

УДК: 517.958:57

Репрессиллятор с запаздывающей экспрессией генов. Часть II. Стохастическое описание

Д. А. Брацун^а, М. Д. Бузмаков

Пермский национальный исследовательский политехнический университет,
Россия, 614990, г. Пермь, ул. Профессора Поздеева, д. 11, к. В

E-mail: ^а DABracun@pstu.ru

Получено 20.03.2021, после доработки —13.04.2021.

Принято к публикации 27.04.2021.

Репрессилатором называют первую в синтетической биологии генную регуляторную сеть, искусственно сконструированную в 2000 году. Он представляет собой замкнутую малоразмерную цепь из трех генов: *lacI*, *λcl* и *tetR*, которые в такой комбинации в природе не встречаются. Промотор каждого гена контролирует следующий за ним цистрон по принципу отрицательной обратной связи, подавляя экспрессию соседнего гена. Ранее в работе [Брацун и др., 2018] была предложена математическая модель запаздывающего репрессилатора и изучены ее свойства в рамках детерминистского описания. Предполагается, что запаздывание может быть как естественным, т. е. возникать во время процессов транскрипции/трансляции в силу многоступенчатого характера этих процессов, так и искусственным, т. е. специально вноситься в работу регуляторной сети с помощью методов геной инженерии. Данная работа посвящена стохастическому описанию динамических процессов в запаздывающем репрессилаторе, которое является важным дополнением детерминистского анализа из-за сильных флуктуаций и небольшого числа молекул, принимающих обычно участие в геной регуляции. Стохастическое исследование было проведено численно с помощью алгоритма Гиллеспи, модифицированного для систем с запаздыванием. Приводятся описание алгоритма, его программная реализация и результаты тестовых расчетов для одногенного авто-репрессора с запаздыванием. При исследовании репрессилатора обнаружено, что стохастическое описание в ряде случаев дает новую информацию о поведении системы, которая не сводится к детерминистской динамике даже при усреднении по большому числу реализаций. В подкритической области, где детерминистский анализ предсказывает абсолютную устойчивость системы, было обнаружено возбуждение квазирегулярных колебаний, вызываемых нелинейным взаимодействием шума и запаздывания. Выше порога возникновения неустойчивости обнаружено спонтанное изменение фазы колебаний из-за внезапной временной деградации этих колебаний. Ранее в детерминистском анализе был обнаружен долгоживущий переходный режим, который отвечает движению фазовой траектории по медленному многообразию и отражает процесс длительной синхронизации пульсаций в работе отдельных генов. Показано, что в стохастическом случае переход к кооперативному режиму работы генов репрессилатора происходит в среднем на два порядка быстрее. Построено распределение вероятности соскока фазовой траектории с медленного многообразия и определено наиболее вероятное время такого перехода. Обсуждается влияние внутреннего шума химических реакций на динамические свойства репрессилатора.

Ключевые слова: репрессилатор, запаздывание, геной регуляция, стохастическое описание

Постановка задачи и модификация алгоритма стохастического расчета выполнены при финансовой поддержке Пермского НОЦ «Рациональное недропользование». Написание авторской программы и проведение численных расчетов выполнены при поддержке Министерства науки и высшего образования РФ (грант № FSNM-2020-0026).

UDC: 517.958:57

Repressilator with time-delayed gene expression. Part II. Stochastic description

D. A. Bratsun^a, M. D. Buzmakov

Perm National Research Polytechnic University,
11B Professor Pozdeev st., Perm, 614990, Russia

E-mail: ^a DABracun@pstu.ru

*Received 20.03.2021, after completion — 13.04.2021.
Accepted for publication 27.04.2021.*

The repressilator is the first genetic regulatory network in synthetic biology, which was artificially constructed in 2000. It is a closed network of three genetic elements *lacI*, λcI and *tetR*, which have a natural origin, but are not found in nature in such a combination. The promoter of each of the three genes controls the next cistron via the negative feedback, suppressing the expression of the neighboring gene. In our previous paper [Bratsun et al., 2018], we proposed a mathematical model of a delayed repressilator and studied its properties within the framework of a deterministic description. We assume that delay can be both natural, i.e. arises during the transcription / translation of genes due to the multistage nature of these processes, and artificial, i.e. specially to be introduced into the work of the regulatory network using gene engineering technologies. In this work, we apply the stochastic description of dynamic processes in a delayed repressilator, which is an important addition to deterministic analysis due to the small number of molecules involved in gene regulation. The stochastic study is carried out numerically using the Gillespie algorithm, which is modified for time delay systems. We present the description of the algorithm, its software implementation, and the results of benchmark simulations for a one-gene delayed autorepressor. When studying the behavior of a repressilator, we show that a stochastic description in a number of cases gives new information about the behavior of a system, which does not reduce to deterministic dynamics even when averaged over a large number of realizations. We show that in the subcritical range of parameters, where deterministic analysis predicts the absolute stability of the system, quasi-regular oscillations may be excited due to the nonlinear interaction of noise and delay. Earlier, we have discovered within the framework of the deterministic description, that there exists a long-lived transient regime, which is represented in the phase space by a slow manifold. This mode reflects the process of long-term synchronization of protein pulsations in the work of the repressilator genes. In this work, we show that the transition to the cooperative mode of gene operation occurs a two order of magnitude faster, when the effect of the intrinsic noise is taken into account. We have obtained the probability distribution of moment when the phase trajectory leaves the slow manifold and have determined the most probable time for such a transition. The influence of the intrinsic noise of chemical reactions on the dynamic properties of the repressilator is discussed.

Keywords: repressilator, time delay, gene regulation, stochastic description

The problem formulation and the modification of the stochastic algorithm were carried out with the financial support of the Perm Science Center “Rational Subsoil Use”. The program code was written and numerical simulations have been supported by the Ministry of Science and Higher Education of the Russian Federation (Grant No. FSNM-2020-0026).

Citation: *Computer Research and Modeling*, 2021, vol. 13, no. 3, pp. 587–609 (Russian).

1. Введение

Во второй половине XX века в молекулярной биологии начался процесс математизации — превращения ее из описательной науки в точную. Основная причина, почему этот процесс так затянулся, заключается в том, что долгое время экспериментальные технологии, применяемые в молекулярно-генетическом анализе, были недостаточными для того, чтобы производить точные измерения белковых полей в пространстве и времени. Важным этапом было выделение в 90-е годы прошлого века зеленого флуоресцентного белка GFP, ген которого в настоящее время широко используется в качестве оптической метки в клеточной и молекулярной биологии для изучения экспрессии клеточных белков [Prasher et al., 1992]. Возможность регистрировать, измерять, а следовательно, и изменять белковые поля дала начало новому научному направлению, которое получило название «синтетическая биология». Считается, что развитие этой науки в перспективе должно привести к конструированию новых организмов с заранее заданными свойствами. Первый шаг синтетической биологии был сделан в работе [Elowitz, Leibler, 2000], в которой было теоретически смоделирована, а затем и экспериментально сконструирована плазида из трех разных генов, которые в естественных условиях в такой комбинации в одном геноме не встречаются. Цепь была сконфигурирована таким образом, что белок каждого гена ингибировал экспрессию следующего за ним соседа. Придуманная схема генерировала колебательную экспрессию генов, которая была измерена экспериментально с помощью GFP и описана в рамках простой шестимодовой динамической модели. Успех работы [Elowitz, Leibler, 2000] показал, что таким путем действительно возможно конструировать генные цепи с предписанной динамической регуляцией, что и стало предметом синтетической биологии [Hasty et al., 2001]. В экспериментальной работе [Potvin-Trottier et al., 2016] было показано, что точность молекулярных часов репрессиллятора может быть улучшена за счет размещения всех генов на одном векторе и контроля за концентрацией белка с помощью специального накопителя.

В ряде работ [Muller et al., 2006; Bennett et al., 2007; Buse et al., 2010; Лихошвай и др., 2010; Strelkowa, Varahona, 2011; Иванченко и др., 2013; Брацун и др., 2018] рассматривались различные теоретические аспекты проблемы. Более сложная модель репрессиллятора с дополнительными связями в контуре была изучена в [Волков, Журов, 2013], ансамбли взаимодействующих репрессилляторов изучены в [Ullner et al., 2007; Потапов, Волков, 2010], обнаружены пространственные волны активности в клеточных ансамблях [Danino et al., 2010]. Последняя работа интересна также тем, что колебания были заданы путем искусственного внесения запаздывания в синтетическую цепь, разработанную авторами. Запаздывающие осцилляторы имитировали в каждой бактерии работу ее биологических часов, а транспортный белок осуществлял глобальную координацию поведения всей колонии. В результате было обнаружено спонтанное возникновение бегущих концентрических волн свечения GFP в колонии бактерий.

Вопросы запаздывания в процессах генной регуляции вызывают все больший интерес начиная примерно с 2000 года. До этого запаздывание в биохимических реакциях генной регуляции рассматривалось как малоинтересная экзотика. Однако экспериментальное подтверждение того, что механизм запаздывания может играть принципиальную роль в генных цепях, впервые полученное в работе [Lee et al., 2000], изменило ситуацию. Вообще говоря, процессы экспрессии генов состоят из многоэтапных реакций, в ходе которых последовательно образуются ансамбли сложных органических соединений. Следовательно, эти процессы распределены по пространству, растянуты по времени и идут с неким характерным временем запаздывания. Строго говоря, эти процессы уже не могут считаться марковскими [Лихошвай и др., 2003; Bratsun et al., 2005a, 2005b; Брацун, 2011]. Если время запаздывания мало по сравнению со всеми остальными характерными временами в системе, этими эффектами можно пренебречь. Однако если величина времени запаздывания одного порядка с другими процессами или обратная связь сильная, то эффекты запаздывания не могут более игнорироваться [Lee et al., 2000; Denault et al., 2001; Лихошвай и др., 2003; Sriram, Gopinathan, 2004; Демиденко и др., 2004; Брацун, Захаров, 2011; Bratsun et al., 2016].

В нашей предыдущей работе [Брацун и др., 2018] мы предложили модифицированную модель репрессилатора с учетом запаздывания в его контурах отрицательной обратной связи. В работе было представлено детерминистское описание процессов генной регуляции в запаздывающем репрессилаторе, что включало в себя как аналитическое изучение устойчивости состояния равновесия, так и численное исследование нелинейной динамики. Были рассмотрены случаи симметричного (все гены одинаковые) и несимметричного контура. Обнаружено существование долгоживущего переходного колебательного режима на медленном многообразии. Как известно, для изучения свойств динамической системы можно использовать как детерминистское, так и стохастическое описание. Если система обладает огромным числом степеней свободы, то стохастическое описание, которое включает влияние флуктуаций, является избыточным. Системы генной регуляции выделяются тем, что число молекул, участвующих в процессе, невелико, а возможные флуктуации могут приводить к качественному изменению динамики в рамках стохастического описания [Thattai, van Oudenaarden, 2001; Kepler, Elston, 2001]. В работах одного из авторов [Bratsun et al., 2005a, 2005b] вопросы взаимодействия запаздывания и стохастических флуктуаций впервые были рассмотрены для процессов транскрипции/трансляции генов. Обнаружен эффект возбуждения в подкритической области колебаний за счет взаимодействия шума и запаздывания обратной связи. Кроме того, в работе [Bratsun et al., 2005b] впервые была предложена модификация классического алгоритма Гиллеспи [Gillespie, 1977], используемого для прямого численного моделирования стохастических дискретных процессов, на случай систем с запаздыванием.

