

# 機械学習を用いて量子多体系を表現する



野村 悠祐

東京大学大学院工学系研究科  
nomura@ap.t.u-tokyo.ac.jp



山地 洋平

東京大学大学院工学系研究科  
yamaji@ap.t.u-tokyo.ac.jp



今田 正俊

東京大学大学院工学系研究科  
imada@ap.t.u-tokyo.ac.jp

量子多体系とは多数の自由度が互いに相互作用しあう系を指し、そこにおける各自由度の運動は多体の波動関数によって支配される。多体波動関数さえわかれば問題は解決であるが、多体系のハミルトニアンは自由度の数に対して指数関数的に増大するため、自由度の数が増えたと多体波動関数を厳密に求めることは不可能になる。そのため、多体波動関数をいかに精度よく表現できるか？という問題は、非常に重要な課題になる。

この問題に対して新しい風が吹いている。これまで物理的洞察に基づいて波動関数を近似する試みは多くなされているが、それとは逆に対称性などの明らかな拘束条件以外には直観に頼らず、機械学習の力を借りて波動関数を構築しようという動きである。機械学習は膨大なデータセットからその本質的なパターンを抽出する。すなわち、本来の波動関数は指数関数的に大きな次元を持つベクトルとみなせるが、その本質的なパターンを機械学習によって見つけ出し、波動関数の次元よりもはるかに少ない数のパラメータを用いて波動関数を精度よく近似しようという試みである。

機械学習の手法の中でも、本稿では、人工ニューラル・ネットワークの一種であるボルツマンマシンを用いた変分波動関数について議論する。ボルツマンマシンは可視（入力）層の自由度である可視ユニットに加え、仮想的自由度である隠れ層の不可視ユニットで構成され、ユニット間が結合（相互作用）を持つ構造をしている。ボルツマンマシンは可視ユニットの状態配置を入力とし、その生成確率分布を学習する機械である。それに対し、多体波動関数は物理的なハミルトニアンは自由度の状態配置

それぞれに対して値を与える。波動関数の値が、それらの状態配置に対する（複素数に一般化された）確率であると考え、物理的ハミルトニアンの自由度と可視ユニット自由度を同一視することで、多体波動関数をボルツマンマシンによって書き下すことができる。すなわちボルツマンマシンをユニット間結合定数などを変分パラメータとする変分波動関数とみなし、物理的ハミルトニアンはエネルギー期待値を最小化するように波動関数を最適化し（機械学習の言語ではこれが学習に対応）、未知の基底状態を探索する。

最初に量子多体系に適用されたのは、最も単純な構造をしている制限ボルツマンマシン（Restricted Boltzmann machine, 略してRBM）である。RBMを用いた変分波動関数は量子スピン系に適用され、その精度の良さが実証された。

この研究を皮切りにRBM波動関数の特性が以下のようにわかってきた。i) 不可視ユニットの数を指数関数的に増やすとどんな波動関数も表現可能。ii) 相互作用が短（長）距離だと、エンタングルメント・エントロピーが表面（体積）則を満たすこと。iii) 従来の波動関数法を組み合わせるとフェルミオン系などにも適応可能となりより良い精度が達成できること。iv) 隠れ層を一層増やした深層ボルツマンマシン（DBM）にすると関数表現能力が格段に向上する。その性質を用いると、DBMを用いて基底状態の波動関数を厳密に構築できること。

ボルツマンマシン波動関数の研究は始まったばかりである。その有用性、より深い学理に近い将来明らかになるだろう。

## —Keywords—

### 多体波動関数：

多数の粒子が相互作用しあうハミルトニアンに対するシュレーディンガー方程式（あるいはディラック方程式）の解である多変数関数。自由度の数が増えていくと、ハミルトニアンの次元は指数関数的に増え、厳密な多体波動関数を得ることは一般には不可能である。

### 変分波動関数法：

多体波動関数を有限個のパラメータに依存する形で書き下し、パラメータを最適化することによって多体波動関数を精度よく近似する方法。パラメータの最適化は変分原理（波動関数のエネルギー期待値は基底状態エネルギーより小さくなることはないことを保証）を用いて、エネルギーを最小化するように行うことが多い。

### ボルツマンマシン：

確率生成モデルの一種。可視ユニットと不可視ユニット（物理における補助場のような役割を果たす）からなる仮想的な系に対して、それらの相互作用などに依存する仮想的なエネルギーを定義する。そのエネルギーに応じたボルツマン重みを用い、不可視ユニットの状態のみについて部分状態和を取ると、可視ユニットの状態を入力として、それに対して確率を与える機械となる。

### 制限ボルツマンマシン (RBM)：

ボルツマンマシンの中でも隠れ層が1層のみ存在し、層間の結合しかない[同一層内(可視ユニット間や不可視ユニット間)に結合がない]構造を持つもの。

## 1. はじめに

近年、物理の分野でも機械学習の有用性の吟味が始まっている。本稿では、その手法の中でも人工ニューラル・ネットワークを物性物理に適用した例を取り上げる。まずどのような適用例があるのかをここで簡単に紹介する。

機械学習は大量のデータからデータの分布に関する何らかのモデルを構築する(=本質的なパターンを抽出する)。モデルはその目的によって、識別モデルと生成モデルに大別される。識別モデルではあるデータが特定のクラスに所属する確率を直接モデル化し、そのモデルに基づいてクラスの境界を学習するものである。犬と猫の画像データセットに対して、犬・猫である確率をモデル化して、犬か猫かに分類するようなことを考えていただければ良い。このように識別モデルでは与えられた観測データがどのような生成確率に基づいているかは考慮しない。逆に生成モデルでは、それぞれのクラスにおけるデータ分布がどのような確率分布で生成されるかをモデル化する。

識別モデルの物理への適用例としては量子相の分類、相境界決定・相転移検出<sup>1-10)</sup>が代表的である。例えば系の波動関数の実空間分布を画像だと思って、知られた画像解析手法を使い、量子相を分類する。

生成モデルの適用例としては、量子多体系のソルバー<sup>11-23)</sup>や、モンテカルロ法の効率化<sup>24,25)</sup>などが挙げられる。前者では、人工ニューラル・ネットワークを用いて、多体系の波動関数を表す。後者では、通常メトロポリス法などに則って生成したモンテカルロサンプルを学習機械に入力し、背景にある生成確率分布を学習させ、効率的なクラスター更新を提案する試みなどがある。

本稿では、生成モデルに注目し、量子多体系の波動関数を表現し、多体系のソルバーとして用いる研究を紹介する。様々な人工ニューラル・ネットワークの中でもボルツマンマシンを使った波動関数表現に関して解説する。まず次節でボルツマンマシンについて紹介する。3節ではRBMを使った波動関数表現の手法を導入し、4節において、その性質や数値結果について議論する。5節と6節では、ボルツマンマシン波動関数を改良するための試みを紹介する。

## 2. ボルツマンマシン

波動関数表現への応用を議論する前に、人工ニューラル・ネットワークの一種であるボルツマンマシン<sup>26)</sup>特にRBM<sup>27)</sup>と深層ボルツマンマシン(DBM: deep Boltzmann machine)の構造と表現能力についてまとめておこう。

### 2.1 制限ボルツマンマシン (RBM)

図1(a)にRBMの構造を示す。RBMは、人工ニューロンの自由度である $N$ 個の可視ユニットと $M$ 個の不可視ユニットから構成される。それぞれの自由度が存在する領域を可視層、不可視(隠れ)層と呼ぶ。可視、不可視ユニットの状態を表す変数をそれぞれ $v_i, h_j$  ( $i=1, \dots, N, j=1, \dots, M$ )とし、各ユニットは2つの状態を取り得るとする。その状

シリーズ「人工知能と物理学」 機械学習を用いて量子多体系を表現する

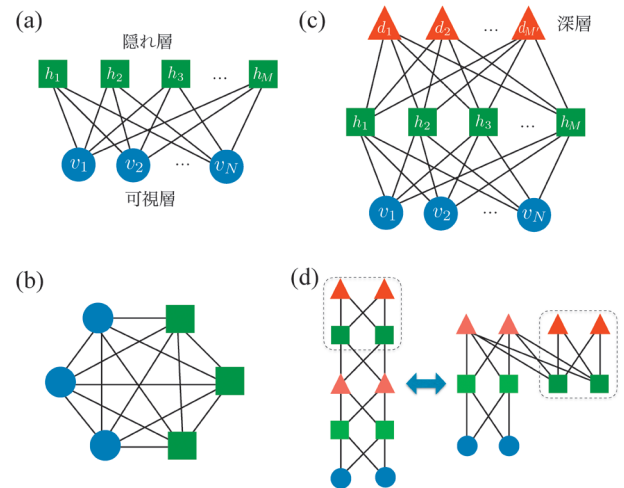


図1 (a) 制限ボルツマンマシン (RBM), (b) 一般のボルツマンマシン, (c) 深層ボルツマンマシン (DBM) の構造. (d) 一見隠れ層が2層以上ある構造でも不可視ユニットの位置の並び替えによって隠れ層が2層の構造に帰着する。

