

量子アニーリング法を用いたクラスタ分析

田中宗^{A 1}、栗原賢一^B、宮下精二^{C,D}

^A 東京大学物性研究所、^B Google、^C 東京大学理学系研究科物理学専攻、^D CREST JST

本原稿は、論文“Quantum Annealing for Clustering” [1] を元に作成した。

1 組み合わせ最適化問題への新しいアプローチ：量子アニーリング

組み合わせ最適化問題は幅広い分野で現れる極めて重要な問題である。よく知られた例として、巡回セールスマン問題が挙げられる。巡回セールスマン問題とは、都市の集合と都市間の移動コスト²が与えられたとき、全ての都市を1回ずつ訪れ、出発地に戻る時のコストが最小のルートを求めよ、という問題である。また、ナップサック問題もよく知られた組み合わせ最適化問題の例である。これは、ナップサックに N 個（価値 p_i 、容積 c_i ）の品物を詰め込むとき、ナップサックの容量を超えずに、価値を最大にするにはどの品物を選べば良いか、という問題である。これらはNP困難のクラスに属する問題である。NP困難のクラスに属する問題は他にも、最小頂点被覆問題、最大独立集合問題、最大クリーク問題など多くの問題が知られている。これらの問題はいずれも、

- 対象とする系の要素が非常に多い。
- 特定の集合上で定義されたある実数値関数の最小値（または最大値）を求める。

という共通点を持っている。「ある実数値関数の最小値（または最大値）」を「自由エネルギーの最小値」と読み替えれば、マクロな個数の要素から構成されるシステムの平衡状態の性質の解明に広く成功を収めている統計物理学と考え方が共通していることがわかる。実際、統計物理学の知見を活かした最適化問題の解法が多く開発されている。統計物理学の発想に立てば、最適解を求める際には取り扱う問題のハミルトニアンを定義し、それを何らかの方法で解析することにより、その安定状態（平衡状態）を得るという過程を経る。最適化問題を解く一つの手法としてモンテカルロ法が挙げられる。モンテカルロ法はマルコフ過程を実行していることに他ならないので、エルゴード的であれば、無限回の試行で確実に平衡状態に行くことが保証されている。しかし実際には、無限回の試行をするわけにはいかず、工夫の無いモンテカルロ法を素直に実行しても、なかなか平衡状態を見つけられないという問題が起こる。最適解を求めたい問題の多くは、系の内部エネルギー構造の複雑性 [2–6]、あるいはエントロピー効果 [7–9] による緩和遅延が起こるのである。そのため、緩和遅延を抑制する効果を取り入れたアルゴリズムの構築は最適化問題にとって必要不可欠なものである。そのような背景の中、1983年にKirkpatrick氏らによって提案されたシミュレーテッド・アニーリング法 [10, 11] が注目を浴びようになる。温度を徐々に下げるというプロセスを用い、最終的に基底状態を得るという方法である。熱ゆらぎの効果を巧みに利用し、自由エネルギーの形状を操作することにより、基底状態の性質を得るこの方法は、その後様々な分野の最適化問題に適用されてきた。またそれぞれが異なる温度にあるいくつかのレプリカを用意し、レプリカ間の状態を巧みに交換させることにより平衡状態を得る方法として、福島孝治氏、根本幸児氏により提案された交換法 [12, 13] が挙げられる。この方法もシミュレーテッドアニーリング法と同様、広く用いられている。

これらの方法に対し、元のハミルトニアンに量子項（非対角項）を導入し、それを徐々に弱めることにより、元のハミルトニアンの安定状態を得る方法として量子アニーリング法 [14–21] が確立された。この方法

¹E-mail: shu-t@issp.u-tokyo.ac.jp

²移動距離でも、移動に要する料金でも、あるいは移動時間でも構わない。

は外部パラメータを導入し、ハミルトニアンをあえて一般化させ、基底状態が自明に分かる点から出発し、求めたいハミルトニアンを表現するパラメータへと変化させることで安定状態を得る方法である。これまでの方法は熱ゆらぎの効果により状態遷移を促していたのに対し、量子アニーリング法は量子ゆらぎの効果を用いて、状態遷移を促すという方法である。図1はシミュレーテッド・アニーリング法 (SA) と従来型の量子アニーリング法 (QA)、また本研究で我々が考察したハイブリッド型量子アニーリング法 (SA+QA)³の概念図である。ハイブリッド型量子アニーリング法は、シミュレーテッド・アニーリングにおける温度及び量子アニーリング法における量子項を同時に制御することにより最適解を得る、両者の合わせ技となっている。

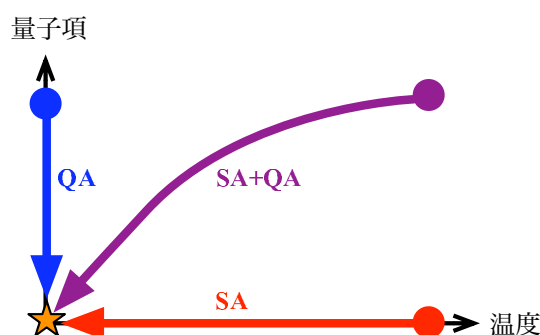


図 1: シミュレーテッドアニーリング (SA)、量子アニーリング (QA)、ハイブリッド型量子アニーリング法の概念図 (SA+QA)。いずれの方法も解析が容易な点 (丸印) から始め、解きたい問題を表す点 (星印) まで徐々に制御パラメータを変化させている。

2 量子アニーリング法の実装方法

量子アニーリング法は、量子統計力学の原理を計算手法として利用した方法である。量子アニーリング法は大きく分けて以下の3つの実現法がある (表1)。

確率的手法 (理論・数値計算)	決定論的手法 (理論・数値計算)	実験的手法
量子モンテカルロ法	実時間ダイナミクス (断熱量子計算) 平均場近似 変分ベイズ推定	光学格子

表 1: 量子アニーリング法の様々な実現方法

確率的手法は、主に量子モンテカルロ法 [22] で実装される。これは量子強相関係の平衡状態を得るために確立された手法であり、多くの工夫されたアルゴリズムが知られている [23–25]。また近年、量子アニーリング法に適した効率の良いアルゴリズムが提案された [26, 27]。量子モンテカルロ法は大規模な系を扱う上で有力な手法の一つである。本原稿で紹介する研究では量子モンテカルロ法を用いた。

決定論的手法にはいくつかの実装方法がある。一番目の例として、実時間ダイナミクスを追跡する手法がある。量子力学の基礎方程式であるシュレディンガー方程式を直接解くことになる [17, 28–31]。この手法は、実験的に見られる時間発展を直接計算機内で (厳密に解ける場合ならば、解析的に) 追跡する方法である。そのため量子情報理論の分野でも活発に議論されており、量子断熱計算とも呼ばれる。この方法の欠点は、ハミルトニアンを対角化するというプロセスを経るので、一般に系のサイズに対し、必要メモリサイ

³本原稿では以降、QA と SA+QA を区別せず、量子アニーリングと呼ぶ。

ズが指数関数的に増大してしまい、最適化問題で必要とされる要素数をシミュレートするには遠く及ばない。また、時間発展密度行列繰り込み群の手法を用いた時間発展の場合 [32,33] は、直接ハミルトニアンを対角化する手法に比べて大きいサイズが出来るが、系の形状がある程度限定されてしまうという問題点がある。実時間ダイナミクスを追跡する方法は常に、系のサイズと形状の問題が付きまとう。二番目の例として、画像修復の問題を中心に田中和之氏、堀口剛氏によって進められた平均場近似を用いた方法が挙げられる [34]。この手法では量子項を含む局所場において、セルフ・コンシステント方程式を数值的に解くことにより、安定解を得る手法である。三番目の例として、変分ベイズ推定法が挙げられる。これは本原稿で述べる研究とは独立に我々が進めてきた研究の1つであり、変分原理を元にしたアルゴリズムである。詳細は論文 [35] を参考されたい。

