arXiv:2601.04592v1 [cs.LG] 8 Jan 2026

密度矩阵循环神经网络（DM-RNN）：一种用于建模音乐语境与复调的量子信息理论框架

Joonwon Seo[1][0009-0000-5162-9915] 和
Mariana Montiel[1][0000-0002-7859-1547]✉

数学与统计学系，乔治亚州立大学，美国乔治亚州亚特兰大，邮编 30303 {jseo27, mmontiel}@gsu.edu

摘要。经典循环神经网络（RNN）将音乐语境概括为确定性的隐藏状态向量 $h_t \in \mathbb{R}^d$，形成的信息瓶颈无法捕捉音乐固有的歧义。我们提出密度矩阵循环神经网络（DM-RNN），这是一种采用密度矩阵 $\rho_t \in \mathcal{D}(\mathcal{H}_d)$ 的新型理论架构。该架构使模型能够维持音乐诠释的统计系综（混合态），同时捕捉经典概率和量子相干性。我们使用量子信道（完全正迹保持映射）严格定义时间动态。尤为重要的是，我们详细介绍了一种基于 Choi-Jamiołkowski 同构的参数化策略，从构造上确保学习到的动态在物理上有效（完全正迹保持）。我们引入一种分析框架，使用冯·诺依曼熵 $S(\rho)$量化音乐的不确定性，并使用量子互信息（QMI）测量声部之间的纠缠。DM-RNN 为建模复杂且含有歧义的音乐结构提供了一个数学上严谨的框架。

关键词： 计算音乐理论 · 密度矩阵 · 量子信息理论 · CPTP 映射 · Choi 矩阵 · 循环神经网络 · 冯 · 诺依曼熵。

1 引言

对音乐序列进行计算建模是一项极具挑战性的任务。音乐具有复杂的时间依赖性和内在的模糊性 [11]。在任意给定的时间步 $t \in \mathbb{Z}^+$，一段音乐往往并不意味着只有一种后续走向，而是对应着多种可能性的概率分布。

传统序列模型（如长短期记忆网络 [8]或 Transformer [18], ）会将先前的语境压缩为单个隐藏状态向量， $h_t \in \mathbb{R}^d$。这种向量表示迫使模型采用点估计，从而使丰富多彩的音乐可能性消弭于一处，也难以捕捉细微的模糊性。

我们提出一种新颖的理论框架——密度矩阵循环神经网络（DM-RNN）。该框架受量子力学的数学形式主义启发，天然适合描述具有统计不确定性与干涉特征的系统。

1.1 密度矩阵方法

我们方法的基石是密度矩阵 ρ。在复希尔伯特空间 $\mathcal{H}$上运算， ρ能够完整描述系统的统计状态。通过将 ρ 用作隐藏状态，我们不将音乐语境建模为确定的量子态，而是建模为一组可能性的统计系综（混合态）。

DM-RNN 将传统的向量 $h_t \in \mathbb{R}^d$ 替换为密度矩阵$\rho_t \in \mathcal{D}$（$\mathcal{H}_d$）。这种形式主义提供了更丰富的表示：

1. **布居（对角元素）：** 表示系统处于特定潜在基态的经典概率。2. **相干性（非对角元素）：** 编码这些状态之间稳定的相位关系，从数学上捕捉不同音乐诠释之间的干涉关系或结构关联。

1.2 严格的时间动态与参数化

我们使用量子信道（定义为完全正迹保持（CPTP）映射）来形式化时间演化，$\rho_{t-1} \mapsto \rho_t$。确保学习到的神经网络变换遵循这些约束并非易事。本研究的一项关键贡献，是通过 Choi-Jamiołkowski同构严格定义模型的参数化，从而能够通过一种应用于 Kraus算子的可微归一化方案来强制满足 CPTP 约束（第 4.4 节）。

1.3 一种新的分析框架

DM-RNN 框架支持基于矩阵理论 [9] 和量子信息度量的分析。我们使用冯·诺依曼熵（VNE）S（ρ）来量化音乐不确定性。对 ρ_t 进行谱分析，可揭示占主导地位的音乐本征态。为处理复调，我们引入量子互信息（QMI）来量化纠缠，前提是采用特定的张量积结构（第 5.3 节）。

1.4 主要贡献

本研究的主要贡献在于理论方面：

1. DM-RNN架构的提出。

2. 使用 Choi 矩阵对其参数化进行严格的数学定义，以确保完全正迹保持动力学。 3. 基于量子信息度量（冯·诺依曼熵、量子互信息）和矩阵理论的计算音乐学新型分析框架。

