

2024年4月5日

報道機関 各位

国立大学法人東北大学

製造容易性に優れた確率論的コンピュータを開発 — 半導体とスピントロニクスを組み合わせて超省エネAI計算 —

【発表のポイント】

- 人工知能（AI）で多く用いられる確率的なアルゴリズムを効率的に実行できるコンピュータの開発が喫緊の課題となっています。
- この用途に適し、製造が容易な確率論的^{（注1）}コンピュータ^{（注2）}を半導体回路と少数スピントロニクス^{（注3）}素子の組み合わせで実現しました。
- 将来的には現行の半導体技術と比べて面積を約4桁小さくでき、エネルギー消費を3桁減らせることを確認しました。

【概要】

昨今の人工知能や機械学習では、確率的なアルゴリズムが様々な形で利用されています。一方でその演算自体は物理的には決定論的^{（注1）}に動作する半導体回路で実行されており、ソフトとハードに不整合があります。人工知能や機械学習による社会変革が進展する中、この不整合を解消して更なる高度化、消費エネルギーの大幅な低減を実現することが期待されています。

今回、東北大学電気通信研究所の小林奎斗（けいと）大学院生（研究当時）と深見俊輔教授らは、米国カリフォルニア大学サンタバーバラ校の Kerem Camsari 博士らと共同で、確率的なアルゴリズムの実行に適し、かつ製造容易性にも優れた近未来版の確率論的コンピュータを開発し、その動作を検証しました。自然の熱で確率的に揺らぐスピン素子が生成する物理乱数^{（注4）}で疑似乱数^{（注4）}生成半導体回路を駆動することで、優れた計算性能が得られることを確認しました。加えて確率動作スピン素子を主体とする最終形態の確率論的コンピュータでは、現行の半導体回路で確率的な計算を行う場合と比べて4桁程度の小面積化と3桁程度の省エネ化がもたらされることを明らかにしました。

今回開発された近未来版の半導体・スピン融合確率論的コンピュータは、現行の人工知能、機械学習向け半導体回路の課題を克服するものです。今後研究開発が進展し、計算性能と省エネ性に優れたスピントロニクス確率論的コンピュータの先陣に立って社会実装されていくことが期待されます。

本研究成果は、2024年3月27日（英国時間）に科学誌 Nature Communications に掲載されました。

【詳細な説明】

研究の背景

複雑なプロセスが生み出す結果の予測や、結果の元となった原因の推定、組合せ最適化、特徴の抽出、自然現象のシミュレーションなど、複雑性の高い計算タスクを処理するためのソフトウェア技術として確率的なアルゴリズムが有用であることが知られており、昨今の人工知能や機械学習でも様々な場面で利用されています。一方で確率的なアルゴリズムが実行されているハードウェアに目を向けると、そこでは決定論的に動作する半導体回路が用いられ、回路が生成する疑似乱数を用いて演算が行われています。決定論的な回路で予測不能な乱数を生成するのは容易ではなく、結果として大きな回路面積と消費エネルギーが必要となります。ここで、確率的なアルゴリズムと整合性の高いハードウェアが実現されれば、人工知能や機械学習の更なる性能向上を可能とし、また消費されるエネルギーも大幅に低減できると期待されます。

こうした中、熱で確率的に抵抗が変化するスピントロニクス素子を用いた確率論的コンピュータが確率的なアルゴリズムを効率的に実行するハードウェアとして期待されています。今回研究を行った東北大学と米国カリフォルニア大学サンタバーバラ校、及び米国パデュー大学の共同研究チームは、これまでにスピントロニクス確率論的コンピュータのデモシステムを構築して組合せ最適化、機械学習、量子多体計算、ベイズ推定などの原理実証を行うとともに、並行して東北大学のチームは高性能化に不可欠な確率動作スピントロニクス素子の性能向上に向けた物理的な理解の促進と素子技術開発を行ってきました。これらを通して、その有用性が次第に明らかになる中で、比較的少ない開発コストで製造が可能で、現行の半導体回路の計算性能向上を可能とし、かつエネルギー効率を高めるハードウェア技術の開発・実証が待ち望まれていました。

