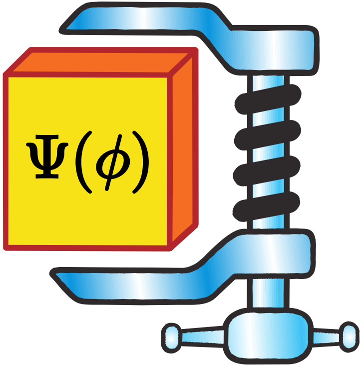
Tensor network states

to compress the (continuum) many body problem



Antoine Tilloy

January 7th, 2022

Séminaire

Modern Physics and the Many Body Problem

Quantum Mechanics in a nutshell

Principles

1. **System** – vector $|\psi\rangle$ in a Hilbert space $\mathscr{H}$
2. **Evolution** – generated by H self-adjoint on $\mathscr{H}$

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi_t\rangle = -\frac{i}{\hbar}H|\psi_t\rangle$$

3. **Measurement** – Ideal measurement $\equiv$ self-adjoint $\mathcal{O} = \sum_k \lambda_k P_k$

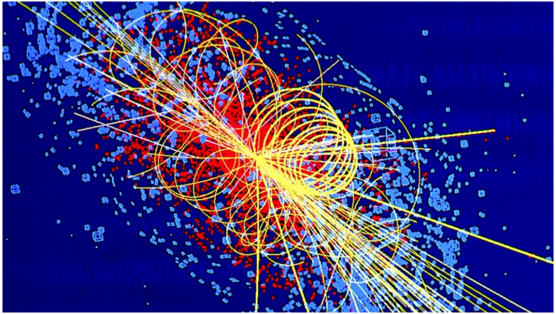
   Result k with probability $p_k = \langle\psi|P_k|\psi\rangle$ followed by collapse

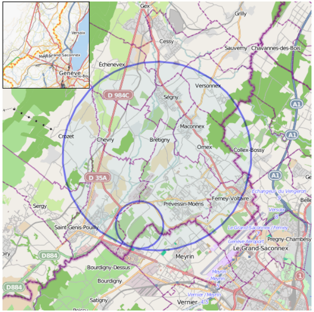
$$|\psi\rangle \longrightarrow \frac{P_k|\psi\rangle}{\|P_k|\psi\rangle\|}$$

Remark:

- Conceptual subtleties in $2 \implies 3$, "measurement problem"

The standard model of particle physics





The **standard model** is an instantiation of quantum mechanics that is potentially fundamental

1. Hilbert space $\mathscr{H}$ (the fundamental particles and their statistics)
2. Hamiltonian H (all the forces/interactions between the particles)

Completed by the independent detection of the Higgs boson by the Atlas and CMS collaborations at the LHC

The standard model: theory and practice

Theory: The standard model cannot be exact even in principle

1. no gravity
2. mathematically breaks down at (insanely) short distances
   (H does not exist as a self-adjoint operator on the $\mathscr{H}$ we would like, *even non-rigorously* with state of the art renormalization group theory)

The standard model: theory and practice

Theory: The standard model cannot be exact even in principle

1. no gravity
2. mathematically breaks down at (insanely) short distances
   (H does not exist as a self-adjoint operator on the $\mathscr{H}$ we would like, *even non-rigorously* with state of the art renormalization group theory)

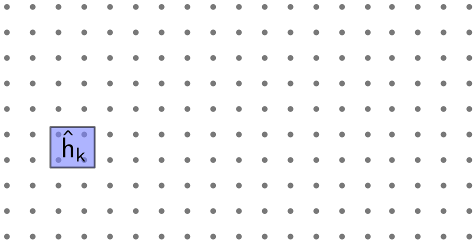
Practice: The standard model is exact for all practical purposes

- Electron - photon physics tested to 12 digits

$$g = 2.10011596521808 5(76)$$

- Ultra high energy subtleties irrelevant for realizable phenomena

Quantum many-body problem on the lattice



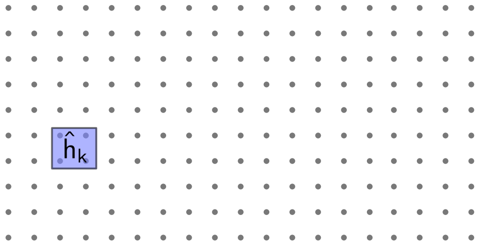
Typical many-body problem

N spins on a lattice

$\mathscr{H} = \bigotimes_{j=1}^{n} \mathscr{H}_j$ with $\mathscr{H}_j = \mathbb{C}^2$

$|\psi\rangle = \sum c_{i_1, i_2, \cdots, i_n} | i_1, i_2, \cdots i_N \rangle$

Quantum many-body problem on the lattice



Problem:

Finding the low energy states of

$$H = \sum_{k=1}^{N} h_k$$

is **hard** because $\dim \mathscr{H} = 2^N$ for spins

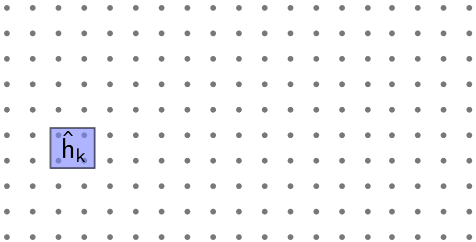
Typical many-body problem

N spins on a lattice

$\mathscr{H} = \bigotimes_{j=1}^{n} \mathscr{H}_j$ with $\mathscr{H}_j = \mathbb{C}^2$

$|\psi\rangle = \sum c_{i_1, i_2, \cdots, i_n} |i_1, i_2, \cdots i_N\rangle$

Quantum many-body problem on the lattice



Typical many-body problem

N spins on a lattice

$\mathscr{H} = \bigotimes_{j=1}^{n} \mathscr{H}_j$ with $\mathscr{H}_j = \mathbb{C}^2$

$|\psi\rangle = \sum c_{i_1, i_2, \cdots, i_n} |i_1, i_2, \cdots i_N\rangle$

Problem:

Finding the low energy states of

$$H = \sum_{k=1}^{N} h_k$$

is **hard** because $\dim \mathscr{H} = 2^N$ for spins

Fugaku – 2 EFLOPS – 150 PB
cannot do $4 \times 4 \times 4$ spins

Perturbation theory

Main idea: usuall $H = H_0 + \lambda V$ where H_0 is diagnalizable exactly
$\rightarrow$ Taylor expand in λ

Perturbation theory

Main idea: usuall $H = H_0 + \lambda V$ where H_0 is diagnalizable exactly
$\rightarrow$ Taylor expand in λ

Field theoretic perturbation theory

1. Interaction representation: $|\psi_t\rangle_I = e^{-iH_0 t}|\psi_t\rangle$ and $V_t = e^{-iH_0 t} V e^{iH_0 t}$

Perturbation theory

Main idea: usuall $H = H_0 + \lambda V$ where H_0 is diagnalizable exactly
$\rightarrow$ Taylor expand in λ

Field theoretic perturbation theory

1. Interaction representation: $|\psi_t\rangle_I = e^{-iH_0 t}|\psi_t\rangle$ and $V_t = e^{-iH_0 t} V e^{iH_0 t}$
2. Schrödinger equation

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi_t\rangle_I = \frac{-i\lambda}{\hbar} V_t |\psi_t\rangle_I$$

Perturbation theory

Main idea: usuall $H = H_0 + \lambda V$ where H_0 is diagnalizable exactly
$\rightarrow$ Taylor expand in λ

Field theoretic perturbation theory

1. Interaction representation: $|\psi_t\rangle_I = e^{-iH_0 t}|\psi_t\rangle$ and $V_t = e^{-iH_0 t} V e^{iH_0 t}$
2. Schrödinger equation

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi_t\rangle_I = \frac{-i\lambda}{\hbar} V_t |\psi_t\rangle_I$$

3. Formal integration

$$|\psi_t\rangle_I = T \exp\left[-\frac{i\lambda}{\hbar}\int_0^t \mathrm{d}u\, V_u\right] |\psi_0\rangle_I$$

Perturbation theory

Main idea: usuall $H = H_0 + \lambda V$ where H_0 is diagnalizable exactly
$\rightarrow$ Taylor expand in λ

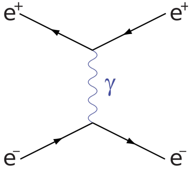
Field theoretic perturbation theory

1. Interaction representation: $|\psi_t\rangle_I = e^{-iH_0 t}|\psi_t\rangle$ and $V_t = e^{-iH_0 t} V e^{iH_0 t}$
2. Schrödinger equation
$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi_t\rangle_I = \frac{-i\lambda}{\hbar} V_t |\psi_t\rangle_I$$
3. Formal integration
$$|\psi_t\rangle_I = T \exp\left[-\frac{i\lambda}{\hbar}\int_0^t \mathrm{d}u\, V_u\right] |\psi_0\rangle_I$$
4. Dyson expansion
$$|\psi_t\rangle_I = \sum_{n=0}^{+\infty} \left(-\frac{i\lambda}{\hbar}\right)^n \int_{u_1 > u_2 > u_n} V_{u_1} V_{u_2} \cdots V_{u_n} |\psi_0\rangle_t$$