本文着重于理论发展。实际实现面临重大的计算挑战，因此需要采用高效方法，例如第 6.3 节讨论的张量网络。

2 背景与相关工作

我们首先严格考察经典循环神经网络的数学框架，并将 DM-RNN 置于现有方法的背景下进行阐释。

2.1 序列建模形式体系

在标准序列建模形式体系中，定义一个有限词汇表 $\mathcal{V}$ 。一个序列表示为 $X = (x_1, \ldots, x_T)$，其中 $x_t \in \mathcal{V}$。目标是估计 $P(X)$，并将其自回归地分解。当 $x_{<t}$ 表示前驱序列（$x_1, \ldots, x_{t-1}$）时，概率 $P(X)$ 可表示为：

$$P(X) = \prod_{t=1}^{T} P(x_t | x_{<t}) \tag{1}$$

其中， $t = 1$ （即 $x_{<1}$）的条件为空。

2.2 经典循环神经网络架构

经典循环神经网络维护一个隐藏状态 $h_t \in \mathbb{R}^d$。输入被映射为嵌入向量 $e_t \in \mathbb{R}^m$。其演化由参数化的转移函数f 控制： $\mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$：

$$h_t = f(h_{t-1}, e_t; \theta) \tag{2}$$

输出分布由 softmax函数 σ给出：

$$P(x_{t+1} | x_{<t}; \theta) = \sigma(W_{\text{out}} h_t + b_{\text{out}}) \tag{3}$$

其中 $W_{\text{out}} \in \mathbb{R}^{|\mathcal{V}| \times d}$ 和 $b_{\text{out}} \in \mathbb{R}^{|\mathcal{V}|}$。

2.3 表征瓶颈

经典循环神经网络方法的关键局限在于，隐藏状态 h_t是 $\mathbb{R}^d$中的点估计。这在对歧义进行建模时形成了瓶颈。如果考虑两种不同的音乐诠释 h_A， $h_B \in \mathbb{R}^d$，循环神经网络可能会学到一种插值， $h_t \approx \alpha h_A + (1 - \alpha) h_B$，其中包含 $\alpha \in [0, 1]$。这种插值与统计系综有着本质区别，可能导致信息丢失。

2.4 不确定性的经典方法

贝叶斯循环神经网络 [7] 和变分循环神经网络（VRNNs） [5] 引入了随机性。这些方法主要关注捕捉认知不确定性（关于参数）或偶然不确定性（数据变化）。它们通常将隐藏状态建模为向量空间 $\mathbb{R}^d$上的经典概率分布。

2.5 量子启发模型与张量网络

将量子形式体系应用于机器学习是一个新兴领域 [2]。密度矩阵形式体系已在多个领域得到探索。在自然语言处理（NLP）中，研究者使用密度矩阵表示词嵌入，并将多义性和下义关系建模为混合态 [1]。

量子声化。 在听觉领域，与量子形式体系的交叉也是一个活跃的研究方向。Christie 和 Trayford [4]近期提出了“开放量子声化”框架，将密度矩阵元素及其相位映射为听觉信号。具体而言，他们将非对角相干项呈现为双耳信号，以直观地展示退相干等量子现象。尽管这些方法侧重于通过声音分析物理量子系统，DM-RNN 则反过来采用密度矩阵形式体系和完全正迹保持动力学作为严谨的生成框架，对音乐结构本身隐含的歧义与相干性进行建模。

此外，张量网络（TN）最初源自凝聚态物理，用于高效表示高维量子态 [13], ，如今也已应用于经典机器学习，包括序列建模 [16]。张量网络方法通常利用矩阵乘积态（MPS）对序列建模，着重于高效捕捉相关性。

DM-RNN 的方法独具特色。尽管与张量网络模型相关 [16], ，DM-RNN 会明确强制状态演化遵循量子有效动力学（完全正迹保持映射），而不是优化经典 MPS 表示。与音乐 Transformer 等基于注意力的模型不同，后者通过注意力分数计算长程依赖关系 [10],；DM-RNN 侧重于状态表示本身（ρ_t）的严格演化，从而天然捕捉歧义与相干性。密度矩阵比经典概率分布提供了更丰富的表示，因为它能明确捕捉相干性（非对角元素），以表示不同解释之间的相位关系（第 5.1.2 节）。

3 音乐中的密度矩阵形式体系

我们采用密度矩阵形式体系来表示音乐语境。

3.1 数学预备知识与定义

我们在一个 d 维复希尔伯特空间 $\mathcal{H}_d \cong \mathbb{C}^d$ 中定义状态空间。使用复空间而非实空间 $\mathbb{R}^d$ 至关重要，因为

非对角元素（相干性）需要用复数来编码基态之间的相位关系，这对于区分经典混合态和相干叠加态（第5.1.2节）是必不可少的。令 $\mathcal{L}(\mathcal{H}_d)$ 表示 $\mathcal{H}_d$ 上线性算子的空间。