Различают два вида шума: внутренний шум химических реакций генной регуляции (intrinsic noise) и внешний шум клеточного ансамбля, члены которого слегка отличаются друг от друга (extrinsic noise). Первый тип шума, как правило, исследуется с помощью алгоритма Гиллеспи, который иногда в ранних работах называют методом Монте-Карло [Elowitz, Leibler, 2000; Thattai, Oudenaarden, 2001; Bratsun et al., 2005a; Yoda et al., 2007; Loinger, Biham, 2007; Potapov et al., 2015]. Модификации алгоритма постоянно предлагаются в литературе. Мотивацией может служить либо физика процесса регуляции [Bratsun et al., 2005b], либо желание упростить вычислительную процедуру [Tuttle et al., 2005], либо создание удобного инструмента для теоретического рассмотрения [Gillespie, 2000]. Внешний шум ансамбля может задаваться с помощью гораздо более разнообразных методов, которые включают уравнение Ланжевена, использование распределение Пуассона при делении клеток, рассмотрение сетей из взаимодействующих генных осцилляторов.

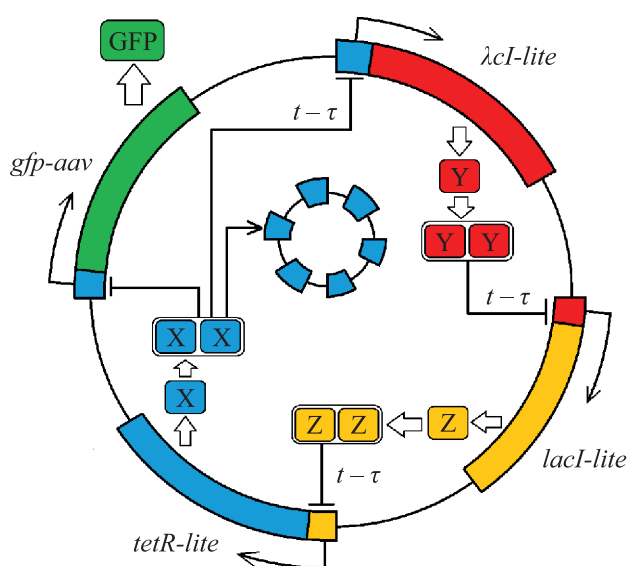


Рис. 1. Схематическое изображение генной сети репрессилатора с запаздыванием

Данная работа посвящена стохастическому анализу предложенной ранее модели запаздывающего репрессиллятора. Описывается алгоритм численного решения задачи, который тестируется на примере модели одногенного авторепрессора с запаздывающей обратной связью. Дано описание программного комплекса, который был использован для численного анализа, и представлены результаты численного моделирования. Внимание концентрируется на тех результатах, где обнаружено качественное различие поведения репрессиллятора в рамках детерминистского и стохастического анализа.

2. Описание регуляторной цепи репрессиллятора с запаздыванием

Схема цепи репрессиллятора [Elowitz, Leibler, 2000] с учетом важных модификаций, предложенных в работе [Potvin-Trottier et al., 2016], представлена на рис. 1. Как следует из рисунка, плазида репрессиллятора включает в себя четыре цистрона. Первый цистрон содержит ген *tetR-lite* (ген и его промоторы отмечены на схеме синим цветом), который находится под управлением промотора *lac* гена *lacI-lite* (оба отмечены на схеме желтым цветом), выступающего в качестве репрессора. При экспрессии *tetR-lite* синтезируется соответствующий белок, который для краткости будем обозначать X. Белок X димеризуется и в этом виде может выполнять три функции. Во-первых, он через промотор *tetR* может служить репрессором для λ фага (соответствующий ген *λcl-lite* отмечен на схеме красным цветом). Во-вторых, он также репрессирует экспрессию флуоресцентного белка GFP (выделен зеленым цветом). В исходной версии репрессиллятора 2000 года ген *gfp-aav*, кодирующий GFP, был размещен на отдельной плазмиде. Как было продемонстрировано в [Potvin-Trottier et al., 2016], это приводит к возникновению сильных флуктуаций в работе репрессиллятора из-за удаленной регуляции. Поэтому в новой версии ген, кодирующий сигнальный белок, был перенесен на ту же самую плазмиду, где располагаются основные игроки репрессиллятора (второй цистрон). В-третьих, димеры белка X могут связываться с оператор-сайтом специальной ДНК плазмиды (синее кольцо в центре схемы), в которой *m* одинаковых позиций для связывания. Это второе важное усовершенствование репрессиллятора, которое приводит к тому, что белок X скорее накапливается в системе, чем деградирует, а цепь с оператор-сайтами выступает как хранилище для белка, кодируемого геном *tetR-lite*. Это было сделано для того, чтобы поддержать концентрацию белка X, который в обычных условиях сильно деградирует при регуляции. Ген *λcl-lite*, находящийся под контролем промотора *tetR*, входит в состав третьего цистрона и кодирует белок, концентрацию которого будем обозначать Z. Как видно на схеме, промотор λ фага выступает как репрессор экспрессии гена *lacI-lite* (четвертый цистрон). Белок, кодируемый последним геном, будем обозначать Y. Таким образом, основное ядро репрессиллятора, определяющее его временную динамику, составлено тремя генами, которые связаны между собой через промоторы отрицательной обратной связью.

Перечислим основные допущения в данной работе:

- регуляция репрессиллятора осуществляется только димерной формой белка;
- образование промежуточной формы мРНК в модели игнорируется, так как считается, что эта стадия может быть учтена введением запаздывания;
- промотор каждого гена может регулировать экспрессию одного другого гена;
- рассматривается случай симметричного репрессиллятора, т. е. все характеристики генов считаются одинаковыми;
- отрицательная обратная связь является запаздывающей для любой пары генов, причем время запаздывания будем полагать одинаковым;
- работа накопителя белка X игнорируется;
- работа гена *gfp-aav* игнорируется;
- любое время запаздывания много больше остальных характерных времен, определяющих нелинейную динамику репрессиллятора.

Таблица 1. Список реакций при генной регуляции репрессилатора с запаздыванием (см. рис. 1)

1	Димеризация белков X, Y, Z
	$X + X \xrightarrow{k_1} X_2, Y + Y \xrightarrow{k_1} Y_2, Z + Z \xrightarrow{k_1} Z_2$
2	Дедимеризация белков X, Y, Z
	$X_2 \xrightarrow{k_{-1}} X + X, Y_2 \xrightarrow{k_{-1}} Y + Y, Z_2 \xrightarrow{k_{-1}} Z + Z$
3	Связывание димера репрессора с оператор-сайтом
	$D_0^X + Z_2 \xrightarrow{k_2} D_1^X, D_0^Y + X_2 \xrightarrow{k_2} D_1^Y, D_0^Z + Y_2 \xrightarrow{k_2} D_1^Z$
4	Освобождение димера репрессора оператор-сайтом
	$D_1^X \xrightarrow{k_{-2}} D_0^X + Z_2, D_1^Y \xrightarrow{k_{-2}} D_0^Y + X_2, D_1^Z \xrightarrow{k_{-2}} D_0^Z + Y_2$
5	Синтез белков X, Y, Z при открытом оператор-сайте, протекающий с запаздыванием
	$D_0^X(t) \xrightarrow{A} D_0^X + X(t + \tau), D_0^Y(t) \xrightarrow{A} D_0^Y + Y(t + \tau), D_0^Z(t) \xrightarrow{A} D_0^Z + Z(t + \tau)$
6	Дегградация белков X, Y, Z
	$X \xrightarrow{B} \emptyset, Y \xrightarrow{B} \emptyset, Z \xrightarrow{B} \emptyset$

Учитывая сделанные выше замечания, полный список биохимических реакций, протекающих при работе репрессилатора, приведен в таблице 1. Соответствующие скорости реакций указаны над стрелками. Наличие оператор-сайтов в таблице учтено с помощью введения специальной дискретной функции $D \in \{D_0, D_1\}$, которая может принимать значение D_0 в случае открытия оператор-сайта и D_1 в случае его закрытия. С формальной точки зрения состояние оператор-сайта можно рассматривать как дополнительный реагент в системе. В случае связывания димера репрессора оператор-сайт закрывается и транскрипция гена не происходит. Таким образом, в системе реализуется отрицательная обратная связь.

В первой части работы [Брацун и др., 2018] была получена следующая система дифференциальных уравнений с запаздыванием, моделирующая работу симметричного репрессилатора в рамках детерминистского описания:

$$\frac{dx}{dt} = \frac{1}{1 + 4\epsilon x} \left(\frac{A}{1 + \epsilon \delta z^2(t - \tau)} - Bx \right), \quad (1)$$

$$\frac{dy}{dt} = \frac{1}{1 + 4\epsilon y} \left(\frac{A}{1 + \epsilon \delta x^2(t - \tau)} - By \right), \quad (2)$$

$$\frac{dz}{dt} = \frac{1}{1 + 4\epsilon z} \left(\frac{A}{1 + \epsilon \delta y^2(t - \tau)} - Bz \right), \quad (3)$$

где $x(t)$, $y(t)$, $z(t)$ — непрерывные функции времени, обозначающие соответственно концентрации белков X, Y, Z в мономерной форме. В уравнениях (1)–(3) введены следующие обозначения:

$$\epsilon = \frac{k_1}{k_{-1}}, \quad \delta = \frac{k_2}{k_{-2}}. \quad (4)$$

Мотивация рассмотрения репрессилатора с запаздывания следующая. Мы до сих довольно мало знаем о реальных молекулярных механизмах генной регуляции. Появление малоразмерных синтетических генных цепочек как раз и является попыткой максимально упростить систему для изучения базовых механизмов регуляции, действующих в них. Даже после появления

технологии GFP эта задача остается нетривиальной. Поэтому любые математические модели генной регуляции реальных систем до сих пор являются приблизительным отражением протекающих там процессов.

В случае генной регуляции циркадианного ритма нейроспоры и дрозофилы, как было показано в [Lee et al., 2000], запаздывание составляет порядка 6 часов, что позволяет данным организмам формировать циркадианный ритм, близкий к 24 часам (учетверенное время запаздывания). Период колебаний репрессиллятора составляет около 5 часов [Elowitz, Leibler, 2000]. Если колебания в репрессилляторе обусловлены механизмом временной задержки, то время запаздывания должно быть порядка 1–1.5 часа. Это вполне реалистичный сценарий, причем модель (1)–(3) в этом случае имеет хорошее обоснование, так как абсолютное большинство реакций генной регуляции (димеризация, динамика оператор-сайта) протекает гораздо быстрее.

Еще одно важное обстоятельство. Если в уравнениях (1)–(3) запаздывание положить равным нулю, т. е. предположить мгновенное срабатывание отрицательных обратных связей, то колебаний в системе нет [Buse et al., 2009]. Есть два пути для получения колебаний в модели: либо повышение олигомерности белков до тетрамеров и выше (что несколько экзотично), либо включение в цепь обратной связи еще одного посредника в виде мРНК. Вторым путем пошли авторы пионерской работы [Elowitz, Leibler, 2000], пытаясь объяснить экспериментально обнаруженные колебания. На самом деле оба пути означают неявное введение запаздывания в цепи обратных связей: в первом случае временная задержка срабатывания обратной связи происходит благодаря ожиданию формирования белковой молекулы большей олигомерности, во втором случае система ожидает окончания процесса трансляции. Мы же в своей модели предлагаем ввести запаздывание явным образом. Введение запаздывания, которое подразумевает протекание многоэтапных реакций, в ходе которых последовательно образуются ансамбли сложных соединений, позволяет понизить порядок системы вдвое по сравнению с моделью [Elowitz, Leibler, 2000].

