

## ОТЗЫВ

официального оппонента Бутусова Дениса Николаевича  
на диссертационную работу  
Борданова Ильи Алексеевича

«Модели и алгоритмы оценки функциональной корректности искусственных нейронных сетей на базе мемристоров», представленную на соискание ученой степени кандидата технических наук по специальности 2.3.1 Системный анализ, управление и обработка информации, статистика (технические науки)

Диссертационная работа Борданова И.А. посвящена решению важной научно-технической задачи повышения точности оценки функциональной корректности искусственных нейронных сетей, аппаратно реализуемых на базе мемристивных устройств. Создание средств моделирования и проектирования мемристивной техники соответствует ключевому направлению Стратегии научно-технологического развития Российской Федерации - переход к передовым технологиям проектирования и создания высокотехнологичной продукции, основанным на применении интеллектуальных производственных решений, роботизированных и высокопроизводительных вычислительных систем, новых материалов и химических соединений, результатов обработки больших объемов данных, технологий машинного обучения и искусственного интеллекта (п.21а Указа Президента Российской Федерации от 28 февраля 2024 г. № 145). Это обуславливает высокую значимость и *актуальность* работы.

### **Актуальность темы исследования**

В условиях быстрого развития нейроморфных вычислителей и перехода к парадигме «вычислений в памяти», аналоговые пассивные элементы с памятью – мемристоры - рассматриваются как один из наиболее перспективных кандидатов для аппаратной реализации синаптических связей и искусственных нейронов. Их уникальные свойства, а именно возможность энергонезависимое хранение данных в виде сопротивлений, способность к многоуровневому хранению данных и совместимость с КМОП-технологией — позволяют создавать нейроморфные вычислители, выполняющие операции матрично-векторного умножения за один такт непосредственно в массиве памяти. Это дает возможность на порядки повысить энергоэффективность и быстродействие аппаратных реализаций искусственных нейронных сетей (ИНС) по сравнению с традиционными решениями на графических и тензорных процессорах. Однако их широкому внедрению препятствует ряд факторов, среди которых ключевыми являются вариации сопротивлений от цикла к циклу (межцикловая вариабельность) и разница циклов от устройства к устройству, а также нелинейность процесса резистивного переключения и другие. Указанные факторы приводят к тому, что веса синапсов нейронов в аппаратно реализованной сети отличаются от значений, полученных при обучении программной модели. Как следствие, функциональная корректность (ФК) искусственных нейронных сетей на базе мемристоров (ИНСМ) при переходе от программных моделей к физической аппаратной реализации может существенно снижаться, что препятствует применению стандартных подходов к проектированию ИНС для ИНСМ.

Существующие подходы к оценке функциональной корректности ИНСМ либо оперируют моделями вольтамперных характеристик, требующими значительных вычислительных ресурсов при моделировании сетей, включающих большое количество нейронов, а, следовательно, и мемристоров, либо используют упрощенные вероятностные модели разброса весов, не привязанные к реальным параметрам управляющих сигналов. Кроме того, при оценке функциональной корректности ИНСМ с применением таких моделей отсутствует связь с параметрами сигнала задания сопротивления, что усложняет синтез конкретных значений параметров сигналов после анализа функциональной корректности ИНСМ.

Таким образом, тема диссертации, направленная на разработку моделей и алгоритмов, для оценки ФК ИНСМ на основе взаимосвязи между параметрами сигнала задания

сопротивления и получаемыми для конкретных мемристивных устройств сопротивлениями является актуальной задачей.

### **Объем и структура диссертации**

Диссертационная работа состоит из введения, 4 глав, заключения и двух приложений, содержащих акты внедрения и свидетельства о регистрации программ для ЭВМ. Объем основного текста диссертации 120 страниц. Список литературы насчитывает 115 источников.

