

The background features a complex, abstract composition. It includes several overlapping circles in shades of gray, yellow, and teal. A network of thin, curved lines connects various points, some of which are marked with small colored dots. Faint mathematical symbols and equations are scattered throughout, such as $\gamma = 10^\circ$, $\gamma = 7^\circ$, $\gamma = 5^\circ$, $\gamma = 12^\circ$, $\gamma = 15^\circ$, $\gamma = 18^\circ$, $\gamma = 21^\circ$, $\gamma = 24^\circ$, $\gamma = 27^\circ$, $\gamma = 30^\circ$, $\gamma = 33^\circ$, $\gamma = 36^\circ$, $\gamma = 39^\circ$, $\gamma = 42^\circ$, $\gamma = 45^\circ$, $\gamma = 48^\circ$, $\gamma = 51^\circ$, $\gamma = 54^\circ$, $\gamma = 57^\circ$, $\gamma = 60^\circ$, $\gamma = 63^\circ$, $\gamma = 66^\circ$, $\gamma = 69^\circ$, $\gamma = 72^\circ$, $\gamma = 75^\circ$, $\gamma = 78^\circ$, $\gamma = 81^\circ$, $\gamma = 84^\circ$, $\gamma = 87^\circ$, $\gamma = 90^\circ$, $\gamma = 93^\circ$, $\gamma = 96^\circ$, $\gamma = 99^\circ$, $\gamma = 102^\circ$, $\gamma = 105^\circ$, $\gamma = 108^\circ$, $\gamma = 111^\circ$, $\gamma = 114^\circ$, $\gamma = 117^\circ$, $\gamma = 120^\circ$, $\gamma = 123^\circ$, $\gamma = 126^\circ$, $\gamma = 129^\circ$, $\gamma = 132^\circ$, $\gamma = 135^\circ$, $\gamma = 138^\circ$, $\gamma = 141^\circ$, $\gamma = 144^\circ$, $\gamma = 147^\circ$, $\gamma = 150^\circ$, $\gamma = 153^\circ$, $\gamma = 156^\circ$, $\gamma = 159^\circ$, $\gamma = 162^\circ$, $\gamma = 165^\circ$, $\gamma = 168^\circ$, $\gamma = 171^\circ$, $\gamma = 174^\circ$, $\gamma = 177^\circ$, $\gamma = 180^\circ$.

NQS: Neural-Network Quantum States

Нейросетевые квантовые состояния

Схема доклада

- История вопроса
- Немного о квантовой физике
- Что же такое NQS?
- Про RBM
- Идеи метода NQS

Откуда взялись NQS?

Первое упоминание — статья Carleo, Troyer (2017) о **решении квантовой задачи многих тел при помощи искусственных нейронных сетей** (*Carleo G., Troyer M. Solving the quantum many-body problem with artificial neural networks // Science. – 2017. – T. 355. – №. 6325. – С. 602-606.*)

Ранее нейронные сети использовались для других задач: распознавание изображений, обработка естественного языка и другое.

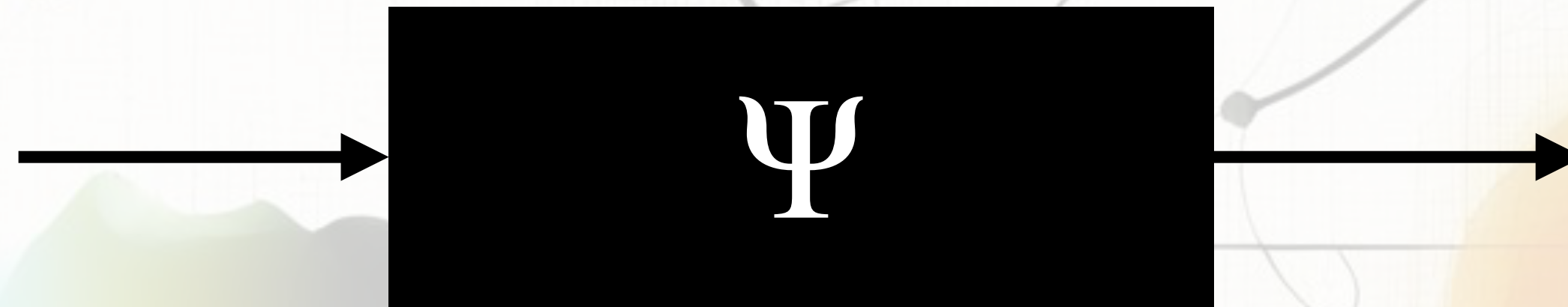
Теперь нейронные сети нашли применение при решении задач статистической физики.

