ЯДЕРНАЯ ФИЗИКА И ИНЖИНИРИНГ, 2020, том 11, № 5, с. 266–270

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ЯДЕРНЫХ ТЕХНОЛОГИЯХ

УДК 004 + 519.245 + 530.145

КВАНТОВОМЕХАНИЧЕСКАЯ МОДЕЛЬ ЭЛЕМЕНТОВ НЕЙРОННОЙ СЕТИ

© 2020 г. А. А. Новоселов[a], *, О. В. Павловский[a, b], **

[a]Физический факультет МГУ им. М.В. Ломоносова, Москва, 119991 Россия

[b]НИЦ "Курчатовский Институт" – ИТЭФ, Москва, 117218 Россия

*E-mail: novoselov@goa.bog.msu.ru

**E-mail: ovp@goa.bog.msu.ru

Поступила в редакцию 07.08.2020 г.
После доработки 12.08.2020 г.
Принята к публикации 12.08.2020 г.

Рассматривается модель квантовомеханической системы, организованной по принципу нейронной сети. Роль нейронов играют квантовомеханические частицы эволюционирующие под действием внешнего потенциала с двумя минимумами. В качестве спайка выступает инстантон. Связь между нейронами обеспечивается потенциалом взаимодействия $V_{int}\left(\hat{\phi}_i, \hat{\phi}_j\right)$. Несимметричность потенциала взаимодействия обеспечивает направленность связи. Таким образом полный гамильтониан системы имеет вид $\hat{H} = \sum_i\left(\frac{1}{2}\hat{p}_i^2 + V_0(\hat{\phi}_i)\right) + \sum_{i>j}V_{int}(\hat{\phi}_i, \hat{\phi}_j)$. Система рассматривается при помощи вычисления интегралов по траекториям методом Монте-Карло. Показано, что при определенном выборе параметров в данной системе возможна передача активности в длинных цепочках нейронов.

Ключевые слова: искусственные нейронные сети, интеграл по траекториям, метод Монте-Карло

DOI: 10.1134/S2079562920040144

1. ВВЕДЕНИЕ

1.1. Постановка задачи

Рассмотрим модель квантовомеханической системы, организованной по принципу нейронной сети. Нейронная сеть состоит из узлов (нейронов) и связей между ними (аксонов). Роль нейронов в модели будут играть квантовомеханические частицы $\hat{p}_i$, эволюционирующие под действием потенциала $\hat{H}_i = \frac{1}{2}\hat{p}_i^2 + V_0(\hat{q}_i)$. Роль связей будет выполнять потенциал взаимодействия: $V_{\text{int}} = V_{\text{int}}\left(\hat{q}_i, \hat{q}_j\right)$. Таким образом полный гамильтониан системы будет иметь вид

$$\hat{H} = \sum_i\left(\frac{1}{2}\hat{p}_i^2 + V_0(\hat{q}_i)\right) + \sum_{i>j}V_{int}(\hat{q}_i, \hat{q}_j). \tag{1}$$

Будем выбирать собственный потенциал нейрона так, чтобы он выполнял свою основную функцию – генерировать спайки. Потенциал взаимодействия выберем таким, чтобы он вызывал инстантон в некотором узле, если инстантон произошел в соседнем, но не наоборот – для этого он должен быть несимметричным.

Так как в общем случае нам потребуется иметь дело с достаточно сложными квантовомеханическими системами, то для описания их свойств мы будем использовать хорошо известный формализм Монте-Карло интегрирования континуальных интегралов. В евклидовом времени статистическая сумма системы имеет вид

$$Z = \int\prod_i \mathcal{D}q_i(\tau)\exp\left(-\frac{S(\phi_i(\tau))}{\hbar}\right), \quad q_i(0) = q_i(B), \tag{2}$$

где $q_i(\tau)$ – евклидова траектория i-ой частицы, $\beta \in [0, B]$ – евклидово время, а $S(q_i)$ – классическое действие:

$$S = \int_0^B dt\left[\sum_i\left(\frac{1}{2}\hat{p}_i^2 + V_0(\hat{q}_i)\right) + \sum_{i>j}V_{\text{int}}(\hat{q}_i, \hat{q}_j)\right]. \tag{3}$$

Наблюдаемые в данном формализме вычисляются как

$$\mathbb{O}(q_1, \ldots, q_i) = \\ = \frac{1}{Z}\int\prod_i\mathcal{D}q_i(\tau)\mathbb{O}(q_1, \ldots, q_i)\exp\left(-\frac{S(q_i)}{\hbar}\right). \tag{4}$$