また量子アニーリング法は計算機内のシミュレーションだけではなく、実験的な実現可能性も示唆されている。近年の実験技術の進展により、量子強相関係を人工的にシミュレートする方法が提案されている。その中で最も代表的なものが、光格子と呼ばれるものである [36]。我々が性質（基底状態や相）を知りたいハミルトニアンがあったとする。そのハミルトニアンを実現する格子系を光を用いて実現し、その格子系にある粒子の振る舞いを観測することで、知りたいハミルトニアンの性質を得るという手法である。

このように量子アニーリング法は様々な方法で実装が可能であり、最適化問題に対する新しい解法として広く期待されている手法である。

3 経路積分量子モンテカルロ法

我々の研究内容に触れる前に、経路積分量子モンテカルロ法について紹介する。簡単のため、周期的境界条件を課した n 個の 1 次元強磁性イジングモデルに一樣な横磁場が引加されているという状況を考える。ハミルトニアンは、

$$\mathcal{H} = -J \sum_{i=1}^n \sigma_i^z \sigma_{i+1}^z - \Gamma \sum_{i=1}^n \sigma_i^x \equiv \mathcal{H}_c + \mathcal{H}_q, \quad (\sigma_{n+1} = \sigma_1) \quad (1)$$

とする。ただしここで、 σ_i^α は、サイト i でのパウリ行列の α 成分である。具体的にパウリ行列を書き下すと、

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2)$$

である。

物理量 A の熱平衡量 $\langle A \rangle$ を求めるには、

$$\langle A \rangle = \frac{\text{Tr} A e^{-\beta \mathcal{H}}}{\text{Tr} e^{-\beta \mathcal{H}}} \quad (3)$$

を計算する必要がある。ここで重要なのが $e^{-\beta \mathcal{H}}$ の取り扱いである。仮に \mathcal{H} が対角行列であった場合、 $e^{-\beta \mathcal{H}}$ も対角行列となり、各行列要素は、

$$(e^{-\beta \mathcal{H}})_{ii} = e^{-\beta E_i} \quad (4)$$

となる。ここで \mathcal{H} の対角要素を E_i とした。しかし、 \mathcal{H} が非対角項を含む場合、一般に

$$(e^{-\beta \mathcal{H}})_{ij} \neq e^{-\beta \mathcal{H}_{ij}} \quad (5)$$

となる。それは行列の指数関数が、

$$e^A = \sum_{m=0}^{\infty} \frac{1}{m!} A^m \quad (6)$$

となっているためである。 $e^{-\beta\mathcal{H}}$ を求めるには、任意の自然数 m に対し、行列の冪乗を計算しなければならない。 \mathcal{H} のサイズがそれほど大きくなければ計算機上で $e^{-\beta\mathcal{H}}$ を計算できる。一方でモンテカルロ法を用いるのは、系のサイズが大きいため、数値対角化によって正確に $e^{-\beta\mathcal{H}}$ を求めることが出来ない場合である。そのため何らかの方法で $e^{-\beta\mathcal{H}}$ を計算可能な形にしなければならない。そこで用いられるのが経路積分表示である。

3.1 鈴木トロツタ分解による経路積分表示

鈴木トロツタ分解 [22, 37] による経路積分表示を行う。式 (1) で与えられるハミルトニアンを用いて、分配関数は

$$Z = \text{Tr} e^{-\beta\mathcal{H}} = \text{Tr} e^{-\beta(\mathcal{H}_c + \mathcal{H}_q)} = \sum_{\sigma} \langle \sigma | e^{-\beta(\mathcal{H}_c + \mathcal{H}_q)} | \sigma \rangle \quad (7)$$

となる。ここで、十分大きな m を用いて

$$\exp\left[-\frac{1}{m}\beta\mathcal{H}\right] = \exp\left[-\frac{1}{m}\beta(\mathcal{H}_c + \mathcal{H}_q)\right] = e^{-\frac{1}{m}\beta\mathcal{H}_c} e^{-\frac{1}{m}\beta\mathcal{H}_q} + \mathcal{O}\left(\left(\frac{\beta}{m}\right)^2\right) \quad (8)$$

と表されることを用いると、式 (7) は

$$\begin{aligned} Z = & \sum_{\sigma_k = \pm 1} \langle \sigma_1 | e^{-\beta\mathcal{H}_c/m} | \sigma'_1 \rangle \langle \sigma'_1 | e^{-\beta\mathcal{H}_q/m} | \sigma_2 \rangle \\ & \times \langle \sigma_2 | e^{-\beta\mathcal{H}_c/m} | \sigma'_2 \rangle \langle \sigma'_2 | e^{-\beta\mathcal{H}_q/m} | \sigma_3 \rangle \\ & \times \dots \\ & \times \langle \sigma_n | e^{-\beta\mathcal{H}_c/m} | \sigma'_n \rangle \langle \sigma'_n | e^{-\beta\mathcal{H}_q/m} | \sigma_1 \rangle \end{aligned} \quad (9)$$

となる。ここで、 $|\sigma_k\rangle$ は、 L 個のスピンの直積空間を表す⁴。

$$|\sigma_k\rangle = |\tilde{\sigma}_{1,k}\rangle \otimes |\tilde{\sigma}_{2,k}\rangle \otimes \dots \otimes |\tilde{\sigma}_{L,k}\rangle \quad (10)$$

式 (9) は、以下の2つの値を計算すれば求められる。

$$\langle \sigma_k | e^{-\beta\mathcal{H}_c/m} | \sigma'_k \rangle \quad (11)$$

$$\langle \sigma'_k | e^{-\beta\mathcal{H}_q/m} | \sigma_{k+1} \rangle \quad (12)$$

式 (11) については、 \mathcal{H}_c が対角行列であることから、

$$\tilde{\sigma}_j^z |\sigma_k\rangle = \tilde{\sigma}_{j,k} |\sigma_k\rangle \quad (13)$$

となり、

$$\langle \sigma_k | e^{-\beta\mathcal{H}_c/m} | \sigma'_k \rangle = \exp\left(\frac{\beta J}{m} \sum_{j=1}^n \tilde{\sigma}_{j,k} \tilde{\sigma}_{j+1,k}\right) \prod_{j=1}^n \delta(\tilde{\sigma}_{j,k}, \tilde{\sigma}'_{j,k}) \quad (14)$$

が得られる。一方、式 (12) については、

$$\langle \sigma'_k | e^{-\beta\mathcal{H}_q/m} | \sigma_{k+1} \rangle = \left[\frac{1}{2} \sinh\left(\frac{2\beta\gamma}{m}\right)\right]^{n/2} \exp\left[\frac{1}{2} \log \coth\left(\frac{\beta\Gamma}{m} \sum_{j=1}^n \tilde{\sigma}_{j,k} \tilde{\sigma}_{j,k+1}\right)\right] \quad (15)$$

⁴本原稿では $\tilde{\sigma}_1$ (チルダ付き) とした場合は1つの要素を表し、 σ (チルダ無し) は n 個の要素全体の状態を表す。

が成立する。以上から、式 (7) によって表される分配関数はトロツタ数 m を用いて、以下のように表すことができる。

$$Z = \lim_{m \rightarrow \infty} \left[\frac{1}{2} \sinh \left(\frac{2\beta\gamma}{m} \right) \right]^{n/2} \sum_{\{\sigma_{j,k} = \pm 1\}} \exp \left[\sum_{j=1}^n \sum_{k=1}^m \left(\frac{\beta J}{m} \tilde{\sigma}_{j,k} \tilde{\sigma}_{j+1,k} \right) + \frac{1}{2} \log \coth \left(\frac{\beta\gamma}{m} \right) \tilde{\sigma}_{j,k} \tilde{\sigma}_{j,k+1} \right] \quad (16)$$

と表すことができる。これは元々考えていた1次元横磁場イジングモデルの分配関数と2次元磁場無しイジングモデルの分配関数とが等価であることを示している。自由エネルギーは分配関数から直ちに求められるので、1次元横磁場イジングモデル(量子系)の安定状態の性質を調べるには、2次元磁場無しイジングモデル(古典系)の安定状態の性質を調べれば良いことが分かる。追加した次元の方向をトロツタ方向と呼ぶことにする。マップされた2次元磁場無しイジングモデルは、トロツタ方向に周期境界条件が課されている。ここで式(16)の末項は横磁場 Γ に対し単調減少の関数になっている。

