arXiv:2611.31916v1 [gr-qc] 24 Nov 2026

机器学习使自旋极端质量比旋近的完整贝叶斯推断成为可能

Bo Liang[1,2,3†], Minghui Du[1,3*†], Zhenwei Lyu[4*†], Hanlin Song[3], Jiakai Zhang[5], Weichen Qin[5], Chia-Jui Chou[6], Yi Yang[6], Jingyi Yu[5], Ziren Luo[1,3,8], Mingming Sun[2*], Peng Xu[1,3,7,8*], Yue-Liang Wu[1,7*]

[1]中国科学院大学引力波宇宙太极实验室（北京/杭州），中国北京，100049。[2]北京数学与应用数学研究院 AGI实验室，中国北京，101408。[3]中国科学院力学研究所国家微重力实验室引力波实验中心，中国北京，100190。[4]大连理工大学莱斯特国际学院，中国盘锦，124221。[5]上海科技大学信息科学与技术学院，中国上海，201210。[6]上海科技大学物理科学与技术学院，中国上海，201210。[7]中国科学院大学亚太理论物理国际中心，中国北京，100049。[8]中国科学院大学杭州高等研究院浙江省引力波精密测量重点实验室，中国杭州，310024。

*通讯作者。电子邮箱：mmsun@bimsa.cn；xupeng@imech.ac.cn；ylwu@ucas.edu.cn；

†这些作者对本研究作出了同等贡献。

摘要

极端质量比旋近（EMRI）是恒星质量致密天体旋入大质量黑洞的系统，有望通过毫赫兹引力波观测测量中心黑洞的质量和自旋，揭示其周围的强场环境，并估测宇宙距离。然而，迄今为止，对这些系统的参数推断，大多数预测都假定源参数已知，而宽先验分析仅处理了这一问题的简化版本。基于完整先验对自旋源的全部14个参数进行联合反演，至今仍未解决。在这里，我们针对增强解析简化模型解决了这一问题：生成网络缩小了完整先验的范围，并揭示出一圈简并自旋方向；精确的探测器似然则确定了后验分布。对于模拟的一年期LISA–Taiji 观测，在一个环形扇区内，我们首次从完整先验获得了自旋极端质量比旋近的14个参数后验分布，这是迄今从宽先验获得的此类源最精确的结果

。大质量黑洞的红移质量和自旋被确定到 10^5, 中的两份，其质量精度比任何电磁测量高出两个数量级。致密天体的红移质量被测量到 10^4, 中的一份，其偏心逆行轨道能够区分不同的俘获通道。0.8%的距离精度和小于一平方度的定位精度足以支持暗标准汽笛宇宙学。为期一年的相干分析解析出了由辐射反作用本身驱动的内旋过程。

关键词： 极端质量比旋近、LISA、太极计划、生成式推断、基于模拟的推断、并行回火马尔可夫链蒙特卡罗、强场引力

在广义相对论中，一个孤立的、静态的黑洞完全由其质量、自旋和电荷决定 [1]。由于预计天体物理黑洞所带的电荷可以忽略不计，因此仅凭六十多年前 Kerr 发现的旋转真空解，黑洞就可由其质量和自旋描述 [2, 3]。在星系核中，这两个量记录了大质量黑洞的成长历程 [4]，但二者都仍然难以测量。视界尺度成像和恒星轨道只能测定少数邻近天体的质量，精度约为百分之十至百分之几 [5, 6],而通过电磁观测推断自旋则取决于对吸积流的模型 [4]。极端质量比旋近（EMRIs）为测量这两个量提供了独立途径，既适用于静默星系，也适用于活动星系；此外，它们还能在不依赖宇宙距离阶梯的情况下确定源的距离。一个恒星质量的致密天体旋入大质量黑洞，为在强引力场中开展测试粒子实验提供了天然条件。在数万次相对论性轨道运行中，其引力辐射的相位编码了大质量黑洞时空的多极矩，Ryan 首先展示了这一点 [7]。LISA 等空间探测器 [8, 9] 和太极计划 [10] 能够在一年或更长时间内相干地跟踪这一相位，在此期间相位会累积约 10^5弧度。由于相位是轨道频率对时间的积分，频率的相对误差为 10^{-5} 会在观测过程中累积约一弧度的失相。对于预期信噪比达到 20 或更高的已探测极端质量比旋近双星 [11]„ 量级为信噪比倒数的相位差是可分辨的，因此这种失相很容易被识别，轨道频率也可被约束到每 10^5。因此，单个极端质量比旋近双星就能将大质量黑洞的质量和自旋确定到 10^{-4} 或更高精度 [11, 12], 并揭示致密天体曾经如何捕获的 [13, 14], 可探测黑洞 [15–17]周围的暗物质和新的基本场，并可作为宇宙学中的暗标准汽笛 [18–20]。没有其他可观测现象能达到这样的探测范围：双星脉冲星探测的是以光速的千分之几运行的轨道 [21]„ 而正在合并的恒星级质量黑洞在强场区只经历几个周期 [22]。这些前景使极端质量比旋近双星成为LISA和太极计划的标志性目标。根据两项任务已发布的计划和路线图，预计两项任务的科学运行将从2035 [8–10],左右开始重叠，因此在此期间观测到的每个极端质量比旋近双星事件，都将由两台轨道和响应特性各异的独立仪器记录，从而能够交叉核验每一项推断出的特征 [23]。波形建模也在同步发展 [24–26]。例如，增强解析拼接模型（AAK）将其轨道频率映射到 Kerr测地线的轨道频率上 [27]。从数据中推断源参数这一逆问题的进展则相对滞后，尽管所有这些科学目标都取决于能否解决这一问题。

对于旋进至自旋黑洞的普遍情形，这一逆问题极具挑战性。必须联合推断 14 个参数。克尔轨道的三个基本频率受到探测器运动的调制，形成密集的谱森林 [28]。由于仅 10^{-5} 的分数偏移就会在一年内累积出一弧度的相位，似然被限制在广阔先验范围内的狭窄弯曲流形上，而该先验中遍布遥远的次级极大值 [29], ，因此，完全相干的模板搜索需要 $\sim 10^{40}$ 个模板 [30]。自旋既使问题更加复杂，也令其值得求解。参考系拖曳只能在地球轨道上测量到约百分之十 [31, 32], ，但它在此处塑造了每一条轨道，并使倾角、自旋方向与轨道相位彼此纠缠。若错过似然盆地，或错误地匹配到一部分谐波，就会使推断出的质量、自旋和

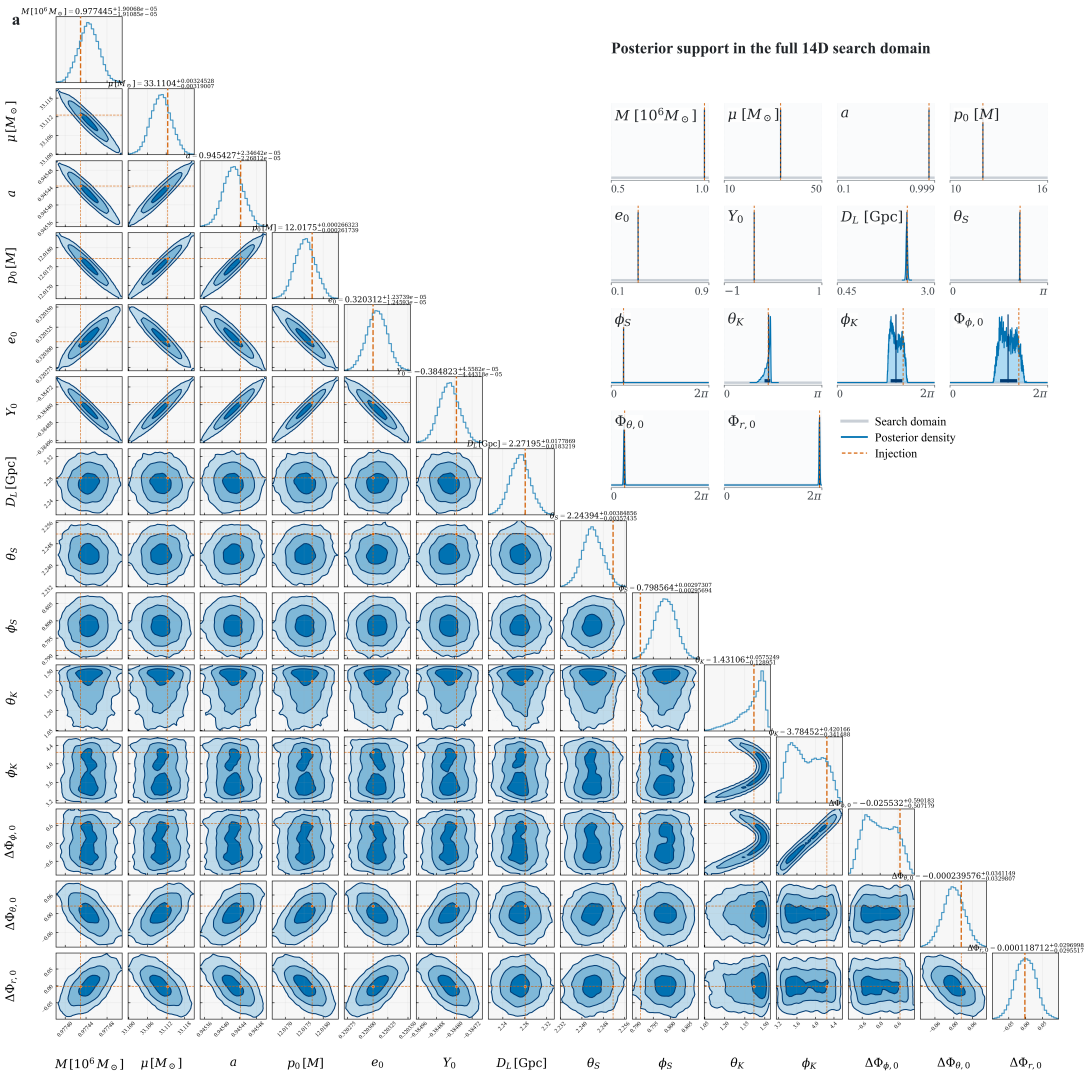


图1 一年联合 LISA–Taiji 数据中的自旋极端质量比旋近十四参数后验分布。左下：通过并行退火马尔可夫链蒙特卡罗得到的一维和二维边缘概率密度。采样时采用精确的双探测器似然，并从已完成的十三参数分析（方法）出发，对全部十四个参数（包括光度距离）进行采样。等高线分别包含二维概率的39%、86%和99%，标签给出各参数的中位数和中心68%区间。轨道相位以相对于其圆均值的偏移量表示。右上：在完整先验范围（灰色带）内绘制各一维边缘分布，展示先验到后验分布的收缩。橙色虚线表示注入值，这些数值在采样后添加，采样器并未使用。

轨道参数 [29], , 并可能掩盖或模拟黑洞天体物理环境所引起的微小去相位效应 [33, 34]。如今，神经推断能够实时分析地面探测器探测结果 [35, 36], , 而传统搜索算法和机器学习方法已稳步拓宽了可用于恢复极端质量比旋近双星的先验范围 [37–41]。然而，这些分析要么只返回点估计，要么得到后验分布

对于问题的简化版本，可以忽略黑洞自旋、仅从宽先验中推断部分参数，或至多分析半年的数据（方法和扩展数据表 1）。实际观测不允许进行这些简化：预计大质量黑洞会自旋 [4],事先不知道任何源参数，而要达到极端质量比旋近的完整精度，需要一年的数据

