

Моделирование частотных сканов

Богачев А.М. 13.04.2022

Оглавление

1. Цели и задачи.....	2
2. Моделирование частотных сканов с одной экспоненциальной составляющей без показателя p	3
2.1 Подготовка данных	3
2.2 Идентификация параметров модели	5
3. Моделирование частотных сканов с учётом показателя p	7
4. Моделирование частотных сканов с несколькими экспоненциальными составляющими	9
4.1 Подготовка данных	9
4.2 Пример удачной идентификации	10
4.3 Пример неудачной идентификации №1	11
4.4. Пример неудачной идентификации №2.....	12
5. Выводы	14

1. Цели и задачи

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 $g(\lambda)$.

Таким образом, цель работы по моделированию частотных сканов в самом общем виде можно описать следующим образом: **найти способ определять спектральную функцию $g(\lambda)$, её параметры, а также зависимость этих параметров от температуры.**

Для достижения поставленной цели нужно решить предположительно следующие задачи:

1. Разработать алгоритм идентификации частотного скана моноэкспоненциального сигнала релаксации.
2. Разработать алгоритм идентификации частотного скана неэкспоненциального сигнала релаксации с показателем p , характеризующим неэкспоненциальность.
3. Разработать программу идентификации частотного скана сигнала релаксации, состоящего из суммы моноэкспоненциальных сигналов. Программа должна определять параметры каждого моноэкспоненциального сигнала.
4. Разработать программу идентификации группы частотных сканов при разных температурах. На данном этапе предполагается, что сигнал релаксации образован суммой моноэкспоненциальных сигналов, количество которых не меняется в зависимости от температуры, меняются только их параметры.
5. Разработать программу идентификации спектральной функции $g(\lambda)$ и её параметров для отдельного частотного скана.
6. Разработать программу идентификации спектральной функции $g(\lambda)$ и её параметров для группы частотных сканов при разных температурах, полагая, что от температуры зависят только параметры $g(\lambda)$.

2. Моделирование частотных сканов с одной экспоненциальной составляющей без показателя p

2.1 Подготовка данных

Выходной сигнал коррелятора спектрометра DLS-82E имеет следующий вид:

$$S(\tau, C_A, F_0) = C_A K_{BS} K_{LS} \varphi(\tau, F_0, t_1), \quad (1)$$

где C_A – амплитуда емкостного релаксационного сигнала,

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

K_{LS} – масштабный коэффициент селектора,

τ – постоянная релаксации ГУ,

F_0 – частота следования импульсов заполнения,

t_1 – длительность импульса заполнения,

$$\varphi(\tau, F_0, t_1) = M \tau F_0 e^{-\frac{0.05}{\tau F_0}} \left(1 - e^{\frac{t_1 F_0 - 0.45}{\tau F_0}} - e^{-\frac{0.5}{\tau F_0}} + e^{\frac{t_1 F_0 - 0.95}{\tau F_0}} \right), \quad (2)$$

где M – масштабный множитель.

Введём коэффициент A характеризующий амплитуду сигнала релаксации ёмкости:

$$A = C_A K_{BS} K_{LS}. \quad (3)$$

При моделировании будем считать масштабный коэффициент и длительность импульса заполнения постоянными:

$$t_1 = 20 * 10^{-6} \text{ с},$$

$$M = 5.861.$$

При идентификации параметров модели будем определять коэффициент A , характеризующий амплитуду сигнала релаксации, и постоянную времени сигнала релаксации τ .

Так как выражения (1) и (2) при моделировании будут зависеть только от трёх параметров: τ , A , F_0 , их можно переписать следующим образом:

$$S(\tau, A, F_0) = A \varphi(\tau, F_0) \quad (4)$$

$$\varphi(\tau, F_0) = M \tau F_0 e^{-\frac{0.05}{\tau F_0}} \left(1 - e^{\frac{t_1 F_0 - 0.45}{\tau F_0}} - e^{-\frac{0.5}{\tau F_0}} + e^{\frac{t_1 F_0 - 0.95}{\tau F_0}} \right) \quad (5)$$

Подготовим исходные данные для моделирования, на первом этапе предлагается использовать расчётные данные вместо экспериментальных, чтобы можно было оценить качество идентификации параметров модели.

Подготовим массив значений частоты, распределённых равномерно по логарифму частоты. Каждое значение вычисляется по следующей формуле:

$$F_{0i} = 10^{(i-1)d}, \quad (6)$$

где i – номер элемента в массиве (от 1 до N ; N – количество элементов в массиве), d – шаг по логарифму частоты, который вычисляется автоматически в программе моделирования.

