

機械学習の量子多体問題への応用

東北大学 金属材料研究所

野村 悠祐



研究室ウェブサイト <https://www.nomura-lab.imr.tohoku.ac.jp>



E-mail yusuke.nomura@tohoku.ac.jp



IMR 東北大学 金属材料研究所
Institute for Materials Research, Tohoku University

資料 2 の内容

1. 量子状態トモグラフィー
2. 基底状態解析と変分法：導入
3. 制限ボルツマンマシン（RBM）を用いた変分法
 - 最急降下法
 - Stochastic Reconfiguration法（自然勾配法）
 - 変分モンテカルロ法

資料 2 の内容

1. 量子状態トモグラフィー

2. 基底状態解析と変分法：導入

3. 制限ボルツマンマシン（RBM）を用いた変分法

- 最急降下法
- Stochastic Reconfiguration法（自然勾配法）
- 変分モンテカルロ法

量子状態トモグラフィー

未知なる量子状態が実験的に実現されている状況で、その具体的な表現を求めるタスク

例：スピン1/2の量子スピン系、実験的にパウリ基底での射影測定が可能と仮定

基底： $\mathbf{b} = (b_1, \dots, b_N)$ ($b_i \in \{x, y, z\}$)

サンプルが従う確率分布： $p(\sigma^{[\mathbf{b}]}) \propto |\psi(\sigma^{[\mathbf{b}]})|^2$

モデルによる近似： $q_\theta(\sigma^{[\mathbf{b}]}) \sim p(\sigma^{[\mathbf{b}]})$

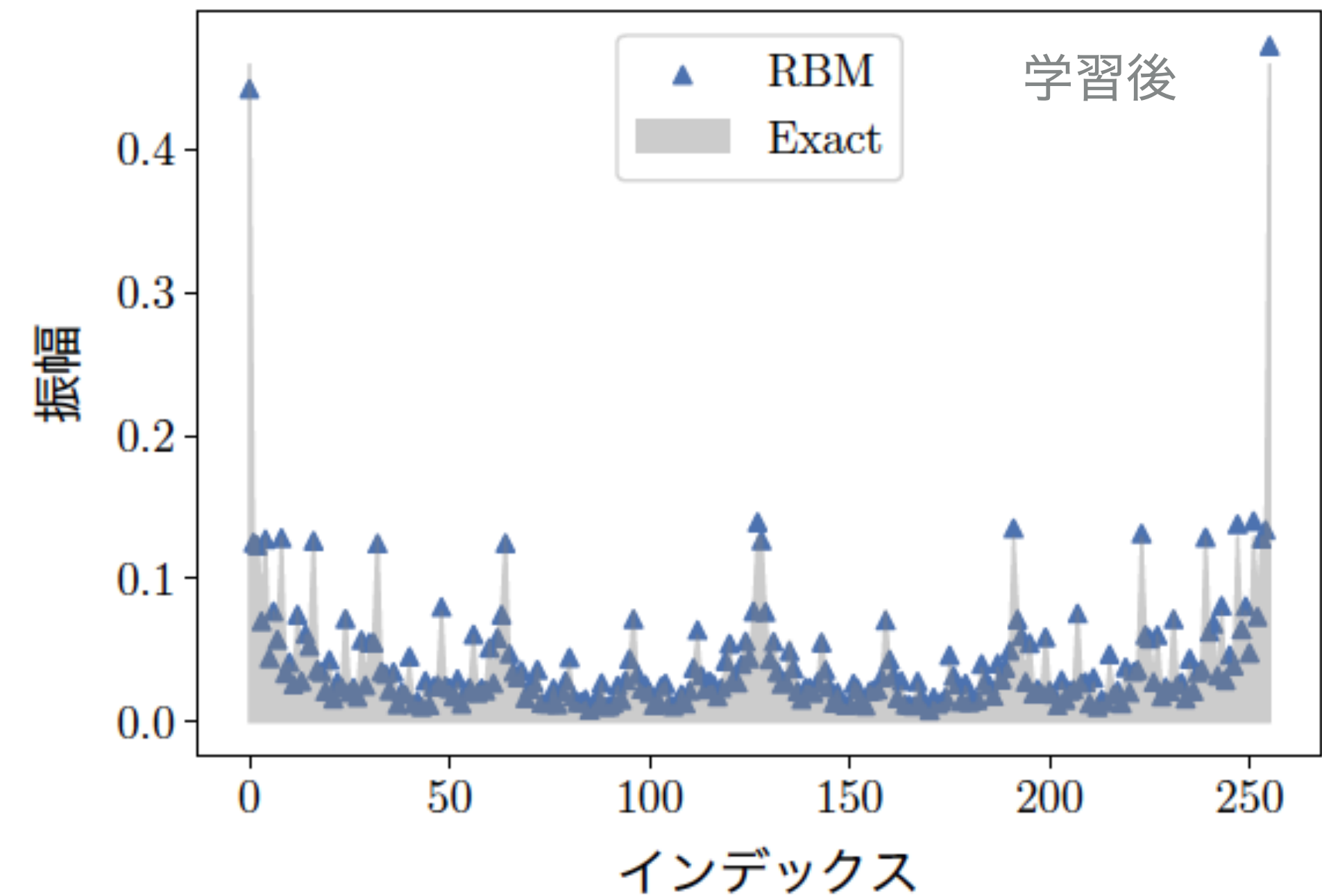
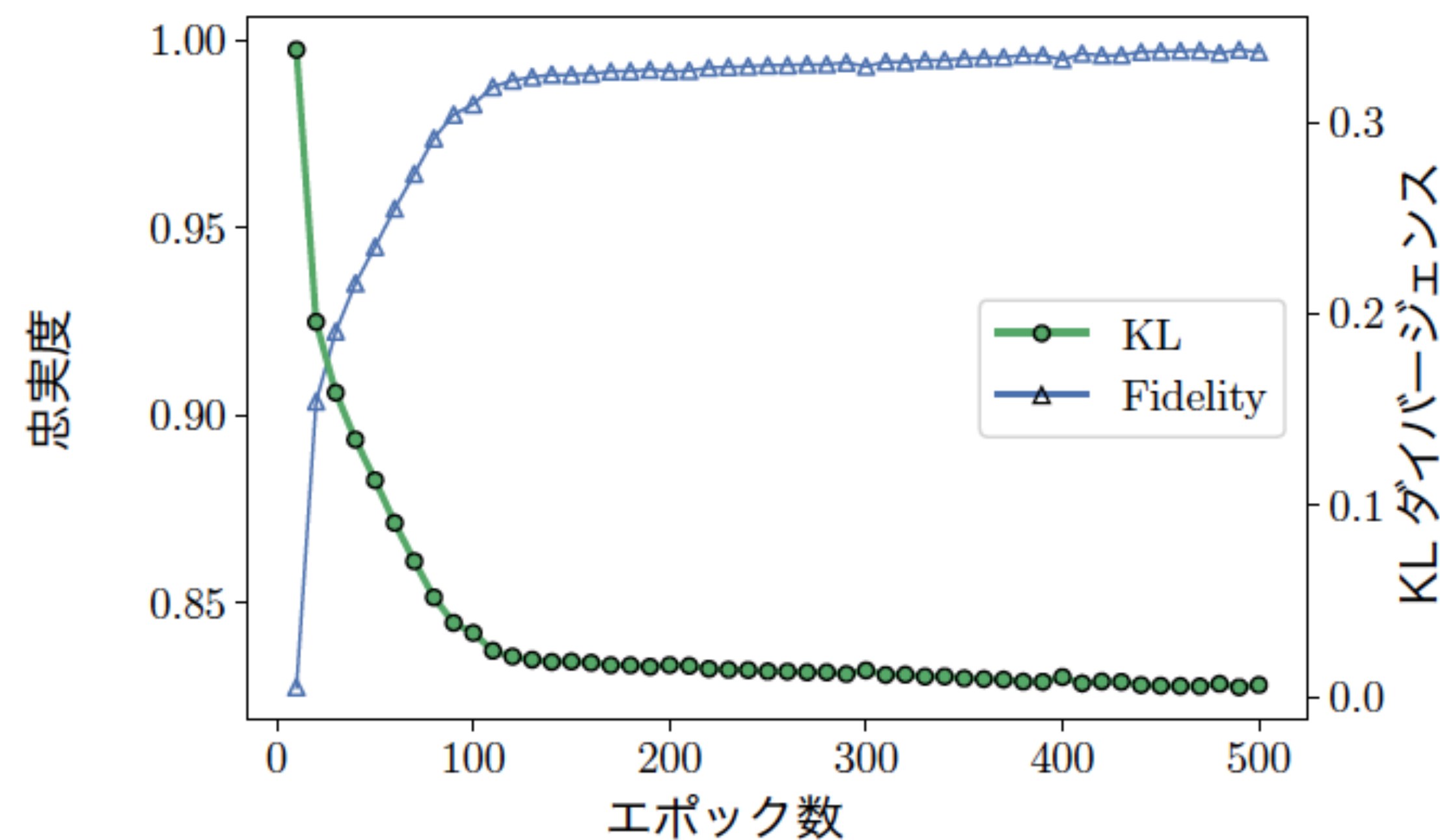
コスト関数（複数の測定軸に対して同時に教師なし学習）：

$$\text{Cost}(\theta) = \sum_{\mathbf{b}} D_{\text{KL}}(p^{[\mathbf{b}]} \| q_\theta^{[\mathbf{b}]}) = \sum_{\mathbf{b}} \sum_{\sigma^{[\mathbf{b}]} \in \mathcal{D}^{[\mathbf{b}]}} p(\sigma^{[\mathbf{b}]}) \log \frac{p(\sigma^{[\mathbf{b}]})}{q_\theta(\sigma^{[\mathbf{b}]})}$$

量子状態トモグラフィー：例として横磁場イジングモデルの基底状態

1次元横磁場イジングモデル： $\hat{\mathcal{H}} = - \sum_i Z_i Z_{i+1} - h \sum_i X_i$

8サイト、 $h=1$ 、RBMを用いて学習、Z軸測定のみ（波動関数が正であるため）、サンプル数 10^4



資料 2 の内容

1. 量子状態トモグラフィー

2. 基底状態解析と変分法：導入

3. 制限ボルツマンマシン（RBM）を用いた変分法

- 最急降下法
- Stochastic Reconfiguration法（自然勾配法）
- 変分モンテカルロ法

量子多体問題をどう解析するか？

時間依存しないシュレーディンガー方程式

$$\hat{\mathcal{H}}|\psi\rangle = E|\psi\rangle$$

計算基底 $\{|x\rangle\}$ で展開

$$\hat{\mathcal{H}} = \sum_{x,x'} |x\rangle \langle x| \hat{\mathcal{H}} |x'\rangle \langle x'|$$

