

Random Magnetic Tunnel Junctions for Probabilistic Computing

Ki Hyuk Han^{1,2}, Yongjin Kim^{2,3}, Hyun Cheol Koo^{2,1}, and OukJae Lee^{2,*}

¹KU-KIST Graduate School of Converging Science and Technology, Korea University, Seoul 02841, South Korea

²Center for Semiconductor Technology, Korea Institute of Science and Technology, Seoul 02792, South Korea

³School of Electrical Engineering, Korea University, Seoul 02841, South Korea

(Received 23 June 2025, Received in final form 1 July 2025, Accepted 2 July 2025)

Despite continued transistor scaling in accordance with Moore's Law, conventional general-purpose computing architectures based on the von Neumann model face significant limitations in addressing energy-intensive tasks such as artificial intelligence (AI). To overcome these challenges, increasing attention is being directed toward specialized computing paradigms optimized for specific problem domains. Among these, probabilistic computing has emerged as a promising alternative that lies between traditional deterministic computing and quantum computing. This paper presents a detailed overview of the theoretical foundation of probabilistic computing, centered on the Ising model, and describes its hardware implementation through a representative device: the stochastic magnetic tunnel junction (s-MTJ). Furthermore, the potential application areas of this technology are discussed, along with key research directions and technical challenges that must be addressed in future developments.

Keywords : Probabilistic computing, Ising model, Stochastic magnetic tunnel junction, Optimization

확률연산 컴퓨팅을 위한 랜덤 자기 터널 접합

한기혁^{1,2} · 김용진^{2,3} · 구현철^{2,1} · 이억재^{2,*}

¹고려대학교 KU-KIST 융합대학원, 서울특별시 성북구 안암로 145, 02841

²한국과학기술연구원 반도체기술연구단, 서울특별시 성북구 화랑로 14길 5, 02792

³고려대학교 전기전자공학부, 서울특별시 성북구 안암로 145, 02841

(2025년 6월 23일 받음, 2025년 7월 1일 최종수정본 받음, 2025년 7월 2일 게재확정)

무어의 법칙(Moore's Law)에 따른 트랜지스터의 지속적인 스케일링에도 불구하고, 기존 폰 노이만(von Neumann) 구조에 기반한 범용 목적 컴퓨팅 방식은 인공지능(AI)과 같이 높은 에너지 효율성을 요구하는 계산 집약적 문제에 대응하는 데 한계에 직면하고 있다. 이러한 제약을 극복하기 위한 방안으로, 최근에는 특정 문제 해결에 최적화된 특수 목적 컴퓨팅(specialized computing) 기법에 대한 연구가 활발히 이루어지고 있다. 그 중에서도 확률 컴퓨팅(probabilistic computing)은 전통적인 결정론적 컴퓨팅과 양자 컴퓨팅(quantum computing) 사이의 중간 영역에 위치한 유망한 대안으로 주목받고 있다. 본 논문에서는 확률론적 컴퓨팅의 핵심 이론적 기반인 이징 모델(Ising model)을 소개하고, 이를 하드웨어적으로 구현하는 대표적 소자인 랜덤 자기 터널 접합소자(stochastic magnetic tunnel junction, s-MTJ)의 구조와 동작 원리를 상세히 설명한다. 아울러, 해당 기술이 적용될 수 있는 주요 응용 분야를 조망하고, 향후 연구에서 해결해야 할 기술적 과제와 발전 방향에 대해 논의한다.

주제어 : 확률 연산 컴퓨팅, 이징 모델, 랜덤 자기 터널 접합, 최적화

I. 서 론

현대의 컴퓨터 시스템은 폰 노이만(von Neumann) 아키텍처를 기반으로 하는 범용 프로세서의 형태로 발전해 왔으며,

무어의 법칙(Moore's Law)에 따라 트랜지스터의 집적도가 지속적으로 증가함에 따라 전반적인 연산 성능, 특히 순차적 처리 능력이 비약적으로 향상되어 왔다. 이러한 구조는 주어진 입력에 대한 단일 출력을 계산하는 정방향 연산(forward computation)에는 매우 효율적이며, 수치 해석이나 정밀하게 설계된 명령어 기반 프로그램 실행에 탁월한 성능을 발휘한다. 그러나 폰 노이만 구조의 본질적 특성은 역방향 연산(inverse

computation) 수행에 근본적인 제약을 내포하고 있다. 예컨대 Fig. 1(a)에 도시된 바와 같이, 특정 입력 벡터($x_1, x_2, \dots, x_n$)에 대해 출력 y 를 결정론적으로 산출하는 것은 가능하지만, 반대로 주어진 출력값 y 를 생성할 수 있는 입력 조합을 효율적으로 역추론하거나 확률적으로 샘플링하는 데에는 구조적 한계가 존재한다. 이와 같은 역연산의 영역에는 NP-난해 최적화 문제, 암호 해독, 파라미터 추정, 이미지 재구성 등 다양한 계산적 난제가 포함된다. 이러한 문제를 해결하기 위한 대안으로 최근에는 확률적 하드웨어 소자를 기반으로 하는 확률론적 컴퓨팅(probabilistic computing)[1-6]과 같은 새로운 비 폰 노이만(non-von Neumann) 컴퓨팅 아키텍처가 주목받고 있다.

