

НОВЫЕ РАДИОЭЛЕКТРОННЫЕ СИСТЕМЫ И ЭЛЕМЕНТЫ

УДК 519.876.5,004.89

КОМПАКТНАЯ ПОВЕДЕНЧЕСКАЯ МОДЕЛЬ НАНОКОМПОЗИТНОГО МЕМРИСТОРА

© 2023 г. И. А. Суражевский^{а, *}, В. В. Рыльков^{а, б}, В. А. Демин^а

^а Национальный исследовательский центр “Курчатовский институт”,
пл. Академика Курчатова, 1, Москва, 123182 Российская Федерация

^б Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова РАН,
пл. Введенского, 1, Фрязино Московской области, 141190 Российская Федерация

*E-mail: isurazhevsky@yandex.ru

Поступила в редакцию 31.10.2022 г.

После доработки 31.10.2022 г.

Принята к публикации 15.11.2022 г.

Предложена компактная поведенческая модель нанокompозитного мемристора $(\text{Co}_{40}\text{Fe}_{40}\text{B}_{20})_x(\text{LiNbO}_3)_{100-x}$, которая количественно описывает динамику изменения проводимости лабораторных образцов, а также реализует механизмы конечного времени хранения резистивных состояний, разбросов по напряжениям переключения от цикла к циклу и от устройства к устройству. Показана возможность реализации импульсной нейронной сети с синаптическими мемристорными связями на основе данной модели.

DOI: 10.31857/S0033849423100170, EDN: DNNGTF

ВВЕДЕНИЕ

В настоящее время невозможно представить создание интегральной микросхемы с большим количеством транзисторов без системы автоматизированного проектирования. Данный факт связан с крайне высокой скоростью усложнения процесса разработки из-за увеличения количества элементов на единицу площади кристалла, что стало очевидно начиная с 70-х годов XX в., когда данная работа велась в ручном режиме с помощью рубелита и лавсановой бумаги. Вместе с тем новые технологические процессы требовали все больших экономических затрат, что привело к пониманию необходимости компьютерной автоматизации как в области разработки интегральных микросхем (ИС), так и в области создания технологий микроэлектроники.

Необходимым компонентом для этого шага являлась разработка компактных моделей базовых компонентов ИС — транзисторов. Термин “компактная” предполагает в первую очередь вычислительную простоту, что крайне актуально для моделирования одновременной работы миллионов и миллиардов транзисторов. При этом данный подход повышает уровень удобства разработки ИС, обладает большей сходимостью при работе алгоритмов САПР и имеет приемлемый рабочий диапазон параметров по сравнению со сложными физическими моделями транзисторов [1].

Сами компактные модели подразделяются на два класса: физические и поведенческие. Первые

строятся на основе физического анализа работы транзистора с упрощенным рядом параметров, геометрией, распределением примесей, в одно-, двух- и трехмерном приближениях. Несмотря на введенные упрощения, данный тип моделей сохраняет физический смысл, и именно они сейчас активно используются при проектировании ИС, так как позволяют прогнозировать поведение транзистора при изменении его геометрии и других параметров [2, 3]. Второй тип компактных моделей является поведенческим: они представляют собой “черный ящик”, выходные сигналы которого связаны с входными через некоторые уравнения, которые выводятся не из физических зависимостей, а через подбор параметров для наилучшей аппроксимации поведения реального устройства [4]. Необходимо отметить, что в декабре 1995 г. в США был создан Совет по компактным моделям (ныне — Коалиция, compact model coalition), основная цель которого — стандартизация и решение проблем качества моделей.

