

# Реферат

Отчет 28 стр., 5 частей, 12 рис., 8 источников

Ключевые слова: Метод полностью параллельной разностной эволюции, Dialogue for Reverse Engineering Assessment and Methods, Поиск параметров генной регуляторной сети.

Цель работы — применение и исследование сходимости метода полностью параллельной разностной эволюции (далее ППРЭ) для нахождения параметров моделей по данным проекта DREAM (испытание 6).

Метод ППРЭ [2, 5] является модификацией стохастического метода оптимизации, предложенного в [1].

Учитывая, что для решения задачи минимизации произвольной целевой функции не существует универсального алгоритма, разработка и усовершенствование методов их решения остается актуальной задачей.

В ходе выполнения работы были реализованы три модели генных регуляторных сетей на языке R. Были проведены численные эксперименты по подбору параметров генных сетей с помощью метода ППРЭ. Проведён анализ динамики и скорости сходимости метода ППРЭ.

Генные сети важны для моделирования организма, для изучения влияния мутаций на организм, клетку, на саму сеть. Генные сети являются способом дешевой проверки воздействия веществ (лекарств) и других факторов на биосистему.

# Содержание

Реферат . . . . .	1
Введение . . . . .	3
Основная часть . . . . .	5
1.1 Метод полностью параллельной разностной эволюции . . . . .	5
1.2 Экспериментальные данные (DREAM) . . . . .	7
1.2.1 Постановка задачи . . . . .	7
1.2.2 Модели генных сетей: представление данных . . . . .	7
1.2.3 Параметры генных сетей . . . . .	8
1.2.4 Уравнения генных сетей . . . . .	10
1.2.5 Начальные данные и возмущения . . . . .	13
1.3 Численные эксперименты . . . . .	14
1.3.1 Постановка задачи в терминах метода ППРЭ . . . . .	14
1.3.2 Прогонка управляющих параметров ППРЭ . . . . .	15
1.4 Выводы . . . . .	25
Заключение . . . . .	26
Литература . . . . .	27

# Введение

Целью данной работы является применение метода ППРЭ (см. 1.1) к экспериментальным данным проекта DREAM6 (Dialogue for Reverse Engineering Assessment and Methods, см. 1.2). Областью исследования DREAM6 были параметры, оптимальные для данных генных регуляторных сетей. В качестве критерия качества алгоритмов поиска параметров была использована специальная функция расстояния между параметрами, предлагаемыми участниками, и параметрами, заранее определёнными организаторами DREAM6. Сама функция расстояния дана в описании DREAM6 [9], в работе (см. 1.3.1) она обозначена как  $\rho$ .

Динамика генной регуляторной сети описывается системой обыкновенных, и, вообще говоря, нелинейных дифференциальных уравнений (ОДУ). Так как начальные данные ОДУ определены, решение этого уравнения есть динамика концентраций мРНК и соответствующих им белков. При фиксированном интервале времени и его разбиении решение есть матрица концентраций. Эти матрицы концентраций предлагались участникам DREAM6 для поиска параметров. Под разбиением интервала здесь понимается конечная последовательность вида:  $t_1 < t_2 < t_3 < \dots < t_n$  где  $t_1$  и  $t_n$  есть первый и последний моменты времени.

Ответом или решением задачи считается набор параметров ОДУ, которые, будучи подставленными в уравнение, дадут максимально похожую матрицу концентраций на ту, что предоставлена организаторами DREAM6.

В работе для оценки ППРЭ используются две характеристики:

1. Расстояние  $\Sigma$  (Евклидово) между известной и полученной матрицами концентраций
2. Расстояние  $\rho$  между известными и полученными параметрами

Важно использовать обе эти характеристики, так как матрица концентраций, предоставленная в DREAM6, была получена с помощью подстановки параметров в ОДУ, соответствующее генной сети, с произвольно происходящими делециями гена (не более одного), мРНК нокдаунами и изменениями активности сайтов связывания рибосом (см. 1.2.5). А это значит, что набор параметров, предоставленный как ответ, будучи подставленным в систему ОДУ, не даст матрицы концентраций, совпадающей с данной.

По этой причине метод ППРЭ может подобрать набор параметров, близкий к оптимальным, т.е. минимизирующий  $\Sigma$ , но при этом не являющийся тем самым набором параметров, который был дан в DREAM6.

Такое вполне возможно в силу большого количества параметров. Более формальное об этом сказано в разделе 1.3.

Таким образом, динамика первой характеристики ( $\Sigma$ ) будет отражать скорость сходимости метода ППРЭ к оптимальному значению, а вторая ( $\rho$ ) — близость найденных и известных параметров для каждого шага метода ППРЭ.

# Основная часть

## 1.1 Метод полностью параллельной разностной эволюции

Метод ППРЭ успешно применялся в различных задачах [3, 4]. Совершенствование методов минимизации для нахождения параметров генных регуляторных сетей требует наличия набора тестов для оценки новых алгоритмов и реализаций и сравнения с предшествующими.

В самом общем смысле класс рассматриваемых задач можно назвать задачами поиска глобального минимума некоторого функционала качества (или функции). Способов решения таких задач достаточно много. В работе рассматривается модификация стохастического итерационного метода разностной эволюции (РЭ).

Идея метода РЭ, предложенного Р. Сторном [1], заключается в моделировании популяции индивидуумов (а точнее, векторов их определяющих). Популяция меняется от поколения к поколению, при этом индивидуумы скрещиваются и мутируют.

Метод РЭ имеет набор управляющих параметров (например, размер популяции или количество старейших индивидуумов, заменяемых на новые, способ рекомбинации при получении нового вектора-индивидуума из нескольких существующих), от которых зависит скорость работы и сходимость. Возраст индивидуума — количество итераций, которые он существует. В [7] была предложена адаптивная схема выбора управляющих параметров метода РЭ. Суть адаптивной схемы заключается в том, что процесс создания нового вектора-индивидуума есть не фиксированная функция, а зависящая от значений самого вектора, т.е. от функционала качества для данного индивидуума.

В работе [8] введено тригонометрическое преобразование (мутация) вектора-индивидуума, зависящее от функционала качества.

В данной работе рассматривается реализация метода РЭ, поддерживающая распараллеливание процесса вычислений — метод ППРЭ [2, 5]. Под распараллеливанием здесь понимается вычисление в несколько потоков на многопроцессорных системах. Каждый индивидуум помещается в очередь, из которой он извлекается и независимо обрабатывается каждым потоком, что позволяет существенно увеличить скорость поиска при вычислениях на нескольких процессорах.

В методе ППРЭ предусмотрено несколько вариантов рекомбинации, за которые отвечает параметр *recombination\_strategy*. Спустя определённое количество итераций осуществляется замена старых индивидуумов на новые, за количество заменяемых индивидуумов отвечает параметр *es\_lambda*. Параметр *population\_size* отвечает за общее количество индивидуумов в популяции.

На мой взгляд, эти параметры существенно влияют на скорость сходимости, поэтому именно их вариация рассматривается в работе.

## 1.2 Экспериментальные данные (DREAM)

Проект DREAM предоставляет унифицированные экспериментальные данные для тестирования алгоритмов. Каждое «испытание» — некая формализованная задача, которую предлагается решить независимым группам исследователей. Лучшие решения и результаты публикуются. [6].

В рамках этой работы требуется подобрать близкие к оптимальным значения параметров ППРЭ, используя в качестве тестовых задач результаты DREAM6.

### 1.2.1 Постановка задачи

Задача принадлежит области обратной инженерии генных регуляторных сетей. Предполагается, что топология генной сети уже определена с достаточным уровнем правдоподобия. Под топологией здесь понимается только структура сети. Задача заключается в подборе параметров (при определении которых будет определена и кинетика) этой сети.

Здесь есть два ключевых аспекта, которые требуют внимания: задача оценки (подбора) параметров модели, при известной топологии модели, а также биологическая задача проектирования наиболее информативных экспериментов для получения неизвестных параметров.

Итак, даны структуры трёх генных регуляторных сетей, от участников требуется разработать и/или применять методы оптимизации, чтобы точно оценить параметры моделей, а так же прогнозировать результаты возмущений в этих моделях.

Эти две задачи и являются областью исследования DREAM6. Однако, для тестирования ППРЭ потребуется рассмотреть лишь первую задачу, т.е. подбор параметров при известной топологии.

### 1.2.2 Модели генных сетей: представление данных

В DREAM6 представлены топологии трёх моделей генных сетей, для которых требуется оценить параметры. Структуры (только топология, без параметров) генных сетей представлены в форматах sbml, tic (текстовые форматы для хранения структуры генных сетей), и в графическом формате. Пример графического представления для первой генной сети приведён на рисунке 1.1.

Для каждой сети предоставляется файл (.m) с описанием модели в синтаксисе MATLAB. Все переменные помечены в соответствии с их типом. Например, переменные, означающие концентрацию белка, помечены как  $p1$ ,  $p2$ , ...  $p6$ .