4 クラスタ分析とは

研究内容について述べる前にまず、クラスタ分析とは何か、簡単に説明する。世の中には多くのデータがあり、色々な場面でそのグループ分けが必要になることがある。例えば Web サイトや新聞などには非常に多くの記事(データベース)が含まれている。これらの記事を政治、経済、スポーツ、芸能などカテゴリーに分類したいときなどがそれに該当する。別の例として、アンケート調査が挙げられる。ある集団に対してアンケートをとったとき、その集団がいくつかのグループに分類できる。グループに分類するというのは、別の見方をすればそのグループのトピックを抽出することに相当する。そこで用いられるのがクラスタ分析と呼ばれる手法である。クラスタ分析は元々は生物分類学などの分野から発展してきたものであり、確立された手法として、認知科学、心理学、社会学や経済工学など幅広い分野で用いられている手法である。上に挙げた2つの例で分かるように、クラスタ分析は実社会においても多くの応用例がある。クラスタ分析を端的に表すと、非常に多くの要素から構成される全体集合を部分集合に分割することである。ここで分割された各部分集合をクラスタと呼ぶ。

簡単な場合についてみてみよう。全体集合を2次元平面とし、その上で多くの点が散らばっている状況を考える。これを「自然な」4つの部分空間に分けることを考える。全体集合をクラスタに分割する方法として、図2(a),(b),(c)の3つを考えよう。ここで、図2(c)は「自然な」4つの部分空間への分割である。これを自由エネルギー描像で考えると、図2(d)のようになっていると理解できる。我々の目的は図2(c)の解を求めることである⁵。

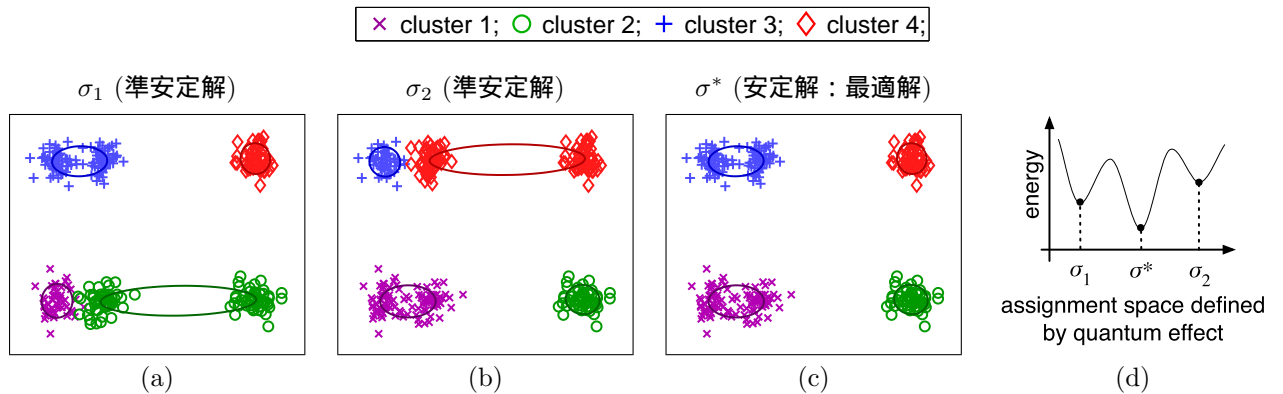


図 2: 4つの正規分布からなる混合正規分布によるクラスタ分析の結果

⁵ここでいうクラスタ分割の「自然さ」は、後に出てくる式(18)の $p_{\text{prob-model}}(X, \sigma)$ の関数形の取り方に依存する。例えば、図2の場合、4つの正規分布の和で与えられる確率分布関数で分割したと思えば、図2(c)が「自然な」解となっていることがわかるだろう。

5 クラスタ分析に対する量子アニーリング法

量子統計物理学の知見を用いてクラスタ分析を取り扱うために、ハミルトニアンを定義する。以降では、 k をクラスタ数、 n を要素の数とする。簡単のため、 $k = 2, n = 2$ の場合を考える。要素 i がクラスタ A に含まれる場合を $\tilde{\sigma} = (1, 0)^T$ 、クラスタ B に含まれる場合を $\tilde{\sigma} = (0, 1)^T$ とする。 $\sigma = \tilde{\sigma}_1 \otimes \tilde{\sigma}_2$ で各状態を表す。

通常のクラスタ分析において、ハミルトニアンは

$$\mathcal{H}_c = \begin{pmatrix} E(\sigma^{(1)}) & 0 & 0 & 0 \\ 0 & E(\sigma^{(2)}) & 0 & 0 \\ 0 & 0 & E(\sigma^{(3)}) & 0 \\ 0 & 0 & 0 & E(\sigma^{(4)}) \end{pmatrix} \quad (17)$$

と表せる。ただし、

$$E(\sigma^{(i)}) \equiv -\log p_{\text{prob-model}}(X, \sigma^{(i)}) \quad (18)$$

である。ここで、 $E(\sigma^{(i)})$ は状態 i にあるときの「内部エネルギー」を表す。ただしここで、表 2 に $\sigma^{(i)}$ に対応する状態を示した。また式 (18) で、 X はデータ列であり、図 2 の各点に相当する。 $p_{\text{prob-model}}(X, \sigma)$ とは、確率モデルを表す。 $p_{\text{prob-model}}(X, \sigma)$ は我々が考えたい問題に応じて用意する必要がある。最も簡単な例としては、混合ガウス分布が挙げられる。これはガウス分布の和で与えられる確率分布関数である。また文書のトピックを抽出するのによく用いられる確率モデルとして Latent Dirichlet Allocation (LDA) [38] がよく知られている。本研究ではこれら 2 つの確率モデルを用いて解析を行った。

状態	要素 1	要素 2
$\sigma^{(1)} = (1, 0, 0, 0)^T$	クラスタ A	クラスタ A
$\sigma^{(2)} = (0, 1, 0, 0)^T$	クラスタ B	クラスタ A
$\sigma^{(3)} = (0, 0, 1, 0)^T$	クラスタ A	クラスタ B
$\sigma^{(4)} = (0, 0, 0, 1)^T$	クラスタ B	クラスタ B

表 2: 各状態におけるクラスタ分析との対応

式 (18) で与えられたハミルトニアンに量子項（非対角項）を加える。どの非対角項に値を入れるかについては任意性があるが、我々は以下のように量子項を導入した。

$$\mathcal{H}_q = \sum_{i=1}^n \chi_i, \quad (19)$$

$$\chi_i = \mathbb{E}^{(k)} - \mathbb{1}^{(k)} \quad (20)$$

ただしここで、 $\mathbb{E}^{(k)}$ は $k \times k$ の単位行列であり、 $\mathbb{1}^{(k)}$ は全要素が 1 の $k \times k$ 行列である。 $k = 2, n = 2$ の場合、ハミルトニアンは

$$\mathcal{H} = \mathcal{H}_c + \Gamma \mathcal{H}_q = \begin{pmatrix} E(\sigma^{(1)}) & -\Gamma & -\Gamma & 0 \\ -\Gamma & E(\sigma^{(2)}) & 0 & -\Gamma \\ -\Gamma & 0 & E(\sigma^{(3)}) & -\Gamma \\ 0 & -\Gamma & -\Gamma & E(\sigma^{(4)}) \end{pmatrix} \quad (21)$$

となる。

ハミルトニアンを定義したので、次にクラスタ分析に対するモンテカルロ法について考えよう。まず始めにハミルトニアンが対角行列のときについて考える。この場合、先ほど述べたように通常の古典モンテカル