Однако, как показала непрерывная миниатюризация, для решения определенных задач все более актуальным становится разработка специализированных аппаратных средств с использованием качественно иной компонентной базы, основанной на новых физических принципах [5]. Одним из примеров подобных устройств являются аппаратные ускорители нейросетевых алгоритмов. Последние представляют собой высокопараллельную систему из элементарных вычислителей (ней-

ронов), которые объединены в связанные между собой слои (с помощью синаптических весовых параметров). Размеры современных нейросетевых алгоритмов достигают миллионов [6, 7] и миллиардов параметров [8], при этом, ограниченность архитектуры фон Неймана, связанная с высокими энергетическими затратами на передачу информации между памятью и арифметико-логическим устройством, не позволяет приблизиться к уровню энергоэффективности, наблюдаемому в биологических прототипах. Таким образом, для обработки подобных алгоритмов за “разумное время” ведутся разработки специализированных вычислительных систем, основанных на многоядерном подходе для параллельной (и что крайне важно — локальной) обработки данных.

Особо перспективным в данной направлении является использование мемристоров — электрических сопротивлений с эффектом памяти резистивных состояний. Их способность изменять и сохранять значение проводимости при определенном внешнем воздействии (эффект резистивного переключения) позволяет рассматривать мемристоров в качестве аналогов биологических синапсов, на основе которых на аппаратном уровне можно моделировать работу многослойных нейронных сетей [9, 10] и существенно увеличить производительность систем их аппаратного ускорения [11]. На этапе разработки интегральной схемы в данной ситуации возникает необходимость создания модели, которая будет как количественно, так и качественно описывать поведение мемристора. При этом было продемонстрировано, что мемристорные структуры возможно получить из различных комбинаций материалов с различными механизмами резистивного переключения и набором параметров (граничными сопротивлениями, напряжениями переключения и т.д.). Описания работы данных устройств и проведение экспериментов впоследствии выполнялось с помощью ряда разработанных моделей [12–15], в основе которых лежат как физические механизмы, так и просто поведенческие представления работы мемристора. Тем не менее для точного описания работы системы с конкретным типом мемристора необходима разработка модели, которая будет точно описывать вольт-амперные характеристики (ВАХ) данного устройства и динамику его переключения.

Цель данного исследования — разработать компактную поведенческую Verilog-A модель нанокompозитного (НК) мемристора $(\text{Co}_{40}\text{Fe}_{40}\text{B}_{20})_x(\text{LiNbO}_3)_{100-x}$ [16, 17] и сравнить результаты работы простого нейросетевого алгоритма с синаптическими связями двух типов: 1) в виде модели мемристора VTEAM [13], настроенной под параметры НК-мемристора; 2) в виде оригинальной модели, разработанной в данном исследовании, которая учитывает механизмы изменения пороговых напряжений при циклических пере-

ключениях, конечного времени хранения граничных резистивных состояний и разброса напряжений переключения.

1. МОДЕЛЬ НАНОКОМПОЗИТНОГО МЕМРИСТОРА

Модель мемристора VTEAM, используемая ранее на предыдущих этапах исследований [18–20], обладает рядом преимуществ, среди которых: возможность гибкой настройки параметров пороговых напряжений переключения, минимального и максимального сопротивлений, параметров, связанных с динамикой переключения мемристора, наличие оконной функции для управления характером переключения вблизи граничных резистивных состояний. В то же время в модели можно отметить недостатки: кроме того, что нет введенных разбросов от цикла к циклу и от устройства к устройству, сама модель представляет собой, по сути, потенциометр, ползунок которого сдвигается при превышении разности потенциалов на некоторых контактах. Это приводит к тому, что сопротивление устройства в стабильном резистивном состоянии никак не зависит от прикладываемого к мемристору напряжения (рис. 1а), в то время как реальные нанокompозитные мемристорные устройства демонстрируют явное наличие диодной характеристики (нелинейная зависимость протекающего тока от прикладываемого напряжения). Для реализации подобного поведения, а также устранения указанных выше недостатков, была проведена модификация модели мемристора VTEAM, описанная на языке Verilog-A.