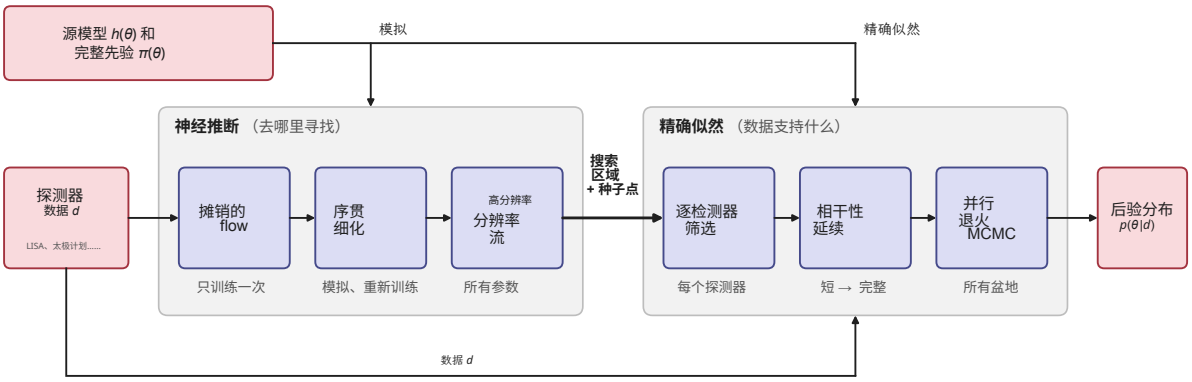


图 2 NEBULA-EMRI 框架。来自探测器网络的数据首先经过神经推断阶段，该阶段决定搜索位置；随后进入精确似然阶段，该阶段判断数据支持什么。网络在源模型和先验生成的模拟数据上进行训练，仅将搜索区域和阶段之间的起始点（种子点）传递过去。精确阶段直接评估数据的似然。候选项在每个探测器中分别进行筛选，随着相干积分时间从较短片段延长至完整观测时长持续跟踪，并通过并行退火马尔可夫链蒙特卡罗在各盆地之间进行合并，其冷链给出后验分布。神经密度既不参与似然计算，也不对后验分布加权。

或更长时间的相干相位跟踪。然而，基于完整先验、涵盖一般自旋源全部14个参数、持续一年的贝叶斯后验分布，至今仍无法获得。

在此，我们使用 AAK 波形模型描述的自旋极端质量比旋近双星，借助面向极端质量比旋近双星的神经到精确贝叶斯统一似然分析（NEBULA-EMRI）求解这一反演问题。该方法分为两个阶段（图 2）。在神经推断阶段，一个基于完整先验训练的摊销网络可在不到一秒内约束质量、自旋大小、轨道要素和距离。在中间似然搜索的辅助下，利用在依次缩小的区域内重新训练的网络，进一步将约束扩展至全部14个参数，并揭示自旋方向中的环状简并。精确似然阶段沿着该环状简并的一个扇区进行搜索。首先分别使用每个探测器的30天数据对候选项进行筛选，仅保留同时得到 LISA 和太极计划支持的候选项。随后将相干积分时间从30天逐步延长至整整一年，并通过并行退火马尔可夫链蒙特卡罗，在精确的一年期双探测器似然下对全部14个参数进行采样。神经密度既不参与似然计算，也不对后验分布加权。网络仅定义搜索区域和起始点。

我们将 NEBULA-EMRI 应用于一个模拟的、为期一年的 LISA–Taiji 观测，该观测针对一个极端质量比旋近，其参数从完整先验中随机抽取，位于光度距离 2.3 吉秒差距处（红移 $z \approx 0.4$；扩展数据表2）。在此距离下，信号在 LISA 中的信噪比达到 46，在太极计划中达到 123，双探测器网络的信噪比则为 131。NEBULA-EMRI 得出了涵盖全部 14个参数、历时一年的后验分布（图1）。据我们所知，这是首次从完整先验中恢复出的旋转极端质量比旋近全部 14个参数的联合贝叶斯后验分布，也是迄今为止利用宽先验对此类源获得的最精确结果。这种精度赋予了反演以科学价值。仅一年的数据，就能以远超电磁方法所能达到的精度确定大质量黑洞的红移质量和自旋，约束致密天体被捕获的方式，并将源定位到足以开展暗标准汽笛宇宙学研究的程度。

一年的数据能测量什么。

借助 NEBULA-EMRI，历时一年的后验分布将大质量黑洞的探测器参考系质量和自旋分别测定到 19 和 24 ppm 的相对精度。它还将致密天体的探测器参考系质量以及三个轨道要素（初始半通径、偏心率和倾角余弦）的相对精度约束在 2×10^{-5}至 1.2×10^{-4}之间，将光度距离约束到 0.8%，并将天空位置约束到不足一平方

表 1 为期一年的联合 LISA–Taiji 后验分布的物理探测范围。精度为中心 68% 分位数半宽度，除非另有说明（方法）。

量	精度	物理学赋能	参考点
大质量黑洞质量 M (探测器参考系)	1.9×10^{-5}	人口统计学与增长大质量黑洞	最佳电磁测量， 2.5×10^{-3} (人马座A*) [6]
自旋 a	2.4×10^{-5}	参考系拖曳与增长历史（吸积还是并合） [4]	来自电磁观测的自旋，这些观测取决于吸积流模型 [4]
致密天体质量 μ （探测器参考系）	9.7×10^{-5} （$0.003\,M_\odot$）	恒星级质量黑洞位于星系核	并合双星的组分质量并合双星，几十百分比 [42]
轨道要素 (p_0, e_0, Y_0)	$(2.2, 3.9, 12) \times 10^{-5}$	俘获通道：两体弛豫与吸积盘迁移 [13, 14]	通道预测出不同的偏心率和倾角分布 [14]
光度距离 D_L	8×10^{-3}	暗标准汽笛测量哈勃常数 [20]	百分比级 H_0 预报来自高信噪比极端质量比旋近源 [20]
天空位置（θ_S， ϕ_S）	4×10^{-3}, 3×10^{-3}弧度	宿主星系识别与电磁后续观测	面积仅为一小部分的平方度区域
联合分布 14 参数协方差	—	所有参数的相关不确定性参数，包括盆地之间的离散程度	无法从点估计中获得估计值

度（68%可信度；图1 和表1）。假设源参数已知的预测长期以来一直表明，极端质量比旋近双星可以达到这样的精度 [11, 12]。在此，我们通过从完整先验出发的全局分析实现了这一精度。将全部14个参数的先验范围与68%后验宽度之比相乘，其结果接近 10^{40}, ，与此前估计的完全相干搜索所需模板数量相当 [30]。此前针对自旋极端质量比旋近的宽先验分析，尚无一项达到相当的后验精度。唯一更早的自旋源宽先验后验分布来自我们先前的研究 [40], ，该研究利用60天的单探测器数据推断出四个内禀参数；而迄今为止，14个参数的搜索仅给出了点估计。最近的一项研究采用原始解析简化波形进行为期半年的单探测器相干搜索，其报告的质量和自旋误差比本文结果大一至两个数量级 [41] （方法和扩展数据表1）。14个注入值中有10个落在各自的68%区间内，其中包括决定轨道动力学和信号振幅的全部七个参数，这符合校准区间的预期（方法）。

每一项测量都有独特的物理含义（表1）。在十亿秒差距距离上，黑洞的探测器参考系质量测定精度比目前对任何大质量黑洞所进行的最佳电磁测量——银河系中心的黑洞——高两个数量级 [6]。黑洞的自旋会通过参考系拖曳塑造每一条轨道，并记录黑洞主要是通过相干吸积还是通过并合成长起来的 [4], ，其测量精度达到2.4×10^{-5}。致密天体的探测器参考系质量测量精度约为 10^{-4}, 比并合双黑洞的组分质量精确三个数量级 [42]。结合红移估计，这使得我们能够直接测量星系核中恒星级质量黑洞的源参考系质量。致密天体轨道的偏心率和倾角测量精度达到 10^{5}–10^{4}, ，可以区分通过两体弛豫俘获形成的系统（其轨道可具有任意倾角且偏心），以及通过吸积盘迁移形成的系统（其轨道近乎圆形、顺行，并与吸积盘对齐） [13, 14]。距离

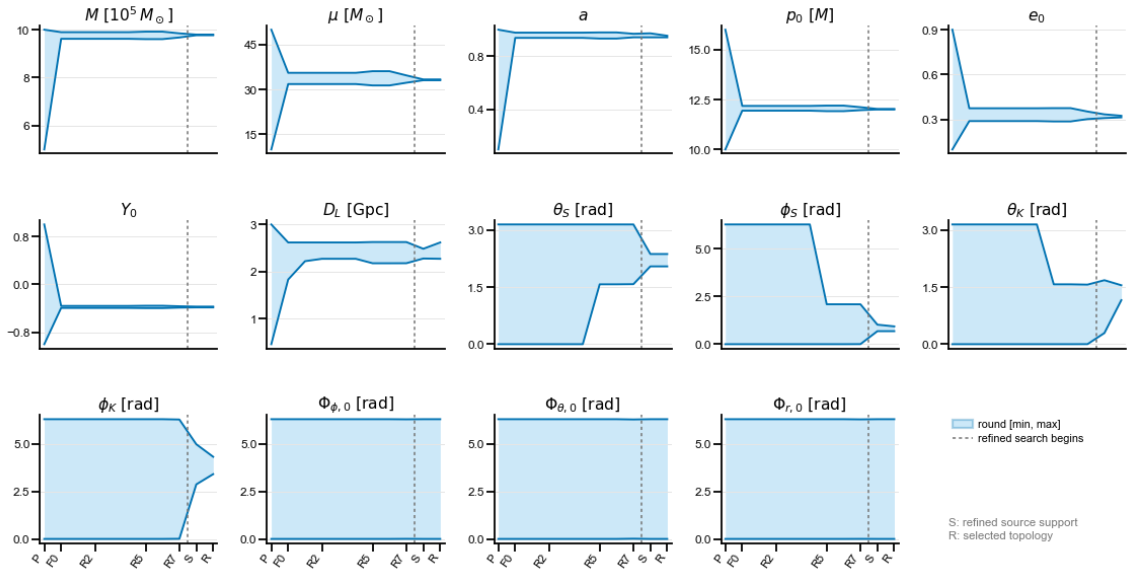


图 3 从完整先验到传递给精确似然阶段的区域的逐坐标支持范围。阴影带表示各参数模拟支持范围的下限和上限，从完整先验（P）经摊销网络（F0）和序贯轮次 R1–R7，到高分辨率模拟域（S）以及传递给精确似然阶段的区域（R）。六个本征参数和光度距离在 F0 阶段收窄，而天空位置和自旋方向直到 R5 阶段仍保持完整范围。R5 阶段这些角度的阶跃反映了对一个方向单元的选择（方法）。轨道相位在整个过程中保持完整的周期范围。虚线标示高分辨率细化的开始。角度单位为弧度。

和天空位置可将候选宿主星系限定在一个很小的体积内，达到电磁后续观测以及独立于宇宙距离阶梯测量哈勃常数所需的精度 [18, 20, 43]。数据还可约束自旋轴与视线之间的夹角。对于活跃星系中的极端质量比旋近双星，将这一角度与望远镜测量得到的吸积盘和喷流的倾角进行比较 [4]，或可揭示黑洞及其吸积盘是否对齐，这是理解黑洞如何成长并塑造其宿主星系的关键问题。

