

Моделирование частотных сканов

Богачёв А.М.

16 октября 2022 г.

Аннотация

В отчёте представлены результаты разработки и апробации модели идеального частотного скана, модели частотного скана, учитывающей нелинейность измерительного тракта и неэкспоненциальность сигнала релаксации, модели многоэкспоненциального частотного скана. В отчёте также приведены теоретические выражения, лежащие в основе разработанных моделей, особенности технической реализации, некоторые методические вопросы идентификации параметров моделей. Модели апробованы на экспериментальных данных, приведены сравнение и анализ результатов. Дополнительно выполнена идентификация многоэкспоненциальной модели с шестью экспоненциальными составляющими и для полученных результатов построены графики в координатах Аррениуса. Показано, что использование модели, учитывающей нелинейность и неэкспоненциальность, или многоэкспоненциальной модели позволяют повысить точность моделирования частотных сканов.

Содержание

1	Введение	3
2	Математические модели частотных сканов	5
2.1	Модель сигнала релаксации ёмкости	5
2.2	Модель идеального частотного скана	6
2.3	Модель, учитывающая нелинейность и неэкспоненциальность	7
2.4	Модель многоэкспоненциального частотного скана	8
3	Реализация моделей	9
3.1	Интерфейс и функционал	9
3.2	Алгоритм идентификации параметров моделей	11
3.3	Оценка качества модели и оптимизация параметров	15
4	Апробация моделей на экспериментальных данных	22
4.1	Экспериментальные данные	22
4.2	Частотный скан при температуре 263 К	24
4.2.1	Моноэкспоненциальная модель с показателем p	24
4.2.2	Моноэкспоненциальная модель с показателем $p = 1$	30
4.2.3	Многоэкспоненциальная модель с $n_exps > 1$	32
4.2.4	Сравнение результатов	36
4.3	Частотный скан при температуре 283 К	37
4.3.1	Моноэкспоненциальная модель с показателем p	37
4.3.2	Моноэкспоненциальная модель с показателем $p = 1$	39
4.3.3	Многоэкспоненциальная модель с $n_exps > 1$	42
4.3.4	Сравнение результатов	45
4.4	Частотный скан при температуре 303 К	46
4.4.1	Моноэкспоненциальная модель с показателем p	46
4.4.2	Моноэкспоненциальная модель с показателем $p = 1$	48
4.4.3	Многоэкспоненциальная модель с $n_exps > 1$	51
4.4.4	Сравнение результатов	54
4.5	Графики в координатах Аррениуса	55
5	Заключение	60
	Список литературы	61

1 Введение

Релаксационная спектроскопия глубоких уровней (РСГУ) – метод исследования электрически активных дефектов полупроводниковых материалов. Данный метод обладает высокой чувствительностью к малым концентрациям ловушек носителей заряда и является спектроскопическим. Существуют различные вариации РСГУ, например токовая и емкостная, так метод емкостной релаксационной спектроскопии глубоких уровней основан на исследовании процесса (сигнала) релаксации ёмкости барьерной структуры. Определив значения постоянной времени сигнала релаксации для разных температур образца, исследователь может определить энергию активации дефекта, вызвавшего релаксацию.

В идеальном случае релаксация ёмкости носит экспоненциальный характер, однако так бывает не всегда. Согласно обзору [5] сигналы релаксации ёмкости барьерных структур можно условно разделить на три группы:

1. Моноэкспоненциальный сигнал релаксации, обусловленный одним единственным энергетическим уровнем в запрещённой зоне полупроводника.
2. Сигнал релаксации, состоящий из суммы нескольких моноэкспоненциальных сигналов релаксации.
3. Сигнал релаксации, характеризуемый непрерывным распределением скоростей эмиссии, представленным спектральной функцией $g(\lambda)$.

В последних двух случаях сигнал не является экспоненциальным.

Задача определения постоянной времени одной или нескольких экспоненциальных составляющих процесса релаксации ёмкости или спектральной функции $g(\lambda)$ относится к задачам экспоненциального анализа. Подходы к решению таких задач, а также фундаментальные ограничения и особенности сбора и обработки экспериментальных данных рассмотрены в исчерпывающем обзоре [5].

Во Владимирском Государственном университете создан измерительно-вычислительный комплекс релаксационной спектроскопии глубоких уровней, основным измерительным прибором которого служит спектрометр DLS-82E фирмы Semilab. Измерительно-вычислительный комплекс реализует метод емкостной РСГУ с частотным сканированием при постоянной температуре. В спектрометре аппаратно реализована корреляционная обработка сигнала релаксации с опорной функцией lock-in.

Технические решения, заложенные в названном измерительном оборудовании, требуют особых методов анализа полученных экспериментальных данных, в частности, разработки и идентификации моделей частотных сканов (экспериментальных данных, полученных на измерительно-вычислительном комплексе).

В следующих разделах отчёта будут рассмотрены модели частотных сканов, их апробация на экспериментальных данных, а также некоторые технические и методические вопросы их реализации и идентификации.

В современной учебной и технической литературе набирают популярность термины «машинное обучение» и «статистическое обучение» (например в книгах [3, 10, 4]). За ними, как правило, скрывается процесс создания моделей, которые после идентификации их параметров на тренировочной выборке (экспериментальных данных с известными целевыми значениями), способны с некоторой точностью определять целевые значения для новых, ранее не встречавшихся данных. Таким образом, регрессия и классификация являются одними из самых распространенных задач машинного обучения [3]. В данном отчёте термин «машинное обучение» будет использоваться именно для обозначения процесса разработки и идентификации моделей, решения задачи регрессии.

2 Математические модели частотных сканов

В данном разделе представлены описания математических моделей сигнала релаксации ёмкости и моделей частотного скана.

2.1 Модель сигнала релаксации ёмкости

Согласно обзору [5], зависимость значения ёмкости от времени $f(t)$ для моноэкспоненциального сигнала релаксации имеет вид выражения 1.

$$f(t) = A \exp(-\lambda t), \quad (1)$$

где

A – амплитуда сигнала релаксации ёмкости;

λ – скорость экспоненциального спада, обратнопропорциональная постоянной времени сигнала релаксации τ (выражение 2).

$$\lambda = \tau^{-1} \quad (2)$$

Спектр моноэкспоненциального сигнала релаксации имеет вид, представленный на рисунке 1.

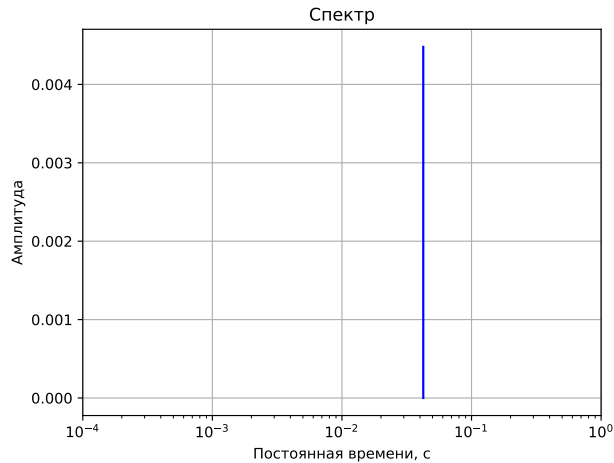


Рис. 1: Пример спектра моноэкспоненциального сигнала релаксации ёмкости. На графике амплитуда указана в пФ.

Согласно источнику [5], зависимость сигнала релаксации ёмкости от времени $f(t)$ для сигнала, образованного несколькими дискретными экспоненциальными сигналами, определяется выражением 3.

$$f(t) = \sum_{i=1}^n A_i \exp(-\lambda_i t), \quad (3)$$

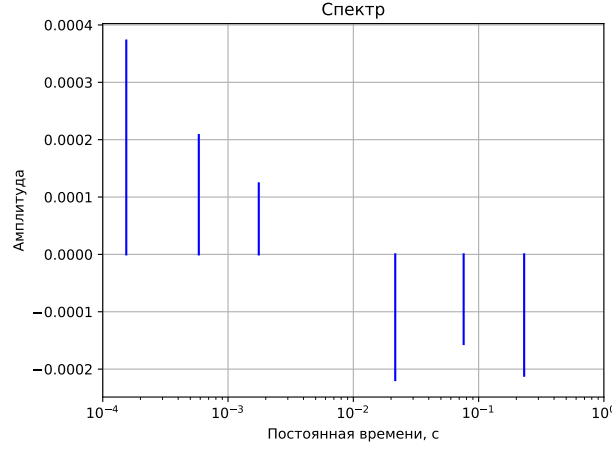


Рис. 2: Пример спектра сигнала релаксации ёмкости, содержащего несколько экспоненциальных составляющих. На графике амплитуда указана в пФ.

где n – количество экспоненциальных составляющих в спектре. Пример спектра такого сигнала показан на рисунке 2.

На спектрах амплитуды могут быть отрицательными. В данном случае знак при амплитуде указывает лишь на направление изменения ёмкости [6].

2.2 Модель идеального частотного скана

Если предполагается, что сигнал релаксации ёмкости содержит одну экспоненциальную составляющую, а измерительный тракт спектрометра линеен, то в таком случае модель частотного скана описывает аппаратные преобразования спектрометра DLS-82E, учитывая только форму его опорной функции и ширину импульса заполнения.

В спектрометре DLS-82E реализована корреляционная обработка сигнала релаксации ёмкости, таким образом сигнал на выходе аналогового тракта спектрометра определяется выражением 4, согласно публикации [5].

$$S[g(\lambda), t_c, t_d] = \frac{1}{t_c} \int_{t_d}^{t_d+t_c} f(t) W(t - t_d) dt, \quad (4)$$

где

$W(t)$ – весовая функция, определённая на интервале времени $[0, t_c]$,

t_c – период (длительность) весовой функции $W(t)$,

t_d – время задержки между началом сигнала релаксации и началом корреляционной обработки. Согласно обзору [5], время задержки t_d , обычно, вводится для улучшения избирательности или для снижения искажения сигнала из-за перегрузки измерительной системы.

$g(\lambda)$ – распределение скоростей экспоненциальных спадов, составляющих релаксационный сигнал.

Модель аппаратных преобразований (корреляционной обработки), учитывающая форму весовой функции, реализованной в спектрометре DLS-82E, для моноэкспоненциального сигнала определяется выражением 5, согласно источнику [6].

$$S(\tau, C_A, F_0, t_1) = C_A K_{BS} K_{LS} \phi(\tau, F_0, t_1), \quad (5)$$

где

C_A – амплитуда емкостного релаксационного сигнала,

K_{BS} – масштабный коэффициент, зависящий от чувствительности емкостного моста,

K_{LS} – масштабный коэффициент селектора,

τ – постоянная времени релаксации глубокого уровня,

F_0 – частота сканирования импульсов заполнения,

t_1 – длительность импульса заполнения,

$\phi(\tau, F_0, t_1)$ – функция определяемая выражением 6.

$$\phi(\tau, F_0, t_1) = M \tau F_0 e^{-\frac{0.05}{\tau F_0}} \left(1 - e^{\frac{t_1 F_0 - 0.45}{\tau F_0}} - e^{-\frac{0.5}{\tau F_0}} + e^{\frac{t_1 F_0 - 0.95}{\tau F_0}} \right), \quad (6)$$

где M – масштабный множитель.

Масштабный множитель M определяется выражением 7.

$$M(\tau, F_0, t_1) = \frac{1}{\max \left[\tau F_0 e^{-\frac{0.05}{\tau F_0}} \left(1 - e^{\frac{t_1 F_0 - 0.45}{\tau F_0}} - e^{-\frac{0.5}{\tau F_0}} + e^{\frac{t_1 F_0 - 0.95}{\tau F_0}} \right) \right]} \quad (7)$$

Введём коэффициент A (выражение 8), характеризующий амплитуду сигнала релаксации ёмкости и перепишем выражение 5 с учётом того, что длительность импульса заполнения t_1 является неизменной величиной, и получим выражение 9.

$$A = C_A K_{BS} K_{LS}. \quad (8)$$

$$S(\tau, A, F_0) = A \phi(\tau, F_0) \quad (9)$$

2.3 Модель, учитывающая нелинейность и неэкспоненциальность

Для одновременного учёта нелинейности аппаратного тракта и неэкспоненциальности сигнала релаксации, связанной с присутствием нескольких

экспоненциальных составляющих, в модель вводят коэффициент нелинейности-неэкспоненциальности p , после чего выражение 9 приобретает вид выражения 10 [6].

$$S(\tau, A, F_0, p) = A [\phi(\tau, F_0)]^p. \quad (10)$$

В случае моноэкспоненциального сигнала релаксации и линейного измерительного тракта коэффициент $p = 1$, в остальных случаях он отклоняется от 1. Данный коэффициент эмпирический и не имеет физического смысла. Он комплексно учитывает нелинейность и неэкспоненциальность, что позволяет повысить точность моделирования, но он не позволяет дать ответ на вопрос какой именно из этих двух факторов имел место.

2.4 Модель многоэкспоненциального частотного скана

Если предположить, что сигнал релаксации ёмкости состоит из нескольких экспоненциальных составляющих и определяется выражением 3, то опираясь на выражения 3, 4, 6 и 9, можно сделать вывод, что частотный скан, созданный таким сигналом релаксации ёмкости определяется выражением 11.

$$Y = \sum_{i=1}^n A_i \phi(\tau_i, F_0), \quad (11)$$

где n – количество экспоненциальных составляющих в сигнале релаксации. Данная модель предполагает, что измерительный такт линейен, а сигнал релаксации создан ловушками носителей заряда, дающими линейчатый спектр, аналогичный спектру, представленному на рисунке 2.

3 Реализация моделей

В данном разделе приводится краткое описание программной реализации моделей, примеры результатов идентификации частотных сканов, рассматриваются некоторые методические вопросы.

3.1 Интерфейс и функционал

Модели оформлены в виде пакета модулей на языке программирования Python 3. Интерфейс моделей совместим с одной из самых популярных библиотек для машинного обучения `scikit-learn` [7]. Такое техническое решение позволяет использовать имеющиеся в библиотеке инструменты preprocessing данных, оптимизации параметров и оценки качества моделей, а также интегрировать модели в другое программное обеспечение.

Разработанные модели (выражения 9, 10, 11), выполняют две функции:

1. Вычисление частотного скана с заданными параметрами.
2. Идентификация параметров модели частотного скана по экспериментальным данным.

Идентификация параметров моделей реализована методом градиентного спуска. Имеется возможность вывода значений параметров модели на каждой итерации. Реализована автоматическая остановка идентификации при достижении заданной точности, определяемой как модуль разности между значениями функции потерь на текущей и предыдущей итерации.

Модели реализуют единообразный интерфейс. При инициализации каждая модель получает длительность импульса заполнения (параметр `filling_pulse`), которая считается неизменной при измерениях) и параметры алгоритма идентификации. Параметры алгоритма идентификации варьируются:

- модель идеального частотного скана и модель частотного скана с показателем нелинейности-неэкспоненциальности (выражения 9, 10):

fit_p_coef – если параметр принимает значение `True` (логическую единицу), тогда при идентификации параметров модели определяется показатель нелинейности-неэкспоненциальности p , иначе (`False` – логический ноль) p считается равным 1 и не изменяется при идентификации параметров модели;

learning_rate – скорость градиентного спуска;

n_iters – максимальное количество итераций при идентификации;

stop_val – точность модели, при достижении которой происходит остановка идентификации;

verbose – если параметр принимает значение True, то в стандартный поток вывода печатаются значения параметров модели на каждой итерации.

- модель многоэкспоненциального частотного скана (выражение 11):

n_exps – количество экспоненциальных составляющих в сигнале релаксации (значение n в выражении 11);

learning_rate – скорость градиентного спуска;

n_iters – максимальное количество итераций при идентификации;

stop_val – точность модели, при достижении которой происходит остановка идентификации;

verbose – если параметр принимает значение True, то в стандартный поток вывода печатаются значения параметров модели на каждой итерации.

При идентификации параметров модели идеального частотного скана определяется амплитуда и десятичный логарифм постоянной времени ($\log_{10}(\tau)$) сигнала релаксации. Причины замены постоянной времени её десятичным логарифмом объясняются в следующем разделе.

При идентификации модели с показателем нелинейности-неэкспоненциальности к амплитуде и десятичному логарифму постоянной времени сигнала релаксации добавляется значение показателя нелинейности-неэкспоненциальности p .

При идентификации многоэкспоненциальной модели определяется n пар амплитуда — десятичный логарифм постоянной времени сигнала релаксации. По паре для каждой из n экспоненциальных составляющих.

Параметры модели частотного скана могут быть не только идентифицированы, но и заданы пользователем.

Чтобы выполнить идентификацию параметров модели, нужно передать ей набор экспериментальных (тренировочных) данных, состоящих из вектора значений десятичных логарифмов частоты опорной функции ($\log_{10}(F_0)$) и соответствующего вектора значений сигнала DLTS – сигнала на выходе аппаратного тракта спектрометра DLS-82E. При идентификации параметров модели по умолчанию их начальные параметры выбираются случайным образом, однако могут быть и определены пользователем.

При идентификации модели значения её параметров на каждой итерации сохраняются в атрибуте `fit_results_`, что позволяет пользователю получить и проанализировать данные о работе алгоритма идентификации.

При вычислении частотного скана с заданными параметрами модели передаётся вектор значений десятичных логарифмов частоты опорной функции, для каждого из которых модель вычислит значение сигнала DLTS.

