# 水中气泡的非线性振荡赋予音乐创造力

Ivan S. Maksymov*Artificial Intelligence and Cyber Futures Institute, Charles Sturt University, Bathurst, NSW 2795, Australia

创作原曲和改编现有音乐作品是一门艺术，需要多年学习和练习才能掌握。然而，尽管人工智能驱动的音乐创造力领域不断取得进展，创作高质量音乐作品仍是人类的专长。本文展示，当一枚水中气泡在编码了古典音乐乐曲的声压信号作用下发生非线性振荡时，可以用它创作出富有创造性的音乐作品。气泡响应的音频信号听起来像是用电吉他演奏的原曲。我们提出，并给出合理的理论依据，说明气泡的这一特性可用于创建受物理学启发的人工智能系统，从而模拟人类在编曲和音乐创作方面的创造力。

arXiv:2304.00822v1 [cs.SD] 3 Apr 2023

## I. 引言

液体中的气泡是许多重要自然现象的基础 [1–4], , 包括空化 [1] 和流水声 [5, 6]。声压波驱动的气泡振荡也与生物大脑的行为相似，因为大脑 [7]和气泡[2]都是非线性动力系统 [8]。因此，研究振荡气泡或许有助于理解负责声音与音乐感知的某些大脑功能。以下实验性证据支持这一观点。

首先，研究已证明，支撑神经信号传递的电荷脉冲伴随着本质上具有非线性的声机械（声音）波 [9–12]。其次，研究表明，人类的音乐体验由非线性声学过程介导 [13]，而同样的过程也支撑着某些动物的听觉处理能力 [14]。具体而言，在一项实验中，研究人员让猫头鹰聆听一段古典音乐，其中的音调经过处理，基频谐波被刻意去除；猫头鹰的大脑恢复了缺失的基频谐波 [15]。虽然这种行为在非线性物理学领域颇具研究价值 [16], , 但实际上，猫头鹰的大脑将音乐构思从原来的位置移至较低频率，这是一种常见的八度移调示例 [14]。值得注意的是，音乐移调是一项并非轻而易举的任务，通常只有受过相关正规教育的人才能做到；而生物大脑却能自然地完成这类音频处理。第三，众所周知，非线性声学过程支撑着许多乐器的运作 [13] [17] ；音乐家以及许多热爱音乐、却没有相关正规背景的人，都能自然而然地理解非线性效应，尽管他们对非线性物理学知之甚少 [14]。这一事实也表明，生物大脑能够自然地处理非线性声信号。

最近，我们提出，水中的一簇振荡气泡可以充当人工神经网络，表现出复杂的非线性行为，并且能够经过训练，预测许多实际情境中出现的高度非线性混沌时间序列 [18] ，例如金融市场分析、天气预报以及自动驾驶车辆的控制 [19–22]。因此，既然振荡气泡能够高效模拟某些类型的人工神经网络——回声状态网络（ESN） [19] 和液体状态机（LSM）[23]——而这些网络也能复现生物大脑的某些功能，那么可以设想，振荡气泡或许也能复现大脑的一些功能，包括与音乐感知相关的功能。因此，在本研究中，我们提出，振荡气泡的高度非线性行为可用于完成一些需要音乐创造力的任务。

音乐创造力可以定义为一种运用现有音乐知识来产生新颖音乐成果的过程，这些成果可以采取即兴演奏、乐曲和编曲等形式。对于机器学习系统而言，创作高质量的音乐成果仍然是最具挑战性的任务之一[24] ，尽管人工智能驱动的音乐创造力近期已取得重大进展（参见 [25–28] ，仅举数例）。因此，认为一个简单的气泡能够产生具有创造性的音乐输出，这一想法不仅在根本上令人着迷，也可能为人工智能领域带来新的知识。

作为一个代表性示例，我们合成了爱德华·格里格创作的乐曲《在山魔王的宫殿里》的一个简单版本，并将其用作驱动水中单个气泡非线性振荡的声学信号。通过严格的数值模拟，我们证明，气泡产生的输出信号会被感知为原曲的“重金属音乐”翻唱，并加入了电吉他所特有的温暖而粗粝的音色 [29]。

本文结构安排如下。主要研究发现见第 II 节，随后在第 III 节进行讨论。讨论基于对声驱动振荡气泡物理性质的全面理论分析和数值建模

* imaksymov@csu.edu.au

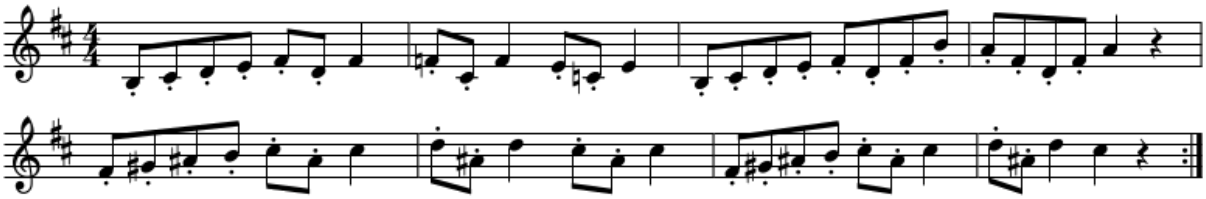

图 1. 音　乐曲中作为驱动气泡振荡的声学信号的音符　　e.

气泡，其结果见第 A 节和第 B 节。由于记忆是创造力的先决条件之一， [30, 31], 我们采用第 C 节所述的回声状态网络算法，在第 D 节证明振荡气泡具有适用于人工智能领域应用的记忆容量。

## II. 结果

我们选用爱德华·格里格《在山魔王的宫殿里》的一个简单钢琴版本，作为驱动气泡振荡的声学信号（图 1）。这首乐曲广为大众熟知，音乐家也常用它来制作自己的录音。依据 8位计算机音乐编曲 的原理 [32], ，我们将旋律中的每个音符编码为一串方波脉冲，其重复频率对应理想化原声钢琴的第 $n$ 个琴键:

$$f(n) = 2^{\frac{n-49}{12}} \times 440\,\mathrm{Hz}\,. \tag{1}$$

旋律采用每小节四拍的拍号，速度为 120 bpm，因此我们也能计算出每小节的秒数。虽然这种方法无法重现钢琴的准确声音，但足以生成易于辨认的乐曲版本（参见补充音频文件 out.in），其格式适合用于本研究采用的数值模型进行处理。

在数值模型中，我们考虑一个单个毫米级气泡。第 B 节讨论了毫米级气泡与单个方波声压脉冲相互作用所依据的基本非线性物理学。这里，我们分析气泡对一系列相同方波脉冲的响应，这些脉冲编码了旋律。我们还将气泡的响应录制为音频文件（见补充文件 out.wav），并讨论所得录音的审美特征。

虽然本节所呈现的结果是通过正方形压力脉冲逐步提高气泡的环境压力而获得的，但使用负方波脉冲时也观察到了气泡类似的物理行为。我们还考察了峰值振幅和持续时间各不相同的脉冲。我们发现，脉冲类型以及其具体持续时间和峰值振幅的选择会影响气泡所产生输出的审美特征。因此，我们选择模型参数的依据并非气泡与声压脉冲之间相互作用的物理学原理，而是实现令人愉悦的审美特征这一目标。

