

ЧИСЛЕННЫЕ МЕТОДЫ | NUMERICAL METHODS

Научная статья | Original paper

УДК 519.6:004.85

Встраивание физически-информированных нейронных сетей в численные схемы для моделирования динамических систем

И.А. Великоречанин¹ ✉, Т.В. Лазовская², Д.А. Тархов²

¹ Московский физико-технический институт (государственный университет) (МФТИ ГУ)
г. Долгопрудный, Московская область, Россия

² Санкт-Петербургский политехнический университет Петра Великого (СПбПУ)
г. Санкт-Петербург, Россия

✉ veligoran2001@gmail.com

Резюме

Контекст и актуальность. Физически-информированные нейронные сети (Physics-Informed Neural Networks, PINN) рассматриваются как перспективный инструмент математического моделирования динамических систем, описываемых дифференциальными уравнениями. Однако классические PINN приемлемой точности имеют большое число весов. Их подбор (обучение сети) является задачей нелинейной оптимизации, решение которой требует многократного вычисления производных, что приводит к значительным вычислительным затратам и ограничивает их применение в задачах моделирования. **Цель.** Разработка и экспериментальная проверка подхода к математическому моделированию динамических систем, основанного на встраивании физически-информированных нейронных сетей в классические численные схемы. **Гипотеза.** Встраивание численных схем в архитектуру физически-информированных нейронных сетей позволяет радикально уменьшить число весов и тем самым, повысить вычислительную эффективность моделирования динамических систем при сохранении сопоставимой точности по сравнению с классическими PINN. **Методы и материалы.** Предложена архитектура нейросетевой модели, интегрирующая обобщенную вычислительную схему метода трапеций непосредственно в архитектуру PINN. Компактная параметрическая сеть с 13 обучаемыми параметрами адаптивно выбирает баланс между явной и неявной численными схемами в различных



точках пространственно-временной области. Экспериментальное исследование проведено для четырех дифференциальных уравнений: уравнения теплопроводности, волнового уравнения, уравнения реакции-диффузии и уравнения Бюргерса. **Результаты.** Показано, что при фиксированном времени обучения предложенная модель с 13 параметрами достигает сопоставимой или более высокой точности по сравнению с классическими PINN с большим числом параметров, при этом демонстрируя существенно меньшую дисперсию результатов между запусками. **Выводы.** Интеграция численных схем в архитектуру физически-информированных нейронных сетей позволяет повысить эффективность построения моделей динамических систем и обеспечивает интерпретируемость архитектуры через используемые численные методы. Предложенный подход может рассматриваться как основа для построения компактных моделей в задачах суррогатного моделирования.

Ключевые слова: физически-информированные нейронные сети, численные схемы, адаптивные методы, динамические системы, дифференциальные уравнения, суррогатное моделирование

Для цитирования: Великоречанин, И.А., Лазовская, Т.В., Тархов, Д.А. (2026). Встраивание физически-информированных нейронных сетей в численные схемы для моделирования динамических систем. *Моделирование и анализ данных*, 16(3), 169—185. <https://doi.org/10.17759/mda.2026160308>

Embedding physics-informed neural networks into numerical schemes for modeling dynamical systems

I.A. Velikorechanin¹ ✉, T.V. Lazovskaya², D.A. Tarkhov²

¹ Moscow Institute of Physics and Technology (MIPT), Moscow, Russia

² Peter the Great Saint Petersburg Polytechnic University (SPbPU), Saint Petersburg, Russia

✉ veligoran2001@gmail.com

Abstract

Context and relevance. Physics-Informed Neural Networks (PINNs) are considered a promising tool for mathematical modeling of dynamical systems described by differential equations. However, classical PINN approaches require repeated computation of high-order derivatives, which leads to significant computational costs and limits their applicability in modeling tasks. **Objective.** To develop and experimentally validate an approach to mathematical modeling of dynamical systems based on embedding physics-informed neural networks into classical numerical schemes. **Hypothesis.** Embedding numerical schemes into the architecture of physics-informed neural networks can improve the computational efficiency



of dynamical system modeling while maintaining accuracy comparable to classical PINN approaches. **Methods and materials.** A neural network architecture integrating a generalized θ -scheme (trapezoidal method) directly into the PINN architecture is proposed. A compact parametric network with 13 trainable parameters adaptively selects the balance between explicit and implicit numerical schemes at different points of the space–time domain. The experimental study was conducted on four types of differential equations characteristic of various dynamical processes: the heat equation, wave equation, reaction–diffusion equation, and Burgers equation. **Results.** In the experiments, under a fixed training time budget, the proposed 13-parameter model achieves comparable or higher accuracy than a 67-parameter classical PINN, while exhibiting significantly lower variance across runs. **Conclusions.** Integrating numerical schemes into the architecture of physics-informed neural networks improves the efficiency of constructing dynamical system models and provides interpretability of the architecture through the underlying numerical methods. The proposed approach can be considered a basis for building compact models in surrogate modeling tasks.

Keywords: physics-informed neural networks, numerical schemes, adaptive methods, dynamical systems, differential equations, surrogate modeling

For citation: Velikorechanin, I.A., Lazovskaya, T.V., Tarkhov, D.A. (2026). Embedding physics-informed neural networks into numerical schemes for modeling dynamical systems. *Modelling and Data Analysis*, 16(3), 169—185. (In Russ.). <https://doi.org/10.17759/mda.2026160308>

Введение

Традиционно приближенные методы решения дифференциальных уравнений подразделяются на численные (методы Эйлера, Рунге–Кутты и др.) и аналитические (разложения в ряд Тейлора или Фурье, методы малого параметра и др.). Эти методы широко используются при построении моделей динамических систем, описываемых дифференциальными уравнениями. В конце XX века к приближенным аналитическим методам были отнесены и нейронные сети. В настоящее время данный подход получил развитие в рамках концепции Physics-Informed Neural Networks (PINN) — метода решения дифференциальных уравнений, сочетающего универсальность нейронных сетей с учетом физических законов (Antonov, Tarkhov, Vasilyev, 2018; Raisi, Perdikaris, Karniadakis, 2019; Tarkhov, Vasilyev, 2019; Vasilyev, Tarkhov, 2014).