3.2 Алгоритм идентификации параметров моделей

Идентификация параметров модели производится методом градиентного спуска, при этом минимизируется среднеквадратическая ошибка между значениями, полученными в результате измерений, и результатами моделирования (выражение 12).

$$E = \frac{1}{m} \sum_{i=1}^m (y_i - y_i^*)^2, \quad (12)$$

где

y_i – значения, полученные в результате измерений,

y_i^* – значения, полученные в результате моделирования,

m – количество измерений.

При каждом обновлении постоянной времени сигнала релаксации (как при идентификации модели, на каждой итерации, так и при обновлении постоянной времени пользователем) вычисляется масштабный множитель $M(\tau, F_0, t_1)$, определяемый выражением 7. Таким образом модель всякий раз вычисляет значение $\max \left[\tau F_0 e^{-\frac{0.05}{\tau F_0}} \left(1 - e^{\frac{t_1 F_0 - 0.45}{\tau F_0}} - e^{-\frac{0.5}{\tau F_0}} + e^{\frac{t_1 F_0 - 0.95}{\tau F_0}} \right) \right]$. В текущей реализации данный максимум вычисляется приблизительно с помощью градиентного спуска. Не смотря на то, что применяемый итеративный алгоритм находит максимум очень быстро, эти вычисления занимают довольно много времени при идентификации моделей, потому что выполняются на каждой итерации (при каждом обновлении постоянной времени τ). В случае многоэкспоненциальной модели данный процесс повторяется еще и для каждой экспоненциальной составляющей, что заметно снижает скорость идентификации модели при значениях параметра `n_exps` больше 10. Таким образом в будущем вычисления можно будет оптимизировать за счёт одного из следующих решений:

1. найти точное аналитическое выражение для расчёта $M(\tau, F_0, t_1)$,
2. найти для $M(\tau, F_0, t_1)$ аппроксимирующую функцию, позволяющую вычислять масштабный множитель с приемлимой точностью.

В моделях градиентный спуск реализован при помощи библиотеки TensorFlow [8], главное преимущество которой в том, что она реализует алгоритм автоматического дифференцирования на графе вычислений. Таким образом, при расчёте градиентов сначала производные берутся символьно (точно), а затем вычисляются их значения, поэтому точность вычисления градиента ограничена только разрядностью чисел [3, 10, 8].

Во всех моделях используется алгоритм идентификации с постоянной и одинаковой для всех параметров скоростью градиентного спуска, поэтому

для ускорения идентификации параметров моделей и улучшения сходимости алгоритма необходима нормализация (приведение к единому масштабу) тренировочных данных.

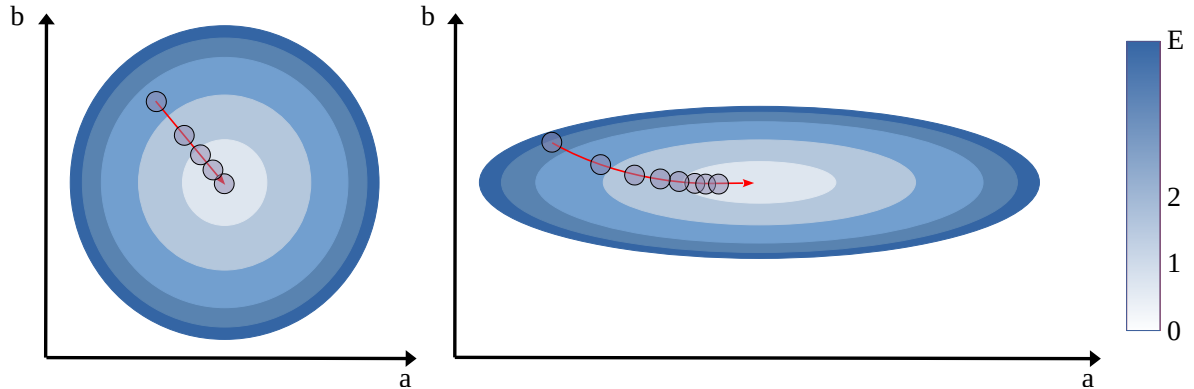


Рис. 3: Градиентный спуск с нормализацией тренировочных данных (график слева) и без неё (график справа) для модели с двумя параметрами. a и b – параметры модели, E – функция потерь.

Рассмотрим примеры на рисунке 3. В идеальном случае функция потерь (среднеквадратическая ошибка – выражение 12) должна иметь форму «симметричной чаши», а изменение обоих параметров должно вносить одинаковый вклад в значение градиента. В таком случае алгоритм градиентного спуска достигнет минимума функции потерь кратчайшим путём, как на левом графике рисунка 3. Если же тренировочные данные имеют очень разные масштабы, функция потерь (график справа на рисунке 3) будет иметь форму «вытянутой чаши», а алгоритм идентификации будет долго подстраивать одни из параметров без существенного изменения функции потерь [3]. Чтобы минимизировать этот эффект, тренировочные данные нормализуют перед идентификацией параметров модели. После идентификации параметры модели и тренировочные данные можно вернуть в исходный масштаб.

Для разработанных моделей тренировочные данные имеют существенно отличающиеся масштабы: значение сигнала DLTS находятся в диапазоне долей пикофарад, а диапазон изменения частоты опорной функции спектрометра покрывает промежуток от 1 Гц до 2500 кГц. Таким образом, желательно выполнять нормализацию исходных данных перед идентификацией параметров модели. Кроме этого, изменение параметров модели (постоянной времени и амплитуды) вносят существенно различающийся вклад в градиент функции потерь. Если снова обратиться к выражениям 6 и 9, можно заметить, что значение функции потерь пропорционально первой степени амплитуды A сигнала релаксации и экспоненте, возведённой в степень скорости экспоненциального спада, то есть $e^{\frac{1}{\tau}}$. Поэтому на вход моделей при идентификации подаётся десятичный логарифм частоты опор-

ной функции ($\log_{10}(F_0)$), а идентификация производится не по постоянной времени τ , а по десятичному логарифму постоянной времени $\log_{10}(\tau)$. На рисунке 4 показан пример идентификации идеального частотного скана. «Путь» параметров модели в процессе идентификации показывает красная линия с маркерами, а изолинии и градиент – значения функции потерь.

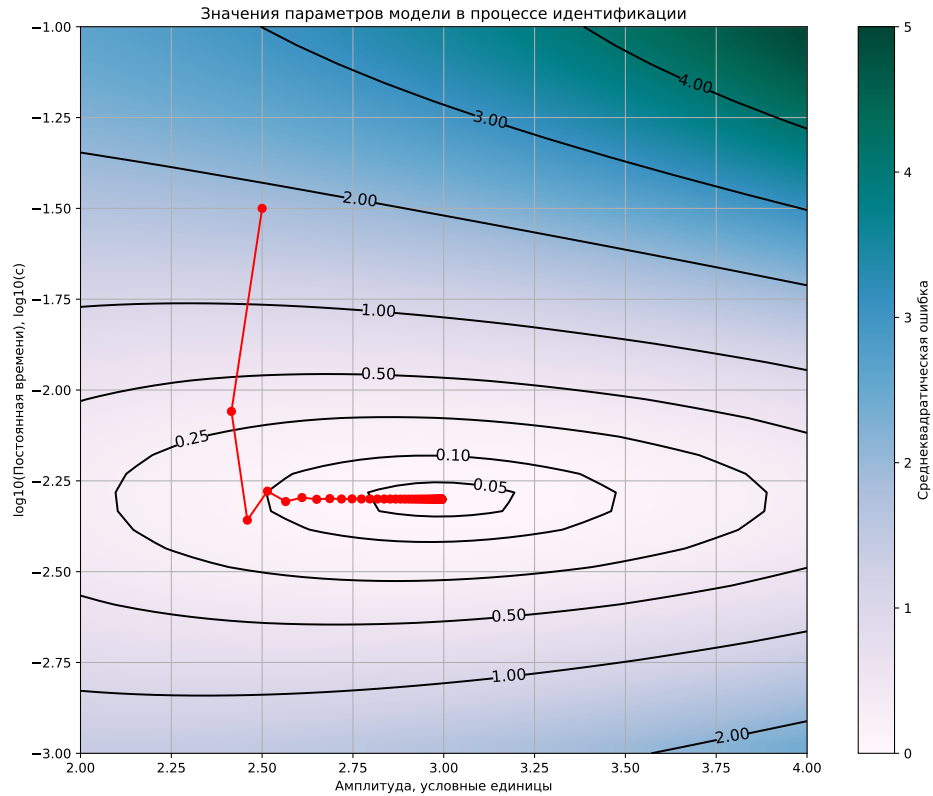


Рис. 4: «Путь» значений параметров при идентификации.

На рисунке 4 форма изолиний отличается от концентрических окружностей, а «путь» параметров не похож на прямую линию, что говорит о том, что возможна дальнейшая оптимизация алгоритма идентификации, однако, названные выше меры позволили значительно смягчить эффект от различных масштабов в экспериментальных данных и в параметрах модели. Также возможно использование одного из алгоритмов адаптивного градиентного спуска, то есть алгоритмов изменяющих скорость по мере приближения к оптимальным значениям параметров. Такое решение не изменит форму функции потерь, но оптимизирует «путь» и количество итераций.

Рисунки 5 и 6 показывают примеры результатов идентификации модели идеального частотного скана и модели многоэкспоненциального частотного скана соответственно на рассчитанных тренировочных данных.

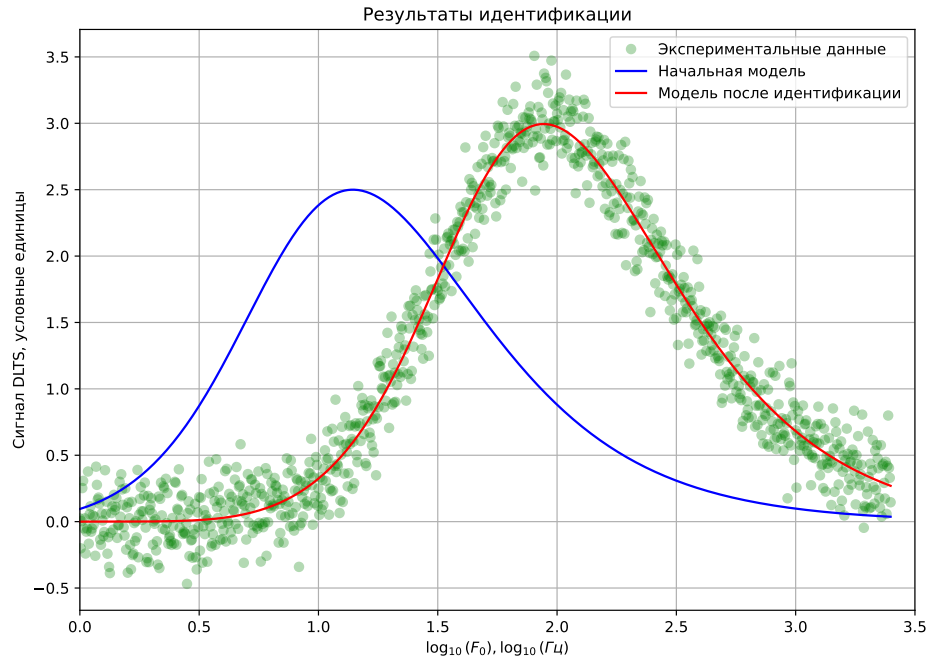


Рис. 5: Пример результата идентификации модели идеального частотного скана.

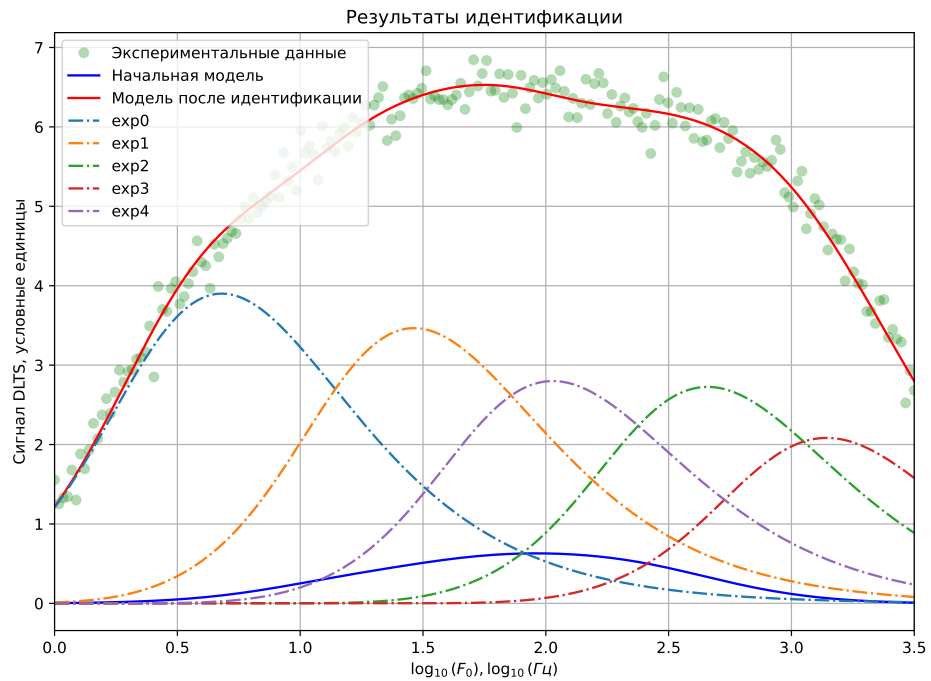


Рис. 6: Пример результата идентификации модели многоэкспоненциального частотного скана. Штрихпунктирными линиями показаны частотные сканы для отдельных экспоненциальных составляющих сигнала релаксации.

3.3 Оценка качества модели и оптимизация параметров

Вычисление корня из среднеквадратической ошибки (англ. root-mean-square error – RMSE) между экспериментальными данными и данными, полученными с помощью моделирования, – один из самых известных и распространенных способов оценки качества регрессии данных разработанной моделью. Однако, одного этого показателя часто бывает недостаточно, более того, **оценка этого показателя только на тренировочных данных** при разработке модели (**особенно эмпирической модели**) – методическая ошибка [3, 2]. Данные утверждения обусловлены следующим:

- сам по себе корень из среднеквадратической ошибки не отвечает на вопрос о «фундаментальном» соответствии данных модели, а также не показывает какова вероятность того, что полученное значение – результат случайного совпадения;
- оценивая данную метрику только на тренировочных данных, есть риск получить слишком оптимистическую оценку и столкнуться с проблемой, называемой «оверфитинг» (от англ. overfitting) [3, 2, 10].

Для начала, для иллюстрации приведённых аргументов обратимся к набору данных, известному под названием «Квартет Энскомба». Это набор данных был создан математиком Фрэнком Энскомбом специально для того, чтобы показать важность визуализации данных в графиках. Особенность этого набора в том, что он состоит из четырёх групп данных, имеющих одинаковые статистические характеристики, но абсолютно разные графики [9, 1]. Графики приведены на рисунке 7, статистические характеристики – в таблице 1.

Таблица 1: Статистические свойства наборов данных в Квартете Энскомба

Характеристика	Значение
Среднее значение переменной x	9.0
Дисперсия переменной x	10.0
Среднее значение переменной y	7.5
Дисперсия переменной y	3.75
Корреляция между переменными x и y	0.816
Уравнение линейной регрессии	$y = 3 + 0.5x$
Коэффициент детерминации линейной регрессии (R^2)	0.67

Что примечательно, все 4 подгруппы данных имеют одно и то же уравнение линейной регрессии (см. таблицу 1) и, как показано в таблице 2, очень близкие значения корня из среднеквадратической ошибки.

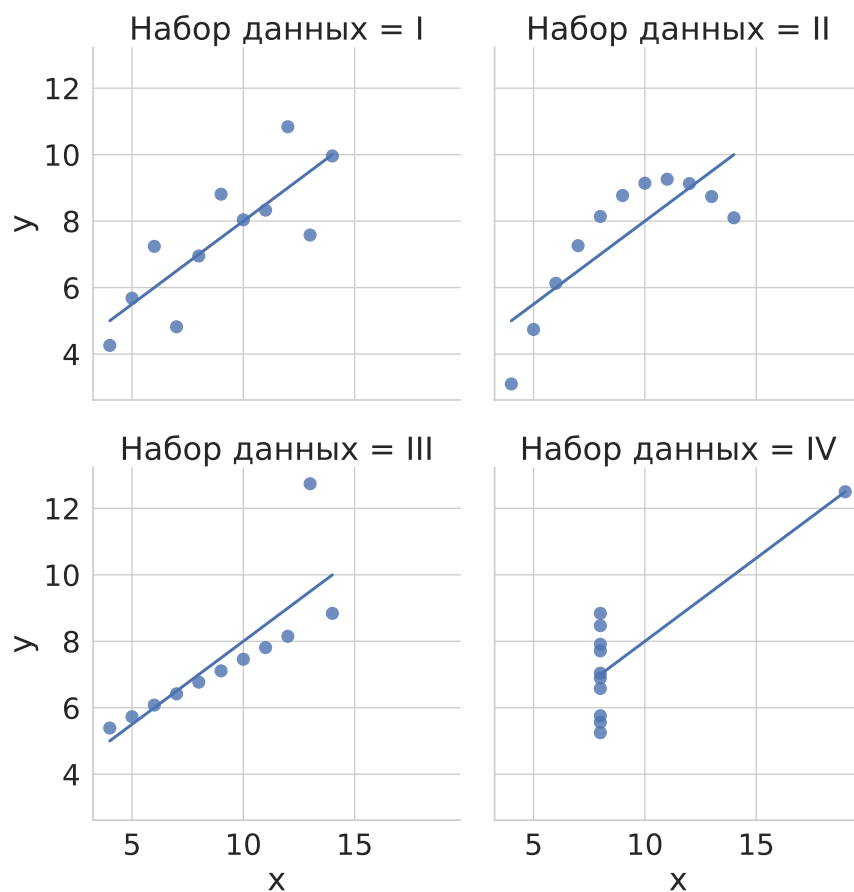


Рис. 7: Квартет Энскомба.

