一种用于大型强子对撞机鲁棒异常检测的量子启发式脉冲变分自编码器

Dean Brand, [1, 2, *] Donovan Slabbert, [1, 2] Chanelle Matadah Manfouo, [1] 和 Francesco Petruccione[2, 3] [1] *Department of Physics, Stellenbosch University, Stellenbosch, 7600, South Africa* [2] *National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, 7600, South Africa* [3] *School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, 7600, South Africa*（日期：2026年11月25日）

大型强子对撞机上的新物理模型无关搜索会采用无监督异常检测，量子和神经形态探测器是潜在的解决方案。我们开发了首个全脉冲变分自编码器——一种基于量子启发式漏积分发放（QLIF）神经元构建的量子启发式脉冲变分自编码器（QSVAE）。在这项工作中，我们将该神经元从基于测量的电路重新构建为可微计算单元。我们在大型强子对撞机双喷注潜在空间上，将该神经元与其他经典、脉冲和量子模型进行了基准测试。结果表明，在噪声扰动下，QLIF网络保留了0.95的干净准确率，而12种经典模型的准确率保留率则降至0.74–0.85。与对照模型的任何分数相比，QLIF异常分数对噪声扰动的响应幅度低3.5倍。我们将其与能够复现QLIF效应的周期性经典对照进行比较，从而将该效应归因于正弦^2 膜映射。这种膜映射是一种原生正则化器，可将鲁棒性与量子比特态空间所提供的模型容量解耦。我们通过多项基准测试和消融研究展示这一点，以刻画量子神经形态模型空间的性能状况。

arXiv:2611.32261v1 [quant-ph] 24 Nov 2026

I. 引言

寻找标准模型之外的物理，需要采用不依赖预先设定信号假设的方法。这些模型无关的方法不会针对预先设定的信号，而是基于标准模型背景进行训练，并标记检测器判定为异常的任何事物。因此，这属于无监督异常检测问题。该问题催生了若干社区基准测试以及大量方法学文献，涵盖基于重构的生成模型、距离与密度估计器，以及弱监督分类器 [1, 2]。

在这种设置下，已有两种计算范式成为潜在的解决方案，各自可能在效率和性能方面具有优势。量子机器学习提供了基于特征映射的核函数，这些核函数可能难以通过经典方式计算 [3, 4]，；与此同时，近期也有人提出将量子自编码器及其他量子异常探测器用于对撞机数据 [5–7]。另一方面，神经形态脉冲神经网络可在专用硬件上实现事件驱动、低功耗推理 [8, 9]，，近期也有人提出将脉冲异常探测器用于大型强子对撞机触发系统这一受延迟和功耗约束的环境 [10]。这两种方案都容易提出，却难以证实，因为模型的灵感可以源自量子或生物学，但其执行的每项操作实际上都可能通过经典方式高效模拟。因此，与强大的经典基线直接比较，对于确定这些方法在此类任务中的相对性能至关重要 [11]。近期，其他研究也报告称，与非脉冲网络相比，某些脉冲神经网络具有更强的对抗鲁棒性。已有研究提出的一些机制包括离散脉冲编码、编码选择、膜泄漏，以及发放阈值和时间窗口等结构参数 [12–15]。量子启发式脉冲检测器是否真正具有竞争力，以及其动态特性是否能在一般的脉冲鲁棒性之外带来额外价值，正是本文要探讨的问题。

我们通过一项围绕 QSVAE 构建的受控基准测试来开展这项工作。QSVAE 是首个完全脉冲式变分自编码器 [16]，构建于 QLIF 神经元之上。我们将该神经元从 Brand–Petruccione QLIF 电路 [17]改造为可微、可梯度训练的组件，并首次将其用于生成任务（图 1）。QLIF 是一种可直接替换使用的脉冲非线性，具有量子比特布居膜动力学，且可通过经典方式模拟。其经典形式是一种局部脉冲点神经元更新规则，因此可编程神经形态基底是一个合理的实现目标 [8, 9]。本文不对硬件实现作出任何声明。我们在 Belis 等人的 [5]，双喷注潜空间数据上对其进行评估，同时纳入经典浅层和深层检测器，以及已发表量子基线的复现结果。该基准构成了研究的主干，固定了数据、信号，以及我们必须实现并匹配的真正量子方法。图—

URE 2a 概述了干净数据的分布，图 2b 则预示了本文的核心结果。在公平且保持几何结构的扰动下，尽管 QSVAE 的重建能力较低，它仍保持鲁棒性，而我们的十二个经典脉冲对照组均未达到这一水平。这些对照涵盖了最接近的替代解释：在五种发放率下进行硬钳位、在三种设置下使用平滑饱和的膜电位（默认增益、QLIF 映射自身的斜率界以及匹配发放），以及采用权重衰减的 FSVAE

* dean.brand@nithecs.ac.za

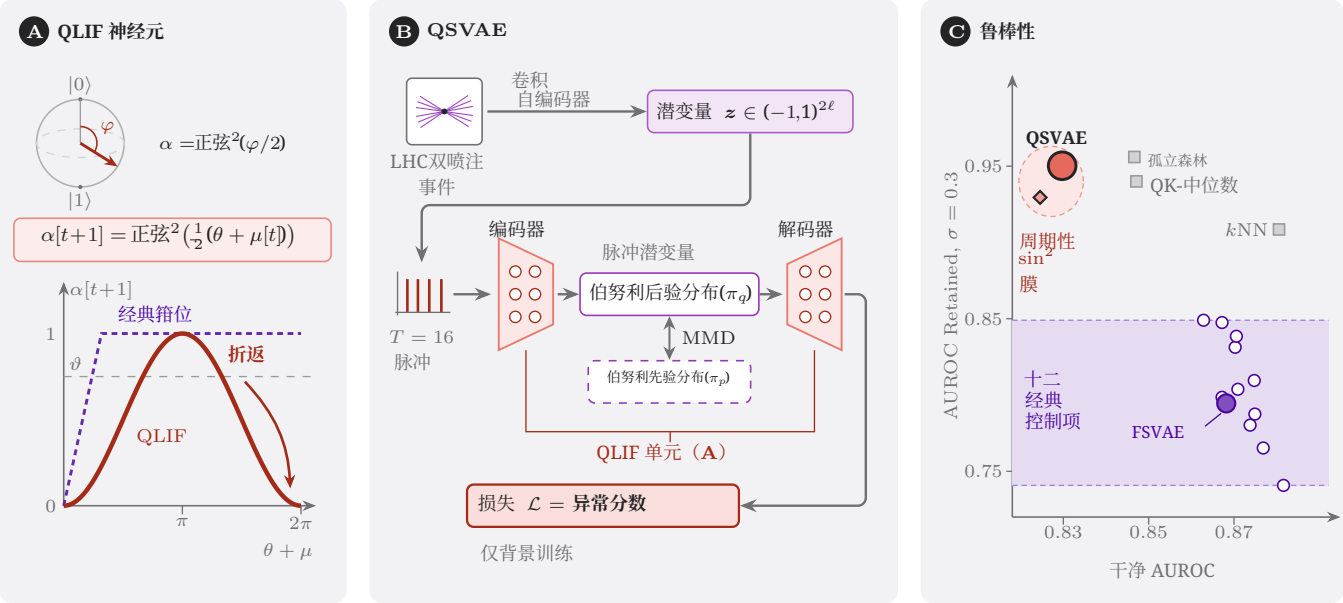

图 1. **量子启发脉冲动力学将噪声鲁棒性与模型容量解耦。(A)**QLIF 神经元将膜电位表示为量子比特的激发态布居，并通过一次旋转进行更新，该旋转结合了驱动 θ 与泄漏 $\mu[t]$，并在阈值 ϑ处发放脉冲并重置。更新映射具有周期性：强驱动会将膜电位折回至 0，而有界经典钳位则将其钳定在 1。 **(B)** QSVAE 将该神经元置于全脉冲变分自编码器 (FSVAE) 的每个脉冲单元中 [16], , 并应用于 Belis 的 LHC 双喷注潜空间 *et al.* [5]。脉冲网络无法对高斯分布进行采样，因此潜在编码本身就是一个二值脉冲序列。在 T 个步骤中的每一步，自回归脉冲网络都会根据此前的脉冲输出下一个潜在脉冲的发放概率，并从这些伯努利输出中抽取编码。最大均值差异 (MMD) 项使后验分布 Ber(π_q) 与先验分布 Ber(π_p) 匹配，以替代难以求解的 Kullback–Leibler散度。我们采用单类方法，在标准模型背景（或正常）数据上训练，其中每个事件的损失（重构加上 MMD 项）就是异常分数。 **(C)** 干净 AUROC 与保持几何结构的噪声下保留的 AUROC 之间的关系（$\sigma = 0.3$， $\ell = 16$）。12 个经典脉冲对照组（包括 FSVAE）落在同一频带内（0.74–0.85）。阴影区分了两个类别。注意：这不是置信区域。两个周期性正弦^2 膜位于该频带之上，且其干净准确率是所有脉冲模型中最低的：QSVAE 保留了 0.95，而周期性（外部泄漏）对照组将泄漏置于旋转之外，保留了 0.93。灰色标记表示非脉冲参考模型，其鲁棒性相当，因此图中所示的差异存在于脉冲模型系列内部。

将权重衰减提高到默认值的一百倍。一个经典对照确实达到了这一水平：在旋转之外施加泄漏的周期性 $\sin^2$ 膜电位，其鲁棒性为 0.93，而 QSVAE 为 0.95。鲁棒性源自 $\sin^2$ 膜映射，这是量子比特态空间通过其构造自然具备的特性。这种差异伴随着直接测得的结果：异常分数对输入扰动的响应强度降低了 3.5 倍。本文始终报告与此相反的证据：简单统计方法在干净数据上表现更好，各向同性噪声反演是噪声模型伪象，而信号衰减对照则表明，这种鲁棒性确实体现了选择性。

本文其余部分的结构如下：在第 II 节中，我们将推导并改写 QLIF 模型，使其成为可通过梯度训练的神经元，并以用于异常检测的脉冲自编码器形式呈现该模型。我们将介绍其结果，并与上述其他模型进行基准测试，展示该神经元为何具有鲁棒性且在此任务中表现出色。在第III节中，我们将进一步阐述所呈现的结果，重点指出该模型的优势与局限性，以及当前任务类别的特点。在第 IV 节中，我们将详细介绍基准测试实验所采用的方法。

II. 结果

A. QLIF 神经元

按照 Brand 和 Petruccione [17], 的做法，QLIF 神经元将其膜电位存储为单量子比特的激发态布居。将步骤 t时的量子比特状态写作 $|\psi[t]\rangle = \cos(\varphi[t]/2)|0\rangle + \sin(\varphi[t]/2)|1\rangle$，布居

$$\alpha[t] \;=\; |\langle 1|\psi[t]\rangle|^2 \;=\; \sin^2\!\big(\varphi[t]/2\big) \in [0,1] \qquad (1)$$

是测量激发态的概率，它是经典膜电位的量子类比物。每一步都通过记忆旋转 $\varphi[t] = 2\arcsin\sqrt{\alpha[t]}$ 重构前一步存储的状态，其中 $\varphi[t] \in [0,\pi]$。随后，两个操作会使量子比特演化。首先，加权输入脉冲使其旋转一个可训练的

角度 θ （积分）。经过延迟 τ 后，激发态布居会衰减至零，类似于振幅阻尼。其缩放因子为 $\lambda = e^{-\tau/T_1}$ （其中 T_1 是弛豫时间），并由衰减角 $\gamma[t] = -2\arcsin\sqrt{\lambda\,\alpha[t]}$ 表示。根据二进制输入脉冲 $X[t] \in \{0,1\}$，这会生成原始更新规则 [17],

$$\alpha[t{+}1] \;=\; \sin^2\!\Big(\tfrac{1}{2}(\theta{+}\varphi[t])\,X[t]\Big) + \sin^2\!\Big(\tfrac{1}{2}\gamma[t]\,(X[t]{-}1)\Big), \tag{2}$$

当神经元群体超过发放阈值 ϑ 时，就会发放一个脉冲，并将量子比特重置为 $|0\rangle$。

我们在一个略有不同的工作区间中实现该神经元，同时保留模型的核心行为。深层脉冲网络单元接收的输入不是单个二元脉冲，而是其突触前脉冲的连续批归一化加权和。我们将这个加权和作为旋转驱动 θ，并在每一步施加该驱动，同时始终启用泄漏，以模拟经典漏积分发放神经元 [18]的行为。通过 $\mu[t] = 2\arcsin\sqrt{\lambda\,\alpha[t]} = -\gamma[t]$将始终启用的泄漏纳入记忆角，可以把积分与泄漏合并为一次旋转，

$$\alpha[t+1] \;=\; \sin^2\!\Big(\tfrac{1}{2}\big(\theta + \mu[t]\big)\Big), \tag{3}$$

从而统一了方程（2）的两个分支。当 $\theta = 0$时，它恰好简化为 $\alpha[t+1] = \lambda\,\alpha[t]$ （纯 T_1 衰减），非零驱动则会进一步旋转神经元群体。其动力学在两方面不同于经典漏积分发放神经元：膜电位被限制在 $[0,1],$，且非线性项为周期性的 $\sin^2,$，因此强驱动会使神经元群体旋转越过阈值后再返回，而不是无限制地积分。与方程（2）一样，每当神经元群体越过阈值 ϑ时，就会发放一个脉冲，并将神经元群体重置为零，然后再形成下一步的记忆角。

这种映射的量子属性有一个限定条件。神经元仅存储激发态布居，并通过记忆旋转从中重新制备纯态。泄漏在纯态假设下复现振幅阻尼通道的布居动力学，但并未实现该通道本身。真正的振幅阻尼会将纯态变为混合态，其相干性$\sqrt{\lambda\alpha}(1-\alpha)$ 和行列式 $\rho = (1-\lambda)\lambda\alpha^2 > 0$。重构后的状态是纯态，其相干性为 $\sqrt{\lambda\alpha}(1-\lambda\alpha)$，因此标量记忆不携带相位。两者在激发态布居上完全一致，这就是为何零驱动时约化为上文的 $\lambda\,\alpha[t]$ 是精确的。此外，创建记忆角时，重置会保持 $\alpha < \vartheta$ ，因此在这些参数下，它们在驱动布居上的单步差异小于 1.1×10^{-2}。因此，方程（3）描述的是每一步都从 α 重新制备状态的神经元映射，正如原始电路在每次测量后所做的那样，而不是一个量子比特跨步持续演化的过程。出于同样的原因，标量α 形式体系仅描述独立神经元。纠缠寄存器中单个量子比特的约化态通常是混合态，无法通过任何记忆角重构。相比之下，重置作为通道表述是精确的因为将每个状态都变为 $|0\rangle$ 就是全强度的振幅阻尼。