Мотивация использования НС: экспоненциальная сложность волновой функции многих тел (об этом далее).

Немного о квантовой физике

Ψ — квантовая волновая функция. ~ наиболее неуловимый объект к.ф.

Конфигурация
многих тел S



Вычислительный чёрный ящик

Фаза и амплитуда
в соответствии с Ψ

Например!

$$|S_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; |S_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Psi = \frac{1}{\sqrt{2}} |S_1\rangle + \frac{1}{\sqrt{2}} |S_2\rangle$$

Спин вверх;
Спин вниз

Ψ -функции выражают **СОСТОЯНИЕ** квантовых систем.
Представляют собой векторы в Гильбертовом пространстве.

Базис векторов: «чистые» потенциально наблюдаемые состояния систем.
(Вектора обыкновенным образом раскладываются по базису)

$$P(S_1) = a_{S_1}^2 = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

Амплитуды разложения — вероятность встретить каждое из «чистых» состояний!

$$|\Psi\rangle = a_1 \cdot |S_1\rangle + a_2 \cdot |S_2\rangle + \dots = \sum_n a_n \cdot |S_n\rangle$$

Амплитуды разложения

$$\sum_n |a_n|^2 = 1$$

Но не все «чистые» состояния разрешены природой! —>

Немного о квантовой физике

Уравнение Шрёдингера — основное уравнение к.ф.

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = \hat{H} |\Psi\rangle$$

\hat{H} — *Гамильтониан*, оператор энергии

Для системы из N спинов \hat{H} имеет размерность $2^N \times 2^N$

Природой разрешены только те «чистые» состояния, которые являются собственными векторами \hat{H} !

$$\hat{H} |\Psi\rangle = E |\Psi\rangle$$

Допустимые Ψ — собственные векторы \hat{H} .

А, значит, собственные значения \hat{H} — значения **энергий**, которые соответствуют этим состояниям.

Основное состояние квантовой системы — такое состояние, при котором $|\Psi\rangle$ соответствует минимальному собственному значению E оператора \hat{H} .

Основное состояние — наиболее вероятное;
Квантовая система всегда стремится в своё основное состояние!

Нахождение основного состояния — одна из центральных задач квантовой физики

Про RBM

RBM — Restricted Boltzmann Machine

(Ограниченная машина Больцмана).
Нейроны одного слоя не соединены

Вероятностная графическая модель.

Впервые построена в 1986 году (P.Smolensky), начала использоваться в середине 2000-х годов, когда Hinton изобрёл быстрые алгоритмы обучения.