Таблица 2: Значения корней из среднеквадратической ошибки

Набор данных	I	II	III	IV
Значение корня среднеквадратической ошибки (RMSE)	1.119	1.119	1.118	1.118

Таблица 3: Квартет Энскомба

I		II		III		IV	
x	y	x	y	x	y	x	y
10,0	8,04	10,0	9,14	10,0	7,46	8,0	6,58
8,0	6,95	8,0	8,14	8,0	6,77	8,0	5,76
13,0	7,58	13,0	8,74	13,0	12,74	8,0	7,71
9,0	8,81	9,0	8,77	9,0	7,11	8,0	8,84
11,0	8,33	11,0	9,26	11,0	7,81	8,0	8,47
14,0	9,96	14,0	8,10	14,0	8,84	8,0	7,04
6,0	7,24	6,0	6,13	6,0	6,08	8,0	5,25
4,0	4,26	4,0	3,10	4,0	5,39	19,0	12,50
12,0	10,84	12,0	9,13	12,0	8,15	8,0	5,56
7,0	4,82	7,0	7,26	7,0	6,42	8,0	7,91
5,0	5,68	5,0	4,74	5,0	5,73	8,0	6,89

В таблице 3 для справки приведены данные из «Квартета Энскомба».

Квартет Энскомба – не единственный набор данных, имеющий такие свойства, но, пожалуй, самый известный [9].

Конечно, достаточно просто построить график и попробовать идентифицировать разные виды моделей на одном и том же наборе данных, чтобы исключить такую ошибку и выбрать наиболее подходящие модели.

Далее, обратимся к проблеме, называемой «оверфитинг». Как уже было сказано, оценка RMSE только на тренировочных данных – методическая ошибка. Если исследователь на этапе выбора модели (например, выбора между линейной и полиномиальной регрессией) или на этапе выбора параметров модели (параметров, не зависящих от данных, например степени полинома) выполняет оценку данной метрики на том же наборе, на котором выполняет идентификацию параметров модели, появляется риск настроить модель таким образом, что она будет демонстрировать превосходные результаты на тренировочных данных (показатель RMSE может буквально равняться нулю), но будет абсолютно бесполезна на новых не входивших в тренировочный набор. Проще говоря, появляется высокий риск выбрать слишком сложную модель, которая примет случайные отклонения в тренировочных данных за закономерность, либо же излишне адаптировать модель под тренировочный набор [3, 10].

Для того чтобы проиллюстрировать эту проблему, подготовим небольшой набор данных, имитирующий результаты измерения величин x и $y(x)$, связанных выражением 13.

$$y(x) = 1 + 0.5x + \epsilon, \quad (13)$$

где ϵ – нормально распределённая случайная величина со средним, значением равным 0, и среднеквадратическим отклонением равным 0.4. В данном случае, зависимость $y(x)$ «фундаментально» линейна. Расчитаем для данной зависимости 10 точек в диапазоне от $x = 0.5$ до $x = 3.5$, затем, выполним регрессию полученного набора данных линейной функцией и алгебраическим полиномом десятой степени, в конце, оценим корень из среднеквадратической ошибки для обеих моделей. Рисунок 8 демонстрирует результаты регрессий.

После идентификации для модели линейной регрессии получились следующие параметры:

a – коэффициент, характеризующий наклон, равен приблизительно 0.43;

b – коэффициент, характеризующий смещение, равен приблизительно 0.92;

RMSE – корень из среднеквадратической ошибки, равный приблизительно 0.48.

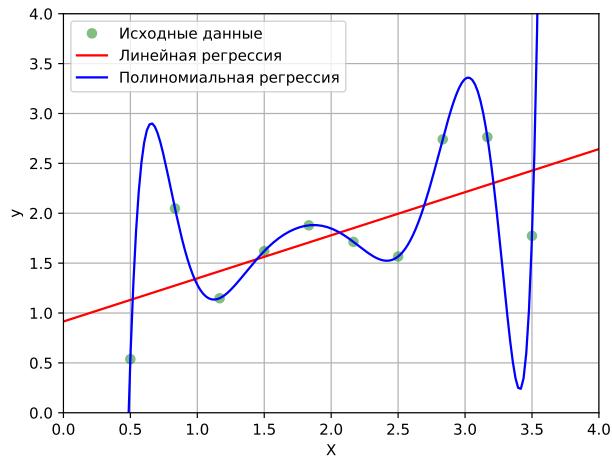


Рис. 8: Иллюстрация оверфиттинга в случае полиномиальной модели. Использовался полином десятой степени.

Как и ожидалось, полученные значения очень близки к параметрам, использованным при расчёте исходного набора данных (выражение 13).

Не будем приводить все коэффициенты полиномиальной регрессии, корень из среднеквадратической ошибки для данной модели получился равным $4.24 \cdot 10^{-11}$, на рисунке 8 видно, что график полиномиальной регрессии проходит через все точки исходных данных, при этом крайне маловероятно, что данная модель будет адекватно предсказывать значения $y(x)$, на значениях x , которых не было в тренировочном наборе.

Для сравнения попробуем идентифицировать многоэкспоненциальную модель на том же наборе данных, примем параметр `n_exps` равным 10. Результаты идентификации такой модели представлены на рисунке 9.

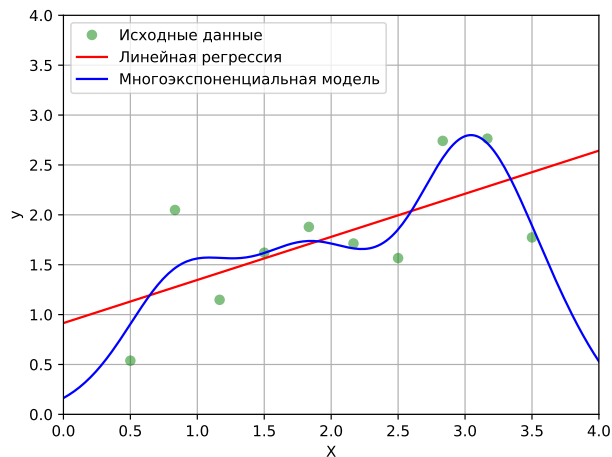


Рис. 9: Иллюстрация оверфиттинга в случае многоэкспоненциальной модели. Использовалась модель с параметром `n_exps` = 10.

Для полученной многоэкспоненциальной модели с параметром `n_exps`, равным 10, корень из среднеквадратической ошибки получается приблизительно равным 0.28, что значительно больше, чем в случае полиномиальной

регрессии, и всё еще меньше, чем в случае линейной регрессии. При этом, хочется подчеркнуть, что данный набор точек является линейной функцией с добавлением нормально распределённой случайной составляющей (выражение 13).

Существует несколько способов борьбы с оверфитингом, ниже приведены некоторые из них:

- собрать больше данных (для того, чтобы предыдущие примеры были показательны, был намеренно сделан набор из 10 точек, получить аналогичный эффект на наборе из 1000 точек сложнее);
- использовать модель попроще (например, выбрать полином меньшей степени);
- использовать регуляризацию;
- оценивать качество модели не на тренировочном наборе, а на отдельном валидационном наборе (англ. validation set), случайно отобранной контрольной группе;
- использовать приём, называемый кроссвалидацией (англ. cross-validation) [3, 2].

В задачах машинного обучения для того, чтобы оценить как точно модель будет предсказывать целевые значения на новых данных, исходный набор разделяют на две группы: тренировочную и тестовую. На тренировочном наборе данных проводится идентификация модели, а на тестовом только оценивают качество самой финальной модели. Таким образом, применительно к задаче регрессии, ожидается, что на новых, невиданных в тренировочный набор данных модель будет демонстрировать среднеквадратическую ошибку, близкую к полученной на тестовом наборе.

Если у модели есть параметр, не зависящий от данных, например, степень алгебраического полинома или количество экспоненциальных составляющих многоэкспоненциальной модели (параметр `p_exps`), и необходимо найти оптимальное значение данного параметра, то из тренировочного набора выделяют еще один набор данных – валидационный набор, на котором оценивают модель при каждом значении подстраемового параметра, в итоге выбор останавливают на параметре с которым модель показала лучший результат (наименьшее значение среднеквадратической ошибки). При этом модель не тренируют ни на валидационном наборе, ни на тестовом, то есть тренировочный набор уменьшается.

Подход с применением валидационного набора позволяет исключить оверфитинг и прогнозировать эффективность модели. Однако, данный приём приводит к сокращению тренировочных данных, что может вызывать проблемы в случаях работы с наборами данных, содержащими небольшое

количество наблюдений. В таких случаях прибегают к технике, называемой кроссвалидацией [3, 2].

При использовании кроссвалидации тестовый набор всё также применяется только для прогнозирования точности будущих предсказаний модели, на нем никогда не производится обучение, и не производится оценка модели во время настройки её параметров. Тренировочный набор разбивается на k подгрупп (в случае использования алгоритма «k-fold cross-validation»), затем модель в цикле из k шагов обучают на $k-1$ подгруппах, а на оставшейся оценивают модель (в случае регрессии считают среднеквадратическую ошибку), при этом на каждом шаге цикла для оценки модели берут новую подгруппу. В конце, считают среднее значение метрики качества модели, полученной на каждом шаге цикла, это и будет результат кроссвалидации. Таким образом, модель каждый раз оценивается на данных, которые она «не видела» во время обучения, за счёт этого получается исключить использование валидационного набора, однако приходится тренировать модель k раз [2].

Рисунок 10 демонстрирует схему работы с данными, при применении кроссвалидации. Зелёным отмечены наборы данных, на которых проводится обучение модели, синим – наборы, на которых выполняется оценка среднеквадратической ошибки [2].

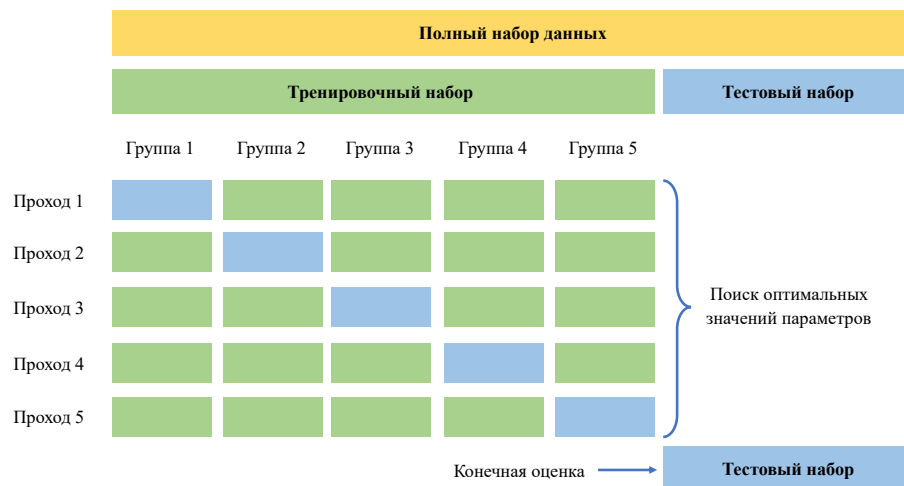


Рис. 10: Иллюстрация кроссвалидации.

При оптимизации параметров модели, не зависящих от данных (по сути, параметров алгоритма идентификации) кроссвалидацию выполняют для каждого комплекта значений параметров и выбирают комплект, с которым модель показала лучший результат (такая техника оптимизации в англоязычной литературе называется «grid search», то есть поиск по сетке). В конце, модель с победившим набором параметров еще раз идентифицируют на полном тренировочном наборе данных, после этого, если разработка модели закончена можно выполнить оценку на тестовом наборе [2].

Таким образом, при оценке модели не стоит ограничиваться только расчётом среднеквадратической ошибки, целесообразно строить графики, применять дополнительные приёмы и метрики, которые на самом деле не ограничиваются названными в данном разделе.

В следующем разделе мы апробуем разработанные модели на экспериментальных данных и сравним результаты для разных моделей.

4 Апробация моделей на экспериментальных данных

Для апробации разработанных моделей выберем результаты измерений на интегральной схеме 1564ЛЕ1, состоящие из трёх частотных сканов, полученных при температурах 263 К, 283 К, 303 К. К сожалению, полученные частотные сканы содержат относительно немного точек: 340, 340 и 34, соответственно. Поскольку требуется только идентификация параметров модели, а не вычисление сигнала DLTS по новым значениям частоты опорной функции, предлагается отказаться от оценки идентифицированной модели на тестовом наборе, и ограничиться только оценками корня из среднеквадратической ошибки (RMSE) на тренировочных данных и кроссвалидацией.

В данном разделе приводятся схема подключения образца, условия эксперимента, результаты измерений (частотные сканы), результаты идентификации частотных сканов разными моделями и их сравнение, а также анализ графиков Аррениуса, построенных по вычисленным параметрам моделей.

4.1 Экспериментальные данные

Цель исследования: регистрация частотных сканов интегральной схемы 1564ЛЕ1 №1 в двухполюсном включении при разных температурах. Схема подключения с эквивалентным диодом показана на рисунке 11.

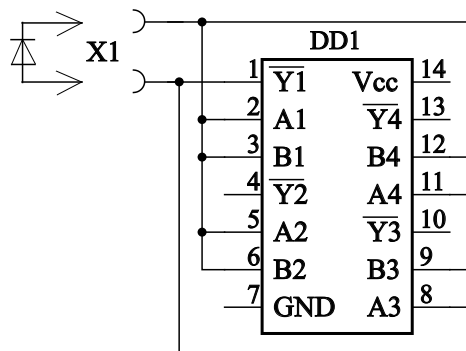


Рис. 11: Схема двухполюсного включения ИС 1564ЛЕ1.

Условия, одинаковые для всех сканов перечислены в таблице 4. Графики с результатами измерений представлены на рисунках 12, 13 и 14.

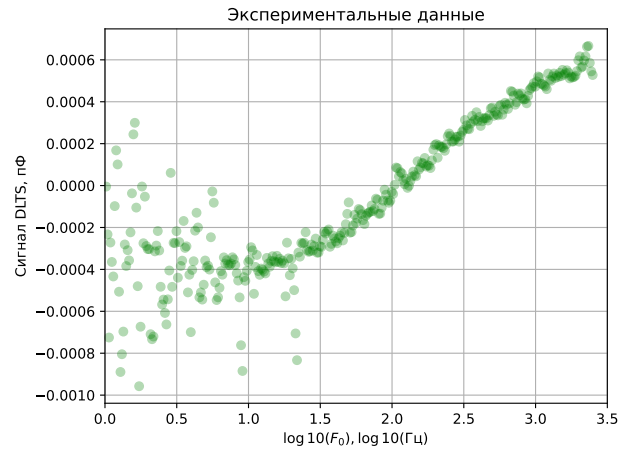


Рис. 12: Частотный скан, полученный при температуре 263 К.

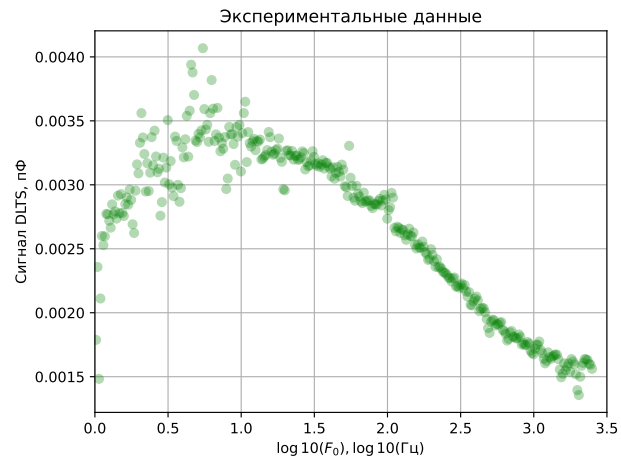


Рис. 13: Частотный скан, полученный при температуре 283 К.

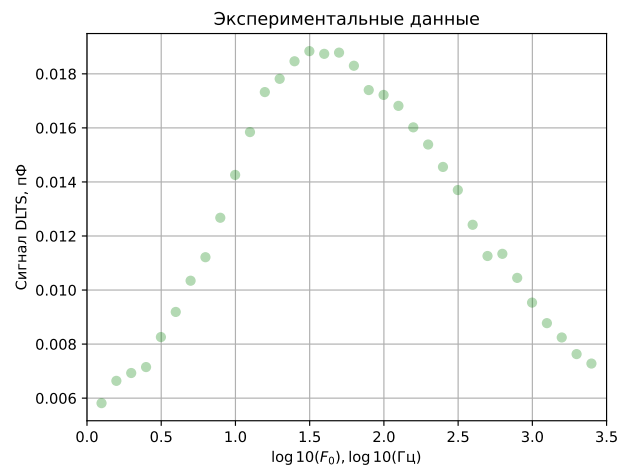


Рис. 14: Частотный скан, полученный при температуре 303 К.

Таблица 4: Условия, одинаковые для всех сканов.