$= H(x,x') : \text{ハミルトニアン行列}$

$$|\psi\rangle = \sum_x |x\rangle \langle x|\psi\rangle$$

$= \psi(x) : \text{波動関数}$

$$\sum_{x'} \mathcal{H}(x, x') \psi(x') = E \psi(x)$$

行列の固有値、固有ベクトルを求める問題

量子多体ハミルトニアン の例：ハイゼンベルク模型（4 サイト）

4 サイトの（ゲージ変換後の）ハイゼンベルク模型のハミルトニアンを行列表示&厳密対角化

(total $S_z = 0$ のブロックのみを考慮)

$$\hat{\mathcal{H}} = \sum_i (-\sigma_i^x \sigma_{i+1}^x - \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z)$$

基底： $|\uparrow\uparrow\downarrow\downarrow\rangle, |\uparrow\downarrow\uparrow\downarrow\rangle, |\uparrow\downarrow\downarrow\uparrow\rangle, |\downarrow\uparrow\uparrow\downarrow\rangle, |\downarrow\uparrow\downarrow\uparrow\rangle, |\downarrow\downarrow\uparrow\uparrow\rangle$

$$\mathcal{H} = \begin{pmatrix} 0 & -2 & 0 & 0 & -2 & 0 \\ -2 & -4 & -2 & -2 & 0 & -2 \\ 0 & -2 & 0 & 0 & -2 & 0 \\ 0 & -2 & 0 & 0 & -2 & 0 \\ -2 & 0 & -2 & -2 & -4 & -2 \\ 0 & -2 & 0 & 0 & -2 & 0 \end{pmatrix}$$

基底状態のエネルギーと波動関数

$$E = -8$$

$$\psi(\sigma) = \frac{1}{2\sqrt{3}}(1, 2, 1, 1, 2, 1)^T$$

Nサイトハイゼンベルク模型のハミルトニアンの次元(total Sz=0)

N	D_N
10	252
20	184756
30	155117520
40	137846528820
50	126410606437752

厳密対角化ができる最大のサイズは40~50サイト程度

変分原理

$$E_\theta = \frac{\langle \psi_\theta | \mathcal{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \geq \underline{E_0}$$

基底状態エネルギー

簡易証明：

$$|\psi_\theta\rangle = \sum_i c_i \underbrace{|\phi_i\rangle}_{\text{固有状態}}$$

$$E_\theta = \sum_i |c_i|^2 E_i \geq \sum_i |c_i|^2 E_0 = E_0$$

変分波動関数： $\psi_\theta(x)$

パラメータセットである θ に依存する

変分法の適用の“本丸”：量子モンテカルロ法が負符号で適用が困難な量子多体模型

フラストレーションのある量子スピン系やフェルミオン系などの物性解明に重要な役割を果たしてきた

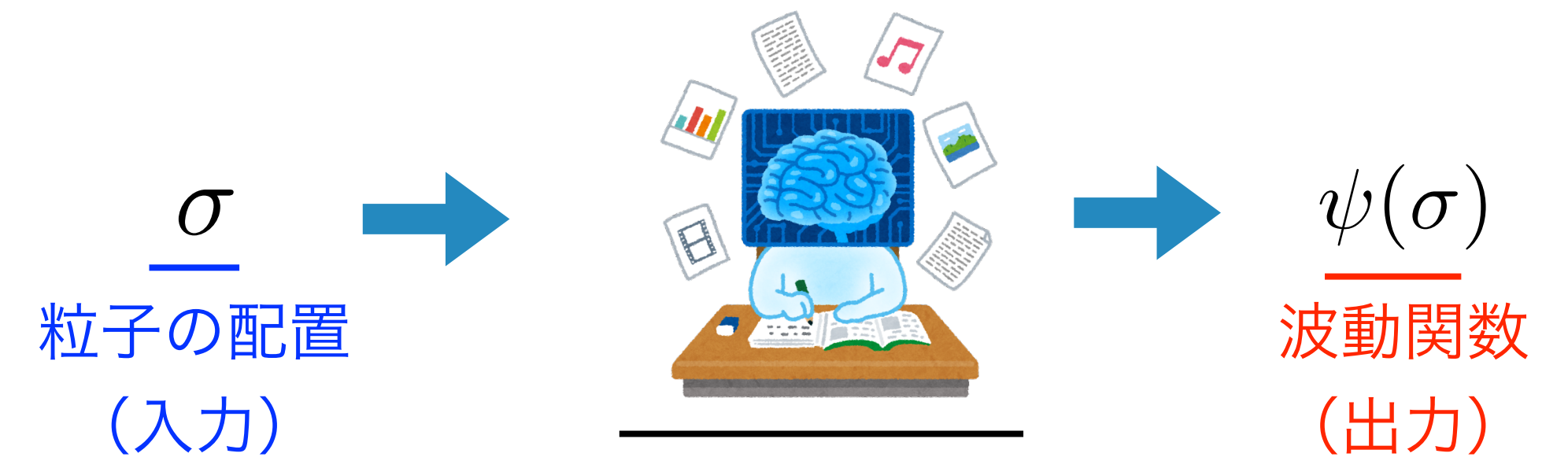
人工ニューラルネットワーク手法と従来の手法の比較

従来の変分法

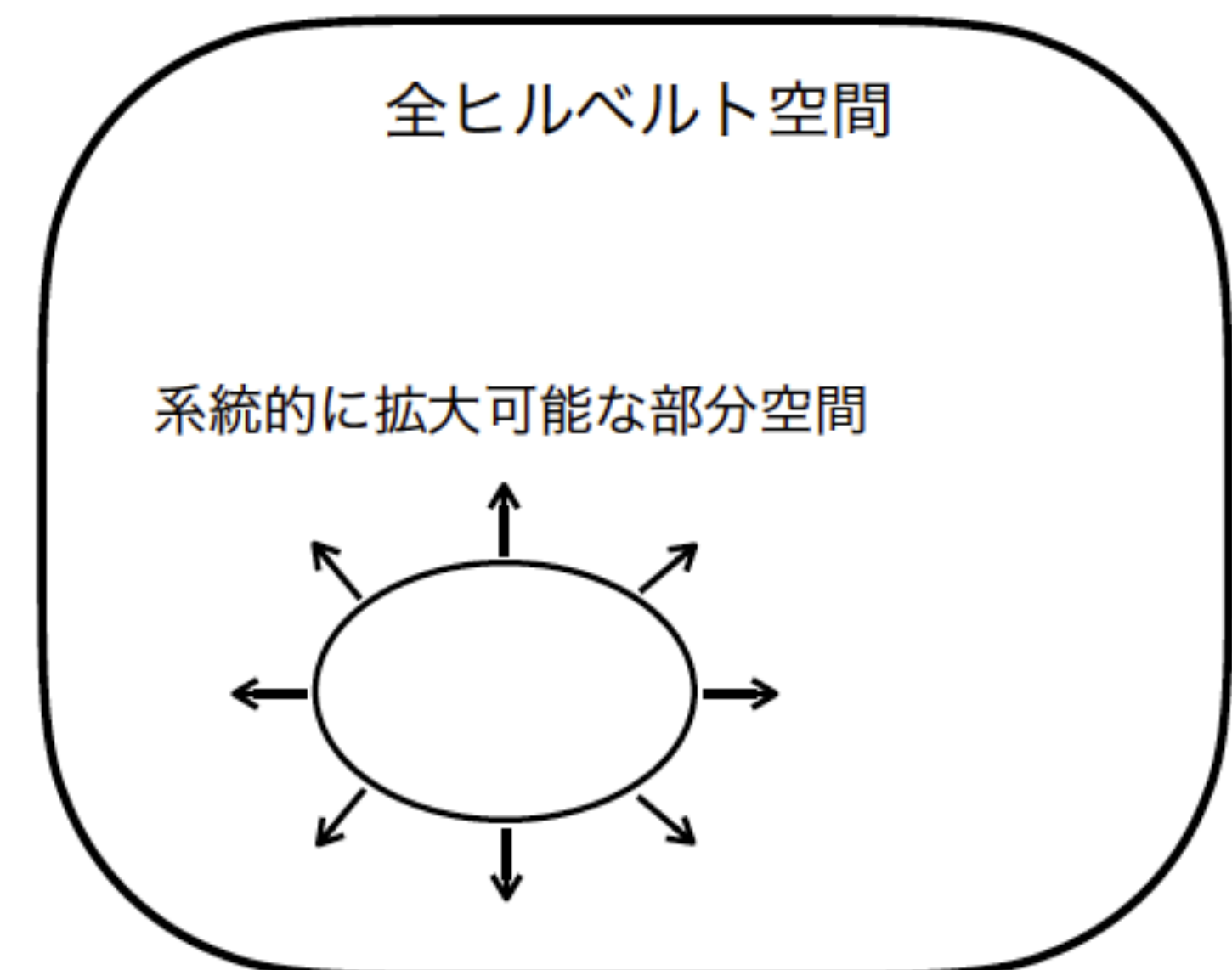
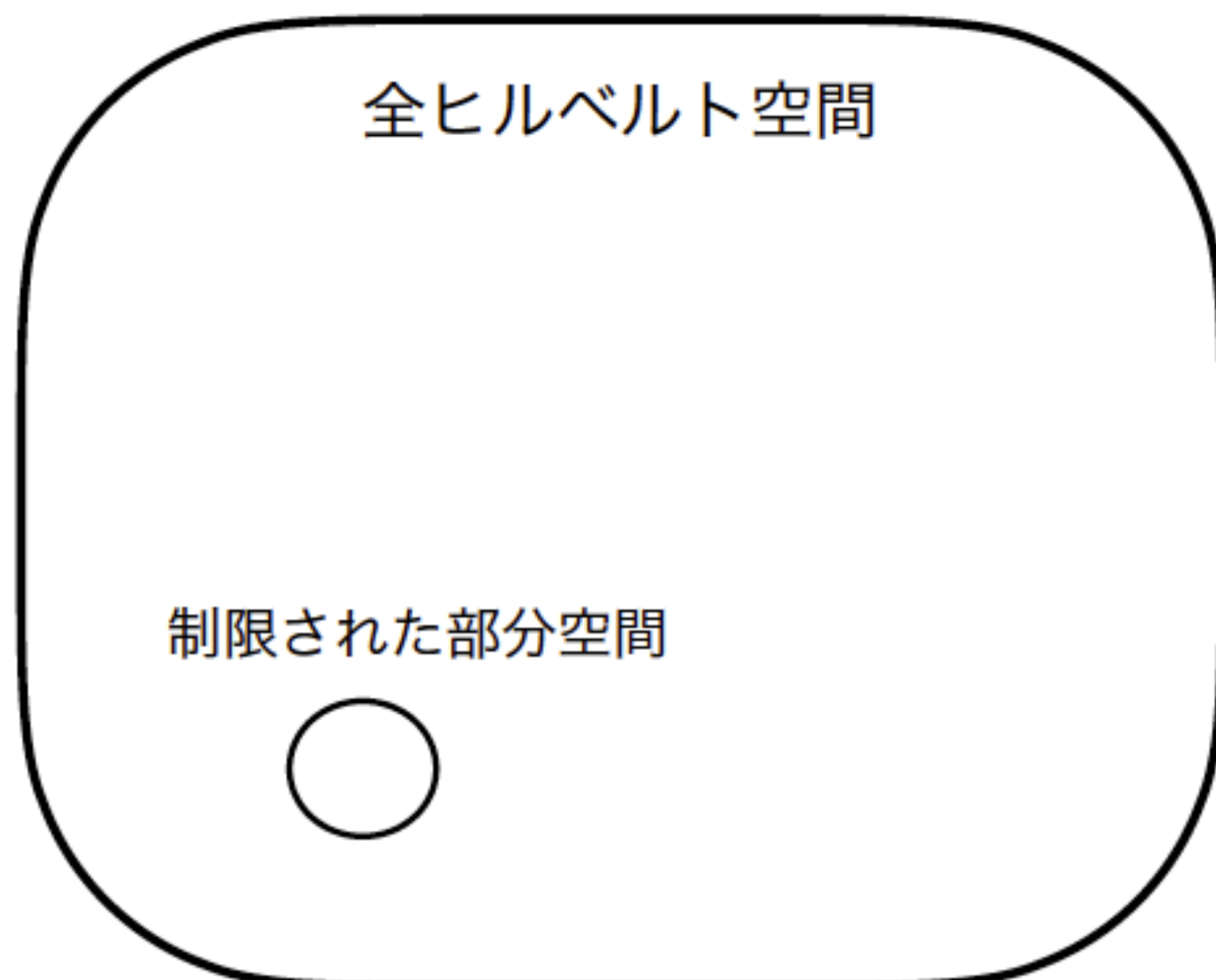