Значения каждого символа в генной сети объясняются в легенде (рис. 1.2). В скобках перечислены префиксы к переменным модели. Линии, соединяющие кодирующую белок последовательность с белком, обозначены префиксом «pp». Транскрипция и трансляция (рис. 1.3) не показаны на основной диаграмме.

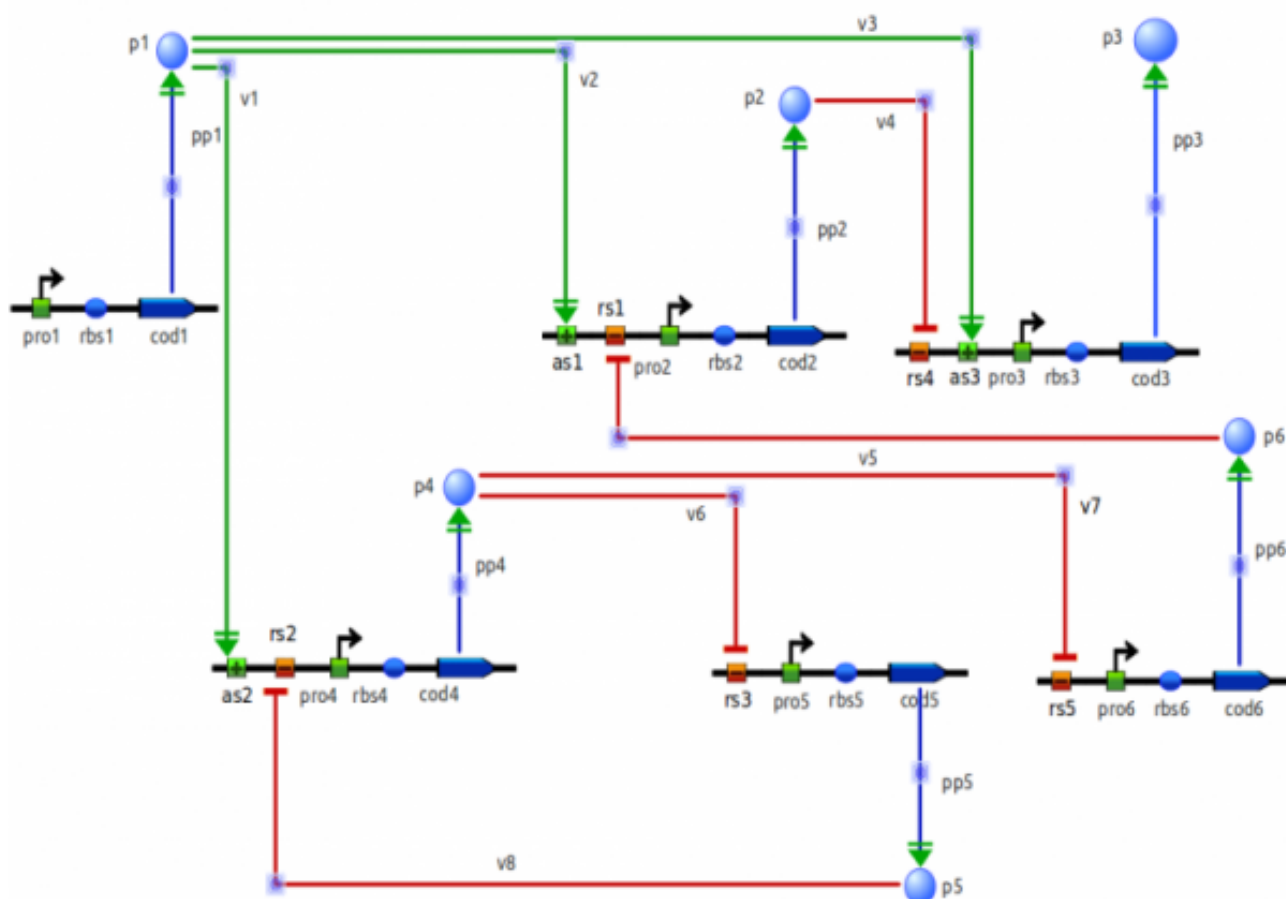


Рисунок 1.1: Пример графического представления для первой генной сети

Имена переменных для концентраций мРНК, результата транскрипции кодирующей последовательности имеют соответствующий префикс «pp». Например, переменная, соответствующая концентрации мРНК с номером 3 будет именована как *pp3\_trna*.

### 1.2.3 Параметры генных сетей

Генная сеть характеризуется топологией (структурой), о которой говорилось выше, и набором параметров — скорость трансляции, транскрипции, и параметров, отвечающих за сайты связывания рибосом. Если все эти параметры и начальные данные (начальные концентрации мРНК, белков) опре-



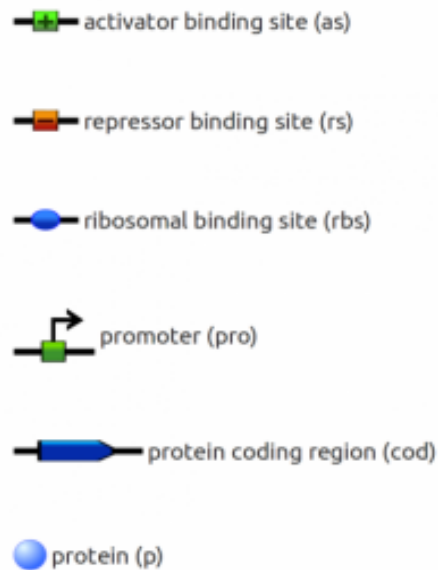


Рисунок 1.2: Аннотация к графическому представлению

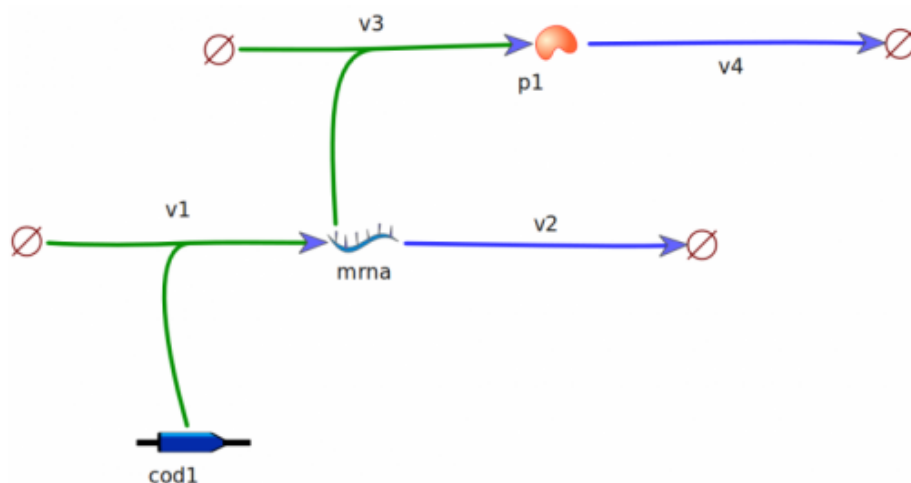


Рисунок 1.3: Транскрипция и трансляция, не показанные на схеме генной сети

делены, рассматривается динамика концентраций мРНК и белков каждого типа.

Таким образом, каждой генной сети с заданными параметрами соответствует ОДУ, решение которого в конкретном интервале времени порождает матрицу, содержащую набор концентраций для фиксированных моментов времени.

### 1.2.4 Уравнения генных сетей

Все параметры деградации мРНК и белков, имена переменных которых имеют вид:  $pp\{\dots\}_{degradation\_rate}$  было принято считать одинаковыми в рамках проекта DREAM. Поэтому в ОДУ для моделей эти параметры обозначены как  $degradation\_rate$ .

Система ОДУ для модели 1:

$$\begin{aligned}
\frac{d}{dt}[pp1\_mrna] &= pro1\_strength - [pp1\_mrna]; \\
\frac{d}{dt}[pp2\_mrna] &= pro2\_strength \cdot \frac{\left(\frac{[p1]}{v2\_Kd}\right)^{v2\_h}}{1 + \left(\frac{[p1]}{v2\_Kd}\right)^{v2\_h}} \cdot \frac{1}{1 + \left(\frac{[p6]}{v5\_Kd}\right)^{v5\_h}} \\
&\quad \cdot degradation\_rate - [pp2\_mrna]; \\
\frac{d}{dt}[pp3\_mrna] &= pro3\_strength \cdot \frac{\left(\frac{[p1]}{v3\_Kd}\right)^{v2\_h}}{1 + \left(\frac{[p1]}{v3\_Kd}\right)^{v3\_h}} \cdot \frac{1}{1 + \left(\frac{[p2]}{v4\_Kd}\right)^{v4\_h}} \\
&\quad \cdot degradation\_rate - [pp3\_mrna]; \\
\frac{d}{dt}[pp4\_mrna] &= pro4\_strength \cdot \frac{\left(\frac{[p1]}{v1\_Kd}\right)^{v2\_h}}{1 + \left(\frac{[p1]}{v1\_Kd}\right)^{v1\_h}} \cdot \frac{1}{1 + \left(\frac{[p5]}{v8\_Kd}\right)^{v8\_h}} \\
&\quad \cdot degradation\_rate - [pp4\_mrna]; \\
\frac{d}{dt}[pp5\_mrna] &= pro5\_strength \cdot \frac{1}{1 + \left(\frac{[p4]}{v6\_Kd}\right)^{v6\_h}} \cdot degradation\_rate \\
&\quad - [pp5\_mrna]; \\
\frac{d}{dt}[pp6\_mrna] &= pro6\_strength \cdot \frac{1}{1 + \left(\frac{[p4]}{v7\_Kd}\right)^{v7\_h}} \cdot degradation\_rate \\
&\quad - [pp6\_mrna]; \\
\frac{d}{dt}[p1] &= rbs1\_strength \cdot [pp1\_mrna] - degradation\_rate \cdot [p1]; \\
\frac{d}{dt}[p2] &= rbs2\_strength \cdot [pp2\_mrna] - degradation\_rate \cdot [p2]; \\
\frac{d}{dt}[p3] &= rbs3\_strength \cdot [pp3\_mrna] - degradation\_rate \cdot [p3]; \\
\frac{d}{dt}[p4] &= rbs4\_strength \cdot [pp4\_mrna] - degradation\_rate \cdot [p4]; \\
\frac{d}{dt}[p5] &= rbs5\_strength \cdot [pp5\_mrna] - degradation\_rate \cdot [p5]; \\
\frac{d}{dt}[p6] &= rbs6\_strength \cdot [pp6\_mrna] - degradation\_rate \cdot [p6];
\end{aligned}$$

Система ОДУ для модели 2:

$$\begin{aligned}
\frac{d}{dt}[pp1\_mrna] &= pro1\_strength - degradation\_rate \cdot [pp1\_mrna]; \\
\frac{d}{dt}[pp2\_mrna] &= pro2\_strength \cdot \left( \frac{(\frac{[p1]}{v1\_Kd})^{v1\_h}}{1 + (\frac{[p1]}{v1\_Kd})^{v1\_h}} + \frac{(\frac{[p2]}{v3\_Kd})^{v3\_h}}{1 + (\frac{[p2]}{v3\_Kd})^{v3\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp2\_mrna]; \\
\frac{d}{dt}[pp3\_mrna] &= pro3\_strength \cdot \left( \frac{(\frac{[p1]}{v9\_Kd})^{v9\_h}}{1 + (\frac{[p1]}{v9\_Kd})^{v9\_h}} + \frac{(\frac{[p2]}{v10\_Kd})^{v10\_h}}{1 + (\frac{[p2]}{v10\_Kd})^{v10\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp3\_mrna]; \\
\frac{d}{dt}[pp4\_mrna] &= pro4\_strength \cdot \frac{1}{1 + (\frac{[p3]}{v2\_Kd})^{v2\_h}} \\
&\quad - degradation\_rate \cdot [pp4\_mrna]; \\
\frac{d}{dt}[pp5\_mrna] &= pro5\_strength \cdot \frac{(\frac{[p2]}{v4\_Kd})^{v4\_h}}{1 + (\frac{[p2]}{v4\_Kd})^{v4\_h}} \cdot \frac{1}{1 + (\frac{[p5]}{v5\_Kd})^{v5\_h}} \\
&\quad - degradation\_rate \cdot [pp5\_mrna]; \\
\frac{d}{dt}[pp6\_mrna] &= pro6\_strength \cdot \frac{1}{1 + (\frac{[p4]}{v6\_Kd})^{v6\_h}} \\
&\quad - degradation\_rate \cdot [pp6\_mrna]; \\
\frac{d}{dt}[pp7\_mrna] &= pro7\_strength \cdot \frac{(\frac{[p7]}{v8\_Kd})^{v8\_h}}{1 + (\frac{[p7]}{v8\_Kd})^{v8\_h}} \cdot \frac{1}{1 + (\frac{[p6]}{v7\_Kd})^{v7\_h}} \\
&\quad - degradation\_rate \cdot [pp7\_mrna]; \\
\frac{d}{dt}[p1] &= rbs1\_strength \cdot [pp1\_mrna] - degradation\_rate \cdot [p1]; \\
\frac{d}{dt}[p2] &= rbs2\_strength \cdot [pp2\_mrna] - degradation\_rate \cdot [p2]; \\
\frac{d}{dt}[p3] &= rbs3\_strength \cdot [pp3\_mrna] - degradation\_rate \cdot [p3]; \\
\frac{d}{dt}[p4] &= rbs4\_strength \cdot [pp4\_mrna] - degradation\_rate \cdot [p4]; \\
\frac{d}{dt}[p5] &= rbs5\_strength \cdot [pp5\_mrna] - degradation\_rate \cdot [p5]; \\
\frac{d}{dt}[p6] &= rbs6\_strength \cdot [pp6\_mrna] - degradation\_rate \cdot [p6]; \\
\frac{d}{dt}[p7] &= rbs7\_strength \cdot [pp7\_mrna] - degradation\_rate \cdot [p7];
\end{aligned}$$

Система ОДУ для модели 3:

$$\begin{aligned}
\frac{d}{dt}[pp1\_mrna] &= pro1\_strength - degradation\_rate \cdot [pp1\_mrna]; \\
\frac{d}{dt}[pp2\_mrna] &= pro2\_strength \cdot \left( \frac{(\frac{[p1]}{v1\_Kd})^{v1\_h}}{1 + (\frac{[p1]}{v1\_Kd})^{v1\_h}} \cdot \frac{1}{1 + (\frac{[p9]}{v13\_Kd})^{v13\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp2\_mrna]; \\
\frac{d}{dt}[pp3\_mrna] &= pro3\_strength \cdot \left( \frac{1}{1 + (\frac{[p2]}{v2\_Kd})^{v2\_h}} \cdot \frac{1}{1 + (\frac{[p3]}{v3\_Kd})^{v3\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp3\_mrna]; \\
\frac{d}{dt}[pp4\_mrna] &= pro4\_strength \cdot \left( \frac{1}{1 + (\frac{[p3]}{v15\_Kd})^{v15\_h}} \cdot \frac{1}{1 + (\frac{[p2]}{v14\_Kd})^{v14\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp4\_mrna]; \\
\frac{d}{dt}[pp5\_mrna] &= pro5\_strength \cdot \frac{(\frac{[p4]}{v4\_Kd})^{v4\_h}}{1 + (\frac{[p4]}{v4\_Kd})^{v4\_h}} \\
&\quad - degradation\_rate \cdot [pp5\_mrna]; \\
\frac{d}{dt}[pp6\_mrna] &= pro6\_strength \cdot \left( \frac{(\frac{[p5]}{v5\_Kd})^{v5\_h}}{1 + (\frac{[p5]}{v5\_Kd})^{v5\_h}} + \frac{(\frac{[p6]}{v6\_Kd})^{v6\_h}}{1 + (\frac{[p6]}{v6\_Kd})^{v6\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp6\_mrna]; \\
\frac{d}{dt}[pp7\_mrna] &= pro7\_strength \cdot \left( \frac{(\frac{[p6]}{v8\_Kd})^{v8\_h}}{1 + (\frac{[p6]}{v8\_Kd})^{v8\_h}} + \frac{(\frac{[p5]}{v9\_Kd})^{v9\_h}}{1 + (\frac{[p5]}{v9\_Kd})^{v9\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp7\_mrna]; \\
\frac{d}{dt}[pp8\_mrna] &= pro8\_strength \cdot \left( \frac{(\frac{[p7]}{v7\_Kd})^{v7\_h}}{1 + (\frac{[p7]}{v7\_Kd})^{v7\_h}} \cdot \frac{1}{1 + (\frac{[p8]}{v11\_Kd})^{v11\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp8\_mrna]; \\
\frac{d}{dt}[pp9\_mrna] &= pro9\_strength \cdot \left( \frac{(\frac{[p7]}{v10\_Kd})^{v10\_h}}{1 + (\frac{[p7]}{v10\_Kd})^{v10\_h}} \cdot \frac{1}{1 + (\frac{[p8]}{v12\_Kd})^{v12\_h}} \right) \\
&\quad - degradation\_rate \cdot [pp9\_mrna]; \\
\frac{d}{dt}[p\{i\}] &= rbs\{i\}\_strength \cdot [pp\{i\}\_mrna] - degradation\_rate \cdot [p\{i\}];
\end{aligned}$$

Для  $\{i\} = 1, \dots, 9$ .

### 1.2.5 Начальные данные и возмущения

Наборы данных, которые предоставляются в качестве входных для поиска параметров, были сформированы искусственно, путем моделирования, с учётом различных возмущений (зашумлений) в генной сети — делеции гена, мРНК нокдаун и изменение активности сайтов связывания рибосом.

Оговорено, что во всех случаях возмущения могут затрагивать только один ген. Удаление гена приводит к полной ликвидации как мРНК, так и белка целевого гена. В случае миРНК, мРНК деградирует (фиксированное уменьшение в 5 раз), что приводит к уменьшению как мРНК, так и концентрации соответствующего белка.