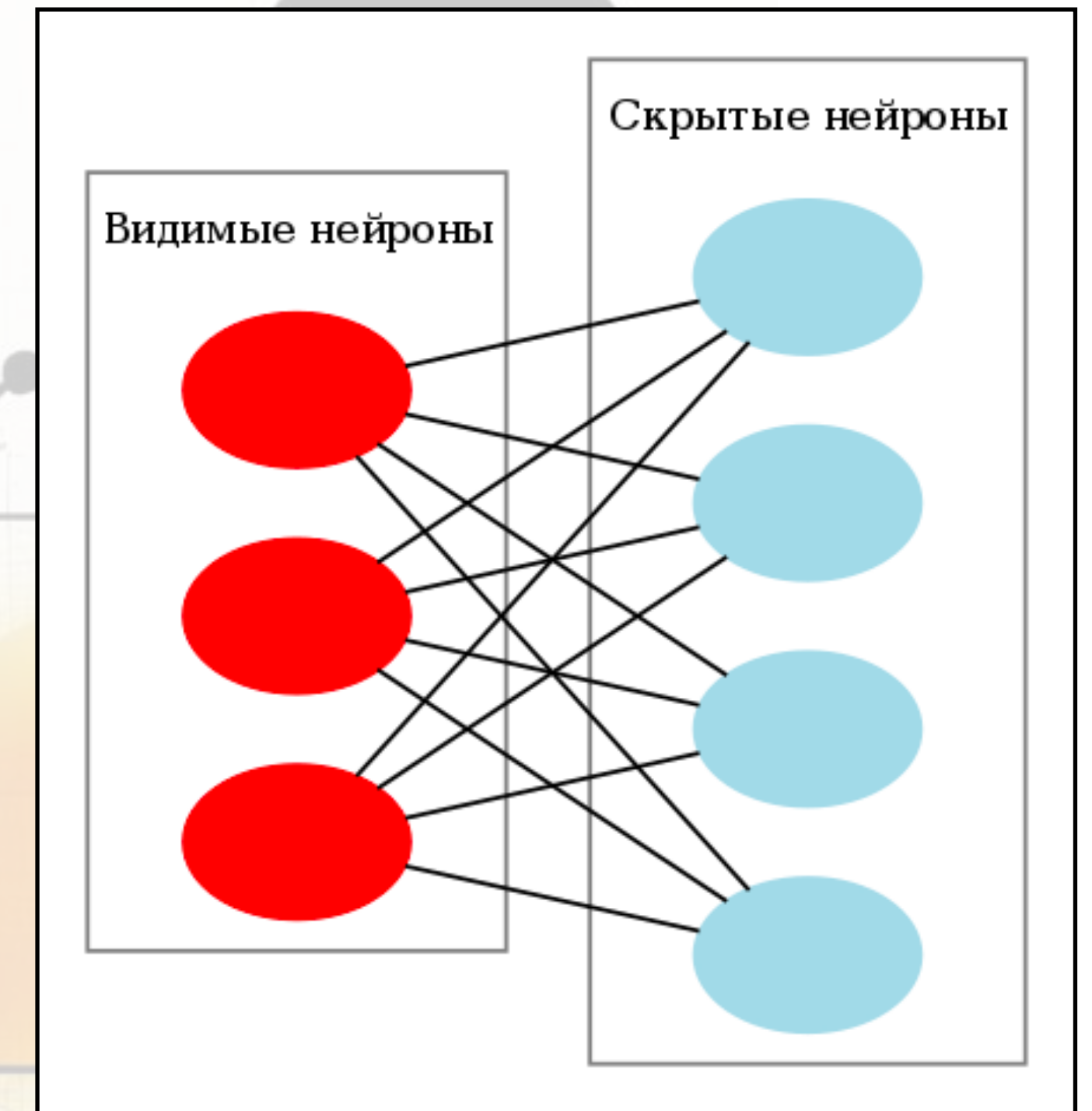
Архитектура: 1 видимый слой, 1 скрытый слой.

Двудольный полносвязный граф

Обучение: без учителя (но может и с учителем).

Метод градиентного спуска (с контрастивной дивергенцией)

Приложения: коллаборативная фильтрация, *снижение размерности данных*, задачи классификации, выделение признаков, тематическое моделирование.



Ограниченная машина Больцмана, wiki

Про RBM-2

Для нас Ψ — чёрный ящик.

Слой *видимых* нейронов — аналог S .