変換の仕方は人（物理的洞察）が決める（計算科学）

人工ニューラルネットワークによる変分法



変換の仕方はデータが決める（データ科学）



資料 2 の内容

1. 量子状態トモグラフィー
2. 基底状態解析と変分法：導入
3. 制限ボルツマンマシン（RBM）を用いた変分法
 - 最急降下法
 - Stochastic Reconfiguration法（自然勾配法）
 - 変分モンテカルロ法

このパートの目標

- ・ 人工ニューラルネットワークによる変分法の最低限の原理を理解する
- ・ (少なくとも) 最急降下法による最適化を、コードを一から組めるくらいの理解度にする

https://github.com/yusukenomura/School_2024 一次元ハイゼンベルグ模型 (Fortran)

https://github.com/ryuikaneko/textbook_2024_machine_learning_physics 一次元横磁場イジング模型 (Python)



人工ニューラルネットワークによる変分法

機械学習でプラクティカルに行われていること＝非線形関数（ここでは人工ニューラルネットワークを使う）の最適化

機械学習タスクで設定しなければならないこと（変分法の場合）

1. 人工ニューラルネットワークへのデータの入力方法を定義する

→ 粒子の実空間配置

2. 人工ニューラルネットワークの出力を定義する

→ 波動関数

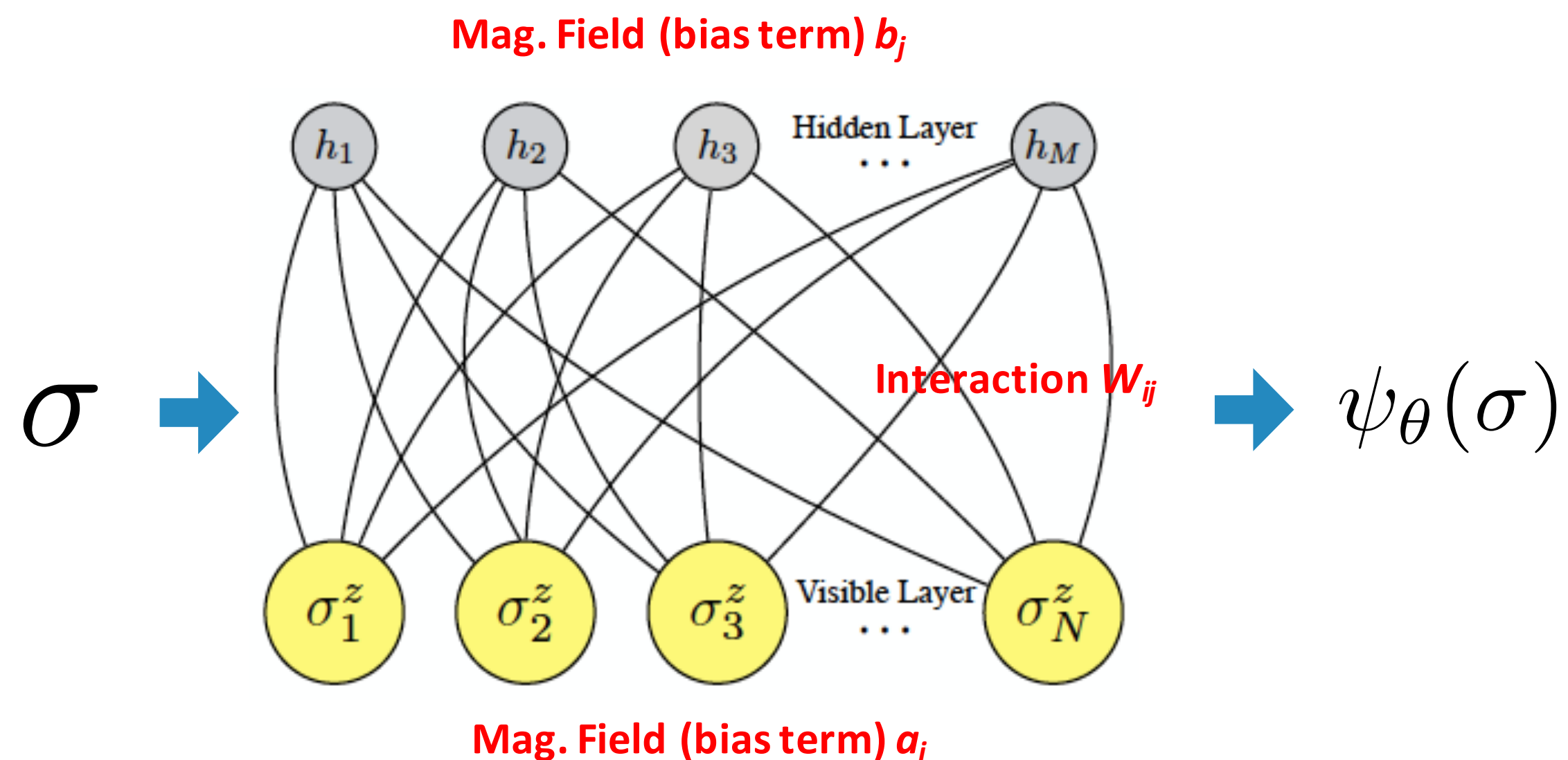
3. コスト関数を定義する

→ エネルギー期待値

4. パラメータの最適化方法を選択する

→ 変分法ではnatural gradient法が使われることが多い（後述）

設定1, 2の例：制限ボルツマンマシン(RBM)による量子状態表現



Carleo and Troyer Science 355, 602 (2017)

RBM (restricted Boltzmann machine) 波動関数

$$\psi_{\theta}(\sigma) = \frac{\sum_h \exp \left(\sum_i a_i \sigma_i^z + \sum_{i,j} W_{ij} \sigma_i^z h_j + \sum_j b_j h_j \right)}{e^{-E_{\text{RBM}}(\sigma, h)} \quad (\text{Boltzmann weight})}$$

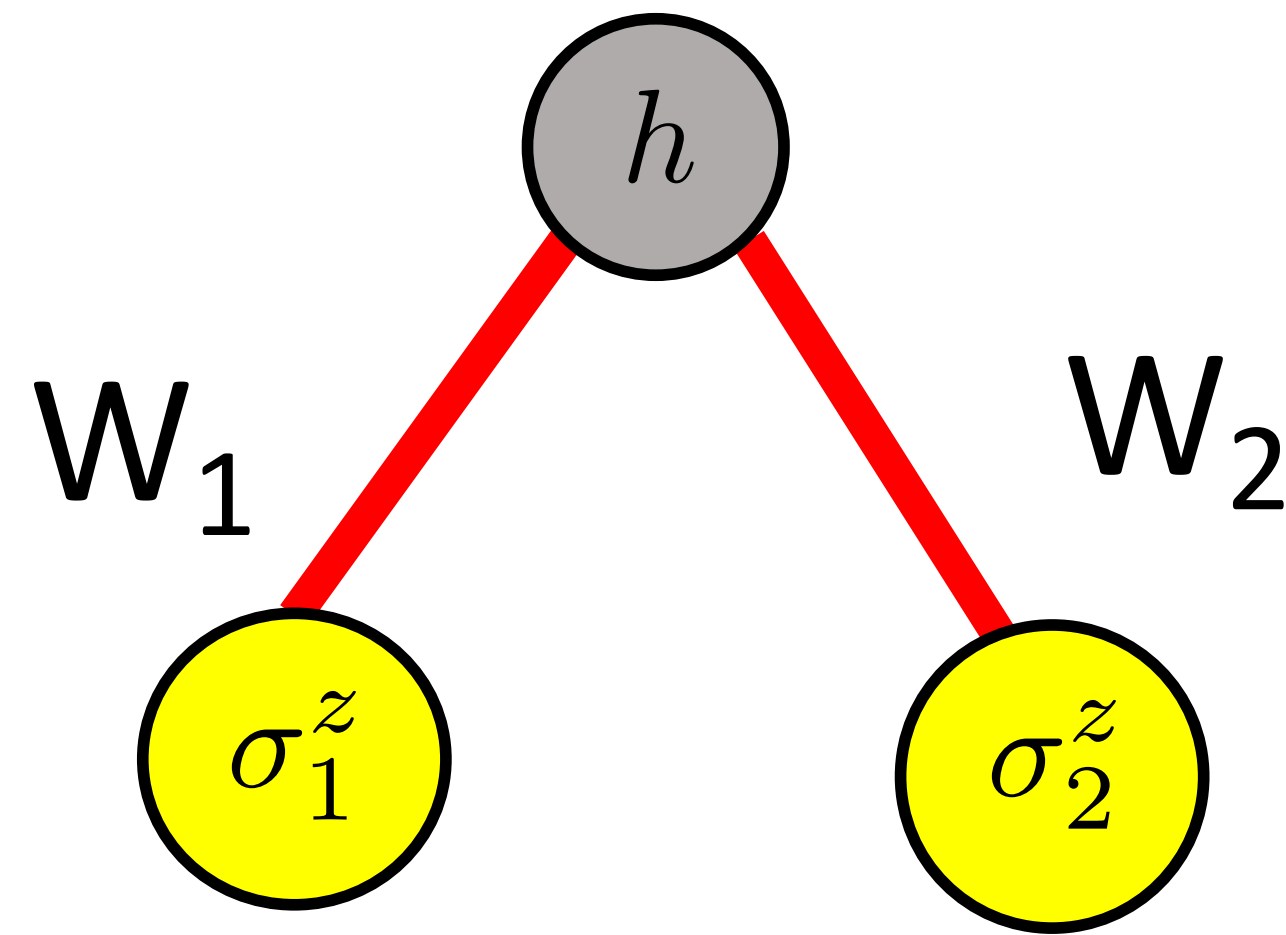
$\sigma = (\sigma_1^z, \sigma_2^z, \dots, \sigma_N^z)$: スピン配置