Нелинейность. По своему определению модель VTEAM определяет протекающий через нее ток в виде закона Ома, при этом проводит непрерывный анализ разности потенциалов на своих контактах и выполняет соответствующую подстройку проводимости. Таким образом, для реализации поведения НК-мемристора зависимость тока от напряжения будет определяться как

$$I = \frac{U}{R} f(U), \quad (1)$$

где функция $f(U)$ описывает некоторую нелинейность. На первоначальном этапе исследования предполагалось, что данная функция будет также зависеть от текущего значения синаптического веса W :

$$W = \frac{G - G_{\min}}{G_{\max} - G_{\min}}, \quad (2)$$

$$f = \begin{cases} U^{0.2+(1-W)/2}, & U \geq 0, \\ -U^{0.2+(1-W)/2}, & U < 0, \end{cases} \quad (3)$$

где G , $G_{\min}$ и $G_{\max}$ — текущие, минимальное и максимальное значения проводимости мемристора. Однако это не принесло желаемого результата:

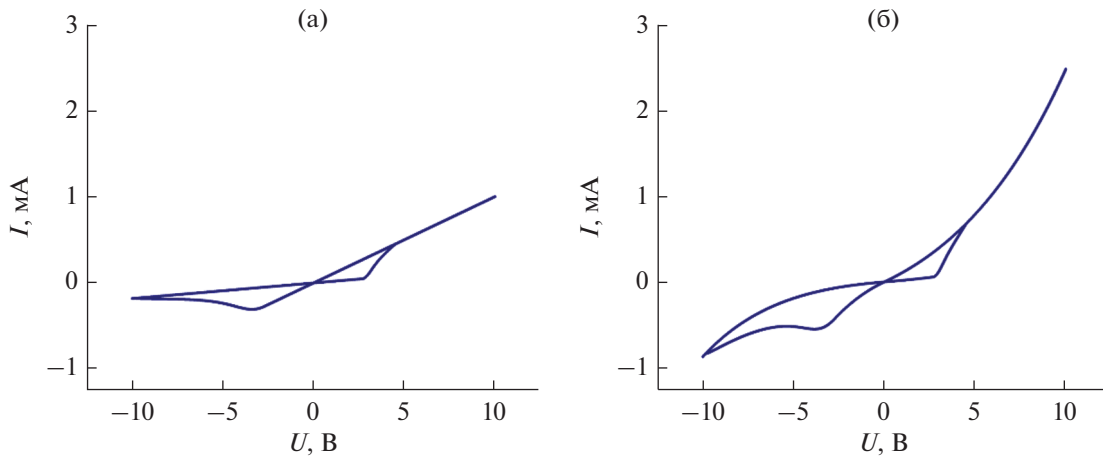


Рис. 1. Вольтамперные характеристики, измеренные со скоростью развертки 2 В/с: настроенной под нанокompозитный мемристор модели VTEAM (а) и предложенной в данной работе модели нанокompозитного мемристора (б).

наклон кривой (по типу параболического) не соответствовал полученным ранее экспериментальным значениям. Эмпирическим путем было найдено решение в виде системы уравнений, в которой разность потенциалов на контактах мемристора находится в показателе степени:

$$f = \begin{cases} A^U, & U \geq 0, \\ B^{-U/2}, & U < 0, \end{cases} \quad (4)$$

где A и B — это подгоночные параметры, подставляя в которые числа 1.12 и 1.18 удалось получить ВАХ (рис. 1б), крайне близкую к экспериментальным данным для лабораторного образца НК-мемристора. При этом необходимо отметить, что задаваемые параметры сопротивлений в граничных состояниях в таком случае теряют смысл, так как это сопротивление будет зависеть от используемого напряжения считывания. Тем не менее наклон кривой ВАХ продолжает регулироваться через данные параметры.

Разброс по напряжению переключения. Реализация данного типа разброса была выполнена на основе вызова генератора случайных чисел (ГСЧ):

— при *разбросе от устройства к устройству* на этапе запуска моделирования происходит вызов функции `$rdist_normal`, в которую передаются три аргумента: начальное значение генератора случайных чисел (SEED), математическое ожидание и дисперсия. Важным моментом на данном этапе является условие корректной работы при реализации массивов мемристорных структур с использованием данной модели: если в качестве SEED использовать числовой параметр модели, задаваемый на старте, то все кроссбар-мемристоры

