измерения и полученные на сегодняшний день результаты, обосновывается необходимость и основные направление дальнейшей работы; наконец, приводится качественная модель процесса регистрации ШАЛ на этом эксперименте.

3.1 Метод регистрации отражённого черенковского света ШАЛ

TBD: идея Чудакова, Наварро, Тянь-Шань, см. [23] (эта же работа на англ. [3]). Написать про мотивацию для метода: статьи с теоретическими критериями для определения массового состава ШАЛ, например [1], [8]

3.2 Предварительные результаты

TBD: см. [2],

3.3 Качественная модель работы детектора

Опишем на качественном уровне теоретические и методические представления, которые лежат в основе интерпретации данных СФЕРА-2.

Для этого проследим, что происходит от развития каскада в атмосфере и до получения файлов с экспериментальными данными.

3.3.1 Развитие ШАЛ в атмосфере

Сколько-нибудь полный обзор теории ШАЛ, и даже черенковского света ШАЛ в отдельности, лежит за рамками данной работы, поэтому изложим только самые общие представления. Широкий атмосферный ливень – каскад вторичных частиц, вызванный взаимодействием первичной частицы большой энергии с атмосферой – развивается в виде тонкого диска заряженных частиц – адронов, мезонов, лептонов, гамма-квантов. Черенковский свет – одно из сопутствующих излучений ШАЛ, возникающее при движении заряженных частиц со скоростью, превышающую скорость света в среде распространения. Черенковский свет ШАЛ также распространяется в виде тонкого – до нескольких метров в толщину – диска, ориентированного перпендикулярно оси ливня. Для практических задач моделирование процессов развития ШАЛ проводится численно, например, с помощью программы CORSIKA [15].

Уже на этапе моделирования ШАЛ закладывается ряд неопределённостей реконструкции параметров первичной частицы. Во-первых, модели ядерного взаимодействия при энергиях $\gtrsim 10^{16}$ эВ не проверены экспериментально, являясь, в сущности, экстраполяциями, и разные такие модели могут приводить к несколько разным картинам развития ливня для одной и той же первичной частицы; во-вторых, развитие ШАЛ – принципиально стохастический процесс, что приводит к статистической неопределённости любой реконструкции. Эти неопределённости будем в дальнейшем называть *модельными*, так как они напрямую связаны с моделью описываемого явления, в противоположность *инструментальным* неопределёностям, связанным с процессом измерения характеристик ливня.

3.3.2 Отражение света от снега

Эксперименты, использующие отражённый черенковский свет, предъявляют особые требования к земной поверхности на уровне наблюдения – она должна служить «экраном», который бы равномерно и предсказуемо рассеивал свет. В случае СФЕРЫ-2 таким экраном служит заснеженный лёд озера Байкал.

Рассеивающие свойства снега могут существенно повлиять на работу эксперимента, поэтому для измерения и мониторинга этих условий были приложены определённые усилия. В частности, были проведены прямые измерения коэффициента отражения в зависимости от угла [3, рис. 11], и было обнаружено, что зависимость хорошо согласуется с законом рассеяния Ламберта для идеальной диффузной поверхности: *яркость* рассеянного света не зависит от угла, то есть интенсивность имеет чисто геометрическую зависимость $I \propto \cos \theta_n$ [4]. Альбеда a – отношение падающего и отражённого потоков – было принято независимым от длины волны [21] и равным 0.9.

3.3.3 Сбор отражённого света

Установка СФЕРА-2 поднималась аэростатом на высоту 400 – 900 м, для сбора света с поверхности использовалась оптическая система из диафрагмы, сферического зеркала и мозаики ФЭУ, которая схематично изображена на рис. 3.1. В результате каждый из ФЭУ обозревал область на поверхности диаметром 10 – 50 м.

Качественно оценим коэффициент сбора света: пусть на участок поверхности в окрестности точки (x, y) в системе координат с центром в проекции детектора падает δN_{gr} фотонов, и пусть горизонтально ориентированная диафрагма радиусом R_d , поднята на высоту H . Обозначая угол рассеяния света от нормали θ_n , можно оценить число фотонов, которое достигнет диафрагмы из результатов предыдущего раздела и геометрических соображений

$$\delta N_d = \frac{R_d^2 \cos \theta_n}{H^2 + x^2 + y^2} K \cos \theta_n \delta N_{gr} \quad (3.1)$$

Из этого простого расчёта при характерных значениях $H = 600$ м, $R \approx 0.45$ м получается оценка общего коэффициента сбора установки: $(0.5 \div 1) \cdot 10^{-6}$.

Для расчёта количества фотонов, достигающего каждого отдельного ФЭУ, требуется численное моделирование распространения света внутри мозаики, отражения от зеркала, поглощения тыльной стороной мозаики и элементами конструкции. Результатом такого моделирования является коэффициент сбора как функция координат на поверхности, или *чувствительность* каждого ФЭУ $k_i(x, y)$. Более строго, если флюенс че-

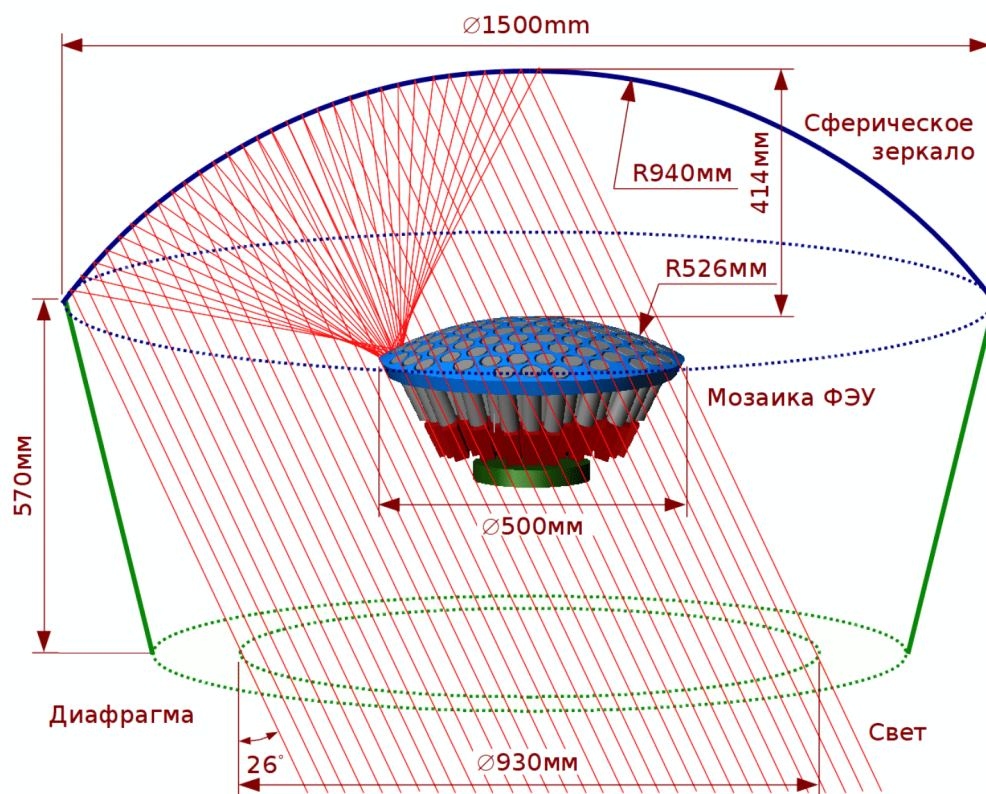


Рис. 3.1: Оптическая схема детектора СФЕРА-2

ренковских фотонов равен $n(x, y)$, то число ожидаемое число фотонов, попавших в i -тый ФЭУ, будет равно $\int_{\infty} k_i(x, y)n(x, y)dxdy$.

Функции k_i имеют форму пятен, соответствующих полям зрения ФЭУ на поверхности льда. Пример нескольких таких пятен, а также суммарной чувствительности всей мозаики для одного из экспериментальных событий изображён на рис. 3.2. Ясно, что поля зрения зависят от высоты подъёма и ориентации установки, поэтому при расчёте учитываются данные GPS и инклинометра соответственно.

3.3.4 Регистрация собранного света

Финальный этап процесса регистрации события включает несколько подэтапов, которые мы изложим особенно подробно, поскольку именно они обуславливают многие исследуемые в настоящей работе эффекты.

3.3.4.1 Мозаика ФЭУ

Вблизи фокальной поверхности сферического зеркала расположена мозаика из 109 ФЭУ (см. рис. 3.1), собранных в приближённо гексагональную сетку. В центре расположен ФЭУ Hamamatsu R3886, характеризующийся большим коэффициентом усиления, остальные – ФЭУ 84-3. Hamamatsu использовался как референсный ФЭУ в процессе калибровки детектора [5].

3.3.4.2 Рождение фотоэлектронов на фотокатоде ФЭУ

Попадая на фотокатод ФЭУ, фотон с определённой вероятностью порождает фотоэлектрон, который затем под действием приложенной разности напряжений устремляется к первому диноду, на котором рождается ещё несколько электронов, и так далее, – таким образом до анода доходит лавина фотоэлектронов, создающая заряд, достаточный для регистрации. Процесс фотоэффекта характеризуется функцией квантовой эффективности, которая характеризует вероятность, с которой фотон данной длины волны породит фотоэлектрон в системе.

Однако, поскольку отражающая способность снега в принятой модели (см. 3.3.2) не зависит от длины волны, спектр черенковских фотонов на фотокатоде оказывается таким же, как и в самом ливне. Поэтому для

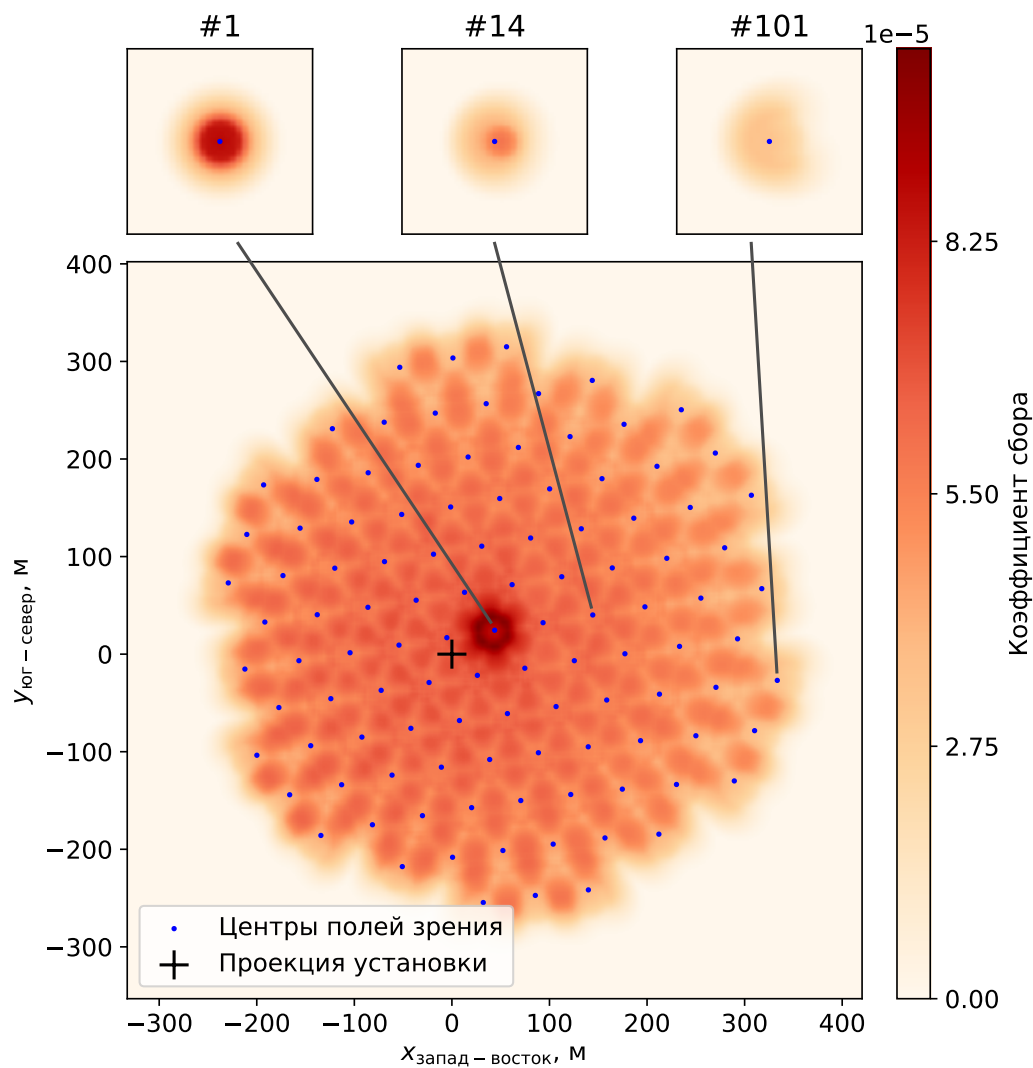


Рис. 3.2: Пример моделирования сбора света с поверхности установкой СФЕРА-2 для экспериментального события #10699. Показаны чувствительности трёх отдельных ФЭУ и суммарная чувствительность всей мозаики. Учтены данные о высоте подъёма и наклоне установки – поэтому картина чувствительности сдвинута относительно проекции установки и слегка вытянута.