Параметр	Значение
Образцы	1564ЛЕ1 №1
Длительность импульса заполнения, мкс	10
U_1	-4 В
U_r	-5 В
Автоматическая компенсация моста	включена
Начальная частота сканирования, Гц	1
Конечная частота сканирования, Гц	2500
Направление сканирования	вниз по частоте
Интервал между измерениями, с	3.5
Постоянная интегрирования, с	3

4.2 Частотный скан при температуре 263 К

4.2.1 Моноэкспоненциальная модель с показателем p

Как показано на рисунке 12, частотный скан имеет два пика: положительный и отрицательный, поэтому сначала выполним идентификацию модели с показателем p с начальными параметрами близкими к положительному пику. Результат идентификации представлен на рисунке 15.

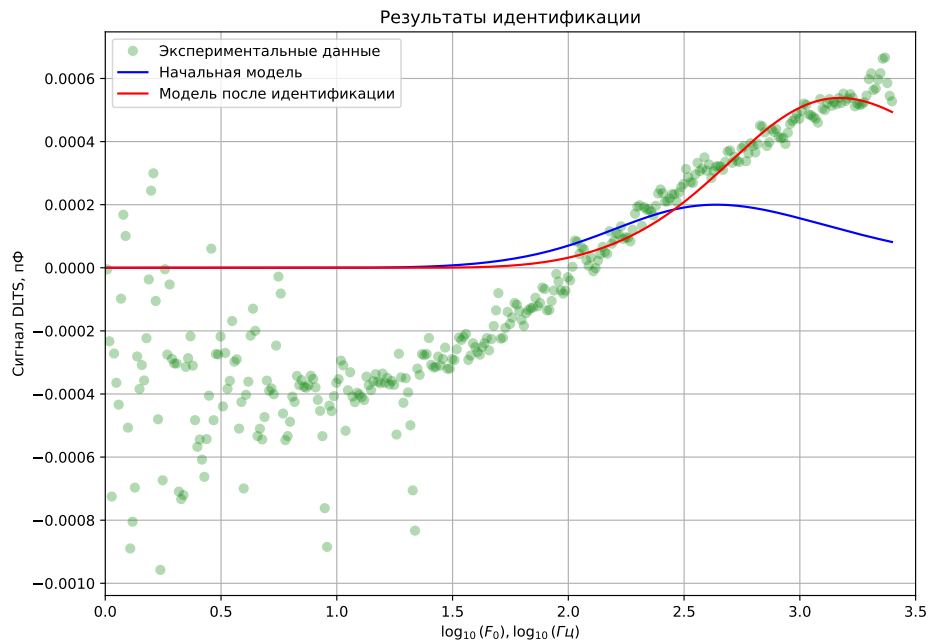


Рис. 15: Результат идентификации положительного пика частотного скана при $T = 263K$.

Для того, чтобы убедиться, что алгоритм идентификации достиг минимума функции потерь, построим график значений среднеквадратической

ошибки на каждой итерации, график нормируем относительно максимального значения метрики. Получившийся график представлен на рисунке 16.

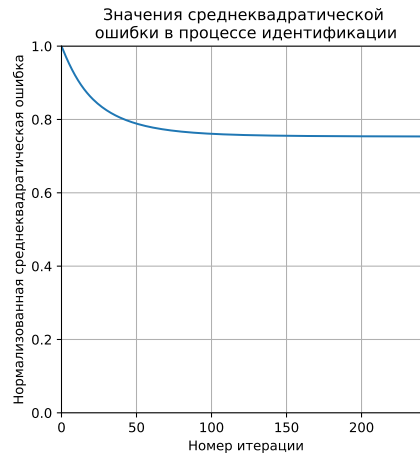


Рис. 16: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

График среднеквадратической ошибки выходит на «плато», что свидетельствует о том, что решение сошлось, однако довольно высокое значение среднеквадратической ошибки может указывать на то, что алгоритм нашёл локальный минимум функции потерь.

Построим график отклонений результатов моделирования от экспериментальных данных (рисунок 17), чтобы оценить систематическую ошибку.

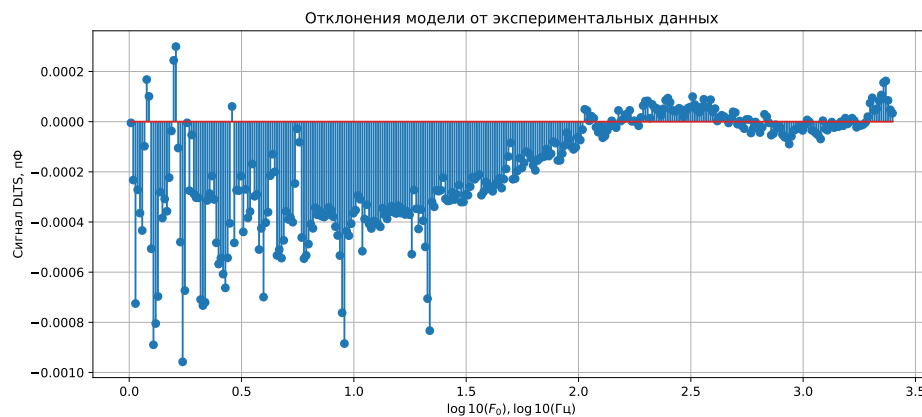


Рис. 17: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

Рисунки 15 и 17 демонстрируют, что модель не охватывает половину частотного скана с отрицательными значениями сигнала DLTS, что закономерно для однопиковой модели.

Построим гистограмму распределения отклонений результатов моделирования от экспериментальных данных. В идеальном случае данное распределение должно стремиться к распределению погрешностей измерительного прибора и быть близко к нормальному распределению, в данном случае

(рисунок 18) это не так. На рисунке 18 сплошной линией показан результат сглаживания гистограммы с гауссовым усредняющим ядром (в англоязычной литературе данный алгоритм называется «kernel density estimation»).

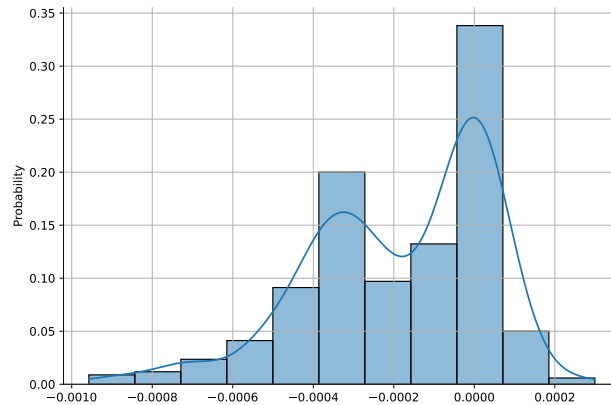


Рис. 18: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

В таблице 5 приведём параметры идентифицированной модели и результаты оценок точности моделирования.

Таблица 5: Результаты идентификации модели

Параметр	Значение
Постоянна времени τ , с	0,000286
Амплитуда, пФ	0.000538
Показатель p	0.802
Результат кросвалидации (RMSE), пФ	0.000290
RMSE, пФ	0.000291

Построим полученный спектр частотного скана (рисунок 19), чтобы предствить результаты в более наглядной форме.

Идентифицируем эту же модель с начальными параметрами близкими к параметрам отрицательного пика. На рисунке 20 представлен график с результатами идентификации.

На рисунке 21 представлен график значений нормированной среднеквадратической ошибки на каждой итерации. На итерациях с номерами выше 200 значения среднеквадратической ошибки практически не изменяются, очевидно, алгоритм нашёл минимум.

На гафике отклонений результатов моделирования от экспериментальных данных (рисунок 22) видно, что теперь модель «не охватывает» диапазон высоких частот опорной функции, где располагается положительный пик.

Гистограмма отклонений (рисунок 23) далека от нормального распределения. Сплошной линией показан результат сглаживания гистограммы.

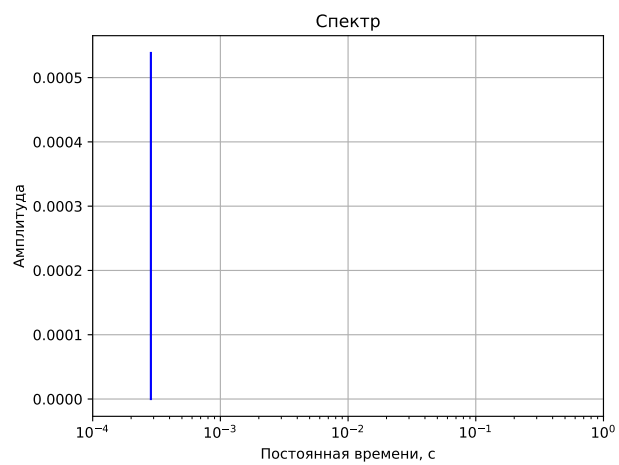


Рис. 19: Спектр сигнала релаксации ёмкости. Амплитуда указана в пФ.

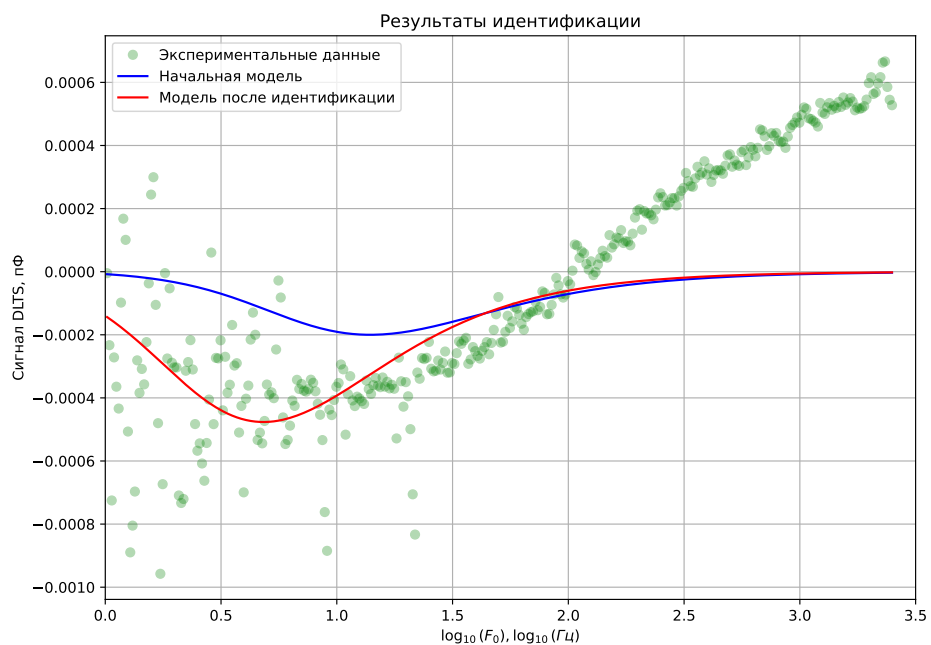


Рис. 20: Результат идентификации отрицательного пика частотного скана при $T = 263K$.

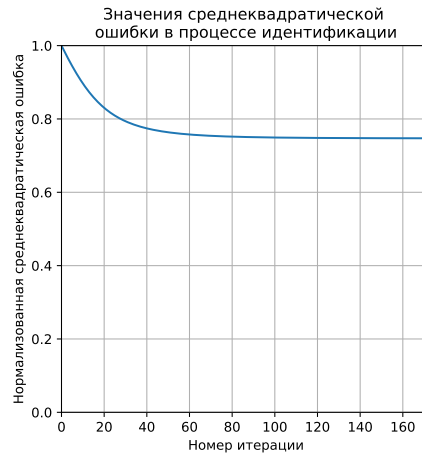


Рис. 21: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.



Рис. 22: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

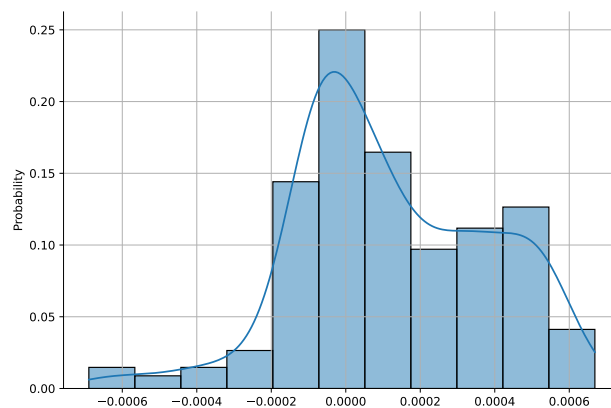


Рис. 23: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

В таблице 6 приведём параметры идентифицированной модели и результаты оценок точности моделирования.

Таблица 6: Результаты идентификации модели

Параметр	Значение
Постоянная времени τ , с	0.091434
Амплитуда, пФ	-0.000476
Показатель p	1.036
Результат кроссвалидации (RMSE), пФ	0.000287
RMSE, пФ	0.000286

Построим полученный спектр частотного скана (рисунок 24), чтобы представить результаты в более наглядной форме.

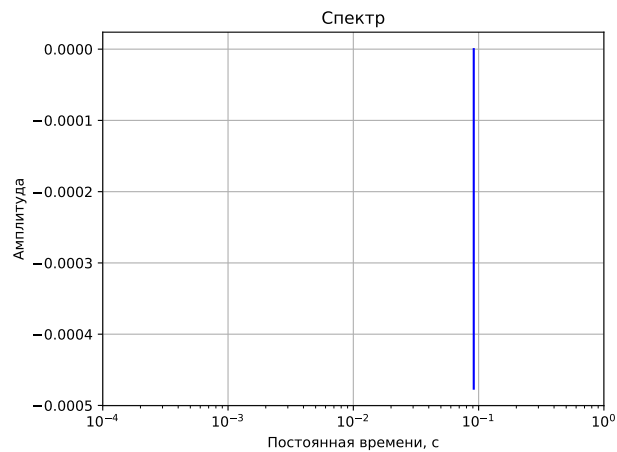


Рис. 24: Спектр сигнала релаксации ёмкости.

4.2.2 Моноэкспоненциальная модель с показателем $p = 1$

Для сравнения, на тех же данных выполним идентификацию модели идеального частотного скана. В этот раз инициализацию модели выполним случайными значениями. Результаты идентификации представлены на рисунке 25.

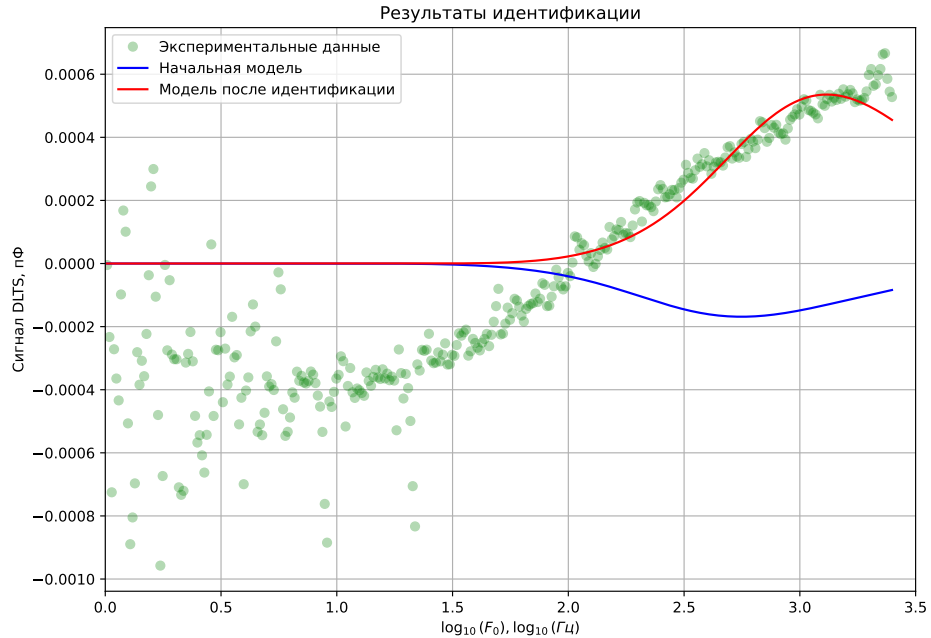


Рис. 25: Результат идентификации модели частотного скана при $T = 263K$.

В данном случае алгоритм идентификации нашёл локальный минимум, соответствующий положительному пику. Полученные результаты качественно похожи на результаты идентификации из предыдущего раздела. На рисунках 26, 27, 28 показаны график значений среднеквадратической ошибки на каждой итерации, график отклонений результатов моделирования от экспериментальных данных и гистограмма отклонений соответственно.

В таблице 7 приведены параметры идентифицированной модели и результаты оценок точности моделирования.

Таблица 7: Результаты идентификации модели

Параметр	Значение
Постоянная времени τ , с	0.000327
Амплитуда, пФ	0.000536
Показатель p	1.0
Результат кроссвалидации (RMSE), пФ	0.000279
RMSE, пФ	0.000291

На рисунке 29 приведён полученный спектр частотного скана.

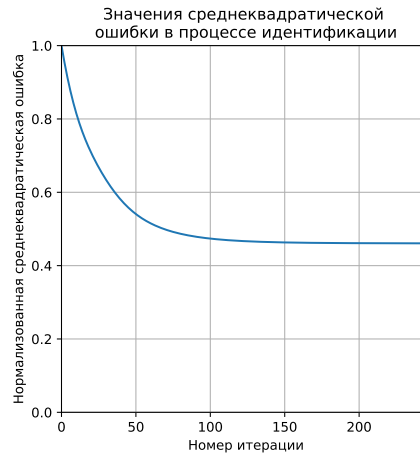


Рис. 26: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.



