



CONSERVATION LAWS IN A NEURAL NETWORK APPROACH TO NUMERICAL SOLVING OF THE NONLINEAR SCHRÖDINGER EQUATION

Yu. Gurieva, E. Vasiliev, L. Smirnov

Lobachevsky State University,
603022, Nizhny Novgorod, Russia

DOI: 10.24412/2073-0667-2023-2-5-20

EDN: LFBNWY

We consider a possible modification of a neural network approach to numerical solving of nonlinear partial differential equations (PDE), describing physical systems with integrals of motion. In this approach, we approximate solutions of the equations by deep neural networks, using physics-informed method.

Physics-informed neural network (PINN) approach proposes nonlinear function approximators that integrate the observational data, initial and boundary conditions and description of physical system in form of PDE by embedding the corresponding residuals into the loss function of a neural network. Therefore, the problem of solving nonlinear differential equations turns into the problem of minimizing the squared residuals over domain which is achieved by automatic differentiation and stochastic gradient descent.

The proposed modification of this method means consideration and implementation of corresponding conservation laws for training of the neural networks, and is expected to improve the physical properties of the trained nonlinear regression models. The purpose of this work is to modify a neural network using the conservation law constraint, such that the predicted solution will satisfy the continuity equation better and faster as well as speed up the rate of convergence and provide better accuracy. Improvement of the conservative properties of the approximation is provided by the specific loss function regularization: addition of the conserved quantities' residuals to the loss function to train the neural network.

To test this method, we considered one-dimensional nonlinear Schrödinger equation and its conservation laws in integral form. Number of quants and energy were used as conserved physical quantities. In our experiments, their values were calculated in several equidistant time moments and compared with reference to find the corresponding residuals and implement the conservation constraint in the loss function. Therefore, the average residuals of number of quants and energy for the prediction are considered as quality metrics in the problem, as well as pointwise difference from the predicted and reference solution (validation error). Reference functions for validation datasets are derived from the analytical expressions for the exact solutions.

This modified neural network approach is applied to the different classes of analytic solutions of the nonlinear Schrödinger equation: one soliton, interaction of two solitons (in breather form), first-order rogue wave. For each solution, we apply three forms of the conservative regularization: quants' number constraint, energy constraint and the sum of them. The training curves and predictions are compared with the solution obtained with the initial loss function (baseline).

It is shown that introduction of the additional conservative constraints to loss function reduces the conserved quantities' residuals for training and prediction in all cases. For the simplest one-soliton

solution, the regularizations improve not only conservation quality metrics, but also pointwise difference with the reference in the same training time. The best result was obtained by the combination of constraints: validation error is reduced by more than three times. However, for more complex solution forms, such as two solitons and rogue wave, the results are not as good. The conservative constraints significantly change the form of loss function, so the training curves start to plateau, and the training process becomes more unstable. For the most complex two soliton interaction, it requires about two times more optimization steps to converge. The validation error is improved only for the energy constraint for both cases: for two-soliton solution, validation error is reduced by 13 %; for rogue wave, it is reduced by 67 %. Therefore, the effect of conservative modification of the deep learning approach for nonlinear partial differential equations is individual for different systems and conserved quantities. Generalization ability of such neural networks should be further investigated and tested for different problems.

Key words: deep learning, neural networks, nonlinear Schrödinger equation, conservation laws, solitons.

References

1. Moxley F., Chuss D., Dai W. A generalized finite-difference time domain scheme for solving nonlinear Schrödinger equations // Computer Physics Communications. 2013. N 184 (8). P. 1834–1841. [El. Res.]: <https://doi.org/10.1016/j.cpc.2013.03.006>.
2. Vasiliev E. P., Bolotov D. I., Bolotov M. I., Smirnov L. A. Neyrosetevoy podhod k resheniyu zadachi samovozdejstviya volnovykh polej v nelinejnyh sredah // Problemy Informatiki. 2022. N 1. P. 5–16. [El. Res.]: <https://doi.org/10.24412/2073-0667-2022-1-5-16>.
3. Karniadakis, G. E., Kevrekidis, I. G., Lu, L. et al. Physics-informed machine learning // Nature Reviews Physics. 2021. N 3. P. 422–440. [El. Res.]: <https://doi.org/10.1038/s42254-021-00314-5>.
4. Raissi M., Perdikaris P., Karniadakis G. E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations // Journal of Computational Physics. 2019. N 378. P. 686–707. [El. Res.]: <https://doi.org/10.1016/j.jcp.2018.10.045>.
5. Pu J. C., Li J., Chen Y. Soliton, breather, and rogue wave solutions for solving the nonlinear Schrödinger equation using a deep learning method with physical constraints // Chinese Physics B, 2021. N 30 (6), 060202. [El. Res.]: <https://dx.doi.org/10.1088/1674-1056/abd7e3>.
6. Wu G. Z., Fang Y., Wang Y. Y., Wu G. C., Dai C. Q. Predicting the dynamic process and model parameters of the vector optical solitons in birefringent fibers via the modified PINN // Chaos, Solitons and Fractals. 2021. N 152, 111393. [El. Res.]: <https://doi.org/10.1016/j.chaos.2021.111393>.
7. Wang L., Yan. Z. Data-driven rogue waves and parameter discovery in the defocusing nonlinear Schrödinger equation with a potential using the PINN deep learning // Physics Letters A. 2021. V. 404. P. 127408. [El. Res.]: <https://doi.org/10.1016/j.physleta.2021.127408>.
8. Jagtap, A., Kharazmi, E., Karniadakis. G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems // Computer Methods in Applied Mechanics and Engineering. 2020. N 365, 113028. [El. Res.]: <https://doi.org/10.1016/j.cma.2020.113028>.
9. Wu G. Z., Fang Y., Kudryashov N., Wang Y. Y., Dai C. Q. Prediction of optical solitons using an improved physics-informed neural network method with the conservation law constraint // Chaos, Solitons and Fractals, 2022. 159(C), 112143. [El. Res.]: <https://doi.org/10.1016/j.chaos.2022.112143>.
10. Lin S., Chen Y. A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions // Journal of Computational Physics, 2022. 457(C), 111053. [El. Res.]: <https://doi.org/10.1016/j.jcp.2022.111053>.

11. [El. Res.]: https://docs.nvidia.com/deeplearning/modulus/user_guide/theory/recommended_practices.html.
12. Zakharov V., Manakov S. On the complete integrability of a nonlinear Schrödinger equation // Theoretical and Mathematical Physics. 1974. N 19 (3). P. 551–559. [El. Res.]: <https://doi.org/10.1007/BF01035568>.
13. [El. Res.]: https://docs.nvidia.com/deeplearning/modulus/user_guide/foundational/scalar_transport.html.
14. Stein M. Large sample properties of simulations using Latin hypercube sampling, Technometrics. 1987. N 29. P. 143–151. [El. Res.]: <https://www.jstor.org/stable/1269769>.
15. Schraudolph N.N., Yu J., Gunter S. A Stochastic Quasi-Newton Method for Online Convex Optimization // International Conference on Artificial Intelligence and Statistics, 2007. P. 436–443. [El. Res.]: <http://proceedings.mlr.press/v2/schraudolph07a/schraudolph07a.pdf>.



