arXiv:2611.29888v1 [cs.AI] 23 Nov 2026

含噪循环神经网络中的二阶动力学在重放中产生探索与时间压缩

Josue Casco-Rodriguez电气与计算机工程系 莱斯大学，美国 jc135@rice.edu

Nanda H. Krishna
Mila - Quebec AI 研究所 tute & 蒙特利尔大学，加拿大

Richard G.Baraniuk电气与计算机工程系
莱斯大学，美国

摘要

生物神经网络（如海马体）能够在内部生成类似刺激驱动活动的“重放”。近期的重放计算模型采用经过路径积分训练的噪声循环神经网络（RNN）。这些网络中的重放被描述为朗之万采样，但仅凭这一理论，无法解释真实生物重放中观察到的环境探索和时间压缩。为弥补这一不足，我们重新审视自发的含噪 RNN 活动（“重放”），将其视为从类似真实经历的时间序列分布中逐个时间步进行的朗之万采样（这不同于先前将采样定义为对静态分布进行蒙特卡洛推断的研究）。我们提出一个规范性二阶朗之万采样框架：脉冲频率适应以牺牲速度为代价驱动空间探索，而树突输入微分产生的动量则驱动时间压缩。我们通过数值模拟验证了这一框架，模拟对象包括在二维三角形迷宫和 T 形迷宫导航活动上训练的含噪循环神经网络，以及在合成大鼠的位置细胞活动上训练的其他网络。值得注意的是，我们发现适应与动量同时存在会提升重放的探索程度和速度，并为基于放电率的 RNN 重放中符合生物学合理性的时间压缩提供了首批解释之一。我们的研究结果促使人们开展实验，探究树突输入微分或其他机制是否能够支持海马体等记忆回路中的时间压缩。

1 引言

在睡眠或清醒静息等静息时段，一些神经回路会在内部产生类似于活跃时段的活动（Tingley 和 Peyrache，2020年）。这类“重放”现象已在前额叶（Euston 等人，2007，Peyrache 等人，2009）、感觉皮层（Kenetet al.，2003，Xu 等人，2012）、运动皮层（Hoffman 和 McNaughton，2002）和内嗅皮层（Gardner 等人，2022）；前丘脑（Peyrache 等人，2015）；以及海马体（

Buzsáki，1986，Skaggs 和 McNaughton，1996，Nádasdy 等人，1999年，Lee 和 Wilson，2002年，Foster，2017）。在这些神经回路中，海马体尤其值得关注，因为它在活跃状态下执行导航（O’Keefe 和 Nadel，1978，Burgess 等人，1994年，McNaughton 等人，1996）和规划（Pfeiffer 和 Foster，2013，Miller 等人，2017）等任务时表现出的稳健性，似乎与其在静息状态下的自发活动密切相关（Buzsáki，1989，2015，Tononi 和 Cirelli，2014，Ólafsdóttir 等人，2015，2018）。

虽然有些研究通过监督式生成模型实现了重放（Deperrois 等人，2022），但大多数现有的海马活动模型都将重放视为精心设计网络所产生的涌现副产物。相关网络参数包括连接结构（Shen 和McNaughton，

预印本。

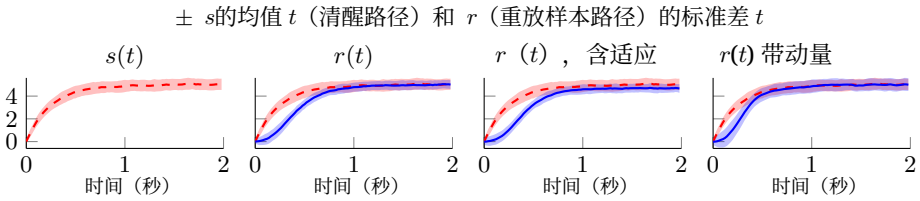

图1：动量加速离线重放，适应则使其减慢。此处，我们模拟一个带噪循环神经网络 $r(t)$，它能够对奥恩斯坦-乌伦贝克过程 $s(t)$的速度 $s'(t)$进行最优路径积分。我们假设 $r(t)$ 最小化公式（8）中的损失函数，因此按照其得分函数$\nabla_{r(t)} \log p(r(t))$（公式 11 和 47）演化；在没有输入时，它会执行一种朗之万采样。上方比较了三种对循环神经网络活动的调节方式：不加调节（亦称过阻尼）、尖峰频率适应（负反馈），以及我们提出的动量机制（可直接实现欠阻尼动力学）。每种调节方式对重放时间序列分布 $p(r(t))$的影响各不相同：适应会减慢 $p(r(t))$，并使其偏离自然稳态；而动量则会加快 $p(r(t))$向其稳态收敛。

1996，Milstein 等人，2023）、局部可塑性机制（Hopfield，2010，Litwin-Kumar 和 Doiron，2014，Theodoni 等人，2018，Haga 和 Fukai，2018，Asabuki 和 Fukai，2025）、放电率适应（Chu 等人，2024，Azizi 等人，2013，Itskov 等人，2011，Dong 等人，2021，Li 等人，2024）以及输入调制（Kangand DeWeese，2019）。虽然这些模型再现了重放的某些特征，但其通常以实证发现为依据，缺乏理论上的论证。

一种更具原理性的海马功能模型能够自然产生重放，即路径积分（Krishna等人, 2024年, Levenstein et al., 2024），在该模型中，神经回路根据对动态环境或任务变量的不完美观测来预测这些变量，例如根据速度测量追踪位置。这种将海马体描述为序列预测器的规范性描述（Levy, 1989, Stachenfeld 等人, 2017）与海马对即将出现的刺激（Davachi 和 DuBrow, 2015年）和预测误差（Aitken 和 Kok, 2022, Miller 等人, 2023）的编码，以及表征可能的未来轨迹的神经活动扫描（Kay 等人, 2020, Johnson 和 Redish, 2007）相吻合。与传统的、手工设计的海马神经回路模型（即连续吸引子网络（Samsonovich 和 McNaughton, 1997, Battaglia 和 Treves, 1998））不同，序列预测学习模型是通过数据训练的。尽管如此，这些模型仍能解释位置细胞（Levenstein 等人, 2024, Recanatesi 等人, 2021, Chen et al., 2024）、网格细胞（Cueva 和 Wei, 2018年, Sorscher 等人, 2019）和头方向细胞（Cueva 等人, 2020, Uria 等人, 2022）的涌现；能够纳入 θ振荡等现象（Levenstein et al., 2024）；并呈现静息重放活动（Krishna等人, 2024年, Levenstein et al., 2024, Chen et al., 2024）。

Krishna等人（2024年）首次为噪声路径积分网络中的重放奠定了理论基础，证明其静息（即由噪声驱动、没有输入）活动是扩散性重放（Stella等人，2019年），通过朗之万采样（Besag，1994年），从清醒活动的时间序列分布中一次生成一个时间步。至关重要的是，这种针对重放的时间序列采样定义，不同于现代人工智能模型生成时间序列的方式——RNN 重放一次生成一个时间步，且不修改其他时间步，而人工智能模型会同时修改所有时间步；它也不同于计算神经科学中其他采样定义——重放从具有时间结构、类似于亲身经历的时间序列分布中采样，而其他类型的采样（Dong 等人，2022年，Aitchison 和 Lengyel，2016年，Hennequin 等人，2014年，Chen 等人，2025年）则使用外部输入生成时间结构无序的时间序列，并从中收集各个时间步用于蒙特卡洛计算（图2）。

随后，Levenstein 等人(2024年提出了一种通过脉冲频率适应（即负反馈）诱发探索的机制，使其发生在重放过程中。探索指的是重放呈现出多种行为序列（Davidson等人，2009年，Pfeiffer，2020年），并与神经空间中的长轨迹、对多个吸引子盆地的访问，以及清醒活动中未曾出现的转移相关。适应可以使吸引子失稳，并诱发重放活动中的突然转移（Itskov 等人，2011年，Dong 等人，2021年，Li 等人，2024年，Levenstein 等人，2024年），从而促进探索，并被认为在体内(Levenstein 等人，2019年)重放动力学中发挥关键作用。然而，现有关于噪声路径积分网络的理论并未考虑这些机制，也无法解释相较于清醒序列的重放序列时间压缩。

体内活动体内(Nádasdy 等人，1999年，Buzsáki，2015年，Michelmann 等人，2019年，Farooq 和 Dragoi，2019年）。据我们所知，任何经过训练的循环神经网络模型所产生的重放都未能体现时间压缩，因此迄今为止，我们对其机制的理解仅限于手工构建的脉冲网络中的短时程易化和相位进动（Leibold 等人，2008年，Jaramillo 和 Kempter，2017年）。

总的来说，尽管噪声路径积分是海马功能与重放的一种有前景的模型，但其理论未能解释生物重放的关键细节：探索和时间压缩。我们通过为自发神经活动（重放）建立一个规范性二阶朗之万采样框架，弥补了这一空白。我们表明，脉冲频率适应驱动空间探索，却会带来动力学阻力，使重放减慢。为了恢复更接近生物学实际的重放速度，我们通过树突输入微分引入动量，并将其与欠阻尼朗之万采样联系起来。利用含噪循环神经网络完成若干连续导航任务，我们证明，在自发循环神经网络动力学中同时引入适应和动量，可使重放轨迹访问多个吸引子盆地（探索），同时保持快速移动（时间压缩）。这些结果首次为基于放电率的预测网络中的时间压缩提供了形式化解释，并对海马体等记忆回路中的时间压缩和树突处理提出了可检验的预测。

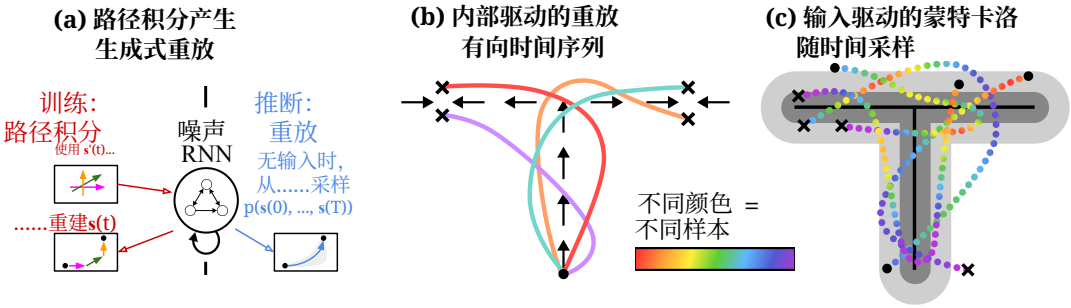

图2： **经过路径积分训练的含噪循环神经网络能够在内部生成与经历过的体验相似的重放轨迹。** **(a)** 带噪循环神经网络学会了 路径积分：它根据速度的（非线性）观测 $\boldsymbol{s}(t)$估计信号 $\boldsymbol{s}'(t)$。没有输入时，这个循环神经网络受噪声驱动的活动 $\boldsymbol{r}(t)$会逐个时间步地从清醒活动的时间序列分布中采样（“重放”）。**(b)**每条内部驱动的重放轨迹，都是从具有严格时间依赖关系的时间序列分布中抽取的单个样本：清醒轨迹会按特定顺序从某些位置开始和停止，而重放轨迹则再现这些时间结构。 **(c)** 我们将重放定义为采样，这与神经采样的其他定义不同；在那些定义中，随机的输入驱动轨迹用于收集样本，以便对不含时间依赖关系的静态分布进行蒙特卡罗推断（见第2.3节）。

2 背景：将重放视为朗之万采样

大脑多个区域的神经回路，尤其是海马体等记忆区域，会自发重放生物体活动或清醒时的活动。传统模型认为，重放是由手工设计的网络产生的；而我们关注的是近期模型，这些模型证明，重放是路径积分网络（如海马体）学习过程的副产物（Krishna等人，2024年）。在没有外部输入（例如感觉线索）的情况下，这些网络会经历朗之万动力学，从而对其训练过的清醒活动分布进行采样。本文将：(1) 定义朗之万动力学；(2) 概述路径积分噪声RNN如何隐式学习生成式朗之万动力学；(3) 阐明重放中的采样与计算神经科学其他领域中的采样之间的关键区别；(4) 说明一阶朗之万采样的局限性，这些局限性促使我们采用二阶动力学来解释神经回路中的重放。在第 A.5 节中，我们将讨论如何将非平稳的含噪循环神经网络活动（重放）刻画为朗之万采样时所涉及的确切细节。

2.1 经典朗之万动力学

在此，我们将不含（方程 1）和含（方程 2 和 3）动量的朗之万动力学定义为一种迭代过程，智能体可借此尝试从一个未知分布中采样 $p(x)$

仅使用其得分函数 $\nabla \log\, p(x)$。从 $p(x)$ 进行朗之万采样，就是通过其得分函数，沿着 $p(x)$ 的对数似然对迭代量 $x(t)$ 执行随机梯度上升：

$$\dot{x}(t) = \nabla \log p(x) + \sqrt{2}\eta(t), \tag{1}$$

其中 $\eta(t)$ 是高斯白噪声。虽然方程 1 描述的是 过阻尼 动力学，但也存在欠阻尼 （即具有动量）的朗之万动力学[1] （方程 2 或 3，参见 Pavliotis(2014年第6章），它们能更快地收敛到目标分布 $p(x)$，并且能更好地利用噪声梯度（Cheng等人，2018年）：

$$\ddot{x}(t) = \nabla \log p(x) - \gamma \dot{x}(t) + \sqrt{2\gamma}\eta(t), \quad \text{or} \tag{2}$$

$$\dot{x}(t) = v(t), \;\; \dot{v}(t) = \nabla \log p(x) - \gamma v(t) + \sqrt{2\gamma}\eta(t) \tag{3}$$

2.2 路径积分含噪循环神经网络中的重放

海马体等生物神经回路在清醒时能够在环境中导航，并在休息期间回忆导航经历。因此，我们概述Krishna等人(2024年的证明：在与生物学相关的导航任务（如路径积分）上训练的带噪循环神经网络可以充当生成模型，如图 2所示。

定义 2.1 和 2.2 表明，我们的循环神经网络具有非线性动力学和内部噪声，并且学会通过整合对 $\boldsymbol{s}(t)$ 的观测来估计信号 $\boldsymbol{s}'(t)$。这种能力称为路径积分，是导航以及根据运动线索估计位置的基础（Seelig和Jayaraman，2015年、Chrastil，2025年）。

定义 2.1。 一个 噪声循环神经网络（*RNN*）具有神经活动 $\boldsymbol{r}(t)$，它在每个时间步 t 通过一个可学习的非线性函数 $f(\cdot)$ 根据其先前的隐藏状态、输入信号$\boldsymbol{u}(t)$和白噪声 $\boldsymbol{\eta}(t)$演化；除非另有说明（见第 A.1.1 节）， $\boldsymbol{u}(t) = \boldsymbol{s}'(t)$ 或者是其非线性观测值：

$$\boldsymbol{r}(t + \Delta t) = f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)), \quad \sigma_r \boldsymbol{\eta}(t) \sim \mathcal{N}(0, \boldsymbol{I}\sigma_r^2 \Delta t) \tag{4}$$

定义 2.2。 一个路径积分目标 $\mathcal{L}(t)$ 对状态变量$\boldsymbol{s}(t)$ 与循环神经网络活动 $\boldsymbol{r}(t)$ 的可学习线性投影之间的差异进行惩罚：

$$\mathcal{L}(t) = \mathbb{E}_{\boldsymbol{\eta}} \|\boldsymbol{s}(t) - \boldsymbol{D}\boldsymbol{r}(t)\|_2 \tag{5}$$

