arXiv:2611.32393v1 [cs.LG] 24 Nov 2026

分形神经合成：通过复二次边界动力学进行常数存储过程化权重推导

Volkan Dağlı[1,2,*] **Zerrin Dağlı**[3,4] **Dağhan Dağlı**[5]

[1]阿纳多卢大学，埃斯基谢希尔，土耳其 • [2]*ITouch Systems*，丘库罗瓦大学科技园，阿达纳，土耳其[3]梅尔辛大学，梅尔辛，土耳其 • [4]优素福·卡尔卡万安纳托利亚高中，国家教育部，梅尔辛，土耳其[5]托罗斯科学高中，*MEV*（托罗斯大学），梅尔辛，土耳其

ORCID：V. Dağlı 0009-0000-1587-8703 • Z. Dağlı 0000-0001-9490-6425 • D.Dağlı0009-0003-2492-8313*通讯
作者：ask@answerr.me • 代码与数据存档：https://doi.org/10.5281/zenodo.22774934

2026年11月

摘要

传统神经网络将突触参数存储为显式张量，其持久性闪存占用随参数数量线性增长$O(W)$，这限制了资源受限的边缘微控制器。我们提出并系统评估分形神经合成（FNS），这是一种程序化权重推导架构，以恒定的 24 字节坐标种子$\Theta = (c_x, c_y, \ell) \in \mathbb{F}_{64}^3$ 取代隐藏层矩阵，该种子锚定在 $\partial\mathcal{M}$ $f_c(z) = z^2 + c$的复二次分岔边界附近。我们引入大退出阈值 $R = 256$平滑逃逸核、四种零开销几何采样孔径——笛卡尔网格、互素 3×3 三分法则$\mathbb{Z}/9\mathbb{Z}$、 ϕ-Vogel 螺旋和多尺度谐波三脚架——并证明条件 Kolmogorov$K_{\mathrm{Kolm}} \leq 192 + c_0$位与香农熵界。在 $\zeta \in [10^0, 10^8]$, C^1 倍放大下，傅里叶和晶格噪声基均退化为标准化后秩 ≈ 2.25，而 SHA-256 伪随机数生成器呈现雪崩效应式的不连续性$\rho \approx 0.019$（微步保留率为 50.6%）；相比之下，$\partial\mathcal{M}$ 是唯一能同时保持秩≈ 10.0和局部连续性$\rho \in [0.81, 0.97]$（保留率为 89.2%）的方法。在四个缩放器字节数为零的非线性基准测试中进行无数据泄漏的十折交叉验证后，采用 INT8 读出层的 FNS 比零隐藏层线性 INT8 高出 +3.6 至 +10.8 个百分点，同时相较于已收敛的 Float32多层感知机，总闪存占用减少 10.7×–129×（隐藏层减少725×–2,773× ）：285 B 时 Ionosphere 达到 90.30%；2,598 B 时光学数字达到 97.22%；2,598 B 时手写数字达到 97.90%；1,570 B 时 Satimage 达到 87.00%。

关键词：分形神经合成、程序化权重生成、复二次边界动力学、TinyML、模型压缩、随机特征、进化策略。

1 引言

现代神经网络通过将每个突触权重都存储为一个显式的密集张量数组中的数值标量 [1], ，实现了出色的经验泛化能力，但会导致持久内存随标量参数总数 W 线性增长$O(W)$。这种线性存储占用在数据中心加速器上尚可管理，但对于资源受限的边缘微控制器、嵌入式遥测节点和自主传感器而言，却会造成严重的非易失性闪存瓶颈；这些设备的片上闪存只读存储器容量严格受限（32–256 KB），且静态内存读取占据了系统能耗的主要部分 [2, 3]。压缩技术——包括训练后量化、1.58位三值模型 [4],、稀疏子网络剪枝 [5],、结构化核变换 [6], 和哈希桶权重共享 [7]——虽然降低了每个权重所需的固定比特数，却仍保留 $O(W)$ 或 $O(B_{\mathrm{hash}})$ 的持久存储扩展规律。

生物神经系统采用了截然不同的发展策略：紧凑的基因组编码依靠递归、自相似的形态发生规则，协调生成远超逐点显式存储所能容纳数量的突触连接。在一维复动力学中，首一二次族 $z(= z^2 + c$ [8) 11] 的连通性轨迹集 $\mathcal{M} \subset \mathbb{C}$ —— f_c提供了程序化结构的一个解析极端：无查找递推会生成一条不可求长的分岔边界 $\partial\mathcal{M}$，其局部豪斯多夫维数为 $\dim_H(\partial\mathcal{M} \cap U) = 2$，对每个与 $\partial\mathcal{M}$ [12] 相交的开集 U 均成立。

在本文中，我们提出并系统评估**分形神经合成（FNS）**，这是一种程序化权重推导框架，以固定的 24 字节 IEEE-754 Float64 坐标种子 $\Theta = (c_x,\ c_y,\ \ell) \in \mathbb{F}_{64}^3$ 取代显式隐藏层存储，并将其锚定在 $\partial\mathcal{M}$附近。鉴于 Shannon 和 Kolmogorov 信息极限的基本约束 [13, 14], ，我们并不声称能够将任意高熵的前沿语言模型压缩为 24 B。相反，我们聚焦于三个具体的表示学习与 TinyML 系统问题： (*i*) 在无数据泄漏的评估中，针对真正的非线性、零缩放器字节真实世界基准测试，*FNS* 与量化线性分类器、带种子伪随机数生成器随机特征、哈希储备池以及已收敛的量化多层感知机相比表现如何？ (*ii*) 哪些可测量的结构属性能够区分 $\partial\mathcal{M}$与加密

伪随机数生成器、平滑傅里叶和以及固定晶格值噪声？（*iii*）几何采样孔径和无梯度优化器如何影响谱条件化与零读出合成？

在我们初步归档研究的基础上进一步推进 [15]（该研究仅在均匀网格上评估了二维合成任务），本文作出五项贡献：

1. **平滑边界核与多尺度生成器消融实验：**我们构建了一种大退出阈值（$R = 256$）平滑逃逸时间核，并在八个数量级的放大尺度（$\zeta = 10^{\ell}$、 $\ell \in \{0, \ldots, 8\}$）下，将$\partial\mathcal{M}$与 SHA-256 伪随机数生成器、平滑二维傅里叶基及固定晶格值噪声进行评估对比。实证结果表明，平滑 C^1 生成器在标准化后会坍缩至有效矩阵秩 ≈ 2.25（证实了平面泰勒斜坡秩界）；而 SHA-256 伪随机数生成器则表现出完全的密码学雪崩效应（$\rho \approx 0.019$，在 1% 微步下准确率降至 50.60%）。相比之下，$\partial\mathcal{M}$独特地兼具持续的标准化后秩（≈ 10.0）、局部连续性（$\rho \in [0.81, 0.97]$），并在零查找表条件下达到 89.15% 的微步稳定性。

2. **四种几何采样孔径：**我们形式化定义并比较四种零开销坐标模板——基线笛卡尔网格、与模剩余索引及动态互素子步长对齐的 3×3 三分法节点模板、保持面积的 ϕ-沃格尔叶序螺旋 [16]，以及多尺度谐波三脚架（$\{0.60, 1.00, 1.60\}$缩放级联）——结果显示，准确率显著提升（$p \leq 0.023$ 在 UCI 手写数字数据集上），且在不增加持久存储的情况下，谱秩利用率相对提高最高达 $+25.5\%$ 。

3. **零缩放器字节非线性基准与公平基线：**在统一的 10 折分层交叉验证下，使用四个真正的非线性真实世界基准（UCI 电离层数据集、光学数字、手写数字和 Statlog Satimage），这些基准无需持久化任何特征缩放器（0 B），我们透明地将 FNS 与零隐藏层线性 INT8、8 字节带种子的伪随机数生成器 RFF/ELM、固定的二维 SplitMix64 哈希极限学习机，以及已收敛的 FP32/INT8/INT4 多层感知机进行基准测试，并精确核算各层的缩放。

4.**信息论容量界与等预算 ES 扩展：**我们证明了条件柯尔莫哥洛夫复杂度（$K_{\mathrm{Kolm}}(\mathbf{W},\ \mathbf{b} \mid d,\ H) \leq 192 + c_0$ 比特）和香农熵界（$H(\mathbf{W},\ \mathbf{b}) \leq 192$比特），从而界定 FNS 的系统 1 工作区间；并评估了四种等预算（每代 32 次评估）的零读出多补丁优化器，其中包括受 EGGROLL [17]启发的混合精英策略 + 秩一进化策略。

5.**预热启动合成与苹果对苹果的静态随机存取存储器性能分析：**我们量化了单调递增的启动时合成延迟（8.05–23.29 毫秒），并将其与相同的纯NumPy单样本（13.6–21.0 μ 秒）和批量（5.6–8.4 μ 秒）静态随机存取存储器 推理进行比较，结果表明，一旦在易失性存储器中物化，其前向传播 FLOPs 与稠密多层感知机相当。

2 相关工作与比较定位

程序化权重推导涉及五条成熟的研究脉络，概述见表 1。

间接编码与超网络。组合模式生成网络（CPPN）和 Hy- perNEAT [18] 通过查询进化坐标神经网络来生成空间连接矩阵，而超网络 [19] 则通过反向传播训练辅助网络来预测目标权重。这两种方法都保留了经学习得到的神经生成器，其大小从数千字节到数兆字节不等。FractalNet [20] 采用自相似宏观路由，但不使用残差连接；不过，它会显式训练并存储所有张量权重。FNS 则通过无需查表的复数二次递推，彻底消除了经学习得到的生成器参数。

随机特征、极限学习机、Fastfood 和哈希储备池。 极限学习机（极限学习机） [21, 22]和随机傅里叶特征（随机傅里叶特征） [23, 24] 将固定的非线性隐藏层投影与闭式 L_2正则化线性读出层相结合。Fastfood [6]通过结构化沃尔什-哈达玛变换，将显式高斯投影从 $O(dH)$ 压缩至 $O(H)$ 的存储需求，而 HashedNets [7] 则在反向传播过程中通过符号哈希将权重映射到 B_{hash} 个共享桶中。重要的是，固定的极限学习机/随机傅里叶特征隐藏层本身也可以在启动时通过一个8字节的伪随机数生成器种子重新生成（$O(1)$ 闪存）。因此，我们在第8节中明确纳入 5.5 字节的带种子伪随机数生成器随机傅里叶特征和极限学习机，以及显式随机傅里叶特征和未经训练的二维哈希极限学习机5.5，从而区分伪随机数生成器储备池与连续 $\partial\mathcal{M}$边界合成之间的精确权衡。

隐式表示与 TinyML 压缩。诸如 SIREN [25]之类的周期性隐式神经表示将连续信号压缩为网络权重，其运行方向与 FNS 相反。在 TinyML 中，MCUNet [2], 彩票票据稀疏子网络 [5], 和 1.58 位三值权重 [4] 在保留 $O(W)$ 渐近闪存缩放的同时，推动显式权重占用向低于字节的精度发展。

内在维度与无反向传播的进化。Li 等人 [26] 表明，许多学习目标都具有紧凑的内在维度。黑盒进化策略（OpenES） [27] 以及 EGGROLL [17] 等近期的超大规模秩-r扰动方案，通过对并行低秩扰动取平均，完全消除了反向传播，从而支持纯整数和不可微训练。FNS 则通过将搜索直接映射到低维坐标空间 $\mathbb{F}_{64}^{P\times 3}$，与这一范式相辅相成。