度会影响气泡所产生输出的审美特征。因此，我们选择模型参数的依据并非气泡与声压脉冲之间相互作用的物理学原理，而是实现令人愉悦的审美特征这一目标。

图 2(a) 中绘出了对应于旋律各音符的若干输入方波脉冲（黑色曲线），并将其与气泡相应的声学响应（红色曲线）进行比较。振荡气泡的非线性响应已成为许多理论和实验研究的主题（例如，见 [1, 2, 33–35] ），我们证实，气泡对旋律声学特征的响应也具有高度非线性。此外可以看到，在与各个音符相关的脉冲之间，气泡仍在持续振荡。这一结果表明，振荡气泡具有记忆能力；记忆是适用于回声状态网络（ESN）的非线性处理单元的一项基本属性，也是实现人工创造力的先决条件（见第 D 节）。

具体而言，我们认为，乐曲声音激发气泡产生的持续振荡，相当于电吉他和钢琴等乐器中使用的延音效果，即刻意延长音符的可听共鸣时长[29]。这种行为也可对应于电吉他失真效果的高级数字实现中采用的所谓记忆非线性[36, 37]。在这种实现中，具有记忆的理想化非线性系统可用沃尔泰拉级数进行解析表示，其输出取决于系统在其他所有时刻的输入，从而能够形成渐忘记忆 [38]。

确实，如图 2(b) 所示，在频域中，气泡的非线性表现为频谱中高次谐波的增多——将表示旋律声学特征频谱的黑色曲线与表示气泡响应的红色曲线进行比较即可看出。在电吉他演奏中，高次谐波的出现与模糊、粗粝的音调有关[29, 36, 39]。随后，录制的输出信号（见补充文件 out.wav）在听感上被认为是对原曲的电吉他翻奏。

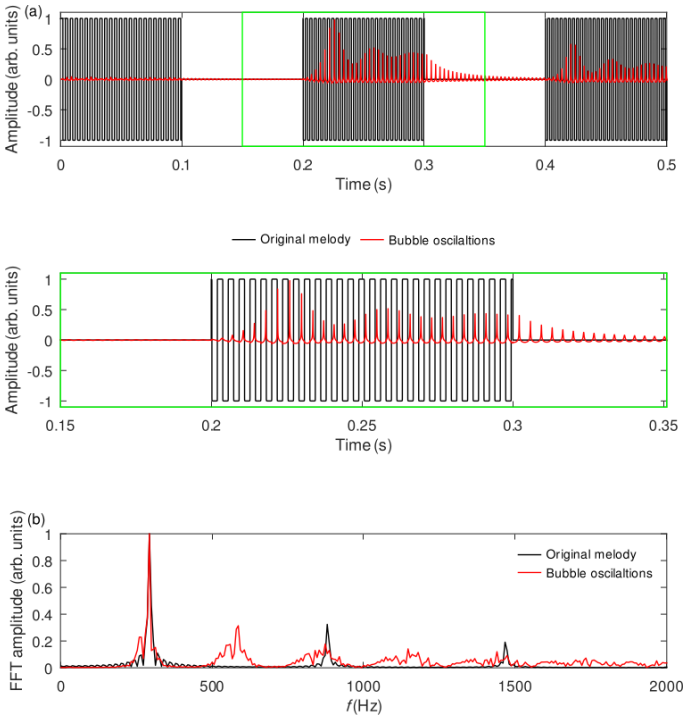


图 2. (a) 输入旋律中三个音符的波形（黑色曲线）以及气泡相应的声学响应（红色曲线）。绿色矩形标示了信号的时间范围；这些信号的放大部分绘于主图下方的插图中。(b) 面板 (a) 插图所示波形的傅里叶功率谱。

为了检验我们的听感判断，我们邀请接受过正规音乐教育的人聆听输出信号，他们确认这段旋律确实与原始旋律的电吉他版本十分相似。我们还使用了 Audacity 音频处理软件，对原始旋律施加标准的静态非线性音频处理效果，以模拟电吉他的效果 [36, 39]。这一流程旨在再现气泡响应的线形和频谱，同时重现气泡对原曲产生的电吉他效果。

数字效果的应用结果如图3所示。我们施加了混响——这一过程既与人工智能中的记忆 [40]有关，也与认知记忆 [41], 有关——以及吉他失真效果；该效果以基频 [29, 39] 的高阶谐波丰富频谱（这种失真效果的非线性是无记忆的 [36, 37]„ 这也证明了在我们的模型中加入混响效果是合理的）。数字生成信号（蓝色曲线）的线形和频谱组成与气泡的响应（红色曲线）相似。由于失真是重金属音乐风格的典型特征 [39], ， 气泡能够再现这种效果，客观地证实了它具备创造性处理音乐的能力。

## III. 讨论

因此，我们表明，单个振荡气泡执行复杂非线性任务的能力，使其能够像人类一样编排现有音乐作品。利用液体特性进行创造性音乐创作，是一种成熟的音乐风格 [42], ； 气泡被困在液体中所产生的非线性声学效应 [5, 6] 很可能早已在某种形式上得到利用。

然而，我们从另一个角度考察气泡的非线性特性。在先前的工作中，我们展示了振荡气泡能够像人工神经网络一样预测混沌时间序列 [18], ， 而这项任务不仅需要非线性特性，也需要记忆。最重要的是，与水中的气泡产生声音、再由艺术家利用这些声音创作音乐不同，在人工神经网络中，气泡发挥着主动作用，充当模拟数据处理单元，模拟生物神经元的运行。尽管与典型的数字计算机相比，这种单元的记忆容量和数据处理速度较低，但已有研究表明，在解决人工智能领域中的某些特定问题时，模拟计算机系统可能比数字计算机更高效 [18, 22, 43, 44]。鉴于此，我们认为，单个气泡可用作模拟人工智能系统的构建模块，在无需或仅需少量人工输入的情况下生成音乐输出。

不同于音乐转录——将为一种乐器创作的乐曲逐音符准确地改编为由另一种乐器演奏（即从钢琴改为吉他）——编曲是一种更具创造性的过程，其中音乐风格会发生变化，并加入新的复杂音调。因此，这项创造性任务对人工智能系统而言尤其具有挑战性，因为它要求机器具备人类智能的一些关键特征，例如联想思想、感知、思考、寻找答案以及批判自身工作成果的能力 [45]。然而，创造力依赖于认知记忆 [30, 31] ，并且与文化背景和个性密切相关，同时还受到艺术家的动机和情绪影响 [45]。有趣的是，欣赏、编曲和创作重金属音乐的能力也被认为与较高的智力相关 [46], ， 这意味着重金属风格音乐的制作对于人工智能而言应当是一项尤其具有挑战性的任务。

从机器学习系统的角度出发，研究人员尝试创建模型，将实现人工创造力的训练输入表示为受扰动和噪声影响、定义模糊的数据集 [47, 48]。对这类模型的研究表明，实现人工创造力可能与训练人工神经网络的标准方法相矛盾，因为与创造力相关的扰动会干扰网络的运行，例如改变其连接权重的值