假设 2.3 是指，最优循环神经网络动力学能够准确地进行路径积分，使得给定环境信号（例如位置）时网络活动的条件概率分布为高斯分布，且所有变异性均来自循环神经网络动力学中的内在噪声。假设2.4 是指，循环神经网络动力学可分解为加性成分。最后，假设2.6 是指，循环神经网络动力学在路径积分方面是贪心最优的——假设在每个先前的时间步都已做到尽可能好，网络在每个时间步都尽其所能。

假设 2.3。 最优的 $\boldsymbol{r}(t)$ 在满足 $p(\boldsymbol{r}^*(t))$ 关于 $\boldsymbol{D}^\dagger \boldsymbol{s}(t$ 呈正态分布的条件下，使 $\mathcal{L}(t)$ 最小):

$$p(\boldsymbol{r}^*(t)|\boldsymbol{s}(t)) \sim \mathcal{N}(\boldsymbol{D}^\dagger \boldsymbol{s}(t),\, \boldsymbol{I}\sigma_r^2 \Delta t) \tag{6}$$

假设 2.4。 带噪循环神经网络的动态可分解为两个函数[2] 和白噪声：

$$\Delta \boldsymbol{r}(t) = \boldsymbol{r}(t + \Delta t) - \boldsymbol{r}(t) \approx \Delta \boldsymbol{r}_1(t) + \Delta \boldsymbol{r}_2(t) + \sigma_r \boldsymbol{\eta}(t), \quad \sigma_r \boldsymbol{\eta}(t) \sim \mathcal{N}(0, \sigma_r^2 \Delta t) \tag{7}$$

引理 2.5。 根据假设 *2.4* 和 *2.3*， $\mathcal{L}(t + \Delta t)$ 的上界为 $\mathcal{L}(t + \Delta t)$:

$$\widehat{\mathcal{L}}(t + \Delta t) = \|\boldsymbol{s}'(t) - \boldsymbol{D}\Delta \boldsymbol{r}_2(t)\| + \mathbb{E}_{\boldsymbol{\eta}} \|\boldsymbol{D}\sigma_r \boldsymbol{\eta}(t)\|_2 + \|\boldsymbol{D}\|_F \mathbb{E}_{\boldsymbol{\eta}} \|\Delta \boldsymbol{r}_1(t) + \sigma_r \boldsymbol{\eta}(t - \Delta t)\|_2 \tag{8}$$

假设 2.6。 最优的 $\boldsymbol{r}(t)$ 在 t 时仅贪心地最小化 $\mathcal{L}$ （T 是最后一个时间步）：

$$\{\boldsymbol{r}^*(t)\}_{t=0}^T = \underset{\{\boldsymbol{r}(t)\}_{t=0}^T}{\arg\min} \sum_{t=0}^{T} \mathcal{L}(t + \Delta t)\Delta t \approx \left\{ \underset{\boldsymbol{r}(t)}{\arg\min}\, \widehat{\mathcal{L}}(t + \Delta t) \right\}_{t=0}^T \tag{9}$$

[1]离散化的欠阻尼朗之万动力学与哈密顿蒙特卡洛（Cheng 等人，2018年）相关。[2]每个函数 $\Delta \boldsymbol{r}_1(t)$ $\Delta \boldsymbol{r}_2(t)$ 都可以依赖 t以外的变量，但为简洁起见，这些变量被省略。

定理2.7 指出，路径积分循环神经网络会利用其学到的清醒活动得分函数来消除内在噪声，并利用输入来更新对环境信号的估计。定理2.8 表明，在没有输入的情况下（例如睡眠时），这些循环神经网络会从清醒活动的分布中进行朗之万采样，从而重放清醒活动。

定理 2.7。 在假设 2.6 下，最优更新 $\Delta \boldsymbol{r}^*(t) = \arg\min_{\Delta \boldsymbol{r}_2(t)} \mathcal{L}(t+\Delta t) + \arg\min_{\Delta \boldsymbol{r}_1(t)} \mathcal{L}(t+\Delta t) + \sigma_r \boldsymbol{\eta}(t)$ 遵循 $\boldsymbol{s}'(t)$，以及 $p(\boldsymbol{r}(t))$ 的得分函数：

$$\Delta \boldsymbol{r}^*(t) = \boldsymbol{D}^\dagger \boldsymbol{s}'(t)\Delta t + \sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) + \sigma_r \boldsymbol{\eta}(t) \tag{10}$$

定理 2.8。 在没有输入（静息状态）的情况下，已经训练为从 $\boldsymbol{s}'(t)$ 对 $\boldsymbol{s}(t)$ 进行路径积分的带噪循环神经网络，将沿着 $\boldsymbol{r}(t)$ 的得分函数执行梯度上升。如果 $p(\boldsymbol{r}(t))$ 和 $p(\boldsymbol{s}(t))$ 是平稳的，那么这种上升就是朗之万采样（公式 1）。如果将 $\sigma_r \boldsymbol{\eta}(t)$ 的方差按 2 倍缩放，则保证 $p(\boldsymbol{r}(t))$ 收敛到稳态分布 $p(\boldsymbol{r}) = p(\boldsymbol{D}^\dagger \boldsymbol{s})$。

$$\textit{If } \Delta \boldsymbol{r}(t) = \sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) + \sqrt{2}\sigma_r \boldsymbol{\eta}(t), \textit{ then } \lim_{t\to\infty} p(\boldsymbol{r}(t)) = p(\boldsymbol{D}^\dagger \boldsymbol{s}) \tag{11}$$

2.3 重放采样与其他类型的采样

我们研究路径积分噪声RNN，它们在没有外部输入的静息状态下重放清醒活动片段。我们的含噪循环神经网络以因果、实时的方式生成轨迹，每次生成一个时间步；这与现代人工智能模型不同，后者可以同时生成时间序列中的所有时间步。

重要的是，我们将采样定义为每次生成一个时间步的时间序列，不应与计算神经科学其他研究中用于解释大脑如何执行贝叶斯推断等概率计算的朗之万采样相混淆（Dong 等，2022；Aitchison 和 Lengyel，2016；Hennequin 等，2014；Chen 等，2025）。我们研究的是时间序列分布：重放 $p(\{\boldsymbol{r}(t)\}_{t=0}^T)$、清醒活动 $p(\{\boldsymbol{s}(t)\}_{t=0}^T)$（$T$ 是最后一个时间步），以及二者之间的比较——每条时间序列都是从具有严格时间结构和非平稳性的时间序列分布中抽取的单个样本（$p(\boldsymbol{r}(t_1)) \neq p(\boldsymbol{r}(t_2)) \; \forall t_1 \neq t_2$）。与此同时，推断网络使用朗之万采样来生成活动序列 $\{\boldsymbol{r}(t)\}_{t=0}^T$，其中每个时间步都是从静态、平稳分布中单独抽取的样本：$\boldsymbol{r}(t) \sim p(\boldsymbol{s}) \; \forall t$。在推断网络中，时间只是收集样本的手段：只要时间足够，就能获得足够多的样本，通过蒙特卡洛计算静态期望、协方差或后验分布。重放网络生成清醒活动片段，而推断网络计算静态统计量。图2进一步阐释了这两种采样概念之间的关键区别。

尽管如此，人们可能会认为，现有用于加速推断网络中蒙特卡洛采样的技术（Dong 等人，2022年，Aitchison 和 Lengyel，2016年，Hennequin 等人，2014年）也应能轻松应用于重放采样网络，但适应现象清楚地表明事实并非如此。Dong 等人（2022年）表明，适应能够加速对静态恒定输入进行的蒙特卡洛采样——他们使用外部输入锚定网络的目标状态，使负反馈（适应）能够迅速破坏初始状态的稳定性，同时在轨迹到达目标后又不会将其排斥。相比之下，我们的网络在没有输入的情况下内部生成重放。由于不存在外部驱动来促使网络到达目标，负反馈会抵消分数驱动的、朝向学习到的吸引子的收敛，从而导致我们随后观察到的减速。适应能够加速输入钳制推理，但不一定能加速内部驱动的重放生成。

2.4 脉冲频率适应与二阶动力学

脉冲频率适应。生物神经元可以通过负反馈缓解持续活动或低频活动，即脉冲频率适应（Benda，2021年；Gutkin 和 Zeldenrust，2014年）。这种反馈对于描述体内海马活动和重放至关重要（Itskov 等人，2011年；Levenstein 等人，2019年）；在重放的计算模型中，它还能防止神经活动陷入吸引子盆地，从而促进较长的重放轨迹（探索）（Dong 等人，2021年；Li 等人，2024年；Levenstein等（2024））。与Levenstein等（2024）类似，我们在基于发放率的神经网络中，将脉冲频率适应（通常针对脉冲神经元定义）实现为添加到循环神经网络活动中的负反馈 $\boldsymbol{c}(t)$（公式 4），并在训练[3]后进行：

$$\boldsymbol{r}(t+\Delta t) = f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) - \boldsymbol{c}(t), \quad \Delta \boldsymbol{c}(t) = \frac{1}{\tau_a}(-\boldsymbol{c}(t) + b_a \boldsymbol{r}(t)) \tag{12}$$

[3]移动项的减法 平均 $\boldsymbol{c}(t)$ 也自然地源自贪心最小化 的 $\mathcal{L}(t) + \frac{1}{2}\|\boldsymbol{c}(t)\|_2^2$。

静息状态（睡眠）期间适应增强的生物学基础。在我们的主要实验中，我们训练循环神经网络在没有适应（$b_a = 0$）的情况下对输入进行路径积分（处于活跃、清醒状态），随后在推断期间，模拟已训练的循环神经网络在没有输入且具有适应（$b_a \geq 0$）的情况下的表现（处于静息、睡眠状态）。我们调节适应强度的依据来自生物学：神经调质去甲肾上腺素、乙酰胆碱（ACh）以及在一定程度上的血清素，会抑制适应（McCormick, 1992, Cui and Strowbridge, 2019, Wotton et al., 2020, O'Donnell et al., 2012）；与清醒状态相比，睡眠期间这些神经调质的含量较低，而这种差异与记忆有关（Gais and Born, 2004, Inayat et al., 2020, Becchetti et al., 2016, Watson et al., 2010, Portas et al., 2000, Bjorvatn et al., 2002）。换言之，大脑本就会利用这些神经调质，在睡眠期间增强适应强度（b_a），并在清醒期间降低适应强度。

3 用于神经重放的二阶朗之万采样

基于上一节的内容，我们在此假设，训练后的循环神经网络已准确估计得分函数，并在静息（由内部驱动）活动期间遵循该函数，通过朗之万动力学生成重放。我们在此探讨，调节此类循环神经网络的动力学会如何影响或偏置重放分布。我们发现，适应（负反馈）会导致一种非理想形式的二阶朗之万采样，因此提出一种互补的替代方案：利用动量，并可通过输入微分在神经元中实现。

3.1 适应作为欠阻尼朗之万动力学

首先，我们考察将适应（负反馈，如第 2.4 节所定义）引入训练后的路径积分循环神经网络的动力学后，会如何影响其静息（由内部驱动）动力学。我们的分析从理论上解释了为何适应会引发探索和迟缓，而我们稍后会在训练后的循环神经网络重放中观察到这两种现象。定理3.1 定义了训练后引入适应的训练后循环神经网络的一阶动力学；定理 3.2 则将其改写为二阶动力系统（推导见第 B.1 节），以进一步阐明定理 3.3 中的见解。

命题 3.1。 在静息状态下（公式 (12)），向经过训练的路径积分循环神经网络添加适应（公式 (11)[4]）会引入带负反馈的朗之万采样：

$$\Delta \boldsymbol{r}(t) = \sigma_r^2 \Delta t \, \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) + \sigma_r \boldsymbol{\eta}(t) - \boldsymbol{c}(t), \quad \Delta \boldsymbol{c}(t) = \frac{1}{\tau_a}(-\boldsymbol{c}(t) + b_a \boldsymbol{r}(t)) \tag{13}$$

定理 3.2。 在静息状态下向经过训练的路径积分循环神经网络添加适应（定理 *3.1*）会产生以下二阶随机动力学（证明见第 *B.1*节）：

$$\begin{aligned} \boldsymbol{r}''(t) = \left(\sigma_r^2 \Delta t \frac{d^2}{d\boldsymbol{r}(t)^2} \log p(\boldsymbol{r}(t)) - \frac{1}{\tau_a} \boldsymbol{I}\right) \boldsymbol{r}'(t) - \frac{1}{\tau_a} \sigma_r \boldsymbol{\eta}(t) + \sigma_r \boldsymbol{\eta}'(t) \\ + \frac{1}{\tau_a} \Big(\underbrace{\sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) - b_a \boldsymbol{r}(t)}_{\text{score conflicts with negative feedback}} \Big) \end{aligned} \tag{14}$$

注释 3.3。将公式 (14) 解读为欠阻尼（动量）朗之万动力学（公式 (2)），可知驱动动力学的有效得分梯度 $\sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t))$ 会被 $-b_a \boldsymbol{r}(t)$ 平移。这种负反馈会惩罚持续活动，使活动远离 $p(\boldsymbol{r}(t))$ 中概率较高的吸引子状态。例如，高斯 $\boldsymbol{r}(t) \sim \mathcal{N}(\boldsymbol{\mu}(t), \boldsymbol{\Sigma}(t))$[5] 的得分为 $\nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) = \boldsymbol{\Sigma}(t)^{-1}(\boldsymbol{\mu}(t) - \boldsymbol{r}(t))$，从中减去 $b_a \boldsymbol{r}(t)$ 会推动朗之万动力学远离其目标 $\boldsymbol{\mu}(t)$，减缓向目标的收敛。如果 $\boldsymbol{\mu}(t) = \boldsymbol{0}$，得分与 $-\boldsymbol{r}(t)$ 可能方向一致，从而加快收敛（取决于 $\boldsymbol{\Sigma}(t)^{-1}$）。

我们的定理 3.3 预测，适应通常会使重放活动远离吸引子状态。这应能防止活动困在吸引子盆地中，从而促进重放过程中的探索，但也会降低重放到达新吸引子盆地的速度。我们随后通过实证观察到，适应会在重放中引发这两种现象——探索和缓慢，这也与我们在第 2.3 节中的推理相符：自发活动中的适应通常会减慢而非加速重放。这引出了一个问题：如何设计一种生物学上合理的网络，通过欠阻尼朗之万动力学加速重放？

[4]我们研究的是非平稳系统，因此省略了 $\sqrt{2}$ 平稳收敛所需的白噪声缩放因子。[5]有关带噪循环神经网络在训练中对 $\boldsymbol{\mu}(t)$ $\boldsymbol{\Sigma}(t)$ 进行路径积分时应采用何种设置的详细信息，其中信号 $\boldsymbol{s}(t)$ 服从任意高斯过程，请参见第C.2节。

3.2 通过动量显式实现欠阻尼朗之万动力学

为了在重放中明确执行欠阻尼朗之万采样，我们提出一种基于方程 3的替代方法。该方法使用动量[6]：一个速度项 $\boldsymbol{v}(t)$ 会在摩擦项 $\boldsymbol{r}(t)$ 低于 $\lambda_v \in [0,1]$ 时，累积 1[7]的先前取值。这里我们对其加以定义，并说明如何通过树突输入微分，以生物学上合理的方式实现它。