表 1：分形神经合成（FNS）与间接编码、随机特征、TinyML 压缩及无梯度范式的比较定位。

范式／参考	持久存储	生成器／机制	无尺度	无反向传播	局部连续性
CPPN / HyperNEAT [18]	$O(\lvert G_{\text{CPPN}}\rvert)$（KB）	进化坐标神经网络	有界	Yes	Yes
超网络 [19]	$O(\lvert G_{\text{hyper}}\rvert)$（MB）	辅助神经生成器	No	No	Yes
FractalNet [20]	$O(W)$（MB）	仅自相似宏拓扑	仅拓扑	No	N/A
HashedNets / Fastfood [6, 7]	$O\ (B_{\text{hash}})$ / $O\ (H)$	哈希桶 / Hadamard 块	No	否 / 模式 II	No
带种子伪随机数生成器 ELM / 随机傅里叶特征 [21, 23]	$O(1)$ (8 B) + $O(HC)$ 头部	计数器伪随机数生成器高斯分布 / 余弦映射	常量变量	是（模式 II）	否（$\rho \approx 0$）
SIREN / 隐式神经表示 [25]	$O\ (W_{\text{INR}})$（KB–MB）	Sine-MLP 编码信号	带限	No	Yes
MCUNet / BitNet [2, 4]	$O(W \cdot b/8)$	量化 / 三值张量	N/A	No	N/A
OpenES / EGGROLL [17, 27]	$O(W)$ 稠密权重	全量或秩-r 扰动	参数空间 $\mathbb{R}^W$	Yes	Yes
FNS（本文工作）	**$O(1)$ 隐藏层（24 B）** +可选头部	**二次 $\partial\mathcal{M}$ + 4孔径**	**是（$\zeta \le 10^8$）**	**Yes**	**是（$\rho \approx 0.84$）**

3 数学表述与理论界限

图1展示了 FNS 的端到端流程：从持久性24字节闪存存储，到大逃逸半径边界评估、四个几何孔径、启动时易失性 SRAM 实例化，以及目标执行环境。

3.1 大逃逸半径平滑逃逸时间核

考虑一个在临界点 $z_0 = 0$ 处初始化的首一复二次递推：

$$z_{n+1} = f_c(z_n) = z_n^2 + c, \qquad c = c_x + ic_y \in \mathbb{C}. \tag{1}$$

复参数平面可不相交地划分为$\mathbb{C} = \text{int}(\mathcal{M}) \sqcup \partial\mathcal{M} \sqcup (\mathbb{C} \setminus \mathcal{M})$，其中 $\mathcal{M} = \{c \in \mathbb{C} : \sup_{n\ge 0} |z_n(c)| \le 2\}$ 由内部双曲分支和分岔边界 $\partial\mathcal{M}$ [10, 28]构成。对于迭代截断值为 $M_{\max}$（默认值为$M_{\max} = 70$）、逃逸半径较大（ $R = 256.0$）的有界求值，我们定义逃逸时间指标，以避免哨兵值冲突：

$$N_{\text{esc}}(c) = \begin{cases} \min\{n \le M_{\max} : |z_n(c)| > R\}, & \text{if escaped,} \\ \infty, & \text{otherwise,} \end{cases} \tag{2}$$

通过 $n \in \{1, \dots, M_{\max}\}$。由式（2）可直接得到两个确定性场：

- **二元有界指示函数** $I(c) = \mathbb{I}[N_{\text{逃逸}}(c) = \infty] \in \{0, 1\}$，用于模式 I 象限布尔单元中的空间面积积分；

- **归一化平滑逃逸时间势**$\tilde{\nu}(c) \in [0, 1]$，若 $N_{\text{逃逸}}(c) = \infty$，则设为 $\tilde{\nu}(c) = 1.0$；否则通过标准平滑迭代次数 [10, 28] 定义：

$$\nu(c) = N_{\text{esc}}(c) + 1 - \log_2 \log_2 |z_{N_{\text{esc}}(c)}(c)|,$$
$$\tilde{\nu}(c) = \text{clip}\left(\frac{\nu(c)}{M_{\max}},\ 0,\ 1 - 10^{-6}\right). \tag{3}$$

关键在于，采用较大的退出阈值 $R = 256.0$ 可确保 $|z_n^2 + c| \approx |z_n|^2$ 在逃逸时，从而使 $\log_2\log_2 |z_{n+1}| - \log_2 \log_2 |z_n| = 1 + O(R^{-2})$，消除宏观的 $O(1)$ 跨越逃逸勒姆尼斯凯特曲线的跳跃不连续性 $\{c : |z_k(c)| = R\}$，这些不连续性出现在最小退出阈值 $R = 2.0$处。请注意，$\tilde{\nu}(c)$ 是 $-\log_2 G(c)$ 的截断、归一化单调替代量，其中 $G(c) = \lim_{n\to\infty} 2^{-n} \ln |z_n(c)|$ 是连续的外部 Douady–Hubbard 格林势 [10]。

3.2 分岔前沿 $\partial\mathcal{M}$：抛物型与 Misiurewicz 缩放

生物运动适应通常依赖方向性误差边界，而非各向同性的梯度步[29]。在复参数空间中，$\partial\mathcal{M}$ 是稀疏集，其拓扑维数为 1；然而，Shishikura [12] 证明，每个与 $\partial\mathcal{M}$相交的非空开集 U 都具有最大的度量豪斯多夫维数 $\dim_H(\partial\mathcal{M} \cap U) = 2$。

为避免与复迭代符号 z_n 冲突，令 $\ell \in \mathbb{R}$ 表示对数放大，$\zeta = 10^\ell$ 表示放大率，$s = 10^{-\ell}$ 表示窗口半宽。$\partial\mathcal{M}$ 附近的逃逸动力学在很大程度上取决于锚点类型：

- **抛物型尖点与山谷锚点：**在倍周期抛物颈点 $c = -3/4$（乘子 $\lambda = -1$）处，垂直方向的逃逸按 $N_{\text{esc}}(-3/4 + i\epsilon) \sim \pi/\epsilon$ [30] 缩放，因此固定截断值 $M_{\max} = 70$ 在 $\zeta \approx 22$ 时就会饱和（$I(c) \equiv 1$）。向海马谷外部边界 $\Theta_0 = (-0.743643887,\ 0.131825904,\ 2.15)$ 偏移，可在 $M_{\max} = 70$ 下，为实际缩放 $\ell \in [0.5, 3.5]$ 提供丰富且不会饱和的边界对比度。

- **Misiurewicz 锚点 ($O(\log\zeta)$ 深度)：**在前周期 Misiurewicz 点 $c_{\text{Mis}} = -0.10109636384562 + 0.95628651080914i$ ($M_{3,\ 1}$) 处，临界轨道落在乘子为 $|\lambda| \approx 1.3283$ 的排斥不动点上，并且根据 Tan Lei 定理 [31]，$\partial\mathcal{M}$ 具有渐近自相似性。对于距离 c_{Mis} 为 $\sim \zeta^{-1}$ 的典型外部点，逃逸时间会按对数规律增长，为 $\Theta(\log\zeta)$；适度的自适应截断 $M_{\max}(\zeta) = 70 + 40\ell$ 可在八个数量级（$\zeta \in [10^0, 10^8]$）范围内精准追踪这一规律，且不增加任何持久化字节。

3.3 四种几何采样孔径

给定一个 24 字节的坐标种子 $\Theta = (c_x,\ c_y,\ \ell) \in \mathbb{F}_{64}^3$ ($s = 10^{-\ell}$)，合成隐藏权重 $\mathbf{W} \in \mathbb{R}^{d\times H}$ 与偏置

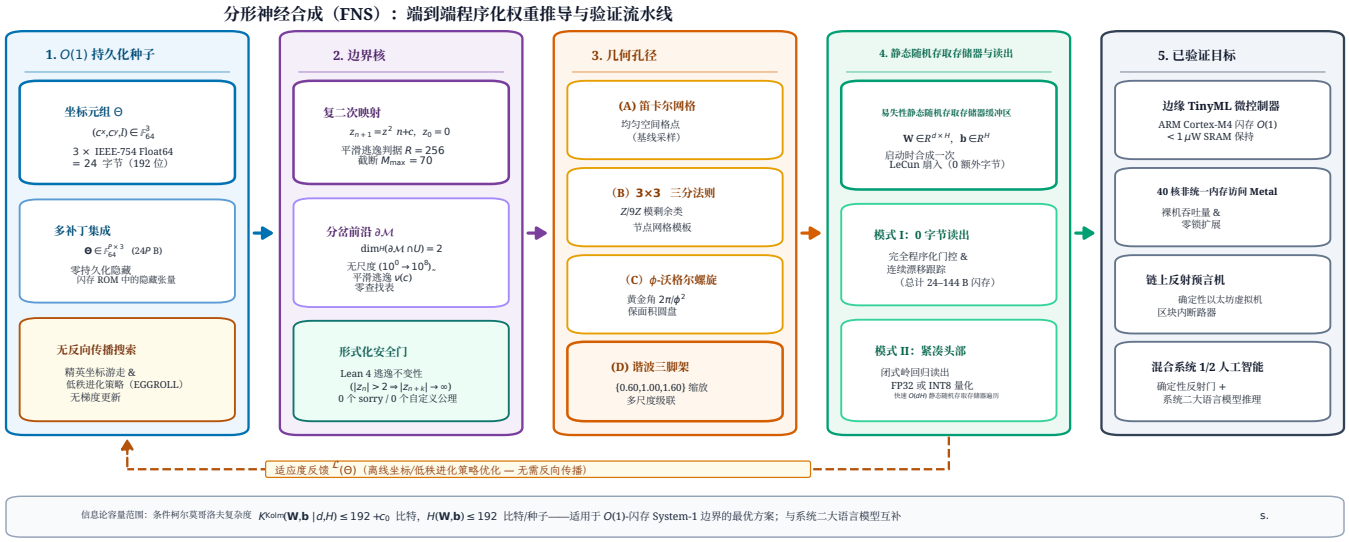

图1：端到端分形神经合成（FNS）流程。一个 24 字节坐标元组 $\Theta = (c_x, c_y, \ell) \in \mathbb{F}_{64}^3$ 驱动无查表二次核 $f_c(z) = z^2 + c$ $(R = 256)$，沿着 $\partial\mathcal{M}$运行，通过四个几何孔径进行采样，写入启动时易失性静态随机存取存储器缓冲区，支持 0 字节程序化模式 I 反射或紧凑的 FP32/INT8 模式 II 读出。

$\mathbf{b} \in \mathbb{R}^H$ 需要 $M = (d+1)H$ 确定性复坐标 $\{C_k\}_{k=0}^{M-1}$。我们定义并评估四种零开销几何孔径:

1. 笛卡尔网格（基线）： 在截断到 M 个点的 $[c_x - s, c_x + s] \times [c_y - s, c_y + s]$ 上评估均匀 $\lceil\sqrt{M}\rceil \times \lceil\sqrt{M}\rceil$ 晶格。对于最简双输入布尔门（实验1），将一个 128×128 网格划分为四个象限（Q_1, Q_2, Q_3, Q_4），将内部暗区比例 $\mathbf{R}_Q \in [0,1]^4$ 直接映射为 $\boldsymbol{\theta} = (w_1, w_2, w_3, b) = 6.0(\mathbf{R}_Q - \frac{1}{2}\mathbf{1}_4)$（保留 w_1, w_2, b）。

2. 3×3 三分法则节点网格（$\mathbb{Z}/9\mathbb{Z}$）： 将窗口划分为九个子盆地，索引为 $r_k = k \bmod 9 \in \mathbb{Z}/9\mathbb{Z}$，宏观偏移量为 $u_k = \lfloor r_k/3 \rfloor - 1$，$v_k = (r_k \bmod 3) - 1$，层索引为 $q_k = \lfloor k/9 \rfloor$。在每个宽度为 $L_{\text{sub}} = \max(2, \lceil\sqrt{\lceil M/9 \rceil}\rceil)$ 的 $1/3 \times 1/3$ 盆地内，我们选择与 L_{sub} $(\gcd(p_r, L_{\text{sub}}) = \gcd(p_c, L_{\text{sub}}) = 1)$ 互素的最小整数 $p_r \geq 5, p_c \geq 7$，从而确保没有子行或子列退化:

$$C_k = \left[c_x + \tfrac{2}{3}s\left(v_k + \frac{(\lfloor q_k/L_{\text{sub}} \rfloor p_c + 3r_k) \bmod L_{\text{sub}}}{L_{\text{sub}} - 1} - \tfrac{1}{2}\right)\right] + i\left[c_y + \tfrac{2}{3}s\left(u_k + \frac{q_k p_r \bmod L_{\text{sub}}}{L_{\text{sub}} - 1} - \tfrac{1}{2}\right)\right]. \quad (4)$$

