

混合时钟驱动自旋神经元器件激活特性和计算性能

袁佳卉 杨晓阔 张斌 陈亚博 钟军 危波 宋明旭 崔焕卿

Activation function and computing performance of spin neuron driven by magnetic field and strain

Yuan Jia-Hui Yang Xiao-Kuo Zhang Bin Chen Ya-Bo Zhong Jun Wei Bo Song Ming-Xu Cui Huan-Qing

引用信息 Citation: *Acta Physica Sinica*, 70, 207502 (2021) DOI: 10.7498/aps.70.20210611

在线阅读 View online: <https://doi.org/10.7498/aps.70.20210611>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于倾斜纳磁体翻转倾向性的与(或)逻辑门应力模型

Modeling of stress-regulated AND (OR) logic gate based on flipping preference of tilted nanomagnet

物理学报. 2019, 68(1): 017501 <https://doi.org/10.7498/aps.68.20181621>

动态突触、神经耦合与时间延迟对神经元发放的影响

Effects of dynamic synapses, neuronal coupling, and time delay on firing of neuron

物理学报. 2017, 66(20): 200201 <https://doi.org/10.7498/aps.66.200201>

变频正弦混沌神经网络及其应用

Frequency conversion sinusoidal chaotic neural network and its application

物理学报. 2017, 66(9): 090502 <https://doi.org/10.7498/aps.66.090502>

神经网络中局部同步引发的各种效应

Various effects induced by local synchronization in neural networks

物理学报. 2021, 70(17): 170502 <https://doi.org/10.7498/aps.70.20210142>

基于抑制性突触可塑性的神经元放电率自稳态机制

Neural firing rate homeostasis via inhibitory synaptic plasticity

物理学报. 2019, 68(7): 078701 <https://doi.org/10.7498/aps.68.20182234>

石墨烯沟道全自旋逻辑器件开关特性

Switching characteristics of all-spin logic devices based on graphene interconnects

物理学报. 2017, 66(20): 208501 <https://doi.org/10.7498/aps.66.208501>

混合时钟驱动自旋神经元器件激活特性和计算性能*

袁佳卉¹⁾ 杨晓阔^{1)†} 张斌¹⁾ 陈亚博²⁾ 钟军³⁾
危波¹⁾ 宋明旭¹⁾ 崔焕卿¹⁾

1) (空军工程大学基础部, 西安 710051)

2) (国防科技大学计算机学院, 长沙 410005)

3) (空军指挥学院, 北京 100097)

(2021年4月1日收到; 2021年6月16日收到修改稿)

自旋神经元是一种新兴的人工神经形态器件, 其具有超低功耗、强非线性、高集成度和存算一体等优点, 是构建新一代神经网络的强有力候选者. 本文提出了一种磁场辅助磁弹应变驱动的混合时钟自旋神经元, 利用 OOMMF 微磁学仿真软件建立了该神经元器件的微磁学模型, 基于 LLG 方程建立了其数值仿真模型, 利用所设计的自旋神经元构建了 3 层神经网络, 研究了不同纳磁体材料 (Terfenol-D, FeGa, Ni) 神经元器件的激活特性及其对 MNIST 手写数字数据集识别性能的影响. OOMMF 仿真和数值模拟发现, 设计的混合时钟结构能够成功驱动纳磁体发生随机磁化翻转, 有效模拟生物神经元的激活行为和特性. MNIST 手写数字识别结果表明: 当输入不同范围的磁场使得 3 种材料的自旋神经元都达到饱和识别精度时, 该自旋神经元器件具有与 Sigmoid 神经元器件相同的识别能力, 有望替代传统的 CMOS 神经元, 并且选择合适的磁致伸缩层材料能够进一步降低智能计算的整体功耗; 当输入相同范围的磁场时, Ni 构成的自旋神经元的识别速度较慢. 研究结果可为新型人工神经网络和智能电路的设计及应用奠定一定的理论基础.

关键词: 纳磁体, 自旋神经元, 磁化翻转, 神经网络计算

PACS: 75.78.-n, 85.75.-d, 87.19.ll

DOI: 10.7498/aps.70.20210611

1 引言

纵观人类发展历史, 人类一直致力于创造出拥有像人类智慧的“智能机器”^[1]. 研究人员采用互补金属氧化物半导体 (complementary metal-oxide-semiconductor, CMOS) 器件来搭建神经形态电路. Bernabe 等^[2]利用 CMOS 电路模拟了 Fitz-Hugh-Nagumo 神经元模型, 实现了神经振荡器的仿真; Lont 等^[3]实现了在 3 μm CMOS 制造工艺下的具有非线性突触的多层感知器, 并且该芯片被

成功集成到计算机中实现了相应功能. 尽管基于 CMOS 的神经形态器件已经逐渐发展成熟, 并且可以制造出成熟的仿生电路, 但是它需要消耗大量的能量, 且集成度较低, 无法实现有效的仿生功能^[4]. 因此, 不适合进行类脑计算. 传统的冯诺依曼计算体系也是制约 CMOS 神经形态芯片发展的主要障碍. 这是由于冯诺依曼计算体系存储与计算分离, 当数据在存储器和处理器之间传输时, 将会消耗大量的时间和能量, 无法实现实时语音、图像识别等运算^[5].

近几年, 随着材料科学和电子科学的发展, 一

* 国家自然科学基金 (批准号: 11975311) 和陕西省自然科学基金基础研究计划项目 (批准号: 2021JM-221, 2020JQ-470) 资助的课题.

† 通信作者. E-mail: yangxk0123@163.com

些新型材料被发现能够直接模拟人脑中神经元和神经突触的计算能力, 如铁电材料^[6]、阻变材料^[7]、相变材料^[8]和自旋电子材料^[9]等. 这些新型器件打破了传统的冯诺依曼计算瓶颈, 具有存算一体, 功耗极低等优点, 能够实现生物神经元或者突触的神经态计算功能^[10].

自旋电子材料作为制作新型神经形态器件的材料之一, 不但具备低功耗、高可靠性等优点, 而且容易制备, 成为学术界和产业界的重点关注对象. 2019年, Yue等^[11]设计了一种基于磁畴壁运动的自旋忆阻器, 并与CMOS器件共同使用, 构建了一种脉冲神经网络, 实现了实时动态学习; Fukami等^[12]利用36个自旋突触器件成功搭建了基于Hopfield模型的人工神经网络. 除自旋突触器件外, 自旋神经元器件也受到了广泛关注. 如基于磁畴壁运动的自旋神经元^[13]、基于自旋矩纳米振荡器的自旋神经元^[14]和基于随机翻转行为的自旋神经元等^[15].