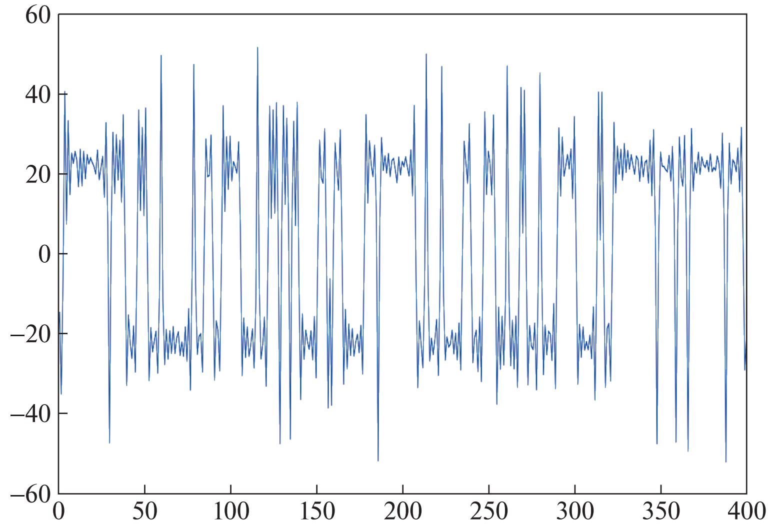

Рис. 1. Траектория свободного нейрона при малых Λ. Нейрон испытывает самопроизвольные спайки.

1.2. Методы решения

Работа нейронной сети основана на распространении активности от входных узлов (сенсоров) к выходным узлам. Сеть может состоять из многих узлов, поэтому для исследования такой сложной квантовой системы естественно применить метод Монте-Карло [1]. С помощью марковского процесса будут генерироваться траектории всех узлов сети q_i с статистическим весом, пропорциональным $\exp\left(-\frac{S(q_i)}{\hbar}\right)$. Для генерации траекторий в данной работе использовался многоуровневый алгоритм Метрополиса. Его использование обеспечивает подавление автокорреляций и позволяет повысить эффективность вычислений.

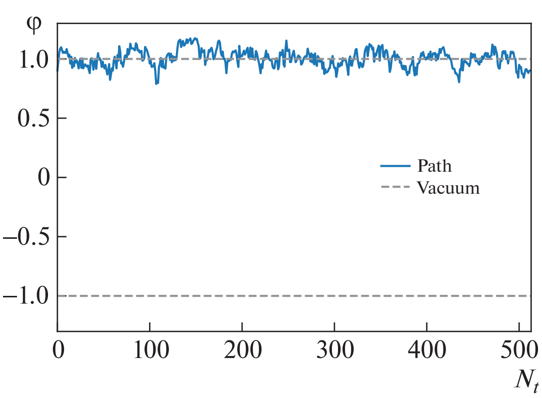

Рис. 2. Траектория свободного нейрона при больших Λ. Нейрон не испытывает самопроизвольных спайков.

2. КВАНТОВЫЕ НЕЙРОНЫ И ИХ СВЯЗИ

2.1. Один нейрон

Начнем рассмотрение нашей модели со случая одного нейрона. В качестве искусственного нейрона будет выступать частица, находящаяся в двухямном W-потенциале: $V_0(\varphi) = \frac{\lambda}{4}\left(\varphi^2 - 1\right)^2$. Выпишем лагранжиан такой системы:

$$\mathcal{L}_0 = \frac{1}{2}\dot{\varphi}^2 + \frac{\Lambda}{4}\left(\varphi^2 - 1\right)^2. \quad (5)$$

Минимумы потенциальной энергии находятся в точках $\varphi = \pm 1$. Типичное квантовомеханическое поведение такой частицы – флуктуации в окрестности одного из вакуумов и быстрая смена вакуума – инстантон. Инстантоны сопровождаются пиком плотности действия, похожим на потенциал спайка биологического нейрона. Поэтому спайком такого нейрона будем называть инстантон. Также запишем величину классического действия для инстантона: [2]

$$S_{\text{cl}} = \frac{2\sqrt{2\Lambda}}{3}. \quad (6)$$

Эта величина понадобится нам для того, чтобы подобрать подходящее значение параметра Λ. Во-первых, мы хотим, чтобы флуктуации нашей частицы вокруг минимумов энергии были не очень большими. Во-вторых, спонтанные переходы из одного состояния в другое не должны случаться слишком часто. В нашем случае оптимальным оказалось $\Lambda = 5000$.