態値は $v_i, h_j = 0, 1$ で定義されることが多いが、ここではRBM波動関数を導入した論文<sup>11)</sup>に従って、それと等価な $v_i, h_j = \pm 1$ (イジング変数)で定義する。RBMでは一般のボルツマンマシン(図1(b))と異なり、可視ユニット間同士や不可視ユニット間同士には相互作用がなく、層間の結合しかないという制約がある。そのため、制限ボルツマンマシン(**Restricted Boltzmann machine**)と呼ばれている。RBMは可視ユニットのイジングスピン変数の任意の配置が生じる確率を近似する生成モデルである。下の議論で明らかになるが、不可視ユニットがあることによって確率を近似する際の柔軟性が生まれている。

可視・不可視ユニットの状態の配置をそれぞれ $v = (v_1, \dots, v_N), h = (h_1, \dots, h_M)$ とし、それぞれの状態の配置に応じて仮想的なエネルギー関数 $E(v, h)$ を

$$E(v, h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} W_{ij} v_i h_j \quad (1)$$

と定義する。ここで $a_i, b_j$ はバイアスと呼ばれる。それぞれのユニットの状態をイジングスピン自由度だと考え、式(1)を仮想的なイジング模型のエネルギーと考えると、これらの項は磁場に対応する。 $W_{ij}$ は隠れ層と可視層の間の結合(スピン間の相互作用)である。分配関数 $Z = \sum_{v,h} e^{-E(v,h)}$ を用いて、ボルツマン重み(統計力学的確率分布関数) $p(v, h)$ を $p(v, h) = e^{-E(v,h)} / Z$ と定義する。 $v, h$ に関する和はイジング自由度に対する状態和(トレース)をとる(トレースアウト)ことを意味する。この分布はボルツマン分布に他ならないので、任意の確率分布を表現するには能力が足りない。しかし、不可視ユニットの状態配置 $h$ に対する部分状態和をとった $v$ に対する確率分布関数( $v$ のみを変数とする確率を与えるので、周辺分布とよばれる)

$$\tilde{p}(v) = \sum_h p(v, h) \quad \left( \sum_v \tilde{p}(v) = 1 \right) \quad (2)$$

を考えると、 $\tilde{p}(v)$ は変分パラメータ $a_i, b_j, W_{ij}$ をうまく選

ぶことにより  $v$  に対する任意の確率分布を精度よく近似できる。数学的には  $v$  が表す  $2^N$  個の配置に対する確率分布を近似する際、指数関数的に大きな数 ( $M \sim 2^N$ ) の不可視ユニットを用意すれば厳密に表現できる<sup>28,29</sup>ことが知られている。しかし、指数関数的に大きな数の不可視ユニットは現実的な計算では扱えないため、実際の計算では  $N$  の冪関数で増える数の不可視ユニットを用いる。

このようにRBMは一般のボルツマンマシン同様、多変数関数を近似する「機械」であるが、不可視ユニット間の相互作用がないために、不可視ユニットに関して一体問題に帰着し、 $h$  に対する部分状態和を解析的に行えることがRBMの特長である。<sup>\*1</sup> 実際、周辺分布  $\tilde{p}(v)$  に対して、

$$\tilde{p}(v) = e^{\sum_i a_i v_i} \times \prod_j 2 \cosh \left[ b_j + \sum_i W_{ij} v_i \right] \quad (3)$$

という解析的な表式を簡単に得ることができ、数値計算の際のコストを大幅に削減できる。

## 2.2 深層ボルツマンマシン (DBM)

次にDBMの説明をしよう。構造が図1(c)に示されている。DBMは、RBMのネットワークに加えて、もう一つの隠れ層が加わり、隠れ層が二層になった構造をしている。隣り合う層の間のみ結合が存在し、層内のユニット間には結合が存在しないという特徴はRBMと共通している。隣接層間のみ結合があるという条件のもと、新たな隠れ層をさらに加えていったとしても、図1(d)のように隠れ層をうまく並び変えてやれば、隠れ層が二層しかない構造に帰着できる。従って、DBMは二層隠れ層がある構造を考えれば十分であり、三層以上を考える必要はない。これ以降は新たに加わった“深い”隠れ層を深層と呼び、もともとあった“浅い”隠れ層を単に隠れ層と呼んで二つの層を区別する。またそれらの層の不可視ユニットも区別するために、それぞれ隠れユニット、深層ユニットと名付ける。

深層の  $M'$  個の深層ユニットの状態を  $d = (d_1, \dots, d_{M'})$ ,  $d_k = \pm 1$  と表すこととすると、DBMにおける仮想的なエネルギー関数は、

$$E(v, h, d) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_k b'_k d_k - \sum_{i,j} W_{ij} v_i h_j - \sum_{j,k} W'_{jk} h_j d_k \quad (4)$$

と表せる。 $b'_k$ は深層のバイアス、 $W'_{jk}$ は隠れ層と深層の間の結合である。RBMと同じやり方で分配関数は  $Z = \sum_{v,h,d} e^{-E(v,h,d)}$ 、ボルツマン重みは  $p(v, h, d) = e^{-E(v,h,d)} / Z$  で与えられ、周辺分布は

$$\tilde{p}(v) = \sum_{h,d} p(v, h, d) \quad \left( \sum_v \tilde{p}(v) = 1 \right) \quad (5)$$

で与えられる。このDBMによる周辺分布はRBMの周辺分布よりも、深層の自由度が増えたことにより確率分布の

<sup>\*1</sup> 結合に制約のない一般のボルツマンマシンの場合、不可視ユニットを厳密にトレースアウトするには  $M$  に対して指数関数的に大きな計算コストがかかる。

表現能力が大きく向上するという長所<sup>30</sup>がある一方、隠れユニット  $h_j$  と深層ユニット  $d_k$  の自由度を同時に解析的にトレースアウトすることができないという短所（そのため学習コストも増える）がある。ただ、隠れ層と深層のどちらか一方の自由度だけについては、解析的にトレースアウトが可能である。例えば深層の自由度だけをトレースアウトすると、RBMのように可視層に加えて隠れ層が一層の構造に帰着するが、RBMとは違い、隠れユニット間にも相互作用が生じてしまう。

## 3. RBMによる波動関数法

いよいよ本稿の本題に移る。まず、カルレオ (Carleo) とトロイヤ (Troyer) によって導入されたRBMを用いた多体波動関数表現について説明する。<sup>11</sup> 一般の量子状態  $|\psi\rangle$  はフォック状態 (数状態) やスピン状態  $\{|x\rangle\}$  を基底に用いて、

$$|\psi\rangle = \sum_x |x\rangle \langle x|\psi\rangle = \sum_x |x\rangle \psi(x) \quad (6)$$

と展開できる。本稿では基底は実空間で定義する。例えばスピン系であれば、取り得る全てのスピン配置によって決まる完全系によって量子状態を展開したことになり、 $\psi(x)$  が各スピン配置に対する波動関数の振幅を表す。

### 3.1 RBM 波動関数

波動関数が実数かつ正定値 ( $\psi(x) > 0$ ) である場合、波動関数を確率分布と解釈することができる。<sup>\*2</sup> この解釈に基づき、 $\psi(x)$  をRBMによる周辺分布 [式(2)] で表す。一般の波動関数は負や複素数の値を取るが、その際は、RBMのパラメータ  $a_i, b_j, W_{ij}$  を複素数に取れば良い (次節で詳しくこのことに触れる)。この節では機械学習の分野との対応を議論する部分があるので、その議論を平易にするために、波動関数が何かしらの確率分布関数として解釈できる ( $\psi(x) > 0$ ) と仮定しておく。

周辺分布 [式(2)] によって波動関数を表す際、可視ユニットの状態変数  $v_i$  は、解析するハミルトニアンを決める物理変数であるとみなす。 $S=1/2$  の量子スピン系 (例えば4.2.1節で扱うハイゼンベルグ模型) の場合、可視ユニットの数  $N$  は量子スピンのサイトの数  $N_{\text{site}}$  と同一に取り ( $N = N_{\text{site}}$ )、 $v_i$  として  $v_i = 2S_i^z = \pm 1$  と定義する。この定義では、可視ユニットの状態が  $z$  方向を対角化する表現での量子スピンの向きに対応しており、スピン状態は可視ユニットの状態  $|x\rangle = |v_1, v_2, \dots, v_N\rangle$  で表現できる。波動関数は、

$$\begin{aligned} \psi_\gamma(x) &= \sum_h e^{\sum_i a_i (2S_i^z) + \sum_j b_j h_j + \sum_{ij} W_{ij} (2S_i^z) h_j} \\ &= e^{\sum_i a_i (2S_i^z)} \times \prod_j 2 \cosh \left[ b_j + \sum_i W_{ij} (2S_i^z) \right] \end{aligned} \quad (7)$$