сокращения объёма данных при моделировании ШАЛ паспортная квантовая эффективность ФЭУ была заложена уже на этапе прослеживания ливня в атмосфере – вместо спектра черенковского света сохранялась его свёртка с кривой квантовой эффективности, дающая ожидаемое число фотоэлектронов. Таким образом, в настоящей работе *фотоны* ШАЛ всегда подразумеваются уже преобразованными в фотоэлектроны описанным способом, поэтому эти термины во многих случаях оказываются взаимозаменяемы, несмотря на фундаментальные физические различия.

3.3.4.3 Лавинное усиление ФЭУ

Процесс развития электронной лавины на системе динодов ФЭУ – принципиально статистический процесс. Это подтверждается прямыми лабораторными измерениями флуктуаций заряда, собранного на аноде Hamamatsu [5, рис. 9], а также общими представлениями о механизме усиления: если динодная система насчитывает $N \approx 10$ динодов, а общий коэффициент размножения составляет в среднем $K \approx 10^6$, то средний коэффициент размножения на одном диноде будет составлять $\sqrt[N]{K} \approx 4$. Стохастический характер фотоэффекта приводит к тому, что истинный коэффициент умножения на каждом диноде будет иметь пуассоновское распределение с математическим ожиданием $\sqrt[N]{K}$. Эта простая модель позволяет получить распределение коэффициента усиления для ФЭУ84-3, для которого недоступны лабораторные измерения. Результаты этого моделирования и их следствия подробно обсуждаются в разделе 4.5.1 (см. рис. 4.7).

Ясно, что статистический характер коэффициента усиления ФЭУ не играет большой роли при измерении больших потоков, так как происходит эффективное усреднение этой величины. Не важен он и при малых потоках, когда ФЭУ работает в режиме счёта фотонов и на осциллограмме наблюдаются отдельные хорошо разрешённые импульсы. Однако характерная интенсивность потока фотонов в эксперименте СФЕРА-2 такова, что режим работы попадет между этими двумя – поток уже слишком велик, чтобы нельзя было разрешить отдельные фотоны, но ещё недостаточен, чтобы произошло эффективное усреднение. Именно этим обусловлена необходимость статистической деконволюции, описанной в главе 4.

3.3.4.4 Оцифровка импульса анодного тока

Учитывая описанные условия, можно качественно охарактеризовать анодный ток ФЭУ. С одной стороны, постоянный поток фоновых фотонов – преимущественно звёздного и зодиакального света, будут приводить к наличию стационарного тока. С другой стороны, фотоны ШАЛ, приходящие короткой сконцентрированным во времени пакетом, будут давать более и менее яркий импульс на этом фоне. Важный качественный результат лабораторных измерений состоит в том, что другие источники шума – темновой ток или другие электронные шумы – пренебрежимо малы по сравнению с сигналом от фоновых фотонов, поэтому в дальнейшем они учитываться не будут.

Подробное описание электроники детектора может быть найдено в работе [6], здесь же ограничимся качественной картиной: постоянная и импульсная компоненты анодного тока разделяются RC-фильтром и записываются отдельными АЦП. Частота дискретизации для импульсной компоненты составляет 12.5 нс, для постоянной – показания АЦП сохраняются поминутно.

3.3.4.5 Линейность

Линейность системы регистрации света (связки ФЭУ и считывающей аппаратуры) была отдельно исследована по калибровочным кадрам. Интенсивность света в них намного выше даже самого яркого события ШАЛ, но даже в этих условиях амплитудная характеристика остаётся весьма близка к линейной, и может быть подвергнута дополнительной коррекции [5, рис. 6]. Поэтому в дальнейшем система регистрации света считается линейной.

Глава 4

Байесовская деконволюция

В этой главе будут подробнее рассмотрены физические интуиции, связанные с системой регистрации отражённого черенковского света в эксперименте СФЕРА-2; затем на их основе будет формально поставлена задача байесовской (статистической) деконволюции и описан метод её решения в общем виде; наконец, разработанный метод будет применён к экспериментальным данным эксперимента СФЕРА-2.

4.1 Предпосылки

Как описано в разделе 3.3.4.3, процесс генерации импульса анодного тока из фотоэлектронов является существенно стохастическим. Степень влияния этой неопределённости на оценку параметров ШАЛ сложно оценить приближённо, однако ясно, что этот вопрос заслуживает отдельного исследования. Яркие импульсы легко выделяются даже визуально, но сигнал в каналах, наблюдающих удалённые от оси ливня области, будет куда слабее, и его выделение на фоне фотонного шума будет требовать всё большей внимательности к статистическим свойствам системы регистрации.

Сначала была предпринята попытка построить процесс оценки параметров пакета фотонов ШАЛ (числа фотонов, среднего времени прихода и дисперсии фотонов пакета во времени) напрямую из записанного сигнала, однако она не увенчалась успехом из-за вычислительной сложности моделирования детектора. Вместо этого возникла идея провести полноценную деконволюцию, то есть извлечь из записанного сигнала данные о

потоке фотонов в каждый момент времени. Ясно, что эта деконволюция не может быть точной – требуется процедура, учитывающая стохастические свойства системы. Подобные проблемы рассматриваются в других областях, например, в обработке спектральных измерений [17] и изображений [22].

В таких задачах, как и во многих других задачах статистики, плодотворно применяется байесовская статистика, основанная на интерпретации распределения случайной величины как меры информации о ней, или уверенности в её значении [11] – она была выбрана в качестве теоретической рамки в настоящей работе.

Имея в виду эти соображения, построим математическое описание системы регистрации света и используем его, чтобы поставить и решить задачу статистической, или байесовской, деконволюции.

4.2 Структура входного сигнала

Представим входной сигнал системы набором δ -функций, сдвиги которых определённым образом распределены во времени. Рассмотрим N последовательных равновеликих временных бинов. Для простоты будем считать единицей времени длительность одного бина. Тогда мы можем задать границы бинов парами значений $[i - 1, i]$, $i = 1 \dots N$. Обозначим число δ -функций в каждом бине n_i , $i = 1, \dots, N$. В дальнейшем при деконволюции будем интересоваться именно величинами n_i , а не положениями отдельных δ -функций.

Следуя байесовской интерпретации вероятности как *меры уверенности* в данном событии и учитывая, что нас не будут интересовать положения отдельных δ -функций на масштабах меньше 1 временного бина, охарактеризуем это положение случайной величиной $t_{inbin} = \{t\}$. Эта величина может быть, вообще говоря, распределена произвольным образом в интервале $[0, 1)$, однако мы воспользуемся физической интуицией об отсутствии *слишком быстрых вариаций* в интенсивности входного сигнала и положим $t_{inbin} \sim U(0, 1)$. Это оправдано для независимых друг от друга фоновых фотонов, и может служить хорошим приближением для фотонов ШАЛ в случае, если дисперсия времён прихода фотонов внутри «пакета» превышает длительность временного бина.

ТВД: дозаменим фотоны на дельта-функции

Предположим, что импульсная характеристика системы – случайная

функция $\tilde{h}(t)$ в том смысле, что для входного сигнала, состоящего из конечного числа дельта-функций на единицу времени, каждая из этих дельта функций сворачивается с отдельной независимой реализацией $h(t) \sim \tilde{h}(t)$. Будем также считать, что любая реализация удовлетворяет условию каузальности, то есть $\forall t < 0 \quad \forall h \sim \tilde{h} \quad h(t) = 0$, и 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 $\lfloor \tilde{L} \rfloor$ – наибольшее целое число, не превышающее \tilde{L} (округление вниз). Обозначим $L \equiv \lfloor \tilde{L} \rfloor$. Тогда полный сигнал от фотонов гарантированно содержится в отсчётах с 1 по $N + L$.

Таким образом, фотоны, приходящие на вход системы в течение времени $[0, N]$ будут приводить к регистрации отсчётов s_j в моменты времени $j = 1, \dots, N + L$.

Для иллюстрации в течение всей главы 4 будем использовать следующую рандомизированную импульсную характеристику:

$$\begin{aligned} \tilde{h}(t) &= \tilde{A} \exp(-t/2) \\ \tilde{A} &\sim U(0.75, 1) \\ L &= 5 \end{aligned} \tag{4.1}$$

В качестве входных данных сгенерируем сигнал, качественно соответствующий фотонному шуму в эксперименте СФЕРА-2:

$$\begin{aligned} N &= 30 \\ \vec{n} &= \{n_i \sim \text{Poisson}(\lambda) \quad \forall i \in [1, N]\} \\ \lambda &= 15 \end{aligned} \tag{4.2}$$

Иллюстрация \vec{n} и \vec{s} для этой демонстрационной задачи приведена на рис. 4.1.

4.2.1 Эффект окна

В реальном эксперименте входные фотоны не ограничены интервалом $[0; N]$, но приходят постоянно. Модифицируем постановку задачи так, чтобы устранить эффект окна, то есть влияние – плавный рост сигнала в начале и затухание в конце.

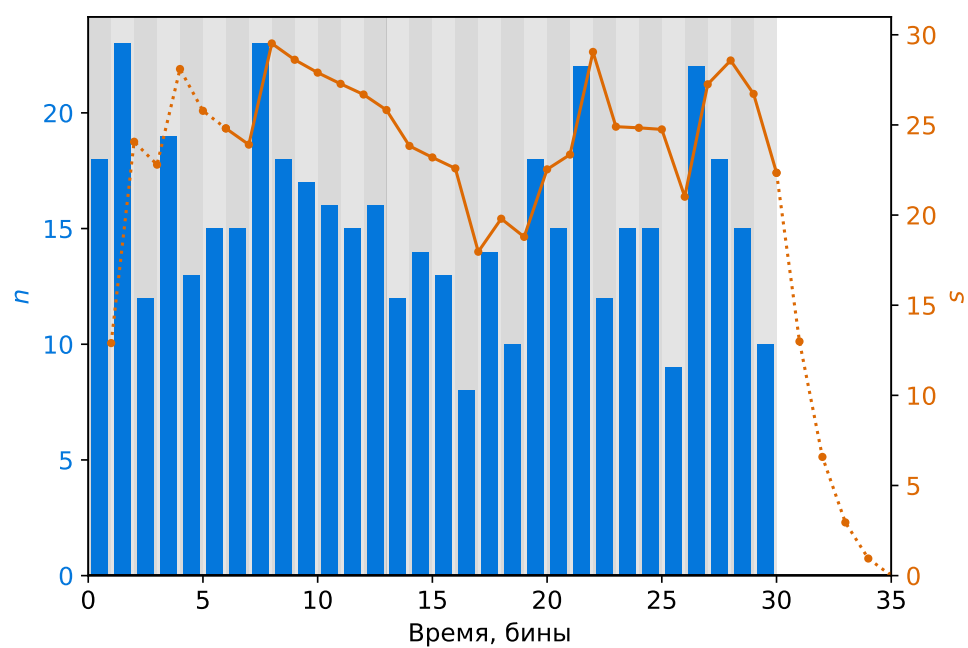


Рис. 4.1: Пример входных и выходных данных для задачи байесовской деконволюции. РИХ определена выражениями (4.1); пунктиром показаны участки, исключаемые из рассмотрения для устранения эффекта окна.

Для этого достаточно исключить из всего дальнейшего рассмотрения эти участки сигнала. Так, в реальной ситуации фотоны продолжают приходить после $t = N$, и вносят соответствующий вклад в отсчёты начиная с $N - 1$. Фотоны, пришедшие до $t = 0$, вносят вклад в отсчёты до L .

Поэтому для восстановления значений n_i , $i = 1, \dots, N$ мы будем использовать только отсчёты S_j при $j = L + 1, \dots, N$. Этот участок изображён на рис. 4.1 сплошной линией, участки по краям, исключаемые из рассмотрения – пунктиром.

4.3 Постановка задачи

Поставим задачу статистической деконволюции следующим образом, используя байесовскую терминологию (поэтому будем также называть эту процедуру байесовской деконволюцией):

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 , $i = 1, \dots, N$.

Заметим, что, в отличие от обычной деконволюции, мы не ставим задачу оценить исходный сигнал сам по себе, представляющий собой сумму δ -функций, но только его несколько обобщённую характеристику.

4.4 Решение

4.4.1 Выходной сигнал как реализация случайного процесса

Ясно, что в силу случайного характера отклика системы, а также агрегирования фотонов в бины, значения отсчётов $\{s_j\}$ являются реализациями некоторых случайных величин. Обозначим сами эти случайные величины как $\{S_j\}$.