УЧЕТ ЗАКОНОВ СОХРАНЕНИЯ ПРИ НЕЙРОСЕТЕВОМ ПОДХОДЕ К ЧИСЛЕННОМУ РЕШЕНИЮ НЕЛИНЕЙНОГО УРАВНЕНИЯ ШРЕДИНГЕРА

Ю. В. Гурьева, Е. П. Васильев, Л. А. Смирнов

Нижегородский государственный университет им. Н. И. Лобачевского,
603022, Нижний Новгород, Россия

УДК 517.957

DOI: 10.24412/2073-0667-2023-2-5-20

EDN: LFBNWY

В работе рассматривается один из возможных вариантов модификации нейросетевого подхода к численному решению нелинейных уравнений в частных производных, у которых благодаря физическим свойствам описываемых явлений имеются интегралы движения. Представленный метод подразумевает учет и непосредственное использование соответствующих законов сохранения при построении и обучении нейронных сетей, аппроксимирующих решения такого класса задач, что позволяет улучшить характеристики и качество получаемых нелинейных регрессионных моделей. Более точное выполнение консервативных свойств физических систем для аппроксиматора обеспечивается регуляризацией функции потерь: добавлением невязки сохраняющейся величины нейросетевого решения. Данная концепция рассмотрена и апробирована на примере нелинейного уравнения Шредингера и двух его интегралов движения, отвечающих законам сохранения числа квантов и энергии. Для вычисления невязки этих сохраняющихся величин и реализации консервативной регуляризации функции потерь был использован метод плоскостей непрерывности (вычисление величин в фиксированные моменты времени). Полученные результаты показывают улучшение консервативных свойств, а также в некоторых случаях точности нейросетевого решения по сравнению с регрессионной моделью, построенной с помощью глубокого обучения без учета предложенной в работе модификации.

Ключевые слова: глубокое обучение, нейронные сети, нелинейное уравнение Шредингера, законы сохранения, солитоны.

Введение. Нелинейное уравнение Шредингера (НУШ) представляет собой одно из ключевых уравнений, описывающее волновые процессы в различных областях физики, включая распространение электромагнитных волн в нелинейных оптических волокнах и волноводах, динамику возбуждений конечной амплитуды в плазме и бозе-эйнштейновском конденсате и др. Для моделирования распространения оптического импульса в нелинейном волокне на основе НУШ был разработан целый ряд численных схем. Одним из наиболее популярных и эффективных алгоритмов является псевдоспектральный метод расщепления [1]. Тем не менее, данная задача остается достаточно сложной с точки зрения

Работа проведена при поддержке Проекта № 0729-2021-013, в рамках Государственного задания на выполнение научно-исследовательских работ лабораториями, прошедших конкурсный отбор в ходе реализации национальной программы «Наука и университеты».

вычислений, особенно когда в системе появляются все более мелкие характерные масштабы, например, вследствие самофокусировки волнового поля.

Одно из основных ограничений, возникающих при моделировании динамики нелинейных систем, состоит в том, что в какой-то момент размеры неоднородностей волнового поля могут стать сопоставимы с шагом дискретизации пространственной координаты, который, в свою очередь, ограничивает шаг по времени в конечно-разностных схемах [2]. Поэтому во избежание численной неустойчивости необходимый объем обрабатываемых данных и число итераций по эволюционной переменной сильно возрастают. Таким образом, существует потребность в альтернативных методах расчета нелинейных волновых процессов в различных областях физики. Нейросетевой подход позволяет снять ограничения классических алгоритмов решения дифференциальных уравнений в частных производных, которые подразумевают дискретизацию пространственно-временной области.

Методы машинного обучения представляются перспективной альтернативой прямому численному моделированию благодаря своей эффективности для многомерных, обратных и некорректно поставленных задач [3]. Методы глубокого обучения, в частности, позволяют выделять закономерности из большого объема неоднородных данных и делать предсказания динамики волновых полей на основе физических законов (дифференциальных уравнений в частных производных) и ограничений (в том числе законов сохранения: массы, импульса, энергии и др.).

Метод глубокого обучения Physics-Informed Neural Network (PINN) был разработан и применен для решения прямых и обратных (восстановление параметров уравнений по данным) задач нелинейной динамики [4]. Построенная при этом нейронная сеть агрегирует наблюдаемые данные, начальные и граничные условия, а также описание физической системы (уравнения), включая соответствующие остатки в функцию потерь. В итоге задача численной аппроксимации решения нелинейного дифференциального уравнения в частных производных сводится к минимизации суммы квадратов невязок в точках пространственно-временной области [3, 4]. Отметим, что в настоящее время применимость и эффективность PINN активно изучается в рамках модельных уравнений, к которым, в частности, относится НУШ. Например, рассматривались возможность предсказания разных сценариев поведения комплексных волновых полей [5], в том числе различные варианты взаимодействия светлых и темных солитонов [6], а также возникновение волн-убийц [7].

Было предложено множество улучшений оригинальной версии PINN [4], основанных на законах сохранения, с целью улучшить качество и характеристики аппроксимации решений нелинейных дифференциальных уравнений в частных производных. В частности, была предложена консервативная модификация метода PINN с разделением области (геометрии) задачи на части [8]. Дополнительная регуляризация функции потерь использует закон сохранения потока на соприкасающихся границах областей и улучшает точность решения. В [9] наряду с основными динамическими уравнениями рассматривалось включение законов сохранения в дифференциальной форме в функцию потерь. В представленной нами работе выбран способ учета законов сохранения в интегральной форме [10, 11].

Целью данного исследования является проверка гипотетической возможности улучшения фундаментальных физических свойств нейросетевого решения НУШ с помощью консервативной модификации функции потерь, состоящей в специфической регуляризации с использованием интегралов движений. Соответствующий анализ проведен для законов сохранения числа квантов и энергии, а нейросетевые решения НУШ, полученные методом PINN, сравнивались с известными аналитическими решениями в виде одного солитона, су-

