arXiv:2512.18290v2 [astro-ph.IM] 26 Dec 2025

针对存在缺口的引力波信号，稳健且可扩展的基于模拟的推断

Ruiting Mao，[1] Jeong Eun Lee，[1] 以及 Matthew C. Edwards[1]

[1] *Department of Statistics, University of Auckland, Auckland 1010, New Zealand*

激光干涉仪空间天线（LISA）的数据流不可避免地会因维护和环境扰动而出现缺口，从而引入非平稳性和频谱泄漏，影响标准的频域似然评估。我们提出了一种可扩展的基于模拟的推断（SBI）框架，能够直接从含间隙时间序列数据中稳健地进行参数估计。我们采用流匹配后验估计（FMPE），以学习得到的数据摘要为条件，并通过端到端训练策略对其进行优化。为应对长持续时间信号带来的计算挑战，我们提出了一种双路径摘要器架构：一条路径采用处理时域数据的一维卷积神经网络（CNN），以实现高精度；另一条路径采用新颖的基于小波的二维 CNN，利用非对称膨胀核，实现对跨越数月的数据集的可扩展处理。我们在模拟的类银河系双星信号上验证了该框架的有效性，结果表明，与两阶段重建流程相比，我们的联合训练方法能够得到更紧致且无偏的后验分布。此外，我们首次进行了系统比较，证明在存在严重数据伪影时，FMPE 相较于传统归一化流具有更优的稳定性和覆盖率校准。

关键词：激光干涉仪空间天线、基于模拟的推断、缺口、流匹配、降维。

I. 引言

激光干涉仪空间天线（LISA）[1, 2]，于 2024 年经欧洲空间局（ESA）正式批准，是首个旨在探索毫赫兹频带引力波宇宙的空间观测站（0.1 毫赫兹–1 赫兹）[3]。与观测赫兹至千赫兹频段事件的现有地基探测器（如 LIGO、Virgo 和 KAGRA）不同，LISA 将在由长持续时间信号主导的工作范围内运行，这些信号会在其敏感频带中持续数月甚至数年 [4]。因此，产生的数据流中将在时域和频域同时包含大量重叠源，包括大质量黑洞双星（MBHBs）的并合 [5, 6]，极端质量比旋近（EMRIs） [7, 8]，以及银河系双星（GBs）形成的连续前景 [9–11]。因此，提取科学信息需要进行“全局拟合”分析，以同时分辨这些相互交织的信号 [12]，；这项任务尤其容易受到不可避免的数据缺口和毛刺所引发的频谱泄漏与非平稳性的影响。 [4, 13–15]。

在长期运行过程中，偏离连续、不中断的观测数据采集这一理想标准是不可避免的。LISA 探路者任务[16, 17]，表明，数据流会因计划内维护（如天线重新指向和机构调整）以及非计划的环境干扰（如微流星体撞击或仪器故障）而出现中断 [4, 14]。在标准的引力波推断中，通常假设噪声平稳，并在频域进行分析，因此可以高效地将噪声协方差矩阵近似为对角矩阵 [14]。然而，这些间断会破坏噪声的时间平移不变性，使噪声变为非平稳，并在频域中引发频谱泄漏 [13, 18]。因此，噪声协方差矩阵不再具有对角结构，这意味着标准的傅里叶域方法（具体而言，即 Whittle 似然）在统计上不再一致 [4, 19]。若不加以处理，这些伪影可能导致天体物理参数估计出现偏差，并造成信噪比损失 [20],；这一风险已在银河系双星 [21] 和大质量黑洞双星中得到充分记录，尤其是在缺口出现在信号并合附近时 [22, 23]。

