使用磁性随机突触的机器学习

Matthew O. A. Ellis, [1, a)] Alexander Welbourne, [2, a)] Stephan J. Kyle, [2] Paul W. Fry, [3] Dan A. Allwood, [2] Thomas J. Hayward, [2] 和 Eleni Vasilaki[1]

[1)]*Department of Computer Science, University of Sheffield, Sheffield, S1 4DP, United Kingdom*

[2)]*Department of Materials Science and Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom*

[3)]*Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom*

（日期：2023年3月3日）

人工神经网络的出色性能是以高能耗和 CO_2排放为代价实现的。非常规计算架构以磁性系统为候选方案，有望成为替代性的节能硬件，但在实现过程中仍面临随机行为等挑战。在此，我们提出一种方法，用于利用纳米线中磁畴壁运动所产生的传统上有害的随机效应。我们展示了功能性二元随机突触，以及一种梯度学习规则，使其能够接受训练，并适用于一系列随机系统。该规则利用神经元输出分布的均值和方差，根据每个突触的测量次数，在突触随机性与能源效率之间取得权衡。对于单次测量，该规则产生具有最小随机性的二元突触，以牺牲潜在性能换取稳健性。对于多次测量，突触分布较为宽广，更接近性能更佳的连续突触。这一观察结果使我们能够根据所需性能、器件的运行速度和能耗来选择设计原则。我们在物理硬件上验证了其性能，结果表明该性能与标准神经网络相当。

arXiv:2303.01886v1 [cs.ET] 3 Mar 2023

引言

人工智能（AI）作为现代生活一部分的迅猛发展带来了诸多优势。然而，随着人工智能程序日益复杂，其能耗也不断增加[1,2]；如今，训练一个最先进的自然语言处理模型所需的能耗，已与一名普通美国公民从出生到成年的能耗相当[3]。多种非传统计算架构旨在降低这种能耗，其中包括非互补金属氧化物半导体技术[4–7]。然而，要使非互补金属氧化物半导体技术具备竞争力的性能，就必须克服互补金属氧化物半导体经过多年发展所形成的潜在优势。

在生物神经网络中，突触被认为是全或无型或分级型且具有非确定性，这与人工网络中建模的完全模拟突触不同[8]。受生物学启发，一些方法考虑了由二元突触和神经元构成的网络，认为二元运算更易于计算，因此能耗更低[9–12]。然而，尽管这些二值化神经网络对噪声的鲁棒性更强，其性能却低于模拟版本。相较之下，具有随机突触的网络可为概率模型提供采样机制[13]，并能与模拟网络一较高下，但代价是采样时间较长[14–19]。要以更少的样本获得更高的性能，就需要采用经过调整的训练方法；而在实现方面，则需要能够原生地（以较低能耗）提供所需随机性的硬件，作为非常规计算的潜在途径之一，磁性架构能够原生地（以较低能耗）提供所需的随机性。长期以来，人们一直期待磁性架构能在计算逻辑领域发挥作用；数据存储市场的发展也使该领域备受关注[6, 7, 20–26]。磁性元件的非易失性天然适用于数据存储，而超低功耗控制机制（例如自旋极化电流或施加应变）[27,28] 则为高能效存内逻辑计算提供了实现途径。持续不断的发展已展示了如何操纵磁畴，以同时移动和处理数据[22, 29–31]。然而，与互补金属氧化物半导体计算中较低的错误容忍度相比，磁畴壁逻辑受随机效应的限制尤为明显[32,33]。

在此，我们提出一种方法：不再试图消除随机效应，而是将其作为计算架构的关键组成部分。作为概念验证，我们展示了如何将纳米线用作随机磁性突触，并通过复用一个硬件突触来实现手写数字识别。

我们开发了一种学习规则，通过考虑突触分布，能够有效训练由这类“有噪声”的突触构成的人工神经网络。假设我们允许通过单次测量来识别突触状态。在这种情况下，学习规则会调整其参数，即传播畴壁所用的场，以降低突触随机性。如果允许多次测量，梯度规则就会找到能够形成宽广突触分布的参数，从而模拟连续突触并提升性能。若没有随机性，这种运算将

a)这些作者对本项工作作出了同等贡献 .

仅限于二元运算，无法达到模拟突触的分辨能力。有了随机性，我们便拥有了一个灵活的系统，可在快速运行时间近似与长运行时间高性能之间进行调节。尽管噪声环境较强或变化多端，我们的学习规则仍能高效训练网络；这不同于其他采用基于平均场的学习规则进行计算的随机神经网络方案[14,16,19]。在此，纳入网络方差使训练能够在低采样区间找到更优解，从而在运行速度／能耗与测试准确率之间实现权衡。

我们通过将训练权重迁移到一个采用此类硬件突触的网络中，实验验证了模型的性能；实验结果与模拟网络的性能高度一致。我们的观察结果为设计提供了一个框架，使我们能够针对给定的期望精度确定所需的测量次数（以及相应的能耗），反之亦然。

这项工作开创了利用基于自旋电子学的逻辑低能耗优势的前景[5–7, 34]。具体而言，它使基于磁畴壁的纳米线器件得以应用[24,31,35,36]，并将迄今为止被视为障碍的噪声运行[32,33]转化为高性能随机机器学习范式的基础。

结果

硬件随机突触

我们提出的基本计算单元是一种基于铁磁性纳米线的二元随机突触，该纳米线具有两种有利的磁化方向。磁化方向不同的区域之间的边界称为磁畴壁（DW）。虽然磁畴壁有不同的形态，但在此处它们呈现出具有周期性磁化纹理的“涡旋”图案。我们的突触是一根宽 400 nm、厚 54 nm 的坡莫合金纳米线，其长度中部刻有凹口，以形成一个人工缺陷位点。图 1.a 显示了该系统的扫描电子显微镜图像，插图放大显示了凹口。通过在一条金电流线上施加电压脉冲，在导线左侧（以蓝色伪彩色显示）产生磁畴壁；该电流线以橙色伪彩色显示。

该系统作为随机突触的工作原理如图 1.b 所示。通过在导线上施加电流脉冲，可以将涡旋磁畴壁[37]注入纳米线中。这相当于向突触输入 1；若不注入磁畴壁，则相当于输入 0。施加磁场可使磁畴壁沿纳米线传播。若传播场足够高，磁畴壁不会钉扎在缺陷位点，而是能够传播至纳米线末端，从而产生输出 1。若传播场较低，磁畴壁则会钉扎在缺口处，从而产生输出 0。对于场的中间值，行为会变得随机，但钉扎概率明确。我们可以将场控制视为对二元突触中权重的控制，并将检测纳米线右侧是否存在磁畴壁作为突触的输出。

随着传播场的调节，磁畴壁通过的概率会发生变化。图1.c展示了使用聚焦磁光克尔效应（FMOKE）测得的通过概率随传播场变化的关系。通过概率呈现类似 sigmoid 的变化特征，橙色虚线表示使用逻辑斯蒂S形函数 $f(h_{ij})$ 拟合的结果（见方法部分）。

因此，二元随机突触由以下方式确定：

$$w_{ij} = \begin{cases} 1 & \text{with probability } f(h_{ij}) \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

其中，$f(h_{ij})$ 是磁畴壁通过概率函数，h_{ij}是连接输入神经元 j 与输出神经元 i 的突触的传播场。通过这一定义，我们的突触完全是兴奋性的，这对应于磁畴壁被钉扎或未被钉扎的物理表示，而不是取值为 $\{-1, 1\}$ 的互补二元方案；后者无法由该物理系统自然地表示。

与二元突触相比，具有模拟或分级突触的神经网络由于状态范围更广，往往表现更好[38,39]。在这里，我们采用了类似随机计算的方案，即使用一系列二元测量或样本的均值来表示一个值。因此，我们允许进行 $K \geq 1$ 次测量以确定突触的状态，并将等效平均权重表示为

$$\bar{w}_{ij} = \frac{1}{K} \sum_k w_{ij}^{(k)}, \tag{2}$$

其中，K 是采样总次数，上标（k）表示根据公式1对突触权重进行的单次采样。平均突触有$1 + K$ 种状态，例如，对于 $K = 1$，两种状态为0和1；对于 $K = 2$，状态则为0、0.5和1。因此，对于 $K \to \infty$，$\bar{w}_{ij}$ 相当于一个呈S形、范围限定在0到1之间的连续突触。图1.d展示了平均权重随采样次数变化的示例，其中我们绘制了 $K = 1$（紫色方块）、4（蓝色菱形）和128（绿色圆形）对应的公式2。每个示例都是通过对 w_{ij} 进行指定次数的采样计算得到的，采样时使用随机选取的固定 h_{ij}。每种情况下都只有离散的取值，但当 $K = 128$ 时，采样足以提供近乎连续的表示。

