

# 1 Аннотация

Данный доклад фокусируется на методе численного интегрирования функции одной переменной с применением нейросетевого подхода. Суть нейросетевого подхода сводится к обучению нейросетевой модели аппроксимировать подынтегральную функцию и дальнейшему использованию параметров обученной нейросетевой модели для численного интегрирования данной функции на любом отрезке в области определения функции.

Применение нейросетевого подхода должно позволить повысить точность вычисления интегралов с большим количеством переменных. При вычислении значения численных интегралов функций многих переменных большинством классических численных методов (например, формула Симпсона, метод трапеций и др.) погрешность значения результатов увеличивается тем сильнее, чем больше число переменных подынтегральной функции. Растет также и сложность вычислений. Применение нейросетевого подхода должно решить данные проблемы, так как основные вычислительные ресурсы в нем тратятся на обучение сравнительно небольшой модели – многослойного перцептрона с одним скрытым слоем – один раз, а затем данную модель можно использовать для получения значений численных интегралов данной функции на разных отрезках много раз. При этом возможность гибкой настройки гиперпараметров обучаемой модели, теоретически позволяет достичь требуемой в прикладной задаче точности.

Данный доклад основан на бакалаврской работе, целью которой является разработка программного пакета численного интегрирования на основе нейросетевого подхода. Данный программный пакет должен будет применяться для решения физической задачи описания поведения лёгких и тяжёлых мезонов. Основу данной задачи будет составлять моделирование процессов образования и свойств частиц для эксперимента NICA. В контексте данной задачи, образование и взаимодействие кварк-антикварковых пар описывается при помощи формул, включающих вычисление определенных интегралов функций многих переменных, а также двух-, трех- и более кратных интегралов.

На данный момент автором доклада разработана программа, которая позволяет вычислять численным методом значения интегралов функции одной переменной (в основном трансцендентных функций).

**Ключевые слова:** численное интегрирование, нейронные сети, легкие и тяжелые мезоны.

## 2 Введение

Задача получения значения определенного интеграла функции встречается во многих как теоретических, так и прикладных вычислениях и сопряжена с рядом проблем, одна из которых – не для всех функций существует аналитическая форма интеграла, а если и существует, получение её может быть очень сложно. Для решения этой проблемы существуют численные методы интегрирования, которые позволяют определить с некоторой заданной погрешностью значение определенного интеграла функции не зная аналитической формы интеграла.

Но здесь возникает новая проблема: так называемое «проклятие размерности» приводит к значительному увеличению вычислительной сложности и снижению точности вычисления значений определенных интегралов при увеличении числа переменных подынтегральной функции. Одним из методов решения данной проблемы предлагается использовать нейросетевой подход для численного интегрирования функций многих переменных, который должен позволить получить более высокую точность и снизить сложность

вычислений.

Нейросетевой подход для численного интегрирования подразумевает собой использование нейронных сетей для получения численных значений интегралов, в том числе интегралов функций многих переменных. Анализ различных способов применения нейросетевой модели в данной задаче сам по себе может стать отдельным исследованием. Основой данной работы послужил подход к использованию нейросетевых моделей для численного интегрирования, описанный в статье «*Using neural networks for fast numerical integration and optimization*» за авторством *Lloyd et al.* [1], однако это, как уже сказано, не единственный вероятный способ применения нейросетевой модели в задаче численного интегрирования.

Данная работа разделена на несколько частей. Сначала будет вкратце описан нейросетевой подход к численному интегрированию. Затем будет приведена физическая задача описания поведения лёгких и тяжелых мезонов, которая также послужила вдохновением для данной работы, так как в рамках данной задачи возникает необходимость вычисления интегралов функций многих переменных, например функции с 7-ю, 10-ю и более переменными. Наконец будет продемонстрировано применение разработанной на данный момент реализации нейросетевого подхода к одному из интегралов вышеупомянутой задачи. В заключении будет вкратце описаны планы дальнейшего развития проекта.