今回の取り組み

今回、東北大学の小林奎斗大学院生（研究当時）、金井駿准教授、大野英男教授、深見俊輔教授らと、米国カリフォルニア大学サンタバーバラ校の Kerem Camsari 博士らからなる共同研究チームは、比較的容易に製造が可能と見込まれ、また計算性能に優れ、加えて小面積性（≒大規模化容易性）とエネルギー効率に優れた、半導体回路とスピントロニクス素子が融合した近未来版確率論的コンピュータを開発し、その動作を実証しました。加えて、スピン素子を主体とした確率論的コンピュータの最終形態で実現される性能を、プロセス設計キット^(注5)を用いて評価しました。

図 1 に現行の半導体コンピュータ、今回開発・実証した近未来版の半導体・スピン融合確率論的コンピュータ（中段）、及び最終形態のスピン素子を主体とした確率論的コンピュータ（下段）の構造と特徴を示します。現行の半導体コンピュータで確率的アルゴリズムを実行する場合には疑似乱数生成器が利用されますが、これは大きな回路面積とエネルギーを要します（図 1 上段）。一

方、最終形態の確率論的コンピュータは多数の確率的に振る舞うスピン素子が必要となり、この実現には今後いくらかの技術開発を要します。しかしこれが実現できた場合、半導体回路を大幅に削減できるため現行技術と比較してトータル回路面積と消費エネルギーは大幅に低減されます（図1下段）。一方で、今回開発・実証した近未来版の確率論的コンピュータは、従来型の疑似乱数生成器を内蔵した半導体回路と少数の確率動作スピン素子から構成されます。十分に成熟した半導体回路製造技術に少数のスピン素子を追加するだけで実現でき、製造が容易であると予測されます（図1中段）。

図2は作製した半導体・スピン融合確率論的コンピュータのプロトタイプ実機の写真です。本研究ではプログラムで回路の書き換えが可能な半導体のFPGA（Field-Programmable Gate Array）と確率動作スピン素子（超常磁性磁気トンネル接合）をケーブルで接続した簡易的な構造が用いられました。当システムは、確率動作スピン素子からなる確率ビット（Pビット）^{（注2）}5つが生成する物理乱数がFPGA内の大量の疑似乱数生成器を駆動して演算が行われるように設計されています。

研究チームはこのシステムを用いて、確率的サンプリングによる推定と深層ボルツマン機械学習の性能評価を行いました。その結果が図3に示されています。今回の実験では、疑似乱数生成器として広く利用されている線形帰還シフトレジスタ（LFSR）^{（注4）}、及びLFSRが出力するビット列に数学的な処理を施してそのランダム性を高めた高級な疑似乱数生成器（Xoshiro）^{（注4）}、とLFSRを確率動作スピン素子で駆動した場合（MTJ+LFSR）を比較しました。その結果、確率動作スピン素子でLFSRを駆動した場合は、高級な疑似乱数生成器（Xoshiro）と同様に計算を重ねるにつれて解の精度、及び学習の効果が高まっていくことが確認され、すなわち高い計算性能を少ないリソースで実現できることが分かりました。

続いて図4にはプロセス設計キットを用いて、LFSR、Xoshiro、及びスピン素子からなる確率ビットを形成するのに必要なトランジスタ数と、1つの乱数を生成するのに必要なエネルギーを比較した結果が示されています。最終形態のスピントロニクス確率ビットは、同等の計算性能が実現されるXoshiro確率ビットと比べてトランジスタ数は0.0003倍、消費エネルギーは0.007倍で済むことが分かりました。

今後の展開

本研究によって、スピントロニクス確率論的コンピュータの最終的な形態でもたらされるメリットと、その前段階として比較的容易に製造できるであろう近未来版の半導体・スピン融合確率論的コンピュータの実現方法が明らかになりました。今後近未来版の活用を通してその有用性を明らかにしつつ、より省エネ性と大規模化容易性に優れた最終形態のスピントロニクス確率論的コンピュータの実現に向けた本格的な開発が進展することが期待されます。