3. Алгоритм Гиллеспи

3.1. Алгоритм Гиллеспи для марковских процессов

Одним из самых известных методов численного исследования малоразмерных стохастических реакций является метод Гиллеспи [Gillespie, 1977], который за время своего появления стал поистине классическим из-за своей простоты и надежности. Математически метод Гиллеспи является разновидностью алгоритмических методов Монте-Карло для численного стохастического анализа и воспроизводит решение для распределения вероятности, задаваемого основным кинетическим уравнением (или *мастер-уравнением* в англоязычной литературе). Вообще говоря, усреднение по реализациям, полученным с помощью алгоритма Гиллеспи, должно сходиться к решению, полученному в рамках детерминистского описания. Любопытно, что статья Гиллеспи, вышедшая в 1977 году, была не замечена большинством исследователей, так как в то время считалось, что малоразмерные химические реакции являются экзотикой и реальных приложений у алгоритма нет. Кардинальное изменение произошло после 2000 года, когда наблюдался взрывной рост цитируемости работы 1977 года, который обычно связывают с появлением направления синтетической биологии. Объектом исследований в новом направлении являются малоразмерные искусственные цепи генной регуляции, состоящие из небольшого числа генов. Для таких систем алгоритм Гиллеспи является идеальным инструментом исследования.

Опишем кратко главную идею метода Гиллеспи, которая была разработана для марковских процессов. Предположим, имеется K химических соединений X_v , вступающих в многоэтапную реакцию между собой. Каждый тип реактивного контакта, случающийся между молекулами, будем называть вслед за Гиллеспи *химическим каналом* R_μ . Пусть таких каналов имеет-

ся M штук. Согласно алгоритму многосоставная реакция распадается на элементарные акты. При совершении каждого акта нужно вычислить время, через которое произойдет следующая элементарная реакция, а также определить канал, по которому соединения вступают в контакт. Для марковских процессов распределение промежутков времени Δt , проходящих между реакциями, подчиняется распределению Пуассона:

$$P(\Delta t) = \sum_{\mu=1}^M a_{\mu} \exp\left(-\Delta t \sum_{\mu=1}^M a_{\mu}\right). \quad (5)$$

Здесь

$$a_{\mu} = c_{\mu} h_{\mu} \quad (6)$$

является вероятностью того, что в данный момент реализуется μ -й химический канал. В англоязычной литературе эта величина называется *propensity*, т. е. «склонность» или «тенденция», однако эта величина имеет смысл ненормированной вероятности наступления реакции. Как видно из определения (6), значение a_{μ} формируется благодаря двум факторам. С одной стороны, вероятность реализации реакционного канала тем выше, чем больше скорость реакции c_{μ} . С другой стороны, вероятность зависит от кинетики элементарной реакции, которая определяет количество различных молекулярных соединений h_{μ} из набора X_v . Как правило, величина h_{μ} может быть легко вычислена (см. примеры, разобранные в работе [Gillespie, 1977]). Распределение вероятностей для химических каналов является дискретным и нормированным на величину достоверного события, которая определяется суммой вероятностей всех каналов (какая-то реакция из набора все равно неизбежно наступит):

$$P(\mu = \mu^*) = \frac{a_{\mu^*}}{\sum_{\mu=1}^M a_{\mu}}. \quad (7)$$

Таким образом, алгоритм включает в себя предварительное вычисление по заранее известной кинетике химических реакций (каналов) значений a_{μ} . После этого на каждом шаге по времени генерируются два случайных числа из отрезка $[0, 1]$. Первое число, согласно распределению (5), определяет время наступления ближайшего реакционного акта Δt , а второе число, согласно распределению (7), определяет конкретный канал μ^* , через который система реализует свою динамику. После выполнения шага Δt вектор состояния системы X_v обновляется, и вся процедура повторяется.

3.2. Алгоритм Гиллеспи для немарковских процессов

Модификация алгоритма на случай немарковских процессов была предложена одним из авторов в работах [Bratsun et al., 2005a, 2005b]. Главная идея обновления алгоритма заключается в следующем (рис. 2). Предположим, один из каналов R_{μ} является запаздывающим. Если выпадает очередь реакции с запаздыванием, то она не выполняется, а запись о ней помещается в стек памяти, с тем чтобы быть выполненной в момент времени $\Delta t + \tau$, где τ — время запаздывания. Этот шаг выглядит простым и логичным. Сложнее определить, что делать, если следующей должна выполняться обычная реакция, но в отрезке времени $[t, t + \Delta t]$ запланировано выполнение отложенной реакции. В работе [Bratsun et al., 2005a] была предложена простая модельная система, для которой основное кинетическое уравнение может быть решено точно. На основании этого было показано, что правильно поступить следующим образом: последняя процедура отбора игнорируется, а система совершает прыжок к моменту времени, которое ранее было уже запланировано, но отложено из-за запаздывания.

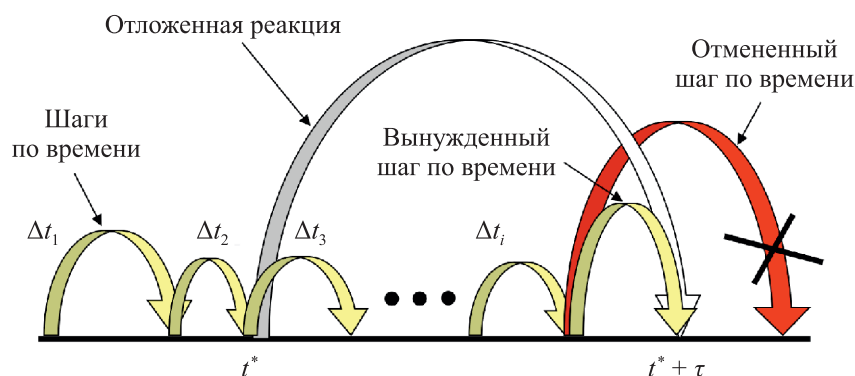


Рис. 2. Схема, поясняющая модифицированный алгоритм Гиллеспи [Bratsun et al., 2005a]

Приведем общую схему работы модифицированного алгоритма Гиллеспи.

1. Задание начальных значений для вектора состояния системы X_v .
Количество химических соединений равно K .
Количество химических каналов равно M .
Установка начального времени: $t = 0$.
Установка начального значения счетчика шагов по времени: $J = 1$.
2. Вычисление вероятности реализации для каждого канала a_μ , $\mu = 1, \dots, M$.
Вычисление значения полной вероятности для каналов: $a_0 = \sum_{\mu=1}^M a_\mu$.
3. Генерирование двух случайных чисел: $r_1, r_2 \in [0, 1]$.
4. Вычисление промежутка времени до следующей реакции: $\Delta t_J = -\ln \frac{r_1}{a_0}$.
5. Проверка на наличие отложенной (запаздывающей) реакции, которая должна произойти внутри промежутка времени: $t^{\text{delay}} \in [t, t + \Delta t_J]$:
 - А. Если в указанном промежутке $[t, t + \Delta t_J]$ в момент t^{delay} должна произойти отложенная реакция, то
 - а) пункты алгоритма 2–4 игнорируются;
 - б) система совершает шаг по времени до момента t^{delay} ;
 - в) запаздывающая реакция совершается;
 - г) счетчик обновляется: $J = J + 1$;
 - д) переход к пункту 2.
 - Б. Если в указанном диапазоне времени событие A не происходит, то переход к пункту 6.
6. Нахождение канала для следующей реакции μ^* согласно $\sum_{\mu=1}^{\mu^*-1} a_\mu < r_2 a_0 \leq \sum_{\mu=1}^{\mu^*} a_\mu$.
 - А. Если выпавший канал μ^* соответствует реакции без запаздывания, то
 - а) система делает шаг по времени до момента $t + \Delta t_J$;
 - б) вектор решения обновляется в соответствии с выпавшим каналом;
 - г) счетчик реакций обновляется: $J = J + 1$;
 - д) переход к пункту 2.
 - Б. Если выпавший канал μ^* отвечает реакции с запаздыванием, то
 - а) выполнение реакции откладывается до момента времени $t + \tau$;
 - б) переход к пункту 2.

Стоит отметить, что любой вид немарковского поведения серьезно усложняет процедуру вычислений, так как в ходе расчета необходимо держать в памяти огромный массив данных. Даже для простейшего типа немарковости, динамики с фиксированным временем запаздывания, в стеке памяти быстро накапливаются сотни тысяч отложенных реакций. Дело в том, что в генной регуляции интересны реакции с большими временами запаздывания, которые приводят к нетривиальным эффектам.

Отметим, что алгоритм стохастического расчета, предложенный в [Bratsun et al., 2005b], хорошо работает, если время запаздывания существенно превышает другие характеристические времена в системе, так как это позволяет разорвать обусловленные вероятности в мастер-уравнении. Более подробно об этом можно прочитать в [Bratsun et al., 2005b]. Как отмечено выше, период колебаний репрессилатора составляет порядка 5 часов, что сигнализирует о вероятной задержке в срабатывании обратной связи в 1–1.5 часа. Это гораздо больше, чем типичное время быстрых реакций димеризации и динамики оператор-сайта. Поэтому предложенная методика стохастического анализа хорошо подходит к модели репрессилатора.

3.3. Программная реализация алгоритма

Для численного исследования запаздывающего репрессилатора была разработана программа на языке программирования C++ в среде *Visual Studio*. Она представляет собой консольное приложение, в котором выводятся результаты расчетов через каждый предварительно указанный шаг по времени. Параллельно с выводом на экран происходит запись значений в текстовый файл. Непосредственно код программы не использует сторонние библиотеки и направлен на простоту исполнения в первую очередь из-за необходимости совершения большого количества простых последовательных вычислений. Их общее количество может достигать нескольких миллионов. Например, один расчет для симметричного запаздывающего репрессилатора на 30000 единиц времени может занимать два часа (AMD Ryzen 5 4600H).

В коде программы каждая реакция описывается классом, который наследует свою основу от родительского класса *Generic_reaction*. Он содержит в себе переменную *speed*, характерную для любой реакции, а также набор методов, позволяющих получить текущую вероятность (*GetProb*), время запаздывания (*GetDelay*) и изменить значения концентраций в системе, то есть непосредственно совершить реакцию (*React*). Наполнение перечисленных методов и добавление дополнительных переменных происходят при описании каждой конкретной реакции. Например, помимо наследованной переменной *speed*, класс *Gene_synthesis* содержит ссылки на переменные, отвечающие за концентрацию мономеров белка, статус оператор-сайта и время запаздывания (рис. 3). Подобная структура позволяет использовать код программы в качестве основы для других подобных систем. Для этого необходимо лишь создать новые классы на основе биохимических реакций или использовать существующие шаблоны. Затем инициализировать соответствующие экземпляры с параметрами, которые будут помещены в вектор *reactions*. В свою очередь, основная часть алгоритма будет работать на основе информации, заложенной в него.

При запуске программы происходит считывание параметров из файла *settings.txt*, который должен находиться в одном каталоге с исполняемым файлом. Функция *ReadSettings* наполняет массив *set* всеми необходимыми параметрами: скоростями реакций, начальными значениями концентраций белка, временами запаздывания и временем окончания расчета. На основе полученных данных заполняются соответствующие переменные. Затем происходит заполнение вектора с реакциями для каждого гена в цепи. Сами по себе классы являются универсальными, поэтому при инициализации экземпляров используются необходимые переменные в соответствии с приведенным списком реакций. В качестве генератора псевдослучайных чисел используется алгоритм *mt19937* из стандартного набора C++. Концентрации мономеров и димеров белка каждого гена в цепи представлены в виде общего массива. При совершении реакций происходит обращение к нему с целью изменения соответствующих значений.