と表される (規格化因子は省略した)。ここで波動関数の値は変分パラメータ  $\gamma = \{a_i, b_j, W_{ij}\}$  に依存しており、これ

<sup>\*2</sup> 量子力学的には  $|\psi(x)|^2 = \psi(x)^2$  が確率分布だが、本稿でのRBM波動関数は  $\psi(x)$  そのものを確率分布と考えている。

らのパラメータをいかに最適化するかが問題になる。

### 3.2 学習 (= 変分パラメータの最適化) 方法

それでは、実際にどのような方針に従って、パラメータを最適化するのか？ 以下3つの例を考える。

i) もし、すべての取り得る状態  $x$  に対して、真の波動関数の値 (= 真の生成確率分布)  $\Psi_{\text{true}}(x)$  がわかっている場合、それを教師にして、RBM 波動関数が与える分布を真の生成確率分布に可能な限り近づけることができる。この学習は、例えば、カルバック・ライブラー情報量  $D(\Psi_{\text{true}}\|\psi_\gamma) = \sum_x \Psi_{\text{true}}(x) \log(\Psi_{\text{true}}(x)/\psi_\gamma(x))$  を最小化することによって実現する。<sup>\*3</sup> すなわち教師あり学習である。

ii) 真の波動関数はわからないものの、真の波動関数の確率分布に従って生成された実空間配置  $x$  の (波動関数の次元よりも少ない数の) サンプルのセットが与えられたとする。例えば、反強磁性的な波動関数であれば反強磁性的な実空間スピン配置が多く観測されるであろう。このような場合は、生成された配置から、それが従う真の確率分布を当てるという密度推定を行うことになる。これは教師なし学習として分類される。

iii) 真の波動関数を知らず、その真の確率分布に従う実空間配置のサンプルのセットも得ることができない。3つの例の中で一番挑戦的で、本稿の主題の多体系の基底状態を求めるという課題が典型例である。この場合の学習 (= パラメータの最適化) 法は変分波動関数法として物性物理の分野で培われてきた。本手法では、その学習法に学習機械 (ボルツマンマシン) が融合し、機械学習を遂行する。具体的な最適化は変分原理に基づく。変分原理は、ヒルベルト空間の任意の状態  $|\psi\rangle$  に対し、そのエネルギー期待値は基底状態エネルギー  $E_{\text{GS}}$  より小さくなることはないことを保証する:  $\langle\psi|\mathcal{H}|\psi\rangle/\langle\psi|\psi\rangle \geq E_{\text{GS}}$  (等号は  $|\psi\rangle$  が基底状態の時に成立)。RBM を使って量子状態  $|\psi\rangle$  を書き表すと [式 (7)], そのエネルギーは変分パラメータ  $\gamma = \{a_i, b_j, W_{ij}\}$  に依存する。従って基底状態を探す問題は、波動関数のエネルギーを一番下げるような変分パラメータのセットを見つけるという最適化問題に帰着する。

### 3.3 変分原理に基づく RBM パラメータの最適化の詳細

では上述の iii) の場合のパラメータの最適化をもっと詳細に見ていこう (数値手法詳細に興味のない方はこの節を読み飛ばしても概略を掴める)。RBM 波動関数  $|\psi_\gamma\rangle$  [式 (7)] のエネルギー期待値  $E_\gamma = \langle\psi_\gamma|\mathcal{H}|\psi_\gamma\rangle/\langle\psi_\gamma|\psi_\gamma\rangle$  は

$$E_\gamma = \sum_x p_\gamma(x) E_\gamma^{\text{loc}}(x) \quad (8)$$

と完全性の関係 ( $\sum_x |x\rangle\langle x| = 1$ ) を挿入して書き直すことで数値的に求める。重み  $p_\gamma(x)$  は  $p_\gamma(x) = |\langle x|\psi_\gamma\rangle|^2/\langle\psi_\gamma|\psi_\gamma\rangle$ ,

局所エネルギー  $E_\gamma^{\text{loc}}(x)$  は  $E_\gamma^{\text{loc}}(x) = \sum_x \langle x|\mathcal{H}|x\rangle \langle\langle x'|\psi_\gamma\rangle/\langle x|\psi_\gamma\rangle\rangle$  で与えられる。ただし、ヒルベルト空間の次元が指数関数的に大きくなると、 $x$  の和は厳密に計算できないので、 $p_\gamma(x) \propto |\psi_\gamma(x)|^2$  に比例する確率で実空間配置  $x$  を発生させる重み付きモンテカルロサンプリングに置き換える。その上でこのエネルギー期待値を最小化するように、変分パラメータを最適化する。数値最適化法に関してはいくつも方法が提案されているが、本研究では確率的再配置 (SR: stochastic reconfiguration) 法<sup>31)</sup>を採用したので、それについて簡単に紹介する。

SR法では、最適化ステップの際、 $N_\gamma$  個の変分パラメータ  $\gamma_m$  ( $m = 1, 2, \dots, N_\gamma$ ) (RBM 波動関数の場合  $\gamma_m$  は  $a_i, b_j, W_{ij}$  がそれぞれにあたる) の更新量  $\Delta\gamma_m$  ( $\gamma_m \rightarrow \gamma_m + \Delta\gamma_m$ ) は微小虚時間幅 (この用語の意味は後述) を  $\Delta\tau$  として、

$$\sum_{n=1}^{N_\gamma} S_{mn} \Delta\gamma_n = -\Delta\tau \mathbf{g}_m \quad (\Delta\gamma = -\Delta\tau S^{-1} \mathbf{g}) \quad (9)$$

で与えられる。  $S$  は正定値行列で  $S_{mn} = \langle\partial_{\gamma_m} \bar{\psi}_\gamma | \partial_{\gamma_n} \bar{\psi}_\gamma\rangle$  と規格化された波動関数  $|\bar{\psi}_\gamma\rangle = |\psi_\gamma\rangle/\sqrt{\langle\psi_\gamma|\psi_\gamma\rangle}$  とその変分パラメータによる偏微分  $|\partial_{\gamma_m} \bar{\psi}_\gamma\rangle = (\partial/\partial\gamma_m)|\bar{\psi}_\gamma\rangle$  で定義される。  $\mathbf{g}$  ベクトルは変分パラメータに対するエネルギー勾配で  $\mathbf{g}_m = \partial E_\gamma/\partial\gamma_m = (\partial/\partial\gamma_m)\langle\bar{\psi}_\gamma|\mathcal{H}|\bar{\psi}_\gamma\rangle$  と与えられる。  $S$  行列や  $\mathbf{g}$  ベクトルは、エネルギー同様にモンテカルロ法によって各最適化ステップ毎に数値的に求める (計算法の詳細は他に譲る<sup>32,33)</sup>)。なお、参考までに、式 (9) における  $S$  行列の部分が単位行列の場合が最急降下法、エネルギーのヘッセ行列の場合がニュートン法に対応する。

SR法の性質を考察する。変分パラメータの微小変化が加わった波動関数  $|\bar{\psi}_{\gamma+\Delta\gamma}\rangle = |\bar{\psi}_\gamma\rangle + \sum_m \Delta\gamma_m |\partial_{\gamma_m} \bar{\psi}_\gamma\rangle$  ともとの波動関数  $|\bar{\psi}_\gamma\rangle$  の違いのノルムの二乗が  $\| |\bar{\psi}_{\gamma+\Delta\gamma}\rangle - |\bar{\psi}_\gamma\rangle \|^2 = \sum_{mn} S_{mn} \Delta\gamma_m \Delta\gamma_n$  で与えられることから、  $S$  行列は変分波動関数の空間における距離を定義する計量テンソルになっている。よって、式 (9) において  $S^{-1}$  を作用させるということは、  $\gamma_m$  の微小変化によって波動関数が大きく変わってしまうような場合、  $\gamma_m$  の更新の値を小さくする (= 波動関数が大きく変わらない) ようにして数値的不安定化を抑制していることがわかる。また、SR法は変分波動関数が張ることのできる空間内で、ハミルトニアンによる微小虚時間発展<sup>\*4</sup>した状態  $(1 - 2\mathcal{H}\Delta\tau)|\psi\rangle$  と、更新された試行波動関数の差が一番小さくなるような最適化法であることもわかっており、一意的に基底状態に漸近できる。<sup>34),\*5</sup>