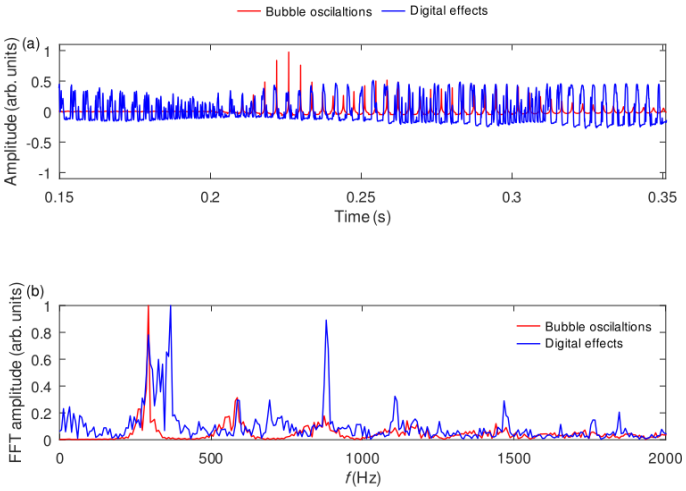


图3。（a）比较气泡的时间响应（红色曲线）与通过对原始旋律进行数字处理、施加混响和非线性声学效应后得到的信号（蓝色曲线）。（b）面板（a）中各信号的傅里叶谱。采用与图2（a）相同的时间尺度。为清晰起见，未显示原始旋律的波形。

[47]。随后，有人提出应改变传统神经网络的组织原则，以实现人工智能驱动的创造力 [48]。

然而，这种对传统人工神经网络在人工智能创造力领域适用性的评估，并未考虑到神经网络模拟（非数字）形式的最新发展。在这些形式中，具有非线性动力学行为的硬件和真实物理系统被用作人工神经元。尽管迄今为止，这类物理计算系统模仿的只是某些数字神经网络架构的运行方式，[22], 但已有研究表明，在旨在模拟生物大脑功能的运算中，它们有望超越计算机程序的能力 [18, 22, 49]。本文的研究结果有助于证明这一潜力。

最后，假设人工智能系统能够创造性地生成可与人类创作的音乐相媲美的音乐输出，研究人员还面临另一个问题：人工智能生成的音乐输出难以评估，因为这一过程依赖于受训听众的鉴赏能力 [25], ，而他们可能持有文化偏见 [50], 或对计算机创作的音乐抱有偏见 [51]。

由于很难对气泡创作的电吉他编曲进行审美质量的批判性评判，我们找到了几段由专业电吉他手演奏的《在山魔王的宫殿里》翻奏版本 [52, 53]。尽管这些乐曲听起来比气泡创作的版本更悦耳，但我们仍能分辨出专业音乐家在演奏中创造的那种温暖而粗粝的音色。因此，我们把判断气泡输出质量的任务留给读者，请大家聆听引用的乐曲并进行比较。不过，我们注意到，人类需要多年练习才能掌握电吉他，而水中的单个气泡似乎天生就具备重现这种乐器声音的能力。

## IV. 结论

通过对水中声驱动单气泡的非线性振荡进行严格的数值建模，我们证明，气泡能够产生听起来在审美上如同人类创作的音乐成果般富有创造力的音频输出。鉴于以往研究已证明，振荡气泡可以形成一种基于物理原理的人工神经网络，能够模拟生物大脑的某些功能，我们认为，单个振荡气泡或由此类气泡组成的网络（簇）都可以用作一种装置，来重现某些形式的人工创造力。迄今为止，如何在音乐中实现创造力一直是现代人工智能系统面临的难题。因此，我们相信，这些发现有助于推动这一重要基础与应用研究领域的进一步发展，也将引起那些探索液体声学性质的艺术家的兴趣。

## 致谢

ISM 感谢 Sergey Suslov 教授和 Andrey Pototsky 博士（斯威本科技大学）提供宝贵的讨论，并感谢 Mikhail Kostylev 教授（西澳大利亚大学）协助计算记忆容量。

[1] C. E. Brennen, *Cavitation and Bubble Dynamics*（牛津大学出版社，纽约，1995）。[2] W. Lauterborn 和 T. Kurz，《气泡振荡的物理学》，《物理学进展报告》**73**，106501（2010）。[3] I. S. Maksymov、B. Q. H. Nguyen、A. Pototsky 和 S. A. Suslov，声学、声子学、布里渊光散射和基于法拉第波的频率梳：物理基础与应用，《传感器》**22**, 3921 (2022年)。[4] I. S. Maksymov、B. Q. H. Nguyen 和 S. A. Suslov，利用液体中的气泡振荡进行生物力学传感及相关技术：理论与实际应用，《生物传感器》**12**, 624 (2022年)。