### 3 Нейросетевой подход к численному интегрированию

Описание теоретической основы нейросетевого подхода можно свести к следующим основным тезисам:

- Для некоторой подынтегральной функции  $f(x)$  можно создать функцию-оценщик (*estimator function*)  $\hat{f}(x)$  – нейросеть с архитектурой *MLP* (*Multi-Layered Perceptron* – многослойный перцептрон) с функцией активации *логистическая сигмоида* на единственном скрытом слое.
- $\hat{f}(x)$  можно обучить аппроксимировать подынтегральную функцию  $f(x)$  с некоторой заданной погрешностью  $\epsilon$ .
- $\hat{f}(x)$  можно описать в аналитическом виде.
- Интегрируя аналитический вид  $\hat{f}(x)$  можно вывести формулу для численного интегрирования функции  $f(x)$ , которая зависит от параметров нейронной сети (весов и смещений) и границ интегрирования:

$$\hat{I}(f) = b^{(2)} \prod_{i=1}^n (\beta_i - \alpha_i) + \sum_{j=1}^k w_j^{(2)} \left[ \prod_{i=1}^n (\beta_i - \alpha_i) + \frac{\Phi_j}{\prod_{i=1}^n w_{ij}^{(1)}} \right], \quad (1)$$

$$\Phi_j = \sum_{r=1}^{2^n} \xi_r \text{Li}_n(-\exp[-b_j^{(1)} - \sum_{i=1}^n w_{ij}^{(1)} l_{i,r}]), \quad (2)$$

$$\xi_r = \prod_{d=1}^n (-1)^{[r/2^{n-d}]}, \quad (3)$$

$$l_{i,r} = \begin{cases} \alpha_i, & \text{если } [r/2^{n-d}] \% 2 = 0 \\ \beta_i, & \text{если } [r/2^{n-d}] \% 2 \neq 0, \end{cases} \quad (4)$$

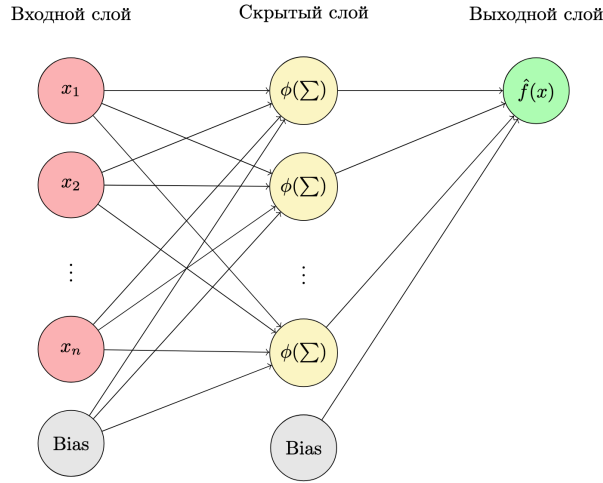


Рисунок 1. Архитектура  $MLP$  описываемой нейросети

где  $b^{(1)}, b^{(2)}, w^{(1)}, w^{(2)}$  – веса нейросети,  $\alpha_i, \beta_i, 1 \leq i \leq n$  – границы интегрирования при  $n$  – число аргументов функции  $f(x)$ .

Упомянутая архитектура  $MLP$  заслуживает отдельного рассмотрения.  $MLP$  можно описать как *полносвязную* нейросеть с несколькими слоями, то есть каждый нейрон слоя  $l_i$  связан с каждым нейроном слоя  $l_{i+1}$ . В данном конкретном случае будет использоваться один скрытый слой нейросети с функцией активации *логистическая сигмоида*:

$$\phi(\bullet) = \frac{1}{1 + \exp(-\bullet)}. \quad (5)$$

Входной и выходной слои не будут иметь функции активации. Входной слой будет включать такое количество входных нейронов, от скольки переменных подынтегральная функция  $f(x)$  зависит. Размер скрытого слоя необходимо подбирать в процессе обучения, а выходной слой содержит один нейрон – значение  $\hat{f}(x)$  (см. рис. 1).

Таким образом, описанная нейросеть (см. рис. 1), имеет аналитический вид:

$$\hat{f}(x) = b^{(2)} + \sum_{j=1}^k w_j^{(2)} \phi(b_j^{(1)} + \sum_{i=1}^n w_{ji}^{(1)} x_i), \quad (6)$$

где  $b^{(1)}, b_j^{(2)}, w_{ji}^{(1)}, w_j^{(2)}$  – веса нейросети,  $k$  – число нейронов скрытого слоя,  $n$  – число нейронов входного слоя,  $\phi$  – логистическая сигмоида (5).