其中由于基于磁化随机翻转行为的自旋神经元器件具有与人脑神经元的随机性相似^[16], 制备过程中具有较高的误差容忍度^[17], 能够进一步降低功耗^[18]等优点, 成为研究人员关注的重点. 日本Fukushima等^[19]提出在自旋转移力矩 (spin transfer torque, STT) 和磁场作用下, 磁化能够发生随机翻转, 产生一系列随机信号; 美国Roy团队^[18]提出基于自旋轨道力矩 (spin orbit torque, SOT) 的自旋神经元器件, 利用SOT驱动磁化翻转至空态, 撤去电流后, 磁化翻转至任一稳态, 实现了随机翻转. 相比利用STT驱动磁化翻转, 它的能耗更低, 集成度更高; Ostwal等^[20]提出SOT和磁场相互作用下的自旋神经元器件并且成功模拟了神经元的激活特性. 可以看出目前自旋神经元激活方式基本采用的是电流驱动时钟, 耗能较大, 相比基于CMOS的神经元器件失去了能耗优势^[18-20].

此外, 基于随机翻转特性的自旋神经元器件与逻辑器件已经成熟的制备工艺相似, 有望将自旋神经元用于实际应用中^[9]. 但由于自旋器件结构性能对其计算功能会产生一定的影响^[21,22], 因此需要从器件的结构尺寸及制作材料等方面出发, 筛选出性能相对更佳的器件. Kurenkov等^[23]提出了一种反铁磁/铁磁异质结, 其利用SOT时钟调控磁化翻转, 通过调节器件尺寸实现了突触和神经元的关键性功能; 曾中明课题组^[24]研究了不同结构的磁性

隧道结 (magnetic tunnel junction, MTJ) 中磁畴壁钉扎作用的影响, 实现了多阻态变化的自旋突触器件; 游龙课题组^[25,26]通过调控磁畴壁中固定层的状态, 利用SOT驱动Ta/CoFeB/MgO异质结中磁矩发生变化; 王开友课题组^[27,28]提出利用层间交换偏置作用, 在所设计的材料体系中实现了无外部磁场辅助的电流诱导的确定性磁化翻转. 除此之外, 2019年, 该课题组^[29]利用电流脉冲驱动方式成功模拟了突触尖峰放电时间依赖可塑性. 这些研究都是基于调控材料结构的方式来实现自旋神经态器件特性的, 关于材料选择方面的研究还较少, 而选择合适的器件材料将直接影响到器件的性能, 并且对自旋神经器件以及其神经网络计算能力的影响目前也还需要进一步探索.

基于此, 本文提出了一种磁场辅助磁弹应变的自旋神经元结构, 与前述的电流驱动的自旋神经元不同, 所设计器件引入了能耗较低的磁弹时钟^[30,31], 实现了电压驱动的自旋神经元器件, 有效降低了能耗. 利用OOMMF (object oriented micromagnetic framework) 微磁学仿真软件构建了器件的微磁学模型验证了所设计的自旋神经元的结构可行性. 随后构建了器件在热噪声环境下的数值模型, 仿真结果表明所设计的器件能够有效工作. 最后利用所提出的自旋神经元构建了三层神经网络用于识别手写数字, 验证了器件的神经网络计算能力. 除此之外, 本文研究了磁致伸缩层材料 (Terfenol-D, FeGa, Ni) 对基于随机磁化翻转特性的自旋神经元激活特性及其神经网络计算能力的影响.

2 神经元结构和磁化动力学模拟

2.1 磁场辅助磁弹应变驱动的神元结构

本文设计的自旋神经元结构如图1所示. 主要分为3部分, 由下到上分别是硅基底、厚度为100 nm的压电层 $\text{Pb}[\text{Zr}_x\text{Ti}_{1-x}]\text{O}_3$ (PZT)^[32] 以及磁致伸缩层. 其中磁致伸缩层是尺寸为 $120 \text{ nm} \times 100 \text{ nm} \times 8 \text{ nm}$ 的椭圆形纳磁体. 如图1所示的笛卡尔坐标系, 红色箭头代表纳磁体磁化方向, θ 是极角, 又称面外角, φ 是方位角, 又称面内角.

磁弹应变驱动纳磁体磁化状态改变的具体原理是: 当在PZT层上施加电压 V 时, 便会在层间形成电场. 因为PZT材料具有逆压电效应, 所以在电场的作用下会发生形变, 从而产生应力. 当

PZT 层上的形变传递到磁致伸缩层时, 由于磁致伸缩层具有逆磁致伸缩效应, 磁化状态会在应力的影响下发生改变. 具体过程是纳磁体磁化方向由稳定状态 (y 轴方向) 转到不稳定状态 (x 轴方向), 即待写入状态.

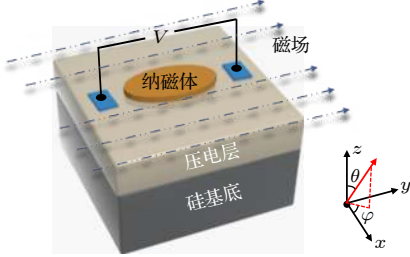


图 1 磁场辅助磁弹时钟神经元

Fig. 1. Magnetic field assisted strain-mediated neuron.

定义纳磁体磁化方向沿 y 轴负方向代表数值“0”, 沿 y 轴正方向代表数值“1”^[33], 设该自旋神经元结构中纳磁体初始磁化方向沿 y 轴负方向, 如图 2 所示. 由于实现信息传递, 纳磁体必须完成 0—1 转换. 但磁弹应变只能令磁化翻转 90° , 无法实现 0—1 转换. 因此, 在利用磁弹应变将磁化翻转 90° 后, 需要施加一定的弱磁场, 推动纳磁体磁化完成 180° 翻转.

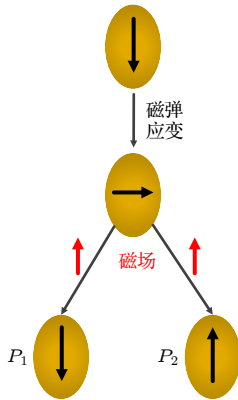


图 2 磁化翻转原理图

Fig. 2. Schematic of magnetization switching.

但是磁化翻转行为是随机的, 会出现 0° 翻转和 180° 翻转. 0° 翻转是指最终磁化方向和初始磁化方向相同, 180° 翻转是指最终磁化方向和初始磁化方向相反. 设施加磁场方向向上, 如图 2 中红色箭头所示, 则 0° 翻转的概率为 P_1 , 180° 翻转的概率为 P_2 . 这种随机翻转行为已经被证实能够模拟生物神经元的随机性行为^[34], 因此提出了磁场辅助磁弹应变驱动的自旋神经元.