通过这种方式，我们提出的二元随机突触可用于构建神经网络，在获取多个样本后，该网络将趋近于有界模拟网络。在物理层面上，这是通过反复运行硬件设备来累积平均值实现的。

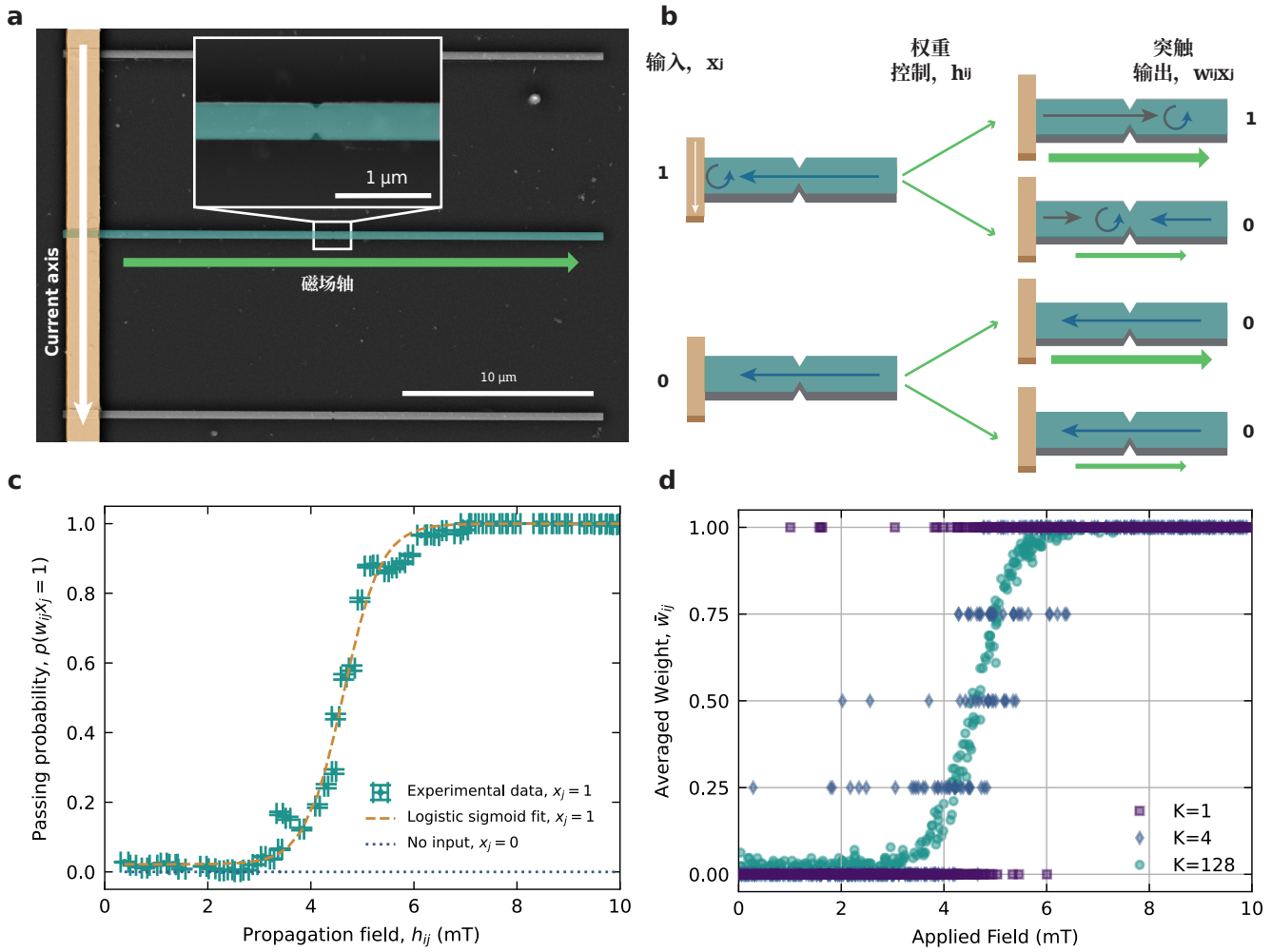

图 1。**作为随机突触的带缺口坡莫合金纳米线（NW）的表征。a,** 坡莫合金纳米线（蓝色）和电流注入线（橙色）的伪彩色扫描电子显微镜图像。插图显示了人工缺口的细节。图中标出了场（绿色）和电流（白色）的轴。**b,** 随机突触工作原理示意图。电流线允许输入（x_j）为 1（电流脉冲，注入磁畴壁）或 0（无脉冲，无磁畴壁）。与纳米线同向的场驱动磁畴壁（若存在）穿过系统：高场使磁畴壁通过缺口并产生 1 的输出，低场则使缺口阻挡磁畴壁并产生 0 的输出。中间场（未显示）提供通过缺口的中间概率。**c,** 实验测得的注入磁畴壁通过缺口的概率。调节传播场可通过类似逻辑 sigmoid 函数的方式，在整个范围内控制该概率。数据点为 1000 个样本的平均值，x 误差条表示传播场选择的精度，y 误差条给出为 $p(1-p)/\sqrt{1000}$。逻辑 sigmoid 拟合方法见方法部分。无输入时（无注入，$x_j = 0$）的成核场为（10.74 ± 0.07）mT。因此，在 10 mT 以下，无输入的通过概率为零。**d,** 平均突触权重，如公式 2 所定义。根据样本数，即 b 中操作的重复次数，有效突触权重可从纯二元（一轮重复/样本，$K = 1$）变化到近乎连续（$K = 128$ 个样本）。

随机网络

我们将这些突触嵌入人工神经网络中，其中神经元 i 的输出为

$$y_i = \sum_j \bar{w}_{ij} x_j, \tag{3}$$

其中 j 是输入维度上的索引。

我们将该网络训练为一个分类器，用于解决一个包含 C 个类别的问题，其中有 C 个彼此独立的神经元（感知器），每个神经元代表一个类别。此任务基于广为人知的 MNIST 数据集，但将每张图像下采样至 14×14 像素，而非标准的 28×28 像素。这样做是为了缩短在原型实验硬件上运行时的运算时间（见方法部分）。在图 2.a 中，我们描绘了对应的感知器

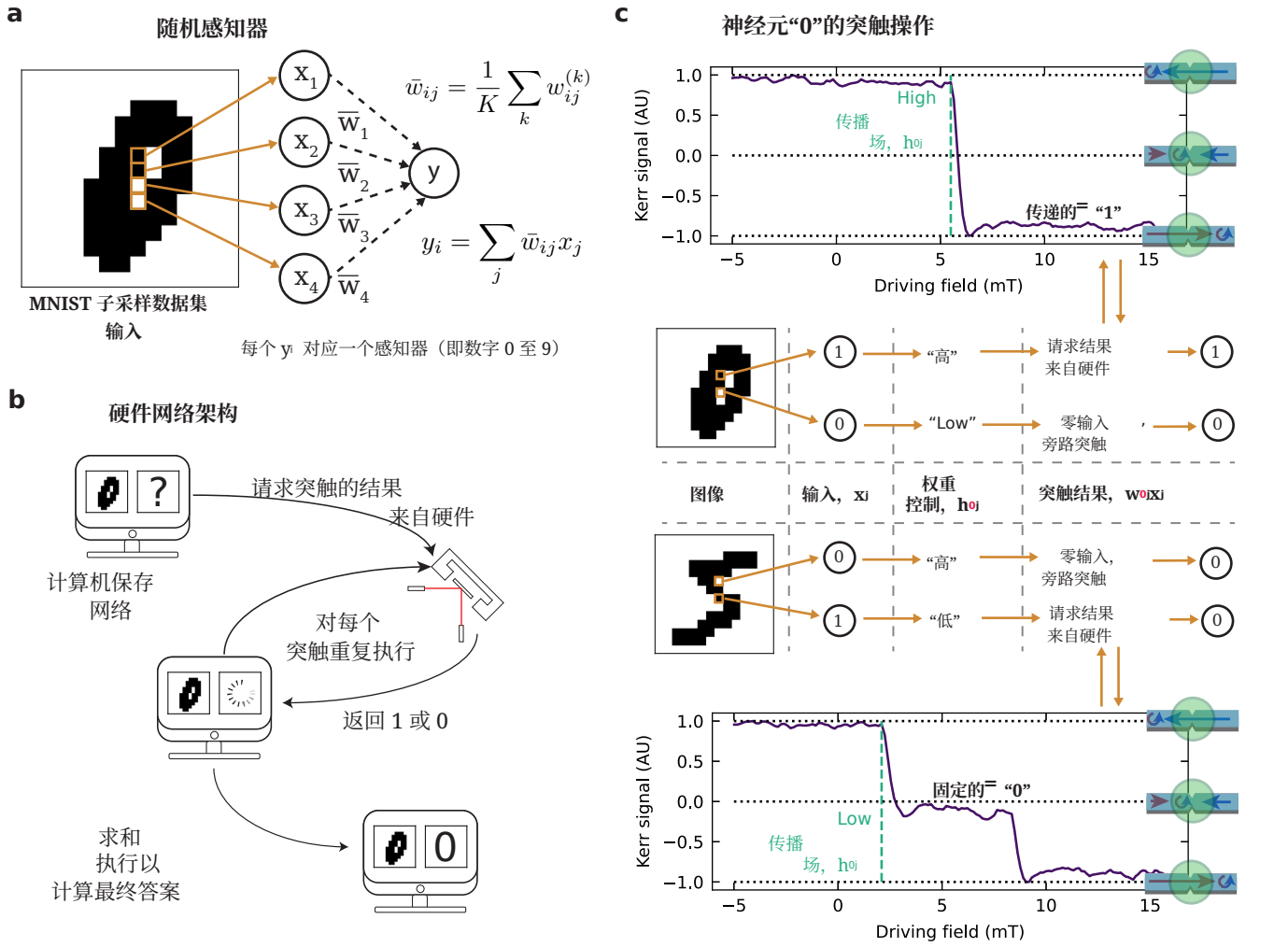

图2。**随机网络运行**。**a**，随机感知器示意图。图像中的每个输入值通过一个平均权重馈送至每个类别对应的神经元。在这里，对加权输入进行求和，以得到神经元的活动值（如公式3所示）。对于MNIST任务，10个类别（数字“0”到“9”；y_0 至 y_9）各对应一个神经元。训练完成后，与正确输入对应的类别神经元（此处为 y_0）应具有最高的活动值。网络中的每个平均权重（$\bar{w}_{ij}$）都是对具有单个权重 w_{ij} 的突触输出进行多次测量（$K \geq 1$）后得到的均值，该单个权重由训练得到的传播场设定（见公式1和2）。需要明确指出的是，这些权重不同于传统神经网络中的权重：它们具有随机性，并且会随网络每次运行而变化。单个权重以概率f（h_{ij}）（磁畴壁通过概率，如图1.c所表征）取值“1”，否则取值“0”。因此，平均权重的取值服从图1.d所示的分布。**b**，硬件网络的架构。为展示成功的性能，仅将随机突触直接运行在硬件上。感知器存储在计算机中，计算机针对给定的突触参数（训练得到的传播场，h_{ij}）向磁性随机突触请求结果（“1”或“0”）。对每个突触重复此过程后，通过求和来预测输入的正确类别。**c**，单个突触 *in materia* 对于神经元 y_0的理想化运行过程。对于两个输入（即像素），图中展示了正确类别（“0”）图像和错误类别（“5”）图像的路径。权重控制量，即传播场 h_{ij}，预计与正确类别图像中的像素相关：对于正确图像中处于“开启”状态的像素，权重控制量应较高；处于“关闭”状态时，权重控制量应较低。如果输入像素值为“0”，则突触被旁路，因为按照构造结果为“0”。但是，如果输入值为“1”，则使用相应的权重控制量从硬件请求结果。如上方图表所示，高传播场会使磁畴壁直接通过缺口（仅观察到一个步骤），这被解释为输出“1”。如下方图表所示，低传播场会产生一个包含两个步骤的过程：磁畴壁最初被钉扎在缺口处，随后在更高的驱动场下解钉扎。这被解释为“0”。在实际情况下，突触的结果会以随机方式变化，从而反映通过概率 f（h_{ij}）。

归为类别“0”。如果向神经元输入一个对应类别的代表样本（本例中为数字“0”的图像），神经元应产生较高的活动，以识别输入为零。

实验流程如图 2.b 所示。为便于演示，仅使用单个硬件突触，并按时间顺序串行执行操作。实际设备可能会有多个突触并行运行，并在测量过程中进行求和。感知器参数存储在计算机上，计算机将输入和突触参数发送到外部硬件突触，并请求结果。重复此过程，直到每个突触收集到 K 个样本（见公式 2）。结果由计算机求和，并额外加上一个偏置项。为避免冗余测量，省略了对应于输入“0”的像素（示例图像中的白色像素），因为按照设计，其输出必定为“0”。接收到黑色像素（$x_j = 1$）的突触会在场设为高值时输出“1”，或在场设为低值时输出“0”，见图 2.c。中间场值会产生随机变化的输出，反映通过概率 $f(h_{ij})$。

随机学习规则的分析

现在，我们简要推导应用于神经网络突触的学习规则。每个突触 $w_{ij}^{(k)}$ 都是从伯努利分布中独立抽取的一个样本，因此这些样本的总和将服从泊松-二项分布。每个输出神经元的均值 μ_i 和方差 σ_i^2（由公式 3 计算）如下：

$$\mu_i = \sum_j f(h_{ij}) x_j, \tag{4}$$

$$\sigma_i^2 = \frac{1}{K} \sum_j f(h_{ij}) x_j \left[1 - f(h_{ij}) x_j\right]. \tag{5}$$

