

# 機械学習を用いて量子多体系を表現する



野村悠祐 <sup>東京大学大学院工学系研究科</sup> 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

#### -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に、 がラ メータの最適化は空か明待 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.

量子多体系とは多数の自由度がお互いに 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ってしま えば問題解決であるが、多体系のハミルト ニアンの次元は自由度の数に対して指数関 数的に増大するため、自由度の数が増える 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 エニューラル・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の自由度の状態配置 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)確率であると考えると、 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 的ハミルトニアンの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 い学理が近い将来明らかになるだろう.

# 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.大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、通常のメトロポリス 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<sub>i</sub>, h<sub>j</sub>(i=1, ..., N, j=1, ..., 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





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  $v_i$ ,  $h_j$ =±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= ( $v_1, ..., v_N$ ),  $h = (h_1, ..., h_M$ )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

$$E(v, h) = -\sum_{i} a_{i}v_{i} - \sum_{j} b_{j}h_{j} - \sum_{i,j} W_{ij}v_{i}h_{j}$$
(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)  $p(v,h) & p(v,h) = e^{-E(v,h)}/Z$ 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 スアウトする)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のには能 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)

$$\tilde{p}(v) = \sum_{h} p(v, h) \qquad \left(\sum_{v} \tilde{p}(v) = 1\right)$$
(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<sup>N</sup>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~2<sup>N</sup>)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 *p*(*v*) に対して,

$$\tilde{p}(v) = e^{\sum_{i} a_{i} v_{i}} \times \prod_{j} 2 \cosh\left(b_{j} + \sum_{i} W_{ij} v_{i}\right)$$
(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に結合が存在し,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. 隣接層間のみに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 $d = (d_1, ..., d_{M'}),$  $d_k = \pm 1$ 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、

$$E(v, h, d) = -\sum_{i} a_{i}v_{i} - \sum_{j} b_{j}h_{j} - \sum_{k} b_{k}'d_{k}$$
$$-\sum_{i,j} W_{ij}v_{i}h_{j} - \sum_{j,k} W_{jk}'h_{j}d_{k}$$
(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

$$\tilde{p}(v) = \sum_{h,d} p(v,h,d) \qquad \left(\sum_{v} \tilde{p}(v) = 1\right)$$
(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**  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*<sub>j</sub>と深層ユニット*d*<sub>k</sub>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 トロイヤー (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,

$$|\psi\rangle = \sum_{x} |x\rangle\langle x|\psi\rangle = \sum_{x} |x\rangle\psi(x)$$
(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、ψ(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.\*2 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 $N_{site}$ と同一に取り( $N=N_{site}$ ), $v_i$ 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 $|x\rangle = |v_1, v_2, ..., v_N\rangle$ で表現できる.波動関数は,

$$\psi_{\gamma}(x) = \sum_{h} e^{\sum_{i} a_{i}(2S_{i}^{z}) + \sum_{j} b_{j} h_{j} + \sum_{ij} W_{ij}(2S_{i}^{z}) h_{j}}$$
$$= e^{\sum_{i} a_{i}(2S_{i}^{z})} \times \prod_{j} 2 \cosh\left[b_{j} + \sum_{i} W_{ij}(2S_{i}^{z})\right]$$
(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 $\gamma = \{a_i, b_i, W_i\}$ に依存しており,これ

<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.

<sup>\*2</sup> 量子力学的には|ψ(x)|<sup>2</sup>=ψ(x)<sup>2</sup>が確率分布だが、本稿での RBM 波動 関数はψ(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) $\Psi_{true}(x)$ 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 $D(\Psi_{true} || \psi_{\gamma}) = \sum_{x} \Psi_{true}(x) \log(\Psi_{true}(x)/\psi_{\gamma}(x))$ 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|w>に対し、そのエネルギー期待値 は基底状態エネルギー EGS より小さくなることはないこと を保証する:  $\langle \psi | \mathcal{H} | \psi \rangle / \langle \psi | \psi \rangle \geq E_{GS}$  (等号は  $| \psi \rangle$  が基底状態 の時に成立). RBMを使って量子状態 | ψ > を書き表すと [式 (7)], そのエネルギーは変分パラメータ $y = \{a_i, b_i, W_{ii}\}$ 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

$$E_{\gamma} = \sum p_{\gamma}(x) E_{\gamma}^{\text{loc}}(x) \tag{8}$$

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 $p_{y}(x)$  は $p_{y}(x) = |\langle x|\psi_{y}\rangle|^{2}/\langle\psi_{y}|\psi_{y}\rangle$ , 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 $x \in \mathcal{R}$ 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、 $N_v$  個の変分パラメー タ $\gamma_m$  ( $m=1, 2, ..., N_v$ ) (RBM 波動関数の場合 $\gamma_m$  は $a_i, b_j, W_{ij}$ がそれにあたる)の更新量 $\Delta \gamma_m$  ( $\gamma_m \rightarrow \gamma_m + \Delta \gamma_m$ ) は微少虚時 間幅 (この用語の意味は後述) を Δτ として,