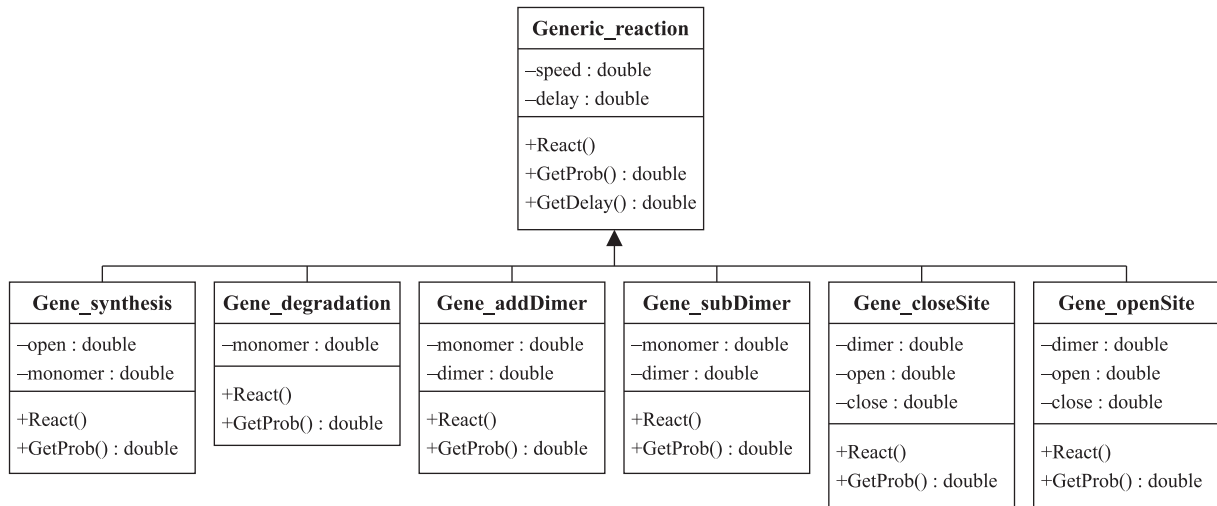


Рис. 3. Диаграмма классов для описания реакций. В верхней секции каждого прямоугольника находится название класса, по центру располагаются названия соответствующих им переменных, в нижней части — названия методов. Линия с пустой стрелкой обозначает наследование переменных и методов от родительского класса

Основной цикл работает следующим образом. Каждую итерацию происходит вычисление суммы вероятностей ($GetA0$), определение двух псевдослучайных чисел и нахождение шага по времени. Функция $GetA0$ в качестве параметра принимает вектор $reactions$ с экземплярами классов, описывающих различные реакции. В теле функции происходит последовательный вызов метода $GetProb$ каждого элемента вектора для суммирования вероятностей и возвращения значения a_0 . В каталоге с программой создается файл $output.txt$, в котором фиксируются значения концентраций белка в определенные моменты времени. Каждую итерацию цикла $while$ происходит сравнение между текущим временем системы t и вспомогательной меткой $timer$. Когда t становится больше $timer$, происходит вывод информации в консоль, запись в текстовый файл, а также инкрементирование переменной $timer$. Сделано это в угоду облегчения файла для последующего считывания сторонними программами. За весь период работы в системе происходят миллионы реакций. Изучение динамики концентраций белка не требует запоминать их значения на каждом шаге. Перед тем как узнать, какая реакция должна выполняться следующей, проводится проверка на необходимость немедленного выполнения отложенной реакции вместо какой-либо другой. Это делается при помощи функции $findDelayReaction$, возвращающей логическую переменную. Записи о запланированных реакциях хранятся в виде структуры, содержащей всего два пункта — идентификатор и время, в которое она должна совершиться. После проверки на наличие в массиве каких-либо записей происходит сравнение текущего значения t с прибавкой к нему планируемого шага и времени первой заготовленной реакции в векторе $delR$. Как только не выполняется хотя бы одно из условий, функция возвращает $false$. Если функция возвращает положительный ответ, то по идентификатору, хранящемуся в каждой записи стека отложенных реакций, совершается соответствующая операция с текущими значениями концентраций в системе. Переменной t присваивается значение, которое было указано в отложенной реакции, и она удаляется из вектора. Если же выполнять отложенную реакцию на данном шаге не нужно, то необходимо найти идентификатор следующей реакции. При помощи функции $GetID$ можно получить идентификатор реакции, которая должна совершиться на текущем шаге. Последовательное суммирование вероятностей и инкрементирование идентификатора продолжается до тех пор, пока значение суммы не станет больше, чем произведение r_2 и a_0 . По окончании цикла возвращается значение идентификатора реакции, которая должна совершиться на текущем шаге. Если у реакции указано время запаздывания, то ее необходимо положить в соответствующий вектор (рис. 4).

```

int id = getID(reactions, r2 * a0);
if (reactions[id]->GetDelay() != 0)
{
    delR.push_back(Reaction{ id,
        t + reactions[id]->GetDelay() });
    t += dt;
}

```

Рис. 4. Фрагмент кода программы, написанной на языке программирования C++: проверка выпавшей реакции на наличие времени запаздывания и добавление элемента в вектор

В случае обычной реакции она сразу же исполняется. В обоих случаях совершается прыжок по времени, равный значению переменной Δt . Вышеописанный цикл продолжает работать до тех пор, пока значение переменной t не сравняется или станет больше, чем указанное пользователем время завершения расчета — *endTime*.

3.4. Стохастическое исследование авторепрессора с запаздыванием

Рассмотрим динамические свойства авторепрессора с запаздывающей обратной связью, который представляет собой один ген x , регулирующий свою работу с помощью молекулярного поля димеров (рис. 5). Список реакций авторепрессора приведен в табл. 2. В предельном случае быстрой динамики оператор-сайта и процессов димеризации/дедимеризации в сравнении с процессами синтеза и деградации белка детерминистское описание авторепрессора сводится к одному уравнению с запаздыванием [Bratsun et al., 2005a; Брацун, 2011]:

$$\frac{dx}{dt} = \frac{1}{1 + 4\delta x + \frac{4\epsilon\delta x}{(1 + \epsilon\delta x^2)^2}} \left(\frac{A}{1 + \epsilon\delta x^2(t - \tau)} - Bx \right), \quad (8)$$

где $x(t)$ — непрерывная функция времени, обозначающая концентрацию мономеров X .

Таблица 2. Список реакций при генной регуляции запаздывающего авторепрессора (рис. 5)

1	Димеризация белка X	$X + X \xrightarrow{k_1} X_2$
2	Дедимеризация белка X	$X_2 \xrightarrow{k_{-1}} X + X$
3	Связывание димера авторепрессора с оператор-сайтом	$D_0 + X_2 \xrightarrow{k_2} D_1$
4	Освобождение димера авторепрессора оператор-сайтом	$D_1 \xrightarrow{k_{-2}} D_0 + X_2$
5	Синтез белка X при открытом оператор-сайте, протекающий с запаздыванием	$D_0(t) \xrightarrow{A} D_0 + X^{t+\tau}$
6	Деградация белка X	$X \xrightarrow{B} \emptyset$

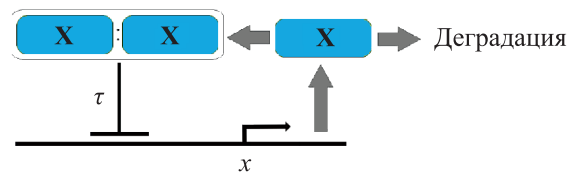


Рис. 5. Схематическое изображение авторепрессора с запаздыванием

Таблица 3. Список вероятностей (склонностей) реакционных каналов авторепрессора, перечисленных в таблице 2, необходимых при реализации алгоритма Гиллеспи

μ	1	2	3	4	5	6
a_μ	$k_1 X^2$	$k_{-1} X_2$	$k_2 D_0 X_2$	$k_{-2} D_1$	$A D_0(t - \tau)$	$B X$

Если в уравнении (8) положить запаздывание $\tau = 0$, то оно с точностью до префактора превращается в модельное уравнение Гудвина [Goodwin, 1965] с небольшим значением коэффициента функции Хилла, в котором колебания не наблюдаются. В фазовом пространстве имеется одно-единственное состояние равновесия x^* , которое для случая малых ε , δ может быть записано как

$$x^* = \frac{1}{2\varepsilon\delta} \left(\sqrt{\frac{B^2}{A^2} + 4\varepsilon\delta} - \frac{B}{A} \right) \approx \frac{A}{B}. \quad (9)$$

Рис. 6. Нейтральная кривая колебательной неустойчивости авторепрессора на плоскости параметров производства и деградации белка, полученная в рамках детерминистского описания системы (8). Отношения скоростей быстрых реакций фиксированы: $\varepsilon = 0.1$, $\delta = 0.2$. Неустойчивость находится выше кривой. Точки соответствуют значениям параметров, для которых выполнялись стохастические вычисления (см. рис. 7–9): $B = 4$, $\tau = 20$, $A = 50$ (ниже кривой), $A = 70$ (выше кривой)

Предположение о малости констант ε и δ является вполне разумным, так как для стабильности процесса скорость обратной реакции должна быть на порядок выше. В итоге равновесие определяется балансом производства и деградации белка. Можно показать, что стационарное решение (9) в отсутствие запаздывания абсолютно устойчиво, что отражает точную работу регуляторной функции авторепрессора: отрицательная обратная связь блокирует работу оператор-сайта ровно настолько, чтобы поддерживать нужный уровень экспрессии гена. Префактор, который появляется из-за корректного подсчета белковых молекул в системе, не влияет на состояния равновесия, но существенно влияет на характер движения фазовых траекторий.

Нелинейная динамика запаздывающего авторепрессора (8) впервые была исследована в [Bratsun et al., 2005a]. Было показано, что равновесие (9) может потерять устойчивость только

колебательным образом, порождая предельный цикл. Формула для инкремента неустойчивости может быть получена аналитически [Брацун, 2011]:

$$\lambda = \frac{1}{\tau} \Lambda(-2\tau A \varepsilon \delta X^* e^{B\tau}) - B, \quad (10)$$

где Λ — функция Ламберта, которая по определению удовлетворяет выражению $\Lambda(x) \exp(\Lambda(x)) = x$. Нейтральная кривая для бифуркации Андронова–Хопфа, которая определяется условием $\lambda = I\omega$, где ω — частота осцилляций предельного цикла в момент его ответвления, представлена на рис. 6 на плоскости параметров деградации и синтеза. Так как положительная петля обратной связи в модели (8) отсутствует, а внутренний источник энергии системы за счет процессов транскрипции/трансляции гена балансируется регуляцией через отрицательную обратную связь, то ничего более сложного, чем предельный цикл в надкритической области, система не демонстрирует. Сами колебания возникают из-за запаздывания и отражают неспособность авторепрессора вовремя реагировать на изменения своего состояния. Это стандартная причина развития колебаний в теории управления.