为应对……带来的分析挑战，

鉴于引力波信号中相位相干的复杂特性，通用信号处理中的标准技术（例如线性插值或均值插补）通常不足以应对这些数据缺口 [24]。因此，信号处理中常见的标准方法是加窗，即应用平滑窗函数来抑制高频频谱泄漏 [19, 21, 22]。然而，Burke等[4]近期的研究对这一做法进行了批判性再评估，指出加窗会从根本上改变噪声的随机性质。由于破坏了标准 Whittle 似然所要求的平稳性（Toeplitz）条件，加窗会引入理论上的不一致，从而可能使参数估计产生偏差。此外，尽管平滑渐变窗可以减轻这些统计误差，但代价是抹去渐变窗旁瓣中编码的信息 [4]。其他方法则尝试重建缺失数据，例如贝叶斯数据增广方法 [13], ，该方法将缺失数据视为辅助变量，与物理参数一同采样。尽管在统计上是一致的，但对于长持续时间信号，由于需要对条件分布执行密集矩阵运算，该方法的计算成本高得难以承受。近期，Pearson和Cornish [14] 提出了一种在小波域中更高效的数据增广方案；然而，该方法的有效性依赖于一个严格假设，即噪声在小波时频像素内保持局部平稳。在我们此前的工作中 [23], ，我们采用深度学习框架解决了这些计算瓶颈：该框架结合去噪自编码器以捕捉信号的全局结构，并使用双向门控循环单元（Bi-GRU）提取时间特征。该方法通过两阶段流程处理数据缺口：首先插补缺失片段，随后执行马尔可夫链蒙特卡罗（MCMC）采样以进行参数估计。然而，该方法对插补的有效实施高度敏感，而这要求输入数据预先去噪，因此需要一个中间模型或流程，将缺口插补与噪声实现解耦，导致效率和灵活性不足。

随着深度学习模型的发展，基于模拟的推断（SBI）取得了显著进展；当似然函数难以处理或计算成本过高时，它提供了一种稳健的替代方案 [25]。它无需显式计算似然函数，转而通过前向模拟直接学习后验分布。在引力波天文学领域，这种方法已成功应用于多种生成架构，包括条件变分自编码器（CVAEs） [26] ，以及尤为重要的归一化流（NFs） [27–30]，并最终催生了 DINGO [31] 等框架。这些模型学习将简单的基分布映射到复杂的高维后验分布，从而实现快速、摊销式的学习，其准确度可与随机采样器相媲美 [32]。至关重要的是，由于 SBI 依赖数据生成过程，而非僵化的数学噪声模型，因此在处理复杂仪器伪影时具有独特优势。近期研究表明，基于流和基于得分的估计器可直接利用实测探测器数据进行训练，即使存在非平稳噪声、毛刺和功率谱密度（PSD）漂移，也能稳健地推断参数；而在这些情况下，标准的高斯似然假设必然会失效[33–36]。此外，近期引入的连续时间生成模型，如流匹配（FM） [37, 38] 和连续归一化流（CNF） [39, 40], ，有望进一步提升这些推断流程的可扩展性和表达能力。

由于具有可处理的后验密度，且能够自由设计非扩散路径，与神经后验分数估计（NPSE）等其他迭代生成方法相比，流匹配后验估计（FMPE）在采样效率方面表现突出 [34, 38]。与地面探测器不同，时域 LISA 数据的维度极高，因此必须进行有效降维。主成分分析（PCA）和自编码器等传统降维方法，往往更重视噪声特征，而非底层信号成分 [41], ；当信噪比较低时，这一问题会更加严重[42]。频域表示结合谱白化，使一维卷积降维编码器得以应用[40]；文献 [41]提出了一种旨在捕获具有物理意义特征的摘要统计框架。然而，由于依赖快速傅里叶变换（FFT），该方法在存在数据缺口时的实际适用性受到限制。对于特征提取，可以考虑将.....分离