缩小完整先验范围。

要从未知信号中获得此类测量，首先必须在完整先验范围内定位信号源；这长期以来被视为极端质量比旋近双星科学的核心计算难题 [44]。摊销推断将这一搜索转化为一次网络评估：网络只需在涵盖完整先验的模拟数据上训练一次，便能在不到一秒内将一年的 LISA–Taiji 数据映射为关于质量、自旋大小、轨道要素和距离的概率分布，其中心区间则界定了参数区域与数据相符的太空范围。在具有不同源参数和信噪比的 2,000 个独立测试信号上，随着信噪比增加，该区域收缩得更明显；这一点可通过网络目标参数的完整取值范围与该区域范围之比来衡量（方法），而区域的中心3σ 部分仍在 98% 的情况下包含真实信号源（扩展数据图 1）。对于此处分析的信号源，该网络能在一秒内识别系统性质：一个自旋快速、质量约为一百万太阳质量的黑洞，以及一个质量为 32–36 太阳质量、处于偏心逆行轨道上的致密天体，距离为 1.8–2.6 吉秒差距（图 3）。这种摊销网络可以扫描 LISA 和太极计划的数据流，其中将包含成千上万个重叠信号 [8]，按信噪比对极端质量比旋近候选项排序，并将其交由序贯细化和精确似然分析处理，从而为编目大质量黑洞及其捕获的致密天体轨道的处理流程提供前端。

序贯细化随后将神经推断阶段扩展到全部14个参数，加入天空位置、自旋方向和轨道

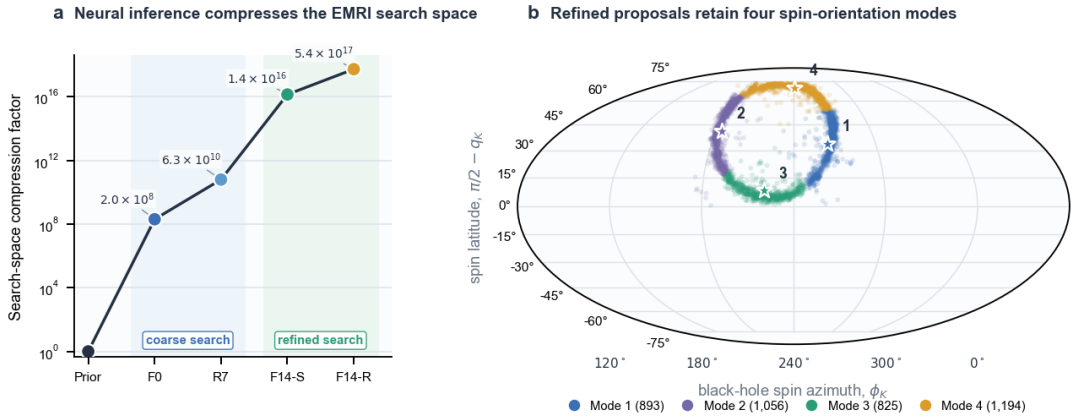


图 4 神经推断阶段缩小了搜索区域，并揭示出一圈自旋方向。a，相对于完整的 14参数先验，搜索区域的收缩程度，即先验体积除以区域体积；分别针对高分辨率模拟域（F14-S）以及传递至精确似然阶段的区域（F14-R），给出摊销网络（F0）和最后一轮序贯迭代（R7）之后的结果。**b**，3,968 个位于自旋边界内的高分辨率提议分布样本的自旋方向，以黄道坐标表示（摩尔威德投影）。这些提议分布样本形成一个环，其中心靠近视线；k-均值聚类将其划分为四个扇区，分别包含 893、1,056、825 和 1,194 个样本，星号标示扇区中心。精确似然阶段采用扇区 3（方法）。

相位。网络在当前区域内的新模拟数据上经过多轮重新训练，随后使用两个高分辨率网络，其中心设在中间似然搜索的结果处（方法）。这些步骤共同使待搜索区域缩小至原来的 5.4×10^{17} 倍（图3 和4a）。这一收缩因子等于完整先验的体积除以搜索区域的体积。由于二者都是由14个参数构成的盒体，它等于每个参数的先验范围与其保留范围之比的乘积（方法）。输出结果是与数据一致的所有解的分布图，而非单一的最佳估计，并揭示出一个清晰的规律：允许的自旋方向构成一个环，其中心接近视线（图4b）。这一环有着简单的物理成因。如果将整个系统绕视线旋转，波的偏振会发生变化，而这种变化很大程度上可以通过移动轨道相位来抵消。因此，数据能够约束自旋轴与视线之间的夹角，却基本无法约束绕其旋转的方位角。网络从模拟中学到了这一对称性，尽管没有人明确告知它这一点。方向各异的探测器能够更准确地测量偏振，因而可以约束环上的位置；这也是探测器网络具有价值的原因之一 [23]。在最终分析之前绘制这类简并结构十分重要，因为未被察觉的结构会使自旋方向产生偏差，并通过相关性影响质量、自旋和轨道参数 [29]。

探测器之间的相干性与一年的数据。

使极端质量比旋近双星如此精确的特性，也使它们难以被发现。在一年的时间里，信号的相位累积约为 10^5 弧度，因此，如果模板的轨道频率与真实值相差几分之 10^5，就会失去与数据的相位相干性。因此，直接从探测器网络所确定的区域初始化一年搜索，会陷入次级似然极大值，而这些极大值中没有一个比单独的噪声更能拟合数据（方法）。相反，该分析从30天数据开始，并针对任何声称存在的信号提出其所需条件：独立确认。差分进化这一基于似然的优化器分别在LISA和太极计划数据上运行，只有同时比单独噪声更好地拟合两者的候选项才会被保留（图5）。具有不同轨道、取向和噪声特性的两个探测器，不太可能因错误匹配信号的一部分而以相同方式受到误导。

随后，相干积分时间从30天逐步增加到整整一年（图5）。在这一范围内，六个内禀参数的不确定性缩小了13至69倍（方法），而仅由信噪比增加所带来的缩小幅度大约只有

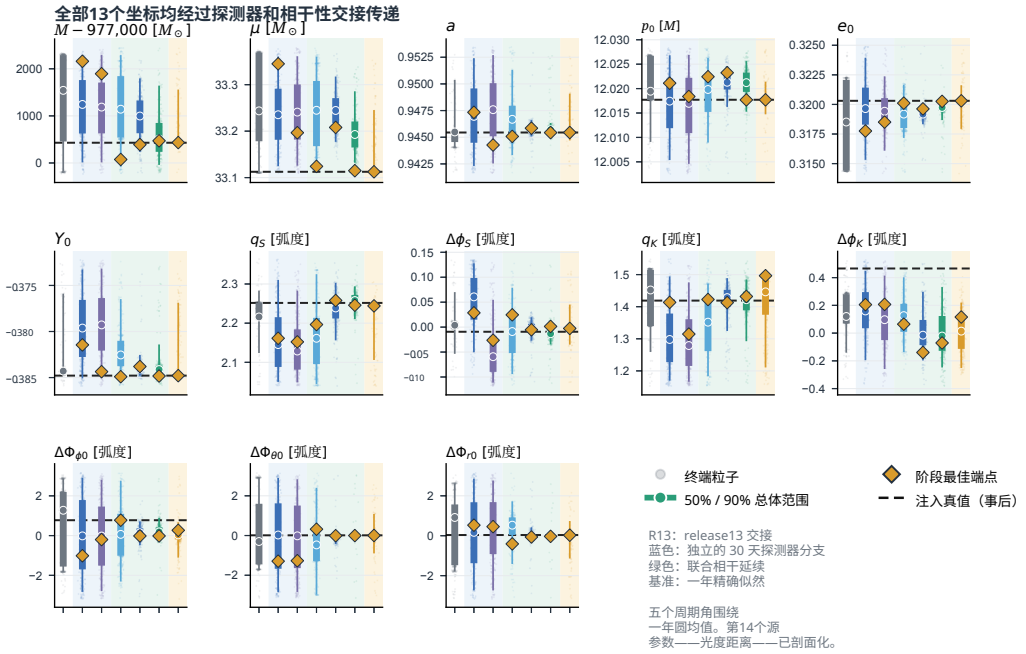


图5 探测器检验和更长的相干性将候选盆地延续至一年。图中展示了每个13个采样参数的精确似然搜索最终总体：从13参数细化（R13）开始，经过在LISA和太极计划中分别进行的30天搜索（蓝色）、在30天、60天和180天数据上进行的联合搜索（绿色），以及一年的精确似然搜索。粗线和细线分别表示每个总体的中间50%和90%，菱形标出各阶段的最佳点。虚线表示注入值，在分析完成后添加。五个周期角均以其一年圆均值为参照显示，光度距离则已剖面化。为清晰起见，未显示360天阶段。

3.5。额外的增益源于致密天体因向外辐射引力辐射而损失能量和角动量，导致轨道频率缓慢演化；只有跟踪整整一年的相位，才能测量这种演化。黑洞时空的多极结构、周围的暗物质云或吸积盘，或者新的基本场，都会在这些轨道频率及其变化率中留下印记 [7, 15, 17, 34]。

最后，并行退火马尔可夫链蒙特卡罗在一年的精确似然下，整合所跟踪扇区中所有保留下来的盆地。所得链通过了标准收敛检验，随后对光度距离进行显式采样，从而得到图1 所示的14参数后验分布（方法）。由于这些盆地被合并而不是缩减为最佳盆地，所报告的不确定性包含了它们之间的离散程度。

讨论。

我们已经完成了对由 AAK 波形描述的自旋极端质量比旋近系统的联合贝叶斯反演：从完整先验出发，并在遵循自旋扇区的条件下，全部 14个参数均在为期一年的双探测器后验分布中得到恢复。据我们所知，此前尚无分析完成过这项反演（扩展数据表 1）。这很重要，因为迄今为止，LISA 和太极计划的极端质量比旋近双星科学案例一直依赖于假设源参数已知的预测。要测量大质量黑洞的质量和自旋、绘制星系核捕获致密天体的过程、利用暗标准汽笛测量哈勃常数，并最终检验强场引力，都必须先从数据中恢复源。在 AAK 模型中，这种预测级精度如今已通过全局分析实现，而不再只是一个假设。

该方法以一种简单的方式划分工作：神经网络决定在哪里寻找，而精确似然决定数据支持什么。搜索区域相当于对均匀先验的截断，在该区域内，精确的探测器似然决定后验分布，如标准贝叶斯分析中所示 [45]。除自旋方向和相关的方位相位外，该

后验分布完全位于该区域的边界之内（方法）。由于神经网络只需要模拟，而优化和采样步骤只需要进行似然评估，因此，同一框架可以纳入更精确的相对论波形 [24, 26]，并应用于不同的探测器，例如天琴 [46]。它还应适用于其他推断问题：长时间观测能够带来高精度，但解隐藏在庞大的参数空间中。

当前演示分析了平稳高斯噪声中的一个模拟源。为使计算易于处理，精确似然阶段仅沿用了四个自旋扇区中的一个（方法），因此测量结果以该扇区为条件。下一步是传播所有候选扇区并比较其证据，在众多源和噪声实现上检验可信区间的校准情况，并纳入现实噪声、数据缺口和重叠信号。这些步骤都不会改变该框架的逻辑。