该神经元最初是一种紧凑电路，包含两个旋转门且不含纠缠门；其中，泄漏由一个旋转实现，重现振幅阻尼引起的布居衰减，而脉冲则通过测量实现，用于浅层量子脉冲分类器的前向推理。我们将其升级为深度生成模型的可训练构件，将方程（3）视为可微且可经典模拟的映射。我们使用快速Sigmoid替代函数对网络进行反向传播，并通过精确的泄漏梯度处理循环膜

$$\frac{\partial\mu[t]}{\partial\alpha[t]} \;=\; \sqrt{\frac{\lambda}{\alpha[t]\,(1-\lambda\,\alpha[t])}}, \tag{4}$$

其在 $\alpha[t] \to 0$ 处发散。将其展开应用于深层、按时间步运行的网络时，会使时间反向传播变得不稳定，因此我们在反向传播过程中对循环乘数进行约束。反向传播规则还会通过硬重置传递有界的循环梯度，而硬重置的精确导数为零。这些反向传播近似不会改变前向动力学。经过这种稳定化处理，QLIF 神经元能够通过梯度下降，在任意深度的脉冲网络中可靠训练，从而完成从单神经元量子电路到深度学习组件的转变。

值得注意的是，其他量子脉冲神经元采用了与此方法不同的构造。Kristensen等人演化两个哈密顿量并进行局部测量，因此其神经元的输入和输出均为量子态 [19]。Chen 等人以多量子比特随机记忆替代单量子比特布居记忆，并通过局部学习规则对其进行训练；他们报告称，在图像分类基准上，其准确率高于原始的基于电路的 QLIF 网络 [20]。

B. 量子启发式脉冲自编码器

QSVAE 将 QLIF 神经元置于 Kamata等人 [16], 提出的全脉冲变分自编码器中；这是首个完全由脉冲层构成的变分自编码器。我们在此概述该模型，随后介绍替换过程，以及如何将该模型调整用于异常检测任务。

a. 全脉冲*VAE*。在该模型中，每个张量都是跨越 T 个时间步的二值脉冲序列。脉冲编码器将输入映射为空间中的脉冲特征，镜像脉冲解码器则将潜变量映射回去。编码器由带有依赖阈值的批归一化（tdBN） [21] 和经典LIF神经元的脉冲线性层构成。我们采用多层感知机变体，该变体消耗逐事件特征向量，而不是图像。对正态变量进行采样需要浮点运算，而脉冲网络不执行浮点运算，这意味着潜变量不可能是连续高斯分布。

相反，它是由自回归脉冲模型生成的伯努利脉冲序列。后验分布 $q(z_{1:T} \mid x)$ 和先验分布 $p(z_{1:T})$ 分别是自回归脉冲网络。在每个时间步，这些网络根据截至当前的脉冲，输出下一个潜在脉冲发放的概率，并从这些伯努利输出中采样潜在编码。这种构造在保持编码为二值的同时，使其具有变分性质。基于变分自编码器的标准证据下界 [22]，，FSVAE用后验与先验脉冲序列样本之间的最大均值差异（MMD）替代脉冲序列分布之间难以处理的Kullback–Leibler项[16]，

$$\mathcal{L} = \underbrace{\mathbb{E}_q\big[-\log p(x \mid z)\big]}_{\text{reconstruction}} + \underbrace{\mathrm{MMD}\big(q(z), p(z)\big)}_{\text{prior matching}}. \quad (5)$$

重构结果是从解码器在 T 个时间步内累积的膜电位输出中读出的，而不是从脉冲中读出，因此它是可与输入相比较的实值信号。

b. 从 *FSVAE* 到 *QSVAE*。QSVAE 是这样一种网络：将其中的每个 LIF 神经元替换为可微QLIF单元。由于神经元的输入和输出形状相同，这种替换在接口处是精确的。每一层已经构成了 θ =tdBN（加权输入），即依赖阈值的批归一化突触前求和，它就是式（3）中的 QLIF旋转驱动。神经元自身不携带参数，因此 QSVAE 的权重数量与 FSVAE 相当；潜变量仍为二值，因此模型始终是完全脉冲式的。网络结构完全相同，使得下文的消融能够清晰地将效果归因于神经元动力学本身。

c. 适配异常检测。我们将模型从分类适配到异常检测，方法是将其作为单类模型进行训练，使其拟合标准模型本底事件，学习重构这些事件并匹配脉冲先验分布。训练集中不包含任何信号。随后，若某个分布外事件的单事件损失较高（式（5），即重构损失加潜变量 MMD），便将其标记出来，并将该损失用作无监督异常分数。据我们所知，这是量子启发式脉冲神经元首次用于生成式模型，也是首次用于完全脉冲式变分模型。

C. 干净数据景观

我们基于 Belis 等人 [5] 的双喷注潜变量空间构建了本文的基准测试。经典卷积自编码器将每个碰撞事件压缩为维度为 ℓ的逐喷注潜在向量。我们在标准模型背景（QCD多喷注）上训练单类检测器，并在由背景和三个超出标准模型的信号组成的留出混合数据集上进行评分。这些信号包括：衰变为 WW 的窄型和宽型 Randall–Sundrum引力子，以及一个标量 $A \to HZ \to ZZZ$。需要指出的是，检测器在训练期间从未见过信号事件。在本文中，我们给出每种信号的 ROC曲线下面积（AUROC），以及在固定信号效率下的背景效率倒数；后者是对撞机搜索中所采用的优值。

在这个基准测试中，拟合检测器确实优于无模型统计量，因此比较不同拟合方法更能揭示模型动力学。在 $\ell = 16$总体表现最强的拟合检测器达到 0.891 的 AUROC（三种信号合并计算；第IVA节），而表现最好的无模型统计量为 0.791。总体差距为 0.100，在每种信号上方向一致，但幅度差异很大：窄质量引力子上的差距为 0.025（0.998 对比 0.973），.108（0 $A \to HZ$ (0.987 对比 0.879），而宽质量引力子上仅为 0.022（0.689 对比 0.667）。无模型统计量是值得参考的基线，同时又不会让这项研究变得多余。这一前提是该工作区间的特性；第IIE节表明，随着潜在维度降低，这一优势会减弱。

图 2a 和表 I 展示了干净潜在数据集上的 16 种检测器（表 I 中有 14 种；持久量子自编码器和瓶颈量子自编码器仅出现在图 2a 中）。在这一排名中，最简单的方法拔得头筹：最近邻距离（0.891）和背景的马氏距离（0.879）在所有组别中均名列前茅，而深度单类基线（深度支持向量数据描述，0.684）垫底，因此在此情形下，复杂性显然并无优势。脉冲 FSVAE 是准确度最高的学习模型之一（0.877）。它与经典自编码器和非纠缠混合量子自编码器表现相当，与最佳距离方法的差距也在 0.015 以内。量子启发式 QSVAE 排名居中（0.827）。脉冲自编码器的表现可与真正的量子方法相竞争。FSVAE 优于量子核和量子聚类检测器，而 QSVAE 则与它们不相上下。要进行最公平的比较，需将训练集大小调整至与量子核的数据预算一致；其 N^2 格拉姆矩阵将数据量限制在约 600 个事件。在这一匹配规模下，FSVAE（0.842）的表现可与纠缠量子核（0.855）相竞争，甚至在最难的信号上也不逊于它们，而且二者都远优于径向基函数单类支持向量机（0.755）。由于窄引力子信号较容易（大多数检测器的得分高于 0.99），而宽引力子信号对所有方法来说都很难（最佳得分为 0.689），因此没有任何方法能在每种信号上都胜出。汇总的总体得分可能掩盖这一现象：例如，QSVAE 在宽引力子信号上的得分仅为 0.544（接近随机水平，且低于无模型最大 $|z|$ 统计量 0.667），因此其总体得分（0.827）依赖于两个较容易的信号。我们始终报告各信号的得分（表 S4），并将在 II E 节再次讨论这一难信号。

D. 复现量子基线

本研究的跨家族比较是否公平，取决于其量子基线，因此我们在相同潜变量上重新实现了原始量子套件并进行了验证

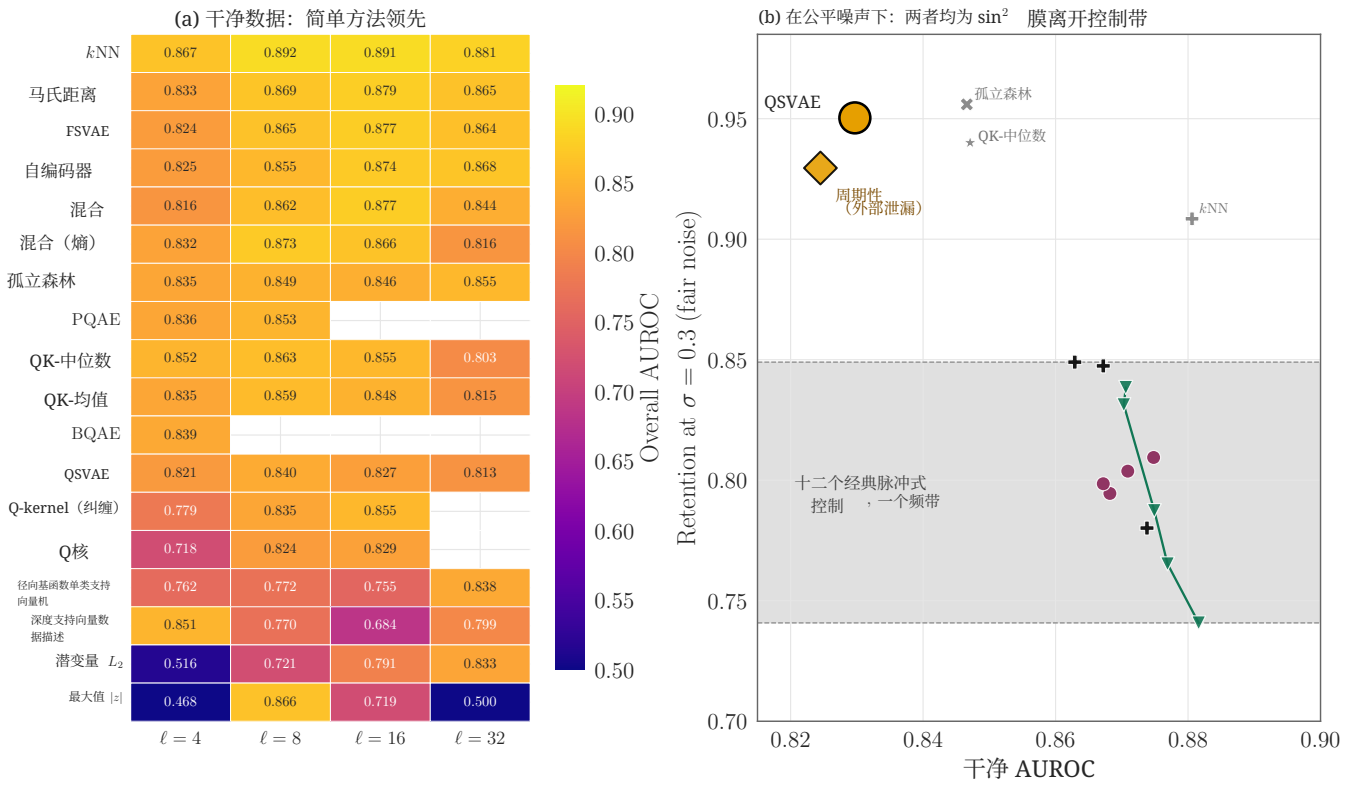

图 2. **干净数据方法比较与公平噪声鲁棒性结果。(a)**干净数据全景：在 LHC双喷注潜空间上，四种自编码器潜在维度 $\ell \in \{4, 8, 16, 32\}$（列）下每种方法（行，按排名排列）的总体 AUROC。脉冲模型使用约 13k 个本底事件进行训练（从 15k 个样本池中扣除 15% 的验证留出集），经典方法使用约 25k 个，量子核和聚类方法使用 600 个，量子自编码器使用 400 个。空白单元格表示态矢量不可行配置：量子核在 $\ell = 32$处，持久量子自编码器（PQAE）超过 $\ell = 8$后，以及瓶颈变体（BQAE）超过 $\ell = 4$后。最简单的距离方法（kNN）在每个维度上均领先，而量子启发式 QSVAE 位居中游：复杂性在干净数据上并无优势。**(b)** 核心结果见 $\ell = 16$：在标准差为 $\sigma = 0.3$的公平、保持几何结构的扰动下，第 II G 节中的 12 个经典脉冲基线神经元的保留 AUROC 全部落在一个窄频带内（阴影区域，0.74–0.85），而量子启发式 QSVAE（大型橙色圆点）和第 IIG 节的周期性机制控制（橙色菱形）则脱离该频带，分别保留 0.95 和 0.93。图中以灰色标出鲁棒参考检测器以供参考。图 7 将完整展开这一面板。

其量子核与 RBF 参考子集，并将结果与原始结果进行了对照。我们的重新实现复现了种子论文的两项主要发现。纠缠量子核优于其经典 RBF 参考（0.855 对比 0.755，取 $\ell = 16$）。随着数据重上传深度增加，量子的性能及其纠缠优势均有所提升（图 3），复现了报告中的深度缩放结果。与 Belis 等人的逐信号 AUROC 值相比，我们对量子核及其 RBF 参考的重新实现，在三个信号和两个潜在维度上的差异均不超过 0.065。最大差距出现在接近随机水平的宽质量引力子信号上，二者均约为 0.5（表S2）。有趣的是，在完整比较中，量子核优于其自身的经典基准，但不及表现最佳的经典检测器或脉冲检测器；而在最困难的信号上，它又被最大 $|z|$ 统计量击败。我们还加入了三个量子自编码器。持久型和瓶颈变体 [23] 分别达到总体 AUROC 0.836 和 0.839，取 $\ell = 4$；持久型变体达到 0.853 时 $\ell = 8$。它们位于干净数据域的中间，而宽 G 仍是这些模型最难识别的信号（0。573–0。584）。单随机种子输入噪声检查在。812（PQAE）和 0。767（BQAE）处保持稳定，而在 $\ell = 4$时，PQAE 为 0。814；当受到 $\ell = 8$的扰动时，$\sigma = 0.3$。混合变体 [24], 将量子编码器与经典解码器配对，在所有四种宽度下均保持可行，并在 $\ell = 8$ 和$\ell = 16$ （图4）处领先量子自编码器系列。它还显示出纠缠交叉：纠缠环在 $\ell \leq 8$时有所助益，在 $\ell \geq 16$时则会造成阻碍，这与核随深度增加而单调提升的优势正好相反。这是因为在高维情况下，过多的重复上传模块会使四量子比特寄存器过度混乱。