口法を用いて解析することが容易に可能である。このとき、温度 T において σ という状態をとる確率は、

$$p_{\text{SA}}(\sigma; \beta) = \frac{\exp[-\beta E(\sigma)]}{\sum_{\sigma} \exp[-\beta E(\sigma)]} = \frac{1}{Z} \langle \sigma | e^{-\beta \mathcal{H}_c} | \sigma \rangle \quad (22)$$

と表せる。ただしここで分配関数は、

$$Z = \text{Tr} e^{-\beta \mathcal{H}_c} = \sum_{\sigma} \langle \sigma | e^{-\beta \mathcal{H}_c} | \sigma \rangle = \sum_{\sigma} e^{-\beta E(\sigma)} \quad (23)$$

と書けることに注意しておく。この系の状態を熱浴法を用いて更新するには、

$$p_{\text{SA}}^{\text{update}}(\tilde{\sigma}_i | \sigma \setminus \tilde{\sigma}_i) = \frac{\exp[-\beta E(\sigma)]}{\sum_{\tilde{\sigma}_i} \exp[-\beta E(\sigma)]} \quad (24)$$

を計算すれば良い。ここで、 $\sigma \setminus \tilde{\sigma}_i$ は、 $\{\tilde{\sigma}_j | j \neq i\}$ を表す。分母は $\mathcal{O}(k)$ で計算できる量である。

一方量子項を導入した場合、温度 T 、量子ゆらぎの強さ Γ において σ という状態を取る確率は、式 (22) と同様に、

$$p_{\text{QA}}(\sigma; \beta, \Gamma) = \frac{\langle \sigma | e^{-\beta \mathcal{H}} | \sigma \rangle}{\sum_{\sigma} \langle \sigma | e^{-\beta \mathcal{H}} | \sigma \rangle} = \frac{\langle \sigma | e^{-\beta \mathcal{H}} | \sigma \rangle}{Z} \quad (25)$$

となる。3節で述べたように、 $\langle \sigma | e^{-\beta \mathcal{H}} | \sigma \rangle$ は $\langle \sigma | e^{-\beta \mathcal{H}_c} | \sigma \rangle = \exp[-\beta E(\sigma)]$ とは異なり、容易に計算できない。そのため、鈴木トロツタ分解による経路積分表示を用いて計算可能な形にする。式 (25) を鈴木トロツタ分解すると(いま σ を σ_1 とおいた)

$$p_{\text{QA}}(\sigma_1; \beta, \Gamma) = \frac{1}{Z} \langle \sigma_1 | \left(e^{-\frac{\beta}{m} \mathcal{H}_c} e^{-\frac{\beta \Gamma}{m} \mathcal{H}_q} \right)^m | \sigma_1 \rangle + \mathcal{O}\left(\frac{1}{m}\right) \quad (26)$$

$$= \frac{1}{Z} \sum_{\sigma'_1} \sum_{\sigma_2} \cdots \sum_{\sigma_m} \sum_{\sigma'_m} \langle \sigma_1 | e^{-\frac{\beta}{m} \mathcal{H}_c} | \sigma'_1 \rangle \langle \sigma'_1 | e^{-\frac{\beta \Gamma}{m} \mathcal{H}_q} | \sigma_2 \rangle \cdots \\ \times \langle \sigma_m | e^{-\frac{\beta}{m} \mathcal{H}_c} | \sigma'_m \rangle \langle \sigma'_m | e^{-\frac{\beta \Gamma}{m} \mathcal{H}_q} | \sigma_1 \rangle \quad (27)$$

$$= \frac{1}{Z} \sum_{\sigma'_1} \sum_{\sigma_2} \cdots \sum_{\sigma_m} \sum_{\sigma'_m} \prod_{j=1}^m \langle \sigma_j | e^{-\frac{\beta}{m} \mathcal{H}_c} | \sigma'_j \rangle \langle \sigma'_j | e^{-\frac{\beta \Gamma}{m} \mathcal{H}_q} | \sigma_{j+1} \rangle \quad (28)$$

となる。ここで、

$$s(\sigma_j, \sigma_{j+1}) \equiv \frac{1}{n} \sum_{i=1}^n \delta(\tilde{\sigma}_{j,i}, \tilde{\sigma}_{j+1,i}), \quad (29)$$

$$f(\beta, \Gamma) \equiv n \log \left(1 + \frac{k}{e^{\frac{k\beta\Gamma}{m}} - 1} \right) \quad (30)$$

と定義すると、

$$\langle \sigma_j | e^{-\frac{\beta}{m} \mathcal{H}_c} | \sigma'_j \rangle = \exp \left[-\frac{\beta}{m} E(\sigma_j) \right] \delta(\sigma_j, \sigma'_j) \propto p_{\text{SA}} \left(\sigma_j; \frac{\beta}{m} \right) \delta(\sigma_j, \sigma'_j), \quad (31)$$

$$\langle \sigma'_j | e^{-\frac{\beta \Gamma}{m} \mathcal{H}_q} | \sigma_{j+1} \rangle \propto \exp [s(\sigma'_j, \sigma_{j+1}) f(\beta, \Gamma)] \quad (32)$$

と書ける。式 (32) を示そう。そのために、 \mathcal{H}_q の冪乗を計算する。 $\mathcal{H}_q = \sum_i \chi_i$ であるため、 χ_i の冪乗を計算すれば良い。

$$\chi_i^l = (\mathbb{E}_k - \mathbb{1}_k)^l = \sum_{j=0}^l \binom{l}{j} \mathbb{E}_k^j (-\mathbb{1}_k)^{l-j} = \mathbb{E}_k + \frac{1}{k} [(1-k)^l - 1] \mathbb{1}_k \quad (33)$$

となるから、式 (32) は、

$$\langle \sigma'_j | e^{-\frac{\beta\Gamma}{m} \mathcal{H}_q} | \sigma_{j+1} \rangle = \prod_{i=1}^n \sum_{l=0}^{\infty} \frac{1}{l!} \left(-\frac{\beta\Gamma}{m} \right)^l \langle \tilde{\sigma}'_{j,i} | \mathbb{E}_k + \frac{1}{k} [(1-k)^l - 1] \mathbb{1}_k | \tilde{\sigma}_{j+1,i} \rangle \quad (34)$$

$$= \prod_{i=1}^n \left[e^{-\frac{\beta\Gamma}{m}} \delta(\tilde{\sigma}'_{j,i}, \tilde{\sigma}_{j+1,i}) + \frac{1}{k} e^{-\frac{\beta\Gamma}{m}(1-k)} - \frac{1}{k} \right] \quad (35)$$

$$\propto e^{s(\sigma'_j, \sigma_{j+1}) f(\beta, \Gamma)} \quad (36)$$

となる。以上より式 (28) は、

$$p_{\text{QA}}(\sigma_1; \beta, \Gamma) = \frac{1}{Z} \langle \sigma_1 | \left(e^{-\frac{\beta}{m} \mathcal{H}_c} e^{-\frac{\beta}{m} \mathcal{H}_q} \right)^m | \sigma_1 \rangle = \frac{1}{Z} \sum_{\sigma_2} \cdots \sum_{\sigma_m} \prod_{j=1}^m p_{\text{SA}} \left(\sigma_j; \frac{\beta}{m} \right) e^{s(\sigma_j, \sigma_{j+1}) f(\beta, \Gamma)} \quad (37)$$

となる。横磁場イジングモデルの例と同様、鈴木トロツタ分解による経路積分表示により、次元を1つ上げた古典系と等価であることが示された。先ほどの横磁場イジングモデルの例と同様、量子項の強さ Γ に対し、 $f(\beta, \Gamma)$ は単調減少関数となっている。つまり、量子項を弱めると $f(\beta, \Gamma)$ は強くなり、トロツタ方向に相関が生まれることになる。式 (37) より鈴木トロツタ分解による経路積分表示を用いた量子モンテカルロ法では、

$$p_{\text{QA-ST}}^{\text{update}} \left(\tilde{\sigma}_{j,i} | \{\sigma_j\}_{j=1}^m \setminus \tilde{\sigma}_{j,i}; \beta, \Gamma \right) = \frac{\exp \left[-\frac{\beta}{m} E(\sigma_j) + (s(\sigma_{j-1}, \sigma_j) + s(\sigma_j, \sigma_{j+1})) f(\beta, \Gamma) \right]}{\sum_{\tilde{\sigma}_{j,i}} \exp \left[-\frac{\beta}{m} E(\sigma_j) + (s(\sigma_{j-1}, \sigma_j) + s(\sigma_j, \sigma_{j+1})) f(\beta, \Gamma) \right]} \quad (38)$$

