文章编号: 1007-4252(2025)05-0387-10 DOI: 10.20027/j.gncq.2025.0049 CSTR: 32180.14.j.gncq.2025.0049

磁存储器的翻转机理及先进电学表征研究进展

马小倩[1], 高世凡[2], 曲益明[2,*]

(1. 华东师范大学通信与电子工程学院, 上海 200241; 2. 浙江大学信息与电子工程学院, 杭州 310027)

摘要: 为突破冯·诺伊曼架构的能效与数据吞吐瓶颈, 基于脉冲神经网络的存算一体化架构对底层非易失性存储器的性能提出严苛要求。在众多新型存储器中, 自旋转移矩磁性随机存储器(spin-transfer torque magnetic random access memory, STT-MRAM)凭借纳秒级读写速度、极高耐久性、出色的数据保持能力、低功耗以及与互补金属-氧化物-半导体(complementary metal-oxide-semiconductor, CMOS)工艺的后端兼容性, 成为极具潜力的候选器件。STT-MRAM 的核心元件为磁隧道结(magnetic tunnel junction, MTJ), 其通过电流直接驱动磁化翻转的物理机制是实现快速读写和低功耗操作的关键。系统综述 STT-MRAM 快速翻转动力学机理与性能优化策略的最新研究进展, 重点分析 MTJ 核心结构参数(如自由层材料、垂直磁各向异性能等)对其翻转性能和可靠性的影响规律, 并对先进电学表征技术进行详细介绍, 为面向存算一体应用的高性能 STT-MRAM 设计提供理论依据与技术路径。

关键词: 自旋转移矩磁性随机存储器; 磁隧道结; 快速电学表征; 翻转动力学; 翻转速度; 临界翻转电流

中图分类号: TN47 **文献标志码:** A

Research progresses on switching mechanism and advanced electrical characterizations of magnetic random access memory

MA Xiaoqian[1], GAO Shifan[2], QU Yiming[2,*]

(1. School of Communication and Electronic Engineering, East China Normal University, Shanghai 200241, China; 2. College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: To address the energy efficiency and data throughput limitations of the Von Neumann architecture, computing-in-memory (CIM) systems based on spiking neural networks (SNNs) impose

收稿日期: 2025−09−29; **修订日期:** 2025−10−29

基金项目: 新一代人工智能国家科技重大专项(No. 2020AAA0109001).

作者简介: 马小倩(1995—), 女, 博士研究生, 主要研究方向为 MRAM 的可靠性测试(E-mail: 1183689705@qq.com).

通信作者: 曲益明(1993—), 女, 博士, 研究员, 主要研究方向为先进工艺节点器件、新型存储器件的工艺及可靠性(E-mail: quyiming@zju.edu.cn).

rigorous demands on the performance of non-volatile memory technologies. Among the various emerging memory solutions, spin-transfer torque magnetic random access memory (STT-MRAM) is regarded as a highly promising candidate for low-power CIM applications. This is attributed to its nanosecond-scale read/write speeds, high endurance, excellent data retention, low power consumption, and compatibility with complementary metal-oxide-semiconductor (CMOS) back-end-of-line processes. The core component of STT-MRAM is the magnetic tunnel junction (MTJ), which operates based on the physical principle of current-driven magnetization switching, enabling both fast and energy-efficient functionality. This review summarize recent advancements in the fast switching dynamics and performance optimization of STT-MRAM, focus on analyzing the influence law of key structural parameters of MTJ (such as free layer materials, perpendicular magnetic anisotropy energy, etc.) on its switch performance and reliability. We also introduce the advanced electrical characterization techniques in details, offering a valuable guidance and technical path for designing high-performance STT-MRAM tailored to computing-in-memory applications.

Key words: STT-MRAM; MTJ; fast electrical characterization; switching dynamics; switching speed; critical switching current

0 引言

在传统冯·诺伊曼架构中，处理器与存储器之间频繁的数据搬运会导致高延迟和巨大功耗问题，这已成为制约片上系统能效进一步提升的关键瓶颈[1]。为突破这一限制，存算一体架构作为一种颠覆性的电路与系统设计范式，近年来受到人们的广泛关注[2]。该架构的核心理念是将计算功能直接嵌入存储阵列，通过原位数据操作消除或大幅减少数据迁移，从而在架构层面实现能效与吞吐量的显著提升[3-4]。

在众多可用于实现存算一体的非易失性存储技术中，自旋转移矩磁性随机存储器（spin-transfer torque magnetic random access memory, STT-MRAM）因与标准互补金属-氧化物-半导体（complementary metal-oxide-semiconductor, CMOS）工艺的高度兼容性、亚纳秒级的读写速度、极高的耐久性以及非易失存储特性，被视为极具潜力的技术路径[5-7]。STT-MRAM 的核心是具有多层磁性薄膜结构的自旋转移矩磁隧道结（spin-transfer torque magnetic tunnel junction, STT-MTJ），其工作原理为：向 MTJ 中注入自旋极化电流可改变自由层的磁化方向，当自由层与参考层的磁化方向平行时，MTJ 呈现低阻态；反之则呈现高阻态。基于这一机制，STT-MRAM 不仅能实现二进制位“0”和“1”的存储，也为逻辑运算与数据存储之间的高效融合提供了物理基础[6]。

在过去的十几年间，人们在 STT-MRAM 的快速翻转性能方面取得了显著进展，其翻转时间从早期的数十纳秒缩短至如今的皮秒量级[8-10]，图 1 展示了部分关键研究所报道的 MTJ 翻转时间[8-14]。然而，研究[11]也表明，存储器的翻转速度越快，功耗越高。因此，如何在提高翻转速度的同时降低功耗，成为亟待解决的关键问题。在此过程中，快速电学测试方法对于研究 MTJ 的翻转性能至关重要。相较于直流测试，超高速脉冲测试能够更准确地反映 MTJ 在高速应力环境下的翻转特性[9]。