в матрице получат одинаковый SEED на старте, так как все они работают по одинаковому коду, а значит, и один и тот же результат работы ГСЧ и разброс. Одним из вариантов, как этого можно избежать, является подготовка для каждой мемристорной структуры файла с уникальным именем, связанным с именем мемристора и записью в него уникального числа, которое будет использоваться как SEED. А первым шагом после начала моделирования является выполнение функции `$fopen("path_to_file/Mem_input_data_%I", "r")`, аргументами которой является путь к уникальному файлу и основой его имени с приставкой `%I`, которая берет имя мемристора и дополняет им полное имя файла. Таким образом, на старте все структуры получают уникальный параметр для запуска ГСЧ и получают свой уникальный разброс;

— при *разбросе от цикла к циклу* также необходимо иметь уникальный SEED, чтобы мемристоры не вели себя одинаково на протяжении всего эксперимента. Сам же разброс подобного типа был реализован на основе единократного вызова ГСЧ: когда разность потенциалов на контактах мемристора становится меньше порогового напряжения после завершения переключения структуры, то происходит повторный вызов ГСЧ для обновления переменных с данными по пороговым напряжениям. Полученный результат с циклической ВАХ представлен на рис. 2а.

Время хранения резистивного состояния. Данный механизм в модели НК-мемристора был реализован с использованием следующей пары уравнений, которые непрерывно выполняются в процессе моделирования:

$$W = \begin{cases} \text{endState} + (W - \text{endState}) \exp\left(-\frac{t_2 - t_1}{\text{retTau}}\right), & W \geq \text{endState}, \\ \text{endState} + (W - \text{endState}) \left(1 - \exp\left(-\frac{t_2 - t_1}{\text{retTau}}\right)\right), & W < \text{endState}, \end{cases} \quad (5)$$

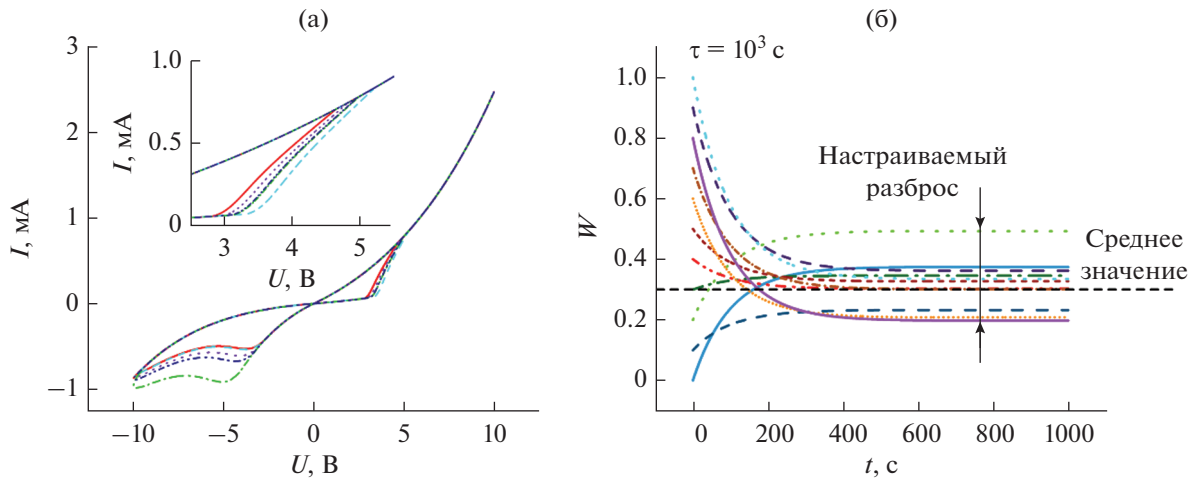


Рис. 2. Циклические ВАХ модели нанокompозитного мемристора со скоростью развертки 2 В/с при включенном разбросе по напряжениям переключения от цикла к циклу (на вставке — переключение из высокоомного состояния в низкоомное состояние) (а); процесс релаксации синаптического веса мемристора из различных начальных состояний к конечному состоянию, выбранному случайным образом с учетом настраиваемого разброса (б).