0.0	0.0	0.0	0.041	0.16	0.189	0.048
2.0	2.754	4.01	4.531	0.30	0.221	0.006
4.0	2.958	2.96	0.911	0.06	0.522	0.39
6.0	4.058	2.18	0.457	0.07	1.609	1.266
8.0	3.41	1.06	0.649	0.08	2.627	2.253
10.0	3.459	0.68	4.398	0.07	2.979	3.811
12.0	2.453	0.67	6.734	0.27	2.618	2.983
14.0	1.234	0.43	5.971	0.02	2.443	3.025
16.0	2.385	0.43	4.606	0.0	1.821	2.823
18.0	3.691	0.52	5.827	0.0	3.444	2.386
20.0	3.252	0.4	8.947	0.0	4.358	3.666

Таблица 1.1: Пример таблицы концентраций мРНК для первой генной сети, предоставленной в DREAM6. В таблице отражена динамика концентраций шести мРНК (соответствуют вертикальным графам), в интервале времени от 0 до 20 секунд (первая графа) с разбиением по 0.2с

## 1.3 Численные эксперименты

### 1.3.1 Постановка задачи в терминах метода ППРЭ

Метод ППРЭ ищет минимум функционала качества по списку параметров. Параметры — неопределённые параметры генной сети, о которых говорилось в предыдущей главе.

За функционал качества выбирается расстояние между заранее определённой матрицей концентраций  $W$  (см. 1.2.3) и матрицей концентраций, полученной с текущими параметрами. При этом расстояние понимается как сумма квадратов поэлементных разностей двух матриц.

Как уже было сказано, в качестве оценки работы ППРЭ используются две характеристики:

1. Расстояние между известной и полученной матрицами концентраций. Т.е. функционал качества.
2. Расстояние между известными и полученными параметрами

Более формально: генная сеть, при выборе вектора параметров  $p$  и задании вектора начальных условий  $e$  порождает дифференциальное уравнение  $ODE(p, e)$ . Решение этого уравнения есть динамика концентраций мРНК и соответствующих им белков. При фиксированном интервале времени и разбиении решение есть матрица концентраций  $M^{(p,e)}$ .

Разбиение интервала времени состоит из 11 точек на расстоянии 0.2с от 0с до 20с включительно:  $0, 0.2, 0.4, \dots, 20$ .

В матрице  $M$  столбцы соответствуют концентрации мРНК и их количество совпадает с количеством мРНК в генной сети, а строки соответствуют моментам времени, их число фиксировано (11).

$$(p, e) \rightarrow ODE(p, e) \rightarrow M^{(p,e)}$$

Так как вектор начальных условий  $e$  неизменен и определён, конструкция упрощается:

$$p \rightarrow ODE(p) \rightarrow M^p$$

Функционал качества для метода ППРЭ есть Евклидово расстояние между матрицами концентраций, а  $\rho$ , определённое организаторами DREAM6, есть:

$$\rho(p, p^*) = \frac{1}{N} \sum_{i=1}^N \ln(p_i/p_i^*)^2$$

Где  $N$  — размерность векторов  $p$  (текущие параметры) и  $p^*$  (известные параметры). Теперь каждый вектор параметров  $p$  порождает два числа (две характеристики, о которых говорилось выше):

$$p \rightarrow ODE(p) \rightarrow M^p \rightarrow \left\{ \sum_{i,j} (M_{i,j}^p - W_{i,j})^2, \rho(p, p^*) \right\}$$

### 1.3.2 Прогонка управляющих параметров ППРЭ

В качестве критерия остановки было выбрано время, прошедшее с момента начала работы. Для каждого запуска выделялось 900 секунд. Для всех трёх моделей был зафиксирован параметр  $population\_size = 150$  ( $p.size$ ) и варьировалось  $es\_lambda = 2, 15, 45$  ( $es\_l.$ ). Всего было проведено 12 запусков для каждой модели.

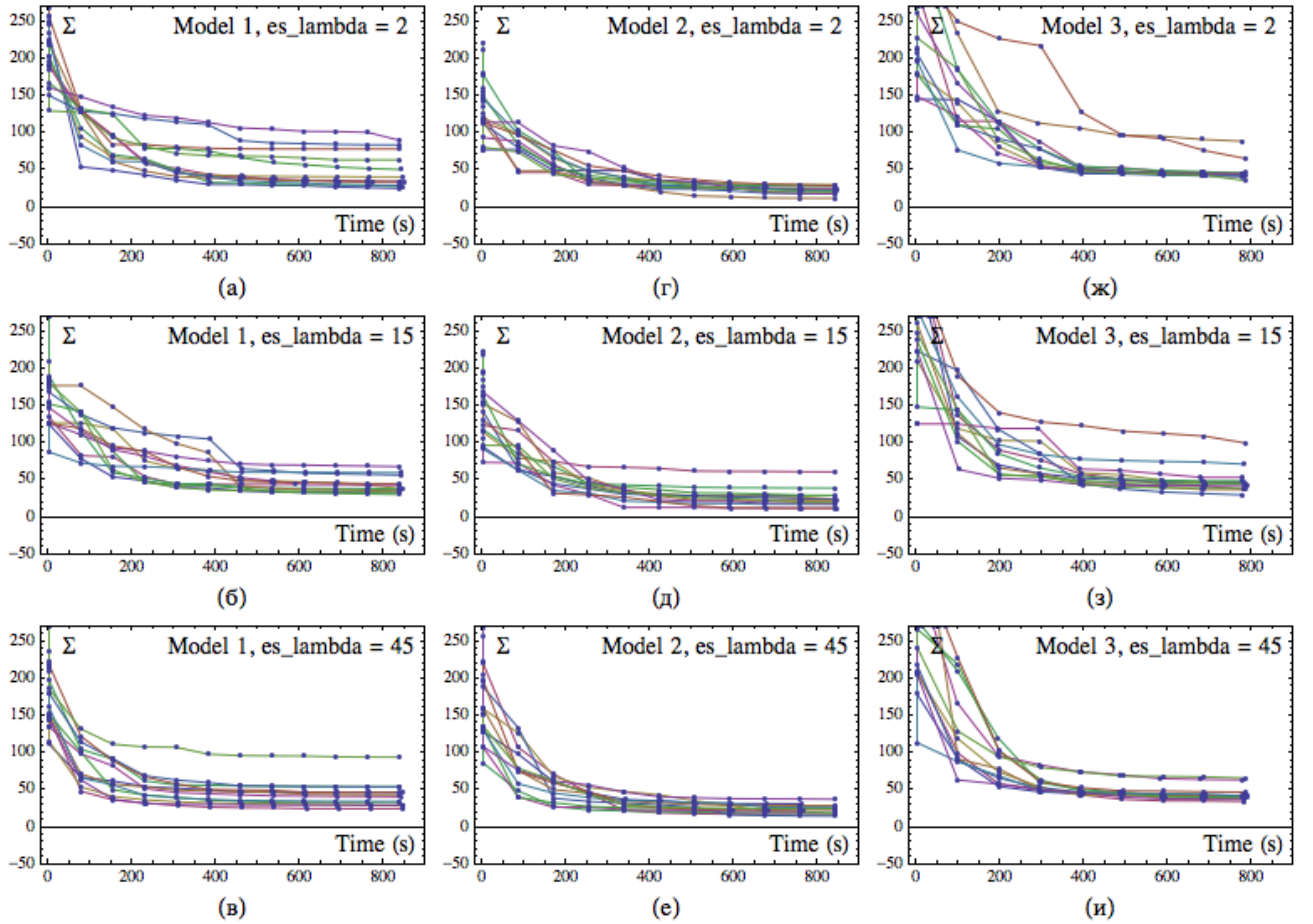


Рисунок 1.4: Динамика изменения  $\Sigma$ . Три модели. 12 независимых запусков.  $population\_size = 150$ . Модели генных сетей (1, 2, 3),  $es\_lambda = 2, 15, 45$ .

p.size	es_1.	Модель 1		Модель 2		Модель 3	
		Среднее	Мин.	Среднее	Мин.	Среднее	Мин.
150	2	49.3667	25.4139	22.8473	12.1157	48.6207	35.7457
	15	44.0293	30.755	25.8365	12.1114	49.8181	30.1344
	45	44.3788	24.9287	24.1793	15.0964	44.8327	34.6884
$\Sigma$ для $p^*$		46.51614		20.43136		55.38312	