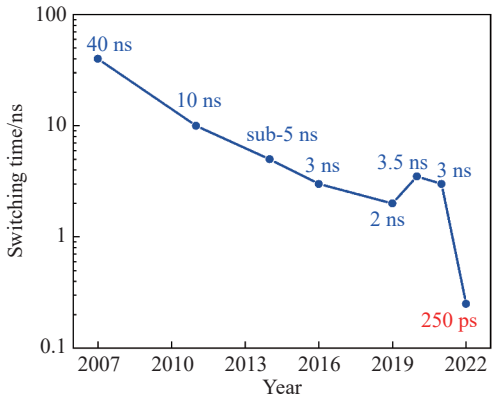

图 1 近年来 MTJ 翻转时间不断朝着更快的读写速度迈进，目前已达亚纳秒水平[8-14]

Fig. 1 In recent years, the switching time of MTJ has been progressing toward faster read and write speeds, currently achieving sub-nanosecond levels[8-14]

本文系统梳理 MTJ 快速翻转方面的研究进展，

并探讨其在推动高能效、高性能存算一体芯片发展中的关键作用。首先, 阐述 MTJ 的翻转机制及其内在随机性在概率计算中的应用, 分析 MTJ 结构优化对翻转性能的影响; 其次, 分析快速电学测试方法的演进及其在 MTJ 翻转机制研究中的作用; 最后, 展望 MTJ 快速翻转的发展前景并讨论其面临的挑战。

1　MTJ 的快速翻转及研究进展

1.1　MTJ 的翻转原理

MTJ 的快速翻转也被称为快速写入, 其翻转原理如图 2 所示。图 2(a)展示了 MTJ 从反平行态切换到平行态的过程。在此过程中, 电流从自由层流向参考层, 携带自旋角动量的电子则从参考层流入自由层。大多数自旋方向平行于自由层磁化方向的电子会穿过势垒层进入自由层; 而少数自旋方向与自由层磁化方向相反的电子则被反射回参考层。隧穿电子携带的自旋角动量作用于自由层的磁矩, 使其方向发生偏转, 当隧穿电子的数量达到阈值时, 自由层磁化方向发生翻转, MTJ 切换至平行态。

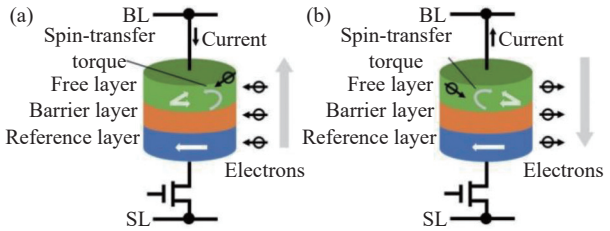

图 2　MTJ 的经典翻转原理示意图。(a)反平行态到平行态的翻转; (b)平行态到反平行态的翻转

Fig. 2 Schematic diagrams of the conventional switching mechanism in MTJ. (a) Switching from the antiparallel state to the parallel state; (b) switching from the parallel state to the antiparallel state

相反, 当 MTJ 从平行态切换到反平行态时(图 2(b)), 电流从参考层流向自由层, 自旋极化电子从自由层流入参考层。此时, 大多数与参考层方向相同的电子隧穿进入参考层, 而少数与参考层方向相反的电子反射回自由层。这部分电子携带的自旋角动量作用于自由层磁矩, 使其方向发生偏转, 当反射电子数量达到阈值时, 自由层磁化方向发生翻转, MTJ 呈现反平行态。

1.2　MTJ 翻转动力学及其在概率计算中的应用

相较于传统的磁场诱导翻转, STT-MRAM 通过自旋极化电流写入的翻转机制具有本征随机性, 因此显示出更高的失败率[15]。在存算一体架构中执行常规计算任务时, 这种随机性会成为主要的误差来源[16]。尽管通过增加翻转电流的幅值和脉宽可以缓解这一问题, 但同时也需要权衡功耗和翻转速度方面的限制[15]。研究[17]表明, 缩小 MTJ 的尺寸有助于在低电流条件下实现快速读取操作。因此, 实现更小尺寸的 MTJ 是提高器件翻转速度并降低磁存储器芯片功耗的有效途径之一。

近年来, MTJ 的随机性逐渐被转化为一种计算资源, 并得到广泛研究与应用[16,18]。描述 MTJ 磁化翻转的经典物理模型为 Landau-Lifshitz-Gilbert (LLG)方程。传统的 LLG 方程仅能描述 MTJ 的确定性翻转行为[19]。然而, 考虑到翻转动力学过程中的内在随机性, 需在经典模型中引入表征热扰动的随机热涨落场 H_{thermal}, 从而构成随机 LLG 方程[15, 19-20]:

$$\mathrm{d}M/\mathrm{d}t = -\gamma M H_{\mathrm{eff}} - \alpha(M\cdot \mathrm{d}M/\mathrm{d}t) + (\gamma/M_{\mathrm{s}})\tau_{\mathrm{STT}} \quad (1)$$

$$H_{\mathrm{eff}} = H_{\mathrm{ext}} + H_{\mathrm{thermal}} + H_{\mathrm{other}} \quad (2)$$

式中, M 为自由层的磁化矢量, t 为时间, γ 为旋磁比, α 为 Gilbert 阻尼常数, M_{s} 为自由层饱和磁化强度, τ_{STT} 为自旋转移矩项, H_{eff} 为有效磁场, H_{ext} 为外加磁场, H_{other} 为各向异性场、退磁场、交换场及安培场的总和。

热涨落场 H_{thermal} 满足以下方程:

$$\langle H_{i,\ \mathrm{thermal}}(t) H_{j,\ \mathrm{thermal}}(t+\mathrm{d}t)\rangle = \delta_{i,\ j}(2\alpha k_{\mathrm{B}}T)/(\mu_0\gamma V M_{\mathrm{s}}\mathrm{d}t) \quad (3)$$

式中, $\delta_{i,j}$ 为克罗内克尔符号, k_{B} 为玻尔兹曼常数, μ_0 为真空磁导率。可以看出, 热涨落场的强度与温度 T 成正比, 与自由层体积 V 成反比。这从物理机理上解释了纳米尺度 MTJ 中热涨落效应如此显著的原因, 也说明热涨落已成为其翻转动力学的关键影响因素。