Данный аналитический вид  $\hat{f}(x)$ , как описано выше, можно интегрировать и на его основе получить формулу численного интегрирования на основе параметров нейросети (1-4). Для интегрирования аналитического вида нейросети (5) можно воспользоваться удобным соответствием логистической сигмоиды полилогарифму нулевого порядка:

$$-Li_0(-\exp(\bullet)) = \frac{1}{1 + \exp(-\bullet)} \quad (7)$$

который достаточно легко интегрировать аналитически. На основе аналитического интеграла нейросети (5), заменив сигмоиду полилогарифмом, как показано выше (7), можно получить формулы (1-4). Нейросетевой подход, таким образом, сводится к следующим шагам:

1. Создание набора данных на основе значений подынтегральной функции и её переменных для обучения нейросетевой модели.

2. Обучение и тестирование нейросетевой модели (5) аппроксимировать подынтегральную функцию.
3. Извлечение параметров нейросетевой модели и применение их в формулам (1-4) для получения численного значения интеграла.

## 4 Описание физической задачи

Основная задача: моделирование процессов образования и свойств частиц для эксперимента *NICA*. В рамках задачи необходимо создать модель, описывающую основные свойства как лёгких, так и тяжелых мезонов. Для этого будет использоваться модель с нелокальными взаимодействиями [2][3].

Суть модели состоит в том, что образование мезонов (как кварк-антикварковых пар) сопровождается обменом глюонов, однако обмен глюонов не описывается с помощью суммирования диаграмм, так как теория возмущений неприменима (разложение по малой константе связи невозможно) (см. рис. 2).

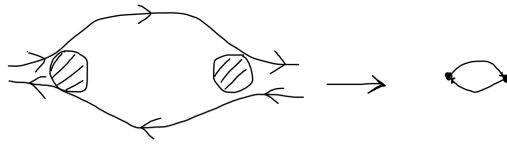


Рисунок 2. Взаимодействие с обменом глюонов

Назовём  $\bullet$  форм-фактором и будем использовать модель Гауссовского типа

$$F(p^2) = \exp\left[\frac{-p^2}{\lambda^2}\right], \quad (8)$$

так что размер форм-фактора будет определяться описанной  $F(p^2)$ . Для описания частиц нам требуется знать массы, константы распадов и процесс распадов. Обозначим взаимодействие:



и тогда можно выразить:

$$\bullet \text{ (loop) } + \bullet \text{ (loop) } + \bullet \text{ (loop) } + \dots = \frac{1}{1 - \bullet \text{ (loop) }}. \quad (9)$$

$$\bullet \text{ (loop) } \rightarrow \int_0^\infty \frac{dp}{2\pi^4} F(p^2) \frac{1}{u_1 u_2} = \int_0^\infty \frac{dp}{2\pi^4} F(p^2) \frac{1}{[(p+q_1)^2 + m_1^2][(p+q_2)^2 + m_2^2]}. \quad (10)$$

Для расчета данных интегралов используются формулы:

$$\frac{1}{(p+q_1)^2 + m_1^2} = \int_0^\infty dt \{ \exp(-t[(p+q_1)^2 + m_1^2]) \} \quad (11)$$

$$F(p^2) = \int_0^\infty ds \{ \exp(-sp^2) F(s) \} \quad (12)$$

Случаи взаимодействия более чем двух мезонов описываются трех- и четырёхкратными интегралами (см. рис. 3).

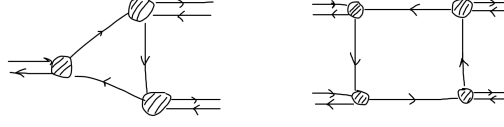


Рисунок 3. Взаимодействие большого числа мезонов

Амплитуды конкретных распадов и их квадраты будут вычисляться с помощью программы аналитических вычислений FORM. Основной проблемой данной задачи является исчисление интегралов функций многих переменных.