実際には式 (9) に基づく最適化を繰り返し (典型的には数千回程度)、変分パラメータを最適化する。初期変分パラメータ  $a_i, b_j, W_{ij}$  は RBM 波動関数にバイアスをかけないよう、小さな乱数を振る場合が多い。幾つかの乱数のセッ

\*3 カルバック・ライブラー情報量  $D(\Psi_{\text{true}}\|\psi_\gamma)$  と波動関数の重なるの差を比較する。波動関数の重なるの差は  $\langle\Psi_{\text{true}}|\Psi_{\text{true}}\rangle - \langle\Psi_{\text{true}}|\psi_\gamma\rangle = \sum_x (\Psi_{\text{true}}(x)\Psi_{\text{true}}(x) - \Psi_{\text{true}}(x)\psi_\gamma(x))$  なのに対し、カルバック・ライブラー情報量は  $\sum_x (\Psi_{\text{true}}(x) \log(\Psi_{\text{true}}(x)) - \Psi_{\text{true}}(x) \log(\psi_\gamma(x)))$  で与えられる。両者とも値は非負であることが示せ、  $\Psi_{\text{true}}$  と  $\psi_\gamma$  が厳密に等しい時のみ0になる。

\*4 基底状態と直交しない状態を初期状態  $|\psi_0\rangle$  として、ハミルトニアンによる虚時間発展  $e^{-\mathcal{H}t}|\psi_0\rangle$  をすると充分長い虚時間  $\tau$  で基底状態に到達する。  $2\Delta\tau$  だけの微小虚時間発展  $e^{-2\mathcal{H}\Delta\tau}$  は  $1 - 2\mathcal{H}\Delta\tau$  と近似できる。  
\*5 微小虚時間発展を誤差を含む数値計算によって行うため、どれだけ基底状態に漸近できるかは例えば初期状態の波動関数に依存してしまう。

トを試し、一番エネルギーの下がった波動関数を採用する。

実は、物性物理の分野で提案されたSR法とまったく同じ最適化法（計量テンソルを利用した最適化）が機械学習の分野でも甘利らによって提案されており、自然勾配法と呼ばれている。<sup>35,36</sup> 物理・機械学習の分野の交流がこれまで疎かったため、お互いの分野で独立に最適化法が提案されたことになる。

### 3.4 物性物理から見た本手法の位置付け

本手法におけるSR法などのパラメータの最適化法や物理量計算の仕方は変分モンテカルロ（VMC: variational Monte Carlo）法<sup>37</sup>とほとんど同じである。VMC法との違いは、試行波動関数の形にある。これまでのVMC法は例えば、反強磁性や超伝導などの平均場的な解も表せるような一般的な一体ハミルトニアン<sup>38</sup>の波動関数（スレーター行列と呼ばれる）や二体波動関数（一般化されたBCS（Bardeen-Cooper-Schrieffer）波動関数）から出発し、これにグッツヴィラー因子と呼ばれる電子相関を取り入れる因子などを作用させて、それらが含む変分パラメータをエネルギーが最小化するように最適化していた。このように物理的洞察に基づく波動関数を用いており、パラメータが物理的直観を表現してはいるが、その最適化によってどこまで良い近似になるか、どう系統的に改良しうるかはずしも明確でなかった。それに対し、RBM波動関数は物理的直観と直接結びつかないユニバーサルな表現能力を持つ関数で、物理のハミルトニアンを学習することによって自らの関数形を整えていく。3.2節でも触れたが、物性物理ではすでに機械学習のうちの、学習のノウハウに関して技術を持っており、それに学習専用の機械が加わったという解釈を我々はしている。将来的には、機械学習をホワイトボックス化することで新たな物理的直観が生まれる可能性もある。

### 3.5 計算手順のまとめ

1. RBMを用いて試行波動関数を構成する[式(7)]. 不可視ユニットの数 $M$ を決め、初期パラメータ $a_i, b_j, W_{ij}$ を値の小さい乱数に取る。
2. 変分パラメータ $a_i, b_j, W_{ij}$ を変分原理に基づき、エネルギーを最小化するように最適化する。その際SR法などを使用する。複数の初期状態に対し最適化を行い、最低エネルギーを与える波動関数を採用する。
3. 最適化された波動関数をもとに物理量を計算する。エネルギー以外の物理量 $\langle O \rangle$ も式(8)において $\mathcal{H}$ を $O$ に置き換えることで求められる。

## 4. RBM波動関数の性質及び数値結果

前節ではRBM波動関数の手法を紹介した。ここではRBM波動関数の能力や数値結果について議論する。

### 4.1 RBM波動関数の性質

RBM以外にも、 $N_{\text{site}}$ に対して冪関数で表される数の変分パラメータを用いて基底状態の高精度の表現を得ようという研究が進んでいる。そのような研究の中でも、特にテ

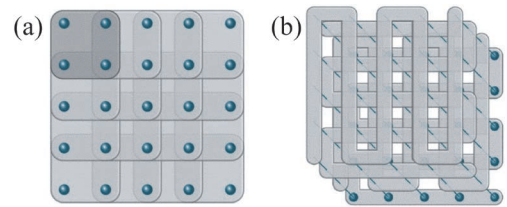


図2 テンソル・ネットワークの波動関数. (a) エンタングルド・プラケット状態 (EPS). (b) スtring・ボンド状態 (SBS). 文献16より転載. Copyright 2018 by American Physical Society.

ンソル・ネットワークとの比較は興味深い。そのため、まず準備としてテンソル・ネットワークを簡単に紹介する。

#### 4.1.1 テンソル・ネットワーク

テンソル・ネットワークの波動関数は、物理自由度の脚を持つたくさんのテンソルの間の縮約<sup>\*6</sup>を取ることで状態を表現する。詳しい説明は他の記事に譲るとして、<sup>38,39</sup> 具体的な例を見ることでその雰囲気を味わおう。

**1. 行列積状態 (MPS: matrix product states).** 波動関数を $\psi(x) = \text{Tr}(\prod_{i=1}^N A^{(i)}[v_i])$ ,  $|x\rangle = |v_1, v_2, \dots, v_N\rangle$ と表す。 $A^{(i)}[v_i]$ は可視ユニット（物理自由度）の状態 $v_i$ に依存した $D$ 行 $D$ 列の行列である。それらの積 $\prod_{i=1}^N A^{(i)}[v_i]$ も $D$ 次元行列となり、その対角和（トレース）を波動関数とみなす。 $D$ をボンド次元と呼び、 $D \rightarrow \infty$ の極限ではどんな波動関数も厳密に表現できる。行列積法は密度行列繰り込み群法（DMRG: density matrix renormalization group）<sup>40</sup>と等価である。

**2. エンタングルド・プラケット状態 (EPS: entangled-plaquette states).**<sup>41,42</sup> 図2(a)のように、システムを $n_p$ 個の物理自由度で構成される $P$ 個の小さなプラケット（灰色の領域）で敷き詰めて、そのプラケット上での可視ユニット（物理自由度）の状態 $\mathbf{v}_p = (v_{p1}, \dots, v_{pn_p})$ に依存した係数 $C^{(p)}[\mathbf{v}_p]$ の積（ボンド次元1の行列の積）によって波動関数を表す。

$$\psi(x) = \prod_{p=1}^P C^{(p)}[\mathbf{v}_p] \quad (10)$$

係数 $C^{(p)}[\mathbf{v}_p]$ は $\mathbf{v}_p$ の全ての状態に対して独立に定義されるので、プラケット内の自由度の数 $n_p$ は大きくは取れない。

**3. スtring・ボンド状態 (SBS: string-bond states).**<sup>43</sup> SBSはMPSの積として表される。図2(b)のように、<sup>\*7</sup> システム上に複数のstring（灰色の紐）を用意し、各string上でMPSを構成し、その積として波動関数を以下のように定義する。

$$\psi(x) = \prod_j \text{Tr} \left( \prod_{i \in S_j} A^{(i,j)}[v_i] \right) \quad (11)$$

ここで $S_j$ は $j$ 番目のstringが通るサイトの集合を表している。

\*6 テンソルの積を縮約ともいう。例えば $C_{ik} = \sum_j A_{ij} B_{jk}$ における $j$ に関する和はテンソル $A, B$ に含まれる情報自由度を圧縮している。

\*7 この例では各stringは全てのサイトを通るようになってはいるが、必ずしも全サイトを通る必要はなく短いstringがあっても良い。

これ以外にもテンソル・ネットワークの波動関数には様々な種類がある。テンソルがネットワークを組んでその縮約によって波動関数が表現される場合は一般にテンソル・ネットワーク法に分類される。