3.保面积 ϕ-沃格尔叶序螺旋： 使用 M 在圆盘上以均匀面密度分布 $D(c, s)$ 的沃格尔费马螺旋 [16]和黄金角 $\alpha_\phi = 2\pi/\phi^2 \approx 2.39996$弧度 $(\phi = (1+\sqrt{5})/2)$:

$$C_k = \left(c_x + s\sqrt{\frac{k+0.5}{M}}\cos(k\alpha_\phi)\right) + i\left(c_y + s\sqrt{\frac{k+0.5}{M}}\sin(k\alpha_\phi)\right). \quad (5)$$

根据赫尔维茨定理，[32] ϕ 最难用有理数逼近，从而避免连续索引 k，$k+1$ 出现低周期角度聚簇（尽管在斐波那契列滞后$F_m \in \{8, 13, 21, 34, \ldots\}$处仍会出现副晶格对齐）。

4. 多尺度谐波三脚架（$\{0.60, 1.00, 1.60\}$）： 通过交错三个同心Vogel层级（$t = k \bmod 3, j_k = \lfloor k/3 \rfloor$，$M_{\text{tier}} = \lceil M/3 \rceil$），打破斐波那契滞后列对齐；这些层级具有缩放倍数$\tau_t \in \{0.60, 1.00, 1.60\}$（生成半径 s/τ_t）和 120°相位偏移:

$$C_k = \left(c_x + \frac{s}{\tau_t}\sqrt{\frac{j_k+0.5}{M_{\text{tier}}}}\cos\left(j_k\alpha_\phi + \frac{2\pi t}{3}\right)\right) + i\left(c_y + \frac{s}{\tau_t}\sqrt{\frac{j_k+0.5}{M_{\text{tier}}}}\sin\left(j_k\alpha_\phi + \frac{2\pi t}{3}\right)\right). \quad (6)$$

设 $\mathbf{u} = [\tilde{\nu}(C_0) \ldots, \tilde{\nu}(C_{M-1})]^T \in [0,1]^M$，其总体均值为 $\mu(\mathbf{u})$，标准差为 $\sigma(\mathbf{u})$。增加 **0 个持久化字节**后，启动时合成会进行逐张量标准化（$\varepsilon = 10^{-8}$），并按行优先（C 顺序）重塑—

以及 LeCun 扇入方差缩放 [33]:

$$\tilde{\mathbf{u}} = \frac{\mathbf{u} - \mu(\mathbf{u})}{\sigma(\mathbf{u}) + \varepsilon}, \tag{7}$$

$$\mathbf{W} = \frac{\gamma}{\sqrt{d}} \operatorname{reshape}^{\mathrm{C}}_{d \times H}(\tilde{\mathbf{u}}_{0:dH}), \qquad \mathbf{b} = \frac{1}{2}\, \tilde{\mathbf{u}}_{dH:M}.$$

3.4 存储复杂度、临时 SRAM 与信息论界

区分持久化的非易失性闪存 ROM 与易失性的运行时静态随机存取存储器，有助于明确 FNS 的系统特性:

- **持久化闪存 ROM（$O(1)$隐藏存储）：** 单个种子 $\Theta \in \mathbb{F}_{64}^3$ 占用恰好 24 B，与 d 或 H无关。模式 I 多补丁网络总共存储 $24P$ B（读出为 0 B）。模式 II 网络存储 $24+B_{\text{头部}}$字节，其中 $B_{\text{头部}} = 4(H+1)C_{\text{输出}}$ (FP32)，或 $(H+1)C_{\text{输出}} + 4$ （对称 INT8 $+$，加一个 4 字节 Float32 缩放因子），其中二分类的 $C_{\text{输出}} = 1$，多分类任务的 $C_{\text{输出}} = C$。

- **临时易失性 SRAM 暂存区：** 合成仅在设备上电时执行一次，在易失性 SRAM 中填充（$\mathbf{W}$，$\mathbf{b}$）。之后，运行时推理执行标准点积，转义时间迭代次数为零；在微控制器占空比睡眠期间，现代 ARM Cortex-M4 的待机保持可将数千字节的 SRAM 保持在低于 $1\,\mu$W [3]的功耗下。

命题1 （信息论容量上限）。设 $\mathcal{F}_{\text{FNS}}$: $\mathbb{F}_{64}^3 \to \mathbb{F}_{64}^{d\times H} \times \mathbb{F}_{64}^{H}$ 表示公式（1）–（7）在固定维度（d，H）、截断值 $M_{\max}$、孔径 $\mathcal{A}$以及固定程序长度c_0位下的确定性 *FNS* 合成映射。则:

*1.*条件柯尔莫哥洛夫复杂度：每个合成的隐藏张量都满足 $K_{\text{Kolm}}(\mathbf{W}、\mathbf{b} \mid d、H、M_{\max}、\mathcal{A}) \le 192 + c_0$ 位，且与 d 和 H 无关。

*2.*像集基数与香农熵：像集包含严格少于 2^{192} 个不同的隐藏张量，$|\mathrm{Im}(\mathcal{F}_{\text{FNS}})| < 2^{192}$。对于有限种子 $\Theta \in \mathbb{F}_{64}^3$ 上的任意分布，香农熵满足$H(\mathbf{W}、\mathbf{b}) \le H(\Theta) \le 192$ 位 *[13]*。

证明。（1）给定 $(d、H、M_{\max}、\mathcal{A})$ 作为旁信息后，一个长度恒定为 c_0 的前缀无歧义程序解析 Θ 的 192 位 IEEE 754 表示，并确定性地执行公式（1）–（7），输出 (W,b)，从而得到 $K_{\text{Kolm}}(\mathrm{W,b} \mid d、H、M_{\max}、\mathcal{A}) \le 192 + c_0$。（2）由于 $\mathbb{F}_{64}^3$ 至多具有 2^{192} 种位模式（排除非数/无穷大后严格更少），且 $\mathcal{F}_{\text{FNS}}$ 是确定性的，因此 $|\mathrm{Im}(\mathcal{F}_{\text{FNS}})| \le |\mathbb{F}_{64}^3| < 2^{192}$；根据离散熵的数据处理不等式，$H(\mathcal{F}_{\text{FNS}}(\Theta)) \le H(\Theta) \le \log_2 |\mathbb{F}_{64}^3| \le 192$ 位。 □

备注1 （模式 II 的表达能力与开放的稠密性问题）。 在模式 *II* 中，添加可训练线性头 $\mathbf{W}_{\text{out}} \in \mathbb{R}^{(H+1)\times C_{\text{out}}}$ 可将 *192* 位隐藏上限提升至$O(HC_{\text{out}})$ 的总存储量，同时将隐藏闪存保持在 $O(1)$。根据经典脊函数理论，*[22, 34, 35]*, 若隐藏特征字典 $\{1,$双曲正切$(\mathbf{w}_j^T\mathbf{x} + b_j)\}_{j=1}^{H}$ 在 $C(K)$ 中张成一个稠密子空间，则在紧子集上 $K \subset \mathbb{R}^d$ 可实现通用逼近。由于 *FNS* 各行是对 $\tilde{\nu}$ 的确定性求值，而非从全支撑测度中进行独立同分布抽样，因此 *FNS* 不会自动继承极限学习机/随机傅里叶特征的概率通用性定理；矩阵满秩秩$_{\text{eff}}(\mathbf{W})$ 本身也不能推出 $C(K)$稠密性。证明由确定性 $\partial\mathcal{M}$生成的字典具有解析稠密性，仍是一个开放的理论问题；本文通过有效谱秩 *[36]*和留出交叉验证对条件性进行实证评估。

命题 1和注释1 界定了 FNS 的认识论范围：这与内在维度理论相符 [26]，也符合无免费午餐界 [37], FNS 面向紧凑、确定性的 System-1 分类器和边缘控制器，而非高熵的 System-2 语言模型。

4 无反向传播合成与优化算法

FNS 在两种运行模式下进行训练，无需计算穿过隐藏层权重张量的反向梯度。这两种模式统一于算法1:

模式 I（零读出坐标搜索）。对于布尔门（24 B/单元）、两层异或（72 B）和紧凑型零读出分类器（48–144 B），所有权重均由坐标矩阵 $\mathbf{\Theta} \in \mathbb{F}_{64}^{P\times 3}$ 合成。我们在严格相等的评估预算下（每代 N_{pop} 个候选）比较四种无梯度优化器:

- 精英坐标游走：原样保留最佳锚点，并以尺度耦合的方差进行变异（c_x，c_y），$\sigma_c^2 s^2$，并以 σ_ℓ^2 对 ℓ 进行变异。

- 反向全秩进化策略 *[27]*：采样 $N_{\text{pop}}/2$个独立同分布矩阵$[\boldsymbol{\epsilon}_i]_{p,\ j} \sim \mathcal{N}(0,\ 1)$，评估反向对$\mathbf{\Theta} \pm \boldsymbol{\sigma} \odot \boldsymbol{\epsilon}_i$，并根据适应度差异进行秩形更新。

- 受*EGGROLL*启发的秩一进化策略 *[17]*：将全秩扰动替换为秩一外积 $\mathbf{E}_i = \mathbf{u}_i\mathbf{v}_i^T (\mathbf{u}_i \sim \mathcal{N}(\mathbf{0},\ \mathbf{I}_P), \mathbf{v}_i \sim \mathcal{N}(\mathbf{0},\ \mathbf{I}_3))$，保持逐元素单位方差，同时将每个工作进程的噪声维度从 $3P$ 降至 $P+3$。

- 混合精英策略 $+$ 秩一进化策略：通过凸组合拉回 $\mathbf{\Theta} \leftarrow 0.55\,\mathbf{\Theta}^* + 0.45\,\mathbf{\Theta}_{\text{ES}}$，使秩一进化策略迭代 $\mathbf{\Theta}_{\text{ES}}$ 朝向已发现的最佳边界状态$\mathbf{\Theta}^*$，将结构化低秩更新与边界悬崖保留相结合。

模式 II（程序化储备池 + 紧凑型岭回归头）。单个24字节种子 Θ 通过公式（($\mathbf{W}$，$\mathbf{b}$）（7）一次性合成（$\mathbf{H}$ = 双曲正切（$\mathbf{X}_{\text{tr}}\mathbf{W} + \mathbf{1}_{N_{\text{tr}}}\mathbf{b}^T$）

算法1 FNS 启动时权重合成与双模式推理 输入：持久化种子 $\Theta = (c_x, c_y, \ell) \in \mathbb{F}_{64}^3$ (24 B) [或 $\boldsymbol{\Theta} \in \mathbb{F}_{64}^{P\times 3}$ 用于模式 I], 维度 (d, H)、孔径 $\mathcal{A}$、截断 $M_{\max}$、退出阈值 $R = 256$。输出：易失性静态随机存取存储器缓冲区 (W, b) 和预测结果 $\hat{y}$。阶段 A：一次性启动合成（闪存 $\rightarrow$ 静态随机存取存储器）1：设置 $s \leftarrow 10^{-\ell}$ 和 $M \leftarrow (d+1)H$。2：通过孔径 $\mathcal{A}$ 计算 $\{C_k\}_{k=0}^{M-1}$ （公式（4）–（6））。3：评估平滑逃逸势 $u_k \leftarrow \tilde{\nu}(C_k)$（公式（1）–（3））。4：通过公式（7）将 (W, b) 标准化并重塑后写入静态随机存取存储器（额外占用 0 B）。阶段 B：离线无反向传播训练 5：如果是模式 I（0 B 读出），则通过精英游走或混合秩一进化策略优化 $\boldsymbol{\Theta}$。6：否则（模式 II），拟合岭回归头 $\mathrm{W}_{\mathrm{out}} \leftarrow (\tilde{\mathrm{H}}^T \tilde{\mathrm{H}} + \lambda \mathrm{I})^{-1} \tilde{\mathrm{H}}^T \mathrm{Y}$；可选择将 $\mathrm{W}_{\mathrm{out}}$ 量化为对称 INT8。阶段 C：运行时静态随机存取存储器推理（$O(dH)$ 点积，0 次分形运算）7：计算 $\tilde{\mathrm{h}} = [\text{双曲正切}(\mathrm{W}^T \mathrm{x} + \mathrm{b});\ 1] \in \mathbb{R}^{H+1}$ 并计算得分 $\mathrm{s} = \mathrm{W}_{\mathrm{out}}^T \tilde{\mathrm{h}}$。8：如果 $C_{\mathrm{out}} = 1$，则返回 $\hat{y} = \mathbb{I}[s_1 \geq 0]$；否则返回 $\hat{y} = \arg\max_c s_c$。