定义 3.4。 我们通过动量 显式地对重放动力学进行欠阻尼，其受摩擦 λ_v控制。当 $\lambda_v = 1$时，动力学恢复为默认状态（过阻尼，方程（4））：

$$\Delta \boldsymbol{r}(t) = \boldsymbol{v}(t+\Delta t) = (1-\lambda_v)\boldsymbol{v}(t) + \underbrace{f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) - \boldsymbol{r}(t)}_{\Delta \boldsymbol{r}(t) \text{ from overdamped case } (\lambda_v = 1)} \tag{15}$$

尽管上述动力学有效地实现了动量，但神经回路如何实现它们尚不清楚。通过仔细考察 $\boldsymbol{v}$（$t+\Delta t$）与 $\boldsymbol{r}$（$t+\Delta t$）$=\boldsymbol{r}$（t）$+\boldsymbol{v}$（$t+\Delta t$）之间的关系，我们在第B.2节证明，神经回路可以通过用近期导数增强输入（突触前）活动来实现动量。

定理 3.5。 定理*3.4* 中的动量动力学可通过输入微分以生物学上合理的方式近似实现（证明见第*B.2*节）：

$$\boldsymbol{r}(t+\Delta t) = f(\underbrace{\boldsymbol{r}(t) + (1-\lambda_v)(\boldsymbol{r}(t-\Delta t) - \boldsymbol{r}(t-2\Delta t))}_{\text{effective incoming } \boldsymbol{r}(t)}, \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) \tag{16}$$

输入微分的生物学合理性。 我们通过回答三个关键问题来评估定理3.5的合理性：

1. 单个神经元如何对其输入进行微分？单个神经元计算其输入近期导数的一种直接方法是利用树突延迟：两条不同的树突在不同的时间延迟下传递相同的输入信号副本，并在两条分支汇合处将其相减。另一种对输入进行微分的方法是通过高通滤波。

2. 大脑如何调节输入微分的增益（强度）？大脑中的神经调质，例如乙酰胆碱（ACh），可以调节神经元的特性。我们提出，它们或许也能调节特定树突分支的强度（即计算输入近期导数的分支），尤其是当这些分支具有可被神经调质激活的特殊受体时。例如，ACh 可调节的 SK（小电导钙激活钾通道）通道和毒蕈碱型 M1 受体（Giessel和Sabatini，2010，Buchanan等，2010）可能集中分布在计算导数的树突分支中。

3. 增益调制如何对应于活动/静息（即觉醒/睡眠）周期？如第2.4节所述，大脑在觉醒/睡眠周期中本就会改变 ACh 等神经调质的水平；这正是我们有理由在 b_a 静息活动期间相比清醒活动期间提高适应强度的原因。

总的来说，我们的目的并非断言大脑必须实现动量，且必须通过输入微分来实现，而是将其作为未来实验研究的一个方向。事实上，我们对大脑如何对自发活动进行时间压缩知之甚少。

4 数值结果

现在，我们考察适应和生物学上合理的动量（方程 12 和16）如何影响经训练以进行路径积分的模拟循环神经网络中的重放分布。我们首先发现，适应通常会减慢重放，而动量会加快重放。随后，我们确认适应会引发探索，并着重说明动量不会阻止探索，反而会通过增加重放路径的行进距离来补充探索。

[6]其他循环神经网络也尝试过动量（Nguyen 等人，2020年），但并未将其用于欠阻尼朗之万采样或计算神经科学。[7]关于 λ_v 与 γ 项的比较，参见方程3 和2的第A.5节。

实验。 我们设置了五项任务，在这些任务上训练循环神经网络，随后考察训练后引入采样机制（适应强度 b_a 和摩擦 λ_v）如何影响重放（静息活动）统计量。为了促进连贯的重放，我们训练循环神经网络进行路径积分，并 $s'(t)$周期性地隐藏（掩码训练，见第 A.1.1节）。我们通过初始化循环神经网络的隐藏状态（见第A.2节），随后在没有外部输入但仍存在内部噪声的情况下，收集多个展开时间步内的活动，从而获得静息活动。根据第 A.1节中的论证，这里我们使用 ReLU 循环神经网络；但在第 F 节和 E 节中，我们确认，我们的大部分结果同样适用于双曲正切循环神经网络的重放[8] 以及带有非合理动量的循环神经网络。在第 A 节中，我们定义所测量的重放统计量，并介绍循环神经网络的实现细节。第 D节还报告了重放统计量的其他测量结果。我们的五项任务如下：

1. 一维奥恩斯坦-乌伦贝克（OU）过程（图1）：关于方程（8）的最优路径积分循环神经网络具有闭式解（方程（47））， 因此我们用该解代替训练后的循环神经网络。2. 二维T迷宫和三角形（图3）：我们将每个行进方向（T形迷宫：两个；三角形：六个）模拟为二维OU过程，并训练一个ReLU循环神经网络，对来自各个方向的二维路径进行积分。3. 大鼠轨迹（图3）：与Krishna等人(2024)类似，我们模拟512个位置细胞，用它们编码来自RatInABox（George等人，2024年）的二维有向（偏置）或无向（无偏置）随机游走；训练ReLU循环神经网络对位置细胞活动进行路径积分，并在二维空间中解码活动。我们使用计算机模拟的大鼠活动来模拟生物数据，因为体内重放事件稀少，且难以从背景噪声活动中提取，因此不在本研究的讨论范围内。

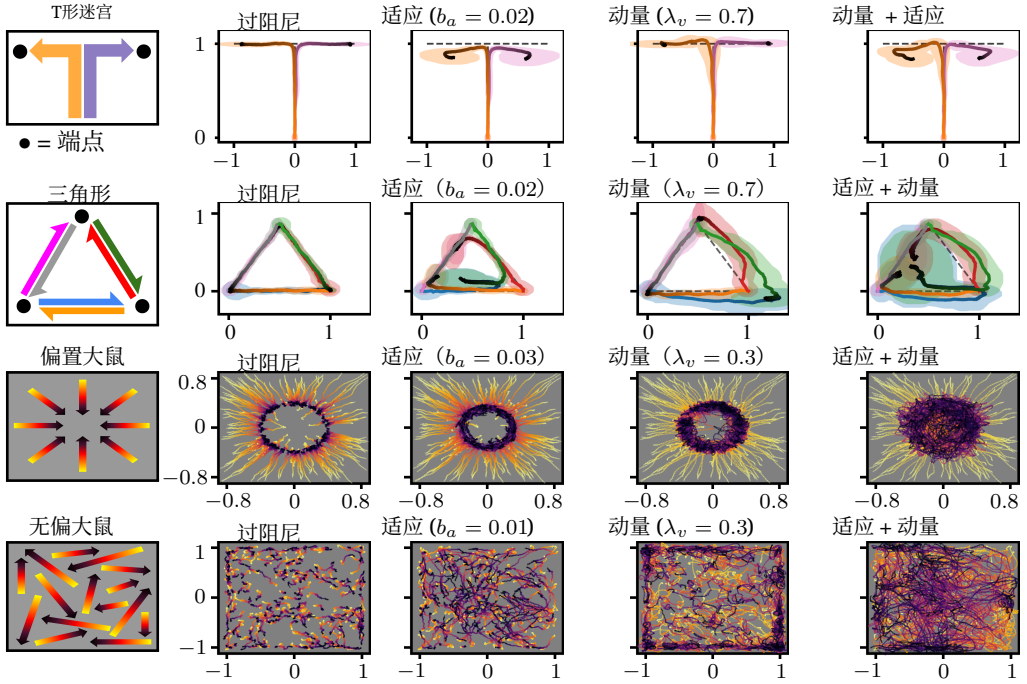

图3： **动量与适应相辅相成，使轨迹能够持续穿越吸引子盆地。**从在四种空间导航任务之一上训练的含噪循环神经网络中解码得到的二维回放轨迹：T形迷宫、三角形、偏置大鼠随机游走和无偏大鼠随机游走。回放轨迹随时间推移按时间顺序逐渐变深。对于T形迷宫和三角形任务，图中显示了每个轨迹方向的均值（实线）和标准差（浅色）；对于大鼠任务，则显示了许多条独立轨迹。标准动力学（过阻尼， $\lambda_v = 1$， $b_a = 0$）能够重建清醒轨迹，但会停滞在终端吸引子处。如图1所示，适应（$b_a > 0$）会使吸引子失稳，从而促进探索，但也会导致轨迹未能抵达目标；而通过输入微分实现的生物学上合理的动量（$\lambda_v < 1$）则会加速轨迹向目标吸引子收敛。两者结合时，动量会抵消适应诱导的阻力，使回放轨迹比单独采用任一机制时都能更快地穿越其环境。

[8]在双曲正切循环神经网络中，我们观察到动量诱导的压缩，但出于某种原因并非适应 ntum-induced compression，但出于某种原因并未适应 离子诱导减速。

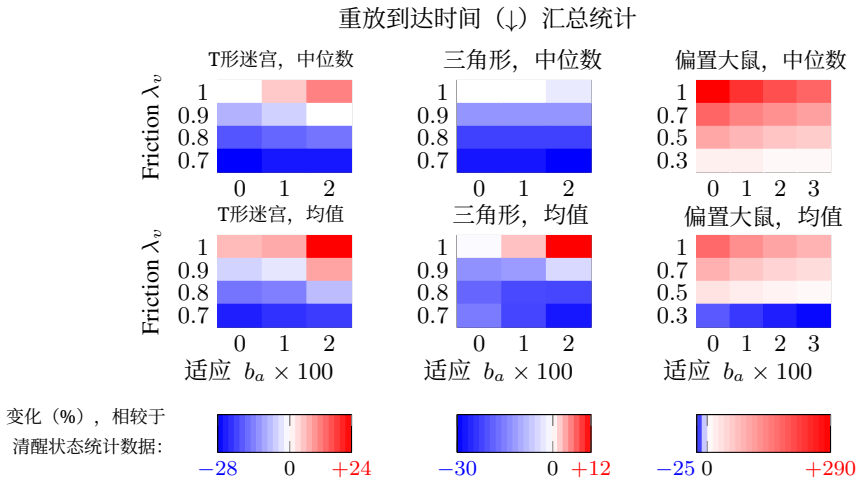

图4： **动量在时间上压缩重放。** 我们计算重放轨迹到达终点所需的时间，并将其与清醒轨迹进行比较（蓝色 = 更快，红色 = 更慢）。通过输入微分实现的生物学上合理的动量（$\lambda_v < 1$）不仅缩短了到达时间，还使其短于清醒路径的到达时间，从而在时间上压缩清醒活动，最多压缩30%。我们未纳入无偏大鼠轨迹，因为它们没有明确的终点；但我们在第D 节确认，动量会加快这些轨迹。对于具有 tanh 非线性的循环神经网络（图14）和非合理动量（图10），也呈现相同趋势。

适应会减慢重放，动量则会加速重放。我们最初在图1 和3中观察到，适应会扰乱重放轨迹所指向的吸引子盆地，从而减慢重放；而通过输入微分产生的动量会加快重放轨迹到达这些目标的速度。在图4 中，我们通过测量重放轨迹（无偏置大鼠任务除外，因为该任务没有明确定义的端点目标）进入理想终点一定半径范围内所需的时间，对这一观察结果进行了量化（详情见第A 节）。在T形迷宫和三角形任务中，适应会减慢重放向目标的收敛，而动量会加快收敛。一个显著例外是有偏置大鼠任务，其中适应也会加快收敛：这是因为所有偏置路径分布都指向同一个中心，类似于具有稳态均值$\mu = 0$的奥恩斯坦-乌伦贝克过程——正如定理3.3所解释的那样，负反馈可以加快向均值0收敛。尽管如此，在这三种任务中，动量都能使重放轨迹比清醒轨迹更快地到达目标，将后者的时间压缩时间压缩最多达30%。

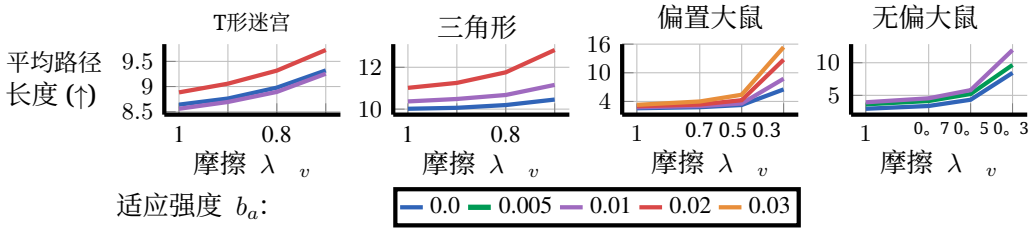

图5： **动量和适应都通过回放路径长度增加探索。** 在这里，我们测量了四个经过训练的循环神经网络任务中，回放轨迹里的重放路径的累积路径长度，以量化动量和适应如何影响重放中的探索。在所有任务中，我们都观察到，这两种机制都会增加重放路径长度。在具有 tanh 非线性的循环神经网络（图 15）和采用非合理动量的循环神经网络（图 11）中，也呈现出相同的趋势。

动量补充了适应带来的探索。我们已经证明，就速度而言，适应和动量似乎相互抵消。然而，体内重放同时具备探索和时间压缩特性。众所周知，适应会通过使吸引子失稳来诱导探索（Levenstein 等人，2024年），但动量是否也会抵消适应的这一方面，尚不清楚。进一步观察图3，我们发现这两种机制似乎会

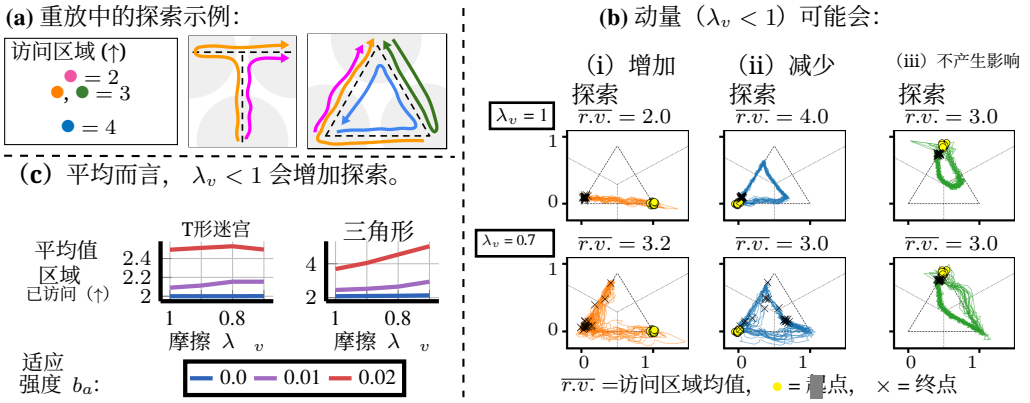

图6：**平均而言，动量通过增加访问区域来促进探索**。探索的另一个相关方面是清醒活动中不存在的转移。**(a)** 在这里，我们通过访问区域($r.v.$)来量化这些转移。每个区域都是一个区域，其中的每个点距离同一个端点最近。在我们的T形迷宫和三角形任务中，清醒路径从起点前往终点并停留在那里，总共访问两个区域（起始区域和终止区域）。与此同时，适应（$b_a > 0$）可能使重放路径访问多个终点，或再次访问先前到达过的终点（$r.v. \geq 3$）。**(b)**有了适应和动量（$\lambda_v < 1$），重放路径可能访问（i）更多、（ii）更少或（iii）相同数量的区域。**(c)** 然而，平均而言，动量会增加访问区域的数量，从而促进探索。关于双曲正切循环神经网络和非合理动量的版本，请参见图16 和12 。