*В первой главе* диссертации автор проводит обзор литературы и состояния предметной области. Соискатель утверждает, что доминирующая вычислительная архитектура фон Неймана обладает фундаментальными энергетическими ограничениями, тогда как мемристивные кроссбар-массивы, реализующие концепцию «вычислений в памяти», позволяют выполнять матрично-векторное умножение за один такт, обеспечивая существенный выигрыш в скорости и энергопотреблении. Показано, что основным препятствием к внедрению ИНСМ являются вариации сопротивлений мемристоров, что обуславливает необходимость разработки моделей быстрой оценки функциональной корректности на этапе проектирования. Автор подчеркивает, что существующие подходы к оценке ФК либо требуют существенных вычислительных ресурсов (физически мотивированные модели), либо не учитывают связь с параметрами сигнала записи весов (высокоуровневые вероятностные модели), а универсальная обобщенная модель мемристора, пригодная для инженерных задач, еще не создана. Сформулирована гипотеза исследования, заключающаяся в том, что применение теории планирования эксперимента позволяет установить статистическую взаимосвязь между параметрами сигналов задания сопротивления, погрешностями весов и метриками качества ИНСМ, обеспечивая повышение оценки функциональной корректности. Проанализированы аппаратные архитектуры (SRAM, DRAM, ReRAM, STT-MRAM, PCM), классифицированы факторы прямого и косвенного влияния на ФК, рассмотрены методы обучения и архитектурной оптимизации. Выявлен ключевой пробел в процессе проектирования и моделирования ИНСМ: отсутствие единой методологии, связывающей физический уровень (параметры импульсов записи, ВАХ) с информационным (веса синапсов, метрики точности). Обоснована целесообразность применения теории планирования эксперимента (ТПЭ) и статистического подхода «черного ящика».

*Во второй главе* автором описана разработанная модель и алгоритм моделирования зависимости сопротивления мемристивного устройства от параметров сигналов его задания, позволяющие как прогнозировать результирующее сопротивление по известным параметрам сигнала, так и решать обратную задачу синтеза — определять требуемые параметры программирования для достижения заданного сопротивления с известной погрешностью. Разработана модель и алгоритм моделирования зависимости веса синапса нейрона от сопротивления мемристора и схемы формирования веса, обеспечивающие возможность прямого анализа и обратного синтеза, при этом объединение обеих моделей позволяет установить сквозную взаимосвязь между параметрами сигнала и итоговым весом, сокращая промежуточные вычислительные операции. Предложен оригинальный алгоритм оценки функциональной корректности ИНСМ, который, в отличие от существующих подходов, учитывает ограничения максимально допустимых напряжений на входах кроссбар-массива и использует процедуру масштабирования выходов нейронов для повышения точности вычислений. Проведенная апробация разработанных моделей и алгоритмов в среде LTSpice.

*В третьей главе* описывается разработанный программно-аппаратный комплекс (ПАК). Детализированы исследуемые металл-оксидные мемристивные структуры Au/Ta/ZrO<sub>2</sub>(Y)/Pt/Ti. Представлена аппаратная часть (стековая архитектура на МК, 12-разрядные ЦАП/АЦП, коммутационная плата для кроссбаров 32×8 1T1R) и программное обеспечение на Python с графическим интерфейсом, поддержкой COM-порта, параллельным запуском LTSpice и модулями для построения моделей, оценки ФК и аппаратного инференса. Программная часть ПАК реализована в виде кроссплатформенного приложения с графическим интерфейсом, предоставляющего инструментарий для исследования мемристивных устройств, построения моделей зависимости веса от сопротивления, а также сопротивления мемристора от параметров сигнала.

В четвертой главе содержатся результаты практического применения разработанных средств. На реальных мемристорных кроссбарах построены модели вариаций сопротивлений и весов. Проведена верификация на задачах классификации (Iris, Secondary Mushroom, Parkinson) и распознавания изображений (Fashion-MNIST) с использованием различных архитектур ИНС. Продемонстрировано, что предложенный алгоритм обеспечивает расхождение между моделированием и аппаратным экспериментом не более 3%, тогда как традиционные подходы на базе моделей ВАХ дают ошибку до 25%. Показана эффективность экспериментального подбора коэффициента масштабирования  $K$  для компенсации ограничений напряжений.

В *Заключении* автором сформулированы выводы по работе, сделаны предложения по направлению дальнейших исследований.

**Достоверность и обоснованность научных положений и выводов, сформулированных в диссертации** обеспечивается согласованностью результатов моделирования с экспериментальными данными, полученными на реальных кроссбар-массивах мемристорных устройств  $32 \times 8$  1T1R, а также совпадением модельных и экспериментальных данных с доверительной вероятностью 0,95. Высокая степень достоверности полученных данных подтверждается результатами апробации разработанных моделей и алгоритмов для трех различных архитектур искусственных нейронных сетей, где расхождение между прогнозируемой и экспериментально полученной долей правильных исходов не превышало 3 %, что существенно ниже погрешности существующих подходов. Полученные результаты базируются на применении методов системного анализа, теории планирования эксперимента и математической статистики, а также подкреплены выполнением исследований в рамках 6 научно-исследовательских работ. Основные положения диссертации также прошли апробацию на шести международных и всероссийских конференциях, а результаты исследования опубликованы в 14 научных трудах, включая издания, рекомендованные ВАК Минобрнауки РФ, а также журналы, индексируемые в международных наукометрических базах Scopus и Web of Science, приравняемые к изданиям перечня ВАК.