Запишем S_j как сумму вкладов от фотонов разных бинов

$$S_j = \sum_{l=0}^L C(n_{j-l}, l) \quad (4.3)$$

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. Из выбранной схемы индексации бинов и условий каузальности и ограниченности во времени РИХ легко видеть, что $l \in [0, L]$, поскольку вклад от фотонов более ранних бинов равен нулю.

Охарактеризуем распределение $C(n, l)$. Проще всего сделать это через Монте-Карло-сэмплирование распределения этой величины. Получим сначала с произвольной точностью эмпирическую функцию плотности распределения для $C(1, l)$. Для этого сгенерируем значения $t_k \sim t_{inbin}$; и функции $h_k(t) \sim \tilde{h}(t)$ для $k = 1 \dots N_{sample}$. Выборка для $C(1, l)$ тогда будет состоять из значений $h_k(l + 1 - t_k)$. Выборка для $C(n, l)$ легко получить, проделав описанную процедуру n раз и сложив все n реализаций N_{sample} -мерных векторов выборок.

4.4.2 Грубая оценка

Перед тем, как решать задачу деконволюции в статистическом смысле, сделаем грубую оценку \vec{n} , основанную только на соотношениях между математическими ожиданиями случайных величин. Для этого применим операцию вычисления математического ожидания к обеим частям равенства (4.3).

Фотоны независимы друг от друга в пределах одного бина, из чего следует $\mathbb{E} C(n, l) = n \mathbb{E} C(1, l)$. Нетрудно заранее вычислить для данной РИХ выборки значений $C(1, l)$ для $l = 0 \dots L$. Тогда получим, обозначая $c_l \equiv \mathbb{E} C(1, l)$,

$$\bar{S}_j = \mathbb{E} S_j = \sum_{l=0}^L \mathbb{E} C(n_{j-l}, l) = \sum_{l=0}^L n_{j-l} \mathbb{E} C(1, l) = \sum_{l=0}^L n_{j-l} c_l \quad (4.4)$$

Суммирование можно записать в матричном виде для $j = 1, \dots, N + L$:

$$\begin{pmatrix} c_0 & 0 & 0 & \cdots & 0 \\ c_1 & c_0 & 0 & \cdots & 0 \\ c_2 & c_1 & c_0 & & \vdots \\ c_3 & c_2 & c_1 & \ddots & 0 \\ \vdots & c_3 & c_2 & \ddots & c_0 \\ c_L & \vdots & c_3 & \ddots & c_1 \\ 0 & c_L & \vdots & \ddots & c_2 \\ \vdots & & c_L & & c_3 \\ 0 & \cdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & c_L \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_N \end{pmatrix} = \begin{pmatrix} \bar{S}_1 \\ \bar{S}_2 \\ \vdots \\ \bar{S}_N \\ \bar{S}_{N+1} \\ \vdots \\ \bar{S}_{N+L} \end{pmatrix} \quad (4.5)$$

Для исключения граничных эффектов, и тем самым перехода к случаю непрерывного потока фотонов достаточно, как указано в разделе 4.2.1, ограничиться рассмотрением строк $j = L+1, \dots, N$, то есть убрать первые и последние L уравнений системы.

Эта возможно несовместная система линейных уравнений, допускает решение в смысле наименьших квадратов с помощью псевдообратной матрицы Мура-Пенроуза [16]. 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. Приближённый численный расчёт такой матрицы можно провести, например, с помощью функции `pinv` модуля `numpy.linalg` в Python [14].

Однако для решения системы необходимо знать $\mathbb{E} S_j$, в то время как в эксперименте мы имеем всего лишь единственную реализацию этой случайной величины s_j . Для грубой оценки остаётся положить $\mathbb{E} S_j \cong s_j$. Результат описанной процедуры приведён на рис. 4.2 фиолетовым, пунктиром отмечена область, в которой оценка искажена эффектом окна.

Строгое исследование свойств такой «one-shot» оценки \vec{n} находится за рамками данной работы. Однако для простой численной проверки можно провести описанную процедуру для большего числа входных бинов и сравнить распределения истинных и оцененных значений \vec{n} . На рис. 4.3 приведено такое сравнение для 1000 бинов. Видно, что распределения практически совпадают, а значит отсутствует по крайней мере система-

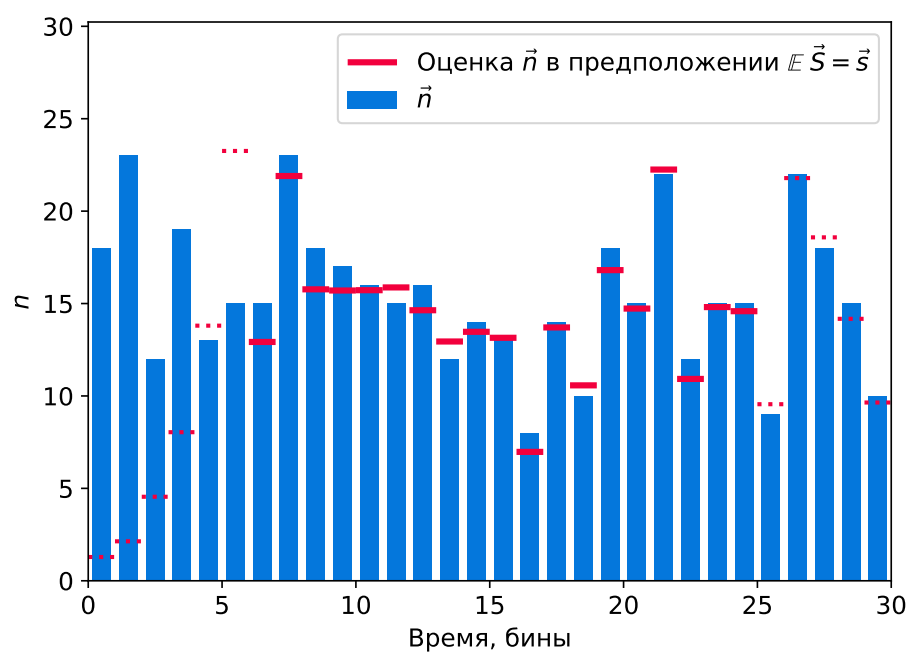


Рис. 4.2: Оценка \vec{n} в предположении, что выходной сигнал \vec{s} равен своему математическому ожиданию. Пунктиром показана область, где оценка искажена эффектом ограниченности выборки во времени.

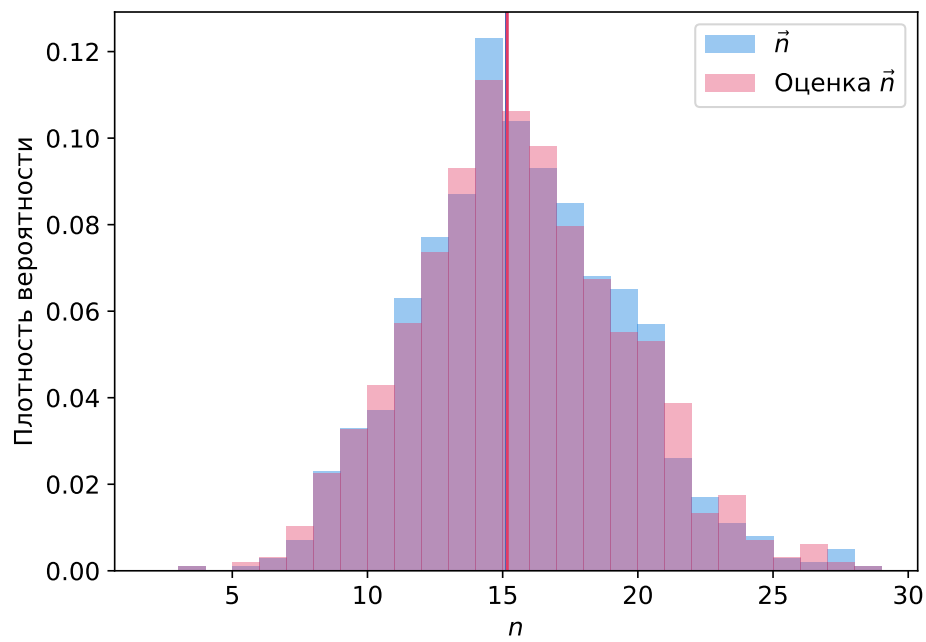


Рис. 4.3: Распределения истинных и грубо восстановленных значений \vec{n} для 1000 значений. Истинные значения выбраны из пуассоновского распределения с $\lambda = 15$, аналогично рис. 4.1 и 4.2.

тическая ошибка. В дальнейшем эта оценка будет играть роль первого приближения, или стартовой точки, на основе которой уже полными статистическим методом можно найти полное решение.

4.4.3 Решение задачи байесовской деконволюции

Теперь перейдём к решению полноценной статистической задачи, используя метод, описанный в предыдущем разделе, как первое приближение.

Запишем сначала теорему Байеса в общем виде, учитывая, что наблюдаемыми значениями является сигнал \vec{s} , а неизвестными параметрами, которые задают распределение наблюдаемых – \vec{n} :

$$P(\vec{n}|\vec{s}) = \frac{P(\vec{s}|\vec{n}) P(\vec{n})}{P(\vec{s})} \quad (4.6)$$

Поясним вероятности, входящие в выражение:

1. $P(\vec{n}|\vec{s})$ – искомое *апостериорное* распределение, описывающее (в байесовском определении вероятности) наши знания о \vec{n} после проведения измерений
2. $P(\vec{s}|\vec{n}) \equiv \mathcal{L}(\vec{s}, \vec{n})$ – функция правдоподобия, описывающая, насколько вероятна регистрация определённых значений \vec{s} при заданных параметрах \vec{n}
3. $P(\vec{n}) \equiv \pi(\vec{n})$ – *априорное* распределение \vec{n} , не требующее знаний о конкретном выходном сигнале
4. $P(\vec{s})$ – полная вероятность регистрации данного сигнала при всех возможных значениях \vec{n} . Она также называемая маргинальной вероятностью или нормировочным множителем, по определению $P(\vec{s}) = \int_{\infty} P(\vec{s}|\vec{n}) P(\vec{n}) d\vec{n}$. Эту величину можно использовать для сравнения моделей, например, если бы мы имели в распоряжении альтернативную модель генерации значений \vec{s} и хотели бы понять, какая из двух лучше описывает данные. Однако для поставленной задачи нет нужды ни вычислять, ни даже учитывать этот множитель.

Учитывая введённые обозначения, запишем

$$P(\vec{n}|\vec{s}) \propto \mathcal{L}(\vec{s}, \vec{n}) \pi(\vec{n}) \quad (4.7)$$

4.4.3.1 Выбор априорного распределения

В качестве априорного распределения будем использовать неограниченное равномерное, или *неинформативное* распределение. Методы байесовской статистики работают даже в ситуации, когда у нас нет вообще никакой информации о распределении \vec{n} до начала измерений. Формально для этого нужно положить $\pi(\vec{n}) = \text{Const} \forall \vec{n}$. Такое распределение нельзя использовать напрямую, поскольку его невозможно нормировать на 1. Однако тогда в выражении 4.7 можно просто исключить $\pi(\vec{n})$ из правой части.

4.4.3.2 Полное вычисление функции правдоподобия методом Монте-Карло

Функция правдоподобия $\mathcal{L}(\vec{s}, \vec{n})$ определяется как вероятность того, что данный сигнал \vec{s} получился в результате преобразования системой входного сигнала \vec{n} .

Опишем сначала Монте-Карло метод оценки правдоподобия, и затем введём упрощения, которые позволят эффективнее вычислять эту функцию, а также обоснуем корректность этих упрощений. Для этого при фиксированных \vec{n} необходимо смоделировать выборку из большого числа реализаций случайной величины \vec{S} , а затем оценить плотность вероятности в точке \vec{s} по этой выборке.

Существует несколько методов для такой оценки, самый простой из которых – многомерная гистограмма. Для этого пространство реализаций \vec{S} нужно поделить на ячейки, подсчитать число элементов выборки, попавших в каждую ячейку, и поделить на общее число элементов выборки и на объём ячейки. Нетрудно видеть, что полученное число даст оценку плотности вероятности в произвольной точке – а значит, искомую оценку функции правдоподобия.

Однако оказывается, что для поставленной задачи метод прямого вычисления функции правдоподобия плохо подходит из-за «проклятия размерности». Размерность пространства, в котором нужно оценить эмпирическую функцию плотности вероятности, равна ширине окна, в котором рассматривается входной сигнал, в бинах. Соответственно, общее число N -мерных бинов составляет g^N , где g – мощность бинирования каждого отсчёта. В этой ситуации для надёжной оценки придётся генерировать выборку сравнимого объёма.

Поэтому вместо прямого использования описанного метода аппроксимируем распределение \vec{S} более вычислительно эффективной функцией.

4.4.3.3 Аппроксимация функции правдоподобия многомерным нормальным распределением

Проанализируем результат полной Монте-Карло симуляции, чтобы исследовать распределение $p(\vec{S}|\vec{n})$ и показать, что его в самом деле можно эффективно и с высокой точностью аппроксимировать многомерным нормальным распределением.