LISA 和太极计划计划从大约 2035 年起同时开展观测，预计仅 LISA 每年就能探测到几十至几千个极端质量比旋近双星 [11]。一旦扩展到包含重叠源的真实数据流，这类分析就能为每个分辨良好的事件提供大质量黑洞的质量和自旋、致密天体抵达该黑洞时所处的轨道以及其距离。对于众多事件而言，持续数年的相位跟踪将因此成为一份关于大质量黑洞及其在宇宙时间尺度上由星系核供给方式的普查，并独立测量宇宙膨胀率。

由于 AAK 描述的是处于真空中的 Kerr 黑洞，且采样过程中未考虑 Kerr四极矩的偏离，因此当前后验分布在 Kerr假设下约束了源参数，但无法确定时空是否为 Kerr 时空。此类检验需要将偏离参数以及相关情况下的环境项，与 14 个源参数一同进行采样 [47]。尽管这会引入额外的简并 [48]，，但此扩展只需对扩展后的波形模型进行模拟，无需改变从神经网络到精确模型的架构，因此长达一年的极端质量比旋近观测有望检验大质量黑洞是否为广义相对论中的 Kerr 黑洞。

数据可用性。本研究分析的数据采用 LISA 数据挑战（LDC） [49] 和太极数据挑战（TDC） [50]模拟框架生成，使用的波形来自 FastEMRIWaveforms [26]。注入参数列于扩展数据表2。补充信息提供了网络架构与训练方法、每轮细化所用的模拟域、搜索与采样设置，以及每个延续阶段的区间和边缘分布，可在以下网址获取：https://github.com/aspirinliang/NEBULA-EMRI。

代码可用性。 NEBULA-EMRI 代码，包括神经网络、精确似然搜索和采样程序，以及用于复现图表的脚本，将公开发布于 https://github.com/aspirinliang/NEBULA-EMRI。波形使用公开可用的 FastEMRIWaveforms软件包 [26]生成。

致谢。 我们感谢 Alvin Chua，他在新加坡国立大学组织了 EMRI研讨会。本研究得到国家重点研发计划（项目编号：2025YFE0217300）和中国科学院国际合作伙伴计划（项目编号：025GJHZ2023106GC）的支持。

利益冲突。 作者声明不存在利益冲突。

参考文献

[1] Chru´sciel, P. T.、Lopes Costa, J. 和 Heusler, M. 稳态黑洞：唯一性及其拓展。《相对论生活评论》 **15**，7（2012年）。

[2] Kerr，R. P. 自旋质量的引力场：代数特殊度规的一个示例。《物理评论快报》**11**，237–238（1963）。

[3] Carter，B. 轴对称黑洞只有两个自由度。《物理评论快报》**26**，331–333（1971）。

[4] Reynolds, C. S. 黑洞自旋的观测约束。《天文学与天体物理学年鉴》**59**, 117–154 (2021年)。

[5] 事件视界望远镜合作组。首批 M87 事件视界望远镜观测结果。I. 超大质量黑洞的阴影。《天体物理学杂志快报》**875**, L1 (2019年)。

[6] GRAVITY合作组。基于多个恒星轨道的干涉测量天体测量确定银河系中心的质量分布。《天文学与天体物理学》**657**，L12（2022年）。

[7] Ryan，F. D. 任意多极矩的巨大轴对称天体中致密天体内旋产生的引力波。《物理评论*D*》**52**，5707–5718（1995年）。网址：https://doi.org/10.1103/PhysRevD.52.5707。

[8] Colpi，M.等。LISA定义研究报告。*arXiv*预印本（2024年）。网址：https://arxiv.org/abs/2402.07571。

[9]欧洲空间局。欧洲空间局雄心勃勃的LISA任务开始建设。新闻稿，6月17日（2025年）。网址：https://www.esa.int/Science Exploration/Space Science/LISA/Constructionof ESAs ambitious LISAmission begins。

[10] Hu, W.-R. & Wu, Y.-L. 太极计划在太空中的引力波物理学与引力的本质。《国家科学评论》**4**，685–686（2017年）。网址：https://doi.org/10.1093/nsr/nwx116。

[11] Babak, S.等。空间干涉仪 LISA 的科学研究。V. 极端质量比旋近。《物理评论*D*》**95**（2017年）。网址：https://doi.org/10.1103/PhysRevD.95.103012。

[12]Barack, L. & Cutler, C. LISA 捕获源：近似波形、信噪比与参数估计精度。《物理评论*D*》**69**（2004年）。网址：https://doi.org/10.1103/PhysRevD.69.082005。

[13] Amaro-Seoane，P. 相对论动力学与极端质量比旋近。《相对论生活评论》**21**，4（2018年）。

[14] Pan，Z. & Yang，H. 活动星系核中极端质量比旋近的形成率。《物理评论*D*》**103**，103018（2021年）。[15] Eda，K.、Itoh，Y.、Kuroyanagi，S. & Silk，J. 暗物质性质的新探针：嵌入暗物质迷你尖峰中的中等质量黑洞发出的引力波。《物理评论快报》**110**，221101（2013年）。

[16]Duque, F., Macedo, C. F. B., Vicente, R. & Cardoso, V. 超轻暗物质中的极端质量比旋近。《物理评论快报》**133**，121404（2024年）。

[17] Maselli,A. 等。 利用LISA对极端质量比旋近产生的引力波进行观测，以探测基本场。 《自然·天文学》**6**, 464–470 (2022年)。

[18] Schutz, B. F. 根据引力波观测确定哈勃常数。《自然》**323**, 310–311 (1986)。

[19] Cai, R.-G., Cao, Z., Guo, Z.-K., Wang, S.-J. & Yang, T. 引力波物理学。《国家科学评论》**4**, 687–706 (2017年)。

[20] Laghi, D. 等。 极端质量比旋近的引力波宇宙学。 《皇家天文学会月报》**508**, 4512–4531 (2021年)。

[21] Kramer, M. 等。 通过双脉冲星计时检验广义相对论。 《科学》**314**, 97–102 (2006)。

[22] LIGO科学合作组织 & Virgo 合作组织。观测到双黑洞并合产生的引力波。《物理评论快报》**116**，061102（2016）。

[23] Ruan, W.-H.、Liu, C.、Guo, Z.-K.、Wu, Y.- L. 和 Cai, R.-G. LISA–太极网络。《自然·天文学》 **4**，108–109（2020）。网址 https://doi.org/10.1038/s41550-019-1008-4。

[24] Wardell,B. 等。 基于二阶自力理论的致密双星引力波形。《物理评论快报》**130**，241402（2023年）。

[25] Chua, A. J.、Katz, M. L.、Warburton, N. 和 Hughes, S. A. 面向 LISA 毫赫兹引力波数据分析的完全相对论极端质量比内旋波形模板快速生成。《物理评论快报》 **126**（2021年）。网址 https://doi.org/10.1103/PhysRevLett.126.051102。

[26] Katz, M. L.、Chua, A. J.、Speri, L.、Warburton, N. 和 Hughes, S. A. 快速极端质量比内旋波形：用于毫赫兹引力波数据分析的新工具。《物理评论*D*》 **104**（2021年）。网址 https://doi.org/10.1103/PhysRevD.104.064047。

[27] Chua, A. J. K., Moore, C. J. 与 Gair, J. R. 用于探测极端质量比旋近的增强型简化波形。《物理评论*D*》**96**,044005（2017年）。

[28] Drasco, S. 与 Hughes, S. A. 一般极端质量比内旋的引力波快照。《物理评论*D*》**73** （2006年）。网址https://doi.org/10.1103/PhysRevD.73.024027。

[29] Chua, A. J. K. 与 Cutler, C. J. 引力波信号在极端质量比旋近内禀空间中的非局域参数退化。《物理评论*D*》**106**, 124046（2022年）。

[30] Gair, J. R.等。LISA 极端质量比捕获源的事件率估计。《经典与量子引力》**21**, S1595–S1606（2004年）。网址 https://doi.org/10.1088/0264-9381/21/20/003。

[31] Ciufolini, I. & Pavlis, E. C. 对广义相对论关于Lense–Thirring效应预言的确认。《自然》431，958–960（2004年）。

[32] Everitt, C. W. F.等Gravity Probe B：检验广义相对论的空间实验最终结果。《物理评论快报》**106**, 221101 (2011).[33]Cardoso, V., Destounis, K., Duque, F., Panosso Macedo, R. & Maselli, A. 天体物理环境中极端质量比系统产生的引力波。《物理评论快报》**129**, 241103 (2022年)。

[34] Speri, L. 等利用引力波探测吸积物理。《物理评论*X*》**13**, 021035 (2023).

[35] Dax, M. 等。利用神经后验估计开展实时引力波科学。《物理评论快报》127，241103（2021年）。

[36] Dax, M. 等。利用机器学习对双中子星并合进行实时推断。《自然》639，49–53（2025年）。网址 https://doi. org/10.1038/s41586-025-08593-z。

[37] Strub, S. H.、Speri, L. & Giardini, D. 在LISA中寻找极端质量比旋近：从识别到参数估计。arXiv预印本（2025年）。网址 https://arxiv.org/abs/2505.17814。

[38] Cole, P. S. 等 极端质量比旋近的序贯模拟推断。《物理评论*D*》**113**, 063030(2026年)。

[39] Liang,B. 等 利用机器学习快速估计极端质量比旋近参数。 *arXiv*预印本 (2024年)。网址 https://arxiv.org/abs/2409.07957。

[40]Liang, B. 等 迈向高效、准确的极端质量比旋近参数估计：机器学习增强的马尔可夫链蒙特卡罗框架。《*Research*》**9**, 1055(2026年)。

[41] Zou, X., Zhong, X., Han, W.-B. & Mohanty, S. D. 通用极端质量比旋近的相干端到端搜索。 *arXiv*预印本(2026年)。网址 https://arxiv.org/abs/2608.05974。

[42] LIGO科学合作组织、Virgo合作组织与KAGRA合作组织。GWTC-3：LIGO和Virgo在第三次观测运行后半段观测到的致密双星并合。《物理评论*X*》**13**, 041039 (2023年)。

[43] LIGO科学合作组织、Virgo合作组等引力波标准汽笛测量哈勃常数。《自然》**551**, 85–88（2017年）。

[44] Babak, S. 等LISA数据挑战：从挑战3到挑战4。经典与量子引力 **27**，084009 （ 2010 年）。网址：https://doi.org/10.1088/0264-9381/27/8/084009。

[45] LIGO科学合作组织与Virgo合作组织。双黑洞并合GW150914的性质。《物理评论快报》**116**，241102（2016年）。

[46] Luo, J. 等。 天琴：一种空间引力波探测器。《Class. Quantum Grav.》**33**，035010（2016年）。

[47] Barack, L. 与 Cutler, C. 利用LISA极端质量比旋近源检验大质量黑洞几何中的偏离克尔解。《物理评论D》**75**，042003（2007年）。

[48] Zou, X.、Zhong, X.、Han, W.-B. 与 Mohanty, S. D. 通过凸起参数化和粒子群优化约束克尔度规在极端质量比旋近中的偏离。《物理评论D》**112**，084075（2025年）。

[49]Baghi, Q. LISA数据挑战。莫里昂会议第56届引力学会议论文集（意大利拉图伊勒）（2022年）。网址https://arxiv.org/abs/2204.12142。arXiv:2204.12142。