[5] M. Minnaert，关于音乐气泡和流水声，《哲学杂志》16, 235 (1933)。[6] T. G. Leighton 和 A. J. Walton，液体中气泡发声的实验研究，《欧洲物理学杂志》8, 98 (1987)。[7] T. M. McKenna、T. A. McMullen 和 M. F. Shlesinger，大脑作为动态物理系统，《神经科学》60, 587 (1994)。[8] S. H. Strogatz, *Nonlinear Dynamics and Chaos With Applications to Physics Biology Chemistry and Engineering* ，，，2019（CRC出版社，）。[9] T. Heimburg 和 A. D. Jackson，生物膜和神经中的孤子传播，《美国国家科学院院刊》102, 9790 (2005)。[10] A. Gonzalez-Perez、R. Budvytyte、L. D. Mosgaard、S. Nissen 和 T. Heimburg，无脊椎动物中线巨轴突和侧巨轴突在碰撞期间动作电位的穿透，《物理评论X》4, 031047 (2014)。[11] A. E. Hady 和 B. B. Machta，生物膜和神经中的孤子传播，《自然·通讯》6, 6697 (2015)。[12] I. S. Maksymov 和 A. Pototsky，振动的活蚯蚓中类法拉第体波的激发，《科学报告》10, 8564 (2020)。[13] I. S. Maksymov 和 A. D. Greentree，光与声的耦合：振荡气泡和液滴产生的巨大非线性效应，《纳米光子学》8, 367 (2019)。[14] D. J. Levitin, *This is Your Brain on Music: The Science of Human Obsession* 2006（达顿出版社，）。[15] P. Janata, *Electrophysiological studies of auditory contexts* 1996（博士论文，俄勒冈大学，）。[16] I. Maksymov 和 A. Pototsky，振动液滴表面的谐波与次谐波波，《物理评论E》100, 053106 (2019)。[17] N. H. Fletcher，音乐管乐器的非线性理论，《应用声学》30, 85 (1990)。[18] I. S. Maksymov、A. Pototsky 和 S. A. Suslov，利用水中气泡振荡的神经回声状态网络，《物理评论E》105, 044206 (2021)。[19] M. Luko˘sevi˘cius 和 H. Jaeger，循环神经网络训练的储备池计算方法，《计算机科学评论》3, 127 (2009)。[20] K. Nakajima，物理储备池计算——入门视角，《日本应用物理学杂志》59, 060501 (2020)。[21] G. Tanaka、T. Yamane、J. B. H´eroux、R. Nakane、N. Kanazawa、S. Takeda、H. Numata、D. Nakano 和 A. Hirose，物理储备池计算的最新进展：综述，《神经网络》115, 100 (2019)。[22] K. Nakajima 和 I. Fisher, *Reservoir Computing*（施普林格，柏林，2021）。[23] W. Maass、T. Natschl¨ager 和 H. Markram，无需稳定状态的实时计算：一种基于扰动的神经计算新框架，《神经计算》14, 2531 (2002)。[24] A. I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity* 2019（麻省理工学院出版社，）。[25] F. Pachet，音乐演奏技巧与创造力，载于*Computers and Creativity* ，J. McCormack 和 M. d'Inverno 编（施普林格柏林海德堡，柏林，海德堡，2012年），第115–146页。[26] A. Huang 和 R. Wu，音乐深度学习，arXiv:1606.04930 (2016)。[27] J.-P. Briot，从人工神经网络到用于音乐生成的深度学习：历史、概念与趋势，《神经计算与应用》33, 39 (2021)。[28] C. Hernandez-Olivan 和 J. R. Beltr´an，《深度学习与音乐创作：综述》，载于 *Advances in Speech and Music Technology: Computational Aspects and Applications* ，由 A. Biswas 、 E. Wennekes 、 A. Wieczorkowska 和 R. H. Laskar 编（施普林格国际出版公司，钱，2023年），第 25–50 页。[29] M. Ross, *Getting Great Guitar Sounds*（Hal Leonard 公司，密尔沃基，1998年）。[30] M. Benedek、R. E. Beaty、D. L. Schacter 和 Y. N. Kenett，《记忆在创造性构思中的作用》，《自然评论：心理学》。https://doi.org/10.1038/s44159-023- 00158-z（2023年）。[31] C. R. Gerver、J. W. Griffin、N. Dennis 和 R. Beaty，《记忆与创造力：记忆系统与创造性认知关系的荟萃分析考察》，PhysArXiv（2022年）。[32] J. Waugh, *Commodore 64 Music: Making Music with Your Micro* 1985（阳光图书公司，）。[33] A. Prosperetti，《液体中气泡的非线性振荡：稳态解》，J. Acoust. Soc. Am. 56, 878 1974年( )。[34] S. A. Suslov、A. Ooi 和 R. Manasseh，《刚性壁面附近微气泡的非线性动态行为》，《物理评论E》85, 066309（2012年）。[35] B. Q. H. Nguyen、I. S. Maksymov 和 S. A. Suslov，《利用液体中气泡簇振荡产生声学频率梳：概念验证》，Sci. Reps. 11, 38（2021年）。[36] D. T. Yeh、J. S. Abel、A. Vladimirescu 和 J. O. Smith，《吉他失真电路的仿真》，Comput. Music J. 32, 23（2008年）。[37] D. T. Yeh, *Digital Implementation of Musical Distortion Circuits by Analysis and Simulation*（博士论文，斯坦福大学，2009年）。[38] S. Boyd 和 L. O. Chua，《渐忘记忆与用沃尔泰拉级数近似非线性算子的问题》，IEEE Trans. Circuits Syst. CAS-32, 1150（1985年）。[39] U. Z¨olzer, *DAFX: Digital Audio Effects*（威利出版社，2011年）。[40] K. G. Kirby，《神经序列学习中的上下文动态》，佛罗里达人工智能研究研讨会（FLAIRS）论文集，66（1991年）。[41] S. Ribeiro 和 M. A. L. Nicolelis，《睡眠期间记忆的混响、存储与突触后传播》，Learn. Mem. 11, 686（2004年）。[42] J. Grant、J. Matthias 和 D. Prior, *The Oxford Handbook of Sound Art* 2021（牛津大学出版社，）。[43] W. Cao、Y. Zhao、A. Boloor、Y. Han、X. Zhang 和 L. Jiang，《神经处理器内存一体化：通过外围设备的神经近似实现高效的存内处理》，IEEE Trans. Comput. 71, 2142（2022年）。[44] D. Ivanov、A. Chezhegov、M. Kiselev、A. Grunin 和 D. Larionov，《神经形态人工智能系统》，Front. Neurosci. 16, 10.3389/fnins.2022.959626（2022年）。[45] M. A. Boden，《创造力与人工智能》，Artif. Intell. 103, 347（1998年）。[46] S. M. Cadwallader, *The Darker Side of Bright Students: Gifted and Talented Heavy Metal Fans*（英国国家资优青少年学院、华威大学，英国，2007年）。[47] S. L. Thaler，《突触扰动神经系统中的模式更替》，Procedia Comput. Sci. 88, 21（2016年）。

[48] A. Oleinik, 《神经网络不擅长什么？论人工创造力》, Big Data Soc. 6, 1（2019年）。[49] M. Cucchi、S. Abreu、G. Ciccone、D. Brunner 和 H. Kleemann, 《储备池计算实践指南：动手实现教程》, Neuromorph. Comput. Eng. 2, 032002（2022年）。[50] S. M. Demorest、S. J. Morrison、V. Q. Nguyen 和 E. N. Bodnar, 《语境线索对音乐记忆中文化偏见的影响》, Music Percept. 33, 590（2016年）。[51] P. Pasquier、A. Burnett 和 J. B. Maxwell, 《探究听众对音乐元创造力的偏见》, 载于 *International Conference on Innovative Computing and Cloud Computing* 2016 （）。

[52] A. Parker, *Hall of the Mountain King invented heavy metal*, 可在 https://www.youtube.com/watch?v=lFuR7rq598 查看。[53] M. Klett, *Hall of the Mountain King*, 可在 https: //www.youtube.com/watch?v=G4C9O7HA9os 查看。[54] J. B. Keller 和 M. Miksis, 《大振幅气泡振荡》, J. Acoust. Soc. Am. 68, 628（1980年）。[55] A. E. Mudrov, *Numerical Methods for Personal Comput-ers in Basic, Fortran, and Pascal (in Russian)*（Rasko, 托木斯克, 1990年）。[56] N. A. Pelekasis 和 J. A. Tsamopoulos, 《两个气泡之间的别克涅斯力。第一部分：对压力阶跃变化的响应》, J. Fluid Mech. 254, 467（1993年）。[57] I. Akhatov、N. Gumerov、C. D. Ohl、U. Parlitz 和 W. Lauterborn, 《表面张力在稳定单泡声致发光中的作用》, Phys. Rev. Lett. 78, 227（1997年）。[58] H. Jaeger, *A tutorial on training recurrent neural net-works, covering BPPT, RTRL, EKF and the "echo state network" approach*（GMD 报告 159, 德国国家信息技术研究中心, 2005年）。[59] N. Bertschinger 和 T. Natschl¨ager, 《循环神经网络在混沌边缘的实时计算》, Neural Comput. 16, 1413（2004年）。[60] S. Watt 和 M. Kostylev, 《基于钇铁石榴石薄膜的自旋波延迟线有源环形谐振器储备池计算》, Phys. Rep. Appl. 13, 034057（2020年）。