#### 4.1.2 性質及びテンソル・ネットワークとの比較

1. 普遍的な関数近似機械. RBM 波動関数 [式(7)] はパラメータ  $a_i, b_j, W_{ij}$  を複素数に取ることによって、複素振幅を持つ波動関数を表現できる。また、RBM 波動関数は関数近似能力が高く、無限個 (正確には可視ユニットに対して指数関数的に増大する数) の不可視ユニットを用意すれば、どんな波動関数  $\psi(x)$  も任意の精度で表現可能である。<sup>17,23)</sup> 同様にテンソル・ネットワークも指数関数的にパラメータを増やすと任意の波動関数を厳密に表現できる。

2. 短距離型 RBM 波動関数: EPS と等価であること, エンタングルメント・エントロピーが面積則に従うこと. 不可視ユニットが全ての可視ユニットと結合  $W_{ij}$  を持つのではなく一部の隣接したサイト (サイト数が  $\mathcal{O}(1)$ ) とのみ結合を持つ場合, これを短距離型 RBM と呼ぶことにする.  $j$  番目の不可視ユニットが結合を持つ可視ユニットサイトの集合を  $p_j$  とすると, 短距離型 RBM 波動関数は (単に一体のポテンシャル項であるバイアス  $a_i$  は簡単のため無視する)

$$\psi(x) = \prod_j 2 \cosh \left[ b_j + \sum_{i \in p_j} W_{ij} v_i \right] = \prod_j C^{(j)}(\mathbf{v}_j) \quad (12)$$

と書き表すことができる。ここで  $\mathbf{v}_j$  は  $j$  番目の不可視ユニットと結合している可視ユニットの状態を表す。式(10)と見比べると、短距離型 RBM 波動関数は EPS の一種として分類することができるがわかる [図 2(a)].<sup>16)</sup> この時、系を 2 つの空間的に分かれた部分系に分けた際にそれらの間の量子もつれの強さを表す指標であるエンタングルメント・エントロピーは表面則に従うことが示せる。<sup>16)</sup>

3. 長距離型 RBM 波動関数: SBS と等価であること, エンタングルメント・エントロピーが体積則に従うこと. 不可視ユニットが  $\mathcal{O}(N_{\text{site}})$  の数の可視ユニットと結合を持つ場合これを長距離型 RBM と呼ぼう。長距離型 RBM の波動関数は  $j$  番目の不可視ユニットが結合を持つ  $n_j$  個の可視ユニットサイトの集合を  $S_j$  とし (再びバイアス  $a_i$  は簡単のため無視する),

$$\psi(x) = \prod_j \text{Tr} \left[ \prod_{i \in S_j} A^{(i,j)}[v_i] \right] \quad (13)$$

と書き直すことができる。ここで,

$$A^{(i,j)}[v_i] = \begin{pmatrix} e^{b_j/n_j + W_{ij}v_i} & 0 \\ 0 & e^{-b_j/n_j - W_{ij}v_i} \end{pmatrix}. \quad (14)$$

式(11)と見比べて、長距離型 RBM 波動関数はボンド次元  $D=2$  の SBS に分類することができるがわかる。<sup>16)</sup>

長距離型 RBM が表す状態のエンタングルメント・エントロピーは体積則に従うことがわかっている。<sup>16)</sup> ボンド次元が  $D$  の MPS が表すことのできるエンタングルメント・エントロピーは最大で  $\ln D$  であることを考えると、このような体積則のエンタングルメント・エントロピーを持つ長

距離型 RBM (SBS: MPS の積) を単一の MPS で表そうとすると必要なボンド次元はシステムサイズ  $N_{\text{site}}$  に対して指数関数的に大きなボンド次元が必要となることがわかる。<sup>15)</sup>

#### 4.1.3 テンソル・ネットワークに対する優位性

テンソル・ネットワークは主にエンタングルメント・エントロピーが表面則に従う状態を対象にしている (典型的には MPS によってうまく表される, 空間 1 次元系で基底状態からの励起エネルギーにギャップのある系). それに対し, RBM が表現することのできる量子もつれの度合いは, 長距離型の場合体積則を許容するので, テンソル・ネットワークでは表現に指数関数的に大きなボンド次元  $D$  が必要な状態を効率的に表すことができる可能性がある。

ただ, どちらが少ないパラメータで効率的に量子状態を表現できるかは, 場合によると思われる。例えば, Affleck-Kennedy-Lieb-Tasaki (AKLT) 状態<sup>44)</sup> は  $D=2$  の MPS で効率的に表すことができるものの,<sup>16)</sup> RBM で AKLT 状態を表そうとすると隠れた反強秩序があるために長距離型の RBM が必要になる。<sup>15)</sup> 従ってエンタングルメント・エントロピーだけが表現能力を決めるわけではない。一方, 長距離のジャストロー型の相関を持つ形で表される (格子上で定義された) ラフリン波動関数は長距離型 RBM を使うことによって効率的に表現が可能である。<sup>16,17)</sup> その他にも RBM 波動関数を用いた, グラフ状態,<sup>17)</sup> カイラル  $p$  波超伝導状態,<sup>23)</sup> 1 次元の対称性によって守られたトポロジカルクラスター状態,<sup>13)</sup> 2・3 次元のトーリックコード状態<sup>15,17)</sup> などの効率的な表現が議論されている。

### 4.2 RBM 波動関数のベンチマークと応用例

#### 4.2.1 簡単な例: 1 次元ハイゼンベルグ模型 (8 サイト)

まず, 1 次元  $S=1/2$  の 8 サイトスピン鎖上の反強磁性ハイゼンベルグ模型の結果を示す。ハミルトニアンは  $\mathcal{H} = J \sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+1}$  ( $J > 0$ ) で表される。bipartite な格子上的フラストレーションのないスピン模型の場合, 格子を A 副格子と B 副格子に分割でき (1 次元ではそれぞれ偶数サイトと奇数サイト), どちらか一方のスピン量子化軸を  $z$  軸の周りに 180 度回転させると, 基底状態波動関数の振幅を正定値にできる (マーシャル符号則)。今の場合, 奇数サイトの量子化軸を回転させると (ゲージ変換に対応), ハミルトニアンは  $\mathcal{H} = J \sum_i (-S_i^x S_{i+1}^x - S_i^y S_{i+1}^y + S_i^z S_{i+1}^z)$  と等価であり, この変換で基底状態波動関数を正定値にできるので, RBM の変分パラメータを実数に取ることができる。

可視ユニットの数はサイト数と等しく ( $N=8$ ), その状態は  $v_i = 2S_i^z$  と定義する。今回の計算では, 不可視ユニットの数もサイト数と等しく取り ( $M=8$ ), バイアス  $a_i, b_j$  を無視し (そうするとスピンに関して上向きと下向きとの対称性が自然と満たされる),  $W_{ij}$  のパラメータに対しては並進対称性を課している。その際, 独立な  $W_{ij}$  ( $=W_{i-j}$ ) の数は並進対称性より  $64 (=M \times N)$  から  $8 (=M)$  まで減る。

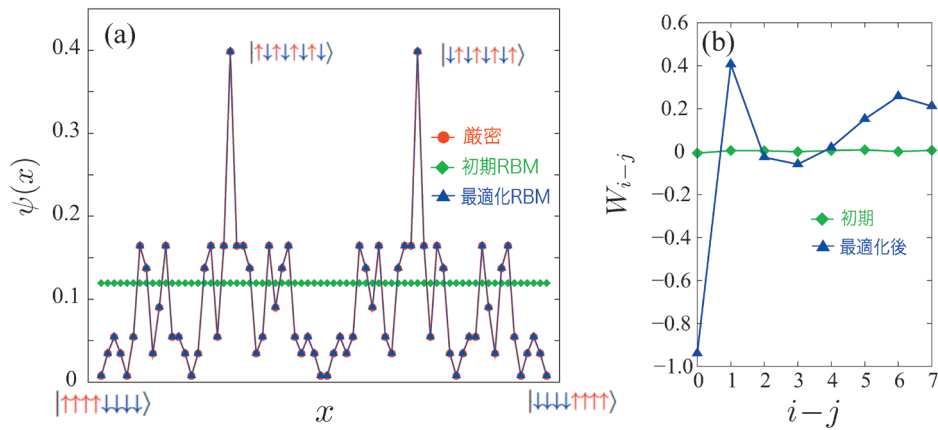


図3 1次元反強磁性ハイゼンベルグ模型 (8 サイト, 周期境界条件) に対する (a) 厳密な波動関数とRBM波動関数の比較. 厳密解は最適化RBM解と重なっている. (b) RBMの変分パラメータ  $W_{i-j}$  の値.