В следующей таблице приведены параметры массива частот:

Таблица 1 – Параметры массива частот

Параметр	Начальная частота, Гц	Конечная частота, Гц	Количество значений
Значения	1	2500	1000

Для каждого значения частоты вычислим значения выходного сигнала коррелятора по следующей формуле:

$$S^*(\tau, A, F_0) = A\varphi(\tau, F_0) + n(\mu, \sigma), \quad (6)$$

где n – нормально распределённый шум с математическим ожиданием μ и среднеквадратическим отклонением σ .

Для расчётов будем использовать параметры, перечисленные в следующей таблице:

Таблица 2 – Параметры для вычисления исходных данных

Параметр	τ , с	A	μ	σ
Значение	0.005	3.0	0	0.2

На рисунке ниже показан рассчитанный частотный скан.

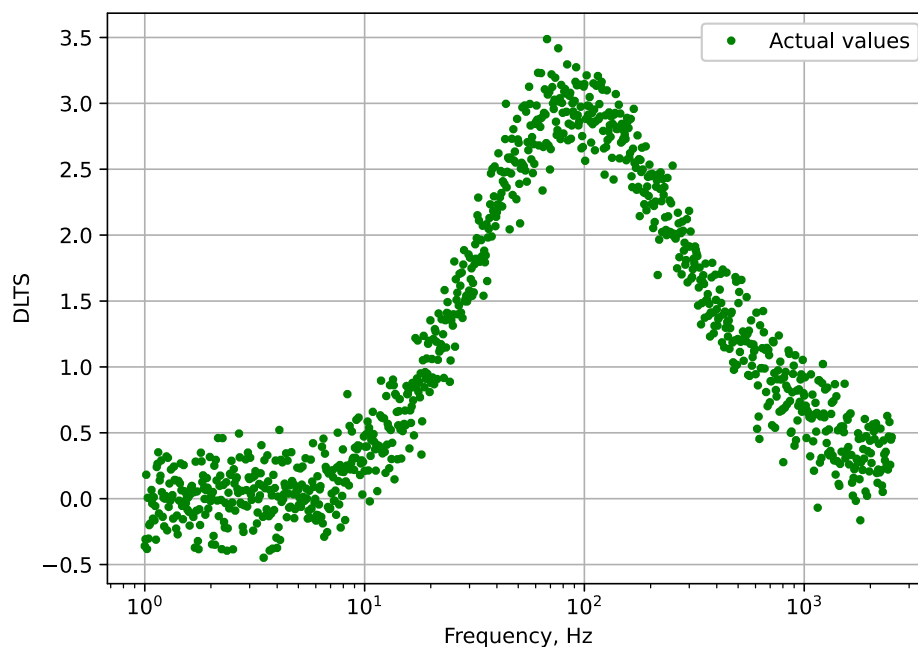


Рисунок 1 – Рассчитанный частотный скан. По горизонтали отложена частота в Гц, по вертикали – сигнал DLTS.

2.2 Идентификация параметров модели

Идентификацию параметров модели будем проводить нелинейным методом наименьших квадратов.

В качестве целевой функции при оптимизации параметров модели выберем минимум среднеквадратической ошибки между исходными данными и данными, вычисляемыми моделью (формула 7).

$$E = \frac{1}{N} \sum_{i=1}^N (S_i^* - S_i)^2 \quad (7)$$

Оптимизацию параметров модели будем выполнять методом стохастического градиентного спуска по мини-батчам (выборкам), так как данный алгоритм обладает высокой скоростью и меньше подвержен риску остановиться в локальном минимуме целевой функции. Размер выборки (параметр “*batch_size*”) зададим равным 100, количество итераций (параметр “*epochs*” – «эпохи») также зададим равным 100, скорость градиентного спуска (параметр “*learning_rate*”) зададим равным 0.1. Данный алгоритм работает следующим образом:

- 1) из всех исходных данных 100 точек (параметр “*batch_size*”) выбираются случайным образом;
- 2) по выбранным точкам выполняется градиентный спуск;
- 3) пункты 1 и 2 повторяются $N/\text{batch_size}$ раз (N – количество точек в массиве исходных данных), при каждом повторении фиксируются лучшие параметры модели;
- 4) пункты 1 – 3 повторяются заданное количество раз (параметр “*epochs*”), в конце каждой итерации фиксируются лучшие значения параметров.

Стохастический градиентный спуск находит решение быстрее и работает стабильнее, если значения параметров модели находятся примерно в одном и том же масштабе. Чтобы выполнить данное условие, будем оптимизировать не постоянную времени τ , а её логарифм, т.е. в выражениях (4) и (5) заменим τ следующим выражением (8) и будем искать оптимальное значение степени ρ вместо τ .

$$\tau = 10^\rho \quad (8)$$

Данная замена также исключит возможность появления отрицательных значений τ в результате работы алгоритма оптимизации.

Начальные значения параметров будем выбирать случайными из диапазонов, приведённых в следующей таблице.

Таблица 3 – Диапазоны начальных значений параметров.

