

先端測定データ解析における超並列モンテカルロ計算

Massively parallel Monte Carlo calculation in advanced measurement data analysis

鳥取大院工¹ KEK 物構研低速陽電子² 東大院総合文化³ 星健夫^{1,2} 岩本晴道¹ 福島孝治³

Tottori Univ¹, KEK-IMSS-SPF², Univ Tokyo³,

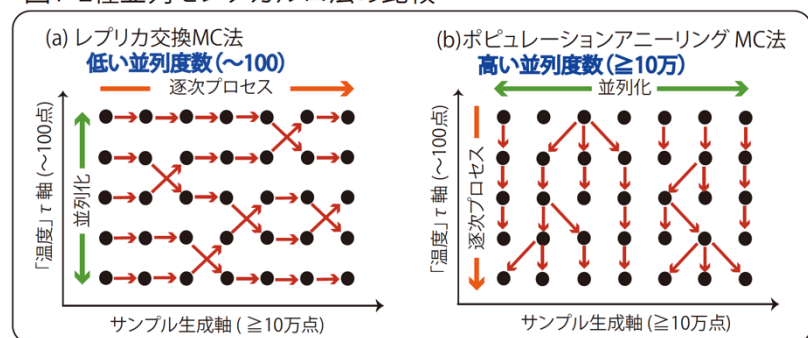
Takeo Hoshi^{1,2}, Harumichi Iwamoto¹, Koji. Hukushima³

E-mail: hoshi@tottori-u.ac.jp

高速高信頼な測定データ解析は計測インフォマティクス の中心であり、我々は高速計算技術(HPC)活用を特徴として、解析手法を開発している。数理基盤は探索型逆問題解析であり、目的状態量 X に対して、測定量 D のモデル(順問題モデル $D_{exp} = D_{cal}(X)$)を想定する。実測量 D_{exp} との残差関数 $R = |D_{cal}(X) - D_{exp}|$ を最小化する数値的最適化問題と考えることができるが、不確か(分布)解析も必須である。現状では、物質表面の原子スケール非破壊測定の新手法である、全反射高速陽電子回折(TRHEPD, トレプト)実験[1]に適用している[2]。Python と Message Passing Interface (MPI) に基づいてソフトウェア「2DMAT」として、2021年3月に公開された[3]。逐次最適化・グリッド型探索・ベイズ最適化・レプリカ交換モンテカルロ(MC)法[5]が実装されている。表面エックス線回折など、対象実験追加にも取り組んでいる。

さらなる手法拡張として、超並列アルゴリズムであるポピュレーションアニーリング MC 法を導入しており、本講演で報告する。MC 法によるベイズ推定は、高信頼データ解析の数理基盤である。目的状態量空間 X におけるサンプリングに相当し、サンプルデータ $\{X_1, X_1, \dots, X_N\}$ のヒストグラムがベイズ事後確率分布を与えるため、複数(局所)解の検出、不確かさ(分布)計算が可能となる。伝統的アルゴリズムは、マルコフ過程に基づく多数の逐次ステップが必要である。並列計算手法としてはレプリカ交換モンテカルロ法[5](図 1(a))が広く使われているが、「温度パラメータ」 τ (許容不確かさの尺度)軸方向の並列アルゴリズムであるため、並列度数が $P=100$ 程度に限られ、超並列計算機(現行・次世代スーパーコンピュータ)の性能をフルに引き出すことが難しい。一方、PAMC 法(図 1(b))は、サンプルデータ生成軸方向の並列計算に相当し、 $P=10^5-10^6$ もの超並列計算が可能であり、超並列計算機に理想的と言える。本講演では、富岳などの超並列スーパーコンピュータでの実例をあげて、その有用性を示す。

図1 2種並列モンテカルロ法の比較



[1] <https://www2.kek.jp/imss/spf/>; 総説として, Fukaya, *et al.*, J. Phys. D 52, 013002 (2018). [2] Tanaka, *et al.*, Acta. Phys. Pol. A 137, 188 (2020); JJAP Conf. Ser., in press; (arxiv:1910.05743); Hoshi *et al.*, <https://arxiv.org/abs/2103.04875> [3] <https://www.pasums.issp.u-tokyo.ac.jp/2DMAT/> (本山裕一・吉見一慶(東大 ISSP)らとの共同開発による) [4] Hukushima and Nemoto, J. Phys. Soc. Jpn. 65, 1604 (1996). [5] Hukushima and Iba, AIP Conf. Proc. 690, 200 (2003).