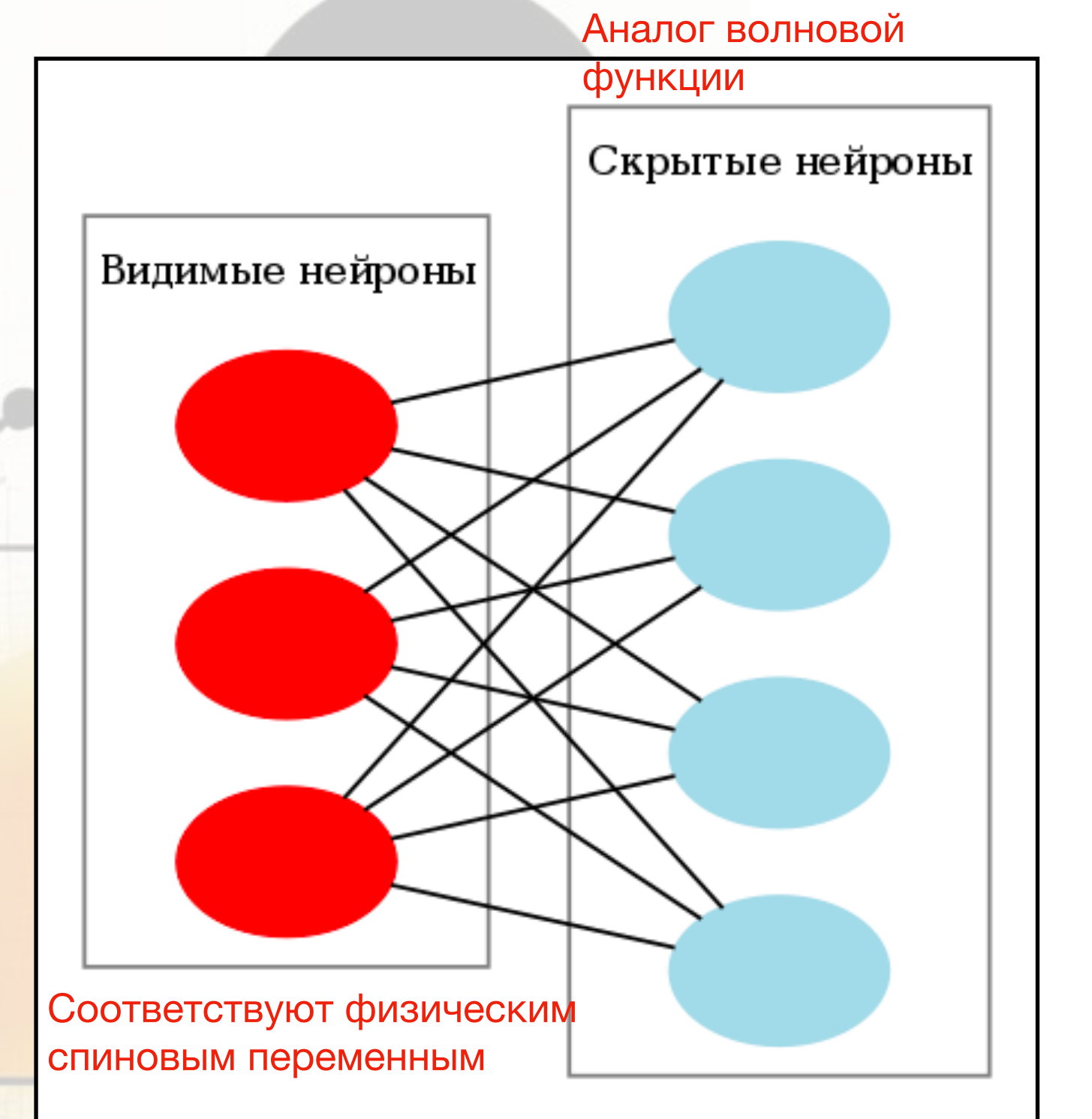
Слой *скрытых* нейронов — некий аналог Ψ .

Наша цель — **аппроксимировать вычислительный «чёрный ящик» нейросетью**, обученной наилучшим образом представлять Ψ .

Авторы наиболее популярной статьи делают это с помощью RBM, однако архитектура сети может быть ~произвольной.

Обращаясь снова к к.ф.: очень малое число возможных состояний квантовых систем физично, то есть встречается в природе.

Наша задача сводится к тому, чтобы **найти ту плотную область Гильбертова пространства с реальными состояниями**.



Ограниченная машина Больцмана, wiki

NQS

Идея:

Спины кодируются 0 и 1. $S = \{s_1, s_2, \dots, s_j\}$

$\text{RBM} := \{W, a, b\}$ **НС описывается своими весами и смещениями**

Состояние RBM $:= |S\rangle = \{h, v\}$. **Состояние описывается видимой и скрытой частью**

Строим волновую функцию системы:

$$\Psi(S, W) = \sum_{h_i} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$

Пользуясь *свойствами* RBM, **упрощаем:**

(Юниты каждой из долей графа не взаимодействуют)

$$\Psi(S, W) = e^{\sum_i a_i \sigma_i^z} \prod_i 2 \cosh[b_i + \sum_j W_{ij} \sigma_j^z]$$

NQS

Как получить веса и смещения для RBM?

Логика: минимальное состояние энергии соответствует основному состоянию, *основное состояние является самым вероятным состоянием.*

Сэмплируем их из волновой функции при помощи MCMC (взяв некоторое начальное случайное состояние системы). Оцениваем среднюю энергию сэмплов. Находим градиент энергии $\frac{\partial E}{\partial \Psi}$ по весам машины и обновить их.

Таким образом энергия всё время *уменьшается*, и мы **точно придём к основному состоянию.**