纯态是一个归一化向量 $|\psi\rangle \in \mathcal{H}_d$。对应的密度矩阵为 $\rho_{\text{pure}} = |\psi\rangle\langle\psi|$。

混合态。 混合态 表示一个统计系综 $\{(p_i, |\psi_i\rangle)\}_{i=1}^N$，其中 $p_i \geq 0$ 和 $\sum_i p_i = 1$。密度矩阵 $\rho \in \mathcal{L}(\mathcal{H}_d)$ 是以下凸组合：

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| \tag{4}$$

密度矩阵集合。 所有有效密度矩阵的集合 $\mathcal{D}(\mathcal{H}_d)$定义如下：

$$\mathcal{D}(\mathcal{H}_d) := \{\rho \in \mathcal{L}(\mathcal{H}_d) \mid \rho = \rho^\dagger, \mathrm{Tr}(\rho) = 1, \rho \geq 0\} \tag{5}$$

这是一个由厄米性、迹为一和半正定性（PSD）定义的凸集。

3.2 音乐诠释与基的作用

当 $\rho_t \in \mathcal{D}(\mathcal{H}_d)$ 作为隐藏状态时，其解释依赖于 $\mathcal{H}_d$ 的一组正交归一基 $\{|u_i\rangle\}_{i=1}^d$。在神经网络中，这个基是在训练期间形成的。虽然布居（ρ_{ii}）和相干性（ρ_{ij}）的具体解释取决于基，但关键度量（例如熵）在酉基变换下保持不变。

训练期间的基稳定化。 尽管密度矩阵形式具有酉规范自由度，DM-RNN架构会通过学习得到的POVM算子 $\{M_v\}$（第4.5节）稳定解释所用的基。优化目标（例如基于玻恩规则的负对数似然）促使模型采用一个特定的基，以便将内部状态 ρ_t 最优地映射到输出词汇表 $\mathcal{V}$，从而确保形成对音乐结构的一致内部模型。

3.3 量化音乐歧义

密度矩阵提供了内在且与基无关的度量，可用于量化歧义。

纯度。 纯度 $\gamma(\rho) := \mathrm{Tr}(\rho^2)$ 的范围为 $1/d \leq \gamma(\rho) \leq 1$。当且仅当量子态是纯态时，$\gamma(\rho) = 1$。

冯·诺依曼熵（VNE）。衡量混合度最全面的指标是冯·诺依曼熵，$S: \mathcal{D}(\mathcal{H}_d) \to \mathbb{R}_{\geq 0}$：

$$S(\rho) := -\mathrm{Tr}(\rho \log_2 \rho) = -\sum_{i=1}^{d} \lambda_i \log_2 \lambda_i \tag{6}$$

其中， $\{\lambda_i\}$ 是 ρ的特征值。 $S(\rho)$量化了混合态中固有的不确定性（以比特为单位）。

4 提出的架构：DM-RNN

我们介绍 DM-RNN 架构，重点阐述其时间演化和参数化的严格数学构造。

4.1 架构概述

DM-RNN 处理输入嵌入 $e_t \in \mathbb{R}^m$ 并更新其隐藏状态 $\rho_{t-1} \in \mathcal{D}(\mathcal{H}_d)$。递归关系由参数化的转移函数 $\mathcal{F}$ 定义： $\mathcal{D}(\mathcal{H}_d) \times \mathbb{R}^m \to \mathcal{D}(\mathcal{H}_d)$:

$$\rho_t = \mathcal{F}(\rho_{t-1}, e_t; \theta) \tag{7}$$

4.2 通过量子信道实现时间动态（完全正迹保持映射）

转移函数 $\mathcal{F}$ 必须实现为量子信道， $\mathcal{E}: \mathcal{L}(\mathcal{H}_d) \to \mathcal{L}(\mathcal{H}_d)$，且必须是完全正且保迹（CPTP） [12]。

CPTP映射的必要性。

– 保迹（TP）： 确保 $\mathrm{Tr}(\mathcal{E}(\rho)) = 1$。**– 完全正（CP）：** 确保对于任意辅助希尔伯特空间$\mathcal{H}_A$，诱导映射（$\mathcal{E} \otimes \mathbb{I}_A$）（其中 $\otimes$ 表示张量积， $\mathbb{I}_A$是作用于 $\mathcal{H}_A$上的恒等映射）仍为正，从而保证其物理有效性。