확률 연산 컴퓨팅(probabilistic computing)은 무작위성을 내재한 연산 단위인 확률 비트(probabilistic bit, p-bit)를 기반으로 하며, 이들은 외부 입력에 따라 0과 1 사이에서 확률적으로 진동하는 특성을 지닌다. 각 p-bit는 다른 p-bit들과 상호 연결된 네트워크 내에서 작동하며, 개별 p-bit의 상태는 주변 노드들과의 상호작용 및 외부 입력 신호에 따라 동적으로 결정된다. 이러한 네트워크는 전체 시스템의 에너지가 최소화 되도록 볼츠만 열평형 분포(Boltzmann thermal equilibrium law)를 따르도록 설계되며, 관련된 동작 메커니즘은 Fig. 1(b)에 도식화되어 있다. 이러한 접근은 Richard Feynman이 1982년 “Simulating Physics with Computers”라는 논문에서 제안한 이론적 틀에 그 기원을 두고 있다[7]. 그는 자연계의 물리 법칙이 본질적으로 확률적 특성을 지니고 있으므로, 이를 효과적으로 모사하기 위해서는 계산 장치 또한 확률론적 특성을 내포해야 한다고 주장하였다.

확률론적 컴퓨팅(probabilistic computing)은 현재의 반도체 공정 기술로 구현이 가능하며, 양자 컴퓨팅과 유사한 확률 기반 연산 특성을 가지면서도, 상용화 관점에서 보다 현실적인 대안으로 인식되고 있다. 확률 컴퓨팅의 기본 동작 원리는 이징 모델(Ising model)[8,9]에 기반하며, 각 확률 비트(p-bit)는

독립적인 클럭 신호에 따라 비동기적으로 상태를 갱신하는 구조를 갖는다. 개별 p-bit는 인접한 p-bit들과의 상호작용 및 외부 입력에 의해 상태를 확률적으로 업데이트하며, 이 과정은 물리적 무작위성(thermal noise 등)을 활용하여 전체 에너지를 최소화하는 해를 탐색하는 방식으로 작동한다. 이러한 연산 방식은 순차적이며 결정론적인 처리를 기반으로 하는 기존 폰 노이만 컴퓨터와는 본질적으로 상이하다.

본 논문에서는 확률론적 컴퓨팅의 개념적 기반과 함께, 이를 물리적으로 구현하기 위한 핵심 소자인 랜덤 자기터널접합소자(stochastic magnetic tunnel junction, s-MTJ)에 대해 소개한다. 아울러 s-MTJ를 이용한 단위 연산 소자인 p-bit 회로의 구조와 동작 원리를 설명하고, 나아가 확률론적 컴퓨팅의 다양한 응용 가능성과 향후 연구 방향에 대해서도 논의하고자 한다.

II. 확률 연산 컴퓨팅의 이론적 배경: 이징 모델

확률연산 컴퓨터를 포함한 이징 머신(Ising machines)은 주로 조합 최적화(combinatorial optimization) 문제의 효율적인 해결에 초점을 맞춘다. 조합 최적화 문제란, 주어진 제약 조건 하에서 가능한 조합들 중 목적 함수를 최적화하는 해를 탐색하는 것으로, 변수의 수가 증가함에 따라 계산 복잡도가 지수적으로 증가하여, 시스템 규모가 커질수록 기존의 범용 컴퓨터로는 제한된 시간 내에 해를 구하기 어려운 것이 일반적이다. 이러한 계산 복잡도를 효과적으로 완화하기 위한 방법 중 하나로 이징 모델(Ising model)이 제안되었다. 이징 모델은 1925년 빌헬름 렌츠(Wilhelm Lenz)가 처음 제안하고, 그의 제자인 에른스트 이징(Ernst Ising)이 발전시킨 통계역학적 모델로, 원래는 자성체의 상전이(phase transition) 현상을 설명하기 위한 이론적 틀로 고안되었다. 이후 이 모델은 조합 최적화, 양자 어닐링(quantum annealing), 확률 신경망 모델링 등 다양한 분야에 적용되었으며, 최근에는 이를 실제 하

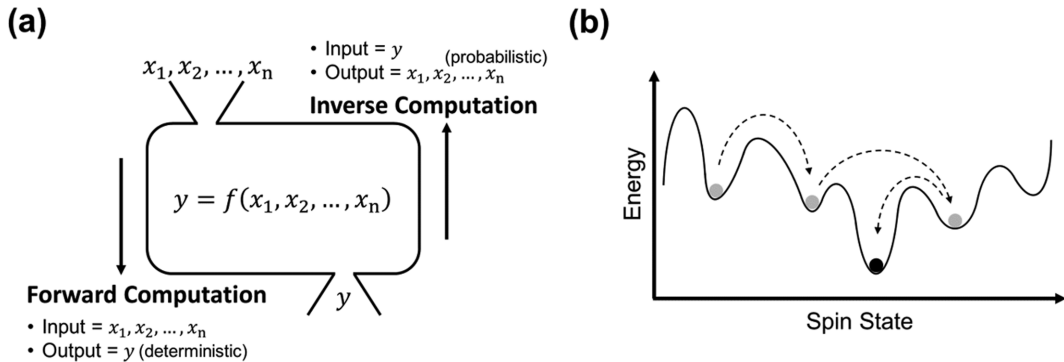


Fig. 1. (a) Conceptual diagram of inverse computation. (b) Schematic diagram of energy minimization, avoiding local minima traps for a given system.

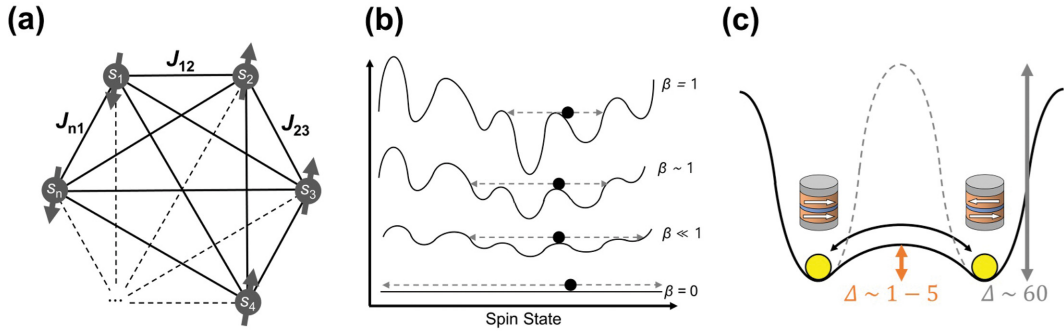


Fig. 2. (Color online) (a) Schematics of the Ising model. (b) Hamiltonian landscape experienced by spins as a function of β . (c) Comparison of energy barriers between a conventional MTJ (gray dashed line) and an s-MTJ fabricated using a lowbarrier magnet (LBM, black solid line).