Таблица 1.2: Средние и минимальные значения  $\Sigma$

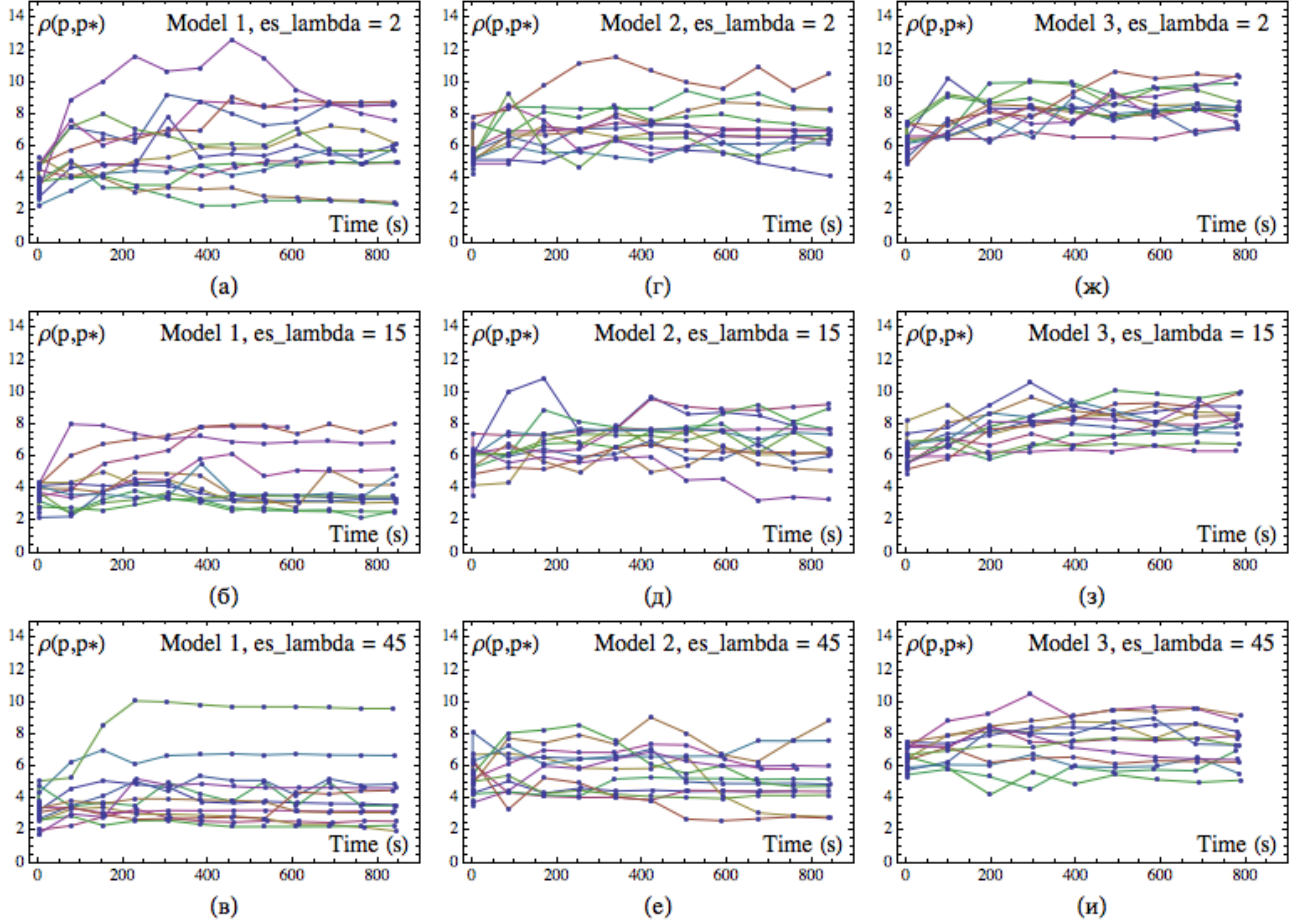


Рисунок 1.5: Динамика изменения  $\rho$ . Три модели. 12 независимых запусков.  $population\_size = 150$ . Модели генных сетей (1, 2, 3),  $es\_lambda = 2, 15, 45$ .



Приведённое в таблице (1.2)  $\Sigma$  — значение функционала качества для известных параметров  $p^*$ , т.е. значение функционала качества, которое считается приемлемым в DREAM6. Следовательно, всякий вектор параметров  $p$ , для которого в данной модели  $\Sigma$  меньше определённого в таблице, можно считать оптимальным. (динамика изменений  $\Sigma$  отражена на графиках 1.4).

Из графиков и таблицы очевидно, что в среднем значение минимизируемого функционала сходится своему минимальному значению. Кроме того, всегда находился набор параметров — вектор  $p$ , значение  $\Sigma$  для которого меньше значения  $\Sigma$  для  $p^*$ .

Рассмотрим изменение расстояния  $\rho$  между известными параметрами генной сети (предоставленными) и параметрами, найденными методом ППРЭ — графики 1.5.

Из графиков видно, что расстояние  $\rho$  между  $p^*$  и  $p$  не стремится к нулю. Следовательно, вектор, который находит ППРЭ не совпадает с вектором  $p^*$ . Надо полагать, что это происходит из-за большого количества (свободных) параметров ОДУ.

Посмотрим, что из себя представляют матрицы (таблицы) концентраций при использовании вектора параметров  $p$  и  $p^*$ . Однако, для наглядности рассмотрим не сами матрицы  $M^p$ , а их отличие (разности) от матрицы  $W$ , т.е.  $|M - W|$ .

Далее приведены графики модуля разности  $|M - W|$ , а так же таблицы найденных  $p_{min}$  и известных  $p^*$  параметров, при которых достигается минимум функционала:

0.	0.	0.041	0.166	0.189	0.048	0.	0.	0.041	0.166	0.189	0.048
0.16	0.093	0.94	0.15	0.02	0.235	0.163	1.119	0.258	0.151	0.221	0.028
0.013	0.035	0.103	0.035	0.563	0.695	0.017	0.391	0.117	0.127	0.521	0.
1.065	0.874	1.204	0.074	0.852	1.195	1.069	0.258	0.275	0.011	0.307	0.315
0.411	0.366	3.889	0.087	0.286	0.66	0.415	0.203	0.183	0.075	0.079	0.256
0.459	0.142	1.301	0.073	0.009	0.823	0.463	0.103	0.193	0.071	0.166	1.009
0.547	0.173	0.807	0.277	0.38	0.015	0.543	0.034	0.583	0.277	0.209	0.122
1.766	0.067	0.003	0.024	0.557	0.025	1.762	0.177	0.452	0.024	0.386	0.155
0.615	0.065	1.369	0.	1.179	0.177	0.611	0.168	1.86	0.	1.009	0.049
0.691	0.027	0.15	0.	0.444	0.614	0.695	0.075	0.647	0.	0.614	0.486
0.252	0.096	2.97	0.	1.358	0.666	0.256	0.197	2.472	0.	1.528	0.794

Рисунок 1.6: Модель 1. Матрица концентраций получена параметрами  $p^*$ , предлагаемыми DREAM6 в качестве ответа.  
 $|W - M^{p^*}|$

Рисунок 1.7: Модель 1. Матрица концентраций получена подстановкой параметров  $p_{min}$ , полученных методом ППРЭ.  $|W - M^{p_{min}}|$

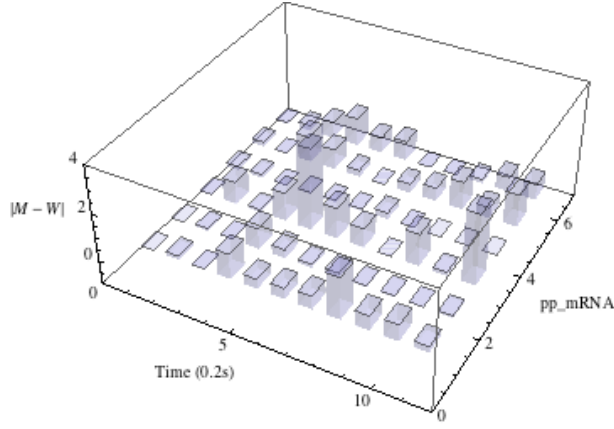


Рисунок 1.8: График таблицы 1.6 для  $p^*$

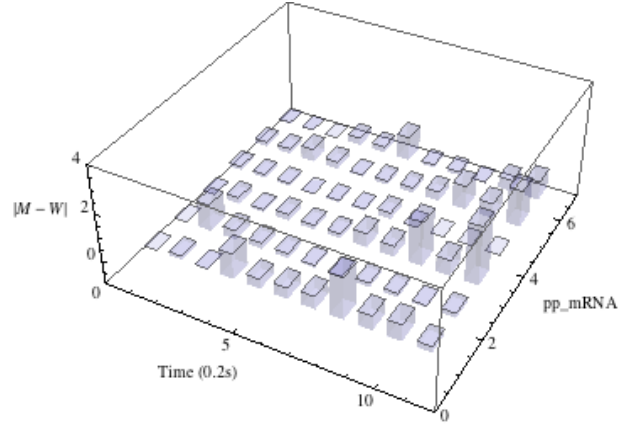


Рисунок 1.9: График таблицы 1.7 для  $p_{min}$

0.	0.	0.044	0.	0.	0.114	0.
0.132	0.319	0.177	0.896	0.504	0.252	0.219
0.102	0.187	0.303	0.143	0.176	0.173	1.235
0.441	1.121	0.2	0.013	0.01	0.288	1.467
0.152	1.236	0.413	0.003	0.256	1.086	0.444
0.142	0.434	0.428	0.059	0.011	1.122	0.085
0.3	0.722	0.979	0.	0.119	1.893	0.006
0.233	0.39	0.739	0.063	0.593	0.052	0.117
0.251	0.648	0.236	0.063	0.103	0.262	0.203
0.071	0.097	0.487	0.166	0.264	0.574	0.013
0.006	0.574	0.886	0.024	0.067	0.213	0.