并且 $\tilde{\mathbf{H}} = [\mathbf{H}\ \mathbf{1}_{N_{\mathrm{tr}}}] \in \mathbb{R}^{N_{\mathrm{tr}}\times(H+1)}$，我们求解闭式岭读出 $\mathbf{W}_{\mathrm{out}} = (\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} + \lambda \mathbf{I})^{-1} \tilde{\mathbf{H}}^T \mathbf{Y}$（其中 $\mathbf{Y} \in \{-1, +1\}^{N_{\mathrm{tr}}\times 1}$ 用于二分类，或独热 $\{0,1\}^{N_{\mathrm{tr}}\times C}$ 用于多分类）。对称 INT8 量化将 $\mathbf{W}_q$ 存储为 $\mathrm{clip}(\mathrm{round}(\mathbf{W}_{\mathrm{out}}/s_q), -127, 127)$，其中 $s_q = (\max_{i,j} |[\mathbf{W}_{\mathrm{out}}]_{i,j}| + 10^{-12})/127$（4 B）。

5 实证评估与受控消融实验

5.1 实验协议、超参数与统计严谨性

所有实验均在 Intel64 主机（Family 6 Model 78、4 个逻辑 CPU、Windows 10）上使用 Python（NumPy 2.4.6、SciPy 1.17.1、scikit-learn 1.9.1 [38]）以确定性方式执行。所有连续型数据和真实世界评估均采用严格的无数据泄漏的十折分层交叉验证（两个表格和帕累托曲线均统一采用该方法）。尤为重要的是，实验3和4中的全部四个真实世界基准测试均采用原生或固定常数输入缩放，在闪存中要求**0 个持久化缩放器字节**。我们报告平均值 $\pm$ 标准差测试准确率、宏平均 F1、精确的隐藏层/读出/总闪存字节数（包括每个量化多层感知机层用于独立权重和偏置 FP32 缩放因子的 +8B），以及在10个折上的未经校正的双侧配对 Wilcoxon 符号秩 p-值[39]（在适用情况下注明经 Holm–Bonferroni 族错误率校正后的显著性）。固定的模式 II 超参数明确披露如下：锚点 $\Theta_0 = (-0.743643887, 0.131825904, 2.15)$、$M_{\max} = 70$、$R = 256.0$；实验1使用 $H = 64$、$\gamma = 2.8$、$\lambda = 10^{-3}$、RFF 缩放因子 1.8、收敛多层感知机（ReLU、L-BFGS、$\alpha = 10^{-3}$，最大_迭代=1000）；实验3–4使用 $H = 256$、$\gamma = 1.8$、$\lambda = 0.5$、RFF 缩放因子 $1.4/\sqrt{d}$、收敛多层感知机（ReLU、Adam[40]、$\alpha = 10^{-3}$，最大_迭代=400）。

表2：实验1：在二维流形上进行严格的10折分层交叉验证（$H = 64$、收敛的多层感知机基线、 +8 B/层量化多层感知机尺度）。

数据集	模型	准确率 (%)	F1值 (%)	H / R / T (B)	p 与多层感知机相比
双月数据集	FNS-Grid (FP32)	99.80 ± 0.63	99.80	24 / 260 / 284	1.0000
	FNS-Spiral (INT8)	99.00 ± 1.25	99.00	24 / 69 / **93**	0.0312
	Linear-INT8 (0-Hid)	87.40 ± 3.41	87.38	0 / 7 / 7	0.0020
	Seeded-RFF (INT8)	99.90 ± 0.32	99.90	8 / 69 / 77	1.0000
	多层感知机 (FP32, L-BFGS)	99.90 ± 0.32	99.90	768 / 260 / 1,028	—
	多层感知机 (INT8)	99.90 ± 0.32	99.90	200 / 73 / 273	1.0000
	多层感知机 (INT4)	99.60 ± 0.97	99.60	104 / 41 / 145	0.5000
双螺旋	FNS-Grid (FP32)	96.00 ± 5.16	95.98	24 / 260 / 284	0.0625
	FNS-Spiral (FP32)	91.00 ± 6.58	90.95	24 / 260 / 284	0.0078
	FNS-Spiral (INT8)	58.50 ± 14.54	57.30	24 / 69 / 93	0.0020
	Linear-INT8 (0-Hid)	66.00 ± 7.38	65.63	0 / 7 / 7	0.0020
	Seeded-RFF（INT8）	99.50 ± 1.58	99.50	8 / 69 / 77	1.0000
	多层感知机（FP32、L-BFGS）	99.50 ± 1.58	99.50	768 / 260 / 1,028	—
	多层感知机（INT4）	95.50 ± 5.50	95.45	104 / 41 / 145	0.0625

5.2 实验1：零读出布尔逻辑与收敛基线二维流形

完全程序化布尔逻辑（0 字节读出）。在 $N_{\mathrm{runs}} = 10$ 个独立进化种子中（$N_{\mathrm{pop}} = 36$，每次成功运行均在 128×128 分辨率下验证），24 字节四象限合成解决了或（10/10 次运行，中位数 3.0 代，四分位距 1.0）、与（8/10，中位数 13.0，四分位距 11.8）、与非（10/10，中位数 10.5，四分位距 24.0）和或非（10/10，中位数 2.0，四分位距 1.8）。将三个合成的24字节种子组合起来（或和与非作为输入，连接至一个合成的与输出神经元），并采用纯连续 sigmoid 激活函数（固定增益3.0，无硬二值化），解决了非线性异或障碍 [41]，在72**B 总存储和 0B 读出的条件下达到100.**0%的**准确率，**输出概率为 $[0.020, 0.926, 0.932, 0.145]$（最小决策裕度为0.355）。

二维流形与收敛基线的比较。表2和图5(a)–(b)报告了在双月数据集（$N = 1000, \sigma = 0.10$）和双螺旋 [42]（$N = 200, \sigma = 0.04$）上的十折交叉验证结果，采用 $H = 64$。在双月数据集上，FNS-Grid（FP32）达到 99.80% $\pm$ 0.63%（284 B，较收敛多层感知机 p =1.000），FNS-Spiral（INT8）则以 93 B 达到 99.00% $\pm$ 1.25%（比 0 隐藏层 Linear-INT8 高 +11.60 个百分点，后者为 87.40%；同时仅用其 11.1× 更小的闪存空间，就达到收敛 FP32 多层感知机的 99.90%）。在双螺旋上，FNS-Grid（FP32）达到 96.00% $\pm$ 5.16%（比 Linear-INT8 的 66.00% 高 +30.00 个百分点，$p = 0.0625$ 而收敛多层感知机为 99.50%）。同样重要的是，表2显示了两项如实呈现的局限：（i）在双螺旋上，激进的 INT8 读出量化会使 FNS-Spiral 的性能大幅下降（从 91.00% →降至 58.50%），这表明高曲率缠绕边界需要 FP32 读出层；（ii）在固定且未经优化的隐藏层条件下，8 字节的带种子的伪随机数生成器随机傅里叶特征以 77 B 达到与收敛多层感知机相当的性能（99.90 % / 99.50%）。

5.3 实验 2：生成器消融（$\partial\mathcal{M}$ 与伪随机数生成器、傅里叶和晶格噪声的比较）

一个核心的基线问题是： $\partial\mathcal{M}$ 与其他 24 字节确定性生成器有何区别：以 Θ 为密钥的 SHA-256 伪随机数生成器、固定的平滑二维傅里叶基（6 个谐波，0 查找表），以及固定的 64×64 C^1 平滑阶跃值噪声晶格（16 KB FP32 查找表）。我们在$\ell\in\{0,\dots,8\}$ 周围评估了全部四种方法 c_{Mis}（$M_{\max}(\zeta)=70+40\ell$，每个尺度 12 条轨迹），并测量：（a）应用公式（7）后的标准化 16×16 有效矩阵秩；（b）局部坐标相关性 $\rho(\Theta,\Theta+\boldsymbol{\delta})=\text{corr}(\text{u}(\Theta),\text{u}(\Theta+\boldsymbol{\delta}))$，针对微步$\boldsymbol{\delta}=(0.02s,0,0)$；以及（c）一个 48 字节的零读出连续域漂移跟踪任务：在双月数据集旋转过程中，对四种生成器采用相同的零字节 Z 分数标准化（$\theta\in\{0°,15°,30°,45°,60°\}$（$N_{\text{runs}}=10$，每次漂移 10 个在线步），并保留从每个已发现 $\theta=0°$ 锚点施加 1% 微步扰动后的结果）。

图 2 和表 3揭示了三种明显的结构性差异：

1. 标准化后的 C^1 泰勒斜坡坍缩：对于任意 C^1 生成器 g，缩放至 $s\to0$会得到 $g(c_0+s\text{p}_k)=g(c_0)+s\nabla g(c_0)^T\text{p}_k+O(s^2)$。即使根据式（7）除以 $\sigma(\text{u})$，标准化后的补丁也会退化为矩阵秩为 ≤2 的平面斜坡 $\hat{}\,\text{g}^T\text{p}_k$。实证结果表明，平滑傅里叶（秩为 $_{\text{eff}}=2.26$）和平滑步进插值的固定晶格值噪声（2.25，受有限空间带宽限制）均会在 $\ell\ge3$ 时坍缩至此界限，使 48B 漂移跟踪的准确率分别降至 53.54% 和 66.51%。

2. **尺度不变的未归一化方差 $\partial\mathcal{M}$：**归一化方差 Var($\tilde{\nu}$) 会随$\Theta((\log\zeta)^{-2})$ 缓慢下降，这是因为除以了 $M_{\max}(\zeta)$（1.82×10^{-3}在 $\ell=6$，$>2.5\times10^8\times$ 高于晶格噪声且$>1.3\times10^9\times$ 高于傅里叶），而未归一化的逃逸时间方差 $M_{\max}(\zeta)^2\text{Var}(\tilde{\nu})$ 在六个完整数量级（$\ell\in\{2,\dots,8\}$）范围内保持 $[173.6,194.1]$ 不变，使标准化后的有效秩维持在 9.99–10.01。

3. **伪随机数生成器雪崩效应与平滑扰动保留率：**SHA-256 伪随机数生成器保持较高有效秩（12.58），但表现出零局部连续性（$\rho=0.019\pm0.069$）。将已发现的 48B 锚点沿坐标步长扰动 1%，会使伪随机数生成器的准确率瞬间崩溃至 50.60% $\pm$ 20.18% **（机会水平）**，而 FNS 仍保持 89.15% $\pm$3.06% （+38.55 **个百分点的稳定性优势**）。相反，我们透明地报告称，在这个仅含 33 个参数的二维问题上，每个漂移步执行 120 次盲随机抽样，仍可使伪随机数生成器随机搜索达到 89.88%$\pm$3.00%，而 FNS 为 87.07% $\pm$2.46%。因此， $\partial\mathcal{M}$ 的独特优势在于兼具多尺度有效秩和 局部流形连续性，且无需查找表。