4.3 克劳斯算子表示

CPTP映射 $\mathcal{E}$ 总可以用克劳斯分解表示:

$$\mathcal{E}(\rho) = \sum_{k=1}^{K} K_k \rho K_k^\dagger \tag{8}$$

其中 $\{K_k\} \subset \mathcal{L}(\mathcal{H}_d)$ 是满足完备性关系的 Kraus算子（$K \leq d^2$）：

$$\sum_{k=1}^{K} K_k^\dagger K_k = \mathbb{I}_d \tag{9}$$

4.4 完全正迹保持映射的严格参数化

核心挑战在于定义一个参数化函数，以生成通道 $\mathcal{E}_t(e_t;\theta)$，同时严格满足完全正迹保持约束。我们通过利用 Choi-Jamiołkowski 同构来确保完全正性（CP），并采用可微分的归一化方案来确保迹保持（TP）。

Choi-Jamiołkowski同构。线性映射 $\mathcal{E}$ 与 Choi矩阵 $C_{\mathcal{E}} \in \mathcal{L}$（$\mathcal{H}_d \otimes \mathcal{H}_d$）之间存在一一对应关系。该同构通过将通道 $\mathcal{E}$ 应用于最大纠缠态的一部分来建立。具体而言， $C_{\mathcal{E}} :=$ （$\mathcal{E} \otimes \mathbb{I}$）（$|\Phi^+\rangle\langle\Phi^+|$），其中 $|\Phi^+\rangle = \sum_{i=1}^{d} |u_i\rangle \otimes |u_i\rangle$是（未归一化的）最大纠缠态， $\mathbb{I}$ 是恒等映射。 $\mathcal{E}$是 CP 当且仅当 $C_{\mathcal{E}} \geq 0$。 $\mathcal{E}$ 是 TP 当且仅当对输出空间进行偏迹后满足 $\mathrm{Tr}_{\mathrm{out}}$（$C_{\mathcal{E}}$） $= \mathbb{I}_d$。

参数化构造。我们通过直接生成并归一化 Kraus算子来构造参数化，从而以可微方式确保 CP 和 TP。

1. **通过 Cholesky因子 强制满足 CP：** 我们利用 Choi 矩阵的 Cholesky因子 的列对应于向量化的 Kraus算子 这一性质。设 G 为一个由 $\theta_G \subset \theta$ 参数化的神经网络，它生成一个复矩阵 $L_t \in \mathbb{C}^{d^2 \times d^2}$ （即 Cholesky因子）：

$$L_t = G(e_t;\theta_G) \tag{10}$$

由此构造的相应 Choi 矩阵 $C'_{\mathcal{E}_t} = L_t L_t^\dagger$ 是半正定的，从而保证满足 CP。

2. **提取 Kraus算子：** 我们通过对 L_t 的各列进行重塑（逆向量化），获得一组未归一化的 Kraus算子 $\{K'_k\}_{k=1}^{d^2}$：

$$K'_k(t) = \mathrm{mat}(L_t[:,k]) \tag{11}$$

其中，mat(·) 是逆向量化操作（vec^{-1}），用于将 $d^2 \times 1$ 列向量重塑为一个 $d \times d$ 矩阵。

3. **通过可微归一化强制满足 TP：** 我们通过归一化层 [17] 强制满足完备性关系（公式 9）。首先，我们计算归一化矩阵 $S_t \in \mathbb{C}^{d\times d}$:

$$S_t = \sum_{k=1}^{d^2} (K'_k(t))^\dagger K'_k(t) \tag{12}$$

我们对算子进行归一化:

$$K_k(t) = K'_k(t) S_t^{-1/2} \tag{13}$$

为确保 S_t 可逆并维持数值稳定性，我们使用经过正则化的逆平方根（例如，向 S_t 添加一个较小的 $\epsilon\mathbb{I}_d$，通常$\epsilon \approx 10^{-6}$）。此操作可微，因此能够进行端到端训练。

DM-RNN 更新。使用式 8 中归一化后的 Kraus算子进行量子态更新。这种方法可严格保证 $\mathcal{E}_t$ 是完全正迹保持的，并避免对整个 Choi 矩阵进行计算开销巨大的特征分解（其复杂度随 $O(d^6)$增长）。

4.5 通过测量进行预测（正算子值测度）

为预测 $x_{t+1} \in \mathcal{V}$，我们使用正算子值测度（正算子值测度）。我们定义一组可学习的测量算子 $\{M_v\}_{v\in\mathcal{V}} \subset \mathcal{L}(\mathcal{H}_d)$，并满足正性（$M_v \geq 0$）和完备性（$\sum_{v\in\mathcal{V}} M_v = \mathbb{I}_d$）。