を計算して状態を更新していけば良いことがわかる。図3にシミュレーテッドアニーリング、交換モンテカルロ法、ならびに今回我々が用いた量子アニーリング法概念図を表した。シミュレーテッドアニーリング (SA) は、 m 個の独立なサンプルを用いて徐々に温度を下げながらシミュレーションすることに相当する。また交換モンテカルロ法 (EXMC) は独立な m 個のサンプルを異なる温度におき、ある確率で隣接レプリカ間の交換を行う。また量子アニーリング法 (SA+QA) は「並列度」 m の「擬並列化」⁶ をしているとも見なせる。

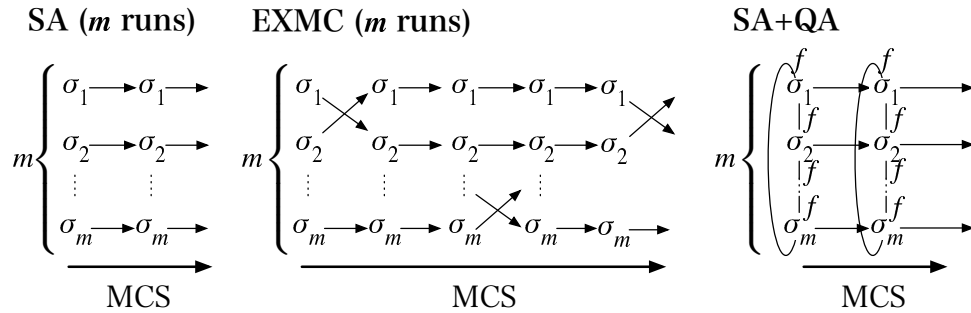


図3: シミュレーテッドアニーリング (SA)、交換モンテカルロ法 (EXMC)、量子アニーリング法 (SA+QA) の概念図

5.1 クラスタ名の問題

式 (29) により定義されている $s(\sigma_j, \sigma_{j+1})$ は、 σ_j と σ_{j+1} がどの程度同じ状態にいるかを表す値である。しかし、図4のように、分割方法は同じだが、クラスタの「名前」が異なる場合、 $s(\sigma_1, \sigma_2) = 0$ となる。

⁶ 「擬並列化」と述べた理由は、非対角項を導入した結果現れる $f(\beta, \Gamma)$ という相互作用があるためである。

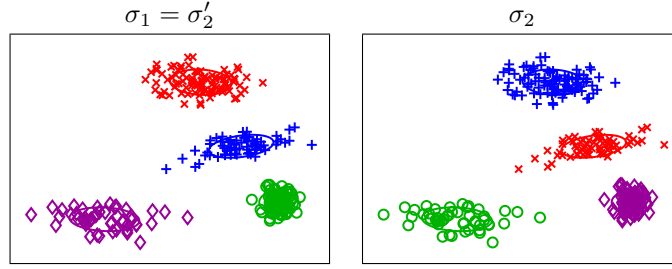


図 4: クラスタ分割の方法は同じだが、クラスタの「名前」が異なる場合

もちろん、量子項の強さ Γ を十分ゆっくり小さくすれば、 $\Gamma = 0$ に到達した際に全て同じクラスタの「名前」になるべきものである。そのため $\sigma_j = \sigma_1, \sigma_{j+1} = \sigma_2$ のような状況が起きてしまうと経路積分表示をしたメリットが十分に活かされなくなってしまう。類似の問題として、 ± 1 の 2 値のみをとる強磁性イジングモデルのドメインウォールの問題がある。磁場無し強磁性イジングモデルでは、全て上向きである状態か、または全て下向きである状態が基底状態となっている。イジングモデルの半分を上向き、もう片方の半分を下向きとしたときに、上向きスピンと下向きスピンの間にはドメインウォールが形成されてしまい、工夫の無いモンテカルロシミュレーションをしてしまうと、安定状態になかなか到達できない。クラスタ分析の場合は、同じクラスタ分割を表す方法が $k!$ 通り存在するので、2 値のみをとるイジングモデルに比べてドメインウォールの問題は更に困難を極める。このドメインウォールの問題を回避するために $s(\sigma_j, \sigma_{j+1})$ の定義を変更した purity という量を導入して近似を行う。purity の定義は

$$\tilde{s}(\sigma_j, \sigma_{j+1}) \equiv \frac{1}{n} \sum_{c=1}^k \max_{c'=1, \dots, k} [Y(\sigma_j) Y^T(\sigma_{j+1})]_{c,c'} \quad (39)$$

となる。ただし、 $Y(\sigma)$ は

$$Y(\sigma) = (\tilde{\sigma}_{j,1}, \tilde{\sigma}_{j,2}, \dots, \tilde{\sigma}_{j,n}) \quad (40)$$

なる、 $n \times k$ 行列を表す。式 (29) により定義されている $s(\sigma_j, \sigma_{j+1})$ は、

$$s(\sigma_j, \sigma_{j+1}) = \frac{1}{n} \text{Tr} [Y(\sigma_j) Y(\sigma_{j+1})^T] \quad (41)$$

と書けることを付記しておく。

purity の性質について簡単な例を通して見てみよう。クラスタ数 $k = 3$ 、要素数 $n = 7$ とする。クラスタの名前を A, B, C とする。まずはじめに、 σ_j 及び σ_{j+1} が表 3 の状態にあるとしよう。

	要素 1	要素 2	要素 3	要素 4	要素 5	要素 6	要素 7
σ_j	A	A	A	B	B	C	C
σ_{j+1}	C	C	B	C	A	A	B

表 3: σ_j と σ_{j+1} の状態

このとき、 $Y(\sigma_j)$ 及び $Y(\sigma_{j+1})$ は、

$$Y(\sigma_j) = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \quad (42)$$

$$Y(\sigma_{j+1}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (43)$$

となる。このとき、式 (39) に従って purity を計算する。

$$Y(\sigma_j)Y(\sigma_{j+1})^T = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad (44)$$

より、purity は $4/7$ となる。一方、 $s(\sigma_j, \sigma_{j+1}) = 0$ である。 σ_j のラベルは表 3 で固定したまま、 σ_{j+1} のラベルを付け替えた場合、 $s(\sigma_j, \sigma_{j+1})$ は、表 4 のようになる。このとき、purity の値と σ_{j+1} のクラスタの名前を付け替えた時の、 $s(\sigma_j, \sigma_{j+1})$ の最大値は一致することが分かった。

	要素 1	要素 2	要素 3	要素 4	要素 5	要素 6	要素 7	$s(\sigma_j, \sigma_{j+1})$	purity
σ_{j+1}	C	C	B	C	A	A	B	0	$4/7$
σ_{j+1}	C	C	A	C	B	B	A	$2/7$	$4/7$
σ_{j+1}	A	A	C	A	B	B	C	$4/7$	$4/7$
σ_{j+1}	A	A	B	A	C	C	B	$3/7$	$4/7$
σ_{j+1}	B	B	C	B	A	A	C	$2/7$	$4/7$
σ_{j+1}	B	B	A	B	C	C	A	$3/7$	$4/7$

表 4: σ_j を固定したもと、 σ_{j+1} の名前を付け替えた時の $s(\sigma_j, \sigma_{j+1})$ の振る舞い

別の例として、 σ_j 及び σ_{j+1} が表 5 の状態にあるとしよう。

	要素 1	要素 2	要素 3	要素 4	要素 5	要素 6	要素 7
σ_j	A	B	B	B	B	C	C
σ_{j+1}	B	A	A	A	C	A	A

表 5: σ_j と σ_{j+1} の状態

このとき、 $Y(\sigma_j)$ 及び $Y(\sigma_{j+1})$ は、

$$Y(\sigma_j) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \quad (45)$$

$$Y(\sigma_{j+1}) = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (46)$$

このとき、

$$Y(\sigma_j)Y(\sigma_{j+1})^T = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix} \quad (47)$$