где переменная `endState` — это конечное резистивное состояние, которое было определено с помощью генератора случайных чисел при первоначальном запуске модели, а переменные $t1$ и $t2$ — это временные метки, разность между которыми равна текущему шагу моделирования. В Verilog-A реализации $t1$ определяется в начальный момент времени как нуль, а $t2$ с помощью функции `$abs-time`. После завершения вычислений значение $t2$ записывается в $t1$, после чего все повторяется на новом временном шаге.

В силу того что для определения конечного состояния использовалась та же функция `$rdist_normal`, данный вариант реализации подразумевает наличие настраиваемого разброса времени хранения резистивных состояний. На рис. 2б продемонстрирован процесс релаксации синаптического веса мемристора из различных начальных состояний с использованием описанного выше метода.

2. ТЕСТИРОВАНИЕ МОДЕЛИ

Была исследована работа простого импульсного нейросетевого алгоритма при использовании модели VTEAM и модели нанокompозитного мемристора с целью поиска отличий и определения актуальности полученных ранее результатов [18–20] в случае перехода на новую поведенческую модель НК-мемристора.

Тестирование проводили на примере импульсной нейросети, основанной на локальных правилах обучения типа STDP (Spike-Timing Dependent Plasticity — пластичность, зависящая от времени прихода импульсов (спайков)), которые являются более перспективными с точки зрения потреб-

ляемой энергии и количества размеченных человеком данных, необходимых для обучения [21–24]. Локальные правила для своей работы требуют только информации об активностях соседних нейронов и значения синаптического веса, который их связывает, поэтому в подобной аппаратной системе отпадает необходимость наличия данных о состоянии всей сети для обновления ее параметров (как это происходит в методе обратного распространения ошибки в формальных нейросетевых алгоритмах). Кроме того, если в качестве синаптического веса выступает мемристорная структура, то для обучения нейросети отпадает необходимость реализации блоков перерасчета весов, так как их настройка происходит путем наложения импульсов от пре- и постсинаптических нейронов (рис. 3а) и зависит от времени прихода этих импульсов (рис. 3б и 3в).

Сравнивая кривые изменения синаптического веса из различных начальных состояний в зависимости от времени прихода спайков для модели VTEAM и модели нанокompозитного мемристора, видим, что они практически совпадают, так как в их основе лежат идентичные уравнения, зависящие от величины превышающего порог напряжения. Однако модель нанокompозитного мемристора, судя по ВАХ, требует для переключения несколько больших токов, что накладывает более строгие критерии для источника сигналов, который будет управлять массивами НК-мемристоров.

Схема используемого в тестировании нейросетевого алгоритма представлена на рис. 4а и состоит из четырех пренейронов и одного постнейрона. Обучение сети происходило с помощью частотно-кодированного метода: логической единице и логическому нулю соответствовали входные сиг-