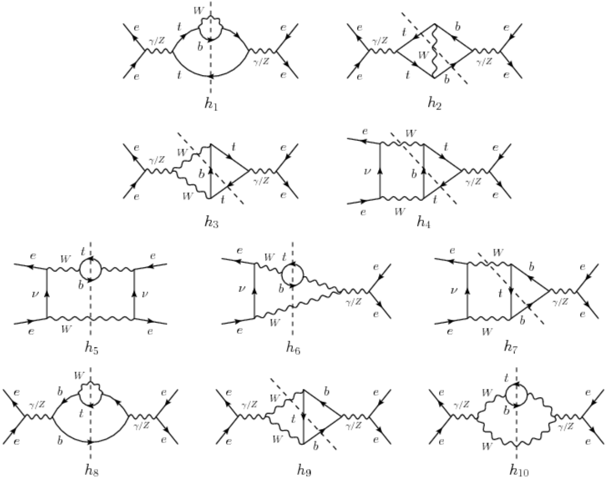
Diagrammatic reorganization of the expansion

Perturbation theory = Feynman diagrams

Order 2 in λ



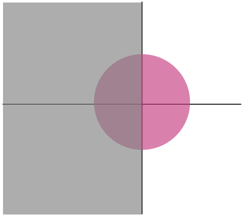
Order 6 in λ



Divergence of the expansion

First noted by Dyson in a 2 page letter

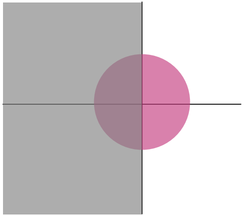
$$\text{any physical quantity} = f(\lambda) \underset{\text{illicit}}{=} \sum_n a_n \lambda^n \;\; \text{diverges} \; \forall \; \lambda$$



Divergence of the expansion

First noted by Dyson in a 2 page letter

any physical quantity $= f(\lambda) \underset{\text{illicit}}{=} \sum_n a_n \lambda^n$ diverges $\forall\ \lambda$



For ϕ_0^4

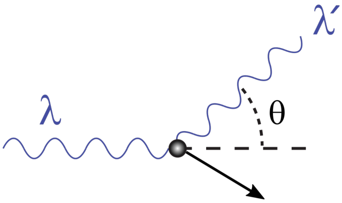
$$\int_{\mathbb{R}} \mathrm{d}\phi \ \exp(-m^2\phi^2 - \lambda\phi^4) \underset{\text{illicit}}{=} \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{n!} \int_{\mathbb{R}} \mathrm{d}\phi\, \phi^{4n} \exp(-m^2\phi^2)$$

$$= \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{n!\, m^{2n+1/2}} \int_0^{+\infty} \mathrm{d}u\, u^{2n+1/2} \exp(-u)$$

$$= \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{m^{2n+1/2}} \frac{\Gamma(2n+3/2)}{\Gamma(n+1)}$$

Successes

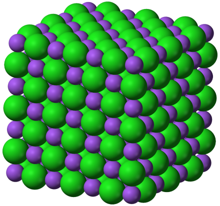
Fundamental Physics



Quantum electrodynamics

Nature is kind: $\lambda \simeq 1/137 \ll 1$
[electron-photon interaction is weak]

Condensed matter

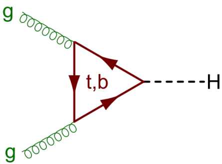


Metals/insulators/superconductors

Nature is kind: taking $\lambda = 0$ is fine
[electron-electron interaction weak]

Open problems

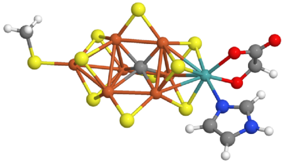
Fundamental Physics



Strong force between quarks and gluons

Nature semi-kind
$\lambda = -\log[E_{\text{kin}}/E_0]$
$E_0 \simeq 200$ MeV
E_{kin} kinetic energy

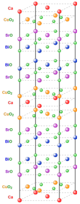
Chemistry



Atoms interacting to form molecules

Nature not kind $\lambda \sim 1$
but common approximations fairly efficient for simple cases

Condensed matter

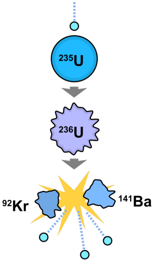


Cuprate perovskites superconductors

Nature not kind $\lambda \sim 1$
electrons interact strongly and get entangled

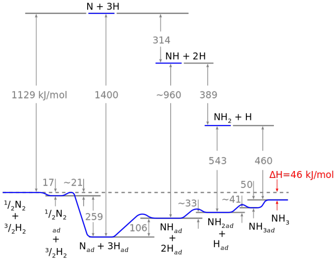
Consequences

Nuclear Physics



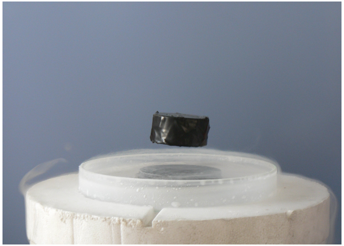
nuclei properties only measured in test reactors

Catalysis



Ammonia costly to produce
1% of CO2 prod.