Классические PINN аппроксимируют решение дифференциального уравнения с помощью нейронной сети, вычисляя необходимые производные через автоматическое дифференцирование и минимизируя остаток уравнения. Несмотря на успешное применение в различных областях, такой подход сопровождается значительными вычислительными затратами и часто обладает ограниченной интерпретируемостью результатов (Karniadakis et al., 2021).

В последние годы было показано, что классические численные методы могут рассматриваться как аналитические при применении к интервалам переменной



длины (Lazovskaya, Tarkhov, 2016; Lazovskaya, Tarkhov, Vasilyev, 2018; Tarkhov, Vasilyev, 2019). Это позволяет по-новому интерпретировать связь между численными и аналитическими методами и использовать ее при построении нейросетевых моделей динамических систем. С этой точки зрения особый интерес представляет концепция многоуровневого моделирования (multi-fidelity modeling), предполагающая использование упрощенных моделей (low-fidelity) для эффективного исследования пространства решений с последующим уточнением результатов с помощью более точных, но вычислительно затратных моделей (high-fidelity) (Cuomo et al., 2022; Jagtap, Kawaguchi, Karniadakis, 2020; Peherstorfer, 2019; Penwarden et al., 2022; Wang, Teng, Perdikaris, 2021).

Использование аналитической модификации численных методов (АМЧМ) с небольшим числом слоев модели (здесь и далее под слоем модели понимается шаг исходного численного метода) позволяет строить компактные модели уровня low-fidelity с приемлемой точностью. Переход к моделям уровня high-fidelity может быть реализован посредством интеграции численных методов непосредственно в архитектуру нейронных сетей (Lazovskaya, Tarkhov, 2026; Tarkhov, Lazovskaya, Malykhina, 2023). Это позволяет существенно сократить число обучаемых параметров и повысить интерпретируемость модели (Tarkhov, Lazovskaya, Malykhina, 2023).

Настоящая работа развивает подход, предложенный в (Lazovskaya, Tarkhov, 2026; Tarkhov, Lazovskaya, Malykhina, 2023), путем построения архитектуры LF-PINN, основанной на обобщенном методе трапеций с адаптивным параметром θ . Компактная параметрическая сеть, подставленная в схему вместо θ автоматически выбирает баланс между явной и неявной схемами для различных точек пространственно-временной области, адаптируясь к локальным особенностям решения и обеспечивая построение эффективных моделей динамических систем.

Целью исследования является демонстрация того, что интеграция численных методов в архитектуру PINN позволяет существенно ускорить обучение при сохранении сопоставимой или более высокой точности моделирования по сравнению с классическим подходом. Дальнейшее повышение точности модели может быть достигнуто за счет увеличения размера используемой нейронной сети.

Материалы и методы

Постановка задачи

Рассмотрим задачу Коши для системы обыкновенных дифференциальных уравнений

$$\begin{cases} \mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)), \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases} \quad (1)$$

Классический метод Эйлера состоит в разбиении промежутка, на котором ищется решение на n частей: $x_0 < x_1 < \dots < x_k < x_{k+1} < \dots < x_n$, и применении итерационной формулы



$$\mathbf{y}_{k+1} = \mathbf{y}_k + h_k \mathbf{f}(x_k, \mathbf{y}_k), \quad (2)$$

где $h_k = x_{k+1} - x_k$; $\mathbf{y}_k$ — приближение к точному значению искомого решения $\mathbf{y}(x_k)$.

Для получения приближенного решения в виде функции по полученным поточечным приближениям обычно проводят ломаную линию (ломаная Эйлера) или сплайн.

Мы предлагаем принципиально другой подход. С помощью формулы (2) будем строить приближенное решение задачи (1) на переменном интервале $\tilde{D} = [\min(x_0, x); \max(x_0, x)]$, где $x \in [a; b]$. При этом $h_k = h_k(x)$, $\mathbf{y}_k = \mathbf{y}_k(x)$, $\mathbf{y}_0(x) = \mathbf{y}_0$. В качестве искомого приближенного решения предлагается использовать $\mathbf{y}_n(x)$.

Для численного решения задачи Коши (1) разработана широкая палитра численных методов (Hairer, Nørsett, Wanner, 1993). Значительная часть из них заключается в применении рекуррентной формулы:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + F(\mathbf{f}, h_k, x_k, \mathbf{y}_k, \mathbf{y}_{k+1}). \quad (3)$$

Здесь оператор F определяет конкретный метод.

Если функция F зависит от $\mathbf{y}_{k+1}$, то равенство (3) можно рассматривать как уравнение относительно $\mathbf{y}_{k+1}$. Данное уравнение может допускать точное решение, тогда вместо (3) получаем соотношение вида

$$\mathbf{y}_{k+1} = B(\mathbf{f}, \mathbf{y}_k, h_k, x_k). \quad (4)$$

Формула (4), как и ранее, позволяет вычислить $\mathbf{y}_n(x)$ и использовать его как приближенное решение задачи (1).

Если уравнение (3) точно решить относительно $\mathbf{y}_{k+1}$ не удастся, то для получения формулы вида (4) можно использовать какой-либо приближенный метод или специально обученную нейронную сеть, как это сделано в (Tarkhov, Lazovskaya, Malykhina, 2023). В итоге получается глубокая нейронная сеть специального вида, которая дает сколь угодно точное решение задачи (1), если использовано достаточно точное решение неявного уравнения и достаточно много слоев n .

Метод распространяется и на дифференциальные уравнения в частных производных вида $\frac{\partial}{\partial t} u(x, t) = F(u(x, t), x, t)$, где $u(x, t)$ — достаточно гладкая функция относительно переменных $(x, t) \in R^p \times R^+$, F — вообще говоря, нелинейное отображение, например, дифференциальный оператор. К этому уравнению можно применить формулы (3).

Адаптивная θ -схема

В работе в качестве (3) рассматривается обобщенная θ -схема, которая представляет собой взвешенную комбинацию явного и неявного методов Эйлера:

$$y_{n+1} = y_n + h \cdot \left[(1-\theta) \cdot f(y_n, t_n) + \theta \cdot f(y_{n+1}, t_{n+1}) \right]$$



где h — размер временного шага, θ — параметр схемы. Различные значения θ соответствуют различным численным методам:

- $\theta = 0$ — явный метод Эйлера
- $\theta = 0,5$ — метод трапеций
- $\theta = 1$ — неявный метод Эйлера

В LF-PINN параметр θ не является константой, а адаптивно вычисляется θ -net для каждой точки пространственно-временной области. Это позволяет модели локально балансировать между устойчивостью (высокие значения $\theta \rightarrow$ неявная схема) и вычислительной эффективностью (низкие значения $\theta \rightarrow$ явная схема). Например, в областях с крутыми градиентами модель может автоматически увеличивать θ для повышения устойчивости.