嵌入网络的训练阶段与推理阶段 [43, 44]分离。文献 [42], 探讨了非线性方法，展现了应用神经网络的灵活性和表达能力。近似贝叶斯计算（ABC）并非通过总结数据来保留其结构特征，而是致力于学习对参数充分的摘要统计量。拒绝采样中的低维摘要统计量最初作为精确拒绝采样的扩展方法提出，旨在解决匹配高维数据向量时效率低下的问题 [45, 46]。随后，Fearnhead 等人 [47]开发了用于选择这些统计量的半自动框架，并从理论上证明，在二次损失函数下，参数的后验均值就是充分摘要统计量。在此基础上，Jiang 等人 [48] 利用神经网络对观测数据进行参数回归，以自动提取这些统计量。我们的流程在 SBI 中运用这一概念，通过联合优化摘要网络和推理引擎来实现这一点。如第IVC节所示，这种端到端训练只专注于最大化与参数相关的信息，因此与相互独立的两阶段特征提取方法相比，能够降低后验不确定性。

鉴于存在数据缺口时基于快速傅里叶变换的方法具有局限性，我们采用时域数据作为所提出模型的输入。对于足够长但计算上仍可处理的信号，我们使用大步幅的一维卷积神经网络（1D CNN）；我们先前的工作已证明，该网络能够直接从时间动态中学习判别特征 [23], ，并直接处理原始时间序列。对于长度极长、计算开销极大的信号，受 [14]和 [49], 的启发，我们首先通过连续小波变换（ CWT）将存在缺口的数据流转换为时频表示。随后，利用具有非对称卷积核的深度扩张二维卷积神经网络（2D CNN）处理该时频图。这一路径擅长大幅降维，使我们能够从超大数据流中稳健、可扩展地进行基于模拟的推断（SBI）参数估计。我们在第IIIA 节和III B节中详细介绍这一双重“摘要器”网络架构，并在第III节全面阐述所提出的端到端、可扩展的基于模拟的推断框架。随后，我们在第IV B节介绍对应于 1 个月观测期的类 LISA 信号模型设计，并在第IVD节介绍对应于 3 个月观测期的设计。最后，我们在第V节总结全文并探讨未来工作的潜在方向。

总之，本研究有两项主要贡献：

- 我们提出了一种非线性降维框架，可扩展至规模大到难以处理的数据集，并与摊销推断参数进行联合训练。通过数值比较，我们证明，与解耦训练程序相比，这种联合训练程序产生的估计偏差更小，表明降维表示保留了与参数恢复相关的信息，从而使所得摘要对这些参数近似充分。

- 通过直接使用采集到的时域数据进行训练，我们以经验方式证明，在高噪声条件下，FMPE 比传统归一化流（NFs）能够实现更稳定、更稳健的推断。据我们所知，这是首次针对使用 FM 进行引力波参数估计，系统地比较后验分布的数值研究，凸显了当数据受到严重噪声污染且存在缺口时，FM 的稳健性。

II. 预备知识

A. 问题描述

许多科学领域都依赖于从含噪时间序列数据中估计物理模型的参数。一个典型例子，也是我们用于阐释的案例研究，来自引力波（GW）天文学。理想的观测数据流通常可以建模为受到噪声干扰的底层信号：

$$d(t) = h(t;\boldsymbol{\theta}) + n(t). \tag{1}$$

这里， $d(t)$ 表示完整的观测时间序列数据， $h(t;\boldsymbol{\theta})$ 表示由具有真实参数 $\boldsymbol{\theta}$ 的模型生成的潜在信号， $n(t)$ 则表示噪声。

长时间序列数据不可避免地会出现数据缺失或无法使用的时段——我们称之为缺口——这是现实世界数据采集中常见且重大的挑战。我们使用二值窗函数 $w(t)$ 对这些缺口进行建模，因此得到用于分析的不完整数据如下：

$$\hat{d}(t) = w(t)d(t), \tag{2}$$

其中

$$w(t) = \begin{cases} 1, & \text{if data at } t \text{ is available} \\ 0, & \text{if data at } t \text{ is unavailable.} \end{cases} \tag{3}$$