Параметры алгоритма Метрополиса были выбраны следующим образом: время (обратная температура) $T = 0.7$, количество временных слоев сетки $N_t = 512$. Для подавления автокорреляций

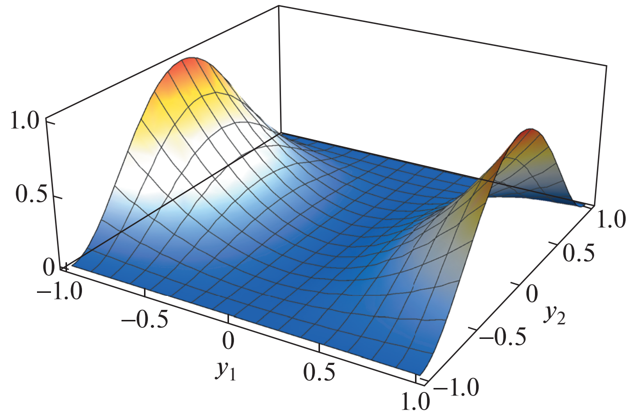

Рис. 3. Потенциал взаимодействия для передачи спайка.

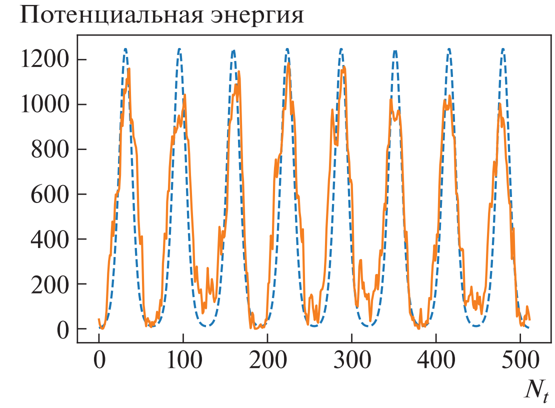

Рис. 4. Потенциальная энергия возбуждаемого нейрона (сплошная линия) коррелирует с потенциальной энергией входного нейрона (пунктирная линия).

использовалась длина термализации в $2 \cdot 10^6$ итераций. Для инициализации траектории использовалась "пила" из нулей и единиц: $\varphi_i^{\text{init}} = i \bmod 2$, где i — номер временного слоя узла сетки. Характерная термализованная траектория представлена на рис. 1 и рис. 2.

2.2. Два нейрона

Перейдем теперь к случаю двух взаимодействующих нейронов. Мы хотим, чтобы спайк в одном из них вызывал спайк в другом, но спайк во втором по возможности никак не влиял на первый. Для этого лагранжиан взаимодействия таких нейронов должен быть несимметричным, кроме того, он должен содержать константу связи, с помощью которой можно регулировать действие одного нейрона на другой. Выберем его в виде:

$$\mathcal{L}_{\text{int}} = \varepsilon_{\text{exc}} \varphi_2^2 \left(\varphi_1^2 - 1\right)^2. \tag{7}$$

Если φ_1 находится в вакууме, то на φ_2 нет никакого воздействия. Однако, если φ_1 испытывает спайк, то и φ_2 тоже имеет тенденцию к спайку. Таким образом активность передается от узла к узлу.

График потенциала такого взаимодействия изображен на рис. 3. По оси y_1 отложен возбуждающий нейрон, а по оси y_2 — возбуждаемый. Если оба нейрона находятся в вакуумах (вакуум соответствует нахождению в одном из углов), то только переход первого нейрона производит воздействие на другой нейрон.

Для представления входной информации мы будем использовать входные нейроны. Каждый входной нейрон может быть пассивным (вообще не влияет на другие, поэтому может быть отброшен) или активным. Активные входные нейроны имеют фиксированную траекторию (в отличие от моделируемых нейронов, траектория которых эволюционирует во время моделирования), который состоит из классических решений—инстантонов. Потенциальная энергия такого входного нейрона изображена на рис. 4 (пунктирная линия). Каждый пик потенциальной энергии соответствует инстантону. Сплошная линия представляет потенциальную энергию одиночного нейрона, возбуждаемого входным нейроном.