図3(a)にRBMの初期状態, 及び最適化された波動関数の結果を示す.\*<sup>8</sup> 磁化が0の量子数セクターのヒルベルト空間の次元は70なので, 容易に厳密対角化できるが, あくまで簡単な例として取り上げる. 初期の  $W_{i-j}$  のパラメータは小さい乱数を取っているためにRBMの初期状態は特徴のない関数形となる.\*<sup>9</sup> それが学習によって基底状態を学んだ結果, 厳密な基底状態を再現していることがわかる. 最適化前後の  $W_{i-j}$  のパラメータを図3(b)に示す.\*<sup>10</sup>  $i-j$  が小さいところで  $W$  の符号反転が見え, 反強磁性な揺らぎを学んでいる.

#### 4.2.2 カルレオ・トロイヤーの数値結果

カルレオとトロイヤーは1次元の横磁場イジング・反強磁性ハイゼンベルグ模型, 及び2次元正方格子上の反強磁性ハイゼンベルグ模型にRBM波動関数を適用した.<sup>11)</sup> これらは全て bipartite 格子上でフラストレーションのないスピン模型のため, ゲージ変換により基底状態波動関数が正定値となる. このことを利用して, 変分パラメータは実数に取っている. また相互作用  $W$  のパラメータにも並進対称を課している (そのため不可視ユニットの数  $M$  は可視ユニットの数  $N$  の整数倍となる).  $\alpha = M/N$  と定義すると,  $\alpha$  がRBM波動関数の精度を制御するパラメータとなり,  $\alpha$  が大きいところで精度が増していく. 代表例として2次元ハイゼンベルグ模型の結果を図4に示す.  $\alpha$  が比較的小さい値の時点ですでにEPSやPEPS (projected entangled-pair states) などのテンソル・ネットワークの精度を上回るような精度が得られている.

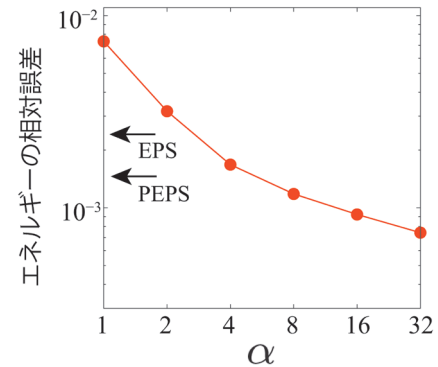


図4 2次元正方格子上の反強磁性ハイゼンベルグ模型 (10×10, 周期境界条件) に対するRBM波動関数のエネルギーの厳密解からの相対誤差. モンテカルロ法によって得られた数値的に厳密なエネルギーと比較している.  $\alpha = M/N$  ( $M$ : 不可視ユニット数,  $N$ : 可視ユニット数). EPS, PEPSはそれぞれテンソルネットワークによる結果. データは文献11による.

### 5. 機械学習と物理が融合した波動関数 (RBM+PP)

この節と次節ではRBM波動関数の精度向上について議論する. この節では  $S=1/2$  スピン系だけではなくフェルミオン系への適用も議論するので, まず可視ユニットの状態  $x = (v_1, \dots, v_N)$  の定義をしておく.  $S=1/2$  スピン系はこれまで同様, 可視ユニットの数が  $N = N_{\text{site}}$  で,  $v_i = 2S_i^z = \pm 1$  と定義する. 格子上で定義されたフェルミオン系においては  $N = 2N_{\text{site}}$  とし, あるサイト  $l$  のスピン  $\sigma$  の状態に対する数演算子を  $n_{l\sigma}$  として  $(v_{2l-1}, v_{2l}) = (2n_{l\uparrow} - 1, 2n_{l\downarrow} - 1)$  と定義する. すなわちサイト  $l$ , スピン  $\sigma$  電子がいるいないで  $\pm 1$  の値を返すように定義する.\*<sup>11</sup>

#### 5.1 RBM+PP 波動関数

柔軟かつ不偏なRBMと, 実空間の非局所的なエンタングルメントを変分パラメータ  $f_{lm}^{\uparrow\downarrow}$  を通じて効率的に取り込むことのできるペア積\*<sup>12</sup> (PP: pair-product) 状態  $|\phi_{\text{PP}}\rangle = (\sum_{l,m=1}^{N_{\text{site}}} f_{lm}^{\uparrow\downarrow} c_{l\uparrow}^\dagger c_{m\downarrow}^\dagger)^{N_e/2} |0\rangle$  ( $N_e$ : 電子数) を組み合わせたRBM+PP波動関数を以下のように導入する<sup>14)</sup> (図5):

\*<sup>8</sup> <https://sites.google.com/view/yusuke-nomuras-webpage-en/home> にてプログラムを公開予定.

\*<sup>9</sup> 全ての  $W$  が厳密に0のとき, RBM初期波動関数はエンタングルメントの全くないプロダクト状態となる  $|\psi\rangle = (|\uparrow\rangle + |\downarrow\rangle)_1 \otimes (|\uparrow\rangle + |\downarrow\rangle)_2 \otimes \dots \otimes (|\uparrow\rangle + |\downarrow\rangle)_g$ . この状態は  $x$  方向にスピンの揃った強磁性状態であり, ゲージ変換する前のスピン量子化軸だと  $x$  軸方向の古典的反強磁性状態となる. よって, 反強磁性模型に対しては悪くない初期状態になっている.

\*<sup>10</sup> 並進対称性があり, かつ不可視ユニットのサイトの位置も任意であるため  $i-j$  の原点は任意だが,  $W$  の絶対値が一番大きくなったところを  $i-j=0$  としている.

\*<sup>11</sup> 定義の仕方には任意性があるが, 一番単純な定義を採用した.

\*<sup>12</sup> 化学の分野ではジェミナルと呼ばれている.

$$\psi_T(x) = \mathcal{N}(x)\phi_{\text{ref}}(x) \quad (15)$$

ここで $|\phi_{\text{ref}}\rangle$ はスピン系に対しては二重占有度を禁止しスピン模型のヒルベルト空間に状態を制限したもの $|\phi_{\text{ref}}\rangle = P_G^\infty|\phi_{\text{PP}}\rangle$  [ $P_G^\infty = \prod_i (1 - n_{i\uparrow}n_{i\downarrow})$ ]を採用する。フェルミオン系に対しては $|\phi_{\text{ref}}\rangle = |\phi_{\text{PP}}\rangle$ とペア積状態をそのまま使う。この $\phi_{\text{ref}}(x)$ は物理的洞察に基づいて構成され、スピン系ではRVB (resonating valence bond) 状態<sup>45)</sup>を記述できたり、フェルミオン系では反強磁性や超伝導の平均場解を記述することもできる柔軟な関数である。 $\mathcal{N}(x)$ はRBMによる相関因子で $\mathcal{N}(x) = e^{\sum_i a_i v_i} \times \prod_j 2 \cosh [b_j + \sum_i W_{ij} v_i]$ で与えられる。 $v_i$ の定義はスピン系とフェルミオン系で違うことに注意(上述)。ペア積状態がフェルミオンの波動関数が満たすべき反対称性を満たしてくれるので、RBM+PPはフェルミオン系にも適用可能な柔軟な関数形になっている。

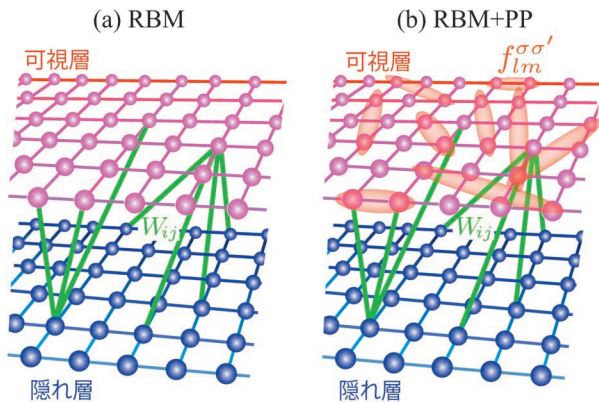


図5 (a) RBM [式(7)] (b) RBM+PP波動関数[式(15)]の概念図。RBM, PPはそれぞれ機械学習、物理的関数。 $W_{ij}$ はRBMにおける可視層と不可視層の間の結合の変分パラメータ。RBM+PPではRBMと、可視層の間の量子相関を直接取り込むことのできるPP波動関数を組み合わせている。 $f_{lm}^{\sigma'}$ はPP波動関数に含まれる変分パラメータ。 $W_{ij}$ や $f_{lm}^{\sigma'}$ は一部のものしか表示していない。文献14より転載。Copyright 2017 by American Physical Society.