より、purity は $6/7$ となる。一方、 $s(\sigma_j, \sigma_{j+1}) = 0$ である。先ほどの例と同様、 σ_j のラベルは表 5 で固定したまま、 σ_{j+1} のラベルを付け替えた場合、 $s(\sigma_j, \sigma_{j+1})$ は、表 6 のようになる。このときは、purity の値と σ_{j+1} のクラスタの名前を付け替えた時の、 $s(\sigma_j, \sigma_{j+1})$ の最大値は一致しないことが分かる。

purity は σ_j と σ_{j+1} がどの程度近い状態にあるかを判定する指標ではあるが、可換ではない⁷。すなわち一般に $\tilde{s}(\sigma_j, \sigma_{j+1}) \neq \tilde{s}(\sigma_{j+1}, \sigma_j)$ である。表 3 の場合は、

$$\tilde{s}(\sigma_j, \sigma_{j+1}) = \tilde{s}(\sigma_{j+1}, \sigma_j) = \frac{4}{7} \quad (48)$$

⁷定義により、 $s(\sigma_j, \sigma_{j+1}) = s(\sigma_{j+1}, \sigma_j)$ となり、これは可換である。

	要素 1	要素 2	要素 3	要素 4	要素 5	要素 6	要素 7	$s(\sigma_j, \sigma_{j+1})$	purity
σ_{j+1}	B	A	A	A	C	A	A	0	6/7
σ_{j+1}	C	A	A	A	B	A	A	1/7	6/7
σ_{j+1}	A	B	B	B	C	B	B	4/7	6/7
σ_{j+1}	C	B	B	B	A	B	B	3/7	6/7
σ_{j+1}	A	C	C	C	B	C	C	4/7	6/7
σ_{j+1}	B	C	C	C	A	C	C	2/7	6/7

表 6: σ_j を固定したもと、 σ_{j+1} の名前を付け替えた時の $s(\sigma_j, \sigma_{j+1})$ の振る舞い

であり可換であるが、表 5 の場合は、

$$\tilde{s}(\sigma_j, \sigma_{j+1}) = \frac{6}{7}, \quad \tilde{s}(\sigma_{j+1}, \sigma_j) = \frac{5}{7} \quad (49)$$

となり等しくない。 $Y(\sigma_j)Y(\sigma_{j+1})^T$ について、各々の c に対し、

$$x_c = \max_{c'=1, \dots, k} \left[Y(\sigma_j)Y(\sigma_{j+1})^T \right]_{c,c'} \quad (50)$$

とし、また各々の c' に対し

$$y_{c'} = \max_{c=1, \dots, k} \left[Y(\sigma_j)Y(\sigma_{j+1})^T \right]_{c,c'} \quad (51)$$

とする。表 3 の例のように、任意の c に対し、 $x_c = y_{c'}$ なる c' が存在するとき、purity は可換となる。またその逆も成立する。なぜなら、

$$\tilde{s}(\sigma_{j+1}, \sigma_j) = \frac{1}{n} \sum_{c=1}^k \max_{c'=1, \dots, k} \left[Y(\sigma_{j+1})Y(\sigma_j)^T \right]_{c,c'} \quad (52)$$

であるからである。一方、表 5 の場合は $Y(\sigma_j)Y(\sigma_{j+1})^T$ が式 (47) となり、 $x_c = y_{c'}$ なる c' が存在しない c が存在する。そのため、purity は可換ではない。

以上見てきたことから、purity の性質をまとめると、

1. purity は一般に非可換である。
2. purity に関する不等式

$$0 \leq s(\sigma_j, \sigma_{j+1}) = \text{Tr} \left(Y(\sigma_j)Y(\sigma_{j+1})^T \right) \leq \tilde{s}(\sigma_j, \sigma_{j+1}) \leq 1 \quad (53)$$

が成立する。

5.2 purity を用いた高速化

鈴木トロッタ分解を用いた経路積分表示による量子モンテカルロ法では、式 (38) で状態を更新していけば良い。しかし $s(\sigma_j, \sigma_{j+1})$ の部分がドメインウォールの問題を引き起こすため、それを回避するための手段として前節で purity を導入した。具体的に purity を用いた高速化は以下のように行う。式 (38) の $s(\sigma_{j-1}, \sigma_j) + s(\sigma_j, \sigma_{j+1})$ の部分を $\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_{j-1}, \sigma_j) + \tilde{s}(\sigma_j, \sigma_{j+1})$ とする。すなわち $p_{\text{QA-ST}}^{\text{update}}(\tilde{\sigma}_{j,i} | \{\sigma_j\}_{j=1}^m \setminus \tilde{\sigma}_{j,i}; \beta, \Gamma)$ の代わりに

$$p_{\text{QA-ST+purity}}^{\text{update}}(\tilde{\sigma}_{j,i} | \{\sigma_j\}_{j=1}^m \setminus \tilde{\sigma}_{j,i}; \beta, \Gamma) = \frac{\exp \left[-\frac{\beta}{m} E(\sigma_j) + \tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) f(\beta, \Gamma) \right]}{\sum_{\tilde{\sigma}_{j,i}} \exp \left[-\frac{\beta}{m} E(\sigma_j) + \tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) f(\beta, \Gamma) \right]} \quad (54)$$

Algorithm 1 Quantum Annealing for Clustering

```
1: Initialize inverse temperature  $\beta$  and quantum annealing parameter  $\Gamma$ .
2: repeat
3:   for  $j = 1, \dots, m$  do
4:     for  $i = 1, \dots, n$  do
5:       Draw the new assignment of the  $i$ -th data point,  $\sigma_{j,i}$ , with a probability given in Eq. (54).
6:     end for
7:   end for
8:   Increase inverse temperature  $\beta$ , and decrease QA parameter  $\Gamma$ .
9: until State  $\sigma$  converges
```

となる。ここで、 $\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_{j-1}, \sigma_j) + \tilde{s}(\sigma_j, \sigma_{j+1})$ としたが、purity は非可換な量なので、実際は以下の4通りの場合が考えられる。

$$\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_{j-1}, \sigma_j) + \tilde{s}(\sigma_j, \sigma_{j+1}) \quad (55)$$

$$\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_{j-1}, \sigma_j) + \tilde{s}(\sigma_{j+1}, \sigma_j) \quad (56)$$

$$\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_j, \sigma_{j-1}) + \tilde{s}(\sigma_j, \sigma_{j+1}) \quad (57)$$

$$\tilde{s}(\sigma_{j-1}, \sigma_j, \sigma_{j+1}) \equiv \tilde{s}(\sigma_j, \sigma_{j-1}) + \tilde{s}(\sigma_{j+1}, \sigma_j) \quad (58)$$

これらのうち我々は式 (55) を採用した。それは以下の理由による。 $[Y(\sigma_j)Y(\sigma_{j+1})^T]$ の (c, c') 要素は σ_j でクラスタ c に属し、かつ σ_{j+1} でクラスタ c' に属する要素の個数である。purity の定義は式 (39) より、 $s(\sigma_j, \sigma_{j+1})$ は、 σ_j を基準として σ_{j+1} の近さを定量化していることがわかる。また式 (54) の分母は $\tilde{\sigma}_{j,i}$ についての和なので、 σ_j を基準としないと良いサンプリングになっていない。そのため、 σ_j について状態を更新する際には、 σ_j を基準とする式 (55) を用いれば良いことが分かる⁸。