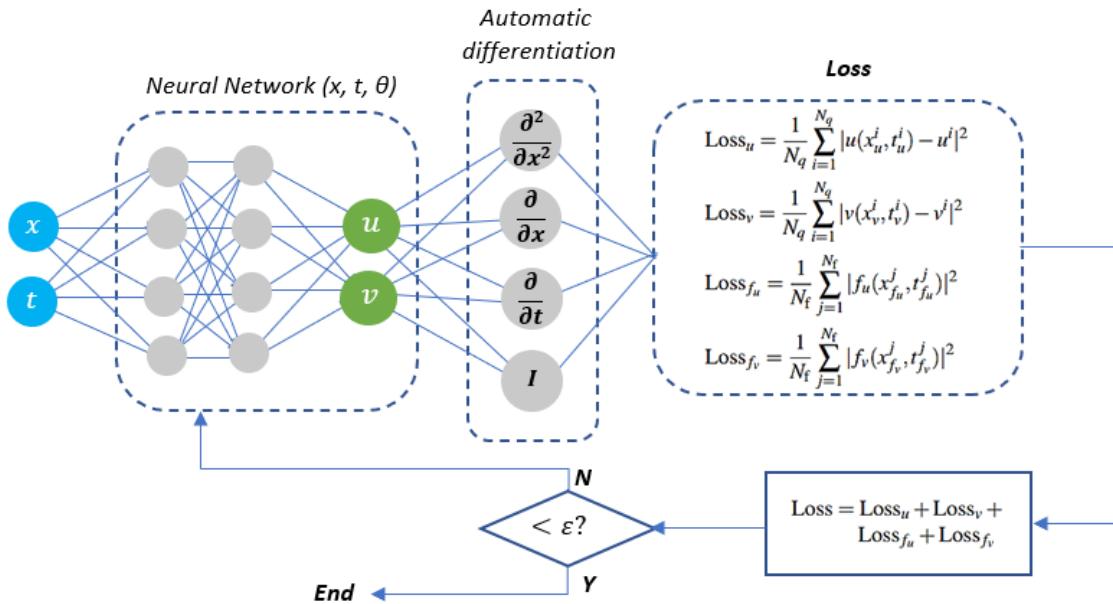


Рис. 1. Общая схема нейросетевого подхода при построении аппроксимации решения нелинейного динамического уравнения в частных производных

перпозиции двух солитонов в форме бризера, а также волны-убийцы первого порядка [12]. Показано, что для модифицированной на основе интегралов движения нелинейной регрессионной модели качество выполнения законов сохранения улучшается, благодаря чему в нескольких случаях точность предсказания (по метрике среднеквадратичного отклонения от точного решения) возрастает при увеличении длительности обучения.

1. Нейросетевой подход к численному решению нелинейного уравнения Шредингера.

1.1. Глубокое обучение для построения аппроксимации решений нелинейных динамических уравнений. В данной работе рассматриваются полносвязные нейронные сети, построенные с помощью методов глубокого обучения и позволяющие предсказывать в указанной области Ω по заданным начальным условиям $\psi_0(x)$ решение НУШ с граничными условиями Дирихле:

$$\begin{cases} i\frac{\partial\psi}{\partial t} + \alpha\frac{\partial^2\psi}{\partial x^2} + \beta|\psi|^2\psi = 0, & x \in \Omega, \quad t \in [t_{min}, t_{max}] \\ \psi(x, t_{min}) = \psi_0(x), \\ \psi(x, t) = \psi_1(x, t), & x \in \partial\Omega, \end{cases} \quad (1)$$

где α и β — действительные постоянные, которые определяются свойствами описываемой с помощью (1) физической системы, а $\psi(x, t)$ — комплексная волновая функция, характеризующая распределение поля. Соответственно, невязка уравнения определяется выражением

$$f := i\frac{\partial\psi}{\partial t} + \alpha\frac{\partial^2\psi}{\partial x^2} + \beta|\psi|^2\psi. \quad (2)$$

Нейронная сеть имеет два входа x, t и два выхода, аппроксимирующие действительную и мнимую часть решения $\psi(x, t) = u(x, t) + iv(x, t)$.

Для задачи (1) функция потерь в оригинальном методе PINN имеет следующий вид:

$$L = L_d + L_f. \quad (3)$$

Здесь L_d обозначает сумму невязок начального и граничных условий:

$$L_d = \sum_{j=0}^1 \frac{1}{N_j} \sum_{i=1}^{N_j} |\psi(x_{j_i}, t_{j_i}) - \psi_j(x_{j_i}, t_{j_i})|^2, \quad (4)$$

где $(x_{j_i}, t_{j_i})_{i=1}^{N_j}$ — набор начальных ($j = 0$, $t_{0i} = t_{min}$) и граничных ($j = 1$, $x_{1i} \in \partial\Omega$) точек. В свою очередь, второе слагаемое L_f в (3) есть сумма квадратов остатков уравнения в случайно сгенерированных точках $(x_{f_i}, t_{f_i})_{i=1}^{N_f}$ пространственно-временной области задачи:

$$L_f = \frac{1}{N_f} \sum_{i=1}^{N_f} |f(x_{f_i}, t_{f_i})|^2. \quad (5)$$

Отдельно отметим то, что в этом случае набор данных фактически состоит из множества точек $(x_{f_i}, t_{f_i})_{i=1}^{N_f}$, случайно сгенерированных в пространственно-временной области $\Omega \times T$, для которых вычисляются значения $f(x, t)$ и считается невязка уравнения. Таким образом, наличие L_d в (3) позволяет отслеживать корректность заданных при постановке задачи (1) начальных и граничных условий для нейросетевого решения, а L_f ответственно за выполнение самого уравнения внутри $\Omega \times T$. Параметры нейронной сети, аппроксимирующую решение $\psi(x, t)$, определяются алгоритмом минимизации суммы квадратов остатков (3). Для вычисления невязок уравнения (2) используется автоматическое дифференцирование выходов сети по входам. Общая схема PINN для одномерного НУШ в постановке (1) показана на рис. 1.

1.2. Консервативная модификация нейронной сети. Основная идея модифицированного подхода PINN для решения нелинейных динамических уравнений состоит во введении ограничений на инвариантность сохраняющихся величин с помощью добавления соответствующих остатков в оригинальную функцию потерь при обучении нейронной сети. Для рассматриваемой нами задачи заведомо имеется целый ряд первых интегралов (сохраняющихся величин) при условии, что $\psi_1(x \in \partial\Omega, t) = 0$. В частности, к ним относятся число квантов (мощность):

$$I = \int_{\Omega} |\psi(x, t)|^2 dx \quad (6)$$

и энергии волнового поля:

$$E = \int_{\Omega} \left(\alpha \left| \frac{\partial \psi}{\partial x} \right|^2 - \frac{\beta}{2} |\psi(x, t)|^4 \right) dx. \quad (7)$$