High T_c superconductors



No room temp. supra
No hovering skateboards
Costly electricity transport

Many options

Many popular approximations to go beyond standard perturbation theory

- Dynamical mean field theory (DMFT) for condensed matter
- Density functional theory (DFT) for chemistry
- Quantum Monte-Carlo
- Diagrammatic Monte-Carlo
- Resummation/resurgence

Many options

Many popular approximations to go beyond standard perturbation theory

- Dynamical mean field theory (DMFT) for condensed matter
- Density functional theory (DFT) for chemistry
- Quantum Monte-Carlo
- Diagrammatic Monte-Carlo
- Resummation/resurgence

2 bleeding edge promising approaches

1. Quantum computing
2. Classical compression (variational method)

The variational method and tensor networks

The direct compression approach

Variational method for ground state search

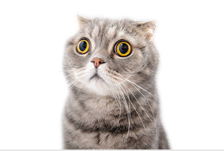
1. Guess a manifold $\mathcal{M} \subset \mathscr{H}$ with few parameters ν *i.e.* $\dim \mathcal{M} \ll \dim \mathscr{H}$
2. Tune ν to minimize energy $\nu = \mathrm{argmin}_{\nu \in \mathcal{M}} \frac{\langle \nu | H | \nu \rangle}{\langle \nu | \nu \rangle}$ and get $|\text{ground state}\rangle \simeq |\nu\rangle$

The direct compression approach

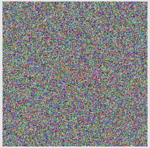
Variational method for ground state search

1. Guess a manifold $\mathcal{M} \subset \mathscr{H}$ with few parameters ν *i.e.* $\dim \mathcal{M} \ll \dim \mathscr{H}$
2. Tune ν to minimize energy $\nu = \text{argmin}_{\nu \in \mathcal{M}} \frac{\langle \nu | H | \nu \rangle}{\langle \nu | \nu \rangle}$ and get $|\text{ground state}\rangle \simeq |\nu\rangle$

Reason for compression (classical)



cat image



"typical" image

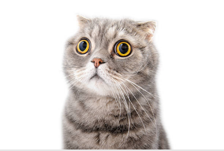
atypical $\implies$ compressible

The direct compression approach

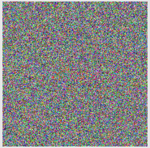
Variational method for ground state search

1. Guess a manifold $\mathcal{M} \subset \mathscr{H}$ with few parameters ν *i.e.* $\dim \mathcal{M} \ll \dim \mathscr{H}$
2. Tune ν to minimize energy $\nu = \text{argmin}_{\nu \in \mathcal{M}} \frac{\langle \nu | H | \nu \rangle}{\langle \nu | \nu \rangle}$ and get $|\text{ground state}\rangle \simeq |\nu\rangle$

Reason for compression (classical)



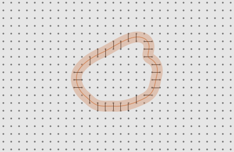
cat image



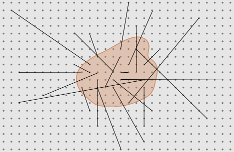
"typical" image

atypical $\implies$ compressible

Reason for compression (quantum)



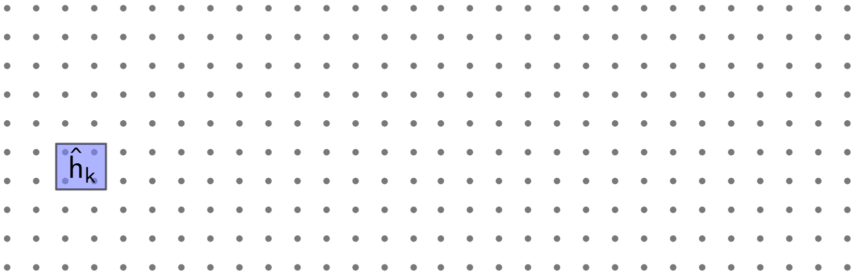
low energy state



random state

area law $=$ atypical $\implies$ compressible

Many-body problem



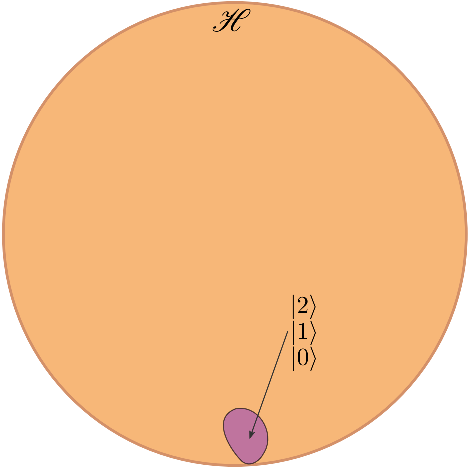
Problem

Finding low energy states of

$$\hat{H} = \sum_{k=1}^{N} \hat{h}_k$$

is **hard** because $\dim \mathscr{H} \propto 2^N$

Variational optimization



Generic (spin $1/2$) state $\in \mathscr{H}$:

$$|\psi\rangle = \sum_{i_1,\cdots,i_n=\pm 1} c_{i_1,i_2,\cdots,i_N} |i_1,\cdots,i_N\rangle$$