长时信号带来的高维性与此类缺口的存在相结合，对传统参数估计技术构成了巨大障碍，尤其是在信号持续数年的天体物理学等领域。

贝叶斯推断是根据观测到的数据流 d估计源参数 $\boldsymbol{\theta}$ 的标准方法。这一过程由贝叶斯定理支配：

$$p(\boldsymbol{\theta}|d) = \frac{p(d|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(d)} \tag{4}$$

$$\propto p(d|\boldsymbol{\theta})p(\boldsymbol{\theta}), \tag{5}$$

其中， $p(\boldsymbol{\theta}|d)$ 是参数的后验概率分布， $p(d|\boldsymbol{\theta})$是似然函数， $p(\boldsymbol{\theta})$ 是先验分布，而 $p(d)$ 是证据。

对于许多应用，在假设噪声平稳且为高斯噪声的条件下，使用 Whittle 似然在频域中计算似然函数十分方便 [50, 51]。似然函数通过噪声加权内积表示[1][52, 53]:

$$p(d|\boldsymbol{\theta}) = -\frac{1}{2}(d - h_m(\boldsymbol{\theta})|d - h_m(\boldsymbol{\theta})), \tag{7}$$

[1] 内积定义如下:

$$(a|b) = 4\mathrm{Re}\int_0^\infty \mathrm{d}f \frac{\tilde{a}(f)\tilde{b}^\star(f)}{S_n(f)}, \tag{6}$$

其中， $S_n(f)$ 是噪声的单边功率谱密度（PSD）。带波浪号的量表示傅里叶变换，$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi i f t}\mathrm{d}t$。

其中， $h_m(\boldsymbol{\theta})$ 是针对给定参数集生成的模板波形。随后通常使用随机采样方法（如马尔可夫链蒙特卡罗（MCMC））来探索该似然函数。

加窗函数 $w(t)$ 会使噪声变为非平稳，从而破坏 Whittle 似然的核心假设。在频域中，带缺口数据的傅里叶变换$\tilde{\hat{}}$ $d(f)$，会成为真实频谱与窗函数频谱的卷积：$\tilde{\hat{}}$ $\tilde{d}(f) = \tilde{w}(f) * d(f)$。这种卷积会导致严重的频谱泄漏，使强信号或尖锐噪声特征中的功率扩散到很宽的频率范围 [4]。这种失真会破坏噪声功率谱密度估计，并使式（6）中的内积无法可靠地衡量数据与模板之间的相似性，进而使最终参数估计产生偏差。虽然已有方法可以缓解这一问题，但这些方法通常复杂且计算开销大。

B. 基于生成流的基于模拟的推断

为避开在频域中计算似然函数所面临的挑战，我们采用 SBI，即无似然推断，通过利用前向生成过程直接学习后验分布 $p(\boldsymbol{\theta}|d)$。这对我们的问题尤其有利，因为我们可以轻松模拟带缺口的时域数据。近年来，随着各种深度生成模型被用于表示这些后验分布，SBI 领域发生了重大变革。

一个**归一化流**通过可解析的雅可比行列式对应的可逆函数，将简单基分布中的样本转换为复杂的目标分布，从而学习该目标分布。这些变换通常并非单一整体，而是由有限个较简单的可逆函数依次复合（或“堆叠”）而成，其中每个函数都作为流中的一个步骤或层发挥作用。在引力波参数估计中，SBI使用的典型例子是**掩码自回归流（MAF）** [27, 54]。MAF的表达能力尤其强，这得益于其自回归特性，使其能够对复杂的条件依赖关系进行建模。MAF架构面临一个关键的权衡——

它能够快速、并行地进行密度估计，因此在训练时（即计算数据的似然函数）效率很高。然而，每个离散变换都必须严格满足可逆性要求，这限制了可用的神经网络架构（例如通过掩码实现），可能使其整体表征能力不及更灵活的生成模型。