## 5.2 RBM 相関因子と従来の相関因子との違い

従来の波動関数法である多変数変分モンテカルロ (mVMC: many-variable VMC) 法<sup>32)</sup>では、式(15)における $|\phi_{\text{ref}}\rangle$  (ペア積状態や、スレーター行列式状態などで与えられる)に対して、グッツヴィラー因子(二重占有度を制御)<sup>46)</sup>・ジャストロー因子(電荷の長距離相関)<sup>47)</sup>等の物理的洞察に基づいた相関因子を作用させた変分関数を使用している。RBM+PPでは、mVMC波動関数のグッツヴィラー・ジャストロー相関因子の部分をもっと柔軟かつ不偏なRBM相関因子に置き換えたという見方を取ることができる。グッツヴィラー・ジャストロー因子は可視ユニット間の直接2体相互作用として与えられる。RBM因子は不可視ユニットが可視ユニット間の2体相互作用を媒介することによりこれらの2体相関を厳密に表せることに加え、3体、4体、5体相関等も取り込むことができるより強力な相関因子になっている。<sup>\*13</sup>

## 5.3 RBM+PP 波動関数の数値結果

図6に2次元ハイゼンベルグ模型とハバード模型の結果を示す。○のデータがRBM部分だけに変分パラメータがあるRBM波動関数で、それに対してRBMとPPを組み合わせたRBM+PP(△,▽)による精度の向上が確認された。<sup>14)</sup>物理により予め重要な相関を取り込むことにより機械学習が容易になったと考えられる。また、ハバード模型においてmVMCとRBM+PP波動関数を比較すると(違いはグッツヴィラー・ジャストロー因子を使うかRBM因子を使うか)、数値的にもRBM相関因子がグッツヴィラー・ジャストロー因子より強力であることが検証された。

## 6. DBM を用いた基底状態波動関数の厳密な構築

最後にDBMを用いてRBM波動関数を改善する方法を議論する。一般に基底状態と直交しない何らかの初期状態 $|\psi_0\rangle$ に対し、ハミルトニアン $\mathcal{H}$ による充分長い虚時間発

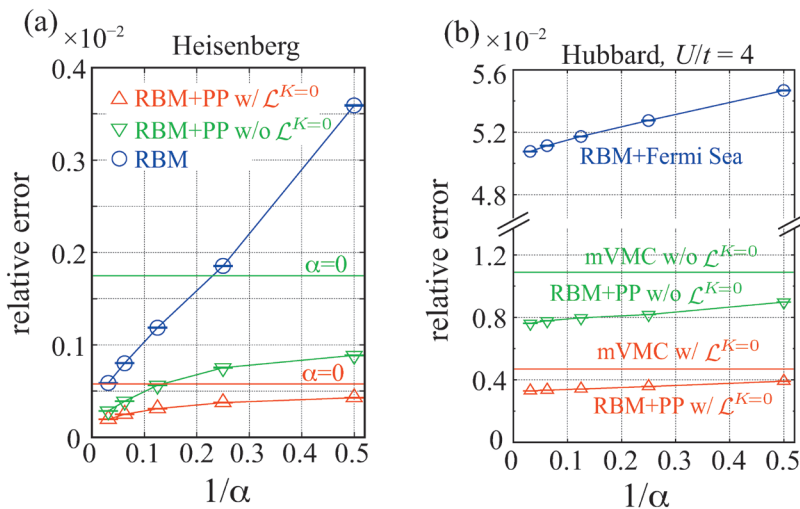


図6 (a) 2次元正方格子上のハイゼンベルグ模型(8×8, 周期境界条件), (b) ハバード模型(8×8, x方向に周期境界, y方向に反周期境界条件, ハーフフィリング,  $U/t=4$ )のRBM及びRBM+PP波動関数のエネルギー。モンテカルロ法により得られた数値的厳密エネルギーからの相対誤差を示している。 $\alpha=M/N$  ( $M$ : 不可視ユニット数,  $N$ : 可視ユニット数)。 $\mathcal{L}^{K=0}$ は総運動量が0の量子数への射影演算子で、 $\mathcal{L}^{K=0}$ をPP状態に作用させることにより、精度が向上する(△と▽のデータの違いは $\mathcal{L}^{K=0}$ の有無)。文献14より転載。Copyright 2017 by American Physical Society.

<sup>\*13</sup> 隠れユニット $h_j$ が可視ユニット $v_k, v_l, v_m$ と相互作用すれば、 $h_j$ の状態和を取ると $\cosh [W_{kj}v_k + W_{lj}v_l + W_{mj}v_m]$ の因子を通じて $v_k, v_l, v_m$ 間の3体相互作用が生じる。



展  $e^{-\tau\mathcal{H}}|\psi_0\rangle$  によって基底状態に到達することができる。3.3節で見たように、SR法によるRBMの波動関数の最適化はRBM波動関数が張ることのできるヒルベルト空間の中で、可能な限りこの虚時間発展を再現する手法であった。しかしながら、DBMはRBMに比べて、確率分布の表現能力が大きく向上するという長所がある(2.2節参照)。この節での一番重要なアイデアは、この表現能力の向上を利用して、鈴木・トロッター分解された虚時間発展を厳密にDBMで追うことができるということである。<sup>48)</sup> 虚時間発展を続けることによって、基底状態を表すDBM波動関数が得られる。この際、虚時間発展をする度にDBMの変分パラメータが変わることになるが、その変化は解析的に求められる。RBMの場合、系のサイズに対して多項式的な数の隠れユニットを用意しただけでは虚時間発展を近似的にしか追えない上に、変分パラメータを数値的に最適化する必要があったのは対照的である。ただし、DBMでは隠れユニット  $h_j$  と深層ユニット  $d_k$  の自由度を同時に解析的にトレースアウトすることができないという短所がある(2.2節参照) ために、物理量を計算する場合は可視ユニット  $v_i$  の状態に加えて、 $h_j$  と  $d_k$  (もしくはそのどちらか) の状態もモンテカルロ法でサンプルする必要がある。

もう少し具体的に見てみよう。鈴木・トロッター公式<sup>14</sup> :  $e^{-\tau\mathcal{H}} \sim (e^{-\delta\tau\mathcal{H}_1} e^{-\delta\tau\mathcal{H}_2})^{N_{\text{slice}}}$ , ( $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$ ,  $\delta\tau = \tau/N_{\text{slice}}$ ) を用いると、DBMで表される任意の波動関数  $|\psi_\gamma\rangle$  に対して、微小虚時間発展させた状態  $e^{-\delta\tau\mathcal{H}_v}|\psi_\gamma\rangle$  を新たな変分パラメータ  $\bar{\gamma}$  を持ったDBM状態  $|\psi_{\bar{\gamma}}\rangle$ <sup>15</sup> で表すことができれば、それを続けることによって虚時間発展が再現できる。すなわち、解くべき式は

$$e^{-\delta\tau\mathcal{H}_v}|\psi_\gamma\rangle = C_v|\psi_{\bar{\gamma}}\rangle \quad (v=1, 2, C_v: \text{const.}) \quad (16)$$

紙面の制約上ここではパラメータ変化の具体的な式をあげることができないが、図7に横磁場イジング模型での式(16)を満たすためのDBMの変化の模式図を示してある。

式(16)の解法を次々と適用し、虚時間発展を続けると最終的に基底状態を記述するDBMネットワークが解析的に求められる。<sup>48)</sup> 微小虚時間発展の度に新たな隠れ・深層ユニットが追加されていき、最終的にその数は虚時間発展の長さで系のサイズそれぞれに比例する。

図8に1次元反強磁性ハイゼンベルグ模型のDBMによるエネルギーの虚時間発展を表している。隠れ・深層ユニットの数が0の初期状態(=プロダクト状態)から出発する場合は、基底状態に到達するには長い虚時間発展が必要だが、この手法では初期状態を工夫することも可能で、数値的に最適化されたRBM波動関数から虚時間発展を開始すると短い虚時間で基底状態に到達できる。<sup>48)</sup>

<sup>14</sup> 実際の計算にはより正確な  $e^{-\delta\tau\mathcal{H}} \sim e^{-(\delta\tau/2)\mathcal{H}_2} e^{-\delta\tau\mathcal{H}_1} e^{-(\delta\tau/2)\mathcal{H}_2}$  を用いているが、簡単のため最も単純な分解の式を示している。