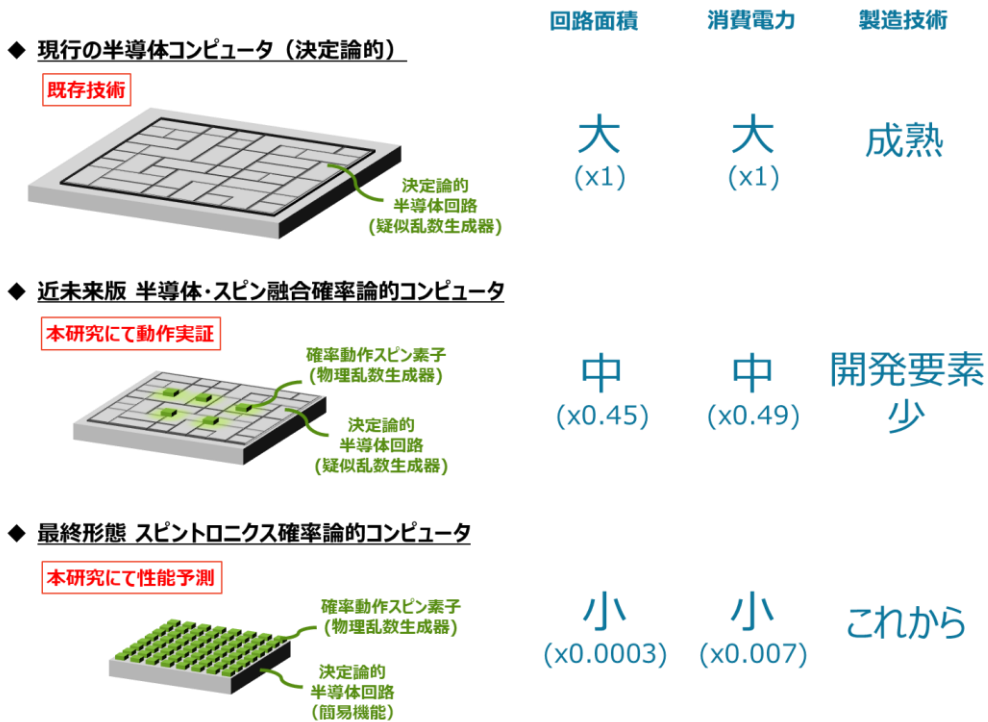


図 1. 本研究の位置付けを整理した模式図。既存技術である決定論的に動作する半導体回路からなるコンピュータ（上段）、本研究にて動作実証した半導体回路と少数の確率動作スピン素子からなる近未来版の確率論的コンピュータ（中段）、及び本研究にて性能を予測した多数の確率動作スピン素子からなる最終形態の確率論的コンピュータ（下段）。右側に回路面積、消費電力、及び製造技術の比較が示されている。

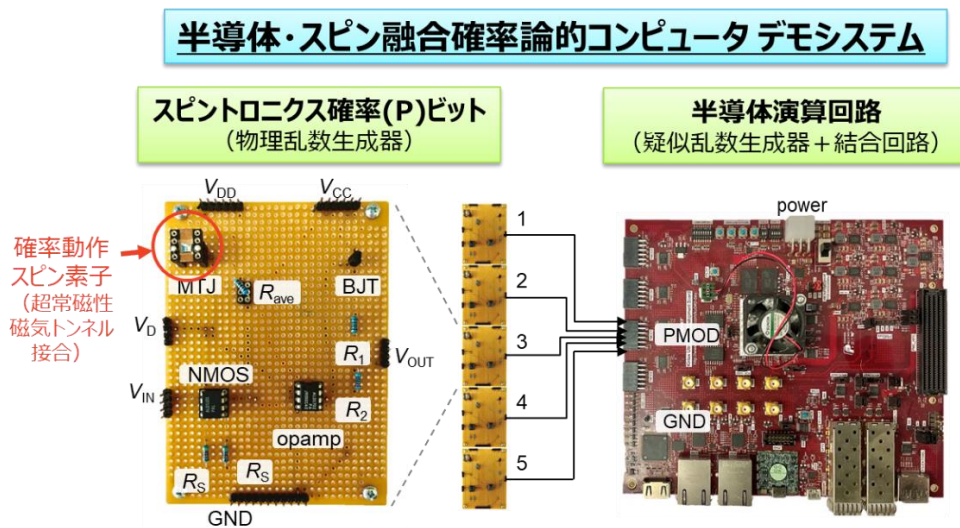
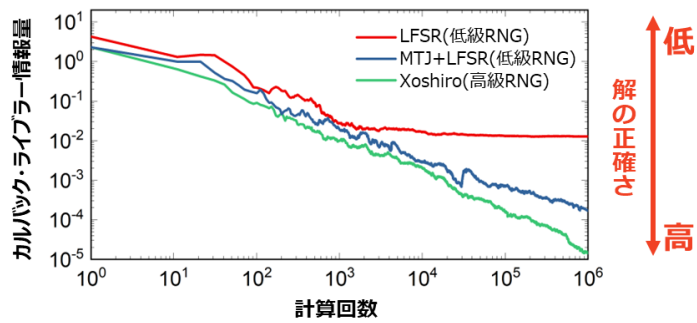


