

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

УДК 53.08; 519.25; 519.876.5

Скакун  
Виктор Васильевич

МОДЕЛИРОВАНИЕ И АНАЛИЗ ПРОЦЕССОВ РЕГИСТРАЦИИ И  
ОБРАБОТКИ СЛУЧАЙНЫХ ПОТОКОВ СОБЫТИЙ В СИСТЕМАХ  
ФЛУОРЕСЦЕНТНОЙ ФЛУКТУАЦИОННОЙ СПЕКТРОСКОПИИ

Диссертация на соискание ученой степени  
кандидата физико-математических наук

по специальности 05.13.01 –  
системный анализ, управление и обработка информации

Научный руководитель доктор  
физико-математических наук  
профессор Апанасович В.В.

Минск, 2009

## ОГЛАВЛЕНИЕ

ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ И СОКРАЩЕНИЙ .....	4
ВВЕДЕНИЕ.....	5
ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ.....	8
ГЛАВА 1 РЕГИСТРАЦИЯ И АНАЛИЗ СЛУЧАЙНЫХ ПОТОКОВ В ФЛУОРЕСЦЕНТНОЙ ФЛУКТУАЦИОННОЙ СПЕКТРОСКОПИИ .....	11
1.1 Флуоресцентная флукуационная спектроскопия.....	11
1.2 Поток случайных событий.....	16
1.3 Преобразование случайных потоков .....	21
1.3.1 Простейшие преобразования .....	22
1.3.2 Регистрация случайных потоков системами с мертвым временем ..	24
1.3.3 Преобразование потоков в системе с последствием .....	27
1.4 Методы анализа данных флукуационной спектроскопии .....	28
1.4.1 Флуоресцентная корреляционная спектроскопия .....	28
1.4.2 Анализ распределения числа фотоотсчетов.....	31
1.4.3 Анализ кумулянтов распределения числа фотоотсчетов.....	36
1.5 Выводы.....	37
ГЛАВА 2 МОДЕЛИРОВАНИЕ СИСТЕМ РЕГИСТРАЦИИ И ОБРАБОТКИ СЛУЧАЙНЫХ ПОТОКОВ.....	39
2.1 Аналитическое моделирование систем преобразования случайных потоков .....	39
2.1.1 Последовательное преобразование случайных потоков.....	39
2.1.2 Слияние потоков .....	40
2.1.3 Разветвление потоков .....	41
2.2 Имитационное моделирование систем регистрации случайных потоков ..	43
2.2.1 Структура имитационной модели .....	43
2.2.2 Программная реализация имитационной модели.....	47
2.2.3 Вычислительный эксперимент .....	52
2.3 Имитационное моделирование систем флукуационной спектроскопии ..	54
2.3.1 Моделирование потока фотонов с заданной автокорреляционной функцией.....	54
2.3.2 Моделирование фотоотсчетов .....	57
2.4 Обработка и хранение данных во флуоресцентной спектроскопии.....	64
2.5 Основные результаты и выводы.....	71
ГЛАВА 3 АНАЛИЗ РАСПРЕДЕЛЕНИЯ ЧИСЛА ФОТООТСЧЕТОВ.....	72
3.1 Производящая функция числа фотоотсчетов с коррекцией профиля засветки .....	72
3.2 Сравнительный анализ методов анализа распределения числа фотоотсчетов .....	75
3.3 Генерация начальных приближений для метода FIDA.....	76
3.3.1 Метод генерации начальных приближений .....	77
3.3.2 Тестирование начальных приближений .....	81
3.4 Анализ распределения числа фотоотсчетов с коррекцией профиля засветки .....	90

3.4.1	Нормализация в методах анализа распределения числа фотоотсчетов .....	90
3.4.2	Метод анализа распределения числа фотоотсчетов с коррекцией профиля засветки .....	91
3.4.3	Тестирование метода .....	94
3.5	Основные результаты и выводы.....	97
ГЛАВА 4	КУМУЛЯНТНЫЙ АНАЛИЗ ФОТООТСЧЕТОВ.....	98
4.1	Кумулянтный анализ фотоотсчетов с коррекцией профиля засветки.....	98
4.1.1	Коррекция профиля засветки в кумулянтном анализе.....	98
4.1.2	Полиномиальная аппроксимация профиля засветки.....	100
4.2	Оценки яркости и количества молекул для различных видов аппроксимации профиля засветки и нормализации.....	100
4.3	Начальные приближения для кумулянтного анализа .....	102
4.4	Результаты тестирования кумулянтного анализа.....	103
4.5	Основные результаты и выводы.....	112
ЗАКЛЮЧЕНИЕ	.....	114
Основные научные результаты диссертации .....	114	
Рекомендации по практическому использованию результатов .....	115	
БИБЛИОГРАФИЧЕСКИЙ СПИСОК .....	116	
Список использованных источников.....	116	
Список публикаций соискателя.....	125	
ПРИЛОЖЕНИЕ А	Примеры применения выражений, описывающих последовательное преобразование потоков .....	128
ПРИЛОЖЕНИЕ В	Расчет факториальных моментов и факториальных кумулянтов .....	130
ПРИЛОЖЕНИЕ С	Решение систем (3.26), (3.28).....	131
ПРИЛОЖЕНИЕ D	Решение системы (3.32).....	132
ПРИЛОЖЕНИЕ E	Сведение трехмерного интеграла в (3.4) к одномерному.....	133
ПРИЛОЖЕНИЕ F	Применение нормализации вида (1.84) в методе РСН.....	134
ПРИЛОЖЕНИЕ G	Структуры данных БД измерений и результатов анализа во время разрешимой флуоресцентной спектроскопии.....	135

## ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ И СОКРАЩЕНИЙ

FCA	–	анализ кумулянтов флуоресценции
FCS	–	флуоресцентная корреляционная спектроскопия
FIDA	–	анализ распределения интенсивности флуоресценции
IFT	–	обратное преобразование Фурье
PCN	–	анализ гистограммы числа фотоотсчетов
S/N	–	отношение сигнал/шум
АКФ	–	автокорреляционная функция
БД	–	база данных
БП	–	блок преобразования
ЛФД	–	лавинный фотодиод
МНК	–	метод наименьших квадратов
НП	–	начальные приближения
ПФ	–	производящая функция
ПФЛ	–	производящий функционал
ПФЧФ	–	производящая функция числа фотоотсчетов
РЧФ	–	распределение числа фотоотсчетов
СРОД	–	система регистрации и обработки данных
СУБД	–	системы управления базами данных
УПФЛ	–	условный производящий функционал
УПФЛ1	–	условный производящий функционал первого рода
УПФЛ2	–	условный производящий функционал второго рода
ФК	–	факториальные кумулянты
ФФС	–	флуоресцентная флуктуационная спектроскопия
$\langle C \rangle$	–	концентрация молекул
$B(\mathbf{r})$	–	функция профиля засветки
$F(t)$	–	интенсивность флуоресценции
$G(\xi)$	–	производящая функция
$g(\tau)$	–	автокорреляционная функция потока фотонов
$K_k$	–	факториальные кумулянты
$L[u, \Omega]$	–	производящий функционал
$N$	–	среднее число молекул в объеме наблюдения
$P(n)$	–	распределение числа фотоотсчетов
$q$	–	характеристическая яркость молекулы
$T$	–	интервал наблюдения
$V$	–	объем наблюдения
$\chi^2$	–	критерий хи-квадрат
$\chi_k$	–	интеграл от функции профиля засветки в степени $k$

## ВВЕДЕНИЕ

В различных областях физики, химии и биологии при изучении свойств веществ, а также при исследовании особенностей протекания химических реакций возникает необходимость в измерении таких характеристик, как концентрация, коэффициент диффузии, скорость переноса энергии и т.д. Экспериментальная физика предлагает для этих целей ряд методов, среди которых заметное место отводится методам, использующим возбуждение вещества с последующим наблюдением за процессом его возвращения в равновесное состояние. Большинство из них сопровождается нарушением термодинамического равновесия, что в некоторых случаях является недопустимым, так как при этом вещество может существенно поменять свои свойства. Более того, такие методы неприменимы для исследования живых клеток.

Для изучения свойств вещества без нарушения термодинамического равновесия наиболее подходящими являются методы одноквантовой флуоресцентной спектроскопии, которые находят все более широкое применение в биофизических, биохимических и медицинских исследованиях. Достоинством этих методов является неинвазивность, высокая чувствительность, а также возможность исследовать процессы в наносекундном диапазоне. За такой промежуток времени происходит ряд молекулярных процессов, способных повлиять на спектральные характеристики флуоресцирующего соединения.

Особый интерес представляют методы флуоресцентной флуктуационной спектроскопии (ФФС), позволяющие исследовать вещество на одномолекулярном уровне. Методы ФФС применяются для исследования динамических и структурных характеристик вещества через регистрацию флуктуаций интенсивности флуоресценции, вызванных изменением числа флуоресцентных молекул в освещенном объеме или (и) их квантового выхода [1]. В этих методах регистрация флуоресценции производится из предельно малого открытого объема (менее одного фемтолитра), определяемого лучом лазера, сфокусированного оптической системой с высоким коэффициентом увеличения ( $1.2NA$  и более) [2, 3].

Для реализации методов ФФС требуется создание совершенных систем регистрации, способных детектировать фотоны от отдельных молекул. Поток фотонов, несущий информацию об исследуемом явлении, подвергается различным преобразованиям, сопровождающим процесс детектирования, регистрации и обработки исходного потока. Преобразование потока обычно связано с прохождением его через некоторые устройства, обладающие линейными или нелинейными характеристиками относительно числа событий потока и их координат, и имеет в общем случае случайный характер.

Основным элементом формализации при создании моделей процессов, происходящих в системе регистрации, выступают потоки случайных событий. Для

полного описания процессов преобразования потока фотонов в системе регистрации предлагается использовать положения теории случайных потоков событий и математический аппарат производящих функционалов [4, 5, 6]. Суть его применения заключается в получении функциональных уравнений, связывающих производящие функционалы исходного и результирующего потока. Найденные уравнения с помощью операций функционального дифференцирования приводятся к обычным алгебраическим, интегральным уравнениям, содержащим искомые характеристики исследуемого потока [4]. В ряде случаев более удобным является применение производящих функций [5, 7].

Системы регистрации, применяемые в ФФС, представляют собой сложно-технические комплексы устройств, каждый из которых может влиять на процесс преобразования потока фотонов. Аналитический аппарат моделирования позволяет описывать отдельные виды преобразований случайных потоков. С его помощью не всегда удается получить аналитически разрешимые модели последовательности нескольких различных типов преобразования исходного потока. В такой ситуации возможным выходом является применение имитационного моделирования. Имитационное моделирование позволяет изучать и адекватно воспроизводить процессы функционирования систем практически любой сложности. Существенным преимуществом имитационного моделирования по сравнению с натурным экспериментом являются минимальные затраты временных и материальных ресурсов. При этом полученные результаты хорошо согласуются с результатами реальных экспериментов [8, 9, 10, 11, 12, 13].

Анализ данных ФФС предполагает создание моделей, соединяющих характеристики исследуемого вещества с характеристиками зарегистрированного потока фотонов, а также разработку методов, алгоритмов и программных средств обработки информации, получаемой в процессе эксперимента. Специфика ФФС состоит в том, что для создания моделей, адекватно описывающих физические процессы в исследуемом веществе, требуется учет физики процессов, происходящих и в системе регистрации, так как неоднородность профиля засветки также влияет на регистрируемые флуктуации интенсивности флуоресценции. Дополнительно система регистрации вносит искажения в регистрируемый поток фотонов. Применение моделей без учета инструментальных искажений с идеализированным представлением профиля засветки ведет к получению значительных ошибок в оценках параметров исследуемого вещества.

Для уменьшения или даже исключения влияния инструментальных искажений необходимо проведение исследований эффективности функционирования системы регистрации в условиях проведения конкретных физических экспериментов и выбору таких параметров измерения, при которых влияние искажений пренебрежимо мало. Данный подход требует детального изучения свойств потоков случайных событий и систем их преобразования.

Аналитическое и имитационное моделирование потоков со схожими характеристиками может дать необходимую информацию для выбора режима измерения или совершенствования измерительной системы путем замены ее отдельных частей более совершенными.

Современные исследования в медицине и микробиологии требуют интенсификации процесса измерения и обработки данных вследствие необходимости проведения большого числа экспериментов с анализируемым веществом. Появился новый термин, определяющий принадлежность некоторого метода к категории быстрых методов исследования, – быстрая обработка измерений (high throughput screening – HTS) [14]. Стремление «попасть» в эту категорию требует разработки не только быстрых методов анализа, но и технологий хранения и обработки данных.

Целью настоящей диссертационной работы является разработка аналитических и имитационных моделей динамики функционирования систем регистрации потоков фотонов, а также создание методов, алгоритмов и программных средств обработки и анализа данных одномолекулярной флуоресцентной флуктуационной спектроскопии.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать аналитические модели системы преобразования потоков случайных событий с использованием аппарата производящих функционалов;
2. Разработать и реализовать программный комплекс имитационного моделирования систем регистрации потоков фотонов во флуоресцентной флуктуационной спектроскопии;
3. Разработать систему обработки и хранения данных флуоресцентной спектроскопии с использованием технологии реляционных баз данных;
4. Разработать методы, алгоритмы и программные средства анализа распределения числа фотоотсчетов с учетом инструментальных искажений системы регистрации;
5. Разработать методы, алгоритмы и программные средства кумулянтного анализа фотоотсчетов для исследования одномолекулярных систем во флуоресцентной флуктуационной спектроскопии.

Решение поставленных задач требует использования методов системного анализа. Исследуемым объектом является сложная физико-техническая система, требующая комплексного изучения процессов взаимодействия излучения с веществом, процессов формирования неоднородной области засветки, процессов детектирования потока фотонов и преобразования его в поток временных событий (фотоотсчетов), а также разработки методов хранения, обработки и анализа данных.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Связь работы с крупными научными программами, темами**

Тема диссертации соответствует направлению 6 «Математическое и физическое моделирование систем, структур и процессов в природе и обществе», пунктам 6.1 «Математические модели и их применение к анализу систем и процессов в природе и обществе» и 6.5 «Аппаратные и программные комплексы и системы для информационного обеспечения» Перечня приоритетных направлений фундаментальных и прикладных научных исследований РБ на 2006-2010 гг., утвержденного Постановлением Совета Министров РБ №512 от 17.05.2005.

Исследования проводились в рамках выполнения в Белорусском государственном университете следующих научно-исследовательских работ: г.б. НИР №719/18 «Разработка методов событийного моделирования динамики функционирования стохастических дискретных процессов и систем» (1998 г., номер госрегистрации 19963416); г.б. НИР № 833/18 «Разработка моделей и методов анализа процессов формирования и преобразования стохастических потоков сигналов» (2000 г., номер госрегистрации 19991554); г.б. НИР № 504/18 «Разработка моделей и методов многомерного анализа кинетики флуоресценции» (2001 г., номер госрегистрации 19992329); г.б. НИР № 466/18 «Разработка теоретических основ статистического анализа временных характеристик случайных потоков в системах автоматизации физического эксперимента» (2003 г., номер госрегистрации 20013802).

### **Цель и задачи исследования**

Целью настоящей диссертационной работы является разработка аналитических и имитационных моделей систем регистрации потоков фотонов, а также создание методов, алгоритмов и программных средств обработки и анализа данных одномолекулярной флуоресцентной флуктуационной спектроскопии.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать аналитические модели системы преобразования потоков случайных событий с использованием аппарата производящих функционалов.
2. Разработать и реализовать программный комплекс имитационного моделирования систем регистрации потоков фотонов в ФФС.
3. Разработать систему обработки и хранения данных флуоресцентной спектроскопии с использованием технологии реляционных баз данных.
4. Разработать методы, алгоритмы и программные средства анализа распределения числа фотоотсчетов с учетом инструментальных искажений системы регистрации.
5. Разработать методы, алгоритмы и программные средства кумулянтного анализа фотоотсчетов для исследования одномолекулярных систем в ФФС.



Объектом исследования являются системы регистрации и обработки данных флуоресцентной флуктуационной спектроскопии. Предметом исследования являются модели, методы, алгоритмы и программные средства регистрации, обработки и анализа данных в системах ФФС.

### **Положения, выносимые на защиту**

1. Программный комплекс имитационного моделирования систем регистрации и обработки данных флуоресцентной флуктуационной спектроскопии, который представляет анализируемую систему в виде последовательности базовых преобразователей и позволяет создавать сложные модели, состоящие из большого числа блоков преобразования, без необходимости изменения программного кода и последующей компиляции программы. Использование комплекса позволило оценить эффективность разработанных методов анализа данных и влияние инструментальных искажений системы регистрации на анализируемые характеристики потоков фотоотсчетов.

2. Метод анализа распределения числа фотоотсчетов, разработанный на основе полученной производящей функции числа фотоотсчетов с коррекцией профиля засветки, характеризующийся отсутствием зависимости яркости и количества молекул в объеме наблюдения от параметров коррекции профиля засветки вследствие предложенной нормализации и позволяющий на порядок уменьшить время анализа данных в сравнении с ранее предложенными методами.

3. Метод генерации начальных приближений для анализа распределения числа фотоотсчетов и кумулянтного анализа, основанный на методе моментов, позволивший находить оценки яркости и количества молекул одно- и двухкомпонентных систем совместно с параметрами коррекции профиля засветки. Метод может быть также использован как самостоятельный метод быстрой оценки параметров исследуемого вещества.

4. Аналитические выражения для факториальных кумулянтов распределения числа фотоотсчетов с коррекцией профиля засветки, характеризующиеся отсутствием зависимости числа оцениваемых параметров от количества анализируемых кумулянтов и позволяющие проводить анализ одиночного набора кумулянтов без постановки калибровочного эксперимента. На основе найденных выражений получены формулы пересчета оценок яркости и количества молекул в объеме наблюдения при использовании различных нормализаций, применяемых в известных и предложенных методах анализа распределения числа фотоотсчетов.

5. Структуры данных измерений и результатов анализа в системах флуоресцентной флуктуационной спектроскопии, на основе которых созданы реляционные базы данных, характеризующиеся возможностью хранения в целостном виде набора анализируемых характеристик совместно с результатами их анализа и позволяющие автоматизировать процесс предварительной обработки и анализа больших объемов данных.

### **Личный вклад соискателя**

Все приведенные в работе результаты получены автором самостоятельно. В публикациях с соавторами вклад соискателя определяется рамками излагаемых в диссертации результатов. Научному руководителю и соавторам в совместных работах принадлежат выбор направлений исследований, предметные постановки целей и задач, проведение натурных измерений и обсуждение результатов, за что автор им очень благодарен: д. ф-м. н., профессору В.В. Апанасовичу за помощь в разработке аналитических и имитационных моделей систем преобразования потоков случайных событий, Prof. Ton Visser (Microspectroscopy Center, WU) и Dr. Oleg Mayboroda (Department of Parasitology, LUMC) за предоставление результатов натурных измерений и обсуждение результатов их анализа.

### **Апробация результатов диссертации**

Основные результаты работы докладывались на: конференции «Анализ и применение систем и сетей массового обслуживания» (Минск, 1994 г.), 3-ем ежегодном международном семинаре «Nonlinear Phenomena in Complex Systems» (Минск, 1995 г.), международной конференции «Computer Data Analysis and Modeling» (Минск, 1995 г.), международной конференции «Исследование систем и сетей массового обслуживания» (Минск, 1996 г.), 7-ом и 9-ом международном семинаре «Carl Zeiss sponsored Workshop on FCS and Related Methods» (Dresden, 2004 г., Stockholm, 2006 г.), международной конференции «Современные информационные компьютерные технологии» (Гродно, 2008 г.), научных семинарах кафедры системного анализа БГУ, расширенных заседаниях кафедр биохимии и молекулярной физики университета Вагенингена (Wageningen, 1998 г., Wageningen 1999 г.), заседании кафедры паразитологии медицинского центра Лейденского университета (Leiden, 2005 г.).

### **Опубликованность результатов диссертации**

Основные результаты диссертации опубликованы в 18 научных работах, из которых 6 статей в соответствии с п. 18 Положения о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь (общим объемом 6,4 авторского листа); 3 статьи в рецензируемых сборниках научных трудов, 7 тезисов докладов на научных конференциях и 2 постера на научных конференциях. Результаты работы вошли в 2 учебных пособия и включены в 4 отчета по НИР.

### **Структура и объем диссертации**

Работа состоит из перечня условных обозначений и сокращений, введения, общей характеристики работы, четырех глав, заключения, 7 приложений и библиографического списка. Полный объем диссертации составляет 138 страниц. Диссертация содержит 31 рисунок на 12 страницах, 13 таблиц на 4 страницах и 7 приложений на 11 страницах. Библиографический список состоит из 161 наименования (включая публикации автора) и занимает 12 страниц.

# ГЛАВА 1

## РЕГИСТРАЦИЯ И АНАЛИЗ СЛУЧАЙНЫХ ПОТОКОВ В ФЛУОРЕСЦЕНТНОЙ ФЛУКТУАЦИОННОЙ СПЕКТРОСКОПИИ

### 1.1 Флуоресцентная флукутационная спектроскопия

Флуоресцентная флукутационная спектроскопия – это совокупность методов, позволяющих исследовать физико-химические параметры вещества на основе анализа спонтанных изменений интенсивности флуоресценции  $F(t)$ . В состоянии термодинамического равновесия флукутации интенсивности  $\delta F(t) = F(t) - \langle F(t) \rangle$  могут происходить вследствие диффузии флуоресцентных молекул через освещенный объем  $V$ , определенный предельно сфокусированным лучом лазера (смотри рисунок 1.1). Хотя флуоресценция возбуждается на всем пути распространения лазерного луча, наибольший вклад в регистрируемый сигнал вносят молекулы, находящиеся в фокальной плоскости и в непосредственной близости от нее. Параллельный лазерный пучок можно сфокусировать в пятно диаметром  $d = 1.22\lambda/NA$  [15], где  $\lambda$  – длина волны лазера,  $NA$  – числовая апертура объектива. Для размеров, указанных на рисунке 1.1, объем наблюдения составляет менее четверти фемтолитра.

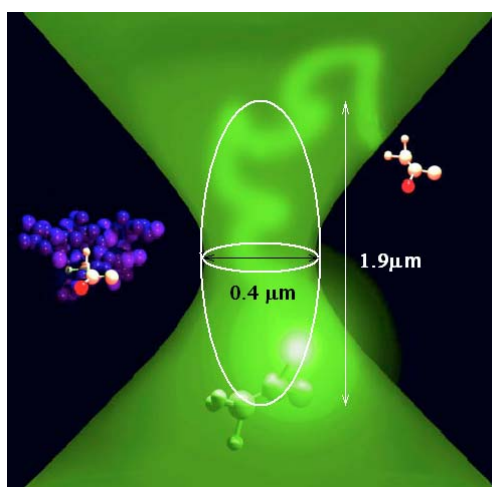


Рисунок 1.1 – Профиль освещенного объема

Интенсивность флуоресценции  $F(t)$  зависит от концентрации  $C$  молекул в объеме наблюдения, их характеристической яркости  $q$  (определяемой как количество фотонов, испущенных одной молекулой в единицу времени, и равной произведению интенсивности засветки в фокусе  $I_0$ , коэффициента поглощения лазерного излучения молекулами вещества  $\sigma_a$ , квантового выхода флуоресценции  $Q$  и эффективности системы регистрации  $\kappa$ ), вероятности возбуждения молекул исследуемого вещества лазерным излучением  $\hat{I}(\mathbf{r})$  и передаточной функции собирающей оптики  $W(\mathbf{r})$  [14]

$$F(t) = I_0 \sigma_a Q \kappa \int_V C(\mathbf{r}, t) \hat{I}(\mathbf{r}) W(\mathbf{r}) d\mathbf{r} = q \int_V C(\mathbf{r}, t) B(\mathbf{r}) d\mathbf{r}. \quad (1.1)$$

Здесь  $C(\mathbf{r}, t)$  – локальная концентрация молекул в точке  $\mathbf{r}$  в момент времени  $t$ . Предполагается, что интенсивность флуоресценции одной молекулы пропорциональна падающему излучению. Удобно ввести общую функцию пространственного распределения зарегистрированной интенсивности флуоресценции  $B(\mathbf{r})$ , равную произведению  $\hat{I}(\mathbf{r})$  и  $W(\mathbf{r})$ . В дальнейшем мы будем называть эту функцию профилем засветки (brightness profile). Наиболее часто  $B(\mathbf{r})$  аппроксимируется трехмерным асимметричным распределением Гаусса, вытянутым по оси  $z$  [23]

$$B(\mathbf{r}) = B(x, y, z) = B_0 \exp\left(-2\frac{(x^2 + y^2)}{\omega_0^2} - 2\frac{z^2}{z_0^2}\right), \quad (1.2)$$

где  $\omega_0$  и  $z_0$  характеризует падение интенсивности флуоресценции соответственно в поперечном и аксиальном направлении и  $B_0 = B(\mathbf{0})$ . Объем наблюдения  $V$  не имеет четко очерченных физических границ, но поскольку интенсивность флуоресценции быстро падает с удалением от фокуса и размеры кюветы с веществом много большие, чем размер освещенной области, интегрирование в выражении (1.1) производится в бесконечных пределах. Для определения количества молекул, участвующих в процессе излучения фотонов, используют величину приведенного объема [16]

$$V_{PSF} = \int_V B(\mathbf{r}) d\mathbf{r}. \quad (1.3)$$

Средняя интенсивность флуоресценции равна произведению средней концентрации молекул  $\langle C \rangle$ , их характеристической яркости в фокусе засветки и величины объема засветки [14, 16]

$$\langle F \rangle = q \langle C \rangle V_{PSF}, \quad (1.4)$$

Для правильной интерпретации выражения (1.4) необходимо иметь в виду, что  $V_{PSF}$  не характеризует закрытый объем с четко очерченными физическими границами и молекулы, находящиеся в различных частях освещенной области, не излучают равное число фотонов (яркость  $q$  определяется как среднее число фотонов, излученных молекулой, находящейся в фокусе засветки). Регистрируемая флуоресценция возникает от большего числа молекул, находящихся в большем по размеру объеме, и молекулы, находящиеся на периферии освещенной области, генерируют меньшее число фотонов.

Каждый раз, когда флуоресцентная молекула входит в освещенный объем, регистрируется всплеск интенсивности. Изменение интенсивности регистрируется даже тогда, когда молекула остается в объеме, но изменяет свое местоположение вследствие сильной неоднородности засветки (максимальная в фокусе и быстро спадающая к периферии). Флуктуации интенсивности могут быть также вызваны

реверсивными химическими реакциями, конформационными изменениями и другими процессами, влияющими на квантовый выход флуоресценции вещества.

Пусть флуктуации интенсивности флуоресценции  $F(t)$  возникают только вследствие изменений локальной концентрации  $\delta C(\mathbf{r}, t)$  молекул одного вида внутри исследуемого объема  $V$ , тогда отклонения интенсивности флуоресценции  $\delta F(t)$  от ее среднего значения могут быть записаны в следующем виде [14, 17]

$$\delta F(t) = \int_V \delta(qC(\mathbf{r}, t))B(\mathbf{r})d\mathbf{r} = q \int_V \delta C(\mathbf{r}, t)B(\mathbf{r})d\mathbf{r}. \quad (1.5)$$

Флуктуации концентрации в точке  $\mathbf{r}$  в момент времени  $t$  относительно среднего значения  $\langle C \rangle$  являются случайной величиной, зависящей от диффузии и равной

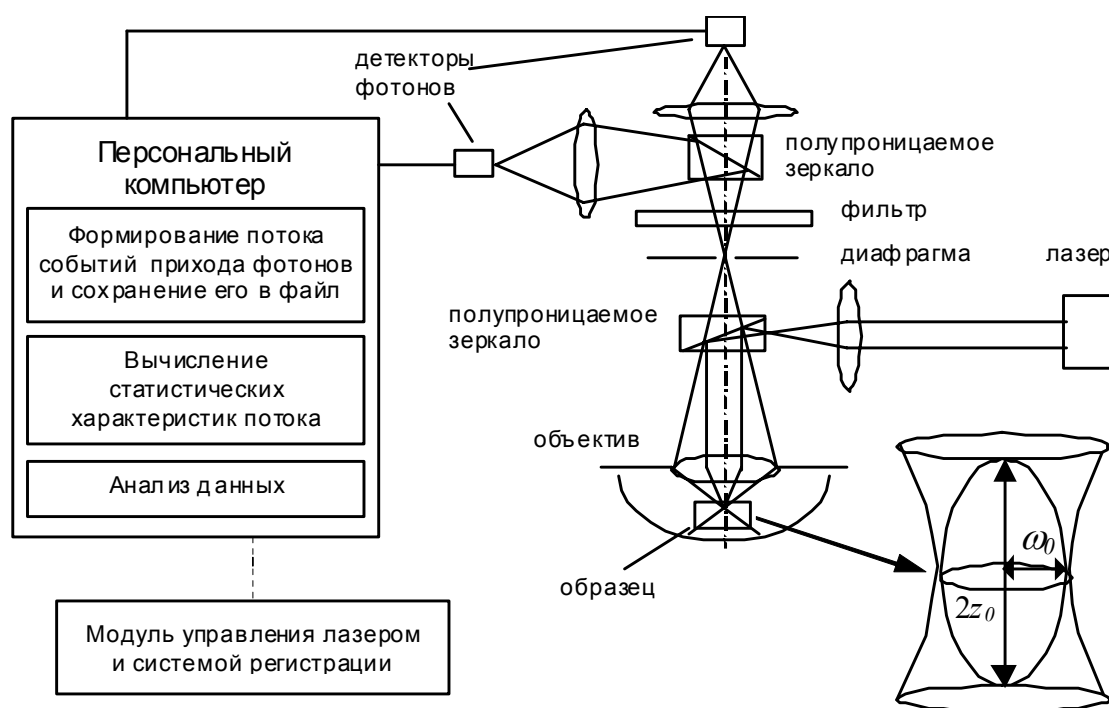
$$\delta C(\mathbf{r}, t) = C(\mathbf{r}, t) - \langle C \rangle. \quad (1.6)$$

При проведении исследований методами ФФС необходимы сигналы от единичных молекул или от их очень малого количества [18]. Наличие большого числа молекул подавляет эффекты флуктуаций, и мы можем наблюдать только среднее значение их ансамбля. Одномолекулярный режим измерения достигается уменьшением размеров освещенного объема до или даже менее предела, определяемого дифракцией [19], и уменьшением концентрации вещества. Уменьшение концентрации возможно до значений, когда интенсивность рассеянного света и фонового сигнала детектора не превысит интенсивности флуоресценции исследуемого вещества.

Хотя технология ФФС была разработана в 1972 году [2], технические проблемы изучения вещества на одномолекулярном уровне были преодолены лишь в последнее десятилетие. Наибольшее содействие этому оказали прогресс в создании источников стабильного лазерного излучения, наличие высокочувствительных фотодетекторов с малым значением темнового тока и, главным образом, появление новых конфокальных микроскопов с очень высоким коэффициентом увеличения и применение точечной диафрагмы в фокальной плоскости для подавления рассеянного света [23].

Рассмотрим структурную схему системы, реализующую технологию ФФС [20, 21] (смотри рисунок 1.2). Луч лазера направляется на вход инвертированного конфокального микроскопа с высоким коэффициентом усиления, отражается от полупрозрачного зеркала и фокусируется на исследуемом образце (чаще всего кювете с водным раствором вещества). Сфокусированный объективом луч освещает небольшую область образца и возбуждает флуоресцентные молекулы, попавшие в эту область. Зеркало изготовлено из материала, отражающего излучение с частотой, равной частоте излучения лазера, но пропускающего флуоресцентное излучение, сдвинутое в сторону больших длин волн. Испускаемое флуоресцентное излучение вместе со значительным количеством рассеянных фотонов собирается тем же объективом микроскопа (конфокальный принцип) и

проходит через диафрагму и фильтр, который задерживает рассеянное излучение лазера. Применение диафрагмы значительно повышает отношение сигнал/шум ( $S/N$ ), так как позволяет регистрировать флуоресценцию только из фокуса освещенной области образца [22, 23]. Размеры сфокусированного луча лазера и размеры диафрагмы определяют тот объем образца, из которого регистрируется флуоресцентное излучение. Испускаемое излучение после прохождения фильтра и диафрагмы разделяется с помощью полупроницаемого зеркала (светоделительной пластинки) и детектируется двумя лавинными фотодиодами (ЛФД), работающими в режиме однофотонного счета [24, 25]. Сигналы детекторов подаются на специальный модуль расширения персонального компьютера и сохраняются в его постоянной памяти. Модуль расширения комплектуется программным обеспечением, обеспечивающим анализ данных и управление подключенными лазерами и системой регистрации. Анализ данных предполагает получение оценок свойств исследуемого вещества путем подгонки некоторой параметрической модели к полученным экспериментальным данным. Представленная схема является типичной для любой системы регистрации и обработки данных (СРОД) в ФФС.



На выноске показан в увеличенном виде профиль освещенного объема

**Рисунок 1.2 – Структурная схема системы регистрации, реализующая технологию ФФС**

СРОД регистрирует поток фотонов, несущий информацию об исследуемом веществе. Процессы, происходящие в веществе, должны рассматриваться в неразрывной связи с физическими процессами, происходящими в системе

регистрации. Флуктуации интенсивности флуоресценции определяются не только характеристиками вещества, например, коэффициентом диффузии, но и профилем функции засветки. Дополнительно, измерительная система вносит искажения в регистрируемый поток фотонов. Поток фотонов преобразуется в поток электрических импульсов, затем в поток временных событий. Фотоны могут теряться вследствие оптических потерь и мертвого времени детектора. Имеется вероятность регистрировать ложные события, вызванные послеимпульсами детектора. При определенных параметрах измерения регистрируемые характеристики могут сильно искажаться. Применение идеальной модели без учета таких искажений ведет к получению значительных ошибок в оценках параметров исследуемого вещества.

Важное значение при исследовании свойства вещества на основе анализа моментов прихода отдельных фотонов играют такие статистические характеристики потока, как интенсивность, автокорреляционная функция (АКФ), распределение числа фотоотсчетов в интервале наблюдения (РЧФ) и кумулянты РЧФ. Интенсивность потока  $F(t)$  дает непосредственное представление о наличии нереверсивных химических реакций и конформационных изменений, процессов фоторазрушения, фотоистощения и т.д. [21, 26]. Автокорреляционная функция  $g(\tau)$  позволяет определять динамические (временные) характеристики вещества, такие как коэффициент диффузии, константы скорости перехода в триплетное состояние, скорости конформационных изменений вещества и реверсивных химических реакций [1, 2, 23]. Распределение числа фотоотсчетов  $P(n)$  и кумулянты распределения  $P(n)$  предоставляют возможность разрешения смеси веществ с различными спектральными (яркостными) свойствами [27].

Наиболее известными методами ФФС являются метод флуоресцентной корреляционной спектроскопии (Fluorescence Correlation Spectroscopy – FCS) [1, 2, 3, 20], в котором анализируется АКФ фотоотсчетов, методы анализа РЧФ (Fluorescence Intensity Distribution Analysis – FIDA и Photon Counting Histogram – PCH) [27, 28] и кумулянтный метод анализа фотоотсчетов (Fluorescence Cumulants Analysis – FCA) [29].

Востребованность изучения свойств вещества на одномолекулярном уровне с использованием методов ФФС предопределило развитие инструментальной базы проведения экспериментов. В настоящее время ряд производителей предлагает коммерчески доступные СПОД. Это системы ConfoCor, ConfoCor 2 [30], ConfoCor 3 фирмы Carl Zeiss GmbH (Germany), ALBA [31] фирмы ISS Inc. (USA), PicoHarp 300 [32] фирмы PicoQuant GmbH (Germany) и другие. Некоторые производители также предлагают карты расширения компьютера, позволяющие самостоятельно собрать СПОД на базе доступных элементов (лазера, инвертированного конфокального микроскопа, детектора и т.д.). Среди них можно выделить Becker

& Hickl GmbH (карты расширения SPC 4xx, SPC 6x0, SPC 830, SPC 134 [33]), PicoQuant GmbH (карта TimeHarp 200) и ISS Inc.

Высокая стоимость натуральных измерений обуславливает необходимость имитационного моделирования СРОД с целью определения режимов измерения, при которых влияние инструментальных искажений системы регистрации минимально. Основным элементом формализации при моделировании системы регистрации выступают потоки случайных событий. Для полного описания процессов преобразования потока фотонов в системе регистрации (потери событий вследствие мертвого времени и появление новых событий вследствие шумов и послеимпульсов) предлагается использовать положения теории случайных потоков событий и математический аппарат производящих функционалов.

## 1.2 Потоки случайных событий

Основным элементом формализации при создании моделей процессов, происходящих в СРОД, выступают потоки случайных событий. Под случайным потоком понимают такой случайный процесс, реализация которого представляет собой набор неразличимых точек, характеризующихся лишь моментами появления на интервале наблюдения. Для целей моделирования потоков случайных событий и систем их регистрации наиболее удобным является параметрический способ описания потоков [4]. Данный способ основан на использовании совместных распределений моментов наступления событий. Суть его применения заключается в получении функциональных уравнений, связывающих производящие функционалы (ПФЛ) исходного и результирующего потока. Найденные уравнения с помощью операций функционального дифференцирования приводятся к обычным алгебраическим, интегральным уравнениям, содержащим искомые характеристики исследуемого потока [4, 34].

Пусть определена область существования потока  $\Omega = [T_1, T_2]$ . Обозначим через  $t_1, t_2, \dots$  – случайные моменты наступления событий в области  $\Omega$ , причем будем полагать, что  $T_1 \leq t_1 \leq t_2 \leq \dots \leq T_2$ . Поток задан, если известны совместные распределения моментов наступления событий  $t_1, t_2, \dots$ , т.е. определена одна из следующих систем плотностей [4]

$$\pi_j(t_1, \dots, t_j; \Omega), j = 0, 1, \dots; f_j(t_1, \dots, t_j), j = 1, 2, \dots \quad (1.7)$$

с помощью которых можно задать любой поток случайных событий, встречающийся в приложениях. Выражение  $\pi_j(t_1, \dots, t_j; \Omega) dt_1 \dots dt_j$  задает вероятность появления в области  $\Omega$  точно  $j$  событий, которые попадают в интервалы  $[t_1, t_1 + dt_1), \dots, [t_j, t_j + dt_j)$ . Функция  $\pi_j(t_1, \dots, t_j; \Omega)$  носит название



плотности распределения первого рода  $j$ -го порядка и для нее выполняется условие нормировки

$$\sum_{i=0}^{\infty} \int_{T_1}^{T_2} \int_{t_1}^{T_2} \dots \int_{t_{i-1}}^{T_2} \pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i = 1. \quad (1.8)$$

Функция  $f_j(t_1, \dots, t_j)$  называется плотностью распределения второго рода  $j$ -го порядка или моментной функцией  $j$ -го порядка. Плотность  $f_j(t_1, \dots, t_j)$  задает совместное распределение моментов появления  $j$  событий потока независимо от того, где выпадали другие события данного потока.

Производящий функционал потока определяется выражением [5, 6]

$$L[u, \Omega] = \left\langle \prod_{j=1}^m [1 + u(t_j)] \right\rangle_{t, m}, \quad (1.9)$$

где:  $u(t)$  – произвольная пробная функция (принимая значения из интервала  $[-1; 0]$ ) и  $\langle \rangle$  – усреднение значений функции по числу появившихся точек и по их координатам. Выполнив усреднение в (1.9), выражение для ПФЛ можно записать в развернутом виде [5]

$$L[u, \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Omega^m} \pi_m(t_1, \dots, t_m; \Omega) \times \prod_{j=1}^m [1 + u(t_j)] dt_1 \dots dt_m. \quad (1.10)$$

Одно из основных преимуществ использования аппарата производящих функционалов заключается в том, что, применяя операцию функционального дифференцирования [5], можно получать произвольные характеристики анализируемого потока. Функциональная производная ПФЛ  $i$ -го порядка при  $u(t) = -1$  дает плотность  $\pi_i(t_1, \dots, t_i; \Omega)$  [5]

$$\pi_i(t_1, \dots, t_i; \Omega) = \left. \frac{\delta L^i[u, \Omega]}{\delta u(t_1) \dots \delta u(t_i)} \right|_{u(t)=-1}, \quad i = 1, 2, \dots \quad (1.11)$$

Функциональная производная от  $L[u, \Omega]$  при  $u(t) = 0$  определяет моментную функцию  $f_i(t_1, \dots, t_i)$  [5]

$$f_i(t_1, \dots, t_i) = \left. \frac{\delta L^i[u, \Omega]}{\delta u(t_1) \dots \delta u(t_i)} \right|_{u(t)=0}, \quad i = 1, 2, \dots \quad (1.12)$$

Среди наиболее часто используемых моментных функций можно выделить интенсивность потока  $f_1(t)$ , представляющую собой моментную функцию первого порядка, и корреляционную функцию случайного потока  $f_2(t_1, t_2)$ , являющуюся моментной функцией второго порядка.

На основе ПФЛ могут быть введены и другие системы плотностей, характеризующих потоки. Одна из важнейших из них определяется  $i$ -кратным функциональным дифференцированием логарифма  $L[u; \Omega]$  при  $u(t) = 0$

$$g(t_1, \dots, t_i) = \frac{\delta^i \ln L[u; \Omega]}{\delta u(t_1) \dots \delta u(t_i)} \Big|_{u(t)=0}. \quad (1.13)$$

Такие функции носят название ковариационных [5] и наряду с корреляционными функциями характеризуют степень взаимосвязи положений отдельных точек. Связь между корреляционной функцией  $f_2(t_1, t_2)$  и ковариационной функцией второго порядка  $g_2(t_1, t_2)$  имеет вид

$$g_2(t_1, t_2) = f_2(t_1, t_2) - f_1(t_1)f_1(t_2). \quad (1.14)$$

В дальнейшем нам понадобятся понятия совместных и условных характеристик потоков. Условные характеристики потоков подразделяются на характеристики с условиями первого и второго рода. Так, условной плотностью распределения первого рода с условием первого рода является функция  $\pi_{m/k}(t_1, \dots, t_m; \Omega | \tau_1, \dots, \tau_k; \Theta)$ , при которой значение выражения  $\pi_{m/k}(t_1, \dots, t_m; \Omega | \tau_1, \dots, \tau_k; \Theta) dt_1 \dots dt_m$  равно вероятности появления в области  $\Omega$  ровно  $m$  событий некоторого случайного потока в окрестностях точек  $t_1, \dots, t_m$ , при условии, что в области  $\Theta$  имелось ровно  $k$  событий случайного потока (того же или совершенно другого), выпавших в моменты  $\tau_1, \dots, \tau_k$  (области  $\Omega$  и  $\Theta$  непересекающиеся) [5]. В свою очередь, условная плотность распределения первого рода с условием второго рода  $\pi_{m/k}(t_1, \dots, t_m; \Omega | \tau_1, \dots, \tau_k)$  есть такая функция, что  $\pi_{m/k}(t_1, \dots, t_m; \Omega | \tau_1, \dots, \tau_k) dt_1 \dots dt_m$  равно вероятности появления в области  $\Omega$  ровно  $m$  событий случайного потока в окрестностях моментов  $t_1, \dots, t_m$  при условии появления не менее  $k$  событий этого же или другого потока, среди которых имеются точки с координатами  $\tau_1, \dots, \tau_k$  [5]. Аналогично определяются условные моментные функции с условием первого  $f_{m/k}(t_1, \dots, t_m | \tau_1, \dots, \tau_k; \Theta)$  и второго  $f_{m/k}(t_1, \dots, t_m | \tau_1, \dots, \tau_k)$  рода. Связь между ними аналогична связи между безусловными характеристиками. На основе условных функций определяются условный производящий функционал первого рода (УПФЛ1)  $L[u; \Omega | \tau_1, \dots, \tau_k; \Theta]$  и условный производящий функционал второго рода (УПФЛ2)  $L[u; \Omega | \tau_1, \dots, \tau_k]$  [5].

Пусть даны два потока случайных событий: поток А и поток В. Совместной плотностью распределения первого рода потоков А и В назовем функцию  $\pi_{mk}(t_1, \dots, t_m; \Omega; \tau_1, \dots, \tau_k; \Theta)$ , такую, что  $\pi_{mk}(t_1, \dots, t_m; \Omega; \tau_1, \dots, \tau_k; \Theta) dt_1 \dots dt_m d\tau_1 \dots d\tau_k$  описывает совместную вероятность появления ровно  $m$  точек потока А в области

$\Omega$  в интервалах  $(t_j - dt_j, t_j]$   $j = 1, \dots, m$  и ровно  $k$  точек потока В в области  $\Theta$  в интервалах  $(\tau_i - d\tau_i, \tau_i]$ ,  $i=1, \dots, k$  [5]. Аналогично вводится совместная моментная функция  $f_{mk}(t_1, \dots, t_m; \tau_1, \dots, \tau_k)$ . При помощи систем плотностей  $\pi_{mk}(t_1, \dots, t_m; \Omega; \tau_1, \dots, \tau_k; \Theta)$  или  $f_{mk}(t_1, \dots, t_m; \tau_1, \dots, \tau_k)$  можно описать совместный производящий функционал  $L[u; \Omega; v; \Theta]$  потоков А и В, где  $u$  – пробная функция потока А, а  $v$  – потока В [5]

$$L[u; \Omega; v; \Theta] = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{m!} \frac{1}{k!} \int_{\Omega^m} \int_{\Theta^k} \pi_{mk}(t_1, \dots, t_m; \Omega; \tau_1, \dots, \tau_k; \Theta) \times \\ \times \prod_{i=1}^m [1 + u(t_i)] \prod_{j=1}^k [1 + v(\tau_j)] dt_1 \dots dt_m d\tau_1 \dots d\tau_k. \quad (1.15)$$

В теории случайных потоков широко используется также производящая функция (ПФ) числа точек потока, определенная как [4, 5, 35]

$$G(\xi) = \langle \xi^n \rangle_n = \sum_{n=0}^{\infty} P(n) \xi^n, \quad (1.16)$$

где  $P(n)$  представляет собой вероятность выпадения на  $\Omega$  ровно  $n$  точек потока и  $\xi$  – некоторая вспомогательная переменная,  $0 < \xi \leq 1$ . Из определения ПФ и выражения (1.10) следует, что ПФ может быть записана с использованием ПФЛ следующим образом:

$$G(\xi) = L[\xi - 1, \Omega]. \quad (1.17)$$

Факториальные кумулянты (семиинварианты) распределения  $P(n)$  получаются как коэффициенты разложения логарифма ПФ по степеням  $\xi - 1$  [35]

$$G(\xi) = \exp \sum_{n=1}^{\infty} \frac{K_n}{n!} (\xi - 1)^n, \quad (1.18)$$

причем

$$K_k = \left. \frac{d^k \ln G(\xi)}{d\xi^k} \right|_{\xi=1}, \quad k = 1, 2, \dots \quad (1.19)$$

Распределение числа событий  $P(n)$  в области  $\Omega$  можно получить путем многократного дифференцирования ПФ

$$P(n) = \left. \frac{d^n G(\xi)}{n! d\xi^n} \right|_{\xi=0} = \frac{1}{n!} \int_{\Omega^n} \pi_n(t_1, \dots, t_n; \Omega) dt_1 \dots dt_n, \quad n = 0, 1, \dots \quad (1.20)$$

Иногда более удобным способом получения  $P(n)$  является использование операции обратного преобразования Фурье (ИФТ). Если в выражении (1.16) формальную переменную  $\xi$  заменить выражением  $e^{i\varphi}$ , где  $i$  – мнимая единица, то связь между производящей функцией и  $P(n)$  примет следующий вид [36]

$$P(n) = IFT(G(e^{i\varphi})), \quad n = 0, 1, \dots, m-1, \quad \varphi = 2\pi n/m, \quad (1.21)$$

где  $m$  – максимальное значение числа выпавших событий.

При создании моделей динамики функционирования СРОД особая роль отводится пуассоновскому и связанному с ним потокам, а также потоку группированных точек. Данные потоки можно использовать в качестве удобных и довольно точных моделей при описании реальных процессов регистрации флуоресценции.

1. Поток Пуассона полностью определяется моментной функцией первого порядка или интенсивностью  $I(t)$ . Его ПФЛ имеет вид [4, 5]

$$L[u; \Omega] = \exp \left\{ \int_{\Omega} I(t)u(t)dt \right\}. \quad (1.22)$$

ПФ числа событий потока Пуассона примет вид

$$G(\xi) = \exp(\Lambda(\Omega)(\xi - 1)), \quad (1.23)$$

где  $\Lambda(\Omega) = \int_{\Omega} I(t)dt$ . Распределение случайной величины  $n$  имеет распределение Пуассона  $P(n) = \Lambda^n(\Omega) \exp(-\Lambda(\Omega))/n!, n = 0, 1, 2, \dots$

Пуассоновский поток является моделью потока фотонов когерентного лазерного излучения. При наличии в освещенном объеме большого количества молекул (или одной неподвижной) регистрируемый поток фотоотсчетов приближается к пуассоновскому потоку [4].

2. Дважды стохастический поток – это поток Пуассона, у которого интенсивность  $I(t)$  является случайным процессом [37]. Такой поток является наиболее адекватной моделью для описания выходного сигнала СРОД. Так как интенсивность не может быть отрицательной, то для любых моментов  $t \in \Omega$   $I(t) \geq 0$ . ПФЛ дважды стохастического потока находится путем усреднения (1.22) по процессу  $I(t)$

$$L[u; \Omega] = \left\langle \exp \left\{ \int_{\Omega} I(t)u(t)dt \right\} \right\rangle_{I(t)}. \quad (1.24)$$

Функционально дифференцируя (1.24) по  $u(t)$  при условии  $u(t) = 0$ , имеем

$$f_i(t_1, \dots, t_i) = \langle I(t_1) \dots I(t_i) \rangle, \quad i = 1, 2, \dots \quad (1.25)$$

Из полученного соотношения следует, что моментные функции потока событий тождественно совпадают с моментными функциями исходного случайного процесса. Распределение числа событий, наступающих в области  $\Omega$ , следует из (1.16), (1.20) и (1.24)

$$P(n) = \left\langle \Lambda(\Omega)^n \exp(-\Lambda(\Omega))/n! \right\rangle_{\Lambda(\Omega)}, \quad n = 0, 1, 2, \dots \quad (1.26)$$

ПФЛ (1.24) может быть записан в явном виде при известных характеристиках случайного процесса  $I(t)$ . Например, если интенсивность потока

Пуассона аппроксимируется нормальным случайным процессом, то поток с такой интенсивностью носит название парно-коррелированного, а его ПФЛ имеет вид [4]

$$L[u; \Omega] = \exp \left\{ \int_{\Omega} f_1(t)u(t)dt + \iint_{\Omega} g_2(t_1, t_2)u(t_1)u(t_2)dt_1dt_2 \right\}. \quad (1.27)$$

3. Поток группированных точек является моделью потока фотоотсчетов при регистрации флуоресценции детектором, обладающим мертвым временем и послеимпульсами. Предположим, что имеется некоторый первичный поток случайных событий (поток  $A$ ), заданный в области  $\Psi$ . Каждое событие потока  $A$  независимо от других событий порождает в  $\Omega$  индивидуальный вторичный поток  $B_i$ ;  $t_i$  – координата наступления порождающего события. Все потоки  $B$  идентичны и отличаются друг от друга лишь координатой рождения. Суперпозиция вторичных потоков образует суммарный поток  $C$  внутригрупповых событий. Как показано в [5], совместный ПФЛ потока  $A$  и потоков  $B_i$  имеет вид

$$L_C[v; \Omega; u; \Psi] = L_A[(1+u(\cdot))L_B[v; \Omega; \Psi] - 1; \Psi], \quad (1.28)$$

где  $L_A[u; \Psi]$  – ПФЛ потока  $A$ ,  $L_B[v; \Omega; \Psi]$  – ПФЛ потока одиночной группы  $B$ . Если в выражении (1.28) положить  $u(\cdot) = 0$ , то получаем формулу для ПФЛ потока  $C$  внутригрупповых событий

$$L_C[v; \Omega] = L_A[L_B[v; \Omega; \Psi] - 1; \Psi]. \quad (1.29)$$

Дифференцируя последнее уравнение, можно установить связь между безусловными плотностями, описывающими суммарный поток, условными характеристиками отдельных вторичных потоков и безусловными плотностями первичного потока.