Таблица 1.3: Модель 2. Матрица концентраций получена параметрами  $p^*$ , предлагаемыми DREAM6 в качестве ответа.  $|W - M^{p^*}|$

0.	0.	0.044	0.	0.	0.114	0.
0.031	0.388	0.171	0.583	0.027	0.432	0.227
0.287	0.602	0.401	0.495	0.2	0.263	0.548
0.629	0.895	0.794	0.004	0.053	0.154	0.005
0.037	1.08	0.1	0.06	0.295	0.597	0.307
0.046	0.568	0.056	0.121	0.028	0.662	0.067
0.112	0.848	0.504	0.061	0.081	1.437	0.009
0.045	0.266	0.266	0.003	0.555	0.575	0.117
0.063	0.525	0.236	0.003	0.065	0.248	0.203
0.259	0.026	0.015	0.226	0.302	0.071	0.013
0.194	0.697	0.414	0.036	0.029	0.713	0.

Таблица 1.4: Модель 2. Матрица концентраций получена подстановкой параметров  $p_{min}$ , полученных методом ППРЭ.  $|W - M^{p_{min}}|$

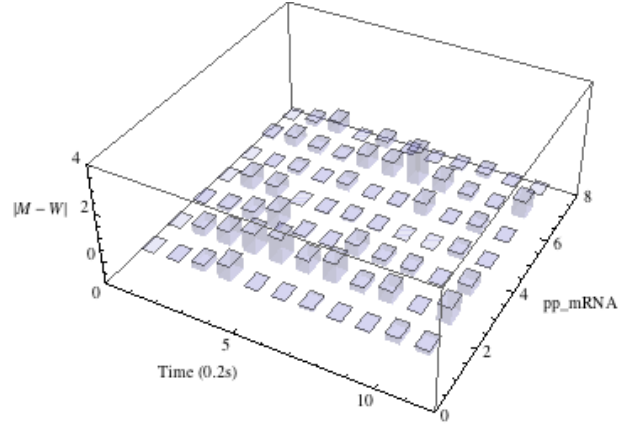
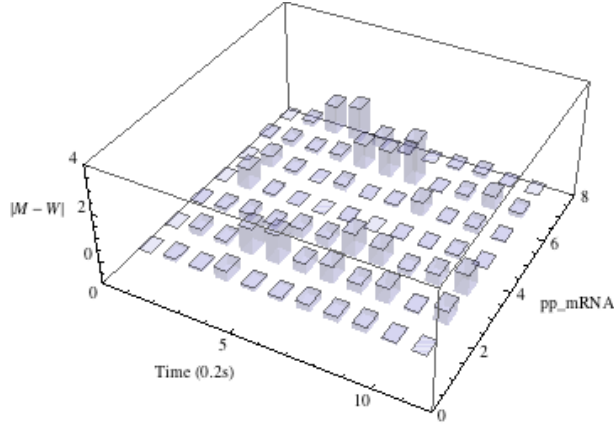


Рисунок 1.10: График таблицы 1.3 для  $p^*$     Рисунок 1.11: График таблицы 1.4 для  $p_{min}$

0.072	0.	0.127	0.074	0.011	0.	0.	0.08	0.
0.19	1.026	0.53	0.882	0.636	0.892	1.003	1.472	1.741
0.052	0.021	3.475	3.462	0.015	0.117	0.258	0.219	0.581
0.058	0.003	1.552	1.577	0.101	0.305	0.562	0.021	0.158
0.484	0.	0.596	0.534	0.46	1.058	0.018	0.042	0.021
0.074	0.104	1.045	0.1	1.25	0.668	0.201	0.154	0.066
0.254	0.01	0.403	0.458	0.486	0.381	0.596	0.146	0.154
0.013	0.001	0.921	0.217	1.75	0.06	0.389	0.009	0.053
0.25	0.039	0.416	1.371	1.002	0.311	0.463	0.154	0.016
0.692	0.	0.385	0.317	0.427	0.489	0.147	0.075	0.101
0.369	0.088	0.42	0.291	0.24	0.193	0.191	0.064	0.06

Таблица 1.5: Модель 3. Матрица концентраций получена параметрами  $p^*$ , предлагаемыми DREAM6 в качестве ответа.  $|W - M^{p^*}|$

0.072	0.	0.127	0.074	0.011	0.	0.	0.08	0.
0.049	0.02	1.397	1.093	0.816	0.432	1.028	1.721	1.602
0.212	0.755	1.551	0.471	0.102	0.353	0.046	0.081	0.522
0.221	0.384	0.444	0.486	0.134	0.099	0.288	0.154	0.126
0.321	0.248	0.301	1.002	0.587	0.856	0.271	0.179	0.06
0.09	0.111	1.077	0.14	1.424	0.869	0.492	0.291	0.026
0.091	0.198	0.429	0.248	0.299	0.18	0.304	0.283	0.113
0.176	0.206	0.959	0.01	1.56	0.261	0.097	0.146	0.012
0.413	0.168	0.376	1.165	0.811	0.512	0.755	0.291	0.057
0.529	0.207	0.344	0.111	0.236	0.288	0.145	0.212	0.06
0.206	0.119	0.379	0.085	0.049	0.008	0.101	0.201	0.019

Таблица 1.6: Модель 3. Матрица концентраций получена подстановкой параметров  $p_{min}$ , полученных методом ППРЭ.  $|W - M^{p_{min}}|$

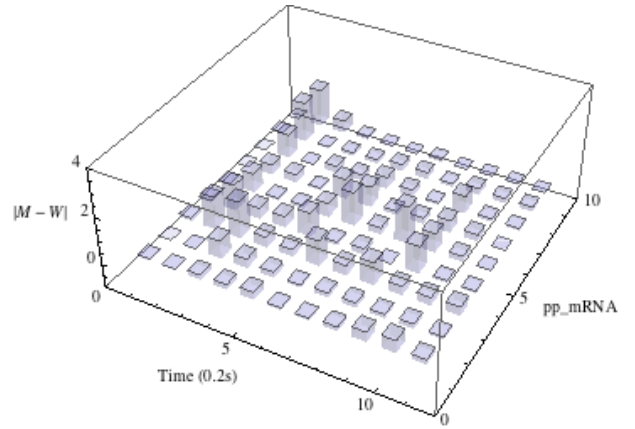
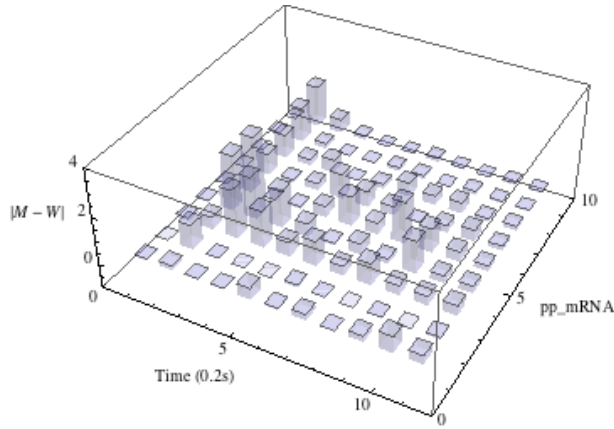


Рисунок 1.12: График таблицы 1.5 для  $p^*$     Рисунок 1.13: График таблицы 1.6 для  $p_{min}$

Параметр	$p^*$	$p_{min}$				
pro1_strength	3	2.9963	2.92585	3.05202	3.00715	2.91818
pro2_strength	8	3.55804	5.93761	3.26835	3.18463	3.70285
pro3_strength	6	6.48589	12.998	6.33644	6.81411	13.
pro4_strength	8	0.192847	0.247468	8.29558	5.69079	1.20639
pro5_strength	3	2.82969	2.76794	2.72905	3.10599	2.83064
pro6_strength	3	2.87218	3.10766	3.1161	2.97107	2.80903
rbs1_strength	3.92	6.36112	12.9358	12.9693	12.989	7.1852
rbs2_strength	5	7.30106	7.80092	7.32912	3.74508	1.4508
rbs3_strength	5	11.197	13.	12.9804	12.9947	12.8981
rbs4_strength	1	0.01	0.0191979	13.	11.6184	12.1075
rbs5_strength	5	3.58838	6.2905	12.689	5.37672	0.0312978
rbs6_strength	5	11.2102	12.9964	8.33923	12.1769	13.
v1_Kd	1	0.100209	6.82445	0.55698	12.1747	12.9971
v1_h	4	9.24381	8.37774	1.86059	6.23978	0.0283795
v2_Kd	1	1.2687	0.01	0.456528	1.46449	5.16301
v2_h	2	3.05743	0.0107374	2.52568	12.6145	7.84212
v3_Kd	0.1	0.981649	13.	0.01	0.0242598	13.
v3_h	2	10.6521	0.0100633	13.	4.07633	0.01
v4_Kd	10	11.6252	13.	12.9998	5.48176	1.31392
v4_h	4	6.5367	7.68906	7.6741	3.80357	2.16893
v5_Kd	1	10.7801	12.9836	7.33553	12.7475	8.96997
v5_h	1	1.46291	1.18206	1.54623	1.61237	2.33514
v6_Kd	0.1	0.01	0.0100392	12.8154	1.99358	0.01
v6_h	2	8.58302	12.0434	0.751893	0.274631	1.03708
v7_Kd	0.1	0.0100073	0.0100161	1.37711	0.402394	0.01
v7_h	2	1.3518	3.00998	1.6421	12.6013	13.
v8_Kd	0.2	0.66476	5.63135	3.55132	2.77787	0.01
v8_h	4	7.38939	12.8872	12.9753	8.25712	12.8461