E. 对潜变量维度的依赖

基准测试的难度并非固定不变，因为自编码器的潜在维度会对其起到调节作用。

表 I. 在 $\ell = 16$ 处的 LHC 双喷注潜在空间总体 AUROC（代表性结果；各信号和各维度的表格以及匹配尺度行见补充材料表 S3 和 S4）。十六行对应十四种不同的检测器：量子核和混合自编码器均分别以有纠缠和无纠缠的形式出现。 σ_{seed} 为跨随机种子标准差（脉冲模型采用五个随机种子，量子方法在重采样训练池上采用五个随机种子，量子自编码器采用三个随机种子）；无模型方法和经典方法均仅拟合一次；对于保存了每事件得分的方法，表 S5 报告测试集上的自助法 95% 置信区间，以替代随机种子间波动。 N_{train} 为近似训练集大小，该数值因类别而异，且已予以说明：无模型统计量无需拟合，量子核受其 ∼600 个事件的 N^2 格拉姆矩阵限制，脉冲模型使用约 13k 个本底事件进行训练，经典方法则使用约 25k 个。 $1/\varepsilon_b$ 为信号效率 $\varepsilon_s = 0.8$下的背景效率倒数，即面向搜索的优值（越大越好）。

类别	方法	N_{train}	总体 AUROC ↑	σ_{seed}	$1/\varepsilon_b$@0.8 ↑
无模型	最大值 $\lvert z\rvert$	–	0.719	–	1.9
无模型	L2 能量	–	0.791	–	2.2
经典（浅层）	kNN	∼25k	0.891	–	5.6
经典（浅层）	马氏距离	∼25k	0.879	–	4.9
经典（浅层）	孤立森林	∼25k	0.846	–	3.2
经典（深层）	自编码器	∼25k	0.874	–	4.7
经典（深层）	深度支持向量数据描述	∼25k	0.684	–	1.8
脉冲	FSVAE（LIF）	∼13k	0.877	0.005	4.4
脉冲	QSVAE（QLIF）	∼13k	0.827	0.001	2.8
量子	量子核（纠缠）	600	0.855	0.003	3.7
量子	量子核（无纠缠）	600	0.829	0.004	2.8
量子	QK-中位数	600	0.855	0.002	3.7
量子	QK-均值	600	0.848	0.001	3.4
量子	径向基函数单类支持向量机（参考）	600	0.755	0.034	1.8
量子自编码器	混合量子自编码器（非纠缠）	400	0.877	0.008	4.4
量子自编码器	混合量子自编码器（纠缠）	400	0.866	0.005	4.2

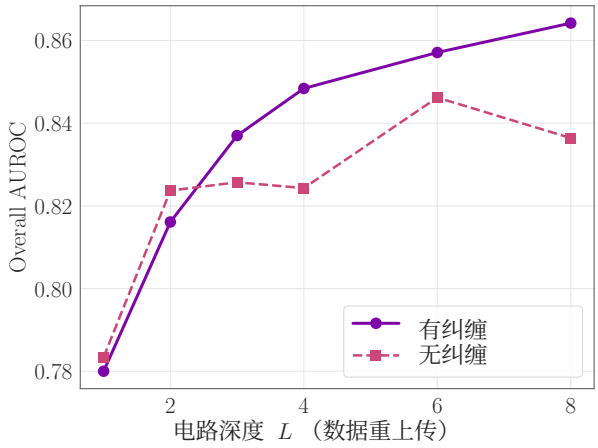

FIG.**复现量子核行为**。无监督量子核机器的总体 AUROC 随数据重上传深度 L 在 $\ell = 8$下的变化，有纠缠和无纠缠两种情况。性能和纠缠优势均随电路深度增加而提升，复现了 Belis *et al.* 报告的行为。即使在最大深度下，量子核的表现也接近但未达到最佳经典方法（kNN，在 $\ell = 8$时为 0.892）。

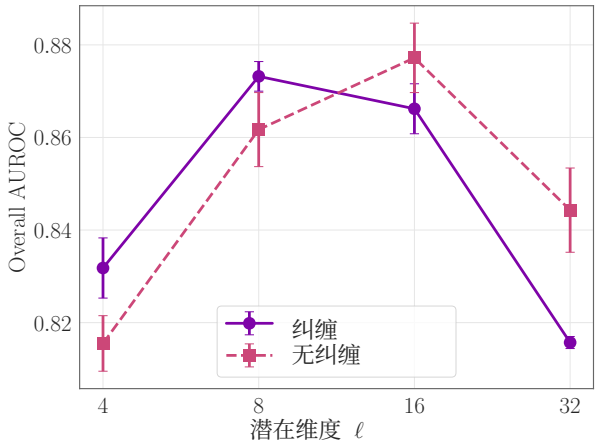

FIG.**混合量子自编码器纠缠交叉**。混合量子自编码器（固定的四量子比特数据重上传编码器搭配经典解码器）在不同潜在维度下，有无纠缠时的总体 AUROC（清理后数据）。纠缠变体在 $\ell = 8$ 处达到峰值，而非纠缠变体在 $\ell \leq 8$ 时有助于（），高维度下则有害（$\ell \geq 16$），这与量子核随深度单调提升的优势正好相反（图3）。

无模型统计量揭示了这一效应（图5）。极值坐标统计量 max $\lvert z\rvert$ 在大多数维度下表现较弱（ 0

。468（ $\ell = 4$处），0。719（ $\ell = 16$处），0。500）但在 $\ell = 8$ 处出现尖峰（0。866），该变体——

量子基线所定义、且量子硬件将用于其中的维度。在该峰值处，无模型统计量在 16 个检测器中排名第四，仅次于两种距离方法和

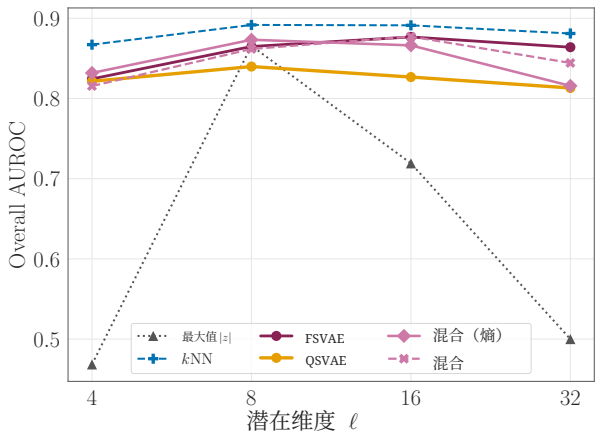

图 5. **总体 AUROC 随潜在维度的变化。** 在每个维度上，学习型方法和距离方法中的领先者（脉冲式 FSVAE/QSVAE、最佳距离方法 kNN，以及纠缠和非纠缠混合量子自编码器）的表现都相差无几，而由无模型最大 $|z|$ 统计量设定的基准在 $\ell = 8$ 处达到最佳，其他情况下表现较弱。混合量子自编码器的表现可与脉冲 VAE 媲美，并呈现出纠缠交叉现象（图 4）。

混合量子自编码器。它的排名也高于量子核（0.835）、两种脉冲式自编码器、深度自编码器和孤立森林方法。这是由于几何结构不同：狭窄的瓶颈会迫使异常事件使一个或多个双曲正切有界潜在坐标达到饱和。极值坐标统计量可直接读取这一特征，而较宽的瓶颈会将信号分散到各个坐标中，使其减弱。因此，学习式与简单方法的先决条件是表征本身的属性，而非任务的属性。最佳距离方法与最佳无模型统计量之间的差距从 $\ell = 16$处的 0.100 缩小到 $\ell = 8$处的 0.026，这限定了任何学习式检测器能够证明其复杂性合理的适用范围。

F. 公平扰动下的鲁棒性

在模拟数据上训练并应用于未见数据的检测器，必须能够容忍扰动（电子噪声、校准漂移和模型失配）。我们在干净背景上拟合每种方法，然后对受到尺度为 σ 的加性噪声污染的测试潜变量进行评分，以检验模型鲁棒性。我们发现，结果在很大程度上取决于噪声的形态。在各向同性高斯噪声下（图6，左），每个潜在坐标受到的扰动幅度相同，原本在干净数据上表现领先的方法都溃败了。我们还发现，在 $\ell = 16$ 和 $\sigma = 0.3$ 下，最近邻距离和马氏距离分别降至 0.501 和 0.485，与随机水平无法区分，仅保留各自干净准确率的 0.57 和 0.55。单看这一结果，似乎表明鲁棒性发生了逆转：最简单的距离检测器反而最脆弱；但这其实是噪声模型造成的假象。各向同性噪声的大部分能量集中在背景数据稀疏分布的方向上。与沿数据分布方向施加的等能量扰动相比，它对最近邻几何的破坏要大得多。

使用保持几何结构的噪声重复测试：噪声按背景数据的离散程度逐坐标缩放，并进行能量匹配，使其注入相同的总功率，同时将功率重新分配到数据实际变化的方向（第IVE节），结果不再出现性能崩溃（图 6，右）。在相同的扰动幅度 $\sigma = 0.3$ 下，最近邻检测器如今保留了其干净准确率的 0。91，跻身最鲁棒的方法之列，而非最不鲁棒的方法，排名逆转现象也消失了。马氏距离是唯一的例外，在两种模型下都仍然很脆弱（公平噪声下保留 0。58），因为它依据背景协方差矩阵的逆进行评分，从而放大低方差方向；在这些方向上，即使扰动与背景数据的离散程度相匹配，相对于背景自身的尺度仍然很大。由此我们得到的方法论警示是：鲁棒性排名可能因所选扰动而逆转。

以这一公平基线为参照，QSVAE 确实表现出鲁棒性。在$\sigma = 0.3$保持几何结构的噪声下，它仍保留了干净准确率的 0.95（在各向同性噪声下为 0.96），因此它是学习型脉冲检测器中最鲁棒的模型，也属于所有类型模型中最鲁棒的模型之一；其表现与量子聚类方法相当，仅略逊于孤立森林（保留 0.96）。保留率会使干净基线较低的模型占优，因此从绝对值来看，QSVAE 在鲁棒区域的 AUROC（0.79）处于领先组之内，而非位居首位（孤立森林 0.81，kNN 0.80，量子聚类 0.80）。这里的关键在于，它远超架构匹配的脉冲孪生模型 FSVAE（0.69，保留 0.80），这表明该效果定位于神经元本身。该鲁棒性既不是反演伪影，也不是经典脉冲式孪生模型所共有的性质，这引出了一个问题：QLIF 神经元究竟是什么机制产生了这种鲁棒性？下一节将通过受控消融回答这一问题。

G. 脉冲动力学的放电匹配消融

QSVAE 与 FSVAE 的区别仅在于其神经元，因此第 II F 节所述的鲁棒性源于 QLIF 神经元的某种性质。其有界 $[0, 1]$ 膜电位、其稀疏放电以及其正弦^2 旋转（式（3））是三个候选性质，我们通过一组经典脉冲控制组将它们加以区分。除神经元外，这些控制组的训练方式均与 QSVAE 完全相同（图 8 展示了相互竞争的更新映射）。钳位神经元具有 QLIF 神经元的有界 $[0, 1]$ 膜电位、阈值和泄漏，但将正弦^2 旋转替换为线性泄漏积分和硬钳位，因此其有界方式与 QLIF 神经元完全相同，但整体上是经典的。由于更稀疏的神经元在机制上具有更高的鲁棒性，我们在五种发放率下扫描钳位范围（约从 0.2 到 0.37，位于所探测的

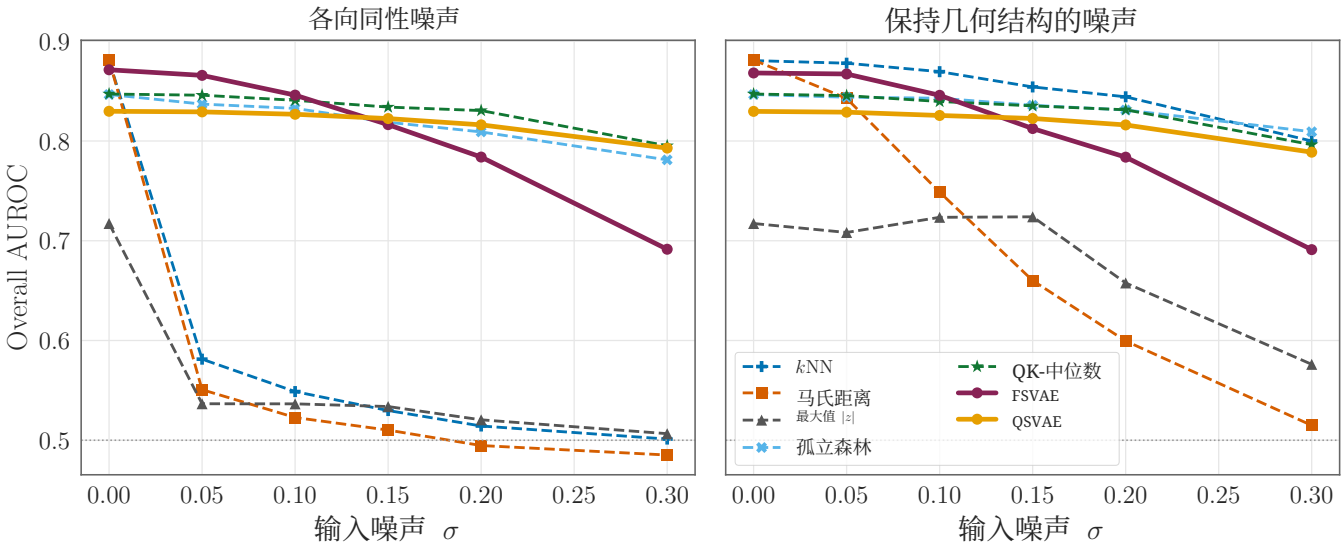

图 6. **各向同性噪声和保持几何结构的噪声下的鲁棒性。** 当测试潜变量受到 σ 的加性噪声扰动时， $\ell = 16$ 处的总体 AUROC。每个检测器均使用干净背景拟合，并在加噪测试集上评分。 *Left:* 各向同性高斯噪声会对每个潜在坐标施加相同扰动。干净数据距离领先方法 kNN 和马氏距离的表现会下降至随机水平附近。 *Right:* 保持几何结构的噪声按背景数据的离散程度逐坐标缩放，并进行能量匹配，使其携带相同的总功率，同时将功率重新分配到数据实际变化的维度。 kNN 的表现不再下降，成为最具鲁棒性的检测器之一，而马氏距离仍然脆弱。量子启发式 QSVAE 在两种噪声模型下都属于最具鲁棒性的检测器之列，而 FSVAE 则不是。