드웨어에서 구현하는 이징 머신 개발이 활발히 이루어지고 있다. 이징 모델은 스핀(spin) 변수들의 상태와 이들 간의 상호작용을 통계역학적 프레임워크 내에서 기술한다. 구체적으로는, 각 스핀 변수의 조합에 따라 전체 시스템의 에너지가 정의되며, 이는 최저 에너지 상태로의 수렴을 통해 최적 해를 탐색하는 방식으로 활용된다. 이러한 에너지 함수 기반 접근은 Fig. 2(a)에 시작적으로 표현되어 있다.

조합 최적화 문제를 이징 모델에 매핑하면, 문제의 목적 함수 또는 비용 함수(cost function)는 해밀토니언 $\hat{H}$ 로 표현될 수 있다.

$$\hat{H} = - \sum_{(i,j)} J_{ij} s_i s_j - \sum_i h_i s_i \quad (s_i \in \{-1, 1\}, \forall i \in V) \quad (1)$$

여기서 s_i 는 각 개별 스핀의 상태를 나타내며, 이 값은 -1 또는 +1의 이진 값을 갖는다. J_{ij} 는 스핀 s_i 와 s_j 사이의 상호작용 강도를 나타내는 계수이며, h_i 는 외부 자기장에 해당하는 바이어스로서, 스핀 s_i 에 인가되는 국소적인 에너지 영향을 반영한다. 이징 모델의 전체 에너지는 이러한 요소들을 종합하여 구성된 해밀토니언(Hamiltonian)에 의해 정의되며, 시스템은 에너지를 감소시키는 방향으로 스핀 상태를 갱신함으로써 점진적으로 최적화된 상태, 즉 최소 에너지 상태에 수렴할 수 있다. 이 과정은 볼츠만 분포(Boltzmann distribution)를 따르는 확률적 열역학 평형 상태에 기반한다. 스핀 변수의 조합은 시간에 따라 동적으로 변화하며 전체 에너지를 최소화하는 해를 탐색한다. 이러한 원리를 통해 이징 모델은 복잡한 조합 최적화 문제의 해를 물리적으로 효과적으로 모사할 수 있다.

확률 연산 컴퓨팅(probabilistic computing)에서는 이징 모델의 스핀 변수 s_i 를 확률 비트(p-bit)로 대체하여 구현한다. P-bit은 매우 짧은 시간 내에서는 이진 상태(-1 또는 +1)를 가지지만, 좀 더 긴 시간 내에서 -1과 +1 사이를 무작위하게 진동하고, 해당 상태가 선호될 확률은 외부 입력 신호(I_i)

에 의해 조절된다. 따라서, 각 p-bit의 평균적인 스핀 상태, 즉 기대값 m_i 는 다음과 같은 시그모이드(sigmoid) 기반 확률 함수에 따라 결정된다.

$$m_i = \text{sign}[\tanh(\beta I_i) + \text{rand}(-1, +1)] \quad (2)$$

$$I_i = - \partial E(m_1, m_2, \dots, m_n) / \partial m_i = \sum_{j=1}^N J_{ij} m_j + h_i \quad (3)$$

여기서 sign 함수는 입력값이 0보다 작을 경우 -1, 0보다 클 경우 +1을 출력하는 이진 임계 함수(binary threshold function)로, 이는 초기 인공지능 신경망에서 사용된 step 함수와 유사한 개념이다. $\text{rand}(-1, +1)$ 은 -1과 +1 사이의 균등 분포를 따르는 난수 값을 의미하며, I_i 는 외부로부터 i번째 p-bit에 인가되는 아날로그 입력 신호를 나타낸다. 한편, β 는 역 온도(inverse temperature)로 정의되며, 시뮬레이티드 어닐링(simulated annealing) 기법[10,11]에서 온도 제어 인자로 작용한다. 초기에는 작은 값을 유지하며 다양한 상태(state)를 이동하여 시스템이 국소 최소값(local minima)에 갇힐 수 있는 가능성을 줄이고, 점차 β 값을 증가시키는 방식(온도 감소)을 통해 보다 전역적인 최적 상태로 수렴할 수 있도록 한다. 이 과정은 Fig. 2(b)에 도시된 바와 같이, 해밀토니언(Hamiltonian)의 영향력이 점진적으로 커지는 열역학적 동역학 과정으로 이해할 수 있다. 이론적으로는 이상적인 어닐링 스케줄 하에서 시스템이 전역 최소 에너지 상태에 도달함이 보장되지만, 실제 구현에서는 유한 시간의 제약으로 인해 근사 최적 해(near-optimal solution)에 도달할 가능성도 배제할 수 없다.

P-bit의 출력 값 m_i 는 입력 신호 I_i 가 0일 경우, 전적으로 무작위 항 $\text{rand}(-1, +1)$ 에 의해 결정되며, 이때 ± 1 사이의 상태 전이가 동등한 확률(50%)로 발생한다. 반면, $I_i \neq 0$ 인 경우, 출력의 기대값은 $\langle m_i \rangle = \tanh(I_i)$ 형태를 따르며, 이는 상호작용이 없는 독립적인 스핀들로 구성된 열역학 시스템에

서 외부 자극(예: 자기장 등)에 의해 나타나는 통계적 분포 특성과 일치한다. 따라서, p-bit 모델은 이러한 확률적 물리 현상을 효과적으로 반영하여, 열역학적 원리에 기반한 계산을 가능하게 한다.