$h_j = \pm 1$: 隠れスピン

- 入力=スピン配置、出力=波動関数
- 隠れスピンとの結合を通じて量子相関を表す
- 隠れスピンの数が無限大の極限で量子状態を任意の精度で表現可能（普遍近似性能）

→ $\alpha = M/N$ (hidden variable density)が精度の制御パラメータ

設定1, 2に関する具体的計算例：RBM波動関数（N=2, M=1の場合）



$$\theta = (W_1, W_2)$$

$$\begin{aligned}\psi_{\theta}(\sigma_1^z, \sigma_2^z) &= \sum_{h=\pm 1} \exp(W_1 \sigma_1^z h + W_2 \sigma_2^z h) \\ &= \exp(W_1 \sigma_1^z + W_2 \sigma_2^z) + \exp(-W_1 \sigma_1^z - W_2 \sigma_2^z) \\ &= 2 \cosh(W_1 \sigma_1^z + W_2 \sigma_2^z)\end{aligned}$$

$$\psi_{\theta}(\uparrow, \uparrow) = \psi_{\theta}(\downarrow, \downarrow) = 2 \cosh(W_1 + W_2)$$

$$\psi_{\theta}(\uparrow, \downarrow) = \psi_{\theta}(\downarrow, \uparrow) = 2 \cosh(W_1 - W_2)$$

一般の場合

$$\psi_{\theta}(\sigma) = \exp\left(\sum_i a_i \sigma_i^z\right) \times \prod_j 2 \cosh\left(b_j + \sum_i W_{ij} \sigma_i^z\right)$$

設定3, 4 : コスト関数の設定と最適化

コスト関数=エネルギー期待値

$$E_{\theta} = \frac{\langle \psi_{\theta} | \mathcal{H} | \psi_{\theta} \rangle}{\langle \psi_{\theta} | \psi_{\theta} \rangle}$$

最適化に必要なもの=勾配ベクトル (コスト関数の微分)

$$g_k^{(t)} = \left. \frac{\partial E_{\theta}}{\partial \theta_k} \right|_{\theta = \theta^{(t)}}$$

$\theta^{(t)}$: t番目のステップ時のパラメータセット

最急降下法 Steepest descent (SD) (or gradient descent)

$$\theta_k^{(t+1)} = \theta_k^{(t)} - \eta g_k^{(t)}$$

η : learning rate ➡ ハイパーパラメータ (人の手で決める)

人工ニューラルネットワークを用いた変分法の計算手順

1. 初期化：人工ニューラルネットワークによって変分波動関数を定義し、パラメータを初期化する

※ パラメータの初期化には小さな値の乱数を用いることが多い

2. 学習：コスト関数（エネルギー）を最小化するようにパラメータを最適化する

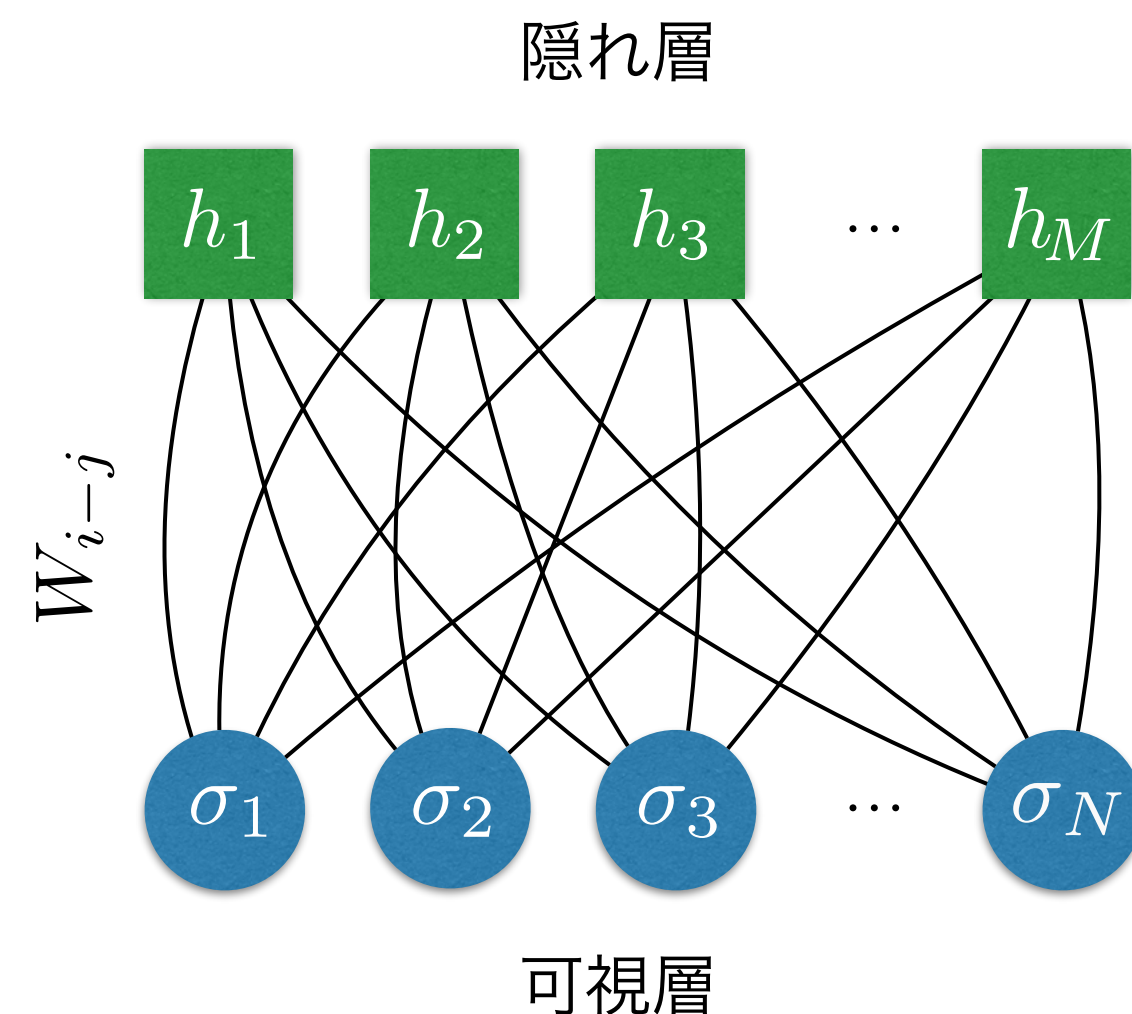
3. ポストプロセス：得られた波動関数を用いて、物理量を計算する

※ 乱数を用いてパラメータを初期化した場合は、複数の初期パラメータを用いて最適化を行い一番エネルギーの下がったものを採用する

1次元ハイゼンベルク模型への適用：SD法による最適化

ハミルトニアン：
$$\mathcal{H} = \sum_i (-\sigma_i^x \sigma_{i+1}^x - \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z)$$

変分波動関数：



$$\alpha(\text{hidden variable density}) = M/N = 1$$

$$\psi_{\theta}(\sigma) = \prod_j 2 \cosh \left(\sum_i W_{i-j} \sigma_i^z \right)$$