θ -net представляет собой компактную нейронную сеть с одним скрытым слоем, которая определяет тип численной схемы для каждой точки пространственно-временной области:

$$\theta \left(x, t, t_{next}, \left\| \frac{\partial y}{\partial x} \right\| \right) = \sigma \left(\sum_{i=1}^d c_i \varphi_i \left(x, t, t_{next}, \left\| \frac{\partial y}{\partial x} \right\| \right) + b_0 \right),$$

где

$$\varphi_i \left(x, t, t_{next}, \left\| \frac{\partial y}{\partial x} \right\| \right) = \tanh \left(w_{i1} x + w_{i2} t + w_{i3} t_{next} + w_{i4} \left\| \frac{\partial y}{\partial x} \right\| + b_i \right), \quad i=1, \dots, d,$$

$$d = 2, \quad w_{ij} \in \mathbb{R}, \quad c_i \in \mathbb{R}, \quad b_i \in \mathbb{R}, \quad b_0 \in \mathbb{R}$$

где σ — сигмоидальная функция активации. Сеть имеет всего 13 обучаемых параметров и получает на вход четыре признака:

x — пространственная координата

t — текущее время

t_{next} — время следующего шага

$\left\| \frac{\partial y}{\partial x} \right\|$ — нормализованная норма пространственного градиента текущего решения (позволяет адаптироваться к локальной жесткости задачи)

Компактность сети (13 параметров) критична для предотвращения переобучения и обеспечения обобщающей способности. Малое количество параметров гарантирует, что модель не будет «запоминать» решение, а будет учиться выбирать правильный тип численной схемы на основе локальных свойств задачи. В связи с этим данная сеть далее называется LF-PINN. При необходимости повысить точность можно нарастить размер сети или использовать отдельную сеть для каждого слоя.

Итерационное решение неявной части схемы

Когда $\theta > 0$, схема становится неявной, то есть значение y_{n+1} входит в правую часть уравнения. Для решения такого уравнения применяется метод простой итерации. На каждом временном шаге выполняется N_{iter} итераций по формуле:

$$y_{n+1}^{(k+1)} = y_n + h \cdot \left[(1-\theta) \cdot f(y_n, t_n) + \theta \cdot f(y_{n+1}^{(k)}, t_{n+1}) \right]$$



где $k = 0, 1, \dots, N_{iter} - 1$, и начальное приближение $y_{(n+1)}^{(0)} = y_n$. В экспериментах использовалось $N_{iter} = 2-4$ итерации, что обеспечивает баланс между точностью и вычислительной эффективностью.

Например, для уравнения теплопроводности на каждой итерации необходимо:

Вычислить $u_x = \partial u / \partial x$ через автоматическое дифференцирование

Вычислить $u_{xx} = \partial^2 u / \partial x^2$ через повторное дифференцирование

Получить $f = \alpha \cdot u_{xx}$ и обновить решение

Таким образом, на каждом временном шаге требуется N_{iter} вычислений пространственных производных второго порядка.

Модель реализует адаптивный механизм выбора числа временных шагов на основе целевого временного горизонта t_{end} :

$$n_{steps} = \min(\lceil t_{end} / h_{max} \rceil, n_{max})$$

где h_{max} — максимально допустимый размер шага, n_{max} — максимальное количество шагов. Такой подход обеспечивает, что размер каждого временного шага $h = t_{end} / n_{steps}$ остается в разумных пределах для обеспечения численной устойчивости.

Функция потерь

Общая функция потерь LF-PINN включает те же три компонента, что и классическая PINN (Vasilyev, Tarkhov, 2014):

$$L_{total} = \lambda_{PDE} \cdot L_{PDE} + \lambda_{BC} \cdot L_{BC} + \lambda_{IC} \cdot L_{IC}$$

Однако вычисление каждого компонента отличается от классического подхода. Остаток PDE вычисляется путем прямой подстановки предсказанного решения $u_\theta(x, t_{end})$ и его производных в исходное дифференциальное уравнение:

$$L_{PDE} = \frac{1}{N} \sum \left| \frac{\partial u_\theta}{\partial t} - F(u_\theta(x, t), x, t) \right|^2$$

где производные вычисляются через автоматическое дифференцирование по выходу прямого прохода модели. Важное отличие от классической PINN (Antonov, Tarkhov, Vasilyev, 2018; Raissi, Perdikaris, Karniadakis, 2019; Tarkhov, Vasilyev, 2019; Vasilyev, Tarkhov, 2014) заключается в том, что производные вычисляются только один раз для финального решения, а не на каждом временном шаге внутри прямого прохода (Tarkhov, Lazovskaya, Malykhina, 2023). Граничные условия L_{BC} проверяются путем подстановки граничных точек в модель, а начальные условия L_{IC} — путем проверки соответствия модели заданным начальным данным при $t = 0$.



Результаты

Экспериментальные результаты показывают, что предложенная модель с 13 параметрами достигает точности, превосходящей или сопоставимой с классической PINN, имеющей в пять раз больше параметров. При этом наблюдаются существенные различия в поведении моделей в зависимости от типа уравнения. Сравнение проводилось в условиях фиксированного времени обучения (60 секунд), что отражает практические ограничения при получении решения за ограниченное время. Сводные результаты сравнения приведены в табл. 1.

Временной бюджет выбран на основании анализа кривых обучения: для всех рассматриваемых уравнений функция потерь обеих моделей выходит на плато в пределах отведенного времени. Дальнейшее увеличение бюджета не приводит к существенному улучшению точности ни одной из моделей. Таким образом, выбранные условия обеспечивают сходимость обеих моделей и корректность сравнения.