수식(2)와 (3)은 각 p-bit가 주변 p-bit들과의 상호작용 및 외부 바이어스의 영향을 반영하여 상태를 갱신하도록 정의되어 있으며, 전체 네트워크가 확률적으로 최소 에너지 상태로 수렴하도록 유도하는 역할을 한다. 개별 p-bit의 상태 전이는 해당 p-bit에 인가되는 바이어스 I_i 에 의해 결정되며, 이 바이어스는 네트워크 내 다른 p-bit들의 상태를 기반으로 계산된다. 구체적으로, I_i 는 해당 노드를 제외한 타 p-bit들의 상태에 대해 가중치를 적용한 선형 결합으로 정의되며, 이는 시스템 에너지 함수에 대한 음의 기울기(negative gradient)를 취함으로써 유도된다. 이와 같은 방식으로 구성된 확률 비트 기반 시스템은 이징 모델의 통계역학적 특성을 물리적으로 구현하며, 이를 통해 조합 최적화 문제의 근사 해를 효과적으로 탐색할 수 있는 연산 플랫폼으로 활용될 수 있다.

III. 확률론적 비트의 구현: 랜덤 자기 터널 접합(s-MTJ)

P-bit 구현을 위한 물리적 시스템 후보 중 하나로 스핀 정보를 활용하는 자기 터널 접합(magnetic tunnel junction, MTJ) 소자가 제안되어 왔다[12-14]. MTJ는 두 개의 강자성 층이 얇은 절연층(터널 배리어, tunneling barrier)으로 분리된 샌드위치 구조로 구성된다. 이 중 하나는 자화 방향이 고정된 고정층(fixed layer), 다른 하나는 자화 방향이 외부 자극에 따라 자유롭게 변할 수 있는 자유층(free layer)으로 이루어져 있다. 두 자성층의 자화 방향이 서로 평행일 때는 낮은 저항 상태(R_P), 반평행일 때는 높은 저항 상태(R_{AP})를 나타내며, 이 두 저항 상태는 각각 디지털 신호 '0'과 '1' 또는 스핀 상태 '-1'과 '+1'에 대응된다. MTJ의 저항 상태 전이는 자유층의 자화 방향이 열적 요동에 의해 활성화 장벽(activation energy barrier)을 넘을 때 발생하며, 이 과정은 열 안정성 지수(thermal stability factor) $\Delta = E_B/k_B T$ 에 의해 정량적으로 정의된다. 여기서 E_B 는 두 자화 상태 간의 에너지 장벽, k_B 는 볼츠만 상수, T 는 절대 온도를 나타낸다. 이 지수 Δ 는 주어진 온도 조건에서 자성 정보가 얼마나 안정적으로 유지될 수 있는지를 나타내며, Δ 값이 높을수록 자화 방향이 외란에 강하고 데이터 유지 특성이 우수하다. 해당 개념은 Fig. 2(c)에 도식화되어 있다. 기존 MTJ 소자는 자동차, IOT 기기용 차세대 저전력 비휘발성 메모리 소자인 MRAM(magnetoresistive random-access memory) 기술에 적용되기 위해 주로 연구되어 왔으며, 약 80°C의 환경에서도 10년 이

상 데이터를 유지할 수 있도록 열 안정성 지수 $\Delta > 60$ 을 확보하는 방향으로 설계되고 최적화되어 왔다.

초집적화 시대에 접어들면서, 소자의 물리적 크기가 소형화됨에 따라 충분한 자기 이방성(magnetic anisotropy)을 확보하는 데 어려움이 커지고 있으며, 그 결과 열 안정성 지수 Δ 가 저하되어 데이터 유지 특성의 불안정성이 증가하는 문제가 보고되고 있다[15]. Δ 의 감소는 외부 자극 없이도 열 잡음(thermal noise)에 의해 자화 방향이 무작위로 전이될 가능성이 높으며, 이는 기존 비휘발성 메모리 응용에서는 성능 저하의 주요 원인이 된다. 그러나 이러한 열적 불안정성은 확률적 연산 소자(probabilistic computing device)로서의 관점에서는 오히려 유리하게 작용할 수 있다. 예를 들어, $\Delta \approx 10$ 인 MTJ의 경우 평균 데이터 유지 시간(retention time)은 수 μs 수준으로 짧아지며, 이로 인해 MTJ의 저항 상태가 일정 시간 주기마다 무작위로 진동하게 된다. 이러한 동작은 Random Telegraph Noise(RTN) 형태의 확률 신호를 자연스럽게 생성하게 되며, 이는 p-bit로 동작하기 위한 필수적인 진성 무작위성(true randomness)을 제공한다. MTJ 기반 p-bit 소자에서 이와 같은 retention time은 확률 연산의 주기를 결정짓는 핵심 요소로 작용하며, 각 p-bit의 상태 전이 속도에 직접적인 영향을 미친다. 또한, bit flipping에 소모되는 에너지는 전체 확률 컴퓨팅 시스템의 성능과 효율성 측면에서 매우 중요한 설계 지표로 간주된다. 따라서 열역학적으로 불안정한 MTJ 구조는 비휘발성 메모리보다는 확률적 컴퓨팅에 더 적합한 특성을 지니며, 이를 기반으로 한 p-bit 구현에 있어 중요한 물리 기반 플랫폼으로 주목받고 있다.

