# 整理（3.29）

Learning Interpretable Representations of Entanglement in Quantum Optics Experiments using Deep Generative Models（**使用深度生成模型学习量子光学实验中纠缠的可解释表示）**

## 主要内容

量子实验会产生有趣的现象，例如干涉或纠缠，这也是许多技术的核心特性，本文提出了第一个量子光学实验的深度生成模型（其中使用了变分自动编码器QOVAE进行训练）。文章主体部分对训练思路和成果进行概述，在补充中详细介绍了训练的具体理论和数学内容。它们解决了科学领域中黑盒模型 [49, 50] 的可解释性问题，而了解这些模型学到的东西可能会导致新的计算机启发的科学见解和发现。

文章在对训练结果进行分析时发现可以通过二维图像对实验是否可以导致纠缠进行一定的解释（这也是我们解释QOVAE如何构建潜在空间的方法）。同时QOVAE学习了实验结构和纠缠之间的关系，能够为具有与其训练数据匹配的特定分布的纠缠量子态生成新的实验。值得注意的是文章的训练数据来自Melvin算法来生成的数据集。

## 模型的理论及结果

* 1. 一些小的点：

1. 为什么关注用机器学习进行量子纠缠相关的实验设计？

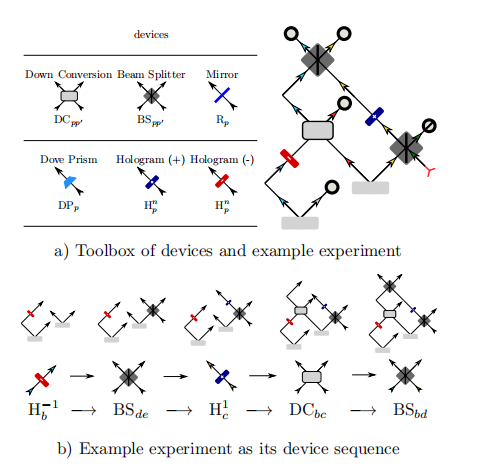
量子纠缠与我们的习惯的经典物理图像不太一样，也很难直观理解相关现象以及现象与实验装置之间的关系，同时纠缠态是许多新兴技术的核心特性。高效设计实验十分有必要。

1. 文中提及现有的一些已有的工作：

其他工作展示了如何使用遗传算法 [18-20]、强化学习 [21] 或参数化优化 [22] 来优化设置。 但是这些努力不会通过使用在实验示例上训练的学习表示来直接生成量子光学实验。

1. 文章使用的训练方法：

深度无监督学习（Deep Unsupervised Learning）



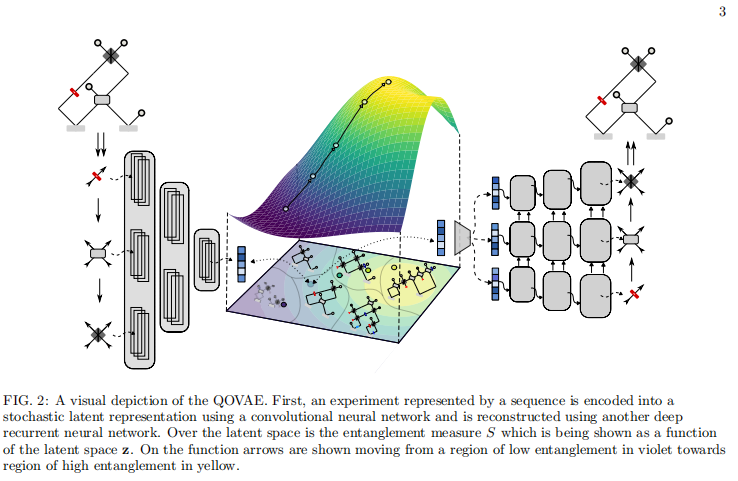
图（1）

* 1. 模型思路

1. 实验中使用了一系列的光学器件（图样如图（1）所示）：

其中各个器件及详情：

* 1. Down-converters
  2. Beam Splitters
  3. Mirrors
  4. Dove Prism
  5. Hologram（+）and (-)