Частные распределения Из самых общих соображения ясно, что распределение $S_j \forall j$ будет более или менее близко к нормальному – просто в силу того, что каждая из S_j является суммой независимых (хотя и не всех одинаково распределённых) вкладов от фотонов в предыдущих бинах. Также сразу можно сказать, что *чем меньше среднее число фотонов в бине, тем хуже будет аппроксимация нормальным распределением*. На рис. 4.4 приведены несколько распределений «лучшего», «медианного» и «худшего» распределений S_j для нескольких значений среднего числа фотонов в бине $\lambda_n \equiv \mathbb{E} n_i$. Из предварительного анализа и независимых соображений известно, что менее 5 фотонов на бин – весьма нечастая ситуация в экспериментальных данных – но даже в этом случае видно, что самое существенное среднеквадратичное отклонение плотности распределения от гауссианы не превысит 10%. Таким образом, *частное распределение* каждой компоненты вектора \vec{S} можно приблизить нормальным распределением с эффективной погрешностью в несколько процентов для практически значимых случаев.

Полное $(N - L)$ -мерное распределение Провести аналогичную процедуру для полного $(N - L)$ -мерного распределения – непростая задача, поэтому ограничимся здесь визуальным исследованием двухмерных распределений для пар $(S_j; S_k)$ для $k - j < L$ (из постановки задачи ясно, что на бóльших расстояниях элементы \vec{S} независимы). На рис. 4.5 представлены такие распределения. Видно, что двухмерные распределения представляют собой овалы симметричной формы, как и ожидается от многомерного нормального распределения.

Таким образом, можно положить, что *распределение вектора \vec{S} при заданных \vec{n} можно приблизить многомерным нормальным распреде-*

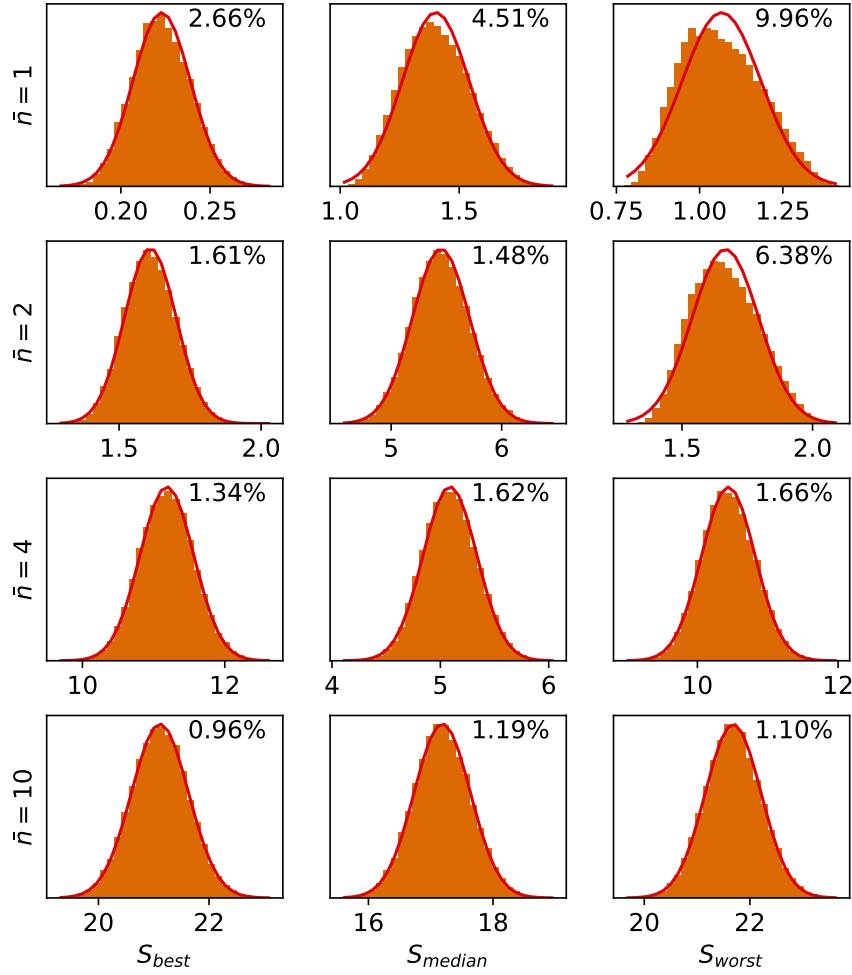


Рис. 4.4: Оценка нормальности частных распределений компонент \vec{S} для разных интенсивностей входного потока: от экстремально низких 1 и 2 до средней ожидаемой 10. «Лучший», «медианный» и «худший» бины определены эвристически по квадрату разницы выборочного среднего и медианы (чем больше эта величина, тем менее симметрично распределение). На каждом графике построена плотность нормального распределения с соответствующими μ и σ , а также приведено среднеквадратичное отклонение гистограммы от гауссианы в процентах относительно максимального значения.

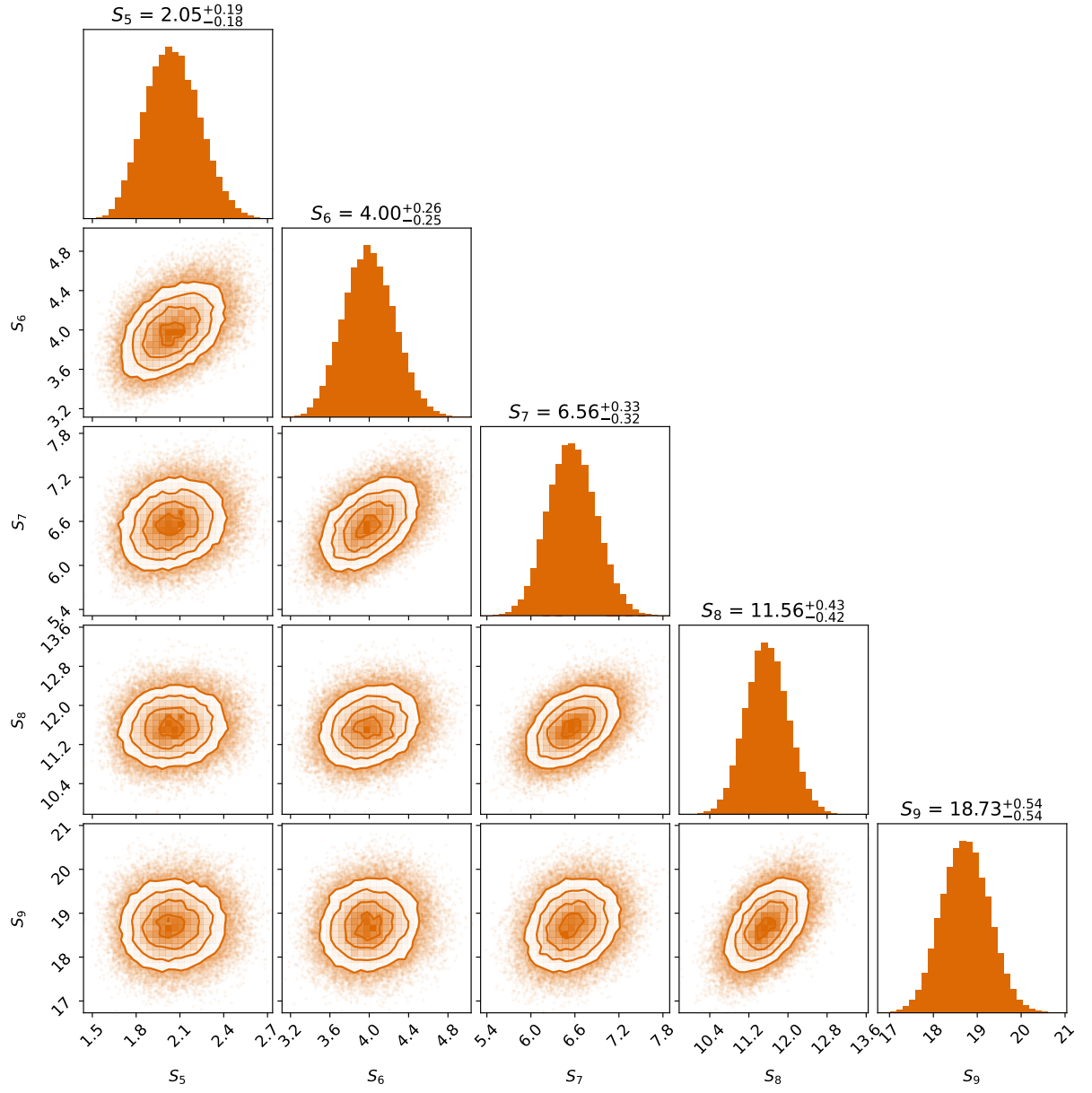


Рис. 4.5: Оценка нормальности совместных двумерных распределений для пар компонент \vec{S} . График построен с помощью библиотеки corner [10].

лением. Это приближение существенно упрощает вычисление функции правдоподобия $\mathcal{L}(\vec{n}) \equiv P(\vec{S}|\vec{n})$, которая и сводится к вычислению плотности вероятности условного распределения \vec{S} в точке, заданной наблюдаемым сигналом. Этот метод оказывается гораздо более устойчивым и вычислительно эффективным, чем наивный метод Монте-Карло, описанный в разделе 4.4.3.2.

Многомерное нормальное распределение в общем виде задаётся вектором математических ожиданий $\vec{\mu}$ и матрицей ковариаций Σ . Функция плотности вероятности $(N + L)$ -мерного случайного вектора \vec{S} тогда вычисляется как

$$p(\vec{S}) = ((2\pi)^{N+L} \det \Sigma)^{-1/2} \exp \left(-\frac{1}{2} (\vec{S} - \vec{\mu})^T \Sigma^{-1} (\vec{S} - \vec{\mu}) \right) \quad (4.8)$$

Вектор $\vec{\mu} \equiv \mathbb{E} \vec{S}$ как функция \vec{n} вычисляется в соответствии с выражением (4.5). Вычислим, аналогично, произвольный элемент матрицы $\Sigma_{ij} \equiv \text{cov}(S_i, S_j)$ при заданном \vec{n} . В силу симметричности матрицы ковариаций можно положить для определённости $i \leq j$. Запишем выражение (4.3) для двух интересующих нас элементов \vec{S} :

$$S_i = \sum_{l=0}^L C(n_{i-l}, l) \quad (4.9)$$

$$S_j = \sum_{k=0}^L C(n_{j-k}, k) \quad (4.10)$$

Источник возможной ненулевой ковариации этих величин – члены, описывающие вклады фотонов из одного бина, но на разных задержках, все прочие слагаемые независимы. Условие, описывающее такие попарно-скоррелированные члены, получается из равенства индексов n в суммах: $i - l = j - k$. Вводя $\Delta \equiv j - i$, отсюда получаем $k = l + \Delta$. Отсюда сразу следует интуитивный вывод о том, что при $\Delta > L$ ковариация будет заведомо нулевой в силу конечности РИХ во времени.

Используем следующее свойство ковариации: если x , y , ϵ , η – случайные величины, из которых только x и y являются зависимыми, то $\text{cov}(x + \epsilon, y + \eta) = \text{cov}(x, y)$. Иначе говоря, заведомо независимые члены при вычислении ковариации можно просто опустить. Прделаем это в

суммах выше, получив таким образом «скореллированные части» \hat{S}_i и \hat{S}_j такие, что $\text{cov}(S_i, S_j) = \text{cov}(\hat{S}_i, \hat{S}_j)$:

$$\hat{S}_i = \sum_{l=0}^{L-\Delta} C(n_{i-l}, l) \quad (4.11)$$

$$\hat{S}_j = \sum_{k=\Delta}^L C(n_{j-k}, k) = [k = l + \Delta] = \sum_{l=0}^{L-\Delta} C(n_{i-l}, l + \Delta) \quad (4.12)$$

Наконец, используем ещё одно свойство ковариации: если среди случайных величин x_1, x_2, y_1, y_2 независимы все пары, кроме x_1 и y_1, x_2 и y_2 , то $\text{cov}(x_1 + x_2, y_1 + y_2) = \text{cov}(x_1, y_1) + \text{cov}(x_2, y_2)$. Индуктивно обобщая на случай сумм с произвольным числом членов, и принимая во внимание, что фотоны в разных бинах независимы, получаем

$$\text{cov}(S_i, S_j) = \text{cov}(\hat{S}_i, \hat{S}_j) = \sum_{l=0}^{L-\Delta} \text{cov}(C(n_{i-l}, l), C(n_{i-l}, l + \Delta)) \quad (4.13)$$

Наконец, принимая во внимание, что сама по себе величина $C(n, l)$ есть сумма n независимых одинаково распределённых случайных величин, а между членами сумм $C(n, l)$ и $C(n, l + \Delta)$ есть только попарные корреляции в отношении 1 к 1, можно записать окончательную формулу для вычисления элемента матрицы ковариации:

$$\text{cov}(S_i, S_{i+\Delta}) = \sum_{l=0}^{L-\Delta} n_{i-l} \text{cov}(C(1, l), C(1, l + \Delta)) \quad (4.14)$$

Из формулы видно, что автокорреляция сигнала описана полностью в терминах автокорреляции РИХ и чисел фотонов в соответствующих бинах, служащих весами для вклада разных участков РИХ.