观测到事件 v 的概率由玻恩规则给出：

$$P(x_{t+1} = v|\rho_t) = \mathrm{Tr}(M_v \rho_t) \tag{14}$$

正算子值测度参数化。我们引入辅助的可学习矩阵 $\{A_v\}_{v\in\mathcal{V}} \subset \mathbb{C}^{d\times d}$，并由 $\theta_M \subset \theta$ 参数化。我们定义 $M'_v = A_v^\dagger A_v$（保证正性）。通过归一化确保完备性：令 $S = \sum_{v\in\mathcal{V}} M'_v$。有效的正算子值测度元素为：

$$M_v = S^{-1/2} M'_v S^{-1/2} \tag{15}$$

为处理 S 可能存在的奇异性并确保数值稳定性，我们采用正则化逆平方根（例如，当 $\epsilon > 0$ 较小时使用 $S + \epsilon\mathbb{I}_d$，或使用摩尔－彭若斯广义逆），这与第 4.4 节采用的归一化策略类似。

5 矩阵理论与音乐分析

DM-RNN 框架为基于矩阵理论和量子信息理论的分析奠定了严谨的基础。

5.1 量化音乐不确定性与相干性

语境 ρ_t 中的歧义通过冯·诺依曼熵$S(\rho_t)$（公式 6）来衡量。

$$S(\rho_t) = -\sum_{i=1}^{d} \lambda_i(t) \log_2 \lambda_i(t) \tag{16}$$

$S(\rho_t)$ 的演化追踪音乐张力（混合）与消解（纯化）。

示例：通过混合态建模歧义我们以 $\mathcal{H}_2 \cong \mathbb{C}^2$ 中的一个示例进行说明，其由 $|A\rangle$（主音）和 $|B\rangle$（属音）张成。

1. t=0： 清晰。 $\rho_0 = |A\rangle\langle A|$。 $S(\rho_0) = 0$。**2. t=1（张力）：** A 和 B 同样合理。

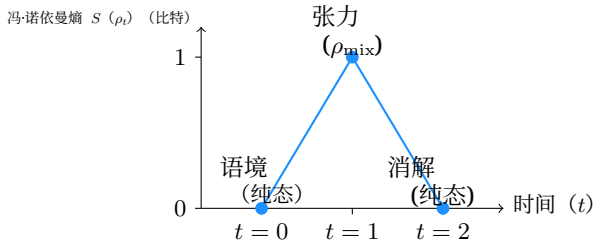

图 1. 冯·诺依曼熵（VNE）的概念演变 $S(\rho_t)$。歧义最大时，熵达到峰值，并表示为混合态（ρ_{mix} 在 $t=1$处）。

- DM-RNN: Maximally mixed state: $\rho_1 = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$. $S(\rho_1) = 1$ bit.
- Classical RNN: Interpolation: $h_1 \approx 0.5h_A + 0.5h_B$. Potential information loss.

3. **t=2:** 已解决为 B。 $\rho_2 = |B\rangle\langle B|$。$S(\rho_2) = 0$。

相干性的作用：超越经典混合态 关键的是，密度矩阵形式主义能够捕捉超越经典统计混合的现象。考虑前一个示例中的混合态 ρ_1 （记作 ρ_{mix}）。它表示经典不确定性。

现在考虑叠加态，例如纯态 $|\psi_+\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$。相应的密度矩阵为：

$$\rho_{\text{sup}} = |\psi_+\rangle\langle\psi_+| = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \tag{17}$$

虽然布居（对角元素）与 ρ_{mix}相同，但非零相干性（非对角元素）表明$|A\rangle$ 与 $|B\rangle$之间存在稳定的相位关系。该量子态是纯态（$\gamma(\rho_{\text{sup}}) = 1$，$S(\rho_{\text{sup}}) = 0$），表示一种确定性但叠置的音乐语境。

音乐诠释：结构性歧义与不确定性。区分 ρ_{mix} 和 ρ_{sup} 对于音乐歧义建模至关重要。以转调过程中使用的枢纽和弦为例，它同时在两个不同的调性中发挥作用。令 $|A\rangle$表示原调中的诠释，$|B\rangle$表示目标调中的诠释。

– 经典混合态（ρ_{mix}）： 表示认知不确定性。模型（或听者）无法确定语境是A还是B。该量子态具有较高的熵（$S(\rho_{\text{mix}}) = 1$ 比特），反映了真实的困惑或信息缺失。**– 相干叠加态（ρ_{sup}）：** 表示作品中固有的结构性歧义。枢纽和弦同时发挥 A和B 的作用。这是一个纯态（$S(\rho_{\text{sup}}) = 0$），反映了这种双重功能的确定性本质。相干性编码了这两种诠释之间稳定的关系。

表示并演化这些相干性的能力，使 DM-RNN 有别于经典概率模型。从 ρ_{sup} 到 ρ_{mix} 的演化存在显著差异，因为动力学（完全正迹保持映射）作用于相干性，使模型能够捕捉结构化音乐关系如何根据学得的干涉模式得到消解。