这些值的详细计算请参见补充材料。

由于输入数量较多，且采样过程意味着该总和涵盖大量事件，因此可以用高斯分布近似泊松-二项分布[40]。利用这一近似，可以对神经元输出重新参数化，使随机性仅出现在一个与可训练参数无关的项中。由此，我们写作

$$\tilde{y}_i = \mu_i + \sigma_i \xi_i, \tag{6}$$

其中 $\tilde{y}_i$ 表示神经元输出的近似值y_i，而 ξ_i 是一个来自均值为零、方差为一的高斯分布的样本。

如果我们假设处于监督学习框架中，并且 E 是我们希望最小化的误差函数（例如均方误差或交叉熵），那么 E 是我们呈现给网络的模式 p 的函数，该模式定义了期望的输出目标。 E也是输出神经元的函数，以向量 y表示，而它也取决于 p。学习规则将根据以下“在线”梯度规则，按 Δh_{ij} 更新施加到每个突触的场值h_{ij}：

$$\Delta h_{ij} = -\eta \frac{\partial E(\tilde{y})}{\partial \tilde{y}_i} \frac{\partial \tilde{y}_i}{\partial h_{ij}} \tag{7}$$

$$= -\eta \frac{\partial E(\tilde{y})}{\partial \tilde{y}_i} \left(\frac{\partial \mu_i}{\partial h_{ij}} + \frac{\partial \sigma_i}{\partial h_{ij}} \xi_i \right), \tag{8}$$

其中，η 是一个表示学习率的小正数。我们根据式 6、4、5 计算 $\frac{\partial \tilde{y}_i}{\partial h_{ij}}$ 的导数。我们还根据式 6 计算 ξ_i 的值，并根据式 4 和 5 计算 μ_i 和 σ_i，再根据式 3 计算 $\tilde{y}_i$（设定$\tilde{y}_i = y_i$）。由此可知，对于 $K \to \infty$， $\sigma \to 0$，我们得到一个“平均场”梯度规则，该规则考虑了输出神经元的均值，但未考虑其方差。

我们在下采样的 MNIST 数据集上测试了该规则的性能。训练期间，重复次数（样本数） K 被设为网络的一个参数，我们将其定义为 K_{train}，因此它会影响训练过程。输出的方差会对分类过程产生重要影响；如果方差较高，就更容易发生误分类，尤其是在各神经元均值相近的类别中。因此，在监督训练期间，网络力求将该方差降至最低。当 K 较低时，这通过改变权重来实现，并由磁场 h_{ij} 控制，使概率接近 1 或 0（施加的磁场较高或较低），因为这会使式 5 中的单样本方差最小化。由此得到的解几乎是确定性的二值网络。然而，当 K 较大时，方差会按因子 $1/K$ 降低，因此与 $K = 1$ 的情况相比，系统能够容忍更高的突触方差。这样便能找到一个伪模拟解。

图 3 描述了学习规则对网络突触的影响。我们绘制了 5 个独立模型在训练前（图 3.a）、使用$K_{\text{train}} = 1$样本训练后（图 3.b）以及使用$K_{\text{train}} = 128$ 样本训练后（图 3.c）所有神经元的传播场 h_{ij}分布。最终分布证实了理论预期： $K_{\text{train}} = 1$会形成二值网络（低方差），而 $K_{\text{train}} = 128$则近似于标准感知器，其突触权重呈连续分布（高方差）。

在图 3.e-f 中，我们展示了在上述三种训练情况下，反复输入同一图像时神经元输出的分布，并发现神经元分布反映了突触分布。现在我们考虑这样一种情况：在测试过程中，计算公式 2 时抽取的样本数量不同，

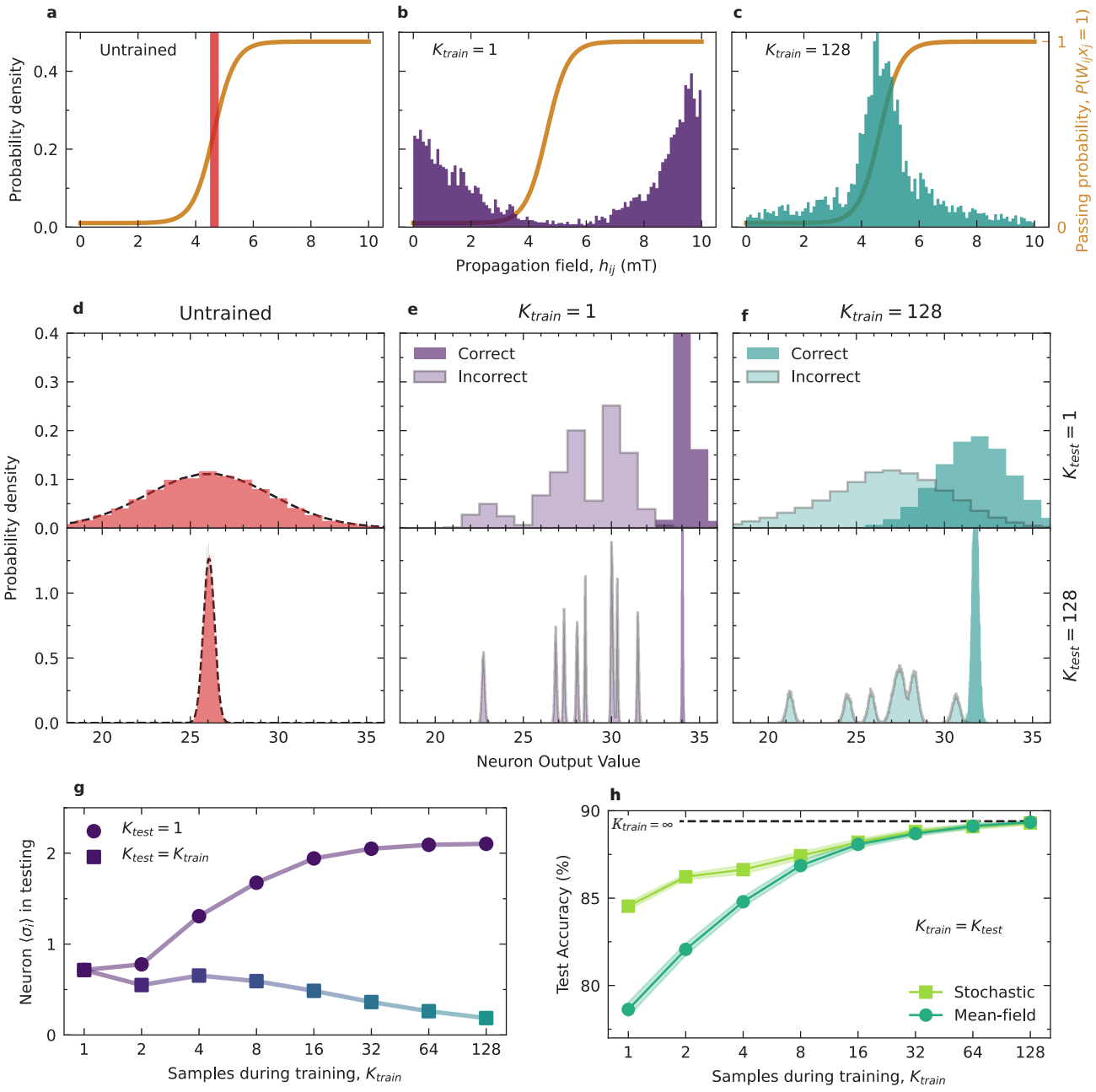

图 3. **随机学习规则分析**。**a-c**，5 个独立模型中突触磁场参数的概率密度直方图。**a** 显示训练前的分布，此时所有磁场均经过初始化，使通过概率（橙色表示，见右侧纵轴）为 0.5。**b** 和**c** 分别显示使用 1 个或 128 个样本训练时的分布。使用 1 个样本时，分布呈双峰，峰值对应的磁场通过概率接近 0 或 1；而使用 128 个样本训练时，分布集中在通过概率函数的中心区域。**d-f**，当向神经元独立展示零图像 10,000 次时，神经元值**y** 的分布；这些神经元分别用于识别正确分类（本例中输出 0）或错误分类（输出 1-9）。在**d** 中，模型尚未训练，因此所有输出具有相同的分布；在**e** 和**f** 中，使用 1 个和 128 个样本训练后，分布分别分为正确输出神经元和错误输出神经元两类。上排显示 $K_{\text{test}} = 1$，下排显示 $K_{\text{test}} = 128$。测试时使用更多样本可降低方差，从而降低误分类的可能性。**g**，随着训练所用样本数增加，对所有神经元的标准差取平均值，$K_{\text{test}} = 1$（圆圈）和 $K_{\text{test}} = K_{\text{train}}$（方块）。这概括了**d-f**中分布图得出的结论。训练时使用更多样本，可提高单个样本对应的标准差，同时降低所有样本的标准差。然而，使用 $K_{\text{test}} < K_{\text{train}}$ 进行测试会导致总体标准差增大。**h**，使用随机（深绿色圆圈）或平均场（浅绿色方块）学习规则时，测试集准确率随训练样本数的变化。各点表示 5 个独立训练模型的平均准确率，阴影区域表示 1 个标准差。样本数较少时，随机学习规则能保持较高的测试准确率。

我们将其定义为 K_{test}。顶行显示 $K_{\text{test}} = 1$时的分布，底行则显示 $K_{\text{test}} = 128$。未训练的神经元值在所有数据样本中呈现高斯分布，且场的初始化使方差达到可能的最大值；见图 3.d。训练后，当使用 $K_{\text{train}} = 1$ 或 $K_{\text{train}} = 128$和 $K_{\text{test}} = K_{\text{train}}$ 时，正确类别神经元与错误类别神经元的分布几乎没有重叠。然而，如果我们使用 $K_{\text{train}} = 128$训练网络后，再用 $K_{\text{test}} = 1$ 对其进行测试，则会出现相当显著的重叠（3.f，上方面板），这表明误分类的概率很高。在所有情况下，使用 $K_{\text{test}} = 128$进行测试（底行）时，方差都会降低 $1/128$（如公式 5 所示），从而能够更准确地分辨均值。在 $K_{\text{train}} = 128$ 的情况下，学习规则利用了额外的采样和方差降低，通过更充分地利用连续范围内的权重来提升性能。然而，当 $K_{\text{test}} < 128$ 时（如 3.f，上方面板所示），方差增加会降低正确分类的概率。在另一种训练情况下（$K_{\text{train}} = 1$，3.e，上方面板），学习规则会调整权重，以找到一种低方差、近乎确定性的二元解。正如预期的那样，测试时进一步采样（3.e，下方面板）会使方差进一步降低，但由于该解已针对较低采样率进行了优化，重叠程度并未显著改变。正如我们将要展示的，这在测试采样（K_{test}）较小时会带来更高的性能；但与较大的 K_{train} 情况相反，当测试采样增加时，性能上限较低。

图3.g比较了测试期间使用 $K_{\text{test}} = 1$ 样本（圆形）和 $K_{\text{test}} = K_{\text{train}}$样本（方形）时的平均方差与训练期间所用样本数量 K_{train}之间的关系。如前所述，使用1个样本进行训练时，通过使通过概率接近0或1，可以将方差保持在较低水平。然而，当训练期间使用更多样本时，单个样本的方差可能会增加，因为平均样本的方差会降低。