Таблица 1 / Table 1

Сравнительные результаты LF-PINN и Classical PINN для различных типов уравнений (среднее $\pm$ стандартное отклонение по 5 прогонам)

Comparative results of LF-PINN and Classical PINN for different equation types (mean $\pm$ standard deviation over 5 runs)

Уравнение / Метод Equation / Method	L2RE	Epochs ($\mu \pm \sigma$)
Heat Equation		
LF-PINN	$1,65 \times 10^{-2} \pm 1,1 \times 10^{-3}$	366 ± 20
Classical PINN	$3,29 \times 10^{-2} \pm 1,4 \times 10^{-2}$	82971 ± 2095
Wave Equation		
LF-PINN	$3,16 \times 10^{-2} \pm 5,7 \times 10^{-4}$	95 ± 6
Classical PINN	$4,05 \times 10^{-2} \pm 1,2 \times 10^{-2}$	65385 ± 582
Reaction-Diffusion		
LF-PINN	$1,24 \times 10^{-1} \pm 2,3 \times 10^{-5}$	130 ± 8
Classical PINN	$1,24 \times 10^{-1} \pm 7,2 \times 10^{-4}$	79700 ± 2875
Burgers Equation		
LF-PINN	$1,62 \times 10^{-1} \pm 1,4 \times 10^{-3}$	150 ± 0
Classical PINN	$2,67 \times 10^{-1} \pm 2,0 \times 10^{-2}$	82574 ± 295

Уравнение теплопроводности

LF-PINN продемонстрировала значительное превосходство с $L2RE = 1,65 \times 10^{-2}$ против $3,29 \times 10^{-2}$ у классической PINN (улучшение на 50%), как показано на рис. 1

Низкая дисперсия результатов ($\sigma = 1,1 \times 10^{-3}$ для LF-PINN против $1,4 \times 10^{-2}$ для Classical) указывает на большую устойчивость метода к выбору начальной инициализации.

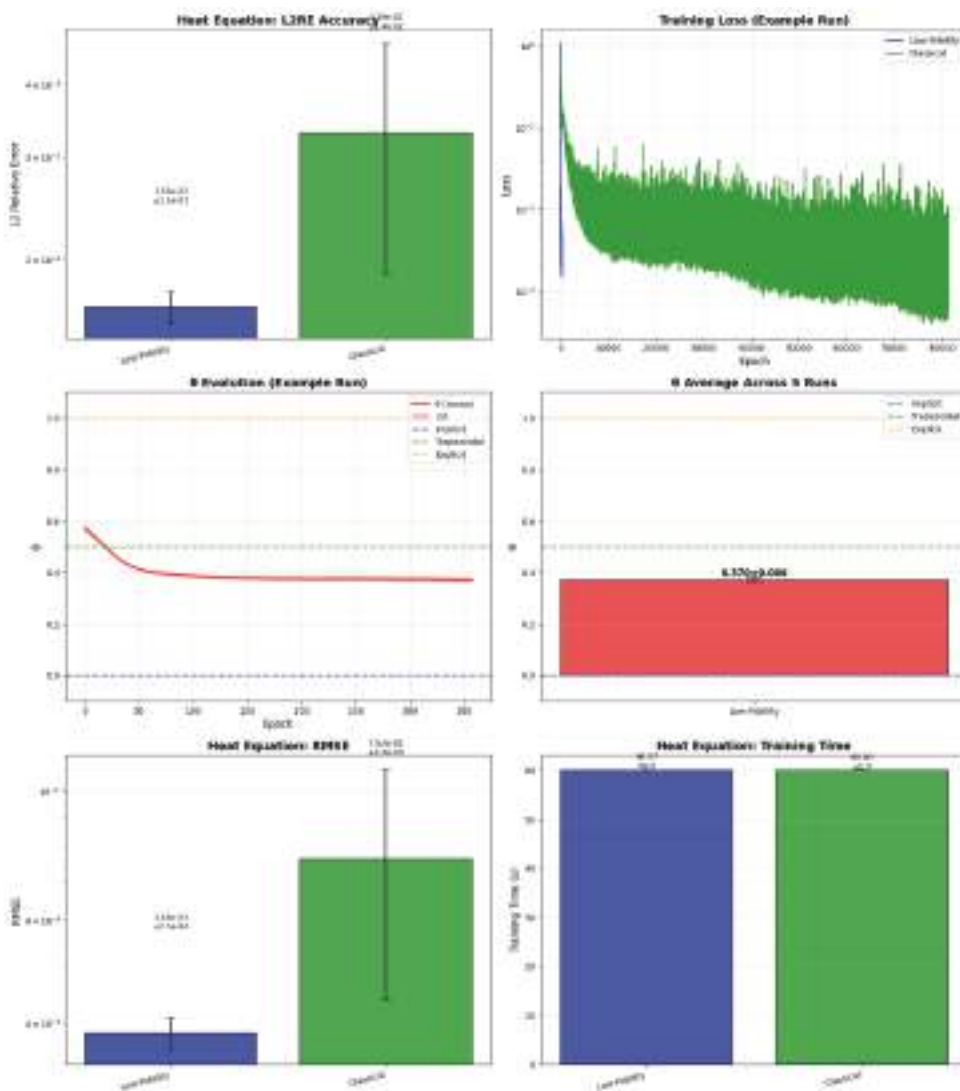


Рис. 1. Сравнение предложенного подхода (Low-Fidelity) с классическим методом на задаче уравнения теплопроводности. (а) Относительная ошибка L2RE. (б) Кривая обучения (Training Loss) для одного репрезентативного прогона. (с) Динамика параметра θ в ходе обучения. (д) Среднее значение θ по 5 запускам. (е) Ошибка RMSE. (ф) Время обучения

Fig. 1. Comparison of the proposed approach (LF-PINN) and Classical PINN on the Heat equation. (a) L2 Relative Error. (b) Training loss curve for a single representative run. (c) Evolution of parameter θ during training. (d) Mean value of θ over 5 runs. (e) RMSE. (f) Training time



Волновое уравнение

Для волнового уравнения LF-PINN также показывает лучшую точность $L2RE = 3,16 \times 10^{-2}$ против $4,05 \times 10^{-2}$, но разница менее выраженная, что видно на рис. 2.