Модификация PINN для НУШ (1), основанная на законе сохранения (6) и (7), подразумевает включение в функции потерь соответствующих невязок: $(I(t) - I_{exact}(t))^2$, $(E(t) - E_{exact}(t))^2$. Для включения законов сохранения в функцию потерь нейронной сети выбирается несколько моментов времени $t_k \in [t_{min}, t_{max}]$, $k = 1, \dots, K$, для них вычисляются значения сохраняющейся величины в текущем нейросетевом решении. В данной работе взято $K = 5$ эквидистантных моментов времени, по аналогии с задачей [13], где рассмотрены интегральные условия на гидродинамическое уравнение непрерывности. Регуляризация в консервативной модификации PINN приближает вычисленные значения к эталонным $I_{exact}(t_k)$, $E_{exact}(t_k)$ за счет минимизации функции потерь. Следовательно, выражение для функции потерь оригинального PINN изменяется:

$$L_1 = L_d + L_f + \varpi_I \Delta I + \varpi_E \Delta E. \quad (8)$$

Здесь ϖ_I и ϖ_E — весовые коэффициенты при соответствующих величинах

$$\Delta I = \frac{1}{K} \sum_{k=1}^K (I(t_k) - I_{exact}(t_k))^2, \quad (9)$$

$$\Delta E = \frac{1}{K} \sum_{k=1}^K (E(t_k) - E_{exact}(t_k))^2, \quad (10)$$

где $I(t_k)$ и $E(t_k)$ вычисляются в процессе обучения для выходов нейронной сети, при этом интегралы в (6) и (7) заменяются суммами на эквидистантной решетке, а производные рассчитываются с помощью автоматического дифференцирования выходов регрессионной модели по входам. Отклонения числа квантов и энергии на предсказанном решении от точных значений также являются метрикой в данной задаче, помимо поточечного отклонения самой аппроксимации от точного решения.

2. Вычислительный эксперимент. В данной работе обсуждаются три вида консервативной модификации PINN через регуляризацию функции потерь, результаты которых сопоставляются с исходным нейросетевым решением с функцией потерь (3). Они получаются из формулы (8), если рассмотреть ограничение только на число квантов $L_I = L_0 + \varpi_I \Delta I$; только на энергию $L_E = L_0 + \varpi_E \Delta E$; и оба ограничения вместе $L_{IE} = L_0 + \varpi_I \Delta I + \varpi_E \Delta E$. В данной работе значения весовых коэффициентов $\varpi_I = \varpi_E = 10^{-2}$ были определены эмпирически (исходя из скорости обучения и получаемых результатов). В дальнейшем можно рассматривать консервативную модификацию PINN совместно с применением известных стратегий взвешивания компонент лосс-функции.

2.1. Односолитонное решение. Наиболее простым аналитическим решением системы НУШ с значениями параметров $\alpha = 1$, $\beta = 2$ является односолитонное решение, которое представляет собой локализованный импульс, распространяющийся с постоянной скоростью и сохраняющий свою форму. Возьмем для примера следующее распределение волнового поля [5, 12]:

$$\psi = 2 \cosh(8t + 2x) e^{-2ix}. \quad (11)$$

Используя выражение (11), можно получить датасет начальных и граничных условий для обучения сети. Случайным образом генерируется $N_0 = 150$ начальных и граничных точек. В качестве тренировочной выборки внутренних точек используется выборка $N_f = 10000$ латинского гиперкуба (Latin hypercube sampling, LHS) [14] для генерации случайных значений внутри области $[(x_{min}, x_{max}), (t_{min}, t_{max})]$. Для сравнения нейросетевого решения с аналитическим внутри области используется дискретизация (11) на равномерной сетке 200×200 (валидационный датасет).

Для моделирования односолитонного решения используется полносвязная сеть с двумя входами (параметры x и t), 4 скрытыми слоями по 100 нейронов в каждом слое, и выходной слой с двумя выходами (действительная и мнимая части решения $u(x, t)$, $v(x, t)$), в качестве функции активации используется гиперболический тангенс. Здесь и далее выбор параметров сделан на основе работ [4–5, 9].

Для регуляризации функции потерь (8) и контроля над законом сохранения берется $K = 5$ эквидистантных моментов времени, в которых интегралы (6), (7) аппроксимируются суммами по 100 точкам на (x_{min}, x_{max}) . Для минимизации функции потерь выбран

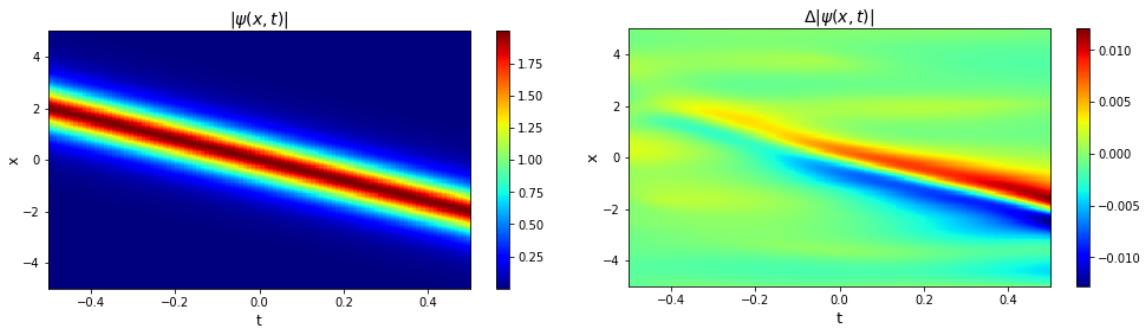


Рис. 2. Точное односолитонное решение (слева) и распределение ошибки предсказания нейронной сети (справа)

Таблица 1

Значения метрик качества односолитонного решения
для консервативных модификаций PINN

Метрика	Baseline	Quants	Energy	Quants + Energy
$MSE(\Delta I_{val})$	0.0316	0.0088	0.0412	0.0044
$MSE(\Delta E_{val})$	0.1095	0.1095	0.0035	0.0023
$\Delta \psi $	0.0136	0.0112	0.0109	0.0044

алгоритм оптимизации L-BFGS [15]. Все эксперименты в данной работе проводились с использованием Python 3.8 и библиотеки глубокого обучения PyTorch 1.13.

Далее представлены результаты обучения модели для односолитонного решения. На рис. 2 показаны точное решение (слева) и разность лучшего предсказанного и точного решений (справа). На рис. 3, а, б, показана динамика невязок (9), (10) числа квантов и энергии, усредненных по выбранным K моментам времени, во время обучения сети. Также представлена разность зависимостей числа квантов и энергии от времени для предсказанного и точного решений ($I(t) - I_{exact}(t)$, $(E(t) - E_{exact}(t))$) на валидационном массиве. Обозначения на рисунках и таблицах следующие: Baseline — обучение с исходной функцией потерь L_0 , Quants/Energy conservation law — обучение с регуляризациями вида $L_0 + \varpi_I \Delta I$ и $L_0 + \varpi_E \Delta E$, соответственно, Quants + Energy — обучение с регуляризацией $L_0 + \varpi_I \Delta I + \varpi_E \Delta E$.

Как и ожидалось, при дополнительных ограничениях в функции потерь закон сохранения выполняется лучше, что наблюдается и в процессе обучения, и для предсказанного решения. Зависимость средней невязки энергии от итерации существенно менее стабильна, чем для числа квантов, что может объясняться более сложным выражением для энергии и наличием производных в нем.