这种通过最小化方差来减少误分类的行为，源于“随机”学习规则中的方差项。其他仅考虑均值项的规则[14,19]在使用随机网络时，无法找到这些确定性解。图3.h显示了我们的随机学习规则（方形）与平均场规则（圆形）的测试准确率随K_{train} 样本数量变化的情况 $K_{\text{test}} = K_{\text{train}}$ ，结果取自五个独立训练的模型的平均值。对于两种规则，增加样本数量都会提高测试准确率，因为突触均值可能取值的级数会随之增加（见图1.d）。然而，在所有情况下，随机学习规则的表现都优于平均场规则；当使用大量样本（$K \geq 8$）进行训练和测试时，两者的表现趋于收敛。虚线表示完全平均场网络的表现，即实际上使用无限数量的样本（也就是连续但有界的突触），代表该网络在给定任务上可能达到的最高准确率。

硬件与运行原理。

现在，我们将展示神经网络在物理硬件上的运行情况，而不仅仅是在模拟中运行。图4.a和b显示了在突触运算经模拟（线条）和由硬件处理（点）时计算出的测试准确率；模型分别使用a、 $K_{\text{train}} = 1$ 或b、$K_{\text{train}} = 128$ 进行训练，并在样本数量递增的情况下进行测试（K_{test}）。由于原型器件的吞吐量有限，我们仅展示截至$K_{\text{test}} = 8$的实验结果。

仿真结果与硬件结果高度一致，并凸显了使用不同采样水平训练的模型所呈现的不同特性。在使用一次重复训练的情况下，网络是确定性的，因此在测试期间对更多样本进行平均并不会显著提升准确率。另一方面，尽管使用128次重复训练的模型在仅进行一次测试重复时性能较低，但随着测试期间样本数量的增加，准确率也会提高。这源于低采样水平下随机性的增加，以及高采样水平下由此带来的精度提升。相应的神经元分布（3.e）和（3.f）也印证了这一行为：使用一个样本进行训练并使用一个样本进行测试时，神经元方差远低于使用 $K = 128$样本进行训练并使用一个样本进行测试的情况。这类似于多数表决，其中分类器必须具有多样性才能提升性能（见[41] 及其中的参考文献）。在这里，当神经元分布具有高方差时，随着 K_{test} （投票者数量）的增加，性能提升幅度也会增加。

图4.c 可进一步探究多数表决行为。它展示了测试准确率随训练样本数和测试样本数变化的彩色图（使用现已验证的仿真模型）。测试时采用不同重复次数会导致性能变化，这揭示了速度与准确率之间一项关键的权衡。初步来看，结果符合随机计算的行为：随着时间推移，近似结果会变得越来越准确（若有需要）。额外重复次数平均也呈现出这一趋势，这意味着累积样本会增加时间和能耗，但也能提升准确率。然而，利用我们的学习规则实现低采样下的确定性解的能力，我们可以在低采样极限下超越朴素的随机计算推理（图3.h中也可见此现象）。若固定K_{train}，低重复次数下的性能与高重复次数下的最终准确率便会相互制衡：例如，若模型在高重复次数下训练，却在测试（推理）时使用较低的重复次数，准确率就会不理想。类似地，若模型在低重复次数下训练，却在测试时使用较多重复次数，最终准确率也会受影响。一种可能的做法是始终将 $K_{\text{test}} = K_{\text{train}}$绑定在一起，但这需要多组

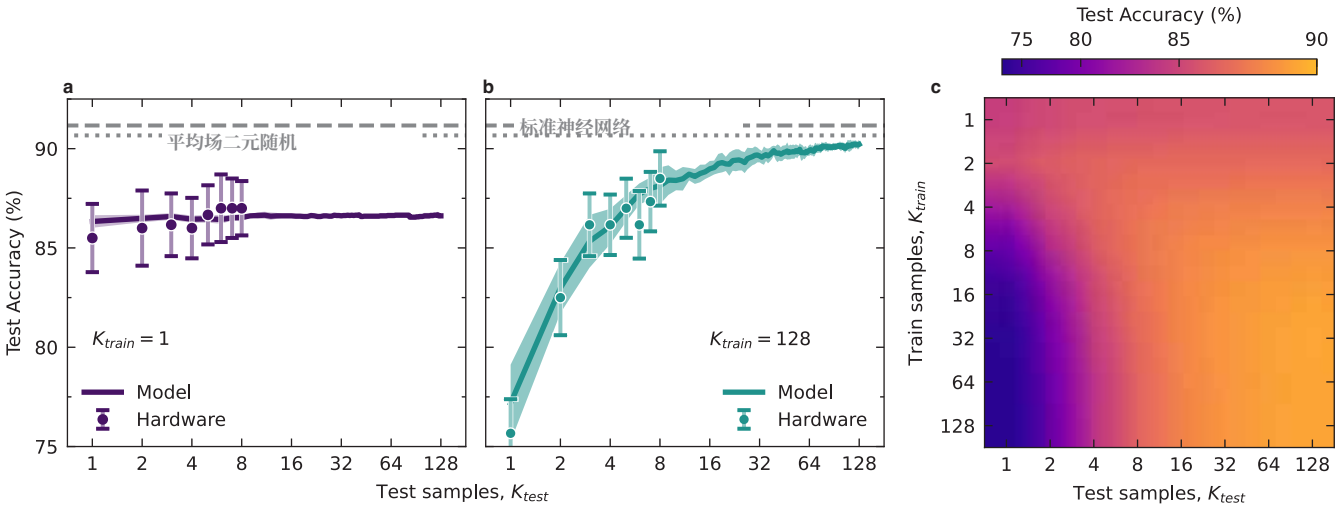

图 4. **硬件验证与采样选择。a、b** 使用物理硬件（数据点）或仿真（曲线）计算得到的测试准确率对比。测试数据集仅包含前 600 张图像，且各数字的数量大致相同。两种情况下，均使用模型进行训练；由于原型器件的吞吐量限制，硬件测试仅使用了 8 个样本。**a** 显示网络仅使用 1 个样本进行训练时的准确率，而**b** 显示使用 128 个样本进行训练时的准确率。如前所述，使用 1 个样本训练会得到近乎确定性的解，因此测试期间重复采样并不能提高准确率。使用 128 个样本训练时，随着测试期间使用的样本数增加，准确率也会提高，并达到更高的峰值性能（尽管初始基准较低）。虚线表示标准神经网络的性能，点划线表示完整平均场二元随机网络的性能。在这两种情况下，硬件性能都与模型计算结果高度吻合。模型准确率是对使用相同训练权重进行的 5 次独立测试取均值，阴影区域表示 1 个标准差。这可以用来表示网络固有随机性导致特定任务性能的变异程度。随着测试样本数增加，变异程度自然会降低；对于更具确定性的 $K_{\mathrm{train}} = 1$ 情况，变异程度也较低。硬件准确率来自对这 600 张图像进行的一次运行，因此误差条表示基于小批次估算准确率的标准误。**c** 子采样 MNIST 任务中，不同训练与测试采样组合对应的测试准确率（由模型测得）。数据采用双线性插值，可视为对数据集中具有不同采样率的各部分取平均。总体而言，测试时使用更多样本会提高准确率，但在 $K_{\mathrm{test}} > K_{\mathrm{train}}$时，这种提升会受到限制。具体而言，在 $K_{\mathrm{train}} = 1$ 情况下，由于权重分布具有确定性，进一步采样几乎不会带来改善。在所有测试情形中，使用 2 个样本训练都优于使用 1 个样本训练；但总体准确率最高的情况是训练和测试均使用 128 个样本。此类数据可为特定任务选择训练和测试样本数提供参考，以满足所需的准确率和运行时间要求。

训练权重。因此，可以利用图**4.c** 等数据作为指导，根据所需准确率和运行时间来训练和测试突触。在保持仅使用一组训练权重这一简洁性的同时（固定 K_{train}），可以选择一系列横向测试值，以达到所需的准确率和能耗范围。

分析 MNIST任务中训练与测试重复次数组合下的性能后，我们发现，在大多数情况下，测试时采用与训练时相近的重复次数，效果良好。一个显著的例外是：使用两个样本进行训练，在所有测试设置下（包括使用一个样本进行测试），性能始终优于使用一个样本进行训练。我们认为，这是因为与使用一个样本相比，使用两个样本时参数空间中的步长更小，因此能够找到更好的解，同时方差仍然很低，并且在使用一个样本进行测试时也保持较小。

讨论

神经形态器件为开发低能耗的机器学习系统提供了一条很有前景的途径，旨在克服传统神经网络的一项主要缺陷。随机二值神经网络因能耗较低且易于实现而展现出这一方面的潜力[9–13]。对这类网络进行多次采样，其性能便可媲美模拟网络[14–19]。然而，仍有一些亟待解决的问题：如何制定训练规则，使网络即使在低采样率下（此时计算速度更快、能耗更低）也能达到高性能；以及如何找到能够原生提供所需随机性的硬件实现。我们开发了一种适用于随机二值神经网络的学习方法，并通过实验加以验证：利用纳米线中磁畴壁的行为作为随机突触。传统上，随机性一直被视为纳米磁逻辑器件的一项局限

因素[32,33], , 但在这里，它却是驱动学习的功能要素。我们证明，该硬件网络的性能可与标准神经网络相媲美，并展示了借助新颖的学习规则，即使在低采样率下也能实现高性能。

实验中，我们观察到，注入带有人工钉扎位点的纳米线中的磁畴壁，可在外加磁场作用下受到随机钉扎，并可调节。随后，我们证明，这种可调的随机钉扎能够为神经网络器件构建突触。由于物理系统的特性，这些突触表现为二元随机突触。训练此类网络的关键在于采用一种学习规则，将网络随机输出的方差纳入考量。这种训练方法通过对网络输出进行多次采样（$K_{\mathrm{train}}/K_{\mathrm{test}}$），计算样本均值和偏差。采样数量较少时，所得结果趋向于确定性的二元解，计算速度较快，但性能不及采样数量较多的情况；后者可近似标准的“模拟”网络，但需要更多时间（和能量）。这种权衡使我们能够根据所需的性能或运行速度，灵活地设计网络。

关键在于，这里提出的学习规则考虑了输出中的随机部分，使我们能够找到一系列工作状态。其他二元随机计算方法，例如 Hirtzlin等[19], , 使用网络的期望值进行训练（我们称之为平均场，这等同于 $K \to \infty$）；当推理（测试）期间使用的样本较少时，这会导致准确率下降。Esser等[42] 也使用高斯近似，在 IBM TrueNorth 神经突触系统上训练具有二元随机突触的网络，但他们认为方差项的贡献可以忽略不计。我们的规则中方差项的贡献，使得训练出的权重在低采样量条件下，比平均场版本表现更好。

其他考虑方差的学习方法源自似然比框架[43–45]，与强化学习中的策略梯度方法[46]相关。虽然这些方法考虑了神经元和突触的随机性，但它们很大程度上依赖于损失函数基线值的选择，而这些基线值需要通过复杂的近似方法求得。此外，这里采用的重参数化方法能够将误差信号直接反馈给突触场参数，并且适用于现有的基于反向传播的学习方法。