図 2. 本研究で作製した半導体・スピン融合確率論的コンピュータのデモシステムの実機の写真。左側が確率動作スピン素子からなるスピントロニクス確率ビット、右側は疑似乱数生成器と結合回路が実装された FPGA (Field-Programmable Gate Array)。

(a) サンプリングによる推定



(b) ボルツマン機械学習

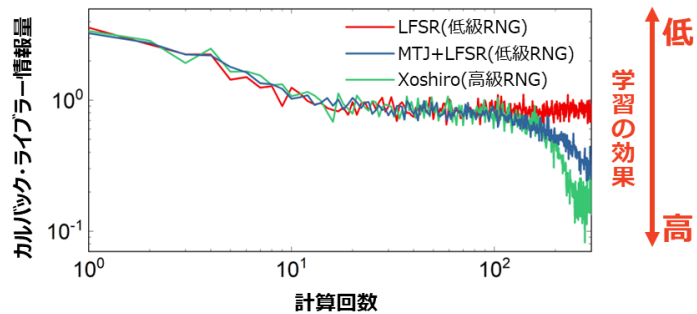
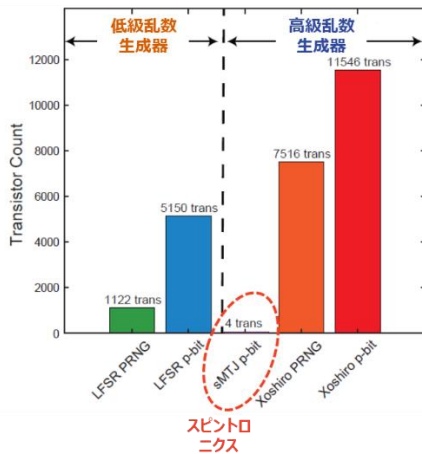


図 3. 作製した半導体・スピン融合確率論的コンピュータ (MTJ+LFSR と表記) でサンプリングによる推定(a)とボルツマン機械学習(b)を行った結果。比較対象として、一般的な乱数生成器 (RNG) として知られている線形帰還シフトレジスタ (LFSR と表記)、及び高級な乱数生成器 (RNG) として知られている Xoshiro で得られた結果も示されている。

(a) トランジスタ数 (≒回路面積)



(b) 消費エネルギー

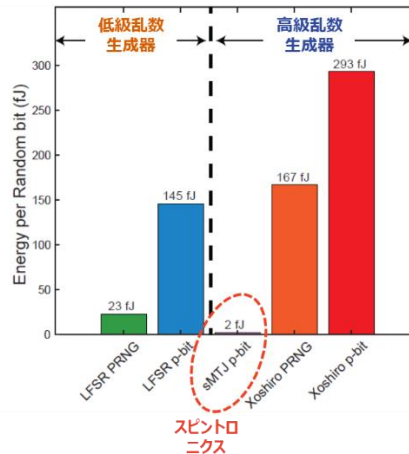


図 4. プロセス設計キットを用いて評価した 1 つの確率ビットの形成に必要なトランジスタ数、及び 1 つの乱数を生成するのに必要なエネルギー。図 3 と同様に、LFSR を用いた確率ビット (LFSR p-bit)、及び Xoshiro を用いた確率ビット (Xoshiro p-bit) の結果が、確率動作スピン素子からなる確率ビット (sMTJ p-bit) とともに示されている。

【謝辞】

本研究は科学技術振興機構（JST）AdCORP（JPMJKB2305）、ASPIRE（JPMJAP2322）、CREST（JPMJCR19K3）、PRESTO（JPMJPR21B2）、文部科学省次世代 X-NICS 半導体創生拠点形成事業（JPJ011438）、村田学術振興財団、丸文財団などの支援の下で行われました。