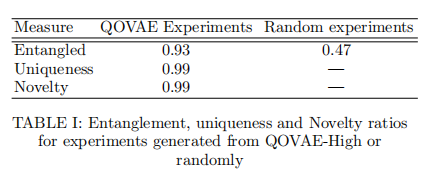
图（2）

1. 思路：

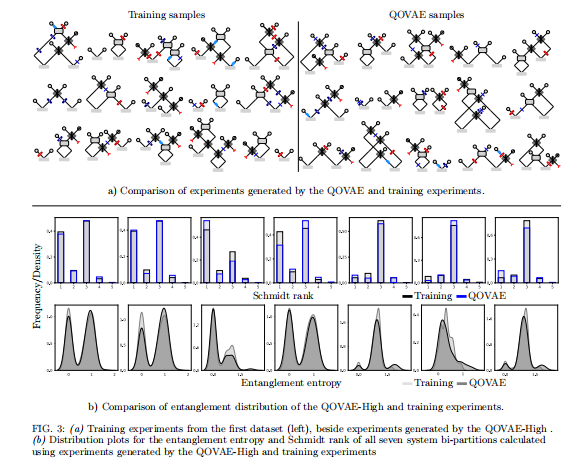
文章选用了6种不同的实验器材（设定实验最多有6种实验器材），并给它们编号。（如图（1）示，每个实验都可以按顺序表示，其中序列中的每个元素都代表一个设备，该设备由其在图中的位置唯一标识，由传播通过的光子及其在序列中的顺序指定。）

文章研究的系统是一个高维四光子量子态，为了量化纠缠，文章使用了纠缠熵。

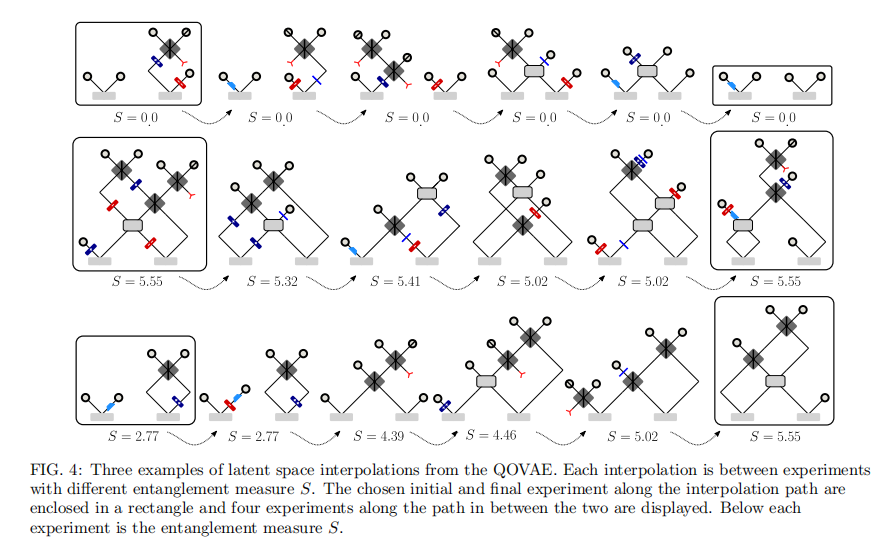
QOVAE的构造：对QOVAE模型，使用变分自动编码器学习量子实验的分布情况并作为序列。如图（2）示，QOVAE将实验的顺序序列作为一系列one hot的向量**x**输入编码器构建潜在空间，最后由解码器重建实验**x**。**x**的构建如图（1）的b所示。