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

で与えられる. Sは正定値行列で  $S_{mn} = \langle \partial_{\gamma_m} \overline{\psi}_{\gamma} | \partial_{\gamma_n} \overline{\psi}_{\gamma} \rangle$ と規格 化された波動関数  $|\overline{\psi}_{\gamma}\rangle = |\psi_{\gamma}\rangle/\sqrt{\langle\psi_{\gamma}|\psi_{\gamma}\rangle}$ とその変分パラメー タによる 偏微分  $|\partial_{\gamma_m} \overline{\psi}_{\gamma}\rangle = (\partial/\partial_{\gamma_m})|\overline{\psi}_{\gamma}\rangle$ で定義される.  $g \sim 2$ トルは変分パラメータに対するエネルギー勾配で $g_m = \partial E_{\gamma}/\partial_{\gamma_m} = (\partial/\partial_{\gamma_m})\langle\overline{\psi}_{\gamma}|\mathcal{H}|\overline{\psi}_{\gamma}\rangle$ と与えられる. S行列や $g \sim 2$ 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  $|\overline{\psi}_{7}+\Delta_{7}\rangle \approx |\overline{\psi}_{7}\rangle + \sum_{m}\Delta\gamma_{m}|\partial_{\gamma_{m}}\overline{\psi}_{7}\rangle$ ともとの 波動関数  $|\overline{\psi}_{7}\rangle$ の違いのノルムの二乗が  $||\overline{\psi}_{7}+\Delta_{7}\rangle - |\overline{\psi}_{7}\rangle||^{2} \approx$  $\sum_{mn}S_{mn}\Delta\gamma_{m}\Delta\gamma_{n}$ 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<sup>-1</sup>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\*4 した状態(1-2H $\Delta\tau$ )| $\psi$ 〉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<sub>i</sub>*, *b<sub>j</sub>*, *W<sub>ij</sub>*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

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

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

<sup>\*4</sup> 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  $e^{-\mathcal{H}\tau}|\psi_0\rangle$  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.

<sup>\*5</sup> 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必ずしも 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<sub>i</sub>*, *b<sub>j</sub>*, *W<sub>ij</sub>*を値 の小さい乱数に取る.

2. 変分パラメータ *a<sub>i</sub>*, *b<sub>j</sub>*, *W<sub>ij</sub>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 〈の〉も式(8) において*H*をのに置 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<sub>site</sub> 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) ストリング・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\*6を取ることに よって状態を表現する.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 $\psi(x) = \text{Tr}(\prod_{i=1}^{N} A^{(i)}[v_i]), |x\rangle = |v_1, v_2, ..., v_N\rangle$ 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→∞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: entangledplaquette 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  $\mathbf{v}_p = (v_{p_1}, ..., v_{p_{n_p}})$  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.

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

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. ストリング・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 ステム上に複数のストリング (灰色の紐) を用意し,各ス トリング上で 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.

$$\psi(\mathbf{x}) = \prod_{j} \operatorname{Tr}\left(\prod_{i \in S_{j}} A^{(i,j)}[v_{i}]\right)$$
(11)

ここで*S<sub>j</sub>はj*番目のストリングが通るサイトの集合を表している.

<sup>\*&</sup>lt;sup>6</sup> テンソルの積を縮約ともいう.例えば $C_k = \sum_j A_{ij} B_{jk}$ 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.

<sup>\*7</sup> この例では各ストリングは全てのサイトを通るようになっているが、 必ずしも全サイトを通る必要はなく短いストリングがあっても良い。

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<sub>ij</sub>を持つので はなく一部の隣接したサイト (サイト数が 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<sub>j</sub> 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<sub>i</sub> は簡単のため無視する)

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

と書き表すことができる. ここで v<sub>j</sub>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ことがわかる [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  $O(N_{site})$ 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<sub>Sj</sub>* 個の可視 ユニットサイトの集合を *S<sub>j</sub>* とし(再びバイアス *a<sub>i</sub>* は簡単の ため無視する).

$$\psi(\mathbf{x}) = \prod_{j} \operatorname{Tr}\left(\prod_{i \in S_{j}} A^{(i,j)}[v_{i}]\right)$$
(13)

と書き直すことができる.ここで,

$$A^{(i,j)}[v_i] = \begin{pmatrix} e^{b/n_{S_j} + W_{ij}v_i} & 0\\ 0 & e^{-b_j/n_{S_j} - W_{ij}v_i} \end{pmatrix}.$$
 (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ことがわかる.<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<sub>site</sub> 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*H*=  $J\sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+1} (J>0)$ で表される.bipartite な格子上のフラス 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*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 $v_i = 2S_i^{-}$ 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の間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.