Для удобства вычислений можно записать это равенство в матричном виде для значений $\Delta \in [0; L]$. Для этого введём обозначение $\xi(l, \Delta) \equiv \text{cov}(C(1, l), C(1, l + \Delta))$ и с помощью него запишем:

$$\begin{pmatrix} \xi(L, 0) & \xi(L-1, 0) & \xi(L-2, 0) & \cdots & \xi(0, 0) \\ 0 & \xi(L-1, 1) & \xi(L-2, 1) & \cdots & \xi(0, 1) \\ 0 & 0 & \xi(L-2, 2) & \cdots & \xi(0, 2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \xi(0, L) \end{pmatrix} \begin{pmatrix} n_{i-L} \\ n_{i-(L-1)} \\ n_{i-(L-2)} \\ \vdots \\ n_i \end{pmatrix} = \begin{pmatrix} \Sigma_{i,i} \\ \Sigma_{i,i+1} \\ \Sigma_{i,i+2} \\ \vdots \\ \Sigma_{i,i+L} \end{pmatrix} \quad (4.15)$$

Это достаточно неинтуитивное выражение позволяет, однако, единожды вычислить матрицу Ξ и эффективно рассчитывать все элементы матрицы ковариаций Σ при каждом заданном \vec{n} . Окончательное единое выражение для всей матрицы Σ будет излишне громоздким, и при реальном вычислении не требуется – достаточно просто составить её из векторов $(\Sigma_{i,i}, \dots, \Sigma_{i,i+L})^T$, учитывая свойство симметрии $\Sigma_{i,j} = \Sigma_{j,i}$, а также доказанный факт $\Sigma_{i,j} = 0 \quad \forall j > i + L$.

После исключения областей влияния эффекта окна – первых и последних L элементов \vec{S} – нас будут интересовать только внутренняя часть матрицы Σ , то есть индексы $i, j \in [L+1; N]$. Это удобно, так как для вычисления интересующей нас части Σ не нужно рассматривать «урезанную» матрицу Ξ , что потребовалось бы, если бы мы не исключали границы.

Таким образом, функция правдоподобия задаётся формулой (4.8), где вектор математических ожиданий и матрица ковариаций являются функциями \vec{n} и вычисляются в соответствии с формулами (4.5) и (4.15).

4.4.3.4 Сэмплирование апостериорного распределения

Имея функцию правдоподобия \mathcal{L} и учитывая выбранные неинформативные априорные распределения, мы получаем функцию плотности апостериорной вероятности $P(\vec{n}|\vec{s}) \propto \mathcal{L}(\vec{n}, \vec{s})$, которую требуется исследовать.

Требуется получить не только наиболее оптимальное значение \vec{n} , но и оценку неопределённости этой величины. Поэтому, например, простой метод максимального правдоподобия, в котором мы нашли бы максимум \mathcal{L} в пространстве параметров, не подходит – хотя существуют методы поиска доверительных интервалов. Мы будем использовать метод Монте-Карло с марковскими цепями (Markov Chain Monte Carlo), который хорошо применим к задачам большой размерности. Суть метода сводится к запуску марковского процесса случайного блуждания в пространстве

параметров \vec{n} с такой специально подобранной вероятностью перехода, зависящей от апостериорной вероятности, что в пределе это блуждание даст выборку из исследуемой функции плотности. Общей характеристикой этих методов является то, что они зависят только от отношения вероятностей в разных точках пространства параметров – именно поэтому в выражении 4.7 и во всех последующих вычислениях мы не интересовались ни полной маргинальной вероятностью наблюдаемых данных, ни нормировкой априорного распределения.

Подробный обзор и теоретическое исследование весьма широкой области МСМС-сэмплирования находятся за рамками данной работы, но могут быть найдены, например, в обзоре [18].

Аффинно-инвариантное МСМС-сэмплирование В данной работе будем использовать популярный алгоритм аффинно-инвариантного сэмплирования, предложенный в 2010 году в работе [13].

В этом алгоритме используется не одна марковская цепь (англ. *walker*), а целый ансамбль из нескольких сотен или даже тысяч цепей, и новый шаг генерируется для каждой цепи в зависимости от состояний остальных. Конкретный алгоритм генерации шага (или просто «шаг», англ. *move*) не фиксирован и может быть выбран в зависимости от конкретной задачи. Каждый из шагов обладает своими особенностями и по-своему влияет на предпочтительное количество цепей и необходимое количество итераций. Самый универсальный – так называемый «шаг-растяжка» (англ. *stretch move*), опишем его вкратце.

Рассмотрим i -тую цепь, находящуюся в состоянии $X_i(t)$ (состояние цепи есть точка в пространстве параметров, иначе говоря, конкретное значение \vec{n}). Выберем случайным образом комплементарную цепь из ансамбля и обозначим её $X_j(t)$, $j \neq i$. Сгенерируем «предложение» (англ. *proposal*) для данной цепи:

$$Y = X_j(t) + Z (X_i(t) - X_j(t)) \quad (4.16)$$

Предложенное новое положение i -той цепи находится на прямой, соединяющей её с другой цепью на данном шаге. Масштабный множитель Z – случайная величина, на которую накладывается важное теоретическое ограничение: для существования стационарного распределения марковской цепи достаточно обеспечить детальное равновесие, то есть равенство вероятностей перехода $X_i \rightarrow Y$ и $Y \rightarrow X_i$. Для этого достаточно потребовать, чтобы функция плотности вероятности величины Z

должна удовлетворять свойству симметрии $g(1/z) = z g(z)$. В качестве конкретного вида этой функции авторы предлагают

$$g(z) \propto \begin{cases} \frac{1}{\sqrt{z}} & \text{при } z \in [\frac{1}{a}, a], \\ 0 & \text{в остальных случаях} \end{cases} \quad (4.17)$$

Наконец, после генерации предложения для нового положения i -той цепи (не важно, описанным шагом-растяжкой или любым другим алгоритмом) требуется принять или отвергнуть предложенное значение Y . Вероятность того или иного исхода получается из условия *частичного ресэмплинга*, которое гласит: чтобы переход ансамбля цепей в новое состояние не менял совместного распределения всего ансамбля (которое полагается равным целевому сэмплируемому распределению), достаточно, чтобы переход одной цепи не менял условную плотность вероятности для данной цепи при фиксированных всех остальных. В конце концов, это сложное условие сводится к обычному детальному равновесию, но записанному для всего ансамбля цепей. Вероятность принять предложение Y , обеспечивающая детальное равновесие, задаётся величиной

$$P(X_i(t+1) = Y) = \min \left\{ 1, Z^{N-1} \frac{p(Y)}{p(X_i(t))} \right\} \quad (4.18)$$

Упрощённо можно сказать, что цепи ансамбля будут с большей вероятностью переходить в области большего значения сэмплируемой функции (то есть исследуемой плотности апостериорной распределения), но при этом также чаще оставаться вблизи «соседей» по ансамблю.

Имплементация МСМС-сэмплирования в пакете emcee Эффективная и устойчивая реализация алгоритма МСМС-сэмплирования представляет собой непростую задачу. В приложении **TBD** приведена авторская реализация процедуры в **MATLAB**, однако в дальнейшем будет использоваться популярная имплементация описанного алгоритма в пакете **emcee** на языке **Python** [10].

Пакет предоставляет возможность не только провести сэмплирование, но и оценить корректность результата. Строгое математическое доказательство позволяет установить только *асимптотическую* сходимость ансамбля цепей к целевому распределению. При этом реальная длина цепей конечна, и степень сходимости требует отдельной оценки – иначе

говоря, количество шагов должно быть *достаточно большим* для использования результата.

Но даже если ансамбль сошёлся с хорошей точностью, из него ещё требуется получить выборку независимых реализаций случайной величины. Состояние ансамбля на конкретном (например, последнем) шаге даёт ряд значений, однако число цепей в ансамбле чаще всего меньше, чем желаемое количество элементов выборки – а значит, приходится брать значения на нескольких шагах. Шаги каждой цепи принципиально зависимы, но корреляция падает с расстоянием – следовательно, для получения выборки нужно выбирать *достаточно редкие* состояния ансамбля цепей.

На практике эти две проблемы выражаются в необходимости выбрать время «разгона» (англ. *burn-in*) и «прореживание» (англ. *thinning*) цепей. Информированный выбор обоих параметров можно провести с помощью оценки *интегрированного времени автокорреляции* τ [19]. Интегрированное время автокорреляции даёт оценку числа шагов, на котором корреляция значений цепи падает в e раз. Соответственно, период разгона можно взять равным нескольким τ , и прореживать значения не реже чем каждые τ шагов. Пакет `emcee` предоставляет готовые средства для оценки величины τ ¹.

Более подробное описание конфигурации сэмплера и контроля качества сэмплирования приведено в разделе **TBD**, так как оно зависит от конкретной используемой РИХ.

Помимо метапараметров (количество цепей, параметр a для «шагарастяжки» или соответствующих параметров для других шагов) важной является инициализация сэмплера. Несмотря на теоретически предсказываемое «забывание» состояния цепей в течении нескольких τ , в большинстве практических задач начальное положение цепей желательно задавать разумным образом. Здесь пригодится результат раздела 4.4.2 – простым и быстрым способом можно получить \vec{n}_{rough} – грубую оценку \vec{n} . Следуя практической рекомендации из статьи [10], инициализируем N_{walker} цепей значениями, близкими к этой оценке, но дополнительно разбросанными в соответствии с нормальным распределением числа фотонов в каждом бине. Стандартное отклонение этого начального разброса должно быть заметно меньше разброса самих данных², например

¹ *Autocorrelation analysis & convergence*. 2019. URL: <https://emcee.readthedocs.io/en/stable/tutorials/autocorr/>.

² «Another general approach is to start the walkers in a very tight N-dimensional ball in

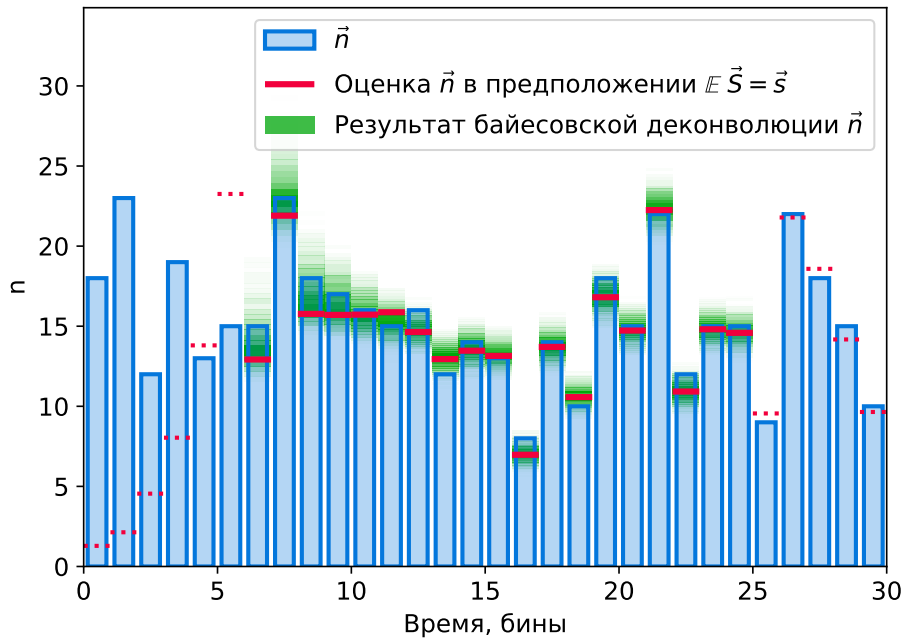


Рис. 4.6: Результат решения задачи байесовской деконволюции. Зелёным показаны *маргинальные апостериорные распределения*, то есть однопараметрические проекции апостериорного распределения как функции n_i для области i , не подверженной влиянию эффекта окна. Следует заметить, что на самом деле многомерное апостериорное распределение не обязательно факторизуемо, поэтому для дальнейшего анализа нельзя рассматривать маргинальные распределения по отдельности – они приведены только для иллюстрации результата.

$\hat{\sigma}_{rough}/10$, где $\hat{\sigma}_{rough}$ – выборочное стандартное отклонение, рассчитанное по элементам \vec{n}_{rough} .

Результат таким образом проведённого сэмплирования для демонстрационной задачи приведён на 4.6.

4.5 Деконволюция экспериментальных данных

Общий метод байесовской деконволюции, изложенный в предыдущем разделе, теперь применим к экспериментальным данным.