[50] Ren, Z.等。 探索引力波宇宙的太极数据挑战。*Front. Phys.* **18**，64302（2023年）。

[51] Prince, T. A.、Tinto, M.、Larson, S. L. 和 Armstrong, J. W. LISA最优灵敏度。《物理评论*D*》 **66**（2002年）。网址 https://doi.org/10.1103/PhysRevD.66.122002。

[52] Tinto, M. 和 Dhurandhar, S. V. 时间延迟干涉测量。 《相对论生活评论》**17**（2014年）。网址 https://doi.org/10.12942/lrr-2014-6。

[53] Vallisneri, M. 合成LISA：在模型LISA中模拟时间延迟干涉测量。《物理评论*D*》**71**（2005）。网址：https://doi.org/10.1103/PhysRevD.71.022001。

[54] Cornish, N. J. 与 Rubbo, L. J. LISA响应函数。《物理评论*D*》**67**（2003）。网址：https://doi.org/10.1103/PhysRevD.67.022001。

[55] Amaro-Seoane, P. 等。 激光干涉空间天线。*arXiv*预印本 （2017年）。网址 https://arxiv.org/abs/1702.00786。

[56] Luo, Z.、Guo, Z.、Jin, G.、Wu, Y. 和 Hu, W. 太极计划简要分析：科学与技术。物理学成果 **16**，102918（2020年）。网址 https://doi.org/10.1016/j.rinp.2019.102918。

[57] Durkan, C., Bekasov, A., Murray, I. & Papa- makarios, G. 神经样条流。 神经信息处理系统进展 **32**，7509–7520（2019年）。网址 https://arxiv.org/abs/1906.04032。

[58] Chua, A. J. K. & Vallisneri, M. 使用神经网络学习引力波推断的贝叶斯后验分布。《物理评论快报》 **124**，041102（2020年）。

[59]˘Sid´ak, Z. 多元正态分布均值的矩形置信区域。《美国统计学会杂志》 **62**，626–633（1967年）。

[60] Gair, J. R., Porter, E., Babak, S. & Barack, L. 在模拟 LISA 数据挑战赛 1B 中针对极端质量比旋近双星的约束 Metropolis–Hastings 搜索。 《经典与量子引力》 **25**，184030（2008年）。网址 https://doi.org/10.1088/0264-9381/25/18/184030。

[61] Babak, S.、Gair, J. R. 与 Porter, E. K.。用于探测 LISA 数据中极端质量比旋近的算法。《经典与量子引力》**26**，135004（2009）。

方法

源模型、先验与模拟观测

我们使用增强解析拼接模型（AAK） [27]，对源进行建模，采用 FastEMRIWaveforms（FEW）中的`Pn5AAKWaveform`实现 [26]，，该实现将 AAK 波形构造与五阶后牛顿内旋轨迹相结合。一个一般的自旋极端质量比旋近由 14个参数描述：大质量黑洞和致密天体的质量 M 和μ ；无量纲自旋 a；初始半通径 p_0 （以 M为单位）、偏心率 e_0 和倾角变量 $Y_0 \equiv \cos\iota_0 = L_z/\sqrt{L_z^2+Q}$，其中 L_z 和 Q 分别为轴向角动量和卡特常数；光度距离D_L；源的黄道坐标系极角和方位角（θ_S、ϕ_S），以及自旋的黄道坐标系极角和方位角（θ_K、ϕ_K），在 FEW 中及某些图中分别记为q_S 和 q_K；以及初始方位、极向和径向轨道相位$\Phi_{\phi,0}$、 $\Phi_{\theta,0}$ 和 $\Phi_{r,0}$。这些质量是探测器参考系中的（红移后的）质量。根据普朗克宇宙学，该源位于 $z \approx 0.40$，因此其源参考系质量小 $1+z \approx 1.40$，且其精度仅与红移的精度相当。自旋和轨道要素均为无量纲且与参考系无关。

我们称之为完整先验的先验在这些坐标中是均匀且可分离的（扩展数据表2）。极角服从均匀分布而非各向同性分布，但在 θ_S 和 θ_K 的 68% 区间内，各向同性权重 $\sin\theta$仅变化 0.6% 和 3.4%；而与 D_L^2 成正比的距离先验会使距离后验移动不到其标准差的 2%。

观测数据使用 LISA数据挑战 (LDC) [49] 和太极计划数据挑战 (TDC) [50] 模拟框架进行模拟。波形以$\Delta t = 10$秒的采样间隔生成，覆盖一个恒星年（$N_{\rm obs} = 3,155,814$ 个采样点），并以第二代时间延迟干涉测量 (TDI) 的形式投影到 LISA（位于欧洲空间局拖曳轨道上、臂长为2.5×10^9 米）和太极计划（位于其任务轨道上、臂长为 3.0×10^9米）随时间变化的响应上，其中包括星座的轨道运动及其随时间变化的臂长 [10, 23, 51–56]。推断使用两个探测器的噪声正交 A 和 E 通道。每个响应均乘以渐变比例为 0.05 的 Tukey窗，补零至 $N_{\rm FFT} = 4,194,304$ 个采样点后进行傅里叶变换，并在四个通道中分别独立加入功率谱密度 (PSD) 为 $S_{n,Ic}(f)$ 的探测器 I 和通道 c 噪声，该噪声为平稳高斯噪声（补充材料第 1 节）。注入信号在 LISA 中的最优信噪比 (SNRs) 为 46.18，在太极计划中为 122.56，网络信噪比为 130.98（扩展数据图 2）。

网络似然与距离剖面分析

例如，对于探测器 I 的 TDI通道 c 中的两个时间序列 x 和 y，即数据和模板，噪声加权内积为

$$(x|y)_{Ic} = \frac{4\Delta t}{N_{\rm FFT}}\,\mathrm{Re}\sum_{k\in\mathcal{K}_{Ic}}\frac{\widetilde{x}^{*}_{Ic,k}\,\widetilde{y}_{Ic,k}}{S_{n,Ic}(f_k)}, \qquad (1)$$

其中，$\widetilde{x}_{Ic,k}$ 是 x 在频率 $f_k = k/(N_{\rm FFT}\Delta t)$ 处的未归一化离散傅里叶变换；星号表示复共轭；$S_{n,Ic}$ 是单边噪声功率谱密度；$\mathcal{K}_{Ic}$ 是屏蔽 TDI传递函数零点后保留的正频率频点集合。X 和 Y 分别是 x 和 y 的连续傅里叶变换，满足 $X(f_k) \approx \Delta t\,\widetilde{x}_{Ic,k}$ 以及 $Y(f_k) \approx \Delta t\,\widetilde{y}_{Ic,k}$；频点间隔为 $\Delta f = 1/(N_{\rm FFT}\Delta t)$。因此，前因子 $4\Delta t/N_{\rm FFT} = 4(\Delta t)^2\Delta f$ 使该求和近似于标准连续内积 $4\,\mathrm{Re}\int_0^\infty X^*(f)\,Y(f)/S_{n,Ic}(f)\,\mathrm{d}f$。模板与数据使用相同的加窗函数和零填充。网络内积 $(x|y)_{\rm net}$ 对两个探测器和两个通道的方程（1）求和。在其他 13 个参数$\boldsymbol{\theta}_{13}$固定时，波形振幅按 $1/D_L$缩放，因此我们写作 $h = \alpha\,h_0(\boldsymbol{\theta}_{13})$，其中 $\alpha = D_{\rm ref}/D_L$和 h_0是在 $D_{\rm ref} = 1$ 吉秒差距处计算的。除一个常数外，高斯对数似然为

$$\ln\mathcal{L}(\boldsymbol{\theta}_{13}, D_L) = -\tfrac{1}{2}\big[d_d - 2\alpha d_h + \alpha^2 h_h\big], \qquad (2)$$

其中，d 表示数据，$d_d = (d|d)_{\rm net}$ 表示数据功率，$d_h = (d|h_0)_{\rm net}$ 表示数据与参考模板之间的重叠，而 $h_h = (h_0|h_0)_{\rm net}$表示参考模板的功率，因此方括号内的量等于噪声加权平方残差 $(d-h|d-h)_{\rm net}$。在距离范围内对距离取最大值

$[D_{\min}, D_{\max}]$ 给出剖面似然

$$\ln \mathcal{L}_{\mathrm{prof}}(\boldsymbol{\theta}_{13}) = -\tfrac{1}{2}\big[d_d - 2\widehat{\alpha} d_h + \widehat{\alpha}^{\,2} h_h\big], \qquad \widehat{\alpha} = \mathrm{clip}\left(\frac{d_h}{h_h}, \frac{D_{\mathrm{ref}}}{D_{\max}}, \frac{D_{\mathrm{ref}}}{D_{\min}}\right), \tag{3}$$

其中 clip(x, a, b) =最小值{最大值{x, a}, b} 将最佳拟合振幅 d_h/h_h 限制在距离边界 $[D_{\min}, D_{\max}]$ 所允许的搜索域范围内。差分进化（DE）搜索和 13 维并行退火马尔可夫链蒙特卡罗（PTMCMC）运行均使用此剖面似然。它保留了精确的D_L^{-1} 振幅依赖关系，同时去除了一个采样维度。在线性信号近似下，对 D_L取最大值和边缘化会为其他参数给出相同的协方差。超出这一近似范围后，两者会因非高斯结构以及距离边界而有所不同。因此，所报告的区间来自一次独立运行，在该运行中对 D_L 进行采样，所用方程为（2）。

摊销全先验网络

第一个模型 F0 在完整先验上训练一次，并将观测数据映射为关于八个目标参数的概率分布：六个本征参数（M、μ、a、p_0、e_0、Y_0），它们决定轨道频率、其长期演化以及谐波的相对强度；D_L，它决定振幅；以及 $\log \rho_{\mathrm{net}}$，其中ρ_{net} 是为期一年的网络信噪比。天空位置、自旋方向和轨道相位通过较弱且更加退化的调制起作用。在这一阶段，它们保留完整取值范围，待内禀参数得到约束后再进入后续模型。训练集包含从完整先验中抽取的 50,000 个源。每次加载波形时，都会将其重新缩放为在 20 到 200 之间均匀抽取的网络信噪比，从而确定其距离；随后加入一次新的噪声实现，因此每个源都会以多种信号强度并在多种噪声实现中被观测到。一个由步幅卷积和十个 MMDiT Transformer模块组成的编码器首先作为异方差回归器进行训练，用于概括数据；随后，带有 20 个耦合变换的条件神经样条流 [57] 根据该概括结果对八个目标参数进行建模。将表征学习与密度估计分离，沿用了相关的引力波分析 [35, 36, 58]。架构和训练计划见补充材料第2节。

我们在 2,000 个独立信号上测试了 F0，这些信号的网络信噪比介于 30 和 200 之间，并为每个信号抽取 1,024 个样本。对于每个信号，八个目标参数按坐标计算的中心 1σ、 2σ 和 3σ 区间定义了一个方框，而注入值位于该方框内的信号比例分别为 16.2%、78.5% 和 98.0%。对于经过校准的高斯后验，这些比例必须介于八个相互独立目标参数对应的数值（4.7%、68.9% 和 97.9%）与单个目标参数对应的数值（68.3%、95.4% 和 99.7%）之间，因为相关性只会提高 [59], 这些比例。观测到的比例位于这些范围内，没有显示出区间过度自信的迹象。完整目标参数范围与 2σ 方框之比的中位数从信噪比 30–50 时的 68.2 上升到 3.43×10^9 信噪比 140–200 时的数值（扩展数据图1）。由于该方框包含对数 ρ_{net}，因此这一比值衡量的是训练目标空间中的收缩，而非物理体积。在一块 NVIDIA H20 GPU 上，编码器的一次前向传播和 1,024 个流样本的生成耗时不到 1 秒。