设置中）。其中一种（钳位（匹配））被调节到接近 QSVAE 的发放率（0.19 对比 0.23）。平滑饱和是下一个候选因素，因为膜电位接近其边界时敏感性趋于消失，这一性质可能正是正弦^2 映射的关键。为评估这一点，我们加入了一个平滑饱和神经元，其有界膜电位遵循逻辑曲线，而非周期性正弦^2,。我们在三种设置下进行训练，包括其默认增益、最大斜率与 QLIF 映射的全局上界二分之一相匹配时的增益，以及与 QSVAE 放电匹配的设置。为考察普通正则化是否也能达到相同状态，我们在三种权重衰减强度下训练 FSVAE，最高达到默认值的 100 倍。这十二个控制组共同独立于正弦^2 动力学，分别改变有界性、发放密度、饱和度平滑性和斜率，以及正则化强度。我们发现它们并未改变异质性，因为包括 QLIF 在内的每个变体，都对每个神经元使用一个固定的泄漏（$T_1 = 60$步数）。已知能够提升脉冲网络鲁棒性的单神经元时间常数多样性 [25] 对它们中的任何一个都不可用，包括 QSVAE。

这些对照实验推翻了用“有界膜电位”简单解释准确率的说法。有界性无需付出代价：钳位神经元的干净 AUROC 与 FSVAE 相当（在 $\ell = 16$ 下，五个随机种子的结果分别为 0.876 和 0.877），平滑饱和神经元也是如此（0.876）。在匹配发放的条件下，有界经典神经元在所有变体中准确率最高（0.883），且均明显高于 QSVAE（0.827）。因此，QSVAE 在干净数据准确率上的代价是由 sin^2 动力学特有的，既与有界膜电位无关，也与放电率无关。

对于鲁棒性分析，控制实验隔离了影响（图 7 和表 S6）。我们训练的每个经典脉冲变体都落在一个单一的窄保持率区间内，介于 0.74 和 0.85 之间。该频带中总共有十二个变体，包括 FSVAE、全部五种放电率下的钳位、全部三种设置下的平滑饱和神经元，以及全部三种正则化强度下的 FSVAE。QSVAE 脱离了该频带，保留率达到 0.95。三项比较进一步缩小了造成差异的因素范围。

发放密度无法解释上述现象。随着放电率变化，钳位沿着单调前沿变化：稀疏放电的准确度更高、鲁棒性更低（保留率最低为 0.74），而密集发放的鲁棒性更高（保留率最高为 0.84）。无论放电率如何，钳位的保留率都无法接近 0.95；而放电匹配钳位的放电甚至比 QSVAE 更稀疏，但保留率仅为 0.76。平滑饱和也没有更好的表现。平滑饱和神经元恰好具有可能被认为能带来鲁棒性的饱和处导数消失性质，但其表现仍落在钳位曲线上（0.85）。将其最大斜率匹配到 QLIF 的二分之一上限后，表现仍停留在该处（0.85）；而进行放电匹配后，表现则沿经典前沿下移（0.78）。这与匹配钳位的移动方向相同，因此在我们测试的任何斜率或放电率下，它都不足以解释这一现象。常规正则化也无法重现该结果：即使将权重衰减放大至 100 倍，FSVAE 的保留率仍为 0.80。综合考察发放密度、饱和形状和正则化强度，十二种经典脉冲对照组均未达到 QSVAE 所处的优势位置，因此这种鲁棒性归因于脉冲动力学^2 。

进一步的控制测试会检验泄漏是否必须位于

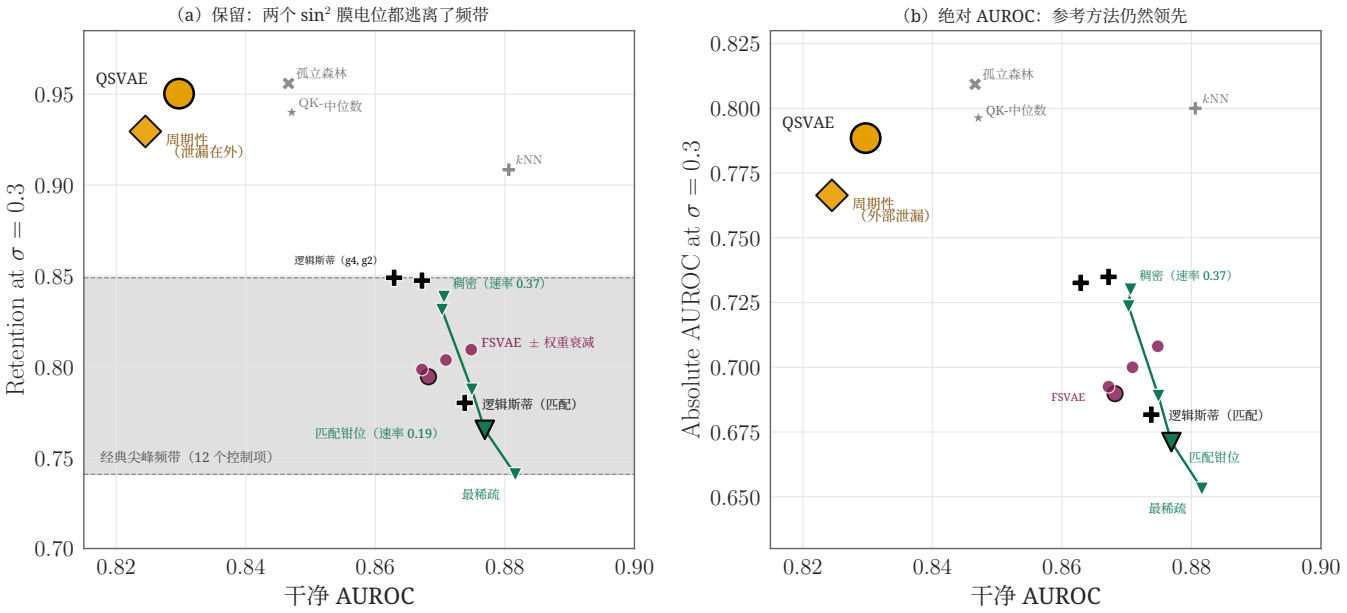

图7。**脉冲模型系列中干净准确率与噪声鲁棒性的关系**。在 $\ell = 16$ 下，几何结构保持噪声作用下扰动为 $\sigma = 0.3$ 时，干净 AUROC（横轴）与鲁棒性的关系。两个面板均限定为脉冲模型系列，窗口内的参考检测器（孤立森林、QK-中位数、 kNN）以灰色显示以供参考。**(a)** 以保留率表示的鲁棒性。十二个经典脉冲基线集中在一个窄频带内（阴影区域，0.74–0.85），其中包括默认权重衰减和三个更高权重衰减下的 FSVAE（紫色圆圈，默认值带轮廓）、从稀疏发放扫描到密集发放的钳位（绿色线，放电匹配设置带轮廓），以及默认增益、QLIF斜率上界和放电匹配条件下的平滑饱和神经元（黑色叉号）。两个变体脱离了该频带，且其干净 AUROC 为所有脉冲模型中最低：量子启发式 QSVAE（橙色圆圈），其保留率为 0.95；以及周期性（外部泄漏）对照（橙色菱形），即将泄漏移至旋转项之外的正弦膜，其保留率为 0.93。**(b)** 噪声下这些变体的绝对 AUROC：两个周期性膜领先于脉冲模型系列，而较强的参考方法位于其所在位置或略高于其位置（马氏距离和最大 $|z|$ 参考方法低于所绘制窗口），因此这一独特分组具体是相对于脉冲模型系列而言的，而不是相对于所有检测器而言。

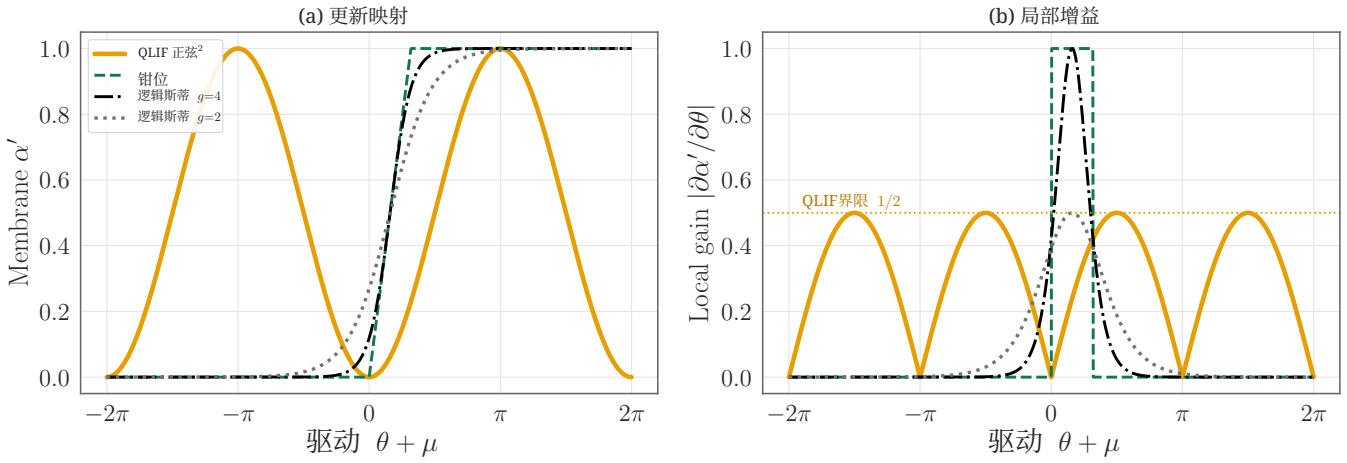

图8. **QLIF更新映射与经典控制变体的对比。(a)** 膜更新 α' 随输入驱动的变化情况，对应于式（² ）中的 QLIF sin3映射，以及作为消融对照的三种有界经典膜：硬钳位，以及增益为 4 和 2 的平滑饱和（逻辑斯蒂）神经元。这四种映射均被限制在单位区间内，其中三种平滑饱和；但只有 $\sin^2$ 映射具有周期性，在强驱动下会使膜电位折返，而非将其钉定在某个值上。**(b)** 对应的局部增益。QLIF 增益始终不超过二分之一（虚线）；钳位在其内部区间内的增益为 1；平滑饱和神经元在中心处达到 $g/4$，因此 $g = 2$ 设置与 QLIF 的界限一致。这些映射构成了消融分析的框架：有界性、平滑性、斜率和发放分别由一种对照复现，而周期性是 QLIF 神经元独有的特性。

记忆角。我们训练一个周期性膜，Periodic

（外部泄漏）中，该测试保留正弦² 旋转，但施加

将泄漏视为外部乘性阻尼因子，而不是将其折叠进记忆角。使用 QSVAE 时，它突破了频带，在 。93 的 $\ell = 16$ 和 。91 的 $\ell = 8$条件下均保持这一水平，且付出了相同的干净准确率代价（0。822，而 QSVAE 为 0。827）。因此，正弦^2 膜映射承载了这一效应，而将泄漏嵌入旋转中只会带来很小的额外优势（0。95，而另一种为 0。93）。这两种突破控制组都采用了完整映射，包括其周期性折返，以及生成该映射的纯态记忆重构，因此这项控制无法将这些组件彼此区分开来。这些控制组还将该效应与通常归因于脉冲网络的鲁棒性区分开来 [12, 15]。所有控制组都采用与 QSVAE 相同的离散脉冲编码和泄漏，放电扫描涵盖了其稀疏度，但差距依然存在。

我们通过另外三项检验进一步考察 QSVAE 的鲁棒性优势。结果并非汇聚三种信号造成的假象：在窄质量引力子、宽质量引力子和 $A \to HZ$ 上，QSVAE 分别保持了 0。97、0。97 和 0。90，且每种情况下都高于控制带。然而，在宽质量引力子上，这一保留值对应的干净 AUROC 接近随机水平，因此意义不大。保留率也会使干净基线较低的模型占优；因此，在所有类别的检测器中，孤立森林、 kNN 和量子聚类方法的鲁棒性相当（第 IIF节）。以绝对鲁棒 AUROC 衡量时，结论也成立：QSVAE 为 0。79，而脉冲控制组为 0。73 至 0。65。在 $\ell = 8$ 处，核心排序同样成立：QSVAE 保持 0。91，而默认钳位、发放匹配钳制、默认平滑饱和神经元和 FSVAE 的数值为 0。81 至 0。85。完整的放电和正则化扫描仅在 $\ell = 16$。

训练稳定性也呈现出相同趋势，不过差异较小。在 $\ell = 16$ QSVAE 的跨随机种子标准差（0.0012）约为 FSVAE（0.0048）和稠密钳制神经元（0.0051）的四分之一。使钳位的发放相匹配后，部分恢复了这一优势（0.0028），但 QSVAE 的稳定性仍是其两倍以上。这一优势仅适用于 $\ell = 16$；在 $\ell = 8$，没有任何脉冲变体表现出可靠的稳定性优势，因此不能将其视为普遍特性。

为探究这种鲁棒性的来源，我们直接测量模型对输入扰动的敏感性。对于每个训练模型，我们对测试输入施加沿几何保持方向迈出一个小步，并记录异常分数相对于其自身分布范围的变化幅度（第 IVE 节）。QSVAE 分数对这种扰动的敏感性比十二个经典脉冲对照组中敏感性最低者还低 3.5 倍；QSVAE 的归一化敏感度为 0.17，而整个对照组的数值范围为 0.60 至 0.70。在我们测试的每个探测尺度上，这种差异都成立（$\sigma \in \{0.05, 0.1, 0.2\}$）。无需归一化的读数也证实，得以保持的是分数排序，而 AUROC 正是取决于这一指标。在相同扰动下，QSVAE 的分数排名几乎不变，而所有受测对照组的排名都乱了（$\tau = 0.40$ 降至 0.51；下降 17% 至 20%）。相同的测试将这种效果定位到动力学上：发放匹配钳制是受测变体中最稀疏的，但其敏感性甚至最高（0.70 ± 0.02。这一结果出现在对照组 0.61–0.70 的范围上限处）。平滑饱和神经元的敏感性与硬钳位相当，因此，无论是发放密度还是平滑饱和，都无法解释这一现象。受限正弦^2 膜电位的重构下限也约为经典神经元的四倍（图 10），而准确率并未因此降低。这与代码受到动力学特定正则化的情况一致。至于这种压缩本身是否是实现鲁棒性的途径，我们暂不作定论。