В этом разделе будем рассматривать преобразование фотоэлектронов, испущенных фотокатодом ФЭУ под воздействием собранных фотонов, в импульс анодного тока, стекающего с его анодной RC-цепочки. На этом этапе преобразований сигнала определяется и временные характеристики (форма размытия импульса во времени зависит от параметров анодной цепи и собственной формы сигнала ФЭУ), и его амплитуда, причём значение последней носит случайный характер для каждого фотона, однако среднее её значение измеряется в процессе калибровки детектора. При этом, как описано в работе [3], система «ФЭУ + анодная схема + АЦП» является линейной с высокой точностью, и небольшая коррекция нелинейности применяется к данным для компенсации тех небольших искажений, что всё же проявляются.

Детальное описание схем подключения и питания ФЭУ, схему преобразования и усиления сигнала, а также характеристики использованных аналого-цифровых преобразователей можно найти в работе [6].

4.5.1 Экспериментальная рандомизированная импульсная характеристика

Определим случайную функцию $\tilde{h}(t)$, которую будем использовать в качестве рандомизированной импульсной характеристики системы. Сначала факторизуем её следующим образом

$$\tilde{h}(t) = C \tilde{C}_{PMT} h_I(t) \quad (4.19)$$

parameter space around one point that is expected to be close to the maximum probability point.» ([10])

Здесь

1. $h_I(t)$ – функция времени, описывающая *форму* импульсной характеристики, то есть безразмерная и нормированная таким образом, что её интеграл равен 1. Эта функция известна из лабораторного измерения, оно же показывает, что форму можно считать одинаковой для всех 109 каналов мозаики. В этой функции мы феноменологически учитываем все физические процессы, искажающие сигнал во времени: фильтрацию импульса анодной RC-цепочкой, влияние длинной линии, по которой сигнал передаётся с мозаики ФЭУ к бортовому компьютеру, а также любые другие частотно-зависимые характеристики операционных усилителей или АЦП. График $h_I(t)$ изображён на верхней панели рис. 4.7.

2. \tilde{C}_{PMT} – случайная величина, описывающая неопределённость коэффициента усиления ФЭУ. Лавинообразный процесс генерации вторичных электронов сильно зависит от случайных процессов, проходящих в области первого динода, и в зависимости от них коэффициент усиления также оказывается случайной величиной – это и есть источник «рандомизированности» импульсной характеристики. \tilde{C}_{PMT} безразмерна и масштабирована так, чтобы $\mathbb{E}\tilde{C}_{PMT} = 1$, плотность и функция распределения показана на нижней панели рис. 4.7. **Как получена – описать процедуру общими словами, сослаться на приложение или private communication**

3. C – масштабный коэффициент, размерность совпадает с размерностью сигнала $s(t)$. Также можно сказать, что с точки зрения экспериментальных данных это *средний абсолютный коэффициент усиления* данного канала. Для иллюстрации этого заметим, что для однородного во времени шумового сигнала в длинном окне ($N \rightarrow \infty$)

$$\begin{aligned}
\mathbb{E} \left[\int_1^N S(t) dt \right] &= \mathbb{E} \left[\sum_{i=1}^N \left(n_i \int_{\infty} h(t) dt \right) \right] = \mathbb{E} \left[\int_{\infty} h(t) dt \right] \sum_{i=1}^N n_i = C \sum_{i=1}^N n_i \\
\frac{1}{N} \mathbb{E} \left[\int S(t) dt \right] &= C \frac{1}{N} \sum_{i=1}^N n_i \\
\bar{S} &= C \bar{n}
\end{aligned} \tag{4.20}$$

Иначе говоря, C – коэффициент пропорциональности между средним значением числа фотоэлектронов в бине и средним значением выходного сигнала. Строго говоря, \bar{S} – среднее значение *интеграла* выходного сигнала за один бин, однако численно они совпадают из-за принятой временной шкалы; различие будет проявляться только в размерности: например, если S выражает анодный ток и измеряется в микроамперах, то C будет иметь размерность заряда (микроампер * бин).

Постоянный коэффициент калибровки C не влияет на статистические свойства системы, и пока можно считать равным единице, его роль будет описана в разделе **TBD**.

Получив РИХ, можно смоделировать отклик установки на определённый входной \vec{n} и провести деконволюцию, чтобы оценить корректность работы метода, аналогично тому, как это было проделано с «игрушечной» РИХ (4.1). На рис. 4.8 приведён результат такой процедуры. По сравнению с решением «игрушечной» задачи, приведённым на рис. 4.6, очевидна худшая точность деконволюции, что вполне объяснимо – в первом случае РИХ быстро спадала во времени, была ограничена всего $L_1 = 10$ отсчётами, а её амплитуда варьировалась от 0.75 до 1; реальная РИХ, напротив, имеет широкий основной пик (около 5 отсчётов) и длинный хвост, а её амплитуда варьируется в пределах от 0 до 2-3 относительно среднего значения. Однако в этом и состоит главное достоинство статистической деконволюции – неполнота информации не ведёт к некорректности метода (что, вероятнее всего, случилось бы в случае с обыкновенной, детерминистичной деконволюцией, например, основанной на вычислении частного от фурье-образов), но предсказуемым образом отражается в результатах.

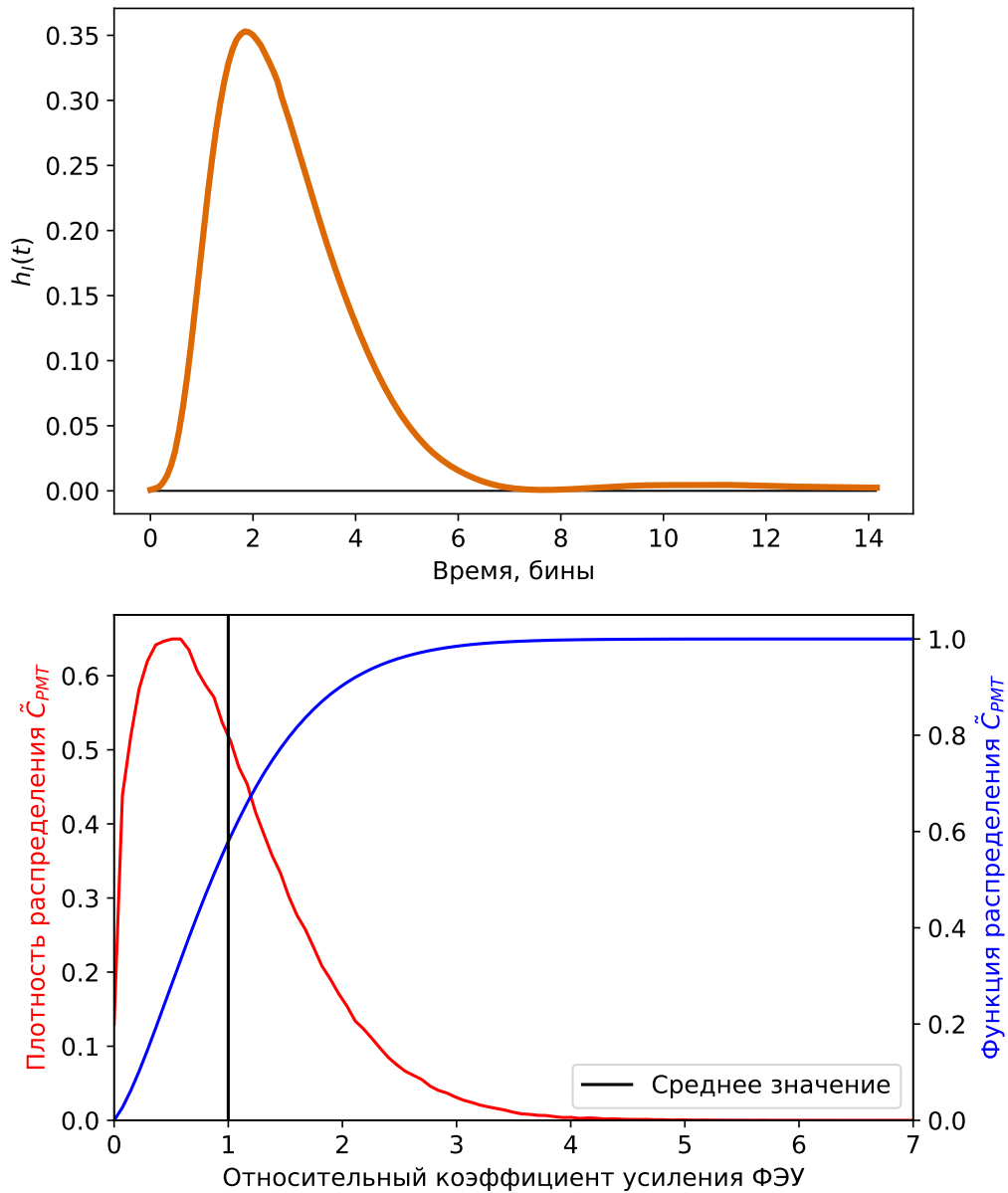


Рис. 4.7: Компоненты экспериментальной рандомизированной импульсной характеристики. Функция $h_I(t)$ (верхняя панель) описывает форму импульса во времени, а случайная величина $\tilde{C}_{РМТ}$ (функции распределения и плотности вероятности на нижней панели) – неопределённость коэффициента усиления ФЭУ.

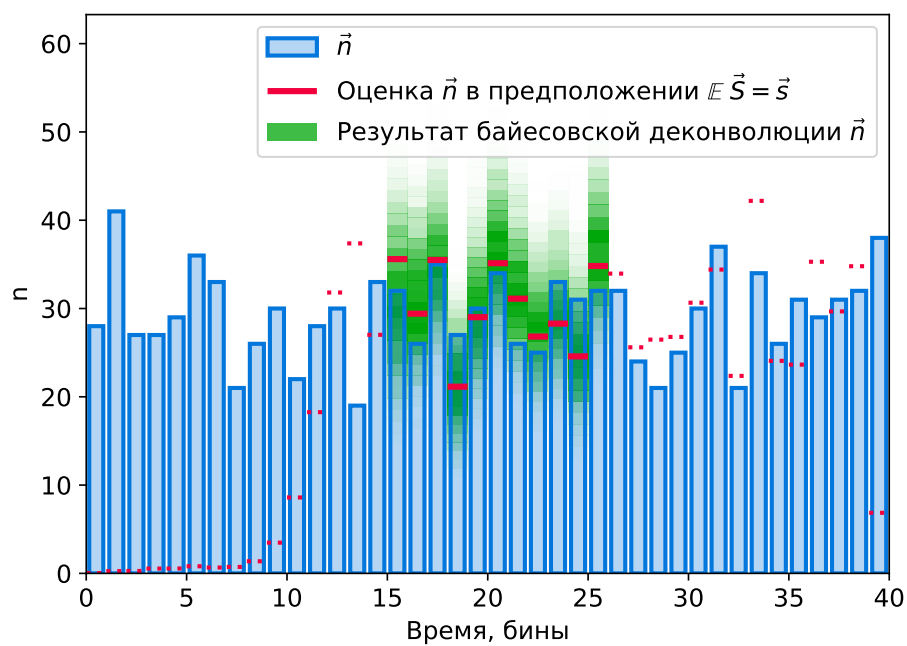


Рис. 4.8: Решение задачи байесовской деконволюции для РИХ, определённой выражением (4.19). Синие столбцы – истинный входной сигнал \vec{n} , розовые полосы – первичная грубая оценка, зелёным изображены маргинальные апостериорные распределения в центральных бинах (см. описание рис. 4.6).

4.5.2 Модификация процедуры для применения к реальным данным

Перед тем как, наконец, перейти от исследования метода к деконволюции реальных данных, необходимо внести две небольшие модификации в описанную процедуру.

4.5.2.1 Учёт неопределённости экспериментального сигнала

Функция правдоподобия, введённая в разделе 4.4.3, и её аппроксимация, описанная в подразделе 4.4.3.3, справедливы для случая, когда сигнал на выходе известен точно. В этом случае не совсем корректно определение правдоподобия как *вероятности* получить данный сигнал \vec{s} при фиксированных \vec{n} – вероятность каждой конкретную реализации непрерывного распределения равна нулю. Поэтому до сих пор под правдоподобием на самом имела в виду *плотность* этой вероятности.