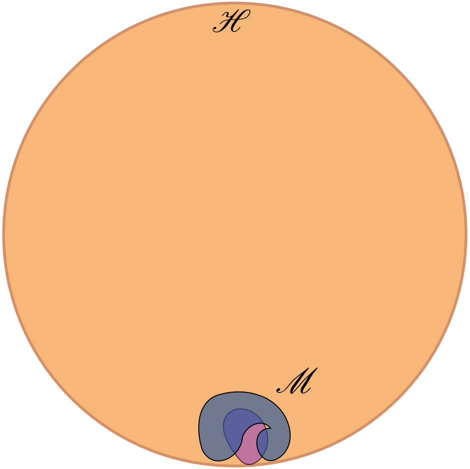
Exact variational optimization

To find the ground state:

$$|0\rangle = \min_{|\psi\rangle\in\mathscr{H}} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

- $\dim \mathscr{H} = 2^N$

Variational optimization



Generic (spin $1/2$) state $\in \mathscr{H}$:

$$|\psi\rangle = \sum_{i_1,\cdots,i_n=\pm 1} c_{i_1,i_2,\cdots,i_N} \, |i_1,\cdots,i_N\rangle$$

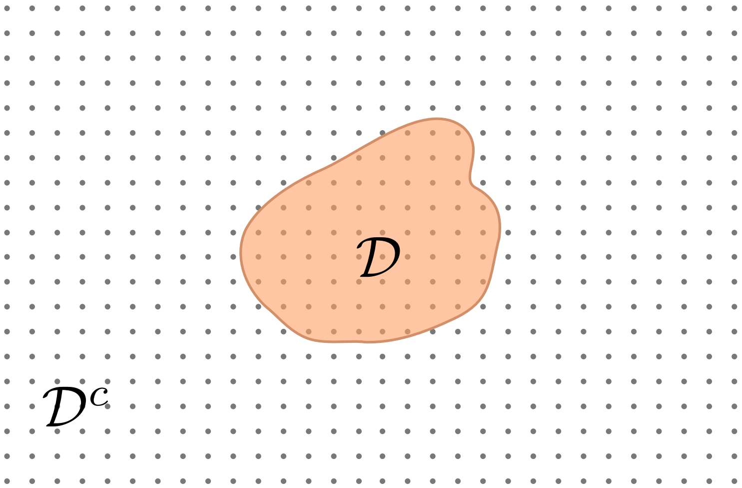
Approx. variational optimization

To find the ground state:

$$|0\rangle = \min_{|\psi\rangle\in\mathscr{M}} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

- $\dim \mathscr{M} \propto \mathrm{Poly}(N)$ or fixed

Interesting states are weakly entangled



Low energy state

$|\psi\rangle = |0\rangle$ or $|1\rangle$ ...

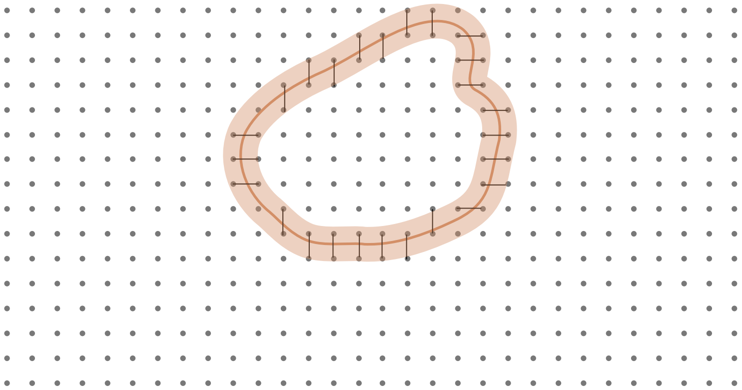
Reduced density matrix
$\rho = \mathrm{tr}_{\mathcal{D}^c}\left[|\psi\rangle\langle\psi|\right]$

Entanglement entropy
$S = -\mathrm{tr}\left[\rho \log \rho\right]$

Area law

$$S \propto |\partial \mathcal{D}|$$

Interesting states are weakly entangled



Low energy state

$|\psi\rangle = |0\rangle$ or $|1\rangle$ ...