表 3：实验 2：多尺度生成器属性（$\ell=6$，$\zeta=10^6$）以及 48B 标准化零读出跟踪（$N_{\text{runs}}=10$）。

生成器（24B）	秩$_{16}$ (ℓ=6)	$\rho(\boldsymbol{\delta})$ (ℓ=6)	1% 步骤回报	漂移准确率	LUT
二次 $\partial\mathcal{M}$	**9.99 ± 0.30**	**0.842 ± 0.043**	**89.15 ± 3.06%**	87.07 ± 2.46%	**0 B**
SHA-256 伪随机数生成器	12.58 ± 0.31	0.019 ± 0.069	50.60 ± 20.18%	89.88 ± 3.00%	0 B
平滑傅里叶	2.26 ± 0.00	1.000 ± 0.000	29.25 ± 0.35%	53.54 ± 0.20%	0 B
晶格噪声	2.25 ± 0.00	1.000 ± 0.000	47.35 ± 17.01%	66.51 ± 5.41%	16 KB

表4：实验3：四个孔径下的谱条件化与10折分层交叉验证准确率（$H=256$，相同的24B种子）。

数据集（d）	孔径	准确率（%）	有效秩 秩利用率	|相关性|	p与网格比较	
Ion球面（34）	笛卡尔网格	91.17 ± 5.10	22.82	67.1%	0.324	—
	3×3 三分法则	90.87 ± 4.64	25.29	74.4%	0.422	1.0000
	ϕ-沃格尔螺旋	**91.71 ± 5.94**	23.19	68.2%	0.353	0.9590
	谐波三脚架	90.02 ± 5.61	**27.55**	**81.0%**	**0.226**	0.3438
最优数字（64）	笛卡尔网格	**97.38 ± 1.42**	36.15	56.5%	0.340	—
	3×3 三分法则	97.16 ± 0.89	43.48	67.9%	0.421	0.6445
	ϕ-沃格尔螺旋	96.33 ± 1.34	37.23	58.2%	0.351	0.0625
	谐波三脚架	97.22 ± 1.11	**45.37**	**70.9%**	**0.221**	0.6406
手写数字（16）	笛卡尔网格	96.65 ± 1.56	12.98	81.1%	0.328	—
	3×3 三分法则	**98.20 ± 0.89**	13.46	84.1%	0.415	**0.0117**
	ϕ-Vogel螺旋	97.15 ± 0.82	12.97	81.1%	0.354	0.5273
	谐波三脚架	97.90 ± 0.84	**14.42**	**90.1%**	**0.260**	**0.0234**
Satimage（36）	笛卡尔网格	86.35 ± 1.90	24.28	67.5%	0.316	—
	3×3 三分法则	86.65 ± 1.75	27.46	76.3%	0.416	0.6992
	ϕ-沃格尔螺旋	85.45 ± 2.33	24.34	67.6%	0.353	0.4766
	谐波三脚架	**87.25 ± 1.89**	**28.79**	**80.0%**	**0.226**	0.1797

5.4 实验3：四种几何孔径与谱条件化

我们在 $H=256$ 使用相同的固定24字节种子，对四个零缩放器字节真实世界数据集的全部四种采样孔径进行了评估（图 3 和表 4），计算有效秩 $\text{rank}_{\text{eff}}(\mathbf{W})=\exp(-\sum_{j=1}^{r}p_j\ln p_j)$ [36]、非对角列相关系数绝对值均值，以及相对于笛卡尔网格基线的配对 Wilcoxon pp 值。

呈现出两种一致的实证模式：

- **谐波三脚架在全部四个数据集上最大化谱条件化：**交错三个同心缩放层级实现了最高的有效秩利用率（电离层 81.0%、光学数字 70.9%、手写数字 90.1%、Satimage 80.0%——在光学数字上相对笛卡尔网格的相对秩增益最高达 +25.5%），以及最低的平均绝对列相关性（0.221–0.260，而笛卡尔网格为 0.316–0.340，单尺度 ϕ-沃格尔螺旋为 0.351–0.354，其略高的相关性反映了斐波那契滞后副螺旋对齐）。

- **多类别轨迹上统计显著的准确率提升：** 在 UCI 手写数字数据集（d =16，$C=10$）上，互素的 3×3 三分法则（98.20% $\pm$ 0.89%，$p=0.0117$）和谐波三脚架（97.90% $\pm$ 0.84%，$p=0.0234$）均显著优于笛卡尔网格（96.65% $\pm$ 1.56%）。在 Statlog Satimage（$C=6$）上，谐波三脚架达到最高准确率（87.25% $\pm$ 1.89%，$p=0.1797$），而在电离层和光学数字数据集上，四个孔径的表现均无统计显著差异（$p\ge0.0625$）。

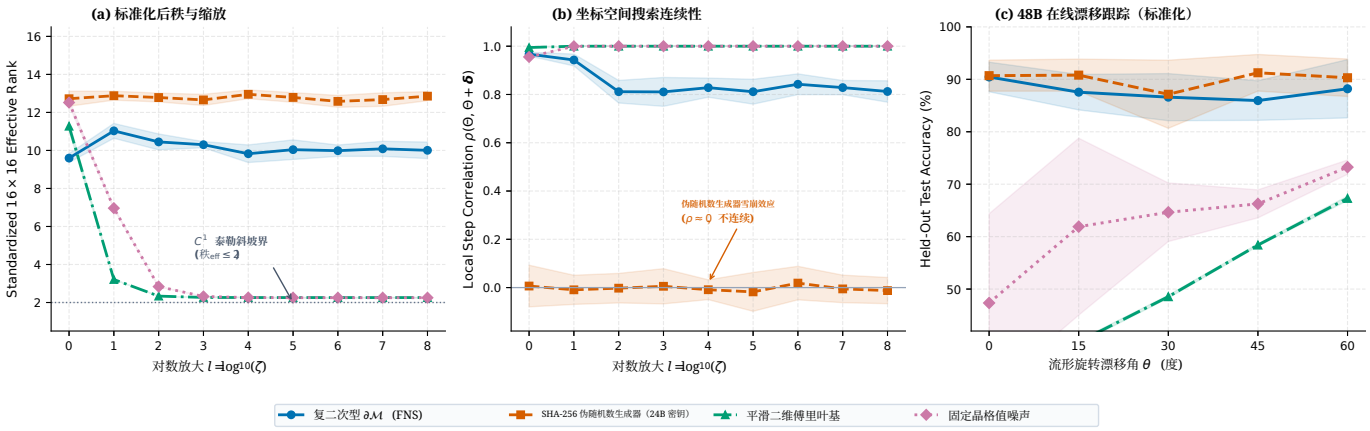

图 2：实验 2 生成器消融（24 字节种子预算，采用相同的 Z 分数标准化）。(a) 标准化后，16×16 在 $\ell \in [0, 8]$上的有效秩：平滑傅里叶和固定晶格噪声均降至C^1 泰勒斜坡界限（秩$_{\text{eff}} \approx 2.25$），而 $\partial\mathcal{M}$ 则稳定保持在 ≈ 10.0。(b) 局部步长相关性 $\rho(\Theta,\ \Theta + \delta)$：SHA-256 伪随机数生成器表现出雪崩不连续性（$\rho \approx 0$），而 $\partial\mathcal{M}$ 保留了平滑可导航性（$\rho \in [0.81, 0.97]$）。(c) 在旋转后的双月流形上进行 48B 零读出在线漂移跟踪（$N_{\text{runs}} = 10$，均值 $\pm$ 标准差）。

5.5 实验 4：零缩放器真实世界基准测试与公平基线

在不进行按数据集选择的情况下锁定统一的谐波三脚架孔径（H =256，24 B 隐藏层种子），我们评估了四个真正非线性、零缩放器字节的真实世界基准测试：UCI 电离层雷达（N =351，$d = 34$，$C = 2$，原生 $[-1, 1]$）、UCI 光学数字（$N = 1797$，$d = 64$，$C = 10$，固定 $X/8 - 1$）、UCI 手写数字（$N = 2000$ 从 10,992 个样本分层抽取，$d = 16$，$C = 10$，固定$X/50 - 1$）以及 Statlog Landsat 卫星图像（$N = 2000$从 6,430 个样本分层抽取，$d = 36$，$C = 6$，预归一化多光谱波段）。我们将其与零隐藏层线性 INT8、8 字节带种子的伪随机数生成器随机傅里叶特征（INT8）、8 字节带种子的伪随机数生成器 ELM（INT8）、显式随机傅里叶特征（FP32）、固定二维 SplitMix64+ 符号哈希未训练哈希极限学习机（256 B 隐藏层），以及已收敛的 FP32/INT8/INT4 多层感知机进行比较（表5及图4中的统一十折帕累托曲线）。

表格中呈现出三个明确的实证发现5 和图 4:

1.相较于线性 INT8 模型，非线性性能显著提升： 在所有基准测试中，FNS-Tripod (INT8) 的表现均在统计学上显著优于 0-hiddenLinear- INT8：在电离层上高出 +3.98 个百分点（90.30% 对比 86.32%，$p = 0.0254$），在光学数字上高出 +3.62 个百分点（97.22% 对比 93.60%，$p = 0.0020$），在手写数字上高出 +10.10 个百分点（97.90% 对比 87.80%，$p = 0.0020$），在 Satimage 上高出 +10.75 个百分点（87.00% 对比 76.25%，$p = 0.0020$——在全部三个多分类任务中，经 Holm–Bonferroni 校正后仍保持显著）。

2. 接近多层感知机的准确率（ 10.7×–129× 总压缩率： FNS-Tripod（INT8）的准确率与收敛后的稠密FP32多层感知机相差不超过0.70–2.00个百分点，同时将隐藏层闪存占用降至 725×–2,773×（24 B，而基线为17.4–66.6 KB），总闪存占用降至 10.7×（手写数字，$p = 0.1094$，在统计上无法区分），（续） 28.1×（Satimage）、29.6×（光学数字）以及 129×（电离层，$p = 0.2031$，无法区分）。与INT4多层感知机相比，FNS-Tripod（INT8）仍小 1.34×–16.2×，同时INT8读出层几乎没有精度损失（光学数字和手写数字上的差异为0.00个百分点；电离层上的差异为 +0.28 个百分点）。

3.与带种子的伪随机数生成器随机傅里叶特征/极限学习机及二维哈希极限学习机进行坦诚比较： 纳入公平的程序化基线揭示了一个重要的系统层面细节。在电离层上（90.30% 对比 88.87%，$p = 0.5312$）和 Satimage 上（87.00% 对比 86.95%，$p = 0.9844$），FNS-Tripod (INT8) 略胜或持平于带种子的伪随机数生成器随机傅里叶特征；而在光学数字上（−1.05 个百分点，p =0.0078）和手写数字上（−0.90 个百分点，$p = 0.0078$），8 字节的带种子伪随机数生成器随机傅里叶特征仅使用少 16 B 的闪存，表现便略优于固定种子的 FNS；此外，正确实现的二维 SplitMix64 哈希极限学习机也具有竞争力（88.89%–98.55%），但隐藏存储量更大（256 B）。因此，在固定种子的模式 II 中，基于计数器的伪随机数生成器储备池仍是一种强有力的轻量级替代方案；FNS 的差异化优势在于无需伪随机数生成器的解析确定性（不受跨编译器伪随机数生成器状态漂移的影响），以及模式 I 中连续坐标空间的可导航性（实验 2）。

5.6 实验 5：等预算多补丁优化扩展

评估一个 6 补丁（18维，总计 144B，读出为 0B）的网络（$2 \to 40 \to 1$），在严格的等预算条件下，对分层双月数据集进行评估，每代进行32次候选评估（60 代，$N_{\text{runs}} = 10$；图5（c））：