### 1.3 Преобразование случайных потоков

Поток фотонов, несущий информацию об исследуемом явлении в веществе, подвергается различным преобразованиям, сопровождающим процесс детектирования, регистрации и обработки исходного потока. Преобразование потока обычно связано с прохождением его через некоторые устройства, обладающие линейными или нелинейными характеристиками относительно числа событий потока и их координат, и имеет в общем случае случайный характер. В процессе регистрации и обработки возможны потери событий, их сдвиг или появление дополнительных событий, порождаемых первичными событиями.

Изменение местоположения событий потока обусловлено неточностью привязок при фиксации момента появления импульса, нелинейным преобразованием время-код и т.д. Потери событий в детекторах происходят из-за отличной от единицы квантовой эффективности регистрации и за счет мертвого времени. Потери событий происходят также и в блоках регистрации. Возможны случаи, когда число событий возрастает по сравнению с исходной реализацией.

Это типичный случай при использовании детекторов с электронным умножением, в которых всегда присутствуют шумы и послеимпульсы.

Для обобщенного описания некоторого преобразования будем использовать метод совместного задания моментов наступления событий исходного и преобразованного потока. Пусть исходный поток задан на  $\Theta = [T_1, T_2]$  с моментами наступления событий  $\tau_1, \tau_2, \dots$ , а выходной поток на  $\Omega = [T_0, T]$  с  $t_1, t_2, \dots$ . Тогда совместный ПФЛ данных потоков можно представить в виде [5]

$$L[v; \Theta; u; \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Theta^m} \pi_m(\tau_1, \dots, \tau_m; \Theta) \times \\ \times \prod_{i=1}^m [1 + v(\tau_i)] \times L[u; \Omega | \tau_1, \dots, \tau_m; \Theta] d\tau_1 \dots d\tau_m, \quad (1.30)$$

где  $\pi_m(\tau_1 \dots \tau_m; \Theta)$  плотности исходного потока и  $L[u; \Omega | \tau_1, \dots, \tau_m; \Theta]$  – УПФЛ зарегистрированного потока при условии, что события исходного потока наступили в моменты времени  $\tau_1, \dots, \tau_m$

$$L[u; \Omega | \tau_1, \dots, \tau_m; \Theta] = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; \Theta) \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_m. \quad (1.31)$$

Условные плотности в выражении (1.31) полностью описывают процесс преобразования исходного потока. Зная выражение для  $L[v; \Theta; u; \Omega]$ , легко восстановить ПФЛ входного потока [4]

$$L^1[v; \Theta] = L[v; \Theta; u; \Omega] \Big|_{u(t)=0} \quad (1.32)$$

и ПФЛ преобразованного потока

$$L^2[u; \Omega] = L[v; \Theta; u; \Omega] \Big|_{v(t)=0}. \quad (1.33)$$

### 1.3.1 Простейшие преобразования

**Тождественное преобразование.** Предположим, что преобразование является тождественным, т.е. входной поток не подвергается никаким изменениям. Пусть области задания потоков равны:  $\Omega = \Theta$ . Тогда очевидно, что

$$\pi_{m|k}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; \Omega) = \begin{cases} m! \prod_{i=1}^m \delta(t_i - \tau_i), & k = m, \\ 0, & k \neq m. \end{cases} \quad (1.34)$$

Так как такое преобразование является покомпонентно независимым, то [4]

$$L[u; \Omega | \tau_1, \dots, \tau_m; \Omega] = \prod_{j=1}^m L[u; \Omega | \tau_j]. \quad (1.35)$$

С учетом (1.34), (1.35) и (1.30) получим [4]

$$L[v; \Omega; u; \Omega] = L^1[(1 + v(\cdot))(1 + u(\cdot)) - 1; \Omega], \quad (1.36)$$

где  $v(\cdot)$ ,  $u(\cdot)$  – пробные функции со «слепым» аргументом – точкой.

**Потери событий.** Рассмотрим процесс прореживания потока, аналогичный потере событий из-за отличной от единицы эффективности регистрации фотонов. Предположим, что всякое событие исходного потока, поступившее в момент  $\tau$  независимо от других событий с вероятностью  $r(\tau)$  остается, а с вероятностью  $1-r(\tau)$  теряется. Функцию  $r(\tau)$  будем называть функцией прореживания потока. Очевидно, что

$$\pi_{k|l}(t_1, \dots, t_k; \Omega | \tau; \Theta) = \begin{cases} r(t_1)\delta(t_1 - \tau), & k = 1, \\ 1 - r(\tau), & k = 0, \\ 0, & k > 1. \end{cases} \quad (1.37)$$

Подставив (1.37) в (1.30) получим

$$L[u; \Omega | \tau_1, \dots, \tau_m; \Theta] = \prod_{i=1}^m [1 + r(\tau_i)u(\tau_i)]. \quad (1.38)$$

На основании (1.30) и того, что такое преобразование является покомпонентно независимым (1.35), находим [4]

$$L[v; \Omega; u; \Omega] = L^1[(1+v(\cdot))(1+r(\cdot)u(\cdot)) - 1; \Omega]. \quad (1.39)$$

Отсюда получаем ПФЛ прореженного потока с функцией прореживания  $q(\tau)$

$$L^2[u; \Omega] = L^1[r(\cdot)u(\cdot); \Omega]. \quad (1.40)$$

Рассмотрим случай, когда  $\Omega \subset \Theta$  и все события исходного потока, попадающие в область  $\Omega' = G \setminus \Theta$ , теряются. Такое преобразование сводится к предыдущему, если положить

$$r(t) = \begin{cases} 1, & t \in \Omega, \\ 0, & t \in \Omega'. \end{cases} \quad (1.41)$$

Подставляя (1.41) в (1.37), а затем полученное выражение в (1.31), находим

$$L[u; \Omega | \tau; \Theta] = 1 + u(\tau) \Big|_{u(t)=0, t \in \Omega'}. \quad (1.42)$$

Соотношение (1.42) позволяет записать ПФЛ потока на подобласти  $\Omega$  [4]

$$L^2[u; \Omega] = L^1[u; \Theta] \Big|_{u(t)=0, t \in \Theta \setminus \Omega}. \quad (1.43)$$

**Сдвиг событий.** Реакция различных блоков и узлов некоторого устройства на поступление на их вход импульса от предшествующих блоков может запаздывать относительно момента их истинного прихода. Смещение это носит чаще всего случайный характер и связано с нестабильностью времени преобразования информации во входных узлах регистрирующей аппаратуры.

Рассмотрим линейное независимое преобразование точек потока, имеющее вид  $\chi_i = \gamma_i + \beta_i$ ,  $i = 1, 2, \dots$ , где  $\beta_i$  – одинаково распределенные независимые

случайные величины с функцией распределения  $B(t)$ , характеризующие смещение точек исходного потока. Условный ПФЛ такого преобразования имеет вид [4]

$$L[u; \Omega | \tau_1, \dots, \tau_m; \Theta] = \prod_{i=1}^m [1 + \int_{\Omega} u(t) dB(t - \tau_i)]. \quad (1.44)$$

Совместный ПФЛ и ПФЛ выходного потока соответственно примут вид [4]

$$L[u; \Omega; v; \Theta] = L^1[(1 + v(\cdot))(1 + \int_{\Omega} u(t) dB(t - \cdot)) - 1; \Theta], \quad (1.45)$$

$$L^2[u; \Omega] = L^1[\int_{\Omega} u(t) dB(t - \cdot); \Theta]. \quad (1.46)$$

На основании формулы (1.46) можно найти связь различных характеристик входного и выходного потоков. Так дифференцируя (1.46) при условии  $u(t) = 0$ , получим выражение для интенсивности  $f_1^2(t)$  выходного потока через интенсивность  $f_1^1(t)$  входного потока и функцию распределения случайной величины  $\beta$

$$f_1^2(t) = \int_{\Omega} f_1^1(\tau) \frac{dB(t - \tau)}{dt} d\tau. \quad (1.47)$$

### 1.3.2 Регистрация случайных потоков системами с мертвым временем

Минимальный промежуток времени между поступающими на вход устройства (системы) событиями, при котором устройство способно их различать и обрабатывать, называется разрешающим временем устройства или мертвым временем [38]. Наличие мертвого времени всегда приводит к появлению просчетов при проведении измерений. Эта ошибка является систематической и результаты измерения можно исправить, если известно мертвое время. Мертвое время представляется наиболее существенным в искажении исходного потока фотонов, особенно для случаев предельно высокой для данного регистратора интенсивности. Преобразования потока, вносимые мертвым временем, являются нелинейными, и их анализ связан со значительными трудностями.

Обычно рассматриваются две группы устройств, характеризующихся мертвым временем [38]. Первую группу составляют устройства с непродлевающим мертвым временем  $\tau$ , возникающим после зарегистрированного события, в течение которого все поступающие события теряются и не удлиняют мертвое время. Такие устройства называют регистраторами I типа, а мертвое время – мертвым временем I рода. Ко второй группе относятся устройства с продлевающим мертвым временем, возникающим после каждого события, независимо, зарегистрировано оно или нет. При этом устройство может зарегистрировать событие лишь в случае, когда окончатся все мертвые времена, порожденные ранее поступившими на устройство событиями. Соответственно



устройства, обладающие продлевающим мертвым временем, называют регистраторами II типа, а само мертвое время – мертвым временем II рода.

**Продлевающее мертвое время.** Продлевающее мертвое время характерно, как правило, для детекторов излучения, а также для комбинации детектор - усилитель - устройство временной фиксации. Встречается оно и в блоках преобразования время-код, имеющих специальные устройства для исключения наложенных импульсов [39].

Исследование влияния продлевающего мертвого времени при общих условиях (на вход поступает произвольный поток событий, мертвое время случайное, характеризующееся условной функцией распределения  $B(x, \tau)$ ) произведено в [4]. Наряду с входным потоком А и выходным С введем фиктивный поток Q, образованный моментами завершения мертвых времен  $x_1, x_2, \dots$ , порожденных событиями потока А. Области задания потоков А и С совпадают и соответствуют интервалу  $\Theta = [T_1 T_2]$ . Для потока Q такой областью является расширенный интервал  $X = [T_1 \infty]$ . УПФЛ преобразования в предположении, что на  $\Theta$  наступило точно  $m$  событий потока А в моменты времени  $\tau_1, \dots, \tau_m$ , а порожденные ими мертвые времена закончатся в моменты времени  $x_1, \dots, x_m$  имеет вид  $(v(), u(), w())$  – пробные функции соответственно потоков А, С, Q) [4]

$$L[u; \Theta | \tau_1, \dots, \tau_m; x_1, \dots, x_m] = \prod_{i=1}^m [1 + u(\tau_i)] \Big|_{u(\tau)=0, \tau \in \Theta'}, \quad (1.48)$$

где  $\Theta' = \bigcup_{k=1}^m [\tau_k, x_k]$ . Преобразуем полученное выражение к виду

$$L[u; \Theta | \tau_1, \dots, \tau_m; x_1, \dots, x_m] = \prod_{i=1}^m [1 + u(\tau_i)] \prod_{\substack{j=1 \\ j \neq i}}^m [1(\tau_j - \tau_i) + 1(\tau_i - x_j)], \quad (1.49)$$

где  $1()$  – функция, принимающая единичное значение для неотрицательного аргумента. ПФЛ выходного потока С запишется как [4]

$$L[u; \Theta] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Theta^m} \int_{X^m} \pi_m(\tau_1, \dots, \tau_m; \Theta) \times \prod_{i=1}^m [1 + u(\tau_i)] \prod_{\substack{j=1 \\ j \neq i}}^m [1(\tau_j - \tau_i) + 1(\tau_i - x_j)] dB(x_i, \tau_i) d\tau_1 \dots d\tau_m. \quad (1.50)$$

Найденный ПФЛ позволяет определить различные характеристики исходного потока С и выразить их через характеристики исходного потока и функцию распределения мертвого времени. Например, дифференцируя выражение (1.50) по пробной функции  $u(t)$  и полагая ее равной нулю, получим выражение, связывающее интенсивности входного и выходного потоков [4]

$$\begin{aligned}
f_1^C(t) &= \frac{\delta L[u; \Theta]}{\delta u(t)} \Big|_{u(t)=0} = \sum_{m=0}^{\infty} \frac{m}{m!} \int_{\Theta^{m-1}} \pi_m(\tau_1, \dots, \tau_{m-1}, t; \Theta) \prod_{i=1}^{m-1} [1(\tau_j - t) + B(t, \tau_i)] d\tau_1 \dots d\tau_{m-1} = \\
&= f_1^A(t) \sum_{m=1}^{\infty} \frac{1}{(m-1)!} \int_{\Theta^{m-1}} \pi_{k-1|1}(\tau_1, \dots, \tau_{m-1}; \Theta | t) \prod_{i=1}^{m-1} [1(\tau_j - t) + B(t, \tau_i)] d\tau_1 \dots d\tau_{m-1}
\end{aligned} \tag{1.51}$$

или

$$f_1^C(t) = f_1^A(t) L[B(t, \cdot) + 1(\cdot - t) - 1; \Theta | t]. \tag{1.52}$$

Искажающее влияние мертвого времени регистрации II типа было исследовано в целом ряде работ [25, 38, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50]. Так в работах [40, 25, 42, 44, 55, 45] рассмотрен наиболее простой случай: простейший входной поток и мертвое время регистрации постоянной длительности. Показано, что интенсивность зарегистрированного потока  $f_1$  выражается через интенсивность входного потока  $\lambda$  и длительность мертвого времени  $\tau$  как

$$f_1 = \lambda \exp\{-\lambda\tau\}. \tag{1.53}$$

В случае малых загрузок регистратора (при условии  $\lambda\tau \ll 1$ ) выражение (1.53) можно представить в виде

$$f_1 = \lambda(1 - \lambda\tau). \tag{1.54}$$

**Непродлевающее мертвое время.** Непродлевающее мертвое время характерно для блоков записи кода момента поступления импульса. Кроме того, мертвое время I типа возможно в некоторых детекторах излучений – таких как детектор Гейгера-Мюллера с внешним гашением [51] и лавинный фотодиод (наиболее часто применяющийся в ФФС) [30]. Для непродлевающего мертвого времени не удалось получить выражение, эквивалентное (1.52). Но в более простых случаях аналитические выражения были получены. Из множества работ, посвященных этому типу мертвого времени, выделим работы [25, 38, 40, 42, 47, 48, 52, 53, 54, 55, 56, 57, 58, 59, 60].

Анализ наиболее простого случая, соответствующего стационарному пуассоновскому входному потоку и мертвому времени постоянной длительности, посвящены работы [38, 40]. Показано, что для этого случая интенсивность зарегистрированного установившегося потока  $f_1$  определяется через интенсивность входного потока  $\lambda$  и длительность мертвого времени  $\tau$  по формуле

$$f_1 = \frac{\lambda}{1 + \lambda\tau}. \tag{1.55}$$

Если  $\lambda\tau \ll 1$  выражение (1.55) представимо в виде

$$f_1 = \lambda(1 - \lambda\tau), \tag{1.56}$$

что полностью совпадает с выражением для продлевающего времени. Из последнего следует, что при малых нагрузках регистратора действие продлевающего и непродлевающего мертвого времени не различаются.

В работах [52, 42, 43] получены аналитические выражения, описывающие зарегистрированную интенсивность для нестационарного пуассоновского потока.

**Обобщенное мертвое время.** Проведение экспериментов с реальными устройствами показало, что практически ни одно устройство нельзя точно отнести к одному из традиционных типов мертвого времени. Модель, удовлетворительно описывающая этот промежуточный тип мертвого времени, была впервые предложена Альбертом и Нельсоном в 1953 году [61] и получила свое развитие в работах [62, 63, 64, 65, 66]. В этой модели предполагается существование конечной вероятности  $\theta$  того, что событие, прибывающее на заблокированное вследствие мертвого времени устройство, будет продлевать это мертвое время. Полагая  $\theta$  равным 0 и 1, получаем соответственно непродлевающее и продлевающее мертвое время.

В [62] для постоянного мертвого времени получено простое выражение, связывающее входную и выходную интенсивности

$$f_1 = \frac{\theta\lambda}{\exp\{\theta\lambda\tau\} + \theta - 1}. \quad (1.57)$$

Нетрудно убедиться, что полагая  $\theta = 0$  и 1, мы получаем (1.55) и (1.53) соответственно.

### 1.3.3 Преобразование потоков в системе с последствием

При регистрации случайных потоков событий с помощью таких детекторов, как ЛФД, наряду с информационными электрическими сигналами наблюдается поток послеимпульсов. Присутствие послеимпульсов ухудшает отношение сигнал/шум на выходе детектора, существенно искажает форму регистрируемого распределения, вносит в исходный поток дополнительные корреляционные зависимости между моментами появления событий. Послеимпульсные явления обнаружены для всех детекторов, основанных на использовании явления вторичного электронного умножения [67, 68, 69].

Модель регистрации потока фотонов детектором с послеимпульсами, позволяющая учитывать послеимпульсы всех поколений, рассмотрена в [70]. Обозначим через  $L^A[u; \Omega]$  ПФЛ поступающего на вход детектора потока событий (поток А), заданного на интервале  $\Omega = [T_0, T]$ . Каждое событие исходного потока А в процессе его регистрации детектором порождает независимо от других событий поток послеимпульсов одного поколения, ПФЛ которого запишем в виде  $L^*[u; \Theta | \tau]$ , где  $\tau$  – момент наступления порождающего послеимпульсы события;  $\Theta = [T_0, T_1]$  – интервал регистрации случайного потока,  $T_1 \geq T$ ,  $T_1 < +\infty$ . Будем

использовать пробную функцию  $v(t)$  для индикации событий на выходе детектора, которые уже образовали послеимпульсы, а  $u(t)$  – для событий, которые еще будут инициировать новые послеимпульсы. Обозначим через  $L^C[v; \Theta]$  ПФЛ потока событий на выходе детектора (поток С). Так как поток С является совокупностью событий исходного потока и потока послеимпульсов всех поколений, то можно записать [70]

$$L^C[v; \Theta] = L^A[(1+v(x_1))L^*[(1+v(x_2))L^*[(1+v(x_3)) \dots] - 1; \Omega]. \quad (1.58)$$

Рассмотрим поток послеимпульсов всех поколений (поток В), образованный одним событием, наступившим в момент времени  $\tau$ . Обозначим его ПФЛ через  $L^B[v; \Theta | \tau]$ , тогда справедливо соотношение [70]

$$L^C[v; \Theta] = L^A[(1+v(\cdot))L^B[v; \Theta | \cdot] - 1; \Omega]. \quad (1.59)$$

Сравнивая (1.59) и (1.58), получим

$$L^B[v; \Theta | \tau] = L^*[(1+v(x_2))L^*[(1+v(x_3)) \dots] - 1; \Theta | \tau]. \quad (1.60)$$

Откуда следует, что

$$L^B[v; \Theta | \tau] = L^*[1+v(\cdot))L^B[v; \Theta | \cdot] - 1; \Theta | \tau]. \quad (1.61)$$

Функциональные соотношения (1.60) и (1.61) определяют математическую модель регистрации потока детектором с послеимпульсами [70]. Наличие функционального уравнения (1.61) принципиально важно для рассмотренной модели. Это уравнение позволяет выразить характеристики суммарного потока послеимпульсов при использовании аналитического описания потока послеимпульсов только одного поколения.

Вычислив функциональные производные от соотношений (1.60) и (1.61) при  $v(t)=0$ , можно найти выражение, связывающее интенсивности потоков А, В, С

$$f_1^C(t) = f_1^A(t) + \int_{\Omega} f_1^A(x) f_1^B(t | x) dx, \quad (1.62)$$

$$f_1^B(t | \tau) = f_1^*(t | \tau) + \int_G f_1^*(x | \tau) f_1^B(t | x) dx, \quad (1.63)$$

где верхние индексы С, А, В и \* используются для обозначения характеристик соответственно потоков С, А, В и послеимпульсов одного поколения.

## 1.4 Методы анализа данных флукуационной спектроскопии

### 1.4.1 Флуоресцентная корреляционная спектроскопия

Флуоресцентной корреляционной спектроскопией (Fluorescence Correlation Spectroscopy – FCS) называется метод, в котором анализируется временная автокорреляционная функция флукуаций интенсивности флуоресценции [1, 2, 3, 20, 23, 71, 72]. Данный метод применяется для измерения коэффициента вращательной и пространственной диффузии, кинетических констант химических

реакций, молекулярных весов, концентрации вещества, кинетики перехода в триплетное состояние и реакций гибридизации.

Нормированную АКФ  $g(\tau)$  флуктуаций интенсивности флуоресценции  $\delta F(t)$  можно определить следующим образом [1]

$$g(\tau) = \frac{\langle F(t+\tau)F(t) \rangle}{\langle F(t) \rangle^2} = \frac{\langle \delta F(\tau)\delta F(0) \rangle}{\langle F \rangle^2} + 1. \quad (1.64)$$

Подставляя выражение (1.5) в (1.64), получим

$$g(\tau) = 1 + \frac{\iint B(\mathbf{r})B(\mathbf{r}')g_C(\mathbf{r},\mathbf{r}',\tau)d\mathbf{r}d\mathbf{r}'}{[\langle C \rangle \int B(\mathbf{r})d\mathbf{r}]^2}, \quad (1.65)$$

где  $g_C(\mathbf{r},\mathbf{r}',\tau) = \langle \delta C(\mathbf{r},\tau)\delta C(\mathbf{r}',0) \rangle$  – корреляционная функция флуктуаций концентрации, что следует из предположения стационарности процесса  $C(\mathbf{r},t)$ .

Для случая свободной диффузии в однородной среде имеем [73]

$$g_C(\mathbf{r},\mathbf{r}',\tau) = \langle C \rangle (4\pi D\tau)^{-\frac{3}{2}} \exp\left(-\frac{(\mathbf{r}-\mathbf{r}')^2}{4D\tau}\right). \quad (1.66)$$

Величина  $D$  является диффузионным коэффициентом вещества (не зависящим от используемого оборудования) и фактически является основной искомой величиной в методе флуоресцентной корреляционной спектроскопии. Однако обычно пользуются характеристическим диффузионным временем  $\tau_{diff}$ , описывающим среднее время пребывания молекулы в освещенной области. Различают поперечное и аксиальное диффузионное время, так как размер освещенного объема различен в этих направлениях. Поперечное диффузионное время в случае применения (1.2) для аппроксимации  $B(r)$  определяется как [74]

$$\tau_{diff} = \omega_0^2 / (4D). \quad (1.67)$$

Значение  $\omega_0$  может быть найдено из калибровочного измерения вещества с хорошо известным  $D$ . Основанием для использования выражения (1.2) при разработке модели является то, что для других аппроксимаций профиля засветки аналитический вид автокорреляционной функции найти очень сложно [23, 74, 75].

Значение  $g(0)$  характеризуется корреляцией флуктуаций концентрации в точках  $\mathbf{r}$  и  $\mathbf{r}'$  в один и тот же момент времени. Если не учитывать взаимодействие между молекулами раствора, то  $g_C(\mathbf{r},\mathbf{r}',0) = \langle C \rangle \delta(\mathbf{r}-\mathbf{r}')$ , где  $\delta()$  есть дельта-функция Дирака. Из последнего выражения и (1.65) получим

$$g(0) = 1 + \frac{1}{V_{eff} \langle C \rangle}. \quad (1.68)$$

$V_{eff}$  в выражении (1.68) это эффективный объем освещенной области, равный

$$V_{eff} = \chi_1^2 / \chi_2, \quad (1.69)$$

где  $\chi_k$  – интеграл от функции профиля засветки в степени  $k$

$$\chi_k = \int_V B^k(\mathbf{r}) d\mathbf{r}. \quad (1.70)$$

Для трехмерного гауссовского распределения

$$V_{eff} = \pi^{3/2} \omega_0^2 z_0. \quad (1.71)$$

Следовательно, амплитуда АКФ определяется средним количеством молекул  $N_{eff} = V_{eff} \langle C \rangle$  в эффективном объеме

$$g(0) = 1 + 1/N_{eff}. \quad (1.72)$$

Кроме пересчета концентрации  $C$  к среднему числу молекул в эффективном объеме иногда применяют пересчет к среднему числу молекул в объеме засветки  $V_{PSF}$  [75], определенному в разделе 1.1. Тогда выражение (1.72) примет вид  $g(\tau) = 1 + \gamma_2/N_{PSF}$ , где  $\gamma_2 = \chi_2/\chi_1$  – это так называемый гамма-фактор, равный отношению интегралов от профиля засветки.

Подставляя выражение (1.66) и (1.67) в (1.65), получим

$$g(\tau) = 1 + \frac{1}{N_{eff}} \frac{1}{\left(1 + \frac{\tau}{\tau_{diff}}\right) \sqrt{\left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \frac{\tau}{\tau_{diff}}\right)}}. \quad (1.73)$$

Для смеси невзаимодействующих молекул двух и более видов выражение (1.5) переписывается в виде

$$\delta F(t) = \sum_{i=1}^M q_i \int \delta(C_i(\mathbf{r}, t)) B(\mathbf{r}) dV, \quad (1.74)$$

где  $M$  – количество компонент смеси. Соответственно, выражение АКФ для смеси  $M$  веществ примет вид [74]

$$g(\tau) = 1 + \sum_{i=1}^M \frac{q_i^2 N_{eff i}}{\left(\sum_{j=1}^M q_j N_{eff j}\right)^2} \frac{1}{\left(1 + \frac{\tau}{\tau_{diff i}}\right) \sqrt{\left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \frac{\tau}{\tau_{diff i}}\right)}}. \quad (1.75)$$

Таким образом, АКФ смеси невзаимодействующих молекул нескольких типов будет представлять линейную комбинацию АКФ отдельных компонент, взвешенных по квадрату их удельных яркостей.

Вывод формул (1.73), (1.75) проведен в предположении, что флуктуации интенсивности определяются только свободной диффузией молекул в однородной среде. Однако почти всегда имеют место дополнительные процессы, способные приводить к флуктуациям интенсивности флуоресценции. К ним можно отнести быстрые химические реакции, переход молекул в триплетное состояние и некоторые другие. Когда характеристическое время этих процессов и

характеристическое время диффузии различаются на несколько порядков, формула для вычисления АКФ имеет следующий вид [76, 77]

$$g(\tau) = 1 + \frac{1}{N_{eff}} g_{process}(\tau) g_{diff}(\tau), \quad (1.76)$$

где

$$g_{diff}(\tau) = \sum_{i=1}^M \frac{F_i}{\left(1 + \frac{\tau}{\tau_{diff i}}\right) \sqrt{1 + \left(\frac{\omega_0}{z_0}\right)^2 \frac{\tau}{\tau_{diff i}}}} \quad (1.77)$$

и  $g_{process}(\tau)$  – автокорреляционная функция одного из вышеперечисленных процессов. В случае учета перехода молекул в триплетное состояние [78]

$$g_{process}(\tau) = 1 + F_{trip} e^{-\tau / \tau_{trip}} / (1 - F_{trip}), \quad (1.78)$$

где  $F_{trip}$  и  $\tau_{trip}$  – соответственно фракция и среднее время пребывания молекул в триплетном состоянии. Более полный обзор выражений для АКФ, полученных с учетом различных процессов, можно найти в [14, 74, 79, 80, 81].

АКФ измеренных данных рассчитывается либо аппаратным, либо программным способом. Для расширения временного диапазона АКФ вычисляют, используя квазилогарифмическую шкалу [82, 83, 84]. Анализ данных в FCS производится путем подгонки теоретической модели (1.76) к измеренной АКФ по методу наименьших квадратов (МНК) [85]. В качестве критерия качества подгонки применяют взвешенный (приведенный к единице) критерий  $\chi^2$  [85] и визуальную оценку взвешенных остатков. При невозможности вычисления среднеквадратических отклонений экспериментальных точек применяют невзвешенный критерий  $\chi^2$ .

#### 1.4.2 Анализ распределения числа фотоотсчетов

При анализе смеси веществ с помощью (1.73), (1.75) необходимо учитывать то, что существующие методы и программные средства, позволяющие проводить анализ экспериментальных данных с целью получения оценок параметров теоретической модели, не позволяют эффективно разрешать диффузионные времена молекул, незначительно отличающихся по массе (и соответственно, по их характерному диффузионному времени). Экспериментально показано [86], что при анализе можно успешно разделить вещества, отличающиеся по диффузионному времени в 1.6 раз, что, принимая во внимание кубическую зависимость коэффициента диффузии от массы молекул приводит к более чем четырехкратному различию их в массе. Это сильно сужает сферу применения данной модели, так как часто приходится изучать смеси веществ с незначительно отличающимися массами. Данное ограничение стимулировало разработку

методов, которые позволили бы разрешать смесь веществ по другим параметрам, нежели по их характерному диффузионному времени.

В подразделе 1.4.1 при выводе формулы для расчета АКФ было показано, что вещества могут также различаться по их характерной яркости (квантовому выходу). Число фотонов, излученных одной молекулой в единицу времени, также может служить характеристикой вещества, так как напрямую зависит от его спектральных свойств. Для разрешения смеси веществ, различающихся по их удельным яркостям, было предложено несколько методов. Вначале была предложена модификация метода моментов [87, 88], сводящаяся к расчету нескольких первых моментов распределения флуктуаций интенсивности флуоресценции (распределения числа фотоотсчетов) и последующему решению системы нелинейных уравнений. Этот подход был впоследствии обобщен и получил название кумулянтного метода анализа фотоотсчетов (FCA) [29]. Затем две научные группы независимо и практически одновременно предложили теоретические модели и методы для анализа РЧФ. Эти методы получили названия РСН [27] и FIDA [28].

Перед выводом теоретического выражения для РЧФ сделаем несколько предположений, не ограничивающих значительно сферу применения данного выражения, зато упрощающих его вывод. Во-первых, предположим, что молекулы вещества независимы между собой, так же как и их вклад в зарегистрированный поток фотонов. Это предположение позволит нам в дальнейшем использовать операцию свертки для вычисления совместного РЧФ от различных молекул, типов молекул и даже от отдельных элементов объема наблюдения. Во-вторых, предположим, что яркостные характеристики вещества не изменяются в течение достаточно короткого времени наблюдения  $T$ . В-третьих, будем считать, что интенсивность флуоресценции отдельной молекулы, находящейся в некоторой области наблюдения, является произведением ее характеристической яркости  $q$  и значения функции профиля засветки  $B(\mathbf{r})$  в этой области. Другие эффекты, например, насыщение и переход в триплетное состояние, учитываться не будут.

Разделим объем наблюдения на небольшие пространственные области  $dV_i$  с одинаковыми значениями  $B_i$  функции распределения засветки  $B(\mathbf{r})$ . Тогда вероятность получения  $n$  фотонов в  $i$ -той области за время интервала наблюдения  $T$  можно выразить как  $P_i(n) = \sum_{m=0}^{\infty} P(m)P(n/m)$ . Здесь  $P(m)$  – вероятность обнаружения  $m$  молекул в области  $B_i$  и  $P(n/m)$  – условная вероятность получения  $n$  фотонов от  $m$  молекул, находящихся в области  $B_i$ . Распределение вероятности обнаружения  $m$  невзаимодействующих между собой молекул в некоторой открытой подобласти  $B_i$  исследуемого объема хорошо описывается распределением Пуассона с параметром  $\langle C \rangle dV_i$ , где  $\langle C \rangle$  средняя концентрация молекул в объеме наблюдения. Закон распределения вероятности регистрации  $n$



фотонов от  $m$  молекул, находящихся в области  $B_i$  за время  $T$  также является пуассоновским по своей природе [7] и будет описываться распределением Пуассона с параметром  $m q B_i T$ . Таким образом, результирующее распределение будет двойным пуассоновским с параметрами  $\langle C \rangle dV_i$  и  $m q B_i T$  [28]

$$P_i(n) = \sum_{m=0}^{\infty} \frac{(\langle C \rangle dV_i)^m}{m!} e^{-\langle C \rangle dV_i} \frac{(m q B_i T)^n}{n!} e^{-m q B_i T}. \quad (1.79)$$

Совместное распределение вероятности  $P(n)$  получения  $n$  фотонов за время  $T$  от всего объема наблюдения (РЧФ) будет представлять собой свертку отдельных распределений  $P_i$  от каждого элемента объема  $dV_i$

$$P(n) = P_1(n) \otimes P_2(n) \otimes \dots \otimes P_i(n). \quad (1.80)$$

Учитывая, что свертка распределений вероятности заменяется простым произведением их производящих функций, для области  $dV_i$  получим [28]

$$G(\xi, dV_i) = e^{-\langle C \rangle dV_i} \sum_{m=0}^{\infty} \frac{(\langle C \rangle dV_i)^m}{m!} e^{-m q B_i T} \sum_{n=0}^{\infty} \frac{(m \xi q B_i T)^n}{n!} = \exp[\langle C \rangle dV_i (e^{(\xi-1) q B_i T} - 1)],$$

где тождество  $\sum_n \frac{x^n}{n!} = e^x$  использовано дважды. Окончательно ПФ для всего объема наблюдения и для случая смеси молекул нескольких типов будет представлена перемножением ПФ распределений от всех элементов объема и различных типов молекул. Принимая во внимание, что умножение экспонент заменяется суммой их показателей, можно перейти к интегралу по всему объему

$$G(\xi) = \exp\left(\sum_i \langle C \rangle_i \int_V \left(e^{(\xi-1) q_i T B(\mathbf{r})} - 1\right) d\mathbf{r}\right). \quad (1.81)$$

Для учета рассеянного света и фонового сигнала детектора в выражение (1.81) можно включить слагаемое вида  $\exp[(\xi-1)\lambda T]$ , представляющего собой ПФ пуассоновского распределения с параметром  $\lambda$ , равным средней интенсивности фонового сигнала [28]

$$G(\xi) = \exp\left(\lambda T (\xi - 1) + \sum_i \langle C \rangle_i \int_V \left(e^{(\xi-1) q_i T B(\mathbf{r})} - 1\right) d\mathbf{r}\right). \quad (1.82)$$

**FIDA.** Ключевой особенностью метода FIDA является аппроксимация профиля засветки экспоненциальной функцией одного аргумента с преобразованием элемента объема в виде полинома этого аргумента. Это позволило перейти от пространственного интеграла в выражении (1.82) к интегралу по одной переменной. Сделаем замену переменных [28, 89]

$$d\mathbf{r}/dx = A_0 (x + ax^2 + bx^3), \quad x = \ln[B_0/B(\mathbf{r})], \quad (1.83)$$

где  $a, b$  – параметры, зависящие от используемой в эксперименте аппаратуры. Такое приближение позволяет учесть разброс значений, вызванный различием

оптических параметров измерительной системы и систематических ошибок, возникающих за счет ее разбалансировки. Способ вычисления коэффициентов  $B_0$  и  $A_0$  может быть произвольным и определяется выбором единиц измерения  $V$  и  $q$  [89]. Вычислим  $B_0$  и  $A_0$  из системы уравнений нормировки в виде [28, 89]

$$\begin{aligned} \int_V B(\mathbf{r}) d\mathbf{r} &= 1, \\ \int_V B^2(\mathbf{r}) d\mathbf{r} &= 1. \end{aligned} \quad (1.84)$$

В дальнейшем замену переменных вида (1.83) мы будем называть полиномиальной аппроксимацией профиля засветки, поскольку преобразование элемента объема воздействует на сам вид функции  $B(\mathbf{r})$ . Вид полинома был найден эмпирически для наилучшей подгонки теоретической модели (1.82) к экспериментальным данным [28]. Можно показать, что при  $d\mathbf{r}/dx = A_0 \sqrt{x}$  замена переменных вида (1.83) эквивалентна гауссовской аппроксимации профиля засветки.

Приведенная теория обеспечивает получение несмещенных оценок концентрации и яркости только для времен наблюдения, значительно меньших диффузионного времени (для выполнения условия постоянства интенсивности флуоресценции за время  $T$ ). Для преодоления предела малых  $T$  и одновременного определения яркостных и диффузионных параметров вещества метод FIDA был расширен путем включения коррекции на диффузию [89]. Среди дальнейших модификаций метода FIDA можно отметить 2D FIDA (метод анализа двумерного РЧФ) [90], FILDA (метод совместного анализа РЧФ и кинетики затухания флуоресценции) [91] и ME FIDA (метод нахождения параметров вещества путем решения системы дифференциальных уравнений кинетики процессов возбуждения и релаксации молекул) [92].

**РСН.** Совместное распределение вероятности  $P(n)$  получения  $n$  фотонов от  $m$  молекул, находящихся в закрытом объеме  $V_0$  за время наблюдения  $T$ , вычисляется с помощью свертки  $m$  одномолекулярных распределений  $p^{(1)}(n)$

$$p^{(1)}(n) = \frac{1}{V_0} \int_{V_0} Poi(n, qTB(\mathbf{r})) d\mathbf{r}, \quad (1.85)$$

где  $Poi(n, \lambda)$  обозначает распределение Пуассона с параметром  $\lambda$  [27]. Для открытого объема со средним количеством частиц  $\langle C \rangle > V_0$  общее распределение  $P(n)$  будет взвешенным средним [27]

$$P(n) = \sum_{m=0}^{\infty} p^{(m)}(n) Poi(m, \langle C \rangle V_0), \quad (1.86)$$

$$p^{(0)}(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0, \end{cases} \quad (1.87)$$

$$p^{(m)}(n) = \underbrace{p^{(1)} \otimes p^{(1)} \otimes \dots \otimes p^{(1)}}_{m-1 \text{ раз}}(n).$$

Более удобными для практического применения являются выражения

$$p^{(1)}(n, Q, q) = \frac{1}{QV_{ref}} \int Poi(n, qTB(\mathbf{r}))d\mathbf{r}, \quad n > 0,$$

$$p^{(1)}(0, Q, q) = 1 - \sum_{k=1}^{\infty} p^{(1)}(k, Q, q), \quad (1.88)$$

$$P(n, N, q) = \sum_{m=0}^{\infty} p^{(m)}(n, Q, q) Poi(m, QN), \quad N = \langle C \rangle V_{ref},$$

где осуществлен переход к вычислению интеграла в бесконечных пределах и произведен пересчет концентрации к числу молекул в ссылочном объеме  $V_{ref}$  (объем засветки  $V_{PSF}$  или эффективный объем  $V_{eff}$ ) [93, 94].  $B(\mathbf{r})$  аппроксимируется нормированным к единице в максимуме ( $B_0 = 1$ ) гауссовским распределением и множитель  $Q$  выбирается таким образом, чтобы значение произведения  $QV_{ref}$  было достаточно велико и приближалось к величине объема при полном охвате области засветки.

Модель в виде (1.88), (1.87) в большинстве случаев не способна точно описать экспериментальные данные, полученные с помощью однофотонного возбуждения. Это происходит главным образом вследствие значительного отличия принятой гауссовской аппроксимации от реального профиля засветки [28, 93, 94]. Вычисление интегралов  $\chi_k$  от гауссовской аппроксимации профиля засветки и от наиболее приближенной к реальности модели с учетом оптических искажений показало [94], что наибольшее отличие испытывает интеграл первого порядка. Для учета этого отличия было предложено ввести параметры подгонки  $F_k$ , определенные как относительная разность интегралов  $\chi_k$  от реальной функции профиля засветки  $B(\mathbf{r})$  и от его гауссовской аппроксимации  $\chi_{Gk}$

$$F_k = (\chi_k - \chi_{Gk}) / \chi_{Gk}, \quad k > 0. \quad (1.89)$$

В большинстве случаев для получения наилучшей подгонки модели под экспериментальные данные достаточна коррекция первого порядка (все  $F_k$  равны нулю кроме  $F_1$ ). В данном определении параметр  $F_1$  может трактоваться как отношение внефокусного излучения к излучению из фокуса [93, 94]. При неудовлетворительных результатах можно применить коррекцию второго порядка ( $F_1$  и  $F_2$  отличны от нуля).

Разложим первое выражение в (1.88) в ряд Маклорена

$$p^{(1)}(n, Q, q) = \frac{1}{n!QV_{ref}} \sum_{k=n}^{\infty} \frac{(-1)^{k-n} (qT)^k \chi_k}{(k-n)!}. \quad (1.90)$$

После подстановки (1.89) получим [94]

$$p^{(1)}(n, Q, q) = p_G^{(1)}(n, Q, q) + \frac{1}{n! Q V_{ref}} \sum_{k=n}^{\infty} \frac{(-1)^{k-n} (qT)^k F_k \chi_{Gk}}{(k-n)!}, \quad (1.91)$$

где  $p_G^{(1)}(n, Q, q)$  – одномолекулярное распределение для гауссовской аппроксимации профиля засветки. В дальнейшем мы будем обозначать метод, использующий коррекцию профиля засветки, как OFC (out-of-focus corrected) PCN.

Как и FIDA, метод PCN был модифицирован для анализа двухмерного РЧФ (2D PCN) [95] и учета диффузии (PCMN) [96]. Представленные методы FIDA и PCN не учитывают влияние мертвого времени и послеимпульсов на форму РЧФ. В [97] было показано, что мертвое время при большой интенсивности, а послеимпульсы при малой интенсивности потока фотонов существенно искажают регистрируемую характеристику. Метод PCN, и позднее FIDA, был дополнен возможностью коррекции на послеимпульсы и мертвое время непродлевающего типа [92, 97]. Применение коррекций позволило расширить диапазон изменения концентрации вещества и интенсивности лазера. Недостатком разработанной коррекции является существенное увеличение времени анализа данных.

Методы FIDA и PCN используют МНК для нахождения оценок параметров теоретической модели. В качестве критерия качества подгонки применяется взвешенный критерий  $\chi^2$  и визуальная оценка взвешенных остатков. Весовые коэффициенты вычисляются исходя из предположения биномиального распределения каждой точки экспериментально полученного РЧФ [27, 28]. Поскольку интервал наблюдения обычно выбирается достаточно малым, значения частот выпадения 0, 1, 2, ... фотонов составляют обычно несколько порядков и, следовательно, применение критерия  $\chi^2$  является оправданным.

Существующие методы анализа РЧФ имеют ряд недостатков. Применяемая коррекция профиля засветки в методе PCN характеризуется зависимостью количества молекул и яркости от параметров коррекции, а в методе FIDA – к потере чувствительности при разрешении многокомпонентных систем.

### 1.4.3 Анализ кумулянтов распределения числа фотоотчетов

Исторически первой попыткой получения информации о веществе на основе его яркостных характеристик было применение метода моментов для анализа данных ФФС. Смесь двух видов молекул, значительно отличающихся по яркости, была успешно разрешена путем решения нелинейной системы уравнений, соединяющей теоретические и экспериментальные моменты [87, 88]. Было показано, что кумулянты интегрированной интенсивности флуоресценции или факториальные кумулянты (ФК)  $K_k$  числа фотонов, полученных за время наблюдения  $T$ , пропорциональны произведению концентрации молекул и их характеристической яркости  $q$  [87, 88, 29]

$$K_k = \chi_k \sum_i \langle C \rangle_i (q_i T)^k, \quad (1.92)$$

где  $\chi_k$  вычисляются по формуле (1.70). В кумулянтном анализе множитель  $\chi_k$  определяет влияние профиля засветки и, соответственно, получил название геометрического фактора. На практике применяется следующая формула [29]

$$K_k = \gamma_k \sum_i N_{PSF_i} (q_i T)^k, \quad (1.93)$$

где интегралы  $\chi_k$  заменены на безразмерные гамма-факторы

$$\gamma_k = \chi_k / V_{PSF}, \quad \gamma_1 = V_{PSF} / V_{PSF} = 1, \quad (1.94)$$

и проведен пересчет концентрации к числу молекул  $N_{PSF} = \langle C \rangle V_{PSF}$  в приведенном объеме  $V_{PSF}$ , численно равному значению  $\chi_1$ . Функция профиля засветки аппроксимируется нормированным к единице в максимуме ( $B_0 = 1$ ) гауссовским распределением [29]. Для гауссовской аппроксимации гамма-факторы любого порядка могут быть рассчитаны по формуле

$$\gamma_{Gk} = 1/k^{3/2}. \quad (1.95)$$

Метод моментов не дает возможности вычислить доверительные интервалы оцениваемых параметров модели и оценить качество подгонки. Для преодоления этого недостатка в [29] было предложено вычислять избыточное количество кумулянтов и применять МНК для получения оценок параметров модели. Весовые факторы (среднеквадратические отклонения) кумулянтов находятся по методу моментов от моментов [29, 98].

Выражение (1.93) было получено в предположении постоянства характеристической яркости молекул за время наблюдения  $T$ , что справедливо только для времен наблюдения много меньших среднего диффузионного времени. Теория для времен произвольной длительности (метод Time Integrated Fluorescent Cumulant Analysis – TIFCA) вместе с коррекцией на непродлевающее мертвое время и послеимпульсы была недавно разработана и успешно применена для анализа ФК РЧФ [99, 100]. Для коррекции профиля засветки в методе TIFCA применяется подгонка  $\gamma$ -факторов (интегрированных характеристик профиля засветки) вместе с другими параметрами [99]. Недостатком подгонки  $\gamma$ -факторов является зависимость числа параметров модели от числа анализируемых кумулянтов и невозможность проведения анализа отдельного набора кумулянтов без постановки калибровочного эксперимента.

## 1.5 Выводы

1. Существующие методы анализа распределения числа фотоотсчетов и кумулянтов распределения числа фотоотсчетов имеют ряд недостатков. Недостатком метода FIDA является слабая чувствительность при разрешении

многокомпонентных систем, так как увеличение числа варьируемых параметров увеличивает число степеней свободы. Применение в РСН последовательных сверток значительно увеличивает время анализа данных, а введенная коррекция профиля засветки в предположении  $B_0 = 1$  приводит к зависимости яркости и числа молекул в объеме наблюдения от параметров коррекции. Выбранная коррекция профиля засветки в методе TIFSA приводит к зависимости числа параметров модели от числа экспериментальных кумулянтов. Число параметров становится больше числа экспериментальных точек, что ведет к получению бесконечного количества решений при анализе отдельного набора кумулянтов.

2. Методы ФФС, решая одну и ту же задачу оценивания параметров многокомпонентных систем, различаются в использовании аппроксимаций профиля засветки и нормализаций, что приводит к получению различающихся оценок параметров. Как следствие, глобальный (обобщенный) анализ [101] данных ФФС с применением рассмотренных методов также не представляется возможным. Это требует нахождения формул пересчета параметров, позволяющих корректно сравнивать результаты применения различных методов анализа экспериментальных данных.

3. Существующие методы ФФС работают эффективно лишь в узком диапазоне оцениваемых параметров, что требует априорных знаний об исследуемом веществе и параметрах, характеризующих установку. Для получения достоверных оценок параметров и стандартизации процедуры анализа (особенно важно при пакетной обработке большого числа данных) требуется разработка начальных приближений.

4. Создание адекватных моделей для анализа данных ФФС требует детального изучения физики прибора и влияния отдельных его частей на поток фотонов (регистрируемую характеристику). Аналитическое и имитационное моделирование случайных потоков со схожими характеристиками и систем их преобразования может способствовать правильной интерпретации полученных результатов и поиску оптимальных режимов измерения, минимизирующих влияние инструментальных искажений.

5. Для получения всесторонней информации о веществе требуется проведение большого количества измерений. Объем данных может быть очень велик. Для повышения эффективности анализа данных требуется разработка современных систем обработки и хранения данных.

## ГЛАВА 2

### МОДЕЛИРОВАНИЕ СИСТЕМ РЕГИСТРАЦИИ И ОБРАБОТКИ СЛУЧАЙНЫХ ПОТОКОВ

#### 2.1 Аналитическое моделирование систем преобразования случайных потоков

##### 2.1.1 Последовательное преобразование случайных потоков

Модель СРОД может быть представлена в виде совокупности последовательно расположенных элементарных преобразователей. В этом случае аналитическая модель системы формируется на основе аналитических моделей отдельных ее компонентов. В работе рассматривается метод анализа систем преобразования потоков событий, базирующийся на использовании совместных ПФЛ исходного и преобразованного потоков, с одной стороны, и УПФЛ компонента преобразования, с другой.

Рассмотрим простейшую систему из двух последовательно соединенных преобразователей. На вход первого преобразователя системы поступает поток  $A$ , заданный на интервале  $\Phi$ , ПФЛ которого –  $L^A[u; \Phi]$ . Поток  $B$ , появляющийся на выходе первого преобразователя на интервале  $\Omega$ , будет входным для второго преобразователя. Его ПФЛ обозначим через  $L^B[v; \Omega]$ . ПФЛ потока  $C$ , появляющегося на выходе системы на интервале  $\Theta$ , обозначим соответственно через  $L^C[w; \Theta]$ . Изменение потока, прошедшего первый преобразователь, можно полностью описать УПФЛ  $L^{B|A}[v; \Omega | t_1^A, \dots, t_m^A; \Phi]$ , полученным при условии, что события исходного потока наступили в моменты времени  $t_1^A, \dots, t_m^A$ . Аналогично для второго преобразователя имеем  $L^{C|B}[w; \Theta | t_1^B, \dots, t_n^B; \Omega]$ .