Reduced density matrix

$\rho = \mathrm{tr}_{\mathcal{D}^c}\left[|\psi\rangle\langle\psi|\right]$

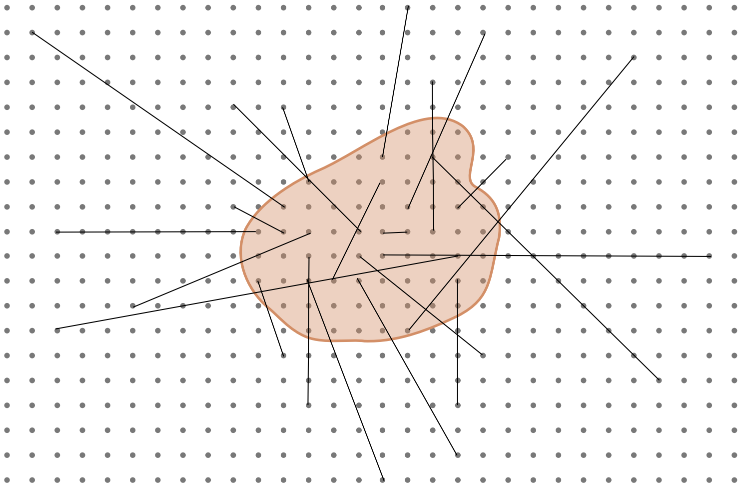
Entanglement entropy

$S = -\mathrm{tr}\left[\rho \log \rho\right]$

Area law

$$S \propto |\partial \mathcal{D}|$$

Typical states are strongly entangled



Random state

$|\psi\rangle = U_{\text{Haar}}|\text{trivial}\rangle$

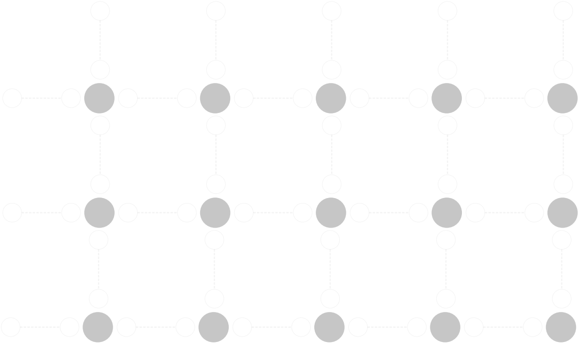
Reduced density matrix
$\rho = \text{tr}_{\mathcal{D}^c}\left[|\psi\rangle\langle\psi|\right]$

Entanglement entropy
$S = -\text{tr}\left[\rho \log \rho\right]$

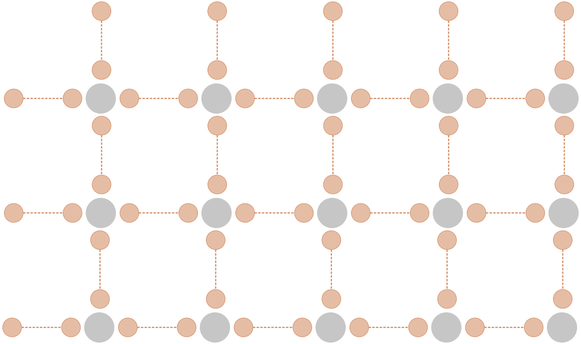
Volume law

$S \propto |\mathcal{D}|$

Constructing weakly entangled states



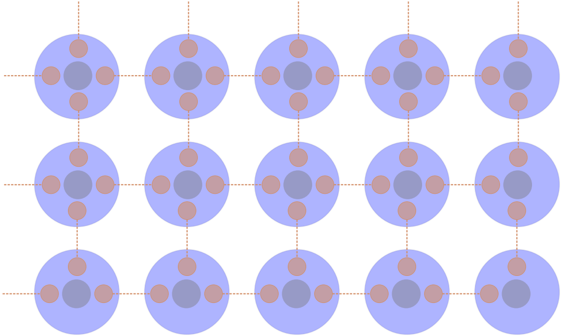
Constructing weakly entangled states



1. Put auxiliary **maximally entangled** states between sites

$$= \sum_{j=1}^{D} |j\rangle|j\rangle$$

Constructing weakly entangled states



1. Put auxiliary **maximally entangled** states between sites

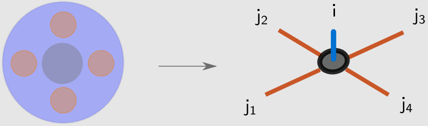
$$= \sum_{j=1}^{D} |j\rangle|j\rangle$$

2. Map to initial Hilbert space on each site

$$= A : \mathbb{C}^{D^4} \to \mathbb{C}^2$$

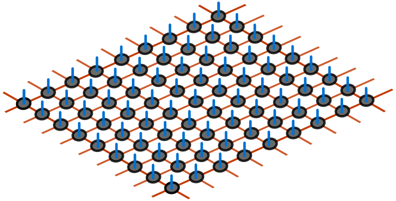
Tensor network states: definition

Why "tensor" network?



$$A : \mathbb{C}^{D^4} \to \mathbb{C}^2 \quad \longrightarrow \quad A^i_{j_1, j_2, j_3, j_4}$$