- **精英坐标游走** 达到 86.00% $\pm$ 2.22% 的测试准确率（BCE 0.2691，1,751 毫秒），显著

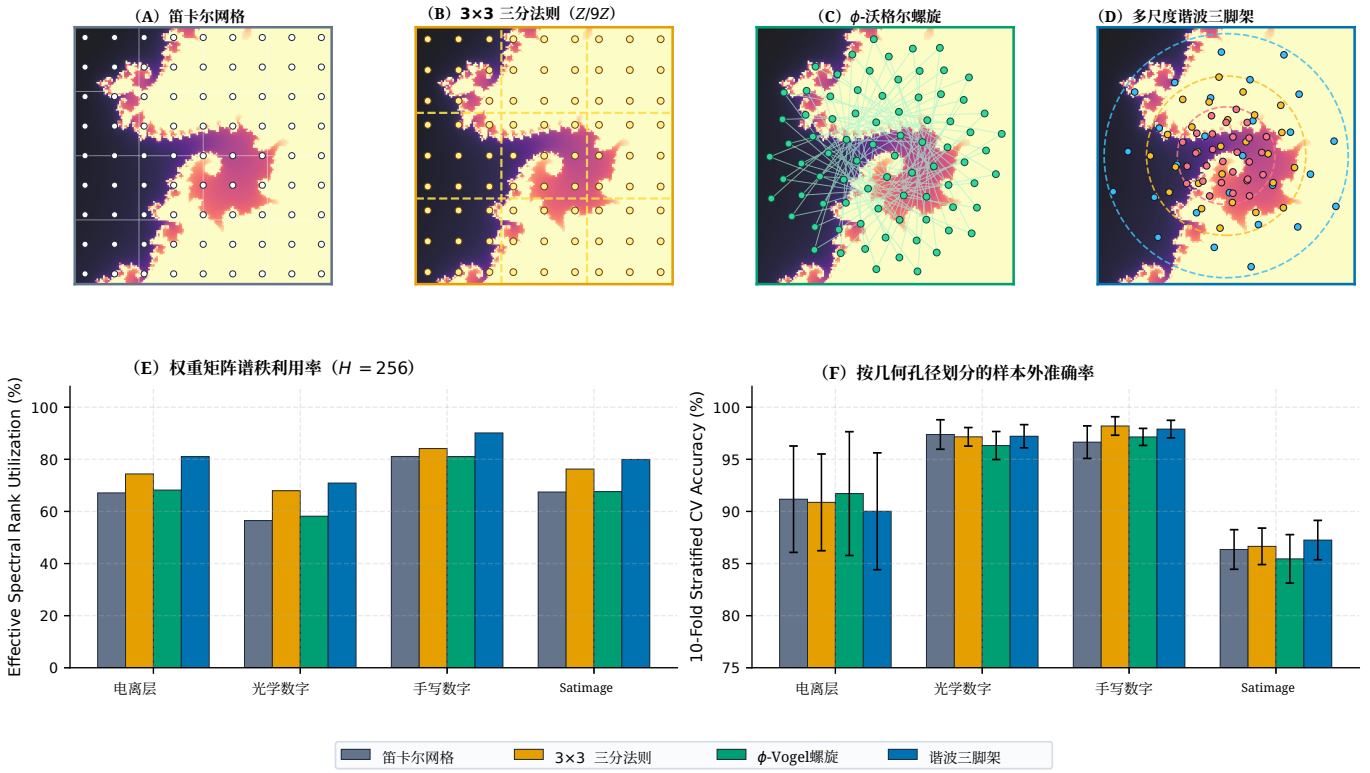

图3：实验3 几何孔径消融实验（$H = 256$，相同的24 B种子）。上部（A–D）：叠加在海马谷逃逸时间场（$R = 256$）上的查询模板。下部：（E）有效谱秩利用率和（F）在四个零缩放器字节真实世界基准测试上的10折分层交叉验证测试准确率。

优于反向全秩进化策略（75.00% ±10.26%，BCE 0.4455, $p = 0.0039$）。

- **EGGROLL 风格秩一进化策略**（75.10% ± 11.01%，二元交叉熵损失0.4622，1,488毫秒）在统计上与反对称全秩进化策略（$p = 0.5156$）持平，证实即使采用大逃逸平滑，对称 $\pm\epsilon$ 平均在陡峭的$\partial\mathcal{M}$边界跃迁附近仍会产生较大方差。
- **混合精英策略 + 秩一进化策略**干净利落地弥合了这一差距，在相同的实际运行速度下（1,502毫秒，$p = 0.0801$ 相较于全秩进化策略，$p = 0.1094$ 相较于精英游走），将测试准确率提升至80.90% ± 11.13%，并使二元交叉熵损失降低29%（$0.4622 \to 0.3278$）。

5.7 预热启动合成与公平对比的静态随机存取存储器性能分析

图6 报告了预热后（预热 10 次 + 50 计时运行）的主机 CPU 性能分析结果，参数数量为 $H = 256$。启动合成随参数数量增加而平稳、单调地扩展$M = (d + 1)H$：8.05 ± 1.88 毫秒（手写数字数据集$d = 16$），12.11 ± 0.56毫秒（电离层数据集$d = 34$），15.64 ± 4.00毫秒（Satimage 数据集$d = 36$），以及23.29 ± 3.74 毫秒（光学数字数据集$d = 64$）。使用完全相同的纯 NumPy 单样本（$b = 1$）前向传播计时器，FNS（13.61–20.96 μ秒）与稠密 FP32 多层感知机（10.78–23.29 μ秒），证实权重物化到易失性存储器后 FLOPs 相同，运行时 $O(dH)$（摊销后，批处理模式下为每个样本 5.58–8.41 μ秒）。如图6(b) 所示，一次性启动合成在 50 μ秒的主机 CPU 时间范围内可在 $N_{\text{cycles}} \approx 280$–750 次推理内完成摊销；直接在裸机 ARM Cortex-M 上进行周期与能耗基准测试仍是未来一项重要工作。

6 相关验证与部署背景

虽然本文严格聚焦于 FNS 的数学基础、消融实验和机器学习基准测试，但专门的配套研究已考察核心二次核在下游部署中的约束：对有限迭代逃逸半径不变性的机器检验 Lean 4 证明（$|z_n| > 2 \Rightarrow |z_{n+k}| \to \infty$，仅使用 0 个 sorry 和标准公理完成验证），以及 40 核裸机非统一内存访问扩展性 [43]；在 System-2 语言模型之前进行亚毫秒级的 System-1 自然语言分诊门控 [44]；以及位精确的定点以太坊虚拟机区块内断路器 [45]。这些研究共同说明了 FNS 强调无查找表、无伪随机数生成器的确定性权重合成的原因。

表5：实验 4：在四个零缩放器字节非线性基准（$H = 256$）上进行严格的10折分层交叉验证。显示隐藏层、读出和持久化闪存总字节数的精确值（缩放器开销为0 B）。 p_{MLP}、 p_{Lin}和 p_{RFF} 表示与稠密FP32多层感知机、Linear-INT8及带种子的伪随机数生成器随机傅里叶特征（INT8）相比，未经校正的双侧配对 Wilcoxon 符号秩检验 p值。

数据集	模型	准确率（均值 ± 标准差）	宏平均 F1	隐藏层（B）	读出（B）	总计（B）	p_{MLP}	p_{Lin}	p_{RFF}
UCI 电离层数据集 ($N = 351,\ d = 34,\ C = 2$) 原生 $[-1, 1]$ （0 B）	**FNS-Tripod（INT8）**	$90.30 \pm 5.09\%$	88.65%	**24**	261	**285** (129×)	0.2031	**0.0254**	0.5312
	FNS-Tripod（FP32）	$90.02 \pm 5.61\%$	88.22%	**24**	1,028	1,052	0.2031	0.0430	0.7422
	Linear-INT8 (0-Hid)	$86.32 \pm 4.84\%$	83.62%	0	39	39	0.0039	—	0.1875
	Seeded-RFF (INT8)	$88.87 \pm 4.38\%$	86.83%	8	261	269	0.0391	0.1875	—
	种子化极限学习机（INT8）	$90.32 \pm 3.34\%$	88.79%	8	261	269	0.0938	0.0703	0.3438
	哈希极限学习机（256B，2D）	$88.89 \pm 3.91\%$	86.89%	256	1,028	1,284	0.0625	0.1797	0.9922
	稠密多层感知机 (FP32)	$92.30 \pm 3.83\%$	91.24%	35,840	1,028	36,868	—	0.0039	0.0391
	量化多层感知机（INT8）	$92.30 \pm 3.83\%$	91.24%	8,968	265	9,233	1.0000	0.0039	0.0391
	量化多层感知机（INT4）	$92.58 \pm 4.10\%$	91.56%	4,488	137	4,625	1.0000	0.0039	0.0234
UCI 光学数字 ($N = 1797,\ d = 64,\ C = 10$) 固定 $X/8 - 1$ (0 B)	**FNS-Tripod (INT8)**	$97.22 \pm 1.11\%$	97.20%	**24**	2,574	**2,598** (29.6×)	0.0039	**0.0020**	0.0078
	FNS-Tripod (FP32)	$97.22 \pm 1.11\%$	97.20%	**24**	10,280	10,304	0.0039	0.0020	0.0078
	Linear-INT8 (0-Hid)	$93.60 \pm 1.33\%$	93.53%	0	654	654	0.0020	—	0.0020
	Seeded-RFF（INT8）	$98.27 \pm 1.00\%$	98.26%	8	2,574	2,582	1.0000	0.0020	—
	种子化极限学习机（INT8）	$98.22 \pm 0.97\%$	98.21%	8	2,574	2,582	0.7773	0.0020	0.5938
	哈希极限学习机（256B，2D）	$97.44 \pm 0.95\%$	97.44%	256	10,280	10,536	0.0078	0.0020	0.0117
	稠密多层感知机 (FP32)	$98.39 \pm 0.81\%$	98.38%	66,560	10,280	76,840	—	0.0020	1.0000
	量化多层感知机（INT8）	$98.33 \pm 0.87\%$	98.32%	16,648	2,578	19,226	1.0000	0.0020	0.8750
	量化多层感知机（INT4）	$97.94 \pm 0.53\%$	97.93%	8,328	1,293	9,621	0.0391	0.0020	0.2969
UCI 手写数字 ($N = 2000,\ d = 16,\ C = 10$) 固定 $X/50 - 1$ (0 B)	**FNS-Tripod（INT8）**	$97.90 \pm 0.97\%$	97.89%	**24**	2,574	**2,598** (10.7×)	0.1094	**0.0020**	0.0078
	FNS-Tripod（FP32）	$97.90 \pm 0.84\%$	97.90%	**24**	10,280	10,304	0.0469	0.0020	0.0039
	Linear-INT8 (0-Hid)	$87.80 \pm 2.46\%$	87.53%	0	174	174	0.0020	—	0.0020
	Seeded-RFF（INT8）	$98.80 \pm 0.95\%$	98.80%	8	2,574	2,582	0.5938	0.0020	—
	种子化极限学习机（INT8）	$98.50 \pm 0.75\%$	98.50%	8	2,574	2,582	0.6406	0.0020	0.1562
	哈希极限学习机（256B，2D）	$98.55 \pm 0.93\%$	98.54%	256	10,280	10,536	1.0000	0.0020	0.4062
	稠密多层感知机（FP32）	$98.60 \pm 0.74\%$	98.58%	17,408	10,280	27,688	—	0.0020	0.5938
	量化多层感知机（INT8）	$98.60 \pm 0.74\%$	98.58%	4,360	2,578	6,938	1.0000	0.0020	0.5938
	量化多层感知机（INT4）	$98.40 \pm 0.57\%$	98.38%	2,184	1,293	3,477	0.2188	0.0020	0.1875
Statlog Satimage ($N = 2000,\ d = 36,\ C = 6$) 预归一化波段（0 B）	**FNS-Tripod（INT8）**	$87.00 \pm 1.94\%$	83.21%	**24**	1,546	**1,570** (28.1×)	0.0215	**0.0020**	0.9844
	FNS-Tripod（FP32）	$87.25 \pm 1.89\%$	83.46%	**24**	6,168	6,192	0.0586	0.0020	0.8691
	Linear-INT8（0-Hid）	$76.25 \pm 2.04\%$	60.54%	0	226	226	0.0020	—	0.0020
	Seeded-RFF（INT8）	$86.95 \pm 1.91\%$	83.11%	8	1,546	1,554	0.0391	0.0020	—
	种子化极限学习机 (INT8)	$88.05 \pm 1.76\%$	84.99%	8	1,546	1,554	0.3184	0.0020	0.0762
	哈希极限学习机（256B，2D）	$87.90 \pm 2.33\%$	84.59%	256	6,168	6,424	0.1699	0.0020	0.5547
	稠密多层感知机（FP32）	$88.80 \pm 2.21\%$	86.56%	37,888	6,168	44,056	—	0.0020	0.0391
	量化多层感知机（INT8）	$88.75 \pm 2.36\%$	86.51%	9,480	1,550	11,030	1.0000	0.0020	0.0430
	量化多层感知机（INT4）	$87.35 \pm 3.22\%$	84.89%	4,744	779	5,523	0.1406	0.0020	0.7656