Запишем совместный производящий функционал потоков  $A$  и  $B$

$$L^{AB}[u; \Phi; v; \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Phi^m} \pi_m^A(t_1^A, \dots, t_m^A; \Phi) \prod_{i=1}^m [1 + u(t_i^A)] L^{B|A}[v; \Omega | t_1^A, \dots, t_m^A; \Phi] dt_1^A \dots dt_m^A. \quad (2.1)$$

Плотности  $\pi_m^A(t_1^A, \dots, t_m^A; \Phi)$  можно получить при помощи функционального дифференцирования ПФЛ потока  $A$

$$\pi_m^A(t_1^A, \dots, t_m^A; \Phi) = \frac{\delta^m L^A[u; \Phi]}{\delta u(t_1^A) \dots \delta u(t_m^A)} \Big|_{u(t)=-1}. \quad (2.2)$$

Тогда справедливо следующее уравнение [1-А, 2-А]

$$L^B[v; \Omega] = L^{AB}[u; \Phi; v; \Omega] \Big|_{u(t)=0} = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Phi^m} \frac{\delta^m L^A[u; \Phi]}{\delta u(t_1^A) \dots \delta u(t_m^A)} \Big|_{u(t)=-1} \times \quad (2.3)$$

$$\times L^{B|A}[v; \Omega | t_1^A, \dots, t_m^A; \Phi] dt_1^A \dots dt_m^A.$$

Уравнение (2.3) является базовым при анализе систем преобразования случайных потоков. Оно позволяет полностью описывать процесс преобразования произвольных случайных потоков. Проведя аналогичные рассуждения, получим выражение для ПФЛ выходного потока

$$L^C[w; \Theta] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Omega^n} \pi_n^B(t_1^B, \dots, t_n^B; \Omega) L^{C|B}[w; \Theta | t_1^B, \dots, t_n^B; \Omega] dt_1^B \dots dt_n^B =$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Omega^n} \frac{\delta^n L^B[v; \Omega]}{\delta v(t_1^B) \dots \delta v(t_n^B)} \Big|_{v(t)=-1} L^{C|B}[w; \Theta | t_1^B, \dots, t_n^B; \Omega] dt_1^B \dots dt_n^B. \quad (2.4)$$

Интенсивность и корреляционная функция преобразованного потока получаются при помощи функционального дифференцирования ПФЛ потока. Таким образом, зная ПФЛ исходного потока и условные производящие функционалы цепочки преобразователей, можно получить все необходимые характеристики преобразованного потока. Если преобразование потока является покомпонентно независимым (смотри (1.35)), то для нахождения моментных функций более удобной является следующая форма записи выражения (2.3)

$$L^B[v; \Omega] = L^A[(1 + u(\cdot))L[v; \Omega | \cdot] - 1; \Phi] \Big|_{u(t)=0} = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{\Phi^m} f_m(t_1^A, \dots, t_m^A) \prod_{i=1}^m ([L[v; \Omega | t_i^A] -$$

$$- 1) dt_1^A \dots dt_m^A = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{\Phi^m} \frac{\delta^m L^A[u; \Phi]}{\delta u(t_1^A) \dots \delta u(t_m^A)} \Big|_{u(t)=0} \prod_{i=1}^m (L[v; \Omega | t_i^A] - 1) dt_1^A \dots dt_m^A. \quad (2.5)$$

Примеры использования формул (2.3) и (2.5) приведены в приложении А.

### 2.1.2 Слияние потоков

Рассмотрим вначале случай, когда смешиваются два независимых потока  $A^1$  и  $A^2$  с производящими функционалами соответственно  $L^1[u^1; \Omega^1]$  и  $L^2[u^2; \Omega^2]$  (области определения могут быть разными), как показано на рисунке 2.1.

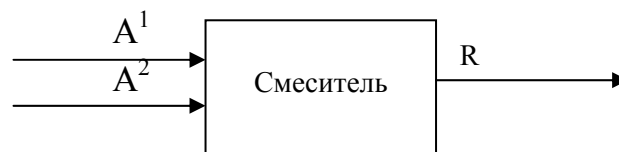


Рисунок 2.1 – Слияние независимых потоков

Так как результирующий поток  $R$  с ПФЛ  $L^R[u^R; \Omega^R]$  полностью составлен из событий смешиваемых потоков, то ПФЛ результирующего потока равен совместному ПФЛ входных потоков. Поскольку потоки 1 и 2 независимы, то  $\pi_{nm}^{12}(t_1^1, \dots, t_n^1; \Omega^1; t_1^2, \dots, t_m^2; \Omega^2) = \pi_n^1(t_1^1, \dots, t_n^1; \Omega^1) \pi_m^2(t_1^2, \dots, t_m^2; \Omega^2)$  и



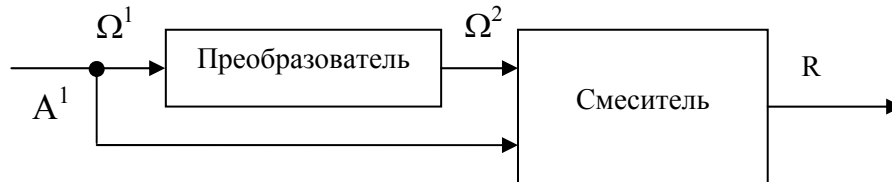
$$L^R[u; \Omega^1 \cup \Omega^2] = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{n!} \frac{1}{m!} \int_{(\Omega^1)^n} \int_{(\Omega^2)^m} \pi_n^1(t_1^1, \dots, t_n^1; \Omega^1) \pi_m^2(t_1^2, \dots, t_m^2; \Omega^2) \times$$

$$\times \prod_{i=1}^n [1 + u(t_i^1)] \prod_{j=1}^m [1 + u(t_j^2)] dt_1^1 \dots dt_n^1 dt_1^2 \dots dt_m^2 = L^1[u; \Omega^1] L^2[u; \Omega^2]. \quad (2.6)$$

Эти рассуждения остаются справедливыми и для любого числа смешиваемых потоков. В итоге получим, что ПФЛ суммарного потока равен произведению ПФЛ смешиваемых потоков

$$L^R[u; \Omega^1 \cup \Omega^2 \cup \dots \cup \Omega^k] = \prod_{i=1}^k L^i[u; \Omega^i]. \quad (2.7)$$

Для зависимых входных потоков рассмотрим случай, когда на первый вход смесителя поступает поток  $A^1$  с ПФЛ  $L^1[u^1; \Omega^1]$ , прошедшим какое-либо преобразование, заданным УПФЛ  $L|_1[u^2; \Omega^2 | t_1^1, \dots, t_n^1; \Omega^1]$ , а на второй вход поступает исходный поток до преобразования, как показано на рисунке 2.2.



**Рисунок 2.2 – Слияние зависимых потоков**

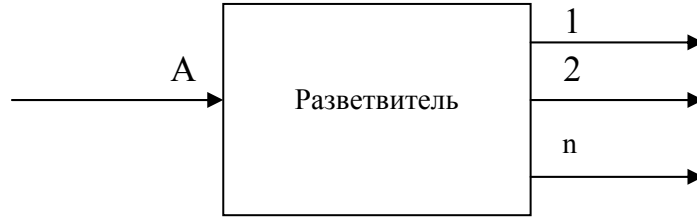
Так как выходной поток  $R$  составлен из точек смешиваемых потоков, то ПФЛ результирующего потока будет равен совместному ПФЛ входных потоков. Область задания выходного потока будет равна области задания преобразованного потока  $\Omega^2$ , которая, в свою очередь, в общем случае является расширенной областью задания входного потока. ПФЛ выходного потока примет вид

$$L^R[u; \Omega^2] = L[u; \Omega^1; u; \Omega^2] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{(\Omega^1)^n} \pi_n^1(t_1^1, \dots, t_n^1; \Omega^1) \prod_{i=1}^n [1 + u(t_i^1)] \times$$

$$\times L|_1[u; \Omega^2 | t_1^1, \dots, t_n^1; \Omega^1] dt_1^1 \dots dt_n^1. \quad (2.8)$$

### 2.1.3 Разветвление потоков

Рассмотрим общий случай разветвления потока случайных событий. На вход блока разветвления поступает поток событий, наступающих во времена  $\tau_1, \dots, \tau_m$ , заданный ПФЛ  $L[v; \Theta]$ , а на выходе мы получаем  $n$  потоков с производящими функционалами  $L_1[u_1; \Omega_1], L_2[u_2; \Omega_2], \dots, L_n[u_n; \Omega_n]$ , смотри рисунок 2.3. Каждое событие исходного потока попадает в один из выходных с некоторой вероятностью  $q_i(t)$ , в общем случае зависящей от времени появления события, причем  $q_1(t) + q_2(t) + \dots + q_n(t) = 1$ . Область задания потоков при таком преобразовании не изменяется, поэтому  $\Omega_1 = \Omega_2 = \dots = \Omega_n = \Omega = \Theta$ .



**Рисунок 2.3 – Разветвление потоков**

Запишем выражение для совместного ПФЛ входного и выходных потоков

$$L[v; u_1, u_2, \dots, u_n; \Theta] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Theta^m} \pi_m(\tau_1, \dots, \tau_m; \Theta) \prod_{i=1}^m [1 + v(\tau_i)] \times \\ \times L[u_1, \dots, u_n | \tau_1, \dots, \tau_m; \Theta] d\tau_1 \dots d\tau_m. \quad (2.9)$$

Разветвление событий является покомпонентно независимым преобразованием, поэтому

$$L[u_1, \dots, u_n | \tau_1, \dots, \tau_m; \Theta] = \prod_{j=1}^m L[u_1, \dots, u_n | \tau_j; \Theta]. \quad (2.10)$$

Плотность вероятности того, что событие, пришедшее на вход блока разветвления в момент времени  $\tau$  попадет на первый выход в момент времени  $t$  можно записать следующим образом  $\pi_{1,1}(t|\tau) = q_1(\tau)\delta(t-\tau)$ . Такие же выражения будут для остальных выходов  $\pi_{1,2}(t|\tau) = q_2(\tau)\delta(t-\tau)$ , ...,  $\pi_{1,n}(t|\tau) = q_n(\tau)\delta(t-\tau)$ . Тогда

$$\pi_1(t|\tau) = \sum_{i=1}^n \pi_{1,i}(t|\tau) = \sum_{i=1}^n q_i(\tau)\delta(t-\tau). \quad (2.11)$$

Принимая во внимание (2.10) и (2.11), получим

$$L[u_1, \dots, u_n | \tau; \Theta] = \int_{\Theta} \pi_1(t|\tau) [1 + u_1(t)] \times \dots \times [1 + u_n(t)] dt = \\ = \int_{\Theta} \pi_{1,1}(t|\tau) [1 + u_1(t)] dt + \dots + \int_{\Theta} \pi_{1,n}(t|\tau) [1 + u_n(t)] dt = \\ = q_1(\tau)(1 + u_1(\tau)) + \dots + q_n(\tau)(1 + u_n(\tau)) = 1 + q_1(\tau)u_1(\tau) + \dots + q_n(\tau)u_n(\tau). \quad (2.12)$$

Подставляя в (2.9) полученное уравнение, запишем окончательное выражение для совместного ПФЛ точек входного и выходных потоков

$$L[v; u_1, u_2, \dots, u_n; \Theta] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Theta^m} \pi_m(\tau_1, \dots, \tau_m; \Theta) \prod_{i=1}^m [1 + v(\tau_i)] \times \\ \times [1 + q_1(\tau)u_1(\tau) + \dots + q_n(\tau)u_n(\tau)] d\tau_1 \dots d\tau_m = \\ = L^A[(1 + v(\cdot))(1 + q_1(\cdot)u_1(\cdot) + \dots + q_n(\cdot)u_n(\cdot)) - 1; \Theta]. \quad (2.13)$$

Полученные модели используются для анализа конкретных ситуаций, встречающихся при проведении экспериментов в ФФС. Аналитические модели флуоресценции вещества, фонового шума детектора и рассеянного света

находятся независимо, а затем строится общая модель с использованием выражения (2.7). Разветвление потока фотонов на два с помощью полупрозрачного зеркала (показанного на рисунке 1.2) может быть учтено с помощью (2.13). Искажения потока фотонов вследствие потерь, сдвига, мертвого времени и послеимпульсов моделируются последовательностью преобразователей на основании выражения (2.4).

## **2.2 Имитационное моделирование систем регистрации случайных потоков**

### **2.2.1 Структура имитационной модели**

Рассмотрим формализованное представление СРОД в виде цепочки блоков генерации (БГ) и блоков преобразования (БП) случайных потоков. Во время работы БП вносит искажения в поток, поступающий на его вход. Искаженный поток с выхода одного БП поступает на вход очередного. Для построения имитационной модели системы регистрации случайных потоков необходимо знание формализованного описания исследуемой системы. На самом первом шаге формализованное описание преобразовывается в математическую модель. Математическую модель системы регистрации потоков можно представить в виде совокупности отдельных математических моделей следующих блоков преобразования [3-А, 4-А, 5-А, 6-А, 7-А]:

**Источник случайных потоков.** Для генерации потока необходимо знать тип потока и в большинстве случаев функцию распределения интенсивности потока (для потока Бернулли необходимо знание парциальных плотностей, а для парнокоррелированного потока и потока Кокса – корреляционную функцию). Алгоритмы генерации потоков случайных событий изложены в [4, 9, 102].

**Сдвиг моментов наступления событий.** Для моделирования смещения момента наступления события достаточно к моменту наступления события прибавить величину смещения. Величина смещения может быть получена из функции распределения по методу обратных функций [9].

**Генератор послеимпульсов.** Используется, как правило, при реализации имитационной модели детектора. Послеимпульсы можно представить потоком случайных событий спадающей либо постоянной (малой) интенсивности. Алгоритм моделирования следующий: после первого события исходного потока генерируется поток послеимпульсов и эти два потока смешиваются в один, затем производится переход к следующему событию в общем потоке и снова генерируется поток послеимпульсов. Этот процесс позволяет смоделировать эффект, когда послеимпульс сам может вызвать послеимпульсы. Процесс генерации потоков послеимпульсов завершается при достижении последнего события в результирующем потоке.

**Генератор шума.** Аддитивный шум обычно является стационарным пуассоновским (подобный случай типичен для реального физического эксперимента). Поэтому для его генерации можно выбрать генератор стационарного потока Пуассона.

**Потери событий.** Потери событий могут происходить за счет отличной от единицы вероятности регистрации события потока, поступающего на его вход (эффективность блока) или за счет мертвого времени. Потери вследствие отличной от единицы эффективности регистрации моделируются следующим образом: если реализация случайной величины, равномерно распределенной на интервале  $[0, 1]$ , меньше вероятности потерь, то событие теряется, в противном случае – нет.

Для воспроизведения мертво-временных потерь необходимо знать значение параметра  $p$  обобщенного мертвого времени, а также вид функции распределения длительности мертвого времени в данном блоке. Обобщенное мертвое время определяется следующим образом: существует конечная вероятность  $p$  того, что событие, прибывающее на блок, заблокированный вследствие мертвого времени, будет продлевать это мертвое время [63, 64]. Соответственно, если  $p = 0$ , получаем непродлевающее мертвое время, если  $p = 1$  - продлевающее. Для непродлевающего мертвого времени все события, попадающие в мертвовременной интервал, теряются и мертвое время генерируется только после зарегистрированных событий. Для продлевающего мертвого времени все события, попадающие в мертвовременной интервал, теряются, но после каждого из них генерируется мертвое время.

**Блок информации.** В данном блоке можно получить оценки интенсивности и автокорреляционной функции исследуемого потока, а также распределения числа событий на интервале наблюдения. Вычисление характеристик потока событий производится по следующим алгоритмам.

**Построение оценки интенсивности.** Процедура регистрации и накопления данных организуется следующим образом. Интервал регистрации  $[T_0, T]$  разбивается на  $M$  равных отрезков (временных каналов) точками  $t = T_0 + j\Delta t$ ,  $j=1,2,\dots,M$ ,  $\Delta t=(T - T_0)/M$ . После этого производится  $W$  повторений эксперимента, по которым определяется необходимая оценка параметров исследуемого потока. Так, если  $f(t)$  интенсивность потока в момент времени  $t$ , то ее оценка  $f$  является средней интенсивностью в  $j$ -м временном канале,  $j=[t/\Delta t]+1$  ( $[x]$  означает целую часть от  $x$ ). Для получения данной характеристики случайного потока достаточно подсчитать количество событий  $n_j$ , зарегистрированных в соответствующем временном канале. Значение  $f$  определяется по формуле

$$f_j = \frac{n_j}{W\Delta t} \quad (2.14)$$

При малых  $\Delta t$  и большом количестве интервалов разбиения  $M$  можно считать, что  $n_j$  является случайной величиной с пуассоновским распределением. Тогда среднеквадратическое отклонение  $f_j$  можно оценить как

$$\sigma(f_j) = \frac{\sigma(n_j)}{W\Delta t} \cong \frac{\sqrt{f_j W \Delta t}}{W\Delta t} = \sqrt{\frac{f_j}{W\Delta t}}. \quad (2.15)$$

Так как ширина временного канала  $\Delta t$  в ФФС берется достаточно малой, то предположение о пуассоновской статистике выполняется практически всегда.

**Построение оценки автокорреляционной функции.** Оценку АКФ  $f_2(t_1, t_2)$  случайного потока обозначим как  $f_{2ij}$ , где номер временного канала  $i$  соответствует моменту  $t_1$ , а  $j$  - моменту  $t_2$ . Процедура накопления данных и построения оценки АКФ аналогична процедуре построения оценки интенсивности потока случайных событий. Интервал регистрации  $[T_0, T]$ , разбивается на  $M$  временных каналов точками  $t = T_0 + j\Delta t, j = 1, 2, \dots, M, \Delta t = (T - T_0)/M$ . После этого проводится  $W$  повторений эксперимента. Для потоков случайных событий корреляция имеет смысл вероятности совместного выпадения событий в каналах  $i$  и  $j$ . Оценку АКФ  $f_{2ij}$  можно представить в виде

$$f_{2ij} = \frac{m_{ij}}{W\Delta t^2}, \quad (2.16)$$

где  $m_{ij}$  сумма парных произведений числа событий потока, зарегистрированных одновременно на  $i$ -ом и  $j$ -ом интервалах. Для стационарных потоков АКФ зависит только от разности аргументов  $t_1, t_2$ , и оценку АКФ можно получить по следующим формулам

$$f_{2i} = \frac{m_i}{W\Delta t^2}, \quad (2.17)$$

$$m_i = \begin{cases} \frac{1}{M} \sum_{j=1}^M n_j(n_j - 1), & i = 1, \\ \frac{1}{M - i + 1} \sum_{j=1}^{M-i+1} n_j n_{j+i-1}, & i \neq 1. \end{cases} \quad (2.18)$$

Обычно в корреляционном анализе строят только первую часть АКФ, несущую основную информационную нагрузку, поэтому в формуле (2.18) берут  $i < j$ . Поскольку в общем случае посчитать среднеквадратическое отклонение случайной величины  $m_i$  очень сложно, прибегают либо к различного рода аппроксимациям, смотри например [18, 103, 84], либо к вычислению АКФ на каждой реализации потока или на  $L$  неперекрывающихся подобластей потока и последующему вычислению стандартных отклонений стандартным способом. АКФ в виде (2.17) является ненормированной. Большой интерес представляет нормированная АКФ. Существует два вида нормализации. В первом виде из АКФ

вычитают квадрат средней интенсивности (АКФ спадает к нулю – центрированная АКФ)

$$g_{2i} = \frac{m_i}{W \Delta t^2} - \left( \frac{\sum n_i}{MW \Delta t} \right)^2, \quad (2.19)$$

во втором – делят на квадрат средней интенсивности (АКФ спадает к единице)

$$g_i = \frac{m_i}{W \Delta t^2} / \left( \frac{\sum n_i}{MW \Delta t} \right)^2. \quad (2.20)$$

**Построение оценки распределения числа событий на интервале наблюдения.** Оценку распределения числа событий на интервале наблюдения  $T$  случайного потока обозначим как  $P(n)$ , где  $n$  – число выпавших событий. Процедура накопления данных и построения оценки распределения числа событий на интервале наблюдения следующая: проводят  $W$  повторений эксперимента, при каждом повторении считают число выпавших на интервале наблюдения событий  $n$ , затем строят их гистограмму. Если число прогонов равно 1, а интервал наблюдения  $T$  много меньше интервала генерации потока, оценка распределения числа событий на интервале наблюдения производится следующим образом: интервал регистрации  $[T_0, T]$  разбивается на  $M$  равных отрезков (временных каналов), затем подсчитывается число выпавших событий  $n$  в каждом временном канале и строится их гистограмма. При необходимости производится нормализация на количество карманов  $M$  (число прогонов  $W$ ).

Распределение числа событий в  $i$ -том канале гистограммы (т.е. вероятность получить 0, 1, 2, ... событий) подчиняется биномиальному распределению

$$P_B \{ \zeta = i \} = \frac{MW!}{i!(MW-i)!} P^i(i) (1-P(i))^{MW-i}, \quad i = 0, 1, 2, \dots, MW, \quad (2.21)$$

где  $P(i)$  вероятность получения  $i$  событий. Среднеквадратическое отклонение биномиального распределения вычисляется по формуле

$$\sigma(P(i)) = \sqrt{MWP(i)(1-P(i))}. \quad (2.22)$$

Поскольку в большинстве случаев произведение  $MW$  достаточно велико (на практике когда  $MWP(i) > 30$ ), биномиальное распределение хорошо аппроксимируется нормальным распределением со средним  $MWP(i)$  и среднеквадратическим отклонением  $\sqrt{MWP(i)(1-P(i))}$ .

Имитационную модель системы регистрации случайных потоков можно представить в виде цепочки последовательно соединенных вышеупомянутых блоков преобразования. Сигнал с выхода предыдущего блока попадает на вход последующего и так далее. Первым в такой цепочке должен стоять генератор исследуемого потока событий. Если в системе преобразования присутствуют несколько независимых генераторов случайных событий (например, основного

сигнала и шума), то необходим еще один блок преобразования для смешивания событий нескольких параллельно работающих генераторов. Но и такую систему можно представить в виде цепочки блоков преобразования, размещая всякий раз после генератора потока блок смешивания событий. Если необходимо смоделировать разветвление потока, то это можно сделать с помощью блока независимых потерь. Каждый БП и блок информации может работать не на всей области задания потока, а на некоторой его подобласти.

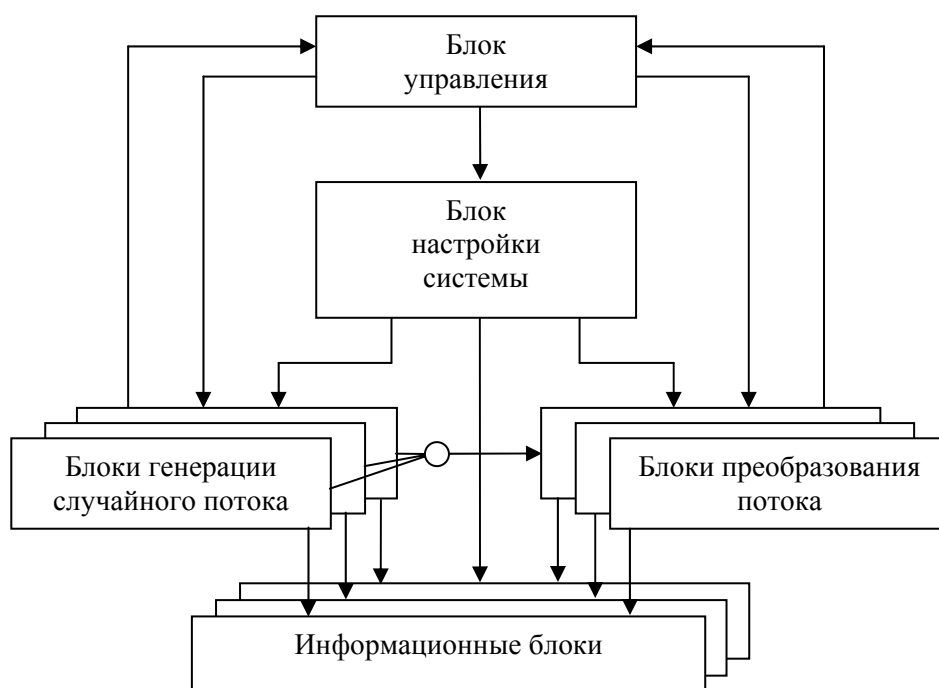
В случае наличия в системе преобразования потоков блоков с независимым действием (например: сдвиг событий и независимые потери) взаимный порядок блоков значения не имеет. Но при присутствии в системе послеимпульсов и мертвого времени порядок следования блоков имеет значение и может определяться из физики протекания реальных процессов в исследуемых системах. Выделение процесса получения оценок интенсивности, корреляционной функции и распределения числа событий в самостоятельные блоки позволяет следить за изменениями в исследуемом потоке случайных событий после каждого блока преобразования, а не только в конце цепочки.

Следующий этап создания имитационной модели заключается в преобразовании математической модели каждого БП в моделирующий алгоритм и реализации его на ЭВМ. Дополнительно необходим алгоритм настройки системы на конкретный режим генерации случайных потоков и режимы обработки потоков различными БП. Объединяющая эти алгоритмы головная программа организует работу модели в целом. На рисунке 2.4 представлена обобщенная структурная схема имитационной модели системы преобразования случайных потоков [3-А, 4-А, 5-А, 6-А, 7-А].

После каждого прогона блок информации рассчитывает заданную характеристику потока. В конце моделирования производится нормализация характеристики на число прогонов. Такой подход позволяет остановить моделирование, просмотреть вычисленные характеристики и при необходимости продолжить моделирование с точки останова.

### **2.2.2 Программная реализация имитационной модели**

На современном этапе развития технологий создания интегрированных программных комплексов, предназначенных для решения сложных научных и технических задач, важную роль играет правильная организация процесса построения программного обеспечения (ПО). Правильность изначально выбранной разработчиком структуры программного комплекса и стратегии программирования во многом предопределяет его простоту и одновременно эффективность при последующем использовании, а также возможность дальнейшего его расширения и совершенствования при минимальном количестве затрат.



**Рисунок 2.4 – Обобщенная структурная схема системы преобразования случайных потоков**

Очевидно, что многообразие областей, где применяются различные комплексы программ, делает практически невозможным выработать единый подход к созданию любого программного обеспечения. Однако можно выделить ряд общих принципов, которые необходимо соблюдать для построения программного комплекса, удовлетворяющего современные требования.

1. Модульность. Создаваемое программное обеспечение должно состоять из набора отдельных приложений и библиотек (модулей), выполняющих относительно обособленные функции в пределах всего комплекса. Данный принцип обеспечивает возможность дальнейшего расширения и наращивания возможностей программного комплекса при одновременном снижении затрат на последующую модификацию.

2. Гибкость. Разрабатываемый программный комплекс должен удовлетворять условию совместимости с другими пользовательскими приложениями, которые наиболее часто могут использоваться совместно с ним.

3. Наличие встроенных средств управления наборами данных. Данное требование становится все актуальнее ввиду необходимости управления наборами данных достаточно большого объема. Наиболее перспективным в данной области является использование баз данных.

4. Дружественный интерфейс. Хорошо продуманный и удобный интерфейс дает пользователю возможность легко и быстро настраивать программное обеспечение на выполнение конкретной задачи, а также воспринимать большое количество результатов проведенных вычислений.



Программная реализация интегрированной системы имитационного моделирования систем регистрации потоков случайных событий (Simulator) была выполнена с использованием методов объектно-ориентированного программирования в инструментальной среде Borland Builder C++ 5.0 с применением компонент библиотеки VCL [6-A, 7-A]. Система предназначена для работы в среде Windows 95/98/Me или Windows NT/2000/XP. Интерфейс программы оформлен в соответствии с широко распространенными программными продуктами фирм Microsoft и Borland. Программный комплекс имеет тип MDI (Multiply Document Interface) и позволяет одновременно загружать и держать в памяти несколько независимых моделируемых конфигураций систем преобразования (в дальнейшем моделей). Максимальное число одновременно обслуживаемых моделей определяется только ресурсами компьютера (в большей степени объемом оперативной памяти). Для активизации определенной модели достаточно выбрать соответствующее ей окно конфигурации. При создании, отладке и выполнении модели можно работать в псевдомногозадачном режиме, т.е. можно переключаться между окнами конфигурации с принадлежащими им моделями. При этом сохраняются все параметры моделей и промежуточные результаты вычислений. Поскольку время моделирования сложных систем может быть велико, головная программа может останавливать моделирование после достижения заданного числа прогонов, сохранять результаты в файл и после повторного запуска программы продолжать моделирование с точки останова. Также можно в любое время прервать и затем при необходимости продолжить моделирование с сохранением всех результатов модели.

Система моделирования позволяет получать характеристики исследуемого потока событий при его преобразовании следующими блоками преобразований:

1. Сдвиг событий (пункт меню Offset). Сдвиг событий может происходить на следующие величины: постоянную, равномерно распределенную на отрезке  $[A, B]$ , заданную функцией распределения вида  $A \exp(-At)$ ;

2. Добавление новых событий за счет аддитивного шума (пункт меню Noise);

3. Добавление новых событий за счет послеимпульсов (пункт меню After Pulses). Послеимпульсы генерируются двух видов: как поток Пуассона с малой постоянной интенсивностью и как поток Пуассона со спадающей интенсивностью вида  $A \exp(-Bt+C)$ ;

4. Потери событий вследствие обобщенного мертвого времени (продлевающее и непродлевающее мертвое время являются предельными случаями) (пункт меню Dead Time). Мертвое время может быть следующих видов: постоянной длительности, равномерно распределенным на отрезке  $[A, B]$  и заданной функцией распределения вида  $A \exp(-At)$ ;

5. Потери событий вследствие отличной от единицы эффективности процесса регистрации (пункт меню Independent Losses);

6. Смешивание потоков. Происходит автоматически при наличии в системе двух и более генераторов потоков;

7. Блок построения оценки интенсивности потока (пункт Intensity меню Information);

8. Блок построения оценки корреляционной функции (пункт Correlation Function меню Information) (в случае стационарности исследуемого потока);

9. Блок построения оценки распределения числа событий на интервале наблюдения (пункт РЧФ меню Information).

Библиотека генераторов потоков случайных событий (пункт Point Processes меню Blocks Library) содержит следующие потоки:

1. Элементарный одноточечный поток (пункт меню SPP);

2. Поток Пуассона (пункт меню Poisson) с интенсивностями: постоянной, вида  $A/(Bt+C)$ , вида  $A\exp(-Bt+C)$ ;

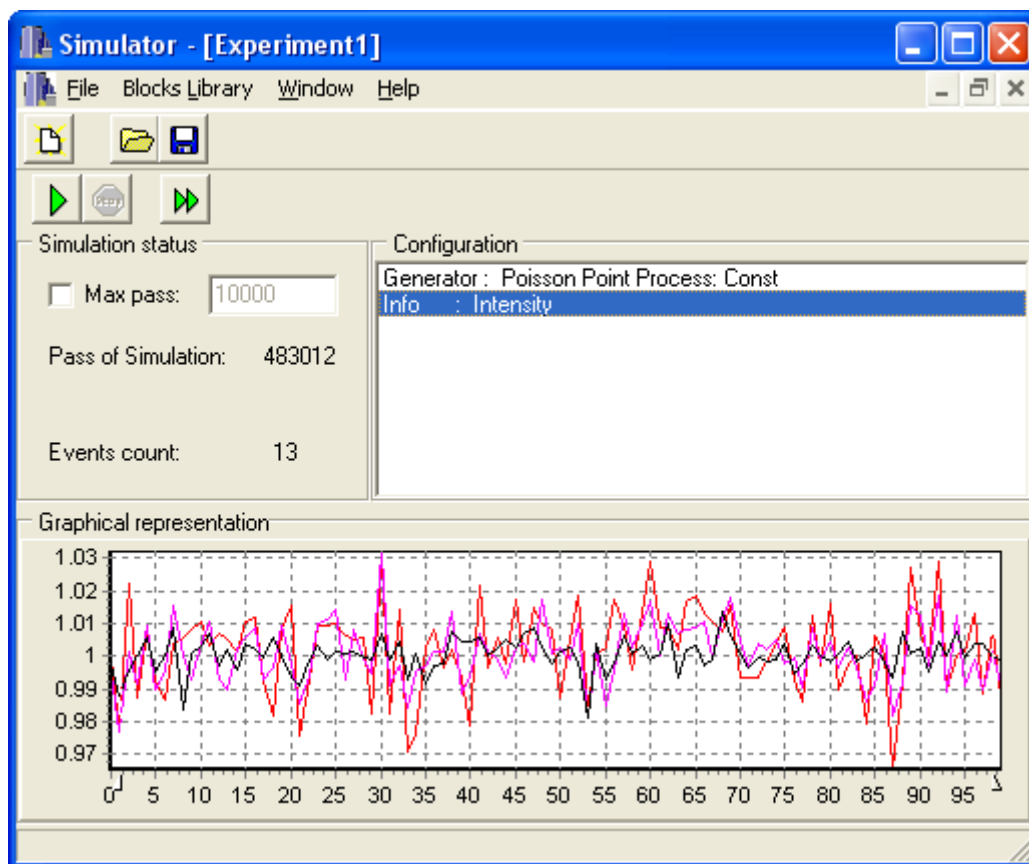
3. Отрицательно-биномиальный поток (пункт меню Negative Binomial) с интенсивностями: постоянной, вида  $A/(Bt+C)$ , вида  $A\exp(-Bt+C)$ ;

4. Поток Бернулли (пункт меню Bernoulli) порождаемого  $n$  источниками с парциальными плотностями:  $\delta$ -функции, вида  $A\exp(-At)$ , вида  $1/t$ ;

5. Поток Кокса (пункт меню Double Stochastic) с постоянной интенсивностью  $A$  и ковариационной функцией вида  $g_2(t_1, t_2) = B\exp(-C(t_2 - t_1))$ , где  $t_2 > t_1$ ,  $A > 2B/C$ , и парнокоррелированный поток с интенсивностью вида  $A^2(t)+B^2(t)$ , где  $A$  и  $B$  нормально распределенные величины с нулевым математическим ожиданием и дисперсией  $D^2$  и ковариационной функцией вида  $g_2(t_1, t_2) = D^2\exp(-Ct)$ .

В начале работы производится настройка модели на исследуемый режим. При нажатии кнопки New появляется окно настройки Configuration Window. Затем, обращаясь к пункту меню Block Library, в окно просмотра вставляются блоки в требуемой последовательности. При настройке указываются тип и параметры случайного потока и каждого блока преобразования. Для каждого блока преобразования можно задать различный диапазон действия этого блока. Таким образом, можно задать переменный режим действия для какого либо блока. Например, послеимпульсы в начале интервала моделирования можно задать одной интенсивности, в конце – другой. Эта возможность особенно необходима для задания интервала построения оценки автокорреляционной функции, так как соответствующий блок дает корректные значения только для стационарных потоков. При наличии переходных процессов рекомендуется вычислять автокорреляционную функцию с момента, когда поток входит в стационарный режим. Затем управляющий алгоритм при нажатии кнопки Run последовательно

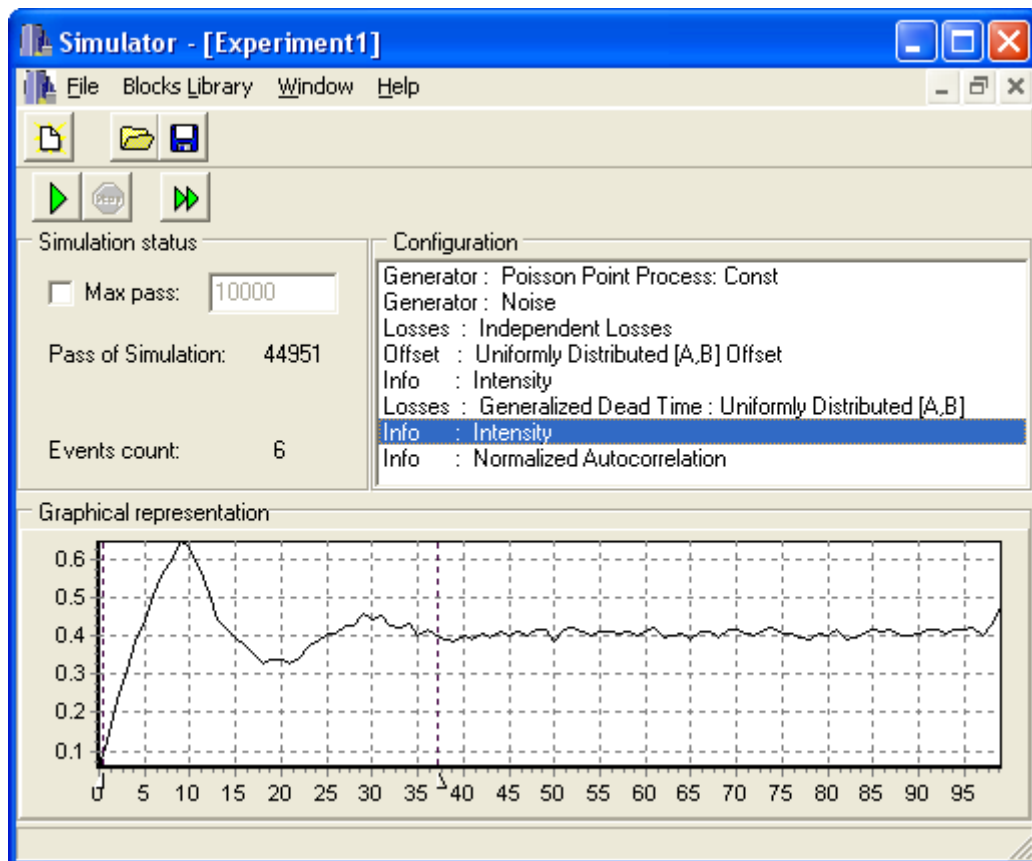
запускает в работу блоки генерации случайных потоков и блоки преобразования этих потоков. На рисунке 2.5 представлено главное окно программы.



**Рисунок 2.5 – Главное окно программного комплекса**

Реализация случайного потока от генератора проходит последовательно все блоки преобразования. Выходной поток предыдущего блока является входным для последующего. Кроме генератора потока можно добавить блок импорта потока из файла. Есть также возможность сохранить последовательность сгенерированных событий в файл.

Разработанное ПО отличается широкой функциональностью. В процессе работы с моделью можно добавлять, удалять блоки и изменять их параметры. При необходимости можно прервать выполнение программы, изменить параметры какого-либо блока, даже добавить либо удалить некоторые блоки и продолжить моделирование с прерванного места с помощью команды Continue меню Run. Например, на рисунке 2.5 представлены графики интенсивности при различном числе прогонов. Результаты эксперимента представляются в нижней части главного окна. Имеется возможность добавления нескольких кривых для их визуального сравнения. В качестве примера на рисунке 2.6 представлено главное окно программного комплекса в режиме отображения графика интенсивности. В графическом окне имеется возможность масштабирования, сдвига и точного замера временных областей с помощью передвигающихся маркеров.



**Рисунок 2.6 – Пример использования программного комплекса**

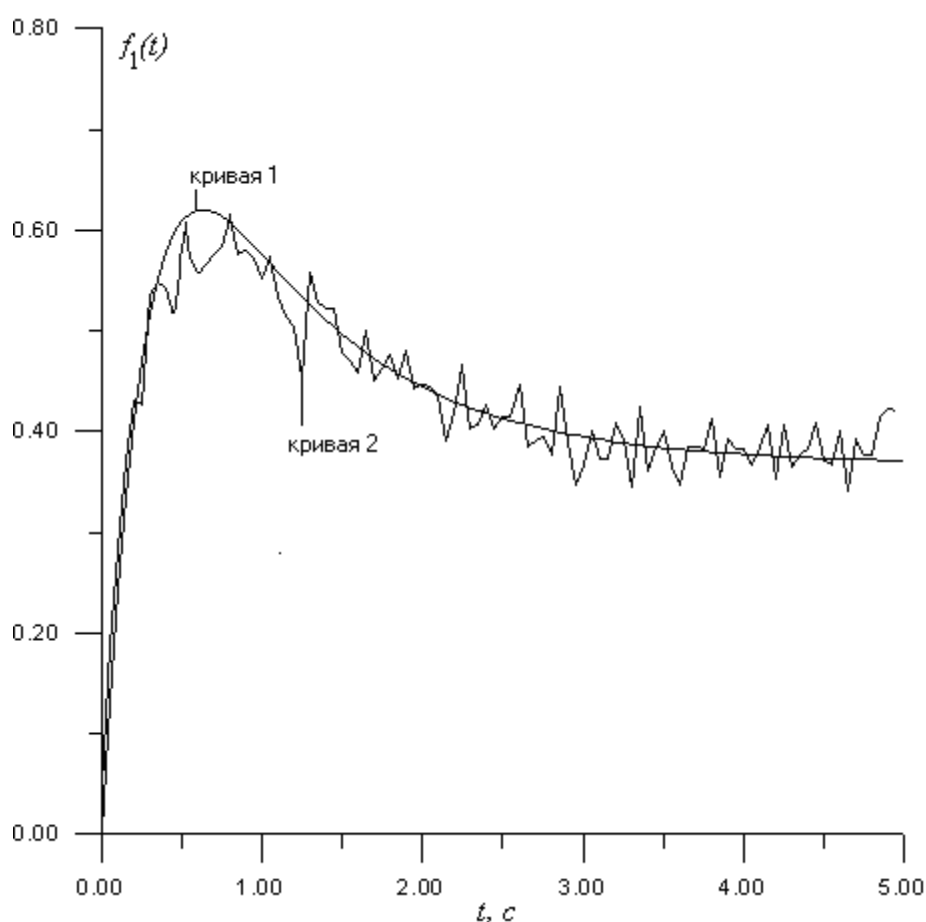
Программный комплекс разрабатывался как универсальный инструмент для имитации потоков и систем их преобразования. Его структура позволяет создавать сложные модели, состоящие из большого числа блоков, без необходимости изменения программного кода и последующей компиляции программы. Классы, отвечающие за вышеупомянутые генераторы потоков и блоки преобразования, создавались с расчетом легкого пополнения соответствующих библиотек, так как невозможно предусмотреть все типы и виды функций, по которым генерируются и преобразуются потоки. В таком виде созданный программный комплекс может служить удобным инструментом в руках исследователя или как основа лабораторного практикума.

### **2.2.3 Вычислительный эксперимент**

Алгоритмы генерации потоков и систем их преобразования прошли различные тесты на корректность их функционирования. Примером применения такой процедуры тестирования является следующий эксперимент.

В приложении А было получено аналитическое выражение (А.9) для интенсивности выходного потока, искаженного системой преобразования потоков случайных событий, состоящей из трех последовательно соединенных блоков преобразования: независимых потерь, линейного сдвига и продлевающего мертвого времени. Было проведено имитационное моделирование такой системы с

помощью разработанного программного комплекса. Для 25000 прогонов на интервале  $[0, 5]$  и  $a = 1, b = 3, c = 1, r = 0.95$  выходная интенсивность показана на рисунке 2.7 (кривая 2).



Кривая 1 - интенсивность, рассчитанная по формуле (А.9);

Кривая 2 – интенсивность, полученная с помощью имитационной модели

**Рисунок 2.7 – Интенсивность выходного потока**

Рассчитанная по формуле (А.9) с этими же параметрами выходная интенсивность  $f_1^D(t)$  представлена кривой 1. Как видно из рисунка, имеется хорошее совпадение результатов, полученных с помощью аналитической модели и имитационного моделирования. С момента времени  $t > 3.5$  сек поток входит в стационарный режим. Таким образом, разработанный программный комплекс позволяет исследовать случайные потоки не только в стационарных, но и на переходных режимах. Сравнение с результатами аналитической модели свидетельствует об адекватности разработанной имитационной модели системы регистрации.

## 2.3 Имитационное моделирование систем флуктуационной спектроскопии

### 2.3.1 Моделирование потока фотонов с заданной автокорреляционной функцией

Для тестирования разрабатываемых моделей и методов анализа данных ФФС необходимо имитационное моделирование потока фотонов на выходе СРОД с целью получения анализируемой характеристики (АКФ, РЧФ, ФК РЧФ) с заданными параметрами. Моделирование позволяет исследовать сходимость и устойчивость метода анализа при различных значениях параметров модели и оценить среднеквадратические отклонения и смещение параметров от их истинных значений при различных уровнях отношения сигнал/шум.

Математической моделью потока случайных событий с постоянной интенсивностью и заданной автокорреляционной функцией является парнокоррелированный поток (смотри раздел 1.2). Парнокоррелированный поток полностью определяется своими моментными функциями первого  $f_1(t)$  и второго порядка  $f_2(t_1, t_2)$ , то есть интенсивностью и корреляционной функцией. Существует метод моделирования такого потока, основанный на представлении его в виде потока группированных точек [102].

Пусть необходимо сгенерировать моменты наступления событий стационарного парнокоррелированного потока с интенсивностью  $f_1(t) = I$  и центрированной АКФ  $g_2(t_1, t_2) = f_2(t_1, t_2) - f_1(t_1)f_1(t_2) = g_2(\tau) = f_2(\tau) - I^2$ ,  $\tau = t_2 - t_1$ , на временном интервале  $[0, T_m]$ . Алгоритм моделирования воспроизводит механизм образования потока группированных точек. Для этого вначале получаем реализацию пуассоновского потока базовых событий с интенсивностью

$$f_1^A(t) = I - \int_0^t g_2(\tau) d\tau. \quad (2.23)$$

Затем для событий базового потока генерируем смещенное событие в соответствии с распределением, имеющим условную плотность, задаваемую выражением

$$\pi(t_2 | t_1) = \frac{g_2(t_1, t_2)}{f_1^A(t_1)}. \quad (2.24)$$

Вероятность наступления смещенного события на интервале  $[t_1, T_m]$  определяется как

$$\varphi(t_1) = \int_{t_1}^T \pi(x | t_1) dx = \frac{1}{f_1^A(t_1)} \int_0^{T-t_1} g_2(\tau) d\tau. \quad (2.25)$$

Следовательно, момент появления смещенного события можно определить по нормированной плотности распределения вероятности наступления смещенного события на интервале  $[t_1, T]$  при условии, что базовое событие произошло в момент  $t_1$ ,

$$p(t_2 | t_1) = \frac{\pi(t_2 | t_1)}{\varphi(t_1)} = \frac{g_2(t_2 - t_1)}{\int_0^{T_m - t_1} g_2(\tau) d\tau}. \quad (2.26)$$

В противоположном случае событие базового потока не порождает коррелированного события в данной реализации. Результирующий поток представляет собой неразличимую смесь событий потока базовых точек и потока смещенных точек.

В FCS используется АКФ с нормализацией вида  $g(\tau) = f_2(\tau)/I^2$  [84, 73, 104], поэтому требуется соответствующая модификация формул (2.23) – (2.26). Выражения для  $g(\tau)$  и  $g_2(\tau)$  связаны следующим соотношением

$$g_2(\tau) = g(\tau)I^2 - I^2 = (g(\tau) - 1)I^2. \quad (2.27)$$

Обозначив  $g(\tau) - 1$  через  $K(\tau)$ , получим

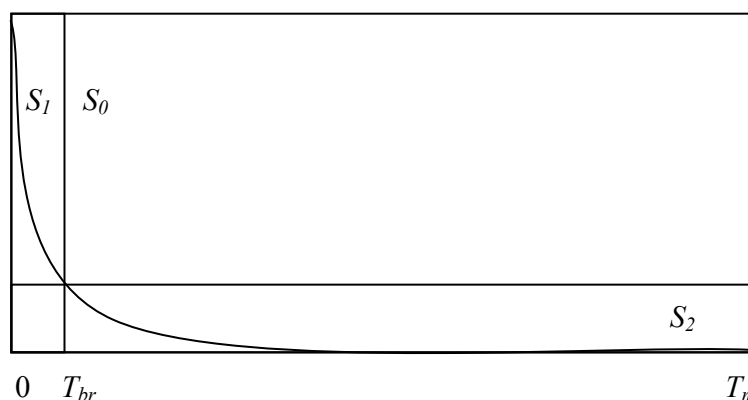
$$f_1^A(t) = I - I^2 \cdot \int_0^t K(\tau) d\tau, \quad \varphi(t_1) = \frac{I^2}{f_1^A(t_1)} \int_0^{T_m - t_1} K(\tau) d\tau, \quad p(t_2 | t_1) = \frac{K(t_2 - t_1)}{\int_0^{T_m - t_1} K(\tau) d\tau}. \quad (2.28)$$

Генерация смещенного события проводится по методу прореживания [9]. Для этого генерируется случайная точка, равномерно распределенная в отрезке  $[0, T_m]$ , затем вычисляется значение функции (2.26) в этой точке и проверяется, меньше ли полученное значение дополнительно генерируемой случайной величины, равномерно распределенной в отрезке  $[0, \max(p(t_2 | t_1))]$ . Если меньше, то временная точка принимается. Если нет, то алгоритм генерации смещенного события повторяется. Эффективность метода прореживания зависит от соотношения площадей ниже и выше кривой функции (2.26). Функция  $K(\tau)$  (смотри (1.73)) имеет гиперболический вид и спадает очень быстро, поэтому прямое применение алгоритма прореживания приводит к большим временным затратам. Найти удобное для генерации (подобно равномерному) распределение, имеющее плотность того же вида, что и (1.73), но лежащее выше на всем интервале моделирования при любых сочетаниях параметров (1.73), не удалось. Для сокращения временных затрат и повышения эффективности метода прореживания использован подход, заключающийся в разбиении временного интервала на два участка с целью минимизации бесполезных попыток генерации смещенного события. Вначале находится такая точка  $T_{br}$  на временной оси, для которой площадь  $S_0$  (смотри схему, поясняющую принцип ускорения

моделирования на рисунке 2.8) максимальна. Для этого производится поиск минимума функции

$$-(\max(p(t_2 | T_{br})) - p(t_2 | T_{br}))(T_m - T_{br}) = 0. \quad (2.29)$$

Затем генерируется реализация базовой случайной величины и, если оно меньше, чем  $S_1/(S_1 + S_2)$ , то производится генерация смещенного события в области  $S_1$ , если нет – то в области  $S_2$ . Таким образом, достигается выигрыш, равный  $1 + S_0/(S_1 + S_2)$ .



**Рисунок 2.8 – Схема, поясняющая метод прерывания**

Отношение  $S/N$  моделируемой характеристики зависит от интенсивности  $I$ , времени  $T_m$  и соотношения  $T_m$  и  $\tau_{diff}$  ( $S/N$  выше, когда смещенное событие в большинстве случаев происходит раньше следующего события исходного потока).

Проверка адекватности алгоритма генерации потока фотонов с заданной АКФ проводилась с помощью программы FCS Data Processor [9-A, 10-A, 11-A, 12-A], разработанной на кафедре системного анализа БГУ в сотрудничестве с центром микроскопии университета Вагенингена (Microscopy Center, WU, Wageningen, The Netherlands). В настоящее время FCS Data Processor используется в мире для анализа данных ФФС рядом научных лабораторий [105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 10-A]. В частности FCS Data Processor позволяет вычислять АКФ потока фотоотчетов и проводить анализ данных по методу FCS.