2.2 磁化动力学模拟

2.2.1 器件的微磁学模型及仿真验证

为了证明本文设计自旋神经元结构的可行性, 利用 OOMMF 微磁学仿真软件建立器件的微磁学模型. 其中磁致伸缩材料选用 Terfenol-D, 各项参数^[35–37]如表 1 所示.

表 1 材料参数表
Table 1. Parameters of different materials.

参数	Terfenol-D	FeGa	Ni
杨氏模量 $Y/(10^{10} \text{ Pa})$	8.0	2.5	21.4
磁致伸缩系数 $\lambda_s/10^{-4}$	6.0	4.0	-0.2
吉尔伯特阻尼系数 α	0.100	0.100	0.045
回磁比 $\gamma/(10^5 \text{ rad}\cdot\text{s}^{-1}\cdot\text{T}^{-1})$	2.21	2.21	2.21
饱和磁化率 $M_s/(10^5 \text{ A}\cdot\text{m}^{-1})$	8.00	13.20	4.84
交换作用常数 $A/(10^{-11} \text{ J}\cdot\text{m}^{-1})$	0.90	1.60	1.05

图 3 展示了纳磁体在应力和磁场共同作用下的磁化过程. 该过程主要分为 3 个阶段, 第 1 阶段是在 0 ns 时, 纳磁体在 30 MPa 的应力作用下, 磁化方向由初始方向翻转 90° 至 x 轴负方向, 到达空态. 第 2 阶段是在 1 ns 时, 撤去应力, 施加 10 mT 磁场, 将继续翻转. 若磁场方向与初始磁化方向一致, 磁化向下翻转回到初态; 若方向相反, 磁化向上翻转达到另一稳态. 第 3 阶段是 2 ns 后, 撤去磁场, 纳磁体将始终位于稳定状态. 由于 OOMMF 软件仿真没有考虑热噪声的影响, 因此无法体现出该自旋神经元结构纳磁体随机翻转的特性. 但是通过施加不同方向的磁场, 对动态磁化过程图进行分析研究后, 发现纳磁体可以在磁场和应力共同作用下完成翻转.

通过 OOMMF 仿真, 还获得了该自旋神经元结构的磁化矢量变化曲线, 如图 4 所示. 结合纳磁体磁化矢量的参数表达式, 见 (5)—(7) 式, 可以对磁化过程做进一步分析. 由于 x 与 y 方向上的磁化矢量 m_x, m_y 与方位角 φ 的大小有关, 因此其值随方位角的改变在不断发生变化. 0 ns 时, 纳磁体磁化翻转至空态, 此时 m_y 由 -1 变化到 0 . 1 ns 时, 纳磁体在只有磁场的作用下发生翻转. 当磁场方向沿 y 轴正方向, m_y 从 0 变化到 $+1$; 当磁场方向沿 y 轴负方向, m_y 则由 0 变化至 -1 . m_x 与 m_y 的变化过程正交.

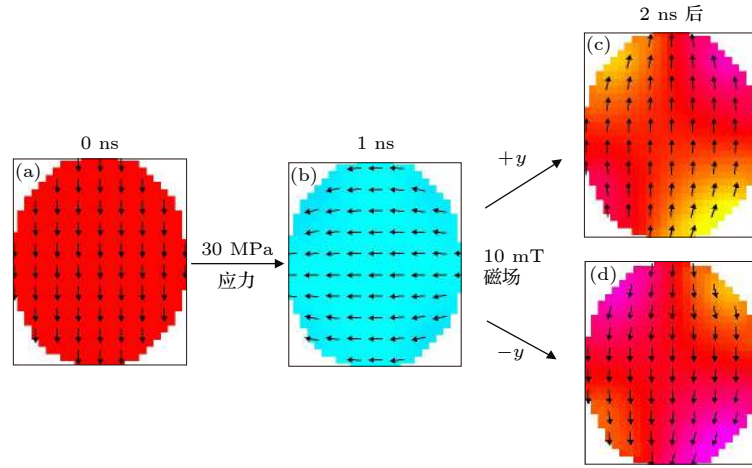


图3 动态磁化过程 (a) 磁化初始方向为 $-y$; (b) 施加 30 MPa 应力, 磁化翻转 90° ; 撤去应力, 施加 10 mT 磁场; (c) 方向相反, 磁化翻转 180° ; (d) 方向相同, 磁化翻转 0°

Fig. 3. Magnetization process: (a) Initial direction of magnetization is $-y$; (b) a 30 MPa strain is applied and then 90° switching is achieved, removing strain and applying a 10 mT magnetic field; (c) 180° magnetization switching; (d) 0° magnetization switching.

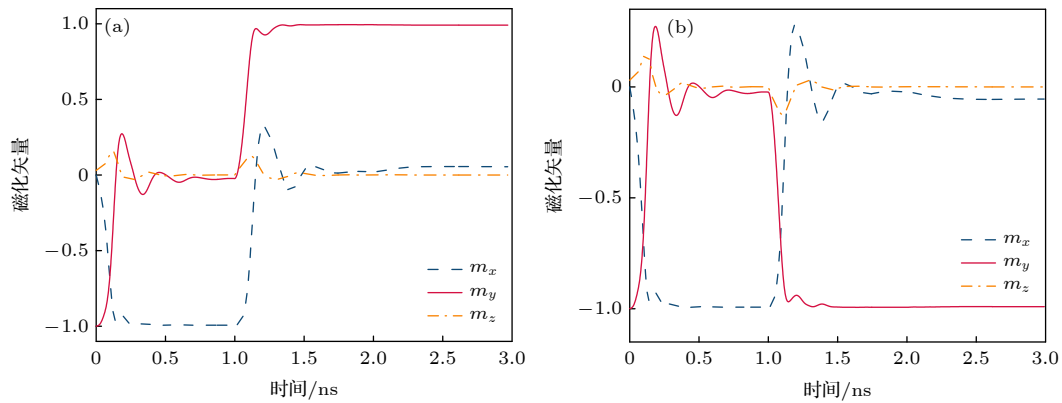


图4 磁化矢量随时间变化曲线图 (a) 1 ns 时, 施加磁场沿 y 轴正方向时的磁化矢量变化图; (b) 1 ns 时, 施加磁场沿 y 轴负方向的磁化矢量变化图

Fig. 4. Magnetization vector with time: (a) When $t = 1$ ns, a magnetic field is applied in the direction along $+y$; (b) when $t = 1$ ns, a magnetic field is applied in the direction along $-y$.