通过大规模微磁模拟求解随机 LLG 方程, 可统计不同写入电流 I 与脉冲宽度 τ 条件下磁化矢量的成功翻转次数, 进而建立翻转概率与写入参数之间的函数关系[21]。理论与实验研究均表明, 翻转概率随写入电流呈周期性波动, 该现象超越了简单的热激活模型的框架, 是磁矩在能量势阱中非线性进动的直接体现[15]。通过精确调控写入脉冲参数, 可将 MTJ 的翻转概率连续设置在 0 与 1 之间。基

于这一物理基础，研究者正尝试将单个 MTJ 器件定义为可编程的概率比特。这种将物理随机性转化为计算资源的能力，标志着 STT-MRAM 研究从致力于抑制随机性以保障存储可靠性，向主动利用随机性以赋能概率计算的重要范式转变。

在该范式推动下，近年来相关研究取得了显著突破。基于 STT-MTJ 翻转动力学的随机计算，已形成从专用架构设计到硬件系统原型的完整技术路径[16, 18]，其核心在于利用 MTJ 的随机翻转特性作为随机数源以支持高效的随机计算。Gu 等[16]通过构建一种存内计算架构，直接在 MRAM 阵列中实现了基于随机计算的贝叶斯神经网络(Bayesian neural network)，从而有效执行不确定性推理任务。作为该方向的硬件验证，Yu 等[18]成功研制了一种 MRAM 随机计算系统的硬件原型，利用 MTJ 的固有随机性生成随机位流，在硬件层面完整实现了随机计算的核心功能。这两项工作共同证明了基于 MTJ 翻转随机性构建从专用架构到完整系统的存内随机计算的可行性与巨大应用潜力。

1.3 结构优化对 MTJ 翻转性能的影响

要实现在存算一体电路中的成功应用，STT-MTJ 需满足一系列严苛的器件与集成指标，例如：电路设计的可靠性要求 MTJ 具备优异的热稳定性和足够长的数据保持时间[7, 22]；而高算力则与其翻转速度有着直接关联[6]。

近年来，通过材料掺杂等策略，MTJ 的翻转速度得到了显著提高，甚至达到皮秒量级[8]；同时，其翻转电流密度也有明显降低，从而推动了这类磁存储器芯片功耗的持续优化。本节系统探讨 MTJ 的结构创新对其快速翻转性能的影响，包括界面各向异性调控、双势垒层和双参考层的引入以及材料掺杂等[12-13, 23-27]。

面内各向异性 MTJ 通常要求器件具有较大的长宽比和椭圆形形状，这使其对先进工艺节点的兼容性受到挑战。基于此，垂直磁各向异性 MTJ (perpendicular magnetic anisotropy MTJ, PMA-MTJ)受到了广泛关注[12]。

Park 等[23]提出了一种具有双 MgO 层的 PMA-MTJ 结构，通过增加界面各向异性和自由层厚度，将 MTJ 直径减小至 20 nm 以下，实现了良好的翻转性能。图 3(a)和(b)为单界面自由层和双界面自由层 MTJ 的器件结构示意图。双界面自由层 MTJ 额外增加了自由层与 MgO 层的下界面，可增强器件的垂直磁各向异性能，从而增大热稳定性因子 Δ，并使归一化翻转电流 I_{c0} 与 Δ 的比值降至单界面自由层 MTJ 的一半，如图 3(c)所示。由于 Δ 直接决定 MTJ 的数据保持特性，因此该结构设计在满足器件高可靠性的同时，通过减小 I_{c0} 降低了功耗。

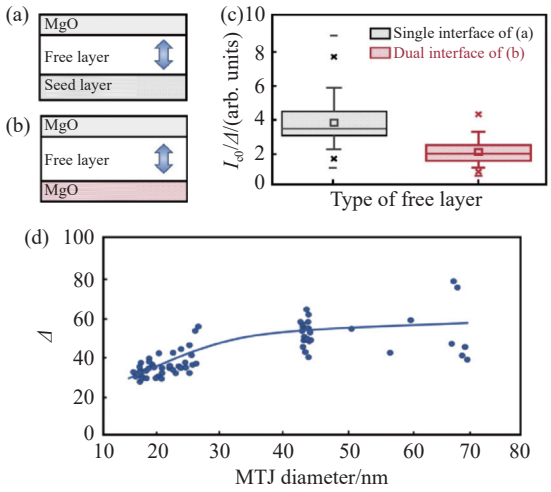

图 3 (a)单界面自由层 MTJ 示意图；(b)双界面自由层 MTJ 示意图；(c)不同脉宽下的 I_{c0}/Δ 值；(d)单界面自由层 MTJ 的 Δ 与器件尺寸的关系[23]

Fig. 3 (a) Schematic diagram of the single interface free layer MTJ; (b) schematic diagram of the dual interface free layer MTJ; (c) I_{c0}/Δ values at different pulse widths; (d) relationship between Δ and device diameter in single interface free layer MTJ[23]

图 3(d)进一步展示了 Δ 与单界面自由层 MTJ 尺寸的关系：当 MTJ 尺寸大于 35 nm 时，Δ 随 MTJ 尺寸的下降趋势较缓，表明在该尺寸区间，MTJ 的翻转过程接近一致相干翻转，能量势垒由整个自由层的体积决定；当 MTJ 尺寸缩减至 35 nm 以下时，边缘损伤导致的 PMA 非均匀性效应加剧，翻转机制逐渐过渡为以畴壁成核和扩散为主导的非一致翻转，能量势垒路径降低，Δ 急剧下降。为解决单界面自由层 MTJ 在亚 35 nm 节点发生的翻转机制转变问题，该研究[23]通过引入双界面结构，增强了整个自由层(尤其是边缘)的 PMA 均匀性，有效抑制了畴壁辅助翻转，使器件在更小尺寸下仍能维持一致翻转，从而实现了高热稳定性和较低的写入电流。