总的来说，本文提出的随机学习规则在高、低采样区间均展现出可调性，并且可以轻松地集成到反向传播式代码中。由于考虑了输出的方差，该规则能够在低采样下的确定性二进制行为与高采样下的随机“类模拟”行为之间进行调节，使我们的系统能够灵活权衡运行速度/能耗与测试准确率。

我们在此展示的磁性畴壁突触是一个原理验证组件，因此有必要着眼于将其改进为更适合“量产”的神经形态器件所需的变革。优化后的器件可能会采用自旋转矩驱动的磁畴壁运动[31] ，并利用局部纳米磁性元件编码权重。我们也可以设想，将我们的学习方法应用于由具有类似随机特性的其他磁性元件构成的网络，例如磁隧道结等[12、14、47–49]。此外，磁畴壁器件已被用作神经元[50] 或激活函数[51] ；而磁性元件总体上也已应用于一系列其他低能耗计算方案[16、47、52–54] ，这些方案利用了磁性器件的随机性。我们的基础元件——磁性随机突触——可以用于此类计算范式，在其中，高效地产生随机比特至关重要。不过需要指出的是，本文的关键成果是展示了在实验硬件上运行所取得的性能，这得益于我们的随机学习规则。进一步优化则有待未来的研究和工程开发。

尽管本文展示的单层网络只能解决线性可分问题，但仍可通过多种方式加以扩展。在保留单层网络简洁性的同时，着眼于全磁架构，它在储备池计算领域具有潜在应用。在储备池计算中，固定储备池会对输入序列进行非线性时空变换，使输出表示具有线性可分性。储备池计算的优势在于，可以将储备池变换交由具有适当特性的物理系统来完成，近期人们对开发基于磁性（自旋电子学）的物理储备池计算产生了浓厚兴趣[7, 55–62]。我们有望将基于磁性畴壁的神经网络连接到这些储备池，从而构建完整的硬件储备池计算系统。此外，还可以采用更传统的途径，将当前方法扩展至多层网络，因为该学习规则与反向传播兼容。这一方向尚待探讨的一个研究问题是，采样过程应在网络的局部还是全局范围内进行。一种方法是利用纳米线互连和逻辑门实现多层结构；但若不局限于全磁架构，也可以设想采用混合磁性—互补金属氧化物半导体专用集成电路（如参考文献 63 所述），这或许能为构建更大规模的网络硬件提供途径。不过，这些实现方案的细节超出了本研究的范围。

总之，我们开发了一种二元随机突触训练方法，该方法在学习过程中考虑网络的随机性，并通过对随机输出进行重新采样，在器件运行时间与所需精度之间实现权衡。该方法已

在基于磁畴壁的随机突触概念验证中得到展示，在推理期间，硬件与模型表现出极佳的一致性。

方法

器件制造

器件采用 CSAR-62 光刻胶，通过两步电子束光刻工艺制备。第一步通过热蒸发沉积坡莫合金（$Ni_{81}Fe_{19}$），形成厚度为 54 nm 的纳米线（基底压力，7×10^{-7} mbar；工艺压力，$\sim 5 \times 10^{-5}$ mbar；速率，0.5 Å s^{-1}）。第二步以钛/金（通过热蒸发标称沉积厚度为 10nm/200nm）的形式沉积电流线和连接焊盘。样本通过银浆与 PCB 器件进行电气连接。

器件操作

器件的操作过程如图 2 所示。使用 AVTECH 脉冲发生器沿电流线（电阻 290 Ω）施加 30 伏、100 纳秒的脉冲。使用电磁铁沿导线长度方向施加磁场。使用美国国家仪器公司的数据采集卡控制两者之间的时序，使脉冲在重复的正弦磁场序列中的特定时刻触发。触发脉冲时的磁场即为传播场。实时校准时序可将触发信号与磁场序列之间因发热产生的任何漂移校正至 $\lesssim 0.1$mT。

使用聚焦磁光克尔效应磁力计（光斑尺寸为 ~ 5 微米）测量纳米线响应。将激光光斑置于缺口上方，获得磁滞回线。磁滞回线中的单个阶跃表示磁畴壁通过缺口（输出为 1）。双阶跃表示经历了两阶段钉扎/去钉扎过程（输出为 0）。一种算法方法可对每条磁滞回线进行自动评估。通过计算克尔信号微分后的峰值数量来进行判断；如果存在两个峰值，则表明磁畴壁曾被钉扎。为消除假阳性，与这些峰值对应的原始信号阶跃必须大于总信号变化的 24%。该阈值经过实验优化，即使激光光斑略微偏离中心（导致步长不等），也能检测到峰值，同时最大限度地减少由噪声引起的错误检测。

磁畴通过概率

研究发现，磁畴壁未被人工缺陷位点钉扎的概率呈现类似 sigmoid 函数的行为。该概率的函数形式被用于模拟磁性随机突触，以训练网络进行计算。我们使用以下方法拟合该概率：

$$f(h) = d + \frac{1-d}{1+\exp(-\Delta(h-h_0))}, \tag{9}$$

其中，$d = 0.0219$ 是低场下的有限通过概率，$h_0 = 4.63$ mT 是场中心，$\Delta = 2.73\,\mathrm{mT}^{-1}$ 是 sigmoid 宽度。需要指出的是，用于训练网络的随机学习规则并不要求拟合函数采用这一确切形式。

随机学习规则

对于由二元随机突触组成的网络，每个神经元的值可近似为一个高斯分布，如公式 (6) 所示，其中每个神经元的均值和方差分别由公式 (4) 和 (5) 定义。利用这一近似，可以推导出基于梯度的学习规则，因为此时随机变量不再依赖于模型参数。该网络的参数是决定突触通过概率的磁场，因此梯度下降更新为

$$\Delta h_{ij} = -\eta \frac{\partial E(\tilde{y})}{\partial \tilde{y}_i} \frac{\partial \tilde{y}_i}{\partial h_{ij}} \tag{10}$$

$$= -\eta \frac{\partial E(\tilde{y})}{\partial \tilde{y}_i} \left(\frac{\partial \mu_i}{\partial h_{ij}} + \frac{\partial \sigma_i}{\partial h_{ij}} \xi_i \right). \tag{11}$$

均值和方差相对于磁场的梯度为

$$\frac{\partial \mu_i}{\partial h_{ij}} = f'(h_{ij}) x_j \tag{12}$$

$$\frac{\partial \sigma_i}{\partial h_{ij}} = \frac{(1-2f(h_{ij})x_j)}{2\sigma_i} f'(h_{ij}) x_j, \tag{13}$$

其中 $f'(h) = \partial f(h)/\partial h$ 是通过概率函数的导数。将这一结果代入公式 (10)，可得更新规则

$$\Delta h_{ij} = -\eta \frac{\partial E(\tilde{y})}{\partial \tilde{y}_i} f'(h_{ij}) x_j \times \left(1 + \frac{1-2f(h_{ij})x_j}{2\sigma_i} \xi_i \right). \tag{14}$$

在这种形式下，该规则包含均值场分量，并乘以一个取决于方差的因子。虽然在推导该规则时，我们将ξ_i 指定为均值为零、方差为一的高斯随机变量，但在训练过程中，它会根据前向阶段利用 $\xi_i = (y_i - \mu_i)/\sigma_i$精确计算得出。因此，若神经元输出高于均值，该值为正；若低于均值，则为负。该值与$1 - 2f(h_{ij})x_j$ 共同决定该因子是增大还是减小权重更新。

模型训练详情

我们使用 MNIST 数据集[64] 作为基准，但为了减少实验硬件上的突触运算次数，使用尺寸为 2x2 的滤波器进行最大池化，对数据集进行下采样。这样得到了一组 14 x 14 像素的图像，并通过将像素强度以 0.5 为阈值进行二值化，将其映射为二值输入。

数据集的训练部分被随机划分为 50,000 个训练样本和 10,000 个验证样本。对仿真的二元突触的输出施加实数偏置，并使用 Softmax 函数将这些值转换为概率；再通过交叉熵损失计算其与图像标签之间的损失。训练以每批 50 张图像的小批次进行，并持续迭代，直到验证损失连续 20 个周期不再下降。训练结束前验证误差最低的模型将作为训练完成的模型返回。根据较低的验证误差，使用 Adam 优化器，学习率分别为 $\eta = 0.001$（用于 $K \geq 2$）和 $\eta = 0.01$（用于$K = 1$）。

设备端机器学习测试

为了演示我们的随机网络在材料中，我们使用自动控制系统，在突触权重指定的磁场下将磁畴壁注入磁性纳米线。首先，我们使用下文详述的方法，在仿真中为网络模型优化突触磁场，针对 $K_{\mathrm{train}} = 1$ 和 128 两种情况进行优化。对于每个 K_{train}，我们训练 5 个模型，然后选择在验证数据集上误差最低的模型。随后，我们将这些模型转移到硬件上，由控制软件从测试数据集中加载像素二元值（x_j），并使用仿真训练得到的磁场（h_{ij}）来控制磁性突触。如图 2 所详述，当像素值为 1 时，控制系统会判断磁畴壁是被钉扎还是通过了缺陷位点，并分别返回 0 或 1。随后，该突触运算的结果被传回运行神经网络推理的程序，由该程序计算神经元值，以预测测试数据的正确类别。

致谢

作者衷心感谢 EP-SRC（资助项目：EP/S009647/1）和 Leverhulme Trust（研究资助项目：RPG-2019-097）提供的资金支持。作者谨向 Luca Manneschi 和 Ian Vidamour 表示感谢，感谢他们对本项工作的帮助与反馈。

代码可用性

与本文相关的代码可在 https://github.com/mattoaellis/binary_stochastic_synapses 获取。

附录 A：泊松-二项分布的均值和方差

对于由二元随机突触构成的网络，假设每个突触的输出都是相互独立的随机二元事件（伯努利试验）。如果这些事件具有相同的概率，那么它们的总和将服从二项分布；但由于每个突触具有不同的输入和突触传递概率，神经元的取值将服从泊松-二项分布。由于完整计算该分布可能较为复杂，我们用高斯分布[40]近似该分布。给定样本数K 时，神经元输出的均值 y_i为

$$\mu_i = \mathbb{E}[y_i] \tag{A1}$$
$$= \mathbb{E}\left[\frac{1}{K}\sum_k\sum_j w_{ij}^{(k)}x_j\right] \tag{A2}$$
$$= \frac{1}{K}\sum_k\sum_j \mathbb{E}[w_{ij}^{(k)}x_j] \tag{A3}$$
$$= \frac{1}{K}\sum_k\sum_j (0f(h_{ij})x_j + 1f(h_{ij})x_j) \tag{A4}$$
$$= \sum_j f(h_{ij})x_j. \tag{A5}$$

其中，在最后一步中，我们注意到突触传递概率 f（h_{ij}）与 k相互独立。该分布的方差为

$$\sigma_i^2 = \mathrm{var}[y_i] \tag{A6}$$
$$= \mathrm{var}\left[\frac{1}{K}\sum_k\sum_j w_{ij}^{(k)}x_j\right]. \tag{A7}$$