#### **Научная новизна результатов диссертации состоит в следующем:**

1. Автором предложены новая модель и авторский алгоритм моделирования зависимости сопротивления мемристорного устройства от параметров задающих сигналов. Новизна данной модели заключается в том, что подобная зависимость описывает не функциональную взаимосвязь между параметрами сигналов и физическими процессами в мемристоре при прохождении сигнала, а статистическую взаимосвязь между параметрами сигнала задания сопротивления и конечным значением сопротивления, что позволяет рассчитать погрешность задания сопротивления мемристора.

2. Разработана новая модель и алгоритм моделирования зависимости веса синапса нейрона от сопротивления мемристорного устройства и схемы формирования веса, отличающиеся от существующих тем, что вес представляется не аналитической зависимостью между электрическими параметрами цепи, а статистической зависимостью между сопротивлением и весом, и позволяющие рассчитать погрешность задания веса ИНСМ.

3. Разработан оригинальный алгоритм оценки функциональной корректности ИНСМ, отличающийся от существующих тем, что погрешности весов различны для каждого номинального значения веса и определяются из моделей зависимости погрешности веса от сопротивления и параметров сигнала задания сопротивления, а также тем, что в нем учитываются ограничения максимально допустимых рабочих напряжений на входе сети, и позволяющий оценить функциональную корректность ИНСМ с высокой степенью приближения к характеристикам реальных устройств.

**Практическая значимость диссертационной работы** заключается в создании программно-аппаратного комплекса, применимого для решения задачи исследования характеристик мемристорных устройств и построения моделей, а также аппаратной реализации ИНСМ для проверки полученных алгоритмов и моделей. Важным достоинством выступает предложенный автором системный подход, позволивший установить взаимосвязь между параметрами сигналов задания сопротивления мемристоров и итоговыми метриками

качества ИНСМ. Отдельно следует отметить апробацию полученных моделей и алгоритмов, выполненную на реальных мемристивных устройствах для трех различных архитектур нейронных сетей и четырех тестовых наборов данных, при этом расхождение между моделью и экспериментом не превышает 3 %, что подтверждает пригодность разработанного инструментария для решения инженерных задач. Практическая значимость также подтверждается внедрением разработанного ПАК и программного обеспечения в научно-исследовательские процессы МИ ВлГУ, НОЦ ФТНС ННГУ им. Н.И. Лобачевского и производственную деятельность ООО «Поликетон»

**Теоретическая значимость** работы состоит в расширении методологического аппарата системного анализа нейроморфных систем на основе мемристоров. Так, автором предложен подход к формализации связи между физическими процессами записи сопротивлений и информационными метриками ФК, что углубляет понимание и интерпретируемость устойчивости различных архитектур ИНС (сверточных, рекуррентных, полносвязных) к аппаратным вариациям. Результаты, полученные соискателем, создают базу для разработки требований к качеству изготовления МУ и стратегий коррекции ошибок на этапе проектирования.

**По тексту диссертации можно сделать следующие замечания:**

1. В разделе 2.4.1 диссертации представлено описание схемы задания сопротивления мемристивного устройства (рис. 2.5), в которой задействована SPICE-модель «Bialek» со ссылкой на ресурс компании Knowm (при этом условное графическое обозначение мемристора соответствует фирменной нотации Knowm). В тексте указано, что «вариации в модель добавлялись путем изменения значения максимального и минимального сопротивления мемристивного устройства по нормальному закону распределения», однако детали программной реализации данного подхода не раскрываются. В связи с изложенным, следовало бы дать разъяснения относительно методологии преобразования детерминированной исполняемой модели в стохастическую. В частности, учитывая отсутствие на схеме внешних источников шумового воздействия, подключенных к мемристивному элементу, каким образом обеспечивается генерация статистических вариаций параметров в рамках моделирования? Требуется уточнить, реализовано ли введение случайных отклонений на уровне параметризации модели (например, через модификацию внутренних переменных или использование стохастических функций в описании модели), либо применен иной подход, и если да — то каким образом обеспечена воспроизводимость и верификация полученных результатов.
2. В разделе 1.2 в качестве критических факторов, влияющих на функциональную корректность, корректно выделены дрейф сопротивлений и случайный телеграфный шум. Однако в практической реализации модели учитывают только мгновенные вариации при записи, игнорируя временную эволюцию параметров. В мемристивных ИНС, предназначенных для непрерывного инференса в edge-устройствах, отсутствие динамической компоненты в оценке функциональной корректности снижает практическую ценность предложенного метода. Возможно ли интегрировать в модель параметр времени/циклов считывания или это явное ограничение представленного в диссертации подхода?
3. Использование кусочно-линейной и кусочно-кубической интерполяции по экспериментальным данным обосновано, однако не ясно, реализован ли автоматический выбор точек разрыва при сильной нелинейности мемристоров вблизи порогов SET/RESET. Выбор способа интерполяции не сопровождается анализом устойчивости и рисками возникновения явления Рунге. Целесообразно было бы обсудить альтернативные базисы (радиальные базисные функции, узлы Чебышева) или адаптивные сетки при создании таких моделей.
4. В разделах 2.4.1 и 3.3.2 диссертации указано, что ускорение вычислительных процедур достигается за счет запуска нескольких независимых копий программы (фоновых процессов) для параллельного моделирования в LTSpice. Однако в тексте отсутствуют количественные показатели производительности разработанного программного обеспечения: не приведены данные о времени выполнения типовых планов эксперимента, масштабируемости ускорения