Тестируемый алгоритм реализован в программе FCS Data Processor. Программа написана на языке C++ в инструментальной среде Borland Builder 5.0 ([www.borland.com](http://www.borland.com)). Тестирование заключалось в генерации случайного потока с заданными свойствами, вычисления АКФ и последующего анализа по методу FCS. Параметры моделирования следующие:  $N_{eff} = 2.25$ ,  $\tau_{diff} = 5E-5$  с,  $a = 10$  (смотри выражение для АКФ в форме (1.73)),  $I = 4000$ ,  $T_m = 600$  с. В результате было получено около  $2.4E6$  событий. Результаты анализа АКФ по методу FCS сведены в таблицу 2.1 и представлены на рисунке 2.9. Среднеквадратические отклонения параметров находились по методу асимптотических стандартных ошибок (АСО) (asymptotic standard errors) [116].



Таблица 2.1 – Результаты анализа АКФ по методу FCS.

Параметр	Истинные значения	Оценки, полученные по методу FCS
$N_{eff}$	2.25	$2.23 \pm 0.02$
$\tau_{diff} (\times 1E-5) \text{ с}$	5	$5.11 \pm 0.08$
$a$	10	$10.07 \pm 0.80$

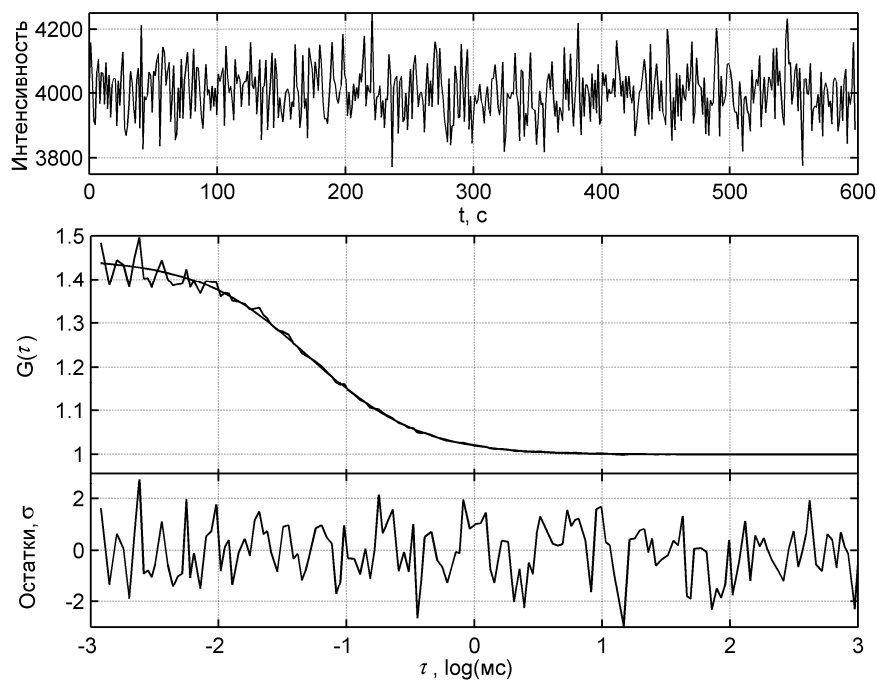


Рисунок 2.9 – Результаты тестирования имитационной модели потока с заданной АКФ ( $\chi^2 = 1.13$ )

Как видно из представленных результатов (интенсивность постоянна со средним значением, равным заданному; оценки параметров близки к истинным; вид остатков и значение критерия  $\chi^2$  свидетельствуют о хорошей подгонке к экспериментальным данным), имитационная модель позволяет генерировать стационарный поток с автокорреляционной функцией вида (1.73).

### 2.3.2 Моделирование фотоотсчетов

Основными процессами, подлежащими моделированию, являются явление флуоресценции и диффузия молекул в растворе. Затухание флуоресценции происходит за время, много меньшее интервала наблюдения  $T$ . Поэтому будем считать, что молекула немедленно переизлучает поглотившийся фотон и интенсивность регистрируемой флуоресценции прямо пропорциональна интенсивности падающего излучения. Интенсивность флуоресценции определяется значением функции профиля засветки  $B(\mathbf{r})$ . Наиболее часто для ее аппроксимации применяют асимметричное гауссовское распределение (1.2). Будем также считать молекулы, находящиеся в области засветки в течение короткого интервала наблюдения, неподвижными.

Количество фотонов, излученных одиночной молекулой за время  $T$ , хорошо аппроксимируется распределением Пуассона [7] со средним, равным произведению характеристической яркости молекулы  $q$  (измеряемой в числе излученных фотонов одной молекулой в единицу времени), интервала наблюдения  $T$  и функции профиля засветки  $B(\mathbf{r})$

$$\lambda_{fluor} = qTB(\mathbf{r}). \quad (2.30)$$

В свою очередь, число молекул, находящихся в растворе в открытом объеме, подчиняется распределению Пуассона со средним равным среднему числу молекул  $\bar{N}$  в этом объеме. Следовательно, для моделирования числа фотонов, излучаемых молекулами одного вида, в течение интервала  $T$  достаточно получить число молекул, попавших в освещенную область, путем генерации случайной величины, подчиняющейся распределению Пуассона с параметром  $\bar{N}$ ; затем для каждой молекулы сгенерировать ее местоположение в объеме засветки (по равномерному закону); и в заключение получить число излученных фотонов (от каждой молекулы) путем генерации случайной величины, подчиняющейся распределению Пуассона с параметром  $\lambda_{fluor}$ . Если моделируется смесь молекул разных видов, то необходимо повторить цикл генерации фотонов для всех видов молекул и сложить все полученные фотоны со всех молекул всех видов вместе. Затем цикл генерации фотонов для всех видов молекул повторяется  $m$  раз для получения РЧФ с заданным отношением  $S/N$  ( $m$  рассчитывается по формуле (2.37)).

При необходимости моделирования потока фотонов с заданной статистикой фотоотсчетов на последнем этапе вместо непосредственной генерации числа фотоотсчетов от одиночной молекулы можно генерировать поток Пуассона с интенсивностью  $\lambda_{fluor}$ . Выходной поток будет представлять собой суперпозицию потоков фотонов от всех молекул каждого вида. Учитывая, что суперпозиция потоков Пуассона является потоком Пуассона с суммарной интенсивностью, можно вначале рассчитать суммарную интенсивность, а затем произвести генерацию потока.

В ФФС наиболее часто применяют нормализацию к эффективному объему засветки  $V_{eff}$ . Тогда исходными параметрами модели будут среднее число молекул  $N_{eff}$ , их характеристическая яркость  $q$  и отношение  $S/N$ . Если моделируются фотоотсчеты от многокомпонентного образца, то указывается  $q_i$  и  $N_{eff_i}$  каждой компоненты. Подлежащим определению в разрабатываемой модели является только среднее количество молекул каждой компоненты  $\bar{N}$  в объеме наблюдения  $V$ , эквивалентное среднему числу молекул  $N_{eff}$  в объеме  $V_{eff}$ . Для расчета  $\bar{N}$  определим величину объема  $V$ . В нашем случае асимметричное гауссовское распределение (1.2) сводится к симметричному распределению (смотри приложение D)

$$B(\mathbf{r}) = B_0 e^{-2r^2}, \quad (2.31)$$

где  $r^2 = x^2 + y^2 + z^2$  длина радиус-вектора  $\mathbf{r}$ . Объем засветки  $V_0$  радиуса  $r_0$  равен

$$V_0 = 4\pi r_0^3 \omega_0^2 z_0 / 3. \quad (2.32)$$

Для приведения значения яркости в соответствие со значением параметра  $q$  в методе РСН положим  $B_0 = 1$ . Тогда среднее количество молекул в объеме наблюдения  $V_0$  будет равно

$$\bar{N} = \langle C \rangle V_0 = \frac{N_{eff} V_0}{V_{eff}} = N_{eff} \frac{4\pi r_0^3 \omega_0^2 z_0}{3\pi^{3/2} \omega_0^2 z_0} = N_{eff} \frac{4r_0^3}{3\sqrt{\pi}}. \quad (2.33)$$

Радиус  $r_0$  должен быть таким, чтобы полностью включать всю область засветки. Но увеличение радиуса приводит к кубическому росту числа моделируемых молекул и к кубическому росту времени моделирования. Для нахождения оптимального радиуса введем относительную ошибку  $\delta$ , определенную как

$$(1 - \delta) = \int_0^{r_0} \iiint B(\mathbf{r}) d\mathbf{r} / V_{eff}. \quad (2.34)$$

Тогда  $r_0$  можно найти из решения уравнения (2.34) при заданной ошибке  $\delta$ .

Генерацию местоположения молекулы внутри освещенного объема (в нашем случае шара радиусом  $r_0$ ) можно провести следующим образом: вначале сгенерировать равномерно распределенные в отрезке  $[-r_0; r_0]$  координаты молекулы  $x, y, z$ ; затем проверить, попадает ли молекула внутрь шара по формуле  $\sqrt{x^2 + y^2 + z^2} \leq r_0$ ; и, при необходимости, повторять генерацию до тех пор, пока молекула не попадет внутрь шара.

Получить РЧФ с заданным отношением  $S/N$  можно также с использованием метода, основанного на вычислении теоретического распределения с заданными параметрами и добавлении подходящего шума к каждой точке полученной кривой. Этот метод широко используется для моделирования кривых затухания флуоресценции [117] и других характеристик при известных статистических свойствах шума моделируемой характеристики.

Для применения данного метода (в дальнейшем будем называть его прямым методом генерации РЧФ) необходимо знать распределение шума моделируемой характеристики. Статистический шум РЧФ хорошо аппроксимируется Биномиальным распределением [27]

$$P\{\zeta = l\} = \frac{m!}{l!(m-l)!} p^l (1-p)^{m-l}, \quad l = 0, 1, 2, \dots, m, \quad (2.35)$$

где  $p$  – вероятность получения  $n$  фотонов в течение интервала наблюдения  $T$  и  $m$  – число интервалов наблюдения. Среднее  $\mu$  и дисперсия  $\sigma^2$  Биномиального распределения вычисляются по формулам  $\mu = mp$  и  $\sigma^2 = mp(1-p)$  соответственно. Таким образом, алгоритм генерации РЧФ будет заключаться в вычислении теоретической кривой  $P(i)$  (для метода FIDA это выражения (1.82), (1.21), для метода РСН – (1.91), (1.86)), затем для каждого  $i$  генерации случайного числа,

имеющего Биномиальное распределение с заданным  $m$  и  $p = P(i)$  [7-А]. Для больших  $m$  применение выражения (2.35) не является целесообразным. Для значений  $mP(i) > 30$  нормальное распределение со средним  $mP(i)$  и стандартным отклонением  $\sqrt{mP(i)(1-P(i))}$  хорошо аппроксимирует Биномиальное распределение [35]. Для воспроизведения реальных измерений полученные значения должны быть округлены к ближайшему целому. Для меньших значений  $mP(i)$ , где  $m$  остается большим, а  $P(i)$  мало [35], можно использовать распределение Пуассона с параметром  $mP(i)$ .

Число интервалов наблюдения  $m$  определяет отношение  $S/N$ . Если определить  $S/N$  как отношение среднего значения в максимуме генерируемой кривой к стандартному отклонению в этой точке  $S/N = \mu/\sigma$  [18], получим

$$S/N = \sqrt{mp_{max}/(1-p_{max})} \quad (2.36)$$

где  $p_{max} = \max_i(P(i))$  и, следовательно,

$$m = (S/N)^2(1-p_{max})/p_{max} \quad (2.37)$$

На практике выражение (2.37) неудобно использовать из-за необходимости предварительного вычисления  $p_{max}$ . Для упрощения можно использовать максимальное значение моделируемой характеристики ( $V_{max}$ ) для грубой оценки (снизу) отношения  $S/N$ . В этом случае  $S/N \approx \sqrt{V_{max}}$ .

Для исследования моделей и методов анализа РЧФ была разработана программа Data Analyser. Программа написана на языке С++ в инструментальной среде Borland Builder 5.0. Главное окно программы представлено на рисунке 2.10. Программа позволяет генерировать РЧФ (прямой метод генерации) и факториальные кумулянты РЧФ (через промежуточную генерацию РЧФ), импортировать измеренные характеристики и проводить их анализ по методу наименьших квадратов. В библиотеку теоретических моделей программы входят модели, полученные для методов FIDA, PCH, FCA с различными видами профилей засветки и нормализации. Соответственно, разработаны и программно реализованы алгоритмы анализа данных по методам FIDA, PCH, FCA. Программа позволяет генерировать начальные приближения (смотри разделы 3.3 и 4.3) и сравнивать результаты двух моделей. Можно задавать начальную и конечную точку анализируемой характеристики, задавать граничные значения для параметров модели и фиксировать их в известные значения.

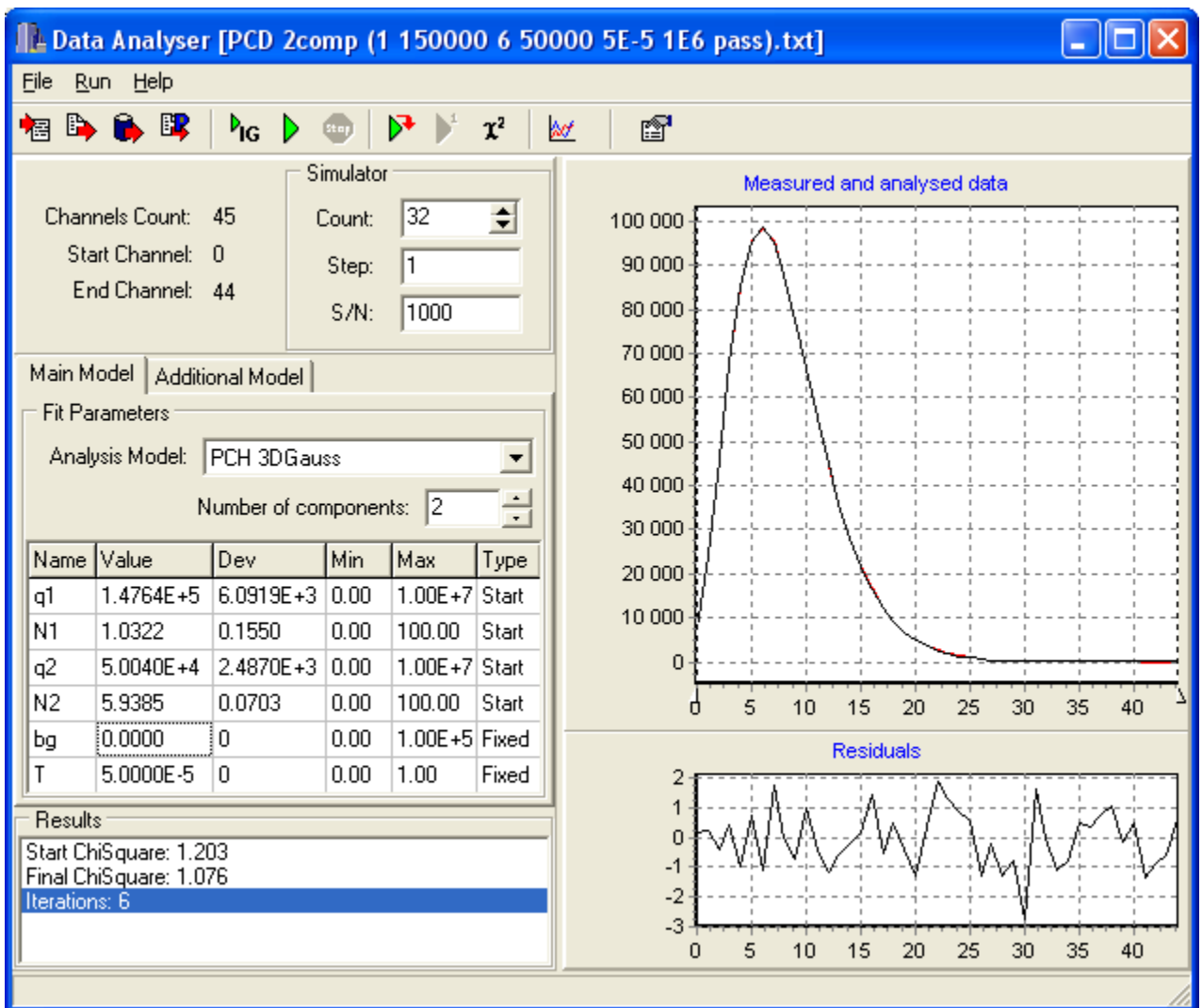


Рисунок 2.10 – Главное окно программы Data Analyser

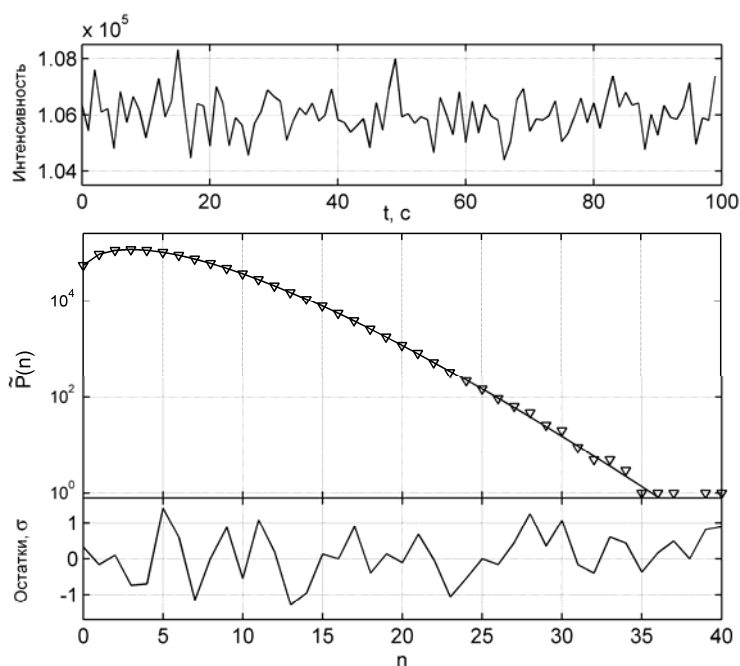
Процедура анализа основана на методе наименьших квадратов с градиентным методом оптимизации Маркварта с коррекцией Левенберга [118]. Качество подгонки проверяется путем вычисления взвешенного критерия  $\chi^2$  (приведенного к единице) и визуальным анализом взвешенных остатков. Поскольку число интервалов наблюдения обычно очень большое, Биноминальное распределение хорошо аппроксимируется нормальным распределением и поэтому применение критерия  $\chi^2$  является оправданным. Среднеквадратические отклонения оцениваемых параметров находятся по методу АСО [116].

Тестирование разработанного алгоритма заключалась в генерации распределений с различным набором параметров и отношения  $S/N$  и последующем их анализе по методу РСН (FIDA). Хорошее совпадение полученных оценок и истинных значений параметров генерируемого распределения, равномерность остатков и значение критерия  $\chi^2$  свидетельствовало об адекватности выбранного распределения шума и разработанного алгоритма генерации.

Имитационная модель фотоотсчетов была реализована в пакете Matlab 6.5. Адекватность модели вначале проверялась в граничных условиях: при  $N$  или  $q = 0$

распределение представляет  $\delta$ -функцию; при  $N \rightarrow \infty$  распределение стремится к Пуассоновскому, как и ожидалось [27]. Затем проводилось тестирование в предположении равномерности засветки и одной молекулы в закрытом объеме. Полученное РЧФ хорошо аппроксимировалось распределением Пуассона со средним, определяемым формулой (2.30). В заключение проводилось моделирование двухкомпонентных данных с гауссовским профилем засветки (1.2). Сгенерированное распределение было проанализировано по методу РСН с помощью программы Data Analyser. Результаты представлены на рисунке 2.10 (параметры моделирования следующие:  $N_{eff1} = 1$ ,  $q_1 = 1.5E5$ ,  $N_{eff2} = 5$ ,  $q_2 = 5E4$ ,  $T = 5E-5$  с, число интервалов  $m = 1E6$ , что эквивалентно отношению  $S/N = 330$ ). Равномерность остатков, значение критерия согласия и близость оценок параметров к их истинным значениям свидетельствуют об адекватности модели.

Модель потока фотонов с заданным РЧФ реализована в программе Simulator имитационного моделирования систем преобразования случайных потоков (смотри раздел 2.2). Это позволило провести моделирование системы регистрации ФФС с учетом фонового шума, мертвого времени и послеимпульсов. Для оценки точности воспроизведения моделируемых процессов число интервалов наблюдения взято равным  $1E6$  (параметры моделирования следующие:  $N_{eff} = 3$ ,  $q = 1E5$ ,  $T = 5E-5$  с). Число фотоотсчетов составило около  $5.3E6$ . Интенсивность потока была стационарна с  $\langle F \rangle = 106021$ , что полностью соответствует ожидаемому значению  $\langle F \rangle = \gamma_2 N_{eff} q = 106080$  (смотри (4.4)). Анализ по методу РСН дал следующие оценки параметров:  $N_{eff} = 2.9996 \pm 0.0022$ ,  $q = 99970 \pm 75$ , что свидетельствует о высокой степени адекватности модели. Результаты анализа и график интенсивности представлены на рисунке 2.11 ( $\tilde{P}(n) = mP(n)$ ).



**Рисунок 2.11 – Результаты анализа РЧФ и график интенсивности смоделированного потока фотонов ( $\chi^2 = 0.76$ )**

Для исследования влияния инструментальных искажений СРОД на регистрируемую характеристику были проведены следующие эксперименты. К смоделированному идеальному потоку фотонов вначале был добавлен шум (блок Noise) – конфигурация I, затем шум и непродлевающее мертвое время (блок Constant Generalized Dead Time) – конфигурация II, затем шум и послеимпульсы (блок Constant Afterpulses) – конфигурация III, затем шум, послеимпульсы и непродлевающее мертвое время – конфигурация IV. Во всех случаях сохранялась указанная последовательность блоков преобразования. Шумовой поток состоит из рассеянных фотонов и темновых импульсов детектора. Все события шумового потока способны генерировать послеимпульсы и мертвое время детектора. В последней конфигурации мертвое время замыкает цепочку блоков преобразования, так как послеимпульсы в свою очередь способны генерировать мертвое время. Типичная интенсивность фонового сигнала для СРОД Confocor 2 при проведении измерений водного раствора красителей при комнатной температуре равна 350 – 1500 отсчетов в секунду (данные представлены кафедрой паразитологии медицинского центра Лейденского университета, LUMC, Leiden, the Netherlands. Интенсивность фонового сигнала была выбрана равной 2000 отсчетов в секунду для повышения вероятности его выделения. Типичное мертвое время наиболее часто используемых детекторов SPCM-AQR (EG&G или Perkin–Elmer) равно  $5E-8$  с и имеет непродлевающий тип, а вероятность послеимпульсов составляет 0.3% [92, 97]. Для моделирования послеимпульсов использовался подход, описанный в разделе 2.2. Вторичный поток моделировался на интервале  $1E-5$  с (вероятность появления послеимпульса быстро падает с увеличением времени, прошедшем с момента регистрации исходного импульса, и уже через несколько микросекунд становится пренебрежимо малой [119]) с интенсивностью 300 отсчетов в секунду, что соответствует 0.3% вероятности появления послеимпульса после каждого фотона исходного потока.

Для всех конфигураций интенсивность выходного потока была стационарной со средним значением близким к ожидаемому. Число сгенерированных послеимпульсов было 16275, что составило  $\sim 0.3\%$  от общего количества сгенерированных фотонов. Потери вследствие мертвого времени составили 38370 (конфигурация II) и 38663 (конфигурация IV) фотонов. Результаты анализа смоделированных данных по методу РСН сведены в таблицу 2.2.

Таблица 2.2 – Результаты моделирования СРОД

Параметр	Ожидаемое значение	Оценки, полученные по методу РСН			
		Конфигурация I	Конфигурация II	Конфигурация III	Конфигурация IV
$N_{eff}$	3	3.014±0.007	3.047±0.007	3.011±0.012	3.040±0.01
$q$	100000	99710±138	97690±233	100180±246	98230±322
$\lambda$	2000	1770±104	2000 (фикс.)	1700±185	2000 (фикс.)
$\langle F \rangle$	108080	108033	107257	108350	107577

Как видно из представленных результатов, для выбранного режима измерения, характерного для ФФС, послеимпульсы и мертвое время не оказывают существенного влияния на оценки яркости и концентрации исследуемого вещества. Моделирование фотоотсчетов с разными параметрами показали, что искажения становятся заметными для случаев малой концентрации и/или малой яркости (послеимпульсы оказывают большее влияние) и большой концентрации и/или большой яркости (мертвое время оказывает большее влияние), как и ожидалось. Аддитивный шум учитывается моделью и его влияние может быть исключено путем фиксации параметра  $\lambda$  в известное значение. Полученные результаты согласуются с выводами, опубликованными в работах [92, 97].

## **2.4 Обработка и хранение данных во флуоресцентной спектроскопии**

Объем данных при проведении флуоресцентных измерений может быть очень большим. Результат измерения представляет собой поток фотоотсчетов, сохраненный в специально разработанном бинарном формате [30, 31, 32, 33]. В ФФС размер одного измерения может достигать 500Мб, да и каждое измерение может повторяться по нескольку раз. Для проведения анализа данных по методам FCS, PCH/FIDA, FCA требуется предварительная обработка данных, заключающаяся в вычислении анализируемых характеристик потока фотонов и формирования целостного набора данных для его последующего анализа. В ряде случаев в ФФС требуется дополнительная обработка, заключающаяся в выборе фрагмента потока с постоянной интенсивностью или в режекции пиков интенсивности, вызванных приходом в освещенную область очень яркой частицы (например, агрегата молекул). Для анализа данных с приемлемой точностью в большинстве случаев бывает недостаточно подгонки только одной характеристики. Во время-разрешимой флуоресцентной спектроскопии требуется хранить набор данных, состоящий из кривой затухания флуоресценции, функции инструментального отклика оборудования и измерения фонового шума [120, 121]. При анализе анизотропии флуоресценции число кривых удваивается [122, 123]. Дополнительно требуется указывать ряд параметров, например, температуру, длину волны испускания, длительность измерения, длительность интервала наблюдения и т.п. Таким образом, готовый к анализу набор данных должен содержать набор характеристик и набор параметров. Результаты анализа также представляют собой набор большого количества взаимосвязанных данных.

Традиционно исследователи сохраняли данные в файлы и заносили параметры в журнал измерений. Подготовка данных могла занимать значительное время. Такой подход не позволяет проводить анализ большого числа данных за небольшое время, автоматизировать и стандартизировать процесс анализа. Это требует применения современных технологий хранения и обработки данных и их



дальнейшей разработки. Технология баз данных (БД) [124, 125, 126, 127, 13-А], при условии разработки эффективной структуры данных, позволяет решить ряд поставленных задач. БД позволяют хранить данные в одном месте в целостном и компактном виде (с минимальной избыточностью), что дает возможность избежать рутинной работы с каталогами, файлами и необходимостью делать записи на бумажных носителях. Обеспечение логической и физической независимости программ от данных, безопасности и секретности хранения, целостности и правильности данных – это основные преимущества технологии БД [13-А]. Удобная навигация, возможность применения быстрого поиска, фильтрации и сортировки дает неоспоримое преимущество в скорости доступа к требуемым данным.

**Разработка структур данных ФФС.** Разработка БД проводилась исходя из методики, опубликованной в [13-А]. Следование указанной методике является обязательным требованием, поскольку разработка БД не является типично программистской задачей. Анализ предметной области, моделирование данных являются важнейшими этапами разработки любой БД. Качество БД напрямую определяется продуманностью структуры данных, являющейся информационной моделью предметной области [13-А].

Перед разработкой структуры данных был произведен сбор и анализ требований к БД, определены задачи и перечень данных, требуемых для решения поставленных задач. В предметной области были выделены следующие объекты: предмет измерения, экспериментатор, измерение, характеристика, параметры измерений и параметры характеристик. Для моделирования предметной области была выбрана реляционная модель данных [124, 125, 13-А]. По требованиям реляционной модели для уменьшения избыточности информации и ликвидации аномалий обновления [125, 13-А] была проведена нормализация данных, заключающаяся в применении нормальных форм до третьего порядка включительно. Нормализация проводилась путем декомпозиции отношений (двумерных множеств, в дальнейшем просто таблиц) с последующей проверкой на существование ловушек разветвления и разрыва. Декомпозиция должна обеспечить равенство результатов выборок из исходного отношения и выборок, основанных на совокупности полученных отношений, т.е. должно соблюдаться следующее соотношение

$$\bigcup_{i=1}^m R_i = R, \quad (2.38)$$

где  $R = \{A_1, A_2, \dots, A_n\}$  – исходное отношение,  $A_i$  – атрибут отношения,  $D = \{R_1, R_2, \dots, R_m\}$  – декомпозиция, т.е. множество нормализованных отношений. По требованиям первой нормальной формы, заключающейся в проверке соответствия исходной таблицы требованиям, предъявляемым к отношениям в

реляционной теории, были исключены составные и многозначные атрибуты и введены ключевые атрибуты [13-А]. Перевод отношения во вторую нормальную форму проводился путем выделения в дополнительные отношения всех атрибутов, не характеризующихся полной функциональной зависимостью от первичного ключа [13-А]. Перевод в третью нормальную форму проводился путем выделения в дополнительные отношения всех атрибутов, транзитивно зависящих от первичного ключа [13-А].

Конечная структура БД измерений в ФФС представлена на рисунке 2.12. Структура данных состоит из 7 таблиц, состоящих в связи «один ко многим» с другими таблицами. Главной является таблица DataSets, введенная специально для формирования готовых для анализа наборов характеристик и параметров. Каждая запись в этой таблице характеризует отдельный набор данных. Таблицы Samples и Guests хранят информацию об объекте измерения и исследователе. Таблица RawData предназначена для хранения измерений, сохраненных в виде потока фотонов, и находится в связи «один к одному» с таблицей DataSets. На основе одного измерения можно вычислить множество характеристик, что и отражает связь «один ко многим» с таблицей Observations. Таблицы Parameters и ExtParameters хранят параметры измерений. Таблица ExtParameters выделена специально для целей хранения так называемых внешних параметров типа температуры и т.д.

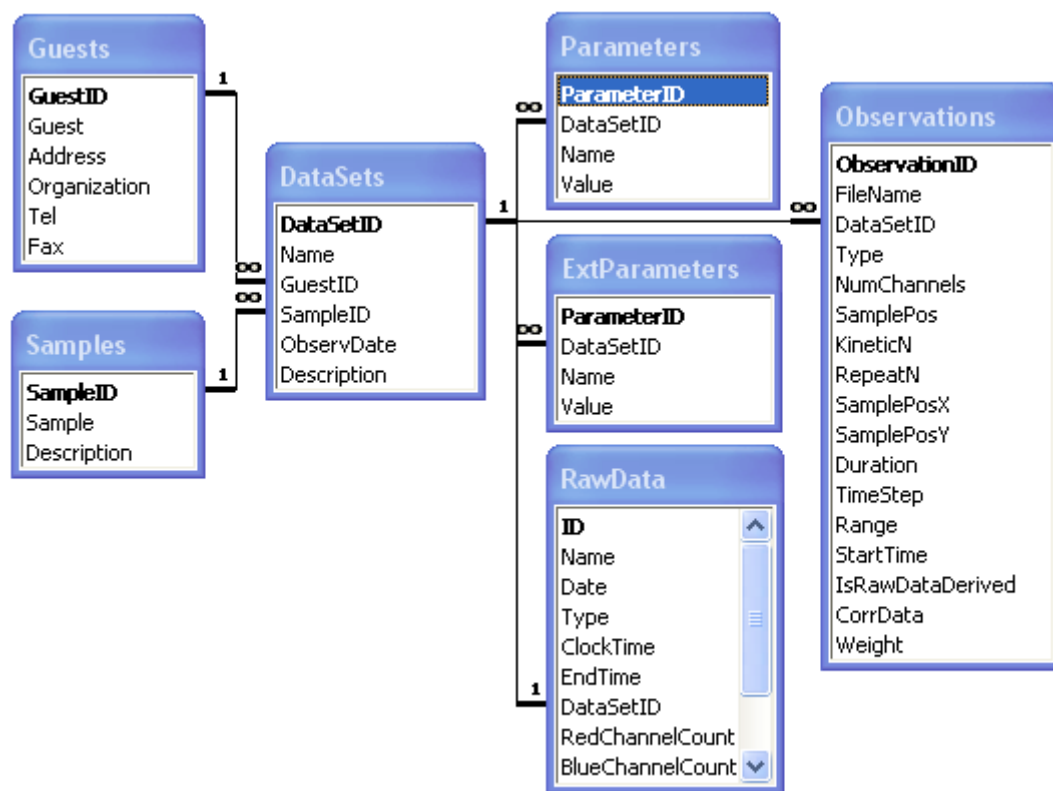


Рисунок 2.12 – Структура данных БД измерений в ФФС

Структура БД результатов анализа данных была разработана с учетом возможности повторного открытия сохраненных результатов анализа, необходимости их быстрого просмотра и построения зависимостей полученных оценок параметров от номера эксперимента и ряда внешних параметров, т.е. с обеспечением возможности статистической обработки результатов. Это, в свою очередь, потребовало построение информационной модели большинства компонентов приложения, выполняющего анализ данных. К базовым компонентам анализирующего приложения можно отнести модели анализа, набор анализируемых характеристик, метод оптимизации и критерии качества подгонки [9-А]. Для каждой компоненты были выделены одно или несколько отношений. Конечная структура данных после выполнения нормализации представлена на рисунке 2.13.

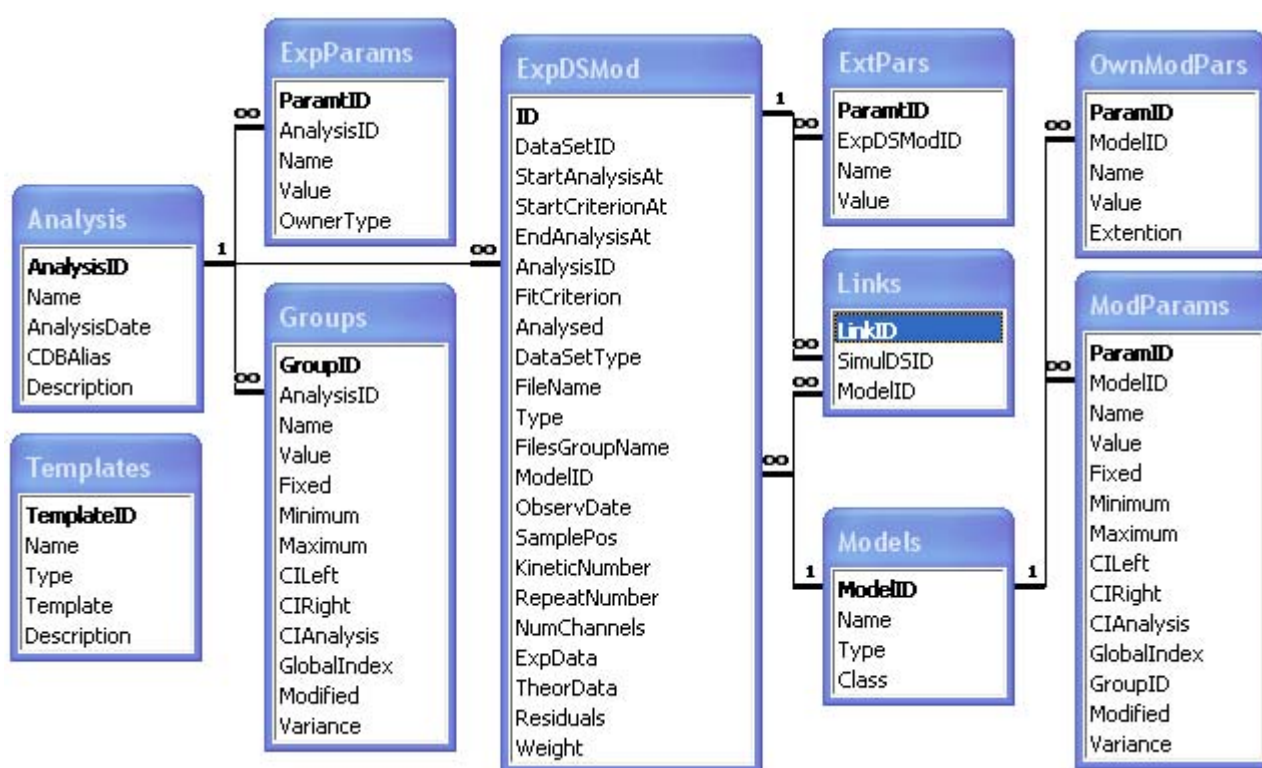


Рисунок 2.13 – Структура данных БД результатов анализа в ФФС

Главной таблицей является таблица Analysis, хранящая записи о проведенных анализах. Каждая запись характеризует отдельный анализ данных независимо от его внутренней сложности. Таблицы ExpParams, ExpDSMod и Groups находятся в связи «много к одному» с таблицей Analysis. Таблица ExpParams хранит параметры метода оптимизации и критериев качества. Таблица ExpDSMod хранит экспериментальные и теоретические кривые, их взвешенную разность (остатки), и связывает модель, анализируемую характеристику и саму запись о проведенном анализе в единое целое. Группа таблиц Models, ModParams, OwnModPars предназначена для хранения информации о моделях. Список оцененных параметров хранится в таблице ModParams, а свойства моделей

хранятся в таблице OwnModPars. Таблицы ModParams и OwnModPars находятся в связи «много к одному» с таблицей Models. Таблица Groups хранит список связанных параметров модели при проведении глобального (обобщенного) анализа [101, 128]. Связь параметра модели и группы производится с помощью внешнего ключа GroupID в таблице ModParams (связь на рисунке не показана). Таблица Links предназначена для хранения смоделированных характеристик и связывает характеристику и модели, лежащие в основе генерации характеристики.

Разработка БД для других методов флуоресцентной спектроскопии проводилась подобным образом. Структуры их данных представлены в приложении G. Разработанные БД вошли в состав трех программных комплексов: FCS Data Processor [9-A, 10-A, 11-A, 12-A], TRFA Data Processor [9-A, 14-A] и TRFD Fitting Software [9-A, 15-A, 16-A]. Time Resolved Fluorescence and Anisotropy (TRFA) Data Processor позволяет проводить анализ кривых затухания флуоресценции, полученных по методу регистрации первого фотона на интервале наблюдения (time correlated single photon counting – TCSPC) [117, 120, 129]. Time Resolved Frequency Domain (TRFD) Fitting Software позволяет проводить анализ данных, полученных по фазово-модуляционному методу [120].

Структура и назначение программных комплексов подробно рассмотрены в [9-A]. В состав каждого программного комплекса входят две БД: база данных результатов измерений и база данных результатов анализа [9-A]. БД результатов измерений выполняет задачи импорта и обработки измеренных данных, а также преобразования их в единый формат, поддерживаемый приложением для анализа. БД результатов анализа предназначена для компактного хранения результатов анализа флуоресцентных данных и позволяет просматривать результаты без необходимости их загрузки в анализирующую программу.

БД реализованы в виде набора таблиц, ядра базы данных и исполняемого модуля, обеспечивающего доступ к хранимой информации и выполнение функций системы управления базами данных (СУБД). В свою очередь исполняемый модуль также организован по модульному принципу. Такая конфигурация позволяет разделить задачи по отдельным независимым модулям и обеспечить требуемую гибкость и логическую независимость системы. В качестве ядра базы данных используется BDE (Borland Database Engine). Таблицы созданы в СУБД Borland PARADOX 7.0. Исполняемые модули системы управления БД состоят из форм, отчетов и компонент доступа к таблицам и запросам. Основное окно каждой БД представляет собой составную форму, отображающую главную таблицу и множество связанной информации, относящейся к выделенной записи в главной таблице (смотри рисунок 2.14). При переходе от записи к записи в главной таблице связанная информация меняется соответственно, что позволяет очень быстро находить требуемые данные. Более подробную информацию можно найти в формах, представляющих отдельные таблицы. Отчеты предназначены для просмотра и вывода на печать выделенных данных.

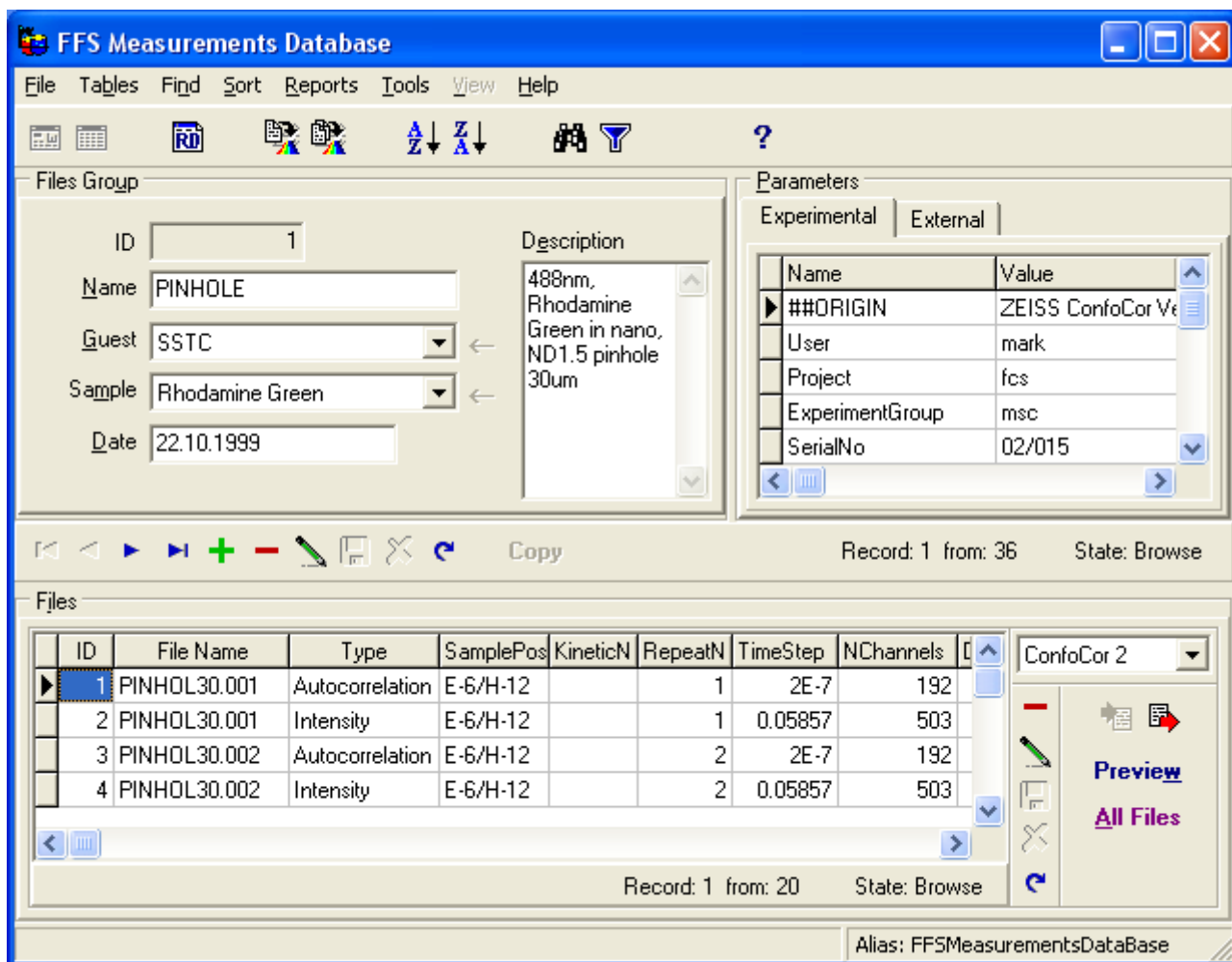


Рисунок 2.14 – Главное окно БД экспериментов в ФФС

**Обработка данных в ФФС.** Для проведения анализа данных ФФС по методам FCS, PCH/FIDA, FCA требуется расчет авто(кросс)корреляционной функции, РЧФ и ФК РЧФ. Применение формулы (2.18) для расчета АКФ в методе FCS не является целесообразным. Сравнительно большое время измерений (обычно 60 с) и малое время  $T$  (обычно  $2E-7$  с) приводят к необходимости резервирования массива размером  $(60/2E-7) \times 4 = 1.2E9$  байт для хранения интенсивности. Для минимизации требований к оперативной памяти компьютера предложен и реализован алгоритм расчета АКФ, позволяющий хранить в памяти только ненулевые каналы интенсивности. Для этого вместо одного массива для хранения интенсивности потока резервируется два массива: один для хранения номера канала интенсивности с ненулевым количеством фотонов, другой для хранения количества фотонов в этом канале. Соответственно произведена модификация формулы (2.18). Поскольку массив интенсивности при  $T = 2E-7$  с является сильно разреженным (большинство каналов нулевые), достигается значительный выигрыш в резервировании памяти. Фактически размер массива определяется числом зарегистрированных фотонов (принимая во внимание, что событие, заключающееся в попадании в канал двух и более фотонов, является

маловероятным). Для расширения диапазона вычисления с одновременной минимизацией количества точек, расчет АКФ производится с квазилогарифмической шкалой по оси  $X$  и симметричной нормализацией [82, 83, 84, 10-А]. Алгоритм тестировался путем сравнения результатов со стандартным алгоритмом вычисления АКФ при больших  $T$ . Весовые факторы АКФ вычисляются соответственно алгоритму, опубликованному в [84].

Для коррекции на конечную вероятность регистрации фотонов синего канала в красном канале  $p_{blue\_to\_red}$  (вследствие перекрытия спектров поглощения и испускания флуорофоров в двухцветовом эксперименте) при вычислении кросскорреляционной функции была проведена модификация разработанного алгоритма, заключающаяся в соответствующем уменьшении интенсивности красного канала  $I_{red} = I_{red} - I_{blue} * p_{blue\_to\_red}$  (отрицательные значения  $I_{red}$  приравнивались к нулю) и переходе к вычислениям в вещественной области.

РЧФ вычислялась как гистограмма интенсивности потока. Весовые факторы для РЧФ вычислялись в соответствии с Биномиальным распределением по формуле  $\sigma_i = \sqrt{mP^*(i)(1-P^*(i))}$ , где  $m$  – число интервалов наблюдения. ФК РЧФ рассчитывались по формулам, представленным в приложении В. Весовые факторы ФК рассчитывались по методу моментов от моментов [98, 99].

Режекция выбросов интенсивности, вызванных, например, приходом агрегата светящихся молекул в освещенную область, проводилась путем вырезания временных интервалов, соответствующих пикам интенсивности, с последующим вычитанием вырезанных интервалов времени из времен прихода всех последующих фотонов (общая длительность потока фотонов уменьшается на сумму всех вырезанных интервалов). Алгоритм режекции следующий: определяется положение пиков интенсивности, что может быть сделано автоматически путем установления порога интенсивности; из времени прихода первого фотона, попавшего в интервал, соответствующий пику интенсивности, вычитается время, прошедшее со времени прихода предыдущего зарегистрированного фотона; фотон игнорируется; из времени прихода второго фотона вычитается уже сумма двух времен задержки. Указанный алгоритм повторяется для всех фотонов, попавших в интервал, соответствующий пику интенсивности. Для всех последующих (полезных) фотонов производится только вычитание накопившейся суммы времен задержек без наращивания самой суммы. Для стационарных потоков такого рода режекция не приводит к искажению вычисленных характеристик, в то время как простое игнорирование фотонов в отдельных областях потока приводит к существенному искажению всех характеристик потока без исключения (эквивалентно влиянию непродлевающего мертвого времени). Алгоритм тестировался на имитационных и измеренных данных и показал свою эффективность.

## 2.5 Основные результаты и выводы

1. Разработаны аналитические модели последовательного преобразования случайных потоков, слияния независимых потоков и потоков, прошедших произвольное преобразование, а также разветвления с указанными вероятностями выхода. Основанные на аппарате производящих функционалов, разработанные модели позволяют исследовать широкий круг процессов, происходящих в системах регистрации во флуоресцентной спектроскопии.

2. Разработаны имитационные модели систем преобразования потоков случайных событий и на их основе создан программный комплекс Simulator имитационного моделирования СРОД. Проведенные вычислительные эксперименты показали высокую эффективность разработанных моделей при исследовании систем регистрации флуоресценции. Разработанные имитационные модели разветвленной системы преобразования потоков случайных событий, обладая высокой степенью общности, обеспечивают возможность моделирования широкого класса систем преобразования случайных потоков.

3. Разработаны имитационные модели потока фотонов со стационарной интенсивностью и заданной автокорреляционной функцией и потока фотонов с заданным распределением числа фотоотсчетов. Разработана программа Data Analyser, позволяющая генерировать распределение числа фотоотсчетов и факториальные кумулянты распределения числа фотоотсчетов с заданными параметрами и проводить анализ данных по методам FIDA, PCH, FCA. Проведенные вычислительные эксперименты с помощью разработанных аналитических и имитационных моделей показали, что искажения вследствие независимых потерь, мертвого времени и послеимпульсов не оказывают существенного влияния на анализируемые характеристики потока фотонов при типичных для ФФС условиях измерения.

4. Разработаны алгоритмы расчета авто(кросс)корреляционной функции, распределения числа фотоотсчетов и факториальных кумулянт распределения числа фотоотсчетов потоков фотоотсчетов, сохраненных в форматах Zeiss Confocor, ISS ALBA, B&H SPC, PicoQuant TimeHarp и PicoHarp с возможностью режекции выбросов интенсивности.

5. Разработаны структуры данных измерений и результатов анализа во флуоресцентной флуктуационной спектроскопии. Разработаны и реализованы БД измерений и результатов анализа. Разработанные БД и алгоритмы обработки данных вошли в состав трех программных комплексов: FCS Data Processor, TRFA Data Processor и TRFD Fitting Software.

## ГЛАВА 3

### АНАЛИЗ РАСПРЕДЕЛЕНИЯ ЧИСЛА ФОТООТСЧЕТОВ

#### 3.1 Производящая функция числа фотоотсчетов с коррекцией профиля засветки

Профиль засветки описывает пространственное распределение зарегистрированной флуоресценции (является сверткой функции распределения интенсивности падающего излучения и передаточной функцией измерительной системы) и в общем случае имеет сложный вид. На его форму оказывает влияние не только объектив конфокального микроскопа, но и ряд других частей измерительной установки. Для точного учета профиля засветки требуется знание параметров конфокальной диафрагмы, размещенной перед детектором, и апертуры самого детектора. Поэтому выбор функции  $B(\mathbf{r})$ , наилучшим образом аппроксимирующей реальный профиль засветки, является ключевым моментом методов анализа данных ФФС. В случае однофотонного возбуждения обычно применяют трехмерное гауссовское распределение в виде (1.2). Гауссовская аппроксимация хорошо описывает пространственное распределение флуоресценции молекул, находящихся в фокусе засветки, но она не учитывает влияние диафрагмы, рассеянное излучение и флуоресценцию молекул, находящихся на пути распространения лазерного излучения. Ее применение является оправданным в методе FCS [75, 130] (рассеянное излучение, фоновый сигнал детектора и флуоресценция молекул, находящихся в областях с однородной засветкой, не вносят заметного вклада в корреляционную функцию), но приводит к неудовлетворительным результатам при анализе распределения числа фотоотсчетов [27, 94].