现在我们利用以下事实：相互独立随机事件之和的方差等于各事件方差之和，并且 $\mathrm{var}\,[y/K] = \mathrm{var}[y]/K^2$，因此

$$\sigma_i^2 = \frac{1}{K^2}\sum_k\sum_j \mathrm{var}\left[w_{ij}^{(k)}x_j\right]. \tag{A8}$$

每个突触作为伯努利事件时的方差为

$$\mathrm{var}\left[w_{ij}^{(k)}x_j\right] = f(h_{ij})x_j\left(1 - f(h_{ij})x_j\right), \tag{A9}$$

同样，由于这与 k 无关，因此神经元输出的方差为

$$\sigma_i^2 = \frac{K}{K^2}\sum_j f(h_{ij})x_j\left(1 - f(h_{ij})x_j\right) \tag{A10}$$

$$= \frac{1}{K}\sum_j f(h_{ij})x_j\left(1 - f(h_{ij})x_j\right). \tag{A11}$$

[1]N. C. Thompson、K. Greenewald、K. Lee 和 G. F. Manso，“深度学习的边际收益递减：改进的代价已变得难以为继”，IEEE Spectru m**58**，50–55（2021年）。[2]M. M. Sabry Aly、T. F. Wu、A. Bartolo、Y. H. Malviya、W. Hwang、G. Hills、I. Markov、M. Wootters、M. M. Shulaker、H. Philip Wong 和 S. Mitra，“面向能效型海量数据计算的n3xt 方法 ”，IEEE 会刊 **107**，19–48 （ 2019 年）。[3]E. Strubell 、 A. Ganesh 和 A. McCallum，“深度学习中的能源与政策考量”，（2019年）。[4]M. Ziegler，“非常规计算的新型硬件与概念”，《科学报告》**10**，1–3（2020年）。[5]M. T. Niemier、G. H. Bernstein、G. Csaba、A. Dingler、X. S. Hu、S. Kurtz、S. Liu、J. Nahas、W. Porod、M. Siddiq 和 E. Varga，“纳米磁体逻辑：迈向系统级集成”，Journal of Physics: Condensed Matter**23**，493202（2011年）。[6]G. Finocchio、M. Di Ventra、K. Y. Camsari、K. Everschor-Sitte、P. K. Amiri 和 Z. Zeng，“自旋电子学在非常规计算中的前景”，《磁学与磁性材料杂志》**521**，167506（2021年）。[7]J. Grollier、D. Querlioz、K. Y. Camsari、K. Everschor-Sitte、S. Fukami 和 M. D. Stiles，“神经形态自旋电子学”，Nature Electronics**3**，360–370（2020年）。[8]C. C. Petersen、R. C. Malenka、R. A. Nicoll 和 J. J. Hopfield，“CA3-CA1 突触处的全或无增强”，《美国国家科学院院刊》**95**，4732–4737（1998年）。[9]T. Simons 和 D.-J. Lee，“二值化神经网络综述”，《电子学》**8**，661（2019年）。[10] A. Laborieux、M. Bocquet、T. Hirtzlin、J.-O. Klein、L. H. Diez、E. Nowak、E. Vianello、J.-M. Portal 和 D. Querlioz，“基于电阻式随机存取存储器突触的三值神经网络低功耗存内实现”，in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)* 2020 136 140 ( ) pp. – 。[11]S. Yu、B. Gao、Z. Fang、H. Yu、J. Kang 和 H.-S. P. Wong，“用于神经形态计算的氧化物二值突触器件中的随机学习”，《神经科学前沿》**7**，186（2013年）。[12]B. Penkovsky、M. Bocquet、T. Hirtzlin、J.-O. Klein、E. Nowak、E. Vianello、J.-M. Portal 和 D. Querlioz，“面向医疗应用的二值化神经网络存内电阻式随机存取存储器实现”，in *2020 Design Automation & Test in Europe Conference & Exhibition* 2020 690 695 ( ) pp. – 。[13]E. O. Neftci、B. U. Pedroni、S. Joshi、M. Al-Shedivat 和 G. Cauwenberghs，“随机突触助力实现高效的脑启发式学习机器”，《神经科学前沿》**10**，241（2016年）。[14]M. W. Daniels、A. Madhavan、P. Talatchian、A. Mizrahi 和 M. D. Stiles，“利用超顺磁隧道结实现能效型随机计算”，Phys. Rev. Appl.**13**，034016（2020年）。[15]Z. Li、J. Li、A. Ren、R. Cai、C. Ding、X. Qian、J. Draper、B. Yuan、J. Tang、Q. Qiu 和 Y. Wang，“HEIF：面向深度神经网络的高效随机计算推理框架”，IEEE Transactions on Computer-Aided Design of Integrated Circ uits and Systems**38**，1543–1556（2019年）。[16]Y. Shao、S. L. Sinaga、I. O. Sunmola、A. S. Borland、M. J. Carey、J. A. Katine、V. Lopez-Dominguez 和 P. K. Amiri，“利用基于磁阻式随机存取存储器的随机计算单元实现人工神经网络”，IEEE M agnetics Lett 等，**12** 1–5（2021年）。[17]P. K. Muthappa, F. Neugebauer, I. Polian 和 J. P. Hayes，“基于硬件的随机计算快速实时图像分类”，载于 *2020 IEEE 38th International Conference on Computer Design (ICCD)* 2020 340 347 ( ) pp. – 。[18]A. Nisar 、 F. A. Khanday 和 B. K. Kaushik，“ 高效磁隧道结随机神经网络的实现及其在鸢尾花数据分类中的应用”，《纳米技术》31，504001（2020）。[19]T. Hirtzlin、B. Penkovsky、M. Bocquet、J.-O. Klein、J.-M. Por- tal 和 D. Querlioz，“用于二值化神经网络硬件实现的随机计算”，《IEEE开放期刊》7，76394–76403（2019）。[20] M. N. Baibich、J. M. Broto、A. Fert、F. N. Van Dau、F. Petroff、P. Etienne、G. Creuzet、A. Friederich 和 J. Chazelas，“（001）Fe/（001）Cr磁性超晶格的巨磁电阻”，《物理评论快报》61，2472–2475（1988）。[21]G. Binasch、P. Gr¨unberg、F. Saurenbach 和 W. Zinn，“具有反铁磁层间交换的层状磁性结构中磁阻的增强”，《物理评论B辑》39，4828 4830 1989–（）。[22]D. A. Allwood、G. Xiong、C. Faulkner、D. Atkinson、D. Petit 和 R. Cowburn，“磁畴壁逻辑”，《科学》309，1688 1692 2005–（）。[23]W. P. Mccray，“自旋电子学如何从实验室走进iPod”，《自然·纳米技术》4（2009）。[24]S. S. Parkin、M. Hayashi 和 L. Thomas，“磁畴壁赛道存储器”，《科学》320，190–194（2008）。[25]R. Lavrijsen、J.-H. Lee、A. Fern´andez-Pacheco、D. C. M. C. Petit、R. Mansell 和 R. P. Cowburn，“用于三维自旋电子存储器和逻辑的磁棘轮”，《自然》493，647–650（2013）。[26]A. Fern´andez-Pacheco、N.-J. Steinke、D. Mahendru、A. Wel- bourne、R. Mansell、S. Chin、D. Petit、J. Lee、R. Dalgliesh、S. Langridge 和 R. Cowburn，“孤子成核与传播过程中合成反铁磁体的磁态及垂直数据传输”，《先进材料·界面》（2016）。[27]S. Emori、U. Bauer、S.-M. Ahn、E. Martinez 和 G. S. Beach，“手性铁磁畴壁的电流驱动动力学”，《自然·材料》12，611–616（2013）。[28]J.-M. Hu、Z. Li、L.-Q. Chen 和 C.-W. Nan，“室温下以超低电压运行的高密度磁阻式随机存取存储器”，《自然·通讯》2，1–8（2011）。[29]S. S. Parkin、M. Hayashi 和 L. Thomas，“磁畴壁赛道存储器”，《科学》320，190–194（2008）。[30]D. Sanz-Hern´andez、R. Hamans、J.-W. Liao、A. Welbourne、R. Lavrijsen 和 A. Fern´andez-Pacheco，“三维纳米磁导管的制造、检测与运行”，《ACS纳米》11（2017）。[31]Z. Luo、A. Hrabec、T. P. Dao、G. Sala、S. Finizio、J. Feng、S. Mayr、J. Raabe、P. Gambardella 和 L. J. Heyderman，“电流驱动的磁畴壁逻辑”，《自然》579，214–218（2020）。[32]T. Hayward，“磁性纳米线器件中随机畴壁钉扎现象的内在特性”，《科学报告》5，1 12 2015–（）。[33]D. Kumar、T. Jin、S. Al Risi、R. Sbiaa、W. S. Lew 和 S. N. Piramanayagam，“赛道存储器应用中的畴壁运动控制”，《IEEE磁学汇刊》55，1–8（2019）。[34]B. Lambson、D. Carlton 和 J. Bokor，“探索集成系统计算的热力学极限：磁存储器、纳米磁逻辑与朗道尔极限”，《物理评论快报》107，010604（2011）。[35]M. Hayashi、L. Thomas、R. Moriya、C. Rettner 和 S. S. P. Parkin，“电流控制的磁畴壁纳米线移位寄存器”，《科学》320，209–211（2008）。[36]S.-H. Yang、K.-S. Ryu 和 S. Parkin，“合成反铁磁体中由交换耦合力矩驱动的畴壁速度可达750 m s-1”