Carleo and Troyer Science 355, 602 (2017)とほぼ同じ設定

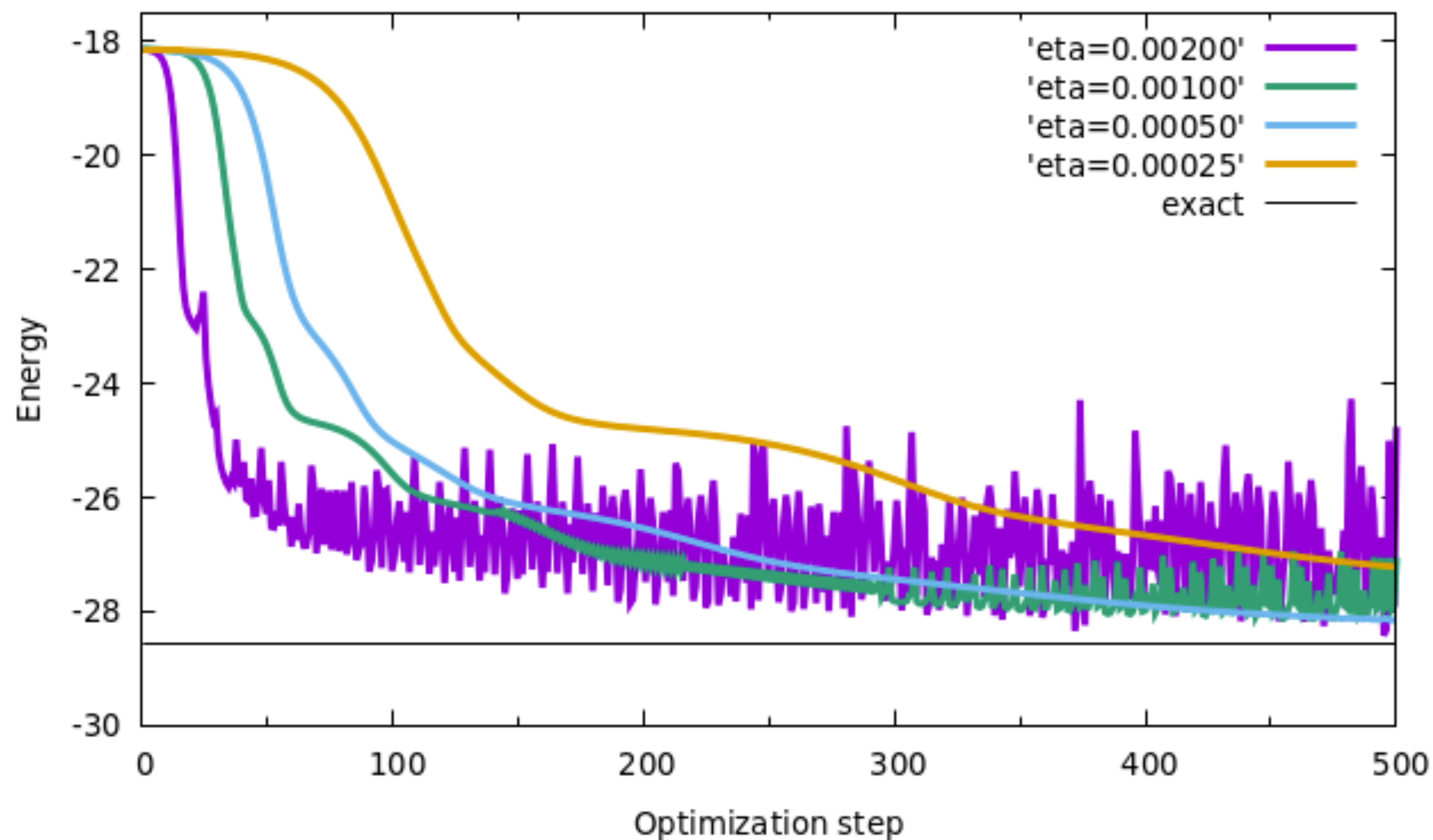
パラメータのアップデートルール：

$$\theta_k^{(t+1)} = \theta_k^{(t)} - \eta g_k^{(t)}$$

SD法：最適化（学習）の安定性

学習率 η を変えながら、(コスト関数) vs (最適化のIteration数) をプロット

16サイトに対する結果



コスト関数 (=エネルギー) の具体的な計算方法+コードの解説

$$\langle \mathcal{H} \rangle = \frac{\langle \psi_\theta | \mathcal{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} = \frac{\sum_{\sigma\sigma'} \psi_\theta^*(\sigma) \mathcal{H}_{\sigma\sigma'} \psi_\theta(\sigma')}{\sum_{\sigma} |\psi_\theta(\sigma)|^2} = \frac{\sum_{\sigma} |\psi_\theta(\sigma)|^2 E_{\text{loc}}(\sigma)}{\sum_{\sigma} |\psi_\theta(\sigma)|^2} \quad E_{\text{loc}}(\sigma) = \sum_{\sigma'} \mathcal{H}_{\sigma\sigma'} \frac{\psi_\theta(\sigma')}{\psi_\theta(\sigma)}$$

```
175 Eloc_x = 0d0
176 do i1 = 1, N
177   i2 = i1 + 1
178   if( i1 == N ) i2 = 1
179   !
180   ! SzSz contribution
181   !
182   Eloc_x = Eloc_x + x(i1)*x(i2)
183   !
184   ! ( SxSx + SySy ) = 2 * ( S+S- + S-S+ ) contribution
185   !
186   if( x(i1) /= x(i2) ) then
187     xp(:) = x(:)
188     xp(i1) = -xp(i1)
189     xp(i2) = -xp(i2)
190     call calc_theta(xp, thetap)
191     call calc_amplitude_RBM(M, thetap, psi_xp)
192     Eloc_x = Eloc_x - 2d0*psi_xp/psi_x
193   end if
194
195 end do ! i1
196 E = E + p_x * Eloc_x
```

$= H_{\sigma\sigma'}$

$= |\psi_\theta(\sigma)|^2$

O(N)でEloc(σ)が計算できる！ → 相互作用が局所的なため

※(コード内のx) = σ

→ 分子の計算 (分母は後で割る)

コスト関数の微分の具体的な計算方法＋コードの解説

$$g_k = \left(\frac{\langle \psi_\theta | \mathcal{H} | \partial_k \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} + \text{c.c.} \right) - \left(\frac{\langle \psi_\theta | \mathcal{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \frac{\langle \psi_\theta | \partial_k \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} + \text{c.c.} \right)$$

$$= 2\text{Re}\langle \mathcal{H} O_k \rangle - 2\langle \mathcal{H} \rangle \text{Re}\langle O_k \rangle$$

$$O_k = \sum_{\sigma} |\sigma\rangle O_k^{\text{loc}}(\sigma) \langle \sigma|$$

$$O_k^{\text{loc}}(\sigma) = \frac{\partial_k \psi_\theta(\sigma)}{\psi_\theta(\sigma)}$$

```

200 k = 0
201 do f = 1, alpha
202 do iw = 0, N-1
203
204     k = k + 1
205     Ovec_loc_x(k) = 0d0
206
207     do jj = 1, N
208         j = (f-1)*N+jj
209         i = jj + iw
210         if( i > N ) i = i - N
211         Ovec_loc_x(k) = Ovec_loc_x(k) + tanh(theta(j))*dble(x(i))
212     end do ! jj
213
214 end do ! iw
215 end do ! f
216 if( k /= Nv ) stop 'k /= Nv'
217 !
218 ! summation over x for gvec and Ovec
219 !
220 do k = 1, Nv
221     gvec(k) = gvec(k) + 2d0 * p_x * Eloc_x * Ovec_loc_x(k)
222     Ovec(k) = Ovec(k) + p_x * Ovec_loc_x(k)
223 end do ! k

```

coshの微分はsinhであることに注意

(coshで割るとtanhがでてくる)

< H O_k >の分子の計算

< O_k >の分子の計算

$$\langle O_k \rangle = \frac{\langle \psi_\theta | \partial_k \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} = \frac{\sum_{\sigma} |\psi_\theta(\sigma)|^2 O_k^{\text{loc}}(\sigma)}{\sum_{\sigma} |\psi_\theta(\sigma)|^2}$$

ここまで：SD法を用いて変分原理に従った計算を行った。

次の課題：安定した最適化を行うには？ → Stochastic reconfiguration (SR)法

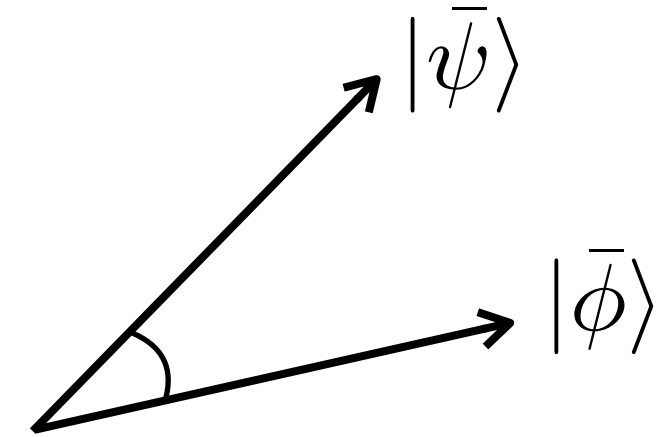
SR法導入の前に・・・ フビニ・スタディ計量（量子状態間の“距離”）

フビニ・スタディ計量

$$\mathcal{F}[|\psi\rangle, |\phi\rangle] := \arccos \sqrt{\frac{\langle\psi|\phi\rangle \langle\phi|\psi\rangle}{\langle\psi|\psi\rangle \langle\phi|\phi\rangle}} \xrightarrow[\substack{\text{波動関数を規格化} \\ |\bar{\psi}\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} \quad |\bar{\phi}\rangle = \frac{|\phi\rangle}{\sqrt{\langle\phi|\phi\rangle}}}]{} \mathcal{F}[|\bar{\psi}\rangle, |\bar{\phi}\rangle] = \arccos (|\langle\bar{\psi}|\bar{\phi}\rangle|)$$

フィデリティ

複素ベクトル間の角度 → quantum angle



フビニ・スタディ計量テンソル（量子幾何テンソルの実部）

$$\mathcal{F}^2[|\psi_{\theta+\delta\theta}\rangle, |\psi_{\theta}\rangle] = \sum_{kl} S_{kl}(\theta) \delta\theta_k \delta\theta_l$$

c.f.

$$ds^2 = dx^2 + dy^2$$

$$ds^2 = dr^2 + r^2 d\theta^2$$

$$\text{量子幾何テンソル: } Q_{kl}(\theta) = \frac{\langle\partial_k \psi_{\theta} | \partial_l \psi_{\theta}\rangle}{\langle\psi_{\theta} | \psi_{\theta}\rangle} - \frac{\langle\partial_k \psi_{\theta} | \psi_{\theta}\rangle}{\langle\psi_{\theta} | \psi_{\theta}\rangle} \frac{\langle\psi_{\theta} | \partial_l \psi_{\theta}\rangle}{\langle\psi_{\theta} | \psi_{\theta}\rangle}$$

$$\text{フビニ・スタディ計量テンソル: } S_{kl}(\theta) = \text{Re}[Q_{kl}(\theta)]$$

(おまけ) 練習問題

ブロッホ球上の量子状態（以下）に対して、量子幾何テンソルとフビニ・スタディ計量テンソルの値を求めよ。
パラメータは θ 、 ϕ の二つなので、それぞれ2x2の行列となる。

$$|\psi(\theta, \phi)\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + e^{i\phi}\sin\frac{\theta}{2}|\downarrow\rangle$$

$$Q = \frac{1}{4} \begin{pmatrix} 1 & i \sin \theta \\ -i \sin \theta & \sin^2 \theta \end{pmatrix} \quad S = \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}$$

$$\mathcal{F}^2 \left[|\psi(\theta + \delta\theta, \phi + \delta\phi)\rangle, |\psi(\theta, \phi)\rangle \right] = \frac{1}{4}(\delta\theta)^2 + \frac{1}{4}(\sin \theta \delta\phi)^2$$

フィッシャー情報行列との関係

波動関数の値が実数かつ正の場合：

$$|\bar{\psi}_\theta\rangle = \sum_{\sigma} \sqrt{p_\theta(\sigma)} |\sigma\rangle$$

$$p_\theta(\sigma) \equiv |\bar{\psi}_\theta(\sigma)|^2$$
$$\sum_{\sigma} p_\theta(\sigma) = 1$$

フビニ・スタディ計量テンソル：

$$S_{kl} = \frac{1}{4} \sum_x p_\theta(x) \frac{\partial \ln p_\theta(x)}{\partial \theta_k} \frac{\partial \ln p_\theta(x)}{\partial \theta_l} = \frac{1}{4} \underline{I_{kl}}$$

フィッシャー情報行列

最適化手法の改善：Stochastic reconfiguration法

S. Sorella, Phys. Rev. B **64**, 024512 (2001)

基底状態探索の有効な手段：虚時間発展

$$e^{-\tau \mathcal{H}} |\psi^{(0)}\rangle = \sum_i e^{-\tau E_i} c_i^{(0)} \underbrace{|\phi_i\rangle}_{\text{固有状態}} = e^{-\tau E_0} \sum_i e^{-\tau(E_i - E_0)} c_i^{(0)} |\phi_i\rangle = e^{-\tau E_0} c_0^{(0)} |\phi_0\rangle \quad (\tau \rightarrow \infty)$$

※ 基底状態に縮退がないと仮定

虚時間発展を変分波動関数の表現能力の範囲内でできるだけ正確に再現する

$$\underbrace{\delta\theta}_{\text{パラメータ変化}} = \arg \min_{\delta\theta} \mathcal{F} \left(\underbrace{e^{-2\delta\tau \mathcal{H}} |\psi_\theta\rangle}_{\text{厳密な発展}}, \underbrace{|\psi_{\theta+\delta\theta}\rangle}_{\text{変分状態による近似}} \right)$$

$$= -\underbrace{\delta_\tau S^{-1}}_{\text{計量テンソル}} \underbrace{\partial_\theta \langle \mathcal{H} \rangle}_{\text{勾配}}$$

cf. Steepest descent (SD) 法

$$\delta\theta_k^{(t)} = -\delta_\tau g_k^{(t)} \quad \text{※ learning rateを}\delta_\tau\text{とおいた}$$

もう少し詳しく書くと


$$\delta\theta_k^{(t)} = -\delta_\tau \sum_l (S^{(t)})_{kl}^{-1} g_l^{(t)}$$

※ 機械学習分野の自然勾配法 (natural gradient) と本質的に等価

S.-I. Amari, Neural Comput. **10**, 251 (1998).