**附录 A：气泡振荡数值模型**

由声压波驱动的单个气泡非线性振荡通常采用 Keller-Miksis（KM）方程 [54] 描述，该方程考虑了黏性耗散和流体可压缩性导致的气泡振荡衰减：

$$\left(1-\frac{\dot{R}}{c}\right)R\ddot{R}+\frac{\dot{R}^2}{2}\left(3-\frac{\dot{R}}{c}\right)=\frac{1}{\rho}\left[1+\frac{\dot{R}}{c}+\frac{R}{c}\frac{d}{dt}\right]\left[P(R,\dot{R})-P_\infty(t)\right], \quad \text{(A1)}$$

其中，上方的点表示对时间求导 $t$和

$$P(R,\dot{R})=\left(P_0-P_v+\frac{2\sigma}{R_0}\right)\left(\frac{R_0}{R}\right)^{3\kappa}-\frac{4\mu\dot{R}}{R}-\frac{2\sigma}{R}. \quad \text{(A2)}$$

表达式 $P_\infty(t)=P_0-P_v+\alpha P_a(t)$ 中， $P_a(t)$ 是任意的驱动声压信号，表示远离气泡处液体中的时变压力。参数 $R_0$, $R(t)$, $\mu$, $\rho$, $\kappa$, $\sigma$, $c$ 和 $\alpha$ 分别表示气泡的平衡半径和瞬时半径、液体的动力黏度和密度、气泡内气体的多变指数、气液界面的表面张力、液体中的声速，以及驱动超声波的振幅。忽略气体通过气泡表面的扩散。

当气泡振荡不受流体可压缩性影响时（本研究即为此情况），气泡在远场区散射的声功率为 [1]

$$P_{scat}(R,t)=\frac{\rho R}{h}\left(R\ddot{R}+2\dot{R}^2\right), \quad \text{(A3)}$$

其中， $h$ 远大于气泡的平衡半径。

由于气泡的振荡由包含方波脉冲的信号激发（见第 II 节），我们引入脉冲重复频率参数 $f_p$ ，并将公式（A1）–（A2）改写为无量纲形式 [34, 35]。其中，我们分别采用气泡的平衡半径 $R_0$ 和 ${\omega_p}^{-1}$ （$\omega_p=2\pi f_p$）作为长度和时间尺度，同时引入无量纲量 $r=R$（$t$）、$/R_0$ 和 $\tau=\omega_p t$。由此得到

$$\begin{aligned} r''[(1-\Omega r')r+\Omega\mathcal{R}] &= (\Omega r'-3)\frac{r'^2}{2}-\frac{\mathcal{W}+\mathcal{R}r'}{r} \\ &+(\mathcal{M}+\mathcal{W})\frac{[1+(1-3\kappa)\Omega r']}{r^{3\kappa}} \\ &-(1+\Omega r')(\mathcal{M}+\mathcal{M}_e P_a(\tau))-\mathcal{M}_e\Omega r {P_a}'(\tau), \end{aligned} \quad \text{(A4)}$$

其中，撇号表示对 $\tau$ 求导，$\Omega=\frac{\omega_p R_0}{c}$, $\mathcal{R}=\frac{4\mu}{\rho\omega_p R_0^2}$, $\mathcal{W}=\frac{2\sigma}{\rho{\omega_p}^2R_0^3}$, $\mathcal{M}=\frac{P_0-P_v}{\rho{\omega_p}^2R_0^2}$, $\mathcal{M}_e=\frac{\alpha}{\rho{\omega_p}^2R_0^2}$ 和 $\mathcal{K}=3\kappa$。参数 $\Omega$ 表征气泡大小相对于声波波长的比例，$\mathcal{M}$表征气体的弹性性质及其可压缩性， $\mathcal{W}$ 和 $\mathcal{R}$ 可视为韦伯数和雷诺数的倒数，分别表示表面张力和黏性耗散的影响，而 $\mathcal{M}_e$ 则衡量声驱动 [34]。

计算采用以下对应于 20℃ 水的流体参数： $c=1484$ m/s、 $\mu=10^{-3}$ kg m/s、 $\sigma=7.25\times10^{-2}$ N/m、$\rho=10^3$ kg/m$^3$ 以及 $P_v=2330$Pa。我们将静止气泡中的空气压力设为 $P_0=10^5$Pa，并将空气的多变指数设为 $\kappa=4/3$。在计算中，脉冲重复频率 $f_p$ 取决于乐曲的音符，且始终低于毫米级气泡的固有振荡频率 $f_0$ （$f_0R_0\approx3.26\,\mathrm{m\,s^{-1}}$ [2]，对于 $R_0=1$mm 的气泡为 $f_0=3260$Hz）。方程（A4）采用定步长四阶龙格-库塔法进行数值求解，该方法由定制子程序 rk4 [55] 实现；此子程序从 Pascal 移植到 Oberon-07 编程语言。