Рис. 27: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

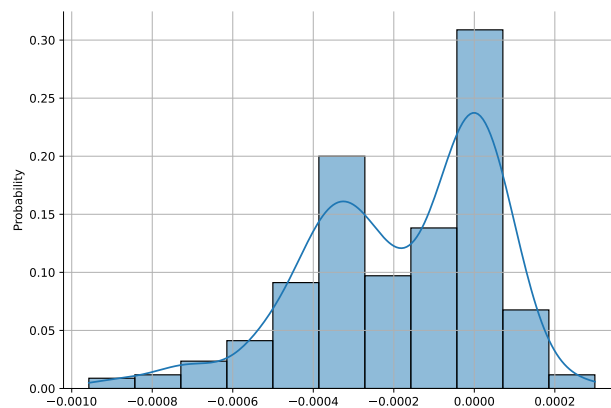


Рис. 28: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

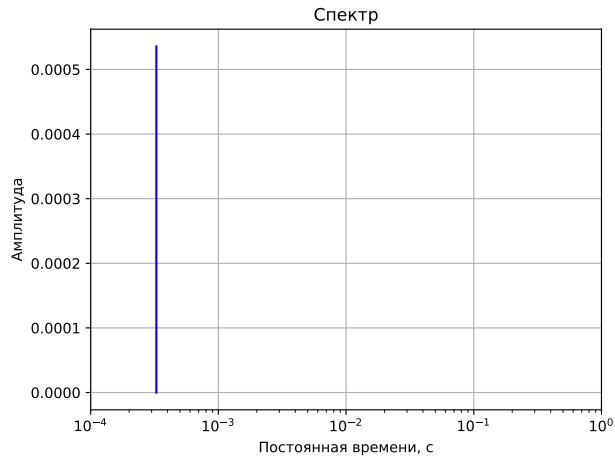


Рис. 29: Спектр сигнала релаксации ёмкости.

4.2.3 Многоэкспоненциальная модель с $n_exprs > 1$

Выполним идентификацию многоэкспоненциальной модели. Особенность данной модели в том, что у неё есть параметр, независящий от данных – количество экспоненциальных составляющих. По скольку у нас нет гипотезы о количестве экспоненциальных составляющих в релаксационном сигнале, воспользуемся приёмом, описанном в разделе 3.3: для моделей с значениями параметра n_exprs равными 5, 6, 7, 8, 9 и 10 будем выполнять кроссвалидацию, выберем ту, которая покажет лучший результат. В таблице 8 предсталены полученные результаты. Лучшей оказалась модель с параметром $n_exprs = 8$, то есть имеющая 8 экспоненциальных составляющих.

Таблица 8: Результаты кроссвалидации для разных значений параметра n_exprs

Значение параметра n_exprs	Результат кроссвалидации
5	0.0001373
6	0.0001361
7	0.0001347
8	0.0001345
9	0.0001353
10	0.0001349

Результаты идентификации модели с восемью экспоненциальными составляющими предствалены на рисунке 30. В данном случае, модель воспроизводит как положительный так и отрицательный пик.

По графику нормированной среднеквадратической ошибки, показанном на рисунке 31, видно, что алгоритм идентификации остановился из-за достижения максимального количества итераций. При этом, форма графика позволяет предположить, что полученные параметры модели близки к оптимальным.

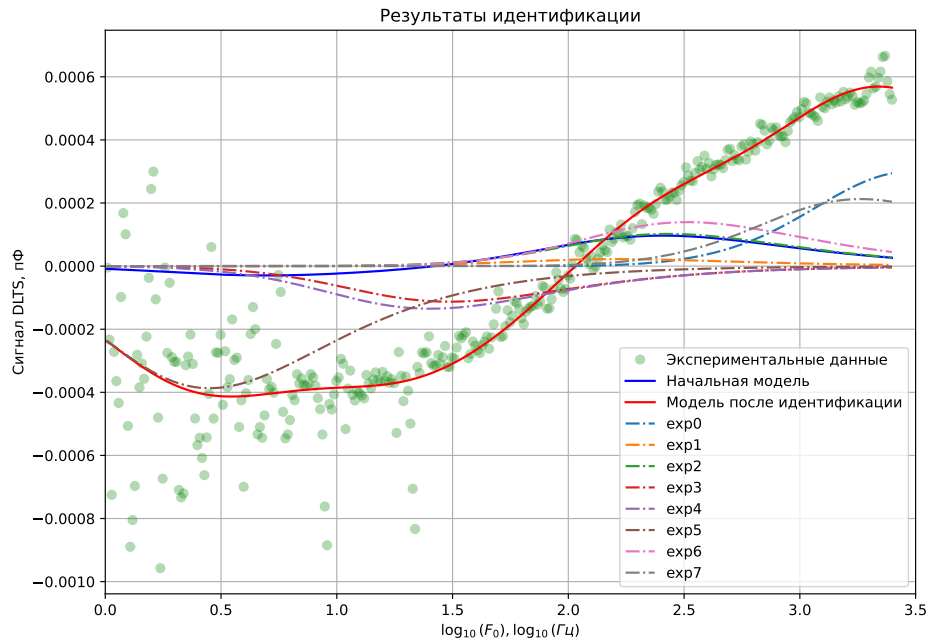


Рис. 30: Результат идентификации многоэкспоненциальной моделью частотного скана при $T = 263K$.

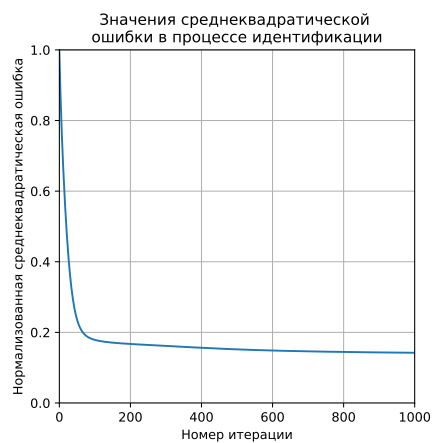


Рис. 31: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

На рисунке 32 показан график отклонений результатов моделирования от экспериментальных данных. В отличие от двух предыдущих моделей, многоэкспоненциальная модель хорошо воспроизводит как положительный, так и отрицательный пик. Увеличение разброса в области низких частот опорной функции связано с особенностями работы аппаратного интегратора спектрометра.



Рис. 32: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

На рисунке 33 показана гистограмма отклонений данных, полученных на идентифицированной модели, от результатов измерений. Форма полученного распределения приближается к форме нормального распределения.

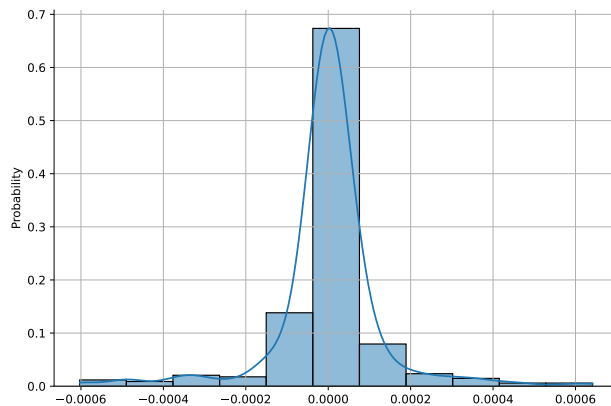


Рис. 33: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

В таблице 9 приведены параметры идентифицированной модели и результаты оценок точности моделирования, τ — постоянная времени, A — амплитуда.

На рисунке 34 приведён полученный спектр частотного скана.

Таблица 9: Результаты идентификации модели.

Параметр	Значение	Параметр	Значение
$\tau_1, \text{с}$	0.000132	$A_1, \text{пФ}$	0.000301
$\tau_2, \text{с}$	0.002744	$A_2, \text{пФ}$	0.000022
$\tau_3, \text{с}$	0.001629	$A_3, \text{пФ}$	0.000102
$\tau_4, \text{с}$	0.014347	$A_4, \text{пФ}$	-0.000113
$\tau_5, \text{с}$	0.017352	$A_5, \text{пФ}$	-0.000135
$\tau_6, \text{с}$	0.156241	$A_6, \text{пФ}$	-0.000387
$\tau_7, \text{с}$	0.001342	$A_7, \text{пФ}$	0.000139
$\tau_8, \text{с}$	0.000230	$A_8, \text{пФ}$	0.000213
Результат кроссвалидации (RMSE), пФ	0.000134		
RMSE, пФ	0.000132		

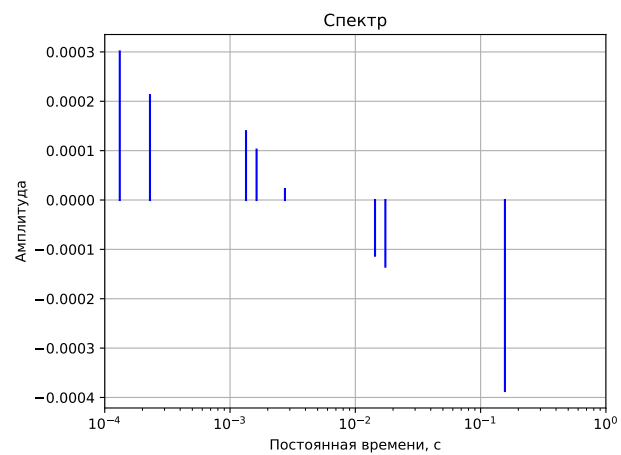


Рис. 34: Спектр сигнала релаксации ёмкости.

4.2.4 Сравнение результатов

В таблице 10 представлено сравнение результатов оценки точности рассмотренных моделей. В то время как модель идеального частотного скана (положительный пик) и модели с показателем p показывают приблизительно одинаковые результаты, многоэкспоненциальная модель показывает вдвое более высокую точность.

Таблица 10: Результаты идентификации моделей.

Модель	Результат кросвалидации, пФ	RMSE, пФ
Идеальный частотный скан	0.000279	0.000291
Модель с показателем p (положительный пик)	0.000290	0.000291
Модель с показателем p (отрицательный пик)	0.000287	0.000286
Многоэкспоненциальная модель	0.000134	0.000132

4.3 Частотный скан при температуре 283 К

4.3.1 Моноэкспоненциальная модель с показателем p

Результаты измерений при температуре 283 К приведены на рисунке 13. Полученный частотный скан имеет один пик, положительной полярности. Результаты идентификации модели с показателем p приведены на рисунке 35.

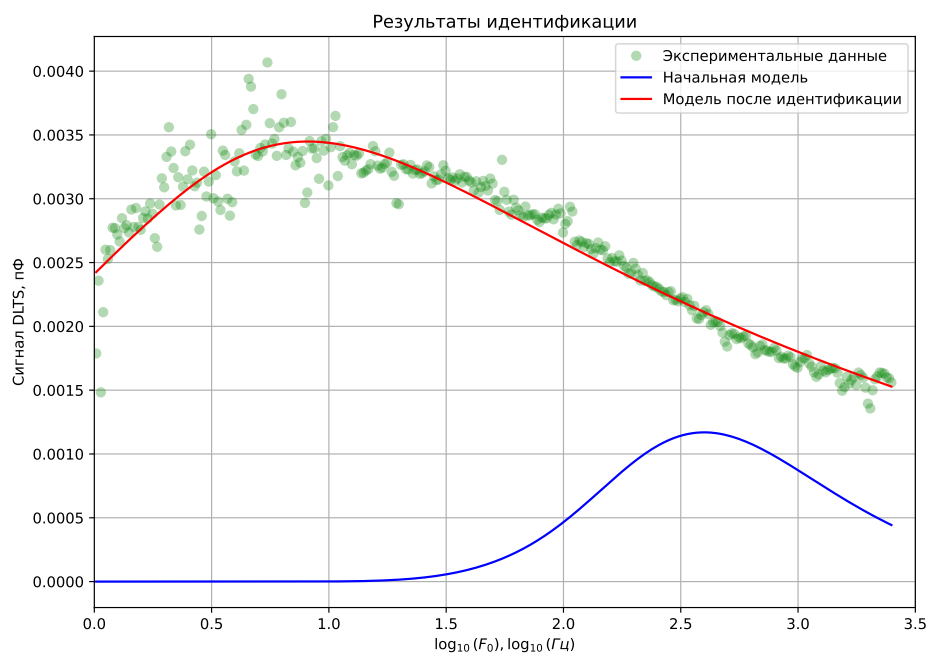


Рис. 35: Результат идентификации частотного скана при $T = 283K$.

График нормированной среднеквадратической ошибки подтверждает на рисунке 36, что алгоритм идентификации нашёл минимум функции потерь.

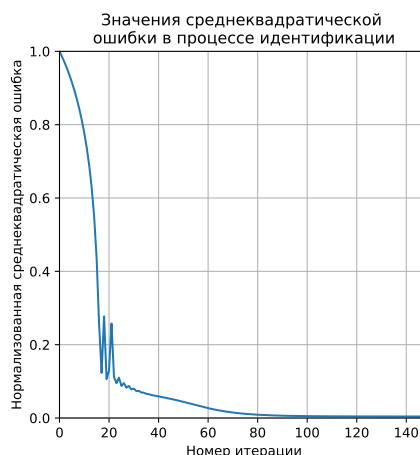


Рис. 36: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

На рисунках 37 и 38 показаны график отклонений результатов моделирования от экспериментальных данных и гистограмма отклонений. Гистограмма по форме близка к нормальному закону.



Рис. 37: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

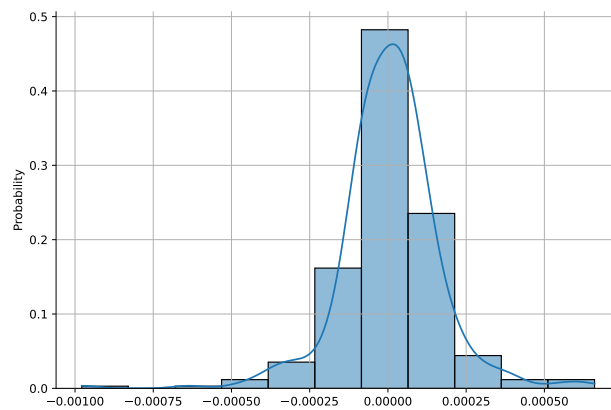


Рис. 38: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

В таблице 11 приведены параметры идентифицированной модели и результаты оценок точности моделирования.

Таблица 11: Результаты идентификации модели

Параметр	Значение
Постоянна времени τ , с	0,054291
Амплитуда, пФ	0.003448
Показатель p	0.173
Результат кроссвалидации (RMSE), пФ	0.000160
RMSE, пФ	0.000161

На рисунке 49 приведён полученный спектр частотного скана.

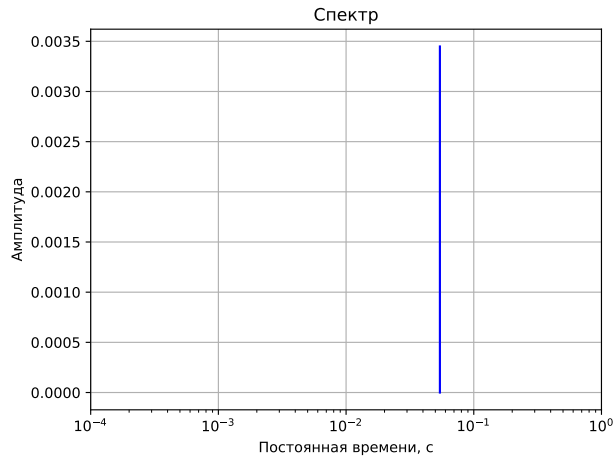


Рис. 39: Спектр сигнала релаксации ёмкости.

4.3.2 Моноэкспоненциальная модель с показателем $p = 1$

Выполним идентификацию модели идеального частотного скана на результатах, полученных при температуре 283 К. Результаты идентификации модели приведены на рисунке 40. На графике заметны существенные отклонения результатов моделирования от экспериментальных данных, при этом, график значений нормированной среднеквадратической ошибки для каждой итерации (рисунок 41) свидетельствует о том, что алгоритм идентификации нашёл минимум функции потерь.

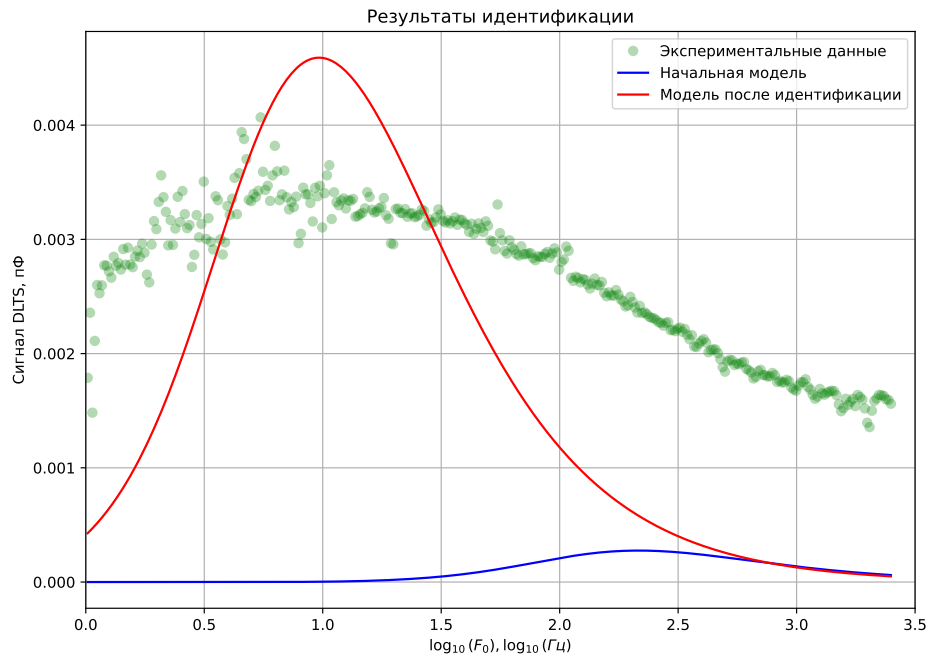


Рис. 40: Результат идентификации частотного скана при $T = 283K$.