Для более адекватной аппроксимации профиля засветки был предложен ряд подходов. Первый подход использует сложную модель профиля засветки, основанную на электромагнитной теории дифракции [94, 131]. Этот подход ведет к разработке сложного численного алгоритма, требующего значительных временных затрат на его выполнение. Ввиду своей сложности данный подход не нашел широкого распространения.

Второй подход основан на применении аппроксимаций, которые позволяют записать модель в явном виде. В методе FIDA реальный профиль засветки аппроксимируется экспоненциальной функцией одного аргумента с преобразованием элемента объема в виде полинома (1.83) с двумя или тремя варьируемыми формальными параметрами [28, 89, 91]. Недостатком данного подхода является необходимость введения дополнительных параметров, что приводит к увеличению числа варьируемых параметров и, следовательно, к потере точности аппроксимации многокомпонентных систем.



Третий подход основан на коррекции гауссовской аппроксимации профиля засветки. Для исправления возможных отличий реального профиля засветки от гауссовской аппроксимации было предложено включить дополнительные параметры подгонки, определенные как относительная разность между интегралами от реального профиля засветки и его гауссовской аппроксимации (1.89) [93, 94]. В большинстве случаев для получения наилучшей аппроксимации экспериментальных данных достаточно коррекции первого порядка. Данный подход представляется наиболее целесообразным. Его достоинством является применение аппроксимации, имеющей явный физический смысл, и необходимость в минимальном количестве дополнительных параметров коррекции.

ПФ числа фотоотсчетов (ПФЧФ) является эффективным инструментом для изучения свойств потока фотонов (смотри раздел 1.2). Зная ПФ, можно получить РЧФ либо путем многократного дифференцирования ПФ, либо через обратное преобразование Фурье от характеристической функции, получаемой из ПФ путем замены формальной переменной  $\xi$  на комплексную экспоненту  $e^{i\varphi}$ . Поток зарегистрированных фотонов можно также характеризовать набором моментов РЧФ. Удобно использовать ФК РЧФ, которые также можно получить из ПФЧФ путем многократного дифференцирования натурального логарифма ПФ.

Производящая функция  $G(\xi)$  распределения  $P(n)$  числа зарегистрированных фотонов, испущенных флуоресцентными молекулами, находящимися в равновесном состоянии в открытом объеме наблюдения  $V$  в течение интервала наблюдения  $T$ , имеет вид (1.82) [28]. Для коррекции профиля засветки разложим экспоненту под интегралом в выражении (1.82) в ряд Маклорена

$$G(\xi) = \exp \left\{ \lambda T (\xi - 1) + \sum_i \langle C \rangle_i \sum_{k=1}^{\infty} \frac{(\xi - 1)^k q_i^k T^k \chi_k}{k!} \right\}. \quad (3.1)$$

Принимая во внимание (1.89), получим

$$G(\xi) = \exp \left\{ \lambda T (\xi - 1) + \sum_i \langle C \rangle_i \sum_{k=1}^{\infty} \frac{(\xi - 1)^k q_i^k T^k (1 + F_k) \chi_{Gk}}{k!} \right\}. \quad (3.2)$$

Представим выражение (3.2) в виде произведения трех ПФ [17-A]

$$G(\xi) = G_{main}(\xi) G_{corr}(\xi) G_{bg}(\xi), \quad (3.3)$$

где

$$G_{main}(\xi) = \exp \left\{ \sum_i \langle C \rangle_i \sum_{k=1}^{\infty} \frac{(\xi - 1)^k q_i^k T^k \chi_{Gk}}{k!} \right\} = \exp \left\{ \sum_i \langle C \rangle_i \int_V \left( e^{(\xi-1)q_i T B_G(\mathbf{r})} - 1 \right) d\mathbf{r} \right\} \quad (3.4)$$

моделирует РЧФ из основной области засветки, аппроксимируемой функцией  $B_G(\mathbf{r})$ ,

$$G_{corr}(\xi) = \exp \left\{ \sum_i \langle C \rangle_i \sum_{k=1}^{\infty} \frac{(\xi-1)^k q_i^k T^k F_k \chi_{Gk}}{k!} \right\} \quad (3.5)$$

моделирует РЧФ из дополнительной (внефокусной) области засветки (коррекция профиля засветки) и

$$G_{bg}(\xi) = \exp \{ \lambda T (\xi - 1) \} \quad (3.6)$$

моделирует фоновый сигнал. Для коррекции первого порядка ( $F_1 \neq 0, F_k = 0, k = 2, 3, \dots$ ) выражение (3.5) примет вид

$$G_{corr1}(\xi) = \exp \left\{ \sum_i \langle C \rangle_i \left( (\xi - 1) q_i T F_1 \chi_{G1} \right) \right\}. \quad (3.7)$$

Выражение (3.7) представляет собой ПФ стационарного потока Пуассона с интенсивностью  $\sum_i \langle C \rangle_i q_i T F_1 \chi_{G1}$ .

Кроме выражения для ПФЧФ в виде (1.82) в последующем нам потребуется также выражение для ПФЧФ одиночной молекулы, свободно перемещающейся в закрытом объеме  $V_0$ . Следуя формализму, использованному при выводе (1.82), запишем вероятность получения  $n$  фотонов от молекулы, находящейся в некотором достаточно малом элементе  $dV_i$  объема  $V_0$  со значением  $B_i$  функции профиля засветки  $B(\mathbf{r})$

$$P_i^{(1)}(n, dV_i) = \frac{dV_i (qTB_i)^n}{V_0 n!} e^{-qTB_i}. \quad (3.8)$$

ПФ одиночной молекулы от элемента объема примет вид

$$G^{(1)}(\xi, dV_i) = \sum_{n=0}^{\infty} P_i^{(1)}(n, dV_i) \xi^n = \frac{dV_i}{V_0} e^{-qTB_i} \sum_{n=0}^{\infty} \frac{(qTB_i \xi)^n}{n!} = \frac{dV_i}{V_0} e^{(\xi-1)qTB_i}.$$

Переходя к полному объему  $V_0$ , получим

$$G^{(1)}(\xi) = \frac{1}{V_0} \int_{V_0} e^{(\xi-1)qTB(\mathbf{r})} d\mathbf{r}. \quad (3.9)$$

ПФЧФ одиночной молекулы с коррекцией профиля засветки может быть получена из (3.9) аналогично (3.2)

$$\begin{aligned} G^{(1)}(\xi) &= \frac{1}{V_0} \int_{V_0} \sum_{k=0}^{\infty} \frac{(\xi-1)^k (qT)^k B^k(\mathbf{r})}{k!} d\mathbf{r} = \frac{1}{V_0} \left[ 1 + \sum_{k=1}^{\infty} \frac{(\xi-1)^k (qT)^k \chi_k}{k!} \right] = \\ &= \frac{1}{V_0} \left[ 1 + \sum_{k=1}^{\infty} \frac{(\xi-1)^k (qT)^k (1 + F_k) \chi_{Gk}}{k!} \right] = \\ &= G_G^{(1)}(\xi) + \frac{1}{V_0} \sum_{k=1}^{\infty} \frac{(\xi-1)^k (qT)^k F_k \chi_{Gk}}{k!}. \end{aligned} \quad (3.10)$$

### 3.2 Сравнительный анализ методов анализа распределения числа фотоотсчетов

Для анализа амплитудного распределения интенсивности флуоресценции и разрешения смеси веществ с разными спектральными свойствами предложен ряд методов, РСН, FIDA и их модификации, каждый из которых чаще всего позиционируется как самостоятельный. Сравнительный анализ этих методов позволяет убедиться в том, что их различие проявляется только в используемой аппроксимации профиля засветки, алгоритмических подходах и деталях численной реализации.

Доказательство эквивалентности методов РСН, FIDA, FCA может быть произведено на основе вывода аналитических выражений, используемых в этих методах, из одной и той же ПФЧФ. Поскольку метод FIDA основан на ПФЧФ, остается показать, что методы РСН (с коррекцией профиля засветки и без нее) и FCA могут быть выведены из одной и той же ПФЧФ. Наиболее просто это сделать для метода FCA. Подставляя (1.82) в (1.19), получим

$$K_1 = \lambda T + \sum_j \langle C \rangle_j q_j T \int_V B(\mathbf{r}) \exp[(\xi - 1)q_j T B(\mathbf{r})] d\mathbf{r} \Big|_{\xi=1} = \lambda T + \chi_1 \sum_j \langle C \rangle_j q_j T \quad (3.11)$$

$$K_k = \chi_k \sum_j \langle C \rangle_j q_j^k T^k, \quad k = 2, 3, \dots,$$

что за исключением слагаемого, отвечающего за фоновый сигнал, полностью совпадает с (1.92).

Метод РСН основан на вычислении одномолекулярного распределения числа фотоотсчетов. Поэтому вначале покажем эквивалентность (1.85) и (3.9), затем покажем, что переход к нескольким молекулам в открытом объеме (т.е. применение (1.86)) приводит к (1.82). Дифференцируя (3.9) по  $\xi$  и полагая  $\xi = 0$ , получим

$$p^{(1)}(0) = \frac{1}{V_0} \int_{V_0} e^{-qTB(\mathbf{r})} d\mathbf{r},$$

$$p^{(1)}(1) = \frac{1}{1!V_0} \int_{V_0} qTB(\mathbf{r}) e^{(\xi-1)qTB(\mathbf{r})} d\mathbf{r} \Big|_{\xi=0} = \frac{1}{V_0} \int_{V_0} \frac{qTB(\mathbf{r})}{1!} e^{-qTB(\mathbf{r})} d\mathbf{r}, \quad (3.12)$$

...

$$p^{(1)}(n) = \frac{1}{V_0} \int_{V_0} \frac{[qTB(\mathbf{r})]^n}{n!} e^{-qTB(\mathbf{r})} d\mathbf{r} = \frac{1}{V_0} \int_{V_0} Poi(n, qTB(\mathbf{r})) d\mathbf{r},$$

что в точности совпадает с первым выражением в (1.88).

Найдем ПФ от (1.86)

$$G(\xi) = \sum_{M=0}^{\infty} (G^{(1)}(\xi))^M \frac{(\langle C \rangle V_0)^M}{M!} e^{-\langle C \rangle V_0} = \exp\{\langle C \rangle V_0 (G^{(1)}(\xi) - 1)\}. \quad (3.13)$$

Подставляя (3.9) в (3.13), получим

$$\begin{aligned}
 G(\xi) &= \exp \left\{ \langle C \rangle \int_{V_0} e^{(\xi-1)qTB(\mathbf{r})} d\mathbf{r} - \langle C \rangle V_0 \right\} = \\
 &= \exp \left\{ \langle C \rangle \left( \int_{V_0} e^{(\xi-1)qTB(\mathbf{r})} d\mathbf{r} - \int_{V_0} d\mathbf{r} \right) \right\} = \exp \left\{ \langle C \rangle \int_{V_0} (e^{(\xi-1)qTB(\mathbf{r})} - 1) d\mathbf{r} \right\}.
 \end{aligned} \tag{3.14}$$

Обобщив полученный результат на  $i$  компонент и введя слагаемое, отвечающее за фоновый шум, получим (1.82). Аналогично, для метода РСН с коррекцией, получим

$$\begin{aligned}
 p^{(1)}(0) &= p_G^{(1)}(0) + \frac{1}{V_0} \sum_{k=1}^{\infty} \frac{(-1)^k (qT)^k F_k \chi_{Gk}}{k!}, \\
 p^{(1)}(1) &= p_G^{(1)}(1) + \frac{1}{1!V_0} \sum_{k=1}^{\infty} \frac{(-1)^{k-1} (qT)^k F_k \chi_{Gk}}{(k-1)!}, \\
 &\dots \\
 p^{(1)}(n) &= p_G^{(1)}(n) + \frac{1}{n!V_0} \sum_{k=n}^{\infty} \frac{(-1)^{k-n} (qT)^k F_k \chi_{Gk}}{(k-n)!},
 \end{aligned} \tag{3.15}$$

что после нормализации к  $V_{eff}$  совпадает с (1.91).

Из полученных результатов очевидно, что методы РСН и FIDA являются полностью эквивалентными в результатах при условии использования одинаковой аппроксимации профиля засветки. Отличия возможны лишь вследствие численных ошибок реализующих их алгоритмов.

Сравнительный численный анализ методов РСН, FIDA, FCA представлен в разделе 4.4.

### 3.3 Генерация начальных приближений для метода FIDA

Для анализа данных ФФС наиболее часто применяется МНК. Данный метод предполагает итерационное изменение параметров теоретической модели посредством некоторого алгоритма оптимизации, обеспечивающего наилучшую аппроксимацию экспериментальных данных. Любой итерационный метод требует задания начальных приближений (НП) для параметров модели. Если НП близки к наилучшим оценкам параметров, МНК требуется всего несколько итераций для достижения глобального минимума. В противном случае, особенно для сложной поверхности целевого критерия, имеющего множество локальных минимумов, сходимость к глобальному минимуму не гарантируется или это требует больших временных затрат. Таким образом, НП в значительной степени увеличивают вероятность достижения глобального минимума, повышая тем самым производительность и устойчивость метода анализа. Автоматическая генерация НП стандартизирует процедуру анализа, минимизируя усилия исследователя при

поиске начальных приближений. Это играет важную роль при последовательном анализе большого числа данных.

Необходимость автоматической генерации НП особенно важна для анализа РЧФ. Специфика используемых в методах FIDA и РСН моделей состоит в том, что допустимый диапазон изменения параметров модели, как правило, очень мал.

В работе предлагается эффективный метод генерации НП для FIDA. Метод основан на статистическом методе моментов и может использоваться как самостоятельный быстрый неитерационный метод анализа распределения числа фотоотсчетов. С небольшой коррекцией данный метод может быть также применен для генерации НП для методов РСН и FCA.

### 3.3.1 Метод генерации начальных приближений

Метод моментов [35] широко используется в статистике для оценки параметров распределения вероятностей, для которого можно подсчитать необходимое количество моментов. Идея метода состоит в приравнивании теоретических моментов  $M_k = M_k(\eta_1, \eta_2, \dots, \eta_m)$  исследуемого распределения к экспериментально полученным моментам  $\tilde{M}_k$  и последующем решении системы уравнений

$$M_k(\eta_1, \eta_2, \dots, \eta_m) = \tilde{M}_k, \quad k = 1, 2, \dots, m \quad (3.16)$$

относительно оцениваемых параметров  $\eta_1, \eta_2, \dots, \eta_m$ .

По такой же схеме могут использоваться факториальные моменты и кумулянты. Их применение имеет значительные преимущества, поскольку теоретические выражения для них легко получить путем последовательного взятия производных от логарифма ПФЧФ (смотри раздел 1.2).

Экспериментальные факториальные кумулянты любого порядка можно вычислить по рекуррентной формуле через факториальные моменты измеренного РЧФ  $P^*(n)$  (расчет кумулянтов через обычные, центральные и факториальные моменты приведен в приложении В)

$$\tilde{F}_k = \langle n(n-1)\dots(n-k+1) \rangle = \sum_{n=k}^m n(n-1)\dots(n-k+1)P^*(n), \quad (3.17)$$

$$\tilde{K}_k = \tilde{F}_k - \sum_{i=1}^{k-1} \binom{k-1}{i} \tilde{K}_{k-i} \tilde{F}_i, \quad \text{где } \binom{k-1}{i} = C_{k-1}^i = \frac{(k-1)!}{i!(k-i-1)!}. \quad (3.18)$$

Подставляя (1.82) в (1.19), запишем выражения для факториальных кумулянтов

$$K_1 = \lambda T + \sum_j \langle C \rangle_j q_j T \int_V B(\mathbf{r}) \exp[(\xi-1)q_j T B(\mathbf{r})] d\mathbf{r} \Big|_{\xi=1} = \lambda T + \chi_1 \sum_j \langle C \rangle_j q_j T \quad (3.19)$$

$$K_k = \chi_k \sum_j \langle C \rangle_j q_j^k T^k, \quad k = 2, 3, \dots,$$

После подстановки уравнений нормировки (1.84) ( $\chi_1 = \chi_2 = 1$ ) получим систему нелинейных уравнений, пригодную для оценки параметров произвольного количества компонент исследуемой молекулярной системы

$$\begin{aligned} K_1 &= (\lambda + \sum_j N_{FIDA_j} q_{FIDA_j}) T \\ K_2 &= \sum_j N_{FIDA_j} q_{FIDA_j}^2 T^2 \\ K_k &= \chi_k \sum_j N_{FIDA_j} q_{FIDA_j}^k T^k, \quad k = 3, 4, \dots, \end{aligned} \quad (3.20)$$

где  $\chi_k$  являются функциями параметров профиля засветки  $a$  и  $b$ . В системе (3.20) сделана замена переменных (произведен переход от концентрации к числу молекул в объеме, определяемом нормализацией)

$$N_{FIDA} = \langle C \rangle \chi_1^2 / \chi_2, \quad q_{FIDA} = q \chi_2 / \chi_1. \quad (3.21)$$

В дальнейшем в данном разделе для упрощения записи индекс FIDA в обозначении  $N$  и  $q$  будет опускаться.

После подстановки (1.83) в (1.70) получим

$$\chi_k = \int_0^\infty (B_0 e^{-x})^k A_0 (x + ax^2 + bx^3) dx = A_0 B_0^k (2ak + 6b + k^2) / k^4. \quad (3.22)$$

Начальные параметры  $A_0$ ,  $B_0$  получим из условий нормализации (1.19)

$$B_0 = 8u/v, \quad A_0 = v/8u^2, \quad u = 2a + 6b + 1, \quad v = 2a + 3b + 2. \quad (3.23)$$

Решение системы (3.20), где теоретические факториальные кумулянты в левой части системы заменены на соответствующие экспериментальные, теоретически позволяет найти оценки для всех параметров распределения в случае двух и более компонент. Однако, с практической точки зрения, нестабильность факториальных кумулянтов высокого порядка не позволяет находить оценки любого количества параметров с удовлетворительной точностью и устойчивостью решения. Более того, численное решение системы нелинейных уравнений (например, методом Ньютона-Рафсона) является итерационным и, следовательно, также требует начальных приближений. Поэтому мы рассмотрим более простые случаи, где решение системы может быть получено неитерационным способом.

**Начальные приближения для однокомпонентной системы.** Запишем по одному уравнению для каждого оцениваемого параметра  $N$ ,  $q$ ,  $\lambda$ ,  $a$  и  $b$

$$\begin{aligned} K_1 &= (\lambda + Nq) T \\ K_2 &= Nq^2 T^2 \\ K_3 &= Nq^3 T^3 (2a + 2b + 3) 64u / (27v^2) \\ K_4 &= Nq^4 T^4 (4a + 3b + 8) 4u^2 / v^3 \\ K_5 &= Nq^5 T^5 (10a + 6b + 25) 4096u^3 / (625v^4). \end{aligned} \quad (3.24)$$

Система нелинейных уравнений (3.24) может быть решена аналитически. Из первых двух уравнений получим выражения для оценки  $q$  и  $N$

$$q = \frac{K_2}{(K_1 - \lambda T)T} \text{ и } N = \frac{(K_1 - \lambda T)^2}{K_2}. \quad (3.25)$$

После подстановки (3.25) в последние три уравнения системы (3.24) получим систему уравнений для оценки  $a$  и  $b$

$$\begin{aligned} \frac{(10a + 6b + 25)(2a + 2b + 3)}{(4a + 3b + 8)^2} &= \frac{16875 K_5 K_3}{16384 K_4^2} \\ \frac{(2a + 3b + 2)(4a + 3b + 8)}{(2a + 2b + 3)^2} &= \frac{1024 K_4 K_2}{729 K_3^2} \end{aligned} \quad (3.26)$$

и выражение для оценки  $\lambda$

$$\lambda = K_1 - \frac{64 K_1^2 (2a + 2b + 3)(2a + 6b + 1)}{27 K_2 (2a + 3b + 2)^2}. \quad (3.27)$$

Система (3.26) может быть сведена к полиному четвертого порядка по  $a$  или  $b$  (смотри приложение С) и решена как численно, так и аналитически. Если определитель полинома положительный, получим два действительных и два комплексно-сопряженных корня [35]. Один действительный корень (всегда представимый решением  $a = -7/2$  и  $b = 2$ ) приводит к сингулярности модели ( $N = 0$ ,  $q = \infty$ , смотри (3.25)) и должен быть опущен. Таким образом, только одна комбинация параметров приемлема как НП. Если определитель отрицательный, получим четыре действительных корня. Один корень (тоже представимый решением  $a = -7/2$  и  $b = 2$ ) должен быть опущен. Три остальные корни физически приемлемы. Множество приемлемых алгебраических решений есть следствие комбинации полиномиальной аппроксимации профиля засветки с выбранной системой уравнений нормализации. Это требует разработки процедуры выбора уникальной комбинации параметров, наиболее подходящей как НП. Для выбора наиболее подходящего решения мы будем опускать все наборы параметров с хотя бы одним физически неприемлемым или явно нереалистичным значением, а для совокупности физически приемлемых решений будем принимать такую комбинацию параметров, которая дает минимальное значение критерия  $\chi^2$ , рассчитанного для анализируемого РЧФ.

Если  $\lambda$  известно (может быть оценено из дополнительного измерения по формуле  $\lambda = K_1/T = \langle n \rangle / T$ , где в качестве образца взят чистый растворитель), система (3.24) упрощается до 4 уравнений. Вследствие уравнений нормализации в форме (1.84) параметры  $N$  и  $q$  могут быть оценены независимо от параметров профиля засветки  $a$  и  $b$  из первых двух уравнений системы (3.24). Оценки для  $a$  и  $b$  можно получить из последних двух уравнений системы (3.24)

$$\frac{64(2a+6b+1)(2a+2b+3)}{27(2a+3b+2)^2} = \frac{(K_1 - \lambda T)K_3}{K_2^2} = \chi_3$$

$$\frac{4(2a+6b+1)^2(4a+3b+8)}{(2a+3b+2)^3} = \frac{(K_1 - \lambda T)^2 K_4}{K_2^3} = \chi_4. \quad (3.28)$$

Проведя обобщение, все  $\chi_k$  могут быть вычислены из факториальных кумулянтов по следующей формуле

$$\chi_k = \frac{(K_1 - \lambda T)^{k-2} K_k}{K_2^{k-1}}. \quad (3.29)$$

Решение системы (3.28) осуществляется аналогично решению системы (3.26). Система (3.28) также имеет несколько решений, но в этом случае выбор набора параметров необходимо производить исходя не из минимизации критерия  $\chi^2$ , а установкой приемлемого диапазона изменения параметров. Для дальнейшего упрощения системы положим  $\lambda = 0$ . В этом случае получим хорошо известные формулы [132, 133] для оценки  $N$  и  $q$

$$q = \frac{K_2}{K_1 T} = \frac{\langle \Delta n^2 \rangle - \langle n \rangle}{\langle n \rangle T}, \quad N = \frac{K_1^2}{K_2} = \frac{\langle n \rangle^2}{\langle \Delta n^2 \rangle - \langle n \rangle}. \quad (3.30)$$

Игнорирование  $\lambda$  приводит к переоценке концентрации и, соответственно, недооценке яркости, что очевидно из уравнений (3.25) и (3.30).

**Начальные приближения для двухкомпонентной системы.** Для двухкомпонентной системы необходимо определить семь уравнений

$$\begin{aligned} K_1 &= (\lambda + N_1 q_1 + N_2 q_2) T \\ K_2 &= (N_1 q_1^2 + N_2 q_2^2) T^2 \\ K_3 &= (N_1 q_1^3 + N_2 q_2^3) T^3 (2a + 2b + 3) 64u / (27v^2) \\ K_4 &= (N_1 q_1^4 + N_2 q_2^4) T^4 (4a + 3b + 8) 4u^2 / v^3 \\ K_5 &= (N_1 q_1^5 + N_2 q_2^5) T^5 (10a + 6b + 25) \frac{4096u^3}{625v^4} \\ K_6 &= (N_1 q_1^6 + N_2 q_2^6) T^6 (2a + b + 6) \frac{4096u^4}{27v^5} \\ K_7 &= (N_1 q_1^7 + N_2 q_2^7) T^7 (14a + 6b + 49) \frac{8^6 u^5}{2401v^6}. \end{aligned} \quad (3.31)$$

Система (3.31) может быть решена численно. Однако, как показывает практика, кумулянты выше пятого порядка в большинстве случаев имеют среднеквадратические отклонения больше значений самих кумулянтов. Поэтому, как и для однокомпонентного случая, рассмотрим более простые варианты.

Известный фоновый шум. Если интенсивность фонового сигнала известна, первое уравнение системы (3.31) примет вид  $K_1 - \lambda T = (N_1 q_1 + N_2 q_2) T$ , а последнее



может быть опущено. Такая система уравнений может быть решена численно и использоваться для нахождения НП при достаточно большом отношении  $S/N$ .

Известный фоновый шум и параметры профиля засветки  $a$  и  $b$ . В этом случае система (3.31) примет вид

$$\begin{aligned} K_1 - \lambda T &= (N_1 q_1 + N_2 q_2) T \\ K_2 &= (N_1 q_1^2 + N_2 q_2^2) T^2 \\ K_3 / \chi_3 &= (N_1 q_1^3 + N_2 q_2^3) T^3 \\ K_4 / \chi_4 &= (N_1 q_1^4 + N_2 q_2^4) T^4, \end{aligned} \quad (3.32)$$

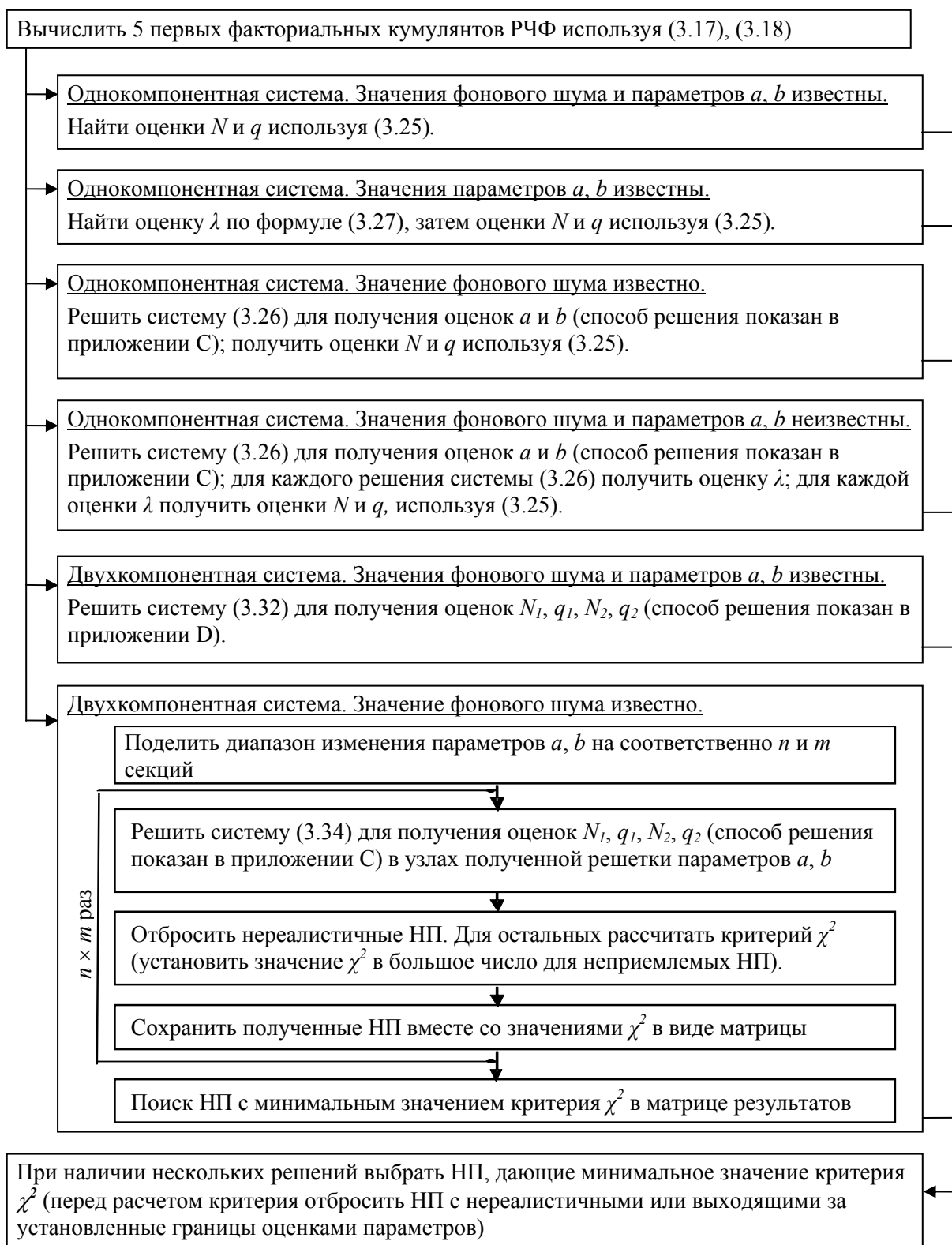
где  $\chi_3$  и  $\chi_4$  могут быть вычислены через известные  $a$  и  $b$ , используя (3.28). Система (3.32) может быть решена аналитически (смотри приложение D). Для данного приближения необходимо знать точные оценки параметров  $a$  и  $b$ . Значения  $a$  и  $b$  определяются параметрами СРОД и не должны меняться существенно от измерения к измерению при сохранении настроек системы. Тогда для случая последовательной обработки некоторого числа измерений достаточно провести точный анализ первого измерения и затем использовать полученные оценки для  $a$  и  $b$  для всех остальных измерений.

Если оценки для  $a$  и  $b$  неизвестны, но известен вероятный диапазон их изменения (определяется несколькими калибровочными измерениями с последующим анализом данных), будем использовать следующий алгоритм. Диапазон изменения параметров делится на несколько частей и производится генерация НП посредством решения системы (3.32) в узлах полученной решетки. Затем для каждого набора параметров вычисляется критерий  $\chi^2$ . Начальными приближениями становится набор параметров с наименьшим значением критерия  $\chi^2$ .

Блок-схема алгоритма генерации НП [18-А, 19-А] представлена на рисунке 3.1. Метод FIDA и алгоритм генерации НП для него реализованы в программе FCS Data Processor.

### 3.3.2 Тестирование начальных приближений

Тестирование НП проводилось по следующему алгоритму: проверка работоспособности на незашумленных данных, затем на смоделированных с известным отношением  $S/N$ , и в заключение на измеренных данных. Для тестирования на смоделированных данных генерировались серия РЧФ с одинаковыми параметрами и различным отношением  $S/N$  (50 в каждой серии). Если вычисление НП было невозможным вследствие численных ошибок, или если сгенерированное приближение не попадало в заранее определенный (достаточно большой) диапазон, генерация повторялась с целью достижения 50 РЧФ в каждой серии. Величина интервала наблюдения  $T$  принималась равной  $5 \times 10^{-5}$  с, если не указывалось иное.



**Рисунок 3.1 – Блок-схема генерации начальных приближений**

Моделирование проводилось по прямому методу генерации РЧФ, рассмотренному в подразделе 2.3.2. Использование прямого метода гарантирует получение несмещенных оценок параметров модели при отсутствии шума или при очень большом отношении  $S/N$ . Начальное значение для отношения  $S/N$  ( $S/N_i$ ) задавалось как корень квадратный из максимума сгенерированной кривой (это

упрощение, так как реальное отношение  $S/N$  зависит от значения РЧФ в максимуме и всегда выше, смотри подраздел 2.3.2). Моделирование проводилось с помощью программы Data Analyser. Теоретическое РЧФ вычислялось по формулам (1.82), (1.83), (1.84) и (1.21). Экспериментальные характеристики (РЧФ, ФК РЧФ) с соответствующими весовыми факторами [27, 28, 99] вычислялись из потока зарегистрированных фотонов, сохраненного в формате ConfoCor 2 (Carl Zeiss, Germany) с помощью программы FCS Data Processor. Алгоритмы вычисления анализируемых характеристик приведены в подразделе 2.4. Анализ по методу FCS также проводился с помощью программы FCS Data Processor [10-A].

**Тестирование начальных приближений для однокомпонентной системы на смоделированных данных.** Оценка алгоритма нахождения НП (уравнения (3.25) – (3.27)) для незашумленных данных показала, что полученные приближения незначительно отличаются от заданных при моделировании. Причину ошибки можно объяснить различием численных алгоритмов, лежащих в основе моделирования РЧФ и вычисления НП. Численное интегрирование, дискретное преобразование Фурье на конечном числе точек и ошибки вычисления вносят свои искажения в смоделированную кривую. Эти систематические ошибки в свою очередь искажают вычисленные ФК РЧФ. Величина ошибки растет с увеличением порядка кумулянта. Поэтому точность НП в большой степени зависит от порядка используемых кумулянтов. В нашем эксперименте относительная ошибка всех параметров кроме  $\lambda$  была порядка 0.01%. Фиксирование  $\lambda$  (в истинное значение) уменьшает количество используемых кумулянтов до 4, причем для нахождения параметров  $N$  и  $q$  необходимы только первых два кумулянта (смотри (3.25)). Это увеличивает точность для параметров  $N$  и  $q$  на четыре порядка, а для параметров  $a$  и  $b$  на два порядка. Относительная ошибка параметра  $\lambda$  (1.4%) является наивысшей среди остальных параметров. Генерация РЧФ с различным значением  $\lambda$  показала, что величина ошибки зависит от отношения  $\lambda$  к произведению  $N$  и  $q$ . Когда  $\lambda$  близко к  $q \cdot N$ , ошибка  $\lambda$  сравнима с ошибкой  $q$  и увеличивается с увеличением  $\lambda$ . Это свойство системы (3.25), где  $\lambda$ ,  $N$  и  $q$  взаимосвязаны. Очевидно, что едва ли возможно оценить вклад фонового шума в зарегистрированный сигнал, если только небольшая часть фотонов происходит вследствие шума.

Результаты тестирования алгоритма НП с использованием уравнений (3.25) – (3.27) для зашумленных данных приведены в таблице 3.1. В таблице приведены точные значения параметров, заданные для модели, а также средние и среднеквадратические отклонения начальных приближений для этих параметров, рассчитанные для серии из 50 РЧФ. Устойчивость алгоритма была достаточно высока для больших отношений  $S/N$ . Меньше чем в 2% случаях были получены физически неприемлемые оценки параметров (отрицательные значения  $\lambda$  или явные выбросы значений  $a$  и  $b$ ). При меньших  $S/N$  (меньше 1000 для  $S/N_i$ ) около

50% сгенерированных НП были неприемлемы. Смещение оценки параметра  $\lambda$  (смотри таблицу 3.1) от его истинного значения является не характерной чертой разработанного метода, а следствием режекции НП, проявляющихся в отрицательных  $\lambda$ .

Таблица 3.1 – НП для однокомпонентных смоделированных данных, вычисленные по формулам (3.25), (3.26), (3.27)

Параметр	Точное значение	НП				
		$S/N_i = 7000$	$S/N_i = 5000$	$S/N_i = 3000$	$S/N_i = 1000$	$S/N_i = 500$
$N$	5	4.994±0.126	4.921±0.145	4.882±0.205	4.327±0.542	4.037±0.728
$q$	20000	20016±252	20165±298	20253±436	21633±1462	22546±2158
$\lambda$	2000	2064±1260	2804±1459	3208±2086	9170±5948	12492±8267
$a$	-1	-0.999±0.019	-0.987±0.023	-0.980±0.034	-0.853±0.152	-0.732±0.274
$b$	0.5	0.500±0.002	0.499±0.002	0.499±0.003	0.495±0.009	0.503±0.020

Значительно более лучшие результаты были получены при генерации НП с фиксированным  $\lambda$  (для тех же смоделированных данных при использовании уравнений (3.25), (3.28)). Результаты представлены в таблице 3.2 ( $\lambda$  фиксировалось в значение, использованное при генерации РЧФ). Даже для отношения  $S/N_i = 20$  все полученные оценки параметров были близки к значениям, взятым при моделировании.

Таблица 3.2 – НП для однокомпонентных смоделированных данных, вычисленные по формулам (3.25), (3.28)

Параметр	Точное значение	НП ( $\lambda = 1000$ )				
		$S/N_i = 1000$	$S/N_i = 500$	$S/N_i = 300$	$S/N_i = 100$	$S/N_i = 50$
$N$	5	5.000±0.006	4.998±0.016	4.999±0.025	5.005±0.079	4.999±0.107
$q$	20000	20002±24	20008±69	20003±107	19995±336	20029±455
$a$	-1	-1.000±0.004	-1.000±0.008	-1.006±0.015	-1.027±0.051	-1.048±0.106
$b$	0.5	0.500±0.008	0.502±0.020	0.514±0.033	0.496±0.108	0.522±0.183

Для проверки влияния фонового шума на качество НП и сравнение алгоритмов с оцениваемым  $\lambda$  (первый алгоритм) или фиксированным  $\lambda$  (второй алгоритм) было сгенерировано несколько серий РЧФ с разным уровнем фонового сигнала. Отношение  $S/N_i$  было принято равным 1000. В первом алгоритме (уравнения (3.25) – (3.27)) оценивались все параметры, включая  $\lambda$ . Во втором алгоритме (уравнения (3.25), (3.28))  $\lambda$  было зафиксировано в 1000 безотносительно к значению, используемому при моделировании. Результаты представлены в таблице 3.3. Поставленный эксперимент показал, что во втором алгоритме среднеквадратические отклонения оцениваемых параметров не зависят от значения  $\lambda$ . Вычитание  $\lambda$  из первого факториального кумулянта не влияет на точность определения параметров, а вносит систематическое смещение, зависящее

от разности используемого при генерации НП  $\lambda$  и реального значения интенсивности фонового сигнала. В первом алгоритме, при одних и тех же значениях параметров  $a$  и  $b$ , точность определения  $\lambda$  возрастает, а точность определения  $N$  и  $q$  уменьшается с уменьшением разности  $\lambda$  и  $q \cdot N$ . Наблюдаемые результаты согласуются с результатами, полученными при анализе незашумленных данных. Такое поведение согласуется с ожидаемым из вида (1.82). Фонový сигнал имеет пуассоновскую статистику и вносит вклад только в первый факториальный кумулянт. И, наконец, точность определения параметров  $a$  и  $b$  не зависит от разности  $\lambda$  и произведения  $q \cdot N$ , что также можно ожидать из вида (1.82).

Таблица 3.3 – НП для однокомпонентных смоделированных данных с различным уровнем фонового шума, вычисленные по формулам (3.25) – (3.27) (алгоритм 1) и (3.25), (3.28) (алгоритм 2)

Параметр	Точное значение	$\lambda$	НП				
			$\lambda = 20000$	$\lambda = 10000$	$\lambda = 5000$	$\lambda = 500$	$\lambda = 200$
$\lambda$		Оценка	19593±9262	14066±6736	9536±6106	8147±5689	7762±5209
$N$	5	Оценка	5.083±0.920	4.623±0.655	4.574±0.571	4.279±0.514	4.287±0.470
		= 1000	7.082±0.010	5.942±0.008	5.409±0.007	4.950±0.007	4.921±0.007
$q$	20000	Оценка	20090±1930	20953±1455	21048±1439	21746±1437	21700±1308
		= 1000	16804±25	18344±26	19226±27	20100±29	20161±30
$a$	-1	Оценка	-0.976±0.167	-0.915±0.124	-0.906±0.145	-0.839±0.186	-0.848±0.143
		= 1000	-1.212±0.005	-1.115±0.005	-1.056±0.004	-0.991±0.004	-0.987±0.004
$b$	0.5	Оценка	0.503±0.010	0.498±0.008	0.497±0.010	0.497±0.013	0.496±0.007
		= 1000	0.527±0.007	0.513±0.009	0.505±0.010	0.496±0.008	0.495±0.008

На практике лучше использовать второй алгоритм (с фиксацией  $\lambda$ ), так как он намного более устойчив и предоставляет более точные оценки параметров без существенных систематических смещений при условии фиксации  $\lambda$  в значение, типичное для проводимых измерений [18-А]. Если  $\lambda$  сравнимо с  $q \cdot N$ , первый алгоритм предоставляет приемлемую точность определения  $\lambda$  и также может быть использован для генерации НП.

**Тестирование начальных приближений для двухкомпонентной системы на смоделированных данных.** Для оценки алгоритма нахождения НП (решение системы (3.31)) для двухкомпонентной системы было проведена генерация нескольких серий РЧФ с различными значениями параметров и отношением  $S/N$ . Удовлетворительные результаты численного решения системы (3.31) (по методу Ньютона-Рафсона) были получены лишь при некоторых комбинациях параметров и очень большом отношении  $S/N$ . Лучшие результаты были получены при фиксировании  $\lambda$ , но численное решение системы оставалось неустойчивым из-за плохой сходимости.

Приемлемые для практического применения НП были получены при использовании системы (3.32).  $\chi_3$  и  $\chi_4$  вычислялись по формулам (3.28) ( $a = -1$ ;  $b = 0.5$ ;  $\lambda = 1000$ ;  $T = 2 \times 10^{-5}$  с). Уменьшение числа используемых кумулянтов до 4 порядка и возможность неитерационного решения системы (3.32) (приложение D) позволили находить НП даже для небольших отношений  $S/N$  (смотри таблицу 3.4).

Таблица 3.4 – НП для двухкомпонентных смоделированных данных, вычисленные путем решения системы (3.32)

Параметр	Точное значение	НП			
		$S/N_i = 1000$	$S/N_i = 500$	$S/N_i = 200$	$S/N_i = 100$
$N_1$	10	9.99±0.13	9.99±0.18	9.86±0.45	10.11±1.70
$q_1$	20000	19883±635	19747±950	19504±2248	18282±5013
$N_2$	2	2.04±0.23	2.07±0.32	2.22±0.75	2.59±1.64
$q_2$	50000	49847±1404	49783±1802	49735±4562	50931±11676

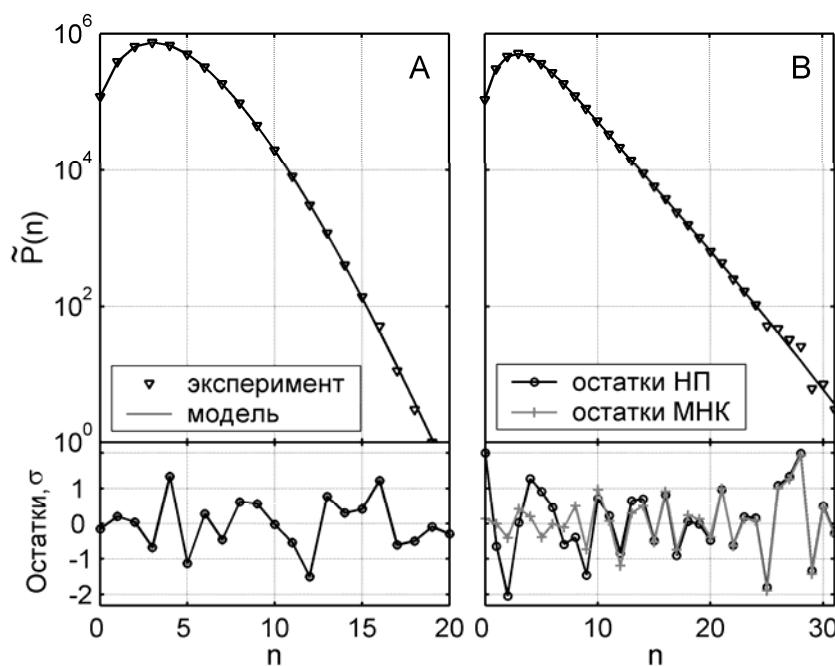
**Тестирование начальных приближений на измеренных данных.** Для тестирования НП на однокомпонентных данных была взята серия измерений красителя Alexa 488 (каждый раз концентрация красителя уменьшалась в два раза, начиная с 40нМ). Интервал наблюдения  $T = 8E-6$  с был выбран в 5 раз меньшим характеристического диффузионного времени для Alexa 488. Диффузионное время было оценено по методу FCS с помощью программы FCS Data Processor [10-A]. Экспериментальные данные были предоставлены кафедрой паразитологии Медицинского центра Лейденского университета (Department of Parasitology, LUMC, the Netherlands). Все измерения были выполнены с помощью СРОД ConfoCor 2 (Carl Zeiss, Germany), оснащенной объективом Zeiss Neofluar (40X 1.2 N.A.) и лавинными фотодиодами (SPCM-AQ, Perkin Elmer, USA). Аргоновый лазер использовался для возбуждения флуоресценции на длине волны 488 нм. В качестве кюветы использовалась стандартная 96 луночная плата с боросиликатным дном (Polyfiltronics). Измеряемый образец растворялся в PBS (фосфатный буфер: 0.035M phosphate, 0.15M NaCl, pH 7.6). Краситель Alexa 488 был приобретен у Molecular Probes (Molecular Probes Europe BV, The Netherlands).

Принимая во внимание, что анализ смоделированных данных показал, что фиксирование  $\lambda$  обеспечивает более точные и устойчивые оценки параметров, НП вычислялись по формулам (3.25), (3.28). Фоновый шум был оценен из дополнительного измерения чистого растворителя. Качество НП оценивалось путем сравнения с результатами последующего анализа по методу МНК. Взвешенный критерий  $\chi^2$  [85, 134], вычисленный для обоих наборов параметров (после генерации НП и после анализа), служил критерием качества. Результаты теста сведены в таблицу 3.5 и показаны на рисунке 3.2А. Оценки параметров, полученные по методу МНК, представлены для сравнения. Среднеквадратические отклонения параметров вычислялись по методу АСО. Интенсивность фонового

шума была оценена из дополнительного измерения чистого растворителя и принята равной 1500 отсчетов в секунду. Как видно из таблицы, НП очень близки к оценкам параметров, полученных по методу МНК. В то же время применение формул (3.25 – 3.27) приводило к нереалистичному завышению оценок для  $\lambda$  и смещенным оценкам для  $N$  и  $q$ .  $\tilde{P}(n) = P(n)T_{meas}/T$  ( $T_{meas}$  – время измерения).

Таблица 3.5 – НП, вычисленные для красителя Alexa 488, измеренного при различных концентрациях по формулам (3.25), (3.28)

Раствор	1		1/2		1/4		1/8	
Параметр	НП	МНК	НП	МНК	НП	МНК	НП	МНК
$N$	24.916	24.923±0.189	12.24	12.246±0.196	6.809	6.793±0.063	3.783	3.783±0.013
$q$	18500	18500±140.5	23400	23400±376.2	24223	24000±226.7	25500	25500±87.71
$a$	-1.463	-1.472±0.001	-1.046	-1.073±0.782	-0.59	-0.912±0.398	-0.665	-0.608±0.196
$b$	0.321	0.324±0.0004	0.224	0.228±0.136	0.196	0.207±0.048	0.182	0.178±0.013
$\chi^2$	0.668	0.666	1.338	1.224	1.538	0.844	0.422	0.412



А) результаты анализа однокомпонентных данных;

В) результаты анализа двухкомпонентных данных

**Рисунок 3.2 – Экспериментальные и теоретические РЧФ (вместе с соответствующими остатками), полученные на основе НП и МНК**

Для оценки алгоритма генерации начальных приближений для двухкомпонентных данных было произведено измерение смеси иммуноглобулина IgG, помеченного красителем Alexa 488, и фрагмента белка Fab, помеченного красителем Cy2. Фрагменты крысиного белка Fab2, помеченного красителем Cy2, были приобретены у Amersham (Amersham, The Netherlands). Крысиный иммуноглобулин IgG был помечен красителем Alexa 488 в соответствие с

рекомендациями производителя красителя. Детали протокола присоединения флуоресцентных меток опубликованы в [135].

НП генерировались путем решения системы (3.32) для набора параметров  $a$  и  $b$ , изменяющихся в заранее определенных пределах, и последующим выбором набора параметров, обеспечивающих минимальное значение критерия  $\chi^2$ . Как и для однокомпонентных данных (смотри таблицу 3.6 и рисунок 3.2В), НП близки к оценкам параметров, полученных по методу МНК. Точность и скорость данного алгоритма зависит от сетки разбиения диапазона изменения параметров  $a$  и  $b$ . С уменьшением шага разбиения точность возрастает, а скорость уменьшается (в нашем случае шаг равнялся 0.1 для  $a$  и 0.05 для  $b$ ).

Таблица 3.6 – НП, вычисленные для смеси IgG, помеченного Alexa 488, и фрагмента белка Fab, помеченного Cy2.  $T = 2 \times 10^{-5}$  с

Параметр	НП	МНК
$N_1$	5.477	5.166±0.089
$q_1$	32099	32768±872
$N_2$	0.387	0.528±0.052
$q_2$	90770	78824±2511
$\lambda$	1000 (фиксирован)	1000 (фиксирован)
$a$	-0.85	-0.769±0.014
$b$	0.25	0.296±0.018
$\chi^2$	1.26	0.78

Для оценки эффективности НП было вычислено отношение количества итераций МНК, требуемых для достижения глобального минимума, без НП и с ними. В первом случае для каждого параметра НП генерировались случайным образом (по равномерному закону распределения), поскольку количество итераций МНК зависит от близости НП от наилучших оценок параметров. Для расчета среднего количества итераций была проведена серия генераций НП. Вычисленное отношение для однокомпонентных данных (смотри таблицу 3.5) составило 5.1. Проведенный тест также позволил оценить увеличение устойчивости достижения глобального минимума при использовании НП. Только в 51.9% случаев (для случайных НП) был достигнут глобальный минимум. В остальных случаях процедура подгонки либо заканчивалась ошибками вычисления модели, либо приводила к попаданию в локальные минимумы.

Для двухкомпонентных данных (смотри таблицу 3.6) отношение составило 8.6 и количество удачных попыток достижения глобального минимума (для случайных НП) было 45.5%.

**Идентифицируемость модели в методе FIDA.** Теоретическая модель должна быть полностью идентифицируема, т.е. должен существовать только один набор параметров, характеризующий данные наилучшим образом, или, иначе



говоря, должен существовать только один глобальный минимум в пространстве целевого критерия. Это не выполняется для модели в методе FIDA. Существование по крайней мере трех минимумов (при фиксированном  $\lambda$ ) было предсказано теоретически (смотри подраздел 3.3.1). Для всех трех наборов параметров физически значимые параметры  $N$  и  $q$  остаются неизменными. Поскольку параметры  $a$  и  $b$  были введены для наилучшей аппроксимации профиля засветки и не имеют строгого физического смысла, и, дополнительно, все три комбинации  $a$  и  $b$  приводят к совершенно одинаковым значениям геометрических факторов  $\chi_k$ , существование нескольких глобальных минимумов не является слабым местом метода FIDA.

