arXiv:2301.00800v2 [cond-mat.stat-mech] 7 Jan 2023

序言：从反常扩散数据表征物理过程

Carlo Manzo†、Gorka Mu˜noz-Gil‡、Giovanni Volpe♯、Miguel AngelGarcia-March♭、Maciej Lewenstein§ 和 Ralf Metzler¶

† 维克大学—加泰罗尼亚中央大学（UVic-UCC）科技与工程学院，C. de la Laura,13, 08500 维克，西班牙

‡ 因斯布鲁克大学理论物理研究所，Technikerstr. 21a, A-6020 因斯布鲁克，奥地利♯ 哥德堡大学物理系，Origov¨agen 6B, SE-41296 哥德堡，瑞典♭ 瓦伦西亚理工大学纯数学与应用数学大学研究所，西班牙§ ICFO—光子科学研究所，巴塞罗那科学与技术研究所，Av. Carl Friedrich Gauss 3, 08860 卡斯特尔德费尔斯（巴塞罗那），西班牙；以及加泰罗尼亚高级研究院，Pg. Llu´ís Companys 23, 08010 巴塞罗那，西班牙

¶ 物理与天文学研究所，波茨坦大学，Karl-Liebknecht-St. 24/25，D-14476 波茨坦-戈尔姆，德国；亚太理论物理中心，大韩民国浦项 37673

电子邮件：carlo.manzo@uvic.cat, rmetzler@uni-potsdam.de

引言

自阿尔伯特·爱因斯坦为 [1] 罗伯特·布朗对花粉粒中微观颗粒运动的观察提供了理论基础以来， [2],人们已在种类繁多的有生命和无生命系统中发现了显著偏离布朗运动规律的现象，涵盖从生物学到股票市场等领域。这种现象被称为反常扩散，它拓展了布朗运动的概念，并与无序系统、非平衡现象、能量与信息流以及生命系统中的输运相关 [3]。反常扩散之所以被称为“非普适”，是因为在物理性质截然不同的系统中，均方位移却具有相同的幂律形式 $\langle x^2(t)\rangle \sim t^{\alpha}$。因此，要正确理解呈现反常扩散的系统，采用可靠的分析方法来揭示影响反常扩散动力学的确切物理机制至关重要。

人们利用经典统计学 [4–10]开发了多种方法，用于检测反常扩散的发生及其背后的机制。然而，近年来，机器学习的蓬勃发展推动了基于数据的方法不断进步，使人们能够从单条轨迹中表征反常扩散，为解决这一问题提供了更精细的工具 [11–14]。

2020年，我们发起了反常扩散（AnDi）挑战赛，旨在通过社区协作，对在各种现实条件下量化反常扩散的经典方法和新方法进行首次评估 [15]。挑战赛以公开竞赛的形式举行，旨在对现有方法进行基准测试，并推动新方法的发明。AnDi挑战赛汇聚了一群活跃且跨学科的科学家，他们致力于研究这一问题；参赛者来自11个国家的22家机构，人数超过30人。最终，对参考数据集 [16] 所得结果的分析，为三项特定任务中从单条轨迹表征反常扩散的方法性能提供了客观评估，这三项任务包括反常扩散指数推断、模型分类和轨迹分割。这项发表于《自然·通讯》的研究分析了社区协作的成果，并指出机器学习显著提升了对扩散粒子性质的估算 [17]。

本专题特刊介绍了参与 AnDi挑战赛的若干方法的详细情况。部分文章描述了最初用于挑战赛的软件的更新版本，并展示了性能提升。这些方法大多依赖最先进的机器学习方法。例如，文献 [18] 将基于经典统计学的特征工程与前馈神经网络相结合。有趣的是，这项工作展示了如何针对挑战赛的数据集构建专门的处理流程，从而在所有任务中取得竞赛中的最佳表现之一。同样，文献 [19] 也将一组统计特征作为极端梯度提升模型的输入，再由该模型负责对扩散模型中的轨迹进行分类。作者还进一步表明，恰当选择特征会极大地

影响分类准确率。

在不同的机器学习方法中，循环神经网络（RNN）因适合处理包含时间信息和长程相关性的数据而受到广泛关注。三项研究展示了基于 RNN 应对挑战任务的不同实现方式 [20–22]。参考文献 [20] 将双向长短期记忆（LSTM）（一种先进的 RNN）与卷积神经网络（CNN）相结合。在一组长短期记忆网络层之前使用 CNN 作为特征提取器，以提升 RNN 的性能，并在反常指数回归方面取得了出色成果。参考文献 [21] 提出了一种架构，将轨迹直接输入长短期记忆网络层。这种实现能够分析任意大小的时间序列，无需进行填充或预处理。该方法在 AnDi挑战赛的多项任务中都展现出顶尖性能，证明了类似架构可以成功用于不同目的。最后，参考文献 [22] 介绍了最具潜力的 RNN 架构之一 WaveNet 的应用。该研究进一步表明，训练数据集的规模（比挑战赛中使用的其他模型大一个数量级）可能是提高方法准确性的关键。