相互补充，共同促进探索。为了量化这两种机制所表现出的协同作用，我们设计了若干探索指标（详见第A 节），其中最显著的两个指标是路径长度和访问区域，用于衡量重放轨迹。前者是重放路径的累积长度（不要与重放轨迹相对于其原点的平均位移混淆；我们也在图8中对后者进行了测量），而后者则是在T形迷宫和三角形任务中更具语义性、针对具体任务的探索度量。在图5 和6 中，我们确认适应和动量都会提高这些探索指标；在第D节中，我们还通过包括方差（图9）和平均位移（图8）在内的指标进一步予以确认。虽然我们的多个探索指标总体上相互一致，但需要注意的是，在T形迷宫和三角形任务中（而在另外两个任务中并非如此），适应会降低平均位移（图8）。

5 结论与未来工作

通过重新审视路径积分噪声RNN中的重放，并将其视为朗之万采样，我们证实一阶动力学不足以解释生物重放的速度与灵活性。我们通过分析和实证表明，脉冲频率适应是一种非理想的二阶采样形式：它虽然能成功削弱局部吸引子的稳定性以促进探索，却也同时引入动力学阻力，减慢重放生成。为解决探索与速度的权衡，我们通过树突输入微分引入了一种生物学上合理的动量机制。两者结合后，动量抵消了适应带来的减速，在保持甚至增强多吸引子空间探索的同时，实现了重放序列的时间压缩。

我们的发现表明，海马体等网络中的时间压缩需要二阶电路机制。尽管序列预测网络的组成部分已被映射到特定的海马亚区（Chen 等人，2024年），未来的研究可以通过纳入生物学约束（例如通过明确引入中间神经元群体来实现 Dale 原理），将我们的朗之万框架扩展至基于脉冲的预测学习模型（Asabuki 和 Fukai，2025年，Saponati 和 Vinck，2023，Bono 等人，2023），或在更复杂的环境中验证我们的发现，从而进一步完善所提出的机制（Levenstein 等人，2024年，Wood 等人，2018年）。

5.1 局限性与可检验预测

我们的理论框架包含若干必要的抽象处理。首先，为了实现精确的分数估计，我们假设生物循环网络中的内在噪声是高斯各向同性白噪声。生物噪声很可能具有复杂的时空相关性，这可能改变采样动力学。其次，将模拟重放直接与体内记录进行比较仍然颇具挑战，因为生物重放较为稀疏，间歇性地夹杂在背景噪声中，且时间压缩程度各不相同。由于对于体内观察中的时间压缩量级尚无共识，比较我们的时间压缩结果与这些观察也颇具挑战；根据建模假设不同，时间压缩率可能从 1–2 倍（体内）到 8–20 倍（Denovellis 等人，2021年）不等（Lee 和 Wilson，2002年，Jeunehomme 和 D'Argembeau，2019年，Buch 等人，2021年）。

尽管如此，我们的框架仍为实验神经科学提出了明确且可检验的预测。由于乙酰胆碱（ACh）等神经调质既能调节脉冲频率适应，也能调节特定的树突受体（例如毒蕈碱型 M1 和 SK通道），我们提出，睡眠期间这些神经调质水平下降，不仅会解除对适应的抑制，也可能促进树突输入微分。我们预测，在静息状态下有针对性地通过药理学增强这些特定神经调质，会显著削弱重放事件的时间压缩。我们还预测，二阶修饰项（动量、适应）将改善对重放的扩散动力学建模（Stella等人，2019年，Wu和Wei，2025年）。

6 致谢

本研究得到美国国家科学基金会（NSF）资助，项目编号为 CCF-1911094 和 IIS-1730574；美国海军研究办公室（ONR）资助，项目编号为 N00014-23-1-2714、N00014-24-1-2225，以及多学科大学研究计划（MURI）项目 N00014-20-1-2787；美国空军科学研究办公室（AFOSR）资助，项目编号为 FA9550-22-1-0060；美国能源部（DOE）资助，项目编号为 DE-SC0020345；美国内政部（DOI）资助，项目编号为 140D0423C0076；以及范内瓦·布什教职奖学金（美国海军研究办公室项目 N00014-18-1-2047）。

代码和数据可用性。 用于复现我们结果的全部代码见
https://github.com/Josuelmet/rnn-replay-second-order/tree/main

参考文献

David Tingley 和 Adrien Peyrache。关于再激活和重放分析的方法。《皇家学会哲学汇刊B辑：生物科学》, 375(1799):20190231, 2020年。

David R Euston、Masami Tatsuno 和 Bruce L McNaughton。睡眠期间前额叶皮层中近期记忆序列的快进回放。《科学》, 318(5853):1147–1150, 2007年。

Adrien Peyrache、Mehdi Khamassi、Karim Benchenane、Sidney I Wiener 和 Francesco P Battaglia。睡眠期间前额叶皮层中规则学习相关神经模式的重放。《自然·神经科学》, 12(7):919–926, 2009年。国际标准连续出版物号 1546-1726。

Tal Kenet、Dmitri Bibitchkov、Misha Tsodyks、Amiram Grinvald 和 Amos Arieli。自发涌现的视觉属性皮层表征。《自然》, 425(6961):954–956, 2003年。国际标准连续出版物号 1476-4687。

Shengjin Xu、Wanchen Jiang、Mu-ming Poo 和 Yang Dan。视觉皮层神经元集群中的活动回忆。《自然·神经科学》， 15(3):449–455，2012年。国际标准连续出版物号 1546-1726。

Kari L Hoffman 和 Bruce L McNaughton。灵长类新皮层中分布式记忆痕迹的协调性再激活。《科学》， 297(5589):2070–2073，2002年。

Richard J Gardner、Erik Hermansen、Marius Pachitariu、Yoram Burak、Nils A Baas、Benjamin A Dunn、May-Britt Moser 和 Edvard I Moser。网格细胞群体活动的环面拓扑。《自然》， 602(7895):123–128, 2022年。国际标准连续出版物号 1476-4687。

Adrien Peyrache、Marie M Lacroix、Peter C Petersen 和 György Buzsáki。头朝向感的内部组织机制。《自然·神经科学》， 18(4):569–575，2015年。国际标准连续出版物号 1546-1726。

György Buzsáki。海马尖波：其原点与意义。《脑研究》，398(2)，1986年。William E Skaggs 和 Bruce L McNaughton。空间经验之后睡眠期间大鼠海马中神经元放电序列的重放。《科学》，271(5257)，1996年。Zoltán Nádasdy、Hajime Hirase、András Czurkó、Jozsef Csicsvari 和 György Buzsáki。海马中重复出现的尖峰序列的重放与时间压缩。《神经科学杂志》，19 (21):9497–9507，1999年。国际标准连续出版物号 0270-6474。Albert K Lee 和 Matthew A Wilson。慢波睡眠期间海马对序列经验的记忆。《神经元》，36(6):1183–1194，2002年。国际标准连续出版物号 0896-6273。David J Foster。重放走向成熟。《神经科学年度评论》，40:581–602，2017年。国际标准连续出版物号 1545-4126。John O'Keefe 和 Lynn Nadel。海马体作为认知地图。牛津大学出版社，1978年。Neil Burgess、Michael Recce 和 John O'Keefe。海马功能模型。《神经网络》，7(6-7)，1994年。Bruce L McNaughton、Carol A Barnes、Jason L Gerrard、Katalin Gothard、Min W Jung、James J Knierim、H Kudrimoti、Y Qin、WE Skaggs、M Suster 等。解读海马多元编码：海马体作为路径积分系统。《实验生物学杂志》，199(1)，1996年。Brad E Pfeiffer 和 David J Foster。海马位置细胞序列描绘通往记忆目标的未来路径。《自然》，497(7447)，2013年。Kevin J Miller、Matthew M Botvinick 和 Carlos D Brody。背侧海马有助于基于模型的规划。《自然·神经科学》，20(9)，2017年。György Buzsáki。记忆痕迹形成的两阶段模型："嘈杂"脑状态的作用。《神经科学》，31(3)，1989年。

György Buzsáki。海马体尖锐波-涟漪：情景记忆与规划的认知生物标志物。*Hippocampus*，25(10)，2015年。

Giulio Tononi 和 Chiara Cirelli。睡眠与可塑性的代价：从突触和细胞稳态到记忆巩固与整合。《神经元》，81(1)，2014年。H Freyja Ólafsdóttir、Caswell Barry、Aman B Saleem、Demis Hassabis 和 Hugo J Spiers。海马位置细胞通过未探索空间构建奖励相关序列。*eLife*，4:e06063，2015年。国际标准连续出版物号 2050-084X。H Freyja Ólafsdóttir、Daniel Bush 和 Caswell Barry。海马重放在记忆与规划中的作用。《当代生物学》，28(1):R37–R50，2018年。国际标准连续出版物号 0960-9822。Nicolas Deperrois、Mihai A Petrovici、Walter Senn 和 Jakob Jordan。通过受扰动和对抗性梦境学习皮层表征。*eLife*，11:e76384，2022年。国际标准连续出版物号 2050-084X。Bin Shen 和 Bruce L McNaughton。模拟睡眠期间特定经验海马细胞集群的自发再激活。海马体，6(6):685–692，1996年。Aaron D Milstein、Sarah Tran、Grace Ng 和 Ivan Soltesz。循环神经元网络中的离线记忆重放源于对在线动力学的约束。《生理学杂志》，601 (15):3241–3264，2023。John J Hopfield。心理探索的神经动力学。《美国国家科学院院刊》，107(4):1648–1653，2010年。Ashok Litwin-Kumar 和 Brent Doiron。通过突触可塑性形成并维持神经元集群。《自然-通讯》，5(1):5319，2014年。国际标准连续出版物号 2041-1723。Panagiota Theodoni、Bernat Rovira、Yingxue Wang 和 Alex Roxin。θ节律调制驱动位置细胞网络模型中重放所依赖连接模式的形成。*eLife*，7:e37388，2018年。国际标准连续出版物号 2050-084X。

Tatsuya Haga 和 Tomoki Fukai。通过反向重放学习目标导向序列的循环网络模型。*eLife*，7:e34171，2018年。国际标准连续出版物号 2050-084X。Toshitake Asabuki 和 Tomoki Fukai。预测性学习规则生成对概率性感觉经验的类皮层重放。*eLife*，13，2025年。Tianhao Chu、Zilong Ji、Junfeng Zuo、Yuanyuan Mi、Wen-hao Zhang、Tiejun Huang、Daniel Bush、Neil Burgess 和 Si Wu。放电率适应使位置细胞能够产生θ扫掠、相位进动和进程。*bioRxiv*，2024年。Amir H Azizi、Laurenz Wiskott 和 Sen Cheng。海马体预演的计算模型。《计算神经科学前沿》，7:161，2013年。Vladimir Itskov、Carina Curto、Eva Pastalkova 和 György Buzsáki。由尖峰阈值适应产生的细胞组装序列能够追踪海马体中的时间。《神经科学杂志》，31(8)，2011年。Xingsi Dong、Tianhao Chu、Tiejun Huang、Zilong Ji 和 Si Wu。噪声适应在吸引子神经网络中产生莱维飞行。神经信息处理系统进展（*NeurIPS*），34，2021年。Yujun Li、Tianhao Chu 和 Si Wu。自适应连续吸引子神经网络的动力学。*arXiv* 预印本 *arXiv:2410.06517*，2024年。Louis Kang 和 Michael R DeWeese。网格细胞吸引子网络中的波前重放与隆起振荡 θ序列。*eLife*，8:e46351，2019年。国际标准连续出版物号 2050-084X。Nanda H Krishna、Colin Bredenberg、Daniel Levenstein、Blake Aaron Richards 和 Guillaume Lajoie。循环神经网络中离线再激活的充分条件。载于国际学习表征会议（*ICLR*），2024年。Daniel Levenstein、Aleksei Efremov、Roy Henha Eyono、Adrien Peyrache 和 Blake Richards。序列预测性学习是海马表征与重放的统一理论。*bioRxiv*，2024年。William B Levy。海马功能的计算方法。载于《学习与动机心理学》，第23卷。爱思唯尔，1989年。Kimberly L Stachenfeld、Matthew M Botvinick 和 Samuel J Gershman。海马体作为预测地图。《自然·神经科学》，20(11)，2017年。Lila Davachi 和 Sarah DuBrow。海马体如何保留顺序：预测与情境的作用。《认知科学趋势》，19(2)，2015年。Fraser Aitken 和 Peter Kok。在预测性联结习得过程中，海马表征从误差转向预测。《自然-通讯》，13(1)，2022年。Adam MP Miller、Alex D Jacob、Adam I Ramsaran、Mitchell L De Snoo、Sheena A Josselyn 和 Paul W Frankland。海马体中预测模型的涌现。《神经元》，111(12)，2023。Kenneth Kay、Jason E Chung、Marielena Sosa、Jonathan S Schor、Mattias P Karlsson、Margaret C Larkin、Daniel F Liu 和 Loren M Frank。海马体中对可能未来的表征以亚秒级间隔持续循环切换。《细胞》，180(3)，2020年。Adam Johnson 和 A David Redish。在决策点，CA3区神经元集群会短暂编码动物前方的路径。《神经科学杂志》，27(45)，2007年。Alexei Samsonovich 和 Bruce L McNaughton。连续吸引子神经网络模型中的路径积分与认知制图。《神经科学杂志》，17(15)，1997年。Francesco P Battaglia 和 Alessandro Treves。存储多种空间表征的吸引子神经网络：海马位置场模型。《物理评论*E*》，58(6)，1998年。

Stefano Recanatesi、Matthew Farrell、Guillaume Lajoie、Sophie Deneve、Mattia Rigotti 和 Eric Shea-Brown。《预测性学习作为提取低维潜在空间表征的网络机制》。《自然-通讯》，12(1)，2021年。Yusi Chen、Huanqiu Zhang、Mia Cameron 和 Terrence Sejnowski。《海马结构中的预测性序列学习》。《神经元》，112(15)，2024年。Christopher J Cueva 和 Xue-Xin Wei。《通过训练循环神经网络执行空间定位，涌现出网格状表征》。载于国际学习表征会议，2018年。Ben Sorscher、Gabriel Mel、Surya Ganguli 和 Samuel Ocko。《通过模式形成视角统一阐释网格细胞的起源理论》。载于 Hanna M Wallach、Hugo Larochelle、Alina Beygelzimer、Florence d'Alché Buc、Emily B Fox 和 Roman Garnett 编，《神经信息处理系统进展》，第32卷。Curran Associates, Inc.，2019年。Christopher J Cueva、Peter Y Wang、Matthew Chin 和 Xue-Xin Wei。《通过优化循环神经网络，涌现出头方向系统的功能与结构特性》。载于国际学习表征会议，2020年。Benigno Uria、Borja Ibarz、Andrea Banino、Vinicius Zambaldi、Dharshan Kumaran、Demis Hassabis、Caswell Barry 和 Charles Blundell。《哺乳动物大脑中从自我中心到异我中心理解的模型》。*bioRxiv*，2022年。Federico Stella、Peter Baracskay、Joseph O'Neill 和 Jozsef Csicsvari。《海马再激活类似于布朗扩散的随机轨迹》。《神经元》，102(2)，2019年。Julian E Besag。对 Ulf Grenander 和 Michael I Miller 所著《复杂系统中的知识表征》的讨论。《英国皇家统计学会杂志*B*辑》，56(4):591–592，1994年。Xingsi Dong、Zilong Ji、Tianhao Chu、Tiejun Huang、Wenhao Zhang 和 Si Wu。《适应加速吸引子神经网络中的基于采样的贝叶斯推断》。神经信息处理系统进展（*NeurIPS*），第35卷，2022年。Laurence Aitchison 和 Máté Lengyel。《哈密顿大脑：利用兴奋-抑制神经回路动力学实现高效概率推断》。《公共科学图书馆·计算生物学》，12(12)，2016年。Guillaume Hennequin、Laurence Aitchison 和 Máté Lengyel。《平衡神经元网络中的快速基于采样的推断》。神经信息处理系统进展（*NeurIPS*），第27卷，2014年。Zimei Chen、Yi Ren、Eryn Sale、Ying Nian Wu 和 Wen-Hao Zhang。《自然梯度贝叶斯采样在经典皮层回路中自动涌现》。*bioRxiv*，2025年。Thomas J Davidson、Fabian Kloosterman 和 Matthew A Wilson。《海马对扩展经历的重放》。《神经元》，63(4)，2009年。Brad E Pfeiffer。《海马"重放"的内容》。《海马体》，30(1)，2020年。Daniel Levenstein、György Buzsáki 和 John Rinzel。《啮齿动物新皮层和海马体中的非快速眼动睡眠反映了可激发动力学》。《自然-通讯》，10(1)，2019年。Sebastian Michelmann、Bernhard P Staresina、Howard Bowman 和 Simon Hanslmayr。《人类情景记忆中时间压缩的前向重放速度灵活变化》。《自然·人类行为》，3(2)，2019年。Usman Farooq 和 George Dragoi。《出生后早期发育中预配置且具有可塑性的时间压缩序列的涌现》。《科学》，363(6423)，2019年。Christian Leibold、Anja Gundlfinger、Robert Schmidt、Kay Thurley、Dietmar Schmitz 和 Richard Kempter。《由短时程突触可塑性介导的时间压缩》。《美国国家科学院院刊》，105(11)，2008年。Jorge Jaramillo 和 Richard Kempter。《相位进动：一种支撑情景记忆的神经编码？》《神经生物学当前观点》，43，2017年。