Для визуализации глобальных минимумов была построена поверхность критерия  $\chi^2$  при варьировании значений параметров  $a$  и  $b$  (параметры  $N$  и  $q$  были зафиксированы в оптимальные значения, смотри таблицу 3.5). Как видно из рисунка 3.3, имеется три четко выраженные минимума (числовые значения параметров  $N$ ,  $q$  представлены в таблице 3.5, а значения параметров  $a$  и  $b$ , полученные при решении системы (3.28), представлены в таблице 3.7), разделенные плоскостью, при которой значение критерия  $\chi^2$  стремится к бесконечности. Подробное изучение поверхности подтверждает соответствие представленных в таблице 3.7 значений  $a$  и  $b$  глобальным минимумам. Наличие плоскости разрыва очевидно из вида (3.28): все геометрические факторы  $\chi_k \rightarrow \infty$  при  $a$  и  $b$  удовлетворяющих уравнению  $2a + 3b + 2 = 0$ . Данная ситуация создана выбранной нормализацией в виде (1.84) и значительно усложняет анализ данных. Следовательно, наличие НП является еще более востребованным.

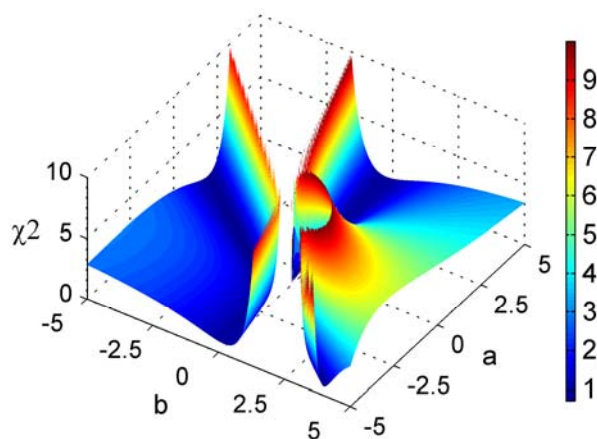


Рисунок 3.3 – Поверхность  $\chi^2$ , построенная при варьировании  $a$  и  $b$

Таблица 3.7 – Результат решения системы (3.28) для 40nM Alexa 488

Параметр	Набор 1	Набор 2	Набор 3
$a$	-1.393	-1.463	-4.643
$b$	-0.861	0.321	3.539
$\chi^2$	0.732	0.668	0.711

Вычисление критерия  $\chi^2$  показало, что имеются небольшие отличия в значениях  $\chi^2$  (теоретически они должны быть одинаковыми). Это может происходить только от погрешности вычислений.

### 3.4 Анализ распределения числа фотоотсчетов с коррекцией профиля засветки

#### 3.4.1 Нормализация в методах анализа распределения числа фотоотсчетов

Рассмотренные методы ФФС требуют нормализации к некоторому известному объему. Например, если для аппроксимации  $B(\mathbf{r})$  выбрать гауссовское распределение (1.2) и положить  $B(\mathbf{0}) = 1$ , все  $\chi_k$  могут быть вычислены как

$$\chi_{Gk} = \int_V B_G^k(\mathbf{r}) d\mathbf{r} = k^{-3/2} (\pi/2)^{3/2} w_0^2 z_0. \quad (3.33)$$

Гауссовское распределение имеет два параметра  $\omega_0$  и  $z_0$ . Оба этих параметра не могут служить в качестве параметров подгонки под реальный профиль засветки, поскольку они не влияют ни на форму распределения числа фотоотсчетов, ни на значения факториальных кумулянтов, а определяют лишь величину объема засветки (и, следовательно, количество молекул в освещенном объеме). В методе FCA форма профиля засветки полностью характеризуется  $\gamma$ -факторами, которые не зависят от  $\omega_0, z_0$  (для трехмерного гауссовского распределения  $\gamma_{Gk} = 1/k^{3/2}$ ). Так как объем наблюдения в ФФС эксперименте не может быть измерен напрямую, значения параметров  $\omega_0$  и  $z_0$  неизвестны, и модель является неидентифицируемой без дополнительных знаний (условий).

Нормализация может быть произведена либо с помощью дополнительных уравнений, либо путем пересчета концентрации к среднему количеству молекул в некотором объеме известного размера. Для определения количества молекул в освещенной области в ФФС используют понятие приведенного объема  $V_{PSF} = \chi_1$ , количественно равного значению интеграла от функции  $B(\mathbf{r})$ . Иногда используют значение так называемого эффективного объема  $V_{eff} = \chi_1^2 / \chi_2$ , традиционно принятого в FCS. Действительно, нормализация к  $V_{PSF}$  в форме  $\langle C \rangle = N_{PSF} / V_{PSF}$  (смотри (1.93) в методе FCA) (или к  $V_{eff}$  в форме  $\langle C \rangle = N_{eff} / V_{eff}$ ) разрешает неопределенность, поскольку теперь  $\omega_0, z_0$  взаимно сокращаются и параметр  $N_{PSF}$  ( $N_{eff}$ ) становится идентифицируемым. Можно убедиться, что такой же результат может быть получен при использовании дополнительного нормировочного условия в виде  $V_{PSF} = 1$  (или  $V_{eff} = 1$ ) и вычислению из него произведения  $\omega_0^2 z_0$ . Выбор нормализации определяется необходимостью сравнения  $N$  с результатами других методов, например, метода FCS, и может быть произвольным.

Связь между  $N_{eff}$  и  $N_{PSF}$  может быть найдена через связь объемов  $V_{PSF}$  и  $V_{eff}$

$$V_{eff} = V_{PSF} / \gamma_2, \quad (3.34)$$

$$N_{eff} = N_{PSF} / \gamma_2. \quad (3.35)$$

Значение  $N_{FIDA}$  (в методе FIDA) количественно равно  $N_{eff}$ , что следует из замены переменных  $N_{FIDA} = \langle C \rangle \chi_1^2 / \chi_2$  в (3.20). Соответственно  $V_{FIDA} = V_{eff} = \chi_1^2 / \chi_2$ .

Нормализация только к объему известного размера решает проблему идентифицируемости модели в методах РСН и FCA, но приводит к возникновению зависимости между параметрами модели в случае введения коррекции профиля засветки. В разделе 4.2 будет показано, что яркость и число молекул становятся функциями параметров коррекции  $F_k$ . Причиной этому служит предположение  $B(\mathbf{0}) = 1$  в выводе моделей в методах РСН и FCA. В реальности  $B(\mathbf{0}) \neq 1$ , что приводит к дополнительному множителю, присутствующему в произведении с  $q$ .

Для определения  $B(\mathbf{0})$  в методе FIDA было введено дополнительное уравнение нормализации (смотри (1.84)). Подставляя (1.83) в (1.82) получим

$$G(\xi) = \exp \left( \lambda T(\xi - 1) + \sum_i A_0 \langle C \rangle_i \int \left( e^{(\xi-1)q_i T B_0 e^{-x}} - 1 \right) (x + ax^2 + bx^3) dx \right). \quad (3.36)$$

Введение начальных параметров  $A_0$ ,  $B_0$  и двух уравнений нормализации делают яркость и число молекул независимыми от параметров профиля засветки, поскольку теперь начальные параметры становятся функциями параметров  $a$  и  $b$

$$B_0 = \frac{8(2a + 6b + 1)}{2a + 3b + 2}, \quad A_0 = \frac{2a + 3b + 2}{8(2a + 6b + 1)^2}. \quad (3.37)$$

Нормализация вида (1.84) удобна также тем (показано в подразделе 3.3.1), что она позволяет находить оценки концентрации и яркости для однокомпонентных данных без знания вида  $B(\mathbf{r})$ .

### 3.4.2 Метод анализа распределения числа фотоотсчетов с коррекцией профиля засветки

Для анализа РЧФ предложены два метода: FIDA и РСН. Основным недостатком метода FIDA является возможность аппроксимации двухкомпонентных данных однокомпонентной моделью, что может привести к неправильной интерпретации результатов. Фиксирование параметров коррекции профиля засветки в значения, полученные в калибровочном эксперименте, помогает исправить ситуацию, но требует точного соответствия условий проведения измерений и калибровочного эксперимента. К другим недостаткам метода FIDA можно отнести наличие линии разрыва в области определения параметров профиля засветки и наличие нескольких глобальных минимумов (смотри подраздел 3.3.2). Метод РСН значительно уступает по скорости

вычислений методу FIDA вследствие применения операции свертки и имеет зависимость параметров  $N, q$  от параметров коррекции профиля засветки.

В работе предлагается метод анализа РЧФ, устраняющий недостатки методов FIDA/РСН. Метод основан на ПФЧФ с коррекцией профиля засветки, полученной в разделе 3.1.

Разработка метода анализа РЧФ включает вывод теоретической модели, соединяющей свойства вещества и анализируемую характеристику, и выбор способа получения оценок параметров модели. Для получения оценок параметров предлагается использовать МНК с градиентным методом оптимизации Марквардта-Левенберга. В качестве целевого критерия предлагается использовать взвешенный критерий  $\chi^2$ . Соответственно потребуется разработка НП для параметров модели и исследование статистических свойств шума РЧФ. Начальные приближения, разработанные для метода FIDA (раздел 3.3), с незначительной модификацией могут быть также использованы и в нашем случае. Исследование шума в экспериментальных данных было произведено в [27, 28].

Запишем выражение для интегралов  $\chi_k$  несимметричного гауссовского распределения с учетом  $B_0 \neq 1$

$$\chi_{Gk} = \int_V B_G^k(\mathbf{r}) d\mathbf{r} = B_0^k k^{-3/2} (\pi/2)^{3/2} \omega_0^2 z_0. \quad (3.38)$$

Введем обозначение  $A_0 = \omega_0^2 z_0$  в выражении (3.38). Значения начальных параметров  $A_0, B_0$  определим из решения системы уравнений

$$\int_V B(\mathbf{r}) d\mathbf{r} = \chi_1, \quad \int_V B^2(\mathbf{r}) d\mathbf{r} = \chi_2. \quad (3.39)$$

После подстановки (3.38) в (1.89) система (3.39) примет вид

$$\begin{aligned} (1 + F_1) \chi_{G1} &= B_0 A_0 (1 + F_1) (\pi/2)^{3/2} = \chi_1, \\ (1 + F_2) \chi_{G2} &= B_0^2 A_0 (1 + F_2) \pi^{3/2} / 8 = \chi_2. \end{aligned} \quad (3.40)$$

Из системы (3.40) получим

$$A_0 = \frac{(1 + F_2) \chi_1^2}{\pi^{3/2} (1 + F_1)^2 \chi_2}, \quad B_0 = \frac{2\sqrt{2} (1 + F_1) \chi_2}{(1 + F_2) \chi_1}. \quad (3.41)$$

ПФ числа фотоотсчетов (3.2) примет вид [20-A]

$$\begin{aligned} G(\xi) &= \exp \left\{ \lambda T (\xi - 1) + \sum_i \pi^{3/2} A_0 \langle C \rangle_i \sum_{k=1}^{\infty} \frac{((\xi - 1) B_0 q_i T)^k (1 + F_k)}{k! (2k)^{3/2}} \right\} = \\ &= \exp \left\{ \lambda T (\xi - 1) + \frac{(1 + F_2)}{(1 + F_1)^2} \sum_i N_{2MNi} \sum_{k=1}^{\infty} \frac{((\xi - 1) 2\sqrt{2} (1 + F_1) q_{2MNi} T)^k (1 + F_k)}{(1 + F_2)^k k! (2k)^{3/2}} \right\}, \end{aligned} \quad (3.42)$$

где введены новые обозначения  $N_{2MN} = \langle C \rangle \chi_1^2 / \chi_2$  и  $q_{2MN} = q \chi_2 / \chi_1$ . Индекс  $2MN$  (two moments normalized) обозначает нормализацию вида (3.39). Произведенная

замена переменных совпадает с (3.21), что означает  $N_{FIDA} = N_{2MN}$  и  $q_{FIDA} = q_{2MN}$ . Перепишем выражение (1.89) в виде  $\chi_k = (1 + F_k)\chi_{Gk}$ . Из последнего выражения и (3.41) следует, что выбранная система нормировочных уравнений приводит к независимости параметров  $N_{2MN}$  и  $q_{2MN}$  от  $F_k$  (параметры коррекции сокращаются).

Вывод формул (3.4), (3.7) для нормализации вида (1.84) приведен в приложении Е. Для сравнения с известными методами анализа РЧФ данный тип нормализации был также применен к методу РСН с коррекцией профиля засветки (смотри приложение F).

Можно показать, что игнорирование значения  $B_0$  в методе OFC РСН и применение нормализации к эффективному объему  $V_{eff}$  приводит к следующим значениям начальных параметров

$$A_0 = (1 + F_2)\chi_1^2 / (\pi^{3/2}\chi_2(1 + F_1)^2), \quad B_0 = 1. \quad (3.43)$$

Из сопоставления (3.41) и (3.43) и принимая во внимание (3.42) следует, что оценка количества молекул в объеме наблюдения не изменится (значения  $A_0$  в обоих случаях одинаковы), а оценка яркости в методе OFC РСН (обозначим как  $q_C$ ) будет связана с  $q_{2MNi}$  формулой

$$q_{Ci} = B_0 q_{2MNi} = 2\sqrt{2}(1 + F_1)q_{2MNi} / (1 + F_2). \quad (3.44)$$

Принимая во внимание, что  $q_{2MN}$  не зависит от параметров коррекции, получим, что оценка  $q_{Ci}$  становится функцией параметров  $F_k$ . Подобная зависимость между параметрами модели не позволяет сравнивать результаты нескольких экспериментов напрямую и проводить глобальный анализ со связыванием  $q_{Ci}$  при различных значениях параметров коррекции.

При нормализации к объему засветки  $V_{PSF}$  в методе OFC РСН выражения для начальных параметров примут вид

$$A_0 = \chi_1 / ((\pi/2)^{3/2}(1 + F_1)), \quad B_0 = 1. \quad (3.45)$$

Сравнивая выражения (3.42) с учетом (3.41) и (3.45) получим, что параметр  $N_{PSFCi}$  (количество молекул в методе OFC РСН) также становится функцией параметров коррекции  $F_i$ .

$$N_{PSFCi} = N_{2MNi} / B_0 = (1 + F_2)N_{2MNi} / ((1 + F_1)2\sqrt{2}). \quad (3.46)$$

Устранение зависимости яркости и количества молекул от параметров коррекции профиля засветки приводит к уменьшению отклонений зависимых параметров от их средних значений, что следует из применения закона распространения ошибки к выражениям (3.44) и (3.46). Например, для яркости  $q_C$  в случае коррекции первого порядка получим

$$(\delta q_{Ci})^2 = 8q_{2MNi}^2 (\delta F_1)^2 + 8(1 + F_1)^2 (\delta q_{2MNi})^2. \quad (3.47)$$

Следовательно,  $\delta q_{Ci} > \delta q_{2MNi}$ .

### 3.4.3 Тестирование метода

Программная реализация метода была выполнена в Borland Builder 5.0 и внедрена в программу Data Analyser. РЧФ находилось путем вычисления ПФ (3.42) и применения преобразования (1.21). Анализ РЧФ основан на МНК и предполагает подгонку теоретической модели (3.42), (1.21) к экспериментально полученному РЧФ. Качество подгонки проверялось путем вычисления значения взвешенного критерия  $\chi^2$  и визуальным анализом остатков. Среднеквадратические отклонения оцениваемых параметров находились по методу АСО. Тестирование метода проводилось на смоделированных и измеренных данных.

**Тестирование метода на смоделированных данных.** Моделирование РЧФ проводилось согласно методам, рассмотренным в подразделе 2.3.2. Как и следовало ожидать, теоретические кривые, вычисленные соответственно тестируемой модели (обозначим как 2MN РСН), и модели OFC РСН полностью совпали. Различие наблюдалось только в оценках параметров модели. Полученные оценки были пересчитаны от одного вида нормализации к другому по формулам (3.44). Например, для РЧФ, смоделированного с параметрами  $q_C = 60000$ ,  $N_{eff} = 5$ ,  $F_1 = 0.4$ ,  $T = 5E-5$  с и отношения  $S/N = 300$ , оценки параметров модели OFC РСН составили:  $q_C = 60680 \pm 910$ ,  $N_{eff} = 5.01 \pm 0.03$ ,  $F_1 = 0.42 \pm 0.02$ . Оценки параметров модели 2MN РСН были соответственно:  $q_{2MN} = 15120 \pm 85$ ,  $N_{2MN} = 5.01 \pm 0.03$ ,  $F_1 = 0.42 \pm 0.02$ . После пересчета  $q_{2MN}$  по формуле (3.44) получили оценку  $q_C = 60680$ , что в точности совпадает с оценкой, полученной для OFC РСН.

Из проведенного эксперимента видно, что среднеквадратические отклонения параметра  $q_{2MN}$  значительно ниже, чем отклонения параметра  $q_C$ . Это ожидаемый результат, поскольку вследствие зависимости параметра  $q_C$  от  $F_k$  в модели OFC РСН среднеквадратические отклонения для  $q_C$  будут дополнительно включать среднеквадратические отклонения  $F_k$  (для нормализации к объему засветки среднеквадратические отклонения параметра  $N_{PSF}$  также будут включать среднеквадратические отклонения  $F_k$ ). Тестирование на большом количестве данных показало увеличение точности оценки параметра  $q_{2MN}$  и его независимость от параметров коррекции  $F_k$ .

Кроме отсутствия зависимости между параметрами модели, разработанный метод характеризуется более высокой производительностью. Время, требуемое для проведения вычислений в разработанном методе и методах OFC РСН и FIDA при различном числе точек РЧФ (для различной длительности интервала наблюдения  $T$ ), представлено на рисунке 3.4. Для увеличения точности сравнения все методы были приведены к одинаковой аппроксимации профиля засветки и нормализации. Соответственно, в тестируемом методе вычисления проводились по формулам (3.42), (1.21), в методе OFC РСН – по формулам (1.91), (1.86), в методе FIDA – по формулам (E.7), (E.8), (1.21) (параметры:  $q_C = 1.2E5$ ,  $N_{eff} = 5$ ,

$F = 0.4$ ,  $T = 5E-6 \div 1E-4$  с). Как видно из рисунка, разработанный метод характеризуется значительно меньшим временем вычисления (в 10 и более раз).

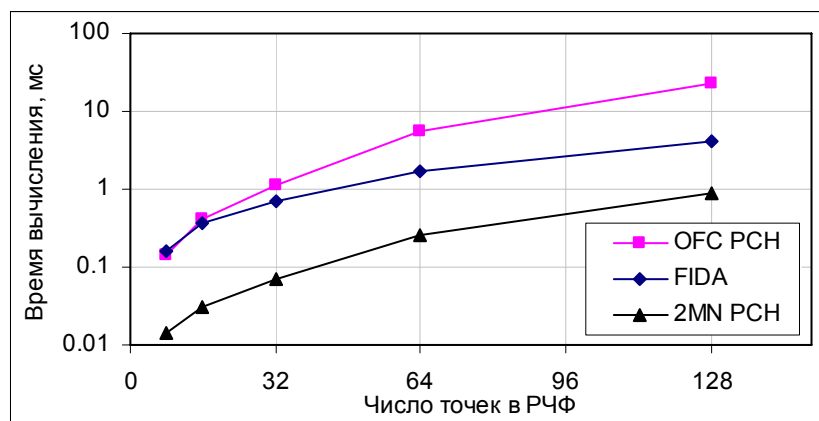
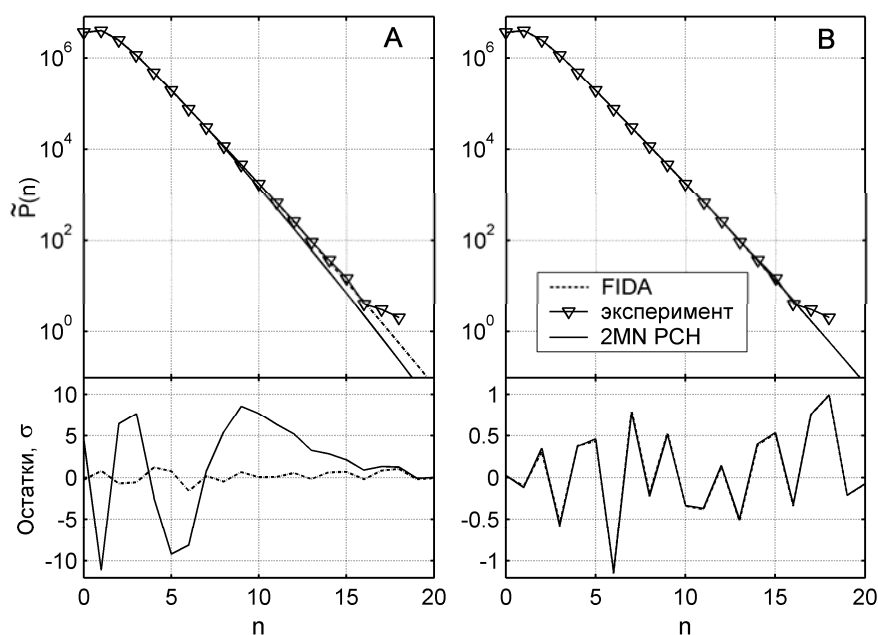


Рисунок 3.4 – Сравнение методов по скорости вычисления

**Тестирование метода на измеренных данных.** Для тестирования метода на измеренных данных было проведено измерение смеси Alexa 488 и Флуоресцеина. Экспериментальные данные были предоставлены кафедрой паразитологии Медицинского центра Лейденского университета (Department of Parasitology, LUMC, the Netherlands). Все измерения были выполнены с помощью СПОД ConfoCor 2 (CarlZeiss, Germany), оснащенной объективом Zeiss Neofluar (40X 1.2 N.A.). Аргонный лазер использовался для возбуждения флуоресценции на длине волны 488 нм. В качестве кюветы использовалась стандартная 96 луночная плата с кварцевым дном (Polyfiltronics). Измеряемый образец растворялся в PBS (0.035 M phosphate, 0.15 M NaCl, pH 7.6). Краситель Alexa 488 и Флуоресцеин были приобретены у Molecular Probes (Molecular Probes Europe BV; The Netherlands). При времени измерения 120 с получили число зарегистрированных фотонов около  $16E+6$ . РЧФ рассчитывалось при  $T = 1E-5$  с, что соответствует  $S/N = 2400$  (смотри (2.36)). Интенсивность фонового шума была оценена из дополнительного измерения чистого растворителя и принята равной 1000 отсчетов в секунду.

На рисунке 3.5 и в таблице 3.8 представлены результаты анализа смеси Alexa 488 и Флуоресцеина. Значения критерия  $\chi^2$  и оценки параметров, полученные с использованием разработанного метода и методов FIDA, OFC PCH (индексы в обозначениях яркости и количества молекул опущены) представлены для сравнения. Вначале был проведен анализ с учетом одной компоненты. Значение критерия  $\chi^2$ , равномерность распределения остатков, малые среднеквадратические отклонения параметров при применении метода FIDA свидетельствовали о хорошем качестве подгонки, в то время как результаты анализа разработанным методом и методом OFC PCH ясно свидетельствовали о наличии нескольких компонент (смотри рисунок 3.5А). Этот пример наглядно демонстрирует возможность переоценки экспериментальных данных с

использованием метода FIDA. Анализ с включением двух компонент показал очень близкие результаты для всех трех методов. Теоретические кривые и, следовательно, остатки стали практически неразличимыми (смотри рисунок 3.5B).



**Рисунок 3.5 – Результаты анализа смеси Alexa 488 и Флуоресцеина однокомпонентной (А) и двухкомпонентной (В) моделью**

Таблица 3.8 – Результаты анализа смеси Alexa 488 и Флуоресцеина одно- и двухкомпонентной моделью

Параметр	1-компонентная модель		2-компонентная модель		
	FIDA	2MN PCH	FIDA	OFC PCH	2MN PCH
$N_1$	4.436±0.005	4.477±2.675	5.521±0.732	6.101±0.158	6.101±0.158
$q_1$	29844±37	29570±18189	16770±852	78750±37130	15060±449
$N_2$	–	–	0.661±0.430	0.634±0.011	0.640±0.010
$q_2$	–	–	60270±15570	331480±29496	63400±717
$F/a$	-1.069±0.004	1.714±3.427	-0.897±0.067	0.849±0.033	0.849±0.033
$b$	0.518±0.008	–	0.244±0.041	–	–
$\chi^2$	0.701	44.01	0.472	0.438	0.438

Как видно из представленных результатов, среднеквадратические отклонения параметров  $q_1$ ,  $q_2$  значительно меньше для разработанного метода, чем для метода OFC PCH, что подтверждает ранее сделанный вывод об увеличении точности метода вследствие устранения зависимости между параметрами модели. Разработанный метод характеризуется большей точностью определения параметров и в сравнении с методом FIDA. Устранение зависимости между параметрами модели предоставило также возможность проведения глобального анализа нескольких распределений со связыванием этих параметров.



### 3.5 Основные результаты и выводы

1. Получено аналитическое выражение для производящей функции числа фотоотсчетов с коррекцией профиля засветки. Показано, что она может быть представлена произведением трех ПФ: производящей функции числа фотонов, регистрируемых из основной области засветки, точно аппроксимируемой функцией  $B(\mathbf{r})$ , производящей функции числа фотонов из дополнительной области засветки (коррекция профиля засветки) и производящей функции фонового сигнала.

Показана эквивалентность методов РСН, FIDA, FCA. Теоретические модели всех вышеперечисленных методов могут быть выведены из полученной производящей функции распределения фотоотсчетов.

2. Предложен быстрый и эффективный метод генерации НП для метода FIDA. С незначительными изменениями (зависящими от примененной аппроксимации профиля засветки) он может быть также использован и для методов РСН, FCA. Его применение позволяет значительно ускорить анализ данных (в тестовых примерах на измеренных данных в среднем в 7 раз) и повысить его устойчивость (в тестовых примерах вероятность получения вычислительных ошибок или неверных результатов была снижена почти на 50%). Кроме генерации НП предложенный метод может быть также использован как самостоятельный метод быстрой оценки параметров исследуемого вещества. Метод FIDA и алгоритм генерации НП для него были реализованы в программе FCS Data Processor.

3. Проведена проверка идентифицируемости модели в методе FIDA. Показано, что существует линия разрыва в области определения параметров и при фиксации параметра  $\lambda$  модель характеризуется тремя эквивалентными наборами параметров. Поскольку физически значимые оценки концентрации и яркости остаются неизменными для всех трех наборов параметров, выбор набора параметров может быть произвольным.

4. Разработан метод анализа РЧФ, основанный на разложении производящей функции числа фотоотсчетов в ряд Тейлора с введением параметров коррекции профиля засветки. Разработанный метод характеризуется большей чувствительностью при разрешении многокомпонентных систем по сравнению с методом FIDA, отсутствием зависимости между параметрами модели, выявленной в методе РСН с однотипной коррекцией профиля засветки, большей точностью определения молекулярной яркости и большей скоростью вычисления по сравнению с методами РСН, FIDA.

## ГЛАВА 4 КУМУЛЯНТНЫЙ АНАЛИЗ ФОТООТСЧЕТОВ

### 4.1 Кумулянтный анализ фотоотсчетов с коррекцией профиля засветки

#### 4.1.1 Коррекция профиля засветки в кумулянтном анализе

Для аппроксимации профиля засветки в кумулянтном анализе применяется трехмерное гауссовское распределение в виде (1.2) с  $B_0 = 1$ . Данная аппроксимация не позволяет получить хорошей подгонки под экспериментальные данные. Для коррекции профиля засветки была предложена подгонка  $\gamma$ -факторов (интегрированных характеристик профиля засветки) вместе с другими параметрами. Предлагается вычислять  $\gamma_1, \gamma_2$  из аппроксимации (1.95) и использовать калибровочный эксперимент для оценки остальных  $\gamma$ -факторов [99]. Это требует проведения дополнительного эксперимента с полным совпадением условий измерения для обеспечения повторяемости значений  $\gamma_3, \gamma_4, \dots$ .

Недостатком методики подгонки  $\gamma$ -факторов является зависимость числа параметров модели от числа анализируемых кумулянтов и невозможность проведения анализа отдельного набора кумулянтов без постановки калибровочного эксперимента. Число  $\gamma$ -факторов на единицу меньше числа экспериментальных точек. Следовательно, общее число параметров модели превышает число экспериментальных точек, что ведет к получению бесконечного числа решений. Для преодоления этой проблемы можно воспользоваться подходами, предложенными для анализа РЧФ. Применение полиномиальной аппроксимации, а также коррекции профиля засветки в виде (1.89) позволит стандартизировать процедуру анализа, устранив зависимость числа параметров модели от экспериментальных данных.

Выведенная в разделе 3.1 производящая функция числа фотоотсчетов позволяет получить выражения для факториальных кумулянтов любого порядка путем дифференцирования ПФ. Подставив (3.2) в (1.19), получим [21-А]

$$\begin{aligned} K_1 &= \lambda T + (1 + F_1) \chi_{G1} \sum_i \langle C \rangle_i q_i T \\ K_k &= (1 + F_k) \chi_{Gk} \sum_i \langle C \rangle_i q_i^k T^k. \end{aligned} \tag{4.1}$$

**Нормализация к эффективному объему.** Нормализация может быть выполнена путем пересчета концентрации к числу молекул в эффективном объеме  $N_{effi} = \langle C \rangle_i V_{eff}$ . Для отличия количества молекул в объеме наблюдения в модели с учетом коррекции от модели с гауссовской аппроксимацией профиля засветки введем параметер  $N_{effC}$ . Принимая во внимание (3.38) и (1.89), получим

$$N_{effCi} = \langle C \rangle_i \frac{(1+F_1)^2 \chi_{G1}^2}{(1+F_2) \chi_{G2}} = \langle C \rangle_i \frac{(1+F_1)^2}{(1+F_2)} \pi^{3/2} \omega_0^2 z_0. \quad (4.2)$$

Для коррекции первого порядка ( $F_1 \neq 0, F_k = 0, k = 2, 3, \dots$ ) выражение (4.1) примет вид ( $\gamma_k = 1/k^{3/2}$ )

$$\begin{aligned} K_1 &= \lambda T + \sum_i N_{effCi} q_{Ci} T / (2\sqrt{2}(1+F_1)) \\ K_k &= \gamma_2 \gamma_k \sum_i N_{effCi} q_{Ci}^k T^k / (1+F_1)^2, \quad k = 2, 3, \dots, \end{aligned} \quad (4.3)$$

где  $q_{Ci} = B_0 q_i$ . Для коррекции второго порядка ( $F_1 \neq 0, F_2 \neq 0, F_k = 0, k = 3, 4, \dots$ ) получим

$$\begin{aligned} K_1 &= \lambda T + \frac{(1+F_2)}{2\sqrt{2}(1+F_1)} \sum_i N_{effCi} q_{Ci} T \\ K_2 &= \frac{(1+F_2)^2}{8(1+F_1)^2} \sum_i N_{effCi} q_{Ci}^2 T^2 \\ K_k &= \gamma_2 \gamma_k \frac{(1+F_2)}{(1+F_1)^2} \sum_i N_{effCi} q_{Ci}^k T^k, \quad k = 3, 4, \dots \end{aligned} \quad (4.4)$$

**Нормализация к объему наблюдения.** Нормализация производится путем пересчета концентрации к числу молекул в объеме наблюдения

$$N_{PSFi} = \langle C \rangle_i V_{PSF} = \langle C \rangle_i (1+F_1) \chi_{G1} = \langle C \rangle_i (1+F_1) B_0 (\pi/2)^{3/2} \omega_0^2 z_0. \quad (4.5)$$

После ввода нового параметра  $N_{PSFCi} = N_{PSFi} / B_0$  получим

$$\begin{aligned} K_1 &= \lambda T + \sum_i N_{PSFCi} q_{Ci} T \\ K_k &= \gamma_k (1+F_k) \sum_i N_{PSFCi} q_{Ci}^k T^k / (1+F_1), \quad k = 2, 3, \dots, \end{aligned} \quad (4.6)$$

Для отличия методов кумулянтного анализа, основанных на полученных выражениях, от метода FCA будем добавлять префикс OFC с порядковым номером коррекции, например OFCII FCA (или просто OFC FCA, если порядок коррекции не имеет значения).

Полученные теоретические выражения (модели) позволяют значительно улучшить качество подгонки под экспериментальные данные вследствие введения коррекции профиля засветки. Другим достоинством разработанных моделей является отсутствие зависимости количества параметров модели от количества анализируемых кумулянтов и возможность анализа одиночного набора кумулянтов без проведения калибровочного эксперимента.

Коррекция профиля засветки ведет к зависимости  $q_C$  и  $N_{PSFCi}$  от параметров коррекции профиля засветки. Как и для аналогичного метода анализа РЧФ (смотри раздел 3.4), зависимость параметров можно устранить, введя нормализацию на объем наблюдения и яркость в фокусе засветки. После подстановки (3.42) в (1.19) (или после подстановки (3.41) в (3.38) и далее в (1.92)) получим

$$K_1 = \lambda T + \sum_i N_{2MNi} q_{2MNi} T$$

$$K_k = \gamma_2 \gamma_k \frac{(1+F_2)(1+F_k)}{(1+F_1)^2} \left( \frac{2\sqrt{2}(1+F_1)}{1+F_2} \right)^k \sum_i N_{2MNi} (q_{2MNi} T)^k. \quad (4.7)$$

Выражения для первых четырех кумулянтов модели с коррекцией профиля засветки и нормализацией вида (1.84) (обозначим как 2MN FCA) примут вид

$$K_1 = \lambda T + \sum_i N_{2MNi} q_{2MNi} T$$

$$K_2 = \sum_i N_{2MNi} q_{2MNi}^2 T^2$$

$$K_3 = \frac{8(1+F_1)}{3\sqrt{3}(1+F_2)^2} \sum_i N_{2MNi} (q_{2MNi} T)^3$$

$$K_4 = \frac{2\sqrt{2}(1+F_1)^2}{(1+F_2)^3} \sum_i N_{2MNi} (q_{2MNi} T)^4. \quad (4.8)$$

#### 4.1.2 Полиномиальная аппроксимация профиля засветки

Подставляя (1.83) в (1.70) и принимая во внимание (1.84), запишем

$$\chi_k = \int_0^\infty (B_0 e^{-x})^k A_0 (x + ax^2 + bx^3) dx, \quad (4.9)$$

где  $B_0 = 8u/v$ ,  $A_0 = v/8u^2$  и  $u = 2a + 6b + 1$ ,  $v = 2a + 3b + 2$ . После интегрирования по  $x$  получим

$$\chi_k = A_0 B_0^k (2ak + 6b + k^2) / k^4. \quad (4.10)$$

Таким образом, модель с полиномиальной аппроксимацией профиля засветки и нормализацией вида (1.84) (обозначим как POLY FCA) примет вид

$$K_1 = \lambda T + \sum_i N_{POLYi} q_{POLYi} T$$

$$K_2 = \sum_i N_{POLYi} q_{POLYi}^2 T^2$$

$$K_k = A_0 B_0^k (2ak + 6b + k^2) / k^4 \times \sum_i N_{POLYi} q_{POLYi}^k T^k, \quad (4.11)$$

где  $N_{POLYi} = \langle C \rangle_i \chi_1^2 / \chi_2$ ,  $q_{POLYi} = q_i \chi_2 / \chi_1$ .

## 4.2 Оценки яркости и количества молекул для различных видов аппроксимации профиля засветки и нормализации

Формулы оценивания  $N$  и  $q$  для различных видов аппроксимации профиля засветки и нормализации могут быть получены из решения систем (1.93), (4.3), (4.4), (4.11), (4.8). Для пересчета оценок от одного вида аппроксимации профиля засветки и нормализации к другому приравняем правые части указанных систем, принимая во внимание равенство экспериментальных факториальных кумулянтов

в левой части уравнений. Полученные формулы затем можно будет использовать не только для кумулянтного анализа, но и для методов анализа РЧФ вследствие эквивалентности этих методов (смотри раздел 3.2).

**Однокомпонентная модель с нормализацией к эффективному объему.** Запишем выражения для факториальных кумулянтов с нормализацией к эффективному объему без коррекции профиля засветки

$$\begin{aligned} K_1 &= \lambda T + N_{eff} q T / 2^{3/2} \\ K_k &= N_{eff} q^k T^k / (2k)^{3/2}, \quad k = 2, 3, \dots \end{aligned} \quad (4.12)$$

Из первых двух уравнений системы (4.12) получим

$$N_{eff} = (K_1 - \lambda T)^2 / K_2. \quad (4.13)$$

Точно такая же формула для оценки  $N_{effC}$  может быть получена из (4.3) и (4.4). Таким образом, коррекция профиля засветки не влияет на оценку количества молекул в эффективном объеме ( $N_{effC} = N_{eff}$ ).

Из решения систем (4.12), (4.3), (4.4) получим

$$q = \frac{2\sqrt{2}K_2}{(K_1 - \lambda T)T}, \quad q_{10C} = \frac{2\sqrt{2}(1+F_1)K_2}{(K_1 - \lambda T)T}, \quad q_{20C} = \frac{2\sqrt{2}(1+F_1)K_2}{(1+F_2)(K_1 - \lambda T)T}. \quad (4.14)$$

Сравнивая формулы для оценки яркости с различным порядком коррекции и без нее, получим

$$q_{10C} = (1+F_1)q, \quad q_{20C} = (1+F_1)q/(1+F_2), \quad q_{20C} = q_{10C}/(1+F_2), \quad (4.15)$$

где  $q, q_{10C}, q_{20C}$  яркость в моделях без коррекции, первого и второго порядка соответственно. Формулы пересчета совпадают с формулами, полученными в подразделе 3.4.2.

Использование формул (4.15) позволяет получить независимые от параметров коррекции значения молекулярной яркости и, следовательно, сравнивать результаты, полученные при различных значениях параметров коррекции.

**Однокомпонентная модель с нормализацией к объему наблюдения.** Из двух первых уравнений системы (4.6) (записанной для коррекции второго порядка) получим

$$N_{PSFC} = \frac{(1+F_2)(K_1 - \lambda T)^2}{(1+F_1)2\sqrt{2}K_2}, \quad q_C = \frac{2\sqrt{2}(1+F_1)K_2}{(1+F_2)(K_1 - \lambda T)T}. \quad (4.16)$$

Вывод формулы пересчета для  $N_{PSFC}$  очевиден из (4.13) и (4.16) (формулы для пересчета  $q_C$  не изменятся)

$$N_{eff} = \frac{(1+F_1)N_{PSFC}}{(1+F_2)\gamma_2}. \quad (4.17)$$

**Однокомпонентная модель с нормализацией на объем наблюдения и яркость в фокусе засветки.** Формулы для оценки количества молекул, полученные из решения систем (4.11), (4.8), имеют такой же вид, как и для случая нормализации к эффективному объему (3.25). Это ожидаемый результат, поскольку формула, выражающая эффективный объем  $V_{eff} = \chi_1^2 / \chi_2$ , является комбинацией уравнений нормализации (1.84). Следовательно,  $N_{POLY} = N_{eff} = N_{2MN}$  и

$$q_{POLY} = q_{2MN}, \quad q_{POLY} = \gamma_2 q,$$

$$q_{POLY} = \gamma_2 q_{1OC} / (1 + F_1), \quad q_{POLY} = \gamma_2 q_{2OC} (1 + F_2) / (1 + F_1). \quad (4.18)$$

**Двух- и более компонентная модель.** В общем случае для пересчета  $N$  и  $q$  требуется решение системы нелинейных уравнений, соединяющих правые части системы (1.93), записанной для требуемых аппроксимаций и нормализаций, т.е. любые комбинации систем (4.3), (4.4), (4.11), (4.8). Например, для пересчета из полиномиальной аппроксимации с нормализацией вида (1.84) в гауссовскую с коррекцией профиля засветки, получим

$$\frac{1}{2\sqrt{2}} \frac{(1 + F_2)}{(1 + F_1)} \sum_i N_{2OCi} q_{2OCi} = \sum_i N_{POLYi} q_{POLYi}$$

$$\frac{1}{8} \frac{(1 + F_2)^2}{(1 + F_1)^2} \sum_i N_{2OCi} q_{2OCi}^2 = \sum_i N_{POLYi} q_{POLYi}^2$$

$$\frac{1}{6^{3/2}} \frac{(1 + F_2)}{(1 + F_1)^2} \sum_i N_{2OCi} q_{2OCi}^3 = (2a + 2b + 3) \frac{64u}{27v^2} \sum_i N_{POLYi} q_{POLYi}^3$$

$$\frac{1}{8^{3/2}} \frac{(1 + F_2)}{(1 + F_1)^2} \sum_i N_{2OCi} q_{2OCi}^4 = (4a + 3b + 8) \frac{4u^2}{v^3} \sum_i N_{POLYi} q_{POLYi}^4. \quad (4.19)$$

Соответственно, требуется знание параметров коррекции профиля засветки с обеих сторон. Система уравнений (4.19) для двухкомпонентной модели ( $i = 1, 2$ ) может быть сведена к полиному третьего порядка (смотри приложение С) и решена как численно, так и аналитически.

### 4.3 Начальные приближения для кумулянтного анализа

Так как в основе метода FCA лежит система уравнений, связывающих экспериментальные и теоретические факториальные кумулянты, для генерации НП требуется решение системы с числом уравнений, равным числу оцениваемых параметров. Вывод некоторого общего выражения для НП не является целесообразным, поскольку точность получаемых оценок параметров сильно зависит от порядка используемых кумулянтов.

Для двухкомпонентных данных найти НП для всех параметров не представляется возможным (смотри раздел 3.3 и [22-А]). В этом случае можно использовать алгоритм, основанный на разбиении области изменения параметров

коррекции профиля засветки и вычисления НП в узлах полученной решетки (смотри раздел 3.3).

НП для метода кумулянтного анализа с полиномиальной аппроксимацией идентичны разработанным в разделе 3.3, поскольку он основан на той же аппроксимации профиля засветки и нормализации, что и метод FIDA. Соответственно, необходимо разработать НП для методов OFC FCA и 2MN FCA. Алгоритм генерации начальных приближений можно разбить на две части: для однокомпонентной модели, и для двух и более компонентной модели. В последнем случае будем использовать систему (3.32) при соответствующей модификации левых частей уравнений. Например, для OFC FCA с нормализацией на  $V_{eff}$  требуется решение следующей системы уравнений

$$\begin{aligned}
 2\sqrt{2}(1+F_1)(K_1 - \lambda T)/(1+F_2) &= (N_{eff\ C1}q_{C1} + N_{eff\ C2}q_{C2})T \\
 8(1+F_1)^2 K_2/(1+F_2)^2 &= (N_{eff\ C1}q_{C1}^2 + N_{eff\ C2}q_{C2}^2)T^2 \\
 6\sqrt{6}(1+F_1)^2 K_3/(1+F_2) &= (N_{eff\ C1}q_{C1}^3 + N_{eff\ C2}q_{C2}^3)T^3 \\
 8\sqrt{8}(1+F_1)^2 K_4/(1+F_2) &= (N_{eff\ C1}q_{C1}^4 + N_{eff\ C2}q_{C2}^4)T^4.
 \end{aligned} \tag{4.20}$$

Систему (4.20) можно свести к полиному третьего порядка относительно  $q_1$  (или  $q_2$ ) и решить как численно, так и аналитически (смотри приложение С). НП для однокомпонентной модели приведены в таблице 4.1 (в колонке слева приведены параметры, для которых заранее известны оценки). Разработанные НП справедливы также и для методов анализа РЧФ с соответствующей аппроксимацией профиля засветки и нормализацией.

Алгоритм генерации НП был протестирован по той же методике, что применялась для тестирования НП для метода FIDA (смотри подраздел 3.3.2). Тестирование НП вначале проводилось на смоделированных данных, затем на измеренных. Для незашумленных данных разница между НП и истинными значениями параметров была в пределах ошибки численных методов, используемых при генерации ФК РЧФ. Результаты тестирования на смоделированных и измеренных данных показали эффективность НП (числовые результаты в работе не представлены, поскольку они аналогичны результатам, приведенным в подразделе 3.3.2).

#### 4.4 Результаты тестирования кумулянтного анализа

**Тестирование на однокомпонентных смоделированных данных.** Для тестирования кумулянтного анализа на смоделированных данных были сгенерированы несколько серий РЧФ и ФК РЧФ (по 100 характеристик в каждой серии, отличающихся только реализацией шума) с различным отношением  $S/N$  (заданного для РЧФ). Генерация факториальных кумулянтов проводилась через промежуточную генерацию РЧФ с последующим расчетом ФК по формулам (3.17)

Таблица 4.1 – Начальные приближения для однокомпонентной модели

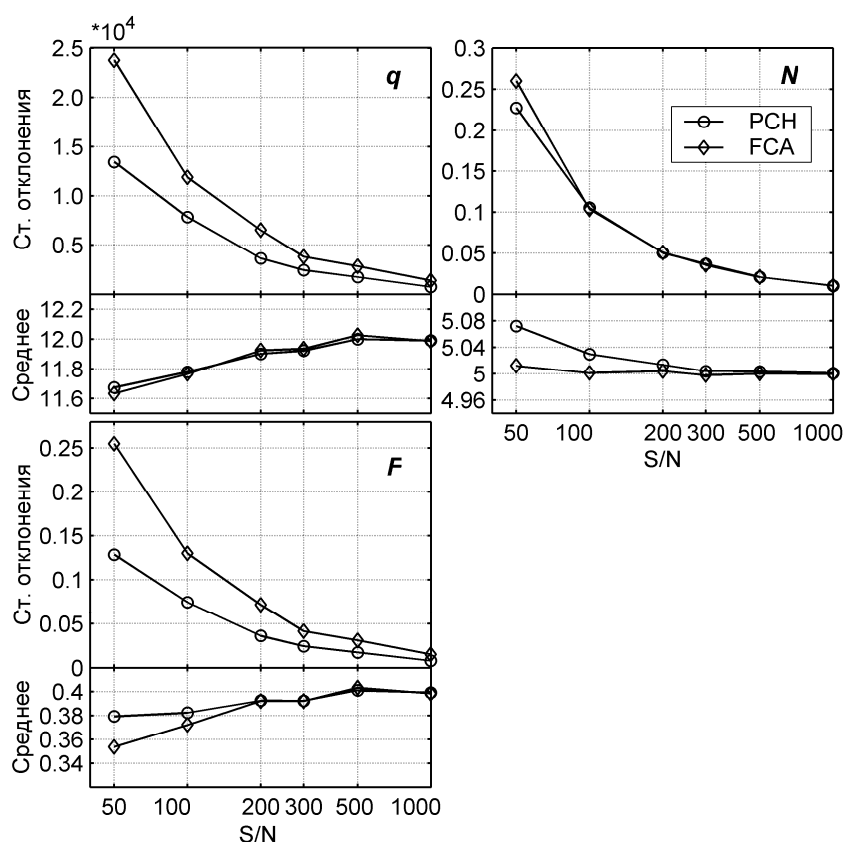
Параметр	OFC FCA (OFC PCH) с нормализацией к $V_{eff}$	2MN FCA (2NM PCH)
$\lambda$ фикс. $F_1$ фикс. $F_2$ фикс.	$N_{effC} = \frac{(K_1 - \lambda T)^2}{K_2}, q_C = \frac{2\sqrt{2}(1+F_1)K_2}{(1+F_2)(K_1 - \lambda T)T}$	$N_{2MN} = \frac{(K_1 - \lambda T)^2}{K_2}, q_{2MN} = \frac{K_2}{(K_1 - \lambda T)T}$
$F_1$ фикс. $F_2$ фикс.	$N_{effC} = \frac{64(F_1+1)^2 K_2^3}{27(F_2+1)^4 K_3^2}, q_C = \frac{3\sqrt{6}(1+F_2)K_3}{4K_2 T}$ $\lambda = \frac{K_1}{T} - \frac{8\sqrt{3}(F_1+1)K_2^2}{9(F_2+1)^2 K_3 T}$	$N_{2MN} = \frac{64(F_1+1)^2 K_2^3}{27(F_2+1)^4 K_3^2},$ $q_{2MN} = \frac{3\sqrt{3}(1+F_2)K_3}{8K_2 T}$ $\lambda = \frac{K_1}{T} - \frac{8\sqrt{3}(F_1+1)K_2^2}{9(F_2+1)^2 K_3 T}$
$\lambda$ фикс. $F_2$ фикс.	$N_{effC} = \frac{(K_1 - \lambda T)^2}{K_2}, q_C = \frac{3\sqrt{6}(1+F_2)K_3}{4K_2 T}$ $F_1 = \frac{3\sqrt{3}(F_2+1)^2 K_3 (K_1 - \lambda T)}{8K_2^2} - 1$	$N_{2MN} = \frac{(K_1 - \lambda T)^2}{K_2}, q_{2MN} = \frac{K_2}{(K_1 - \lambda T)T}$ $F_1 = \frac{3\sqrt{3}(F_2+1)^2 K_3 (K_1 - \lambda T)}{8K_2^2} - 1$
$\lambda$ фикс.	$N_{effC} = \frac{(K_1 - \lambda T)^2}{K_2}, q_C = \frac{8\sqrt{3}K_4}{9K_3 T},$ $F_1 = \frac{64\sqrt{3}K_4^2 (K_1 - \lambda T)}{243K_3^3} - 1,$ $F_2 = \frac{16\sqrt{2}K_2 K_4}{27K_3^2} - 1, F_2 \neq 1$	$N_{2MN} = \frac{(K_1 - \lambda T)^2}{K_2}, q_{2MN} = \frac{K_2}{(K_1 - \lambda T)T},$ $F_1 = \frac{64\sqrt{3}K_4^2 (K_1 - \lambda T)}{243K_3^3} - 1,$ $F_2 = \frac{16\sqrt{2}K_2 K_4}{27K_3^2} - 1, F_2 \neq 1$
$F_2$ фикс.	$N_{effC} = \frac{19683(F_1+1)^2 K_3^6}{4096K_2 K_4^4}, q_C = \frac{8\sqrt{3}K_4}{9K_3 T}$ $F_2 = \frac{16\sqrt{2}K_2 K_4}{27K_3^2} - 1, F_2 \neq 1$ $\lambda = \frac{K_1}{T} - \frac{81\sqrt{3}(F_1+1)K_3^3}{64K_4^2 T}$	$N_{2MN} = \frac{19683(F_1+1)^2 K_3^6}{4096K_2 K_4^4},$ $q_{2MN} = \frac{64\sqrt{3}K_2 K_4^2}{243(1+F_1)K_3^3 T}$ $F_2 = \frac{16\sqrt{2}K_2 K_4}{27K_3^2} - 1, F_2 \neq 1$ $\lambda = \frac{K_1}{T} - \frac{81\sqrt{3}(F_1+1)K_3^3}{64K_4^2 T}$
$F_1$ фикс.	нет решения	нет решения

(3.18). Данный подход позволяет моделировать РЧФ и ФК РЧФ одновременно и сравнивать результаты кумулянтного анализа с результатами хорошо зарекомендовавших себя методов анализа РЧФ (PCH, FIDA).