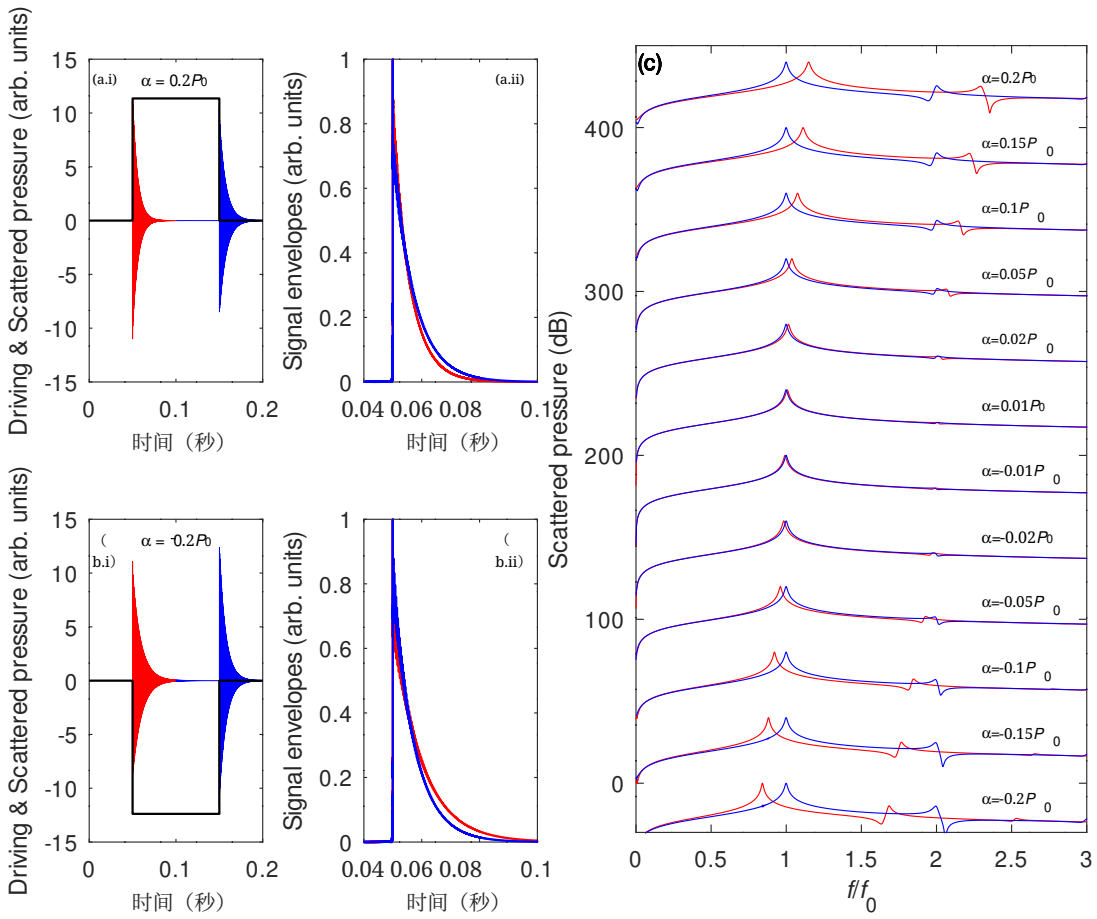


图 4. 水中单个毫米级气泡对 (a.i) 正压力脉冲和 (b.i) 负压力脉冲的声学响应。面板 (a.ii) 和 (b.ii) 分别绘出了气泡对声压阶跃变化的响应包络，并将其归一化为 1。(c) 振荡气泡在不同峰值振幅 $\alpha$ 的方波脉冲驱动下所散射压力的频谱。红色（蓝色）曲线表示驱动方波脉冲压力的第一次（第二次）阶跃变化所引起的气泡振荡频谱。由于第二次阶跃变化，气泡总会回到其平衡状态，且其基振荡频率保持不变。

我们令 $R_0 = 1\text{mm}$、 $\alpha = 0.2P_0$ 和 $f_p = 100\text{Hz}$，计算无量纲参数 $\mathcal{M}$、 $\mathcal{W}$、 $\mathcal{R}$ 和 $\Omega$ 的典型值：

$$\mathcal{R} = \frac{\mathcal{R}_0}{\Omega^2}, \mathcal{W} = \frac{\mathcal{W}_0}{\Omega^3}, \mathcal{M} = \frac{\mathcal{M}_0}{\Omega^2}, \mathcal{M}_e = \frac{\mathcal{M}_{e0}}{\Omega^2}, \quad \text{(A5)}$$

其中

$$\begin{aligned} \mathcal{R}_0 &= \frac{4\mu\omega_p}{\rho c^2} \approx 1.14 \times 10^{-9} \\ \mathcal{W}_0 &= \frac{2\sigma\omega_p}{\rho c^3} \approx 2.79 \times 10^{-11} \\ \mathcal{M}_0 &= \frac{P_0 - P_v}{\rho c^2} \approx 4.43 \times 10^{-5} \\ \mathcal{M}_{e0} &= \frac{\alpha}{\rho c^2} \approx 9.08 \times 10^{-5}, \end{aligned} \quad \text{(A6)}$$

对于 $\Omega = 4.54 \times 10^{-6}$。在这些参数下， $\mathcal{M}$的值比 $\mathcal{W}$ 和 $\mathcal{R}$ 高出几个数量级，这体现了一个众所周知的事实：大气泡的动力学主要由气体弹性决定，而黏性耗散和表面张力的作用则居于次要地位 [34]。

**附录 B：单位阶跃声压变化驱动的气泡**

此前曾在两个振荡气泡相互作用的背景下研究单个气泡对驱动声压单位阶跃变化的响应 [56]。本文则从声音处理的角度进行研究，因为用于激发气泡振荡的乐曲是通过一系列方波脉冲编码的（见第 II 节）。

我们将驱动声学信号定义为 $P_a(\tau) = \Theta(\tau)$,

其中 $\Theta(\tau)$ 是赫维赛德函数

$$\Theta(\tau)=\begin{cases}0 & \text{for } \tau<0\\ 1 & \text{for } \tau>0\,.\end{cases} \tag{B1}$$

我们假设无量纲驱动声音压强振幅 $\mathcal{M}_e$ 很小，且气泡在其平衡状态附近振荡，因此其瞬时无量纲半径为 $r(\tau)=1+r_1(\tau)$，其中 $r_1(\tau)\ll 1$。我们在 $r=1$ 附近对方程 (A4) 进行线性化，得到下列受恒定激励的阻尼线性谐振子方程

$$(\Omega\mathcal{R}+1){r_1}''+(\mathcal{R}+\mathcal{M}_e\Omega+\mathcal{A}\Omega){r_1}'+\mathcal{A}r_1=-\mathcal{M}_e\,, \tag{B2}$$

where $\mathcal{A}\equiv(3\mathcal{W}+3\mathcal{M})\kappa-\mathcal{W}$. Then we denote $Q=\dfrac{\Omega\mathcal{R}+1}{\mathcal{R}+\mathcal{M}_e\Omega+\mathcal{A}\Omega}\approx\dfrac{\Omega}{\mathcal{K}\mathcal{M}_0}$, $\Lambda=\dfrac{\mathcal{M}_e}{\Omega\mathcal{R}+1}\approx\dfrac{\mathcal{M}_{e0}}{\Omega^2}$ and $\omega_0=\sqrt{\dfrac{\mathcal{A}}{\Omega\mathcal{R}+1}}\approx\sqrt{\mathcal{K}\mathcal{M}}=\dfrac{\sqrt{\mathcal{K}\mathcal{M}_0}}{\Omega}$, where we used the fact that $\mathcal{M}\gg\mathcal{W},\mathcal{R}$ established in Sec. A, and we recast Eq. (B2) as an equation of an unforced harmonic oscillator

$${r_1}''+\frac{1}{Q}{r_1}'+\omega_0^2\left(r_1+\frac{\Lambda}{\omega_0^2}\right)=0\,, \tag{B3}$$

其中，参数 $Q$ 是品质因数，用于描述谐振子的共振行为；$\omega_0$ 对应于气泡的无阻尼固有频率。在初始条件 $r_1(0^-)={r_1}'(0^-)=0$ 下，方程 (B3) 的解为

$$r_1=e^{-\frac{\tau}{\tau_0}}\left(\frac{\Lambda\sin\omega'\tau}{2Q\omega_0^2\omega'}+\frac{\Lambda\cos\omega'\tau}{\omega_0^2}\right)-\frac{\Lambda}{\omega_0^2}\,, \tag{B4}$$

其中，$\omega'=\sqrt{\omega_0^2-\dfrac{1}{4Q^2}}$ 是气泡的阻尼固有频率，$\tau_0=2Q\approx 2\Omega\sqrt{\mathcal{K}\mathcal{M}_0}$ 是特征弛豫时间；在此时间内，瞬态状态下振荡的幅度会减小至原来的 $e$。