【用語説明】

注1. 決定論的、確率論的

現在のコンピュータによる演算では、入力情報に対して出力情報が一意に決まる。このような演算の仕組みを「決定論的」演算と言う。それに対して出力を一意に決めず、統計的な手法で決定する演算の仕組みを「確率論的」演算と言う。

注2. 確率論的コンピュータ、確率ビット

確率ビット（P ビット）とは、短時間で 0 と 1 の信号を確率的に出力し、かつ各ビットを電氣的に相関させられる情報処理の基本単位。確率論的コンピュータは確率ビットを用いて演算を行うコンピュータ。

確率ビットは 0 と 1 の重ね合わせ状態を持ち、かつビット間でもつれあい（相関状態）を形成できる量子ビットとは本質的に異なるが一定の類似性があり、確率論的コンピュータは量子コンピュータと並んで新概念コンピュータの一つとして注目されている。1981 年にリチャード・ファインマンが行った講演において、量子コンピュータと並んで、確率的な現象を効率的に計算する仕組みとして紹介されている。

注3. スピントロニクス

物質中の電子が持つ、電氣的な性質（電荷）と磁氣的な性質（スピン）が協調することによって発現する現象を理解し、工学的な応用を目指す学問分野。特に、磁性体のスピンの向き（上・下）を情報（0,1）の担い手として制御する、磁気ランダムアクセスメモリ(MRAM)や磁気センサー等への応用が代表的。

注4. 乱数、物理乱数、疑似乱数、LFSR、Xoshiro

01010101、0000111100001111 のような法則性を持たず、完全にランダムに並んだ数の列を乱数列と言う。物理乱数とは、ランダムな物理現象（例えばサイコロの目など）によって生成される乱数を指し、一方、疑似乱数とは決定論的（確定的）な計算によって出力される乱数を指す。

決定論的に動作する半導体回路を用いて疑似乱数を生成することができる。線形回帰シフトレジスタ（Linear Feedback Shift Register: LFSR）は代表的な疑似乱数生成回路であり、ある初期値を与えるとそれに基づいてある周期内でランダムな乱数列が生成され、またその周期は回路規模で決

まる。Xoshiro は、ある数列に XOR 演算、SHIFT 演算、ROTATION 演算を行うことでそのランダム性を高める回路、あるいは数学的な処理を指す。

多くのプログラミング言語に乱数出力専用の関数（rand()など）が組み込まれているが、その多くはランダム性に問題があることが指摘されている。

注5. プロセス設計キット（PDK）

プロセス設計キット（Process Design Kit: PDK）とは半導体集積回路の設計に必要な情報をまとめたものであり、半導体ファウンドリメーカーから回路設計者に提供される。回路設計者は PDK を用いることで意図した機能の回路を製造するのに必要な面積や消費電力などを正確に評価することができる。

今回の研究では、アリゾナ州立大学が一般に公開している 7-nm FinFET 世代のオープンソース PDK である ASAP7 が用いられた。

【論文情報】

タイトル：“CMOS plus stochastic nanomagnets enabling heterogeneous computers for probabilistic inference and learning”（CMOS と確率的に振る舞うナノ磁性体の融合が可能とする確率的な推測や学習のためのヘテロロジーニアスコンピュータ）

著者： Nihal Sanjay Singh*, Keito Kobayashi*, Qixuan Cao*, Kemal Selcuk, Tianrui Hu, Shaila Niazi, Navid Anjum Aadit, Shun Kanai, Hideo Ohno, Shunsuke Fukami** & Kerem Y. Camsari**

*: 共同筆頭著者

** : 共同責任著者 :

掲載誌： Nature Communications

DOI： 10.1038/s41467-024-46645-6

URL: <https://doi.org/10.1038/s41467-024-46645-6>

【問い合わせ先】

(研究に関すること)

東北大学電気通信研究所

教授 深見 俊輔

TEL: 022-217-5555

Email: s-fukami@tohoku.ac.jp

(兼)東北大学大学院工学研究科電子工学専攻

(兼)東北大学先端スピントロニクス研究開発センター (CSIS)

(兼)東北大学国際集積エレクトロニクス研究開発センター (CIES)

(兼)東北大学材料科学高等研究所 (WPI-AIMR)

(兼)稲盛科学研究機構 (InaRIS)

(報道に関すること)

東北大学電気通信研究所 総務係

TEL: 022-217-5420

Email: riec-somu@grp.tohoku.ac.jp