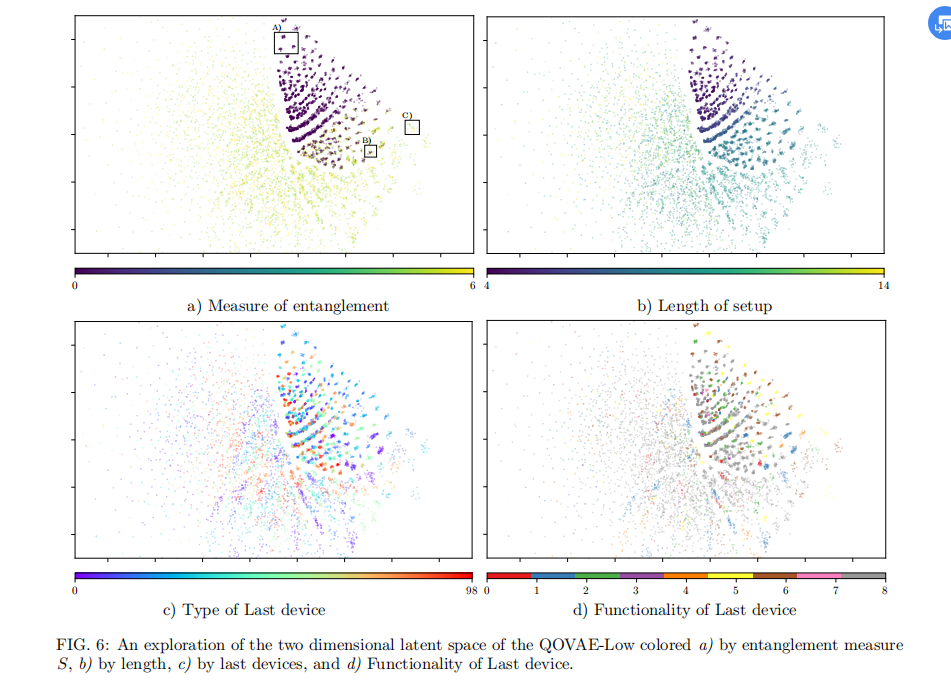
图（3）



图（4）



图（5）



图（6）

3）结果：

由Melvin生成了一个包含大约 200K个实验的数据集，其中一半产生了多方纠缠态。 从生成的实验中，文章仅使用最多具有 6 个设备的装置，这些设备是分束器或下变频器（两个路径设备）。实验中训练了两个模型QOVAE-High（80k个已纠缠的数据集训练）和QOVAE-Low（35k个包含纠缠、未纠缠的实验，并且具有二维潜在空间）。

在训练完QOVAE-High后，随机选取10k个潜在空间的潜在向量，进行解码获得10k个实验。**那么从这种独特的数据集训练出来的模型是否可以使有纠缠的实验在总生成的实验中占比更高呢？**由图（3）第一行可知，比例为0.93，而随机生成实验只能给出比例为0.47的纠缠实验。下两行说明了实验的唯一程度和新颖程度，均为0.99。

接下来，文章评估了 QOVAE-High 学习到的纠缠分布。 假设 QOVAE-High 可以学习创建具有纠缠态分布的实验。 在这种情况下，每个双分区或子系统的纠缠分布在训练数据和模型的实验之间应该是相似的。为了测试这一点，文章从模型和训练数据中抽取 10K 次实验，并计算每个抽样实验的所有七个双分区的纠缠熵和施密特秩等级。绘制其分布用到了核密度估计器 [45]，结果如图（4）示。可以从图中看到，在前四个子系统（前四列）中，QOVAE-High 成功地学习到训练纠缠分布具有两个模式（或峰值）。 对于最后三个子系统图（最后三列），QOVAE-High 能够了解到训练分布中存在单一纠缠模式。 因此文章得出结论，QOVAE 已经学会为系统的每个分区匹配纠缠的训练分布。

之后文章测试了QOVAE-Low是否可以区分其训练数据中的两个分布（纠缠的实验分布与非纠缠的实验分布）。最终发现在10k的随机生成结果里，纠缠与非纠缠实验的比例为0.54。复现了训练数据里纠缠与非纠缠的双峰分布。

之后文章探讨了QOVAE在纠缠实验学习中的准连续嵌入能力。文章根据纠缠度量 S 分析潜在空间的平滑度。具体来说，我们测试在潜在空间中接近的实验是否具有相似的纠缠特性。 这可以通过在潜在空间中从实验的一个潜在表示 z1 到另一个 z2 进行插值，并在从 z1 → z2 的路径上解码实验来完成。 我们对路径使用球面线性插值 [46]。 该实验表明，QOVAE 学习了量子光学实验的相当平滑的表示，该表示编码了实验结构和纠缠特性之间的相似性度量。

同时QOVAE可以用来高效搜索高纠缠度的实验。（理论和上一段差不多，不再详述）

**探究可解释性的部分**：（这部分基本是删减翻译的了，我不太敢乱总结）

由于 QOVAE-Low 只有两个潜在维度，我们可以直接绘制潜在空间并根据结构-性质关系解释潜在空间。 这些图有助于我们解释获得的结果。文章通过将每个训练示例编码到潜在空间来执行这种解释，并观察显着的结构和不同的设置集群（见图 6）。文章使用相应纠缠度量对潜在空间中的点进行颜色编码。 我们在潜在空间中发现了一个非常清晰的未纠缠区域（紫色），以及另一个有纠缠的结构（见图 6a）。因为模型完全没有相应的纠缠信息进行训练，只是使用了设置的结构，该结果表明该模型隐含地区分了纠缠和非纠缠实验。（因为模型完全没有相应的纠缠信息进行训练，只是使用了设置的结构）。 文章分析了对应于三个不同区域的训练示例，未纠缠区 A)，共享纠缠和未纠缠状态区 B)，仅包含纠缠状态区 C)。