7 局限性、认识论范围与未来方向

明确界定 FNS 并非最优的场景，有助于实现审慎采用：

1. **信息论上限：** 根据命题1，每个24字节的种子最多携带192 比特的条件信息。FNS 适用于紧凑型 System-1 分类器、边缘控制器和确定性护栏，而非取代高熵的前沿变换器。

2. **读出层和特征缩放器存储：** 模式 II 将闪存缩放从 $O(dH + HC)$ 转移到 $O(1 + HC)$；当 C 较大时，后者占总闪存的主要部分（在 $H = 256$下，10 类数字数据集的读出层为 2,574 B，而隐藏层为 24 B）。此外，缺少固定输入范围的数据集需要存储每个特征的缩放统计量（$8d$ B FP32）；全连接多层感知机可以将其折叠到 $\mathbf{W}_1$，中，而 FNS 则必须持久化保存这些统计量，除非使用原始输入或恒定缩放的输入。

3. **模式 II：带种子的伪随机数生成器储备池**：当隐藏层种子未经优化时，基于 8 字节计数器的伪随机数生成器随机傅里叶特征/极限学习机储备池在闪存占用少 16 B 的情况下，在光学数字和手写数字数据集上取得了略高的准确率。未来的模式 II 工作应研究 Θ^* 沿$\partial\mathcal{M}$ 进行的内折坐标优化，以测试边界连续性是否能带来数据自适应储备池增益，而这种增益无法通过不连续的伪随机数生成器种子实现。

4. **曲率敏感读出量化与目标微控制器性能分析：** 双螺旋数据集的 INT8 性能下降（91.0% $\rightarrow$ 58.5%）表明，紧密缠绕的边界需要 FP32 或混合精度读出层；此外，主机CPU上的Python计时还必须通过在实体 ARM Cortex-M 和 RISC-V 微控制器上进行原生 C/Q16.16 定点基准测试来补充。

更广泛的研究展望。 展望本文评估的紧凑型分类基准之外的领域，无查表边界范式自然引出了四个有待专门未来研究深入探讨的长期研究方向：(i) 将纯整数、无反向传播的低秩进化策略 [17] 扩展到更大型的多层程序化架构；(ii) 在去中心化多智能体集群和嵌入式科学仪器中实现低带宽策略同步，其中紧凑的 24 字节状态描述符可通过受限信道进行传输；

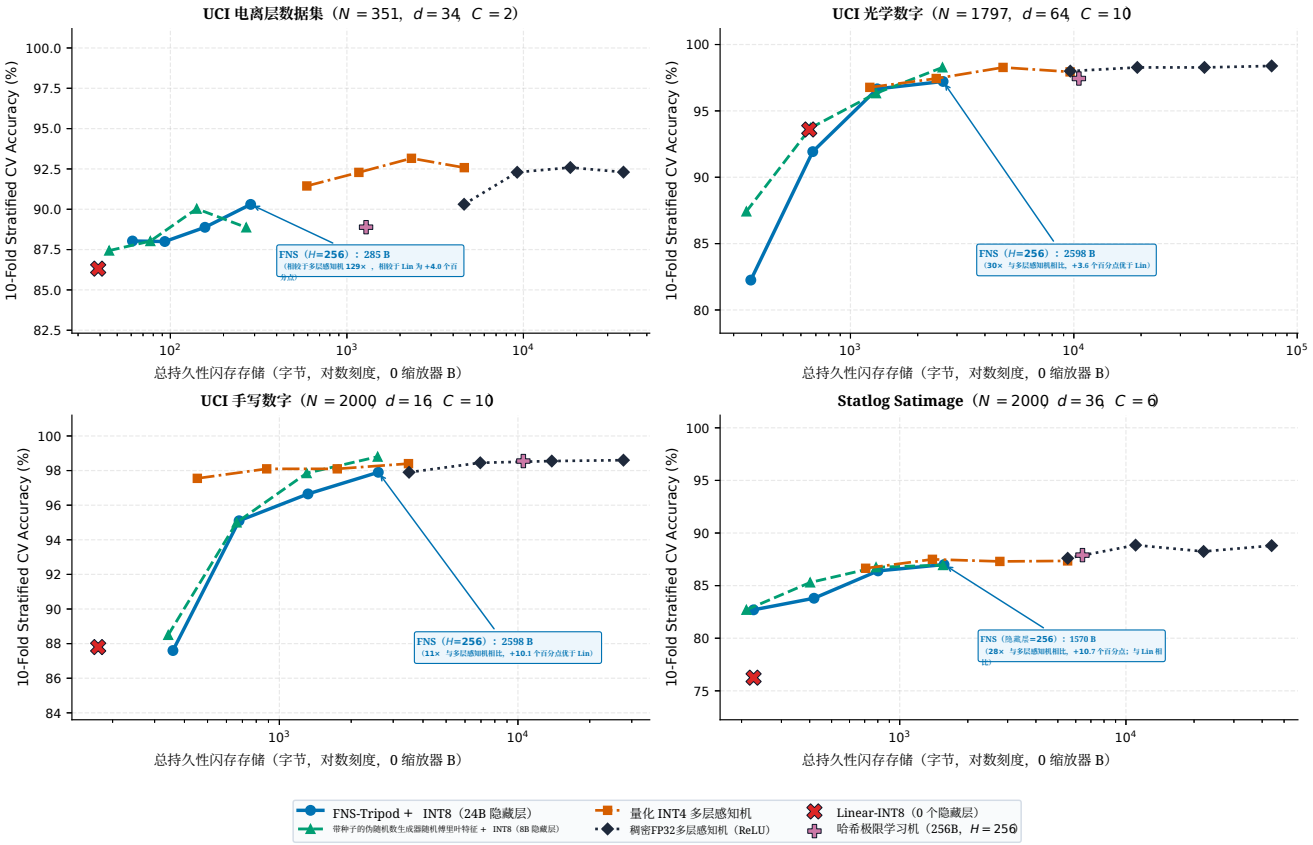

图 4：实验 4：四个非线性真实世界基准测试的统一 10折分层交叉验证内存-准确率帕累托前沿（$H \in \{32, 64, 128, 256\}$，缩放器开销为 0 B）。与已收敛的稠密FP32多层感知机相比，FNS-Tripod + INT8 的表现显著优于 Linear-INT8（+3.6 至 +10.8 个百分点），并实现了 10.7×–129× 的闪存总量缩减（隐藏层缩减 725×–2,773× ）。

；(iii) 开发能够在无需迭代数字循环的情况下计算边界积分权重的无源模拟光学或光子波干涉评估器；以及 (iv) 通过在双层边缘反射至执行架构中采用不会导致性能下降的分析参数先验，缓解合成模型坍缩。

8 结论

我们提出并系统评估了**分形神经合成（FNS）**，这是一种常量存储的程序化隐藏层推导架构，基于大退出阈值（$R = 256$）复二次边界动力学和四种几何采样孔径。通过对布尔逻辑、二维流形以及四个零标度字节非线性现实基准（$d = 16$–64、 $C = 2$–10）进行受控多尺度消融和统一的 10 折分层交叉验证，我们得出了三项结论：（1）在 24 字节无查找生成器中， $\partial\mathcal{M}$独特地兼具无尺度的标准化后矩阵秩（≈ 10.0 跨越 $\zeta \in [10^0, 10^8]$，而 ≈ 2.25 适用于 C^1

泰勒展平傅里叶／格基）与平滑的局部连续性（$\rho \in [0.81, 0.97]$；微步为 1% 时仍保持 89.15% 的准确率，而 SHA-256 伪随机数生成器为 50.60%）；（2）互素的 3×3 三分法则和多尺度谐波三脚架孔径显著改善了谱条件和多类别泛化，且不增加存储成本；（3）将 24 字节程序化隐藏层与 INT8 线性头配对，相较于零隐藏层线性 INT8，准确率高出 +3.6 至+10.8 个百分点；与收敛的 Float32多层感知机相比，在运行时 SRAM 延迟相同的情况下，总闪存用量减少 10.7×–129× （隐藏层缩减725×–2,773× ）。

致谢

支持更广泛验证套件的计算基础设施和服务器资源由 ITouch Systems（土耳其阿达纳的丘库罗瓦大学科技园）提供。作者谨此感谢 ITouch Systems 合伙人 Bekirhan Dağlı 对基础设施支持和资源分配所作的贡献。

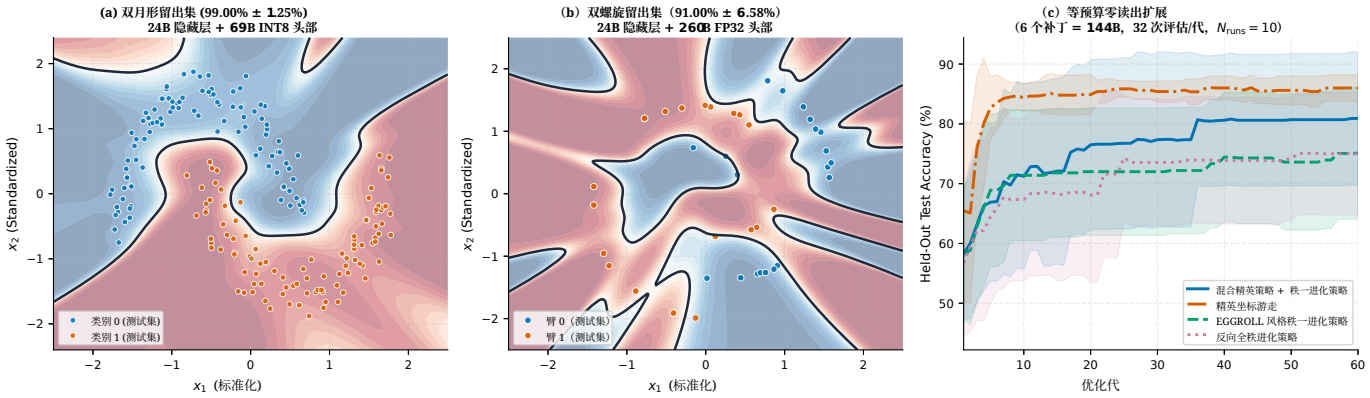

图 5：实验 1 和 5。（a）留出集双月轮廓（24 B 隐藏层 + 69 B INT8 头部，99.00% ± 1.25% 十折交叉验证）。（b）留出集双螺旋轮廓（24 B 隐藏层 + 260 B FP32 头部，91.00% ± 6.58%）。（c）严格等预算（每代 32 次评估）多分区零读出扩展（6 个分区 = 144 B，$N_{\text{runs}} = 10$）：混合精英策略 + 秩一进化策略相比纯秩一进化策略恢复了+5.80 个百分点，而精英游走产生的方差最小。