補足：SR法の導出

“距離”の最小化

$$\delta\tilde{\theta} = \arg \min_{\delta\theta} \mathcal{F}[e^{-2\delta_\tau \mathcal{H}} |\psi_\theta\rangle, |\psi_{\theta+\delta\theta}\rangle] \quad \xleftrightarrow{\text{等価}} \quad \delta\tilde{\theta} = \arg \max_{\delta\theta} F^2[e^{-2\delta_\tau \mathcal{H}} |\psi_\theta\rangle, |\psi_{\theta+\delta\theta}\rangle]$$

フィデリティ最大化

フィデリティ

$$F^2[e^{-2\delta_\tau \mathcal{H}} |\psi_\theta\rangle, |\psi_{\theta+\delta\theta}\rangle] = \frac{\langle \psi_\theta | e^{-2\delta_\tau \mathcal{H}} | \psi_{\theta+\delta\theta} \rangle \langle \psi_{\theta+\delta\theta} | e^{-2\delta_\tau \mathcal{H}} | \psi_\theta \rangle}{\langle \psi_\theta | e^{-4\delta_\tau \mathcal{H}} | \psi_\theta \rangle \langle \psi_{\theta+\delta\theta} | \psi_{\theta+\delta\theta} \rangle}$$

微小量の2次のオーダーまで展開

$$F^2 = 1 - \left(\sum_{k,l} \delta\theta_k S_{kl} \delta\theta_l + 2\delta_\tau \sum_k g_k \delta\theta_k + 4\delta_\tau^2 E_{\text{var}} \right)$$

$$E_{\text{var}} = \frac{\langle \psi_\theta | \mathcal{H}^2 | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} - \left(\frac{\langle \psi_\theta | \mathcal{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \right)^2 = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2$$

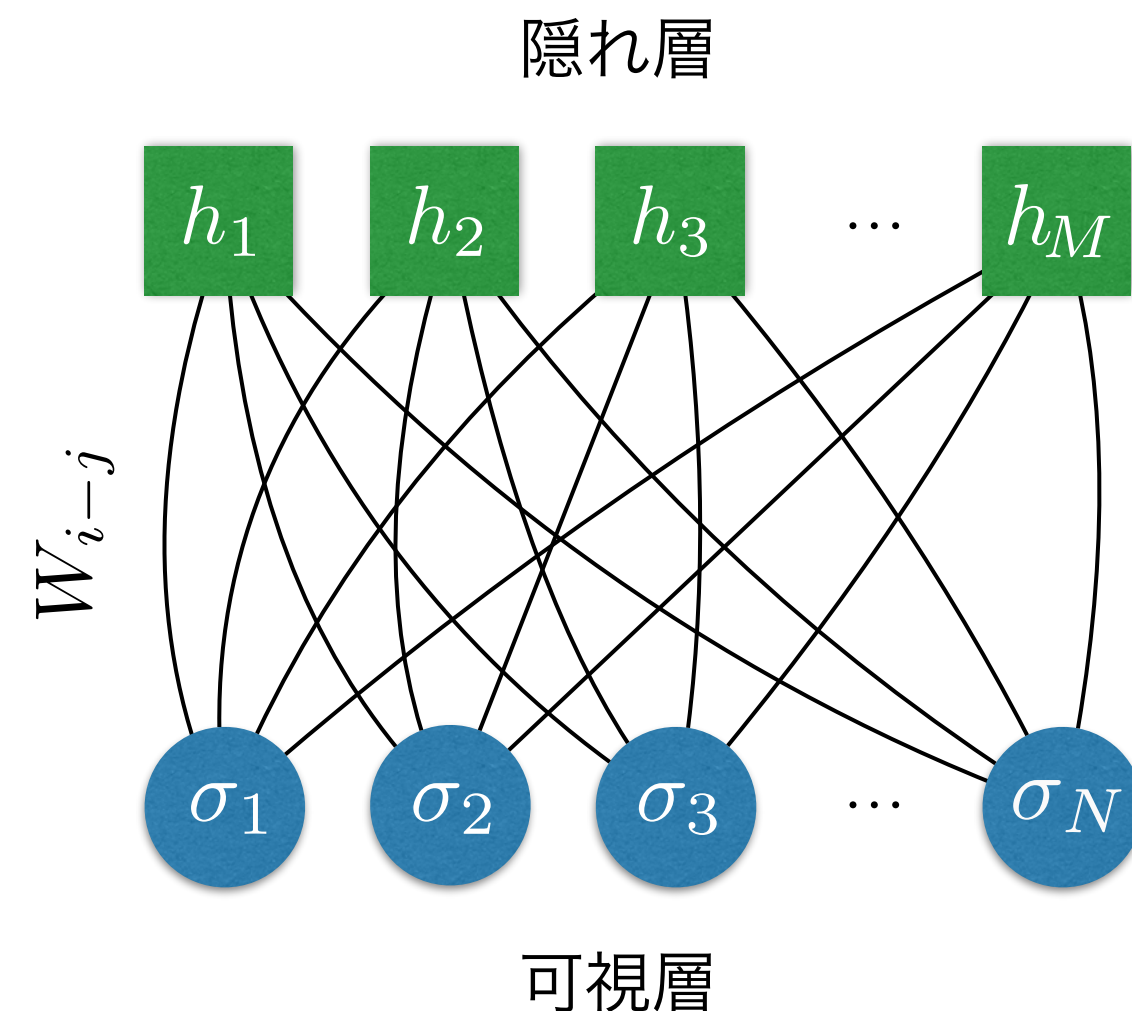
停留条件より

$$\delta\theta_k = -\delta_\tau \sum_l S_{kl}^{-1} g_l$$

1次元ハイゼンベルク模型への適用：SR法による最適化

ハミルトニアン：
$$\mathcal{H} = \sum_i (-\sigma_i^x \sigma_{i+1}^x - \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z)$$

変分波動関数：



$$\alpha(\text{hidden variable density}) = M/N = 1$$

$$\psi_{\theta}(\sigma) = \prod_j 2 \cosh \left(\sum_i W_{i-j} \sigma_i^z \right)$$

Carleo and Troyer Science 355, 602 (2017)とほぼ同じ設定

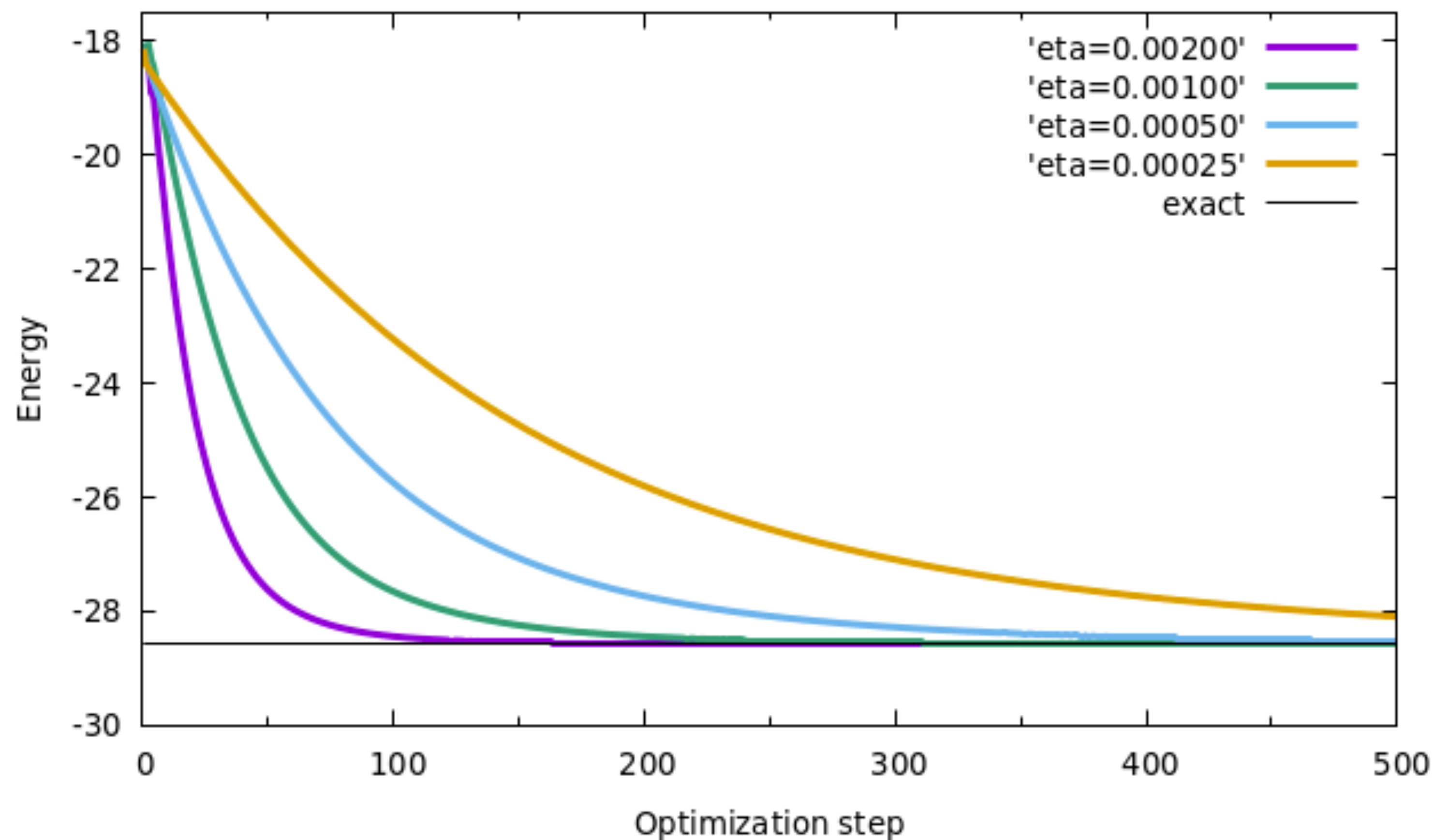
パラメータのアップデートルール：

$$\theta_k^{(t+1)} = \theta_k^{(t)} - \delta_{\tau} \sum_l \left(S^{(t)} \right)_{kl}^{-1} g_l^{(t)}$$