Сформулируем теперь задачу для проведения численного исследования с помощью модифицированного алгоритма Гиллеспи. В случае запаздывающего авторепрессора имеется шесть реакционных каналов (см. табл. 2) и четыре химических соединения X_i : (X , X_2 , D_0 , D_1). Опираясь на кинетику реакций, вероятности (склонности) реализации химических каналов a_μ можно записать так, как это представлено в таблице 3. Как видно из таблицы, вероятности являются динамически меняющимися величинами и в зависимости от текущих значений концентраций реагентов могут меняться в широких пределах. Из экспериментальных наблюдений за процессами генной регуляции известно, что динамика оператор–сайта и процессы димеризации/димеризации (первые четыре канала) протекают на порядок быстрее процессов синтеза и деградации белка (последние два канала):

$$k_1, k_{-1}, k_2, k_{-2} \gg A, B. \quad (11)$$

Если значение характерного времени для первых процессов составляет секунды–минуты, то для вторых — часы. Условие (11) было использовано при выводе модели (8), которая по своему смыслу представляет собой описание медленной динамики системы на многообразии процессов синтеза и деградации белка. Таким образом, чтобы провести сравнение стохастического анализа с детерминистским описанием (8), нужно не просто задавать отношения скоростей быстрых реакций ε и δ , но также и контролировать в расчете сами скорости реакций k_i . Мы использовали значения $k_1 = 100$, $k_{-1} = 1000$, $k_2 = 200$, $k_{-2} = 1000$, которые гарантируют выполнение условия (11). Пятый канал является запаздывающим, и к нему должна быть применена модификация алгоритма [Bratsun et al., 2005a], которая обсуждалась выше. Запаздывание предполагается достаточно большим ($\tau = 20$), чтобы причинно-следственная связь между моментом выпадения запаздывающего реакционного канала и реальным исполнением реакции была бы разорвана (см. обсуждение в [Брацун, 2011]).

Для начала приведем тривиальный результат, полученный для авторепрессора без запаздывания ($\tau = 0$). В этом случае должна наблюдаться стационарная работа экспрессии гена (9). Рис. 7 показывает одну конкретную реализацию стохастического расчета, а также результат усреднения по 50 реализациям, представленным в сравнении с интегрированием детерминистского уравнения (8). Все расчеты выполнены с одинаковых начальных условий. Можно заметить, что алгоритм Гиллеспи хорошо воспроизводит детерминистское решение при усреднении по достаточно большому числу реализаций, демонстрируя стационарный режим производства белка.

Рассмотрим теперь случай авторепрессора с запаздыванием. Демонстрируемая динамика здесь имеет гораздо более сложный характер. Приведем два характерных стохастических расчета, проведенных выше и ниже нейтральной кривой (отмечены точками на рис. 6). Было обнаружено два важных эффекта, которые качественно отличают стохастическую динамику от детерминистской. Рис. 8 показывает две разные численные реализации (не усредненные), прове-

денные для одних и тех же начальных условий. Значения управляющих параметров соответствуют точке чуть выше нейтральной кривой: $B = 4$, $A = 70$ (рис. 6). При той же скорости деградации нейтральное значение скорости производства белка, рассчитанное в рамках детерминистской модели (8), равно $A^* = 57.85$. Выше этого значения в системе развиваются периодические колебания. В целом стохастическое моделирование подтверждает этот вывод: колебания действительно возникают. Представленная на рис. 8 динамика числа отложенных реакций открывает нам детали механизма колебаний: при накоплении числа отложенных реакций концентрация мономеров в системе резко снижается, а когда начинается процесс выполнения этих реакций, число мономеров быстро растет. Количество отложенных в память реакций на пике достигает значения 900 и падает до 300 в случае полной «разрядки» (при указанных значениях параметров).

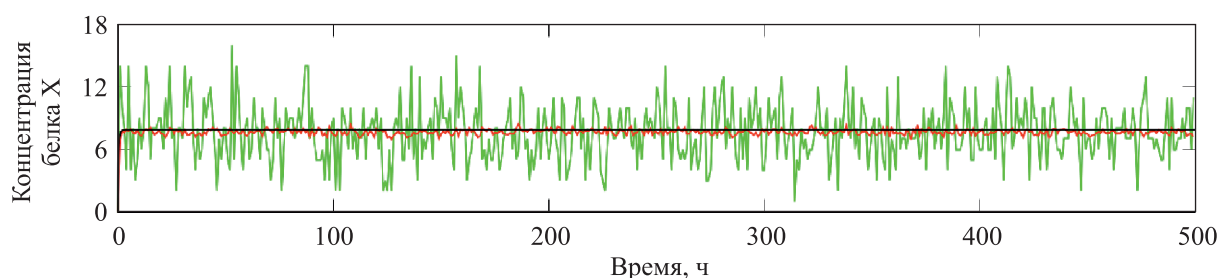


Рис. 7. Сравнение временной эволюции авторепрессора без запаздывания, рассчитанной в рамках детерминистского интегрирования уравнения (8) (черная линия) и стохастического расчета (красная линия), усредненного по 50 реализациям (одна из реализаций приведена явно и выделена зеленым цветом). Параметры расчета: $A = 70$, $B = 4$, $\tau = 0$, $\varepsilon = 0.1$, $\delta = 0.2$. Начальные условия расчета в обоих случаях: отсутствие мономеров белка при открытом оператор-сайте

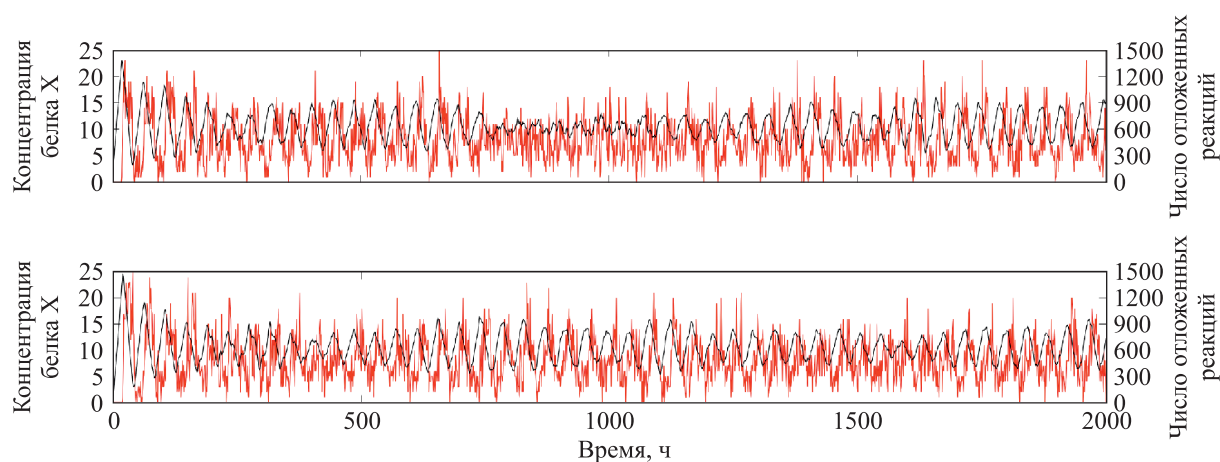


Рис. 8. Две характерные реализации стохастического расчета с одних и тех же начальных условий, иллюстрирующие самопроизвольный сдвиг фазы осцилляций за счет внезапной деградации колебаний в надкритической области параметров. Параметры расчета: $A = 70$, $B = 4$, $\tau = 20$, $\varepsilon = 0.1$, $\delta = 0.2$. Начальные условия расчета в обоих случаях: отсутствие мономеров белка при открытом оператор-сайте. Мгновенное значение числа отложенных реакций отмечено черной линией (шкала справа)

На временной развертке первой реализации, показанной на рис. 8, хорошо виден эффект спонтанного вырождения колебаний, при котором число отложенных реакций практически не меняется: оно какое-то время слабо флуктуирует около значения 600. Можно сказать, что между временными отметками 800 и 1100 авторепрессор успевает вовремя отправлять управляющие сигналы, регулирующие экспрессию, и переходит в квазистационарный режим работы. Через некоторое время колебания восстанавливаются, однако их фаза может измениться случайным образом по отношению к первоначальным осцилляциям. Вторая реализация показыва-

ет, что такое увядание колебаний в надкритической области в указанном диапазоне времен не наблюдается. Таким образом, этот эффект может произойти в ходе эволюции системы случайно в любой момент времени. При усреднении сигнала по большому ансамблю реализаций мы получаем колебательный процесс, который существенно отличается от детерминистских колебаний: например, усредненная амплитуда стохастических колебаний примерно в два раза меньше, чем у колебаний, генерируемых уравнением (8). Но на самом деле, когда мы рассматриваем отдельные реализации, видим, что амплитуда колебаний в среднем одна и та же, но при этом время от времени самопроизвольно сдвигается фаза колебаний (рис. 9).

Похожий процесс, но с обратным знаком, происходит в подкритической области. Как было показано в работах [Bratsun et al., 2005a; Брацун, 2011], авторепрессор с запаздыванием в стохастическом случае демонстрирует в подкритической области квазирегулярные колебания с частотой, которая близка к критической частоте колебаний $\omega = 2\pi / 2\tau$, измеренной на нейтральной кривой. Для запаздывания $\tau = 20$ частота колебаний примерно равна $\omega \approx 0.159$. Механизм спонтанной генерации колебаний связан здесь с одновременным действием запаздывания и шума. Отсылая читателя к цитированным выше работам, мы приводим динамику корреляционной функции стохастического сигнала в подкритической области, которая показывает, что ниже нейтральной кривой вполне комфортно существуют периодические колебания (рис. 10).

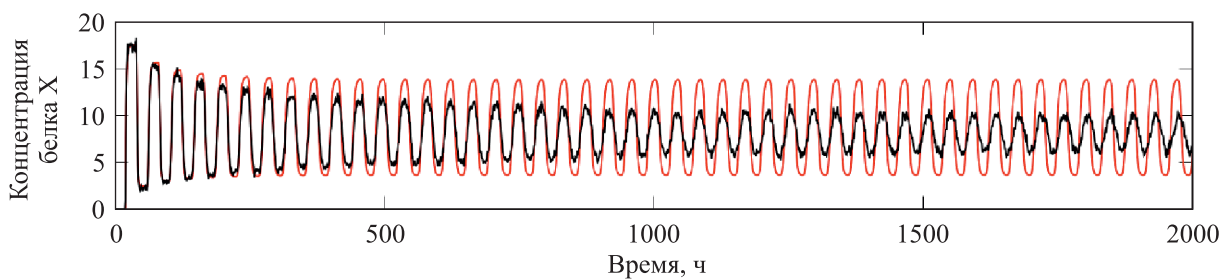


Рис. 9. Сравнение временной эволюции авторепрессора, рассчитанной в рамках детерминистского (красная линия) и стохастического (черная линия) описания, усредненного по 50 реализациям. Параметры расчета и начальные условия такие же, как на рис. 8

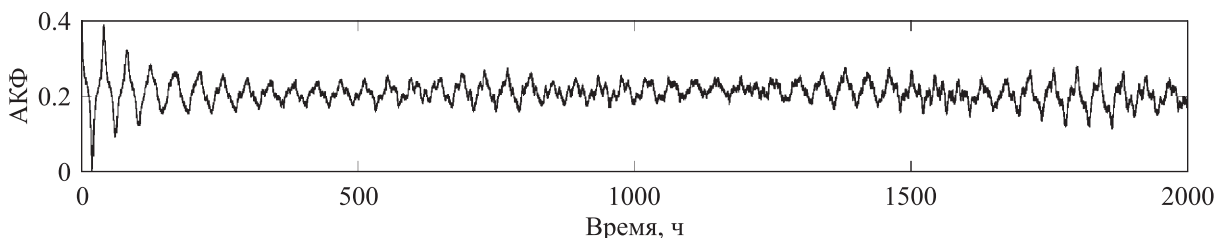


Рис. 10. Автокорреляционная функция стохастического сигнала, вычисленная в подкритической области значений параметров при $A = 50$, $B = 4$, $\tau = 20$, $\varepsilon = 0.1$, $\delta = 0.2$ (точка ниже нейтральной кривой на рис. 6)

4. Стохастическое исследование репрессиллятора с запаздыванием

4.1. Подготовительные процедуры

Согласно принципиальной схеме устройства запаздывающего репрессиллятора (рис. 1) и полного списка реакций (табл. 1), у данной системы всего имеется 18 реакционных каналов. Три из этих каналов являются запаздывающими. В реакциях принимает участие 12 химических соединений.