В качестве метрики качества решения рассматриваются: отношение нормы разности $|\psi(x_{val}, t_{val})|$ с точным решением к норме точного решения, где (x_{val}, t_{val}) — валидационный датасет; среднеквадратичное отклонение зависимостей $I(t)$ и $E(t)$ от эталонных. По результатам обучения были получены следующие значения метрик качества (табл. 1). Видно, что в данном случае консервативная регуляризация функции потерь улучшает результат обучения сети не только по метрике законов сохранения, но и уменьшает среднеквадратичное отклонение от точного решения. При этом во всех случаях распределение

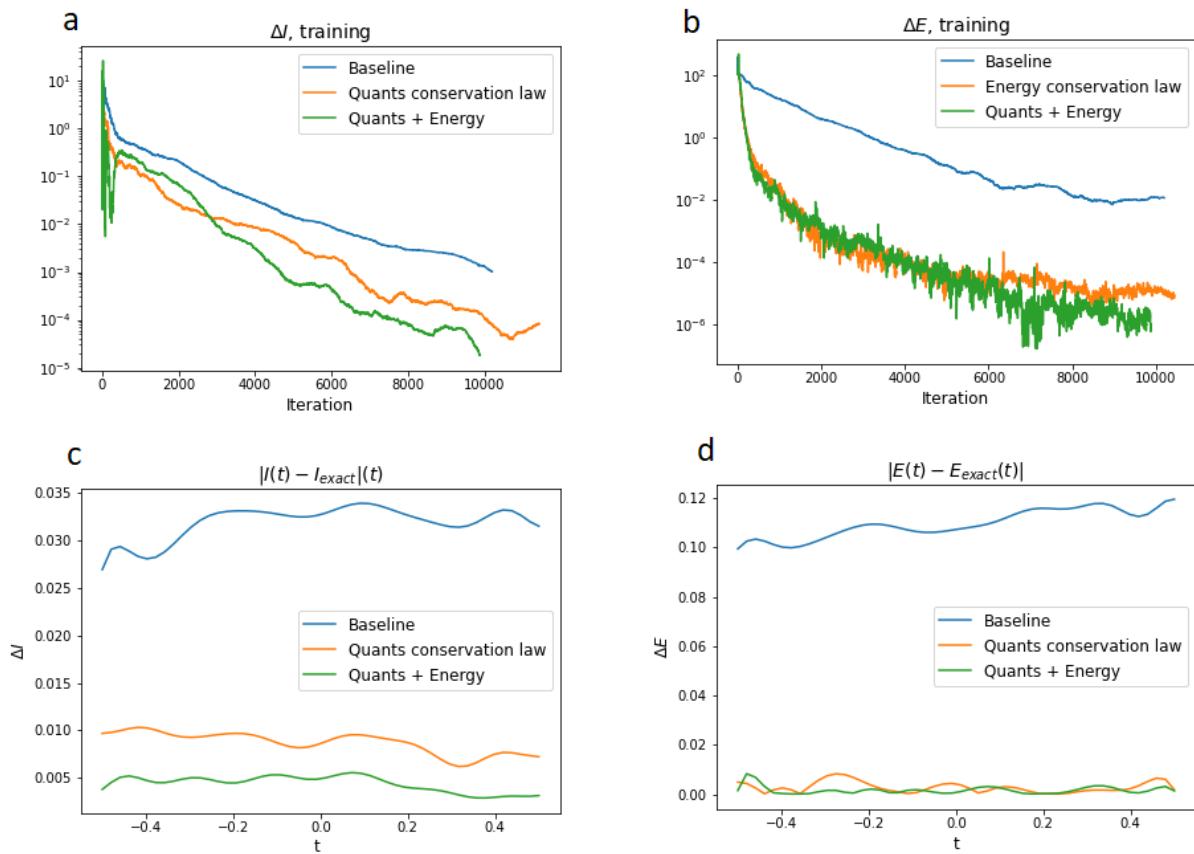


Рис. 3. Сверху: зависимость усредненных невязок числа квантов (а) и энергии (б) от итерации обучения. Внизу: качество выполнения закона сохранения для предсказанного решения: число квантов (с), энергия (д)

погрешности предсказания в пространственно-временной области показывает накопление ошибки во времени (рис. 2). Из рис. 3 можно сделать вывод, что обучение с дополнительной регуляризацией занимает примерно столько же времени (количество итераций оптимизатора сопоставимо). Из табл. 1 также видно, что применение обоих законов сохранения дает наиболее точное предсказание по всем метрикам.

2.2. Двухсолитонное решение. Далее рассмотрено более сложное двухсолитонное решение, точная аналитическая форма которого задается следующим выражением [5], [12]:

$$\psi = \frac{1.848e^{1.96it} \cosh(0.8x) - 1.056e^{0.64it} \cosh(1.4x)}{-2.24 \cos(1.32t) + 2.42 \cosh(0.6x) + 0.18 \cosh(2.2x)}. \quad (12)$$

Данное точное близерное решение и распределение ошибки наилучшего предсказания нейронной сети показаны на рис. 4. В левой части рис. 4 можно отчетливо видеть два солитона одинаковой ширины и амплитуды (в общем случае они отличаются), взаимодействие которых сохраняет энергию системы и дает пик большей амплитуды в центре, который затем симметрично разделяется снова на два одинаковых солитона.

Условия вычислительного эксперимента остаются прежними, за исключением архитектуры сети: для нахождения решения более сложной формы нужно усложнять модель. Было взято 6 скрытых слоев по 100 нейронов на слое. Универсальный алгоритм подбора

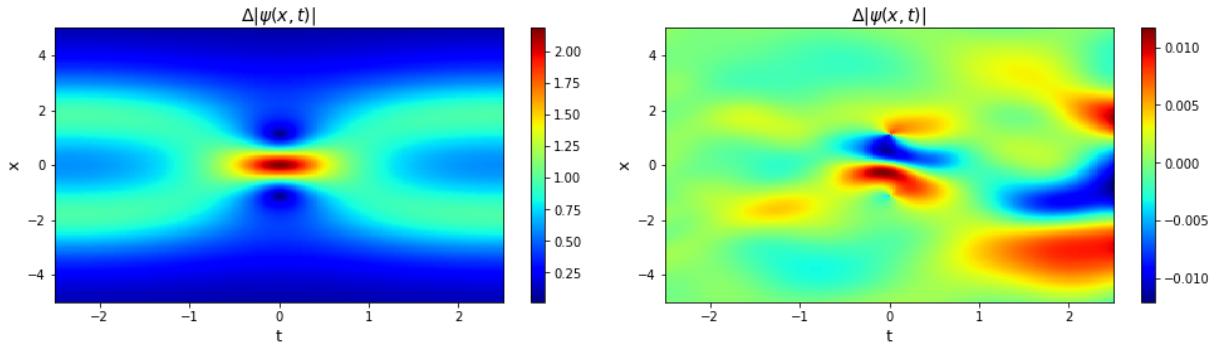


Рис. 4. Точное двухсолитонное решение (слева) и распределение ошибки предсказания нейронной сети (справа)