Введем величину активности моделируемого нейрона как отношение интегральной потенциальной энергии этого нейрона к потенциальной энергии входного нейрона (например, активность нейрона, который никогда не покидает вакуумного состояния, будет равна 0, а активность нейрона, траектория которого повторяет траекторию входного нейрона, будет равна 1). Чтобы исследовать различные схемы, мы будем изучать графики активности некоторых нейронов в зависимости от константы связи ε_{exc}.

Для изучения различных конфигураций нейронов введем модулирующий фактор k. После того, как мы выберем соответствующий параметр $\hat{\varepsilon}$ для каждой связи, мы умножим каждый из них на этот коэффициент, чтобы получить новые константы связей $\varepsilon = k \cdot \hat{\varepsilon}$ и затем отобразим активность интересующего нейрона как функцию одного параметра k.

Было обнаружено, что ε_{exc} может принимать значения в диапазоне от 3000 до 8000 (рис. 5). В случае слишком малых ε_{exc} нейроны почти не взаимодействуют, а если значение ε_{exc} слишком вели-

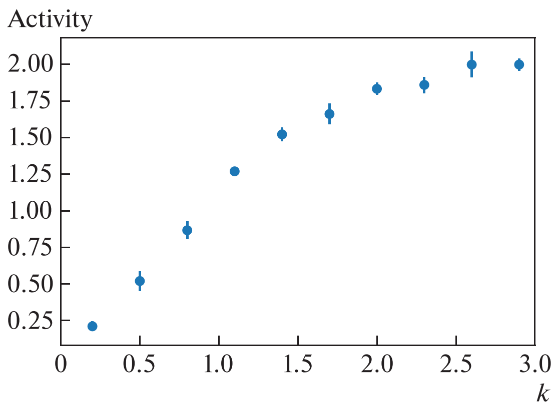

Рис. 5. Активность выходного нейрона возрастает как функция константы связи.

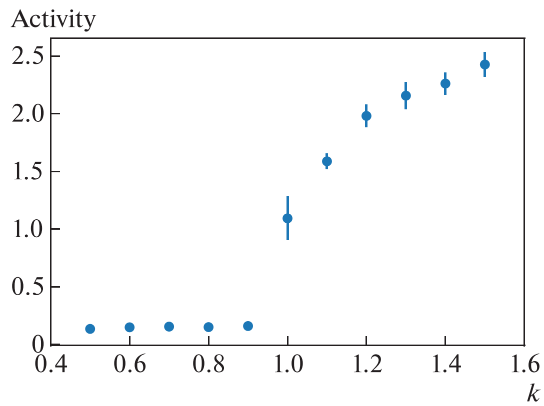

Рис. 6. Активность выходного нейрона (3-го в цепочке) как функция константы связи. $\varepsilon_1 = k \cdot 1.5 \cdot 10^4$, $\varepsilon_2 = k \cdot 1.0 \cdot 10^4$, $\varepsilon_3 = k \cdot 0.5 \cdot 10^4$.

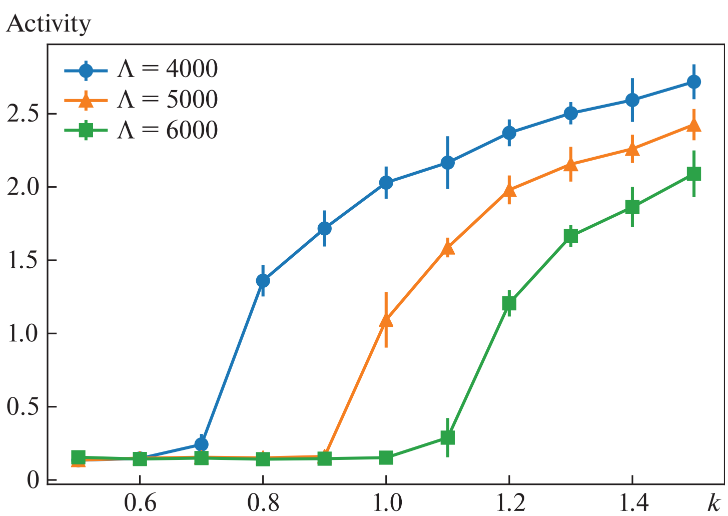

Рис. 7. Активность выходного нейрона (в цепи из трех) как функция Λ.