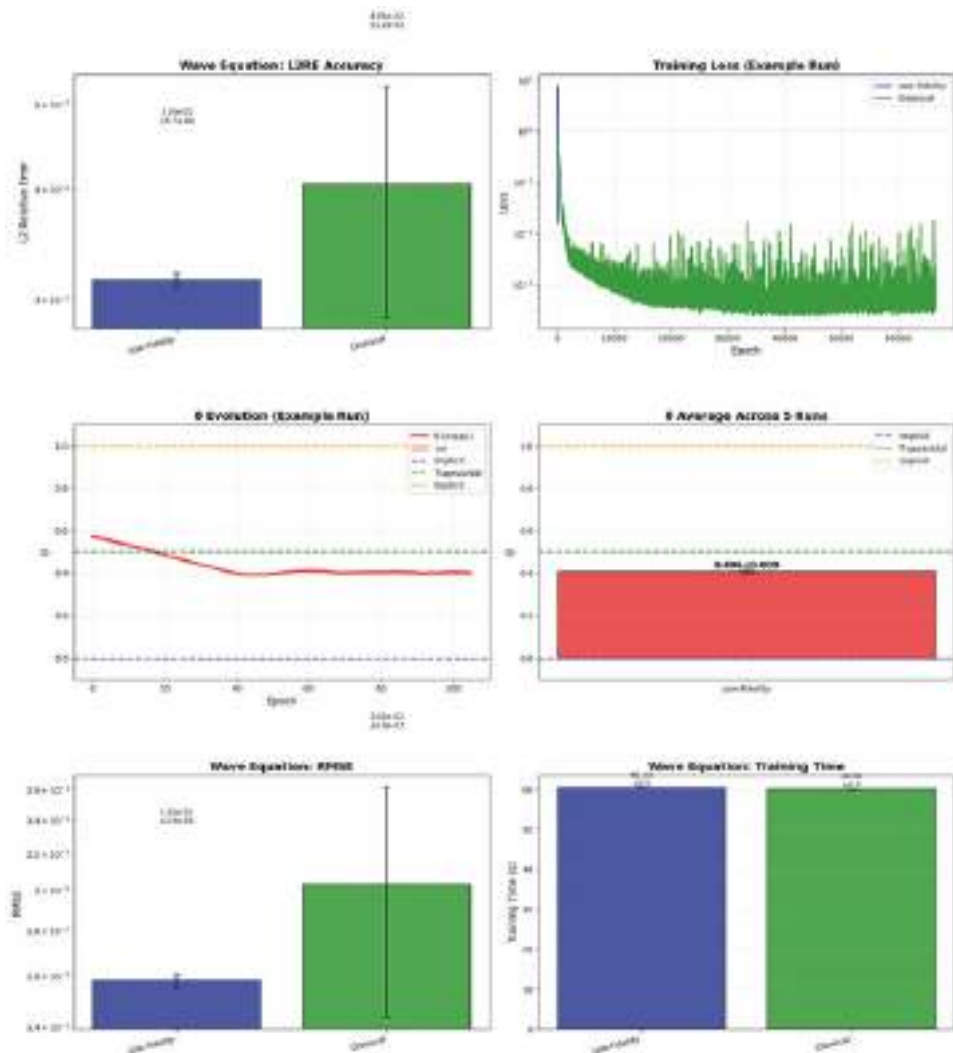


Рис. 2. Сравнение предложенного подхода (Low-Fidelity) с классическим методом на задаче волнового уравнения. (а) Относительная ошибка L2RE. (б) Кривая обучения (Training Loss) для одного репрезентативного прогона. (с) Динамика параметра θ в ходе обучения. (д) Среднее значение θ по 5 запускам. (е) Ошибка RMSE. (ф) Время обучения

Fig. 2. Comparison of the proposed approach (LF-PINN) and Classical PINN on the Wave equation. (a) L2 Relative Error. (b) Training loss curve for a single representative run. (c) Evolution of parameter θ during training. (d) Mean value of θ over 5 runs. (e) RMSE. (f) Training time



Уравнение реакции-диффузии

На этой задаче, как иллюстрирует рис. 3, обе модели показали практически идентичное качество $L2RE \approx 1,24 \times 10^{-1}$.

Высокая относительная ошибка объясняется наличием крутого фронта в решении (характерная ширина фронта $\sqrt{(2 \cdot D)} \approx 0,14$), который сложно аппроксимировать для обеих архитектур при ограниченном количестве точек коллокации.