확률론적 변동 특성을 보이는 자기 터널 접합은 확률론적 자기 터널 접합소자(stochastic magnetic tunnel junction, s-MTJ)라 불리며, 이러한 s-MTJ의 핵심 설계 요소는 낮은 배리어 자석(low barrier magnet, LBM)을 자유층(free layer)에 도입하는 것이다[16-19]. LBM은 열적 요동(thermal fluctuation)만으로도 자화 상태의 확률적 전이가 자발적으로 발생할 수 있도록 설계되며, 이를 통해 소자는 본질적인 무작위성을 획득하게 된다. MTJ에서 자유층의 자화 방향에 따라 열 안정성 지수 Δ 를 조절하는 공정 방식에는 차이가 존재한다. 수평 자화(in-plane magnetic anisotropy, IMA) 구조의 MTJ[20]에서는 주로 형상 이방성(shape anisotropy)을 이용해 자기 이방성을 구현하며, Δ 를 낮추기 위한 주요 전략으로는 포화 자기화 세기(M_S)가 낮은 연자성 물질(soft magnetic material)의 사용, 자유층 유효 두께(t_{eff})의 감소, 소자 면적의 미세화, 그리고 종횡비(anisotropy field, H_K)가 낮은 구조의 구현이 포함된다. 반면, 수직 자화(perpendicular magnetic anisotropy, PMA) 구조의 MTJ[21,22]에서는 수식(4)에서 보듯이 자유층의(최적화된) 두께 조절 및 패터닝 공정을 통한 면적

최적화를 통해 Δ 를 제어한다. IMA-MTJ와 PMA-MTJ 모두 확률적 연산 응용에 적합한 열 안정성 설계를 위해 다양한 박막 구조 설계 및 공정 기반 조정 기법이 연구되고 있다.

$$\Delta = \frac{E_B}{k_B T} = \frac{(K_{eff} t_{FM}) A}{k_B T} \left(K_{eff} = \frac{K_i}{t_{FM}} + K_V - \frac{1}{2} \mu_0 M_s^2 \right) \quad (4)$$

IMA-MTJ와 PMA-MTJ는 각각 랜덤 자기 터널 접합(s-MTJ)으로 활용될 때 고유한 장단점을 지닌다. IMA-MTJ는 탈자화장(demagnetizing field)의 영향을 받아 자화 벡터가 블로흐 구면(Bloch sphere) 적도 부근에서 제한된 범위 내에서 빠르게 변동하는 특성을 보이는 반면, PMA-MTJ는 자화 벡터가 구면 전체에 걸쳐 무작위로 분포하며 변동한다. 이에 따라, 변동 속도(fluctuation speed) 측면에서는 IMA-MTJ가 PMA-MTJ보다 상대적으로 빠른 경향을 나타낸다[24]. 특히, IMA-MTJ는 열 안정성 지수 $\Delta < 1$ 의 조건 하에서, 즉 p-bit의 에너지 장벽이 $k_B T$ 이하로 낮아지는 경우, 연속적인 확률적 변동(continuous stochastic fluctuations)을 나타내는 것이 특징이다. 이는 확률 연산 소자로서의 적합성을 높이는 요인이 될 수 있다. 그러나 전력 소비(power consumption) 측면에서는 IMA-MTJ가 불리한 측면이 존재한다. 탈자화장이 지속적으로 작용함에 따라 자화 방향을 유지하는 데 필요한 에너지가 증가하게 되며, 결과적으로 PMA-MTJ에 비해 더 높은 소비 전력을 요구하게 된다[24]. 또한, $\Delta \leq 1$ 의 저 안정성 영역에서는 MTJ의 상태 전이에 대한 통계적 특성이 더 이상 Arrhenius 법칙을 따르지 않으며, 대신 확률적 동역학을 기술하는 Fokker-Planck 방정식에 의해 지배된다. 이는 기존의 비휘발성 메모리 설계 기준과는 다른 물리적 모델링이 필요함을 의미하며, s-MTJ 기반 확률 연산 회로의 동작 특성 분석에 중요한 이론적 기반을 제공한다.

s-MTJ에서 확률적 상태 전이는 스핀 전류(spin-transfer torque, STT 또는 spin-orbit torque, SOT) 혹은 외부 자기장의 인가를 통해 제어될 수 있으며, 이를 통해 $\Delta_{P \rightarrow AP}$ 와 $\Delta_{AP \rightarrow P}$ 의 비대칭적인 열 안정성 조절이 가능하다. 이러한 열 안정성 조정 메커니즘은 확률 분포의 형태를 바꾸는 데 핵심적으로 작용하며, 스핀 기반 확률 비트의 설계 유연성을 제공한다. 열 안정성은 인가된 전류 I 와 외부 자기장 H 에 따라 변동하며, 다음과 같은 식으로 유효 열 안정성 Δ_{eff} 으로 표현된다[18]:

$$\Delta_{eff}(R_P \leftrightarrow R_{AP}) = \Delta_0 \left(1 \pm \frac{I}{I_C} \right) \left(1 \pm \frac{H}{H_K} \right)^2 \quad (5)$$

여기서 Δ_0 는 전류와 자기장이 모두 0일 때의 초기 열 안정성, I_C 는 자화 반전을 유도하는 임계 전류(critical current), H_K 는 이방 자기장(anisotropy field)을 나타낸다. 이 관계는

MTJ의 확률적 동작 특성이 외부 인가 조건에 따라 연속적으로 조절될 수 있음을 보여준다. 한편, MTJ의 상태 유지 시간(retention time)은 일반적으로 Arrhenius 법칙을 따르며, 특정 상태(예: R_P)에 머물 확률은 입력 변수인 전류 I 와 자기장 H 에 대한 $\tanh(I, H)$ 함수로 근사할 수 있다. 이는 s-MTJ가 수식(2)에서 정의된 확률적 상태 전이 모델을 물리적으로 구현할 수 있는 하드웨어 플랫폼임을 이론적으로 뒷받침한다.