Таблица 2

Значения метрик качества двухсолитонного решения
для консервативных модификаций PINN

Метрика	Baseline	Quants	Energy	Quants + Energy
$MSE(\Delta I_{val})$	0.0237	0.0016	0.0050	0.0011
$MSE(\Delta E_{val})$	0.0374	0.0152	0.0039	0.0024
$\Delta \psi $	0.0053	0.0079	0.0046	0.0079

архитектуры на данный момент на стадии исследований, поэтому эмпирически выбирается вариант, подходящий с точки зрения точности и времени обучения. Аналогично тому, как это было сделано в разделе 2.1, рассмотрим динамику отклонения числа квантов (6) и энергии (7) от истинных значений для нейросетевого решения во время обучения и зависимости этих физических величин от времени для итогового предсказания (см. рис. 5). По кривым обучения можно видеть, что дополнительные регуляризации сильно меняют структуру функции потерь, вследствие чего зависимость лосс-компонент, соответствующих законам сохранения, нестабильна: невязки величин существенно увеличиваются ближе к 20000 итераций, затем снова спадают. При использовании одновременно двух законов сохранения, $\Delta I + \Delta E$, кривая обучения выходит на участок плато, и затем функция все же спадает, но при значительном увеличении числа итераций (обучение почти в 2 раза дольше исходной версии). Таким образом, в отличие от моделирования распространения солитона, для консервативной модификации PINN обучение замедляется, что было ожидаемо для существенно более сложного режима взаимодействия двух солитонов.

Значения величин погрешностей для рассматриваемых величин приведены в табл. 2. В данной серии экспериментов для двух солитонов улучшение по всем метрикам было получено только для регуляризации $L_E = L_0 + \varpi_E \Delta E$. При этом динамика компонент функции потерь во время обучения очень нестабильна (или выходит на плато), что замедляет обучение и приводит к улучшению точности только по метрикам законов сохранения. Следовательно, рассматриваемые регуляризации не являются универсальными для разных систем и законов сохранения.

2.3. *Волна-убийца*. Исследование волн-убийц актуально для многих разделов физики. Такая волна обладает многими характерными свойствами, например, пиковой амплитудой, более чем двухкратно превосходящей общий фон. Наличие волн-убийц также характер-

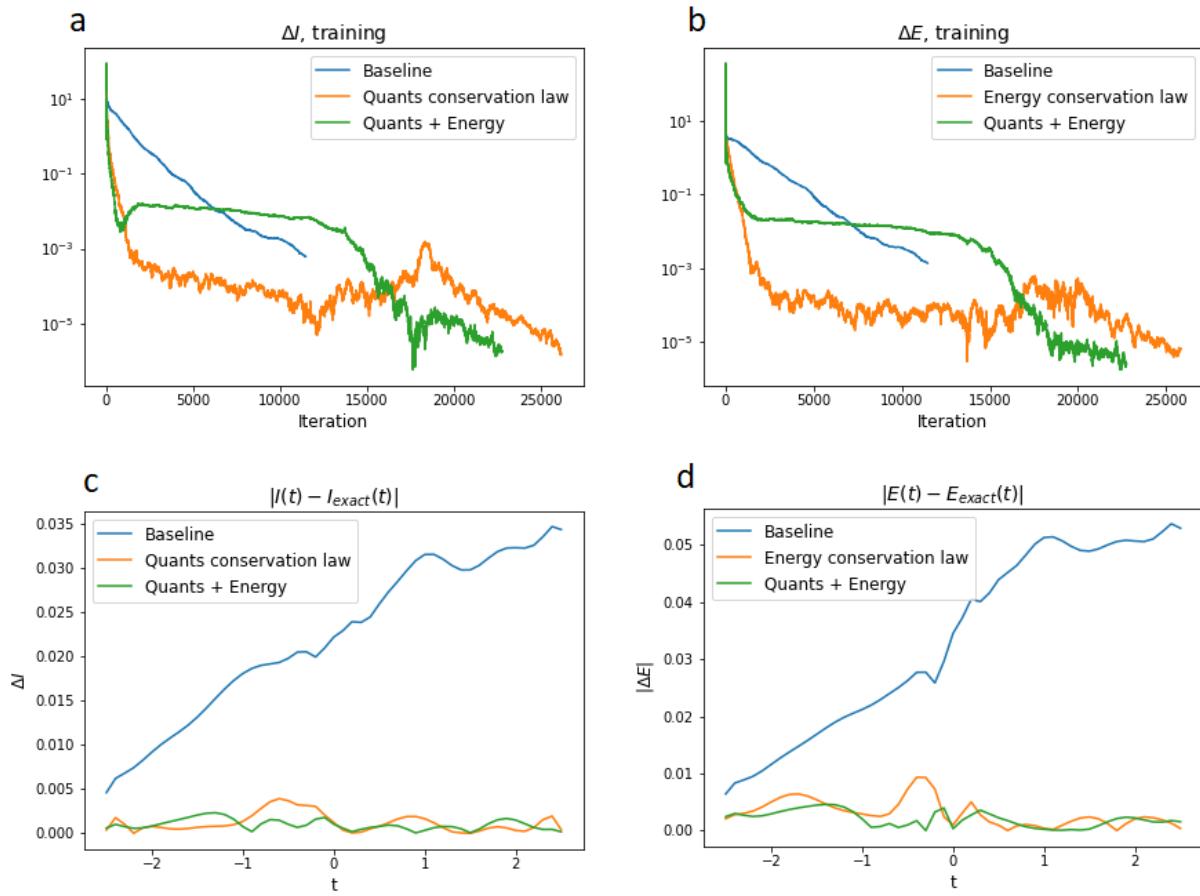


Рис. 5. Сверху: зависимость усредненных невязок числа квантов (а) и энергии (б) от итерации обучения. Внизу: качество выполнения закона сохранения для предсказанного решения: число квантов (с), энергия (д)

ризует степень неустойчивости в системе, поэтому их исследование представляет большой практический интерес. Форма волн-убийц в рамках НУШ описывается рациональными функциями. Для НУШ с коэффициентами $\alpha=0.5$, $\beta=1$ наиболее простое решение первого порядка имеет следующий вид [5], [12]:

$$\psi(x, t) = \left(1 - \frac{4(1 + 2it)}{4t^2 + 4x^2 + 1}\right) e^{it}. \quad (13)$$

На рис. 6 показаны точное решение (13) и разность лучшего предсказанного и точного решений. На левом фрагменте рис. 6 видно, что распределение поля проще по сравнению с близким решением (12), но сложнее по сравнению со случаем распространения солитона (11). С использованием (13) составляется набор данных для моделирования волны-убийцы первого порядка нейронной сетью с 4 слоями по 100 нейронов на слой.