График отклонений результатов моделирования от экспериментальных данных (рисунок 42) также демонстрирует значительные систематические отклонения. Закономерно, гистограмма (рисунок 43) по форме далека от

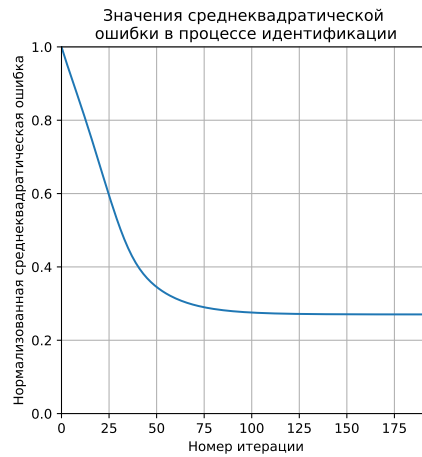


Рис. 41: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

нормального распределения.



Рис. 42: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

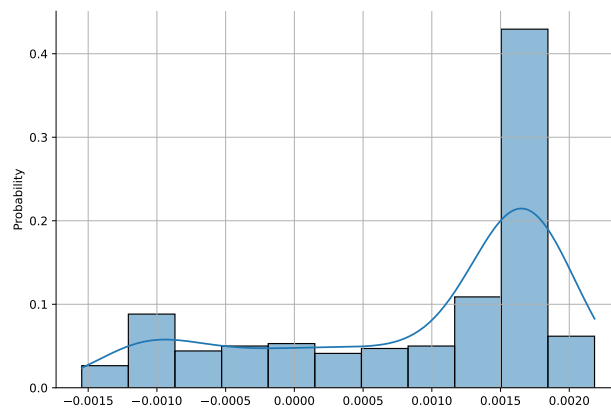


Рис. 43: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

В таблице 12 приведены параметры идентифицированной модели и результаты оценок точности моделирования.

Таблица 12: Результаты идентификации модели

Параметр	Значение
Постоянна времени τ , с	0.045708
Амплитуда, пФ	0.004591
Показатель p	1.0
Результат кроссвалидации (RMSE), пФ	0.001396
RMSE, пФ	0.001389

На рисунке 44 приведён полученный спектр частотного скана.

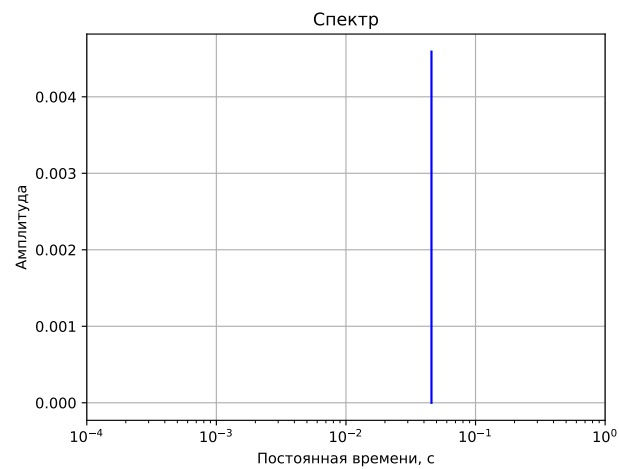


Рис. 44: Спектр сигнала релаксации ёмкости.

4.3.3 Многоэкспоненциальная модель с $n_exprs > 1$

Выполним идентификацию многоэкспоненциальной модели, для этого сравним результаты кроссвалидации для моделей, имеющих значения параметра n_exprs равные 5, 6, 7, 8, 9 и 10, и выберем показавшую лучший результат. В таблице 13 представлены полученные результаты. Лучшей оказалась модель с параметром $n_exprs = 6$, то есть имеющая 6 экспоненциальных составляющих.

Таблица 13: Результаты кроссвалидации для разных значений параметра n_exprs

Значение параметра n_exprs	Результат кроссвалидации
5	0.0001579
6	0.0001530
7	0.0001532
8	0.0001547
9	0.0001537
10	0.0001546

Результаты идентификации модели с шестью экспоненциальными составляющими представлены на рисунке 45.

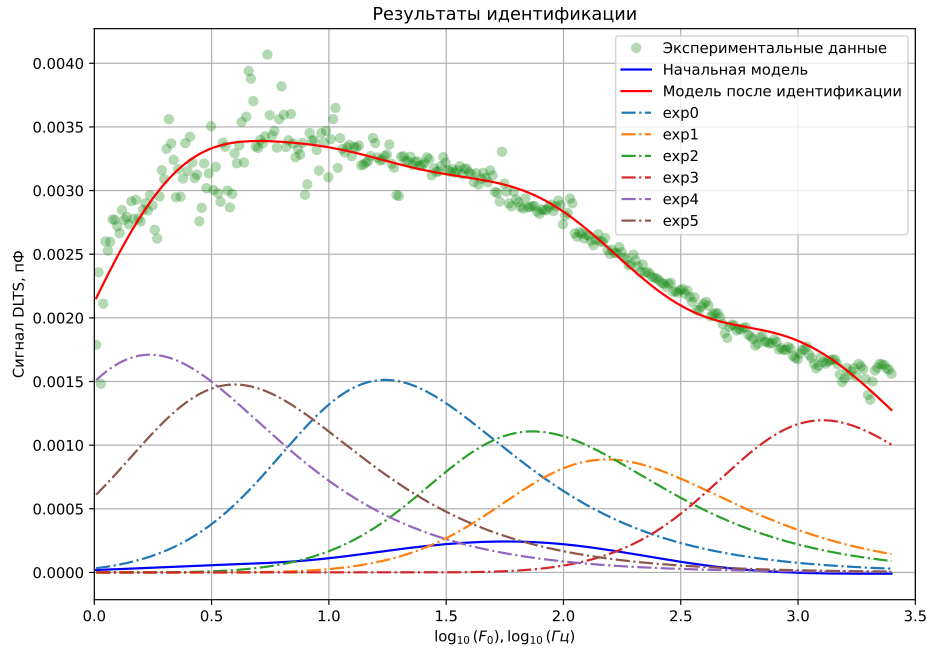


Рис. 45: Результат идентификации многоэкспоненциальной моделью частотного скана при $T = 283K$.

На рисунке 46 показан график нормированной среднеквадратической ошибки. Форма графика говорит о том, что алгоритм идентификации нашёл минимум функции потерь.

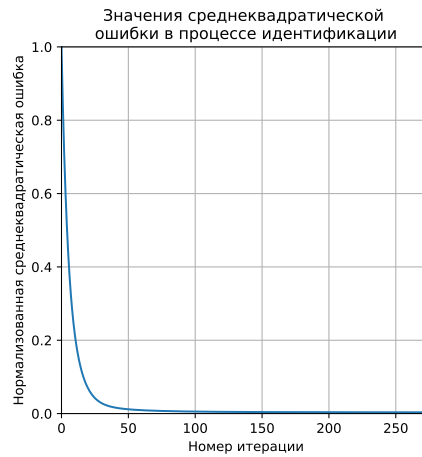


Рис. 46: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

На графике отклонений результатов моделирования от экспериментальных данных (рисунок 47) заметны систематические отклонения в области высоких значений частоты опорной функции. Они же заметны на рисунке 45. Их появление может быть связано с влиянием случайной инициализации модели – алгоритм из-за «неудачных» начальных значений завершил работу вблизи минимума функции потерь, но самого минимума так и не достиг.



Рис. 47: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

На рисунке 48 показана гистограмма отклонений результатов моделирования от экспериментальных данных. Форма гистограммы близка к форме нормального распределения.

В таблице 14 приведены параметры идентифицированной модели и результаты оценок точности моделирования, τ – постоянная времени, A – амплитуда.

На рисунке 49 приведён полученный спектр частотного скана.

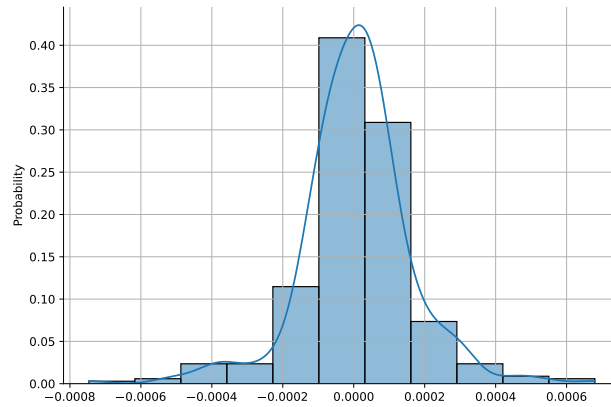


Рис. 48: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

Таблица 14: Результаты идентификации модели.

Параметр	Значение	Параметр	Значение
τ_1 , с	0.025388	A_1 , пФ	0.001513
τ_2 , с	0.002868	A_2 , пФ	0.000887
τ_3 , с	0.005958	A_3 , пФ	0.001108
τ_4 , с	0.000336	A_4 , пФ	0.001196
τ_5 , с	0.255645	A_5 , пФ	0.001711
τ_6 , с	0.111247	A_6 , пФ	0.001476
Результат кроссвалидации (RMSE), пФ	0.000158		
RMSE, пФ	0.000156		

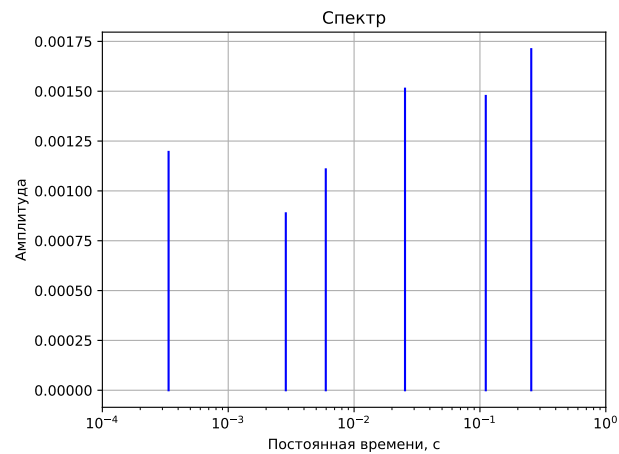


Рис. 49: Спектр сигнала релаксации ёмкости.

4.3.4 Сравнение результатов

В таблице 15 представлено сравнение результатов оценки точности рассмотренных моделей.

Модель идеального частотного скана демонстрирует значительные отклонения от результатов измерений, и, скорее всего, не подходит для описания экспериментальных данных, в то время как модель с показателем p и многоэкспоненциальная модель неплохо согласуются с результатами измерений и показывают примерно одинаковую точность.

Таблица 15: Результаты идентификации моделей.

Модель	Результат кросвалидации, пФ	RMSE, пФ
Идеальный частотный скан	0.001396	0.001389
Модель с показателем p	0.000160	0.000161
Многоэкспоненциальная модель	0.000158	0.000156

4.4 Частотный скан при температуре 303 К

4.4.1 Моноэкспоненциальная модель с показателем p

Результаты измерений при температуре 303 К приведены на рисунке 14. К сожалению, частотный скан содержит всего 34 точки, поэтому гистограммы и графики отклонений будут не так информативны как в предыдущих разделах. Результаты идентификации моноэкспоненциальной модели с показателем p приведены на рисунке 50.

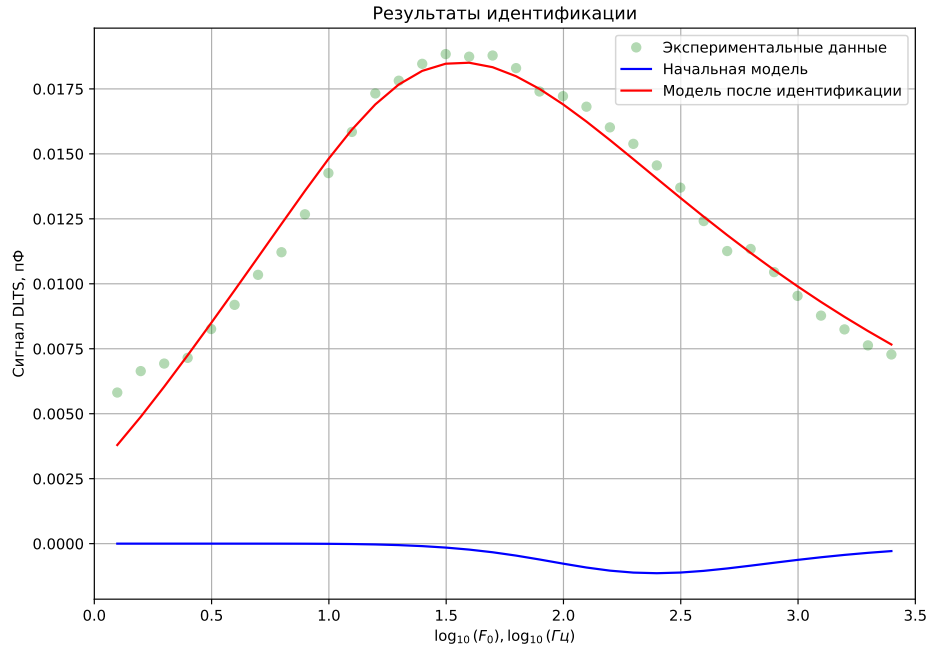


Рис. 50: Результат идентификации положительного пика частотного скана при $T = 303\text{K}$.

График нормированной среднеквадратической ошибки, приведенный на рисунке 51, указывает на то, что алгоритм идентификации нашёл минимум функции потерь.

На графике отклонений результатов моделирования, полученных на идентифицированной модели, от экспериментальных данных (рисунок 52) заметны систематические отклонения. Гистограмма отклонений, приведённая на рисунке 53, по форме существенно отличается от нормального распределения.

В таблице 16 приведены параметры идентифицированной модели и результаты оценок точности моделирования.

На рисунке 54 приведён полученный спектр частотного скана.

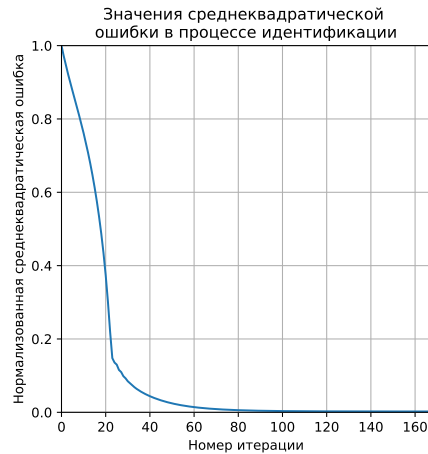


Рис. 51: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

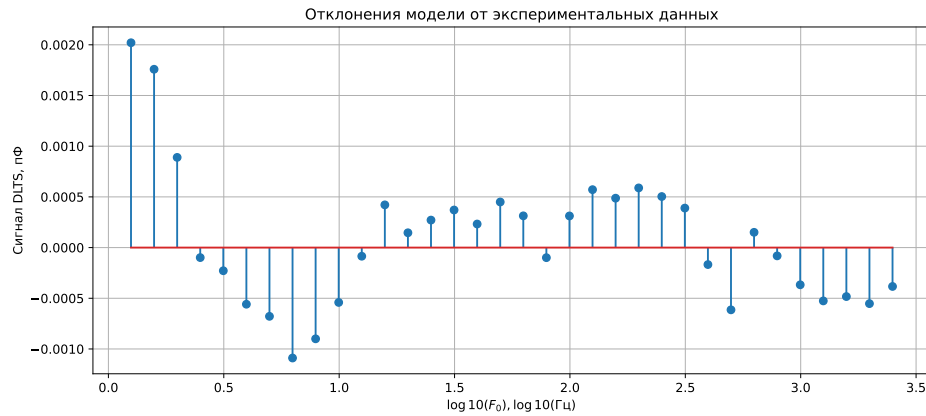


Рис. 52: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

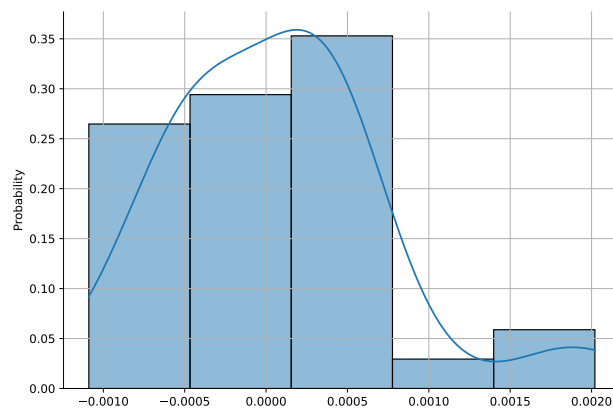


Рис. 53: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

Таблица 16: Результаты идентификации модели

Параметр	Значение
Постоянная времени τ , с	0,012014
Амплитуда, пФ	0.018520
Показатель p	0.275712
Результат кроссвалидации (RMSE), пФ	0.000713
RMSE, пФ	0.000660

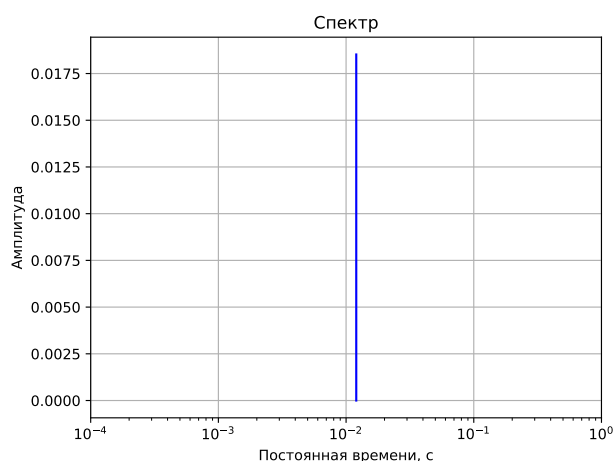


Рис. 54: Спектр сигнала релаксации ёмкости.