其他机器学习方法在处理随机扩散方面也展现出卓越的性能。文献 [23]尝试使用极限学习机（ELM）为机器学习方法建立基准。极限学习机是一种收敛速度快的训练算法，应用于一组统计特征上的单隐层前馈神经网络。文献 [24] 使用预训练的卷积神经网络 ResNet-50 对不同的随机过程进行分类，并将其与其他卷积神经网络进行比较。文献 [25] 提出了一种新的架构——卷积Transformer，用于从轨迹中提取特征，并将其输入两个Transformer编码块，以执行回归或分类。文献 [26]提出了一种基于图神经网络（GNN）的方法，为每个轨迹位置关联一个特征向量，并为每条轨迹构建一个稀疏图结构。与文献 [22], 类似，作者使用表征学习技术研究其模型的潜在空间特征，并提出一种可视化探索方法，用于分析图神经网络从未见过的行走轨迹。事实上，扩散模型的无监督分析是表征未知数据集的一种有前景的工具，甚至可能帮助识别新的机制。沿着这一思路，文献 [27]研究了自编码器作为反常扩散轨迹特征提取器的适用性，并提出了一种利用异常检测对其进行表征的方法。

除了机器学习方法外，研究人员还提出了基于理论的方法来表征反常扩散，并在 AnDi 挑战赛中进行了测试。文献 [28] 讨论了用于求得反常扩散指数的数值方法，并提出了一种基于特征分析的模型选择问卷。文献 [29] 则采用贝叶斯推断来区分缩放布朗运动与分数布朗运动；文献 [30] 介绍了一种处理 L´évy 行走轨迹的方法。当充分了解轨迹信息，并且能够为所要表征的行走类型构建特定的先验分布时，贝叶斯方法尤其有效

可以构建。

本专题特刊还收录了若干理论研究成果，推动了随机过程领域的发展，和／或探讨了其在时间序列中的应用。因此，参考文献 [31]探讨了马尔可夫跳跃-陷阱情景中出现的瞬态反常扩散。在截尾分数积分过程中也观察到了瞬态反常扩散 [32]。参考文献 [33] 提出了经验异常度量，用于衡量反常扩散过程与正常扩散之间的距离。研究表明，L´évy 行走的极限性质有助于识别和验证 L´évy 行走型运动，也有助于采用最大似然方法进行参数估计 [34]。参考文献 [35] 研究了随机扩散率下分数布朗运动中出现的残余非遍历性，这或许有助于区分和归类粒子位移的某些非遍历和非高斯特征。参考文献 [36], 探讨了将单轨迹功率谱方法应用于鸢和鹳的运动数据，并展示了如何利用该方法提取随机模型。

文献 [37], 讨论了量子行走，展示了量子相干性与行走者均方位移之间的相互作用如何为这一过程提供信息。文献 [38] 研究了逆散射变换在非线性方程分数阶版本中的应用，例如科尔特韦格-德弗里斯方程，为具有幂律色散关系的孤子解提供了一个框架。文献 [39]报告称，自避随机行走在均方位移呈现布朗标度的同时，初始非高斯性很强。

本专题特刊报道的研究为理解反常扩散过程及其分析作出了重大贡献。具体而言，文中介绍了一系列工具，有望成为分析各种实验所生成轨迹的标准方法，涵盖从原子物理学到生态学等领域。此外，这些研究成果进一步凸显了社群协作对于推动科学进步的重要性。这项计划取得成功，促使我们围绕这样一个问题组织了 2nd AnDi挑战赛：如何根据单条轨迹检测输运性质的变化以及运动物体之间的相互作用。

致谢

我们感谢所有作者的贡献、期刊编辑团队的支持，以及 AnDiChallenge 参与者的积极参与。

CM 感谢获得由 MCIN/AEI/10.13039/501100011033 和“欧洲社会基金投资于你的未来”资助的 RYC-2015-17896 资助项目，以及由 MCIN/AEI/10.13039/501100011033/ 和“欧洲区域发展基金“建设欧洲之路””资助的 BFU2017-85693-R 和 PID2021-125386NB-I00 资助项目。GMG 感谢欧盟（ERC、QuantAI，项目编号 10105529）以及奥地利科学基金（FWF）通过 SFB BeyondC 提供的支持