2.2.2 室温下器件的数值模型仿真

由于微磁学仿真没有考虑热噪声的影响, 因此还需建立该器件在热噪声影响下工作的数值模型, 从而验证该结构的实际工作效能。

磁场辅助磁弹应变作用到纳磁体上, 其磁化方向随时间变化的过程可以用 LLG(Landau-Lifshitz-

Gilbert) 方程来描述^[38]:

$$\frac{d\mathbf{M}}{dt} = -\gamma \mathbf{M} \times \mathbf{H}_{\text{eff}} - \frac{\alpha\gamma}{M_s} [\mathbf{M} \times (\mathbf{M} \times \mathbf{H}_{\text{eff}})], \quad (1)$$

其中 \mathbf{M} 是磁化矢量, γ 是回磁比, α 是吉尔伯特阻尼系数, M_s 是饱和磁化率, \mathbf{H}_{eff} 是有效场. 将 (1) 式简化为沿 x , y , z 三个方向上的标量方程:

$$(1 + \alpha^2) \frac{dm_x}{dt} = -\gamma (m_y h_z - m_z h_y) - \alpha\gamma [m_x (m_x h_x + m_y h_y + m_z h_z) - h_x], \quad (2)$$

$$(1 + \alpha^2) \frac{dm_y}{dt} = -\gamma (m_z h_x - m_x h_z) - \alpha\gamma [m_y (m_x h_x + m_y h_y + m_z h_z) - h_y], \quad (3)$$

$$(1 + \alpha^2) \frac{dm_z}{dt} = -\gamma (m_x h_y - m_y h_x) - \alpha\gamma [m_z (m_x h_x + m_y h_y + m_z h_z) - h_z], \quad (4)$$

其中, m_x, m_y, m_z 可以用参数方程表示为

$$m_x = \sin \theta \cos \varphi, \quad (5)$$

$$m_y = \sin \theta \sin \varphi, \quad (6)$$

$$m_z = \cos \theta, \quad (7)$$

式中, h_x, h_y, h_z 是 \mathbf{H}_{eff} 在 x, y, z 三个方向上的有效场分量:

$$h_x = -M_s N_{dx} m_x + \sqrt{\frac{2\alpha k T f}{\gamma \mu_0 M_s V_1}} G(0, 1)(t), \quad (8)$$

$$h_y = -M_s N_{dy} m_y + \frac{3\lambda_s}{M_s \mu_0} \sigma m_y + \sqrt{\frac{2\alpha k T f}{\gamma \mu_0 M_s V_1}} G(0, 1)(t) + h_{B_y}, \quad (9)$$

$$h_z = -M_s N_{dz} m_z + \sqrt{\frac{2\alpha k T f}{\gamma \mu_0 M_s V_1}} G(0, 1)(t), \quad (10)$$

其中, μ_0 是真空磁导率, 值为 $4\pi \times 10^{-7}$; V_1 是单个纳磁体的体积; k 是玻尔兹曼常数, 值为 1.38×10^{-23} J/K; T 是室温下的温度, 值为 300 K; f 是热噪声频率, 设为 1 GHz; $G(0, 1)(t)$ 是正态分布的随机热场. N_{dx}, N_{dy}, N_{dz} 是退磁因子, 表达式为^[39]

$$N_{dx} = \frac{\pi}{4} \left(\frac{th}{a} \right) \left[1 + \frac{5}{4} \left(\frac{a-b}{a} \right) + \frac{21}{16} \left(\frac{a-b}{a} \right)^2 \right], \quad (11)$$

$$N_{dy} = \frac{\pi}{4} \left(\frac{th}{a} \right) \left[1 - \frac{1}{4} \left(\frac{a-b}{a} \right) - \frac{3}{16} \left(\frac{a-b}{a} \right)^2 \right], \quad (12)$$

$$N_{dz} = 1 - \frac{\pi}{4} \left(\frac{th}{a} \right) \left[2 + \left(\frac{a-b}{a} \right) + \frac{18}{16} \left(\frac{a-b}{a} \right)^2 \right], \quad (13)$$

其中 a 是纳磁体长轴长度, b 是纳磁体短轴长度, th 是纳磁体厚度.

将各表达式和材料相关参数, 代入磁化标量方程 (2)—(4) 后, 利用经典四阶 Runge-Kutta 算法求解, 即可得到磁化状态的动态改变情况. 其中材料参数如表 1 所示.

选取 Terfenol-D 作为磁致伸缩层材料, 利用 MATLAB 软件, 对该结构数值模型进行仿真. 整个仿真过程为前 3 ns 施加 40 MPa 的应力, 3—6 ns 间撤去应力, 施加 10 mT 的磁场, 方向沿 y 轴正方向, 6 ns 后将磁场也撤去. 仿真后得到该结构随机磁化翻转的动态过程, 如图 5 所示. 图 5(a) 展示了

纳磁体完成 180° 磁化翻转的过程, 图 5(b) 是纳磁体没有完成 180° 翻转, 回到初始状态. 这一过程说明本文设计的自旋神经元结构在常温下能够实现随机翻转, 进一步验证了该器件结构的正确性.

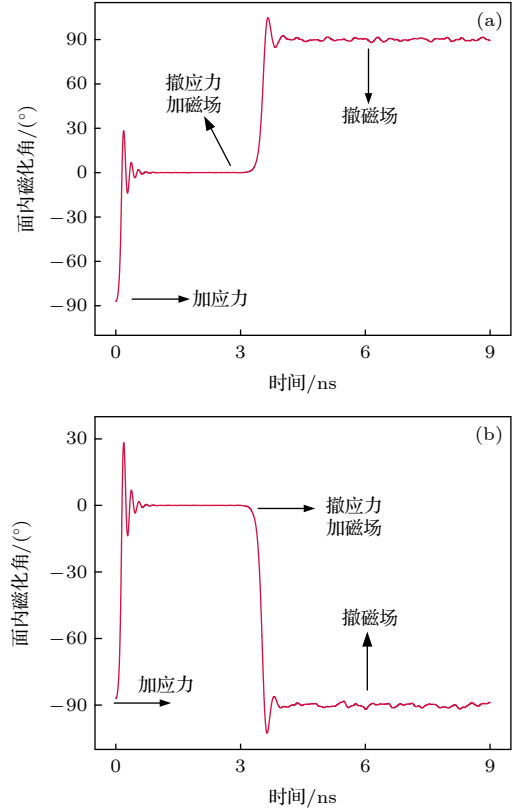


图 5 室温下随机磁化翻转的动态磁化过程 (a) 180° 磁化翻转; (b) 0° 磁化翻转

Fig. 5. Magnetization dynamics at room temperature: (a) 180° magnetization switching; (b) 0° magnetization switching.