在区域 A) 中，我们发现三个不同的集群，其中没有包含纠缠态。 我们发现所有三个集群都包含只有四个设备的实验，而在区域 B) 和 C) 中，所有设置的设备长度均为 9。 因此，我们首先了解 QOVAE 通过光学设备的数量对训练示例进行编码。 我们通过设置长度对潜在空间进行颜色编码来确认这一点，如图 6b 所示。

然而，设置的大小不能单独对潜在空间的非常特殊的结构及其纠缠分布负责。 毕竟，区域 B) 和 C) 具有相同数量的元素，但是 B) 8 包含纠缠和非纠缠设置，而 C) 仅包含纠缠设置。 我们进一步分析这两个区域并发现，区域 B) 的每个设置中的最后一个元素始终是而在 C) 中始终是 。(H、DC是装置的编码，不用在意)。 因此，我们了解设置的最后一个元素是 QOVAE 用于分隔设置的第二个属性，我们可以通过按最终元素对所有点进行颜色编码来确认这一点（见图 6c）。有了这个新见解，我们终于可以解释为什么两个区域 B) 和 C) 具有不同的纠缠特性：每个设置（通过构造）中的初始两个元素由路径 a、b 和 c、d 中的一个非线性晶体组成， 在所有四个路径中创建一个相关状态。

为了在叠加中获得第二个相关状态（纠缠的条件），装置需要在四个路径中创建另一个相关组合。 仅有的另外两种方法是在路径 ac 和 bd 或路径 ad 和 bc 中创建相关性，这可以用图论的方式来理解 [47, 48]。 在区域 C) 中，最后一个元素是路径 ac 中的晶体，它已经创建了必要的相关性。 在这个设置中，前面的八个元素只需要在路径 bd 中创建另一个相关来创建纠缠，而在区域 B）中，需要创建完全相关的状态，这不太可能发生。 我们可以通过将所有可能的最后元素聚集成八个官能团来进一步加强我们的解释，例如，所有全息图对纠缠特性的影响相同，因此我们可以将它们组合成一个官能团。这时我们已经可以利用这些结果解释某些东西了，比如编码在纠缠方面是平滑的，是因为光学元件的数量显着决定了从设置内部建立的可能相关性。

## 思考和注意到的地方

a).在2的3）中提到QOVAE-High是由筛选出来的已经确定有纠缠的实验来训练的，那么这些筛选出的实验数据集会不会有问题？比如说全部都有某一个元素？这个时候会不会导致数据集失去普遍性，导致生成的实验是片面的，没有包含所有可以产生纠缠的情况。进一步，会不会导致最终该模型的可解释性存在问题，我们的模型没有真正的能够区分纠缠和未纠缠实验的特征？

## 相关名词解释和补充：

1. Quantum entanglement（量子纠缠）
2. Melvin算法：它学习使用有用的元素 [16] 扩展自己的工具箱，或基于图形的拓扑优化器，允许提取新的人类可解释的概念 [17]。
3. Deep Unsupervised Learning（深度无监督学习）
4. Deep generative models(深度生成模型)：已被应用于图像、文本、音频、化学生成化合物上。
5. The entanglement entropy（纠缠熵）从离散的施密特秩向量推导出纠缠熵(SRV)[44]。S>0为有纠缠，S=0为无纠缠。
6. Discrete Schmitt rank vector(离散施密特秩向量): SRV 是由所有二元或子系统的施密特秩组成的向量，在4个粒子的情况下，其大小为7。 对于纠缠的整体测量，我们使用所有子系统纠缠熵的总和，我们将其表示为 S，其中 S > 0 的实验是纠缠的，而 S = 0 的实验是未纠缠的。(摘自翻译)
7. Variational Autoencoders（变分自动编码器）
8. One hot
9. 编码器，文中编码器由1D卷积层来进行学习。
10. 解码器，使用循环神经网络生成实验序列
11. 潜在空间
12. bi-partitions（双分区）
13. Schmidt ranks（施密特秩）
14. 核密度估计器 [45]
15. Functional group（官能团）