5.2 识别音乐本征态： ρ_t 的谱分析

由于 ρ_t 是厄米矩阵，谱定理保证它可以对角化。

特征分解。

$$\rho_t = U_t \Lambda_t U_t^\dagger = \sum_{i=1}^{d} \lambda_i(t) \, |\psi_i(t)\rangle \, \langle \psi_i(t)| \tag{18}$$

其中， U_t 是一个酉矩阵，其列为特征向量 $|\psi_i(t)\rangle$； Λ_t是由相应特征值 $\lambda_i(t)$构成的对角矩阵。

– **特征向量（$|\psi_i(t)\rangle$）：** 这些音乐本征态是构成系综 ρ_t的纯态。– **特征值（$\lambda_i(t)$）：** 它们表示本征态上的概率分布。

5.3 分析复调相关性：纠缠

密度矩阵形式体系能够严谨地对各子系统之间的相关性（纠缠）进行建模。

二部系统表示。 考虑两个声部 A 和 B。我们在复合希尔伯特空间 $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ 中分析该系统。联合量子态为 $\rho_{AB} \in \mathcal{D}(\mathcal{H})$。

先决条件：施加张量积结构。 一个关键的先决条件是， $\mathcal{H}_d$ 必须具有明确定义的张量积结构（$\mathcal{H}_d \cong \mathcal{H}_A \otimes \mathcal{H}_B$），并与音乐声部相对应，使得 $d = d_A \times d_B$。这种结构必须通过架构设计明确施加。

结构的架构性强制。 在 DM-RNN 的语境中，这一施加方式决定了参数化量子信道$\mathcal{E}_t$ （第 4.4 节）的设计。神经网络 G必须经过结构设计，以便在这一复合空间中运行。

一种实用的方法是将动力学分解为局部操作和相互作用项：

$$\mathcal{E}_t \approx (\mathcal{E}_A \otimes \mathcal{E}_B) \circ \mathcal{E}_{\text{int}} \tag{19}$$

其中， $\mathcal{E}_A$ 和 $\mathcal{E}_B$ 是分别独立作用于 $\mathcal{H}_A$ 和 $\mathcal{H}_B$上的局域信道，而$\mathcal{E}_{\text{int}}$ 是用于建模相互作用的纠缠信道。 G可以通过专门的模块来构建，分别为这些组成信道生成参数。这样可以确保学习到的隐藏状态 ρ_t （表示 ρ_{AB}）保持与音乐声部相对应且有意义的因子分解。

约化密度矩阵。 假设施加了特定结构，声部 A 的量子态可通过对子系统 B 求偏迹得到，Tr_B：$\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_A)$.

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) \tag{20}$$

此操作实际上“消去”了子系统 B 的影响。具体而言，如果 $\{|j\rangle_B\}$ 是 $\mathcal{H}_B$ 的正交归一基，则偏迹的计算方式如下：

$$\rho_A = \sum_j (\mathbb{I}_A \otimes \langle j|_B)\rho_{AB}(\mathbb{I}_A \otimes |j\rangle_B) \tag{21}$$

其中，$\mathbb{I}_A$ 是作用于 $\mathcal{H}_A$ 上的恒等算子。

通过量子互信息量化关联。我们使用量子互信息（QMI） [20], $I(A;B)$，来衡量子系统之间的总关联（包括经典关联和量子关联）：

$$I(A;B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \tag{22}$$

$I(A;B) \geq 0$，当且仅当系统之间不存在关联（$\rho_{AB} = \rho_A \otimes \rho_B$）时取等号。对于混合态，虽然量子互信息为可蒸馏纠缠提供了上界，但它仍是衡量声部间整体相互依赖关系的可靠指标。

示例：可分态与纠缠态。为说明这一点，考虑一个最简双体系统，其中 $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$。

1. **可分态（无关联）：** 如果各声部彼此独立，量子态就是可分的。设 $\rho_{AB}^{\mathrm{sep}} = \rho_A \otimes \rho_B$。假设 $\rho_A = \rho_B = |0\rangle\langle 0|$。那么 $\rho_{AB}^{\mathrm{sep}} = |00\rangle\langle 00|$。熵为 $S(\rho_A) = S(\rho_B) = S(\rho_{AB}^{\mathrm{sep}}) = 0$。因此，$I(A;B) = 0$。 2. **最大纠缠态（有关联）：** 考虑贝尔态 $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$。联合量子态 $\rho_{AB}^{\mathrm{ent}} = |\Psi^+\rangle\langle\Psi^+|$ 是纯态，因此 $S(\rho_{AB}^{\mathrm{ent}}) = 0$。然而，约化密度矩阵（通过公式 20 得到）是最大混合态：$\rho_A = \mathrm{Tr}_B(\rho_{AB}^{\mathrm{ent}}) = \frac{1}{2}\mathbb{I}_2$。因此，$S(\rho_A) = 1$ 比特（同理，$S(\rho_B) = 1$）。量子互信息为 $I(A;B) = 1 + 1 - 0 = 2$ 比特。这意味着关联达到最大：尽管 A 的局部量子态具有最大不确定性，声部 A 的状态仍能完美预测声部 B 的状态。