F7102。GV 感谢 ERC StG ComplexSwimmers（资助编号 677511）和克努特与爱丽丝·瓦伦堡基金会提供的资助。MAGM 感谢西班牙教育与职业培训部（MEFP）通过 2018 年贝阿特丽斯·加林多计划（BEAGAL18/00203）、西班牙经济与竞争力部 MINECO（FIDEUA PID2019-106901GBI00/10.13039/501100011033）、QuantERA II 联合资助计划 2021 PCI2022-133004、MCIN 项目（由欧盟 NextGenerationEU（PRTR-C17.I1）提供资金）以及瓦伦西亚自治区政府提供的资助，项目编号 20220883（PerovsQuTe）。ML 感谢以下机构提供的支持：ERC AdG NOQIA；国家研究署（PGC2018-097027-B-I00/10.13039/501100011033、CEX2019-000910-S/10.13039/501100011033、国家计划 FIDEUA PID2019-106901GB-I00、FPI、QUANTERA MAQS PCI2019- 111828-2、QUANTERA DYNAMITE PCI2022-132919、I+D+I 项目“协作挑战”QUSPIN RTC2019-007196-7）；MICIIN（由欧盟 NextGenerationEU(PRTR-C17.I1)提供资金）以及加泰罗尼亚自治区政府；Fundaci´o Cellex；Fundaci´o Mir-Puig；加泰罗尼亚自治区政府（欧洲社会基金 FEDER 和 CERCA 计划、AGAUR 资助编号 2017 SGR 134、QuantumCAT/U16-011424，由加泰罗尼亚欧洲区域发展基金业务计划 2014-2020 共同资助）；巴塞罗那超级计算中心 MareNostrum（FI-2022-1-0042）；欧盟地平线 2020 FET-OPEN OPTOlogic（资助编号 899794）；欧盟地平线欧洲计划（资助协议 101080086 — NeQST）；波兰国家科学中心（Symfonia 资助编号 2016/20/W/ST4/00314）；ICFO 内部“QuantumGaudi”项目；欧盟地平线 2020 研究与创新计划下的玛丽·斯克沃多夫斯卡-居里资助协议编号 101029393（STREDCH）和编号 847648（“La Caixa”青年领袖奖学金，ID100010434：LCF/BQ/PI19/11690013、LCF/BQ/PI20/11760031、LCF/BQ/PR20/11770012、LCF/BQ/PR21/11840013）。然而，本研究中表达的观点和意见仅代表作者个人，并不一定反映欧盟、欧洲气候、基础设施与环境执行局（CINEA）或任何其他资助机构的观点和意见。欧盟或任何资助机构均不对这些观点和意见承担责任。RM 感谢德国科学基金会（DFG 资助编号 ME 1535/12-1）提供的资助。

参考文献

[1] Einstein A 1905 物理学年鉴 322 549–560[2] 布朗 R 1828 哲学杂志 4 161–173[3] Metzler R、Jeon J H、Cherstvy A G 和 Barkai E 2014 物理化学与化学物理 16 24128–24164[4] Thapa S、Lomholt M A、Krog J、Cherstvy A G 和 Metzler R 2018 物理化学与化学物理 20 29018–29037 [5] Burnecki K、Kepten E、Garini Y、Sikora G 和 Weron A 2015 科学报告 5 11306[6] Weron A、Burnecki K、Akin E J、Sol´e L、Balcerek M、Tamkun M M 和 Krapf D 2017 科学报告 7 5404[7] Sikora G、Wy loma´nska A、Gajda J、Sol´e L、Akin E J、Tamkun M M 和 Krapf D 2017 Physical Review E 96 062404