<sup>15</sup> 既存の変分パラメータが変化する場合もあれば、新たに隠れ・深層ユニットが追加されそれに伴い新たな有限の結合が生成される場合もある。

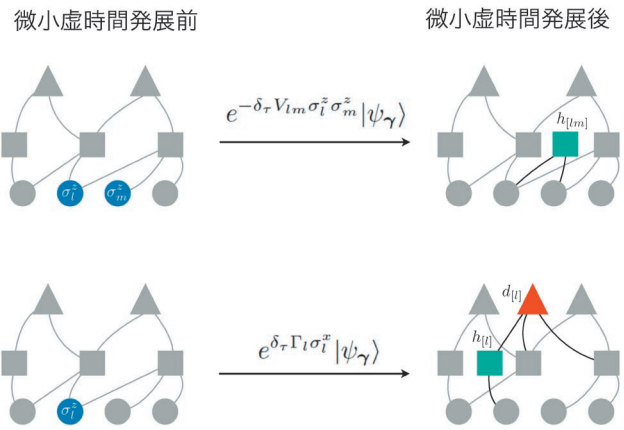


図7 横磁場イジング模型  $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$ ,  $\mathcal{H}_1 = \sum_{i < m} V_{im} \sigma_i^z \sigma_m^z$ ,  $\mathcal{H}_2 = -\sum_j \Gamma_j \sigma_j^x$  の微小虚時間発展によるDBMの構造の変化。物理的スピンの状態  $\sigma_i^z = \pm 1$  と可視ユニットの状態  $v_i = \pm 1$  が同一視されている。

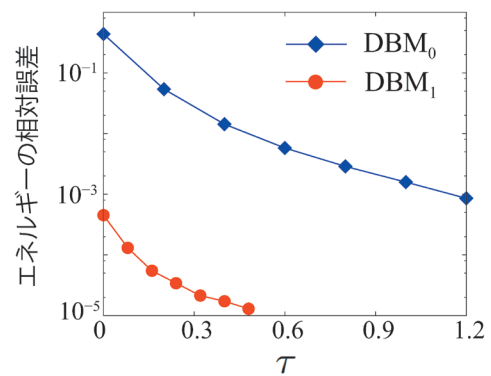


図8 1次元反強磁性ハイゼンベルグ模型(80サイト、周期境界条件)のDBM波動関数のエネルギーの虚時間発展。DBM<sub>0</sub>はプロダクト状態( $M = M' = 0$ )から、DBM<sub>1</sub>は数値的に最適化したRBM波動関数( $M/N = 1, M' = 0$ )を初期状態として虚時間発展させたもの。データは文献48による。

DBMは可視ユニットに加え、隠れ・深層ユニットからなる古典系なので、この手法は量子系の基底状態を古典系にマップする新たな量子古典対応を提供する。<sup>48), 16)</sup>  $D$ 次元の量子系が $D+1$ 次元の古典系に解析的にマップできるように、今回は深層が新たに加わったことによって、RBMではできなかった量子古典マッピングが可能になった。すなわち本手法においては新たな隠れ層(深層)は、統計力学における新たな次元のような役割を果たしていることがわかる。

## 7. おわりに

本稿ではこれまで物理で培ってきた学習のノウハウに機械学習の分野から万能関数近似機械であるボルツマンマシンを導入し、それらを融合することによって、これまででない精度での基底状態探索が可能になってきたことを解説した。最近では“NetKet”(https://www.netket.org)と呼ばれるRBM波動関数を基本とした機械学習による多体波動関数計算のためのオープンソースなども用意されているので

<sup>16</sup> 基底状態を構成するDBMの構成法は複数あり、その特別な場合、経路積分と等価になる。すなわちDBMは通常の経路積分による量子化を包含する、より広い一般化された量子化の枠組みを提供する。

興味ある方は是非試していただきたい。なお、本稿では詳しく触れることができなかったが、ボルツマンマシン以外の人工ニューラル・ネットワークを用いた波動関数なども試されており、<sup>19-22)</sup> 今後ますます強力な波動関数が提案されるものと期待される。また波動関数以外にも実験の観測のデータの特徴を抽出するのにボルツマンマシンを応用したり<sup>49)</sup> など、ボルツマンマシンの応用が期待できる分野は幅広い。これからの発展に期待したい。

本稿の内容の一部は Andrew S. Darmawan 氏、Giuseppe Carleo 氏との共同研究によるものである。

## 参考文献

- 1) J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
- 2) E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
- 3) K. Ch'ng et al., Phys. Rev. X **7**, 031038 (2017).
- 4) Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
- 5) 大槻東巳, パリティ **32**(7), 52 (2017).
- 6) T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
- 7) T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **86**, 044708 (2017).
- 8) A. Tanaka and A. Tomiya, J. Phys. Soc. Jpn. **86**, 063001 (2017).
- 9) F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
- 10) N. Yoshioka, Y. Akagi, and H. Katsura, Phys. Rev. B **97**, 205110 (2018).
- 11) G. Carleo and M. Troyer, Science **355**, 602 (2017).
- 12) D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).
- 13) D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B **96**, 195145 (2017).
- 14) Y. Nomura et al., Phys. Rev. B **96**, 205152 (2017).
- 15) J. Chen et al., Phys. Rev. B **97**, 085104 (2018).
- 16) I. Glasser et al., Phys. Rev. X **8**, 011006 (2018).
- 17) S. R. Clark, J. Phys. A: Math. Theor. **51**, 135301 (2018).
- 18) R. Kaubruegger, L. Pastori, and J. C. Budich, Phys. Rev. B **97**, 195136 (2018).
- 19) H. Saito, J. Phys. Soc. Jpn. **86**, 093001 (2017).
- 20) H. Saito and M. Kato, J. Phys. Soc. Jpn. **87**, 014001 (2018).
- 21) H. Saito, J. Phys. Soc. Jpn. **87**, 074002 (2018).
- 22) Z. Cai and J. Liu, Phys. Rev. B **97**, 035116 (2018).
- 23) Y. Huang and J. E. Moore, arXiv:1701.06246.
- 24) L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
- 25) L. Wang, Phys. Rev. E **96**, 051301 (2017).
- 26) D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, Cognitive Science **9**, 147 (1985).
- 27) P. Smolensky, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, edited by D. E. Rumelhart, J. L. McClelland, and PDP Research Group (MIT Press, Cambridge, 1986).
- 28) N. L. Roux and Y. Bengio, Neural Comput. **20**, 1631 (2008).
- 29) G. Montufar and N. Ay, Neural Comput. **23**, 1306 (2011).
- 30) X. Gao and L.-M. Duan, Nat. Commun. **8**, 662 (2017).
- 31) S. Sorella, Phys. Rev. B **64**, 024512 (2001).
- 32) D. Tahara and M. Imada, J. Phys. Soc. Jpn. **77**, 114701 (2008).
- 33) K. Ido, T. Ohgoe, and M. Imada, Phys. Rev. B **92**, 245106 (2015).
- 34) J. Haegeman et al., Phys. Rev. Lett. **107**, 070601 (2011).
- 35) S.-I. Amari, K. Kurata, and H. Nagaoka, IEEE Trans. Neural Networks **3**, 260 (1992).
- 36) S.-I. Amari, Neural Comput. **10**, 251 (1998).
- 37) W. L. McMillan, Phys. Rev. **138**, A442 (1965).
- 38) F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).
- 39) R. Orús, Ann. Phys. **349**, 117 (2014).
- 40) S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
- 41) A. Gendiar and T. Nishino, Phys. Rev. E **65**, 046702 (2002).
- 42) F. Mezzacapo et al., New J. Phys. **11**, 083026 (2009).
- 43) N. Schuch et al., Phys. Rev. Lett. **100**, 040501 (2008).
- 44) I. Affleck et al., Phys. Rev. Lett. **59**, 799 (1987).
- 45) P. Anderson, Mater. Res. Bull. **8**, 153 (1973).
- 46) M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
- 47) R. Jastrow, Phys. Rev. **98**, 1479 (1955).
- 48) G. Carleo, Y. Nomura, and M. Imada, Nat. Commun. **9**, 5322 (2018).
- 49) Y. Yamaji, T. Yoshida, A. Fujimori, and M. Imada, unpublished, 2019.

## 著者紹介

**野村悠祐氏**： 専門は物性理論、計算物質科学。強相関系に対する第一原理手法、低エネルギーソルバーの開発・応用に興味がある。

**山地洋平氏**： 専門は物性理論。とくに強相関電子系における量子相転移やトポロジカルな性質の研究を行っている。最近では計算物質科学的手法による理論物質設計にも取り組んでいる。

**今田正俊氏**： 専門は物性物理学理論、統計物理学、計算物質科学。関心の対象は強相関量子・電子系、マクロとミクロでの非平衡現象、数値計算手法。

(2018年8月5日原稿受付)

## Representing Quantum Many-Body States by Machine Learning

Yusuke Nomura, Youhei Yamaji, and Masatoshi Imada

abstract: Machine learning is used to extract essential pattern from big data. This technique can be used to extract the essential feature of quantum many-body wave function (=a vector with exponentially large dimensions), and to obtain compact representation of many-body states. In this article, we review representations of many-body states using Boltzmann machine, a type of artificial neural network. We introduce an efficient representation using restricted Boltzmann machines (RBM) and also discuss the efforts to improve the RBM wave functions.