5.3 熱ゆらぎと量子ゆらぎの同時制御

式 (37) より分かるように、我々が用いたアルゴリズムは β/m による熱ゆらぎの効果と $f(\beta, \Gamma)$ による量子ゆらぎの効果が両方働いている。そのため、熱ゆらぎと量子ゆらぎを巧みに制御することにより、より良い解を得られることが期待できる。シミュレーテッドアニーリングと従来型の量子アニーリング法の良いところを相補的に利用しようという試みである。実際のアルゴリズムは Algorithm 1 のようになっている。ここで重要になるのが、Algorithm 1 の 8 行目のプロセスである。ここで温度と量子項の同時制御を行っているのである。

温度と量子ゆらぎの同時制御をどのように行うかを考えるために、まず極端な場合について考える。 $\frac{\beta}{m} \gg f(\beta, \Gamma)$ のとき、それぞれのレプリカがほぼ独立に $\{\sigma_j\}_{j=1}^m$ は式 (22) の分布に従う。また、 $\frac{\beta}{m} \ll f(\beta, \Gamma)$ のとき、 $\{\sigma_j\}_{j=1}^m$ はエネルギー $E(\sigma_j)$ によらず、全てのレプリカについて同じ状態になるうとする。

小さい系について、温度と量子ゆらぎを同時制御したテスト計算を行った。その結果、途中で $\{\sigma_j\}_{j=1}^m$ が準安定状態に到達しているときにより良い解を得ていることが分かった。

以上から、

1. $\frac{\beta}{m}$ は $f(\beta, \Gamma)$ より十分大きくなるようにとり、準安定状態を検出する。
2. $f(\beta, \Gamma)$ が $\frac{\beta}{m}$ を追い抜くようにする。

⁸実際、簡単な系で式 (55) から式 (58) を用いて、同じスケジュールで計算をしたところ、式 (55) が最も良い解を得ることが分かった。

という2段階のプロセスを経るようなスケジュールを考えれば良いことがわかる。図5に概念図を示す。図中の f^* がより良い解を出すためのスケジュールである。図5の f_1 ははじめのうちから β より強い場合に対応し、量子ゆらぎによる状態の混合をほとんどしていないことに相当する。一方、図5の f_2 は量子ゆらぎが強いいため、実質的に m 枚のスライスがほぼ独立に振る舞うことになり、これはシミュレーテッドアニーリング法と本質的に変わらないことになっている。

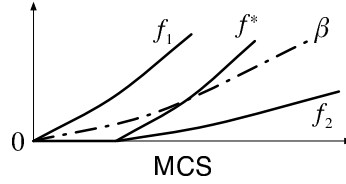


図5: 熱・量子揺らぎの同時制御。良い解を出すスケジュールは f^* である。

我々は温度と量子項のスケジュール関数として、

$$\beta(t) = \beta_0 r_\beta^t \quad (59)$$

$$\Gamma(t) = \infty \quad (t < \tau) \quad (60)$$

$$\Gamma(t) = \Gamma_0 \exp(-r_\Gamma^{t-\tau}) \quad (t \geq \tau) \quad (61)$$

を用いた。ただしここで τ は温度 $\beta(t) = m$ となる t である。 $\frac{k\beta\Gamma}{m} \ll 1$ のとき、トロツタ方向の相互作用 f は

$$f(\beta, \Gamma) \sim -n \log\left(\frac{\beta\Gamma}{m}\right) = nr_\Gamma^t - n \log\left(\frac{\beta\Gamma_0}{m}\right) \quad (62)$$

となることから、 f^* のようなスケジュールを構成するには、十分大きい Γ_0 で、かつ $r_\beta < r_\Gamma$ であれば良いことがわかる。簡単のため、我々は $\beta = m$ になるまで $f(\beta, \Gamma) = 0$ とした ($\Gamma(t) = \infty$)。すなわち、 $\beta = m$ になるまでは独立なシミュレーテッドアニーリングを m 個並列に実行し、その後量子項を弱めることにより ($f(\beta, \Gamma)$ の値を大きくして) 解を得ることを試みた。

6 数値実験結果

我々は以下の3つの問題についてこれまで紹介してきた量子アニーリング法及び、比較実験としてシミュレーテッドアニーリングを用いて数値実験を行った。温度、量子同時制御スケジュールとしては、レプリカ数 $m = 50$ 、初期逆温度 $\beta_0 = 0.2m$ 、初期量子項 $\Gamma_0 = e^{1/2}$ 、温度変化率 $r_\beta = 1.05$ を用いた。また比較実験として行ったシミュレーテッドアニーリングでは、初期逆温度 $\beta_0 = 0.2$ 、温度変化率 $r_\beta = 1.05$ とし、55回の独立な計算結果の平均を取った。

- MNIST データ [39] を用いた、混合正規分布 (Mixture of Gaussian: MoG) の推定 (クラスタ数 $k = 30$)
- Reuters データ [40] を用いた、Latent Dirichlet Allocation(LDA) の推定 (クラスタ数 $k = 20$)
- NIPS コーパス [41] を用いた、Latent Dirichlet Allocation(LDA) の推定 (クラスタ数 $k = 20$)

全ての問題について量子項変化率 $r_\Gamma = 1.05$ を用いた。ただし3番目の問題については、 $r_\beta = 1.02$ を用いた比較実験も行った。

図6に数値実験結果を示す。上から1番目(MNIST データ)、2番目(Reuters データ)、3番目(NIPS データ)の図は $\beta_0 = 0.2$ 、 $\Gamma_0 = e^{1/2}$ 、 $r_\beta = 1.05$ を用い、1番下(NIPS データ)の図は $\beta_0 = 0.2$ 、 $\Gamma_0 = e^{1/2}$ 、