对于所分析的观测，七个物理目标参数的中心 2σ 区间定义了第一个模拟域，即 $M \in [9.62, 9.90] \times 10^5\, M_\odot$、 $\mu \in [31.8, 35.6]\, M_\odot$、 $a \in [0.937, 0.977]$、$p_0/M \in [11.94, 12.18]$、 $e_0 \in [0.289, 0.375]$、 $Y_0 \in [-0.394, -0.361]$ 和 $D_L \in [1.83, 2.62]$ 吉秒差距，其大小是完整先验的 2.02×10^8 分之一。

序贯细化与自旋扇区

每一轮后续模拟都会在当前模拟域内生成新的源，将条件密度重新拟合至观测结果，并更新模拟域。每个模拟域都是一个轴对齐盒 $B_r = \prod_{j=1}^{14} [\ell_j^{(r)}, u_j^{(r)}]$，每个坐标对应一个区间 j，其中 r 表示阶段（F0、R1–R7、F14-S 或 F14-R）。我们通过其相对于完整先验的收缩程度来衡量该模拟域 $\prod_{j=1}^{14} [\ell_j^{\mathrm{prior}}, u_j^{\mathrm{prior}}]$（扩展数据表 2），

$$C_r = \prod_{j=1}^{14} \frac{u_j^{\mathrm{prior}} - \ell_j^{\mathrm{prior}}}{u_j^{(r)} - \ell_j^{(r)}}. \tag{4}$$

每个因子都是某一坐标的先验宽度与其保留宽度之比，因此 C_r 是完整先验的体积除以 B_r的体积，而 10^8 的收缩意味着该盒体

积占据了先验体积的 10^{-8} 。对于周期角，保留宽度是所保留的弧长。量 C_r 是这种参数化下的盒体积比，而不是后验体积或信息度量。在 R1–R7 轮次中，首先缩小了距离区间，随后将天空位置和自旋方向限制在一个单元格内（R5），最后收紧六个本征参数（R7），使收缩率从 F0 时的 2.02×10^{8} 提高到 6.26×10^{10}。在整个过程中，轨道相位始终保留其完整范围（各轮次的规则和边界见补充材料第3节）。

高分辨率细化随后采用基于新模拟训练的新网络，这些模拟以一个候选结果为中心；该候选结果是在R7之后，通过对90天半相干统计量进行粒子群优化和相干细化得到的。在一年似然下，另外三个达到较强半相干极值的分支，其结果比纯噪声模型低3,700–5,000个对数似然单位；而保留的候选结果则高出3,812个单位，因此仅凭精确似然即可将它们区分开来。两个高分辨率网络都使用时频Transformer编码器，分析持续一年响应中最后140.5天的数据，这部分包含信号中频率最高、振幅最大的部分。第一个网络FlowA从以该候选结果为中心的参数域中的50,000次模拟里学习11个非相位参数。第二个网络Formal14使用其样本的第5和第95百分位数定义模拟域（F14-S），并从100,000次新的模拟中学习全部14个参数。设 $\boldsymbol{\eta}_{11}$ 为非相位参数， $\boldsymbol{\Phi}_0$ 为相位， $\boldsymbol{z}$ 为数据的嵌入，Formal14将其密度分解为

$$q_\psi(\boldsymbol{\eta}_{11}, \boldsymbol{\Phi}_0 \mid \boldsymbol{x}) = q_{\psi,\mathrm{np}}(\boldsymbol{\eta}_{11} \mid \boldsymbol{z})\, q_{\psi,\mathrm{ph}}(\boldsymbol{\Phi}_0 \mid \boldsymbol{\eta}_{11}, \boldsymbol{z}), \tag{5}$$

其中， $\boldsymbol{x}$ 表示网络分析的白化数据， ψ 表示网络权重。因子 $q_{\psi,\ \mathrm{np}}$ 是给定嵌入的非相位参数密度，由神经样条流建模；因子 $q_{\psi,\ \mathrm{ph}}$ 是给定非相位参数和嵌入的三个相位的密度，由三维环面上的条件圆样条流建模。这种因子分解并不假设相位与其他参数相互独立。Formal14针对该观测返回4,096个提议分布。

对于每个非相位坐标，搜索边界取这些提议分布与 F14-S 区间相交。自旋方向位于 F14-S 内的 3,968 个提议构成一个环（图 4b），其 k-均值聚类将自旋单位向量划分为四个扇区，分别包含 893、1,056、825 和 1,194 个成员。在这四个扇区内分别运行差分进化后，825 个成员的扇区达到最高似然，其中包含注入的自旋方向。我们采用该扇区，并将（θ_K， ϕ_K）边界设为其中央 99% 区间。所得搜索域 F14-R 比完整先验小 5.41×10^{17} 倍，而 F14-S 则小 1.39×10^{16} 倍（图 4a 和扩展数据表 2）。距离在 FlowA 区间内进行剖面化。在跟踪扇区中，有 697 个提议满足全部十一个非相位边界；从中选出 22 个代表候选源，取其距离和相位以外的十个坐标，并为每个候选源配对 128 个覆盖相位环面的扰乱 Sobol 点，由此得到 2,816 个初始状态。神经密度用于选择这些起始点，但不会参与似然计算。

对于依据似然比较候选区域的情况，例如R7之后的粒子群分支和四个自旋扇区，似然本身决定了要跟踪的区域。

精确似然搜索和双探测器筛选

直接从网络识别出的区域开始的一年搜索无法达到后验分布。在一年剖面似然下进行的三组差分进化（DE）分别先改变三个相位、再改变十个非相位坐标，最后改变全部13个坐标（补充材料第4节），最终停留在次级极大值处：其最佳候选项的对数似然分别比注入参数低9,560.1、9,341.0和9,250.0个单位，尽管每个坐标都接近其注入值。相对于仅含噪声的模型，注入参数高出8,657个单位，而最佳候选项低593.5个单位。因此，分析重新从30天数据开始。

终端13维种群（图中 R135）在每个探测器中初始化独立的30天差分进化搜索，并采用相位放宽统计量；该统计量在共同振幅、距离边界以及每个数据段的一个额外相位上最大化探测器 I 的似然，

并将结果与仅含噪声的情况进行比较：

$$\mathcal{S}_I(\boldsymbol{\theta}) = 2\Big[\max_{\alpha,\{\delta\phi_s\}} \ln \mathcal{L}_I\big(d_I \mid \alpha h_I(\boldsymbol{\theta};\{\delta\phi_s\})\big) - \ln \mathcal{L}_I(d_I \mid h_I = 0)\Big]. \tag{6}$$

此处 d_I 表示探测器 I的数据， $\mathcal{L}_I$ 仅限于该探测器的方程（2）的似然，$h_I(\boldsymbol{\theta};\ \{\delta\phi_s\})$ 是其在参考距离 $D_{\rm ref}$ 处、针对除距离之外的 13 个参数 $\boldsymbol{\theta}$ 的模板，并在数据段 s中施加了额外的相位偏移 $\delta\phi_s$ ，而 $\alpha = D_{\rm ref}/D_L$ 则是受距离边界约束的振幅缩放因子。第二项是仅噪声情况下的对数似然，因此，正分数意味着该模板对该探测器数据的拟合优于仅噪声。由于距离边界使振幅远离零，较差的模板可能得到负分数。我们为每个探测器使用一个 30 天的数据段，额外相位对两个通道均为共同相位，并且不同于三个轨道相位。每个探测器的搜索都会返回 2,048 个候选项，全部 4,096 个候选项都会在两个探测器中进行评估（扩展数据图 3）。其中，622 个在两个探测器中的分数均为正，417 个仅在 LISA 中为正，162 个仅在太极计划中为正，2,895 个在两个探测器中均不为正。这 622 个候选项按照其两个分数中较小者进行排序，每个分数均以其在候选库中的最大值归一化，从而使一个探测器中的高分无法弥补另一个探测器中的低分。作为参考，注入参数在 LISA 和太极计划中的分数为（59.22，257.08），归一化后为（1.016，1.266）。 k-均值聚类将这 622 个候选项分为 16 组，其中排名最佳的成员分别作为 16 个种群的锚点；每个种群包含 128 个粒子，并使用两个探测器共同的一个波形和一个距离，进行 20 代联合 30 天差分进化。由此，筛选过程塑造了联合搜索的初始种群，但既不改变其边界，也不改变其目标。我们尚未针对同等成本下未经筛选的初始化，测量该筛选过程的收益。

相干延续与一年采样

联合分布搜索继续针对持续时间为 30、60、180 和 360 天的相干观测进行，随后再进行一次年。对于持续时间 T，目标是

$$\ell_T^{\rm prof}(\boldsymbol{\theta}_{13}) = \max_{D_{\rm min}\le D_L\le D_{\rm max}} \ln \mathcal{L}_T(d_T \mid \boldsymbol{\theta}_{13}, D_L),$$
$$\widetilde{p}_T(\boldsymbol{\theta}_{13} \mid d_T) \propto \pi_{13}(\boldsymbol{\theta}_{13}) \exp\big[\ell_T^{\rm prof}(\boldsymbol{\theta}_{13})\big], \tag{7}$$

其中， d_T 表示限制在相干时长 T内的 $\mathcal{L}_T$ 在这些数据上计算的公式（2）中的双探测器似然，$\ell_T^{\rm prof}$ 表示对距离取最大值后的 $\ln \mathcal{L}_T$ ，表示固定搜索域上的均匀先验， π_{13} 而$\widetilde{p}_T$ 表示时长为 T时得到的剖面后验分布。每个 T时长下都会重新构建探测器响应和傅里叶变换，两个探测器共用一个波形和一个距离。在每个有限时长下，差分进化会初始化一次简短的并行温度马尔可夫链蒙特卡罗运行；其中按似然排序且彼此分离的冷链代表候选源将作为下一时长的种子。由于数据不同，因此不会比较不同时长下的似然值。对于一年时长，依次进行三次并行温度马尔可夫链蒙特卡罗运行，温度数分别为八、十二和十六；每次运行都以前一次运行的冷链为种子。最后一次运行在每个温度下使用 256 个行走者，进行 2,000 次迭代；舍弃前 1,000 次后，留下 256,000 个相关的冷链状态（所有运行的设置见补充材料第5节）。所有运行均采用带自适应温度阶梯的集合拉伸移动，报告的分布仅使用冷链。

在最终运行中，冷链拉伸接受率为20.4%，相邻温度之间的交换接受率介于27.2%和92.7%之间（扩展数据图4）。预热期之后，冷链似然保持平稳，全部16个种子盆地仍处于占据状态，保留链前后两半部分的盆地占据率相差总变差距离的4.4%，并且13个坐标中最大的分裂R为1.0085。这些盆地的主要差异在于自旋方向和方位相位，它们分别覆盖搜索宽度的52–91%；而联合后验分布的本征边缘分布呈单峰分布（图1）。

延长相干时长对内禀参数的精度提升，远超 SNR 增益所能带来的幅度。在 30天时，($M/10^6 M_\odot$、 $\mu/M_\odot$、 a、 p_0/M、 e_0、 Y_0) 的 68% 半宽为 (5.9×10^{-4}、0。067、 7.2×10^{-4}、 3.4×10^{-3}、 8.5×10^{-4}、 8.0×10^{-4})。一年的显式距离区间窄了 13–69 倍，而 SNR 本身仅增长了 $\sqrt{365/30}\approx3.5$。一年的