反铁磁体”，《自然·纳米技术》**10**，221–226（2015）。[37]Y. Nakatani、A. Thiaville 和 J. Miltat，“软磁纳米条带中的头对头磁畴壁：精细相图”，《磁学与磁性材料杂志》欧洲联合磁学研讨会论文集（JEMS’ 04），**290-291**，750–753（2005）。[38]J. Satel、T. Trappenberg 和 A. Fine，“在随机吸引子网络中，二元突触是否优于分级权重表示？ ”，《认知神经动力学》**3**，243–250（2009）。[39]A. M. Dubreuil、Y. Amit 和 N. Brunel，“具有随机二元突触的网络的记忆容量”，《公共科学图书馆·计算生物学》**10**，e1003727（2014）。[40]Y. Hong，“关于泊松二项分布分布函数的计算”，《计算统计与数据分析》**59**，41–51（2013）。[41]A. Vouros、T. V. Gehring、K. Szydlowska、A. Janusz、Z. Tu、M. Croucher、K. Lukasiuk、W. Konopka、C. Sandi 和 E. Vasi- laki，“对莫里斯水迷宫中游泳路径进行详细分类的通用框架”，《科学报告》**8**，1 15 2018–（）。[42]S. K. Esser、R. Appuswamy、P. Merolla、J. V. Arthur 和 D. S. Modha，“面向能效神经形态计算的反向传播”，《神经信息处理系统进展》**28**（2015）。[43]R. J. Williams，“用于联结主义强化学习的简单统计梯度跟踪算法”，《机器学习》**8**，229 256 1992–（）。[44]S. Gu、S. Levine、I. Sutskever 和 A. Mnih，“Muprop：随机神经网络的无偏反向传播”，arXiv预印本 arXiv:1511.05176（2015）。[45]P. Parmas 和 M. Sugiyama，“似然比梯度与重参数化梯度的统一视角”，载于 *International Conference on Artificial Intelligence and Statistics* 2021 4078 (PMLR,) pp. – 4086。[46]E. Vasilaki、N. Fr´emaux、R. Urbanczik、W. Senn 和 W. Gerstner，“连续状态与动作空间中的基于脉冲的强化学习：策略梯度方法何时失效”，《公共科学图书馆·计算生物学》**5**，e1000586（2009）。[47]M. A. Azam、D. Bhattacharya、D. Querlioz、C. A. Ross 和 J. Atulasimha，“用于能效神经形态器件的磁性纳米线中磁畴壁的电压控制”，《纳米技术》**31**，145201（2020）。[48]D. Sanz-Hern´andez、M. Massouras、N. Reyren、N. Rougemaille、V. Sch´anilec、K. Bouzehouane、M. Hehn、B. Canals、D. Querlioz、J. Grollier、F. Montaigne 和 D. Lacour，“人工自旋网络中的可调随机性”，《先进材料》**33**，2008135（2021）。[49]W. A. Misba、M. Lozano、D. Querlioz 和 J. Atulasimha，“采用低分辨率随机磁畴壁突触的深度神经网络能效学习”，《IEEE开放期刊》**10**，84946 84959 2022–（）。[50]N. Hassan、X. Hu、L. Jiang-Wei、W. H. Brigner、O. G. Akinola、F. Garcia-Sanchez、M. Pasquale、C. H. Bennett、J. A. C. Incorvia 和 J. S. Friedman，“具有侧向抑制的磁性磁畴壁神经元”，《应用物理学杂志》**124**，152127（2018）。[51] W. H. Brigner、N. Hassan、X. Hu、C. H. Bennett、F. Garcia-Sanchez、C. Cui、A. Velasquez、M. J. Marinella、J. A. C. Incorvia 和 J. S. Friedman，“具有基于形状的可配置激活函数的磁畴壁漏积分发放神经元”，《IEEE电子器件汇刊》**69**，2353–2359（2022）。[52]W. A. Borders、A. Z. Pervaiz、S. Fukami、K. Y. Camsari、H. Ohno 和 S. Datta，“利用随机磁性隧道结进行整数分解”，《自然》**573**，390–393（2019）。[53]W. Al Misba、M. Lozano、D. Querlioz 和 J. Atulasimha，“采用低分辨率随机磁畴壁突触的深度神经网络能效学习”，《IEEE开放期刊》**10**，84946–84959（2022）。[54] M. Koo、G. Srinivasan、Y. Shim 和 K. Roy，“Sbsnn：采用片上学习的随机比特赋能二元脉冲神经网络，实现能效神经形态计算……”

边缘”，《IEEE电路与系统汇刊I：常规论文》**67**，2546–2555（2020年）。[55]R. W. Dawidek、T. J. Hayward、I. T. Vidamour、T. J. Broomhall、G. Venkat、M. A. Mamoori、A. Mullen、S. J. Kyle、P. W. Fry、N.-J. Steinke、J. F. K. Cooper、F. Maccherozzi、S. S. Dhesi、L. Aballe、M. Foerster、J. Prat、E. Vasilaki、M. O. A. Ellis 和 D. A. Allwood，“纳米磁性系统中的动态驱动涌现”，《先进功能材料》**31**，2008389（2021年）。[56]R. V. Ababei、M. O. A. Ellis、I. T. Vidamour、D. S. Devadasan、D. A. Allwood、E. Vasilaki 和 T. J. Hayward，“利用单个磁畴壁进行神经形态计算”，《科学报告》**11**，15587（2021年）。[57]A. Welbourne、A. L. R. Levy、M. O. A. Ellis、H. Chen、M. J. Thompson、E. Vasilaki、D. A. Allwood 和 T. J. Hayward，“用于低功耗储备池计算的电压控制超顺磁系综”，《应用物理快报》**118**，202402（2021年）。[58]J. C. Gartside、K. D. Stenning、A. Vanstone、H. H. Holder、D. M. Arroo、T. Dion、F. Caravelli、H. Kurebayashi 和 W. R. Branford，“通过自旋波指纹识别在人工自旋涡旋冰中实现可重构训练和储备池计算”，《自然·纳米技术》**17**，460–469（2022年）。[59]I. T. Vidamour、M. O. A. Ellis、D. Griffin、G. Venkat、C. Swindells、R. W. S. Dawidek、T. J. Broomhall、N. J. Steinke、J. F. K. Cooper、F. Maccherozzi、S. S. Dhesi、S. Stepney、E. Vasilaki、D. A. Allwood 和 T. J. Hayward，“利用涌现磁化动力学量化纳米磁性储备池计算平台的计算能力”，《纳米技术》**33**，485203（2022年）。[60]I. Vidamour、C. Swindells、G. Venkat、P. Fry、A. Welbourne、R. Rowan-Robinson、D. Backes、F. Maccherozzi、S. Dhesi、E. Vasilaki、D. Allwood 和 T. Hayward，“磁性超材料中的涌现动力学储备池计算”，（2022年），arXiv:2206.04446 [cond-mat]。[61]D. A. Allwood、M. O. A. Ellis、D. Griffin、T. J. Hayward、L. Manneschi、M. F. K. Musameh、S. O’Keefe、S. Stepney、C. Swindells、M. A. Trefzer、E. Vasilaki、G. Venkat、I. Vidamour 和 C. Wringe，“关于利用纳米磁性器件进行物理储备池计算的展望”，（2022年），arXiv:2212.04851 [physics]。[62]K. D. Stenning、J. C. Gartside、L. Manneschi、C. T. S. Che-ung、T. Chen、A. Vanstone、J. Love、H. H. Holder、F. Caravelli、K. Everschor-Sitte、E. Vasilaki 和 W. R. Branford，“用于原位神经形态计算的自适应可编程网络”，（2022年），arXiv:2211.06373 [cond-mat]。[63]T. Hirtzlin、M. Bocquet、B. Penkovsky、J.-O. Klein、E. Nowak、E. Vianello、J.-M. Portal 和 D. Querlioz，“利用差分氧化铪阻变存储阵列对二值化神经网络进行数字化生物学可信实现”，《神经科学前沿》**13**，1383（2020年）。[64]Y. LeCun，“手写数字的MNIST数据库”，http://yann.lecun.com/exdb/mnist/（1998年）。

[1]N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso，《深度学习的收益递减：改进的代价正变得难以为继》，IEEE Spectrum**58**, 50–55 (2021年)。[2]M. M. Sabry Aly, T. F. Wu, A. Bartolo, Y. H. Malviya, W. Hwang, G. Hills, I. Markov, M. Wootters, M. M. Shulaker, H. Philip Wong, and S. Mitra，《面向高能效海量数据计算的 n3xt 方法》，IEEE 会刊**107**, 19–48 (2019)。[3]E. Strubell, A. Ganesh, and A. McCallum，《深度学习在自然语言处理中的能源与政策考量》(2019)。[4]M. Ziegler，《非常规计算的新型硬件与概念》，《科学报告》**10**, 1–3 (2020)。[5]M. T. Niemier, G. H. Bernstein, G. Csaba, A. Dingler, X. S. Hu, S. Kurtz, S. Liu, J. Nahas, W. Porod, M. Siddiq, and E. Varga，《纳米磁体逻辑：迈向系统级集成》，Journal of Physics: Condensed Matter**23**, 493202 (2011)。

[6]G. Finocchio, M. Di Ventra, K. Y. Camsari, K. Everschor-Sitte, P. K. Amiri, and Z. Zeng, 《自旋电子学在非常规计算中的前景》, 《磁学与磁性材料杂志》**521**, 167506 (2021年)。[7]J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, 《神经形态自旋电子学》, Nature Electronics**3**, 360–370 (2020)。[8]C. C. Petersen, R. C. Malenka, R. A. Nicoll, and J. J. Hopfield, 《CA3-CA1 突触处的全或无增强》, 《美国国家科学院院刊》**95**, 4732–4737 (1998年)。[9]T. Simons and D.-J. Lee, 《二值化神经网络综述》, 《电子学》**8**, 661 (2019)。[10] A. Laborieux, M. Bocquet, T. Hirtzlin, J.-O. Klein, L. H. Diez, E. Nowak, E. Vianello, J.-M. Portal, and D. Querlioz, 《基于电阻式随机存取存储器的突触对三值神经网络的低功耗存内实现》, 载于 *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)* 2020 136 140 ( )

页码 –。[11]S. Yu, B. Gao, Z. Fang, H. Yu, J. Kang, and H.-S. P. Wong, 《用于神经形态计算的氧化物二值突触器件中的随机学习》, 《神经科学前沿》**7**, 186 (2013)。[12]B. Penkovsky, M. Bocquet, T. Hirtzlin, J.-O. Klein, E. Nowak, E. Vianello, J.-M. Portal, and D. Querlioz, 《面向医疗应用的二值化神经网络存内电阻式随机存取存储器实现》, 载于 *2020 Design Automation & Test in Europe Conference & Exhibition* 2020 690 695

( ) 页码 –。[13]E. O. Neftci, B. U. Pedroni, S. Joshi, M. Al-Shedivat, and G. Cauwenberghs, 《随机突触助力高效的脑启发式学习机器》, 《神经科学前沿》**10**, 241 (2016)。[14]M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, 《利用超顺磁隧道结实现高能效随机计算》, Phys. Rev. Appl. **13**, 034016 (2020)。[15]Z. Li, J. Li, A. Ren, R. Cai, C. Ding, X. Qian, J. Draper, B. Yuan, J. Tang, Q. Qiu, and Y. Wang, 《HEIF：面向深度神经网络的高效随机计算推理框架》, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**38**, 1543–1556 (2019)。[16]Y. S hao, S. L. Sinaga, I. O. Sunmola, A. S. Borland, M. J. Carey, J. A. Katine, V. Lopez-Dominguez, and P. K. Amiri, 《基于磁阻式随机存取存储器的随机计算单元实现人工神经网络》, 《IEEE磁学快报》**12**, 1–5 (2021年)。[17]P. K. Muthappa, F. Neugebauer, I. Polian, and J. P. Hayes, 《基于硬件的随机计算快速实时图像分类》, 载于 *2020 IEEE 38th International Conference on Computer Design (ICCD)* 2020 340 347 ( )