3 基于自旋神经元的神经网络计算

3.1 不同材料自旋神经元的激活特性曲线

自旋神经元的翻转概率与输入量之间存在一定的输入输出关系, 为了获得本文神经元结构的 180° 翻转概率与输入磁场之间的关系, 以 Terfenol-D 自旋神经元为例展开研究. 首先施加 40 MPa 应力, 然后在 0—50 mT 之间随机选取 20 个不同的磁场值输入, 并计算这 20 个输入磁场下纳磁体 180° 磁化翻转概率, 如图 6(a) 中绿点所示, 而施加负磁场时的 180° 翻转概率可利用纳磁体的对称性分析得出^[35]. 即施加大小相同、方向相反的磁场, 实现 180° 翻转的概率之和为 1. 若已知施加正磁场时的 180° 翻转概率 P_2 , 则施加负磁场时的 180° 翻

转概率为 $1-P_2$. 由于该翻转概率函数可以用 Sigmoid 函数 $y = 1/[1 + \exp(-\alpha x)]$ 来拟合^[9], 并且当 $\alpha = 0.13$ 时, 图 6(a) 中橙色的拟合函数曲线与翻转概率数据拟合程度最高. 因此, 所得拟合函数 $y = 1/[1 + \exp(-0.13x)]$ 便是 Terfenol-D 神经元器件的激活函数. 通过分析, 当输入磁场增大时, 纳磁体 180° 翻转概率也随之增大; 当输入磁场达到约 20 mT 时, 概率接近饱和.

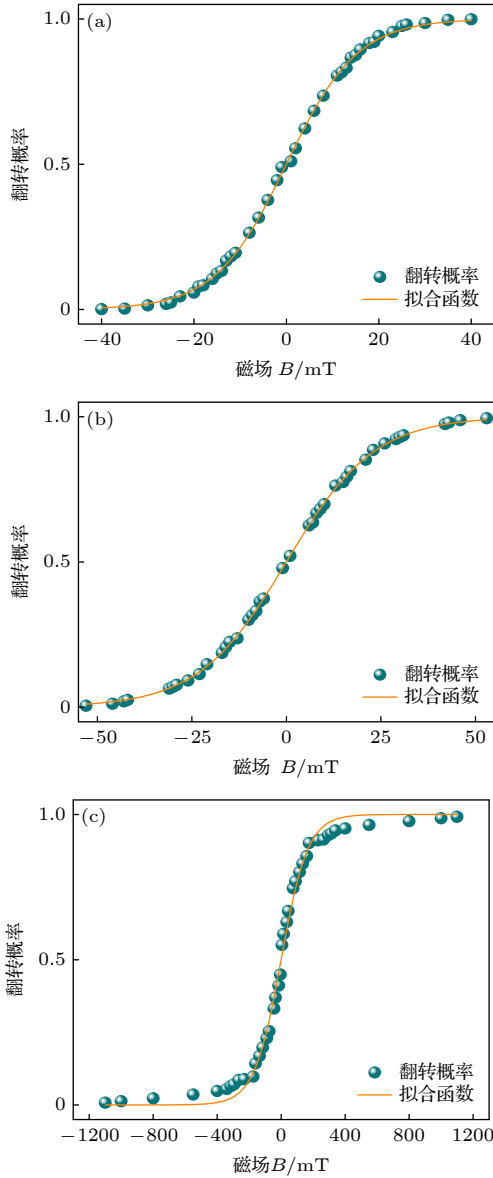


图 6 180° 磁化翻转概率与输入磁场的关系, 磁致伸缩层材料为 (a) Terfenol-D, (b) FeGa, (c) Ni

Fig. 6. 180° magnetization switching probability versus magnetic field, the magnetostrictive layer material is (a) Terfenol-D, (b) FeGa, (c) Ni.

利用这种方式, 对不同磁致伸缩材料下自旋神经元的激活特性进行进一步研究. 将 Terfenol-D

换成同样具有磁致伸缩效应的 FeGa 和 Ni. 当磁致伸缩材料为 FeGa 时, 保持其他条件不变, 施加应力 50 MPa, 磁场从 0 到 60 mT 之间选取 20 组数据, 可以得到 Sigmoid 关系式 $y = 1/[1 + \exp(-0.09x)]$, 如图 6(b) 所示. 当输入磁场为 30 mT 时, 翻转概率接近饱和. 当磁致伸缩层材料为 Ni 时, 仍然保持其他条件不变, 施加大小为 90 MPa 的应力, 磁场由 0 到 1200 mT 之间选取 20 个值, 得到拟合函数关系式为 $y = 1/[1 + \exp(-0.012x)]$, 如图 6(c) 所示. 当输入磁场为 175 mT 时, 概率接近饱和.

上述研究中, 不同磁致伸缩材料的自旋神经元需要施加的应力不同, 这由 (14) 式决定^[35]:

$$\sigma_{y_min} = \frac{|N_{dx} - N_{dy}|}{3\lambda_s} \mu_0 M_s^2, \quad (14)$$

式中 σ_{y_min} 是克服纳磁体形状各向异性能量势垒所要施加的最小应力. N_{dx} , N_{dy} , N_{dz} 由纳磁体长轴长度、短轴长度和厚度决定. 因为均选用规格为 $120 \text{ nm} \times 100 \text{ nm} \times 8 \text{ nm}$ 的纳磁体, 所以 N_{dx} , N_{dy} , N_{dz} 相同. 于是决定 σ_{y_min} 的是材料的磁致伸缩系数 λ_s 和饱和磁化率 M_s . 因为 Terfenol-D, FeGa 和 Ni 三种材料的 λ_s 和 M_s 各不相同, 如表 1 所示, 所以施加的应力也不同. 在磁弹时钟的作用下, 纳磁体磁化方向翻转至空态, 自旋神经元处于一个待写入状态. 此时只需要施加一个弱偏置磁场, 就能驱动磁化随机翻转, 实现信息写入. 利用本文设计的磁场辅助磁弹应变的自旋神经元结构, 可成功得到了 3 种磁致伸缩材料的激活特性曲线.

3.2 材料对自旋神经元计算功能的影响分析

本文利用 Liyanagedera 等^[40] 设计的神经元读取电路, 将纳磁体磁化翻转概率转化为确定性输出电压值读出. 具体过程为: 首先将本文设计的器件结构作为 MTJ 的自由层, 当磁场辅助磁弹应变驱动纳磁体磁化随机翻转时, MTJ 会在高电阻和低电阻两个状态间跳变. 然后给 MTJ 施加恒定电流, 则其两端的电压值将随电阻的变化而变化. 在一定时间内对输出电压值求平均, 则 MTJ 的平均输出电压将和纳磁体 180° 磁化翻转概率一致, 于是平均输出电压和输入磁场大小之间的关系与图 6 中 S 型激活函数相同. 将平均输出电压先转换为电流, 再转换为磁场, 可以继续施加到下一层神经元中^[20], 实现神经元间的信息传递.