剖面样本按下文所述进行了截断，其给出的倍数为 23–120。额外的增益来自轨道频率的长期演化，只有追踪整整一年的相位才能分辨这种演化。自旋方向宽度并非单调缩小，因为它们逐渐接近该扇区搜索域的界限。非周期性边缘分布以其中位数和中央 68.27% 区间概括，而相位则以相对于最终剖面链圆均值的偏移表示，$(\Phi_{\phi,0}, \Phi_{\theta,0}, \Phi_{r,0})_{\mathrm{ref}}$ =(3。780632、0。852990、6。119448)rad。各阶段的区间和边缘分布见补充材料第5节。

显式距离采样与报告区间

剖面化在距离上进行优化，而不是对其积分，且剖面链被限制在搜索域内。为获得所报告的后验分布，我们因此使用 D_L 方程（2）的似然进行显式采样，同时保持数据、波形、噪声加权和先验形式不变。在首次显式距离运行中，样本在 μ、D_L 和ϕ_K的域边界处聚集，因此我们将全部十一个非相位区间以其中心为基准扩大了 1.5 倍（$D_L \in [2.18133, 2.70303]$ 吉秒差距；相位不变），并从最终剖面链重新开始运行。该运行使用 16 个温度，每个温度使用 256 个行走者。在计划的 2,000 次迭代中运行 914 次后停止；我们采用第 451–900 次迭代，这些迭代产生了 115,200 个冷链状态（图 1 和扩展数据表 2）。

六个内禀参数的显式距离区间比剖面链对应区间宽 1.73–1.78 倍，而天球角区间的差异在 2% 以内。在差分信号近似中，剖面化与边缘化给出的协方差相同，因此剖面化无法解释这种差异。其原因在于原始域的边界。其下 μ 界 $33.1108\,M_\odot$ 仅比注入值 $0.0018\,M_\odot$低，但高于显式距离中位数 μ， $33.1104\,M_\odot$，并且约 40% 的显式距离样本低于原始距离上限 2.268 吉秒差距，此处剖面振幅会被截断。由于六个内禀参数高度相关（图1），μ边界沿主退化方向截断了剖面链，这解释了剖面区间的不对称性

以及其相对于显式距离中位数共同偏移约 0.7 个标准差的现象。因此，剖面链用于定位并连接各个盆地；所有报告的区间，包括表1中的精度，均来自显式距离运行。这些精度是显式距离边缘分布中央 68% 区间的半宽除以其中位数；对于天球角，则采用绝对半宽。

由于该次运行提前停止，其收敛情况不如剖面链得到的确认那样可靠。三项观察支持所报告的区间。除自旋方向和 $\Phi_{\phi,0}$ 外，各边缘分布都平滑且呈单峰（图 1）。所有中位数与相应剖面中位数的距离均在 0.75 个标准差以内。除 θ_K 和 ϕ_K 外，没有任何坐标的样本落在扩宽边界 2 % 范围内的比例超过 0.001%。自旋方位角和 $\Phi_{\phi,0}$ 通过环形简并呈强相关，且仍受跟踪扇区的范围限制，因此它们的区间以该扇区为条件，整体测量结果亦然。

全部六个本征参数以及 D_L 均落在其 68% 区间内。天球角 θ_S 和 ϕ_S 分别偏离 2.1 和 2.4 个标准差； ϕ_K 和 $\Phi_{\phi,0}$ 则略微超出各自区间，距中位数分别为 1.2 和 1.3 个半宽。同样这四个坐标也落在剖面链的区间之外。十四个坐标中有四个超出其 68% 区间，这与校准区间预期的 4.4 个相符；对于十四个坐标，其中一个偏离至少 2.4 个标准差的情况约会在 20% 的噪声实现中发生。单次实现无法检验校准，因为这需要重复注入。

与先前极端质量比旋近分析的比较

扩展数据表1 比较了 NEBULA-EMRI 与此前能够给出后验不确定性的端到端EMRI分析。Strub 等人 [37]在 0.44 年的 LISA 数据中，从宽质量和偏心率先验出发，通过时间–频率搜索，随后以搜索最大值为起点，对 11 个参数进行马尔可夫链蒙特卡罗（MCMC）采样，恢复了一个非自旋的偏心极端质量比旋近双星。Cole 等人 [38] 使用截断边缘神经比率估计，将无自旋极端质量比旋近双星的 11 维先验收缩了 10^6–10^7 倍，在

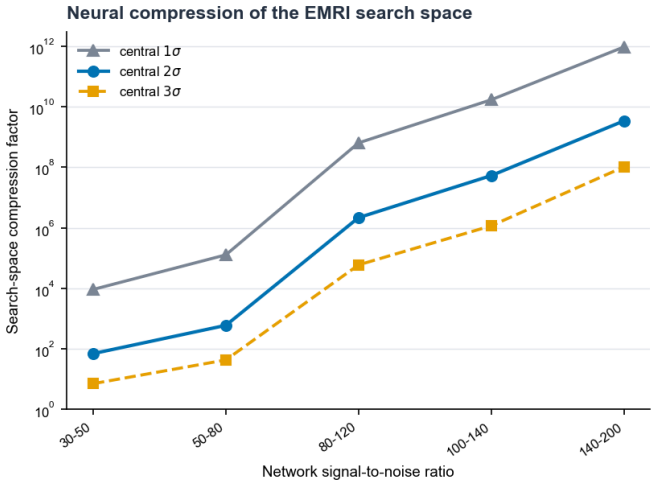
0.1 年的数据中获得了一维边缘提议分布，而不是联合后验分布。我们此前关于流匹配马尔可夫链蒙特卡罗的研究 [40]从太极计划 60 天的数据和宽先验出发，恢复了一个 AAK 源四个内禀参数的联合后验分布；其中，外禀参数初始化在其注入值附近，并表明从宽先验出发的并行温度马尔可夫链蒙特卡罗会被次级极大值困住。这些分析均针对该问题的简化版本：它们忽略了自旋，限制了从宽先验推断的参数数量，或分析的单探测器数据最多只有半年。

返回点估计的搜索也取得了进展 [41, 60, 61]。最近，Zou等人 [41]对采用原始解析简化模型生成的半年期单探测器信号的全部14个参数进行了相干搜索。与AAK的轨道频率不同，该模型的轨道频率并未映射到Kerr测地线上 [27]。他们报告称，10^{-3}–10^{-2} 相位演化参数的相对误差约为3%，距离误差低于3%，天空位置误差小于0.1弧度，但未能恢复自旋方向或初始相位。点估计既不包含可信区间，也不包含参数协方差，因此此类结果未纳入扩展数据表1。此外，点估计也无法保证能够计算出后验分布：在我们的分析中，即使优化得到的候选项在每个坐标上都接近注入信号，在一年似然下，其对数似然值仍比注入信号低9,250–9,560个单位，因此从这些候选项出发的采样器会探索一个次级盆地 [40]。要得到后验分布，需要进行双探测器筛选、相干延续和多盆地采样。不同研究之间的数值比较应考虑其信噪比、观测时长、波形模型和探测器配置的差异。

扩展数据

扩展数据表1 可提供后验不确定性的端到端EMRI分析：所解决的推断问题与可探测的物理学。 上半部分描述推断问题，下半部分列出各后验分布所支持的物理测量。参数是根据宽先验推断得出的。 ✓，可探测； ×，不可探测。

	Strub 等人 [37]	Cole 等人 [38]	Liang等人 [40]	**本研究**
Inference problem				
一般自旋源	×	×	✓	✓
从宽	11	11	4	**14**
先验	宽质量和偏心率先验	宽泛的 11 维维度先验	宽内禀先验；注入附近的注入	**完整先验（扩展数据表2）**
相干数据	0.44 年	0.1 年	60 d	**1 yr**
探测器	LISA	LISA	太极计划	**LISA + 太极计划**
波形	相对论的 ，非自旋的	相对论的 ，非自旋的	AAK	**AAK**
后验分布	联合分布，11参数	一维边缘分布提议分布	联合分布，4个内禀参数	**联合分布和多盆地，全部14个参数**
Physics accessible from the posterior				
黑洞自旋	×	×	✓	✓
俘获通道（偏心率和倾角）	×	×	×	✓
暗标准汽笛（距离和天空位置）	✓	仅边缘分布	×	✓
所有源参数	仅无自旋	×	×	✓
在独立探测器中的确认	×	×	×	✓

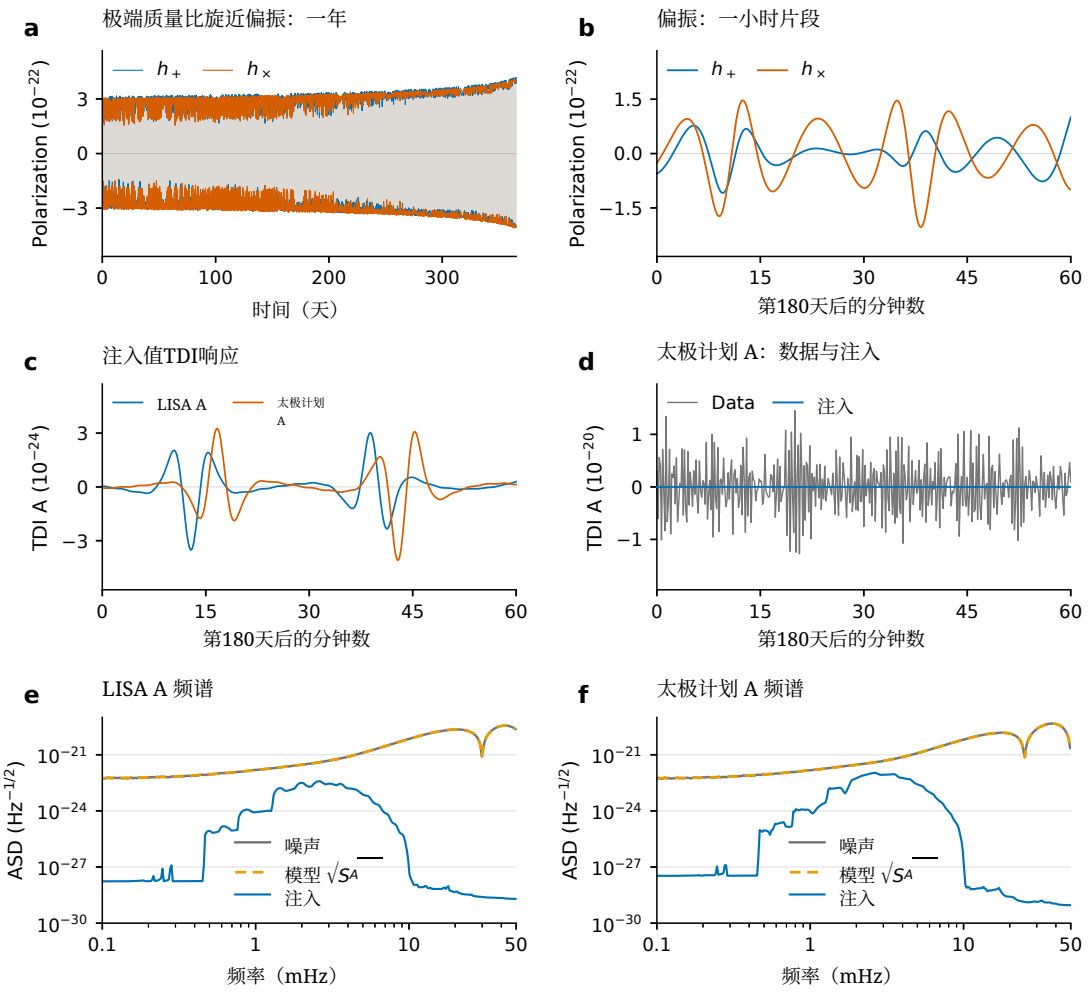


扩展数据图 1 摊销搜索区域的收缩程度随信噪比增大。对于 2,000 个源参数各异且加入了新噪声的独立模拟信号，按网络信噪比（SNR）分窗，给出摊销网络生成的搜索支持域收缩程度的中位数。80–120 和 100–140 两个窗口存在重叠。对于每个信号，从条件流中抽取 1,024 个样本，以八个目标参数（七个物理参数和网络信噪比的对数）各坐标的中央 1σ,2σ 或 3σ 区间确定一个包围盒，收缩程度为完整目标范围与该包围盒的比值。在所有信号中，注入值落在 1σ、 2σ 和 3σ 包围盒内的比例分别为 16.2%、78.5% 和 98.0%。线条仅供视觉参考。