Xiang Cheng、Niladri S. Chatterji、Peter L. Bartlett 和 Michael I. Jordan。《欠阻尼朗之万马尔可夫链蒙特卡罗：非渐近分析》。载于第*31*届学习理论会议论文集，第75卷，2018年。Grigorios A Pavliotis。《随机过程与应用》。应用数学文本，第60卷，2014年。Johannes D Seelig 和 Vivek Jayaraman。《地标定向与角度路径积分的神经动力学》。《自然》，521(7551)，2015年。Elizabeth R. Chrastil。《人类路径积分及其神经基础》。载于《人脑百科全书》（第二版），第157–170页。爱思唯尔，2025年。ISBN 978-0-12-820481-8。doi：https://doi.org/10.1016/B978-0-12-820480-1.00016-4。

Jan Benda。神经适应。《当代生物学》，31(3)，2021年。

Boris Gutkin 和 Fleur Zeldenrust。尖峰频率适应。《学术百科》，9(2)，2014年。David A McCormick。丘脑和大脑皮层中的神经递质作用及其在丘脑皮层活动神经调制中的作用。《神经生物学进展》，39(4)，1992年。

Edward D Cui 和 Ben W Strowbridge。内源性乙酰胆碱选择性减弱 Ether-a-go-go 相关的 K+ 电流，从而降低脉冲频率适应和网络相关性。*eLife*，8，2019年。

Caitlin A Wotton、Cassidy D Cross 和 Lane K Bekar。血清素、去甲肾上腺素和乙酰胆碱对星形胶质细胞钾清除的影响各不相同，从而调节雄性小鼠的躯体感觉信号传导。《神经科学研究杂志》，98(5)，2020年。

John O’Donnell、Douglas Zeppenfeld、Evan McConnell、Salvador Pena 和 Maiken Nedergaard。去甲肾上腺素：一种神经调节剂，可增强多种细胞类型的功能，从而优化中枢神经系统的性能。《神经化学研究》，37(11)，2012年。

Steffen Gais 和 Jan Born。慢波睡眠期间较低的乙酰胆碱水平对陈述性记忆巩固至关重要。《美国国家科学院院刊》，101(7)，2004年。

Samsoon Inayat、Qandeel、Mojtaba Nazariahangarkolaee、Surjeet Singh、Bruce L McNaughton、Ian Q Whishaw 和 Majid H Mohajerani。睡眠早期较低的乙酰胆碱水平对运动记忆巩固很重要。《睡眠》，43(6)，2020年。

Andrea Becchetti、AlidaAmadeo 等人。我们为何会忘记梦境：清醒状态和快速眼动睡眠中的乙酰胆碱与去甲肾上腺素。《行为与脑科学》，39，2016年。

Christopher J Watson、Helen A Baghdoyan 和 Ralph Lydic。《睡眠与清醒状态的神经药理学》。《睡眠医学诊所》，第5卷第4期，2010年。Chiara M Portas、Bjørn Bjorvatn 和 Reidun Ursin。《血清素与睡眠/觉醒周期：特别强调微透析研究》。《神经生物学进展》，第60卷第1期，2000年。B Bjorvatn、J Grønli、F Hamre、E Sørensen、E Fiske、Alvhild Alette Bjørkum、CM Portas 和 R Ursin。《睡眠剥夺对大鼠海马体和额叶皮层细胞外血清素的影响》。《神经科学》，第113卷第2期，2002年。Tan Nguyen、Richard Baraniuk、Andrea Bertozzi、Stanley Osher 和 Bao Wang。《MomentumRNN：将动量整合到循环神经网络中》。神经信息处理系统进展（*NeurIPS*），第33卷，2020年。Andrew J Giessel 和 Bernardo L Sabatini。《M1毒蕈碱受体通过抑制突触后SK通道增强树突棘中的突触电位和钙内流》。《神经元》，第68卷第5期，2010年。Katherine A Buchanan、Milos M Petrovic、Sophie EL Chamberlain、Neil V Marrion 和 Jack R Mellor。《毒蕈碱型M1受体通过抑制SK通道介导长时程增强的促进作用》。《神经元》，第68卷第5期，2010年。

Tom M George、Mehul Rastogi、William de Cothi、Claudia Clopath、Kimberly Stachenfeld 和 Caswell Barry。《RatInABox：用于模拟连续环境中运动和神经元活动的工具包》。*eLife*，13:e85274，2024年。国际标准连续出版物号 2050-084X。Matteo Saponati 和 Martin Vinck。《序列预期和脉冲时序依赖可塑性源自预测学习规则》。《自然-通讯》，第14卷第1期，2023。Jacopo Bono、Sara Zannone、Victor Pedrosa 和 Claudia Clopath。《行为和重放期间，脉冲神经元通过预测编码学习预测性认知地图》。*eLife*，第12卷，2023。Ruth A Wood、Marius Bauza、Julija Krupic、Stephen Burton、Andrea Delekate、Dennis Chan 和 John O'Keefe。《蜂窝迷宫为研究海马依赖性空间导航提供了新型测试》。《自然》，第554卷第7690期，2018年。Eric L Denovellis、Anna K Gillespie、Michael E Coulter、Marielena Sosa、Jason E Chung、Uri T Eden 和 Loren M Frank。《以现实世界速度重放经历》。*eLife*，第10卷，2021年。Olivier Jeunehomme 和 Arnaud D'Argembeau。《记住所需的时间：现实生活事件记忆中的时间压缩及时长判断》。《实验心理学季刊》，第72卷第4期，2019年。Ethan R Buch、Leonardo Claudino、Romain Quentin、Marlene Bönstrup 和 Leonardo G Cohen。《与清醒期海马—新皮层重放相关的人类技能巩固》。《细胞报告》，第35卷第10期，2021年。Zhongxuan Wu 和 Xue-Xin Wei。《海马重放的漂移扩散动力学》。*bioRxiv*，2025年。Chaoning Zhang、Chenshuang Zhang、Junha Song、John Seon Keun Yi 和 In So Kweon。《视觉自监督学习中掩码自编码器的综述》。载于 *IJCAI*，2023。Abdullahi Ali、Nasir Ahmad、Elgar de Groot、Marcel Antonius Johannes van Gerven 和 Tim Chris- tian Kietzmann。《预测编码是循环神经网络能量效率的结果》。《*Patterns*》，第3卷第12期，2022年。Dehong Xu、Ruiqi Gao、Wenhao Zhang、Xue-Xin Wei 和 Ying Nian Wu。《关于网格细胞的保角等距：学习保距位置嵌入》。载于国际学习表征会议（*ICLR*），2025年。Xiaohan Zhang、Xiaoyang Long、Sheng-Jia Zhang 和 Zhe Sage Chen。《兴奋性—抑制性循环动力学产生稳健的视觉网格和稳定吸引子》。《细胞报告》，第41卷第11期，2022年。Mufeng Tang、Helen Barron 和 Rafal Bogacz。《通过预测编码学习网格细胞》。*arXiv*预印本 *arXiv:2410.01022*，2024年。Mufeng Tang、Helen Barron 和 Rafal Bogacz。《通过时间预测编码实现序列记忆》。载于 神经信息处理系统进展（*NeurIPS*），第36卷，2023。Mark M Churchland 和 Krishna V Shenoy。《准备活动与扩展零空间》。《自然综述：神经科学》，第25卷第4期，2024年。Selmaan N Chettih、Emily L Mackevicius、Stephanie Hale 和 Dmitriy Aronov。《食物贮藏鸟海马体中的情景记忆条形码化》。《细胞》，第187卷第8期，2024年。Daniel C McNamee、Kimberly L Stachenfeld、Matthew M Botvinick 和 Samuel J Gershman。《内嗅皮层—海马系统中序列生成的灵活调制》。《自然·神经科学》，第24卷第6期，2021年。Victor M Panaretos 和 Yoav Zemel。《Wasserstein距离的统计学方面》。《统计学及其应用年度评论》，第6卷第1期，2019年。Nicolas Bonneel、Julien Rabin、Gabriel Peyré 和 Hanspeter Pfister。《测度的切片Wasserstein重心与Radon-Wasserstein重心》。《数学成像与视觉杂志》，第51卷第1期，2015年。

Sina Alemohammad、Josue Casco-Rodriguez、Lorenzo Luzi、Ahmed Imtiaz Humayun、Hossein Babaei、Daniel LeJeune、Ali Siahkoohi 和 Richard G Baraniuk。自我消耗式生成模型走向 MAD。载于国际学习表征会议（*ICLR*），2024年。Lorenzo Luzi、Paul M Mayer、Josue Casco-Rodriguez、Ali Siahkoohi 和 Richard Baraniuk。Boomerang：使用扩散模型在图像流形上进行局部采样。机器学习研究汇刊（*TMLR*），2024年。Ling Yang、Zhilong Zhang、Yang Song、Shenda Hong、Runsheng Xu、Yue Zhao、Wentao Zhang、Bin Cui 和 Ming-Hsuan Yang。扩散模型：方法与应用的综合综述。*ACM*计算调查，56(4)，2023。Florinel-Alin Croitoru、Vlad Hondru、Radu Tudor Ionescu 和 Mubarak Shah。视觉中的扩散模型：综述。*IEEE*模式分析与机器智能汇刊，45(9)，2023。Xiang Li、John Thickstun、Ishaan Gulrajani、Percy S Liang 和 Tatsunori B Hashimoto。扩散 LM 改进了可控文本生成。神经信息处理系统进展（*NeurIPS*），第35卷，2022年。Roberto Gozalo-Brizuela 和 Eduardo C Garrido-Merchan。ChatGPT 并非你所需的一切：大型生成式人工智能模型的最新技术综述。*arXiv* 预印本 *arXiv:2301.04655*，2023。Yang Song 和 Stefano Ermon。通过估计数据分布的梯度进行生成建模。载于神经信息处理系统进展（*NeurIPS*），第32卷，2019年。

补充材料

A 方法

A.1 各实验共用的超参数

A.1.1 掩码训练

去噪器和自编码器受益于掩码训练，其中输入数据的某些区域在模型处理之前被设为零（Zhang 等人，2023）。Levenstein 等人（2024年）将掩码训练引入路径积分，定期对输入 $\boldsymbol{u}(t)$（对 $\boldsymbol{s}'(t)$ 的观测）：循环神经网络仅在每隔 k 个时间步时才观测到非零输入（方程17）。

$$\boldsymbol{u}(t) = \begin{cases} \boldsymbol{s}'(t), & t \bmod k = 0 \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{17}$$

Levenstein 等人 (2024年)发现，掩码训练能使重放序列更加连贯，并使神经活动流形与环境的空间布局更加相似。我们发现，掩码训练能够提高重放稳定性，因此在训练所有循环神经网络时，我们都使用该方法（使用 $k \geq 3$）。

A.1.2 激活函数

在我们的研究中，我们主要使用采用 ReLU 或带泄漏 ReLU 激活函数的循环神经网络（第 F 节除外）；这样做有三个原因。首先，与 tanh 激活不同，ReLU 或类似 ReLU 的激活并不总是会饱和。这有助于缓解循环神经网络中的梯度消失（尽管必须注意避免梯度爆炸），但更重要的是，我们观察到，这会减弱路径积分循环神经网络学到的固定吸引子强度，这有利于探索（在第 F 节中，我们观察到，在三角形任务中，饱和型循环神经网络的重放探索程度较低）。其次，与采用乘法门控交互的循环神经网络（例如门控循环单元和长短期记忆网络）相比，ReLU 循环神经网络在生物学上更为合理，因为它们很容易被解释为脉冲网络中非负且稀疏的放电率。第三，其他大多数序列预测编码领域的既有研究往往使用 ReLU 激活，并且表明使用更复杂的非线性函数时，结果并无显著差异：

- Ali 等人(2022年是最早报告循环神经网络中出现预测表征的作者之一，他们使用 ReLU 激活。
- Krishna等人(2024年主要使用 ReLU 网络，并报告了使用 GRU 网络时也能得到类似结果。
- Levenstein 等人(2024年将 ReLU 激活与层归一化结合使用，但报告称移除层归一化后结果也类似。
- Chen 等人(2024年大多数结果使用双曲正切，但在关注稀疏、局部化的非负激活（即位置细胞）的出现时，他们使用 ReLU 激活。
- Sorscher 等人(2019年), Xu 等人(2025年），以及Zhang 等人(2022年）都关注预测表征中网格细胞的出现，且主要使用 ReLU 或类似 ReLU 的激活函数；尝试使用其他非线性函数的研究者未观察到显著的结果差异。
- Tang 等人 (2024年同样对网格细胞感兴趣，同时使用 ReLU 和 tanh激活，并且也报告称两者之间没有显著差异。
- Tang 等人(2023使用线性网络，并注意到使用 tanh激活时没有显著差异。

A.2 实验

所有实验均使用 PyTorch，在单块消费级 GPU（NVIDIA Titan X Pascal 12 GB）上进行。

A.2.1 奥恩斯坦-乌伦贝克

清醒轨迹 $s(t)$ 通过以下参数进行模拟：$\Delta t = 0.02$，$\sigma_s = 0.1$，$\sigma_0 = 0.2$，$\theta = 2$，$\mu = 5$，共 $T = 100$ 次迭代。我们使用公式（47）作为 $\Delta \boldsymbol{r}(t)$ 的确定性部分，因为一维奥恩斯坦–

乌伦贝克过程对于 $\frac{d}{dr}(t)$ 的对数 $p(r(t))$有闭式表达式。在图 1中，动量（欠阻尼）对应于 $\lambda_v = 0.5$，而适应对应于 $b_a = 0.01$，$\tau_a = 20$。