H. 信号衰减控制

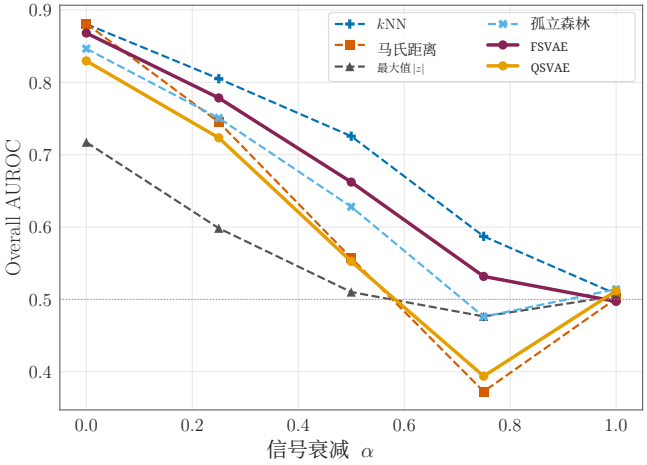

图9. **信号衰减条件下的 AUROC，位于** $\ell = 16$。当测试信号逐渐衰减至背景时，在 $\ell = 16$ 处的总体 AUROC：从真实信号（$\alpha = 0$）到与背景在统计上完全相同的“信号”（$\alpha = 1$）。随着判别结构被移除，包括 QSVAE 在内的所有检测器都失去了区分能力，并在 $\alpha = 1$ 处趋近随机水平。低于 0.5 的中间低谷与马氏距离和孤立森林分数一致，表明部分形态改变的信号事件比典型背景信号更深地嵌入背景中。对输入不敏感的检测器则会保持平稳。

在输入噪声下评分几乎不变的检测器（第IIF节）很容易被解读为一种贬低性结果：或许 QSVAE 之所以不敏感，是因为它丢弃了信息，只给出近乎恒定的评分，而这些评分恰好不受扰动影响。若真是如此，其鲁棒性便是一种退化。我们通过信号衰减对照发现了反驳这一解释的证据。该对照与噪声扫描进行相反的测试：噪声扫描会破坏输入，而该对照则移除信号。我们将每个测试信号事件逐步向背景插值，并在从真实信号（ α ）到统计特性与背景完全相同的信号（$\alpha = 0$）的衰减程度下进行评分（第$\alpha = 1$IV E节）。真正读取

判别结构必须随着$\alpha \to 1$而失去区分度，而忽略输入的检测器则不会。

QSVAE 的表现正如信号跟踪检测器所应有的那样，如图 9 所示。在衰减作用下，其 AUROC 从 0。830 急剧下降；当信号在 $\alpha = 0$ 处完全被背景取代后，降至随机水平 $\alpha = 1$，其对信号的跟踪方式与距离和经典检测器完全相同。随着判别结构被抹去，所有方法的结果都趋近于 0。5。接近$\alpha = 0.75$时低于 0。5 的中间下探，同样出现在马氏距离和孤立森林分数中，这本身表明该分数读取的是潜在几何：部分变形的信号事件落在背景内部更深处，比典型背景事件更深入，因此也按此进行排序。因此，QSVAE 对噪声的不敏感是真实的选择性。它响应的是将信号与背景区分开的物理机制，只有当这种物理机制消失时才停止响应。

III. 讨论

我们研究的核心结果是，QLIF 神经元偏离了非周期经典脉冲对照组所描绘的准确率–鲁棒性前沿（图 7）。高容量检测器能够紧密拟合干净背景并准确评分，而这种紧密拟合也使其分数在输入受到噪声扰动时发生剧烈波动。有界经典箝位继承了这一权衡：随着其向更高准确率方向强化，鲁棒性反而降低；而 QLIF 神经元则不会如此。我们通过直接测量每个模型的输入到分数的敏感度，确定了差异所在。与十二个经典脉冲控制组中敏感度最低者相比，QSVAE 的异常分数对小幅保几何输入扰动的响应低 。5 倍。其分数排序——AUROC 所读取的量——因此保持稳定，而每个对照组的分数都发生了混乱。这种低敏感度与保持差距一致，并且在定量上足以解释该差距。在经典区间内，局部探测无法分辨较小的保持差异，因此我们将其视为区分 QSVAE 与该类别的方法，而不是用于排列该类别内部模型的方法。用鲁棒性文献的语言来说，膜^2 电位充当架构性输入梯度正则化器；低测得敏感度正是双重反向传播和受 Lipschitz 约束的训练通过显式惩罚项 [26–28]所施加的量。这里，这种性质源自神经元的动力学，无需额外的损失项，并在损坏式评估方案 [29]下进行评估。

这些对照实验系统地排除了可能解释 QSVAE 鲁棒性的自然候选因素。饱和本身不足以解释这一点：平滑饱和对照模型在两个边界处同样具有消失导数，但对输入的敏感程度与硬钳位一样。增益收缩也不足以解释这一点。根据式 (3)，QLIF 单步增益处处满足$|\partial\alpha/\partial\theta| = \frac{1}{2}|\sin(\theta+\mu)| \leq \frac{1}{2}$，训练后的各层也遵循这一规律。¯这引出了利普希茨式解释，但斜率匹配的逻辑函数对照模型实现了相同的增益分布，却仍然对输入完全敏感，这表明单靠增益收缩无法解释观察到的鲁棒性。排除这些因素后，剩下的便是 sin^2 旋转独有的周期性折返；它确实在训练网络中发挥作用，同时，纯态记忆重构生成了该映射，并将泄漏置于旋转内部，作为记忆角 μ。已知基于正弦的激活函数和傅里叶特征输入编码在经典网络中具有不同寻常的归纳偏置 [30–32]。相同的 sin^2 结构使数据重上传量子模型关于其输入呈现傅里叶级数 [33], ，因此 QLIF 膜属于一个文献记载充分的类别。第 IIG 节中的周期性（外部泄漏）对照表明，泄漏的位置并非必要因素。将泄漏移到旋转之外后，神经元仍达到了 QSVAE 的鲁棒性优势区间。该对照保留了映射的其余部分，因此，产生这一效果的究竟是折返，还是与之相伴的重构，目前尚无定论。我们承认这是本研究的一项局限：鲁棒性由有界周期性膜电位所致，而经典脉冲网络也可以直接采用这种机制。量子启发的贡献在于映射本身，包括其增益界和折返；这一映射是通过认真考察量子比特动力学得出的，而非通过搜索激活函数得到的。由于大型强子对撞机的异常检测器是在模拟数据上训练、并应用于带有额外探测器噪声、校准漂移和模型失配的数据，因此，对输入扰动具有鲁棒性是一种理想特性。在这对脉冲模型中，FSVAE 在干净数据上的优势在受到扰动后不复存在，因此量子启发神经元的鲁棒性更强。消融实验表明，干净数据准确率的代价是动力学机制所致，而非调参伪象。

这一机制处于更广泛表征的一部分。我们开发了 QSVAE——首个基于升级版 QLIF 神经元构建的全脉冲变分自编码器——并在一项高能物理对撞机任务中，将其与经典、脉冲及真正的量子检测器进行了基准测试，同时在相同数据上复现了已发表的量子基线。在干净数据上，简单方法表现最佳，复杂性并不会带来额外收益。两项方法学对照对此提供了进一步洞见：各向同性噪声下表面上出现的鲁棒性反转是噪声模型伪象，而信号衰减对照则证实这种鲁棒性确实体现了真正的选择性。

此外，这些发现还受到若干局限的约束。在干净数据上，经典脉冲 FSVAE 是准确率更高的模型，而 QSVAE 并不是干净数据准确率的最先进水平竞争者。它的价值在于鲁棒性区间，在这一范围内，搜索最关注低假阳性率指标；而在干净数据上，QSVAE 是其中表现最弱的学习型检测器之一。其压缩高底限编码是直接原因，而在 $\sigma = 0.3$处的公平扰动下，这一排序会在脉冲模型系列内部发生变化。就低假阳性率指标而言，QSVAE 领先于所有受测的脉冲模型对照组，但孤立森林仍然绝对领先。没有任何方法能够在每种信号上都胜出

，而在最接近背景的信号上，一个无模型统计量与所有方法相比都具有竞争力；我们也将此视为此类研究中一个值得注意的有趣方面。

区分 QSVAE 的鲁棒性类似于正则化和压缩型检测器（如孤立森林和量子聚类方法）的鲁棒性。这些方法同样具有鲁棒性。QLIF 神经元的独特之处在于，它在我们测试的每一种非周期性经典脉冲神经元都遵循的前沿之外，达到了这一工作区间。还有一种脉冲替代方案不属于控制类别，即采用不同的单神经元时间常数；这能够提高经典脉冲网络的鲁棒性 [25]。然而，这种多样性不可能是 QSVAE 自身的机制，因为所有 QLIF 神经元都共享同一个固定泄漏。具有学习得到的异质泄漏的经典脉冲检测器，或许仍可通过这一途径达到相当的鲁棒性区间，而测试这种检测器是控制类别的自然扩展。如果可以运行孤立森林，就应当加以使用。不过，在检测器必须发放脉冲的场景中，结果才真正重要，即在神经形态硬件所面向的、受延迟和功耗约束且毗邻触发器的设置中 [8–10]。

我们使用能量匹配且遵循几何结构的高斯噪声来建模分布偏移。这是一种合理但经过简化的代理模型：在真实情况下，能标系统会在整个事件内以及整个测试群体中施加一致的偏移，而我们按事件、按坐标进行的抽样无法捕捉这一点。比较还仅涵盖一个数据集和一种学习得到的表征。我们没有检验异常分数阈值是否会塑造双喷注不变质量谱。这是模型无关搜索的标准可信度检查，因为缓存的潜变量不保留重建该谱所需的组成粒子运动学信息。该检查连同探测器级系统误差，属于专门的部署研究。真正的量子基线是在单个处理器上进行态矢量模拟的，这将核限制为十六个量子比特，并且在潜在维度最大时不再纳入该核。

在整个工作中，我们明确说明哪些内容涉及量子，哪些不涉及量子。QLIF 神经元是量子启发式的，因为其膜电位遵循纯态假设下的单量子比特布居动力学，并且它是一种硬件可实现的脉冲非线性；然而，我们执行的所有操作都可由经典方法模拟，也不声称有计算加速。核方法和自编码器基线都是真正的量子机器学习模型，并在相同数据和计算资源下相互比较。在这一框架内，复现的量子核方法如原报告所述，优于其自身的经典基准，但不及最强的经典或脉冲检测器。其优势也取决于潜在维度。我们的贡献在于公正地定位这些方法，并对量子启发神经元进行机制明确的表征，而非展示任何量子优势。

这项工作带来了三个研究方向。QLIF 神经元的更新规则是一种标量递推，可由可编程神经形态基底合理实现，尽管它并非原生的 LIF 基元。一个自然的下一步可以进行硬件演示，将 QSVAE 的能耗和延迟与经典脉冲基线进行比较。鲁棒性结果为后续测试奠定了基础：以现实且相关的检测器偏移取代高斯代理，并在第二个数据集和另一种表征上开展测试，以确定容量解耦状态的泛化能力。由于 $\sin^2$ 动力学以牺牲容量为代价带来了鲁棒性，因此值得探索一种理想形式的神经元：在需要容量时减弱其饱和特性，在需要鲁棒性时则保留这种特性。我们由此得出的结论是，一种成本低、可由经典方法模拟的量子启发式脉冲神经元，在准确率–鲁棒性平面上占据了一个独特且有用的位置。在我们用于模拟部署中噪声和偏移条件的代理情形下，它表现出鲁棒性，而我们测试过的非周期经典脉冲神经元无法达到这一点，因此值得进一步探索。

IV. 方法

A. 数据与任务

我们使用随 Belis 的量子异常检测研究发布的双喷注潜在表征等人 [5]。在该研究中，经典卷积自编码器将每个大型强子对撞机双喷注事件（包含两个喷注，每个喷注最多由一百个具有特征 $\Delta\eta$， $\Delta\phi$和 p_T的组分描述）压缩为逐喷注潜在向量。我们将两个喷注拼接为一个维度为 2ℓ的单一事件向量，其中 $\ell \in \{4, 8, 16, 32\}$ 是逐喷注潜在维度。自编码器采用双曲正切瓶颈，因此潜在坐标的取值范围为（-1，1）。背景是标准模型量子色动力学多喷注产生。三个标准模型之外的信号分别是窄（3.5 TeV）和宽（1.5 TeV）的 Randall–Sundrum 引力子衰变为 WW，以及质量为 3 $A \to HZ \to ZZZ$*.5 TeV* 的标量。每个检测器都以留出的本底事件样本作为单类模型进行训练，并分别在由背景和各个信号构成的平衡混合样本上评估。我们报告各信号的得分，以及汇总三个信号的总体得分。脉冲模型使用一个含有 *15k* 个事件的背景样本池，其中 *15%* 留作验证。经典检测器使用约 *25k* 个事件进行拟合，量子方法的预算见表I。

B. 异常检测协议与指标

主要指标是 ROC 曲线下面积（AUROC）。为探究对搜索而言至关重要的低假阳性率区间，我们还按照 Belis 等人 [5], 的方法，在固定信号效率 $\varepsilon_s \in \{0.6, 0.8\}$ 下报告背景效率倒数 $1/\varepsilon_b$，并按照 Dark Machines挑战赛 [2]的方法，在固定背景效率 $\varepsilon_b \in \{10^{-2}, 10^{-3}\}$ 下报告显著性改进指标 $\varepsilon_s/\sqrt{\varepsilon_b}$。对于每个检测器，我们都报告这些指标，

表 S5 中保存了每事件得分，表 S4 中给出了每个信号的 AUROC。对于存在重复拟合的情况，我们报告均值和跨随机种子标准差。主要脉冲检测器使用五个随机种子，而消融扫描使用三个或五个，具体如第 IIG 节所述；量子检测器根据维度使用一个或五个随机种子，纯混合模型、PQAE 和 BQAE 的结果则使用三个随机种子。PQAE/BQAE 鲁棒性曲线是固定随机种子的压力检查。单次拟合的无模型检测器和经典检测器不存在随机种子变化，因此我们改为报告其在测试集上的自助法 95% 置信区间。对于量子核和量子自编码器，由于成本限制了训练集大小，每个随机种子都会从固定池中抽取一个仅背景训练子集。