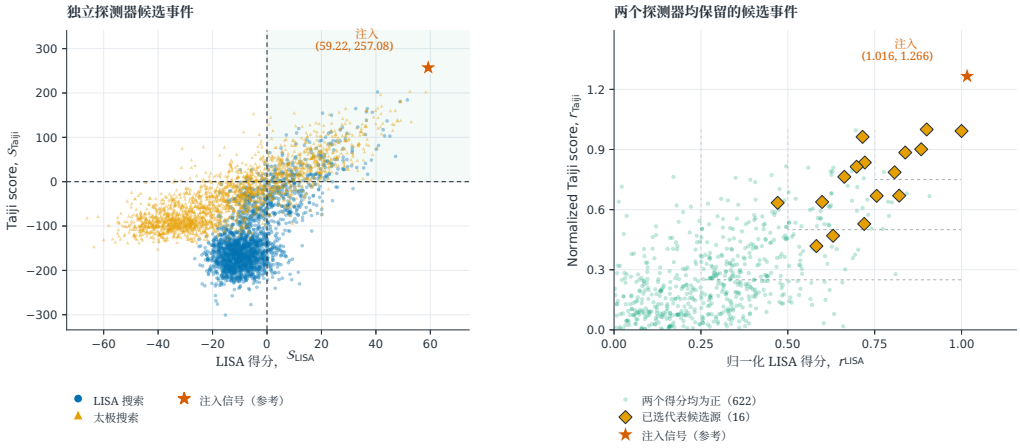
扩展数据表2 先验、搜索域和后验区间。 先验：完整先验，各参数均服从均匀分布。搜索域：精确似然搜索和剖面链所用的 F14-R 区域。注入值：模拟输入。剖面：为期一年、包含十六种温度的剖面似然链（256,000 个状态）。显式距离：在中心相同、范围扩大 1.5 倍的域内对 D_L 进行采样的运行（115,200 个状态；图 1）。表中给出中位数和中央 68.27% 区间。精度：显式距离区间半宽除以 M–D_L的中位数；角度则给出绝对半宽（弧度）。质量的单位与首列相同，角度单位为弧度。相位是相对于 $(3.780632, 0.852990, 6.119448)$ 弧度的偏移量，而其先验和搜索域指的是相位本身。下 μ限和距离上限会截断剖面链（方法）。

参数	先验	搜索域	注入值	剖面（13D）	显式距离（14D）	精度
$M/(10^6 M_\odot)$	$[0.5, 1.0]$	$[0.9767833, 0.9794250]$	0.97742962	$0.9774322^{+0.0000079}_{-0.0000140}$	$0.9774451^{+0.0000188}_{-0.0000192}$	1.9×10^{-5}
$\mu/M_\odot$	$[10, 50]$	$[33.11076, 33.37859]$	33.112520	$33.1126^{+0.0023}_{-0.0014}$	$33.1104^{+0.0033}_{-0.0032}$	9.7×10^{-5}
a	$[0.1, 0.999]$	$[0.941609, 0.953381]$	0.94544326	$0.9454428^{+0.0000160}_{-0.0000098}$	$0.9454266^{+0.0000234}_{-0.0000226}$	2.4×10^{-5}
p_0/M	$[10, 16]$	$[12.002577, 12.028257]$	12.0177123	$12.01771^{+0.00019}_{-0.00011}$	$12.01753^{+0.00027}_{-0.00026}$	2.2×10^{-5}
e_0	$[0.1, 0.9]$	$[0.314074, 0.324895]$	0.32030659	$0.3203026^{+0.0000056}_{-0.0000086}$	$0.3203116^{+0.0000122}_{-0.0000125}$	3.9×10^{-5}
Y_0	$[-1, 1]$	$[-0.385872, -0.372277]$	-0.38479048	$-0.384792^{+0.000032}_{-0.000019}$	$-0.384823^{+0.000046}_{-0.000044}$	1.2×10^{-4}
D_L/吉秒差距	$[0.45, 3.0]$	$[2.26828, 2.61608]$[a]	2.281054	已剖面化	$2.2720^{+0.0179}_{-0.0184}$	8.0×10^{-3}
θ_S	$[0, \pi]$	$[2.04141, 2.36553]$	2.2516818	$2.2442^{+0.0036}_{-0.0037}$	$2.2439^{+0.0038}_{-0.0036}$	3.7×10^{-3}
ϕ_S	$[0, 2\pi)$	$[0.68328, 0.93374]$	0.7913202	$0.7984^{+0.0030}_{-0.0029}$	$0.7986^{+0.0030}_{-0.0030}$	3.0×10^{-3}
θ_K	$[0, \pi]$	$[1.15476, 1.54627]$	1.4189733	$1.451^{+0.040}_{-0.098}$	$1.431^{+0.058}_{-0.129}$	0.093
ϕ_K	$[0, 2\pi)$	$[3.39967, 4.31087]$	4.2494406	$3.79^{+0.36}_{-0.29}$	$3.78^{+0.42}_{-0.34}$	0.38
$\Delta\Phi_{\phi,\,0}$	$[0, 2\pi)$	$[0, 2\pi)$	0.641799	$-0.03^{+0.50}_{-0.41}$	$-0.05^{+0.59}_{-0.51}$	0.55
$\Delta\Phi_{\theta,\,0}$	$[0, 2\pi)$	$[0, 2\pi)$	0.010602	$0.000^{+0.029}_{-0.030}$	$-0.014^{+0.034}_{-0.033}$	0.034
$\Delta\Phi_{r,\,0}$	$[0, 2\pi)$	$[0, 2\pi)$	0.015264	$0.000^{+0.025}_{-0.025}$	$0.017^{+0.030}_{-0.030}$	0.030

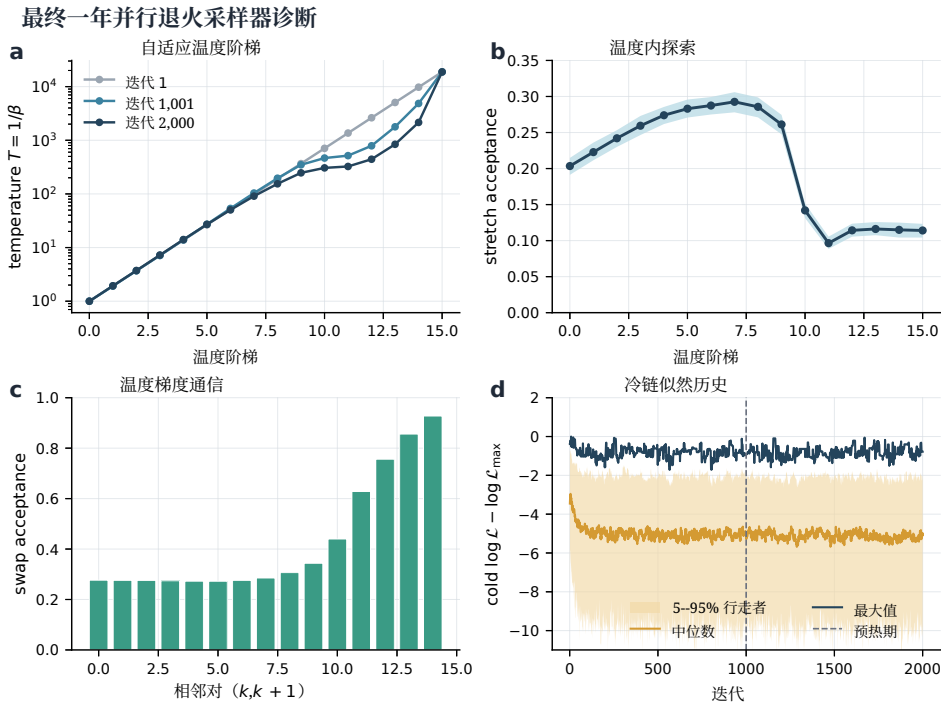
[a]13参数运行的剖面区间。显式距离运行对 $D_L \in [2.18133, 2.70303]$吉秒差距进行采样。



扩展数据图 2 模拟的极端质量比旋近信号与探测器噪声。**a**，注入参数下入射偏振的全年包络 h_+ 和 $h_\times$。色带表示 2,000 个时间区块中每个区块的最小值和最大值。**b**，从第 180 天开始一小时内的两种偏振。**c**，同一时间段内无噪声 LISA 和太极计划 A通道 TDI响应。**d**，太极计划 A通道数据与处于同一振幅缩放因子下的注入响应。**e**、**f**，LISA 和太极计划 A通道注入响应及噪声实现的振幅频谱，并与噪声模型进行比较。曲线表示在 320 个对数频率分箱中取平均后的功率平方根，模型 PSD 也在相同分箱上取平均。对于内部频率分箱，周期图为 $P_x(f_k) = 2\Delta t\,|\widetilde{x}_k|^2/N_{\mathrm{FFT}}$。这里以 A 通道为例。推断同时使用每个探测器中的 A 和 E。未白化时间序列由噪声主导，而相干分析则在全年时间内累积信号。



扩展数据图3 30天候选源的双探测器筛选。左图：独立搜索所得4,096个候选项在LISA和太极计划中的得分（方程式（6））。蓝色圆圈和金色三角形分别标记在LISA搜索和太极搜索中找到的候选项。得分为在噪声单独存在的情况下，额外对一个相位和一个有界振幅进行最大化后，对数似然改进的两倍。零值处的虚线框出在两个探测器中均为正值的622个候选项。右图：分别用每个探测器中的模板库最大值归一化后保留的候选项。金色菱形标记16个代表候选源，L形等高线表示$\min(r_{\rm LISA}, r_{\rm Taiji}) = 0.25$、0.50和0.75。朱红色星号标记注入参数，其得分为（59.22, 257.08），归一化得分为（1.016, 1.266）。注入值不属于用于归一化的模板库，因此其归一化得分可以超过1。



扩展数据图4 最终为期一年的剖面似然PTMCMC运行诊断。**a**，第1次、第1,001次和最后一次迭代时的自适应温度阶梯。**b**，各温度下的拉伸移动接受率。**c**，相邻温度之间的交换接受率。**d**，冷链对数似然相对于其最大值的变化：行走者中的最大值和中位数，以及5–95%的范围。虚线标示丢弃的1,000次迭代结束的位置。