SR法：最適化（学習）の安定性

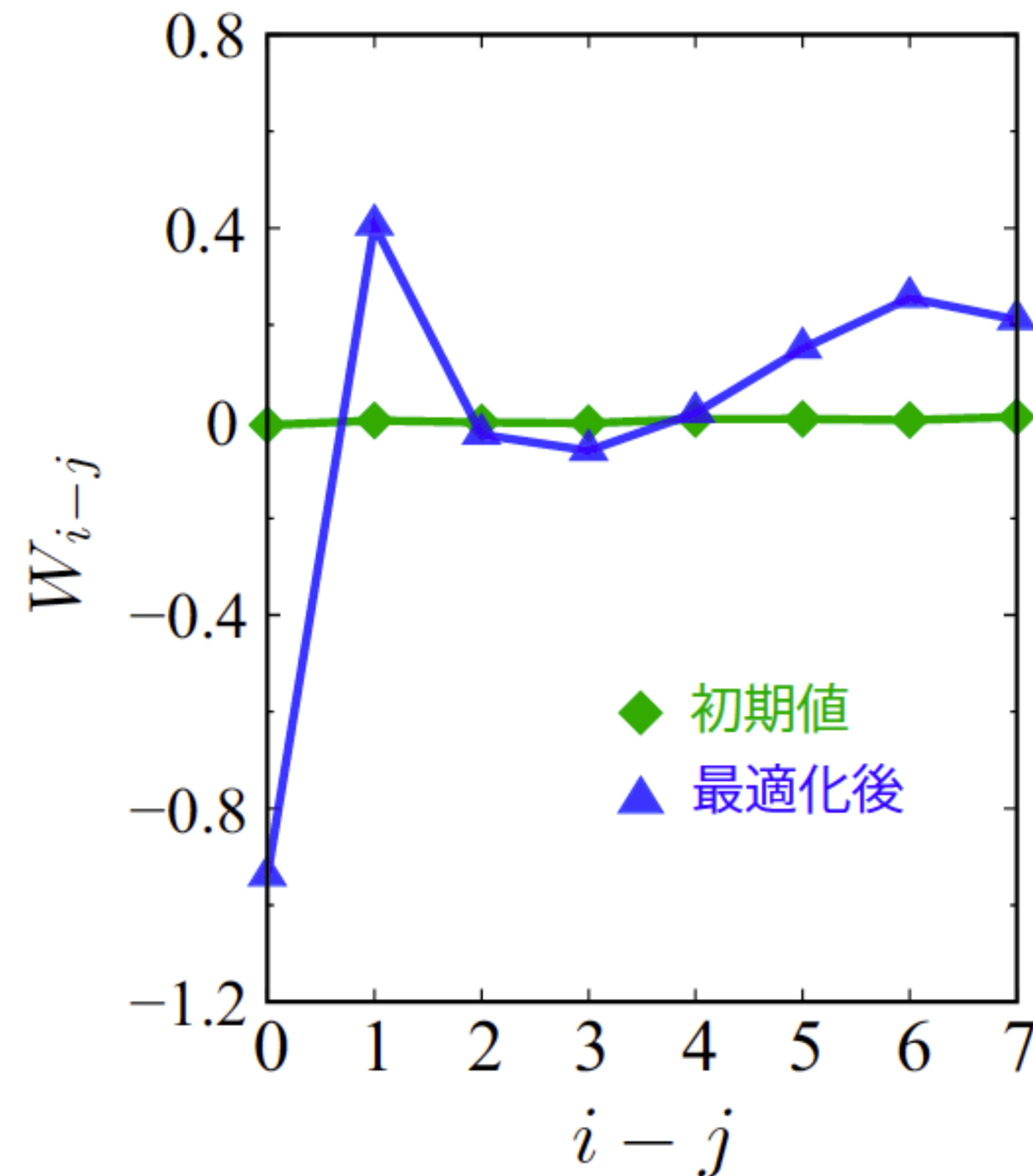
学習率 $\eta=\delta\tau$ を変えながら、(コスト関数) vs (最適化のIteration数) をプロット

16サイトに対する結果

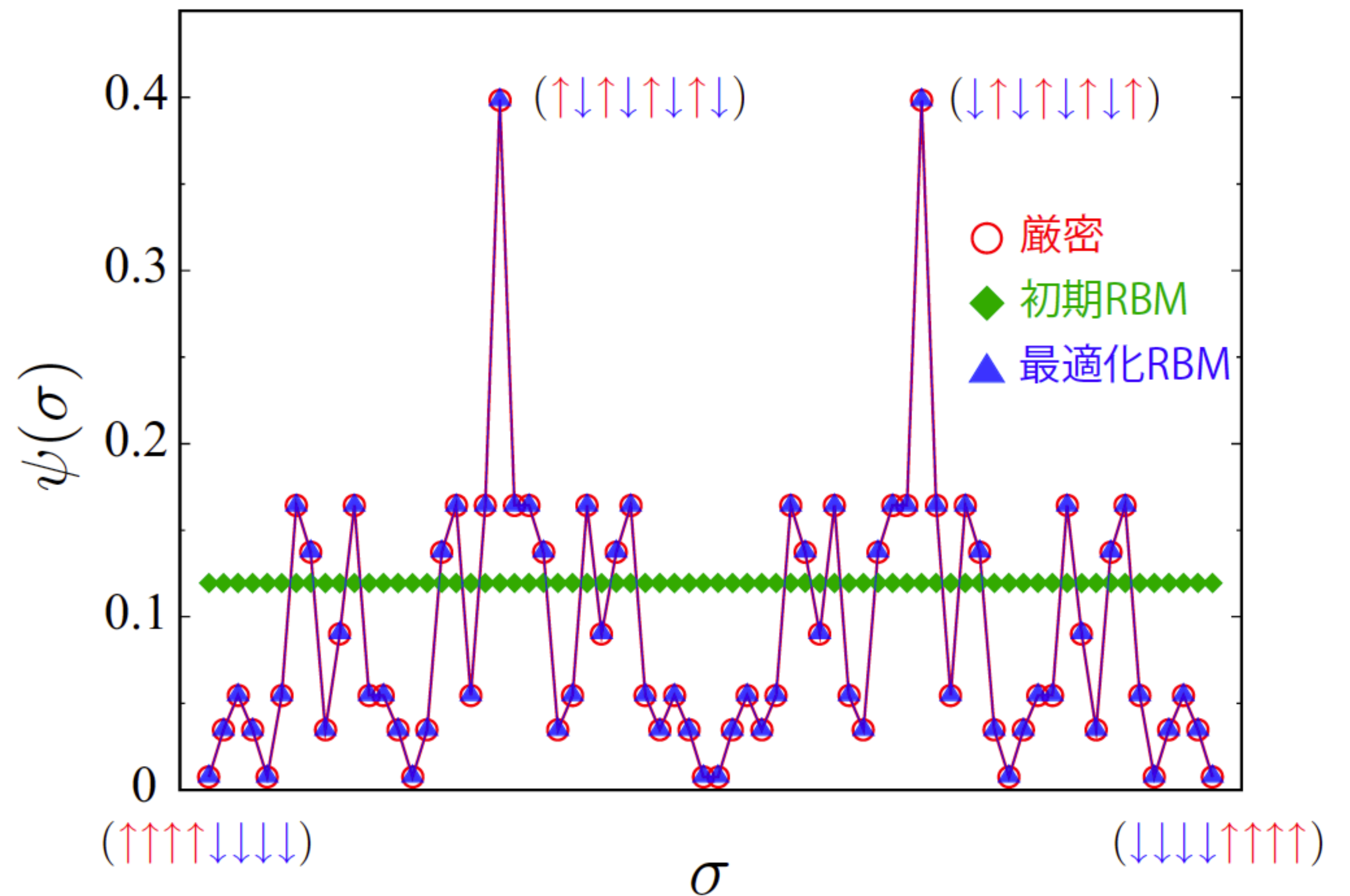


最適化前後の変分パラメータと波動関数（出力）：8サイトに対する結果

RBMのパラメータ



RBMの入力 σ と出力 $\psi(\sigma)$ の関係



ここまで：SR法の導入によって、最適化が安定化する様子を見た。

次の課題：より大きなシステムサイズを計算するには？ → モンテカルロ法の導入

※ これまでの例では、物理量計算に必要なスピン配置の和は厳密に計算していた（下は物理量の例としてエネルギーの計算の仕方を示す）

$$\langle \mathcal{H} \rangle = \frac{\sum_{\sigma} |\psi_{\theta}(\sigma)|^2 E_{\text{loc}}(\sigma)}{\sum_{\sigma} |\psi_{\theta}(\sigma)|^2}$$

変分モンテカルロ法

モンテカルロ法による期待値計算（エネルギーを例にとって説明）

$$\langle \mathcal{H} \rangle = \frac{\sum_{\sigma} |\psi_{\theta}(\sigma)|^2 E_{\text{loc}}(\sigma)}{\sum_{\sigma} |\psi_{\theta}(\sigma)|^2} \approx \frac{1}{N_s} \sum_{s=1}^{N_s} E_{\text{loc}}(\sigma_s) \quad * \quad |\psi_{\theta}(\sigma)|^2 \text{ を重みとしたサンプル生成を行う}$$

```
154 do iupdate = 1, N
155   call spin_flip_candidate(x,i1,i2)
156   xp(:) = x(:)
157   xp(i1) = -xp(i1)
158   xp(i2) = -xp(i2)
159   call calc_theta(xp,thetap)
160   call calc_amplitude_RBM(M,thetap,psi_xp)
161   if( grnd() > (psi_xp/psi_x)**2 ) cycle ! reject
162   !
163   ! accept => update configuration
164   !
165   x(:) = xp(:)
166   theta(:) = thetap(:)
167   psi_x = psi_xp
168 end do ! iupdate
```