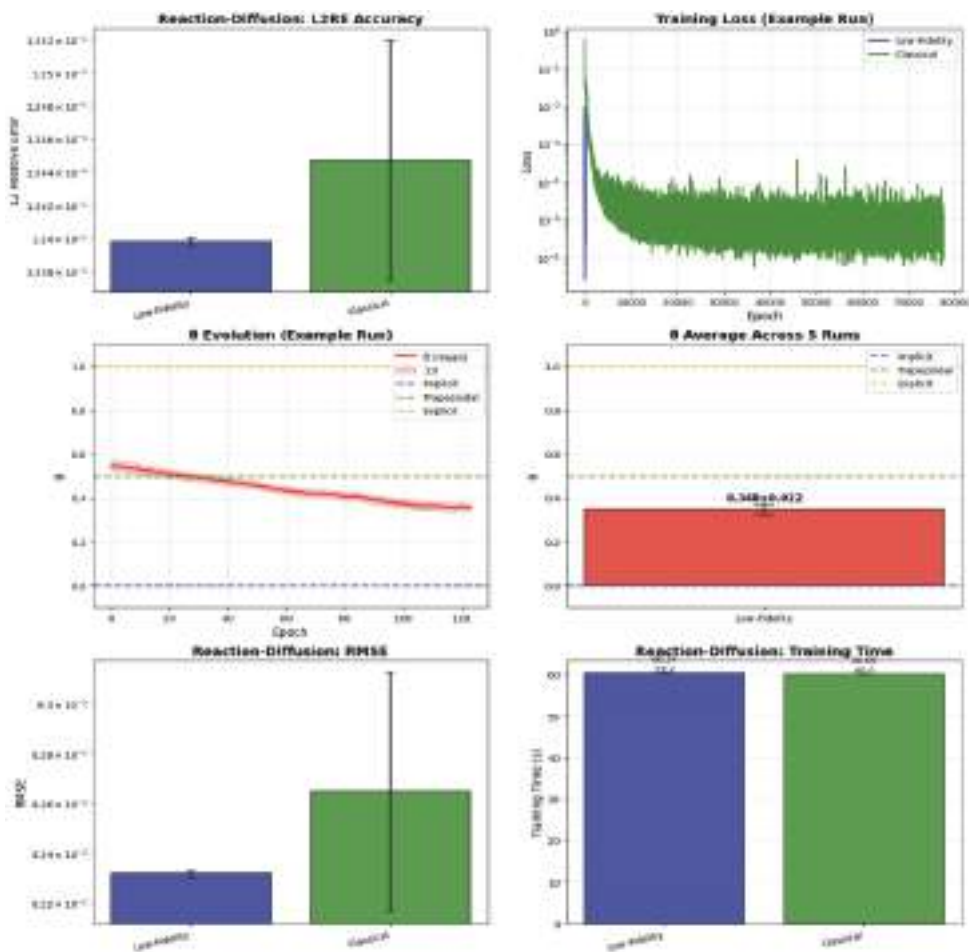


Рис. 3. Сравнение предложенного подхода (Low-Fidelity) с классическим методом на задаче уравнения реакции-диффузии. (а) Относительная ошибка L2RE. (б) Кривая обучения (Training Loss) для одного репрезентативного прогона. (с) Динамика параметра θ в ходе обучения. (д) Среднее значение θ по 5 запускам. (е) Ошибка RMSE. (ф) Время обучения

Fig. 3. Comparison of the proposed approach (LF-PINN) and Classical PINN on the Reaction-Diffusion equation. (a) L2 Relative Error. (b) Training loss curve for a single representative run. (c) Evolution of parameter θ during training. (d) Mean value of θ over 5 runs. (e) RMSE. (f) Training time



Уравнение Бюргерса

На наиболее сложной задаче с сильной нелинейностью (Jagtap, Kawaguchi, Karniadakis, 2020) LF-PINN снова превосходит классический подход: $L2RE = 1,62 \times 10^{-1}$ против $2,67 \times 10^{-1}$ (улучшение на 39%), как демонстрирует рис. 4.

Нелинейный конвективный член $u \cdot \partial u / \partial x$ создает формирование ударных волн и требует особенно аккуратного численного подхода (Cuomo et al., 2022; Jagtap, Kawaguchi, Karniadakis, 2020). Адаптивный выбор схемы через θ -net позволяет LF-PINN эффективно справляться с этими особенностями.

Интересно, что среднее значение $\theta \approx 0,20$ — наименьшее среди четырех рассмотренных задач — показывает предпочтение более явных схем.

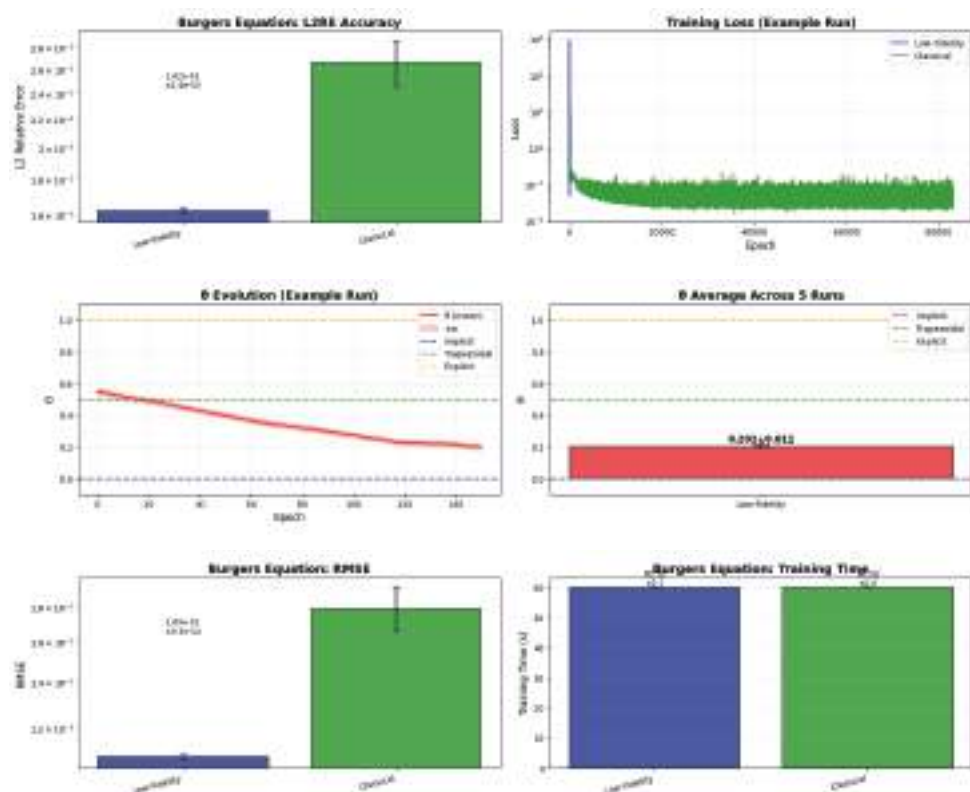


Рис. 4. Сравнение предложенного подхода (Low-Fidelity) с классическим методом на задаче уравнения Бюргерса. (a) Относительная ошибка L2RE. (b) Кривая обучения (Training Loss) для одного репрезентативного прогона. (c) Динамика параметра θ в ходе обучения. (d) Среднее значение θ по 5 запускам. (e) Ошибка RMSE. (f) Время обучения

Fig. 4. Comparison of the proposed approach (LF-PINN) and Classical PINN on Burgers equation. (a) L2 Relative Error. (b) Training loss curve for a single representative run. (c) Evolution of parameter θ during training. (d) Mean value of θ over 5 runs. (e) RMSE. (f) Training time



Обсуждение результатов

Ключевое преимущество LF-PINN заключается в том, что высокая точность достигается при малом числе параметров. В проведенных экспериментах даже компактная архитектура (один скрытый слой с 2 нейронами) обеспечивает качество, сопоставимое или превосходящее классические PINN (Antonov, Tarkhov, Vasilyev, 2018; Raissi, Perdikaris, Karniadakis, 2019; Tarkhov, Vasilyev, 2019; Vasilyev, Tarkhov, 2014), требующие существенно большего числа параметров, что обусловлено использованием встроенной численной схемы.