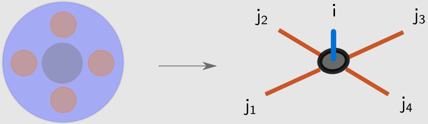
$|A\rangle =$



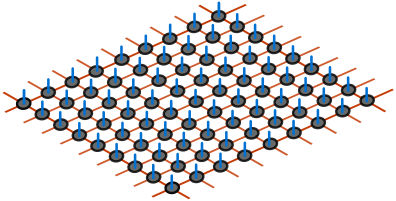
with tensor contractions on links

Tensor network states: definition

Why "tensor" network?



$$A : \mathbb{C}^{D^4} \to \mathbb{C}^2 \quad \longrightarrow \quad A^{i}_{j_1,j_2,j_3,j_4}$$

$|A\rangle =$ 

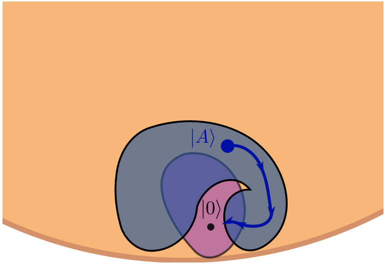
with tensor contractions on links

Optimization

Find best A for fixed Di $(2 \times D^4 \times$ coeff.)

$$E_0 \simeq \min_A \frac{\langle A|\widehat{H}|A\rangle}{\langle A|A\rangle}$$

for example go down $\frac{\partial E}{\partial A^{i}_{j_1,j_2,j_3,j_4}}$



Matrix Product States (MPS) aka tensor trains

Definition

A MPS for a translation invariant chain of N spins/qubits ($\mathscr{H}_k = \mathbb{C}^d$) with periodic boundary conditions is a state

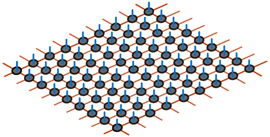
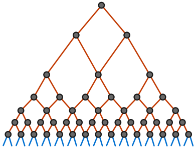
$$|\psi(A)\rangle := \sum_{i_1, i_2, \ldots, i_N} \mathrm{tr}\left[A_{i_1} A_{i_2} \cdots A_{i_N}\right] |i_1, i_2, \ldots, i_N\rangle$$

where A_i are 2 matrices $\in \mathcal{M}_D(\mathbb{C})$.

- The matrices A_i for $i = \pm 1$ are the free parameters
- The size D of the matrices is the **bond dimension** (quantifies freedom)
- Correlation functions (and $\langle H \rangle$) efficiently computable
- Optimizing over A provably gives good results for gapped H

Tensor network states in a nutshell

.zip or **.jpg** for complex quantum states that appear in Nature

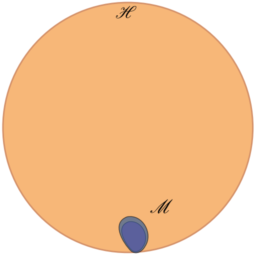
1. **Exponential reduction:** $2^N \longrightarrow N \times D^{2d}$ parameters
   [N number of spins, D amount of entanglement, d space dimension $(1, 2, 3)$]

2. **Efficient compression:** compression error $\leqslant e^{-D}$ or $1/\text{superpoly(D)}$
   [For a large number of *a priori* non-trivial problems]

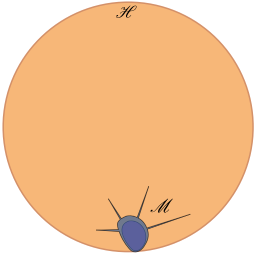
[History: 1992 for $d = 1$, 2004 for $d \geqslant 2$, 2016 for 2d-Hubbard at $T = 0$]

Some facts

$d = 1$ spatial dimension



$d \geqslant 2$ spatial dimension



Theorems (colloquially)

1. For gapped H, TNS $|A\rangle$ approximate well $|0\rangle$ with D fixed
2. **All** $|A\rangle$ are ground states of gapped H

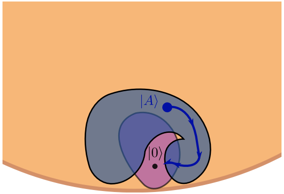
Folklore

1. For gapped H, TNS $|A\rangle$ approximate well $|0\rangle$ with D fixed
2. **Most** $|A\rangle$ are ground states of gapped H

Optimization

To find lowest energy state, with generic TNS, still need to optimize the poly(D) parameters

- Naive gradient descent inefficient (works only for $D \leqslant 10$)
- Riemanian gradient descent highly efficient $(= \text{TDVP})$



Metric on tensor network state manifold

1. $|\psi(A)\rangle \in \mathscr{M}$ a state in the tensor network manifold
2. $|\psi(A), W\rangle = W \cdot \nabla_A |\psi(A)\rangle$ the tangent vector in A along direction W
3. $g_A(V, W) := \mathrm{Re}\,\langle \psi(A), V | \psi(A), W\rangle$ induced Hilbert metric

Note: best is to do Riemanian quasi-Newton, like Riemanian conjugate gradient or Riemanian LBFGS $\rightarrow$ `OptimKit.jl` by Haegeman et al.