Однако сигнал в реальном эксперименте известен с погрешностью – в первую очередь это погрешность дискретизации АЦП. Поэтому требуется модифицировать функцию правдоподобия так, чтобы она показывала не плотность вероятности единичного значения \vec{s} , а вероятность нахождения \vec{s} в некотором наборе состояний. В рамках байесовской интерпретации вероятности будем считать, что неопределённость измерения сигнала выражается некоторым апостериорным распределением (апостериорным относительно процедуры его экспериментального измерения) $p_{exp}(\vec{s})$. Тогда, обозначая плотность вероятности (4.8) $p(\vec{s}, \vec{n})$, можно записать модифицированную функцию правдоподобия в виде

$$\mathcal{L}(\vec{n}) = \int_{\infty} p_{exp}(\vec{s}) p(\vec{s}, \vec{n}) d\vec{s} \quad (4.21)$$

В эксперименте ошибка дискретизации АЦП превосходит другие источники ошибки как минимум на порядок (см. таблицу 4.1). Поэтому можно считать, что, если АЦП округляет сигнал вниз и имеет величину дискретизации δ , то в i -том отсчёте истинное значение s_i равномерно распределено между записанным округлённым значением \bar{s}_i и следующей «ступенькой», отстоящей от него вверх на δ , и распределено независимо от других отсчётов. Иначе говоря, $p_{exp}(\vec{s}) = \prod_i U(\bar{s}_i, \bar{s}_i + \delta_s)$. Это сводит вычисление интеграла (4.21) к более простому

$$\mathcal{L}(\vec{n}) = \int_{s_1}^{s_1+\delta} \int_{s_2}^{s_2+\delta} \dots \int_{s_N}^{s_N+\delta} p(\vec{s}, \vec{n}) d\vec{s} \quad (4.22)$$

Задача вычисления функции правдоподобия сводится к интегрированию плотности вероятности многомерного нормального распределения по многомерному кубу – и всё ещё представляет собой отдельную вычислительную задачу. К счастью, для её решения можно воспользоваться готовым алгоритмом адаптивного интегрирования, описанным в работе [12], имплементация которого доступна в пакете `scipy.stats` [20] в виде **Fortran**-сабрутины с интерфейсом для **Python**.

На рис. 4.9 представлена постановка и решение задачи в этой форме, предельно близкой к задаче деконволюции экспериментальных сигналов. Величина шага дискретизации АЦП положена равной 2, а к исходному сигналу добавлена простейшая имитация сигнала ШАЛ: к шумовым фотонам количество которых, как и раньше, выбиралось из пуассоновского распределения с $\lambda_n = 20$, в трёх центральных бинах добавлено по $2\lambda_n$ «сигнальных» фотонов.

4.5.2.2 Каскадное сэмплирование

Полученная функция правдоподобия (4.22) всё ещё требует существенного времени для вычисления. Большая часть вычислений функции правдоподобия приходится на период «разгона» сэмплера и набора достаточного количества итераций для устойчивой оценки времени автокорреляции τ (см. параграф 4.4.3.4). Этот процесс можно ускорить, если инициализировать сэмплер не точками в области грубой оценки \vec{n} , но некоторым промежуточным распределением, близким к конечному, но более простым в получении. Такое промежуточное распределение можно получить, рассмотрев функцию правдоподобия в приближении *факторизованного* многомерного нормального распределения.

Если бы корреляции между элементами случайного вектора \vec{S} при заданных \vec{n} были бы пренебрежимо малы (то есть попарные распределения на рис. 4.5 имели форму окружностей), вычисление функции правдоподобия даже с учётом ошибок измерения сигнала было бы гораздо проще. А именно, оно свелось бы к перемножению вероятностей получить наблюдаемое значение *в каждом отсчёте*, которые в свою очередь рассчитывались бы из функции одномерного нормального распределения. С вычислительной точки зрения такой расчёт оказывается примерно в

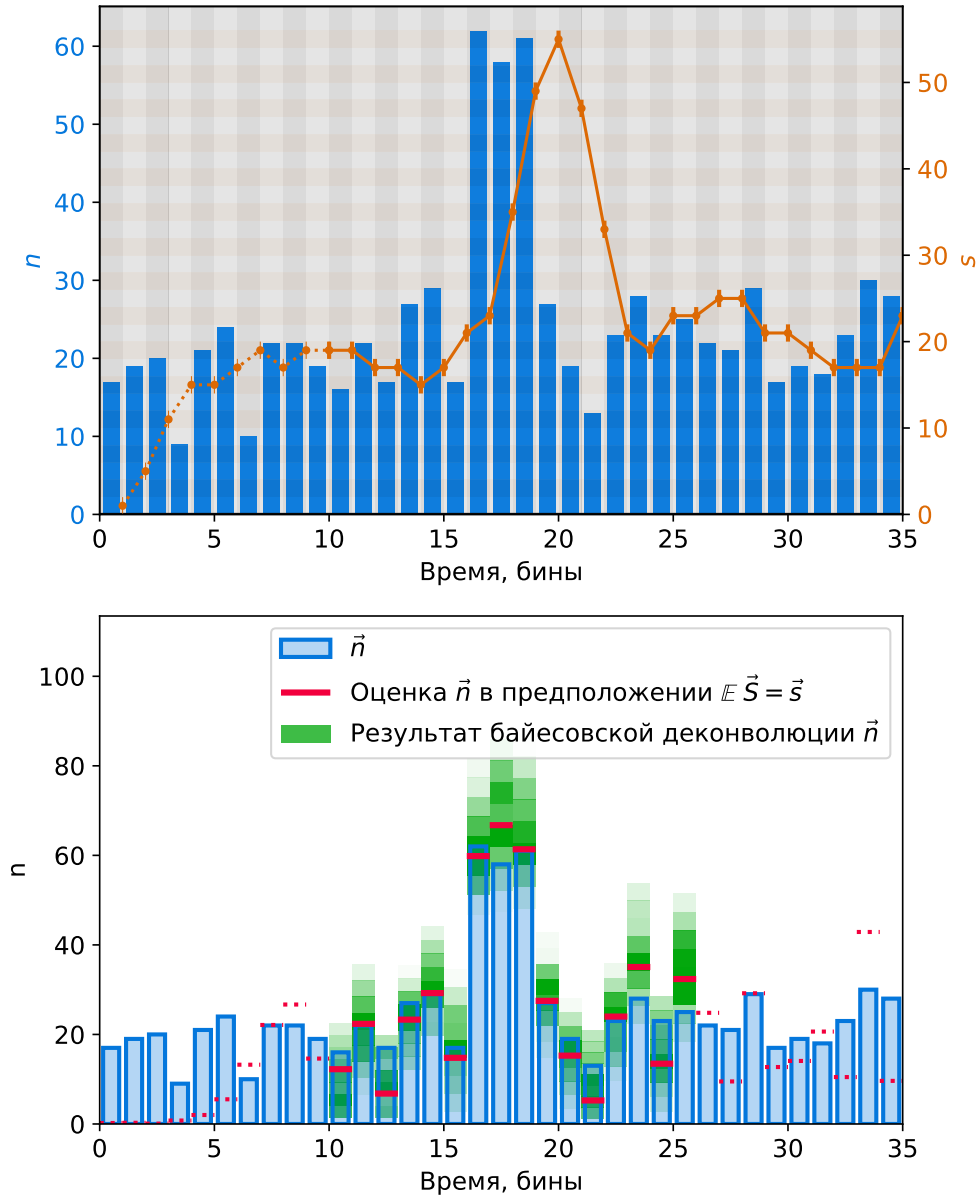


Рис. 4.9: Постановка и решение задачи байесовской деконволюции для экспериментальной РИХ с учётом погрешности измерения сигнала. В центре окна добавлена имитация пакета фотонов ШАЛ, двукратно превышающая фон в течение 3 бинов.

100 раз эффективнее, а получающееся в результате сэмплирования этой упрощенной функции правдоподобия апостериорное распределение качественно подобно истинному.

Всё это позволяет ввести процедуру каскадного сэмплирования:

1. Сгенерировать длинные марковские цепи с упрощённой и вычислительно дешёвой функцией правдоподобия, тем самым получить выборку, близкую к апостериорному распределению, а также оценку времени автокорреляции τ .
2. «Подменить» функцию правдоподобия на истинную, и провести сравнительно короткое финальное сэмплирование, которое сойдётся уже к искомому апостериорному распределению

На практике оказывается достаточным провести предварительное сэмплирование на 20000 шагов с 256 цепями, и затем окончательное – всего $4000 \cong 2\tau$ шагов с 128 цепями. На рис. 4.10 приведено сравнение маргинальных распределений в отдельно взятом бине после предварительного и окончательного сэмплирования. Видно постепенное уточнение оценки. С точки зрения времени вычислений предварительное сэмплирование занимает примерно в 30-50 раз меньше времени, несмотря на большее число шагов и цепей.

4.5.3 Подготовка экспериментальных данных

Данные эксперимента СФЕРА-2 представлены тремя сущностями, из которых нам предстоит сконструировать искомые сигналы анодного тока. Итак, для каждого из 109 каналов известны [6]:

1. \tilde{s} – переменная компонента анодного тока в области триггера, усиленная и записанная АЦП, сохранённая в *единицах кода*. Погрешность дискретизации в этих величинах составляет 1 по определению единицы кода. Также присутствует дополнительная погрешность, связанная с тем, что для перевода сигнала в динамический диапазон АЦП к нему искусственно добавлялся постоянный ток («пьедестал»). Далее из записанных событий этот пьедестал удалялся вычитанием среднего значения по выборке из 200 шумовых

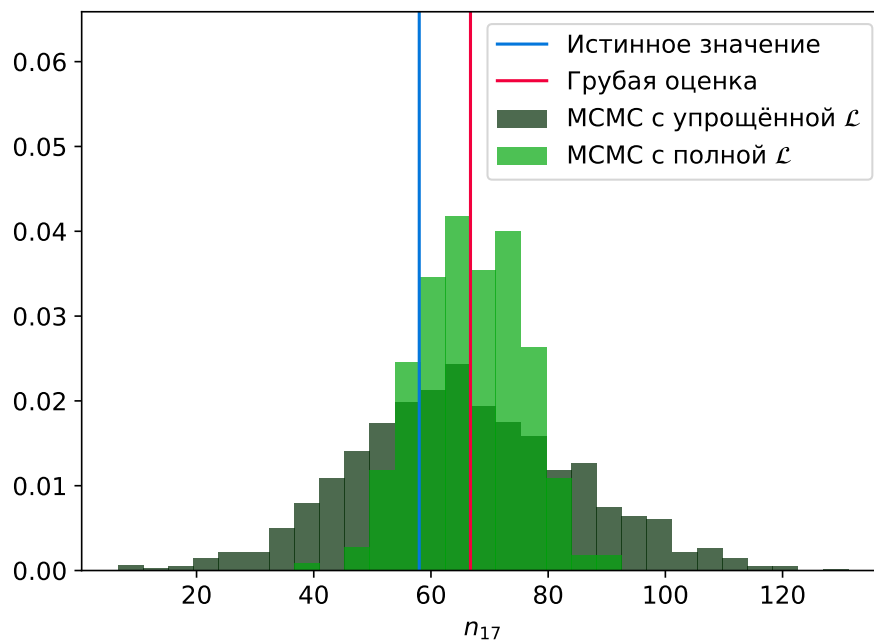


Рис. 4.10: Сравнение маргинальных апостериорных распределений в отдельно взятом бине, полученных с упрощённой функцией правдоподобия (предварительный расчёт, «разгон» сэмплера) и с полной функцией, дающий окончательный результат.

отсчётов. Следовательно, если дисперсия шумовых отсчётов составляет порядка нескольких единиц кода, то погрешность, вносимая процедурой вычитания пьедестала, рассчитывается по формуле для дисперсии выборочного среднего и составит порядка $1/\sqrt{200} \approx 0.07$ единиц кода. Поэтому этой погрешностью можно пренебречь и использовать выражение (4.22) для функции правдоподобия.

2. \bar{s} – постоянная компонента анодного тока, отфильтрованная RC-цепочкой и записанная отдельным АЦП. Показания этого АЦП сразу пересчитывались к значению в микроамперах, погрешность его составляет $0.1 \mu A$, типичные значение – от 7 до $15 \mu A$ для разных каналов, то есть погрешность измерения этой величины также на 2 порядка меньше, чем погрешность дискретизации переменной части сигнала.
3. A – коэффициент *относительной* калибровки канала. Имеет следующий смысл: домножая \tilde{s} и \bar{s} на A , мы получаем сигнал, который записал бы референсный канал – ФЭУ Hamamatsu R3886 при заданных температуре и давлении (**каких?**). Величина A измеряется для всех каналов экспериментально с помощью специального калибровочного кадра, а также паспортных зависимостей коэффициента усиления от температуры и давления для центрального ФЭУ, этот процесс подробно описан в работе [5].

Для референсного канала известно (**откуда?**) значение параметра C в выражениях (4.19 - 4.20) при условии, что сигнал измеряется в единицах кода: $C_{ref} = 0.326$ ед. кода * бин. Таким образом, для произвольного канала $C = C_{ref}/A$.

Остаётся пересчитать \bar{s} из микроамперов в единицы кода, чтобы можно было восстановить постоянную компоненту сигнала так, как она была бы записана АЦП в отсутствии разделения постоянной и переменной компоненты. Для этого необходимо: 1) разделить эту величину на 2 (разделение тока на согласованном сопротивлении кабеля, по которому импульс передаётся от мозаики к АЦП) 2) ещё раз разделить на 2 (разделение тока на два идентичных канала АЦП, в сумме дающих дискретизацию 12.5 нс, см. [6]) 3) умножить на 30 (коэффициент усиления операционного усилителя) 4) преобразовать в падение напряжения на 100Ω резисторе 5) разделить на шаг дискретизации, который

Источник ошибки	Распределение	Характерная величина [ед. кода]
Дискретизация АЦП	U	1
Измерение \bar{s}	N	$0.38 \cdot 0.1 \approx 0.03$
Перевод \bar{s} в ед. кода	N	≈ 0.08
Вычитание пьедестала \tilde{s}	N	≈ 0.07

Таблица 4.1: Сравнение разных источников погрешностей измерения экспериментального сигнала.