Аналогично моделированию одного и двух солитонов рассмотрим динамику отклонения числа квантов и энергии от истинных значений во время обучения и зависимости $I(t)$, $E(t)$ для итогового предсказания (валидация) (см. рис. 7). По кривым обучения видно, что, как и при моделировании двухсолитонного решения, компоненты функции потерь

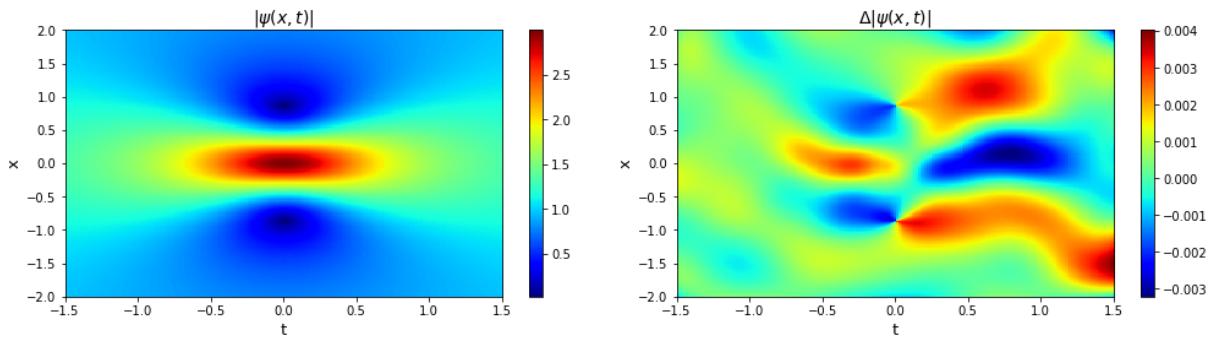


Рис. 6. Точное решение вида «волна-убийца» (слева) и распределение ошибки предсказания нейронной сети (справа)

Таблица 3

Значения метрик качества двухсолитонного решения
для консервативных модификаций PINN

Метрика	Baseline	Quants	Energy	Quants + Energy
$MSE(\Delta I_{val})$	0.0188	0.0084	0.0265	0.0152
$MSE(\Delta E_{val})$	0.0187	0.0591	0.0021	0.0061
$\Delta \psi $	0.0029	0.0057	0.0010	0.0048

выходят на плато при введении дополнительных ограничений на сохраняющиеся величины. Особенно сильно плато проявляется при использовании ограничения на энергию. В отличие от результатов моделирования двух солитонов, в данном случае обучение замедляется только для регуляризации с невязкой энергии, при этом увеличение числа итераций несущественно (рис. 7, б). Для остальных двух случаев делается меньше итераций оптимизатора, чем при минимизации исходной функции потерь (baseline, L_0).

Значения метрик качества для моделирования волны-убийцы приведены в табл. 3. Как и ожидалось, наилучший результат показала регуляризация с ограничением на энергию, поскольку для функции потерь более сложного вида сделано большее число итераций оптимизатора. Для двухсолитонного решения результаты с использованием закона сохранения энергии также были наилучшими, несмотря на сопоставимое число итераций для всех консервативных модификаций, превосходящее исходное в два раза.

Заключение. В данной работе при глубоком обучении PINN нейронных сетей, аппроксимирующих решения НУШ различной сложности, были применены законы сохранения физических характеристик системы: числа квантов и энергии. На их основе рассмотрены три регуляризации функции потерь: ограничение только на число квантов, ограничение только на энергию и совместное ограничение на число квантов и энергию. Были рассмотрены три качественно разных аналитических решения НУШ: односолитонное и двухсолитонное (брозерное) решение, волна-убийца первого порядка. Целью работы было улучшение физических свойств решения с точки зрения выполнения законов сохранения. Показано, что введение дополнительных консервативных ограничений в функцию потерь уменьшает отклонение значений числа квантов и энергии от эталонных. Однако для более сложных ситуаций (взаимодействие двух солитонов и волна-убийца) обучение нейронной сети замедляется, и результаты предсказаний по метрике среднеквадратичного

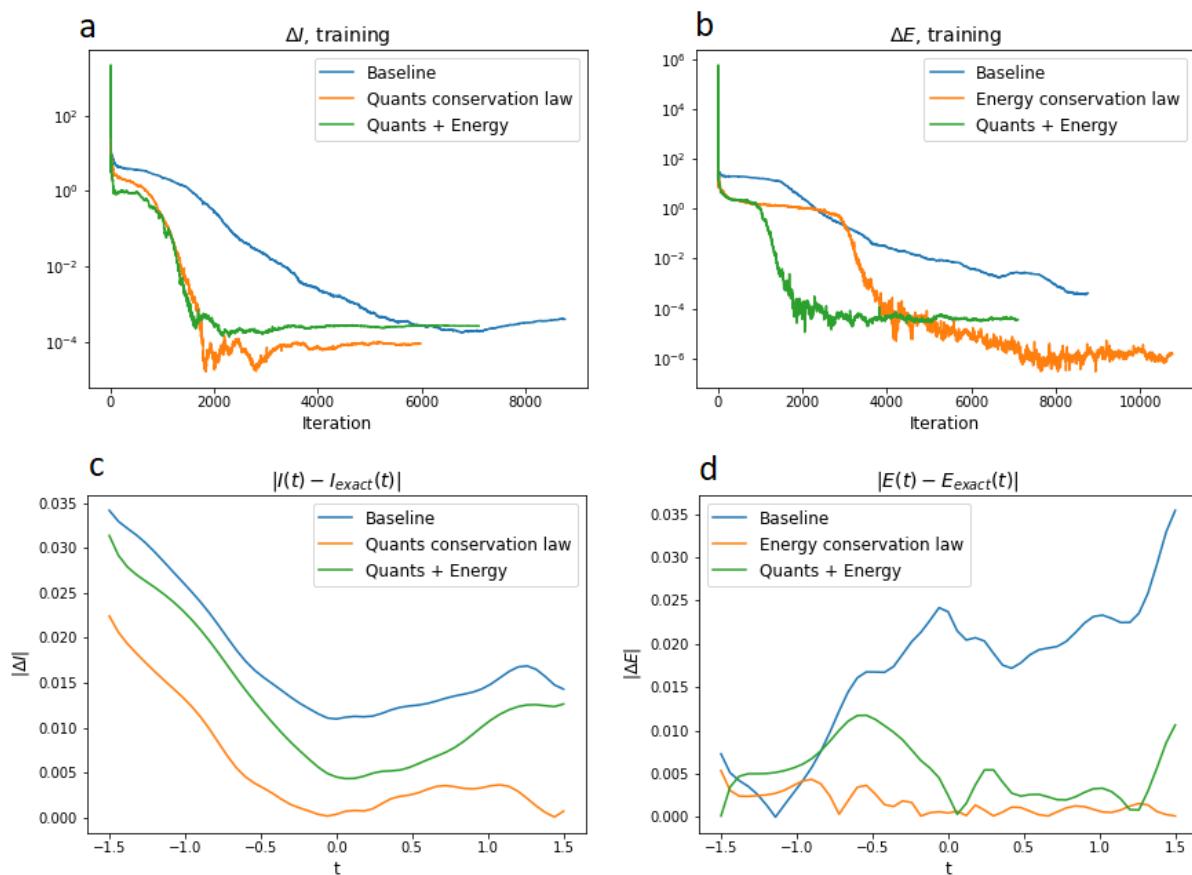


Рис. 7. Сверху: зависимость ΔI (а) и ΔE (б) от итерации обучения. Внизу: качество выполнения закона сохранения для предсказанного решения: $I(t)$ (с), $E(t)$ (д)