Каждый раз генерировалось РЧФ с заданным отношением  $S/N$ , затем вычислялись первые 5 кумулянтов и проводился анализ с помощью методов РСН/FIDA и FCA. Длительность интервала  $T$  принималась равной  $2E-5$  с. НП принимались равными истинным значениям. Полученные оценки затем комбинировались для вычисления среднего и стандартных отклонений. При тестировании методов РСН, OFC РСН, FCA и OFC FCA применялась нормализация к эффективному объему.

Результаты тестирования кумулянтного метода анализа с коррекцией профиля засветки первого порядка (OFCI FCA) (выражение 4.3) на однокомпонентных данных показаны на рисунке 4.1 (параметры моделирования:  $q_C = 120000$ ,  $N_{effC} = 5$ ,  $F = 0.4$ ). В целях сокращения записи индексы в обозначениях количества молекул и яркости на рисунке 4.1 и далее по разделу опущены.



**Рисунок 4.1 – Сравнение результатов анализа однокомпонентных данных методами РСН и FCA с коррекцией первого порядка**

Среднеквадратические отклонения и средние параметров  $q$ ,  $N$ ,  $F$  представлены в зависимости от отношения  $S/N$ . Зависимость ожидаемая: с увеличением отношения  $S/N$  среднеквадратические отклонения оцениваемых параметров уменьшаются и оценки стремятся к их истинным значениям. В проведенном тесте важным является сравнение результатов методов РСН и FCA с однотипной коррекцией профиля засветки. Как видно из рисунка, даже для небольших  $S/N$  результаты очень близки для обоих методов. Как и следовало

ожидать, метод РСН предоставляет большую точность, чем FCA, вследствие использования значительно большего количества экспериментальных точек (16 и выше для РСН против 5 для FCA). Вероятность получения неправдоподобно малых значений  $F$  (меньших истинного значения на 2–3 порядка) при малых  $S/N$  для метода FCA была в несколько раз больше, чем для метода РСН (11% для FCA против 2% для РСН при  $S/N = 50$ ). Существование подобных выбросов для оценок  $F$  проявляется в систематическом смещении оценок  $q$  (в сторону меньших значений), поскольку эти оценки находятся в прямой зависимости друг от друга.

Результаты тестирования кумулянтного метода анализа с коррекцией профиля засветки второго порядка (OFCII FCA) (выражение 4.4) на однокомпонентных данных показаны на рисунке 4.2 (параметры моделирования:  $q = 120000$ ,  $N = 5$ ,  $F_1 = 0.4$ ,  $F_2 = 0.025$ ). Точность и устойчивость метода РСН с коррекцией второго порядка были значительно выше, чем для FCA. Это является следствием неудовлетворительной точности оценивания параметра  $F_2$  в методе FCA. Параметры  $F_2$  и  $F_1$  входят в выражение (4.4) в форме  $(1+F_2)$  и  $(1+F_1)$ . Очевидно, что если  $F_2$  много меньше 1 (принято при моделировании, исходя из типичных значений анализа экспериментальных данных),  $(1+F_2)$  близко к 1 и практически не влияет на результат. Для доказательства было проведено моделирование ФК РЧФ со сравнимыми значениями  $F_1$  и  $F_2$ . Среднеквадратические отклонения этих параметров стали сравнимыми, и устойчивость анализа значительно возросла.

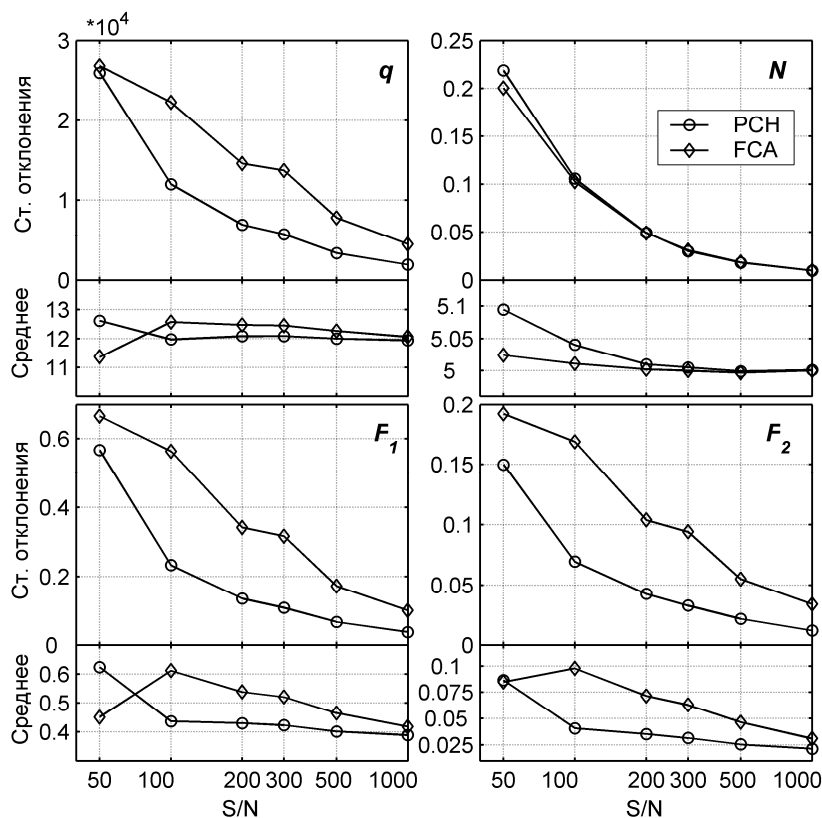
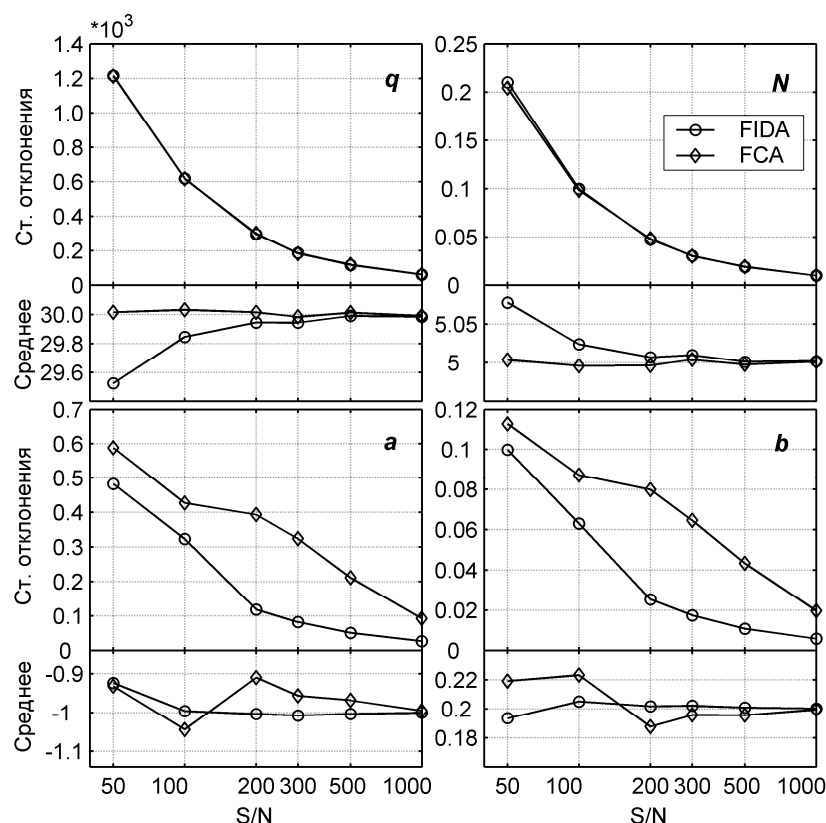


Рисунок 4.2 – Сравнение результатов анализа однокомпонентных данных методами РСН и FCA с коррекцией второго порядка

Результаты тестирования метода FCA с полиномиальной аппроксимацией (выражение (4.11)) представлены на рисунке 4.3 (параметры моделирования:  $q = 30000$ ,  $N = 5$ ,  $a = -1$ ,  $b = 0.2$ ). В сравнении с методами с коррекцией профиля засветки, метод FCA с полиномиальной аппроксимацией показывает лучшие результаты (меньшее смещение при малых  $S/N$ ) в оценке числа молекул и яркости, чем в методе FIDA.



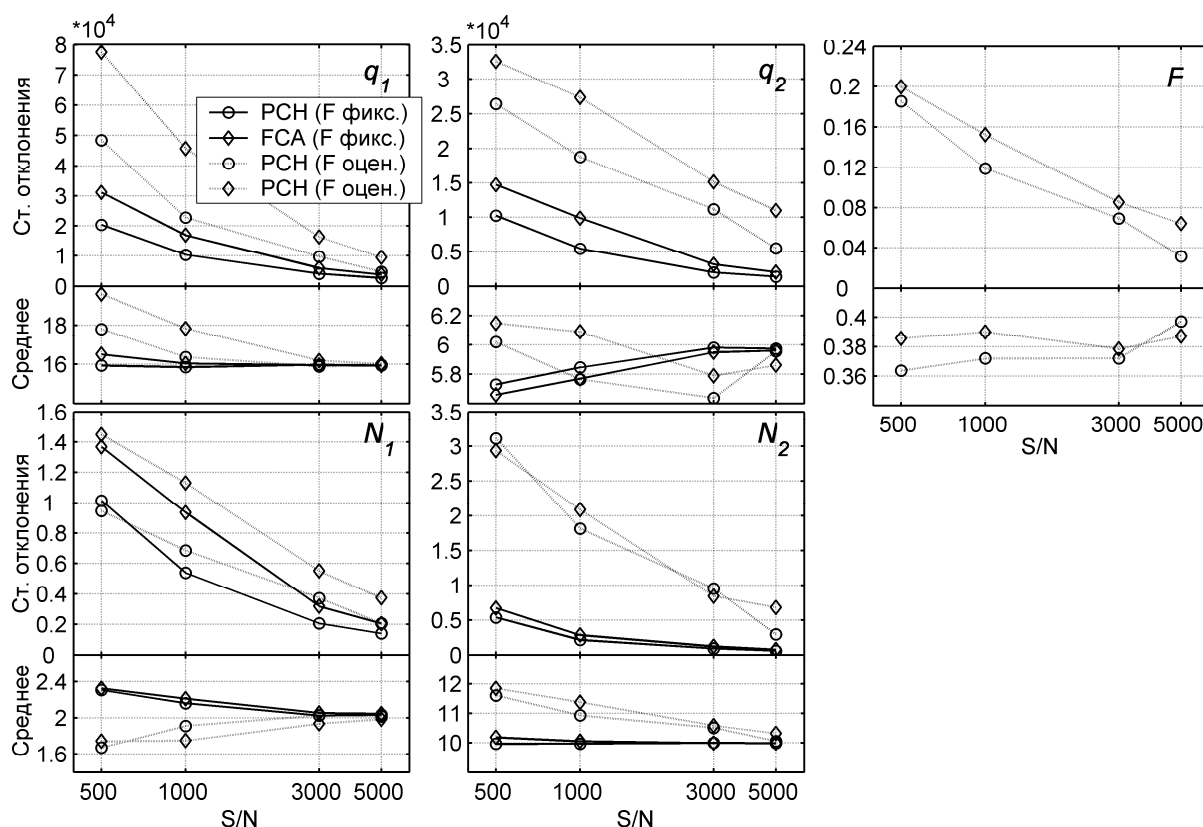
**Рисунок 4.3 – Сравнение результатов анализа однокомпонентных данных методами FIDA и FCA с полиномиальной аппроксимацией**

Среднеквадратические отклонения параметра  $N$  практически одинаковы для всех рассмотренных методов. Это подтверждает наше заключение о независимости  $N$  от типа аппроксимации профиля засветки и выбранной нормализации для однокомпонентной модели.

**Тестирование на двухкомпонентных смоделированных данных.** Тестирование НП для метода FIDA показало, что кумулянт 6 порядка обычно сильно зашумлен. Поэтому, несмотря на то, что для двухкомпонентной модели методы FIDA и FCA с полиномиальной и гауссовской аппроксимацией с коррекцией  $\Pi$  порядка имеют по 6 оцениваемых параметров, при тестировании использовались только 5 первых кумулянтов. Следовательно, оценивались только параметры  $N_i$ ,  $q_i$  при фиксировании остальных параметров в их истинные значения.

Результаты тестирования кумулянтного метода анализа с коррекцией профиля засветки первого порядка на двухкомпонентных данных показаны на

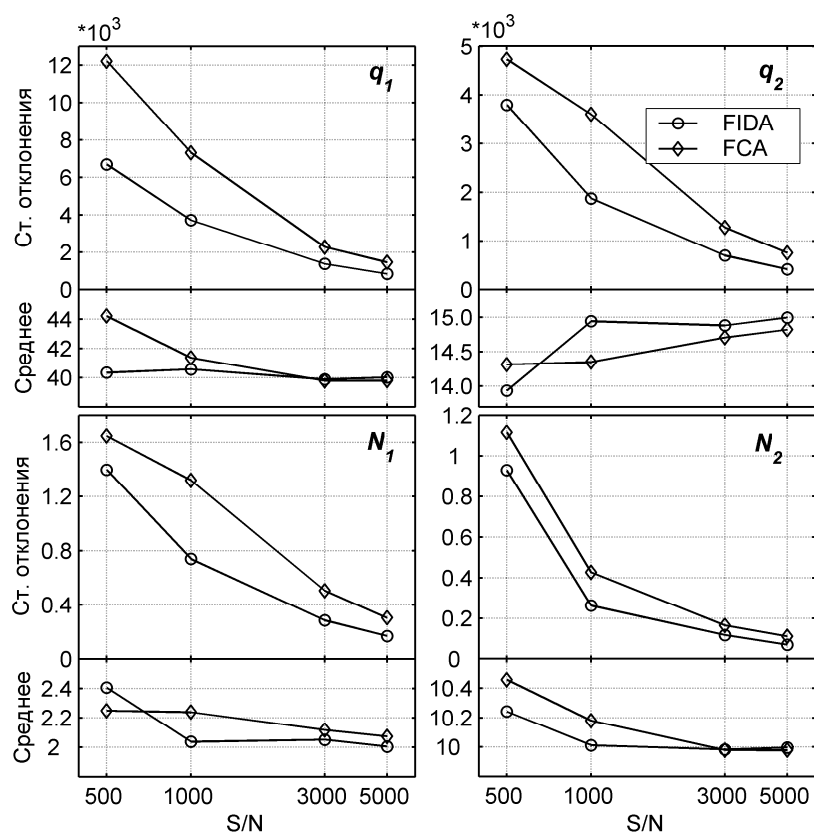
рисунке 4.4 (параметры моделирования:  $q_1 = 160000$ ,  $N_1 = 2$ ,  $q_2 = 60000$ ,  $N_2 = 10$ ,  $F = 0.4$ ). Как и для однокомпонентной модели, метод РСН показал лучшие результаты. Поскольку для коррекции I порядка можно оценивать 5 параметров, при тестировании использовались два сценария. В первом сценарии параметр  $F$  был зафиксирован в истинное значение (оценивались только  $N_1$ ,  $N_2$ ,  $q_1$ ,  $q_2$ ). Во втором, параметр  $F$  оценивался вместе с остальными параметрами. Наибольшие среднеквадратические отклонения и смещение оцениваемых параметров были получены при использовании второго сценария. Тем не менее, параметр коррекции  $F$  может быть оценен с достаточной точностью из одиночного набора кумулянтов по крайней мере при больших отношениях  $S/N$ .



**Рисунок 4.4 – Сравнение результатов анализа двухкомпонентных данных методами РСН и FCA с коррекцией первого порядка**

Результаты тестирования кумулянтного метода анализа с полиномиальной аппроксимацией на двухкомпонентных данных показаны на рисунке 4.5 (параметры моделирования:  $q_1 = 40000$ ,  $N_1 = 2$ ,  $q_2 = 15000$ ,  $N_2 = 10$ ,  $a = -1$ ,  $b = 0.2$ ). В целом, результаты подобны предыдущим для первого сценария. Если параметры, отвечающие за коррекцию профиля засветки, зафиксированы, методы FIDA/PCN и FCA становятся практически независимыми от выбранной аппроксимации профиля засветки.

Устойчивый анализ двухкомпонентных данных возможен при значительно больших отношениях  $S/N$ . В нашем случае число неудачных попыток получить



**Рисунок 4.5 – Сравнение результатов анализа двухкомпонентных данных методами FIDA и FCA с полиномиальной аппроксимацией**

приемлемые оценки параметров, т.е. физически приемлемые и без явных выбросов по крайней мере одного параметра, изменялось от нескольких процентов для больших  $S/N$  ( $S/N = 1000$  и выше) до более чем половины при небольших  $S/N$  ( $S/N = 300$  и менее).

Метод FCA предоставляет меньшую точность и устойчивость в сравнении с методами PCH/FIDA. Но это становится заметным только для небольших отношений  $S/N$ . При типичных для ФФС измерений (около 1000 для РЧФ) результаты становятся сравнимыми. В то же время метод FCA демонстрирует значительно более высокую производительность анализа данных. Время вычисления теоретической модели в методе FCA в среднем на 3 порядка меньше времени вычисления модели в методе PCH. Приемлемый диапазон изменения параметров также намного шире для метода FCA, чем для PCH/FIDA, что делает FCA менее чувствительным к качеству начальных приближений.

**Тестирование на измеренных данных.** Для тестирования предложенных методов кумулянтного анализа на измеренных данных было проведено измерение красителя Alexa 488 и смеси Alexa 488 и Флуоресцеина. Экспериментальные данные были предоставлены кафедрой паразитологии Медицинского центра Лейденского университета (Department of Parasitology, LUMC, the Netherlands). Все измерения были выполнены с помощью СПОД ConfoCor 2 (Carl Zeiss, Germany).

Измеряемый образец растворялся в PBS (0.035M phosphate, 0.15M NaCl, pH 7.6). Краситель Alexa 488 и Флуоресцеин был приобретен у Molecular Probes (Molecular Probes Europe BV; The Netherlands). Для однокомпонентных данных при времени измерения 30 с число зарегистрированных фотонов составило  $8.6E+6$ . Интервал наблюдения  $T$  был выбран равным  $8E-6$  с, что соответствует отношению  $S/N = 1100$ . Для двухкомпонентных данных время измерения составило 120 с и число зарегистрированных фотонов около  $16E+6$ . Интервал наблюдения  $T$  был выбран равным  $1E-5$  с, что соответствует  $S/N = 2400$ . Фоновый сигнал был оценен из дополнительного измерения чистого растворителя.

Анализ данных методами РСН и FCA показал, что для получения достоверных результатов достаточно применения коррекции первого порядка. Для проверки применялась также коррекция второго порядка. Результаты анализа практически не изменились, и значение параметра  $F_2$  составило  $1E-7$ .

Результаты анализа Alexa 488 приведены в таблице 4.2 и показаны на рисунке 4.6. Представлены результаты кумулянтного анализа, а также результаты методов FCA, FIDA и РСН для сравнения. В таблице 4.2 колонка  $F/a$  содержит значения или параметра  $F$  (для коррекции профиля засветки), или параметра  $a$  (для полиномиальной аппроксимации). Параметр  $\lambda$  зафиксирован в значение 1000. Доверительные интервалы находились по методу исчерпывающего поиска (exhausted search) [116, 9-A] при доверительной вероятности 0.67. Яркость, пересчитанная к истинным значениям (соответственно (4.15), (4.18)), показана в скобках.

Таблица 4.2 – Результаты анализа Alexa 488

Метод	$\chi^2$	Параметр			
		$N$	$q*10^4$	$F/a$	$b$
PCN	12.0	12.44 [12.09; 12.85]	6.52 [6.31; 6.73]	–	–
FCA	48.8	12.18 [11.16; 13.37]	6.66 [6.07; 7.24]	–	–
OFCI PCN	1.39	12.29 [12.15; 12.44]	10.48 [9.53; 11.23] (6.60)	0.57 [0.45; 0.70]	–
OFCI FCA	0.59	12.28 [12.05; 12.56]	10.21 [8.74; 11.66] (6.61)	0.55 [0.32; 0.77]	–
FIDA	1.37	12.29 [12.12; 12.44]	2.33 [2.30; 2.36] (6.59)	-1.13 [-1.27; -1.06]	0.24 [0.16; 0.27]
POLY FCA	0.38	12.28 [12.05; 12.50]	2.34 [2.29; 2.38] (6.62)	-1.05 [-1.32; -0.92]	0.23 [0.08; 0.26]
2MN FCA	0.59	12.29 [12.05; 12.56]	2.34 [2.31; 2.36] (6.62)	0.55 [0.32; 0.77]	

Соответственно, таблица 4.3 и рисунок 4.7 отображают результаты анализа смеси Alexa 488 и Флуоресцеина. Как видно из представленных таблиц и рисунков, применение коррекции значительно улучшает качество анализа

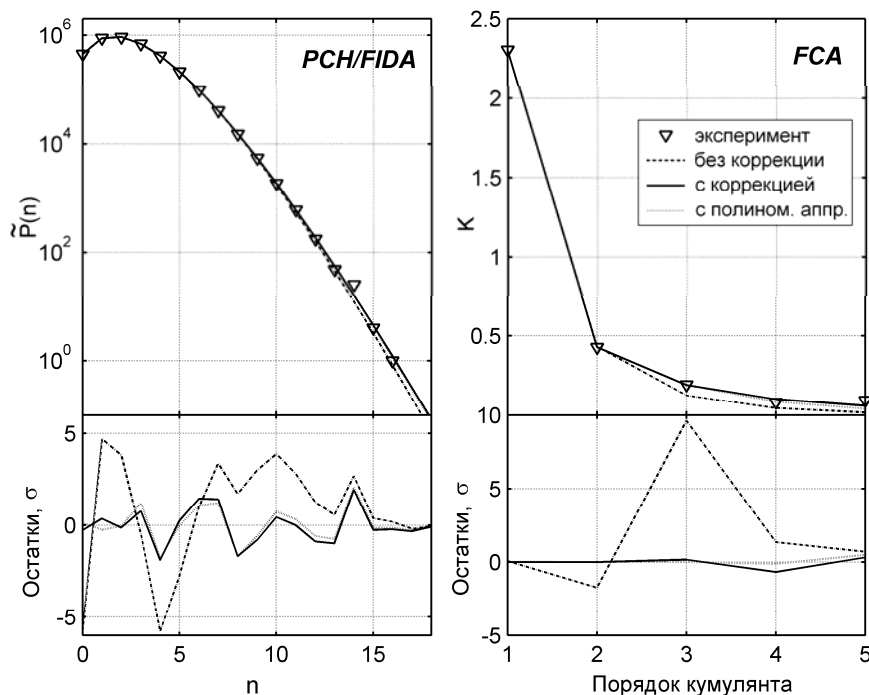
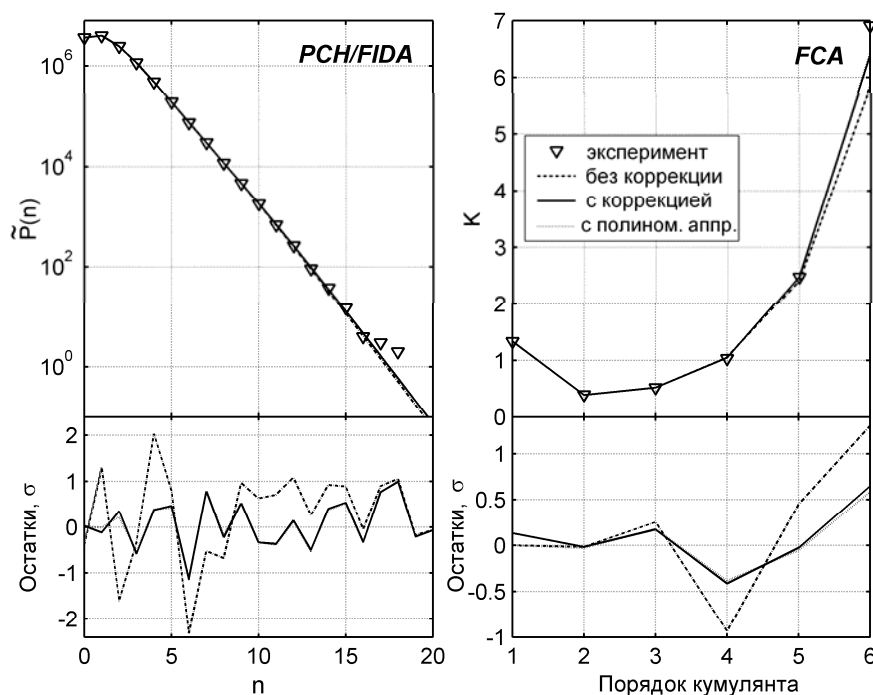


Рисунок 4.6 – Результаты анализа Alexa 488

измеренных данных. Результаты кумулянтного анализа с предложенной коррекцией профиля засветки практически совпадают с результатами методов FIDA/PCH, что подтверждает применимость FCA для анализа измеренных данных и эквивалентность рассмотренных методов.

Таблица 4.3 – Результаты анализа смеси Alexa 488 и Флуоресцеина (прочерк в доверительных интервалах означает отклонение более чем на 100%)

Метод	$\chi^2$	Параметр					
		$N_1$	$q_1 \cdot 10^4$	$N_2$	$q_2 \cdot 10^4$	$F/a$	$b$
PCH	1.58	12.7 [11.7; 13.8]	2.31 [2.07; 2.53]	0.28 [0.25; 0.31]	30.1 [29.0; 31.2]	–	–
FCA	2.82	10.9 [7.5; –]	2.76 [0.28; 4.40]	0.23 [0.12; –]	32.0 [25.6; 38.8]	–	–
OFCI PCH	0.44	6.10 [5.38; 7.64]	7.87 [5.40; 10.2] (4.26)	0.64 [0.56; 0.71]	33.2 [31.5; 35.2] (17.97)	0.85 [0.58; 1.02]	–
OFCI FCA	0.60	5.99 [5.46; 7.38]	8.27 [5.51; 10.2] (4.48)	0.58 [0.34; 0.96]	34.1 [29.7; 39.0] (18.45)	0.85 (фикс.)	–
FIDA	0.47	5.43 [2.71; 7.65]	1.67 [0.86; 2.73] (4.72)	0.71 [0.34; 1.12]	5.85 [2.89; 9.29] (16.55)	-0.89 [-1.21; –]	0.25 [0.19; –]
POLY FCA	0.51	5.39 [5.09; 6.19]	1.75 [1.21; 2.12] (4.95)	0.63 [0.34; 1.11]	6.04 [5.25; 6.97] (17.08)	-0.89 (фикс.)	0.25 (фикс.)
2MN FCA	0.60	5.99 [5.02; 6.17]	1.74 [1.2; 2.23] (4.92)	0.6 [0.35; 1.10]	6.02 [5.21; 6.72] (17.03)	0.85 (фикс.)	–



**Рисунок 4.7 – Результаты анализа смеси Alexa 488 и Флуоресцеина**

В разделе 4.2 были выведены формулы пересчета яркости и количества молекул от одной аппроксимации профиля засветки и выбранной нормализации к другой. Справедливость формул (4.15), (4.17), (4.18), (4.19) была проверена на смоделированных данных и измеренных данных (смотри значения в скобках в таблице 4.2 и таблице 4.3). Пересчет яркости производился к его истинному значению, не зависящему от параметров коррекции (пересчет количества молекул не требовался, так как ранее было показано, что оценки количества молекул в случае нормализации вида (1.84) и нормализации к эффективному объему совпадают). Как видно из таблицы 4.2 и таблицы 4.3, оценки яркости могут быть пересчитаны к одним значениям и, следовательно, имеется возможность практического сравнения результатов, полученных разными методами.

Метод кумулянтного анализа с предложенной коррекцией профиля засветки и НП для него были внедрены в программу FCS Data Processor [17-A].

#### **4.5 Основные результаты и выводы**

1. Получены аналитические выражения для факториальных кумулянтов распределения числа фотоотсчетов с коррекцией профиля засветки. Применение полученных выражений позволяет значительно улучшить качество анализа экспериментальных данных, проводить анализ одиночного набора кумулянтов без постановки калибровочного эксперимента и устранить зависимость числа оцениваемых параметров от количества анализируемых кумулянтов.

2. Показано, что оценки яркости и количества молекул в методах FCA и PCH с коррекцией профиля засветки зависят от параметров коррекции, что не



позволяет проводить глобальный анализ со связыванием зависимых параметров, а также сравнивать результаты нескольких экспериментов. Получены и проверены на смоделированных и измеренных данных формулы пересчета яркости и количества молекул для различных аппроксимаций профиля засветки и нормализаций. Найдены аналитические выражения для факториальных кумулянтов распределения числа фотоотсчетов с коррекцией профиля засветки, характеризующиеся отсутствием зависимости между параметрами и позволяющие проведение глобального анализа со связыванием всех параметров.

3. Разработан алгоритм генерации начальных приближений для кумулянтного анализа. Разработанный алгоритм позволяет также генерировать начальные приближения для анализа распределения числа фотоотсчетов с однотипной аппроксимацией профиля засветки.

4. Метод кумулянтного анализа с предложенной коррекцией профиля засветки был программно реализован и включен в состав программы FCS Data Processor.

## ЗАКЛЮЧЕНИЕ

### Основные научные результаты диссертации

1. Разработаны аналитические и имитационные модели систем преобразования потоков случайных событий, включающие модели последовательного преобразования, слияния и разветвления случайных потоков, а также модели потоков фотонов с заданной автокорреляционной функцией и распределением числа фотоотсчетов. На основе полученных моделей создан программный комплекс имитационного моделирования систем регистрации и обработки данных ФФС, позволяющий создавать сложные модели без необходимости изменения программного кода и последующей компиляции программы. Использование комплекса позволило оценить эффективность разработанных методов анализа данных и влияние инструментальных искажений системы регистрации на анализируемые характеристики потоков фотоотсчетов. [1-А, 2-А, 3-А, 4-А, 5-А, 6-А, 7-А, 8-А, 9-А, 10-А, 18-А]

2. Получено выражение для производящей функции числа фотоотсчетов с коррекцией профиля засветки. Показано, что известные методы анализа распределения числа фотоотсчетов могут быть выведены с использованием полученной производящей функции. Разработан метод анализа распределения числа фотоотсчетов, характеризующийся отсутствием зависимости яркости и количества молекул в объеме наблюдения от параметров коррекции профиля засветки, а также значительно большей скоростью вычислений в сравнении с ранее предложенными методами. [17-А, 20-А]

3. Предложен эффективный метод генерации начальных приближений для анализа распределения числа фотоотсчетов и кумулянтного анализа, позволяющий находить оценки яркости и количества молекул одно- и двухкомпонентных систем, а также параметров коррекции профиля засветки. Показано, что его применение позволяет существенно увеличить скорость схождения к глобальному минимуму, а также снизить вероятность получения вычислительных ошибок. Предложенный метод может быть также использован как самостоятельный метод быстрой оценки параметров исследуемого вещества. [9-А, 18-А, 19-А, 22-А]

4. Получены аналитические выражения для факториальных кумулянтов распределения числа фотоотсчетов с коррекцией профиля засветки. Применение полученных выражений позволяет проводить анализ одиночного набора кумулянтов без постановки калибровочного эксперимента и устранить зависимость числа оцениваемых параметров от количества анализируемых кумулянтов. Получены формулы пересчета яркости и количества молекул в объеме наблюдения при использовании различных аппроксимаций профиля засветки и выбранной нормализации. [17-А, 18-А, 21-А]

5. Разработаны базы данных измерений и результатов анализа во флуоресцентной спектроскопии, а также алгоритмы предварительной обработки данных, включающие расчет авто(кросс)корреляционной функции, распределения числа фотоотсчетов и факториальных кумулянтов распределения числа фотоотсчетов с возможностью режекции выбросов интенсивности. Разработанные базы данных вошли в состав трех программных комплексов: FCS Data Processor, TRFA Data Processor и TRFD Fitting Software. [9-А, 10-А, 11-А, 12-А, 13-А, 14-А, 15-А, 16-А, 23-А]

### **Рекомендации по практическому использованию результатов**

Разработанный программный комплекс имитационного моделирования системы преобразования потоков случайных событий, обладая высокой степенью общности, обеспечивает возможность моделирования широкого круга систем регистрации потоков фотонов во флуоресцентной спектроскопии. Комплекс может быть использован также и как основа лабораторного практикума.

Применение разработанных моделей и методов анализа распределения числа фотоотсчетов и факториальных кумулянтов распределения числа фотоотсчетов позволит получать более достоверные оценки параметров исследуемого вещества вследствие введения коррекции профиля засветки. Устранение зависимости между параметрами модели в случае коррекции профиля засветки позволит повысить точность определения количества молекул и характеристической яркости, а также проводить глобальный анализ различных характеристик со связыванием этих параметров. Применение начальных приближений позволит увеличить производительность и устойчивость анализа данных при минимизации усилий исследователя (экспериментатора).

Применение разработанной на основе технологии реляционных баз данных системы хранения и обработки данных позволит поднять производительность анализа данных ФФС. Разработанные базы данных, модели и методы анализа данных ФФС реализованы в программном комплексе FCS Data Processor. В настоящее время FCS Data Processor используется в мире для исследования сложных биологических образцов рядом научных лабораторий [105 – 115]. В частности, в [115] FCS Data Processor использовался для проведения кумулянтного анализа структуры GFP - CLIP-170, в [112] – для анализа распределения числа фотоотсчетов флуоресцентно помеченных рецепторов BRI1 и AtSERK1(3), в [105 – 114] – для анализа процессов жизнедеятельности клетки по методу FCS.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

### Список использованных источников

1. Elson, E.L. Fluorescence correlation spectroscopy. I. Conceptual basis and theory / E.L. Elson, D. Magde // *Biopolymers*. – 1974. – № 1. – P. 113.
2. Magde, D. Thermodynamic fluctuations in a reacting system: Measurement by fluorescence correlation spectroscopy / D. Magde, E.L. Elson and W.W. Webb // *Phys. Rev. Lett.* – 1972. – № 29. – P. 705.
3. Ehrenberg, M. Rotational Brownian motion and fluorescence intensity fluctuations / M. Ehrenberg, R. Rigler // *Chem. Phys.* – 1974. – № 4. – P. 390–401.
4. Апанасович, В.В. Статистический анализ случайных потоков в физическом эксперименте / В.В. Апанасович, А.А. Коляда, А.Ф. Чернявский. – Минск: Университетское, 1988. – 256 с.
5. Большаков, И.А. Прикладная теория случайных потоков / И.А. Большаков, В.С. Ракощиц. – Москва: Сов. Радио, 1978. – 246 с.
6. Daley, D.J. An Introduction to the Theory of Point Processes / D.J. Daley, D. Vere-Jones. – New York: Springer, 1988. – 386 p.
7. Saleh, B. Photoelectron statistics. With applications to spectroscopy and optical communication / B. Saleh. – Berlin: Springer Verlag, 1978. – 431 p.
8. Апанасович, В.В. Цифровое моделирование измерительно-вычислительных систем анализа случайных потоков / В.В. Апанасович, Е.Г. Новиков // *Управляющие системы и машины*. – 1990. – № 1. – С. 71–74.
9. Апанасович, В.В. Цифровое моделирование стохастических систем / В.В. Апанасович, О.М. Тихоненко. – Минск: Университетское, 1986. – 127 с.
10. Бусленко, Н.П. Моделирование сложных систем / Н.П. Бусленко. – Москва: Наука, 1978. – 400 с.
11. Быков, В.В. Цифровое моделирование в статистической радиотехнике / В.В. Быков. – Москва: Сов. Радио, 1971. – 326 с.
12. Ермаков, С.М. Курс статистического моделирования / С.М. Ермаков, Г.А. Михайлов. – Москва: Наука, 1976. – 320 с.
13. Поляк, Ю.Г. Вероятностное моделирование на электронных вычислительных машинах / Ю.Г. Поляк. – Москва: Сов. Радио, 1971. – 400 с.
14. Pramanik, A. Fluorescence Correlation Spectroscopy (FCS) / A. Pramanik, J. Widengren // *Encyclopedia of Molecular Cell Biology and Molecular Medicine*; Ed. Meyers. – Berlin: Wiley-VCH Verlag, 2004. – P. 461-500
15. Татаркова, С.А. Измерение диффузии флуоресцентных молекул в живых клетках / С.А. Татаркова, К. Ллойд, Д. Берг // *Квантовая электроника*. – 2003. – № 33(4). – С. 357-361.

- 
16. Nagy, A. Observation volumes and  $\gamma$ -factors in two-photon fluorescence fluctuation spectroscopy / A. Nagy, J. Wu, K.M. Berland // *Biophys. J.* – 2005. - №89(3). – P. 2077-2090.
  17. Molecular Dynamics in Living Cells Observed by Fluorescence Correlation Spectroscopy with One- and Two-Photon Excitation / P. Schwille [et al.] // *Biophys. J.* – 1999. - № 77. – P. 2251–2265.
  18. Koppel, D.E. Statistical accuracy in fluorescence correlation spectroscopy / D.E. Koppel // *Phys.Rev. A.* – 1974. – № 10. – P. 1938–1945.
  19. Fluorescence Fluctuation Spectroscopy in Subdiffraction Focal Volumes / L. Kastrup [et al.] // *Phys. Rev. Letters.* – 2005. - № 94. - P. 178104.1–4.
  20. Magde, D. Fluorescence correlation spectroscopy. II. An experimental realization / D. Magde, E.L. Elson, W.W. Webb // *Biopolymers.* – 1974. - № 13. - P. 29–61.
  21. Векшин, Н.Л. Флуоресцентная спектроскопия биополимеров / Н.Л. Векшин. – Пушино: Фотон-век, 2006. – 168 с.
  22. Eigen, M. Sorting single molecules: application to diagnostics and evolutionary biotechnology / M. Eigen, R. Rigler // *Proc. Natl. Acad. Sci. USA.* – 1994. - № 91. – P. 5740–5747.
  23. Fluorescence correlation spectroscopy with high count rates and low background, analysis of translational diffusion / R. Rigler [et al.] // *Eur. Biophys. J.* – 1993. - № 22. - P. 169.
  24. Гулаков, И.П. Метод счета фотонов в оптико-физических измерениях / И.П. Гулаков, С.В. Холондырев. – Минск: Университетское, 1989. – 256 с.
  25. Демчук, М.И. Статистический одноквантовый метод в оптико-физическом эксперименте / М.И. Демчук, М.А. Иванов - Мн.: Из-во БГУ, 1981. - 176 с.
  26. Photobleaching, Mobility, and Compartmentalisation: Inferences in Fluorescence Correlation Spectroscopy / A. Delon [et al.] // *J. of Fluorescence.* - 2004. - Vol. 14, №. 3. - P. 255-267.
  27. The photon counting histogram in fluorescence fluctuation spectroscopy / Y. Chen [et al.] // *Biophys. J.* – 1999. - № 77. - P. 553 - 567.
  28. Fluorescence–intensity distribution analysis and its application in biomolecular detection technology / P. Kask [et al.] // *Proc. Natl. Acad. Sci. USA.* – 1999. - Vol. 96, №. 24. - P. 13756 - 13761.
  29. Muller, J.D. Cumulant analysis in fluctuation spectroscopy / J.D. Muller // *Biophys. J.* – 2004. - № 86. P. 3981–3992.
  30. ConfoCor 2. Application Manual. - Jena: Carl Zeiss AG, 2001. - 142 p.
  31. ALBA. Fluorescence Correlation Spectrometer. Software Manual. - Illinois: ISS Inc, 2001. – 95 p.

- 
32. PicoHarp 300. Time-Correlated Single Photon Counting System with USB Interface. User's Manual and Technical Data. - Berlin: PicoQuant, 2006. – 67 p.
33. Becker, W. The BH TCSPC Handbook. / W. Becker. – Berlin: Becker & Hickl, 2005. – 378 p.
34. Novikov, E.G. Analytical model of the fluorescence fluctuation spectroscopy experiment / E.G. Novikov // J. of Chem. Phys. – 2001. - Vol. 114, No. 4. - P. 1745–1753.
35. Корн, Г. Справочник по математике для научных работников и инженеров. Определения, теоремы, формулы / Г. Корн, Т. Корн. Перевод с англ. под общей редакцией И.Г. Абрамовича. – Москва: Наука, 1970. – 720 с.
36. Fluorescence Intensity Distribution Analysis (FIDA) and related fluorescence fluctuation techniques: theory and practice / P. Kask [et al.] // Fluorescent Spectroscopy, Imaging and Probes. New Tools in Chemical, Physical and Life Sciences. Eds: R. Kraayenhof, A.J.W.G. Visser, H.C. Gerritsen. – Berlin, Heidelberg: Springer-Verlag, 2002. – P. 153–181.
37. Grandell, J. Double stochastic point processes/ Grandell J. - Berlin: Springer, 1978. - 234 p.
38. Ватутин, В.А. Вероятностные методы в физических исследованиях / В.А. Ватутин, Т.М. Телевинова, В.П.Чистяков. - Москва: Наука, 1985. - 207 с.
39. Курочкин, С.С. Многомерные статистические анализаторы / С.С. Курочкин. - Москва: Атомиздат, 1968. - 446 с.
40. Гольданский, В.И. Статистика отсчетов при регистрации ядерных частиц / В.И. Гольданский, А.В. Куценко, М.И. Подгорецкий. - Москва: ГИФМЛ, 1959. - 411 с.
41. Демчук, М.И. Импульсная спектрометрия / М.И. Демчук, М.А. Иванов. - Минск: Университетское, 1986. - 206 с.
42. Электронные методы ядерной физики / Л.А. Маталин [и др.]. – Москва : Атомиздат, 1973. - 520 с.
43. Рехин, Е.И. Измерение интервалов времени в экспериментальной ядерной физике / Е.И. Рехин, А.А. Курашов, П.С. Чернов. - Москва: Атомиздат, 1967. – 382 с.
44. Чернявский, А.Ф. Статистические методы анализа случайных сигналов в ядерно–физическом эксперименте / А.Ф. Чернявский, С.В. Бекетов, А.В. Потапов. – Москва : Атомиздат, 1974. - 352 с.
45. Faraci, G. Experimental dead–time distortions of Poisson processes / G. Faraci, A.R. Pennisi // Nucl. Instrum. and Meth. - 1983. - No. 212. - P. 307–310.
46. Mandel L. Inversion problem in photon counting with dead time / L. Mandel // J. Opt. Soc. Am. - 1980. - Vol. 70, № 7. - P. 873–874.

- 
47. Muller, J.W. Dead-time problems / J.W. Muller // Nuclear. Instr. and Meth. - 1973. - № 112. - P. 45–47.
48. Muller, J.W. Some formulae for a dead-time-distorted Poisson process / J.W. Muller // Nuclear. Instr. and Meth. - 1974. - №. 117. - P. 401–404.
49. Picot, A. Some remarks on the determination of dead-time by the two sources methods / A. Picot // Rev. Sci. Instrum. - 1976. - Vol. 47, № 3. - P. 385–386.
50. Vincent, C.H. Type II dead-time losses between detections occurring at random times within a gate-open time / C.H. Vincent // Nuclear Instruments and Methods in Physical Research. - 1988. - № 269. - P. 335–337.
51. De Lotto, I. Time distribution of events at the output of the system with variable dead-time / I. De Lotto, P.F. Manfredi // Energia nucleare. - 1965. - № 2. - P. 95–97.
52. Апанасович, В.В. Влияние мертвого времени регистрации при статистическом методе исследования затухания люминесценции / В.В. Апанасович, А.Ф. Чернявский // Журн. прикл. спектр. - 1983. -Т. 39, № 5. - С. 853–855.
53. Курочкин, С.С. К оценке просчетов цифровых измерительных систем непуассоновских входящих потоков импульсов / С.С. Курочкин // Приборы и техника эксперимента. - 1965. - № 5. - С. 100–104.
54. Cantor, B. Photocounting distribution with variable dead-time / B. Cantor, L. Matine, M.C. Teich // Applied optics. -1975. - № 12. - P. 2819.
55. De Lotto, I. Counting statistics and dead-time losses. Part 1 / I. De Lotto, P.F. Manfredi, P.Principi // Energia nucleare. - 1964. - № 10. - P. 557–564.
56. De Lotto, I. Statistical behavior of system with variable dead-time / I. De Lotto, P.F. Manfredi // Nucl. Instrum and Meth. - 1965. - V. 32. - P. 175–176.
57. Sabol, J. Approximation of the nonparalyzable dead time by the shortest time interval / J. Sabol // Rev. Sci. Instrum. - 1988. - Vol. 59, № 9. - P. 2086–2087.
58. Schatzel K. Dead time correction of photon correlations functions / K. Schatzel // Applied Physics. - 1986. - № 41. - P. 95–102.
59. Vannucci, G. Equivalent of threshold detections with and without dead-time / G. Vannucci, M.C. Teich // Applied Optics. - 1979. - V. 18, № 23. - P. 3886–3887.
60. Cox, D.R. Point processes / D.R. Cox, V. Isham. - London: Charman and Hall, 1980. - 188 p.
61. Albert, G. E. Contribution to the statistical theory of counter data/ G. E. Albert, L. Nelson // Ann. Math Stat. – 1953. - № 24. - P. 9–22.
62. Muller, J.W. A simple derivation of the Takacs formulae / J. W. Muller // Rapport BIPM–88 [Electronic resource]. – 1988. – Mode of access: [www.bipm.org/utis/common/pdf/rapportBIPM/1988/03.pdf](http://www.bipm.org/utis/common/pdf/rapportBIPM/1988/03.pdf). - Date of access: 15.10.2007.

- 
63. Muller, J. W. Generalized dead times / J. W. Muller // *Nuclear Instr. and Methods.* – 1991. - № A301. - P. 543–551.
64. Correction of detection system dead time effects on photon correlation functions / K. Schatzel [et al.] // *J. Opt. Soc. Am.* – 1989. - Vol 6, No 5. - P. 937–947.
65. Takacs, L. On a probability problem in the theory of counters / L. Takacs // *Ann. Math. Stat.* – 1958. – № 29. - P. 1257–1263.
66. Апанасович, В. В. Computation of extended dead time of photon counters in recording unsteady luminous fluxes / В.В. Апанасович, I.R. Gulakov, E.E. Prolisko // *Journal of Applied Spectroscopy.* – 1987. –Vol. 46, № 2. - P. 208-211.
67. Ветохин, С.С. Последействие фотоумножителей / С.С. Ветохин, И.Р. Гулаков, Ю.В. Писляк // *Приборы и техника эксперимента.* - 1983. - № 4. - С. 173 – 175 .
68. Топе, С. Исследование послеимпульсов в фотоумножителях / С. Топе // *Приборы для научных исследований.* - 1983. - №4. - С. 173 – 175 .
69. Wiza, J.L. Microchannel plate detectors / J.L. Wiza // *Nuclear Instruments and Methods in Physical Research.* – 1979. - No. 162. - P.587 – 601.
70. Апанасович, В.В. Влияние послеимпульсов детектора на характеристики регистрируемого потока событий / В.В. Апанасович // *Радиотехника и электроника.* – 1991. -Т. 36, № 12. - С. 2409. – 2412.
71. Aragón, S. R. Fluorescence correlation spectroscopy and Brownian rotational diffusion / S.R. Aragón, R. Pecora // *Biopolymers.* – 1975. - № 14. – P. 119–138.
72. Aragón, S.R. Fluorescence correlation spectroscopy as a probe of molecular dynamics / S.R. Aragón, R. Pecora // *J Chem Phys.* – 1976. - № 64. – P. 1791–1803.
73. Schwille, P. Kinetic investigations by fluorescence correlation spectroscopy: the analytical and diagnostic potential of diffusion studies / P. Schwille, J. Bieschke, F. Oehlenschläger // *Biophys Chem.* - 1997. - № 66. – P. 211-228.
74. Krichevsky, O. Fluorescence correlation spectroscopy: the technique and its applications. / O. Krichevsky, G. Bonnet // *Rep. Prog. Phys.* – 2002. - № 65. – P. 251–297.
75. Müller, J.D. Fluorescence Correlation spectroscopy / J.D. Müller, Y. Chen, E. Gratton // *Methods Enzymol.* – 2003. - № 361. – P. 69–92.
76. Palmer, A.G. Molecular aggregation characterized by high order autocorrelation in fluorescence correlation spectroscopy / A.G. Palmer, N.L. Thompson // *Biophys J.* – 1987. - № 52. – P. 257–270.
77. Widengren, J. Fluorescence correlation spectroscopy of triplet states in solution: A theoretical and experimental study / J. Widengren, Ü. Mets, R. Rigler // *J. Phys. Chem. B.* – 1995. - № 99. – P. 13368.



- 
78. Widengren, J. Fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces / J. Widengren, R. Rigler // *Cell. Mol. Biol.* – 1998. - № 44. – P. 857.
79. Thompson, N.L. Fluorescence correlation spectroscopy. / N.L. Thompson // *Topics in fluorescence spectroscopy* / Editor J.R. Lakowicz. - New York: Plenum Press, 1991. – P. 337–378.
80. Biological and chemical applications of fluorescence correlation spectroscopy. A review / S. T. Hess [et al.] // *Biochemistry.* – 2002. - № 41. – P. 697–705.
81. Bacia, K. A dynamic view of cellular processes by in vivo fluorescence auto- and cross-correlation spectroscopy / K. Bacia, P. Schwille // *Methods.* – 2003. - № 29. – P. 74–85.
82. Schatzel, K. Photon correlation measurements at large lag times: improving statistical accuracy. / K. Schatzel, M. Drewel, S. Stimac // *J. Mod. Opt.* - 1988. - № 35. – P. 711–718.
83. Schatzel, K. Noise on multiple-tau photon correlation data / K. Schatzel // *Photon Correlation Spectroscopy: Multicomponent System: proc. of the SPIE conference, Los Angeles, January 22, 1991* / SPIE; ed.: K. S. Schmitz. - Los Angeles, 1991. - № 1430. – P. 109–115.
84. Wohland, T. The standard deviation in fluorescence correlation spectroscopy / T. Wohland, R. Rigler, H. Vogel // *Biophys J.* – 2001. - № 80. – P. 2987–2999.
85. Bevington, P.R. *Data Reduction and Error Analysis. For the Physical Sciences* / P.R. Bevington, D.K. Robinson. - Third Edition.– New York: McGraw-Hill, 2003. – 320 p.
86. Resolution of fluorescence correlation measurements / U. Meseth [et al.] // *Biophys. J.* – 1999. - № 76. – P. 1619.
87. Qian, H. Distribution of molecular aggregation by analysis of fluctuation moments / H. Qian, E.L. Elson // *Proc. Natl. Acad. Sci. USA.* – 1990. - № 87. P. 5479.
88. Qian, H. On the analysis of high order moments of fluorescence fluctuations / H. Qian, E.L. Elson // *Biophys. J.* – 1990. - № 57. – P. 375.
89. Fluorescence intensity multiple distribution analysis: concurrent determination of diffusion times and molecular brightness / K. Palo [et al.] // *Biophys. J.* – 2000. – № 79. – P. 2858–2866.
90. Two-dimensional fluorescence intensity distribution analysis: Theory and applications. / P. Kask [et al.] // *Biophys. J.* - 2000. - № 78. – P. 1703–1713.
91. Fluorescence intensity and lifetime distribution analysis: toward higher accuracy in fluorescence fluctuation spectroscopy. / K. Palo [et al.] // *Biophys. J.* - 2002. - № 83. – P. 605–618.
92. Calculation of photon count number distributions via Master equations / K. Palo [et al.] // *Biophys. J.* - 2006. - № 90. - P. 2179–2191.