Важным наблюдением является значительно меньшая дисперсия результатов для LF-PINN по сравнению с классической моделью (Antonov, Tarkhov, Vasilyev, 2018; Raissi, Perdikaris, Karniadakis, 2019; Tarkhov, Vasilyev, 2019; Vasilyev, Tarkhov, 2014). Это указывает на то, что архитектура LF-PINN более устойчива к выбору начальной инициализации весов и менее подвержена попаданию в локальные минимумы. Стандартное отклонение L2RE для LF-PINN в среднем на порядок меньше, чем для классической PINN.

Адаптивность параметра θ демонстрируется его изменением в зависимости от типа уравнения и локальных свойств решения. Для уравнения теплопроводности среднее значение $\theta = 0,37$, что ближе к методу трапеций и обеспечивает баланс между точностью и устойчивостью. Для волнового уравнения $\theta = 0,41$, для уравнения реакции-диффузии $\theta = 0,35$, а для уравнения Бюргерса $\theta = 0,20$, что свидетельствует о предпочтении более явных методов. Эта адаптивность является ключевым преимуществом подхода, так как не требует ручной настройки численной схемы для каждого типа уравнения. Сводное сравнение относительной ошибки по всем четырем типам уравнений приведено на рис. 5.

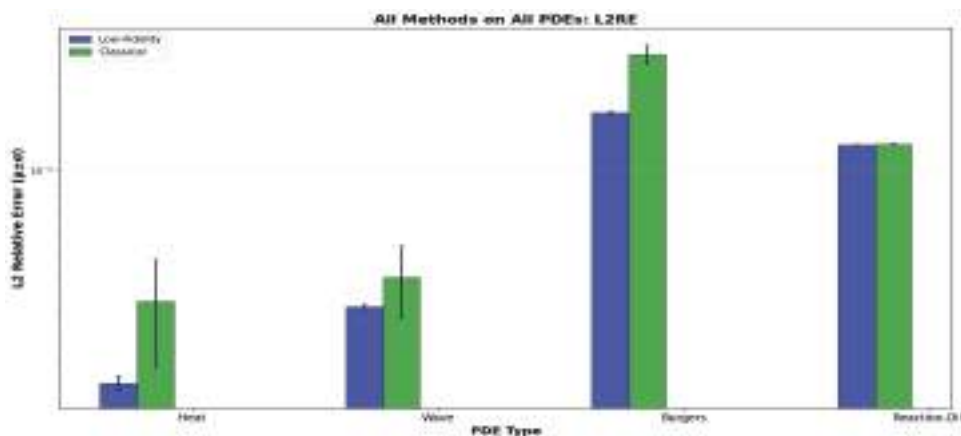


Рис. 5. Сравнение относительной ошибки L2RE ($\mu \pm \sigma$ по 5 запускам) предложенного Low-Fidelity метода с классической реализацией PINN на четырех типах уравнений в частных производных

Fig. 5. Comparison of L2 Relative Error ($\mu \pm \sigma$ over 5 runs) between the proposed Low-Fidelity PINN and Classical PINN across four types of partial differential equations



Заключение

Разработанная архитектура LF-PINN (Lazovskaya, Tarkhov, 2026; Tarkhov, Lazovskaya, Malykhina, 2023) демонстрирует ряд существенных преимуществ по сравнению с классическими PINN (Antonov, Tarkhov, Vasilyev, 2018; Raissi, Perdikaris, Karniadakis, 2019; Tarkhov, Vasilyev, 2019; Vasilyev, Tarkhov, 2014):

- **Вычислительная эффективность.** Интеграция численной схемы непосредственно в архитектуру нейронной сети позволяет избежать многократного вычисления производных высокого порядка через автоматическое дифференцирование. Предложенная модель требует существенно меньшего числа эпох для сходимости — от 95 до 366 против 65 000–83 000 у классической PINN, что делает метод практически применимым для реальных задач.
- **Адаптивность.** Параметрическая сеть θ -net автоматически выбирает оптимальный баланс между явной и неявной схемами для каждой точки пространственно-временной области (Lazovskaya, Tarkhov, 2026). Это позволяет модели адаптироваться к локальным особенностям решения, таким как крутые градиенты или осцилляции. В отличие от традиционных численных методов, где выбор схемы делается вручную и фиксируется на всю область, LF-PINN может использовать различные схемы в различных частях области.
- **Компактность модели.** Использование всего 13 параметров в θ -net предотвращает переобучение и обеспечивает хорошую обобщающую способность. В отличие от классической PINN, которая требует глубоких сетей с десятками или сотнями параметров (Karniadakis et al., 2021; Raissi, Perdikaris, Karniadakis, 2019), LF-PINN полагается на встроенную физическую информацию через численную схему (Lazovskaya, Tarkhov, 2026; Tarkhov, Lazovskaya, Malykhina, 2023). Это также упрощает интерпретацию модели: вместо изучения сложного нелинейного отображения, θ -net учится только выбору оптимальной численной схемы.
- **Устойчивость обучения.** Значительно меньшая дисперсия результатов по сравнению с классической PINN указывает на то, что метод менее чувствителен к инициализации и гиперпараметрам. Это критически важно для практического применения, где требуется воспроизводимость результатов. В экспериментах стандартное отклонение L2RE для LF-PINN составляло 10^{-5} – 10^{-3} , в то время как для классической PINN 10^{-3} – 10^{-2} , что на 1–2 порядка выше.

Рассмотренные LF-PINN удобно использовать для суррогатного моделирования (Jiang, Zhou, Shao, 2020; Manzoni et al., 2020; Ranftl, von der Linden, 2021; Zhou, Mei, Xu, 2023).