State of the art

Dense: all states approximable (trivial)
Efficient: cost $Poly(D)$ error $1/superPoly(D)$ for local gapped
(cost assuming free optimization – theoretical guarantees on optim very bad)

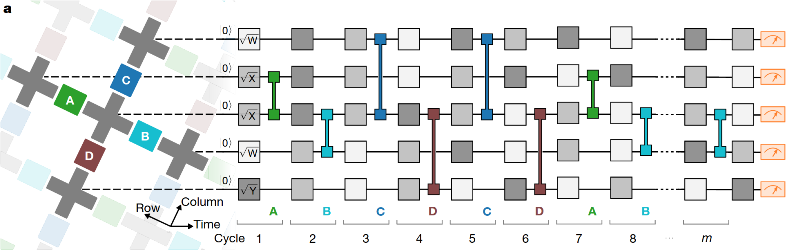
It's all a matter of prefactors and exponents

- 1 space dimension $\rightarrow D \geqslant 1000 \rightarrow$ machine precision
  (MPS results "numerically exact")
- 2 space dimensions $\rightarrow D \sim 10 \rightarrow$ efficient
  (PEPS efficient to $10^{-2} - 10^{-6}$ depending on problems)
- 3 space dimensions $\rightarrow D \sim 3 \rightarrow$ theoretically efficient but too expensive

Going forward

♣ **Extend to hard problems** without theoretical guarantees no area law

1. Quantum chemistry
2. Real-time evolution
   Hardness motivates Google supremacy experiment [Nature, 2020]

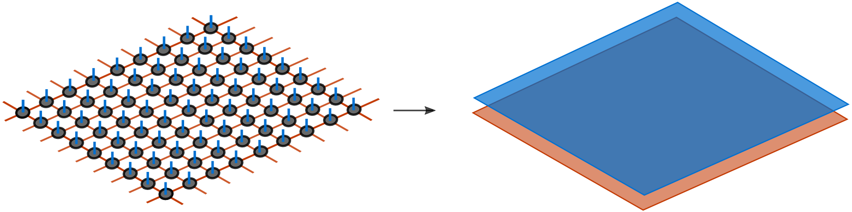


   Still possible compression (Zhou, Stoudenmire, Waintal)

♣ **Extend to problems that should not be so hard** but unadapted

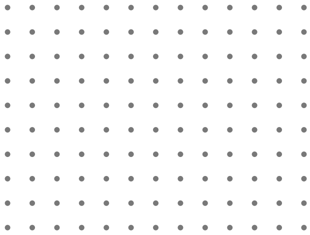
1. Continuum field theories

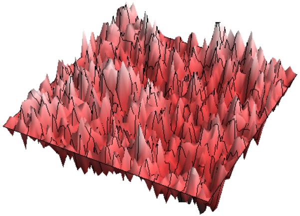
Tensor network states for continuum theories



The quantum many-body problem in the continuum

From the lattice to the continuum and Quantum Field Theory (QFT)





$$|\Psi\rangle = \sum_{i_1, i_2, \cdots, i_N} c_{i_1 i_2 \cdots i_N} |i_1 i_2 \cdots i_N\rangle \quad \longrightarrow \quad |\Psi\rangle = \int \mathcal{D}\phi \; \psi(\phi) \, |\phi\rangle$$

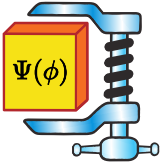
New problem: 2^N $\mathbb{C}$-parameters $\rightarrow$ $\dim \mathscr{H} = \infty^\infty$ even at finite size!

Question Can one compress ∞^∞ down to a manageable number of parameters?

Objective

Continuous tensor networks

Compress field wavefunctions $\psi(\phi)$ and use them to solve the continuous-many-body problem directly leveraging a continuous generalization of tensor networks



First insights in 2010 and recent progress

	non-relativistic	relativistic	critical
$d = 1$ space	Verstraete-Cirac 2010	Tilloy 2021	
$d \geqslant 2$ space	Tilloy-Cirac 2019		

no idea	heuristics	clear definition	fast algorithm

Relativistic CMPS

Definition

$$|R, Q\rangle = \mathrm{tr}\left\{\mathcal{P}\exp\left[\int \mathrm{d}x\, Q \otimes \mathbb{1} + R \otimes a^\dagger(x)\right]\right\}|0\rangle_a$$

$a^\dagger(x)$ Fourier transform of mode creation operator,
$|0\rangle_a$ Fock vacuum annihilated by $a(x)$,
$\mathcal{P}$ path ordering operator

Some properties

1. Expectation values can be evaluated to machine precision at cost D^3
2. R, Q can be optimized with geometric methods

Numerics

For φ^4 theory:

$$H = \int dx \, \frac{:\pi^2:}{2} + \frac{:(\nabla\phi)^2:}{2} + \frac{m^2}{2} :\phi^2: + g :\phi^4:$$