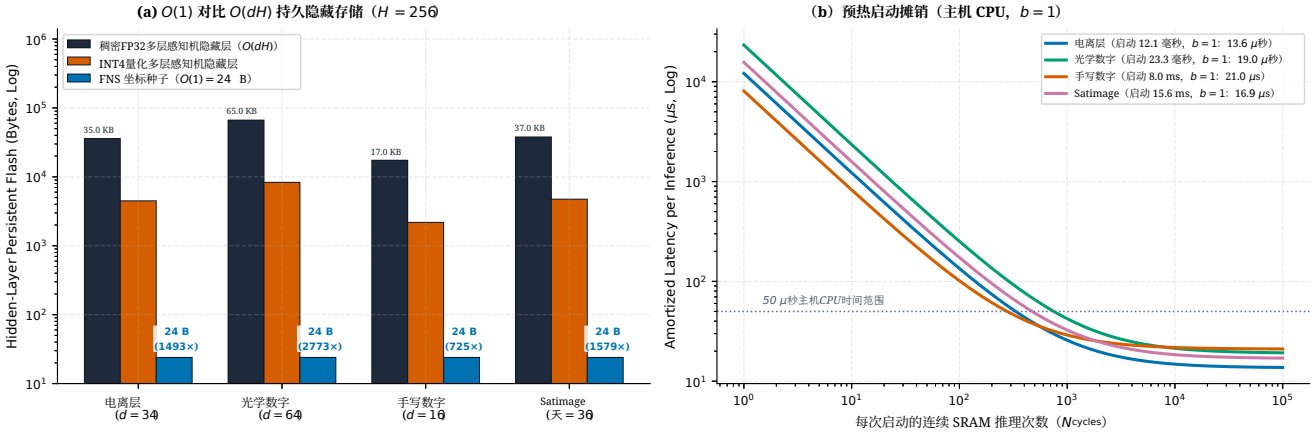

图 6：实验 4 的系统性能分析（$H = 256$，Intel64 主机）。（a）持久化隐藏层闪存（24 B 恒定 FNS 种子，与 $O(dH)$ FP32/INT4 多层感知机相比，725×–2,773× 缩减）。（b）使用同等条件下的单样本静态随机存取存储器推理进行预热启动时间摊销（$\tau_{\text{boot}}/N_{\text{cycles}} + \tau_{\text{SRAM},b=1}$）。

资金支持

本研究未获得公共部门、商业部门或非营利部门资助机构提供的任何专项拨款；ITouch Systems 仅为非商业性开放科学研究免费提供了计算基础设施。

竞争利益与知识产权披露声明

Volkan Dağlı 是 ITouch Systems 的联合创始人及研究人员，也是向土耳其专利商标局（TÜRKPATENT）提交的待审专利申请（TR 2026/016285、TR 2026/016633 和 TR 2026/020266 ）的发明人，这些申请涉及程序化边界权重合成和确定性边缘执行架构。其余作者声明没有其他利益冲突。本文报告的所有基准测试脚本、确定性种子和数据集均已公开发布，以供独立学术验证。

数据与代码可用性

所有基准测试脚本、确定性随机种子、原始 JSON/CSV 遥测数据和矢量图生成器均已通过 CERN Zenodo（https://doi.org/10.5281/zenodo.22774934）和 SSRN 摘要 7659198（https://papers.ssrn.com/abstract=7659198）公开存档。所有真实世界数据集（UCI 电离层数据集、光学数字、手写数字和 Statlog Satimage）均为标准公开基准，可通过 `scikit-learn` 和 OpenML 以可复现的方式获取。

参考文献

[1] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、Ł. Kaiser 和 I. Polosukhin，《注意力就是一切》，载于*Advances in NeuralInformation Processing Systems*，第 30 卷，2017年，第 5998–6008 页。

[2] J. Lin、W.-M. Chen、Y. Lin、J. Cohn、C. Gan 和 S. Han，《MCUNet：在物联网设备上实现微型深度学习》，载于*Advances in Neural Information ProcessingSystems*，第 33 卷，2020 年，第 11711–11722 页。

[3] *STM32L476xx 超低功耗 ARM Cortex-M4 微控制器数据手册*，意法半导体，2021 年，文档编号 025976，修订版 8。

[4] S. Ma、H. Wang、L. Ma、L. Wang、W. Wang、S. Huang、L. Dong、R. Wang、J. Xue 和 F. Wei，《1比特大语言模型时代：所有大语言模型均为 1.58 比特》，arXiv:2402.17764，2024 年。

[5] J. Frankle 和 M. Carbin，《彩票假设：寻找稀疏且可训练的神经网络》，载于国际学习表征会议 *(ICLR)*，2019年。

[6] Q. V. Le、T. Sarlós 和 A. J. Smola，“Fastfood：以对数线性时间近似核展开”，载于第 *30* 届国际机器学习会议（*ICML*）论文集，第 28 卷，2013 年，第 244–252 页。

[7] W. Chen、J. T. Wilson、S. Tyree、K. Q.Weinberger 和 Y. Chen，“使用哈希技巧压缩神经网络”，载于第 *32* 届国际机器学习会议（*ICML*）论文集，第 37 卷，2015 年，第 2285–2294 页。

[8] B. B. Mandelbrot，“复数 λ and z 的迭代的分形方面”，纽约科学院年刊，第 357 卷，第 1 期，第 249–259 页，1980 年。

[9] ——，《自然的分形几何》。纽约：W.H. Freeman and Company，1982年。

[10] A. Douady 和 J. H. Hubbard，《复多项式的动力学研究，第一部分和第二部分》，丛书《奥赛数学出版物》。奥赛：巴黎南大学，1984年，第84-02 / 85-04卷。

[11] H.-O. Peitgen 和 P. H. Richter，《分形之美：复动力系统的图像》。柏林、海德堡：施普林格出版社，1986年。

[12]M. Shishikura，“曼德博集及朱利亚集边界的豪斯多夫维数”，《数学年刊》，第147卷，第2期，第225–267页，1998年。

[13] C. E. Shannon，“通信的数学理论”，贝尔系统技术杂志，第 27 卷，第 3 期，379–423 页，1948年。

[14]M. Li 和 P. M. B. Vitányi，柯尔莫哥洛夫复杂度及其应用导论，第 4 版。Cham：Springer，2019年。[15] V. Dağlı、Z. Dağlı 和 D. Dağlı，“分形神经合成：源代码、实证基准与验证套件”，Zenodo，2026年，另存档于 SSRN 摘要 7659198。 [在线]。可访问：https://doi.org/10.5281/zenodo.22774934

[16] H. Vogel，“构造向日葵花盘的更好方法”，《数学生物科学》，第44卷，第3–4期，第179–189页，1979年。

[17] B. Sarkar、M. Fellows、J. A. Duque、A. Letcher、A. L. Villares、A. Sims、C. Wibault、D. Samsonov、D. Cope、J. Liesen、K. Li、L. Seier、T. Wolf、U. Berdica、V.Mohl、A. D. Goldie、A. Courville、K. Sevegnani、S. White- son 与 J. N. Foerster，《超大规模下的进化策略》，arXiv:2511.16652，2025年。

[18] K. O. Stanley、D. B. D’Ambrosio 与 J. Gauci，《用于演化大型神经网络的基于超立方体的编码》，人工生命，第 15 卷，第 2 期，第 185–212 页，2009年。

[19] D. Ha、A. Dai 和 Q. V. Le，《超网络》，载于国际学习表征会议（*ICLR*），2017年。

[20] G. Larsson、M. Maire 和 G. Shakhnarovich，《Frac- talNet：无残差的超深神经网络》，载于国际学习表征会议（*ICLR*），2017年。

[21] G.-B. Huang、Q.-Y. Zhu 和 C.-K. Siew，《极限学习机：理论与应用》，神经计算，第 70 卷，第 1–3 期，第 489–501 页，2006 年。

[22] G.-B. Huang、L. Chen 和 C.-K. Siew，《使用具有随机隐藏节点的增量式构造前馈网络实现通用逼近》，*IEEETransactions on Neural Networks*，第 17 卷，第 4 期，第 879–892 页，2006 年。

[23] A. Rahimi 与 B. Recht，《用于大规模核机器的随机特征》，载于神经信息处理系统进展，第 20 卷，2007 年，第 1177–1184 页。

[24] D. J. Sutherland 与 J. Schneider，《关于随机傅里叶特征误差的研究》，载于第 *31* 届人工智能不确定性会议（*UAI*）论文集，2015 年，第 862–871 页。

[25] V. Sitzmann、J. N. P. Martel、A. W. Bergman、D. B. Lindell 和 G. Wetzstein，《使用周期激活函数的隐式神经表示》，载于*Advancesin Neural Information Processing Systems*，第 33 卷，2020 年，第 7462–7473 页。

[26] C. Li、H. Farkhoor、R. Liu 和 J. Yosinski，《测量目标函数景观的内在维度》，载于国际学习表征会议 *(ICLR)*，2018 年。

[27] T. Salimans、J. Ho、X. Chen、S. Sidor 和 I.Sutskever，“进化策略作为强化学习的可扩展替代方案”，arXiv:1703.03864，2017年。

[28] J.Milnor，《单复变数动力学》，第3版，数学研究年刊系列。新泽西州普林斯顿：普林斯顿大学出版社，2006年，第160卷。

[29] R. Shadmehr、M. A. Smith 和 J. W. Krakauer，“运动控制中的误差校正、感觉预测与适应”，《神经科学年度评论》，第33卷，第89–108页，2010年。

[30]H.-O. Peitgen、H. Jürgens 和 D. Saupe，《混沌与分形：科学的新前沿》。纽约：施普林格出版社，1992年。[31] L. Tan，“曼德博集与 Julia 集之间的相似性”，数学物理通讯，第 134 卷，第 3 期，第 587–617 页，1990年。[32] A. Hurwitz，“关于用有理分数近似表示无理数”，数学年鉴，第 39 卷，第 2 期，第 279–284 页，1891年。[33] Y. LeCun、L. Bottou、G. B. Orr 和 K.-R. Müller，“高效反向传播”，载于神经网络：实用技巧，计算机科学讲义丛书。施普林格出版社，1998年，第 1524 卷，第 9–50 页。[34] G. Cybenko，“通过 S 形函数的叠加进行逼近”，控制、信号与系统数学，第 2 卷，第 4 期，第 303–314 页，1989年。[35] M. Leshno、V. Y. Lin、A. Pinkus 和 S. Schocken，“具有非多项式激活函数的多层前馈网络可以逼近任意函数”，神经网络，第 6 卷，第 6 期，第 861–867 页，1993年。[36] O. Roy 和 M. Vetterli，“有效秩：有效维数的度量”，载于第 *15* 届欧洲信号处理会议（*EU-SIPCO*）论文集，2007年，第 606–610 页。[37] D. H. Wolpert 和 W. G. Macready，“优化的无免费午餐定理”，*IEEE* 进化计算汇刊，第 1 卷，第 1 期，第 67–82 页，1997年。[38] F. Pedregosa、G. Varoquaux、A. Gramfort、V. Michel、B. Thirion、O. Grisel、M. Blondel、P. Prettenhofer、R. Weiss、V. Dubourg、J. Vanderplas、A. Passos、D. Cournapeau、M. Brucher、M. Perrot 和 É. Duchesnay，“scikit-learn：Python 中的机器学习”，《机器学习研究杂志》，第 12 卷，第 2825–2830 页，2011年。

[39] J. Demšar，“多个数据集上分类器的统计比较”，《机器学习研究杂志》，第 7 卷，第 1 期，第 1–30 页，2006年。

[40]D. P. Kingma 和 J. Ba，“Adam：一种用于随机优化的方法”，载于国际学习表征会议（*ICLR*），2015年。

[41] M. Minsky 和 S. A. Papert，《感知机：计算几何导论》。马萨诸塞州剑桥：MIT出版社，1969年。

[42] K. J. Lang 和 M. J. Witbrock，《学习区分两个螺旋》，载于 *1988* 年连接主义模型暑期学校论文集，1988 年，第 52–59 页。

[43] V. Dağlı、Z. Dağlı 和 D. Dağlı，《通过边界动力学实现零存储程序化神经合成：Lean 4 中的形式化验证与裸机严苛验证》，arXiv:2609.33066，2026年。[44] ——，《通用分形自然语言决策图：异构领域中的实时边缘分诊》，arXiv:2609.25498，2026年。[45] ——，《Werracle：面向以太坊虚拟机智能合约的低于一美分区块内 AI 快速预言机与闪电贷断路器》，arXiv:2609.30719，2026年。