在数据保持特性方面，MTJ 的非易失性是实现高可靠性存算一体的关键。通过增强自由层的

垂直磁各向异性, 可在维持较低写入电流的同时显著延长信息存储时间, 从而更好地适配需长期保留数据的存算一体应用场景。

Tsunoda 等[17] 设计了一种直径为 53 nm 的新型 MTJ, 如图 4 所示。该结构由虚拟自由层和垂直磁各向异性双磁隧道结组成, 其中虚拟自由层(上层)通过薄帽层与堆栈自由层(下层)实现弱耦合。MgO 帽层作为第二势垒, 用于提升各向异性能量并降低自由层的自旋阻尼, 从而减小翻转电流。此外, 虚拟自由层的杂散场可增强自由层的热稳定性。该器件的 Δ 为 52, I_{c0} 为 57 μA, 实现了在保持高热稳定性的同时降低翻转功耗。类似地, Jan等[12] 设计了一种具有双 MgO 层结构的 28 nm PMA-MTJ, 进一步增强了 PMA 效应, 翻转速度可达 1.5 ns, 且低于 30 nm 的尺寸有助于实现低功耗快速写入。

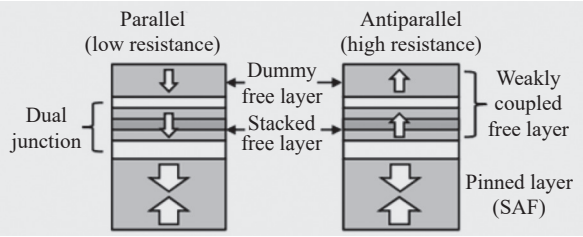

图 4　具有虚拟自由层的 MTJ 器件在平行态和反平行态的结构示意图[17]

Fig. 4　Schematics diagrams of the parallel and antiparallel states of the MTJ with the dummy free layer[17]

除引入自由层外, 研究者们还通过引入双参考层等手段优化 MTJ 的器件特性。图 5 展示了典型的双自旋扭矩磁隧道结器件结构[27], 通过引入双参考层, 该器件实现了低翻转电流和较快的翻转速度。该结构通过增加参考层与 MgO 界面以及第二参考层与间隔层界面的扭矩, 增强了自旋极化电子对自由层磁矩的作用, 提高了自旋极化率, 从而降低了自由层的翻转电流并提高了翻转速度。采用该创新结构的器件翻转电流降低至原有基本结构

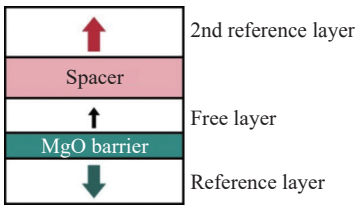

图 5　双自旋扭矩磁隧道结器件的结构示意图[27]

Fig. 5 Schematic diagram of the double spin-torque MTJ device[27]

器件的 1/2, 并实现了 3 ns 的快速翻转[27]。另一研究[8] 表明, 基于双参考层的 MTJ 可实现亚纳秒级翻转和更高的可靠性, 如图 6 所示, 其写入速度达 250 ps, 写入错误率为 1×10^{-6}。

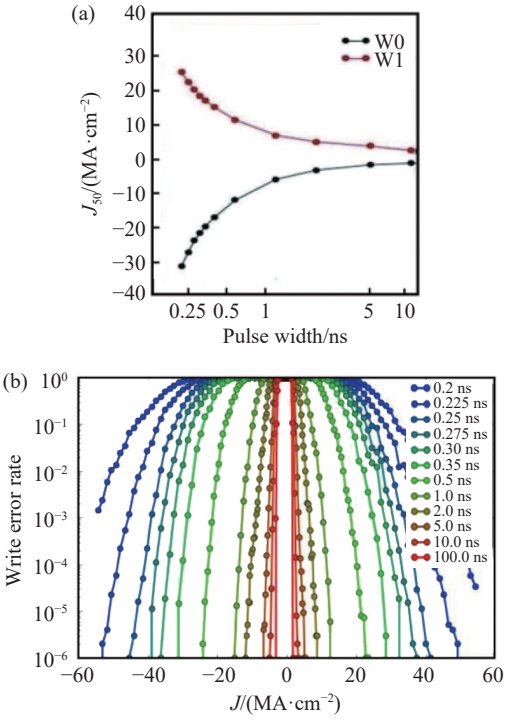

图 6　(a)单个器件写入 1(W1)和写入 0(W0)对应的开关电流密度, 开关概率为 50%; (b)单个器件在多种不同脉冲宽度(0.2~100 ns)下的写入错误率[8]

Fig. 6 (a) Single device write 1 (W1) and write 0 (W0) switching current densities for fifty percent switching probability; (b) write error rate on a single device for various pulse widths ranging from 0.2 ns to 100 ns[8]

为实现更高密度、更快写入速度和更低成本的 MRAM, 进一步缩小 MTJ 的尺寸势在必行。然而, 随着器件尺寸的减小, 其可靠性退化问题日益突出[28], 尤其是小尺寸情况下, 器件的数据保持能力面临严峻挑战。如图 7 所示, 当传统双界面自由层 MTJ 器件的直径减小至 20 nm 以下时, 其 Δ 迅速降低, 难以满足 10 年以上数据保持的高可靠性要求[24]。因此, 对于直径小于 20 nm 的 MTJ, 核心挑战在于缩小直径的同时保持较高的 Δ, 以确保可靠的数据保持能力。

为此, 研究者[24] 在 MTJ 器件的自由层中引入掺杂 Co-Pt 的合金层, 成功开发出 14 nm 高性能 MTJ。该器件有效降低了翻转电流密度, 同时实现了高数据保持特性以及稳定的 5 ns 快速翻转速度。