Список источников / References

1. Antonov, V.I., Tarkhov, D.A., Vasilyev, A.N. (2018). Unified approach to constructing the neural network models of real objects. Part 1. *Mathematical Methods in the Applied Sciences*, 41(18), 9244–9251. <https://doi.org/10.1002/mma.5205>



2. Cuomo, S., Di Cola, V.S., Giampaolo, F., Rozza, G., Raissi, M., Piccialli, F. (2022). Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3), Article 88. <https://doi.org/10.1007/s10915-022-01939-z>
3. Hairer, E., Nørsett, S.P., Wanner, G. (1993). *Solving ordinary differential equations I: Nonstiff problems* (2nd rev. ed.). Berlin: Springer. <https://doi.org/10.1007/978-3-540-78862-1>
4. Jagtap, A.D., Kawaguchi, K., Karniadakis, G.E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404, Article 109136. <https://doi.org/10.1016/j.jcp.2019.109136>
5. Jiang, P., Zhou, Q., Shao, X. (2020). *Surrogate model-based engineering design and optimization*. Singapore: Springer. <https://doi.org/10.1007/978-981-15-0731-1>
6. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440. <https://doi.org/10.1038/s42254-021-00314-5>
7. Lazovskaya, T.V., Tarkhov, D.A. (2016). Multilayer neural network models based on grid methods. *IOP Conference Series: Materials Science and Engineering*, 158, Article 012061. <https://doi.org/10.1088/1757-899X/158/1/012061>
8. Lazovskaya, T.V., Tarkhov, D.A. (2026). Numerics as neural networks: A compact low-fidelity layer for multi-fidelity modelling. In: *Advances in neural computation, machine learning, and cognitive research IX* (pp. 381–391). Cham: Springer. https://doi.org/10.1007/978-3-032-07690-8_31
9. Lazovskaya, T.V., Tarkhov, D.A., Vasilyev, A.N. (2018). Multi-layer solution of heat equation. In: B. Kryzhanovsky, W. Dunin-Barkowski, V. Redko (Eds.), *Advances in neural computation, machine learning, and cognitive research* (Studies in Computational Intelligence, Vol. 736, pp. 17–22. Cham: Springer. https://doi.org/10.1007/978-3-319-66604-4_3
10. Manzoni, L., Papetti, D.M., Cazzaniga, P., Spolaor, S., Mauri, G., Besozzi, D., Nobile, M.S. (2020). Surfing on fitness landscapes: A boost on optimization by Fourier surrogate modeling. *Entropy*, 22(3), Article 285. <https://doi.org/10.3390/e22030285>
11. Peherstorfer, B. (2019). Multifidelity Monte Carlo estimation with adaptive low-fidelity models. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2), 579–603. <https://doi.org/10.1137/17M1159208>
12. Penwarden, M., Zhe, S., Narayan, A., Kirby, R.M. (2022). Multifidelity modeling for physics-informed neural networks (PINNs). *Journal of Computational Physics*, 451, Article 110844. <https://doi.org/10.1016/j.jcp.2021.110844>
13. Raissi, M., Perdikaris, P., Karniadakis, G.E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
14. Ranftl, S., von der Linden, W. (2021). Bayesian surrogate analysis and uncertainty propagation. *Physical Sciences Forum*, 3(1), Article 6. <https://doi.org/10.3390/psf2021003006>
15. Tarkhov, D.A., Lazovskaya, T.V., Malykhina, G.F. (2023). Constructing physics-informed neural networks with architecture based on analytical modification of numerical methods by solving the problem of modelling processes in a chemical reactor. *Sensors*, 23(2), Article 663. <https://doi.org/10.3390/s23020663>



16. Tarkhov, D.A., Vasilyev, A.N. (2019). *Semi-empirical neural network modeling and digital twins development*. Cambridge, MA: Academic Press.
17. Vasilyev, A.N., Tarkhov, D.A. (2014). Mathematical models of complex systems on the basis of artificial neural networks. *Nonlinear Phenomena in Complex Systems*, 17(3), 327–335.
18. Wang, S., Teng, Y., Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5), A3055–A3081. <https://doi.org/10.1137/20M1318043>
19. Zhou, M., Mei, G., Xu, N. (2023). Enhancing computational accuracy in surrogate modeling for elastic–plastic problems by coupling S-FEM and physics-informed deep learning. *Mathematics*, 11(9), Article 2016. <https://doi.org/10.3390/math11092016>

Информация об авторах

Великоречанин Игорь Анатольевич, студент магистратуры, Московский физико-технический институт (МФТИ), г. Долгопрудный, Российская Федерация, ORCID: <https://orcid.org/0009-0007-7889-6811>, e-mail: veligoran2001@gmail.com

Лазовская Татьяна Валерьевна, кафедра высшей математики, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), г. Санкт-Петербург, Российская Федерация, ORCID: <https://orcid.org/0000-0002-3324-6213>, e-mail: tatianala@list.ru

Тархов Дмитрий Альбертович, профессор кафедры высшей математики, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), г. Санкт-Петербург, Российская Федерация, ORCID: <https://orcid.org/0000-0002-9431-8241>

Information about the authors

Igor A. Velikorechanin, Masters student, Moscow Institute of Physics and Technology, Dolgoprudny, Russian Federation, ORCID: <https://orcid.org/0009-0007-7889-6811>, e-mail: veligoran2001@gmail.com

Tatiana V. Lazovskaya, Department of Higher Mathematics, Peter the Great Saint Petersburg Polytechnic University (SPbPU), Saint Petersburg, Russian Federation, ORCID: <https://orcid.org/0000-0002-3324-6213>, e-mail: tatianala@list.ru

Dmitry A. Tarkhov, Professor, Department of Higher Mathematics, Peter the Great Saint Petersburg Polytechnic University (SPbPU), Saint Petersburg, Russian Federation, ORCID: <https://orcid.org/0000-0002-9431-8241>

Вклад авторов

Великоречанин И.А. — программная реализация архитектуры LF-PINN; проведение вычислительных экспериментов; визуализация результатов.

Лазовская Т.В. — постановка задачи; разработка методологии; математическая постановка и теоретический анализ; научное руководство.

Тархов Д.А. — участие в подготовке текста статьи; критическое редактирование и доработка рукописи.

Все авторы приняли участие в обсуждении результатов и согласовали окончательный текст рукописи.



Contribution of the authors

Igor A. Velikorechanin — software implementation of the LF-PINN architecture; computational experiments; visualization of results.

Tatiana V. Lazovskaya — conceptualization of the study; development of the methodology; mathematical formulation and formal analysis; supervision.

Dmitry A. Tarkhov — participation in manuscript preparation; critical revision and editing of the text.

All authors participated in the discussion of the results and approved the final text of the manuscript.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest

The authors declare no conflict of interest.

Поступила в редакцию 02.04.2026

Поступила после рецензирования 10.08.2026

Принята к публикации 10.08.2026

Опубликована 30.09.2026

Received 2026.04.02

Revised 2026.08.10

Accepted 2026.08.10

Published 2026.09.30