C. 脉冲自编码器与消融神经元

QLIF神经元和量子启发式自编码器分别在第II A节和第II B节中推导。这里我们给出相关的复现实验细节以及消融神经元。两个脉冲检测器都采用了我们对Kamata等人 [16]提出的全脉冲变分自编码器进行的特征向量适配。该模型直接处理 2ℓ维潜向量：脉冲多层感知器编码器和解码器在 $T = 16$ 个时间步上运行，并采用伯努利脉冲后验与先验。它遵循变分自编码器 [22], 的证据下界框架，将潜变量Kullback–Leibler项替换为FSVAE特有的最大均值差异项，该项计算后验与先验脉冲序列样本之间的差异 [16]„，并将其用作每事件异常得分。两个检测器的唯一区别在于神经元：FSVAE使用标准LIF单元，而QSVAE使用公式（3）中的QLIF神经元。两个网络都通过带有替代梯度 [34] 的时间反向传播以及依赖阈值的批归一化 [21]进行训练。对于QLIF神经元，我们对反向膜梯度（公式（4））进行钳制，以实现稳定优化，并将依赖阈值的归一化初始化为接近12%的放电率。训练后，平均放电率稳定在更高水平，并且因神经元而异。在三随机种子探测中，FSVAE为0.29，QSVAE为0.23，稠密钳位神经元为0.37，而放电匹配钳位（匹配）为0.19。正是这些训练后的放电率由放电匹配对照进行近似。QLIF脉冲阈值在激发态布居上设为0.75，其泄漏由 $T_1 = 60$ 个时间步设定，与原始QLIF [17]相同。架构和完整超参数集详见表S1。两个检测器均使用AdamW优化器进行60轮训练，学习率为 10^{-3}，权重衰减为10^{-3}，批量大小为256，并在评估前收敛。

为隔离 QLIF 神经元的哪项属性产生了鲁棒性（第 IIG节），我们引入一种有界的经典控制组——钳位神经元。它保留 QLIF 神经元的有界 $[0, 1]$ 膜电位、尖峰阈值和泄漏，但将式（^{2} 中的正弦3）旋转替换为普通的线性泄漏积分，随后进行硬钳位，$\alpha[t+1] = \mathrm{clip}(\lambda\,\alpha[t]+\theta[t],\ 0,\ 1)$，其中 $\theta[t]$ 是加权输入，λ 是相同的泄漏因子。它与 QLIF 神经元完全一样受到界限约束，但其动力学是经典且局部线性的，因此将其与 QSVAE 进行比较，可以区分有界膜电位与正弦^2 动力学。由于更稀疏的神经元在机制上更加鲁棒，我们扫描钳位的阈值依赖的归一化尺度（bn scale $\in \{0.30, 0.45, 0.60, 0.80, 1.0\}$）；在最稠密的设置下，训练后的放电率为 0.37，在 bn scale $= 0.45$ 下为 0.19，在更低的尺度下则更加稀疏。该 bn scale $=_0.45$ 设置，即钳位（匹配），其放电率接近 QSVAE 的放电率（0.19 对比 0.23），从而消除了发放密度这一混杂因素。为检验仅平滑饱和是否足以复现这一效果，平滑饱和神经元用逻辑映射 $\alpha[t+1] = \sigma\big(g(\lambda\,\alpha[t] + \theta[t] - \frac{1}{2})\big)$替代硬钳位，将值限制在（0，1）内；其导数在两个边界处均趋于零，如同正弦^2 映射，但它是单调且非周期性的。我们在三种设置下训练该神经元：增益 $g = 4$（中心斜率为 $g/4 = 1$，与钳位的内部增益匹配）、增益 $g = 2$（最大斜率为二分之一，与 QLIF 映射的全局增益上界匹配），以及 $g = 4$（将归一化降低至 bn scale $= 0.45$，与放电匹配，仿照匹配钳位）。此外，为检验普通正则化的作用，我们以权重衰减 $\{10^{-2}, 3\times10^{-2}, 10^{-1}\}$训练 FSVAE，最高达到默认值的 100 倍。所有控制组均采用 QLIF 神经元的阈值、泄漏、优化器和训练计划。它们的反向传播规则并不完全相同，因为 QLIF 神经元和周期性神经元共享一种穿过硬重置的截断的循环梯度。钳位控制组和平滑饱和控制组通过普通自动微分，在重置处将循环梯度置零；而 FSVAE 保留其已发表的替代梯度和乘法重置。放电、饱和和正则化扫描使用三个随机种子，其他变体使用五个。

D. 基线类别

我们将模型与四类基线进行比较，所有基线都使用相同的潜变量。无模型统计量无需训练，直接在潜向量上计算，指标包括潜能量（mean z^2）、 L_2范数以及最大绝对坐标 max $|z|$。经典检测器取自 PyOD 库 [35]„，包括 a k 最近邻距离、相对于背景的马氏距离、孤立森林 [36], 、深度自编码器以及深度单类检测器深度支持向量数据描述（Deep SVDD） [37]。量子检测器采用Belis等人 [5], 提出并在 PennyLane 中重新实现的量子套件 [38]。量子核由数据重上传特征映射构建，其中每个量子比特通过旋转编码两个潜在坐标，并采用最近邻CNOT纠缠和 L 次重上传。我们将该核评估为态矢量之间的平方态重叠 $k(x,\ x') = |\langle\psi(x)|\psi(x')\rangle|^2$，并将其输入使用预计算格拉姆矩阵的单类支持向量机 [3, 4]。该类别还包括该核的非纠缠变体，即

量子 k均值和 k中位数聚类得分，以及原研究用作参照的经典 RBF 单类支持向量机。格拉姆矩阵的规模随 N^2 增长，因此训练预算设定为约 600 个事件。量子自编码器包括两个全量子变分自编码器：持久型（PQAE）和瓶颈型（BQAE）变体，均以重构误差评分；此外还有混合自编码器，它将量子数据重上传编码器与经典解码器配对，并分为纠缠和非纠缠两种形式。全量子变体仅在较小宽度下可通过态矢量进行模拟（PQAE 可达 $\ell = 8$，BQAE 则为 $\ell = 4$）。我们根据 Belis等人（图3）报告的各信号数值以及纠缠和深度缩放结果，对量子套件进行了验证。

E. 扰动方案：噪声鲁棒性与敏感性

为探究对噪声和分布偏移的鲁棒性，我们在干净背景上拟合每个检测器，并且仅在评估时，在对每个测试潜变量进行评分之前添加噪声。所有方法对完全相同的加噪输入进行评分，因此结果差异反映的是检测器本身，而非扰动。我们采用两种幅度匹配的噪声模型。各向同性噪声会向每个坐标等量添加相互独立的高斯分布噪声，其标准差为 σ。而保持几何结构的噪声则根据背景在相应坐标上的自身离散程度，按坐标缩放噪声，$\sigma_i = \sigma\, s_i$ 其中$s_i = \hat{\sigma}_i^{\mathrm{bkg}} / \sqrt{\langle (\hat{\sigma}^{\mathrm{bkg}})^2 \rangle}$，因此缩放因子经过均方根归一化（$\langle s_i^2 \rangle = 1$）。在相同的 σ 下，注入的总功率与各向同性情形相等，只是重新分配以贴合数据，因此两种模型能够进行公平比较，而不是其中一种仅仅幅度更小。我们对 σ 进行扫描，遍历 $\{0, 0.05, 0.1, 0.15, 0.2, 0.3\}$，并报告每个水平下的 AUROC，以及保留率 AUROC(σ)/AUROC(0)。保持几何结构的模型是公平的基线；保留各向同性模型，则是为了揭示第 IIF节中的噪声模型伪象。鲁棒性和敏感性评估流程会在各自的评估划分上重新对每个检测器评分，因此，鲁棒性各节所引述的干净 AUROC 与表I中的数值相差至多 0。010。计算保留率时，始终以相应流程自身的干净值作为分母。

一项互补的敏感性对照用于检验检测器的噪声鲁棒性究竟源于真正的信号跟踪，还是仅仅源于不敏感。在保持训练模型不变的情况下，我们将每个测试信号事件与随机配对的背景事件进行线性插值，使其逐渐向背景衰减，$x \mapsto (1-\alpha)x + \alpha\, x_{\mathrm{bkg}}$，将 α从真实信号（$\alpha = 0$）扫描 $\{0, 0.25, 0.5, 0.75, 1\}$至统计上与背景相同的信号（$\alpha = 1$）；背景测试事件则保持不变。能够读取区分性物理机制的检测器，其分离能力应随 $\alpha \to 1$趋向于 $= 0.5$ 而减弱；而对输入不敏感的检测器则不会如此。作为进一步的简并性检查，我们记录在 $x \mapsto$ 处经过最小—最大归一化的异常分数的离散范围；对于近似恒定的检测器，该范围会收缩为零。

为直接探究鲁棒性机制，我们测量每个训练模型的异常分数S 对微小输入扰动的响应程度。在固定模型的情况下，我们对每个测试潜变量施加一个遵循几何结构的步长 $\delta = \sigma\, s \odot \eta$（其中 $\eta \sim \mathcal{N}(0, I)$，并采用与上文相同的逐坐标尺度 s，探测尺度为$\sigma \in \{0.05, 0.1, 0.2\}$）。据此，我们记录分数变化的平均绝对值 $\langle |S(x+\delta) - S(x)| \rangle$，该值在多次抽样和测试事件上取平均，并按模型自身干净分数的离散度 std($S(x)$) 进行归一化。分数对输入的这种归一化敏感度，可用于比较分数尺度不同的模型；数值越小，表示异常分数与输入扰动的耦合越弱。由于不同模型的归一化因子本身可能不同，我们还用两项无需归一化的指标对同一扰动进行交叉验证：干净分数排序与扰动后分数排序之间的 Kendall 秩相关系数，以及平均绝对秩位移占测试集的比例。AUROC 取决于排序，而非分数值。我们报告这些指标，以及在三个随机种子上取平均的训练后脉冲层平均放电率；计算对象为 QSVAE 和十二种非周期经典脉冲对照模型，所有结果均由已保存的检查点计算得出。最后，为刻画训练后神经元的工作状态，我们将记录的脉冲层输入重新输入各个膜递归，并统计逐步驱动及其局部增益的直方图。对于 QLIF，驱动角为 $\theta + \mu$，增益为 $\frac{1}{2}|\sin(\theta+\mu)|$；对于平滑饱和对照模型，增益取自其逻辑映射；对于钳位模型，则统计处于线性内部区间的更新比例。

F. 计算

所有实验均在单张 NVIDIA RTX 3080 Ti 上运行。量子探测器通过 PennyLane [38] 使用态矢量进行模拟。量子核在 $\ell = 16$（十六个量子比特）时可行，而在 ℓ =32 时不予采用，因为此时 2^{32} 维态矢量无法处理，因此该位宽下仅保留聚类和参考探测器。全量子自编码器在态矢量可行的位宽下进行模拟：PQAE 在 $\ell = 4$, 8 时使用八个和十六个量子比特，BQAE 在 $\ell = 4$ 时使用十五个量子比特。论文随附发布了代码、配置、整理后的派生结果、已保存的脉冲模型检查点和事件评分，以及量子自编码器参数和事件评分（见代码可用性）。

数据可用性

双喷注潜变量空间数据来自 Belis 等人 [5]，并已在 Zenodo 上公开，10.5281/zenodo.7673769。

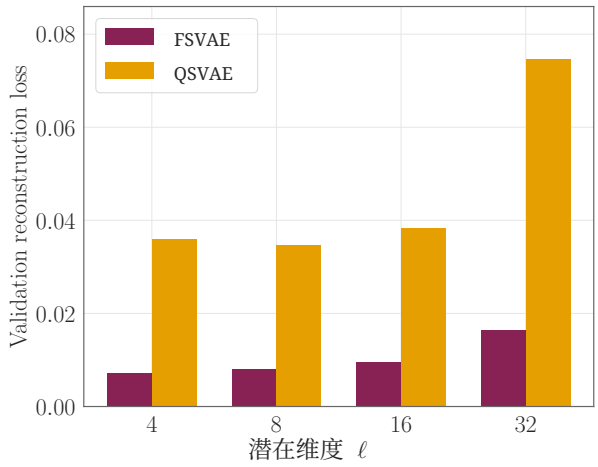

图 10. **重构下限。** FSVAE 和 QSVAE 在不同潜在维度下的最终验证重构损失。QLIF 神经元有界的正弦^2-旋转膜电位收敛到比经典泄漏神经元 ∼4 倍更高的重构下限。发放率匹配消融（第 IIG节）表明，该下限不会带来准确率惩罚，因为同样有界的经典神经元能够达到与 FSVAE 相同的重构效果。这与动力学特定正则化相符；我们认为，这种正则化与 QSVAE 的鲁棒性和稳定性有关。

代码可用性

异常检测流程、QLIF 神经元和 QSVAE 的实现，以及经典、量子和量子自编码器基线的复现版本，均可在以下地址获取：`https://github.com/deanbrand/QSVAE`。该软件的 1.0.1 版本及实验配置、整理后的派生结果、保存的脉冲模型检查点和事件评分，以及量子自编码器参数和事件评分，已存档于 Zenodo，`10.5281/zenodo.24460595`。其他基线的事件评分可使用随附的实验驱动程序和公开输入数据集重新生成。

致谢

本研究得到国家理论与计算科学研究所的支持。C.M.M 感谢卢旺达非洲数学科学研究与创新中心的支持。

作者贡献

D.B. 构思了这项研究，开发了 QLIF 神经元和软件，开展了实验，进行了形式化分析，并撰写了初稿。D.S. 提供了持久型和瓶颈量子自编码器基线，C.M.M. 提供了混合量子自编码器基线。F.P. 负责监督这项工作并筹集经费。所有作者均审阅并编辑了稿件。

人工智能声明

作者使用 OpenAI 的 GPT-6 Astra 和 Anthropic 的 Claude Fable 5.1 协助准备稿件并开发计算代码。这些工具的使用遵循作者的研究决策、指导方针和预定方法。这些工具还协助进行了语言编辑和代码审查。作者对所有人工智能辅助生成的内容和代码进行了严格审查，以确保其准确性，并确保其符合预定的研究目标和方法。作者对最终稿件及所有报告的结果承担全部责任。

利益冲突

作者声明不存在利益冲突。

附录 A：补充材料

1. 架构与训练超参数

FSVAE 和 QSVAE 共用脉冲变分自编码器配置；两者仅在脉冲神经元方面有所不同（表S1）。这些值是 `config.py` 中的运行配置（每个喷注的潜变量宽度为2ℓ，适用于 $\ell \in \{4, 8, 16, 32\}$）。

2. 量子基线的复现验证

表S2 将我们使用 PennyLane 重新实现的量子核及其经典基准，与 Belis 等人发表的逐信号数值并列展示 等 [5]，这些数值取自他们图 3 中印刷的 AUC 图例。每个信号的 AUROC 差异均在 0.065 以内；最大差异出现在接近随机水平的宽质量引力子上，两种实现的结果都在 0.5 两侧。Belis 等 人未报告合并信号后的总体 AUROC，因此正文中引用的总体数值是我们自行汇总的结果，而非复现的已发表数值；我们的经典 RBF 参考模型略弱于他们的模型，因此我们测得的量子模型相对经典模型的优势若有影响，也只会显得更保守。