s-MTJ 소자에서 발생하는 신호의 진성 무작위성(true randomness)은 표준화된 방법인 NIST SP 800-22a 통계 테스트 모음(NIST SP 800-22a statistical test suite)[25,26]를 통해 검증되었다. 이는 난수 생성기의 품질을 평가하기 위해 고안된 것으로, 총 188개의 상이한 통계적 검정을 포함하고 있다. s-MTJ 소자는 상태 유지 시간(retention time) 내에서는 이진 출력 값이 일정하게 유지되기 때문에, 이보다 충분히 긴 시간 간격으로 출력을 샘플링함으로써 새로운 데이터를 생성한다. 그러나 이러한 방식으로 생성된 난수는 통계적으로 완전한 독립성과 균일성을 보장하지 못할 가능성이 있어, 편향된 분포를 갖는 난수가 생성될 수 있다. 이에 따라, 2개, 4개 혹은 8개 난수 값을 XOR 연산을 통해 결합하는 ‘XOR-whitening’ 기법을 적용함으로써 편향을 줄이고, 결과적으로 난수의 품질을 향상시킬 수 있다.

IV. 확률론적 컴퓨팅 회로 구조

P-bit 회로는 앞서 설명한 난수 발생 소자(random number generator)를 기반으로 하여, 여러 개의 트랜지스터 및 주변 소자들과 결합된 형태로 구성된다. 가장 단순한 형태의 회로는 1개의 트랜지스터, 1개의 랜덤 자기 터널 접합 소자(stochastic magnetic tunnel junction, s-MTJ), 1개의 저항, 그리고 1개의 비교기(comparator)로 구성되며, 이는 Fig. 3(a)에 도식화되어 있다. 회로 동작은 V_{DD} 로부터 인가되는 전압을 기반으로 한다. 이때, 각 p-bit에 인가되는 바이어스 I_i 는 해당 p-bit를 제외한 다른 p-bit들의 상태 s_j 와 이들과의 상호작용 계수 J_{ij} 의 가중합으로 정의되며, 이는 수식 $\sum_{(i,j)} J_{ij} s_j$ 로 표현된다. 이 바이어스 I_i 는 게이트 전압 V_{input} 을 조절하며, 이는 결과적으로 MTJ 소자에 흐르는 전류의 크기를 결정하게 된다. 이러한 구조는 각 p-bit의 상태가 주변 p-bit들과의 상호작용에 따라 실시간으로 조정됨을 의미하며, 네트워크의 동적 특성을 반영한다. 게이트 전압의 변화는 스핀 전류(spin current)의 세기에 영향을 주며, 이는 s-MTJ의 자유층(free layer)의 자화(magnetization) 방향에 대한 선호도(preference)를 결정한다.

자화 방향이 변화하면 MTJ의 저항 값이 변하게 되며, 이에 따라 MTJ의 하단 전압도 달라지게 된다. 이 전압은 비교

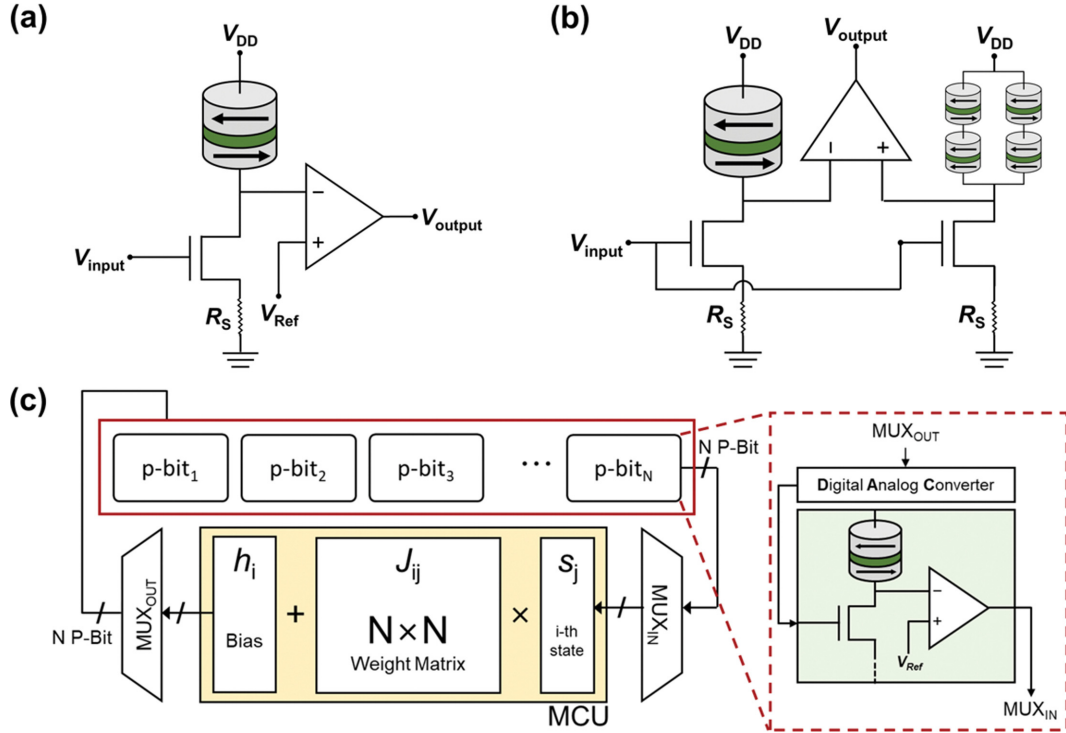


Fig. 3. (Color online) (a) Schematic of a p-bit based on a 1T-1MTJ circuit configuration. (b) Revised p-bit circuit incorporating an adjusted reference voltage (V_{Ref}) to mitigate output errors due to the reduced R_{AP} values. (c) Schematic of the p-computing circuit architecture.