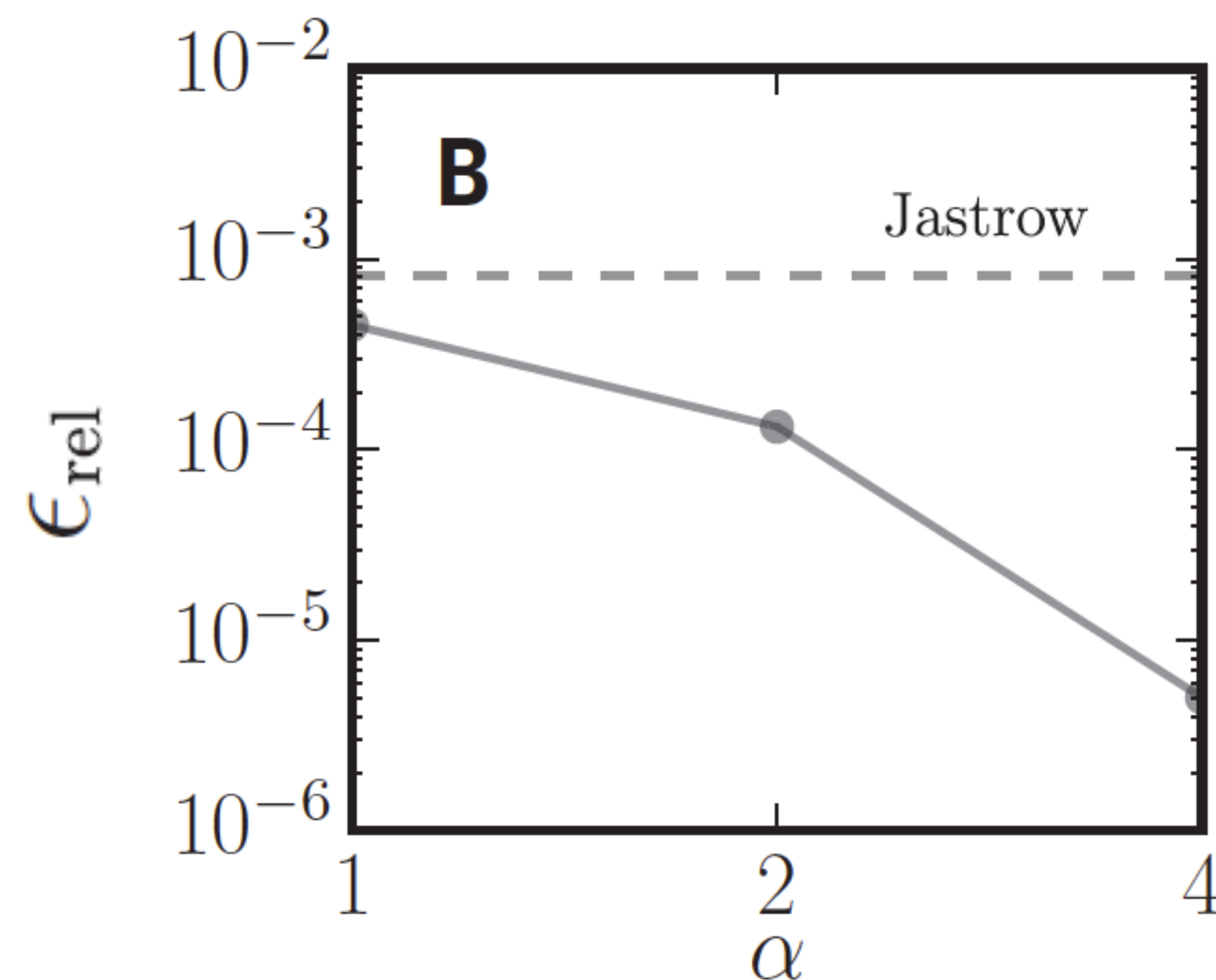
total Sz=0を守るような配置の中でのサンプル更新の提案

詳細釣り合い (grnd()は乱数)

発展課題：カルレオ・トロイヤー論文のFig. 3Bと同じ計算をする

80サイトの1次元ハイゼンベルク模型（周期境界条件）に対して、カルレオ・トロイヤーと同じ計算をする

$\alpha = M/N$ (hidden variable density)



$$\epsilon_{\text{rel}} = \frac{E_{\text{RBM}}(\alpha) - E_{\text{exact}}}{|E_{\text{exact}}|}$$

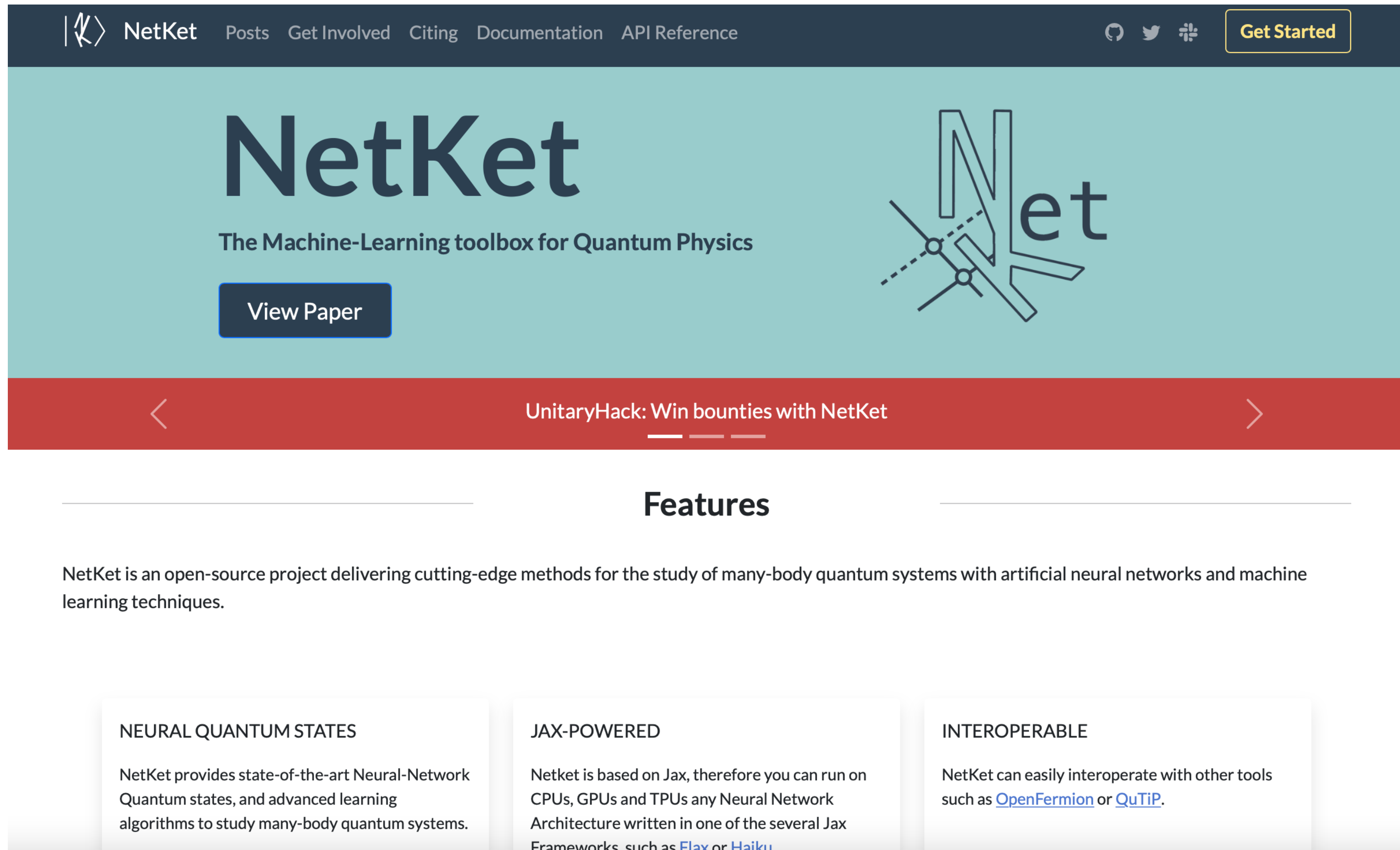
参照となる基底状態エネルギー：

$E_{\text{exact}} = -141.84830424$ （非常に大きなボンド次元のDMRGで計算）

Carleo and Troyer Science **355**, 602 (2017)

※ github上のプログラムを使うとすごい時間がかかる。少なくともファーストアップデートは取り入れるべき。できれば並列化も。

(参考) オープンソースコード：NetKet



The screenshot shows the NetKet website homepage. The header is dark blue with the NetKet logo (a stylized ket vector) and navigation links: Posts, Get Involved, Citing, Documentation, and API Reference. On the right of the header are social media icons for GitHub, Twitter, and a general share icon, along with a yellow 'Get Started' button. The main content area has a light blue background. On the left, the word 'NetKet' is written in large dark blue font, followed by the tagline 'The Machine-Learning toolbox for Quantum Physics' and a dark blue 'View Paper' button. On the right is a large 'Net' logo with a quantum circuit diagram. Below this is a red banner with the text 'UnitaryHack: Win bounties with NetKet' and navigation arrows. The 'Features' section is centered below the banner. It contains three white boxes with shadows. The first box, 'NEURAL QUANTUM STATES', describes state-of-the-art neural network quantum states and advanced learning algorithms. The second box, 'JAX-POWERED', states that NetKet is based on Jax and can run on CPUs, GPUs, and TPUs, mentioning its architecture is written in Jax frameworks like Flax or Haiku. The third box, 'INTEROPERABLE', notes that NetKet can easily interoperate with other tools like OpenFermion or QuTiP.

NetKet

The Machine-Learning toolbox for Quantum Physics

View Paper

UnitaryHack: Win bounties with NetKet

Features

NetKet is an open-source project delivering cutting-edge methods for the study of many-body quantum systems with artificial neural networks and machine learning techniques.

NEURAL QUANTUM STATES

NetKet provides state-of-the-art Neural-Network Quantum states, and advanced learning algorithms to study many-body quantum systems.

JAX-POWERED

Netket is based on Jax, therefore you can run on CPUs, GPUs and TPUs any Neural Network Architecture written in one of the several Jax Frameworks. such as [Flax](#) or [Haiku](#).

INTEROPERABLE

NetKet can easily interoperate with other tools such as [OpenFermion](#) or [QuTiP](#).

プラクティカルな部分のまとめ

Machine learning + Expert knowledge → state-of-the-art

※ 今回は性能を出すための最適化手法の工夫や、大規模化のためのモンテカルロ法について述べた

(個人的意見) 変分波動関数をどう構築かすべきか、という点についてより考察すべき

(おまけ) 近年の面白い動向:

トランスフォーマー型の変分波動関数の導入

[L. L. Viteritti, R. Rende, and F. Becca, Phys. Rev. Lett. **130**, 236401 \(2023\)](#)

(変分パラメータの数) $\gg$ (モンテカルロサンプル数) の場合に効率的にSR最適化を実行する手法の開発 (minSR法)

[A. Chen and M. Heyl, Nat. Phys. **20**, 1476 \(2024\)](#)