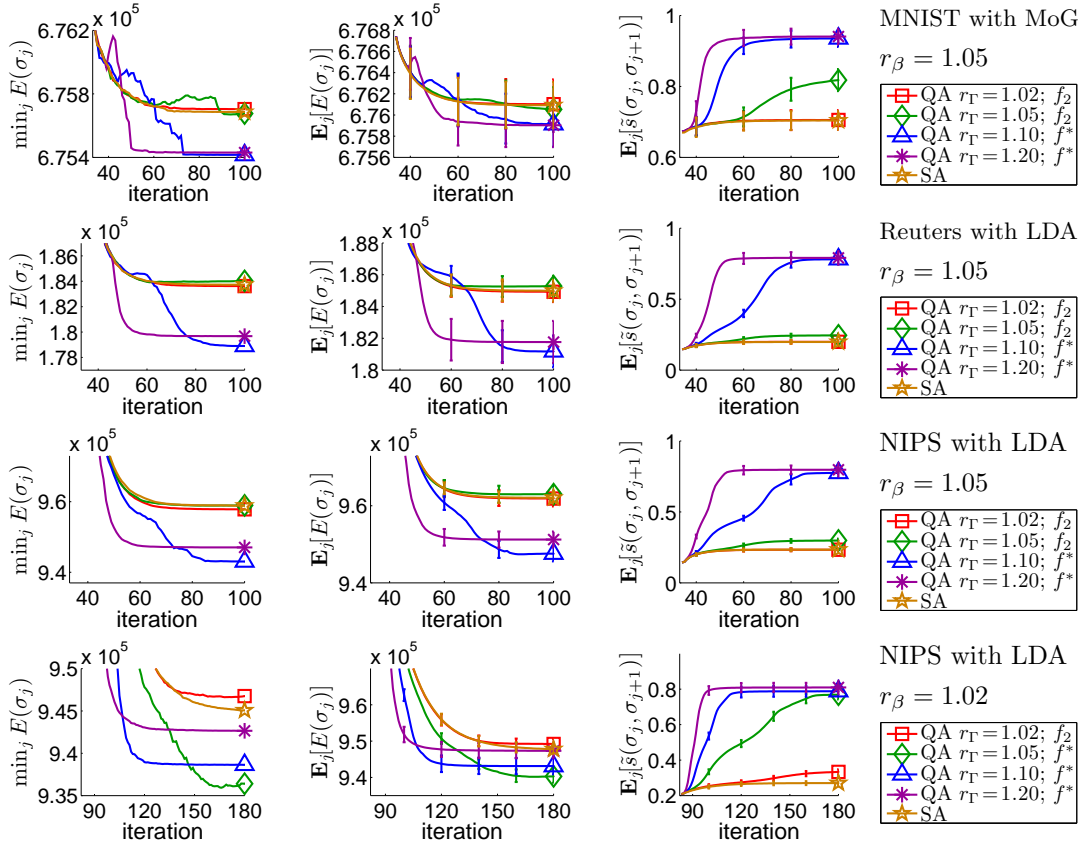


図 6: アニーリングスケジュールを変化させて、シミュレーテッドアニーリングと量子アニーリング法を比較した。1 番上のグラフは MNIST データ、上から 2 番目のグラフは Reuters データ、上から 3 番目及び 4 番目のグラフは NIPS データを用いた。1 番左の図は m 個のレプリカのうち最小エネルギー、左から 2 番目のグラフは m 個のレプリカの平均エネルギー、左から 3 番目は隣接するレプリカ間の purity の平均値。 r_T の値と f_2, f^* との対応は右に記した。いずれの場合も f^* では SA よりも低いエネルギー状態を実現できていることが分かる。 $\beta = 1$ (シミュレーテッドアニーリング) または $\beta = m$ (量子アニーリング) となつてからの値をプロットしてある。

$r_\beta = 1.02$ を用いた結果である。それぞれの問題に対し、左の図は式 (18) の m 個のレプリカのうち、最小値を表している。左から 2 番目は式 (18) の m 個のレプリカの平均値を表している。左から 3 番目は、式 (39) で定義された purity の平均値を表している。いずれのデータも $\beta = 1$ (シミュレーテッドアニーリング) または $\beta = m$ (量子アニーリング) となつてからの値のみをプロットしてある。

左の図から、量子アニーリング法を用いた場合、 f^* に対応する熱・量子同時制御を行うと、シミュレーテッドアニーリングよりもよい解が得られていることが分かる。

7 結論

本原稿で紹介した研究では、

1. クラスタ分析に適した量子アニーリング法の開発
2. 熱・量子最適同時制御の概形の解明

の 2 点を行った。クラスタ分析に量子アニーリング法を適用すると、クラスタの「名前」の問題(ドメインウォールの問題)が起こる。我々は purity と呼ばれる量を導入することにより、クラスタの「名前」の問

題を回避することに成功した。また、初期段階では量子項を強くしておくことにより、各レプリカでほぼ独立に（擬並列化）準安定状態を検出し、その後量子ゆらぎの項を弱くすることにより、より良い解を得ることに成功した。

謝辞

本研究を進めるにあたり、佐藤一誠氏（東京大学）、鈴木大慈氏（東京大学）、ならびに機械学習研究グループ T-PRIMAL の方々には多くの有益な議論をしていただきました。筆者の一人（田中）は東京大学物性研究所スーパーコンピュータセンターを利用させていただきました。また筆者の一人（栗原）は佐藤泰介研究室の計算機を利用させていただきました。この場を借りて感謝いたします。

参考文献

- [1] K. Kurihara, S. Tanaka, and S. Miyashita: Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, 2009. (同じものが arXiv:0905.3527 に掲載されている)
- [2] M. Mezard, G. Parisi, and M.A. Virasoro: Spin Glass Theory and Beyond, World Scientific, 1987.
- [3] K.H. Fisher and J.A. Hertz: Spin Glasses, Cambridge, 1991.
- [4] P. Nordblad and P. Svedlindh: in Spin-glasses and Random Fields, edited by A. P. Young, World Scientific, pp.1-28, 1998.
- [5] J.P. Bouchaud, L.F. Cugliandolo, J. Kurchan, and M. Mezard: in Spin-glasses and Random Fields, edited by A. P. Young, World Scientific, pp.161-224, 1998.
- [6] E. Vincent: in Ageing, Rejuvenation and Memory: The Example of Spin-Glasses, Lecture Notes in Physics, Springer Berlin / Heidelberg, vol. 716, pp.7-60, 2007.
- [7] S. Tanaka and S. Miyashita: J. Phys. Condens. Matter, vol.19, p.145256, 2007.
- [8] S. Tanaka and S. Miyashita: J. Phys. Soc. Jpn., vol.76, p.103001, 2007.
- [9] S. Tanaka and S. Miyashita: arXiv:0711.3261, to appear in J. Phys. Soc. Jpn. vol.78, no.8, 2009.
- [10] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi: Science, vol.220, p.671, 1983.
- [11] S. Kirkpatrick: J. Stat. Phys. vol.34, p.975, 1984.
- [12] K. Hukushima and K. Nemoto: J. Phys. Soc. Jpn., vol.65, p.1604, 1996.
- [13] K. Hukushima: Phys. Rev. E, vol.60 p.3606, 1999.
- [14] B. Apolloni, C. Carvalho, and D. de Falco: Stochastic Processes and their Applications vol.33 p.233, 1989.
- [15] A.B. Finnila, M.A. Gomez, C. Sebenik, C. Stenson, and J.D. Doll: Chem. Phys. Lett. vol.219 p.343, 1994.
- [16] T. Kadowaki and H. Nishimori: Phys. Rev. E, vol.58, p.5355, 1998.
- [17] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda: Science vol.292 p.472, 2001.

- [18] G.E. Santoro, R. Martoňák, E. Tosatti, and R. Car: Science vol.295, p.2427, 2002.
- [19] A. Das and B.K. Chakrabarti: Quantum Annealing and Related Optimization Methods, (Lecture Note in Physics), Springer-Verlag, 2005.
- [20] G.E. Santoro and E. Tosatti: J. Phys. A, vol.39, p.R393, 2006.
- [21] A. Das and B.K. Chakrabarti: Rev. Mod. Phys. vol.80 p.1061, 2008.
- [22] M. Suzuki: Prog. Theor. Phys. vol.56 p.1454, 1976.
- [23] H. G. Evertz, G. Lana, and M. Marcu: Phys. Rev. Lett. vol.70 p.875, 1993.
- [24] A. W. Sandvik: Phys. Rev. B vol.59 p.14157, 1999.
- [25] N. Kawashima and K. Harada: J. Phys. Soc. Jpn. vol.73 p.1379, 2004.
- [26] T. Nakamura: Phys. Rev. Lett. vol.101 p.210602, 2008.
- [27] S. Morita, S. Suzuki, and T. Nakamura: Phys. Rev. E vol.79 p.065701, 2009.
- [28] J. Dziarmaga: Phys. Rev. Lett. vol.95 p.245701, 2005.
- [29] S. Suzuki and M. Okada: Quantum Annealing and Related Optimization Method pp.209-240, Eds. A. Das and B. K. Chakrabarti, (Springer, Heidelberg).
- [30] S. Tanaka and S. Miyashita: J. Magn. Magn. Mater. vol.310 p.e468, 2007.
- [31] S. Miyashita, S. Tanaka, H. de Raedt, and B. Barbara: J. Phys.: Conf. Ser. vol.143 p.012005, 2009.
- [32] S. R. White and A. E. Feiguin: Phys. Rev. Lett. vol.93 p.076401, 2004.
- [33] S. Suzuki and M. Okada: Interdisciplinary Information Sciences, vol.13, p.49, 2007.
- [34] K. Tanaka and T. Horiguchi: Electronics and Communication in Japan, vol.83 p.2117, 2000.
- [35] I. Sato, K. Kurihara, S. Tanaka, H. Nakagawa, and S. Miyashita: Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, 2009. (同じものが arXiv:0905.3528 に掲載されている)
- [36] I. Bloch, J. Daliard, and W. Zwerger: Rev. Mod. Phys. vol.80 p.885, 2008.
- [37] H. F. Trotter: Proc. Am. Math. Soc. vol.10 p.545, 1959.
- [38] D. M. Blei, A. Y. Ng, and M. I. Jordan: Journal of Machine Learning Research vol.3 p.993, 2003.
- [39] <http://yann.lecun.com/exdb/mnist/>
- [40] <http://www.cs.toronto.edu/~roweis/data.html>
- [41] <http://books.nips.cc>