Параметр	Минимальное значение	Максимальное значение
ρ	-3	-0,5
A	0,3	3,3

Не исключено, что вместо произвольных значений правильнее выбирать значения параметров в произвольной точке частотного скана.

Ниже приведён результат работы алгоритма идентификации.

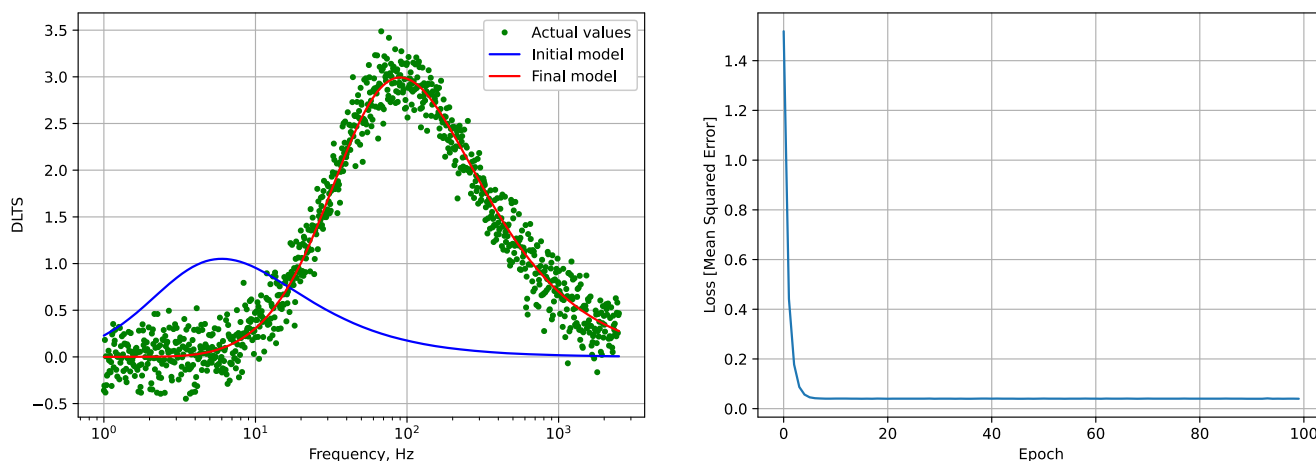


Рисунок 2 – Результаты идентификации модели. На графике слева показаны исходные данные (зелёные точки), частотный скан, полученный на модели до оптимизации параметров (синяя линия), частотный скан, полученный на модели после оптимизации параметров (красная линия); по горизонтали отложена частота в Гц, по вертикали – сигнал DLTS. На графике справа показана зависимость среднеквадратической ошибки от количества «эпох».

В следующей таблице приведены численные значения результатов идентификации модели.

Таблица 4 – Результаты идентификации параметров модели.

Параметр	Значения до оптимизации	Значения после оптимизации	Исходные значения
A	1.0500	2.9973	3.0
ρ	-1.133	-2.312	-
τ	0.0736	0.0049	0.005
E	2.1472	0.0406	-
\sqrt{E}	1.465317	0.201380	-

Алгоритм достигает сходимости очень быстро, таким образом, нет необходимости использовать вариации градиентного спуска с адаптивной скоростью.

3. Моделирование частотных сканов с учётом показателя p

Для комплексного учёта нелинейности аналогового тракта спектрометра и неэкспоненциальности релаксационного сигнала используется показатель p , в этом случае выражение (4) приобретает следующий вид:

$$S(\tau, A, F_0) = A[\varphi(\tau, F_0)]^p \quad (9)$$

Рассчитаем экспериментальные данные аналогично тому как это делалось в предыдущей части, но принимая во внимание коэффициент p . Для расчёта используется та же сетка частот, что и в предыдущем разделе. Параметры, использованные для расчёта приведены ниже.

Таблица 5 – Параметры для вычисления исходных данных

Параметр	τ , с	A	p	μ	σ
Значение	0.005	3.0	1.5	0	0.2

На рисунке ниже приведён рассчитанный частотный скан.