量子互信息的音乐诠释量子互信息的动态变化为分析音乐织体和声部间的相互作用提供了一种新颖的度量方法。

– **高量子互信息（强相关性／纠缠）：** 预计出现在对位相互依存程度较高的段落（例如赋格曲或卡农）中，此时一个声部的量子态会强烈制约另一个声部的量子态。– **低量子互信息（弱相关性／可分性）：** 预计出现在主调织体中，此时声部以节奏齐奏的方式进行，或某一声部明显占据主导地位。联合量子态趋近于可分态。

$I(A;B)$ 的时间演化因此能够追踪整首复调作品中声部之间不断变化的依赖关系。

6 讨论与未来工作

DM-RNN 框架带来了范式转变，不再局限于向量表示 $h_t \in \mathbb{R}^d$，而是采用基于矩阵的状态 $\rho_t \in \mathcal{D}(\mathcal{H}_d)$。

6.1 理论启示

表达歧义性与相干性的能力。通过利用密度矩阵，DM-RNN 在数学上具备表示统计系综（混合态）和纯叠加态（相干性）的能力，从而能够原生地捕捉音乐语境固有的歧义性。

6.2 计算挑战

可扩展性。 DM-RNN 将隐藏状态表示从 O（经典循环神经网络）扩展到 O。更重要的是，更新过程在参数复杂度方面引入了显著瓶颈。如第 4.4 节所述，对一般完全正迹保持映射进行参数化，需要神经网络 G 生成 Cholesky因子 L_t，这在每个时间步都不可避免地涉及 O 个参数。对于适中的希尔伯特空间维度（例如 $d = 32$），每一步需要 $O \approx O$ 个参数，因此对于复杂的音乐结构，稠密形式在实践中难以处理。

施加约束。优化模型需要对 Choi 矩阵施加迹保持约束，并对正算子值测度施加完备性约束，这些操作的计算开销很大 [17]。

6.3 实施策略与未来方向

通过矩阵乘积算子（MPO）实现高效计算。为解决稠密参数化中难以承受的 $O(d^4)$ 参数复杂度，DM-RNN 的实际实现必须利用张量网络因子分解 [13]。我们推测，具有音乐意义的语境通常纠缠程度较低。在物理学中，许多系统遵循“面积定律”，即纠缠受到限制 [6]。如果音乐语境也遵循类似原理，密度矩阵 ρ_t 便可以用矩阵乘积态（MPS） [14] 高效表示；对于混合态，更准确地说，则可以用矩阵乘积密度算符（MPDO） [19]表示。

关键在于，演化算符（量子信道 $\mathcal{E}_t$）必须表示为矩阵乘积算符（MPO）。MPO 可将高维演化映射（或其在 L_t中的对应 $O(d^4)$参数）分解为由低维张量构成的网络，从而实现高效表示；该网络受一个称为键维数的参数控制， χ。键维数衡量了这种表示所能承载的纠缠容量。如果能够用较小的键维数（$\chi \ll d$）准确近似动力学过程，参数复杂度便可从 $O(d^4)$大幅降低至约 $O(d^2\chi^2)$。对于前面的示例（$d = 32$），使用较小的键维数（例如 $\chi = 8$）可将

复杂度降至 $O(32^2 \cdot 8^2) \approx O(65{,}000)$，实现显著改进。实现策略是定义神经网络 G（第 4.4 节），使其以 MPO 分解形式生成动力学过程。因此，使用 MPO 对 DM-RNN 的可行运行至关重要。

纠缠架构。针对复调建模，开发能够施加并利用张量积结构（第 5.3 节）的专用 DM-RNN 架构，是至关重要的下一步。

理论扩展：开放系统。未来的研究可以探索在开放量子系统框架内对音乐语境进行建模。这种方法能够自然地模拟音乐中的“耗散”或“退相干”——即相干性随时间逐渐丧失，这或许可以表示和声张力的消退或对过往语境的遗忘。在假设演化具有马尔可夫性的情况下，连续时间动力学由 Lindblad 主方程 [3] 支配：

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_k \gamma_k \left(L_k \rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\} \right) \tag{23}$$