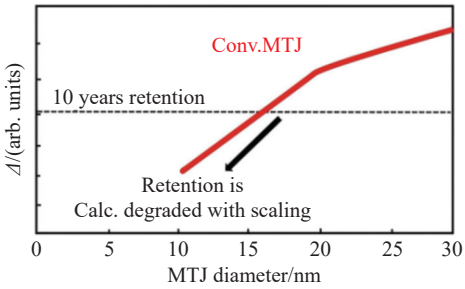

图 7 计算得到的传统 MTJ 直径与 Δ 的关系，其中虚线表示 10 年数据保持特性对 Δ 的要求[24]

Fig. 7 Calculated Δ in conventional MTJ versus diameter, the dashed line represents the requirement of Δ for 10-year data retention[24]

由于 Co-Pt 合金层具有较高的垂直磁各向异性能量密度[29]，即使在较小体积下，器件仍能保持较高的热稳定性（90 ℃ 下 Δ 为 55），如图 8 所示[24]。该微磁模拟研究为开发高性能 sub-20 nm MTJ 提供了重要参考。

当 MTJ 的直径进一步减小至 sub-10 nm 时，翻转能量虽然有所降低，但翻转错误率可能上升。这主要是由于在该尺度下对晶格模型的偏离会导致磁化阻尼的迅速下降，从而损害器件的热稳定性[25]。目前已有研究[13]报道了直径为 5 nm 以下的超小型 MTJ，测试结果表明其兼具高数据保持特性和 10 ns 以内的翻转速度。该超小型 MTJ 实现优异性能的关键在于通过增加多层铁磁体中的 CoFeB/MgO 界面数量来增强有效磁各向异性场强 H_k^{eff}，从而有效调控 MTJ 翻转的特征弛豫时间 τ_D，进而实现 3.5 ns 的稳定翻转。

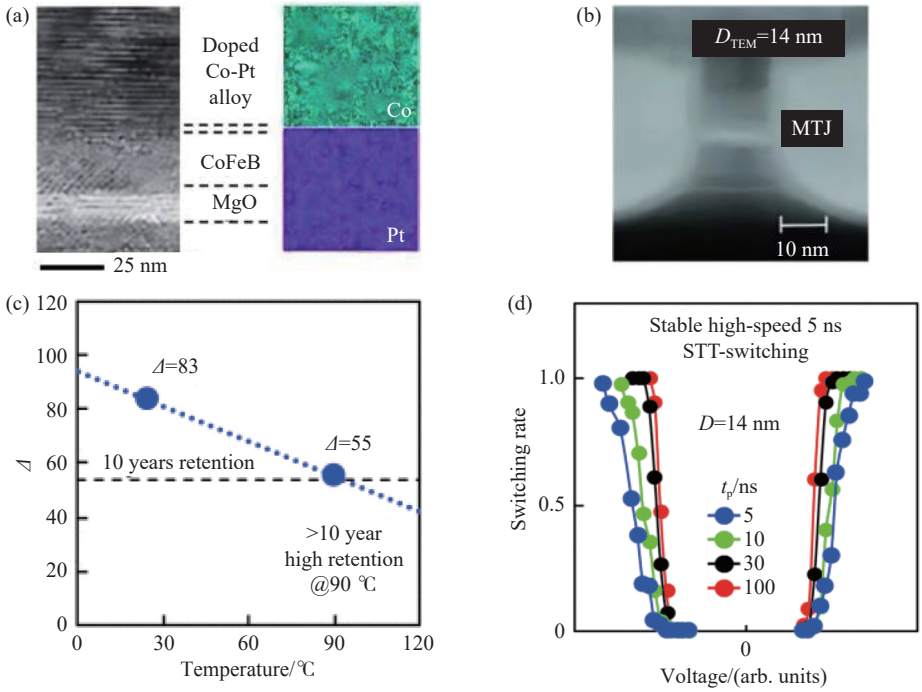

图 8 （a）14 nm 高性能 MTJ 薄膜的透射电子显微镜图像及掺杂 Co-Pt 合金薄膜的元素分布结果；（b）器件横截面透射电子显微镜图像；（c）不同温度下的 Δ 结果；（d）14 nm 高性能 MTJ 在室温下的典型翻转概率，可实现较稳定的 5 ns 高速翻转[24]

Fig. 8 (a) Cross-sectional transmission electron microscopy (TEM) image and corresponding elemental mapping of the Co-Pt alloy-doped thin film in the 14 nm high-performance MTJ; (b) TEM image of the device cross-section; (c) Δ vs Temperature; (d) representative switching probability of the 14 nm high-performance MTJ at room temperature, demonstrating stable high-speed switching with a 5 ns pulse[24]

特征弛豫时间 τ_D 是决定 MTJ 在进动翻转阶段翻转行为的关键因素，更快的弛豫意味着磁化状态能更迅速地稳定至目标方向，从而直接降低翻转随机性和错误率，这对于存算一体架构中要求确定性的二进制神经网络等应用至关重要。τ_D 与 H_k^{eff} 成反比，二者关系[13]可表示为：

$$\tau_D = (1+\alpha^2)/\left(\alpha\gamma\mu_0 H_k^{eff}\right) \tag{4}$$

式中，α 为阻尼常数，γ 为旋磁比，μ_0 为真空磁导率。基于此，可通过改变 MTJ 的磁性特性来调控 τ_D。

同时，翻转电压 V_C（翻转概率 P_{sw} = 0.5）与 τ_D

的关系[13]可表示为:

$$V_C = V_{C0}\{1+(\tau_D/\tau_p)\ln[4\Delta/\ln(2)]/2\} \quad (5)$$

式中, τ_p 为脉冲宽度, V_{C0} 为本征翻转电压。可见, 在较短脉冲下, 随着 τ_D 的减小, V_C 的增长趋缓, 有助于降低翻转功耗。

2　快速电学表征技术的研究进展

STT-MTJ 的翻转动力学行为所决定的读写速度对存算一体架构的实时处理能力具有重要影响。然而, 传统的直流或长脉冲(脉宽>100 ns)测试因显著的焦耳热累积, 难以准确反映器件的真实动态翻转过程。相比之下, 超短脉冲(脉宽<10 ns)测试能够更精确地评估器件的关键性能指标并揭示其翻转物理机制[9]。本节系统探讨快速电学测试方法在研究 MTJ 翻转机制中的作用, 重点分析超短脉冲测试与脉冲设计对 MTJ 翻转动力学的影响。