[8] Kepten E、Weron A、Sikora G、Burnecki K 和 Garini Y 2015 PLoS One 10 e0117722[9] Krapf D、Lukat N、Marinari E、Metzler R、Oshanin G、Selhuber-Unkel C、Squarcini A、Stadler L、Weiss M 和 Xu X 2019 Physical Review X 9 011019[10] Thapa S、Wylomanska A、Sikora G、Wagner C、Krapf D、Kantz H、Chechkin A 和 Metzler R 2020 New Journal of Physics 23 013008[11] Mu˜noz-Gil G、Garcia-March M A、Manzo C、Mart´ın-Guerrero J D 和 Lewenstein M 2020 New Journal of Physics 22 013010[12] Granik N、Weiss L E、Nehme E、Levin M、Chein M、Perlson E、Roichman Y 和 Shechtman Y 2019 生物物理学杂志 117 185–192[13] Bo S、Schmidt F、Eichhorn R 和 Volpe G 2019 Physical Review E 100 010102[14] Kowalek P、Loch-Olszewska H 和 Szwabi´nski J 2019 Physical Review E 100 032410[15] Mu˜noz-Gil G、Volpe G、Garc´ıa-March M A、Metzler R、Lewenstein M 和 Manzo C 2020 反常扩散挑战：单轨迹表征作为竞赛 人工智能新兴主题 2020 vol 11469（国际光学与光子学学会）p 114691C[16]Mu˜noz-Gil G、Manzo C、Volpe G、Garcia-March M A、Metzler R 和 Lewenstein M 2020 反常扩散挑战数据集 URL https://github.com/AnDiChallenge/ANDI_datasets [17] Mu˜noz-Gil G、Volpe G、Garcia-March M A、Aghion E、Argun A、Hong C B、Bland T、Bo S、Conejero J A、Firbas N、Garibo i Orts O、Gentili A、Huang Z、Jeon J H、Kabbech H、Kim Y、Kowalek P、Krapf D、Loch-Olszewska H、Lomholt M A、Masson J B、Meyer P G、Park S、Requena B、Smal I、Song T、Szwabi´nski J、Thapa S、Verdier H、Volpe G、Widera A、Lewenstein M、Metzler R 和 Manzo C 2021 《自然·通讯》 12 6253[18] Gentili A 和 Volpe G 2021 《物理学杂志 A：数学与理论》 54 314003[19] Kowalek P、Loch-Olszewska H、Laszczuk L、Opa la J 和 Szwabi´nski J 2022 《物理学杂志 A：数学与理论》 55 244005[20] Garibo-i Orts ˋO、Baeza-Bosca A、Garcia-March M A 和 Conejero J A 2021 《物理学杂志 A：数学与理论》 54 504002[21] Argun A、Volpe G 和 Bo S 2021 《物理学杂志 A：数学与理论》 54 294003[22] Li D、Yao Q 和 Huang Z 2021 《物理学杂志 A：数学与理论》 54 404003[23] Manzo C 2021 《物理学杂志 A：数学与理论》 54 334002[24] Al-hada E A、Tang X 和 Deng W 2022 《物理学杂志 A：数学与理论》 55 274006[25] Conejero J A、Firbas N、Garibo-i Orts ˋO 和 Garcia-March M A 2022 《物理学杂志 A：数学与理论》 即将出版[26] Verdier H、Duval M、Laurent F、Cass´e A、Vestergaard C L 和 Masson J B 2021 《物理学杂志 A：数学与理论》 54 234001[27] Mu˜noz-Gil G、i Corominas G G 和 Lewenstein M 2021 《物理学杂志 A：数学与理论》 54 504001[28] Meyer P G 、 Aghion E 和 Kantz H 2022 《物理学杂志 A：数学与理论》 55 274001[29] Thapa S、Park S、Kim Y、Jeon J H、Metzler R 和 Lomholt M A 2022 《物理学杂志 A：数学与理论》 55 194003[30] Park S、Thapa S、Kim Y、Lomholt M A 和 Jeon J H 2021 《物理学杂志 A：数学与理论》 54 484001[31] Vitali S、Paradisi P 和 Pagnini G 2022 《物理学杂志 A：数学与理论》 55 224012[32] Sabzikar F、Kabala J 和 Burnecki K 2022 《物理学杂志 A：数学与理论》 55 174002[33] Maraj K、Szarek D、Sikora G 和 Wy loma´nska A 2020 《物理学杂志 A：数学与理论》 54 024001

[34] Magdziarz M 和 Zorawik T 2020年《物理学杂志 A：数学与理论》53 504001[35] Wang W、Cherstvy A G、Chechkin A V、Thapa S、Seno F、Liu X 和 Metzler R 2020年《物理学杂志 A：数学与理论》53 474001[36] Vilk O、Aghion E、Nathan R、Toledo S、Metzler R 和 Assaf M 2022年《物理学杂志 A：数学与理论》55 334004[37] Hegde A S 和 Chandrashekar C 2022年《物理学杂志 A：数学与理论》 55 234006[38] Ablowitz M J 、 Been J B 和 Carr L D 2022 年《物理学杂志 A：数学与理论》55 384010[39] Marcone B、Nampoothiri S、Orlandini E、Seno F 和 Baldovin F 2022年《物理学杂志 A：数学与理论》55 354003