因此，除瞬态响应外，解还包含一个与时间无关的响应 $-\Lambda/\omega_0^2$。因此可以看出，在恒定激励下，气泡在变化后的平衡状态 $1+r_1(\tau\to\infty)=1-\frac{\Lambda}{\omega_0^2}$ 附近振荡。根据所采用的模型，这意味着气泡的平衡半径减小了。这与恒定激励实际上提高了气泡外部静压这一事实相符。因此，气泡的振荡频率升高，因为 $\omega_0\sim\Omega^{-1}$，而弛豫时间缩短，因为 $\tau_0\sim\Omega$，这也解释了较小气泡具有更高刚度的事实 [57]（请注意，参数 Ω 表征气泡的平衡半径）。

通过对式 (A4) 进行直接数值求解，采用单个方波脉冲驱动项 $P_a(\tau)$（定义如下），证实了这一行为。

$$P_a(\tau)=\begin{cases}A & \text{for } \dfrac{-\tau_p}{2}\le\tau\le\dfrac{\tau_p}{2}\\ 0 & \text{otherwise}\,,\end{cases} \tag{B5}$$

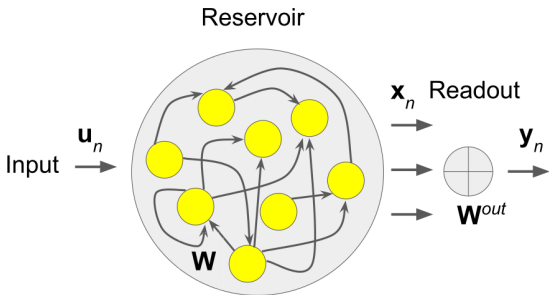


图 5。（a）ESN算法示意图。

其中，脉冲持续时间为 $\tau_p>\tau_0$，脉冲的单位幅值 $A$ 可以是 $-1$ 或 1，从而能够考虑正、负方波脉冲对气泡的驱动效应；驱动强度由参数 $\mathcal{M}_e$ 定义，它是量纲压力幅值 $\alpha$ 的函数。

图 4 展示了气泡对正方波脉冲 [图（a.i）] 和负方波脉冲 [图（b.i）]的响应（根据式（A3），该响应由气泡散射的声压给出）。红色曲线表示气泡受到阶跃压力变化时的响应，蓝色曲线表示气泡恢复到原始平衡状态后的响应。图（a.i）中的压力幅值为 $\alpha=0.2P_0$，图（b.i）中的压力幅值为 $\alpha=-0.2P_0$。图（c）展示了气泡在 $\alpha=0.2P_0$（上方曲线）和 $\alpha=-0.2P_0$（下方曲线）处的振荡频谱，以及在 $-0.2P_0<\alpha<0.2P_0$ 范围内若干取值下计算得到的频谱。

气泡的振荡频率由图（c）中频谱主要峰值的频率给出。与线性化的 KM 方程 (B2) 的分析一致，气泡在其原始平衡状态附近振荡时（蓝色曲线），其频率保持不变；而受声压正、负阶跃变化驱动的气泡（红色曲线），其频率则分别升高和降低。在图 (a.ii) 和 (b.ii) 中，我们还绘出了气泡瞬态响应的包络线，并观察到弛豫时间的变化与理论预测一致。在图（c）中，我们观察到数值模拟预测，由于非线性，气泡自然振荡频率的二次谐波会出现（三次和四次谐波也已被识别，但未在图中显示）。比较基频和二次谐波频率峰在 $\alpha=\pm0.2P_0$ 处的线形，也可以观察到相应的弛豫时间变化。

## 附录 C：回声状态网络

为了支持我们关于气泡能否作为人工神经元运行的讨论，我们采用标准的回声

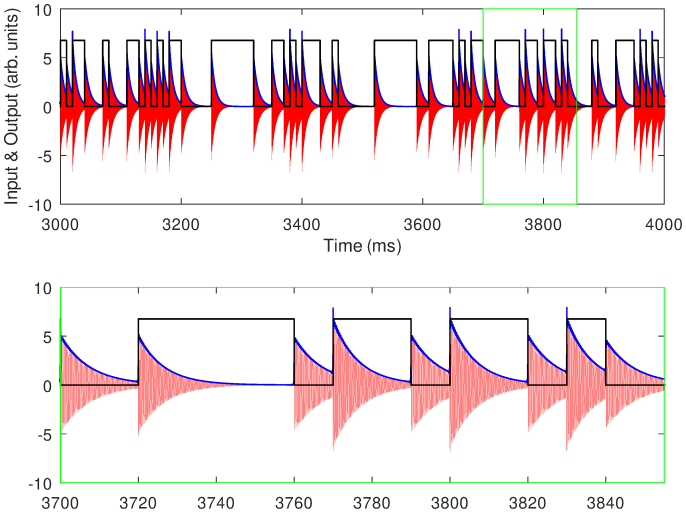


图 6. 单个振荡气泡（红色曲线）对一系列方形压力脉冲的声学响应；这些脉冲编码了正文中讨论的随机生成的二进制时间序列（黑色曲线）。蓝色曲线描绘了气泡响应的包络。上图中，绿色矩形标出了下图所绘信号放大部分对应的时间范围。

状态网络（ESN）算法 [19, 58]。我们选择 ESN，是因为与之相关的液体状态机（LSM）概念 [23], 以及这种神经网络，都可用于模拟生物大脑的某些功能 [49, 59]。

在回声状态网络中，人工神经元构成一个复杂的网络，称为储备池（图 5）。储备池的动态由以下非线性更新方程控制

$$\mathbf{x}_n = (1-\alpha_r)\mathbf{x}_{n-1} + \alpha_r \tanh(\mathbf{W}^{in}\mathbf{u}_n + \mathbf{W}\mathbf{x}_{n-1}) \quad (C1)$$

其中，$n$ 表示等间隔离散时刻 $t_n$的索引，$\mathbf{u}_n$ 是输入值向量，$\mathbf{x}_n$ 是储备池神经元激活值向量，算子 tanh(·) 是神经元的 S 形激活函数 [19], $\mathbf{W}^{in}$ 是包含 $N_x \times N_u$个元素的输入矩阵，$\mathbf{W}$ 是由$N_x \times N_x$ 个元素构成的循环权重矩阵，而 $\alpha_r \in (0, 1]$ 是控制储备池时间动态的参数 [19]。

训练储备池时，通常通过求解线性方程组 $\mathbf{Y}^{target} = \mathbf{W}^{out}\mathbf{X}$ 来计算输出权重 $\mathbf{W}^{out}$，其中 X 和 $\mathbf{Y}^{target}$ 分别是状态矩阵和目标矩阵，它们以 $\mathbf{x}_n$ 和目标输出向量 $\mathbf{y}_n^{target}$作为列，按每个离散时间点$t_n$构建。通常得到的解为 $\mathbf{W}^{out} = \mathbf{Y}^{target}\mathbf{X}^\top(\mathbf{X}\mathbf{X}^\top + \beta\mathbf{I})^{-1}$，其中 I 是单位矩阵，$\beta \geq 0$ 是正则化系数，$\mathbf{X}^\top$ 是 X [19]的转置。在预测阶段，对于新的输入数据 $\mathbf{u}_n$，通过求解方程（C1）并使用常数偏置和拼接 $[\mathbf{u}_n; \mathbf{x}_n]$，来计算输出向量 $\mathbf{y}_n = \mathbf{W}^{out}[1; \mathbf{u}_n; \mathbf{x}_n]$。