4.4.2 Моноэкспоненциальная модель с показателем $p = 1$

Выполним идентификацию модели идеального частотного скана на результатах, полученных при температуре 303 К. Результаты идентификации модели приведены на рисунке 55. На графике заметны существенные отклонения результатов моделирования от экспериментальных данных, при этом, график значений нормированной среднеквадратической ошибки для каждой итерации (рисунок 56) свидетельствует о том, что алгоритм идентификации нашёл минимум функции потерь.

График отклонений результатов моделирования от экспериментальных данных (рисунок 57) также демонстрирует значительные систематические отклонения, а гистограмма (рисунок 58) по форме далека от нормального распределения.

В таблице 17 приведены параметры идентифицированной модели и результаты оценок точности моделирования.

На рисунке 59 приведён полученный спектр частотного скана.

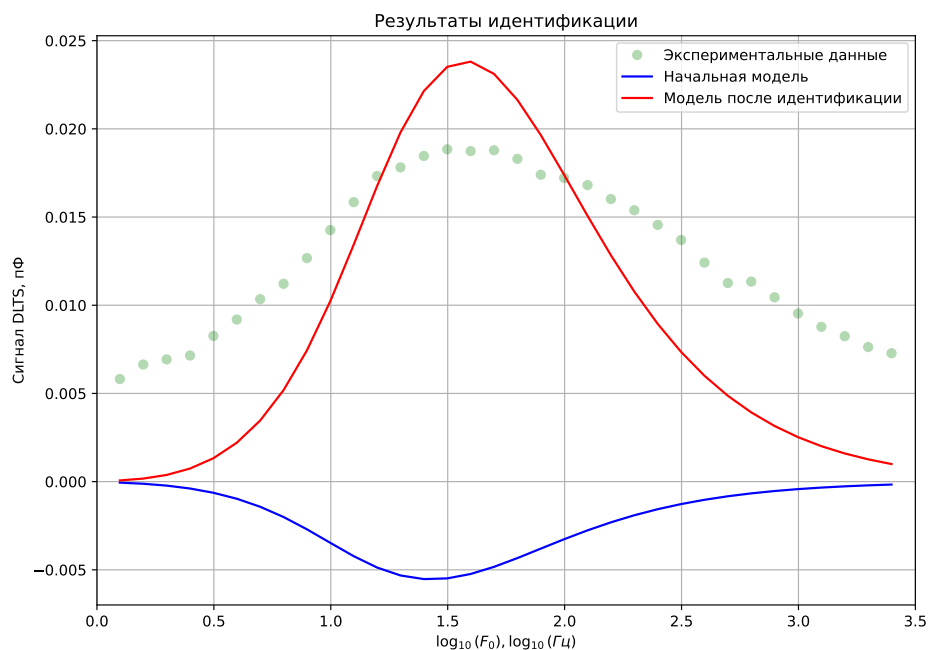


Рис. 55: Результат идентификации положительного пика частотного скана при $T = 303K$.

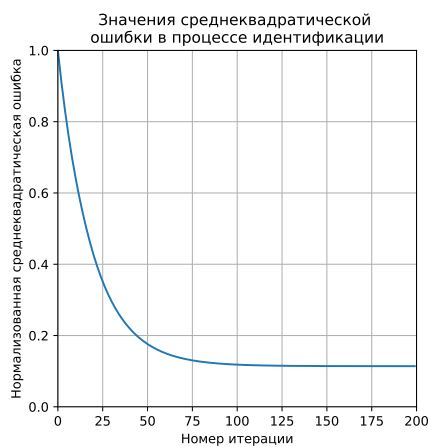


Рис. 56: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

Таблица 17: Результаты идентификации модели

Параметр	Значение
Постоянна времени τ , с	0,011709
Амплитуда, пФ	0.023837
Показатель p	1.0
Результат кроссвалидации (RMSE), пФ	0.005837
RMSE, пФ	0.005436

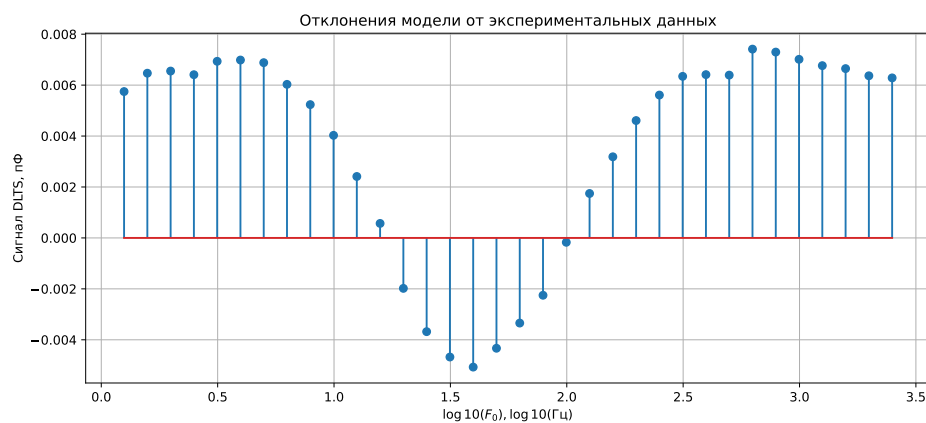


Рис. 57: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

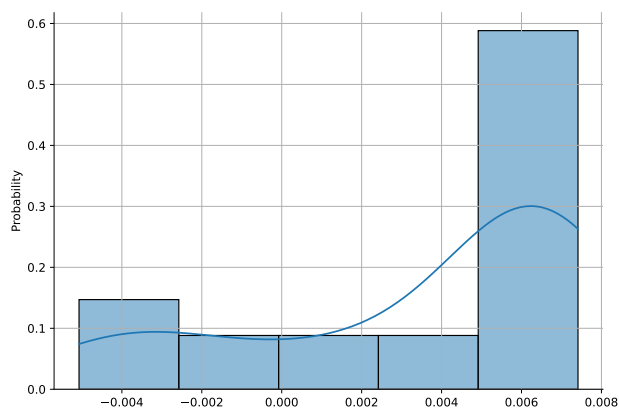


Рис. 58: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных. Сплошной линией показан результат сглаживания гистограммы.

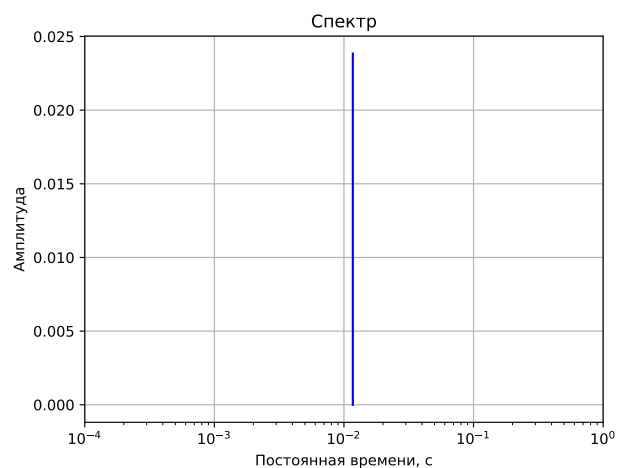


Рис. 59: Спектр сигнала релаксации ёмкости.

4.4.3 Многоэкспоненциальная модель с $n_exprs > 1$

Выполним идентификацию многоэкспоненциальной модели, для этого сравним результаты кроссвалидации для моделей, имеющих значения параметра n_exprs равные 5, 6, 7, 8, 9 и 10, и выберем показавшую лучший результат. В таблице 18 представлены полученные результаты. Лучшей оказалась модель с параметром $n_exprs = 6$, то есть имеющая 6 экспоненциальных составляющих.

Таблица 18: Результаты кроссвалидации для разных значений параметра n_exprs

Значение параметра n_exprs	Результат кроссвалидации
5	0.0005614
6	0.0003977
7	0.0004187
8	0.0004464
9	0.0004572
10	0.0004476

Результаты идентификации модели с шестью экспоненциальными составляющими представлены на рисунке 60.

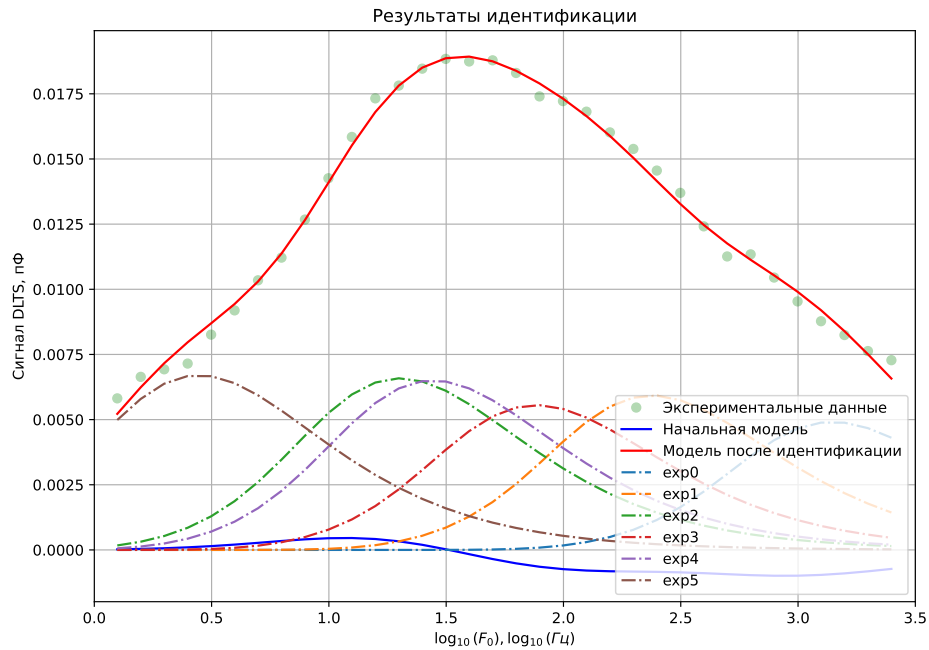


Рис. 60: Результат идентификации многоэкспоненциальной моделью частотного скана при $T = 303K$.

На рисунке 61 показан график нормированной среднеквадратической ошибки. Форма графика говорит о том, что алгоритм идентификации нашёл минимум функции потерь.

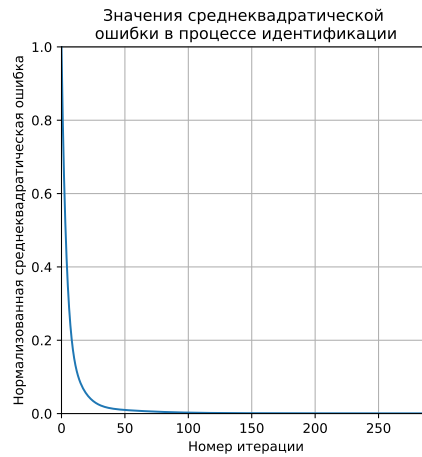


Рис. 61: График среднеквадратической ошибки в процессе идентификации, нормированной относительно её максимального значения.

Из-за малого количества точек на частотном скане по графику отклонений результатов моделирования от экспериментальных данных (рисунок 62) трудно сделать какие-либо выводы. Гистограмма отклонений (рисунок 63) по форме ближе к нормальному распределению, однако, из-за недостаточного количества данных не является показательной.

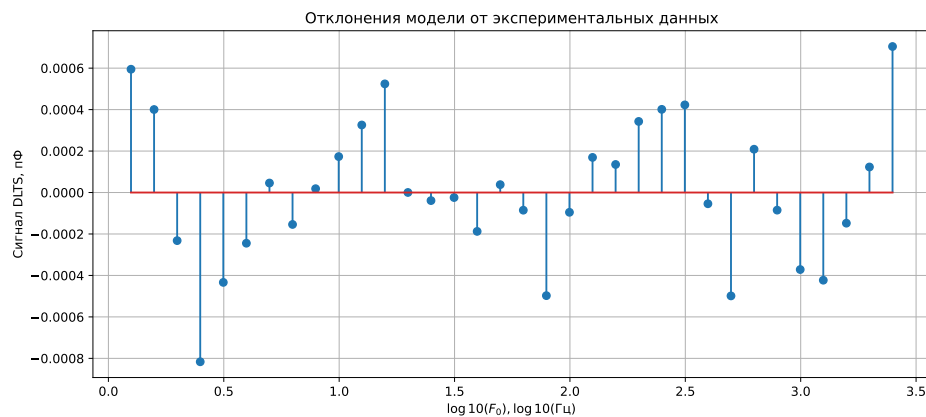


Рис. 62: График отклонений результатов, полученных на идентифицированной модели, от экспериментальных данных.

В таблице 19 приведены параметры идентифицированной модели и результаты оценок точности моделирования, τ – постоянная времени, A – амплитуда.

На рисунке 64 приведён полученный спектр частотного скана.

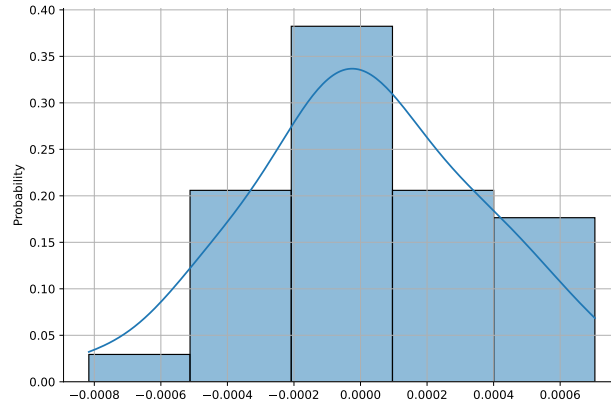


Рис. 63: Гистограмма отклонений данных, полученных на идентифицированной модели, от экспериментальных данных.

Таблица 19: Результаты идентификации модели.

Параметр	Значение	Параметр	Значение
τ_1 , с	0.000304	A_1 , пФ	0.004912
τ_2 , с	0.001845	A_2 , пФ	0.005924
τ_3 , с	0.021940	A_3 , пФ	0.006586
τ_4 , с	0.005750	A_4 , пФ	0.005556
τ_5 , с	0.015893	A_5 , пФ	0.006504
τ_6 , с	0.158127	A_6 , пФ	0.006706
Результат кроссвалидации (RMSE), пФ	0.000442		
RMSE, пФ	0.000338		

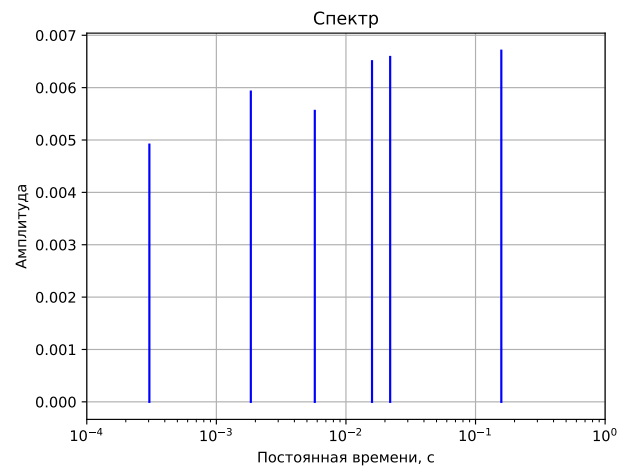


Рис. 64: Спектр сигнала релаксации ёмкости.

4.4.4 Сравнение результатов

В таблице 20 представлено сравнение результатов оценки точности рассмотренных моделей.

Модель идеального частотного скана демонстрирует значительные отклонения от результатов измерений, и, скорее всего, не подходит для описания экспериментальных данных, в то время как модель с показателем p и многоэкспоненциальная модель неплохо согласуются с результатами измерений. Лучшие результаты показывает многоэкспоненциальная модель.

Таблица 20: Результаты идентификации моделей.

Модель	Результат кросвалидации, пФ	RMSE, пФ
Идеальный частотный скан	0.005837	0.005436
Модель с показателем p	0.000713	0.000660
Многоэкспоненциальная модель	0.000442	0.000338

4.5 Графики в координатах Аррениуса

Для завершения анализа и вычисления спектров энергий активации, необходимо построить графики в координатах Аррениуса. В предыдущем разделе многоэкспоненциальные модели для 263 К, 283 К и 303 К имели разное оптимальное количество моноэкспоненциальных составляющих: 8, 6 и 6, соответственно. Для того, чтобы было удобнее строить графики в координатах Аррениуса и интерпретировать результаты, выполним повторную идентификацию всех частотных сканов многоэкспоненциальной моделью с шестью экспоненциальными составляющими.

Результат идентификации частотного скана при $T = 263\text{ К}$ и соответствующий спектр показаны на рисунках 65 и 66 соответственно.