页码 –。[18]A. Nisar, F. A. Khanday, and B. K. Kaushik, 《适用于鸢尾花数据分类的高效磁性隧道结随机神经网络实现》, 《纳米技术》**31**, 504001 (2020)。[19]T. Hirtzlin, B. Penkovsky, M. Bocquet, J.-O. Klein, J.-M. Por- tal, and D. Querlioz, 《用于二值化神经网络硬件实现的随机计算》, 《IEEE开放期刊》**7**, 76394–76403 (2019)。[20]M. N. Baibich, J. M. Broto, A. Fert, F. N. Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, 《(001)Fe/(001)Cr 磁性超晶格的巨磁电阻》, 《物理评论快报》**61**, 2472–2475 (1988)。[21]G. Binasch, P. Gr¨unberg, F. Saurenbach, and W. Zinn, 《具有反铁磁层间交换的层状磁性结构中的磁阻增强》, 《物理评论B辑》**39**, 4828 4830 1989– ( )。[22]D. A. Allwood, G. Xiong, C. Faulkner, D. Atkinson, D. Petit, and R. Cowburn, 《磁畴壁逻辑》, 《科学》**309**, 1688 1692 2005– ( )。[23]W. P. Mccray, 《自旋电子学如何从实验室走进 iPod》, 《自然·纳米技术》**4** (2009)。[24]S. S. Parkin、M. Hayashi 和 L. Thomas，“磁畴壁赛道存储器”，《科学》320，190–194（2008）。[25]R. Lavrijsen、J.-H. Lee、A. Fern´andez-Pacheco、D. C. M. C. Petit、R. Mansell 和 R. P. Cowburn，“用于三维自旋电子存储器和逻辑的磁棘轮”，《自然》493，647–650（2013）。[26]A. Fern´andez-Pacheco、N.-J. Steinke、D. Mahendru、A. Wel- bourne、R. Mansell、S. Chin、D. Petit、J. Lee、R. Dalgliesh、S. Langridge 和 R. Cowburn，“用于垂直数据传输的孤子成核与传播过程中多层合成反铁磁体的磁态”，Advanced Materials Interfaces（2016）。[27]S. Emori、U. Bauer、S.-M. Ahn、E. Martinez 和 G. S. Beach，“手性铁磁畴壁的电流驱动动力学”，《自然·材料》12，611–616（2013）。[28]J.-M. Hu、Z. Li、L.-Q. Chen 和 C.-W. Nan，“室温下以超低电压运行的高密度磁阻式随机存取存储器”，《自然·通讯》2，1–8（2011）。[29]S. S. Parkin、M. Hayashi 和 L. Thomas，“磁畴壁赛道存储器”，《科学》320，190–194（2008）。[30]D. Sanz-Hern´andez、R. Hamans、J.-W. Liao、A. Welbourne、R. Lavrijsen 和 A. Fern´andez-Pacheco，“三维纳米磁性导管的制造、检测与运行”，ACS Nano 11（2017）。[31]Z. Luo、A. Hrabec、T. P. Dao、G. Sala、S. Finizio、J. Feng、S. Mayr、J. Raabe、P. Gambardella 和 L. J. Heyderman，“电流驱动的磁畴壁逻辑”，《自然》579，214–218（2020）。[32]T. Hayward，“磁性纳米线器件中随机畴壁钉扎现象的内在本质”，《科学报告》5，1 12 2015–（）。[33] D. Kumar、T. Jin、S. Al Risi、R. Sbiaa、W. S. Lew 和 S. N. Piramanayagam，“用于赛道存储器应用的畴壁运动控制”，IEEE Transactions on Magnetics 55，1–8（2019）。[34]B. Lambson、D. Carlton 和 J. Bokor，“探索集成系统中计算的热力学极限：磁存储器、纳米磁逻辑与朗道尔极限”，《物理评论快报》107，010604（2011）。[35]M. Hayashi、L. Thomas、R. Moriya、C. Rettner 和 S. S. P. Parkin，“电流控制的磁畴壁纳米线移位寄存器”，《科学》320，209–211（2008）。[36]S.-H. Yang、K.-S. Ryu 和 S. Parkin，“合成反铁磁体中由交换耦合力矩驱动的畴壁速度高达 750 m s-1”，《自然·纳米技术》10，221–226（2015）。[37]Y. Nakatani、A. Thiaville 和 J. Miltat，“软磁纳米条带中的头对头畴壁：精细相图”，《磁学与磁性材料杂志》，欧洲联合磁学研讨会论文集（JEMS’ 04），290-291，750–753（2005）。[38]J. Satel、T. Trappenberg 和 A. Fine，“在随机吸引子网络中，二元突触是否优于分级权重表示？”，《认知神经动力学》3，243–250（2009）。[39]A. M. Dubreuil、Y. Amit 和 N. Brunel，“具有随机二元突触的网络的记忆容量”，《公共科学图书馆·计算生物学》10，e1003727（2014）。[40]Y. Hong，“关于泊松二项分布分布函数的计算”，Computational Statistics & Data Analysis 59，41–51（2013）。[41]A. Vouros、T. V. Gehring、K. Szydlowska、A. Janusz、Z. Tu、M. Croucher、K. Lukasiuk、W. Konopka、C. Sandi 和 E. Vasi- laki，“莫里斯水迷宫中游泳路径详细分类的通用框架”，《科学报告》8，1 15 2018–（）。[42]S. K. Esser、R. Appuswamy、P. Merolla、J. V. Arthur 和 D. S. Modha，“面向节能神经形态计算的反向传播”，Advances in Neural Information Processing Systems 28（2015）。

[43]R. J. Williams，“用于联结主义强化学习的简单统计梯度跟踪算法”，《机器学习》**8**，229 256 1992–（）。[44]S. Gu、S. Levine、I. Sutskever 和 A. Mnih，“Muprop：随机神经网络的无偏反向传播”，arXiv 预印本 arXiv:1511.05176（2015）。[45]P. Parmas 和 M. Sugiyama，“似然比梯度与重参数化梯度的统一视角”，收录于 *International Conference on Artificial Intelligence and Statistics* 2021 4078

（PMLR，）第 – 4086 页。[46]E. Vasilaki、N. Fr´emaux、R. Urbanczik、W. Senn 和 W. Gerst- ner，“连续状态与动作空间中的基于脉冲的强化学习：策略梯度方法何时失效”，《公共科学图书馆·计算生物学》**5**，e1000586（2009）。[47]M. A. Azam、D. Bhattacharya、D. Querlioz、C. A. Ross 和 J. Atulasimha，“磁性纳米线中磁畴壁的电压控制及其在节能神经形态器件中的应用”，《纳米技术》**31**，145201（2020）。[48]D. Sanz-Hern´andez、M. Massouras、N. Reyren、N. Rougemaille、V. Sch´anilec、K. Bouzehouane、M. Hehn、B. Canals、D. Querlioz、J. Grollier、F. Montaigne 和 D. Lacour，“人工自旋网络中的可调随机性”，《先进材料》**33**，2008135（2021）。[49]W. A. Misba、M. Lozano、D. Querlioz 和 J. Atulasimha，“利用低分辨率随机磁畴壁突触实现深度神经网络的节能学习”，《IEEE开放期刊》**10**，84946 84959 2022–（）。[50]N. Hassan、X. Hu、L. Jiang-Wei、W. H. Brigner、O. G. Akinola、F. Garcia-Sanchez、M. Pasquale、C. H. Bennett、J. A. C. Incorvia 和 J. S. Friedman，“具有侧向抑制的磁畴壁神经元”，《应用物理学杂志》**124**，152127（2018）。[51]W. H. Brigner、N. Hassan、X. Hu、C. H. Bennett、F. Garcia- Sanchez、C. Cui、A. Velasquez、M. J. Marinella、J. A. C. Incorvia 和 J. S. Friedman，“具有基于形状的可配置激活函数的磁畴壁泄漏积分发放神经元”，《IEEE电子器件汇刊》**69**，2353–2359（2022）。[52]W. A. Borders、A. Z. Pervaiz、S. Fukami、K. Y. Camsari、H. Ohno 和 S. Datta，“利用随机磁隧道结进行整数分解”，《自然》**573**，390–393（2019）。[53]W. Al Misba、M. Lozano、D. Querlioz 和 J. Atulasimha，“利用低分辨率随机磁畴壁突触实现深度神经网络的节能学习”，《IEEE开放期刊》**10**，84946–84959（2022）。[54]M. Koo、G. Srinivasan、Y. Shim 和 K. Roy，“Sbsnn：随机比特赋能的二元脉冲神经网络，采用片上学习实现边缘端节能神经形态计算”，IEEE电路与系统汇刊I：常规论文**67**，2546–2555（2020）。[55]R. W. Dawidek、T. J. Hayward、I. T. Vidamour、T. J. Broomhall、G. Venkat、M. A. Mamoori、A. Mullen、S. J. Kyle、P. W. Fry，

N.-J. Steinke, J. F. K. Cooper, F. Maccherozzi, S. S. Dhesi, L. Aballe, M. Foerster, J. Prat, E. Vasilaki, M. O. A. Ellis 和 D. A. Allwood，“纳米磁性系统中的动态驱动涌现”，《先进功能材料》**31**，2008389（2021年）。[56]R. V. Ababei, M. O. A. Ellis, I. T. Vidamour, D. S. Devadasan, D. A. Allwood, E. Vasilaki 和 T. J. Hayward，“利用单个磁畴壁进行神经形态计算”，《科学报告》**11**，15587（2021年）。[57]A. Welbourne, A. L. R. Levy, M. O. A. Ellis, H. Chen, M. J. Thompson, E. Vasilaki, D. A. Allwood 和 T. J. Hayward，“用于低功耗储备池计算的电压控制超顺磁系综”，《应用物理快报》**118**，202402（2021年）。[58]J. C. Gartside, K. D. Stenning, A. Vanstone, H. H. Holder, D. M. Arroo, T. Dion, F. Caravelli, H. Kurebayashi 和 W. R. Branford，“通过自旋波指纹识别，在人工自旋涡旋冰中实现可重构训练和储备池计算”，《自然·纳米技术》**17**，460–469（2022年）。[59]I. T. Vidamour, M. O. A. Ellis, D. Griffin, G. Venkat, C. Swindells, R. W. S. Dawidek, T. J. Broomhall, N. J. Steinke, J. F. K. Cooper, F. Maccherozzi, S. S. Dhesi, S. Stepney, E. Vasilaki, D. A. Allwood 和 T. J. Hayward，“量化具有涌现磁化动力学的纳米磁性储备池计算平台的计算能力”，《纳米技术》**33**，485203（2022年）。[60]I. Vidamour, C. Swindells, G. Venkat, P. Fry, A. Welbourne, R. Rowan-Robinson, D. Backes, F. Maccherozzi, S. Dhesi, E. Vasilaki, D. Allwood 和 T. Hayward，“磁性超材料中涌现动力学的储备池计算”（2022年），arXiv:2206.04446 [cond-mat]。[61]D. A. Allwood, M. O. A. Ellis, D. Griffin, T. J. Hayward, L. Manneschi, M. F. K. Musameh, S. O’Keefe, S. Stepney, C. Swindells, M. A. Trefzer, E. Vasilaki, G. Venkat, I. Vidamour 和 C. Wringe，“纳米磁性器件物理储备池计算展望”（2022年），arXiv:2212.04851 [physics]。[62]K. D. Stenning, J. C. Gartside, L. Manneschi, C. T. S. Che-ung, T. Chen, A. Vanstone, J. Love, H. H. Holder, F. Caravelli, K. Everschor-Sitte, E. Vasilaki 和 W. R. Branford，“用于物质内神经形态计算的自适应可编程网络”（2022年），arXiv:2211.06373 [cond-mat]。[63]T. Hirtzlin, M. Bocquet, B. Penkovsky, J.-O. Klein, E. Nowak, E. Vianello, J.-M. Portal 和 D. Querlioz，“利用差分氧化铪阻变存储阵列对二值化神经网络进行数字化生物学可信实现”，《神经科学前沿》**13**，1383（2020年）。[64]Y. LeCun，“手写数字的 MNIST 数据库”，http://yann.lecun.com/exdb/mnist/（1998年）。