Таблица 1.7: Параметры модели 1

Параметр	$p^*$	$p_{min}$				
pro1_strength	1.31	1.12166	1.3428	1.06452	1.34043	1.18295
pro2_strength	7.79	2.57413	2.12451	2.16855	3.33444	3.00753
pro3_strength	7.79	3.63213	1.8939	1.78388	6.27388	2.94564
pro4_strength	4	3.66947	5.23487	11.1304	12.067	3.59045
pro5_strength	5.99	8.80347	8.3395	12.7556	9.5366	4.13528
pro6_strength	5	4.67417	4.29806	4.56755	4.43596	4.61987
pro7_strength	2.93	13.	7.66922	5.91811	8.94231	1.12427
rbs1_strength	5	2.25641	10.4715	10.2654	10.2357	3.9725
rbs2_strength	1	9.78127	10.119	10.4709	0.0153719	5.27953
rbs3_strength	5	11.8777	13.	6.64263	12.2433	9.38997
rbs4_strength	5	0.152162	12.0366	7.55073	12.7734	2.48847
rbs5_strength	5	2.72958	11.4941	12.9739	0.615125	9.13042
rbs6_strength	5	7.72043	1.52188	12.079	12.9991	10.7603
rbs7_strength	5	12.9867	6.64221	11.9133	12.9977	4.19665
v1_Kd	10	4.86011	13.	8.37926	9.51029	8.98448
v1_h	1	1.28737	4.28145	0.0621061	11.8475	0.0435457
v2_Kd	2	12.998	12.9389	3.02416	6.05854	7.13107
v2_h	2	6.0014	11.8375	12.9628	12.947	5.50593
v3_Kd	10	1.45081	0.01	5.63589	0.261099	10.467
v3_h	4	1.17019	12.7148	5.37748	5.84989	0.918297
v4_Kd	0.001	10.4194	12.5022	10.3433	7.53694	5.77231
v4_h	1	11.664	0.950678	1.27285	0.248257	0.0111751
v5_Kd	4.951	2.60774	12.8827	12.96	12.9668	12.9261
v5_h	4	5.50981	7.08617	12.9986	0.166834	6.6113
v6_Kd	0.7093	0.01	0.580914	0.181896	0.651964	0.070362
v6_h	2	2.73495	1.98526	1.26195	5.05768	1.94584
v7_Kd	1	0.0100008	0.0568314	1.11111	0.0181018	11.9347
v7_h	4	12.6789	13.	2.88253	6.11441	6.48009
v8_Kd	0.01	13.	1.99672	13.	0.0108357	0.0100724
v8_h	4	0.01	0.0109494	1.1386	11.0739	0.683375
v9_Kd	10	12.9949	0.0151922	0.0185519	6.69634	12.8928
v9_h	1	8.5502	10.5261	0.668325	0.39999	0.0153574
v10_Kd	10	9.36028	13.	10.9177	12.5016	12.3133
v10_h	4	12.9581	2.40681	9.9295	11.9624	3.31859

Таблица 1.8: Параметры модели 2

Параметр	$p^*$	$p_{min}$				
pro1_strength	2	1.83655	1.75989	1.80524	1.70689	1.73567
pro2_strength	4.5077	2.73106	0.561684	0.0141565	1.79753	10.5839
pro3_strength	5	9.02625	7.7042	12.9842	3.93447	12.4497
pro4_strength	5	4.06825	3.65868	10.0973	3.63505	12.8847
pro5_strength	5	10.1031	5.37787	13.	7.94748	11.5807
pro6_strength	1.31	1.20946	1.61518	1.28631	1.29889	2.63557
pro7_strength	1.31	3.47279	1.87719	1.48501	1.40278	1.20376
pro8_strength	5	0.324542	9.94177	0.01	1.25838	6.92487
pro9_strength	5	1.11254	9.69905	0.01	0.130907	1.31434
rbs1_strength	0.3668	0.011053	0.0125254	8.38407	0.01	0.183802
rbs2_strength	1.4102	12.8511	9.12216	5.48446	0.0628125	10.4756
rbs3_strength	0.8	1.754	1.1403	5.59756	0.688286	4.00149
rbs4_strength	2.21	12.8869	6.68718	0.013985	9.76812	2.21424
rbs5_strength	0.5	0.0100057	12.3662	9.93452	12.8247	9.80061
rbs6_strength	2	12.9981	11.1952	12.6931	13.	0.0509437
rbs7_strength	5	12.6994	4.17751	0.0145791	0.0979573	11.0331
rbs8_strength	3.6377	5.18856	7.91333	6.96911	12.981	12.2507
rbs9_strength	8	12.9951	0.85603	3.52368	13.	9.90043
v1_Kd	11.147	1.27036	0.158631	1.07167	10.6224	6.26708
v1_h	1	0.605086	12.9968	0.0878918	0.452064	2.25854
v2_Kd	1	12.789	12.9998	0.0220842	7.13145	11.5336
v2_h	4	2.694	0.01	6.91033	0.71999	11.7688
v3_Kd	20	4.12552	11.1919	12.9999	9.79581	6.47455
v3_h	1	0.101528	12.9759	1.94343	12.9949	1.25699
v4_Kd	0.2	12.9996	7.10004	12.8009	10.006	11.9616
v4_h	4	0.0395719	12.3663	0.0719414	0.625118	0.412131
v5_Kd	0.2	0.01	3.75671	12.5781	12.0499	12.2468
v5_h	4	8.78179	13.	12.3549	13.	10.3321
v6_Kd	0.04	1.14168	0.010017	2.55974	0.0100053	12.3909
v6_h	4	4.27492	0.0530206	10.347	12.9636	8.99329
v7_Kd	0.02	12.9751	10.1405	0.101414	10.9749	0.690458
v7_h	4	9.31997	0.0254962	10.8923	0.541086	0.0100963
v8_Kd	0.04	0.01	0.967533	12.7922	9.73419	0.01
v8_h	4	0.204643	7.1033	13.	12.9828	4.52434
v9_Kd	0.2	13.	10.9287	12.9697	0.380651	12.9955
v9_h	4	6.80736	0.0836545	12.9986	3.16794	6.93065
v10_Kd	0.02	9.29614	2.99535	0.01	12.7499	12.9959
v10_h	4	0.01	0.0552407	12.9508	0.01	4.06318
v11_Kd	0.1	12.7237	1.12745	2.75496	12.7124	2.74753
v11_h	2	1.00985	13.	8.59118	2.01365	2.0733
v12_Kd	0.1	0.01	0.24689	5.51703	5.98262	0.0100005
v12_h	2	0.273525	9.15347	1.67924	0.0144777	0.279904
v13_Kd	0.01	12.6122	10.8823	0.0108015	4.60434	12.9962
v13_h	2	8.16779	6.52591	3.36257	0.364694	2.47203
v14_Kd	1	9.14653	3.3209	0.951031	0.0170284	3.49404
v14_h	4	11.2696	10.1934	0.156777	13.	1.30948
v15_Kd	20	13.	12.7052	13.	12.8918	12.3816
v15_h	1	13.	11.5342	0.01	13.	7.19648

Таблица 1.9: Параметры модели 3

Была проведена серия экспериментов с изменением параметра, отвечающего за способ рекомбинации при создании нового поколения (*recombination\_strategy*), и фиксированным интервалом времени для каждого запуска.

Результаты приведены на графике 1.14 и в таблице:

Способ рекомбинации	Модель 1		Модель 2		Модель 3	
	Среднее	Мин.	Среднее	Мин.	Среднее	Мин.
de_3_bin_T	46.0306	27.9233	30.4181	14.1278	54.2926	32.8514
de_3_bin	60.3923	33.9588	41.1592	34.2473	76.0278	52.8957
de_3_exp_T	81.3085	44.7055	54.6786	24.7477	81.9631	40.2255
de_3_exp	77.0761	50.8476	44.8638	30.6153	121.066	87.7188
simple	98.0068	49.9085	65.0740	36.8293	99.6021	52.9599
$\Sigma$	46.51614		20.43136		55.38312	

Исходя из этих экспериментов можно сделать вывод, что далеко не каждый способ рекомбинации подходит для рассматриваемой задачи. Например, при значении параметра *simple* или *de\_3\_exp\_T* метод ППРЭ на первых шагах уходит в один из локальных минимумов и в дальнейшем не сходится. При способе рекомбинации *de\_3\_exp* наблюдается плохая сходимость в модели №3.