3. 完整基准测试表

表 S3给出了所有检测器在全部四种潜在维度下的总体 AUROC，包括匹配尺度（$N_{\text{train}} \approx 600$）的脉冲行，以及在态矢量可行宽度下的持久型和瓶颈量子自编码器。表S4 给出了逐信号 AUROC 和面向搜索的低假阳性率指标 $1/\varepsilon_b$ 在$\varepsilon_s = 0.8$ 在 $\ell = 16$。

4. 全量子自编码器

采用第四节 D 中所述的角度域重构，PQAE 在 ℓ =4 时的三随机种子干净总体 AUROC 为0.836 ± 0.005，BQAE 为 0.839 ± 0.004；PQAE 在 $\ell = 8$ 时为 0.853 ± 0.005。在所有三种配置中，宽引力子最难处理（分别为 0.580 ± 0.006、0.573 ± 0.009 和 0.584 ± 0.014）。在固定随机种子的鲁棒性检查中，$\sigma = 0.3$ 时保留的干净总体 AUROC 比例为：PQAE 在 $\ell = 4$ 时为 0.812，BQAE 为 0.767；PQAE 在 $\ell = 8$ 时为 0.814。没有干净坐标超出 $[-1, 1]$ 编码器范围。在$\sigma = 0.3$，$\ell = 4$ 时的裁剪前坐标比率分别为：背景 3.95%，窄 G 为 5.33%，宽 G 为 2.91%，$A \to HZ$ 为 3.42%；在 $\ell = 8$ 时分别为 6.81%、10.27%、6.79% 和 9.23%。我们将这些值裁剪至边界，绝不对其进行周期性回绕。干净和加噪结果来自 runs/hep qae v2_l{4,8} 和 runs/hep robustness qae v2。

[1]G. Kasieczka, B. Nachman, D. Shih, *et al.*, LHC奥林匹克竞赛2020：面向高能物理异常检测的社区挑战，Reports on Progress in Physics 84, 124201 (2021), arXiv:2101.08320 [hep-ph]. [2] T. Aarrestad *et al.*, 暗机器异常评分挑战：大型强子对撞机的基准数据与模型无关事件分类，SciPost Phys. 12, 043 (2022), arXiv:2105.14027 [hep-ph]. [3] V. Havlíˇcek, A. D. C´orcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, 利用量子增强特征空间进行监督学习，Nature 567, 209 (2019), arXiv:1804.11326 [quant-ph]. [4] M. Schuld and N. Killoran, 特征希尔伯特空间中的量子机器学习，Physical Review Letters 122, 040504 2019 1803 ( ), arXiv: .07128 [quant-ph]. [5] V. Belis, K. A. Wo´zniak, E. Puljak, P. Barkoutsos, G. Dis- sertori, M. Grossi, M. Pierini, F. Reiter, I. Tavernelli, and S. Vallecorsa, 大型强子对撞机质子碰撞事件潜在空间中的量子异常检测，Communica- tions Physics 7, 334 (2024), arXiv:2301.10780 [hep-ex]. [6] V. S. Ngairangbam, M. Spannowsky, and M. Takeuchi, 使用量子自编码器进行高能物理异常检测，Phys. Rev. D 105, 095004 (2022), arXiv:2112.04958 [hep-ph]. [7] J. Schuhmacher, L. Boggia, V. Belis, E. Puljak, M. Grossi, M. Pierini, S. Vallecorsa, F. Tacchino, P. Barkoutsos, and I. Tavernelli, 借助经典与量子异常检测揭示标准模型之外的物理，Machine Learning: Science and Technology 4，045031（2023），arXiv:2301.10787 [hep-ph]。[8] M. Davies、N. Srinivasa、T.-H. Lin、G. Chinya、Y. Cao、S. H. Choday、G. Dimou、P. Joshi、N. Imam、S. Jain、Y. Liao、C.-K. Lin、A. Lines、R. Liu、D. Mathaikutty、S. McCoy、A. Paul、J. Tse、G. Venkataramanan、Y.-H. Weng、A. Wild、Y. Yang 和 H. Wang，Loihi：一种具有片上学习能力的神经形态多核处理器，IEEE Micro 38，82（2018）。[9] K. Roy、A. Jaiswal 和 P. Panda，迈向基于脉冲的机器智能与神经形态计算，Nature 575，607（2019）。[10] B. M. Dillon、J. Harkin 和 A. Javed，面向大型强子对撞机物理的脉冲神经网络异常检测，Machine Learning: Science and Technology 7，025019（2026）。[11] S. Han、X. Hu、H. Huang、M. Jiang 和 Y. Zhao，ADBench：异常检测基准测试，载于 *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)* 2022 2022，NeurIPS（Curran Associates, Inc.，）第32142–32159页，arXiv:2206.09426 [cs.LG]。[12] S. Sharmin、N. Rathi、P. Panda 和 K. Roy，深度脉冲神经网络的固有对抗鲁棒性：离散输入编码与非线性激活的影响，载于 *Computer Vision – ECCV 2020*（Springer

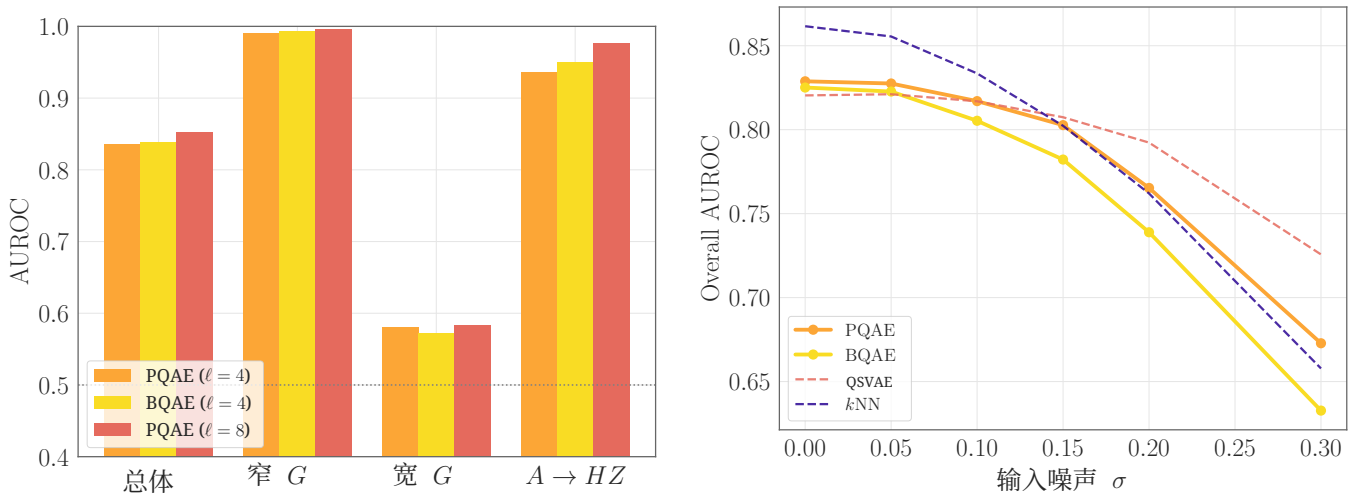

图 S1. **全量子自编码器。** **（a）**PQAE 和 BQAE 在其态矢量可行宽度下的干净数据逐信号 AUROC（三个随机种子均值）。宽 G 仍是最难检测的信号，PQAE 从 $\ell = 4$ 提升至 $\ell = 8$。 **（b）**在 $\ell = 4$ 下，输入噪声对单随机种子 PQAE 和 BQAE 得分的影响，比对正则化 QSVAE 对照模型更为显著； kNN 作为干净数据表现领先的模型纳入比较。所有模型均在干净的量子色动力学背景上训练，然后在相同的含噪测试事件上评分。

国际出版，2020年）第399–414页。[13] A. Marchisio、G. Nanfa、F. Khalid、M. A. Hanif、M. Mar-tina 和 M. Shafique，脉冲神经网络安全吗？脉冲神经网络与深度神经网络安全漏洞的比较研究，载于 *2020 International Joint Conference on Neural Networks (IJCNN)* 2020 1 8 （电气电子工程师学会，）第–页。[14] Y. Kim、H. Park、A. Moitra、A. Bhattacharjee、Y. Venkate-sha 和 P. Panda，频率编码还是直接编码：哪种方式更适合构建准确、鲁棒且节能的脉冲神经网络？，载于 *ICASSP 2022 – 2022 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)* 2022 71 75 （电气电子工程师学会，）第–页。[15] R. El-Allami、A. Marchisio、M. Shafique 和 I. Alouani，借助固有结构参数保护深度脉冲神经网络免受对抗攻击，载于 *2021 Design Automation & Test in Europe Conference & Exhibition (DATE)* 2021 774 779 （电气电子工程师学会，）第–页。[16] H. Kamata、Y. Mukuta 和 T. Harada，完全脉冲变分自编码器，AAAI人工智能会议论文集36，7059（2022年），arXiv:2110.00375 [cs.NE]。[17] D. Brand 和 F. Petruccione，一种量子泄漏积分发放脉冲神经元及网络，npj 量子信息10，125（2024年），arXiv:2407.16398 [quant-ph]。[18] J. K. Eshraghian、M. Ward、E. Neftci、X. Wang、G. Lenz、G. Dwivedi、M. Bennamoun、D. S. Jeong 和 W. D. Lu，借鉴深度学习经验训练脉冲神经网络（2023年），arXiv:2109.12894 [cs]。[19] L. B. Kristensen、M. Degroote、P. Wittek、A. Aspuru-Guzik 和 N. T. Zinner，一种人工脉冲量子神经元，npj量子信息7，59（2021年），arXiv:1907.06269 [quant-ph]。[20] J. Chen、B. Rajendran 和 O. Simeone，具有量子存储器和局部学习能力的随机量子脉冲神经网络，IEEE Journal on Selected Areas in Communications 44，5421（2026年）。[21] H. Zheng、Y. Wu、L. Deng、Y. Hu 和 G. Li，通过直接训练更大型的脉冲神经网络深入探索作品，AAAI人工智能会议论文集 35, 11062 (2021年)，arXiv:2011.05280 [cs.NE]. [22] D. P. Kingma 和 M. Welling，《自编码变分贝叶斯》，载于 *2nd International Conference on Learn-ing Representations, ICLR 2014, Banff, AB, Canada, April 14–16 2014 Conference Track Proceedings* 2014 , , ( ) arXiv:1312.6114 [stat.ML]. [23] D. Slabbert 和 F. Petruccione，基于量子卷积自编码器的异常检测重构（2026年），arXiv:2607.02135 [quant-ph]. [24] C. M. Manfouo、D. Slabbert、P. K. Osei 和 F. Petruccione，用于脉冲星候选体校准分类的混合量子-经典逻辑回归（2026年），arXiv:2605.07866 [quant-ph]. [25] N. Perez-Nieves、V. C. H. Leung、P. L. Dragotti 和 D. F. M. Goodman，神经异质性促进鲁棒学习，《自然·通讯》12，5791 (2021年)。[26] H. Drucker 和 Y. Le Cun，利用双重反向传播提升泛化性能，《电气电子工程师学会神经网络汇刊》3, 991 (1992)。[27] M. Cisse、P. Bojanowski、E. Grave、Y. Dauphin 和 N. Usunier，Parseval网络：提升对抗样本鲁棒性，载于 *Proceedings of the 34th International Conference on Machine Learning (ICML)*，《机器学习研究论文集》第70卷（2017年），第854–863页，arXiv:1704.08847 [stat.ML]. [28] M. Hein 和 M. Andriushchenko，关于分类器抵御对抗操纵的鲁棒性的形式化保证，载于 *Advances in Neural Information Processing Systems (NeurIPS)* 30 2017 2266 2276, , 第( )卷，第–页，arXiv:1705.08475 [cs.LG]. [29] D. Hendrycks 和 T. Dietterich，基准测试神经网络对常见损坏和扰动的鲁棒性，载于 *International Conference on Learning Representations (ICLR)* 2019 1903 12261 [ ] ( ) arXiv: . cs.LG . [30] L. Ziyin、T. Hartwig 和 M. Ueda，神经网络无法学习周期函数及其解决方法，载于 *Advances in Neural Information Processing Systems (NeurIPS)*,

第33卷（2020年），第1583–1594页，arXiv:2006.08195 [cs.LG]. [31] V. Sitzmann、J. N. P. Martel、A. W. Bergman、D. B. Lindell 和 G. Wetzstein，具有周期激活函数的隐式神经表示，载于 *Advances in Neural Information Processing Systems (NeurIPS)* 33 2020，第( )卷，第7462–7473页，arXiv:2006.09661 [cs.CV]. [32] M. Tancik、P. P. Srinivasan、B. Mildenhall、S. Fridovich-Keil、N. Raghavan、U. Singhal、R. Ramamoorthi、J. T. Barron 和 R. Ng，傅里叶特征使网络能够在低维域中学习高频函数，载于 *Advances in Neural Information Processing Systems (NeurIPS)* 33 2020 7537 7547，，第( )卷，第–页，arXiv:2006.10739，arXiv:2006.10739. [33] M. Schuld、R. Sweke 和 J. J. Meyer，数据编码对变分量子机器学习模型表达能力的影响，Physical Review A **103**，032430（2021年）。[34] E. O. Neftci 、 H. Mostafa 和 F. Zenke，脉冲神经网络中的替代梯度学习：带来基于梯度的优化应用于脉冲神经网络的能力，《IEEE信号处理杂志》36，51（2019），arXiv:1901.09948 [cs.NE]。[35] Y. Zhao、Z. Nasrullah 和 Z. Li，PyOD：用于可扩展异常值检测的 Python工具箱，《机器学习研究杂志》20，1（2019），arXiv:1901.01588 [cs.LG]。[36] F. T. Liu、K. M. Ting 和 Z.-H. Zhou，孤立森林，载于 *2008 Eighth IEEE International Conference on Data Mining* 2008 413 422 （电气电子工程师学会，）第 – 页。[37] L. Ruff、R. Vandermeulen、N. Goernitz、L. Deecke、S. A. Siddiqui、A. Binder、E. M¨uller 和 M. Kloft，深度单类分类，载于 *Proceedings of the 35th International Conference on Machine Learning*，《机器学习研究论文集》第80卷，由 J. Dy 和 A. Krause 编（PMLR，2018），第4393–4402页。[38] V. Bergholm、J. Izaac、M. Schuld、C. Gogolin、 *et al.*，Pennylane：混合量子-经典计算的自动微分（2018），arXiv:1811.04968 [quant- ph]，arXiv:1811.04968 [quant-ph]。