при росте числа процессов и характеристиках используемых вычислительных систем и методов моделирования.

5. В работе преимущественно рассматриваются мемристивные устройства на основе диоксида циркония. Не обсуждается, в какой степени предложенные модели и алгоритмы применимы к устройствам с иными физическими механизмами переключения (например, фазовыми или магниторезистивными). Желательно уточнить границы применимости разработанного подхода.
6. После раздела 2.4 Исследование предложенных моделей и алгоритмов моделирования в работе идет раздел 2.7 Выводы по главе. А где разделы 2.5 и 2.6?

### **Общая оценка диссертационной работы.**

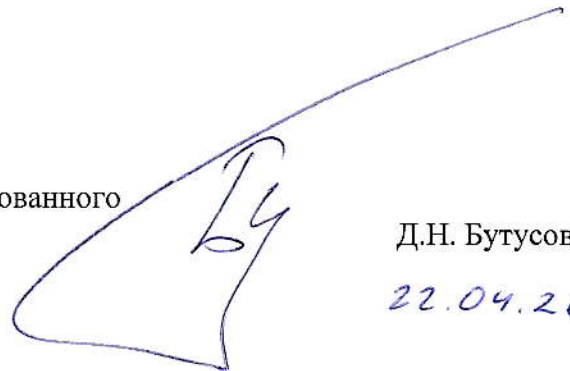
Указанные недостатки не снижают общей положительной оценки диссертации. Диссертация Борданова И. А. является завершенной научно-квалификационной работой, в которой содержится решение научной задачи, имеющей значение для развития методов системного анализа и оценки качества нейроморфных вычислителей на базе мемристивных устройств. Автореферат полностью отражает содержание диссертации, основные положения, выносимые на защиту, научную новизну и практическую значимость. Публикации автора (14 научных трудов, в том числе 4 – в изданиях из перечня ВАК, 4 – в изданиях, индексируемых Scopus/Web of Science, 3 свидетельства о регистрации программ для ЭВМ) в достаточной мере освещают результаты исследования.

### **Заключение**

Диссертационная работа «Модели и алгоритмы оценки функциональной корректности искусственных нейронных сетей на базе мемристоров» соответствует требованиям пп. 9–14 «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства РФ № 842 от 24.09.2013, предъявляемым к кандидатским диссертациям, а ее автор, Борданов Илья Алексеевич, заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика.

Официальный оппонент,  
Доктор физико-математических наук,  
Доцент, заведующий кафедрой Систем автоматизированного  
проектирования СПбГЭТУ «ЛЭТИ»

Подпись Бутусова Д.Н. заверяю.



Д.Н. Бутусов  
22.04.26

Сведения об оппоненте:

Бутусов Денис Николаевич, гражданин Российской Федерации, доктор физико-математических наук (научная специальность 1.2.2. «Математическое моделирование, численные методы и комплексы программ»), доцент, заведующий кафедрой систем автоматизированного проектирования СПбГЭТУ «ЛЭТИ».

Адрес: 197022, субъект Российской Федерации: Санкт-Петербург, город Санкт-Петербург, улица Профессора Попова, дом 5, литера Ф.

Телефон: +7(950) 008-71-90.

E-mail: dnbutusov@etu.ru.



ПОДПИСЬ ЗАВЕРЯЮ  
НАЧАЛЬНИК ОДС  
Т.Л. РУСЯЕВА