- 
93. Photon Counting Histogram for One–Photon Excitation / T.D. Perroud [et al.] // ChemPhysChem. - 2003. - № 4. – P. 1121–1123.
94. Huang B. Photon Counting Histogram: One–Photon Excitation / B. Huang, T.D. Perroud, R.N. Zare // ChemPhysChem. - 2004. - № 5. – P. 1523–1531.
95. Dual–Color Photon–Counting Histogram / Y. Chen [et al.] // Biophys. J. - 2005. - № 88. – P. 2177–2192.
96. Perroud, T. D. Effect of Bin Time on the Photon Counting Histogram for One–Photon Excitation / T. D. Perroud, B. Huang, R. N. Zare // ChemPhysChem. - 2005. - № 6. – P. 905 – 912.
97. Hillesheim, L.N. The photon counting histogram in fluorescence fluctuation spectroscopy with non–ideal photodetectors / L.N. Hillesheim, J.D. Müller // Biophys. J. - 2003. - № 85. – P. 1948–1958.
98. Кендал, М. Теория распределений / М. Кендал – Москва: Наука, 1966. - 587 с.
99. Wu, B. Time–Integrated Fluorescence Cumulant Analysis in Fluorescence Fluctuation Spectroscopy / B. Wu, J.D. Müller // Biophys. J. - 2005. - № 89. – P. 2721–2735.
100. Wu, B. Dual–color time–integrated fluorescence cumulant analysis / B. Wu, Y. Chen, J.D. Müller // Biophys. J. - 2006. - № 91. – P. 2687–2698.
101. Beechem, J.M. Global analysis of biochemical and biophysical data / J.M. Beechem // Methods Enzymol. – 1992. - № 210. – P. 37–54.
102. Апанасович, В.В. Моделирование парнокоррелированного потока случайных событий / В.В. Апанасович, Е.Г. Новиков // Автоматика и телемеханика. - 1989. - № 11. - С. 86–91.
103. Saffarian, S. Statistical analysis of fluorescence correlation spectroscopy: The standard deviation and bias/ S. Saffarian, E.L. Elson // Biophys J. – 2003. - № 84. – P. 2030–2042.
104. A statistical analysis of fluorescence correlation data / E. Van Graenenbroeck [et al.] // J. of Fluorescence. – 1999. - Vol. 9, No. 4. - P. 325–331.
105. Physical Interactions of the Peroxisomal Targeting Signal 1 Receptor Pex5p, Studied by Fluorescence Correlation Spectroscopy / D. Wang [et al.] // J. Biol. Chem. - 2003. - № 278. – P. 43340 – 43345.
106. Pyruvate Carboxylase Is an Essential Protein in the Assembly of Yeast Peroxisomal Oligomeric Alcohol Oxidase / P. Ozimek [et al.] // Mol. Biol. Cell. - 2003. - № 14(2). – P. 786–797.
107. Pleckstrin Homology Domain Diffusion in Dictyostelium Cytoplasm Studied Using Fluorescence Correlation Spectroscopy / Ruchira [et al.] // J. Biol. Chem. - 2004. - № 279. – P. 10013 – 10019.

- 
108. Hansenula polymorpha Pex19p Is Essential for the Formation of Functional Peroxisomal Membranes / M. Otzen [et al.] // *J. Biol. Chem.* - 2004. - № 279. – P. 19181–19190.
109. Hansenula polymorpha Pex20p is an oligomer that binds the peroxisomal targeting signal 2 (PTS2) / M. Otzen [et al.] // *J. Cell Science.* - 2005. - № 118. – P. 3409–3418.
110. The Phosducin-Like Protein PhLP1 Is Essential for G $\alpha$  Dimer Formation in Dictyostelium discoideum / J.C. Knol [et al.] // *Molecular and Cellular Biology.* - 2005. - №25 (18). – P. 8393–8400.
111. Fluorescent T7 display phages obtained by translational frameshift / E. J. Sloatweg [et al.] // *Nucleic Acids Res.* - 2006. - № 34(20). – P. e137 – e137.
112. Fluorescence fluctuation analysis of Arabidopsis thaliana somatic embryogenesis receptor-like kinase and brassinosteroid insensitive 1 receptor oligomerization / M. Hink [et al.] // *Biophys. J.* - 2008. - № 94. – P. 1052-1062.
113. Octameric alcohol oxidase dissociates into stable, soluble monomers upon incubation with dimethylsulfoxide / V. Nina [et al.] // *Arch. of Biochem. and Biophys.* – 2007. - № 459(2). – P. 208-213.
114. In Vivo Hexamerization and Characterization of the Arabidopsis AAA ATPase CDC48A Complex Using Forster Resonance Energy Transfer-Fluorescence Lifetime Imaging Microscopy and Fluorescence Correlation Spectroscopy / J. Aker [et al.] // *Plant Physiology.* – 2007. - № 145. – P. 339-350.
115. Dynamic behavior of GFP-CLIP-170 reveals fast protein turnover on microtubule plus ends / K.A. Dragestein [et al.] // *J Cell Biol.* – 2008. - № 180(4). – P. 729–737.
116. Johnson, M. L. Parameter Estimation by Least-Squares Methods / M. L. Johnson, L. M. Faunt // *Methods in Enzymology.* - 1992. - Vol. 210. - P. 1–37.
117. Demas J.N. Excited State Lifetime Measurements / J.N. Demas - New York: Academic Press, 1983. – 273 p.
118. Marquardt, D.W. Algorithm for least squares estimation of nonlinear parameters / D.W. Marquardt // *J. Soc. Indust. Appl. Math.* - 1963. - № 11. –P. 431–441.
119. Sanchez-Andres, A. Molecular Brightness Determined from a Generalized Form of Mandel's Q-Parameter / A. Sanchez-Andres, Y. Chen, J. D. Muller // *Biophys. J.* – 2005. - № 89. – P. 3531–3547.
120. Lakowicz, J.R. Principles of Fluorescence Spectroscopy / J.R. Lakowicz. - 2nd ed. - New York: Kluwer Academic/Plenum Publishers, 1999. – 698 p.
121. Van Hoek, A. Artefact and distortion sources in time correlated single photon counting / A. van Hoek, A.J.W.G. Visser // *Anal. Instrumentation.* – 1985. - № 14(3&4). – P. 359-378.

- 
122. Vos, K. Application of a reference deconvolution method to tryptophan fluorescence in proteins. A refined description of rotational dynamics/ K. Vos, A. van Hoek, A.J.W.G. Visser // *Eur. J. Biochem.* – 1987. - № 165. – P. 55-63.
123. Van Hoek, A. Ultrasensitive Time-Resolved Polarized Fluorescence Spectroscopy as a tool in Biology and Medicine/ A. van Hoek, A.J.W.G. Visser // *IEEE J. of Quantum Electronics.* – 1987. – Vol. QE-23, № 10. – P. 1812-1820.
124. Elmasri, R. Fundamentals of Database systems / R. Elmasri, S. B. Navathe. – Addison–Wesley, 2000. – 1038 с.
125. Дейт, Дж. Введение в системы баз данных. 8–ое издание. Пер. с англ. / Дж. Дейт. – Москва: Вильямс, 2005. – 1328 с.
126. Конолли, Т. Базы данных. Проектирование, реализация и сопровождение. Теория и практика / Т. Конолли, К. Берг. – Москва, СПб, Киев: Вильямс, 2000. – 1120 с.
127. Хомоненко, А. Д. Базы данных. Учебное пособие / А. Д. Хомоненко. – СПб.: Корона, 2002. – 672 с.
128. The global analysis of fluorescence intensity and anisotropy decay data: second–generation theory and programs / J.M. Beechem [et al.] // *Topics in fluorescence spectroscopy: in 5 vol. /Editor J.R. Lakowicz.* - New York: Plenum Press, 1991. – Vol. 2. - P. 241–305.
129. O’Connor, D.V. Time–Correlated Single Photon Counting / D.V. O’Connor, D. Phillips. – London: Academic Press, 1984. – 288 p.
130. Fluorescence correlation spectroscopy with high count rates and low background, analysis of translational diffusion / R. Rigler [et al.] // *Eur. Biophys. J.* - 1993. - № 22. - P. 169–175.
131. Performance of Fluorescence Correlation Spectroscopy for measuring diffusion and concentration / J. Enderlein [et al.] // *Chem. Phys. Chem.* - 2005. - № 6. – P. 2324–2336.
132. Probing Ligand Protein Binding Equilibria with Fluorescence Fluctuation Spectroscopy. / Y. Chen [et al.] // *Biophys. J.* - 2000. – № 79. – P. 1074–1084.
133. Molecular Brightness Characterization of EGFP in Vivo by Fluorescence Fluctuation Spectroscopy / Y. Chen [et al.] // *Biophys. J.* - 2002. - № 82. - P. 133–144.
134. Di Cera, E. Use of weighting functions in data fitting / E. Di Cera // *Methods Enzymol.* – 1992. - № 210. – P. 68–87.
135. Schistosoma: analysis of monoclonal antibodies reactive with the circulating antigens CAA and CCA / A.M. Deelder [et al.] // *Parasitology.* - 1996. - № 112(1). – P. 21–35.
136. Numerical recipes in C. The art of scientific computing / W.H. Press [et al.]. - 2nd ed. – Cambridge University Press, 1992. – 997 p.

---

137. Lau, H.T. A numerical library in C for scientists and engineers / H.T. Lau - CRC Press, Inc., 1995. – 796 p.

138. Горев, А. Эффективная работа с СУБД / А. Горев. – СПб: ПИТЕР, 1998. - 414 с.

### **Список публикаций соискателя**

1-А. Skakun, V.V. Modeling of discrete stochastic system with a sequence of event transformation / V.V. Skakun, V.V. Apanasovich // Intelligent Technologies in Human-Related Sciences: proceedings of the International Conference ITHURS-96, Leon, Spain, Jul. 5-7, 1996 : in 2 vol. / Univ. de Leon; editor: E.L. Gonzalez. - Leon, 1996. - Vol. 2. - P. 395 – 398.

2-А. Скакун, В.В. Анализ систем последовательного преобразования потоков случайных событий / В.В. Скакун, В.В. Апанасович // Исследование систем и сетей массового обслуживания: материалы Междунар. науч. конф. (BWWQT-96), Гродно, 29 – 31 янв. 1996 / Гродненский. гос. ун-т; редкол.: М. А. Матальцкий [и др.]. - Гродно, 1996. - С. 11.

3-А. Скакун, В.В. Интегрированная система имитационного моделирования систем преобразования случайных потоков / В.В. Скакун // Анализ и применение систем и сетей массового обслуживания: материалы науч. конференции, Минск, 6-8 дек. 1994 г. / Белорус. гос. ун-т. - Минск, 1994. - С. 126.

4-А. Apanasovich, V.V. Nonlinear dead time distortions of random event flow characteristics in detection systems / V.V. Apanasovich, E.E. Prolisko, V.V. Skakun // Nonlinear Phenomena in Complex Systems: Third Annual Seminar: proceedings of the International conference NPC'S'94, Polatsk, Feb. 14-16, 1994 / Institute of Physics; ed.: V. I. Kuvshinov [et al.]. - Minsk, 1995. - P. 374 - 375.

5-А. Skakun, V.V. Program complex for simulation of point processes transformation systems / V.V. Skakun, V.V. Apanasovich // Computer Data Analysis and Modeling: proc. of the International Conference CDAM, Minsk, Sep. 4-8, 1995 : in 2 vol. / Bel. State Univ.; eds.: Y. Kharin [et al.]. - Minsk, 1995. - Vol. 1. - P. 20 - 21.

6-А. Скакун, В.В. Интегрированная среда имитационного моделирования систем преобразования случайных потоков / В.В. Скакун, В.В. Апанасович // Каталог фонда программных средств. Программные продукты для персональных ЭВМ. – Минск: изд. центр БГУ, 1995. - Выпуск 5, ч. 2. - С. 19-20.

7-А. Апанасович, В.В. «TRANS» - интегрированная система имитационного моделирования систем преобразования потоков случайных событий / В.В. Апанасович, В.В. Скакун // Каталог фонда программных средств. Программные продукты для персональных ЭВМ. – Минск: изд. центр БГУ, 1999. - Выпуск 9, ч. 2. - С. 16-17.

8-А. Апанасович, В.В. «VOLNA» - интегрированная система для анализа временных и корреляционных характеристик случайных процессов в электронных

---

приборах / В.В. Апанасович, В.В. Скакун // Каталог Фонда Программных Средств, программные продукты для персональных ЭВМ. – Минск: изд. центр БГУ, 2001. - Выпуск 10, ч. 2. - С. 38.

9-A. Integrated data analysis in time-resolved fluorescence and fluorescence correlation spectroscopy / M.M. Yatskou, A.V. Digris, E.G. Novikov, V.V. Skakun, V.V. Apanasovich // Recent Research Developments in Physical Chemistry. – 2004. - Vol. 7, Part I. - P. 165-183.

10-A. Global Analysis of Fluorescence Fluctuation Data / V.V. Skakun, M.A. Hink, A.V. Digris, R. Engel, E.G. Novikov, V.V. Apanasovich, A.J.W.G. Visser // Eur. Biophys. J. – 2005. - № 34. – P. 323–334.

11-A. Fitting software for fluorescence correlation spectroscopy / A.V. Digris, V.V. Skakun, E.G. Novikov, V.V. Apanasovich, M.A. Hink, A.J.W.G. Visser // Europ. conf. on the spectroscopy of biological molecules: proc. of the Int. conf., Twente, Aug. 29 to Sep. 2, 1999 / Spectroscopy of Biological Molecules: New Directions; editors: J. Greve [et al.]. – Dordrecht: Kluwer Academic Publishers, 1999. - P. 421-422.

12-A. Global Analysis of fluorescence fluctuation data / A.V. Digris, V.V. Skakun, E.G. Novikov, V.V. Apanasovich, M.A. Hink, A.J.W.G. Visser // Biophysical Society meeting: A0 poster, New Orleans, Feb. 12 - 16, 2000 // Biophysical J. – 2000. - № 78. – P. 440A.

13-A. Змитрович, А.И. Базы данных и знаний: учебное пособие. / А.И. Змитрович, В.В. Апанасович, В.В. Скакун. – Минск: Издательский центр БГУ, 2007, - 364 с.

14-A. Thermal stability of a flavoprotein assessed from associative analysis of polarized time-resolved fluorescence spectroscopy / A.V. Digris, V.V. Skakun, E.G. Novikov, A. van Hoek, A.J.W.G. Visser // Eur. Biophys. J. – 1999. - № 28. - P. 526-531.

15-A. Excitation Energy Migration in a Photonic Dye-Zeolite Antenna: Computational Techniques / M. Meyer, M. Yatskou, M. Pfenniger, S. Huber, G. Calzaferri, A. Digris, V. Skakun, E. Barsukov, V.V. Apanasovich // J. Comput. Methods in Sci. and Eng. (JCMSE). – 2003. - № 3. - P. 201-208.

16-A. Анализ фазовых спектров затухания люминесценции фотонной антенны на цеолит-красителях / В.В. Апанасович, Е.А. Барсуков, А.В. Дигрис, В.В. Скакун, Н.Н. Яцков, М. Майер, Ш. Хубер, Д. Калсафери // Радиофизика и электроника: Сб. науч. тр. – Мн.: БГУ, 2003. - Вып. 6. – С. 35-40.

17-A. Skakun, V.V. Correction for out-of-focus emission in fluorescence fluctuation spectroscopy; generalization of the algorithms / V.V. Skakun, E.G. Novikov, O.A. Mayboroda // The 9<sup>th</sup> Carl Zeiss sponsored Workshop on FCS and related methods: A1 poster, Stockholm, Dec. 4 - 6, 2006 / Workshop abstracts [Electronic resource]. – 2006. - Mode of access: <http://www.biomolphysics.kth.se/meetings/FCS-2006.html>. - Date of access: 17.01.07.

---

18-A. Initial Guesses Generation for Fluorescence Intensity Distribution Analysis / V.V. Skakun, E.G. Novikov, V.V. Apanasovich, H.J. Tanke, A.M. Deelder, O.A. Mayboroda // Eur. Biophys. J. – 2006. - № 35(5). – P. 410-423.

19-A. Initial Guesses Generation for Fluorescence Intensity Distribution Analysis / V.V. Skakun, E.G. Novikov, V.V. Apanasovich, H.J. Tanke, A.M. Deelder, O.A. Mayboroda // The 7<sup>th</sup> International Carl Zeiss sponsored Workshop on FCS and Related Methods: proc. of the Int. conf., Dresden, Oct. 5 - 6, 2004 / Workshop abstracts [Electronic resource]. – 2004. - Mode of access: [www.zeiss.de/fcsevents/](http://www.zeiss.de/fcsevents/). - Date of access: 08.03.05.

20-A. Скакун, В.В. Анализ распределения числа фотоотсчетов с коррекцией профиля засветки в одномолекулярной флуктуационной спектроскопии / В.В. Скакун, В.В. Апанасович // Весн. Беларус. гос. ун-та. Сер. 1, Физика. Математика. Информатика. – 2008. – № 2. – С. 31–35.

21-A. Скакун, В.В. Кумулянтный анализ потоков фотонов в флуоресцентной флуктуационной спектроскопии / В.В. Скакун, В.В. Апанасович // Современные информационные компьютерные технологии: сб. науч. ст. Междунар. науч. конф. mcIT-2008, Гродно, 21-24 апр. 2008 г. : в 2 ч. / Гродн. гос. унив.; редкол.: Е. А. Ровба [и др.]. - Гродно, 2008. - Ч. 2. - С. 258-261.

22-A. Шингарев, И.П. Начальные приближения для метода анализа распределения счета фотонов двухкомпонентной системы / И.П. Шингарев, В.В. Скакун // Физика конденсированного состояния: тез. докл. XVI Респ. науч. конф. аспирантов, магистрантов и студентов, Гродно, 23-25 апр. 2008 г.: в 2 ч. / Гродн. гос. унив.; редкол.: Е. А. Ровба [и др.]. – Гродно, 2008. - Ч. 1. - С. 239 - 240.

23-A. Скакун, В.В. Системы управления базами данных: пособие для студентов факультета радиофизики и электроники / В.В. Скакун. – Минск: БГУ, 2008. – 114 с.

**Примеры применения выражений, описывающих последовательное преобразование потоков**

Приведем некоторые примеры использования выражений (2.3) и (2.5) (смотри подраздел 2.1.1), описывающих последовательное преобразование потоков. Рассмотрим цепочку, состоящую из трех последовательно соединенных преобразователей: независимых потерь, линейного сдвига и продлевающего мертвого времени, смотри рисунок А1. Такая система преобразования потоков случайных событий часто встречается в задачах моделирования флуоресцентной спектроскопии, регистраторов слабого оптического излучения и других систем.



**Рисунок А1 – Система преобразования потоков случайных событий**

Пусть независимые потери происходят с вероятностью  $1-r$  ( $r$  принадлежат интервалу  $[0,1]$ ), линейный сдвиг происходит на случайную величину  $\beta$  с функцией распределения  $B(x) = 1 - e^{-bx}$ , а мертвое время характеризуется условной функцией распределения  $D(x) = 1 - e^{-cx}$  и пусть входной поток  $A$  будет Пуассоновским с интенсивностью  $a$ . ПФЛ потока Пуассона имеет вид (1.22).

Независимые потери событий с функцией прореживания  $r(t)$  описываются УПФЛ (1.38). Подставим выражение (1.38) в (2.5)

$$L^B[v; \Omega] = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{\Phi^m} \frac{\delta^m L^A[u; \Phi]}{\delta u(t_1^A) \dots \delta u(t_m^A)} \Big|_{u(t)=0} \prod_{i=1}^m (1 + v(t_i^A)r(t_1^A) - 1) dt_1^A \dots dt_m^A. \quad (A.1)$$

Для потока Пуассона

$$\frac{\delta^m L^A[u; \Phi]}{\delta u(t_1^A) \dots \delta u(t_m^A)} \Big|_{u(t)=0} = \prod_{i=1}^m \lambda(t_i^A). \quad (A.2)$$

Принимая во внимание (A.1) и (A.2), получим

$$\begin{aligned} L^B[v; \Omega] &= 1 + \sum_{m=1}^{\infty} \int_{\Phi^m} \frac{1}{m!} \prod_{i=1}^m (\lambda(t_i^A)v(t_i^A)r(t_1^A)) dt_1^A \dots dt_m^A = \\ &= 1 + \int_{\Phi} \lambda(t_1^A)v(t_1^A)r(t_1^A) dt_1^A + \sum_{m=2}^{\infty} \int_{\Phi^m} \frac{1}{m!} \prod_{i=1}^m (\lambda(t_i^A)v(t_i^A)r(t_1^A)) dt_1^A \dots dt_m^A. \end{aligned} \quad (A.3)$$

Для интенсивности преобразованного потока получим выражение



$$f_1^B(t) = \frac{\delta L^B[v, \Omega]}{\delta v(t_1^A)} \Big|_{v(t)=0} = \lambda(t)r(t). \quad (\text{A.4})$$

В общем случае

$$\frac{\delta^m L^B[v; \Omega]}{\delta v(t_1^B) \dots \delta v(t_m^B)} \Big|_{v(t)=0} = \prod_{i=1}^m r(t_i^B) \lambda(t_i^B). \quad (\text{A.5})$$

Для линейного независимого сдвига событий вида  $\xi_i = \gamma_i + \beta_i$ ,  $i=1, 2, \dots$ , где  $\beta_i$  – одинаково распределенные случайные величины с функцией распределения  $B(x)$ , УПФЛ имеет вид (1.44). Подставим (1.44) и (A.5) в (2.5)

$$\begin{aligned} L^C[w; \Theta] &= 1 + \sum_{m=1}^{\infty} \int_{\Omega^m} \frac{1}{m!} \prod_{i=1}^m r(t_i^B) \lambda(t_i^B) \prod_{i=1}^m \left( 1 + \int_{\Theta} v(t_i^B) dB(t - t_i^B) - 1 \right) dt_1^B \dots dt_m^B = \\ &= 1 + \int_{\Omega} r(t_i^B) \lambda(t_i^B) \int_{\Theta} v(t_i^B) dB(t - t_i^B) + \\ &+ \sum_{m=2}^{\infty} \int_{\Omega^m} \frac{1}{m!} \prod_{i=1}^m \left( r(t_i^B) \lambda(t_i^B) \int_{\Theta} v(t_i^B) dB(t - t_i^B) \right) dt_1^B \dots dt_m^B. \end{aligned} \quad (\text{A.6})$$

Дифференцируя (A.6) при  $u(t)=0$ , получим выражение для интенсивности  $f_1^C(x)$  преобразованного потока через интенсивность  $\lambda(t)$  входного потока

$$f_1^C(t) = \frac{\delta L^C[w, \Theta]}{\delta w(t_1^C)} \Big|_{w(t)=0} = \int_{\Theta} \lambda(t_1^C) r(t_1^C) \frac{dB(t - t_1^C)}{dt} dt_1^C. \quad (\text{A.7})$$

УПФЛ, характеризующийся продлевающимся мертвым временем с условной функцией распределения  $D(x, t)$ , имеет вид (1.49) [4]. Опуская промежуточные вычисления, запишем выражения для интенсивности выходного потока

$$f_1^D(t) = \frac{\delta L^D[s, \Psi]}{\delta s(t_1^C)} \Big|_{s(t_1^C)=1(t_1^C-t)+B(t-t_1^C)-1}. \quad (\text{A.8})$$

Тогда, принимая во внимание (A.7) и (A.8) и (1.22), получим [1-A]

$$\begin{aligned} f_1^C(t) &= ar(1 - e^{-bt}), \\ f_1^D(t) &= \frac{\delta \exp \left\{ \int f_1^C(t_1^C) s(t_1^C) dt \right\}}{\delta s(t_1^C)} \Big|_{s(t_1^C)=1(t_1^C-t)+B(t-t_1^C)-1} = \\ &= ar(1 - e^{-bt}) \exp \left\{ \frac{ar}{c} (e^{-ct} - 1) + \frac{ar}{c-b} (e^{-bt} - e^{-ct}) \right\}. \end{aligned} \quad (\text{A.9})$$

**Расчет факториальных моментов и факториальных кумулянтов**

Расчет первых 4 факториальных моментов  $F_k$  через моменты  $M_k$  и первых 4 факториальных кумулянтов  $K_k$  через обычные, факториальные и центральные моменты представлены ниже

$$F_1 = \langle n \rangle = M_1$$

$$F_2 = \langle n(n-1) \rangle = \langle n^2 \rangle - \langle n \rangle = M_2 - M_1$$

$$F_3 = \langle n(n-1)(n-2) \rangle = \langle n^3 \rangle - 3\langle n^2 \rangle + 2\langle n \rangle = M_3 - 3M_2 + 2M_1$$

$$F_4 = \langle n(n-1)(n-2)(n-3) \rangle = \langle n^4 \rangle - 6\langle n^3 \rangle + 11\langle n^2 \rangle - 6\langle n \rangle = \\ = M_4 - 6M_3 + 11M_2 - 6M_1$$

$$K_1 = F_1 = \langle n \rangle$$

$$K_2 = F_2 - K_1 F_1 = F_2 - F_1^2 = \langle n^2 \rangle - \langle n \rangle^2 = \langle \Delta n^2 \rangle - \langle n \rangle$$

$$K_3 = F_3 - 2K_2 F_1 - K_1 F_2 = F_3 - 3F_1 F_2 + 2F_1^3 = \\ = \langle n^3 \rangle - 3\langle n^2 \rangle \langle n \rangle + 2\langle n \rangle^3 + 2\langle n \rangle - 3\langle n^2 \rangle + 3\langle n \rangle^2 = \\ = \langle \Delta n^3 \rangle - 3\langle \Delta n^2 \rangle + 2\langle n \rangle$$

$$K_4 = F_4 - 3K_3 F_1 - 3K_2 F_2 - K_1 F_3 = F_4 + 12F_1^2 F_2 - 4F_1 F_3 - 3F_2^2 - 6F_1^4 = \\ = \langle n^4 \rangle - 6\langle n^3 \rangle + 11\langle n^2 \rangle - 6\langle n \rangle - 4\langle n \rangle \langle n^3 \rangle + 18\langle n \rangle \langle n^2 \rangle - \\ - 11\langle n \rangle^2 - 3\langle n^2 \rangle^2 + 12\langle n^2 \rangle \langle n \rangle^2 - 12\langle n \rangle^3 - 6\langle n \rangle^4 = \\ = \langle \Delta n^4 \rangle - 6\langle \Delta n^3 \rangle + 11\langle \Delta n^2 \rangle - 3\langle \Delta n^2 \rangle^2 - 6\langle n \rangle$$

**Решение систем (3.26), (3.28)**

Система (3.26) может быть переписана в следующем виде

$$\begin{cases} a^2(16u - 20) + b^2(9u - 12) + 8ab(3u - 4) + 16a(4u - 5) + b(48u - 68) + (64u - 75) = 0 \\ a^2(4v - 8) + b^2(4v - 9) + 2ab(4v - 9) + 12a(v - 2) + b(12v - 30) + (9v - 16) = 0, \end{cases} \quad (C.1)$$

где

$$u = 16875K_5K_3/(16384K_4^2), \quad v = 1024K_4K_2/(729K_3^2). \quad (C.2)$$

Аналогично для системы (3.28) получим

$$\begin{cases} 4a^2(u - 1) + 3b^2(3u - 4) + 4ab(3u - 4) + 8a(u - 1) + 4b(3u - 5) + (4u - 3) = 0 \\ a^2(4v - 8) + b^2(6v - 18) + ab(10v - 30) + a(10v - 20) + b(13v - 51) + (6v - 8) = 0, \end{cases} \quad (C.3)$$

где

$$u = 27K_3(K_1 - \lambda T)/(64K_2^2), \quad v = 16K_4(K_1 - \lambda T)/(27K_2K_3). \quad (C.4)$$

Системы (C.1, C.3) могут быть сведены к полиному четвертого порядка по  $a$  или по  $b$  и решены либо аналитически, либо численно. Обозначим множители перед  $a^2, b^2, ab, a, b, 1$  в системе (C.1) (аналогично для системы C.3) как  $A_1 - A_6$  (первое уравнение) и  $B_1 - B_6$  (второе уравнение). Комбинируя уравнения (C.1), получим выражение для оценки  $b$

$$b = -(D_1a^2 + D_3a + D_5)/(D_2a + D_4), \quad (C.5)$$

где  $D_1 = A_1B_2 - A_2B_1$ ,  $D_2 = A_3B_2 - A_2B_3$ ,  $D_3 = A_4B_2 - A_2B_4$ ,  $D_4 = A_5B_2 - A_2B_5$ ,  $D_5 = A_6B_2 - A_2B_6$ . После подстановки (C.5) во второе уравнение системы (C.1) получим полином 4 порядка

$$\begin{aligned} & (B_1D_1^2 + B_2D_1^2 - B_3D_1D_2)a^4 + \\ & + (2B_1D_2D_4 + 2B_2D_1D_3 - B_3D_1D_4 - B_3D_2D_3 + B_4D_2^2 - B_5D_1D_2)a^3 + \\ & + (B_1D_4^2 + B_2D_3^2 + 2B_2D_1D_5 - B_3D_3D_4 - B_3D_2D_5 + 2B_4D_2D_4 - \\ & - B_5D_1D_4 - B_5D_2D_3 + B_6D_2^2)a^2 + (2B_2D_3D_5 - B_3D_4D_5 + B_4D_4^2 - \\ & - B_5D_3D_4 - B_5D_2D_5 + 2B_6D_2D_4)a + (B_2D_5^2 - B_5D_4D_5 + B_6D_4^2) = 0. \end{aligned} \quad (C.6)$$

Если определитель полинома положительный, получим два действительных и два комплексно-сопряженных корня [35]. Один действительный корень (всегда представимый решением  $a = -7/2$  и  $b = 2$ ) приводит к сингулярности модели ( $N = 0, q = \infty$ ) и должен быть опущен. Если определитель отрицательный, получим четыре действительных корня. Один корень (тоже представимый решением  $a = -7/2$  и  $b = 2$ ) должен быть опущен. Три остальные корни физически приемлемы

и для их поиска необходимо применять численные алгоритмы. Например, можно использовать алгоритмы, опубликованные в [136, 137].

**Решение системы (3.32)**

Система (3.32) может быть сведена к полиному третьего порядка по  $q_1$  (или  $q_2$ ) и решена либо аналитически, либо численно. Обозначим левые части системы (3.32) как

$$\begin{cases} (K_1 - \lambda T)/T = A_1 \\ K_2/T^2 = A_2 \\ K_3/\chi_3 T^3 = A_3 \\ K_4/\chi_4 T^4 = A_4. \end{cases} \quad (D.1)$$

Комбинируя уравнения в (D.1) после подстановки (3.22), получим систему для оценки  $q_1$

$$\begin{cases} A_2 N_2 q_2^4 - 2A_3 N_2 q_2^3 + A_2 N_2 q_2^2 - A_2 A_4 + A_3^2 = 0 \\ A_1 N_2 q_2^4 - A_2 N_2 q_2^3 - A_3 N_2 q_2^2 + A_4 N_2 q_2 - A_1 A_4 + A_2 A_3 = 0. \end{cases} \quad (D.2)$$

Из первого уравнения получим выражение для оценки  $N_2$

$$N_2 = \frac{A_2 A_4 - A_3^2}{A_2 q_2^4 - 2A_3 q_2^3 + A_4 q_2^2} \quad (D.3)$$

После подстановки (D.3) во второе уравнение (D.2) получим полином третьего порядка для  $q_2$

$$(A_2^3 A_3 - A_3^2 A_1) q_2^3 + (2A_1 A_3 A_4 - A_2^2 A_4 - A_2 A_3^2) q_2^2 + (A_3^3 - A_1 A_4^2) q_2 + (A_4^2 A_2 - A_3^2 A_4) = 0. \quad (D.4)$$

Заключительно

$$q_1 = A_3 - N_2 q_2^3 / (A_2 - N_2 q_2^3), \quad N_1 = A_1 - N_2 q_2 / q_1. \quad (D.5)$$

Если дискриминант полинома положительный, можно найти аналитическое решение системы (3.32). Если обозначить

$$\begin{aligned} Z_1 &= A_4^2 A_1^2 - 6A_1 A_2 A_3 A_4 - 3A_3^2 A_2^2 + 4A_1 A_3^3 + 4A_4 A_2^3 \\ Z_2 &= A_1^3 A_4^2 + A_1^3 A_3 (2A_3^2 - 5A_2 A_4) + 3A_1 A_4 A_2^3 - A_3 A_2^4 \\ Z_3 &= A_1^2 A_4 - 2A_1 A_2 A_3 + A_2^3 \end{aligned} \quad (D.6)$$

получим

$$q_1 = \frac{\sqrt{Z_1} - A_1 A_4 + A_2 A_3}{2(A_2^2 - A_1 A_3)}, \quad q_2 = \frac{\sqrt{Z_1} + A_1 A_4 - A_2 A_3}{2(A_1 A_3 - A_2^2)} \quad (D.7)$$

$$N_1 = \frac{\sqrt{Z_1} Z_2 + Z_1 Z_3}{2Z_1 (A_2 A_4 - A_3^2)}, \quad N_2 = \frac{\sqrt{Z_1} Z_2 - Z_1 Z_3}{2Z_1 (A_3^2 - A_2 A_4)}. \quad (D.8)$$

Если дискриминант полинома отрицательный, получим два или три действительных корня. Для их поиска необходимо применять численные алгоритмы.

**Сведение трехмерного интеграла в (3.4) к одномерному**

Функция профиля засветки  $B(\mathbf{r})$  предполагается трехмерной гауссовской с  $B_0 \neq 1$

$$B(x, y, z) = B_0 \exp\left(-\frac{2x^2 + 2y^2}{\omega_0^2} - \frac{2z^2}{z_0^2}\right). \quad (\text{E.1})$$

где  $w_0^2, z_0$  расстояния от фокуса в поперечном и продольном направлениях. В декартовой системе координат (3.4) можно записать как

$$G_{main}(\xi) = \exp\left\{\sum_i \langle C \rangle_i \iiint (e^{(\xi-1)q_i TB(x,y,z)} - 1) dx dy dz\right\}. \quad (\text{E.2})$$

Для сведения трехмерного интеграла в (E.2) к одномерному применим специальное преобразование из декартовой системы координат в модифицированную сферическую

$$\begin{cases} x = r\omega_0 \cos\varphi \sin\alpha \\ y = r\omega_0 \sin\varphi \sin\alpha \\ z = z_0 r \cos\alpha. \end{cases} \quad (\text{E.3})$$

Якобиан такого преобразования равен

$$\sqrt{g} = \omega_0^2 z_0 r^2 \sin\alpha \quad (\text{E.4})$$

и выражение для  $B(\mathbf{r})$  в виде (E.1) упростится до

$$B(r) = B_0 e^{-2r^2}. \quad (\text{E.5})$$

В этой системе координат интеграл по объему будет зависеть от параметров  $w_0, z_0$

$$V(r_0, \omega_0, z_0) = \int_0^{r_0} \int_0^{2\pi} \int_0^\pi r^2 \omega_0^2 z_0 dr d\varphi d\alpha = \frac{4\pi}{3} r_0^3 \omega_0^2 z_0. \quad (\text{E.6})$$

После подстановки (3.41) (нормализация к объему засветки и яркости в фокусе) и интегрирования по  $\varphi$  и  $\alpha$  (E.2) примет вид

$$\begin{aligned} G_{main}(\xi) &= \exp\left\{\sum_i 4\pi\omega_0^2 z_0 N_{2MNi} \int_0^\infty \left(e^{(\xi-1)q_{2MNi} TB_0 e^{-2r^2}} - 1\right) r^2 dr\right\} = \\ &= \exp\left\{\sum_i \frac{4(1+F_2)}{\sqrt{\pi}(1+F_1)^2} N_{2MNi} \int_0^\infty \left(e^{(\xi-1)q_{2MNi} T \frac{2\sqrt{2}(1+F_1)}{1+F_2} e^{-2r^2}} - 1\right) r^2 dr\right\}. \end{aligned} \quad (\text{E.7})$$

Соответственно, выражение (3.5) для коррекции второго порядка примет вид

$$G_{corr}(\xi) = \exp\left\{\sum_i N_{2MNi} \left[ (\xi-1) \frac{q_{2MNi} T F_1}{1+F_1} + \frac{(\xi-1)^2}{2} \frac{q_{2MNi}^2 T^2 F_2}{1+F_2} \right]\right\}. \quad (\text{E.8})$$

**Применение нормализации вида (1.84) в методе РСН**

Применение другого вида нормализации в РСН меняет только вид выражения для одномолекулярной РЧФ. Принимая во внимание (3.41) (и, следовательно,  $V_{ref} = 1$ ), получим для выражения (1.90)

$$p^{(1)}(n, Q, q) = \frac{1 + F_2}{(1 + F_1)^2 n! Q} \sum_{k=n}^{\infty} \frac{(-1)^{k-n} (2\sqrt{2}qT(1 + F_1))^k (1 + F_k)}{(2k)^{2/3} (1 + F_2)^k (k - n)!}. \quad (F.1)$$

Выражение (F.1) может быть представлено суммой одномолекулярной РЧФ  $p_G^{(1)}(n, Q, q)$  с гауссовской аппроксимацией профиля засветки и коррекции гауссовской аппроксимации. Выполняя интегрирование в специальной системе координат, заданной преобразованием (E.3), получим

$$p_G^{(1)}(n, Q, q) = \frac{4(1 + F_2)}{Q\sqrt{\pi}(1 + F_1)^2} \int_0^{\infty} Poi\left(n, \frac{2\sqrt{2}(1 + F_1)qTe^{-2r^2}}{(1 + F_2)}\right) r^2 dr, \quad (F.2)$$

где  $Poi$  обозначает распределение Пуассона. Выражение (F.2) может быть также записано через неполную гамма функцию, т.е. аналогично [27]

$$p_G^{(1)}(n, Q, q) = \frac{(1 + F_2)}{Q\sqrt{\pi}(1 + F_1)^2 n!} \int_0^{\infty} \gamma\left(n, \frac{2\sqrt{2}(1 + F_1)qTe^{-2r^2}}{(1 + F_2)}\right) dr, n > 0. \quad (F.3)$$

Выражение (F.3) можно получить, применяя интегрирование в цилиндрической системе координат и принимая во внимание следующие тождества

$$\gamma(k, x) = \frac{1}{(k - 1)!} \int_0^x t^{k-1} e^{-t} dt, \quad (F.4)$$

$$\gamma(k, x) = 1 - \sum_{j=0}^{k-1} Poi(j, x). \quad (F.5)$$

Таким образом, выражение для одномолекулярного РЧФ с коррекцией II порядка и нормализацией вида (1.84) примет вид

$$\begin{aligned} p^{(1)}(n, q) &= p_G^{(1)}(n, q) + \frac{qTF_1}{Q(1 + F_1)} - \frac{(qT)^2 F_2}{Q(1 + F_2)}, n = 1, \\ p^{(1)}(n, q) &= p_G^{(1)}(n, q) + \frac{(qT)^2 F_2}{2Q(1 + F_2)}, n = 2, \\ p^{(1)}(n, q) &= p_G^{(1)}(n, q), n > 2, \\ p^{(1)}(0, q) &= 1 - \sum_n p^{(1)}(n, q). \end{aligned} \quad (F.6)$$

**Структуры данных БД измерений и результатов анализа во время разрешимой флуоресцентной спектроскопии**

Рассмотрим типичную запись журнала измерений в методах анализа затухания интенсивности флуоресценции (смотри таблицу G1).

Таблица G1 – Журнал измерений

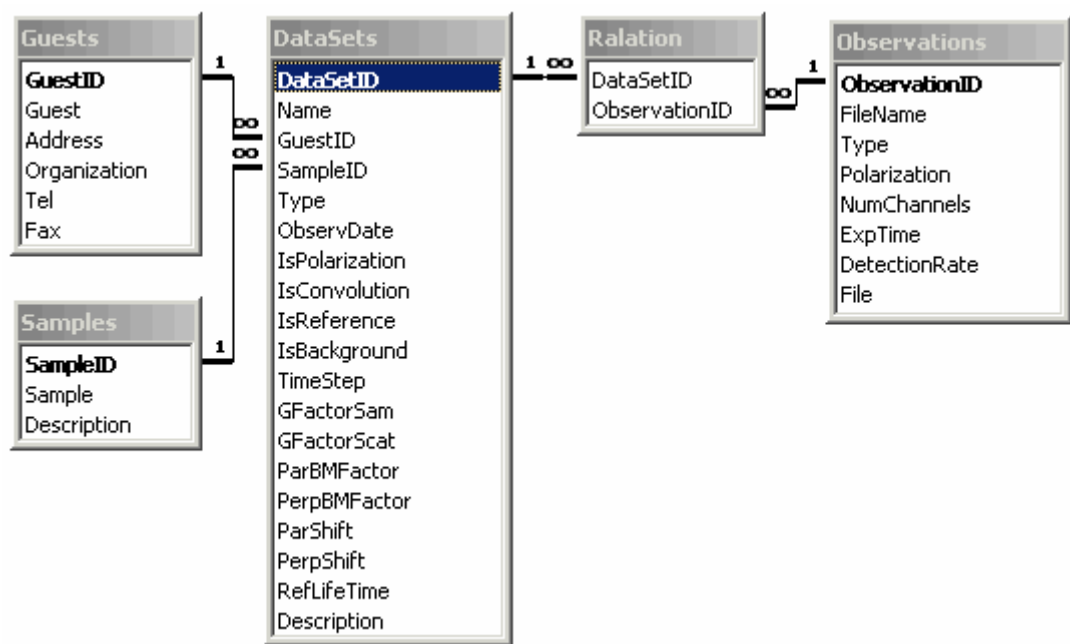
Имя файла	Тип	Поляризация
1. 10oct001	Scatter	Parallel
2. 10oct002	Scatter	Perpendicular
3. 10oct003	Sample	Parallel
4. 10oct004	Sample	Perpendicular
5. 10oct005	Background	Parallel
6. 10oct006	Background	Perpendicular
7. 10oct007	Scatter	Parallel
8. 10oct008	Scatter	Perpendicular
9. 10oct009	Sample	Parallel
10. 10oct010	Sample	Perpendicular
11. 10oct013	Background	Parallel
12. 10oct014	Background	Perpendicular
13. 10oct015	Scatter	Parallel
14. 10oct016	Scatter	Perpendicular

Первые восемь файлов составляют первый набор данных, файлы с 7 по 14 – второй набор данных. Видно, что два файла в середине таблицы принадлежат обоим наборам данных. Таким образом, между наборами данных и принадлежащими им файлами существует связь «много-ко-многим». Такая схема данных требует создания трех таблиц в реляционных базах данных. Проведя нормализацию таблиц по первым трем нормальным формам [138, 13-А, 23-А], получим структуру данных, представленную на рисунке G1.

Каждая запись в таблице Observations описывает отдельное измерение и содержит такие экспериментальные параметры, как скорость накопления, время измерения, а также массив измеренной интенсивности. Информация, которая одинакова для каждого измерения внутри одного набора данных (например, дата измерения, образец исследования и такие параметры, как ширина канала, G – факторы, сдвиг, и т. д.), записываются в таблицу DataSets. Одна запись в ней представляет отдельный набор данных. Связь «многие-ко-многим» между этими таблицами реализована с помощью таблицы Relation. Такая связь, помимо устранения избыточности данных, позволяет выделить весь набор измерений посредством выделения только одной записи в таблице DataSets.

Имеется две вспомогательные таблицы: таблица Guests и таблица Samples. Первая содержит информацию о человеке или организации, заказавшей измерения, а вторая предоставляет детальную информацию об исследуемом образце. Обе таблицы связаны с таблицей DataSets связями «один-ко-многим».

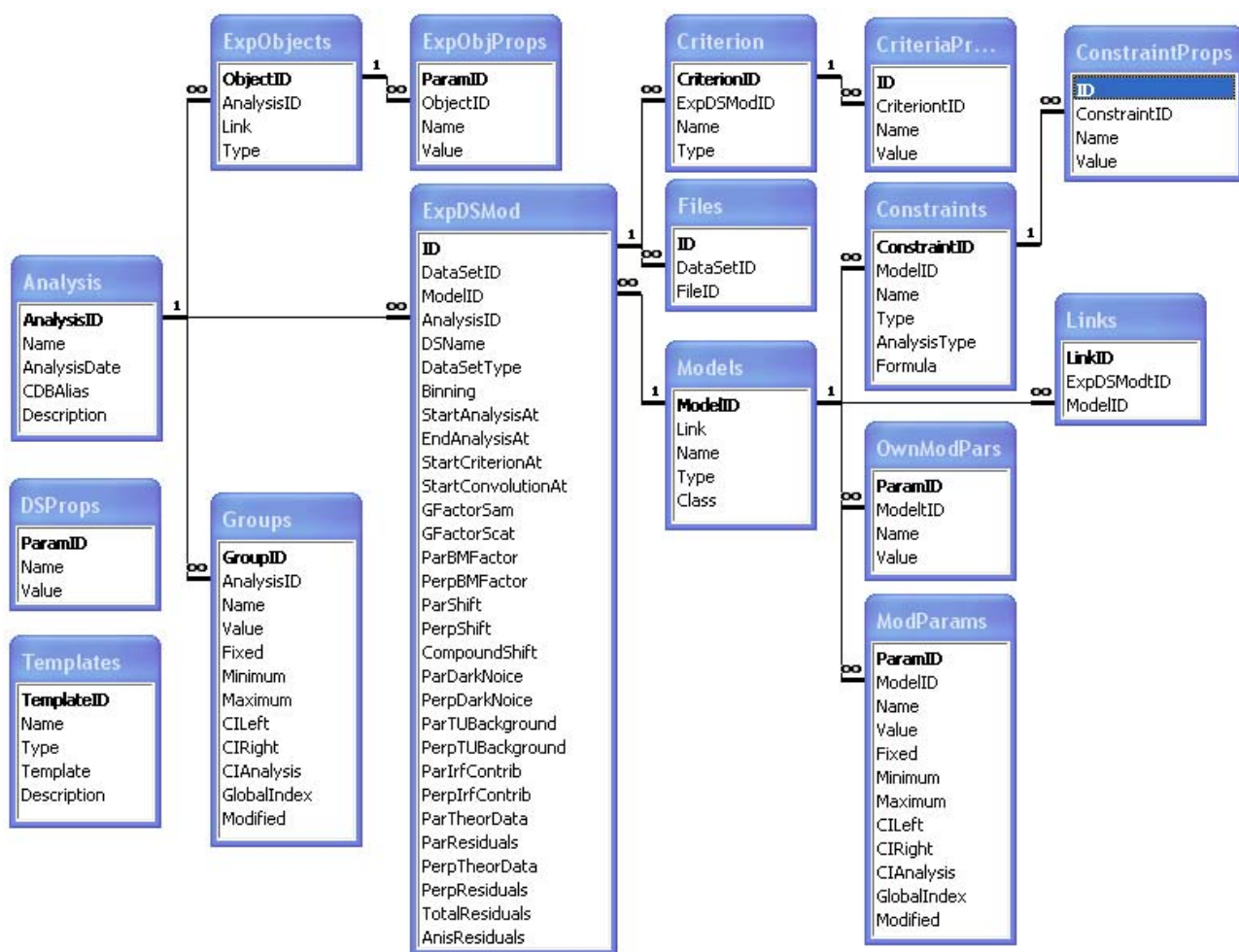




**Рисунок G1 – Структура данных измерений во время разрешимой флуоресцентной спектроскопии**

Такая конфигурация позволяет легко просматривать измерения, произведенные для конкретного лица или над определенным образцом. С другой стороны, вся информация, относящаяся к конкретному набору данных, становится доступной сразу после выделения всего одной записи в таблице DataSets.

Структура данных БД результатов анализа в комплексах TRFA Data Processor и TRFD Fitting Software аналогична рассмотренной структуре данных в подразделе 2.4. На рисунке G2 представлена структура данных базы результатов анализа во время разрешимой флуоресцентной спектроскопии. Представленная структура имеет намного более сложный вид, чем рассмотренная в подразделе 2.4. Это вызвано более сложной внутренней структурой объектов приложения, реализующего специфику анализа данных во время разрешимой флуоресцентной спектроскопии [9-A, 16-A]. Основные отличия состоят в том, что структура моделей имеет иерархический вид (например, для организации ассоциативного анализа затухания анизотропии), что отражено дополнением таблицы Models полем Link, ссылающимся на первичный ключ таблицы, и введением рекурсивных функций для чтения и модификации данной таблицы и совокупности связанных таблиц; для параметров моделей можно устанавливать сложные ограничения, что отражено ведением таблиц Constraints и ConstrainsProps; необходимо хранить набор значений критериев качества подгонки (таблицы Criterion и CriteriaProps); структура методов оптимизации и других объектов конфигурации анализа данных имеет иерархическую структуру (введены таблицы ExpObjects, ExpObjectsProps и таблица ExpObjects дополнена полем Link, ссылающимся на первичный ключ таблицы ObjectID).



**Рисунок G2 – Структура данных базы результатов анализа во время разрешимой флуоресцентной спектроскопии**

Дополнительно структура данных была оптимизирована. Все таблицы, хранящие свойства объектов, получили общий набор полей: первичный ключ, внешний ключ, Name, Value (текстового типа). Это позволяет хранить любые свойства любых объектов, независимо от их типа.

При нормализации данных пришлось несколько отступить от требований теории нормализации в пользу обеспечения целостности данных и удобного просмотра данных в базе данных. Можно видеть, что ряд полей (Name, Value) повторяется во многих таблицах, а таблицы Groups и ModParams вообще отличаются лишь полем связи. Это приводит к некоторой избыточности данных, зато, принимая во внимание, что каждая из таблиц соответствует объектам главного приложения, делает схему данных прозрачной и понятной. Это, в свою очередь, приводит к упрощению кода записи, считывания и визуализации. Таблица Analysis является главной. Каждая запись представляет отдельный эксперимент. Таблица ExpDSMod связывает таблицы Analysis, Models и таблицу DataSets из базы данных измерений. Таким образом, существует отношение «много-ко-многим» между таблицами Analysis, Models и DataSets, что

соответствует реальной схеме анализа. Удобство такой схемы состоит в том, что, выделив одну запись в таблице Analysis, легко выделить информацию, относящуюся ко всему эксперименту. Некоторое неудобство проявляется от необходимости создания гетерогенных запросов, объединяющих информацию из двух отдельных баз данных, но это разделение было продиктовано требованиями надежности.