기의 입력으로 전달되며, 기준 전압 V_{Ref} 과 비교된다. 비교기는 MTJ 하단 전압이 V_{Ref} 보다 클 경우(즉, 저항이 낮은 병렬 상태, R_P)에는 출력값 $V_{output} = 1$ 을, 반대로 작을 경우(즉, 저항이 높은 반평행 상태, R_{AP})에는 $V_{output} = 0$ 을 출력한다. 이를 통해 MTJ의 저항 상태가 디지털 확률 비트의 출력으로 변환되며, 전체 회로가 물리 기반의 확률 연산 유닛으로 동작하게 된다.

MTJ 소자는 전류 바이어스 혹은 온도가 증가함에 따라 병렬 상태(R_P)의 저항은 거의 일정하게 하지만, 반평행 상태(R_{AP})의 저항 값은 바이어스 의존성[27]에 의해 감소하는 경향이 있다. 이로 인해 고정된 기준 전압 V_{Ref} 를 사용하는 경우, 반평행 상태에서도 충분한 전압 강하가 발생하지 않아 비교기 출력이 지속적으로 '1'을 유지하는 문제가 발생할 수 있다. 이러한 현상을 보완하기 위한 방안으로, Fig. 3(b)와 같이 V_{Ref} 를 MTJ의 평균 저항 값에 따라 동적으로 설정하는 방법이 제안된다. 이 방식은 흐르는 전류 혹은 온도 변화에 따라 기준 전압도 자연스럽게 조정되므로, R_P 와 R_{AP} 사이의 정상적인 비교 동작이 유지되며, 출력 오동작 가능성을 줄일 수 있다. 해당 회로는 하나의 p-bit을 구성하는 단위 셀로 설계되며, 각 p-bit의 디지털 출력 $V_{output,i}$ 는 멀티플렉서(MUX)를 통해 행렬 연산 유닛(matrix computation unit) 또는 마이크로컨트롤러(MCU)로 전달된다. MCU는 이를 현재 시스템 상

태로 인식하고, 사전에 정의된 가중치 행렬 J_{ij} 와의 연산을 수행하여 새로운 바이어스 값을 계산한다. 이 결과는 다시 각 MTJ 소자의 게이트 입력 전압(V_{input})으로 피드백되어, 다음 연산 주기에서 p-bit의 상태 갱신에 반영된다. 이러한 구조를 통해 확률 연산 컴퓨팅 시스템은 실시간 상태 갱신과 반복적인 확률 기반 연산을 동적으로 수행할 수 있으며, 관련된 시스템 구성은 Fig. 3(c)에 제시되어 있다.

V. MTJ를 제외한 다른 p-bit 후보군

s-MTJ 소자 외에도 p-bit의 물리적 구현을 위한 다양한 플랫폼이 제안되어 왔다. P-bit은 본질적으로 입력 바이어스에 따라 이진 확률 분포를 조절할 수 있는 난수 발생기이므로, 이러한 특성을 만족하는 물리 시스템은 모두 p-bit 구현에 활용될 수 있다. 대표적인 예로는 멤리스터(memristor), 상전이 소자(Mott insulator-to-metal transition, Mott IMT), OTS (ovonic threshold switch) 소자, 비선형 오실레이터(nonlinear oscillator), 2차원 물질 기반 소자, 전계 효과 트랜지스터(field-effect transistor), 링 오실레이터의 위상 잡음(phase noise), Johnson-Nyquist 열잡음 등이 있으며, 이들 시스템에서 확률적 동작 특성이 보고된 바 있다. 멤리스터 기반 p-bit 구현에서는 휘발성 멤리스터(volatile memristor) 또는 임계

스위칭 소자(threshold switching device)의 확률적 전환 특성을 이용하며, 입력 전압 바이어스에 따라 출력 확률을 시그모이드(sigmoid) 형태로 조절할 수 있음이 실험적으로 입증되었다. 광학적 플랫폼에서는 광파라메트릭 오실레이터(optical parametric oscillator, OPO) 네트워크가 p-bit의 무작위 진동 특성을 모사할 수 있음이 제안되었고, 이는 이징 모델(Ising model) 매핑을 통해 확률 논리 회로 구현이 가능한 것으로 보고되었다. 또한 금속-절연체-금속(metal-insulator-metal, MIM) 구조에서도 열적 요인에 의해 두 개의 저항 상태 간 무작위 전환 현상이 관찰되어, p-bit 구현 가능성을 보여준다.

그러나 이러한 대안적 플랫폼들은 대부분 이진 상태의 안정성, 확률 제어 정밀도, 또는 공정 호환성 측면에서 제약이 존재한다. 예를 들어, MIM 상전이 구조는 확률 제어의 정밀도가 낮고 대면적 균일 제작이 어렵다는 한계가 있으며, 안정적인 동작을 위해 보조 회로가 필요해 구현 효율이 떨어지는 것으로 평가된다. 한편, 기존 CMOS 기반의 선형 피드백 시프트 레지스터(linear feedback shift register, LFSR)나 이진 의사 난수 발생기(pseudorandom bit generator) 또한 p-bit 기능을 수행할 수 있지만, 진성 무작위성(randomness)을 제공하지 못하며, 수천 개의 트랜지스터가 요구되는 복잡한 회로 구성으로 인해 면적과 전력 소모 측면에서 불리하다.