在 MTJ 的早期研究中, 研究人员采用持续时间为 50 μs~1 ms 的长脉冲测试, 揭示了 MTJ 在热激活阶段的磁化翻转机制[30]。如前所述, 该方法在准确度方面存在一定局限性。

短脉冲测试方法的引入突破了对 MTJ 快速翻转机制的研究限制[9, 31]。Zhu 等[9]通过实验明确了在 MgO 基 MTJ 中存在两种翻转机制, 分别为快速进动翻转(发生在数纳秒范围内)和慢速热激活翻转(写入时间超过数十纳秒时发生)。如图 9 所示: 当脉冲持续时间大于 10 ns 时, MTJ 处于热激活翻转阶段, 此时磁矩翻转受到热激活辅助, 因此翻转电流密度较小[21]; 当脉冲持续宽度小于 10 ns 时, MTJ 进入快速进动翻转阶段, 此时的翻转机制变为由电流驱动的进动翻转主导, 翻转电流密度随脉冲宽度的减小而急剧增加。Hu 等[11]的研究进一步验证了这两种翻转机制的存在, 如图 10 所示。究其内在原因, 在长脉冲情况下, 由于时间足够长, 在随机热振动下, 磁矩有充分的时间越过能量势垒, 而无需更多外界能量辅助, 因此写入功耗较低[21]; 然而, 随着脉宽进一步缩短(例如, 从 10 ns 减少至 3 ns), 器件进入进动翻转阶段, 需要足够的角动量来克服磁阻尼以实现在极短时间内完成翻转, 导致所需要的写入电流急剧增加。值得注意的是, 这一电能驱动过程会导致写入功耗的急剧增加[11]。这些研究表明, 短脉冲测试对于揭示 MTJ 翻转动力学起到了重要作用。

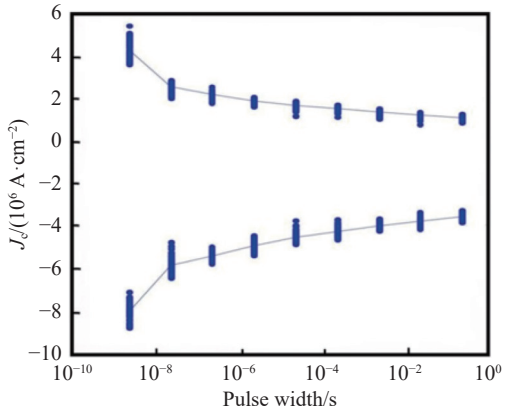

图 9　翻转电流密度与电流脉宽的关系[9]

Fig. 9　Switching current density versus current pulse widths[9]

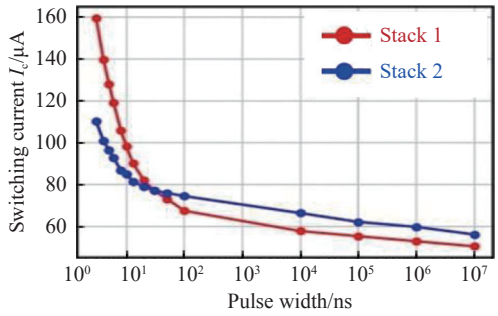

图 10　两种具有不同自由层材料的 MTJ 的翻转电流与脉宽关系曲线[11]

Fig. 10 Switching current versus pulse-width curves of two MTJs with different free layer materials[11]

近期另一项研究[14]结合磁畴壁运动的微磁模拟与高时间分辨率电学测量, 深入探讨了不同写入速度下的翻转动力学模型。与简单地将 MTJ 的翻转过程划分为两个阶段不同, 该研究实现了对不同翻转阶段所占时间比例的实时追踪。如图 11 所示, 微磁模拟仿真与高速示波器追踪的翻转过程中的电学参数变化结果表明, MTJ 的翻转过程可分为两个阶段: 首先, 磁畴壁成核, 对应于图 11(b)中隧道磁阻(tunnel magnetoresistance, TMR)20% 的变化过程; 随后, 畴壁穿过器件并不断扩散, 对应于 TMR 过程从 20% 到 80% 的过渡阶段。这种时间分辨测量技术为研究影响畴壁成核与扩散的关键物理因素, 进而为有针对性地优化 MTJ 的翻转性能提供了可能。

此外, 脉冲形状对 MTJ 的翻转性能也有重要影响。阶跃式脉冲注入和梯度脉冲注入方法被认

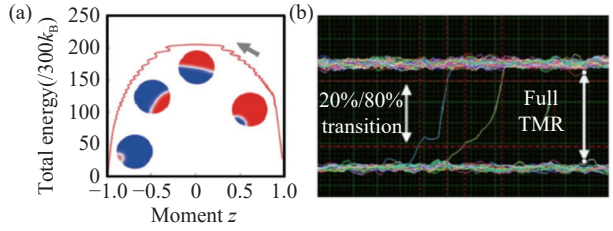

图 11 (a)利用微磁模拟仿真分析 MTJ 器件在不同时间尺度下的翻转动力学特性;(b)通过实验测试验证,翻转过程中成核期对应完整 TMR 达到 20% 的阶段,扩散期则对应从完整 TMR 的 20% 到 80% 的过渡阶段[14]

Fig. 11 (a) The switching dynamics of MTJ devices at different time scales using micro-magnetic simulations; (b) test verification results: nucleation time is defined as the time required to reach 20% of the full TMR during switching, while propagation time refers to the transition from 20% to 80% of the full TMR[14]