设计如图 7 所示的人工神经网络识别 MNIST (mixed national institute of standards and technology database) 手写数字数据集, 以进一步研究材料对自旋神经元神经形态计算能力的影响. MNIST 数据集中每个数字图片为 28×28 像素, 共 784 个像素点, 因此设置神经网络输入层为 784 个节点. 隐藏层设为 300 个节点. 输出层设置为 10 个节点, 分别对应 0 到 9 十个数字. 由于手写数字图片像素是介于 0—255 之间的灰度值, 因此为了将图像输入到神经元中, 需要将灰度值 0—255 线性转换为文中图 6 所示的神经元激活函数非线性区域对应的输入磁场范围. 这里只考虑输入正磁场的情况. 由图 6 知, Terfenol-D, FeGa 和 Ni 对应的输入磁场范围分别是 0—20 mT, 0—30 mT 和 0—175 mT. 传统的 Sigmoid 函数 $y = 1/[1 + \exp(-x)]$ 也采用了同样的设置, 取其非线性区域对应的数值范围 ($0 < x < 5$) 输入. 设置神经网络训练次数为 300, 可得到如图 8 所示的识别精度图. 由图 8 分析知, 对于 Terfenol-D, FeGa 和 Ni 这 3 种材料组成的自旋神经元, 当训练次数分别为 36, 57 和 61 次时, 神经网络达到饱和识别率 98%, 传统 Sigmoid 神经元在 215 次时达到饱和识别精度 98%. 可以看出, 在神经网络训练初期, 本文设计的自旋神经元构成的神经网络的图像识别精度比传统的 Sigmoid 人工神经元高. 这是由于在神经网络计算前对输入磁场进行了一次线性转换, 这种方式有利于提高神经网络的识别速度, 并且更快地达到饱和识别率 [41]. 结果表明本文设计的自旋神经元器件搭建的神经网络能够实现与传统 Sigmoid 神经元器件相同的识别能力.

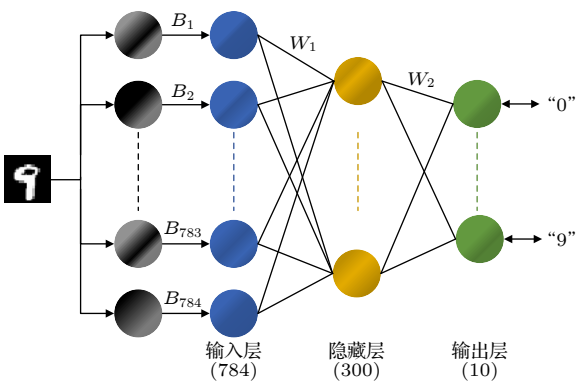


图 7 基于磁场辅助磁弹时钟的自旋神经元的三层神经网络结构示意图

Fig. 7. Three-layer neural network based on magnetic field + strain spin neurons.

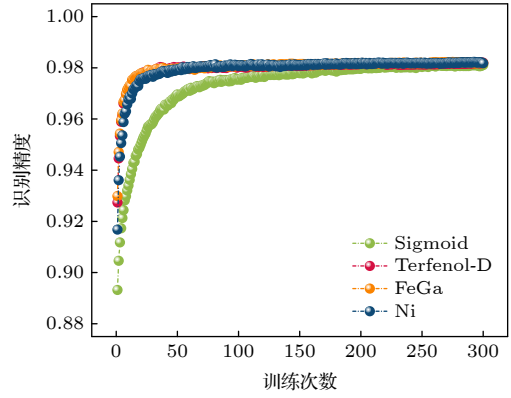


图 8 不同材料神经器件的识别精度

Fig. 8. Recognition rate obtained from the neural network with different materials.

除此之外, 该自旋神经元结构由于引入了低能耗的磁弹应变时钟, 理论上相比其他电流驱动的神经元结构具有能耗优势. 具体计算过程是: 该神经元结构功耗主要来源于磁场和磁弹应变时钟的能量损耗. 其中磁弹应变时钟的能量损耗 E_{strain} 的计算公式为 [42]

$$E_{\text{strain}} = C_{\text{pzt}} V^2, \quad (15)$$

C_{pzt} 是 PZT 层的寄生电容, 计算公式为 [43]

$$C_{\text{pzt}} = (h/w) \varepsilon_0 \varepsilon_r t_{\text{pzt}}, \quad (16)$$

式中 h 和 w 分别是该神经元器件的高度和宽度, h/w 近似为 1; t_{pzt} 是 PZT 层的厚度, 值为 100 nm; 真空介电常数 ε_0 为 8.85×10^{-12} F/m, PZT 层的相对介电常数 ε_r 为 1000. V 是施加在 PZT 层上的电压, 计算公式为 [44]

$$V = \sigma t_{\text{pzt}} / Y d_{31}, \quad (17)$$

其中 σ 是施加在 PZT 层上的电压产生的应力, Y 是磁致伸缩层材料的杨氏模量, $d_{31} \approx -10^{-10}$ m/V. 磁场时钟产生的能量损耗公式为 [45]

$$E_{\text{mag}} = I^2 R t_{\text{clock}}, \quad (18)$$

由毕奥-萨伐尔定律可推导出 $I = 2\pi r B / \mu_0$, 其中 r 设置为 50 nm. B 为该器件结构达到饱和识别率时所需的磁场. 导线电阻 $R = \rho L / S_d$, 其中铂的电阻率 $\rho = 2.22 \times 10^{-7}$ $\Omega \cdot \text{m}$, 铂线的长度 $L = 1$ μm , 铂导线的横截面积 $S_d = 10^{-14}$ m^2 , 施加磁场的的时间 $t_{\text{clock}} = 3$ ns.

通过 (18) 式可估算得出 Terfenol-D, FeGa 和 Ni 多铁神经元的能耗分别为 1.67 pJ, 3.75 pJ 和 12.75 nJ, 而利用电流驱动的自旋神经元器件功耗 [20] 比 Terfenol-D 和 FeGa 多铁神经元的能耗高

了近3个数量级. 这是由材料的磁致伸缩系数决定的, 由表1知, Terfenol-D磁致伸缩系数是FeGa的1.5倍, 是Ni磁致伸缩系数的30倍, 其中Terfenol-D的磁致伸缩系数最小, 因此其功耗最低. 所以, 当该器件结构选择合适的磁致伸缩层材料时, 能够大幅度降低功耗.