VI. 결론 및 전망

본 논문은 기존의 폰 노이만 구조에서 비롯되는 계산 병목 현상과 에너지 소모 문제를 극복할 수 있는 차세대 연산 플랫폼으로서 확률론적 컴퓨팅(probabilistic computing)의 개념을 고찰하였다. 특히 본 연구는 에너지 효율성과 고도의 병렬 처리가 요구되는 역연산 기반 계산 영역에서 확률론적 컴퓨팅이 갖는 높은 적용 가능성에 주목하였다. 이러한 컴퓨팅 아키텍처의 하드웨어 구현을 위한 핵심 소자로는 랜덤 자기터널 접합소자(stochastic magnetic tunnel junction, s-MTJ)가 사용되며, 이는 기존 MRAM 공정과의 호환성을 바탕으로 높은 실용성과 집적 가능성을 제공한다. 더불어 s-MTJ는 상온 동작, 초저전력 소비, 고집적화 등 다양한 물리적 이점을 지니고 있어 실용적 응용에 적합하다. s-MTJ 기반 확률론적 컴퓨팅 시스템은 조합 최적화, 암호 해독, 베이지안 추론 등 복잡한 계산 문제 해결에 효과적으로 활용될 수 있다. 향후 연구에서는 s-MTJ의 자유층 구조 및 자성 재료의 최적화, 소자간 분산 제어, 신뢰성 있는 입출력 회로 설계, 적절한 알고리즘 개발, 그리고 대규모 집적을 위한 시스템 통합 기술 개발이 주요 과제로 남아 있다. 이들이 해결 한다면 고전 컴퓨터에서 양자 컴퓨터로 넘어가는 중간 단계의 역할을 훌륭히 수행할 수 있을 것으로 기대한다.

감사의 글

This work was supported by the KIST Institutional Program, the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. GTL24041-000), and the KU-KIST Institutional Program.

References

- [1] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, *Phys. Rev. X* **7**, 031014 (2017).
- [2] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, *Nature* **573**, 390 (2019).
- [3] J. Kaiser and S. Datta, *Appl. Phys. Lett.* **119**, 150503 (2021).
- [4] N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, J. M. Martinis, G. Finocchio, and K. Y. Camsari, *Nat. Electron.* **5**, 460 (2022).
- [5] S. Chowdhury, A. Grimaldi, N. A. Aadit, S. Niazi, M. Mohseni, S. Kanai, H. Ohno, S. Fukami, L. Theogarajan, G. Finocchio, S. Datta, and K. Y. Camsari, *IEEE J. Explor. Solid-State Comput. Devices Circuits* **9**, 1 (2023).
- [6] S. Misra, L. C. Bland, S. G. Cardwell, J. A. C. Incorvia, C. D. James, A. D. Kent, C. D. Schuman, J. D. Smith, and J. B. AIM-one, *Adv. Mater.* **35**, e2204569 (2023).
- [7] R. P. Feynman, *Int. J. Theor. Phys.* **21**, 22 (1982).
- [8] G. F. Newell and E. W. Montroll, *Rev. Mod. Phys.* **25**, 353 (1953).
- [9] S. G. Brush, *Rev. Mod. Phys.* **39**, 883 (1967).
- [10] R. A. Rutenbar, *IEEE Circuits and Devices Mag.* **5**, 19 (1989).
- [11] D. Bertsimas and J. Tsitsiklis, *Stat. Sci.* **8**, 10 (1993).
- [12] S. Tehrani, B. Engel, J. M. Slaughter, E. Chen, M. DeHerrera, M. Durlam, P. Naji, R. Whig, J. Janesky, and J. Calder, *IEEE Trans. Magn.* **36**, 2752 (2000).
- [13] J.-G. Zhu and C. Park, *Mater. Today* **9**, 36 (2006).
- [14] S. Yuasa and D. D. Djayaprawira, *J. Phys. D: Appl. Phys.* **40**, R337 (2007).
- [15] K. Watanabe, B. Jinnai, S. Fukami, H. Sato, and H. Ohno, *Nat. Commun.* **9**, 663 (2018).
- [16] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, and H. Ohno, *Nat. Mater.* **9**, 721 (2010).
- [17] M. Wang, Y. Zhang, X. Zhao, and W. Zhao, *Micromachines* **6**, 1023 (2015).
- [18] B. R. Zink, Y. Lv, and J.-P. Wang, *IEEE J. Explor. Solid-State Comput. Devices Circuits* **8**, 173 (2022).
- [19] J. Z. Sun, *J. Appl. Phys.* **137**, 013901 (2025).
- [20] K. H. Han, Y. J. Kim, H. C. Koo, O. J. Lee, and S. Hong, *Appl. Phys. Lett.* **125**, 142402 (2024).
- [21] G. Srinivasan, A. Sengupta, and K. Roy, *Proceedings of the 2017 Design, Automation and Test in Europe, DATE 2017*, 530 (2017).

- [22] C. Safranski, J. Kaiser, P. Trouilloud, P. Hashemi, G. Hu, and J. Z. Sun, *Nano Lett.* **21**, 2040 (2021).
- [23] O. Hassan, S. Datta, and K. Y. Camsari, *Phys. Rev. Appl.* **15**, 064046 (2021).
- [24] O. Hassan, R. Faria, K. Y. Camsari, J. Z. Sun, and S. Datta, *IEEE Magn. Lett.* **10**, 4502805 (2019).
- [25] L. E. Bassham III, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, S. D. Leigh, M. Levenson, M. Vangel, and D. L. Banks, *A Statistical Test Suite for Random and Pseudo-random Number Generators for Cryptographic Applications*, National Institute of Standards & Technology Special Publication (NIST SP) - 800-22 Rev 1a (2010).
- [26] T. Kim, H. G. Park, K.-H. Han, Y.-J. Nah, H. C. Koo, B.-C. Min, S. Hong, and O. J. Lee, *AIP Adv.* **12**, 075104 (2022).
- [27] Y. Lu, X. W. Li, G. Xiao, R. A. Altman, W. J. Gallagher, A. Marley, K. Roche, and S. Parkin, *J. Appl. Phys.* **83**, 6515 (1998).