отклонения решения от эталонного имеют однозначную тенденцию к улучшению для регуляризации, учитывающей только закон сохранения энергии. Следовательно, эффект от консервативной модификации подхода PINN является относительно индивидуальным для разных систем и разных сохраняющихся величин, а обобщающие способности таких методов глубокого обучения, принимающих во внимание интегралы движения, требуют дополнительного изучения и анализа для других задач (например, НУШ с нелинейной дисперсией или многомерных вариантов постановки проблемы).

Список литературы

1. Moxley F., Chuss D., Dai W. A generalized finite-difference time domain scheme for solving nonlinear Schrödinger equations // Computer Physics Communications, 2013. N 184 (8). P. 1834–1841. [El. Res.]: <https://doi.org/10.1016/j.cpc.2013.03.006>.
2. Васильев Е. П., Болотов Д. И., Болотов М. И., Смирнов Л. А. Нейросетевой подход к решению задачи самовоздействия волновых полей в нелинейных средах // Проблемы Информатики, 2022. № 1. С. 5–16. [Электрон. рес.]: <https://doi.org/10.24412/2073-0667-2022-1-5-16>.
3. Karniadakis, G. E., Kevrekidis, I. G., Lu, L. et al. Physics-informed machine learning // Nature Reviews Physics. 2021. N 3. P. 422–440. [El. Res.]: <https://doi.org/10.1038/s42254-021-00314-5>.

4. Raissi M., Perdikaris P., Karniadakis G. E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations // Journal of Computational Physics. 2019. N 378. P. 686–707. [El. Res.]: <https://doi.org/10.1016/j.jcp.2018.10.045>.
5. Pu J. C., Li J., Chen Y. Soliton, breather, and rogue wave solutions for solving the nonlinear Schrödinger equation using a deep learning method with physical constraints // Chinese Physics B, 2021. N 30 (6), 060202. [El. Res.]: <https://dx.doi.org/10.1088/1674-1056/abd7e3>.
6. Wu G. Z., Fang Y., Wang Y. Y., Wu G. C., Dai C. Q. Predicting the dynamic process and model parameters of the vector optical solitons in birefringent fibers via the modified PINN // Chaos, Solitons and Fractals. 2021. N 152, 111393. [El. Res.]: <https://doi.org/10.1016/j.chaos.2021.111393>.
7. Wang L., Yan. Z. Data-driven rogue waves and parameter discovery in the defocusing nonlinear Schrödinger equation with a potential using the PINN deep learning // Physics Letters A. 2021. V. 404. P. 127408. [El. Res.]: <https://doi.org/10.1016/j.physleta.2021.127408>.
8. Jagtap, A., Kharazmi, E., Karniadakis, G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems // Computer Methods in Applied Mechanics and Engineering. 2020. N 365, 113028. [El. Res.]: <https://doi.org/10.1016/j.cma.2020.113028>.
9. Wu G. Z., Fang Y., Kudryashov N., Wang Y. Y., Dai C. Q. Prediction of optical solitons using an improved physics-informed neural network method with the conservation law constraint // Chaos, Solitons and Fractals, 2022. 159(C), 112143. [El. Res.]: <https://doi.org/10.1016/j.chaos.2022.112143>.
10. Lin S., Chen Y. A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions // Journal of Computational Physics, 2022. 457(C), 111053. [El. Res.]: <https://doi.org/10.1016/j.jcp.2022.111053>.
11. [El. Res.]: https://docs.nvidia.com/deeplearning/modulus/user_guide/theory/recommended_practices.html.
12. Zakharov V., Manakov S. On the complete integrability of a nonlinear Schrödinger equation // Theoretical and Mathematical Physics. 1974. N 19 (3). P. 551–559. [El. Res.]: <https://doi.org/10.1007/BF01035568>.
13. [El. Res.]: https://docs.nvidia.com/deeplearning/modulus/user_guide/foundational/scalar_transport.html.
14. Stein M. Large sample properties of simulations using Latin hypercube sampling, Technometrics. 1987. N 29. P. 143–151. [El. Res.]: <https://www.jstor.org/stable/1269769>.
15. Schraudolph N. N., Yu J., Gunter S. A Stochastic Quasi-Newton Method for Online Convex Optimization // International Conference on Artificial Intelligence and Statistics, 2007. P. 436–443. [El. Res.]: <http://proceedings.mlr.press/v2/schraudolph07a/schraudolph07a.pdf>.



Гурьева Юлия Вадимовна — лаборант-исследователь научно-исследовательской лаборатории «Искусственного интеллекта и больших массивов данных» кафедры теории управления и динамики систем Института информационных технологий, математики и механики ННГУ им. Н. И. Лобачевского; e-mail: gureva-yulya@list.ru, тел. 8-831-462-33-20. Область научных инте-

ресов: искусственный интеллект, глубокое обучение, нелинейная динамика.

Yulia Gurieva is a research assistant at the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: artificial intelligence, deep learning, nonlinear dynamics.



Васильев Евгений Павлович — младший научный сотрудник научно-исследовательской лаборатории «Искусственного интеллекта и больших массивов данных», старший преподаватель кафедры математического обеспечения и суперкомпьютерных технологий Института информационных технологий, математики и механики ННГУ им. Н.И. Лобачевского; e-mail: eugene.unn@gmail.com, тел. 8-831-462-33-20. Область научных интересов: научная визуализация, компьютерное зрение, глубокое обучение.

Evgenii Vasiliev is a Research Associate at the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System Dynamics and lecturer of Department of Software and Supercomputer Technologies of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: scientific visualization, computer vision, deep learning.

Смирнов Лев Александрович — кандидат физико-математических наук, заведу-

ющий научно-исследовательской лабораторией «Искусственного интеллекта и больших массивов данных» кафедры теории управления и динамики систем Института информационных технологий, математики и механики ННГУ им. Н.И. Лобачевского; e-mail: lev.smirnov@itmm.unn.ru, тел. 8-831-462-33-20. Область научных интересов: нелинейные волны (солитоны и вортексы), неравновесная статистическая механика, турбулентность, высокопроизводительные вычисления.



Lev Smirnov is a PhD in Physical and Mathematical Sciences, the Head of the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: nonlinear waves (solitons and vortices), non-equilibrium statistical mechanics, turbulence, high-performance computing.

Дата поступления — 31.01.2023