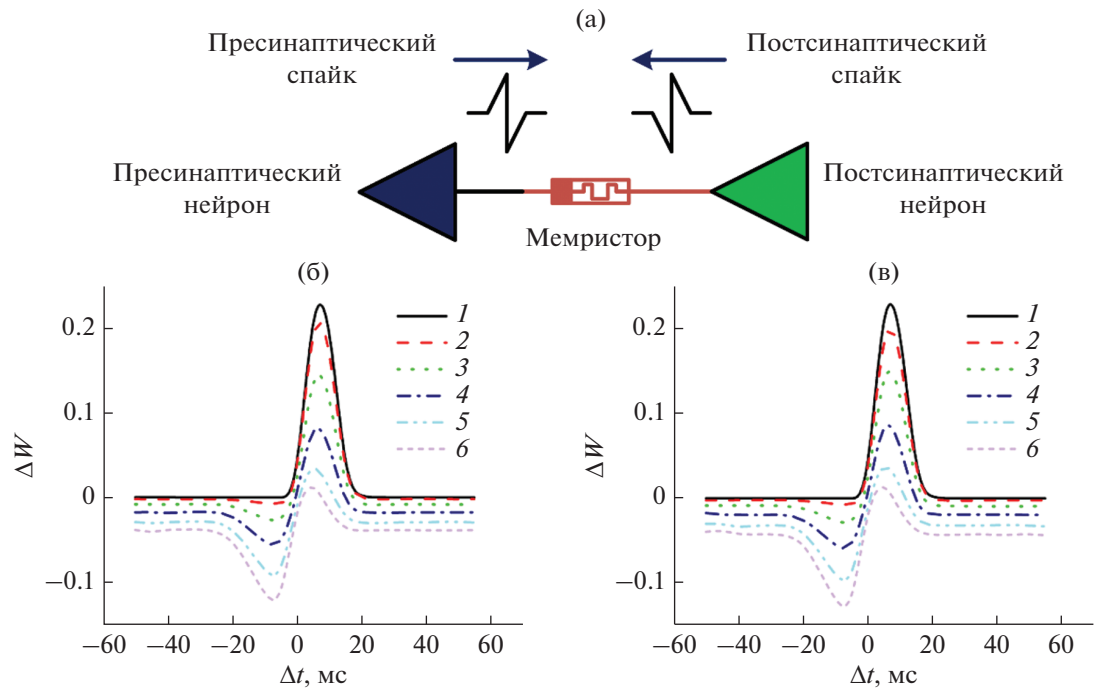


Рис. 3. Пара импульсных нейронов с битреугольным типом импульсов и синаптической связью в виде мемристора (а); семейства кривых STDP для битреугольного типа импульсов, измеренные из различных начальных состояний мемристора для модели VTEAM (б) и модели нанокompозитного мемристора (в) при $W_0 = 0.0$ (1), 0.20 (2), 0.40 (3), 0.60 (4), 0.80 (5), 0.95 (6).