## 5 Применение нейросетевого подхода к физической задаче

На основе описанного ранее нейросетевого подхода была разработана программная реализация на языке *Python* с применением библиотек *PyTorch*, *NumPy*, *mpmath* [4][5][6][7]. В вышеописанной задаче сформулирован повторный интеграл (13), который применяется в различных вычислениях в рамках задачи. Данный интеграл был одним из способов тестирования работоспособности разрабатываемого нейросетевого подхода.

$$\int_0^1 d\alpha \{ \alpha^a (1 - \alpha)^b \} \int_0^\infty dt \{ \frac{t^m}{(1+t)^n} F[z_0] \} \equiv I(a, b, m, n; F[z_0]), \quad (13)$$

$$F[z_0] = \exp[-2z_0], \quad (14)$$

$$z_0 = tD + \frac{t}{1+t} R^2, \quad (15)$$

$$D = \alpha_1(b_1^2 P^2 + m_1^2) + \alpha_2(b_2^2 P^2 + m_2^2), \quad (16)$$

$$R^2 = (\alpha_1^2 b_1^2 + \alpha_2^2 b_2^2 + 2\alpha_1 \alpha_2 b_1 b_2) P^2, \quad (17)$$

$$b_1 = -\frac{m_1}{m_1 + m_2}, \quad (18)$$

$$b_2 = \frac{m_2}{m_1 + m_2}. \quad (19)$$

Предварительно, с использованием языка *FORTRAN*, были получены значения выражения (13) с различными параметрами  $a, b, m, n$ . Также, для выражения (13) заданы неизменные значения  $m_1 = m_2 = 0.7083333$ ,  $P^2 = -1.665046$ .

На первом этапе тестирования были вычислены интегралы по  $d\alpha$  и по  $dt$  по отдельности. Точное значение интеграла функции  $\alpha^a (1 - \alpha)^b$  было легко получить аналитически, поэтому имелась возможность сравнивать значения, получаемые с использованием нейросетевого подхода и точные значения. Точное значение интеграла  $\frac{t^m}{(1+t)^n} F[z_0]$  получить было сложнее, поэтому в данном случае акцент был на значение функции потерь  $MSE$  на тестовом наборе данных для данного интеграла.

На втором этапе тестирования была предпринята попытка интегрирования выражения (13) по двум переменным. Для этого было выполнено следующее преобразование повторного интеграла:

$$\int_0^1 \alpha^a (1 - \alpha)^b d\alpha \int_0^\infty \frac{t^m}{(1+t)^n} F[z_0] dt = \int_0^1 \int_0^\infty \frac{t^m}{(1+t)^n} F[z_0] \alpha^a (1 - \alpha)^b dt d\alpha. \quad (20)$$

Это позволило обучить нейросетевую модель аппроксимировать функцию  $\frac{t^m}{(1+t)^n} F[z_0] \alpha^a (1 - \alpha)^b$  как функцию двух переменных и воспользоваться нейросетевым подходом для получения значения численного интеграла (13).

## РЕЗУЛЬТАТЫ

## 6 Заключение и дальнейшие планы

В настоящее время ведется работа над программной реализацией нейросетевого подхода для функций двух и более переменных, которые будут апробированны на более сложных выражениях из задачи описания свойств мезонов, включающих интегралы функций двух и более переменных. Финальной целью данной работы должен стать программный пакет на языке *Python*, позволяющий применять нейросетевой подход для интегрирования сложных функций как в данной задаче, так и в других задачах как теоретических, так и прикладных.

## Список литературы

1. *Lloyd S., Irani R. A., Ahmadi M.* Using neural networks for fast numerical integration and optimization // IEEE Access. — 2020. — Т. 8. — С. 84519—84531.
2. *Blaschke D., Grigorian H., Kalinovsky Y. L.* Meson form-factor scheme for the chiral lagrangian approach to  $J/\psi$  breakup cross sections motivated by a relativistic quark model // Physics of Particles and Nuclei Letters. — 2012. — Т. 9. — С. 7—17.
3. *Costa P., Ruivo M. C., Kalinovsky Y. L.* Pseudoscalar neutral mesons in hot and dense matter // Physics Letters B. — 2003. — Т. 560, № 3/4. — С. 171—177.
4. Документация по Python. — 2024. — URL: <https://docs.python.org/3/>; (Дата обращения: 11 декабря 2024).
5. Документация по PyTorch. — 2024. — URL: <https://pytorch.org/docs/stable/index.html>; (Дата обращения: 11 декабря 2024).
6. Документация по NumPy. — 2024. — URL: <https://numpy.org/doc/>; (Дата обращения: 11 декабря 2024).
7. Документация по Mpmath. — 2024. — URL: <https://mpmath.org/doc/current/index.html>; (Дата обращения: 11 декабря 2024).