表 S1. **脉冲变分自编码器架构和训练设置** （FSVAE 和 QSVAE），以及 QLIF 神经元参数。FSVAE 使用标准漏积分发放神经元替代 QLIF。

组件	设置
输入宽度（双喷注潜变量）	2ℓ
脉冲变分自编码器潜在维度	64
编码器/解码器隐藏层宽度	256，128
时间步数 T	16
脉冲神经元	漏积分发放 (FSVAE) / QLIF (QSVAE)
目标函数	证据下界（重构 +最大均值差异）
优化器	AdamW，学习率 10^{-3}，权重衰减 10^{-3}
批量大小	256
训练轮数	60
随机种子	5 个脉冲
QLIF neuron	
发放阈值	0.75
T_1 松弛	60
膜时间常数 τ	1
tdBN 角度尺度（发放 ≈12%）	1.0
分级（非二进制）输出	no
反向传播膜梯度截断	1.0

表 S2。 **量子核基线相对于 Belis 的验证复现** 等。 在种子论文的参考工作点（$N_{\text{train}} = 600$， $\nu = 0.01$； ℓ等于量子比特数）下，无监督量子核机器（最近邻纠缠）及其经典 RBF 单类支持向量机参考模型的逐信号 AUROC。已发表数值取自 Belis 的图 3 中印刷的曲线下面积图例 *et al.* [5]：其逐信号面板固定为 $\ell = 8$，其潜在维度面板报告 $A \to HZ$信号，我们在 $\ell = 16$处对其进行匹配。我们的数值是来自 `runs/hepquantuml8`和 `runs/hepquantuml16`运行的五个随机种子的均值。

探测器信号（ℓ）	Belis *et al.*（已发表）	本工作（已复现）	$\lvert\Delta\rvert$
Unsupervised quantum kernel machine (entangled)			
窄 $G \to WW$（ℓ=8）	0.995	0.997	0.002
$A \to HZ$(ℓ=8)	0.947	0.968	0.021
宽 $G \to WW$(ℓ=8)	0.476	0.541	0.065
$A \to HZ$(ℓ=16)	0.968	0.980	0.012
Classical RBF one-class support vector machine (the kernel's own reference)			
窄 $G \to WW$(ℓ=8)	0.993	0.986	0.007
$A \to HZ$(ℓ=8)	0.933	0.889	0.044
宽 $G \to WW$(ℓ=8)	0.456	0.441	0.015
$A \to HZ$(ℓ=16)	0.912	0.848	0.064

表 S3. **不同潜在维度下的总体 AUROC**。 紧凑的不确定性表示法：0.877(5) 表示 0.877 ± 0.005，即以末位数字为单位的跨随机种子标准差（脉冲模型使用五个随机种子；量子方法根据维度使用一个或五个；量子自编码器使用三个）；没有括号表示仅进行了一次拟合。匹配尺度行 FSVAE (600) 和 QSVAE (600) 在量子核的 ~ 600 事件预算下重复脉冲模型的结果。“–”表示该配置在该维度下无法用态矢量实现。来源：runs/hep comparison v2 l{4,8} 和 runs/hep comparison l{16,32}。

类别	方法	$\ell = 4$	$\ell = 8$	$\ell = 16$	$\ell = 32$
无模型	最大 $\lvert z \rvert$	0.468	0.866	0.719	0.500
	L_2 能耗	0.516	0.721	0.791	0.833
经典	kNN	0.867	0.892	0.891	0.881
	马氏距离	0.833	0.869	0.879	0.865
	孤立森林	0.835	0.849	0.846	0.855
	自编码器	0.825	0.855	0.874	0.868
	深度支持向量数据描述	0.851	0.770	0.684	0.799
脉冲	FSVAE	0.824(5)	0.865(2)	0.877(5)	0.864(8)
	QSVAE	0.821(2)	0.840(2)	0.827(1)	0.813(3)
	FSVAE（600）	–	0.841(9)	0.842(3)	–
	QSVAE（600）	–	0.837(3)	0.821(2)	–
量子	Q-kernel（纠缠）	0.779	0.835(2)	0.855(3)	–
	Q-kernel（无纠缠）	0.718	0.824(5)	0.829(4)	–
	QK-中位数	0.852	0.863(3)	0.855(2)	0.803
	QK-均值	0.835	0.859(1)	0.848(1)	0.815
	径向基函数单类支持向量机（参考）	0.762	0.772(30)	0.755(34)	0.838
量子自编码器	混合量子自编码器（非纠缠）	0.816(6)	0.862(8)	0.877(8)	0.844(9)
	混合量子自编码器（纠缠）	0.832(6)	0.873(3)	0.866(5)	0.816(1)
	PQAE	0.836(5)	0.853(5)	–	–
	BQAE	0.839(4)	–	–	–

表 S4。**在 $\ell = 16$ 的逐信号性能**。AUROC（± 多随机种子时的跨随机种子标准差），以及信号效率为 $\varepsilon_s = 0.8$（总体）时面向搜索的背景效率倒数 $1/\varepsilon_b$。窄和宽分别表示窄、宽 Randall–Sundrum 引力子； $A \to HZ$ 标量。来源：runs/hep comparison l16。

类别	方法	窄	宽	$A \to HZ$	$1/\varepsilon_b$@0。8
无模型	最大 $\lvert z \rvert$	0.683	0.667	0.807	1.9
	L_2 能耗	0.973	0.520	0.879	2.2
经典	kNN	0.998	0.689	0.987	5.6
	马氏距离	0.996	0.663	0.978	4.9
	孤立森林	0.996	0.573	0.968	3.2
	自编码器	0.995	0.649	0.977	4.7
	深度支持向量数据描述	0.865	0.441	0.746	1.8
脉冲	FSVAE	0.997±0.001	0.657±0.009	0.977±0.006	4.4
	QSVAE	0.991±0.000	0.544±0.003	0.945±0.001	2.8
	FSVAE（600）	0.991±0.001	0.603±0.011	0.933±0.003	3.1
	QSVAE (600)	0.990±0.000	0.535±0.004	0.938±0.002	2.6
量子	Q-kernel（纠缠）	0.999±0.001	0.587±0.008	0.980±0.003	3.7
	Q-kernel（无纠缠）	0.989±0.002	0.549±0.004	0.949±0.006	2.8
	QK-中位数	0.993±0.001	0.617±0.008	0.956±0.004	3.7
	QK-均值	0.994±0.001	0.598±0.001	0.953±0.003	3.4
	径向基函数单类支持向量机（参考）	0.978±0.012	0.440±0.035	0.848±0.054	1.8
量子自编码器	混合量子自编码器（非纠缠）	0.997±0.001	0.652±0.015	0.983±0.007	4.4
	混合量子自编码器（纠缠）	0.996±0.001	0.630±0.013	0.973±0.004	4.2
	PQAE	–	–	–	–
	BQAE	–	–	–	–

表 S5. **不确定性与低假阳性搜索指标（$\ell = 16$，总体）**。对于单次拟合的检测器，AUROC 附带测试集上的自助法 95% 置信区间（1000 次重采样）；对于脉冲检测器，则报告跨随机种子标准差（五个随机种子）。 $1/\varepsilon_b$ 表示信号效率为 ε_s 时的背景效率倒数；SIC $= \varepsilon_s/\sqrt{\varepsilon_b}$ 表示背景效率为 ε_b 时的显著性提升。根据保存的逐事件测试分数计算。来源：runs/hep score stats。

方法	AUROC	$1/\varepsilon_b$@0.6	$1/\varepsilon_b$@0.8	SIC@10^{-2}	SIC@10^{-3}
FSVAE	0.877 ± 0.005	60.1	4.4	5.5	9.9
QSVAE	0.827 ± 0.001	19.2	2.8	4.0	5.7
钳制	0.876 ± 0.005	63.9	4.4	5.5	9.5
钳位（匹配）	0.883 ± 0.003	54.7	4.9	5.4	9.7
孤立森林	0.847 [0.844, 0.850]	30.9	3.2	4.9	9.0
马氏距离	0.879 [0.876, 0.882]	51.7	5.0	5.3	2.0
L_2 能耗	0.791 [0.787, 0.795]	9.7	2.2	3.1	3.8
最大值 $\lvert z\rvert$	0.719 [0.715, 0.723]	3.8	1.9	1.1	0.9

表 S6。 **经典脉冲消融控制**（$\ell = 16$）。 鲁棒性流程中的干净 AUROC（使用其自身的评估划分；数值与表 I 相差不超过 0。010，见章节 IVE）以及在 $\sigma = 0.3$ 保持几何结构的噪声下的保留比例；训练后的平均放电率（三随机种子探测，若有测量）；探测尺度为 0。05 时归一化的输入到分数的敏感度（数值越低，分数与输入的耦合越弱）；以及干净与扰动后的分数排序之间的肯德尔秩相关 τ（数值越高，排序越稳定）。所有经典脉冲基线的保留比例均为 0。74–0。85；只有 QSVAE 的保留比例达到 0。95，同时具有 $\sim 3.5\times$ 更低的敏感度和稳定得多的排序。来源： `runs/hep robustness whiten e{5,6,7,8}`， `runs/hep stress probes`, `runs/hep mechanism`。

变体	干净 AUROC	保留率@0。3	放电敏感性	肯德尔 τ	
QSVAE (正弦^2)	0.830	0.950	0.23	0.17	0.83
FSVAE（LIF）	0.868	0.795	0.29	0.64	0.48
钳制，批归一化 1.0	0.871	0.839	0.37	0.62	0.43
钳制，批归一化 0.80	0.870	0.831	–	0.61	0.46
钳制，批归一化 0.60	0.875	0.787	–	0.64	0.49
钳制，批归一化 0.45（匹配）	0.877	0.765	0.19	0.70	0.46
钳制，批归一化 0.30	0.882	0.741	–	0.68	0.51
逻辑斯蒂，$g = 4$	0.867	0.847	0.37	0.63	0.42
逻辑斯蒂，$g = 2$（斜率匹配）	0.863	0.849	–	0.63	0.40
逻辑斯蒂，$g = 4$（放电匹配）	0.874	0.780	–	0.66	0.45
FSVAE，权重衰减 10^{-2}	0.871	0.804	–	0.61	0.48
FSVAE，权重衰减 3×10^{-2}	0.867	0.799	–	0.60	0.49
FSVAE，权重衰减 10^{-1}	0.875	0.809	–	0.63	0.48

表 S7. **逐信号保留率：** $\sigma = 0.3$ （$\ell = 16$）。在保持几何结构的噪声下保留的干净 AUROC 比例，针对每种信号分别计算（脉冲变体取三个随机种子均值）。QSVAE 的显著优势并非汇聚造成的假象：它在每种信号上都成立，但需注意，QSVAE 的宽引力子列对应的干净 AUROC 接近随机水平，因此权重很小。来源：runs/hep stress probes/persignal.csv。

变体	总体	窄	宽	$A \to HZ$
QSVAE	0.943	0.965	0.972	0.904
FSVAE	0.814	0.878	0.804	0.756
钳制	0.848	0.914	0.822	0.798
钳位（匹配）	0.786	0.846	0.786	0.725
逻辑斯蒂，$g = 4$	0.851	0.913	0.828	0.802
逻辑斯蒂，$g = 2$	0.854	0.913	0.840	0.803
逻辑斯蒂（脉冲率匹配）	0.800	0.862	0.790	0.744
孤立森林	0.961	0.985	0.962	0.935

表 S8. **完整噪声扫描中的总体 AUROC（$\ell = 16$，符合几何约束）**。QSVAE 与所有经典脉冲基线之间的差距从 $\sigma \approx 0.1$ 开始，并单调扩大；顶点值的 $\sigma = 0.3$ 是这一趋势的终点，而非单点伪象。来源：runs/hep robustness whiten e{5,6,7,8}。

变体	$\sigma = 0$	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.15$	$\sigma = 0.2$	$\sigma = 0.3$
QSVAE (正弦^2)	0.830	0.829	0.826	0.823	0.816	0.788
FSVAE（漏积分发放）	0.868	0.867	0.847	0.812	0.784	0.690
钳制，批归一化 1.0	0.871	0.864	0.849	0.827	0.810	0.730
钳制，批归一化 0.80	0.870	0.869	0.855	0.830	0.810	0.724
钳制，批归一化 0.60	0.875	0.869	0.853	0.818	0.789	0.689
钳制，批归一化 0.45（匹配）	0.877	0.868	0.847	0.807	0.776	0.671
钳制，批归一化 0.30	0.882	0.873	0.852	0.805	0.769	0.653
逻辑斯蒂，$g = 4$	0.867	0.866	0.854	0.828	0.812	0.735
逻辑斯蒂，$g = 2$（斜率匹配）	0.863	0.863	0.851	0.827	0.810	0.733
逻辑斯蒂，$g = 4$（放电匹配）	0.874	0.867	0.850	0.812	0.782	0.682
FSVAE，权重衰减 10^{-2}	0.871	0.867	0.849	0.816	0.791	0.700
FSVAE，权重衰减 3×10^{-2}	0.867	0.862	0.846	0.813	0.787	0.693
FSVAE，权重衰减 10^{-1}	0.875	0.870	0.853	0.818	0.797	0.708

表 S9. **公平扰动（$\ell = 16$，$\sigma = 0.3$，总体）下的搜索指标**。**反**背景效率 $1/\varepsilon_b$（信号效率为 $\varepsilon_s = 0.8$ 时）以及显著性提升 SIC $= \varepsilon_s/\sqrt{\varepsilon_b}$（在 $\varepsilon_b = 10^{-3}$时），均基于符合几何约束的加噪测试分数计算（三个随机种子均值）。在该扰动下，QSVAE 在 $1/\varepsilon_b$ 方面优于所有受测的经典脉冲控制模型，在 SIC 方面则与其中表现最佳者持平；孤立森林仍然遥遥领先。来源：runs/hep stress probes/persignal.csv。

变体	$1/\varepsilon_b$@0。8 SIC@10^{-3}	
QSVAE (正弦^2)	2.1	2.5
FSVAE（LIF）	1.7	1.5
钳制，批归一化 1.0	1.9	2.4
钳制，批归一化 0.80	1.9	1.9
钳制，批归一化 0.60	1.8	1.2
钳制，批归一化 0.45（匹配）	1.7	0.9
钳制，批归一化 0.30	1.6	0.7
逻辑斯蒂，$g = 4$	1.9	2.2
逻辑斯蒂，$g = 2$（斜率匹配）	1.9	2.2
逻辑斯蒂，$g = 4$ (放电匹配)	1.7	1.3
FSVAE，权重衰减 10^{-2}	1.8	1.5
FSVAE，权重衰减 3×10^{-2}	1.7	1.4
FSVAE，权重衰减 10^{-1}	1.8	1.7