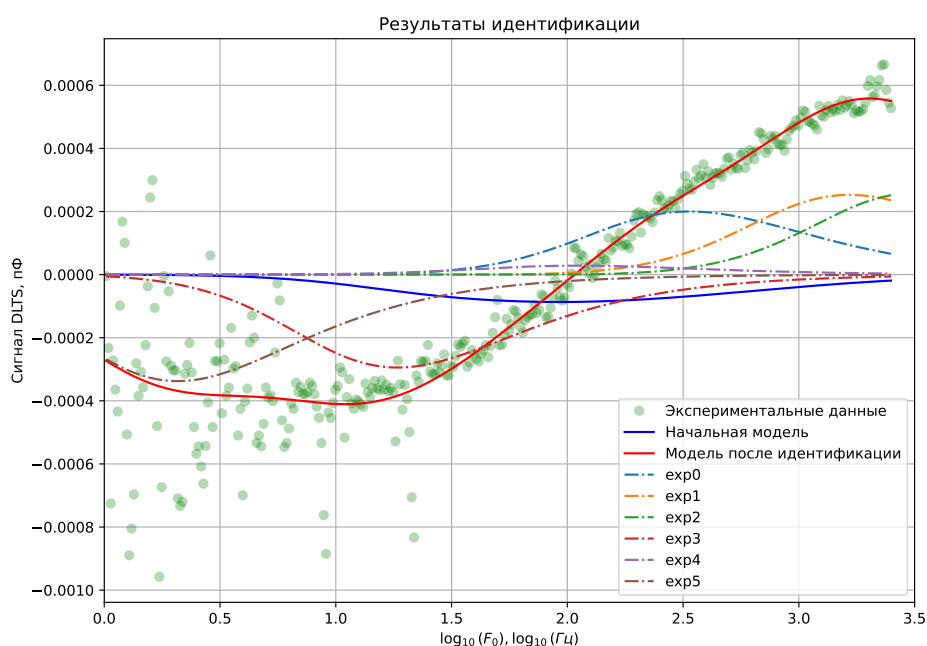


Рис. 65: Результат идентификации многоэкспоненциальной модели на частотном скане, полученном при $T = 263\text{ К}$.

Результат идентификации частотного скана при $T = 283\text{ К}$ и соответствующий спектр показаны на рисунках 67 и 68 соответственно.

Результат идентификации частотного скана при $T = 303\text{ К}$ и соответствующий спектр показаны на рисунках 69 и 70 соответственно.

В таблице 21 для каждого значения температуры приведены параметры идентифицированных моделей, отсортированные по возрастанию значений постоянных времени, и оценки точности моделирования.

На рисунке 71 приведён график с результатами, перенесёнными в координаты Аррениуса. Для каждой группы значений постоянной времени (например, τ_1 для 263 К, 283 К, 303 К) построены графики линейной регрессии. На вставке показана группа точек, соответствующих температуре 303 К.

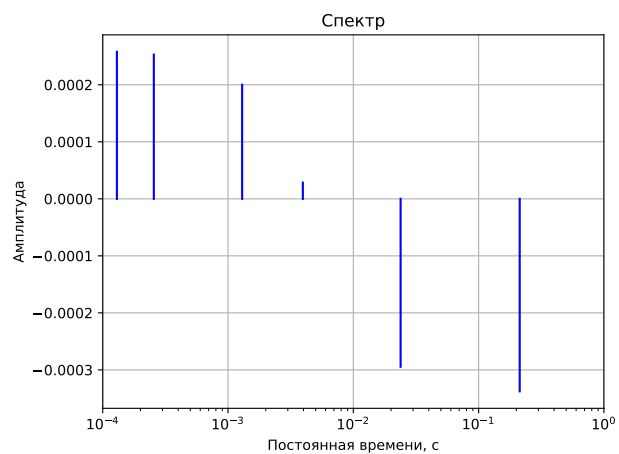


Рис. 66: Спектр сигнала релаксации ёмкости при $T = 263K$.

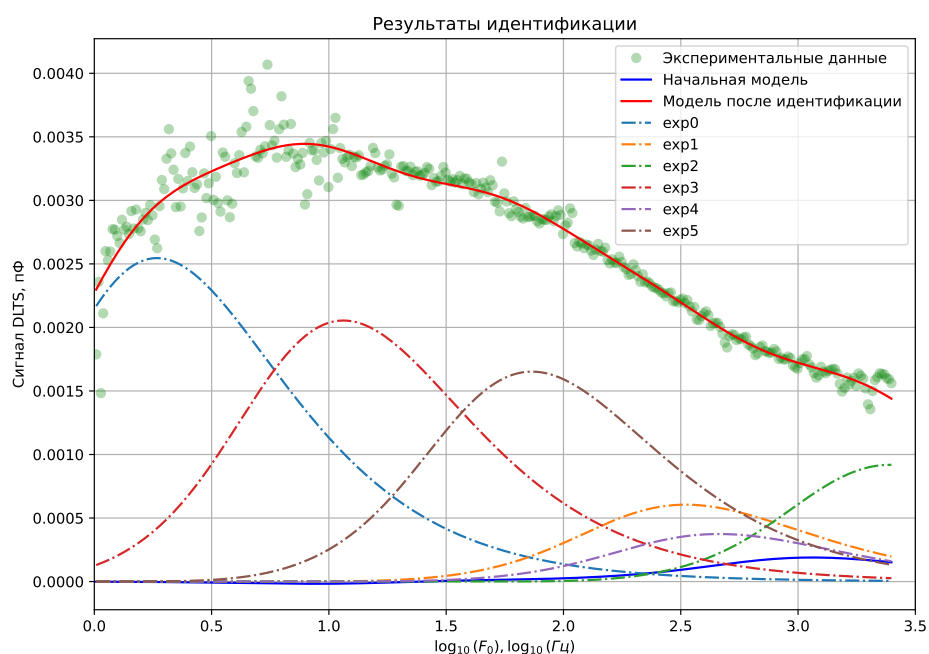


Рис. 67: Результат идентификации многоэкспоненциальной модели на частотном скане, полученном при $T = 283K$.

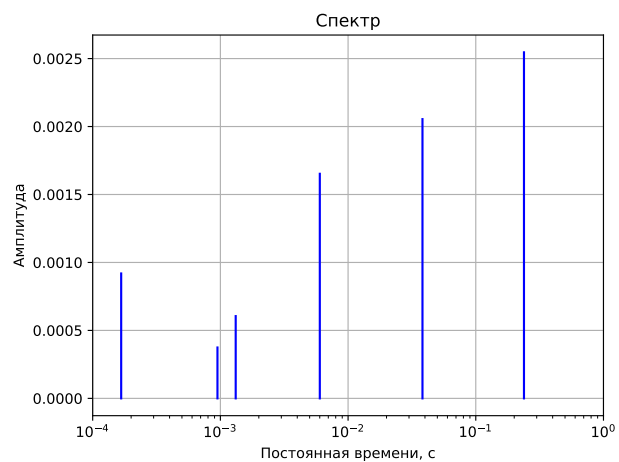


Рис. 68: Спектр сигнала релаксации ёмкости при $T = 283K$.

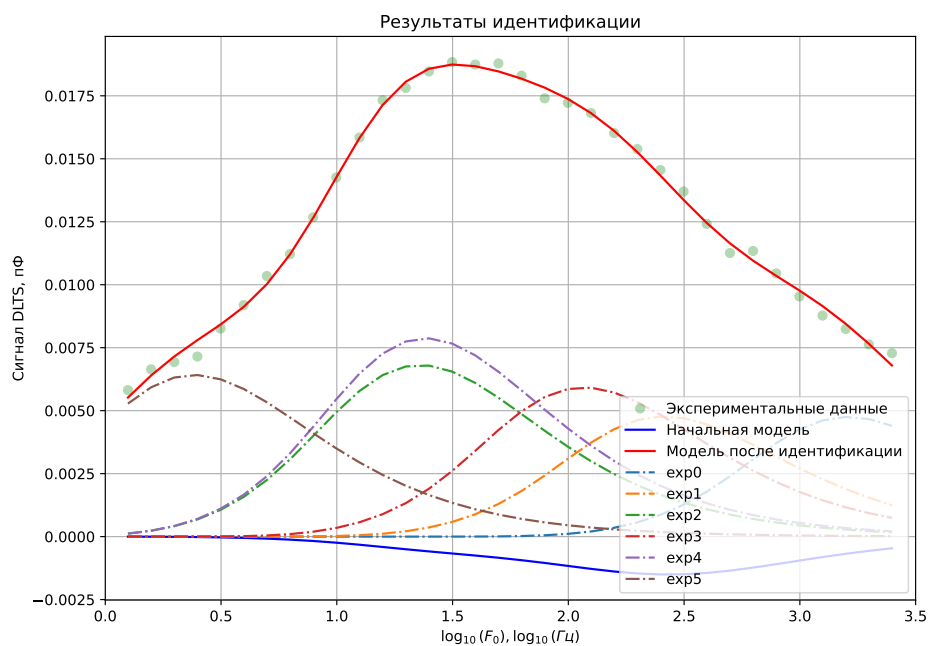


Рис. 69: Результат идентификации многоэкспоненциальной модели на частотном скане, полученном при $T = 303K$.

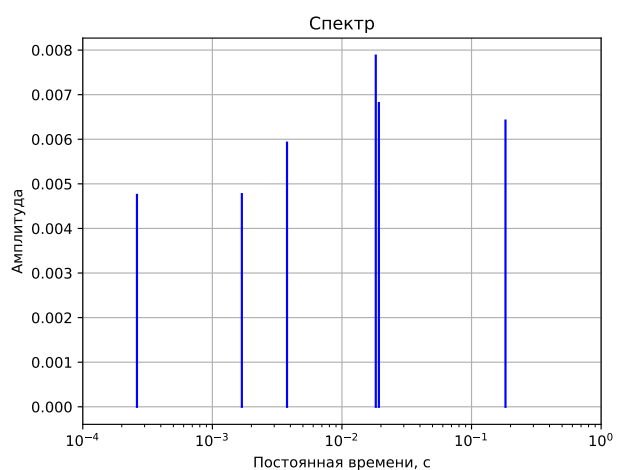


Рис. 70: Спектр сигнала релаксации ёмкости при $T = 303K$.

Таблица 21: Результаты идентификации моделей.

Параметр	263 K	283 K	303 K
τ_1	0.000130	0.000167	0.000262
A_1	0.000258	0.000919	0.004751
τ_2	0.000256	0.000949	0.001689
A_2	0.000253	0.000374	0.004769
τ_3	0.001296	0.001318	0.003768
A_3	0.000200	0.000605	0.005924
τ_4	0.003955	0.006007	0.018241
A_4	0.000028	0.001652	0.007875
τ_5	0.023848	0.038246	0.019270
A_5	-0.000294	0.002055	0.006812
τ_6	0.212794	0.238471	0.182804
A_6	-0.000338	0.002546	0.006419
Результат кроссвалидации (RMSE), пФ	0.000131	0.000150	0.000376
RMSE	0.000131	0.000141	0.000252

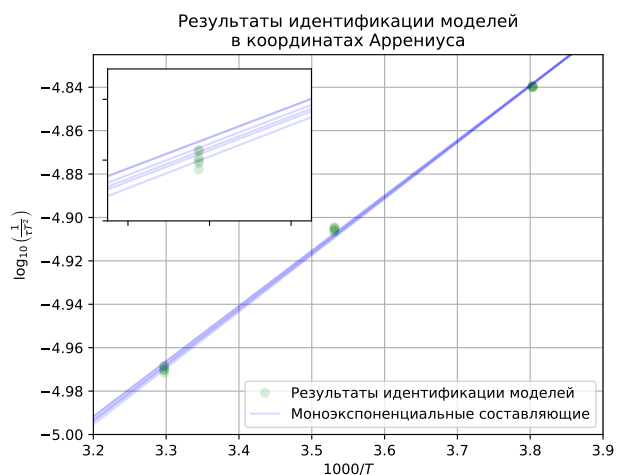


Рис. 71: График в координатах Аррениуса с графиками линейной регрессии.

Для сравнения на рисунке 72 приведён график в координатах аррениуса, где для тех же данных выполнена регрессия полиномом второй степени. На вставке показана группа точек, соответствующих температуре 303 К.

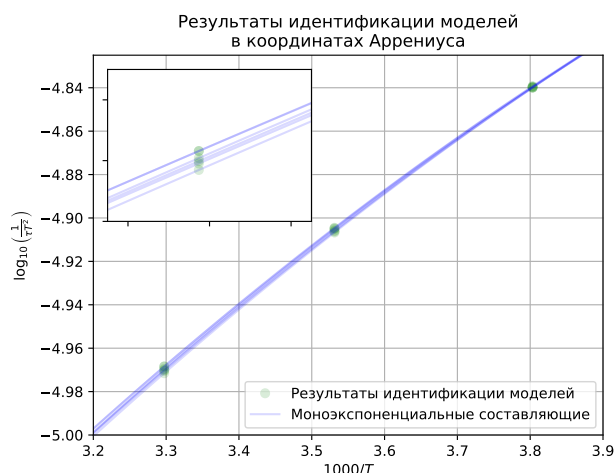


Рис. 72: График в координатах Аррениуса с графиками регрессии полиномом второй степени.

На обоих графиках (рисунки 71 и 72) заметно уширение «корридора», образованного линиями регрессии.

Если воспользоваться рассчитанными параметрами линейных регрессий, приведённых на рисунке 71, то, в соответствии со стандартом ASTM F978-02, возможно вычислить энергии активации глубоких центров. Результаты вычисления энергий представлены в таблице 22.

Таблица 22: Наклон прямых в координатах Аррениуса и рассчитанная энергия активации.

№	Наклон	Энергия активации, эВ
1	0.25467582	-0.0505
2	0.25470544	-0.0505
3	0.25627477	-0.0509
4	0.25695162	-0.0510
5	0.25764135	-0.0511
6	0.25972542	-0.0515

По расчётам энергии активации занимают диапазон примерно от 50 мэВ до 52 мэВ, что позволяет предположить, что получены некорректные результаты. Действительно, по частотным сканам, полученным при трёх температурах в диапазоне от 263 К до 303 К, трудно оценить энергию активации, поэтому необходимо выполнить дополнительные измерения (получить больше частотных сканов) при различных температурах, прежде чем делать выводы.

5 Заключение

Применение многоэкспоненциальной модели, как показано в таблицах 10, 15, 20 позволяет повысить точность моделирования частотных сканов. При этом, важно помнить, что моноэкспоненциальная модель никак не учитывает искажения частотного скана за счёт влияния нелинейностей измерительного тракта.

Инициализация многоэкспоненциальной модели случайными начальными параметрами может влиять на результат решения и приводить к тому, что алгоритм будет останавливаться вблизи оптимальных значений параметров, но оптимальных достигать не будет (рисунок 45), в таких случаях необходимо выполнять идентификацию повторно, либо подстраивать параметры алгоритма идентификации, например скорость градиентного спуска.

Многоэкспоненциальная модель нуждается в доработке алгоритма идентификации, в частности, оптимизации с точки зрения времени расчётов.

На графиках в координатах Аррениуса наблюдается уширение «корридора» образованного линиями регрессии, как для линейных регрессий, так и для регрессий полиномом второй степени (рисунки 71 и 72). Для анализа результатов в координатах Аррениуса и вычисления энергии активации необходимо собрать больше данных: получить больше частотных при различных температурах.

Список литературы

- [1] F. J. Anscombe. «Graphs in Statistical Analysis». В: *The American Statistician* 27.1 (1973), с. 17—21. ISSN: 00031305. DOI: [10.2307/2682899](https://doi.org/10.2307/2682899). URL: <http://www.jstor.org/stable/2682899> (дата обр. 11.10.2022).
- [2] *Cross-validation: evaluating estimator performance*. URL: https://scikit-learn.org/stable/modules/cross_validation.html (дата обр. 12.10.2022).
- [3] Aurelien Geron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd. O'Reilly Media, Inc., 2019. ISBN: 1492032646.
- [4] Trevor Hastie, Robert Tibshirani и Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York, NY, 2017. ISBN: 978-0-387-84858-7. DOI: [10.1007/978-0-387-84858-7](https://doi.org/10.1007/978-0-387-84858-7).
- [5] Andrei A. Istratov и Oleg F. Vyvenko. «Exponential analysis in physical phenomena». В: *Review of Scientific Instruments* 70.2 (1999), с. 1233—1257. DOI: [10.1063/1.1149581](https://doi.org/10.1063/1.1149581). eprint: <https://doi.org/10.1063/1.1149581>. URL: <https://doi.org/10.1063/1.1149581>.
- [6] Vladimir Krylov, Aleksey Bogachev и Т. Pronin. «Deep level relaxation spectroscopy and nondestructive testing of potential defects in the semiconductor electronic component base». В: *Radio industry (Russia)* 29 (май 2019), с. 35—44. DOI: [10.21778/2413-9599-2019-29-2-35-44](https://doi.org/10.21778/2413-9599-2019-29-2-35-44).
- [7] *scikit-learn: Machine Learning in Python*. URL: <https://scikit-learn.org/> (дата обр. 10.10.2022).
- [8] *TensorFlow*. URL: <https://www.tensorflow.org/> (дата обр. 10.10.2022).
- [9] *Квартет Энскомба*. URL: https://en.wikipedia.org/wiki/Anscombe%27s_quartet (дата обр. 12.10.2022).
- [10] С. Николенко, А. Кадури и Е. Архангельская. *Глубокое обучение*. СПб.: Питер, 2018, с. 480. ISBN: 978-5-496-02536-2.