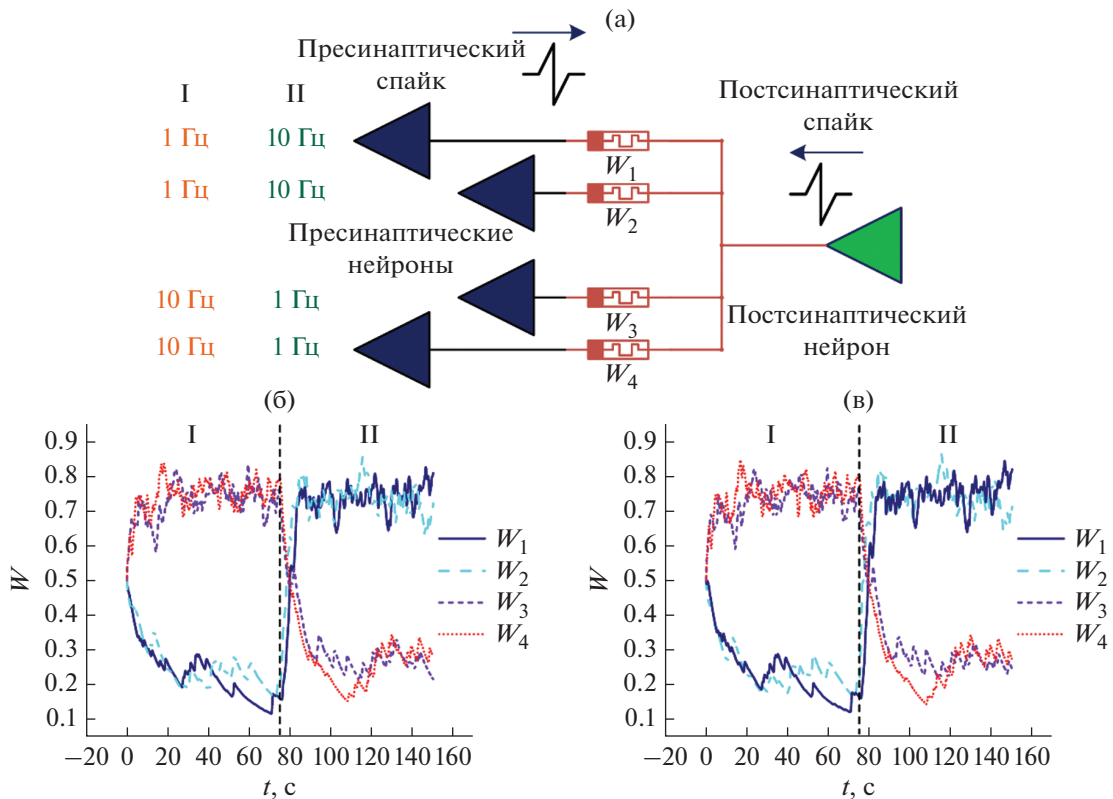


Рис. 4. Схематическое изображение тестового нейросетевого алгоритма из четырех пренейронов и одного постнейрона (а): на этапе I частотно-кодированного обучения на входы W_1 и W_2 подавался сигнал с низкой средней частотой импульсов (1 Гц), распределенных по Пуассону, а на входы W_3 и W_4 — сигнал с высокой средней частотой (10 Гц), на этапе II они менялись. Зависимость синаптического веса от времени для нейросети со связями в виде модели VTEAM (б) и модели НК-мемристора (в).

налы со средней частотой следования импульсов равной 10 и 1 Гц (распределены по Пуассону). Сам процесс состоял из двух этапов: на первом в течение 75 с низкочастотный сигнал подавался на первые два входа, а высокочастотный — на вторые два входа, после чего они менялись для выполнения переобучения связей импульсного нейрона. При этом входные последовательности были одинаковы в обоих случаях. Полученные результаты зависимости динамики синаптических весов от времени представлены на рис. 4б и 4в для модели VTEAM и НК-модели соответственно.

В обоих случаях наблюдалось обучение и переобучение импульсного нейрона, а отличия в динамике крайне незначительны. Однако при переходе к НК-модели процесс обучения не выполнялся сразу: параметр постнейрона, а именно порог на его мембране, был настроен на меньшее количество получаемого заряда, что заставляло его генерировать спайки от каждого входящего импульса, в то время как для обучения по Хеббу активность постнейрона должна быть обусловлена только импульсами от входов нейросети с высокой частотой для создания эффекта накопления заряда именно от них.

Таким образом, качественное изменение ВАХ мемристорной структуры не приводит к значительным изменениям в работе небольших импульсных нейросетевых алгоритмов, что делает актуальным полученные ранее результаты [19] по стабилизации синаптических весов и для реальных структур при соответствующих настройках нейронов. При этом необходимость настройки порогов на нейронах может оказать сильное влияние при исследовании крупномасштабных нейросетей с мемристорными синаптическими связями, так как в этом случае использование VTEAM приведет к значительным отличиям от результатов, полученных на реальных лабораторных образцах.

ЗАКЛЮЧЕНИЕ

Разработана компактная поведенческая модель мемристора, сочетающая в себе возможности введения конечного времени хранения резистивного состояния, а также разбросов по напряжениям от цикла к циклу и от устройства к устройству. Путем моделирования Verilog-A реализации модели продемонстрирована возможность получения вольт-амперных характеристик, наблюдаемых для нанокompозитных мемристорных структур $(\text{Co}_{40}\text{Fe}_{40}\text{B}_{20})_x(\text{LiNbO}_3)_{100-x}$.

Проведено сравнение работы импульсного нейросетевого алгоритма с двумя типами синаптических связей: на основе модели VTEAM и на основе модели НК-мемристора. Показано, что итоговая динамика синаптических весов в обоих

случаях идентична, с той лишь разницей, что для НК-структур необходима дополнительная настройка порогового напряжения на мембране постнейронов из-за нелинейной зависимости тока, протекающего через мемристор, от напряжения в данном случае. Как итог, это сохраняет актуальность полученных ранее результатов [18–20] с помощью модели VTEAM, настроенной под НК-мемристоры, и в то же время позволит в перспективе более точно настраивать взаимодействие между полностью программными и аппаратными реализациями нейросетевых алгоритмов с синаптическими связями на основе нанокompозитных структур.