Для проведения стохастических расчетов необходимо определить вероятности (склонности) a_{μ} наступления реакционного события для всех каналов. Полный список этих величин приведен в таблице 4.

Таблица 4. Список вероятностей (склонностей) реакционных каналов репрессиллятора (см. табл. 1), необходимых для реализации алгоритма Гиллеспи

μ	1	2	3	4	5	6
a_μ	$k_1 X^2$	$k_{-1} X_2$	$k_2 D_0^X Z_2$	$k_{-2} D_1^X$	$A D_0^X(t - \tau)$	BX
μ	7	8	9	10	11	12
a_μ	$k_1 Y^2$	$k_{-1} Y_2$	$k_2 D_0^Y X_2$	$k_{-2} D_1^Y$	$A D_0^Y(t - \tau)$	BY
μ	13	14	15	16	17	18
a_μ	$k_1 Z^2$	$k_{-1} Z_2$	$k_2 D_0^Z Y_2$	$k_{-2} D_1^Z$	$A D_0^Z(t - \tau)$	BZ

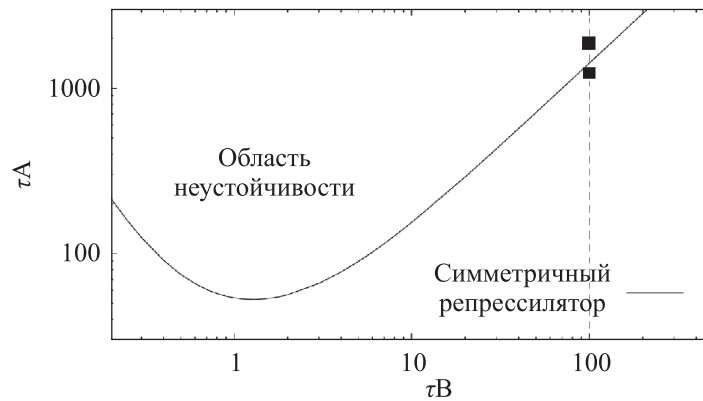


Рис. 11. Нейтральная кривая колебательной неустойчивости симметричного репрессиллятора с запаздыванием на плоскости параметров производства и деградации белка, полученная в рамках детерминистского описания системы (1)–(3). Отношения скоростей быстрых реакций фиксированы: $\varepsilon = 0.1$, $\delta = 0.2$. Неустойчивость находится выше кривой. Точки соответствуют значениям параметров, для которых выполнялись стохастические вычисления (см. рис. 12–15): $B = 5$, $\tau = 20$, $A = 70$ (выше нейтральной кривой), $A = 80$ (выше кривой)

4.2. Устойчивость равновесия

В работе [Браун и др., 2018] было показано, что система уравнений с запаздыванием (1)–(3) имеет единственное состояние равновесия $\{x^*, y^*, z^*\}$, устойчивость которого может меняться только колебательным образом. В случае симметричного репрессиллятора нейтральная кривая для бифуркации Андронова–Хопфа неявно задается следующим алгебраическим уравнением:

$$\operatorname{Re} \left[\Lambda \left(2 \sqrt[3]{-1} - \frac{\tau \varepsilon \delta X^{*3} B^2}{A(1 + 4\varepsilon X^*)} e^{\frac{\tau B}{1 + 4\varepsilon X^*}} \right) \right] - \frac{\tau B}{1 + 4\varepsilon X^*} = 0, \tag{12}$$

где предполагается, что

$$x^* = y^* = z^* \equiv X^*. \tag{13}$$

Рассмотрим характерный срез карты устойчивости на рис. 11. При уровне деградации белка $B = 5$ нейтральное значение скорости производства белка равно $A^* = 70.81$. На рис. 12 представлено сравнение между детерминистским и стохастическим расчетами, выполненными при небольшой надкритичности $A = 80$ (точка выше нейтральной кривой на рис. 11). Как и в случае запаздывающего авторепрессора, вблизи нейтральной кривой было обнаружено явление спонтанной деградации колебаний и произвольной смены фазы колебаний, что находит свое выражение в уменьшении амплитуды колебаний усредненного сигнала. Рисунок показывает, что стандартное усреднение по реализациям, используемое в стохастическом анализе, приводит к искажению картины: график можно интерпретировать как затухание колебаний (рис. 12, a).

Реализация одного конкретного расчета (рис. 12, б) показывает, что в системе продолжают происходить полномасштабные колебания. Между тем процессы в репрессилляторе протекают сложнее, чем в авторепрессоре. Ниже будет показано, что в сильно нелинейной области в системе конкурируют два колебательных режима. В одном режиме все белковые поля осциллируют почти синхронно, одновременно вырастая и затухая с характерным периодом, близким к 2τ (при $\tau = 20$ период колебаний равен 40). Другой режим характеризуется поочередной, т. е. асинхронной, экспрессией генов, с характерным периодом колебаний, близким к 6τ (период колебаний близок к 120). В стохастическом случае вблизи порога можно наблюдать спонтанное возбуждение любого из этих режимов, а также внезапную деградацию колебаний (рис. 12, б).

Сравнение рис. 12, а и рис. 12, б также позволяет оценить уровень флуктуаций в стохастической системе: если в детерминистском случае амплитуда колебаний при указанных значениях параметров имеет размах от 4 до 12 молекул, то в стохастическом случае число молекул белка X меняется от 0 до 25 штук.

В подкритической области, как и в случае авторепрессора, наблюдается существование квазирегулярных колебаний, что иллюстрируется автокорреляционной функцией сигнала, рассчитанной при $A = 70$ (нижняя точка на рис. 11) и приведенной на рис. 13. Автокорреляции свидетельствуют, что динамическая система явно чувствительна к синхронным колебаниям с периодом 2τ . Интересной особенностью процесса является постепенное усиление автокорреляций для все более дальних точек временного ряда, что свидетельствует о постепенной самонастройке репрессиллятора.

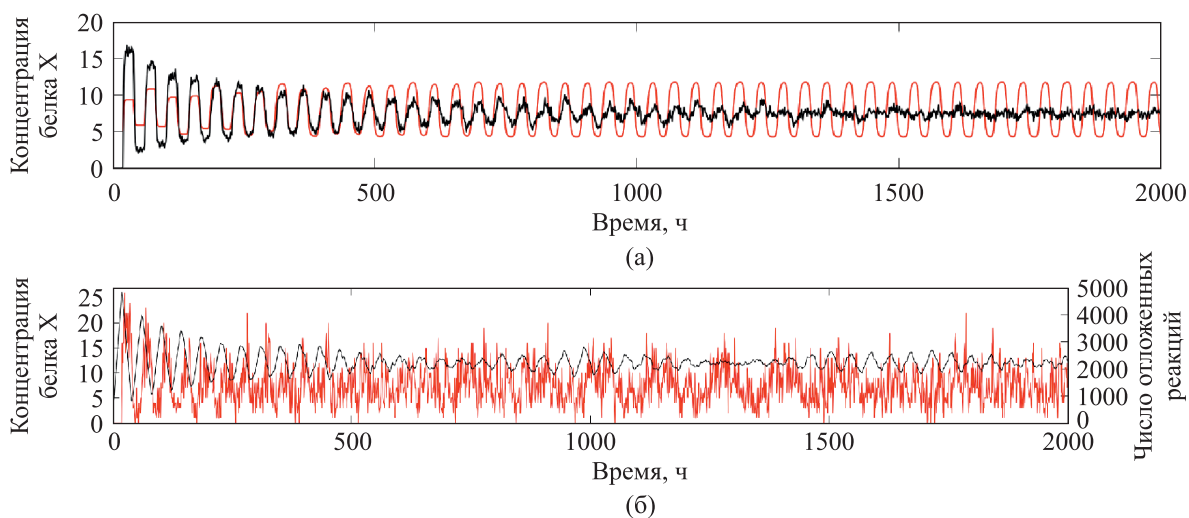


Рис. 12. (а) Сравнение временной эволюции концентрации белка X симметричного репрессиллятора, рассчитанной в рамках детерминистского интегрирования (красная линия) и стохастического расчета (черная линия), усредненного по 50 реализациям. (б) Характерная реализация стохастического расчета, иллюстрирующая спонтанную деградацию и сдвиг фазы осцилляций. Параметры расчета во всех случаях: $B = 5$, $A = 80$, $\varepsilon = 0.1$, $\delta = 0.2$, $\tau = 20$ (выше нейтральной кривой на рис. 11). Начальные условия расчета: (5, 6, 6) при открытом оператор-сайте. Мгновенное значение числа отложенных реакций отмечено черной линией (шкала справа)

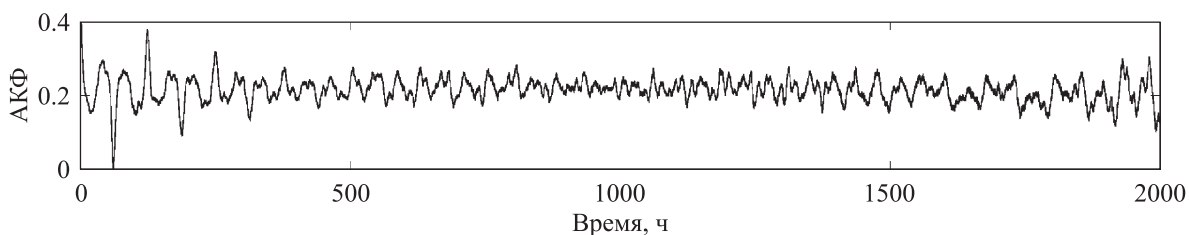


Рис. 13. Автокорреляционная функция стохастического сигнала симметричного репрессиллятора, вычисленная в подкритической области значений параметров при $B = 5$, $A = 70$, $\varepsilon = 0.1$, $\delta = 0.2$, $\tau = 20$

4.3. Переходный колебательный процесс на медленном многообразии

Рассмотрим теперь динамику сильно нелинейных процессов. При детерминистском анализе репрессиллятора с запаздыванием [Браун и др., 2018] было показано, что выход системы к предельному циклу, который отражает кооперативную работу генов, происходит через медленное многообразие. Если задать начальные условия интегрирования таким образом, чтобы концентрации белков были примерно равны и невелики, то фазовая траектория быстро притягивается к этому образованию, а затем медленно эволюционирует по нему, в конце концов покидая и притягиваясь к циклу. На рис. 14, а представлен пример детерминистского расчета с начального условия (5, 6, 6), в котором фазовая траектория периодически меняет свой цвет каждые 100 000 единиц времени, пробегая за это время основные цвета спектра. Оказалось, что общее время интегрирования до момента срыва траектории с многообразия превышает 450 000 единиц времени. Природа притягивающего многообразия связана с симметрией системы, хотя и в случае несимметричного репрессиллятора это явление себя также проявляет [Браун и др., 2018]. Можно заметить, что наиболее долгая эволюция к предельному циклу происходит при более симметричных начальных условиях. Когда начальные концентрации белков близки друг другу, то эволюция системы протекает в долгой конкурентной борьбе между генами за лидерство. Так как белки выступают в качестве репрессора экспрессии соседнего гена, то при равных концентрациях гены работают почти синхронно во времени. В конце концов система приходит к режиму работы, при котором гены включаются и выключаются не все сразу, а по очереди. Последнее и является нормальной работой репрессиллятора. Конечно, переходный процесс длиной в 450 000 единиц времени (по традиции решения задач в математической генетике мы не обезразмеривали задачу, поэтому вполне естественно рассматривать час в качестве единицы времени [Браун, Захаров, 2011]) выглядит малореалистично. Для расчета на рис. 14, а переходный режим составляет порядка 18 750 дней, или более 50 лет!