A.2.2 T迷宫和三角形

任务描述。 在二维T迷宫和三角形任务中，我们模拟沿多个方向进行的有向随机游走（Ornstein-Uhlenbeck过程），并训练循环神经网络根据这些随机游走的速度对其进行路径积分。在T迷宫任务中，有两个移动方向：二者都从原点出发向上移动，随后一个向左、另一个向右（在图3中分别以橙色和紫色表示）。与此同时，在三角形任务中有六个方向：将等边三角形的三个角分别记为 $A = (0,\ 0)$ $B = (1,\ 0)$和 $C = (1/2,\ \sqrt{3}/2)$，这些方向分别为 $\overrightarrow{AB}$，$\overrightarrow{BC}$，$\overrightarrow{CA}$，$\overrightarrow{AC}$，$\overrightarrow{CB}$，$\overrightarrow{BA}$，如图 3 所示，分别以蓝色、红色、灰色、粉色、绿色和橙色表示。清醒和重放路径持续 100 个时间步；除非正在测量探索，在这种情况下，重放路径允许持续 400 个时间步。

架构与训练。 对于两个任务，我们均使用了浅层泄漏ReLU循环神经网络，并采用线性输出投影，具体如第 A.4节所述。在T迷宫任务中，我们使用20个隐藏神经元；在三角形任务中，则使用40个。我们在渐进式课程中使用掩蔽难度 $k = 3$。三角形任务的循环神经网络首先使用 $k = 1$ 训练 20,000 个训练轮次，随后使用 $k = 2$ 和 $k = 3$ 分别训练 $5,000$ 个训练轮次；T迷宫任务的循环神经网络训练方式相同，但训练轮次更少（在 12,000 个训练轮次时使用 $k = 1$，在 5,000 个训练轮次时使用 $k = 2$，并且 $k = 3$ 各自使用相应设置）。三角形任务训练9个循环神经网络随机种子，T形迷宫任务训练10个。

隐藏状态初始化。 在每个任务中，多个方向从空间中的同一点出发（例如，在三角形任务中，$\overrightarrow{AB}$ 和 $\overrightarrow{AC}$），因此我们在初始隐藏状态上添加了一些与二维输出投影正交、且针对每个方向特定的随机向量。例如，$\overrightarrow{AB}$ 和 $\overrightarrow{AC}$ 路径的隐藏状态都初始化为从 A附近开始，但 $\overrightarrow{AB}$ 路径还会在其初始隐藏状态中加入一个固定的随机向量 $\boldsymbol{\eta}_{AB}$，而 $\overrightarrow{AC}$ 则会在其初始隐藏状态中加入另一个固定的随机向量 $\boldsymbol{\eta}_{AC}$。在计算神经科学背景下，已有研究讨论过这种将隐藏状态初始化为与输出投影正交方向的做法（ChurchlandandShenoy，2024年）；近期证据表明，记忆（重放与之类似）具有独特且可识别的、与输出正交的神经激活模式（Chettih 等人，2024年）。

A.2.3 大鼠位置细胞轨迹

我们的无偏（在二维盒子内进行无向随机游走）和有偏（朝二维盒子中心移动的有向随机游走）大鼠位置细胞轨迹实验与（Krishna等人，2024年）中的实验完全相同：每项任务使用 10 个不同的 ReLU 循环神经网络随机种子，但有两处修改：

1. 我们加入了掩码训练，在有偏任务中的掩蔽难度为 $k = 3$，在无偏任务中的掩蔽难度为 $k = 6$。我们在无偏任务中采用更高的掩蔽难度，以鼓励生成更长的重放轨迹；当 k较低时，无偏重放路径远短于清醒轨迹。从朗之万采样的角度来看，我们推测这是因为得分函数 $\nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t))$ 相当弱，因为无偏轨迹呈均匀分布[9]。

2. 我们采用一种速度较慢但更为精细的解码器。（Krishna等人，2024年）他们将给定时间步的位置解码为活动最强的3个位置细胞的位置均值。这种方法速度很快，但实际上会使解码后的位置发生量化。我们用这种方法初始化位置，随后通过梯度下降优化位置，使观测到的位置细胞活动与这些优化后位置所对应的位置细胞活动之间的均方误差最小化（这得益于我们掌握了精确的位置细胞活动函数）。我们的方法耗时更长，但能得到真正连续的重放轨迹，避免显著的量化。

[9]无偏活动得分函数相对较弱（尤其是与有偏活动得分函数相比，后者的轨迹都朝二维盒子中心移动），这也是我们在模拟无偏任务循环神经网络的重放时采用相对较小适应强度 b_a 的原因。

A.3 指标

A.3.1 速度：到达时间

在T形迷宫、三角形和有偏置老鼠任务中，清醒轨迹都有明确的终点，并会到达这些终点。重放轨迹也会朝这些终点前进，尽管适应和动量（欠阻尼）等二阶动力学修饰因素会影响重放路径到达终点的速度或接近程度。我们通过测量重放路径进入终点某个指定半径范围内所需的时间步数，来量化这些变化（例如，$\overrightarrow{AB}$ 路径进入距离 $B0.1|\overrightarrow{AB}|$ 半径范围内所需的时间步数）。

A.3.2 探索：路径长度

我们量化重放路径中的探索程度的一种方式是通过路径长度。我们将其简单计算为速度大小之和。

A.3.3 探索：访问区域

我们量化T形迷宫和三角形任务中的探索程度的另一种方式，是通过访问区域。如图6所示，区域是输入空间中的一部分，其中所有点都距离同一个端点最近。在三角形任务中，有3个端点（它们同时也都是初始点）；而在T形迷宫中，有2个端点和1个共享起点，出于区域分配的目的，我们也将后者视为一个“端点”。因此，每项任务中都有3个端点，也就有3个区域。清醒轨迹从一个区域前往另一个区域，并停留在那里；但重放轨迹可能会继续访问更多区域。我们将“访问”定义为在单个区域内连续存在至少10个时间步；自然地，探索性重放序列可以重新访问先前访问过的区域，这就是其“访问区域”指标能够超过3的原因。

A.3.4 探索：附加指标

为完整起见，我们补充另外两个探索测度：平均位移，由McNamee 等人（2021年）提出，以及方差。它们的定义非常简单：$\mathbb{E}_{p(\boldsymbol{r}(t))}[|\boldsymbol{r}(t) - \boldsymbol{r}(0)|]$ 和$\mathbb{E}_{p(\boldsymbol{r}(t))}[\mathrm{Var}(\boldsymbol{r}(t)]$。较大的平均位移和方差通常与更长的路径长度和更强的探索相关。我们在图8 和图9中绘制了它们。

A.3.5 Wasserstein距离

我们希望比较轨迹分布。在我们的大多数任务中，这些分布是维度为$T \times 2$的对象，其中 T 是时间步的数量（回顾一下，T形迷宫和三角形任务本身就在二维空间中，而大鼠轨迹则是在投影到二维空间后进行分析）。为了比较此类高维分布，我们使用 Wasserstein距离（Panaretos 和 Zemel，2019年）：

- 在 T形迷宫和三角形任务中，我们首先计算属于同一方向的清醒路径与重放路径之间的Wasserstein距离（例如，在三角形任务中为 $\overrightarrow{AB}$）。由于同一方向上的路径应当彼此相似，即使不一定遵循高斯过程，这些路径也应近似服从正态分布，因此我们可以使用两个正态分布之间 Wasserstein距离的闭式方程。计算出每个方向的 Wasserstein距离后，我们取其平均值作为最终距离。在此任务中，KL散度或许也能发挥作用，但实际上，（$T \times 2$）维分布的协方差矩阵是奇异的。
- 在大鼠实验中，解码得到的二维轨迹不易分解为高斯过程组，因此我们改用切片 Wasserstein 距离（Bonneel 等人，2015年）来比较 $(T \times 2)$维分布。切片 Wasserstein 距离的计算本质上是将两个分布随机投影到一维空间多次，在一维空间中Wasserstein距离具有闭式解。若改为计算 KL 散度，过程将颇具挑战，因为大鼠路径分布是高维的，且无法对概率密度进行直接估计。

A.4 含噪循环神经网络的离散化与实现

虽然生物神经网络是连续时间系统，但循环神经网络在实际中以离散形式实现。为了将适应和动量等二阶过程纳入循环神经网络，我们必须对其进行离散化。我们最初采用如下方式，并将其称为“非合理动量”：

$$\Delta\widetilde{\boldsymbol{r}}(t) = f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r\boldsymbol{\eta}(t)) - \boldsymbol{r}(t) \tag{18}$$

$$\boldsymbol{c}(t+\Delta t) = \boldsymbol{c}(t) + \frac{1}{\tau_a}(-\boldsymbol{c}(t) + b_a\boldsymbol{r}(t)) \tag{19}$$

$$\boldsymbol{v}(t+\Delta t) = (1-\lambda_v)\boldsymbol{v}(t) + \Delta\widetilde{\boldsymbol{r}}(t) \tag{20}$$

$$\boldsymbol{r}(t+\Delta t) = \boldsymbol{r}(t) - \boldsymbol{c}(t) + \boldsymbol{v}(t+\Delta t) \tag{21}$$

请注意，与 $\boldsymbol{c}(t+\Delta t)$ 和 $\boldsymbol{v}(t+\Delta t)$ 不同， $\boldsymbol{r}(t+\Delta t)$ 依赖于在 $t+\Delta t$ 处计算的另一项。这是二阶动力学的辛欧拉离散化，我们采用这种方法来确保 $\boldsymbol{r}(t)$ 与其动量 $\boldsymbol{v}(t)$ 之间相互作用的稳定性。负反馈 $\boldsymbol{c}(t)$ 稳定得多，因此其离散化保持不变。

如果适应强度 $b_a = 0$，则不存在适应（$\boldsymbol{c}(t) = \boldsymbol{0}$）。至于我们的摩擦 项 $\lambda_v \in [0,1],$，如果 $\lambda_v = 1$，则不存在动量，因为 $\boldsymbol{v}(t) = \Delta\boldsymbol{r}(t) + \sigma_r\boldsymbol{\eta}(t)$ 不再累积 $\boldsymbol{v}(t)$ 的先前取值。换言之，摩擦项 λ_v 允许在过阻尼和动量采样之间平滑插值。

虽然我们最初采用了上述动力学（结果见第E 节），但它们通过 $\boldsymbol{v}$（t）使用了在生物学上不合理的动量。因此，我们通过输入微分重新定义了上述动力学，引入合理动量，得到最终的主要动力学（推导见第 B.2节）：

$$\boldsymbol{c}(t+\Delta t) = \boldsymbol{c}(t) + \frac{1}{\tau_a}(-\boldsymbol{c}(t) + b_a\boldsymbol{r}(t)) \tag{22}$$

$$\boldsymbol{r}(t+\Delta t) = f(\boldsymbol{r}(t) + (1-\lambda_v)(\boldsymbol{r}(t-\Delta t) - \boldsymbol{r}(t-2\Delta t)), \boldsymbol{u}(t), \sigma_r\boldsymbol{\eta}(t)) - \boldsymbol{c}(t) \tag{23}$$

我们使用带有或不带泄漏积分的 ReLU循环神经网络进行训练和采样：$f(\boldsymbol{r}(t)\ \boldsymbol{u}(t)\ \sigma_r\boldsymbol{\eta}(t)) = \kappa\boldsymbol{r}(t) + \text{ReLU}(\boldsymbol{W}_r\boldsymbol{r}(t) + \boldsymbol{W}_{in}\boldsymbol{u}(t) + \sigma_r\boldsymbol{\eta}(t))$（其中 $\kappa \in [0,1]$ 为 0 或可学习参数），除非另有说明；我们在 A.1 节中说明了选择 ReLU激活函数的理由，并在 F 节中验证结果能否推广到其他激活函数。在没有定向输入 $\boldsymbol{u}(t)$ 的情况下， $f(\boldsymbol{r}(t)\ \boldsymbol{u}(t)\ \sigma_r\boldsymbol{\eta}(t))$ 应大致等同于公式 11 中的噪声得分函数。

在所有训练循环神经网络实验中，我们使用 $b_a = 0$， $\lambda_v = 1$ 用于清醒活动（即训练）。生成重放时，我们始终使用 $\tau_a = 50$，并根据测量速度和保真度（ $T = 100$、400、或500 个时间步；（见图$T = 100$ 3、4、7）或探索（$T = 400$用于 T形迷宫和三角形， $T = 500$ 用于图5、6中的大鼠任务）而定。

A.5 与传统朗之万采样的偏差

We 必须指出，我们的重放循环神经网络与传统朗之万采样之间存在一些细微差别 g.

平稳性。 我们将神经重放视为一系列事件，因此关注所有时间步 $p\left(\{\boldsymbol{r}(t)\}_{t=1}^{T}\right)$ 上重放活动的联合分布。然而，该分布在 t 上不一定是平稳的。平稳性意味着 $p(\boldsymbol{r}(t))$ 无需也无意沿着任何有意义的路径前进。尽管路径积分循环神经网络确实沿着 $\log p(\boldsymbol{r}(t))$ 执行梯度上升，但它是沿 t 分段进行的，而不是同时沿所有 t 进行。这种梯度上升方式的变化，加上 $p(\boldsymbol{r}(t))$ 的非平稳性，意味着朗之万采样的保证并不成立。因此，我们对循环神经网络进行了修改；根据平稳朗之万采样理论，这些修改或许会显得缺乏原则依据。

动量（显式欠阻尼）。 摩擦项 λ_v 作用于速度 $\boldsymbol{v}(t)$ 与 γ 在方程 3 和 2中略有不同。 $\lambda_v \in [0,1]$ 可以解释为试图将 $\gamma \in [0,\infty)$映射到 $[0,1]$。与 $\gamma \to \infty$一样， $\lambda_v \to 1$ 消除了 $\boldsymbol{v}(t+\Delta t)$ 对 $\boldsymbol{v}(t)$的任何依赖。

噪声缩放。 RNN 重放与朗之万采样不同，对所用噪声方差很敏感。我们发现，省略 $\sigma_r \boldsymbol{\eta}(t)$ 前面的 $\sqrt{2}$ 因子效果最好。因此，我们也省略了作为 $\sigma_r \boldsymbol{\eta}(t)$缩放项的$\lambda_v$。如果允许 λ_v 调制 $\sqrt{2\sigma_r}\boldsymbol{\eta}(t)$，我们发现当 $\lambda_v < 1$时很容易显示出重放保真度的提升，但这些提升来自噪声缩放，而非作为机制的动量。

与扩散模型的关系。在本文中，我们将循环神经网络视为使用一种变体朗之万采样的生成模型。表达能力强的生成模型，尤其是扩散模型，其活动模式类似于朗之万采样，因此受到机器学习界的广泛关注（Alemohammad 等人，2024年，Luzi 等人，2024年，Yang 等人，2023，Croitoru 等人，2023，Li 等人，2022年，Gozalo-Brizuela 和 Garrido-Merchan，2023，Song 和 Ermon，2019年）。我们的循环神经网络与扩散模型之间存在两项根本差异。第一，生成的重放轨迹中的每个时间步都是按顺序生成的，并且沿对数似然梯度仅进行一个有效步；而扩散模型会同时生成路径中的所有时间步，并沿对数似然梯度进行多步迭代。第二，扩散模型使用随时间变化的噪声水平（退火动力学），并明确以时间为输入进行条件化；而我们的循环神经网络在所有时间步使用相同的噪声水平，且从不显式以时间为条件。