为能够有效缩短 STT-MRAM 翻转事件发生前的弛豫时间并降低功耗[32]。图 12(a)、(b)展示了梯度脉冲与恒定脉冲的理想波形示意图,并通过微磁模拟仿真,对比了在相同 20 ns 脉冲宽度下恒定脉冲(3 MA·cm^{-2})与梯度脉冲(5 MA·cm^{-2})的磁矩翻转过程,如图 12(c)所示。显然,梯度脉冲作用下的 MTJ 器件的翻转速度显著快于恒定脉冲。在 16 ns 时,梯度脉冲几乎已完成翻转,而恒定脉冲的磁矩翻转仍在进行中。该研究结果还表明,在梯度脉冲作用下,MTJ 器件的弛豫时间和翻转功耗均有所降低。这是由于梯度脉冲在完成初始驱动后会降低电流幅度,从而减少了总能量消耗[10]。这一方法

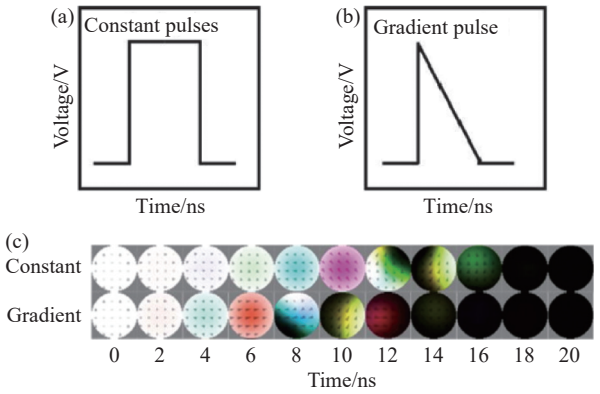

图 12 (a)恒定脉冲和(b)梯度脉冲的理想波形;(c)在 20 ns 恒定脉冲(3 MA·cm^{-2})和梯度脉冲(5 MA·cm^{-2})下模拟的磁矩翻转过程[10]

Fig. 12 Ideal waveforms of (a) constant pulse and (b) gradient pulse; (c) magnetic moment switching process under 20 ns constant (3 MA·cm^{-2}) and gradient (5 MA·cm^{-2}) pulses[10]

从物理层面提升了 MTJ 的翻转性能,有助于降低存算一体架构中单次计算的能量成本。

更进一步地,相关研究[33]揭示了不同脉冲形状对磁性存储器翻转效率和速度的影响。实验结果表明,相较于方波脉冲,下降三角波、上升三角波、对称三角波以及半周期正弦波等 4 种变形脉冲均可在保持相同翻转概率与速度的前提下,将写入能量降低 50%。尽管该实验聚焦于自旋转移矩效应辅助的自旋轨道扭矩磁隧道结,但其研究结果仍可为探究脉冲形状对 MTJ 器件翻转过程的影响提供参考。

3 总结与展望

本文面向高性能存算一体芯片的发展需求,系统综述了 MTJ 快速翻转机制的研究进展,并阐明了其对高能效、高吞吐存算一体系统发展的重要意义。围绕该主题,梳理了近 20 年的关键研究成果,主要涵盖两个方面:一是通过优化 MTJ 结构(如界面工程、自由层材料)来提升其翻转性能;二是利用先进电学表征技术来揭示其微观翻转动力学机制。

尽管该领域已取得显著进展,但随着产业界对 STT-MRAM 在集成度、功耗、写入速度及数据保持特性等方面的要求日益严苛,未来仍面临诸多挑战与机遇。未来可从以下几个方面开展进一步研究:通过掺杂新型材料增强薄膜各向异性、采用不同形状脉冲来缩短弛豫时间,乃至利用飞秒激光脉冲来辅助实现超快翻转。这些新方法有望带来更多突破。

随着对 MTJ 物理机制的深入理解与性能的持续优化,基于 STT-MRAM 的存算一体架构有望突破现有算力瓶颈。本文梳理的研究脉络与前沿洞察可为开发面向下一代计算需求的高性能存储解决方案提供关键参考与设计指导,并推动其在不远的未来实现更广泛的商业应用。

参考文献:

[1] JAIN S, RANJAN A, ROY K, et al. Computing in memory with spin-transfer torque magnetic RAM[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2018, 26(3): 470-483.

[2] JAIN S, SAPATNEKAR S, WANG J P, et al. Computing-in-

memory with spintronics[C]//Proceedings of the 2018 Design, Automation & Test in Europe Conference & Exhibition. New York: IEEE, 2018: 1640-1645.

[3] WANG Y C, LI T, ENDOH T. A high-reliable and high-speed all-digital STT-MRAM based computing-in-memory for binary neural networks[J/OL]. IEEE Transactions on Magnetics, 2025: 1-5(2025-10-08). https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=11197029.

[4] KANG W, ZHANG H, ZHAO W S. Spintronic memories: From memory to computing-in-memory[C]//Proceedings of the 2019 IEEE/ACM International Symposium on Nanoscale Architectures. New York: IEEE, 2020: 1-2.

[5] JIANG X P, BAO J L, ZHANG L, et al. A novel dual-reference sensing scheme for computing in memory within STT-MRAM[J]. Microelectronics Journal, 2022, 121: 105355.

[6] ZHOU R, LIU B, SI X, et al. Complementary series-connected STT-MTJ for time-based computing-in-memory[C]//Proceedings of the 2024 IEEE International Symposium on Circuits and Systems. New York: IEEE, 2024: 1-5.

[7] LI J F. Testing and reliability of computing-in memories: solutions and challenges[C]//Proceedings of the 2022 IEEE International Test Conference in Asia. New York: IEEE, 2022: 55-60.

[8] SAFRANSKI C, HU G, SUN J Z, et al. Reliable sub-nanosecond MRAM with double spin-torque magnetic tunnel junctions[C]//Proceedings of the 2022 IEEE Symposium on VLSI Technology and Circuits. New York: IEEE, 2022: 288-289.

[9] ZHU W Z, LI H, CHEN Y R, et al. Current switching in MgO-based magnetic tunneling junctions[J]. IEEE Transactions on Magnetics, 2011, 47(1): 156-160.

[10] SUN L, GUO L, WANG G C, et al. A high-speed and power-efficient gradient-pulse injection method for spin-transfer torque magnetic random-access memory[J]. Applied Physics Letters, 2023, 123(3): 033502.