Выше подчеркивалось, что процессы генной регуляции принципиально стохастичны, так как в них принимает участие небольшое количество молекул. Поэтому стохастическое моделирование осцилляций репрессиллятора в режиме движения по медленному многообразию представляет особый интерес. Рис. 15 показывает сравнение временной эволюции системы при детерминистском и стохастическом описании. Можно заметить, что стохастический расчет

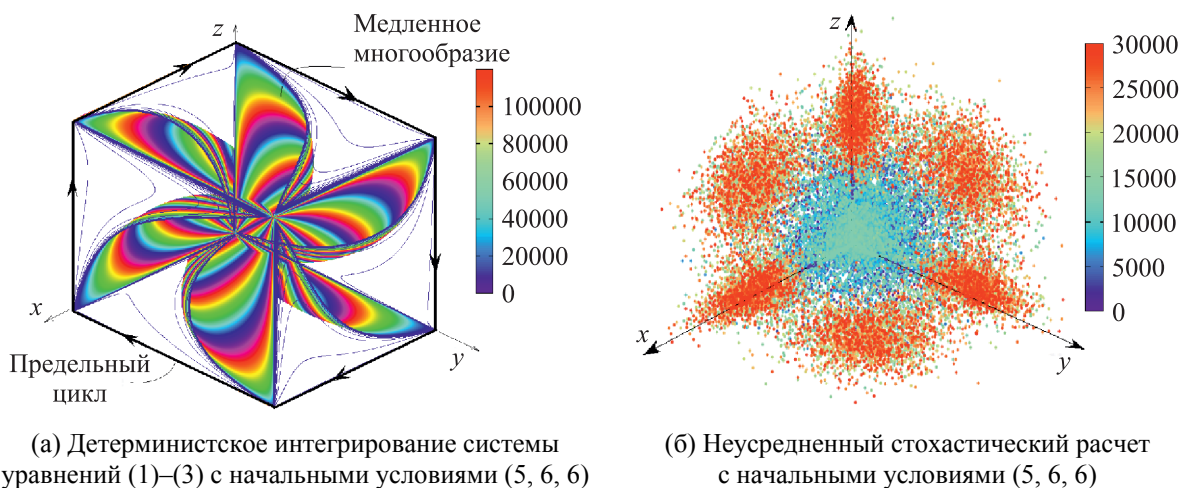


Рис. 14. Фазовый портрет симметричного репрессиллятора, полученный в рамках детерминистского (слева) и стохастического (справа) описания. Параметры расчета в обоих случаях: $B = 5$, $A = 120$, $\tau = 20$, $\varepsilon = 0.1$, $\delta = 0.2$. При любых начальных условиях фазовая траектория притягивается к предельному циклу, который отражает кооперативную (асинхронную) работу трех генов. Для визуализации медленного многообразия цвет фазовой траектории менялся периодически, пробегая спектр за 100 000 (детерминистский расчет) и 30 000 (стохастический расчет) единиц времени. Цветовые шкалы отображают соответствующее изменение цвета в рамках одного цикла

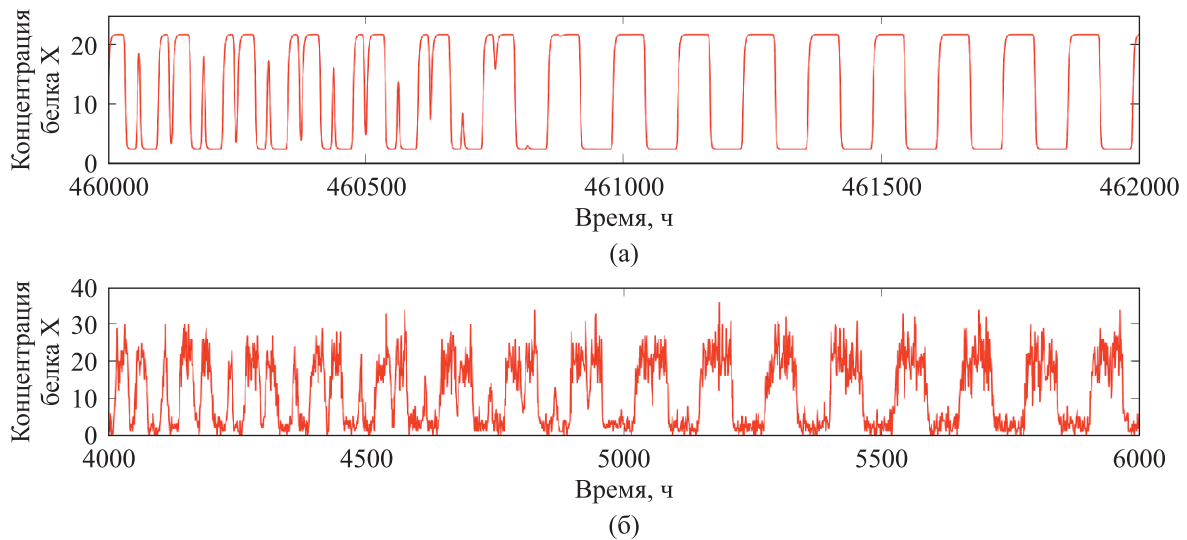


Рис. 15. Сравнение процессов срыва фазовой траектории с медленного многообразия в случае детерминистического (а) и стохастического (б) описания для симметричного репрессиллятора. Начальные условия расчета одинаковы в обоих случаях: (5, 6, 6). Параметры расчета фиксированы: $B = 5$, $A = 120$, $\tau = 20$, $\varepsilon = 0.1$, $\delta = 0.2$

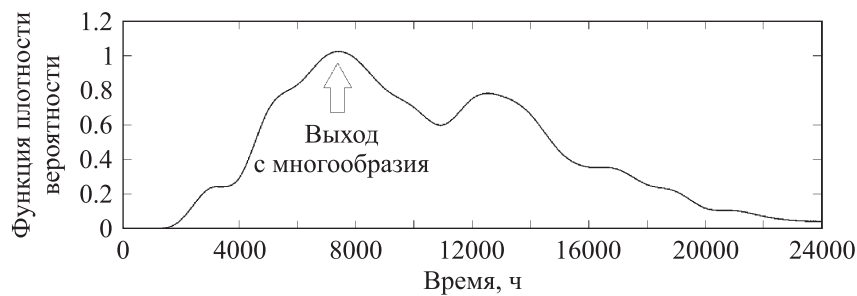


Рис. 16. Распределения вероятности срыва фазовой траектории с медленного многообразия как функция времени от начала стохастического расчета с начальных условий (5, 6, 6). Распределение построено на основе усреднения более 500 отдельных реализаций. Параметры расчета для симметричного репрессиллятора фиксированы: $B = 5$, $A = 120$, $\tau = 20$, $\varepsilon = 0.1$, $\delta = 0.2$

принципиально отличается ранним срывом фазовой траектории с медленного многообразия и переходом к когерентным колебаниям с периодом примерно 120 единиц времени. В этом случае переход происходит в 100 раз быстрее! Рис. 14, б показывает фазовый портрет стохастической динамической системы: если на начальном этапе эволюции колебания происходят вблизи диагонали $x = y = z$, то уже по истечении 15 000–20 000 единиц времени система переходит к поочередным включениям генов. Таким образом, флуктуации помогают стремительной самонастройке системы.

Понятно, что сделанный вывод носит предварительный характер, так как на рис. 15, б представлена только одна конкретная реализация расчета. В условиях случайного характера флуктуаций срыв фазовой траектории с многообразия гипотетически может произойти в любой момент эволюции. Поэтому следующий график, представленный на рис. 16, имеет принципиальное значение: на нем изображено распределение вероятности спонтанного перехода динамической системы к коллективным колебаниям как функция времени, отсчитываемого от начала эволюции. Распределение построено на основе обобщения более 500 отдельных численных реализаций, стартующих с одного и того же начального условия. Хорошо видно, что распределение имеет «двугорбый» характер, а наиболее вероятное время срыва траектории с многообразия составляет порядка 7 тыс. единиц времени. Среднее значение распределения имеет слегка большее значение из-за несимметричной структуры последнего.

5. Обсуждение и заключение

Работа посвящена стохастическому описанию поведения модели репрессиллятора с запаздыванием, предложенной ранее в первой части работы [Брацун и др., 2018]. Так как процессы генной регуляции включают небольшое число белковых молекул, динамика репрессиллятора *in vivo* является принципиально стохастической и система испытывает сильные флуктуации белковых полей. Именно поэтому анализ динамических свойств запаздывающего репрессиллятора, представленный в рамках детерминистского подхода в цитированной выше работе, не является достаточным.

Основным инструментом стохастического анализа малоразмерных генетических систем, испытывающих внутренний шум химических реакций генной регуляции, является алгоритм Гиллеспи [Gillespie, 1977]. Абсолютное большинство работ, цитированных во введении, посвящено исследованию марковских стохастических систем. Модель репрессиллятора без запаздывания для описания колебаний требует как минимум учета процесса трансляции, что увеличивает размерность системы вдвое. Стохастическое описание такой модели было выполнено уже в пионерской работе [Elowitz, Leibler, 2000]. Авторы абсолютного большинства работ, посвященных изучению стохастических свойств репрессиллятора, пошли этим же путем. В данной работе мы вводим в рассмотрение запаздывание с самого начала и в численном анализе используем модификацию алгоритма Гиллеспи для немарковских систем [Bratsun et al., 2005b]. Нелинейная динамика, демонстрируемая в рамках модели запаздывающего репрессиллятора, существенно отличается от модели без запаздывания.

В рамках стохастического анализа симметричного репрессиллятора с запаздыванием показано, что взаимодействие шума и запаздывания может приводить не только к количественным, но и качественным последствиям для системы. Вблизи нейтральной кривой система демонстрирует процесс спонтанной деградации и возбуждения периодических колебаний, при этом каждый раз фаза колебаний восстановленных колебаний случайно меняется. Показано, что стандартное усреднение такого динамического поведения с блуждающей фазой по расчетным реализациям приводит к квазистационарному поведению, что не соответствует действительности. В случае больших закритичностей обнаружено, что шум способствует более эффективной самонастройке симметричного репрессиллятора на совместную работу генов. Если в рамках детерминистского описания система вынуждена долго эволюционировать по медленному многообразию, то флуктуации приводят к быстрому срыву с многообразия и переходу к поочередным циклам экспрессии генов.