для 10-битного АЦП с рабочим диапазоном напряжений 2 В будет равен $2 \text{ В} / 2^{10} \approx 1.95 \text{ мВ}$. В итоге получим

$$\bar{s} [\text{ед. кода}] = \frac{\left(\frac{1}{2} \cdot \frac{1}{2} \cdot 30 \cdot \bar{s}\right) [\text{мкА}] \cdot 100 [\Omega]}{1.95 [\text{мВ/ед. кода}]} \approx 0.38 [\text{ед. кода/мкА}] \cdot \bar{s} [\text{мкА}] \quad (4.23)$$

Дополнительная ошибка, вносимая при этом пересчёте, составляет порядка 1 – 2% – она вызвана неопределённостью параметров цепи, коэффициента усиления операционного усилителя, диапазона работы АЦП, и т.д. Разные источники ошибки измерения сигнала приведены в таблице 4.1. Видно, что доминирующей ошибкой является ошибка дискретизации, все прочие источники неопределённости составляют в сумме около 11% от её величины. Поэтому в дальнейшем, как было описано в разделе 4.5.2.1, будем учитывать только ошибку дискретизации АЦП.

Таким образом, финальная процедура «сборки» сигнала из экспериментальных данных выглядит так:

1. Перевести \bar{s} из мкА в единицы кода домножением на 0.38
2. Сложить \tilde{s} и \bar{s} , получив полный сигнал s в единицах кода, как он был бы записан АЦП в отсутствии разделения на постоянную и переменную компоненты
3. Вычислить масштабный множитель РИХ C для данного канала по формуле C_{ref}/A , где A – коэффициент относительной калибровки

Вообще говоря, деконволюцию можно было бы проводить уже на этом этапе, однако тогда для каждого сигнала придётся использовать свою

РИХ. Заметим, что ситуацию можно облегчить, если добавить ещё один шаг:

4. Привести сигнал s и ошибку дискретизации к относительным единицам, разделив на C

В этих относительных единицах все каналы имеют одну РИХ; все их индивидуальные характеристики оказываются выражены в величине ошибки дискретизации – канал с большим коэффициентом усиления будет давать больший сигнал в единицах кода и как следствие иметь меньшую ошибку округления в приведённом относительном масштабе, и наоборот. Кроме того, как следует из выражения (4.20), в таком масштабе *среднее значение сигнала равно среднему количеству фотоэлектронов в бине.*

Пример экспериментального кадра после всех описанных преобразований приведён на рис. 4.11. Наконец, к полученному сигналу можно применить процедуру деконволюции. Её результат приведён на рис. 4.12.

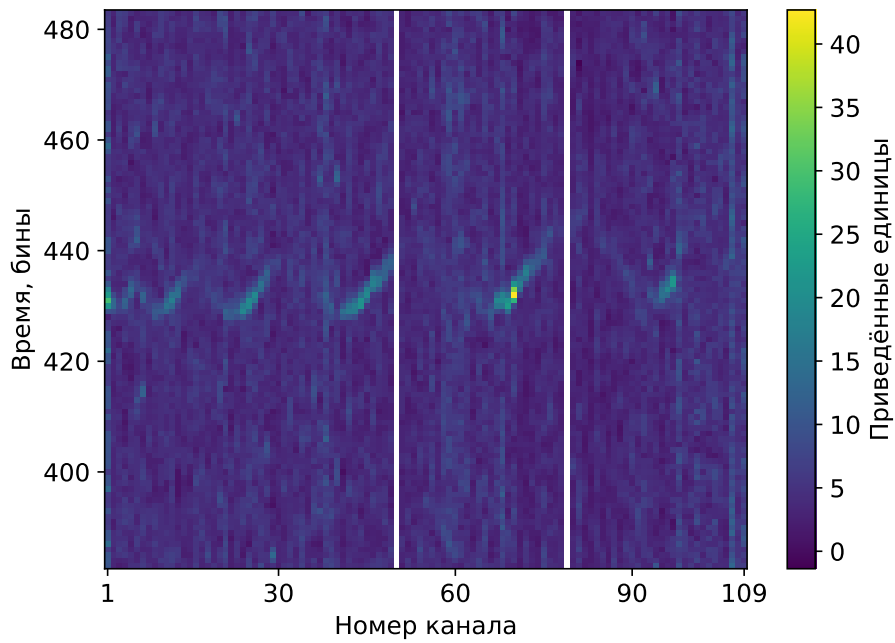


Рис. 4.11: Пример экспериментального события в приведённом масштабе (см. текст). Видна характерная «змейка» фронта черенковского света в центре кадра, и общая однородная яркость фона – её систематическое изменение от канала к каналу объясняется разными коэффициентами сбора света с поверхности для каждого ФЭУ и, как следствие, немного разными интенсивностями фоновых фотонов. Исключены каналы 50 и 79, которые не работали из-за сбоя электроники.

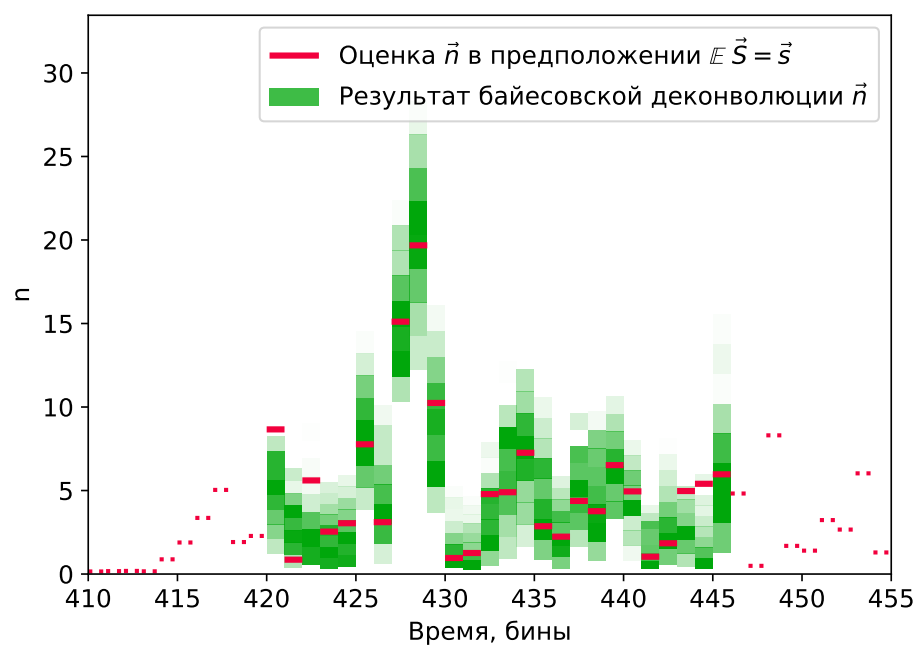


Рис. 4.12: Результат байесовской деконволюции для 2 канала события, изображённого на рис. 4.11.

Глава 5

Оценка параметров ШАЛ

В предыдущей главе описан метод байесовской деконволюции и его применение к данным эксперимента СФЕРА-2. В результате получена безмодельная оценка потоков фотонов на мозаике ФЭУ – безмодельная в том смысле, что она не зависит от предположений о свойствах и источниках света, падающего на мозаику. В этой главе на основе полученного результата, а также качественного представления о пространственно-временной структуре сигнала ШАЛ реконструируется функция пространственного распределения (ФПР) черенковского света. Наконец, на основе полученной ФПР делаются оценки параметров ШАЛ.

Список литературы

- [1] A. M. Anokhina и др. «Method for mass analysis of primary cosmic ray particles with the SPHERE-2 system». в: *Bulletin of the Russian Academy of Sciences: Physics* 71.4 (апр. 2007), с. 506—509. DOI: 10.3103/s106287380704020x.
- [2] R A Antonov и др. «Results on the primary CR spectrum and composition reconstructed with the SPHERE-2 detector». в: *Journal of Physics: Conference Series* 409 (февр. 2013), с. 012088. DOI: 10.1088/1742-6596/409/1/012088.
- [3] R. A. Antonov и др. «Detection of reflected Cherenkov light from extensive air showers in the SPHERE experiment as a method of studying superhigh energy cosmic rays». в: *Physics of Particles and Nuclei* 46.1 (янв. 2015), с. 60—93. DOI: 10.1134/s1063779615010025.
- [4] R.A. Antonov и др. «Spatial and temporal structure of EAS reflected Cherenkov light signal». в: *Astroparticle Physics* 108 (март 2019), с. 24—39. DOI: 10.1016/j.astropartphys.2019.01.002.
- [5] R.A. Antonov и др. «The LED calibration system of the SPHERE-2 detector». в: *Astroparticle Physics* 77 (апр. 2016), с. 55—65. DOI: 10.1016/j.astropartphys.2016.01.004.
- [6] R.A. Antonov и др. «The SPHERE-2 detector for observation of extensive air showers in 1 PeV – 1 EeV energy range». в: *Astroparticle Physics* 121 (сент. 2020), с. 102460. DOI: 10.1016/j.astropartphys.2020.102460.
- [7] *Autocorrelation analysis & convergence*. 2019. URL: <https://emcee.readthedocs.io/en/stable/tutorials/autocorr/>.

- [8] Dmitry Chernov и др. «Cosmic ray study by means of reflected EAS Cherenkov light method with the SPHERE-2 detector». в: *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Sissa Medialab, авг. 2017. DOI: 10.22323/1.301.0537.
- [9] Bruce R. Dawson и др. «The energy spectrum of cosmic rays at the highest energies». в: *EPJ Web of Conferences* 53 (2013). под ред. К.-Н. Kampert и др., с. 01005. DOI: 10.1051/epjconf/20135301005.
- [10] Daniel Foreman-Mackey. «corner.py: Scatterplot matrices in Python». в: *The Journal of Open Source Software* 1.2 (июнь 2016), с. 24. DOI: 10.21105/joss.00024.
- [11] Andrew Gelman и др. *Bayesian Data Analysis*. Chapman и Hall/CRC, нояб. 2013. DOI: 10.1201/b16018.
- [12] Alan Genz. «Numerical Computation of Multivariate Normal Probabilities». в: *Journal of Computational and Graphical Statistics* 1.2 (июнь 1992), с. 141—149. DOI: 10.2307/1390838.
- [13] Jonathan Goodman и Jonathan Weare. «Ensemble samplers with affine invariance». в: *Communications in Applied Mathematics and Computational Science* 5.1 (янв. 2010), с. 65—80. DOI: 10.2140/camcos.2010.5.65.
- [14] Charles R. Harris и др. «Array programming with NumPy». в: *Nature* 585.7825 (сент. 2020), с. 357—362. DOI: 10.1038/s41586-020-2649-2. URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- [15] D. Heck и др. *CORSIKA: a Monte Carlo code to simulate extensive air showers*. 1998.
- [16] R. Penrose. «On best approximate solutions of linear matrix equations». в: *Mathematical Proceedings of the Cambridge Philosophical Society* 52.1 (янв. 1956), с. 17—19. DOI: 10.1017/s0305004100030929.
- [17] Cynthia Rhode и Scott Whittenburg. «Bayesian Deconvolution». в: *Spectroscopy Letters* 26.6 (июль 1993), с. 1085—1102. DOI: 10.1080/00387019308011596.
- [18] Sanjib Sharma. «Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy». в: *Annual Review of Astronomy and Astrophysics* 55.1 (авг. 2017), с. 213—259. DOI: 10.1146/annurev-astro-082214-122339.

- [19] A. Sokal. «Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms». в: *Functional Integration*. Springer US, 1997, с. 131—192. DOI: 10.1007/978-1-4899-0319-8_6.
- [20] Pauli Virtanen и др. «SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python». в: *Nature Methods* 17 (2020), с. 261—272. DOI: 10.1038/s41592-019-0686-2.
- [21] Stephen G. Warren. «Optical properties of snow». в: *Reviews of Geophysics* 20.1 (1982), с. 67. DOI: 10.1029/rg020i001p00067.
- [22] David Wipf и Haichao Zhang. «Revisiting Bayesian Blind Deconvolution». в: *arXiv e-prints*, arXiv:1305.2362 (май 2013), arXiv:1305.2362. arXiv: 1305.2362 [cs.CV].
- [23] Д. В. Чернов и др. «Регистрация отражённого черенковского света ШАЛ в эксперименте «СФЕРА» как метод изучения космических лучей сверхвысоких энергий». в: *Физика элементарных частиц и атомного ядра* 46.1 (2015), с. 115—166. ISSN: 0367-2026.