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

77



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<sub>i-j</sub>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.\*\* 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,反強磁性的な揺 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). α= M/Nと定義すると、α がRBM波動関数の精度を制御するパラメータとなり、α 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.αが比較的少な 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. α=*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 $x=(v_1, ..., v_N)$ 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、 $v_i=2S_i^r=\pm 1$ 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Iのスピン $\sigma$ の状態に対する 数演算子を $n_{I\sigma}$ として  $(v_{2I-1}, v_{2I}) = (2n_{I\uparrow} - 1, 2n_{I\downarrow} - 1)$ と定 義する.すなわちサイトI、スピン $\sigma$ 電子がいるいないで ±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  $|\phi_{PP}\rangle =$  $(\sum_{l,m=1}^{N_{stec}} f_{lm}^{\uparrow\downarrow} c_{m\downarrow}^{\dagger})^{N_c/2} |0\rangle (N_c: 電子数) を組み合わせた$ RBM + PP 波動関数を以下のように導入する<sup>14)</sup> (図 5):

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

<sup>\*\*</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  $|\psi\rangle = (|\uparrow\rangle + |\downarrow\rangle)_1 \otimes (|\uparrow\rangle + |\downarrow\rangle)_2 \otimes ... \otimes (|\uparrow\rangle + |\downarrow\rangle)_8$ . この状態はx方向にスピンが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>\*&</sup>lt;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>\*&</sup>lt;sup>12</sup> 化学の分野ではジェミナルと呼ばれている.

$$\psi_{\gamma}(x) = \mathcal{N}(x)\phi_{\text{ref}}(x) \tag{15}$$

ここで $|\phi_{ref}\rangle$ 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 $|\phi_{ref}\rangle = P_G^{\infty}|\phi_{PP}\rangle$  [ $P_G^{\infty} = \prod_i (1 - n_i \uparrow n_i \downarrow)$ ]を採用する.フェルミオン系 に対しては $|\phi_{ref}\rangle = |\phi_{PP}\rangle$ とペア積状態をそのまま使う.こ の $\phi_{ref}(x)$ 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,物理の関数.  $W_{ij}$ 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.  $f_{in}^{go}$ は PP 波動関数に含まれる変分パラメータ.  $W_{ij} や f_{in}^{go}$ 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.



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 | $\phi_{ref}$ )(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より柔軟かつ不偏な 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 |ψ<sub>0</sub>〉に対し,ハミルトニアン *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 口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,が可視ユニット v<sub>k</sub>, v<sub>l</sub>, v<sub>m</sub>と相互作用すれば, h<sub>j</sub>の状態 和を取ると cosh [W<sub>kj</sub>v<sub>k</sub>+W<sub>ij</sub>v<sub>l</sub>+W<sub>mj</sub>v<sub>m</sub>]の因子を通じて v<sub>k</sub>, v<sub>l</sub>, v<sub>m</sub>間の3 体相互作用が生じる.

展 e<sup>-tH</sup>|w<sub>0</sub>〉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 る必要があったのとは対照的である.ただし、DBMでは 隠れユニットh<sub>i</sub>と深層ユニットd<sub>k</sub>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、 $h_i \ge d_k$ (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>:  $e^{-\tau H} \sim (e^{-\delta, \mathcal{H}_1} e^{-\delta, \mathcal{H}_2})^{N_{\text{slice}}}$ ,  $(\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2, \delta_\tau = \tau/N_{\text{slice}})$ を 用いると, DBMで表される任意の波動関数 | $\psi_7$ 〉に対して, 微小虚時間発展させた状態  $e^{-\delta, \mathcal{H}_v} |\psi_7\rangle$ を新たな変分パラ メータ  $\overline{p}$ を持った DBM 状態 | $\psi_{\overline{p}}\rangle^{*15}$ 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

$$e^{-\delta_{t}\mathcal{H}_{v}}|\psi_{\gamma}\rangle = C_{v}|\psi_{\overline{\gamma}}\rangle \quad (v=1, 2, C_{v}: \text{const.})$$
(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はできないが、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 の長さと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>



図7 横磁場イジング模型 $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$ ,  $\mathcal{H}_1 = \sum_{l < m} V_{lm} \sigma_{l\sigma_m}^{r}$ ,  $\mathcal{H}_2 = -\sum_l \Gamma_l \sigma_l^{r}$ 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 $\sigma_l^{r} = \pm 1$  と可視ユニットの状態 $v_l = \pm 1$ 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、これまでに 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>\*14</sup> 実際の計算にはより正確なe<sup>-á,H</sup>~e<sup>-(á,l2)H2</sup>e<sup>-á,H1</sup>e<sup>-(á,l2)H2</sup>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.

<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 Micro-structure of Cognition: Foundations*, edited by D. E. Rumelhart, J. L. McLelland, 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.