上文讨论了输入不同范围的磁场使得3种材料都达到饱和识别精度的情况, 接下来研究当输入磁场范围相同时, 3种材料的神经网络计算能力. 设低功耗输入时的磁场范围为0—20 mT, 神经网络训练次数设为2000次, 训练完毕后, 可得如图9所示的3种材料的识别率对比图. 从图中可以看出, 三种材料的自旋神经元都能达到饱和识别精度98%. 但Terfenol-D和FeGa神经元的识别速度明显高于Ni的识别速度. 这是由于Terfenol-D, FeGa和Ni激活函数非线性区域对应的磁场范围为0—20 mT, 0—30 mT和0—175 mT, 当输入磁场范围为0—20 mT时, 对应Ni激活函数的非线性区域太小, 导致其并非标准的非线性区域. 因此低功耗输入时, Ni的识别速度不如Terfenol-D和FeGa.

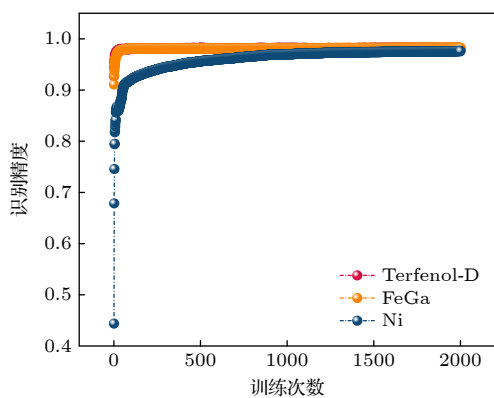


图9 输入磁场为0—15 mT时, 不同材料神经元器件的识别精度

Fig. 9. When the magnetic field is 0—15 mT, the recognition rate obtained from the neural network with different materials.

通过上述对比研究, 可以得到当输入不同范围的磁场使得 Terfenol-D, FeGa 和 Ni 都达到饱和识别精度时, Terfenol-D 自旋神经元的功耗相比 FeGa 和 Ni 较低; 当低功耗输入相同的磁场时, Terfenol-D 材料相识别速度较快. 因此 Terfenol-D 更适合作为该自旋神经元的磁致伸缩层材料. 这也为今后选择多铁神经元材料提供了参考标准, 即

选择磁致伸缩系数较大的材料更利于进行神经网络计算.

4 结 论

本文通过微磁学建模和数值仿真方法, 证明了磁场辅助磁弹应变能够成功驱动纳磁体实现随机磁化翻转, 从而模拟生物神经元的激活行为. 利用该自旋神经元器件, 构建了3层人工神经网络识别 MNIST 手写数字数据集, 并且研究了磁致伸缩材料 (Terfenol-D, FeGa, Ni) 对其激活特性和神经网络计算能力的影响, 结果发现: 当输入磁场对应3种材料激活特性曲线的非线性区域时, 识别率高达98%, 能够达到与传统的 CMOS 神经元相同的识别精度, 并且3种材料中 Terfenol-D 多铁神经元性能最优; 而当低功耗输入时, Ni 构成的自旋神经元识别速度较慢. 这些结论为新型人工神经网络和智能电路的应用奠定了理论基础并提供了重要指导.

参考文献

- [1] Aleksander I 2004 *Nature* **432** 18
- [2] Linares-Barranco B, Sanchez-Sinencio E, Rodriguez-Vazquez A, Huertas J L 1991 *IEEE J. Solid-State Circuits* **26** 956
- [3] Lont J B, Guggenbuhl W 1992 *IEEE Trans. Neural Networks* **3** 457
- [4] Chen Y R, Li H, Chen Y Z, Chen F, Li S C, Liu C C, Wen W J, Wu C P, Yan B N 2018 *AI-View* **2** 46 (in Chinese) [陈怡然, 李海, 陈逸中, 陈凡, 李思成, 刘晨晨, 闻武杰, 吴春鹏, 燕博南 2018 *人工智能* **2** 46]
- [5] Yang R, Terabe K, Yao Y P, Tsuruoka T, Hasegawa T, Gimzewski J K, Aono M 2013 *Nanotechnology* **24** 384003
- [6] Chen C, Yang M, Liu S, Liu T, Zhu K, Zhao Y, Wang H, Huang Q, Huang R 2019 *Symposium on VLSI Technology* (Kyoto: IEEE) p136
- [7] Liu D Q, Cheng H F, Zhu X, Wang N N, Zhang C Y 2014 *Acta Phys. Sin.* **63** 187301 (in Chinese) [刘东青, 程海峰, 朱玄, 王楠楠, 张朝阳 2014 *物理学报* **63** 187301]
- [8] Tuma T, Pantazi A, Gallo M L, Sebastian A, Eleftheriou E 2016 *Nat. Nanotechnol.* **11** 693
- [9] Cai J L, Fang B, Zhang L, Lv W X, Zhang B S, Zhou T J, Finocchio G, Zeng Z M 2019 *Phys. Rev. Appl.* **11** 034015
- [10] Zhu J D, Zhang T, Yang Y C, Huang R 2020 *Appl. Phys. Rev.* **7** 011312
- [11] Yue K, Liu Y Z, Lake R K, Parker A C 2019 *Sci. Adv.* **5** eaau8170
- [12] Fukami S, Ohno H 2018 *J. Appl. Phys.* **124** 151904
- [13] Sengupta A, Choday S H, Y Kim, Roy K 2015 *Appl. Phys. Lett.* **106** 143701
- [14] Fulara H, Zahedinejad M, Khymyn R, Dvornik M, Fukami S, Kanai S, Ohno H, Akerman J 2020 *Nat. Commun.* **11** 4006
- [15] Dong I, Yoon G, Sik H, Park, Wanjun 2015 *J. Appl. Phys.*