第一项通过哈密顿量 H 描述相干演化，第二项（耗散子）则通过跳跃算子 L_k 和速率 γ_k 描述退相干，为音乐预期的演化与消退建模提供了一个严谨的框架。

7 结论

本文介绍了密度矩阵循环神经网络（DM-RNN）这一新颖的理论架构，它采用密度矩阵 $\rho_t \in \mathcal{D}(\mathcal{H}_d)$ 作为隐藏状态表示。这使模型能够自然地捕捉音乐中预期固有的歧义、相干性和统计混合特性。

我们使用完全正迹保持映射严格定义了该架构及其时间动态，并通过Choi-Jamiołkowski同构，为这些映射的参数化提供了具体的数学构造。DM-RNN 框架为基于矩阵理论的计算音乐分析确立了新范式。尽管存在计算方面的挑战，DM-RNN 仍为复杂音乐结构的建模提供了一种数学上严谨的方法。

利益披露。作者声明不存在与本文内容相关的竞争性利益。

参考文献

1. Bankova, D. 等：《意义的几何：语义表示的密度矩阵方法》。载于：国际量子交互会议（QI 2019）。LNCS，第11927卷。Springer（2019）。
https://doi.org/10.1007/ 978-3-030-36701-4_6

2. Biamonte, J. 等：量子机器学习。《自然》549(7671)，195–202（2017 23474 3 10 1038）。https://doi.org/ . /nature . Breuer, H.-P., Petruccione, F.：《开放量子系统理论》。牛津大学出版社，牛津（2002）。https://doi.org/10.1093/acprof:oso/9780199213900.001.0001 4. Christie, R., Trayford, J.：《退相干之声》。arXiv预印本 arXiv:2412.17045（2024）。5. Chung, J. 等：序列数据的循环潜变量模型。载于：神经信息处理系统进展28（NIPS 2015）。（2015）。6. Eisert, J., Cramer, M., Plenio, M.B.：专题讨论：纠缠熵的面积定律。《现代物理评论》82(1)，277–306（2010）。https://doi.org/10. 1103 277 7 82 /RevModPhys. . . Fortunato, M., Blundell, C., Vinyals, O.：贝叶斯循环神经网络。arXiv预印本 arXiv:1704.02798（2017）。8. Hochreiter, S., Schmidhuber, J.：长短期记忆。《神经计算》9 1735 9 8 1735 1780 1997 10 1162 1997 9 8 2（），–（）。https://doi.org/ . /neco. . . . . Horn, R.A., Johnson, C.R.：《矩阵分析》。第nd版。剑桥大学出版社，剑桥（2013 9781139020411 10 10 1017）。https://doi.org/ . /CBO . Huang, C.-Z.A. 等：音乐 Transformer。载于：国际学习表征会议（ICLR 2019）。（2019）。11. Huron, D.：《甜蜜的期待：音乐与期待心理学》。麻省理工学院出版社，马萨诸塞州剑桥（2006）。https://doi.org/10.7551/mitpress/6525.001. 0001 12. Nielsen, M.A., Chuang, I.L.：《量子计算与量子信息》。十周年纪念版。剑桥大学出版社，剑桥（2010）。https: // doi.org/10 9780511976667 13 1017. /CBO . Orús, R.：张量网络实用入门。《物理年鉴》349，117– 158 013 14 2014 10 1016 2014 06（）。https://doi.org/ . /j.aop. . . . Schollwöck, U.：矩阵乘积态时代的密度矩阵重整化群。《物理年鉴》326(1)，96–192（2011）。https://doi.org/10. 1016 012 15 2010 09 /j.aop. . . . Simon, I. 等：演奏 RNN。载于：第19届ISMIR会议论文集。（2018）。https://doi.org/10 1495510 16 5281. /zenodo. . Stoudenmire, E., Schwab, D.J.：张量网络监督学习。载于：神经信息处理系统进展29（NIPS 2016）。（2016）。17. Trouvain, T. 等：神经网络量子通道的参数化与优化。载于：AAAI人工智能会议论文集34 6074 18 04 6108 6115 2020 10 1609 34 04（），–（）。https://doi.org/ . /aaai.v i . . Vaswani, A. 等：注意力就是一切。载于：神经信息处理系统进展30（NIPS 2017）。（2017）。19. Verstraete, F., García-Ripoll, J.J., Cirac, J.I.：矩阵乘积密度算符。《物理评论快报》93(20)，207204（2004）。https://doi.org/10.1103/ PhysRevLett.93 207204 20 . . Wilde, M.M.：《量子信息理论》。第2版。剑桥大学出版社，剑桥（2017）。https://doi.org/10.1017/9781316809976