Список литературы (References)

- Брацун Д. А. Эффект возбуждения подкритических колебаний в стохастических системах с запаздыванием. Часть I. Регуляция экспрессии генов // Компьютерные исследования и моделирование. — 2011. — Т. 3, № 4. — С. 421–438.
- Bratsun D. A. Effekt vozbužhdeniya podkriticheskikh kolebanij v stohasticheskikh sistemah s zapazdyvaniem. Chast' I. Reguljatsija ekspressii genov [Effect of subcritical excitation of oscillations in stochastic systems with time delay. Part I. Regulation of gene expression] // Computer Research and Modeling. — 2011. — Vol. 3, No. 4. — P. 421–438 (in Russian).
- Брацун Д. А., Захаров А. П. Моделирование пространственно-временной динамики циркадианных ритмов *Neurospora crassa* // Компьютерные исследования и моделирование. — 2011. — Т. 3, № 2. — С. 191–213.
- Bratsun D. A. Modelirovanie prostranstvenno-vremennoj dinamiki tsirkadiannyh ritmov *Neurospora crassa* [Modeling spatio-temporal dynamics of circadian rhythms in *Neurospora crassa*] // Computer Research and Modeling. — 2011. — Vol. 3, No. 2. — P. 191–213 (in Russian).
- Брацун Д. А., Лоргов Е. С., Полуянов А. О. Репрессиллятор с запаздывающей экспрессией генов. Часть I. Детерминистское описание // Компьютерные исследования и моделирование. — 2018. — Т. 10, № 2. — С. 241–259.

- Bratsun D. A., Lorgov E. S., Poluyanov A. O.* Repressilyator s zapazdyvayushchej ekspressiej genov. Chast' I. Deterministskoe opisanie [Repressilator with time-delayed gene expression. Part I. Deterministic description] // *Computer Research and Modeling*. — 2018. — Vol. 10, No. 2. — P. 241–259 (in Russian).
- Волков Е. И., Журов Б. А.* Динамическое поведение изолированного репрессилатора с обратной связью // *Известия вузов. Радиофизика*. — 2013. — Т. 56, № 10. — С. 774–786.
- Volkov E. I., Zhurov B. A.* Dynamic Behavior of an Isolated Repressilator with Feedback // *Radiophysics and Quantum Electronics*. — 2014. — Vol. 56, Iss. 10. — P. 697–707. (Original Russian paper: *Volkov E. I., Zhurov B. A.* Dinamicheskoe povedenie izolirovannogo repressiljatora s obratnoj svyaz'ju // *Izvestija vuzov. Radiofizika*. — 2013. — Vol. 56, No. 10. — P. 774–786.)
- Демиденко Г. В., Колчанов Н. А., Лихошвай В. А., Матушкин Ю. Г., Фадеев С. И.* Математическое моделирование регулярных контуров генных сетей // *Ж. вычисл. матем. и матем. физ.* — 2004. — Т. 44, № 12. — С. 2276–2295.
- Demidenko G. V., Kolchanov N. A., Likhoshvai V. A., Matushkin Yu. G., Fadeev S. I.* Mathematical modeling of regular circuits of gene networks // *Computational Mathematics and Mathematical Physics*. — 2004. — Vol. 44, Iss. 12. — P. 2166–2183. (Original Russian paper: *Demidenko G. V., Kolchanov N. A., Likhoshvai V. A., Matushkin Yu. G., Fadeev S. I.* Matematicheskoe modelirovanie reguljarnyh konturov gennyh setej // *Zhurnal vychislitel'noj matematiki i matematicheskoi fiziki*. — 2004. — Vol. 44, No. 12. — P. 2276–2295.)
- Иванченко М. В., Канаков О. И., Котельников Р. А., Крылов И. Б.* Нелинейная динамика синтетических сетей генной регуляции // *Изв. вузов «ПНД»*. — 2013. — Т. 21, № 4. — С. 77–92.
- Ivanchenko M. V., Kanakov O. I., Kotel'nikov R. A., Krylov I. B.* Nelinejnaya dinamika sinteticheskikh setej gennoj reguljacji // *Izvestiya VUZ. Applied Nonlinear Dynamics*. — 2013. — Vol. 21, No. 4. — P. 77–92 (in Russian).
- Лихошвай В. А., Хлебодарова Т. М., Колчанов Н. А.* Компьютерный ресурс «Генетический конструктор» для моделирования молекулярно-генетических процессов в бактериальной клетке: анализ циклической генной цепи // *Роль микроорганизмов в функционировании живых систем: фундаментальные проблемы и биоинженерные приложения*. — Изд-во СО РАН Новосибирск, 2010. — С. 392–404.
- Likhoshvai V. A., Khlebodarova T. M., Kolchanov N. A.* Komp'yuternyj resurs "Geneticheskij konstruktor" dlya modelirovaniya molekulyarno-geneticheskikh protsessov v bakterial'noj kletke: analiz tsiklicheskoj gennoj tsepi // *Rol' mikroorganizmov v funktsionirovanii zhivyh sistem: fundamental'nye problemy i bioinzenernye prilozhenija*. — SO RAN Novosibirsk, 2010. — P. 392–404 (in Russian).
- Лихошвай В. А., Матушкин Ю. Г., Фадеев С. И.* Задачи теории функционирования генных сетей // *Сиб. журн. индустр. матем.* — 2003. — Т. 6, № 2. — С. 64–80.
- Likhoshvai V. A., Matushkin Yu. G., Fadeev S. I.* Zadachi teorii funktsionirovanija gennyh setej [Problems of theory of gene networks functioning] // *Siberian Journal of Industrial Mathematics*. — 2003. — Vol. 6, No. 2. — P. 64–80 (in Russian).
- Потапов И. С., Волков Е. И.* Анализ динамических режимов взаимодействующих синтетических генетических репрессилаторов // *Компьютерные исследования и моделирование*. — 2010. — Т. 2, № 4. — С. 403–418.
- Potapov I. S., Volkov E. I.* Analiz dinamicheskikh rezhimov vzaimodejstvujushchih sinteticheskikh geneticheskikh oscilljatorov [Dynamics analysis of coupled synthetic genetic repressilators] // *Computer Research and Modeling*. — 2010. — Vol. 2, No. 4. — P. 403–418 (in Russian).
- Bennett M. R., Volfson D., Tsimring L., Hasty J.* Transient dynamics of genetic regulatory networks // *Biophys. J.* — 2007. — Vol. 92. — P. 3501.
- Bratsun D. A., Merkuriev D. V., Zakharov A. P., Pismen L. M.* Multiscale modeling of tumor growth induced by circadian rhythm disruption in epithelial tissue // *J. Biol. Phys.* — 2016. — Vol. 42, No. 1. — P. 107–132.
- Bratsun D., Volfson D., Hasty J., Tsimring L.* Delay-induced stochastic oscillations in gene regulation // *Proc. Natl. Acad. Sci. U.S.A.* — 2005a. — Vol. 102, No. 41. — P. 14593–14598.
- Bratsun D., Volfson D., Hasty J., Tsimring L.* Non-Markovian processes in Gene Regulation // *Noise in Complex Systems and Stochastic Dynamics III* / ed. by L. B. Kish, K. Lindenberg, Z. Gingl. Proceeding of SPIE. — 2005b. — Vol. 5845. — P. 210–219.
- Buse O., Kuznetsov A., Perez R. A.* Existence of limit cycles in the repressilator equations // *Int. J. Bifurcat. Chaos*. — 2009. — Vol. 19, No. 12. — P. 4097–4106.

- Buse O., Perez R., Kuznetsov A.* Dynamical properties of the repressilator model // *Phys. Rev. E.* — 2010. — Vol. 81. — Art. 066206.
- Danino T., Mondragón-Palomino O., Tsimring L., Hasty J.* A synchronized quorum of genetic clocks // *Nature.* — 2010. — No. 463. — P. 326–330.
- Denault D., Loros J., Dunlap J.* WC-2 mediates WC-1–FRQ interaction within the PAS protein-linked circadian feedback loop of *Neurospora* // *EMBO J.* — 2001. — Vol. 20. — P. 109–117.
- Elowitz M. B., Leibler S. A.* synthetic oscillatory network of transcriptional regulators // *Nature.* — 2000. — Vol. 403. — P. 335–338.
- Gillespie D. T.* Exact stochastic simulation of coupled chemical reactions // *J. Phys. Chem.* — 1977. — Vol. 81. — P. 2340–2361.
- Gillespie D. T.* The chemical Langevin equation // *J. Chem. Phys.* — 2000. — Vol. 113, No. 1. — P. 297–306.
- Goodwin B. C.* Oscillatory behavior in enzymatic control processes // *Adv. Enzyme Regul.* — 1965. — Vol. 3. — P. 425–438.
- Hasty J., Dolnik M., McMillen D., Collins J. J.* Designer gene networks: Towards fundamental cellular control // *Chaos.* — 2001. — Vol. 11. — P. 207–220.
- Kepler T. B., Elston C.* Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations // *Biophys. J.* — 2001. — Vol. 81. — P. 3116–3136.
- Lee K., Loros J. J., Dunlap J. C.* Interconnected feedback loops in the *Neurospora* circadian system // *Science.* — 2000. — Vol. 289, No. 5476. — P. 107–110.
- Loinger A., Biham O.* Stochastic simulations of the repressilator circuit // *Phys. Rev. E.* — 2007. — Vol. 76, No. 5. — Art. 051917.
- Muller S., Hofbauer J., Endler L., Flamm C., Widder S., Schuster P.* A generalized model of the repressilator // *J. Math. Biol.* — 2006. — Vol. 53. — P. 905.
- Potapov I., Zhurov B., Volkov E.* Multi-stable dynamics of the non-adiabatic repressilator // *J. R. Soc. Interface.* — 2015. — Vol. 12, No. 104. — Art. 20141315.
- Potvin-Trottier L., Lord N. D., Vinnicombe G., Paulsson J.* Synchronous long-term oscillations in a synthetic gene circuit // *Nature.* — 2016. — Vol. 538. — P. 514–517.
- Prasher D. C., Eckenrode V. K., Ward W. W., Prendergast F. G., Cormier M. J.* Primary structure of the *Aequorea victoria* green-fluorescent protein // *Gene.* — 1992. — Vol. 111, No. 2. — P. 229–33.
- Sriram K., Gopinathan M. S.* A two variable delay model for the circadian rhythm of *Neurospora crassa* // *J. Theor. Biol.* — 2004. — Vol. 231. — P. 23–38.
- Strelkova N., Barahona M.* Transient dynamics around unstable periodic orbits in the generalized repressilator model // *Chaos.* — 2011. — Vol. 21. — Art. 023104.
- Thattai M., Oudenaarden A.* Intrinsic noise in gene regulatory networks // *Proc. Natl. Acad. Sci. U.S.A.* — 2001. — Vol. 98. — P. 8614–8619.
- Tuttle L. M., Salis H., Tomshine J., Kaznessis Y. N.* Model-driven designs of an oscillating gene network // *Biophys. J.* — 2005. — Vol. 89, No. 6. — P. 3873–3883.
- Ullner E., Zaikin A., Volkov E. I., Garcia-Ojalvo J.* Multistability and clustering in a population of synthetic genetic oscillators via phase-repulsive cell-to-cell communication // *Phys. Rev. Lett.* — 2007. — Vol. 99. — Art. 148103.
- Yoda M., Ushikubo T., Inoue W., Sasai M.* Roles of noise in single and coupled multiple genetic oscillators // *J. Chem. Phys.* — 2007. — Vol. 126, No. 11. — Art. 03B612.