[11] HU G, NOWAK J J, GOTTWALD M G, et al. Spin-transfer torque MRAM with reliable 2 ns writing for last level cache applications[C]//Proceedings of the 2019 IEEE International Electron Devices Meeting. New York: IEEE, 2020: 2.6. 1-2.6. 4.

[12] JAN G, THOMAS L, LE S, et al. Demonstration of fully functional 8 Mb perpendicular STT-MRAM chips with sub-5 ns writing for non-volatile embedded memories[C]//Proceedings of the 2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers. New York: IEEE, 2014: 1-2.

[13] JINNAI B, IGARASHI J, SHINODA T, et al. Fast switching down to 3.5 ns in sub-5-nm magnetic tunnel junctions achieved by engineering relaxation time[C]//Proceedings of the 2021 IEEE International Electron Devices Meeting. San Francisco. New York: IEEE, 2022: 1-4.

[14] JAN G, THOMAS L, LE S, et al. Achieving sub-ns switching of STT-MRAM for future embedded LLC applications through improvement of nucleation and propagation switching mechanisms[C]//Proceedings of the 2016 IEEE Symposium on VLSI Technology. New York: IEEE, 2016: 1-2.

[15] SAIDA D, JIBIKI Y, TAKAGISHI M, et al. Periodic fluctuations of switching probability in spin-transfer magnetization switching in magnetic tunnel junctions[J]. IEEE Transactions on Magnetics, 2018, 54(9): 1400405.

[16] GU H Y, JIA X T, LIU Y H, et al. CiM-BNN: Computing-in-MRAM architecture for stochastic computing based Bayesian neural network[J]. IEEE Transactions on Emerging Topics in Computing, 2024, 12(4): 980-990.

[17] TSUNODA K, NOSHIRO H, YOSHIDA C, et al. A novel MTJ for STT-MRAM with a dummy free layer and dual tunnel junctions[C]//Proceedings of the 2012 International Electron Devices Meeting. New York: IEEE, 2013: 29.1. 1-29.1. 4.

[18] YU Z K, FEI T H, CAI H, et al. A hardware prototype of an MRAM-based stochastic computing system[C]//Proceedings of the Great Lakes Symposium on VLSI 2025. New Orleans: ACM, 2025: 835-840.

[19] CARPENTIERI M, CONSOLO G, AZZERBONI B, et al. Spin-transfer torque switching in magnetic multilayers[J]. IEEE Transactions on Magnetics, 2007, 43(4): 1677-1680.

[20] SLONCZEWSKI J C. Current-driven excitation of magnetic multilayers[J]. Journal of Magnetism and Magnetic Materials, 1996, 159(1-2): L1-L7.

[21] ITO K. Micromagnetic simulation on dynamics of spin transfer torque magnetization reversal[J]. IEEE Transactions on Magnetics, 2005, 41(10): 2630-2632.

[22] WANG C, WANG Z H, WANG G F, et al. Design of an area-efficient computing in memory platform based on STT-MRAM[J]. IEEE Transactions on Magnetics, 2021, 57(2):

3400504.

[23] PARK J H, KIM Y, LIM W C, et al. Enhancement of data retention and write current scaling for sub-20 nm STT-MRAM by utilizing dual interfaces for perpendicular magnetic anisotropy[C]//Proceedings of the 2012 Symposium on VLSI Technology. New York: IEEE, 2012: 57-58.

[24] NAKAYAMA M, OIKAWA S, KAMATA C, et al. 14 nm high-performance MTJ with accelerated STT-switching and high-retention doped Co-Pt alloy storage layer for 1Z nm MRAM[C]//Proceedings of the 2023 International Electron Devices Meeting. New York: IEEE, 2024: 1-4.

[25] HONG J M, HADJIKHANI A, STONE M, et al. The physics of spin-transfer torque switching in magnetic tunneling junctions in sub-10 nm size range[J]. IEEE Transactions on Magnetics, 2016, 52(7): 1400504.

[26] JINNAI B, IGARASHI J, WATANABE K, et al. High-performance shape-anisotropy magnetic tunnel junctions down to 2.3 nm[C]//Proceedings of the 2020 IEEE International Electron Devices Meeting. New York: IEEE, 2021: 24.6. 1-24.6. 4.

[27] HU G, LAUER G, SUN J Z, et al. 2X reduction of STT-MRAM switching current using double spin-torque magnetic tunnel junction[C]//Proceedings of the 2021 IEEE International Electron Devices Meeting. New York: IEEE, 2022: 2.5. 1-2.5. 4.

[28] SUN J Z. Spin-current interaction with a monodomain magnetic body: A model study[J]. Physical Review B, 2000, 62(1): 570-578.

[29] PENG S Z, ZHANG Y, WANG M X, et al. Magnetic tunnel junctions for spintronics: Principles and applications[EB/OL]. (2014-12-15). https://onlinelibrary.wiley.com/doi/full/10.1002/047134608X.W8231.

[30] LEE J M, YE L X, WENG M C, et al. Spin transfer magnetization switching read/write cycle test in MgO-based magnetic tunnel junctions[J]. IEEE Transactions on Magnetics, 2007, 43(7): 3349-3353.

[31] SAIDA D, KASHIWADA S, YAKABE M, et al. Sub-3 ns pulse with sub-100 μA switching of 1x-2x nm perpendicular MTJ for high-performance embedded STT-MRAM towards sub-20 nm CMOS[C]//Proceedings of the 2016 IEEE Symposium on VLSI Technology. New York: IEEE, 2016: 1-2.

[32] PATHAK S, CHA J, JO K, et al. Fast and efficient STT switching in MTJ using additional transient pulse current[J]. Applied Physics Letters, 2017, 110(23): 232401.

[33] HOFFMANN M, KRIZAKOVA V, KATEEL V, et al. Pulse-shaping strategies for efficient switching of magnetic tunnel junctions by spin-orbit torque[J]. Physical Review Applied, 2024, 22(3): 034052.