B 二阶随机微分方程推导

B.1 将适应表示为二阶随机微分方程

我们首先通过以下替换，对方程（13）进行抽象化处理。该方程描述的是训练后引入适应机制的路径积分循环神经网络的静息动力学：得分函数$\sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log\, p(\boldsymbol{r}(t)) \;\to\; g(\boldsymbol{r}(t))$、缩放后的噪声 $\sigma_r \boldsymbol{\eta}(t) \;\to\; \boldsymbol{\varepsilon}(t)$，以及适应参数$\frac{1}{\tau_a} \;\to\; a$和 $\frac{b_a}{\tau_a} \;\to\; b$。由此，我们将方程（13）改写为如下单个二阶随机微分方程：

$$\boldsymbol{r}'(t) = g(\boldsymbol{r}(t)) - \boldsymbol{c}(t) + \boldsymbol{\varepsilon}(t), \quad \boldsymbol{c}'(t) = -a\boldsymbol{c}(t) + b\boldsymbol{r}(t) \tag{24}$$

$$\begin{aligned}
\boldsymbol{r}''(t) &= g'(\boldsymbol{r}(t))\boldsymbol{r}'(t) - \boldsymbol{c}'(t) + \boldsymbol{\varepsilon}'(t) && (25)\\
&= g'(\boldsymbol{r}(t))\boldsymbol{r}'(t) - (-a\boldsymbol{c}(t) + b\boldsymbol{r}(t)) + \boldsymbol{\varepsilon}'(t), \quad \boldsymbol{c}(t) = -\boldsymbol{r}'(t) + g(\boldsymbol{r}(t)) - \boldsymbol{\varepsilon}(t) && (26)\\
&= g'(\boldsymbol{r}(t))\boldsymbol{r}'(t) - (-a[-\boldsymbol{r}'(t) + g(\boldsymbol{r}(t)) - \boldsymbol{\varepsilon}(t)] + b\boldsymbol{r}(t)) + \boldsymbol{\varepsilon}'(t) && (27)\\
&= g'(\boldsymbol{r}(t))\boldsymbol{r}'(t) - a\boldsymbol{r}'(t) + ag(\boldsymbol{r}(t)) - a\boldsymbol{\varepsilon}(t) - b\boldsymbol{r}(t) + \boldsymbol{\varepsilon}'(t) && (28)\\
&= (g'(\boldsymbol{r}(t)) - a)\boldsymbol{r}'(t) + (ag(\boldsymbol{r}(t)) - b\boldsymbol{r}(t)) - a\boldsymbol{\varepsilon}(t) + \boldsymbol{\varepsilon}'(t) && (29)
\end{aligned}$$

Rep将替换变量还原为其原始值后，可得以下结果（方程（14））：

$$\begin{aligned}
\boldsymbol{r}''(t) = &\left(\sigma_r^2 \Delta t \frac{d^2}{d\boldsymbol{r}(t)^2} \log p(\boldsymbol{r}(t)) - \frac{1}{\tau_a}\boldsymbol{I}\right)\boldsymbol{r}'(t)\\
&+\frac{1}{\tau_a}\left(\sigma_r^2 \Delta t \nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) - b_a \boldsymbol{r}(t)\right)\\
&-\frac{1}{\tau_a}\sigma_r \boldsymbol{\eta}(t) + \sigma_r \boldsymbol{\eta}'(t)
\end{aligned} \tag{30}$$

B.2 生物学上合理的动量推导

在此，我们通过输入微分推导出生物学上合理的动量机制。首先，我们从第 A.4 节中离散化动力学的定义出发，并假设不存在适应（$b_a = 0$，因此 $\boldsymbol{c}(t) = \boldsymbol{0}$）：

$$\boldsymbol{v}(t + \Delta t) = (1 - \lambda_v)\boldsymbol{v}(t) + f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) - \boldsymbol{r}(t) \tag{31}$$

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t + \Delta t) \tag{32}$$

接下来，我们将一个方程代入另一个方程，消去 $\boldsymbol{r}(t)$ 项：

$$\begin{aligned}
\boldsymbol{r}(t + \Delta t) &= \boldsymbol{r}(t) + (1 - \lambda_v)\boldsymbol{v}(t) + f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) - \boldsymbol{r}(t) && (33)\\
&= (1 - \lambda_v)\boldsymbol{v}(t) + f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) && (34)
\end{aligned}$$

根据定义，$\boldsymbol{v}(t) = \boldsymbol{r}(t) - \boldsymbol{r}(t-\Delta t)$。因此：

$$\boldsymbol{r}(t+\Delta t) = (1-\lambda_v)(\boldsymbol{r}(t) - \boldsymbol{r}(t-\Delta t)) + f(\boldsymbol{r}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)) \tag{35}$$

因此，动量等于将按标量加权的近期速度（$(1-\lambda_v)$（$\boldsymbol{r}$（t）$-\boldsymbol{r}$（$t-\Delta t$））加到下一时间步的活动量 $\boldsymbol{r}$（$t+\Delta t$）的计算中。神经元不太可能做到这一点。不过，神经元对输入进行微分并将这些导数作为独立输入是合理的，因此我们提出将微分从神经元的输出（突触前）移至其输入（突触后），本质上是将动量应用于 $\boldsymbol{r}$（$t+\Delta t$）中的 $\boldsymbol{r}$（t）项，而不是直接应用于 $\boldsymbol{r}$（$t+\Delta t$）本身：

$$\boldsymbol{r}(t+\Delta t) = f(\boldsymbol{r}_{in}(t), \boldsymbol{u}(t), \sigma_r \boldsymbol{\eta}(t)), \quad \text{where} \tag{36}$$

$$\boldsymbol{r}_{in}(t) = \boldsymbol{r}(t) + (1-\lambda_v)(\boldsymbol{r}(t-\Delta t) - \boldsymbol{r}(t-2\Delta t)) \tag{37}$$

因此，动量可以通过输入微分的一种特定变体，以生物学上合理的方式实现。当然，如果 $\lambda_v = 1$，$\boldsymbol{r}_{in}$（t）$= \boldsymbol{r}$（t）。

C 高斯分布与过程的得分函数

对于涉及矩阵微积分的部分，我们采用分母布局。

C.1 多元高斯分布的得分

假设 $\boldsymbol{r} \sim \mathcal{N}$（$\boldsymbol{\mu}$，$\boldsymbol{\Sigma}$）。如果 $\boldsymbol{r} \in \mathbb{R}^d$，那么：

$$p(\boldsymbol{r}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{r}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{r}-\boldsymbol{\mu})\right) \tag{38}$$

$$\log p(\boldsymbol{r}) \propto -\frac{1}{2}(\boldsymbol{r}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{r}-\boldsymbol{\mu}) \tag{39}$$

$$\nabla_{\boldsymbol{r}} \log p(\boldsymbol{r}) = -\frac{1}{2}((\boldsymbol{\Sigma}^{-1})^T + \boldsymbol{\Sigma}^{-1})(\boldsymbol{r}-\boldsymbol{\mu}) \tag{40}$$

$$= -\boldsymbol{\Sigma}^{-1}(\boldsymbol{r}-\boldsymbol{\mu}) \tag{41}$$

$$= -\sigma^{-2}(r-\mu) \text{ if } r \in \mathbb{R} \tag{42}$$

C.2 任意高斯过程 r（t） s（t）的得分

回顾 p（$\boldsymbol{r}$（t）$|\boldsymbol{s}$（t）$\sim \mathcal{N}$（$\boldsymbol{D}^\dagger \boldsymbol{s}$（$t$） $\boldsymbol{I}\sigma_r^2 \Delta t$）自公式6。假设 $\boldsymbol{s}$（t）服从正态分布，其均值和协方差为$\boldsymbol{\mu}_{\boldsymbol{s}(t)}$、$\boldsymbol{\Sigma}_{\boldsymbol{s}(t)}$，则可以得到 p（$\boldsymbol{r}$（t）：

$$p(\boldsymbol{r}(t)) \sim \mathcal{N}(\boldsymbol{D}^\dagger \boldsymbol{\mu}_{\boldsymbol{s}(t)},\ \boldsymbol{I}\sigma_r^2 \Delta t + \boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T), \tag{43}$$

将其代入公式41 可得 $\nabla_{\boldsymbol{r}(t)} \log p$（$\boldsymbol{r}$（$t$）：

$$\nabla_{\boldsymbol{r}(t)} \log p(\boldsymbol{r}(t)) = -\left(\boldsymbol{I}\sigma_r^2 \Delta t + \boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T\right)^{-1} \left(\boldsymbol{r}(t) - \boldsymbol{D}^\dagger \boldsymbol{\mu}_{\boldsymbol{s}(t)}\right) \tag{44}$$

此外，我们可以使用上述得分函数来计算公式1中的最优 $\Delta \boldsymbol{r}$（t） 0:

$$\begin{aligned} \Delta \boldsymbol{r}^*(t) = \sigma_r^2 \Delta t \left(\boldsymbol{I}\sigma_r^2 \Delta t + \boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T\right)^{-1} \left(-\boldsymbol{r}(t) + \boldsymbol{D}^\dagger \boldsymbol{\mu}_{\boldsymbol{s}(t)}\right) \\ + \boldsymbol{D}^\dagger \boldsymbol{s}'(t) \Delta t + \sigma_r \boldsymbol{\eta}(t) \end{aligned} \tag{45}$$

泄漏矩阵 $\sigma_r^2 \Delta t \left(\boldsymbol{I}\sigma_r^2 \Delta t + \boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T\right)^{-1}$ 的一些性质包括：

1. 协方差矩阵 $\boldsymbol{\Sigma}_{\boldsymbol{s}(t)}$ 是半正定的（PSD）：其所有特征值均为 $\geq$ 0.
2. $\boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T$ 也是半正定 [10], 对称的，因此可对角化。
3. 因此，$(\boldsymbol{I}\sigma_r^2 \Delta t + \boldsymbol{D}^\dagger \boldsymbol{\Sigma}_{\boldsymbol{s}(t)} (\boldsymbol{D}^\dagger)^T)^{-1}$ 的特征值全都是 $\leq (\sigma_r^2 \Delta t)^{-1}$ [11]。

[10]证明：如果 $\boldsymbol{B}$ 是半正定的，那么 $x^T \boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T x = (\boldsymbol{A}^T x)^T \boldsymbol{B}(\boldsymbol{A}^T x) = v^T \boldsymbol{B} v \geq 0$。[11]证明：对于可对角化的 $\boldsymbol{A}$，$(\lambda \boldsymbol{I} + \boldsymbol{A})^{-1} = (\boldsymbol{Q}(\lambda \boldsymbol{I} + \boldsymbol{\Lambda})\boldsymbol{Q}^{-1})^{-1}$ 的第 i 个特征值等于 $(\lambda + \boldsymbol{\Lambda}_{ii})^{-1}$，当 $\boldsymbol{A}$ 为半正定时，该值不会大于 λ^{-1}。

4. 因此， $\sigma_r^2\Delta t(\boldsymbol{I}\sigma_r^2\Delta t + \boldsymbol{D}^\dagger\boldsymbol{\Sigma}_{\boldsymbol{s}(t)}(\boldsymbol{D}^\dagger)^T)^{-1}$ 的特征值全都是 ≤ 1。 5. 如果上述泄漏矩阵的非对角线元素的绝对值足够小，那么根据 Gershgorin圆盘定理，所有对角线元素都应小于 1。事实上，如果泄漏矩阵是对角矩阵，那么它的所有值都必须小于 1 （这可以通过 Sigmoid函数或谱归一化来实现）。6. 关于其解释，较小的泄漏特征值意味着 $\boldsymbol{D}^\dagger\boldsymbol{\Sigma}_{\boldsymbol{s}(t)}(\boldsymbol{D}^\dagger)^T$的特征值较高，或者本质上意味着噪声更多。泄漏矩阵的最大行列式为 1，此时 $\boldsymbol{s}(t)$中本质上没有噪声。

C.3 若干马尔可夫高斯过程的得分

维纳过程。一种简单的随机过程是维纳过程；从导航的角度来看，它表示无向随机游走（$\theta = 0$）。假设 $s_w(0) = 0$，则 $s_w(t) \sim \mathcal{N}(0, \sigma_s^2 t)$，因此 $p(r_w(t)) \sim \mathcal{N}(0, \sigma_s^2 t + \sigma_r^2 \Delta t)$ 根据方程 6，可得如下得分：

$$\sigma_r^2 \Delta t \, \nabla_{r_w(t)} \log p(r_w(t)) = \sigma_r^2 \Delta t \frac{-r_w(t)}{\sigma_s^2 t + \sigma_r^2 \Delta t} \tag{46}$$

即使对于简单的维纳过程，我们也能观察到，每个时间步的最优得分并不随 t 保持恒定：当 $t = 0$ 时，它等于 $-r_w(t)$；而当 t 接近 ∞ 时，它趋近于 0。

奥恩斯坦-乌伦贝克过程。现在，我们引入非零泄漏（$\theta > 0$），以描述从任意起点 $s_{ou}(0)$ 朝平均目的地 μ 前进的有向随机游走。若 $p(s_{ou}(0)) \sim \mathcal{N}(0, \sigma_0^2)$，则 $p(s_{ou}(t)) \sim \mathcal{N}\left(\mu(1 - e^{-\theta t}), \frac{\sigma_s^2}{2\theta}(1 - e^{-2\theta t}) + \sigma_0^2 e^{-2\theta t}\right)$，其得分函数为：

$$\sigma_r^2 \Delta t \, \nabla_{r_{ou}(t)} \log p(r_{ou}(t)) = \sigma_r^2 \Delta t \frac{-(r_{ou}(t) - \mu(1 - e^{-\theta t}))}{\frac{\sigma_s^2}{2\theta}(1 - e^{-2\theta t}) + \sigma_0^2 e^{-2\theta t} + \sigma_r^2 \Delta t} \tag{47}$$