**附录 D：振荡气泡的记忆容量**

我们证明，非线性振荡的气泡能够保留过去声学激励输入的记忆。我们采用了用于测试回声状态网络 [58, 59]记忆容量的方法。我们用随机生成的时间序列激发单个气泡振荡，该序列包含只能取“0”或“1”值的二进制元素。单个“1”值表示为一个单位长度的正压力方波脉冲，而“0”值则对应一段没有声学激励的时段，其持续时间同样为一个单位长度。例如，输入数据子集“10101”表示为三个单位长度的方波脉冲，中间由两个单位长度的无激励时段隔开。因此，序列“111001111”对应一个三连方波脉冲，之后是两个连续的无激励时段，再之后是一个四连方波脉冲。单个“1”状态对应的单个方波脉冲的实际单位长度根据经验选取，使其与气泡的弛豫时间 $\tau_0$ 处于同一量级。这样可以确保气泡对一个脉冲的响应在接收到随机时间序列中的下一个脉冲之前不会完全衰减，从而使气泡能够记住过去的输入。

按照物理储备池计算的概念，人工神经网络的非线性动力学行为由物理非线性动力系统的响应取代 [20–22, 49], ，我们将式（C1）替换为式（A4）。这意味着人工神经激励$\mathbf{x}_n$ 成为振荡气泡声学响应的函数。然而，与参考文献 [18], 不同，在该文献中，回声状态网络利用水中振荡气泡簇的非线性动力学响应实现，簇中的各个气泡分别充当人工神经元；在本研究中，我们采用单个振荡气泡作为储备池，并对气泡的时间响应进行采样，从而获得 20 个虚拟人工神经元，用以生成$\mathbf{x}_n$的取值（相关讨论见 [60]）。

我们随机生成一个由 2000 个二进制“0”和“1”值组成的时间序列。我们分别将每个二进制值编码为方波脉冲和它们之间无激励间隔。方波脉冲的振幅为 $\alpha = 0.1P_0$。在回声状态网络的框架下，这种编码后的信号起到输入向量$\mathbf{u}_n$的作用。

我们计算气泡对整个输入信号的响应，然后对气泡的响应进行采样，获得对应于 20 个虚拟神经元的神经激活$\mathbf{x}_n$。随后，我们将输入的前半部分及气泡相应响应的前半部分用作训练数据，并按照第 C 节所述的流程计算$\mathbf{W}^{out}$。这些数据集的后半部分用于进行记忆测试。

图 6 展示了气泡对输入时间序列的典型响应，该时间序列由方波脉冲及其间隔表示。每个方波脉冲后都跟着一段无

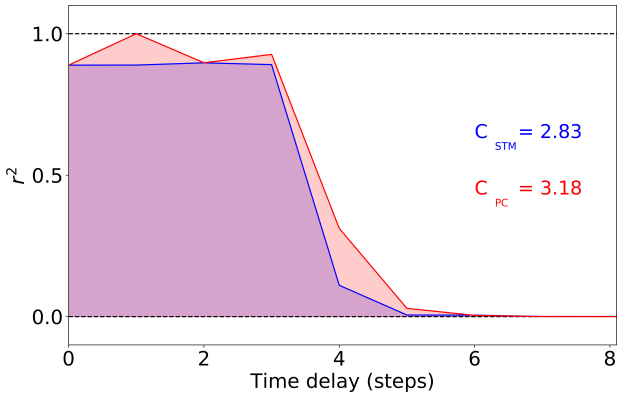


图 7. 振荡气泡的短期记忆（STM）和奇偶校验（PC）任务记忆容量，以针对不同离散时间延迟计算的相关系数平方表示$r^2$ 。

激励时段，导致驱动声波的压力发生阶跃变化，从而触发气泡振荡的两次爆发：一次在脉冲开始时，另一次在脉冲结束时。因此，一个脉冲引起的振荡不会完全衰减，直到气泡受到下一个脉冲的激励，这意味着系统状态取决于过去的输入。进而，这表明系统具有某种记忆 [58, 60], ，下文将对此进行说明。

我们采用两种标准方法来计算振荡气泡的记忆容量。第一种是将 $k$-延迟任务应用于随机二进制时间序列，并计算延迟目标信号 $\mathbf{y}_{n-k}^{target}$ ,$\mathbf{y}_{k,n}$,$\mathbf{y}_{n-k}^{target}$ 与振荡气泡产生的预测输出 $\mathbf{y}_{k,n}$ 之间的相关系数$r(\mathbf{y}_{n-k}^{target}$ ,$\mathbf{y}_{k,n})$。振荡气泡使用训练数据 $\mathbf{u}_{k,n}$ 进行训练，离散时间延迟为 $k$ [58]。函数 $r^2(\mathbf{y}_{n-k}^{target}$ ,$\mathbf{y}_{k,n})$ 的取值范围为 0 到 1，其中 0 表示相关性完全丧失，1 则表示正确目标数据与气泡所作预测之间完全相关。将 $r^2$ 作为离散延迟时间 $k$的函数进行计算，即可确定所谓的短期（亦称衰减）记忆容量（STM） [58, 60] ，其定义为

$$C_{STM} = \sum_{k=1}^{k_{max}} r^2(\mathbf{y}_{n-k}^{target}, \mathbf{y}_{k,n}) , \tag{D1}$$

为 $k_{max}$ 分析中考虑的最大延迟。

随后，我们让振荡气泡执行奇偶校验测试 [59], ：气泡需要预测一个时间序列，该序列定义为 $PARITY$（$\mathbf{u}_{n-k}$, $\mathbf{u}_{n-k-1}$, $\mathbf{u}_{n-k-2}$），对应逐渐增加的延迟 $k$。基于原始随机生成的二进制时间序列，由此生成的新时间序列本身也是一个二进制序列。奇偶校验函数不可线性分离，其预测要求储备池具备记忆能力。与方程（D1）类似，我们通过对截至最大延迟 $k_{max}$的奇偶校验值求和，计算奇偶校验记忆容量 $C_{PC}$ 。

计算得到的记忆容量见图 7，其中我们绘制了 $r^2$ 随离散延迟时间变化的关系。我们得到 $C_{STM} = 2.83$比特和 $C_{PC} = 3.18$比特。这些数值与其他类型物理储层系统所得结果相近 [60] ，并证实振荡气泡具有相当不错的记忆容量。我们还注意到，在图 7 的零延迟处，短期记忆（STM）和奇偶校验（PC）测试得到的记忆均约为 0.9 比特，而非理论上预期的、对应 100% 相关性的 1 比特。在标准回声状态网络（ESN）方法中，这种情况源于储备池中的长期混响 [40] ，它会降低对紧邻过去输入的回忆能力 [58]。