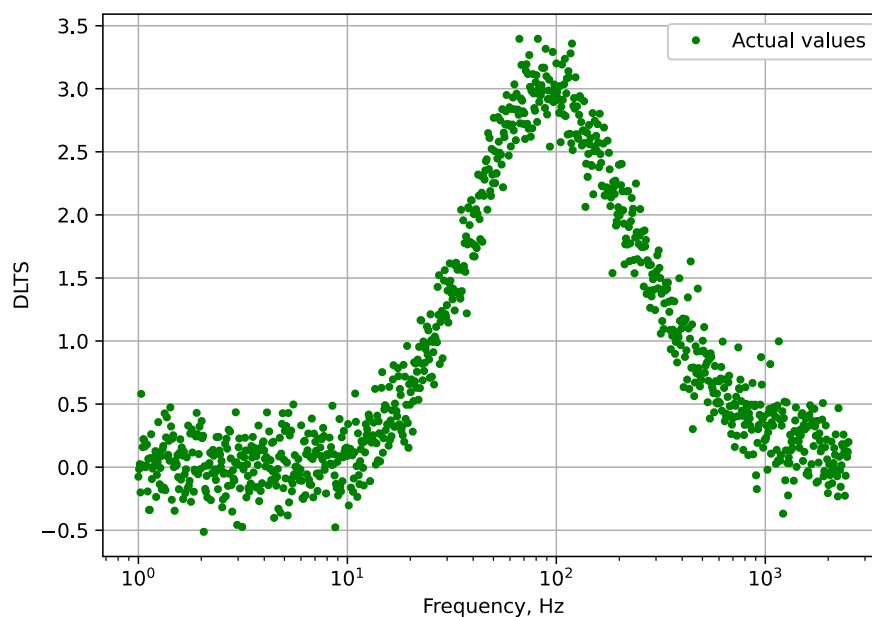


Рисунок 3 – Рассчитанный частотный скан. По горизонтали отложена частота в Гц, по вертикали – сигнал DLTS.

Выполним идентификацию параметров модели тем же способом, что и в предыдущем разделе. Алгоритм и его параметры оставим без изменений, за исключением того, что идентификация теперь проводится по трём параметрам, а параметр p в начале идентификации всегда будет равным 1.

Ниже приведён результат работы алгоритма идентификации.

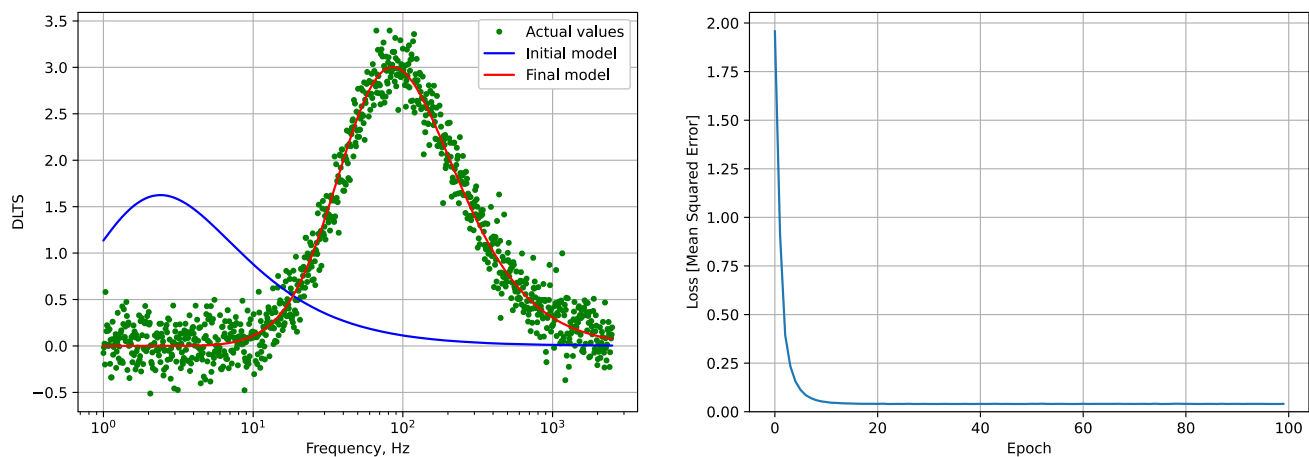


Рисунок 4 – Результаты идентификации модели. На графике слева показаны исходные данные (зелёные точки), частотный скан, полученный на модели до оптимизации параметров (синяя линия), частотный скан, полученный на модели после оптимизации параметров (красная линия); по горизонтали отложена частота в Гц, по вертикали – сигнал DLTS. На графике справа показана зависимость среднеквадратической ошибки от количества «эпох».

В следующей таблице приведены численные значения результатов идентификации модели.

Таблица 4 – Результаты идентификации параметров модели.

Параметр	Значения до оптимизации	Значения после оптимизации	Исходные значения
A	1.6234	3.0084	3.0
ρ	-0.7361	-2.2892	-
τ	0.1836	0.0051	0.005
p	1.0	1.5225	1.5
E	2.2571	0.0404	-
\sqrt{E}	1.5024	0.2010	-

4. Моделирование частотных сканов с несколькими экспоненциальными составляющими

4.1 Подготовка данных

Сигнал релаксации ёмкости может иметь неэкспоненциальную форму из-за того, что он сам является суммой нескольких экспоненциальных релаксационных сигналов.

Такой сигнал можно описать следующим выражением:

$$f(t) = \sum_{i=1}^m A_i e^{-\frac{t}{\tau_i}}, \quad (10)$$

где m – количество экспоненциальных составляющих.