D 其他结果

这里展示补充主文内容的其他结果或图表。

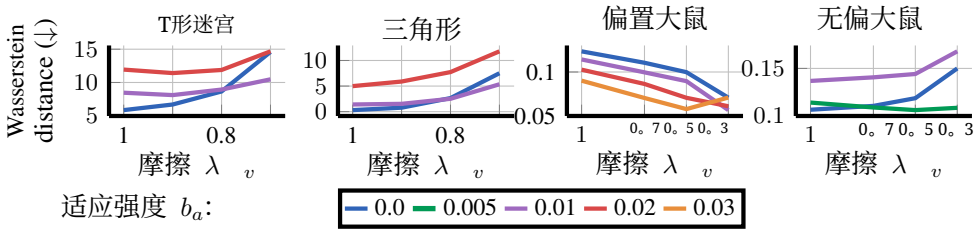

图 7：重放轨迹与清醒轨迹之间的 Wasserstein距离（第 A.3.5 节）。具有 tanh激活和非合理动量的循环神经网络版本见图 17 和 13。

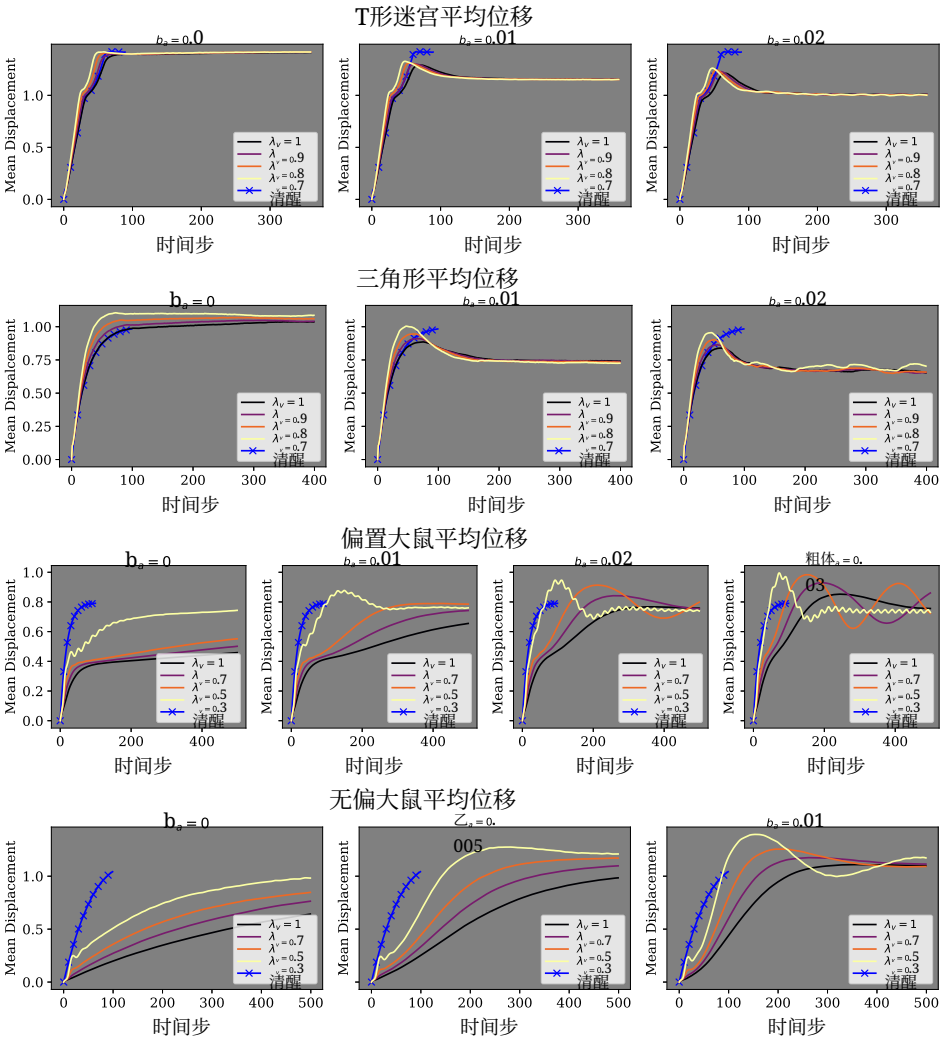

图8： **动量（$\lambda_v < 1$）会增加重放轨迹的平均位移，尤其是在早期时间步。** 与McNamee 等人（2021年）类似，我们在此分析平均位移，即重放轨迹随时间相对于其起点的距离。作为参考，我们还绘制了清醒轨迹随时间的平均位移。在探索实验中，我们模拟重放 $4\times$ 清醒轨迹的持续时间。随时间增加的平均位移通常与路径长度和探索程度的增加相关。请注意，清醒活动始终使用 $b_a = 0$， $\lambda_v = 1$进行计算，因此同一任务中的各图清醒状态统计数据相同。

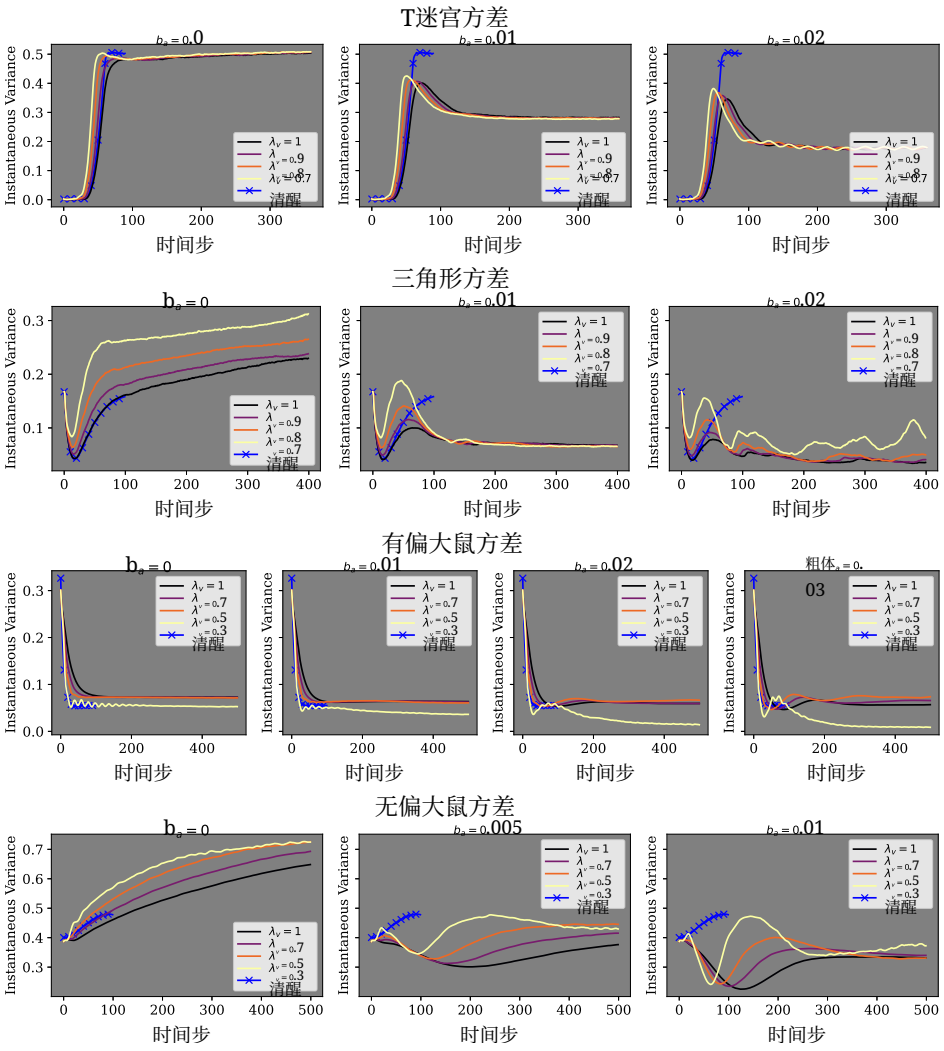

图9：**动量（$\lambda_v < 1$）会增加重放轨迹方差**。与图8中一样，我们在这里分析了轨迹在每个时间步的方差。轨迹是二维的，因此上方绘制的方差只是沿各个坐标方向方差的平均值。动量会增加方差，这进一步证实它能够补充适应诱导的探索。

E 非合理动量的结果

这里我们重现主要发现，但使用非合理动量（第 A.4节），而不是通过输入微分实现的合理动量。

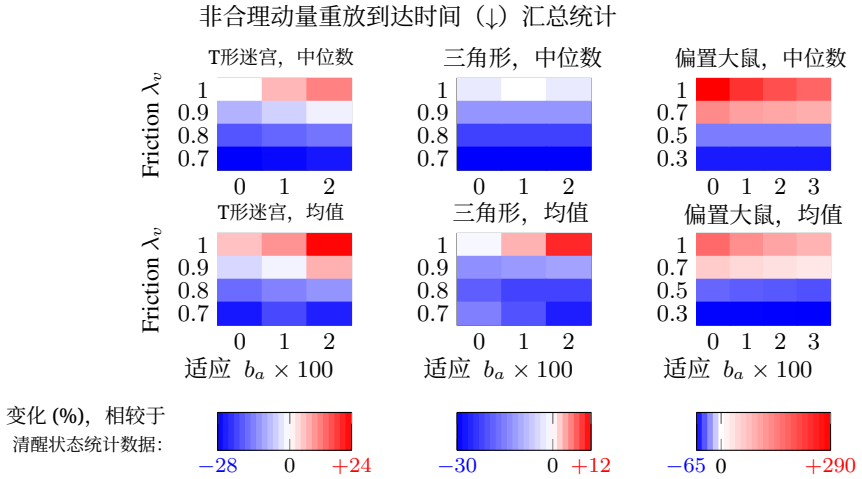

图10： **非合理动量使循环神经网络中的重放在时间上得到压缩**。这是图 4 的另一个版本，但此处使用了非合理动量（见第 A.4 节）。有关双曲正切循环神经网络中合理动量的版本，请参见图 14 。

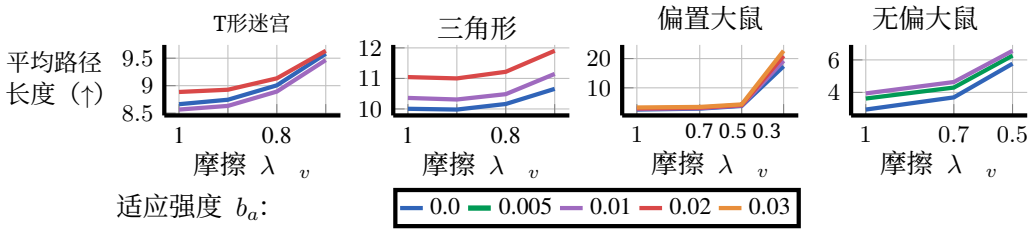

图11： **在具有非合理动量的循环神经网络中，动量和适应都会通过回放路径长度增加探索**。 这是图5的另一个版本，不过这里采用的是非合理动量（见第 A.4节）。关于双曲正切循环神经网络中采用合理动量的版本，请参见图15 。

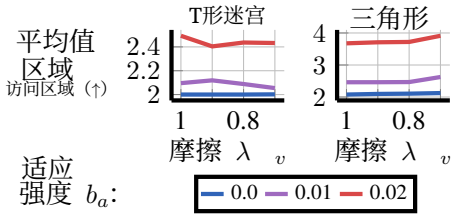

图12： **在具有非合理动量的循环神经网络中，动量不会通过重放期间的访问区域减少探索**。这是图 6 中测量结果的另一个版本，但使用了非合理动量（第 A.4 节）。有关使用双曲正切循环神经网络和合理动量的版本，请参见图 16 。

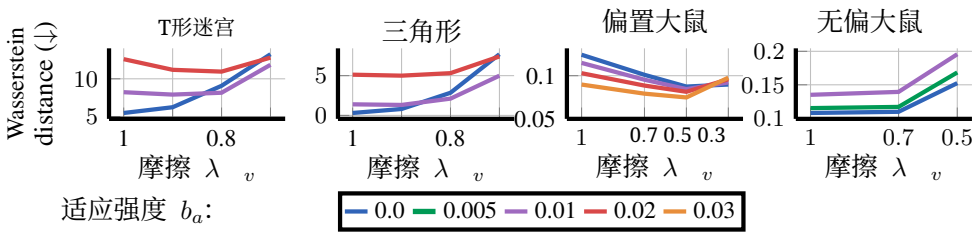

图13：Wasserstein距离（第 A.3.5节）中具有非合理动量的循环神经网络（见第 A.4节）的重放轨迹，与清醒轨迹的比较。具有合理动量和tanh激活的循环神经网络版本见图7 和 17。

F 使用双曲正切激活的结果

在文本的其他所有部分中，我们都使用了分段线性（线性或带泄漏/非泄漏 ReLU）激活函数，相关依据见第 A.1 节。为求完整起见，我们在此重现了研究中的一些关键发现；除循环神经网络现在使用双曲正切激活函数而非分段线性激活函数外，所有超参数均保持不变。我们发现，动量会在时间上压缩重放并增强探索，不过在双曲正切循环神经网络中，适应似乎不会减缓重放。

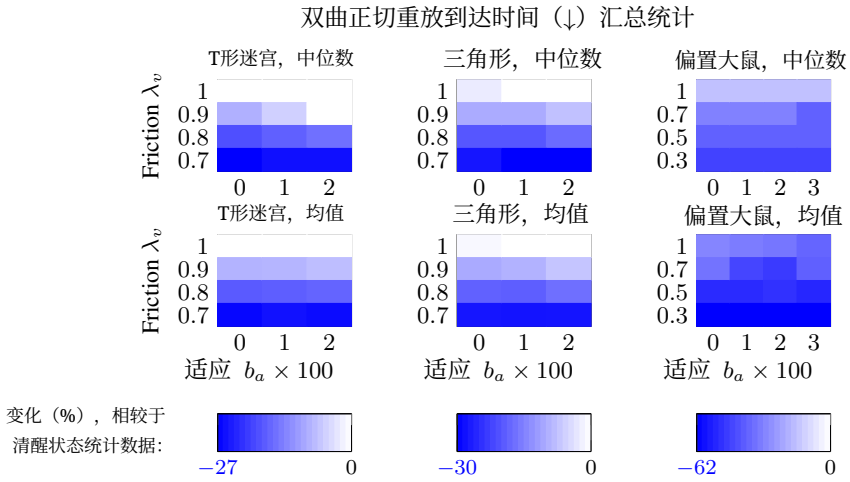

图14：**动量会在时间维度上压缩双曲正切循环神经网络中的重放**。这是图4的另一个版本，但使用的是双曲正切循环神经网络。图10展示了采用非合理动量和 ReLU 循环神经网络的版本。值得注意的是，在双曲正切循环神经网络中，适应并不会显著减慢重放。

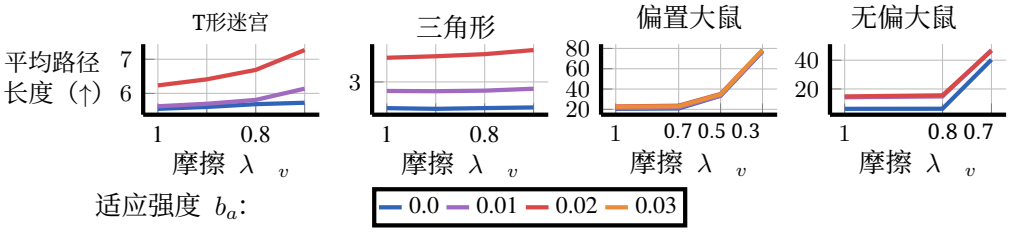

图15：**动量和适应都通过循环神经网络中的回放路径长度增加探索**。这是图5的另一个版本，但使用的是双曲正切循环神经网络。有关非合理动量在ReLU 循环神经网络中的版本，请参见图11。

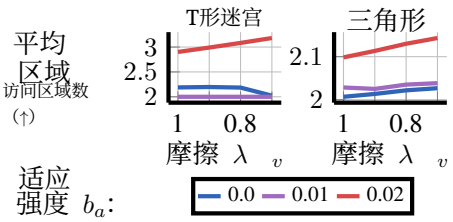

图16：**动量不会通过双曲正切循环神经网络重放中的访问区域减少探索**。这是图6中测量结果的另一个版本，但使用的是双曲正切循环神经网络。有关在 ReLU 循环神经网络中使用非合理动量的版本，请参见图12。

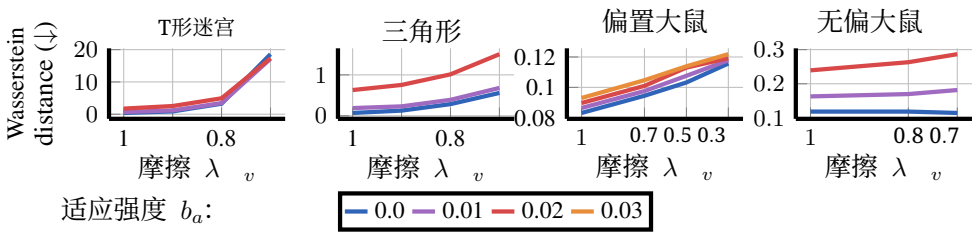

图17：双曲正切循环神经网络中重放轨迹与清醒轨迹相比的 Wasserstein距离（第A.3.5节）。具有 ReLU 激活和非合理动量的循环神经网络版本见图7 和13。