117 17D714

- [16] Vincent A F, Jerome L, Locatelli N, Nesrine B R, Bichler O, Gamrat C, Zhao W S, Klein J O, Galdin-Retailleau S, Querlioz D 2015 *IEEE T. Biomed. Circ. S* **9** 166
- [17] Chen Y B, Song M X, Wei B, Yang X K, Cui H Q, Liu J H, Li C 2020 *IEEE Magn. Lett.* **11** 4504505
- [18] Kim Y, Fong X, Roy K 2015 *IEEE Magn. Lett.* **6** 3001004
- [19] Fukushima A, Seki T, Yakushiji K, Kubota H, Imamura H, Yuasa S, Ando K 2014 *Appl. Phys. Express* **7** 083001
- [20] Ostwal V, Debashis P, Faria R, Chen Z H, Appenzeller J 2018 *Sci. Rep.* **8** 16689
- [21] Yang X K, Cai L, Zhang B, Cui H Q, Zhang M L 2015 *J. Magn. Magn. Mater.* **394** 391
- [22] Carlton D B, Emley N C, Tuchfeldand E, Bokor J 2008 *Nano Lett.* **8** 4173
- [23] Kurenkov A, DuttaGupta S, Zhang C H, Fukami S, Horio Y, Ohno H 2019 *Adv. Mater.* **31** 1900636
- [24] Cai J L, Fang B, Wang C, Zeng Z M 2017 *Appl. Phys. Lett.* **111** 182410
- [25] Zhang S, Luo S J, Xu N, Zou Q M, Song M, Yun J J, Luo Q, Guo Z, Li R F, Tian W C, Li X, Zhou H G, Chen H M, Zhang Y, Yang X F, Jiang W J, Shen K, Hong J M, Yuan Z, Xi L, Xia K, Salahuddin S, Diény B, You L 2019 *Adv. Electron. Mater.* **5** 1800782
- [26] Zhang S, Su Y, Li X, Li R, Tian W, Hong J, You L 2019 *Appl. Phys. Lett.* **114** 042401
- [27] Sheng Y, Edmonds K W, Ma X Q, Zheng H Z, Wang K Y 2018 *Adv. Electron. Mater.* **4** 1800224
- [28] Cao Y, Rushforth A W, Sheng Y, Zheng H Z, Wang K Y 2019 *Adv. Funct. Mater.* **29** 1808104
- [29] Wang Z W, Yang Y C, Cai Y M, Zhu T, Cong Y, Wang Z H, Huang R 2019 *Bulletin of National Natural Science Foundation of China* **33** 656 (in Chinese) [王宗巍, 杨玉超, 蔡一茂, 朱涛, 丛杨, 王志衡, 黄如 2019 中国科学基金 **33** 656]
- [30] Liu J H, Yang X K, Cui H Q, Wei B, Li C, Chen Y B, Zhang M L, Li C, Dong D N 2019 *J. Magn. Magn. Mater.* **491** 165607
- [31] Ma J, Hu J M, Li Z, Nan C W 2011 *Adv. Mater.* **23** 1062
- [32] Yang N N, Chen X, Wang Y J 2018 *Acta Phys. Sin.* **67** 157508 (in Chinese) [杨娜娜, 陈轩, 汪尧进 2018 物理学报 **67** 157508]
- [33] Cowburn R P, Welland M E 2000 *Science* **287** 1466
- [34] Locatelli N, Cros V, Grollier J 2013 *Nat. Mater.* **13** 11
- [35] Chen Y B, Wei B, Yang X K, Liu J H, Cui H Q, Li C, Song M X 2020 *J. Magn. Magn. Mater.* **514** 167216
- [36] Li X, Carka D, Liang C Y, Sepulveda A E, Keller S M, Amiri P K, Carman G P, Lynch C S 2015 *J. Appl. Phys.* **118** 014101
- [37] Wang Q W, Zhang J J, Ma T Y, Yan M 2009 *Rare. Metal. Mat. Eng.* **38** 1234 (in Chinese) [王庆伟, 张晶晶, 马天宇, 严密 2009 稀有金属材料与工程 **38** 1234]
- [38] Bertotti G, Serpico C, Mayergoyz I D 2009 *Nonlinear Magnetization Dynamics in Nanosystems* (Oxford: Elsevier) pp401–445
- [39] Beleggia M, Graef M D, Millev Y T, Goode D A, Rowlands G 2005 *J. Phys. D. Appl. Phys.* **38** 3333
- [40] Liyanagedera C M, Sengupta A, Jaiswal A, Roy K 2017 *Phys. Rev. Appl.* **8** 064017
- [41] Glorot X, Bengio Y 2010 *J. Mach. Learn. Res.* **9** 249
- [42] Fashami M S, Atulasimha J, Bandyopadhyay S 2012 *Nanotechnology* **23** 105201
- [43] Vacca M, Graziano M, Crescenzo L D, Chiolerio A, Lamberti A, Balma D, Canavese G, Celegato F, Enrico E, Tiberto P, Boarino L, Zamboni M 2014 *IEEE Trans. Nanotechnol.* **13** 963
- [44] Liu J H, Yang X K, Zhang M L, Wei B, Li C, Dong D N, Li C 2018 *IEEE Electron Device Lett.* **40** 220
- [45] Das J, Alam S M, Bhanja S 2011 *IEEE J. Emerg. Sel. Top. Circuits Syst.* **1** 267

Activation function and computing performance of spin neuron driven by magnetic field and strain^{*}

Yuan Jia-Hui¹⁾ Yang Xiao-Kuo^{1)†} Zhang Bin¹⁾ Chen Ya-Bo²⁾
Zhong Jun³⁾ Wei Bo¹⁾ Song Ming-Xu¹⁾ Cui Huan-Qing¹⁾

1) (*Fundamentals Department, Air Force Engineering University, Xi'an 710051, China*)

2) (*College of Computer, National University of Defense, Changsha 410005, China*)

3) (*Airforce Command College, Beijing 100097, China*)

(Received 1 April 2021; revised manuscript received 16 June 2021)

Abstract

The spin neuron is an emerging artificial neural device which has many advantages such as ultra-low power consumption, strong nonlinearity, and high integration. Besides, it has ability to remember and calculate at the same time. So it is seen as a suitable and excellent candidate for the new generation of neural network. In this paper, a spin neuron driven by magnetic field and strain is proposed. The micromagnetic model of the device is realized by using the OOMMF micromagnetic simulation software, and the numerical model of the device is also established by using the LLG equation. More importantly, a three-layer neural network is composed of spin neurons constructed respectively using three materials (Terfenol-D, FeGa, Ni). It is used to study the activation functions and the ability to recognize the MNIST handwritten datasets. Results show that the spin neuron can successfully achieve the random magnetization switching to simulate the activation behavior of the biological neuron. Moreover, the results show that if the ranges of the inputting magnetic fields are different, the three materials' neurons can all reach the saturation accuracy. It is expected to replace the traditional CMOS neuron. And the overall power consumption of intelligent computing can be further reduced by using appropriate materials. If we input the magnetic fields in the same range, the recognition speed of the spin neuron made of Ni is the slowest in the three materials. The results can establish a theoretical foundation for the design and the applications of the new artificial neural networks and the intelligent circuits.

Keywords: nanomagnet, spin neuron, magnetization switching, neural network computing

PACS: 75.78.-n, 85.75.-d, 87.19.lj

DOI: 10.7498/aps.70.20210611

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 11975311) and the Natural Science Basic Research Program of Shaanxi, China (Grant Nos. 2021JM-221, 2020JQ-470).

[†] Corresponding author. E-mail: yangxk0123@163.com