ко, их собственные потенциалы становятся незначительными по сравнению с взаимодействием, что приводит к нежелательной задержке нейронов в состоянии $\varphi = 0$. График для $\varepsilon_{exc} = 6000$ представлен на рис. 4. В симуляции, представленной на рис. 4, активность выходного нейрона оказалась равна 0.92.

Также можно передать импульс по цепочке из нескольких нейронов. Рассмотрим линию из трех последовательно соединенных связями нейронов (рис. 6). В этом случае выберем $\varepsilon_1 = k \cdot 1.5 \cdot 10^4$, $\varepsilon_2 = k \cdot 1.0 \cdot 10^4$, $\varepsilon_3 = k \cdot 0.5 \cdot 10^4$. Как видно из рис. 6, при малых значениях константы связи ε, спайки не проходят через цепочку нейронов, но если ε достигает некоторого критического значения, цепочка становится проводящей для спайков. Этот эффект позволяет нам контролировать проводимость нейронной сети путем небольшого изменения константы связи ε. Таким образом, контролируя костанту связи, мы можем реализовать сложные логические соединения в нашей нейронной сети.

Зависимость активности нейронов от параметра Λ показана на рис. 7. Из этого рисунка видно, что критическое значение константы связи ε зависит от Λ. Очевидно, что такая зависимость критического значения ε связана с увеличением действия инстантона.

3. ЗАКЛЮЧЕНИЕ

Была предложена модель нейрона, основанная на поведении частицы в двухуровневой системе. Для моделирования такой квантовомеханической системы использовалось вычисление интегралов по траекториям методом Монте-Карло. Был предложен конкретный вид потенциала частицы, кроме того были предложены потенциала взаимодействия частиц, позволяющий передавать возбуждения. Отметим также и недостатки предложенной модели. Во-первых, так как выходная функция сети не может быть аналитически продифференцирована по величинам связей внутри сети, мы не можем применять алгоритм обратного распространения ошибки для обучения сети. Поэтому обучение сети представляется возможным лишь при помощи куда более медленных генетических алгоритмов. Во-вторых, стоит вообще отметить общую проблему скорости вычислений. Моделирование более-менее сложных сетей потребует привлечения очень больших вычислительных мощностей.

БЛАГОДАРНОСТИ

Работа выполнена с использованием ресурсов суперкомпьютерного комплекса МГУ имени М.В. Ломоносова [5] и при финансовой поддержке Российского научного фонда (грант № 16-12-10059-П).

СПИСОК ЛИТЕРАТУРЫ/REFERENCES

1. *Ceperley D.M.* // Rev. Mod. Phys. 1995. V. 67. P. 279.
2. *Polyakov A.M.* Gauge Fields and Strings. 1987. London: Harwood Academic Publishers.
3. *Voevodin Vl.V., Antonov A.S., Nikitenko D.A., Shvets P.A., Sobolev S.I., Sidorov I.Yu., Stefanov K.S., Voevodin Vad.V., Zhumatiy S.A.* // Supercomput. Front. Innovat. 2019. V. 6. No. P. 4–11. https://doi.org/10.14529/jsfi190201.

Quantum Mechanical Model of Neuron Network Elements

A. A. Novoselov[1, *] and O. V. Pavlovsky[1, 2, **]

[1] *Faculty of Physics, Moscow State University, Moscow, 119991 Russia*

[2] *Alikhanov Institute for Theoretical and Experimental Physics, National Research Center "Kurchatov Institute", Moscow, 117218 Russia*

**e-mail: novoselov@goa.bog.msu.ru*

***e-mail: ovp@goa.bog.msu.ru*

Received August 7, 2020; revised August 12, 2020; accepted August 12, 2020

Abstract—A model of a quantum-mechanical system organized by the principle of a neural network is considered. The role of neurons is played by quantum particles evolving under the action of an external double-well potential. The instanton acts as a spike. The connection between neurons is provided by the interaction potential $V_{\text{int}}\left(\hat{q}_i, \hat{q}_j\right)$. The connection is directional due to asymmetry of the interaction potential. Thus, the complete Hamiltonian of the system is $\hat{H} = \sum_i \left(\frac{1}{2}\hat{p}_i^2 + V_0(\hat{q}_i)\right) + \sum_{i>j} V_{\text{int}}(\hat{q}_i, \hat{q}_j)$. The system is investigated by PIMC metod. It is shown that with a certain choice of parameters in this system, the transfer of activity in long chains of neurons is possible.

Keywords: artificial neuron network, path integral, PIMC