Авторы заявляют об отсутствии конфликта интересов.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Национального исследовательского центра “Курчатовский институт” (№ 86) в части моделирования динамики резистивного переключения нанокompозитных мемристоров и Российского научного фонда (проект № 22-19-00171) в части экспериментальных исследований их мемристивных свойств.

СПИСОК ЛИТЕРАТУРЫ

1. Gasseling T. // Microw. J. 2012. V. 55. P. 74.
2. Денисенко В.В. // Электроника: Наука, технология, бизнес. 2004. № 5. С. 76.
3. Chauhan Y.S., Venugopalan S., Karim M.A. et al. // 2012 Proc. Europ. Solid-State Circuits Conf. Bordeaux. 12–17 Sept. N.Y.: IEEE, 2012. P. 30. <https://doi.org/10.1109/ESSCIRC.2012.6341249>
4. Łuszczek M., Turzynski M., Swisulski D. // Int. J. Electron. Telecommun. 2020. V. 66. № 4. P. 753.
5. Hennessy J.L., Patterson D.A. Computer Architecture: A Quantitative Approach. 7th ed. Cambridge MA: Morgan Kaufmann Publ., 2019.
6. Simonyan K., Zisserman A. // 3rd Int. Conf. Learn. Represent. ICLR San Diego 7–9 May 2015 – Conf. Track Proc. 2015. P. 1.
7. He K., Zhang X., Ren S., Sun J. // 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR). Las Vegas. 27–30 Jun. N.Y.: IEEE, 2016. P. 770.
8. Brown T.B., Kaplan J., Ryder N. et al. // arXiv: 2005.14165. 2020. <https://doi.org/10.48550/arXiv.2005.14165>
9. Yao P., Wu H., Gao B. et al. // Nature. 2020. V. 577. № 7792. P. 641.
10. Hu M., Graves C.E., Li C. et al. // Adv. Mater. 2018. V. 30. № 9. Article No. 1705914.
11. Wan W., Kubendran R., Schaefer C. et al. // Nature. 2022. V. 608. № 7923. P. 504.
12. Kvatinisky S., Friedman E.G., Kolodny A., Weiser U.C. // IEEE Trans. 2013. V. TCAS-I-60. № 1. P. 211.

13. *Kvatinsky S., Ramadan M., Friedman E.G., Kolodny A.* // IEEE Trans. 2015. V. TCAS-II-62. № 8. P. 786.
14. *Abunahla H., Mohammad B.* Memristor Technology: Synthesis and Modeling for Sensing and Security Applications. Cham: Springer, 2018. P. 93.
15. *Теплов Г.С., Горнев Е.С.* // Микроэлектроника. 2019. Т. 48. № 3. С. 163.
16. *Мацукатова А.Н., Никируй К.Э., Миннеханов А.А. и др.* // РЭ. 2020. Т. 65. № 10. С. 1008.
17. *Никируй К.Э., Ильясов А.И., Емельянов А.В. и др.* // ФТТ. 2020. Т. 62. № 9. С. 1562.
18. *Demin V.A., Surazhevsky I.A., Emelyanov A.V. et al.* // J. Comput. Electron. 2020. V. 19. № 2. P. 565.
19. *Surazhevsky I.A., Demin V.A., Ilyasov A.I. et al.* // Chaos, Solitons and Fractals. 2021. V. 146. Article No. 110890.
20. *Surazhevsky I.A., Nikiruy K.E., Emelyanov A.V. et al.* // Nanoindustry Russ. 2022. V. 15. P. 470.
21. *Davies M., Srinivasa N., Lin T.H. et al.* // IEEE Micro. 2018. V. 38. № 1. P. 82.
22. *Sboev A., Serenko A., Rybka R., Vlasov D.* // Math. Methods Appl. Sci. 2020. V. 43. № 13. P. 7802.
23. *Demin V.A., Nekhaev D.V., Surazhevsky I.A. et al.* // Neural Networks. 2021. V. 134. P. 64.
24. *Demin V., Nekhaev D.* // Frontiers in Neuroinformatics. 2018. V. 12. Article No. 79.