Частотный скан полученный в таком случае будет определяться следующим выражением:

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

где $S(\tau_i, A_i, F_0)$ – частотный скан для каждой моноэкспоненциальной составляющей без показателя p , определяемый выражением (4).

Внесём соответствующие изменения в программную реализацию модели и рассчитаем исходные данные так же, как и в предыдущих разделах. В таблице ниже приведены параметры отдельных составляющих рассчитанного частотного скана.

Таблица 5 – Параметры моноэкспоненциальных составляющих частотного скана (нечётные пары столбцов выделены для повышения читаемости).

Параметр	Значение	Параметр	Значение	Параметр	Значение	Параметр	Значение
τ_1 , с	0.2	τ_2 , с	0.008	τ_3 , с	0.0004	μ	0
A_1	3.0	A_2	2.5	A_3	2.0	σ	0.2

Рассчитанный частотный скан будет иметь вид, приведённый на следующем рисунке:

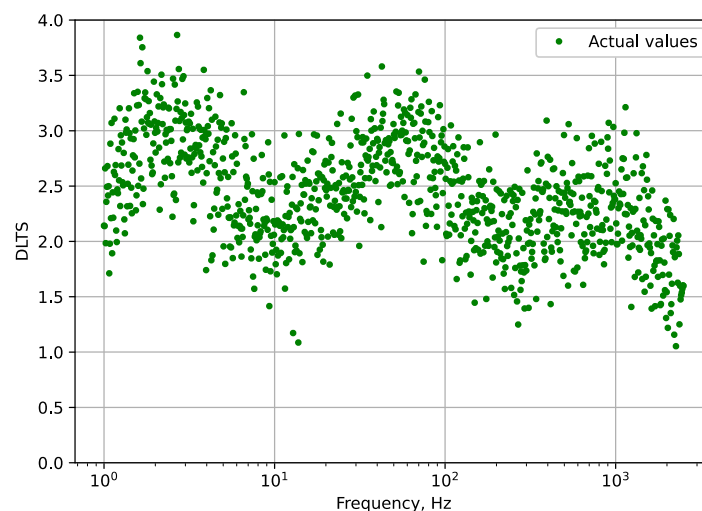


Рисунок 5 – Рассчитанный частотный скан. По горизонтали отложена частота в Гц, по вертикали – сигнал DLTS.

4.2 Пример удачной идентификации

Выполним идентификацию частотного скана аналогично тому, как это делалось в предыдущих двух частях, за исключением того, что модель теперь определяется выражением (11). При этом укажем алгоритму идентификации, что количество экспоненциальных составляющих $m = 3$.

На следующем рисунке представлены результаты идентификации:

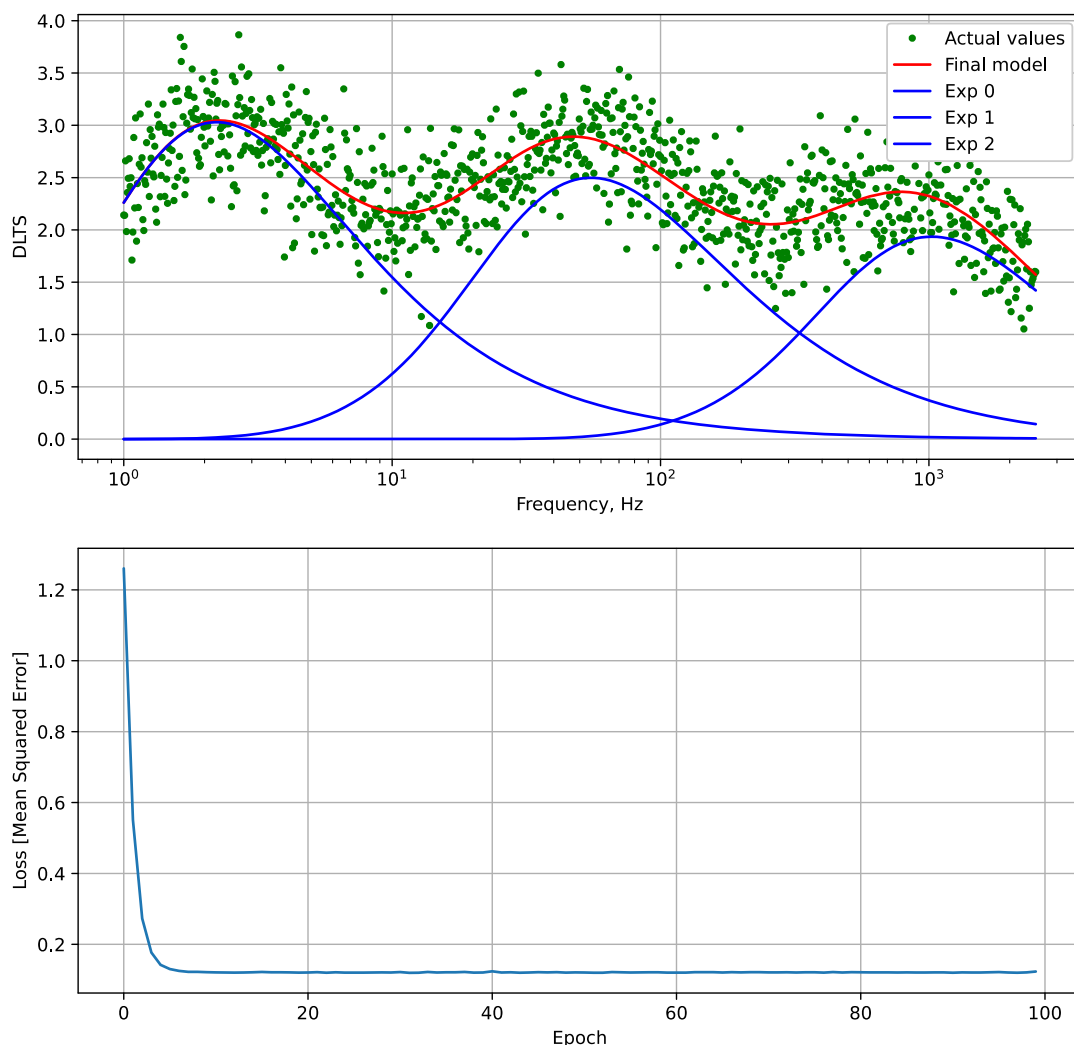


Рисунок 6 – Результаты идентификации модели. На верхнем графике показаны исходные данные (зелёные точки), частотный скан, полученный на модели после оптимизации параметров (красная линия), частотные сканы отдельных экспоненциальных составляющих после оптимизации параметров (синие линии); по горизонтали отложена частота в Гц, по вертикали – сигнал DLTS. На графике справа показана зависимость среднеквадратической ошибки от количества «эпох».

В таблице ниже приведены параметры модели после идентификации.

Таблица 6 – Параметры модели после идентификации (нечётные пары столбцов выделены для повышения читаемости).

Параметр	Значение	Параметр	Значение	Параметр	Значение	Параметр	Значение
τ_1 , с	0.0079898	τ_2 , с	0.0004082	τ_3 , с	0.1990181	\sqrt{E}	0.345909
A_1	2.4998	A_2	1.9873	A_3	3.0273	-	-

Нужно отметить, что алгоритм оптимизации считает все составляющие равноправными, поэтому порядковые номера составляющих в таблицах 5 и 6 не совпадают.

4.3 Пример неудачной идентификации №1

Ниже приведем пример неудачной идентификации. Исходные данные будут теми же, но количество экспоненциальных составляющих $m = 6$. Результаты приведены на рисунке ниже:

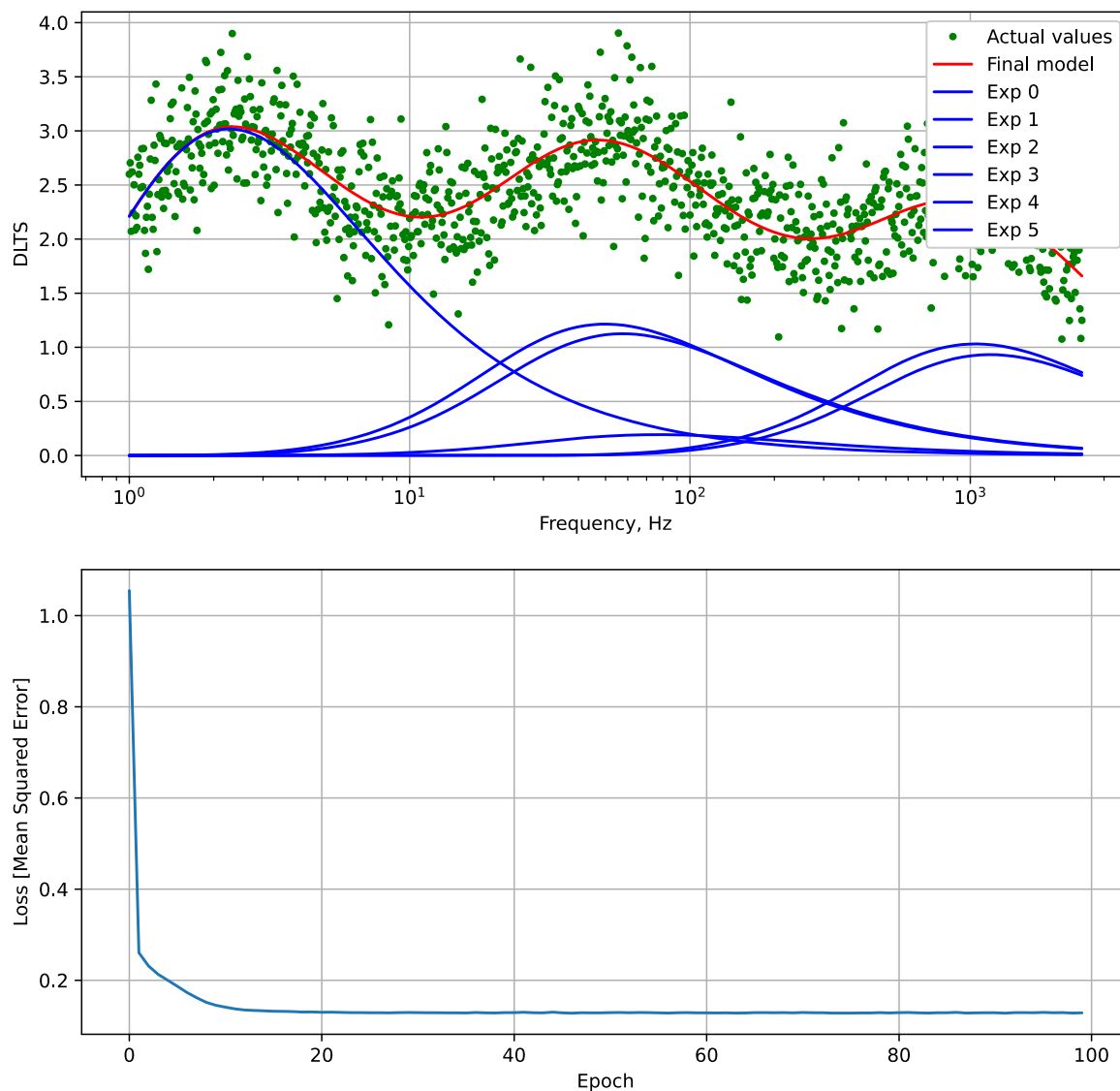


Рисунок 7 – Результаты неудачной идентификации модели. На верхнем графике показаны исходные данные (зелёные точки), частотный скан, полученный на модели после оптимизации параметров (красная линия), частотные сканы отдельных экспоненциальных составляющих после оптимизации параметров (синие линии); по горизонтали отложена частота в Гц, по вертикали – сигнал DLTS. На графике справа показана зависимость среднеквадратической ошибки от количества «эпох».

В следующей таблице приведём численные значения параметров после идентификации модели.

Таблица 7 – Параметры модели после неудачной идентификации

Параметр	Значение
τ_1 , с	0.005769
A_1	0.1933
τ_2 , с	0.007637
A_2	1.1259
τ_3 , с	0.000353
A_3	0.9613
τ_4 , с	0.194380
A_4	3.0155
τ_5 , с	0.008820
A_5	1.2150
τ_6 , с	0.000399
A_6	1.0595
\sqrt{E}	0.357646

Нужно отметить, что в таких случаях появляются частотные сканы с практически одинаковой постоянной времени или с амплитудой близкой к нулевой. Также возможно появление составляющих с отрицательными амплитудами. В дальнейшем данная особенность может служить «индикатором» того, что количество экспоненциальных составляющих нужно сократить.

Также примечательно, что корень из среднеквадратического отклонения в данном (\sqrt{E} в таблице 7) и (\sqrt{E} в таблице 6) предыдущем случаях отличаются незначительно примерно на 0,011.

4.4. Пример неудачной идентификации №2

Приведём еще один пример неудачной идентификации, все параметры исходных данных останутся прежними, но количество экспоненциальных составляющих $m = 2$.

В следующей таблице приведём численные значения параметров после идентификации модели.

Таблица 7 – Параметры модели после неудачной идентификации

Параметр	Значение
τ_1 , с	0.0035643
A_1	2.9990
τ_2 , с	0.1554596
A_2	3.1999
\sqrt{E}	0.789497

Графики, иллюстрирующие данную модель, приведены на следующем рисунке:

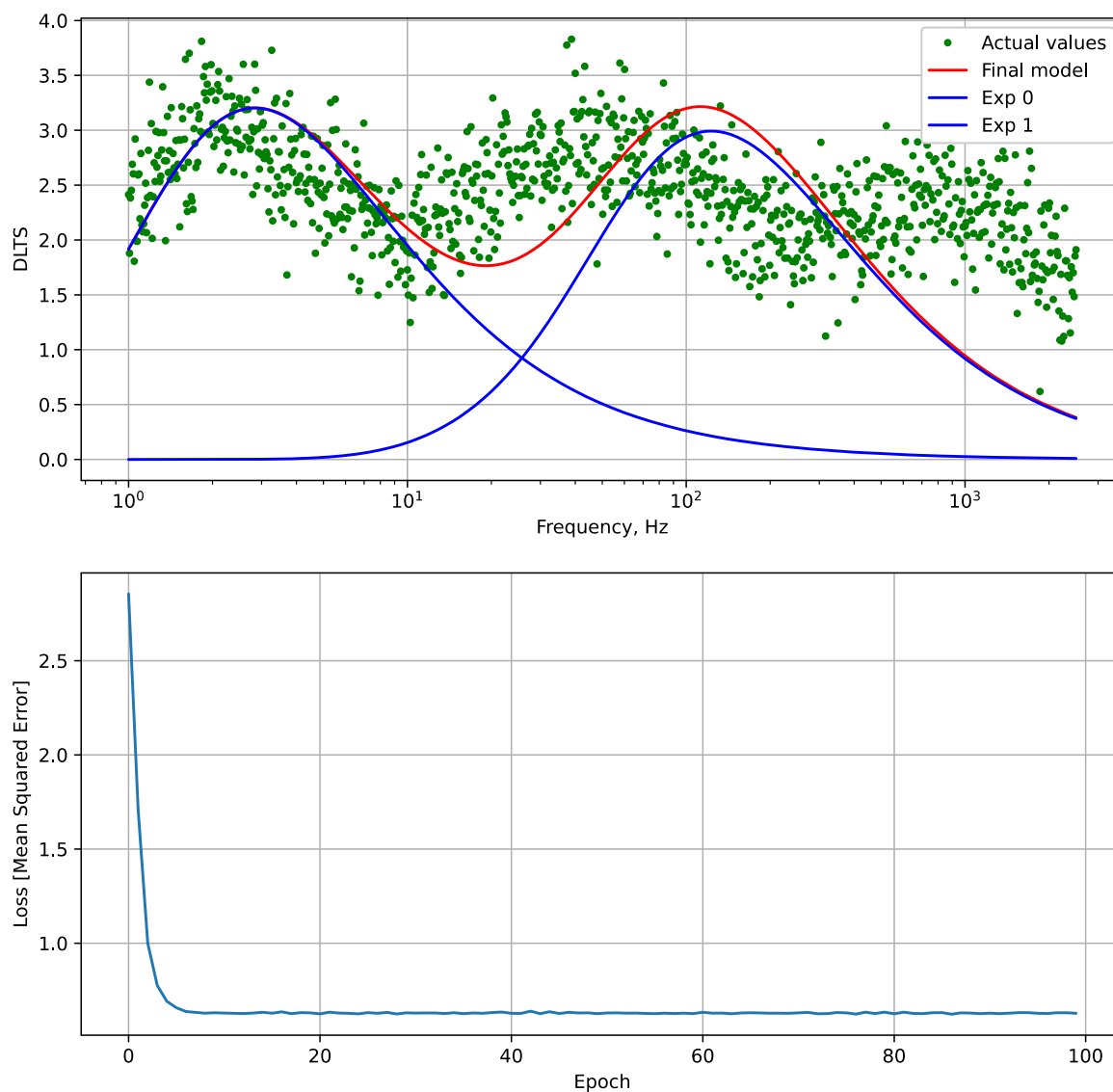


Рисунок 8 – Результаты неудачной идентификации модели. На верхнем графике показаны исходные данные (зелёные точки), частотный скан, полученный на модели после оптимизации параметров (красная линия), частотные сканы отдельных экспоненциальных составляющих после оптимизации параметров (синие линии); по горизонтали отложена частота в Гц, по вертикали – сигнал DLTS. На графике справа показана зависимость среднеквадратической ошибки от количества «эпох».

Важно, что в данном случае корень из среднеквадратического отклонения значительно выше чем в случае с $m=3$.

5. Выводы

Решены первые три задачи из поставленных в начале отчёта. При этом программа идентификации параметров частотного скана, образованного несколькими экспоненциальными составляющими нуждается в доработке, а именно, нужно решить проблему определения количества моноэкспоненциальных составляющих m .

Важно то, что m – параметр алгоритма идентификации, а не модели, к тому же его значения дискретны.

В литературе можно найти следующие рекомендации по выбору такого параметра:

- 1) Ограничить значение параметра руководствуясь какой-либо гипотезой об исходных данных.
- 2) Перебрать несколько значений параметра и сравнить значения среднеквадратических отклонений. Применительно к данной задаче, из двух значений m , дающих близкие среднеквадратические отклонения, меньшее значение предпочтительнее.
- 3) Добавить регуляризацию.
- 4) Комбинировать первые три пункта.

Также в определении параметра m может помочь переход к идентификации по нескольким частотным сканам в диапазоне температур.

Прежде чем переходить к следующему этапу работы нужно решить следующие дополнительные задачи:

- опробовать имеющиеся алгоритмы идентификации на экспериментальных данных,
- определиться с методом выбора значения m ,
- сравнить результаты идентификации модели с показателем p и результаты идентификации модели частотного скана сигнала релаксации, состоящего из суммы моноэкспоненциальных сигналов, на одинаковых тестовых и экспериментальных данных,
- разобраться с масштабным множителем M .