Можно сказать, что ППРЭ в среднем решил задачу для 1 и 3 модели при способе рекомбинации *de\_3\_bin\_T*.

Несмотря на ограничение времени для каждого запуска вычислений, можно сделать вывод, что изменение параметра рекомбинации негативно влияет на сходимость метода. Значение *recombination\_strategy* = *de\_3\_bin\_rand\_T* является оптимальным.

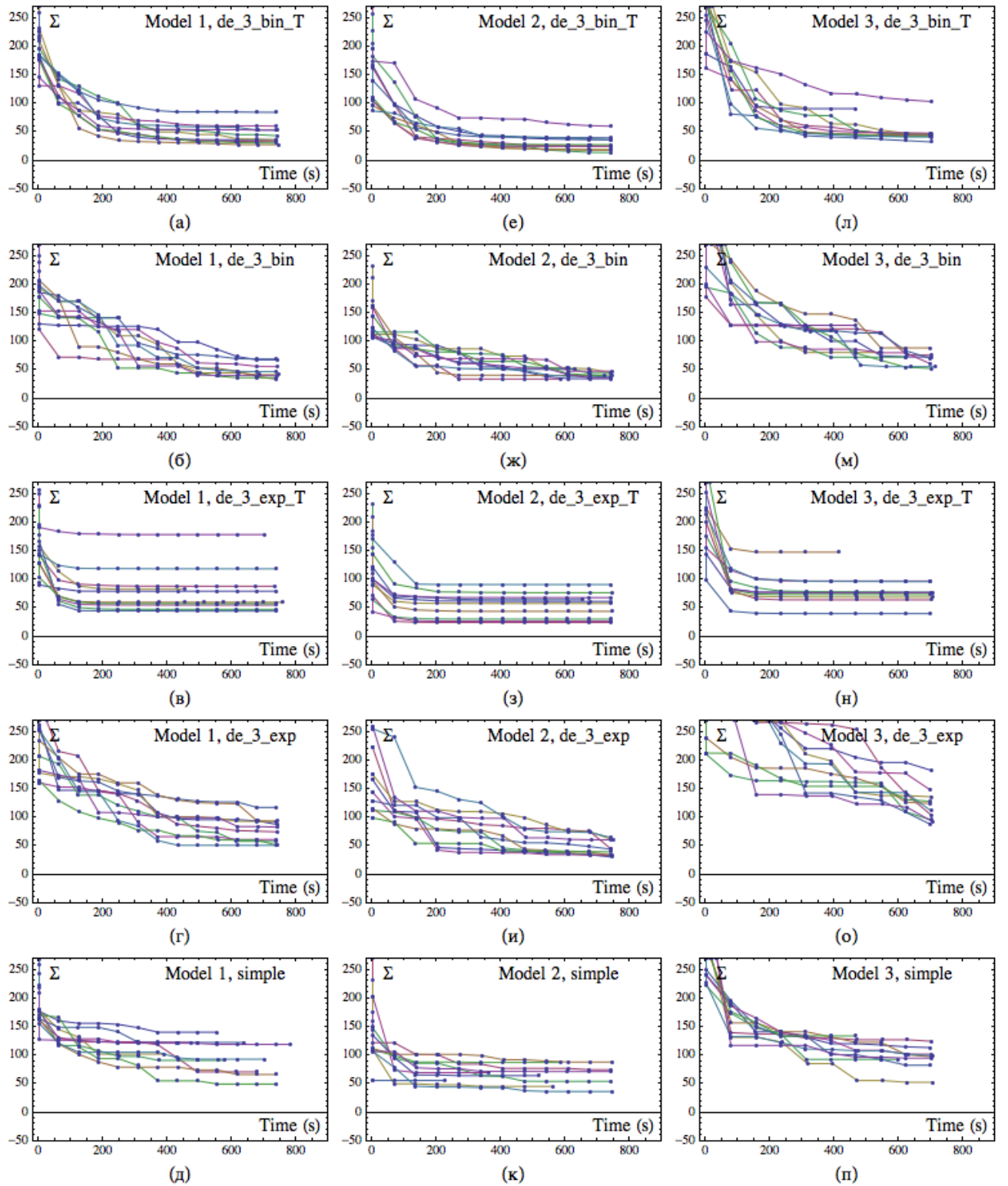


Рисунок 1.14: Графики сходимости метода при различных значениях параметра, отвечающего за способ рекомбинации



## 1.4 Выводы

Исходя из численных экспериментов можно сделать несколько выводов:

1. Алгоритм ППРЭ хорошо решает поставленную задачу поиска глобального минимума. При этом с биологической точки зрения отличий в поведении системы при использовании найденных параметров нет. В результате расчетов получено отклонение модельных концентраций от экспериментальных для первой модели: 0.207112, для второй модели: 0.131785, для третьей модели: 0.102499, где отклонение за принимается средняя нормированная (поделённая на количество элементов матрицы) сумма квадратов разностей, представленная в таблице 1.2.
2. Для данных генных сетей набор параметров обширен, по этой причине существует много вариантов их значений, в которых, возможно, достигается глобальный минимум. Соответственно, если в рамках задачи стоит вопрос поиска конкретного вектора параметров, метод ППРЭ, вероятно, не даст удовлетворительных результатов. Однако, такая задача не может встретиться на практике при решении проблемы обратной инженерии генных сетей.

## Заключение

В ходе работы были проведены численные эксперименты с применением метода ППРЭ к задаче обратной инженерии генных регуляторных сетей.

Был реализован перебор управляющих параметров ППРЭ вблизи рекомендуемых (реалистичных) значений. В частности, было проанализировано влияние на сходимость метода таких параметров как *population\_size*, *recombination\_strategy*, *es\_lambda*. В результате работы получены оптимальные для данных задач значения исследуемых параметров.

Одна из задач, поставленная в DREM6 и сформулированная в определениях метода ППРЭ была успешно решена. Было получено отклонение модельных концентраций от экспериментальных для трёх моделей (0.207112, 0.131785, 0.102499 соответственно).

В дальнейшем планируется рассмотреть вторую задачу DREM6 — предсказание динамики концентраций, поставить её в терминах ППРЭ и решить.

# Литература

1. Storn R, Price K (1995) Differential evolution a simple and efficient heuristic for global optimization over continuous spaces. Technical Report TR-95-012, ICSI
2. Konstantin Kozlov and Samsonov, A. (2011). DEEP – Differential Evolution Entirely Parallel Method for Gene Regulatory Networks. *Journal of Supercomputing*, 57:172–178
3. Н.В. Иванисенко, Е.Л. Мищенко, И.Р. Акбердин, П.С. Деменков, В.А. Лихошвай, К.Н. Козлов, Д.И. Тодоров, М.Г. Самсонова, А.М. Самсонов, Н.А. Колчанов, В.А. Иванисенко (2013) Репликация субгеномного репликона вируса гепатита С в присутствии ингибиторов NS3-протеазы: стохастическая модель, *биофизика*, том 58, вып. 5, с. 758-774
4. Konstantin Kozlov, Surkova, S., Myasnikova, E., Reinitz, J., and Samsonova, M. (2012). Modeling of gap gene expression in drosophila Kruppel mutants. *PLoS Computational Biology*. doi:10.1371/journal.pcbi.1002635
5. Konstantin Kozlov, Nikita Ivanisenko, Vladimir Ivanisenko, Nikolay Kolchanov, Maria Samsonova, and Alexander M.Samsonov (2013), «Enhanced Differential Evolution Entirely Parallel Method for Biomedical Applications», LNCS 7979, *Parallel Computing Technologies*, Victor Malyskin (Ed.), Springer 2013, ISSN 0302-9743, Proceedings of the 12th International Conference, PaCT 2013, St.Petersburg, Russia, September/October 2013, pp. 409-416.
6. Steiert B, Raue A, Timmer J, Kreutz C (2012) Experimental Design for Parameter Estimation of Gene Regulatory Networks. *PLoS ONE* 7(7): e40052. doi:10.1371/journal.pone.0040052
7. D. Zaharie Control of Population Diversity and Adaptation in Differential Evolution Algorithms, in R. Matousek, P. Osmera (eds.), *Proc. of Mendel 2003*, 9th International Conference on Soft Computing, Brno, pp. 41-46, 2003.
8. Fan H-Y, Lampinen J (2003) A trigonometric mutation operation to differential evolution. *J Glob Optim* 27:105–129
9. Описание DREAM6 [the-dream-project.org](http://the-dream-project.org)
10. Модели генных сетей (язык R) [github.com/latur/DREAM6](https://github.com/latur/DREAM6)