连续归一化流（CNFs）通过将变换定义为常微分方程（ODE）的解，对 NFs 进行了推广。通过积分学习得到的向量场 $v(\boldsymbol{z},\ t)$，即可将来自基分布 $\boldsymbol{z}_0 \sim p_0$ 的样本映射为来自目标分布$\boldsymbol{z}_1 \sim p_1$ 的样本：

$$\frac{d\boldsymbol{z}_t}{dt} = v(\boldsymbol{z}_t, t). \tag{8}$$

尽管比离散归一化流更加灵活，[55, 56]，训练 CNFs 可能较慢且数值不稳定 [37, 57]。

流匹配（FM）是一项近期创新，显著改进了 CNFs 的训练 [37,58]。FM 不通过复杂的学习型常微分方程来定义流，而是构造一个更简单的向量场来引导生成过程。对于给定的数据观测 d，定义条件概率路径 $p_t(\boldsymbol{\theta}|d)$，在 $t = 0$ 处将简单的基分布 $p_0(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}|0,\ \mathbf{I})$ 与 $t = 1$处所需的后验分布 $p_1(\boldsymbol{\theta}|d)$ 进行插值。

这一路径的一种特别有效的选择受到最优传输（OT）的启发，它能得到简单的线性插值 [59]。相应的目标向量场为：

$$u_t(\boldsymbol{\theta}|d) = \boldsymbol{\theta}_1 - \boldsymbol{\theta}_0, \tag{9}$$

其中 $\boldsymbol{\theta}_0 \sim p_0(\boldsymbol{\theta})$ 和 $\boldsymbol{\theta}_1 \sim p_1(\boldsymbol{\theta}|d)$。随后，通过稳定的回归目标训练神经网络 $v_\phi(\boldsymbol{\theta}_t, t, d)$，使其逼近该目标向量场：

$$\mathcal{L}_{\text{CFM}} = \mathbb{E}_{t,p_t(\boldsymbol{\theta}|d)}\left[\left\|v_\phi(\boldsymbol{\theta}_t, t, d) - u_t(\boldsymbol{\theta}|d)\right\|^2\right]. \tag{10}$$

这种方法提供了更稳定、高效的训练目标，避免了反向传播穿过常微分方程求解器（如 CNF 中） [38]，或处理低噪声水平下方差爆炸的问题（如扩散模型中） [37]。

向量场 v_ϕ 训练完成后，从后验分布生成样本的推断过程与 CNF 的过程完全相同。我们从 $\boldsymbol{\theta}_0 \sim p_0(\boldsymbol{\theta})$ 开始，并使用数值常微分方程求解器求解初值问题：

$$\frac{d\boldsymbol{\theta}_t}{dt} = v_\phi(\boldsymbol{\theta}_t, t, d), \quad \text{initial condition } \boldsymbol{\theta}_0. \tag{11}$$

$t = 1$ 处的解记为 $\boldsymbol{\theta}_1$，代表从我们对后验分布 $p_\phi(\boldsymbol{\theta}|d)$ 的学习近似中抽取的单个样本。在基于模拟的推断框架中，这一过程会重复多次，每次积分都从 p_0 中抽取一个新样本作为起点。由此得到的样本集合 $\{\boldsymbol{\theta}_1^{(i)}\}_{i=1}^N$ 构成一个经验分布，用于近似真实后验分布，并可据此计算可信区间及其他摘要统计量。

III. 模型结构

我们提出的框架遵循 SBI 的两部分“以例为师”结构，由一个**摘要网络**执行降维，以及一个**推断网络**近似后验分布。整体架构如图1所示。该流程整合了一个前向建模阶段，将噪声和伪影注入模拟信号，随后由摘要器将这些信号压缩为具有信息量的条件向量。这些向量引导条件流匹配网络回归目标向量场，从而将简单的基分布传输至后验分布。通过联合优化摘要器和推断引擎，模型能够自动学习足以进行参数估计的稳健特征。这种方法的一个关键特点是摘要器具有灵活性和充分性，可根据输入信号的长度和计算需求进行调整。

A. 路径1：时域摘要器

对于较长但在时域内仍可进行计算处理的信号，我们采用深层 1D-CNN 作为摘要网络，这与我们先前在 [23] 中的工作类似。该网络专为对受噪声和缺失值影响的归一化时间序列数据进行高效降维而设计，同时提取最显著的特征。由于

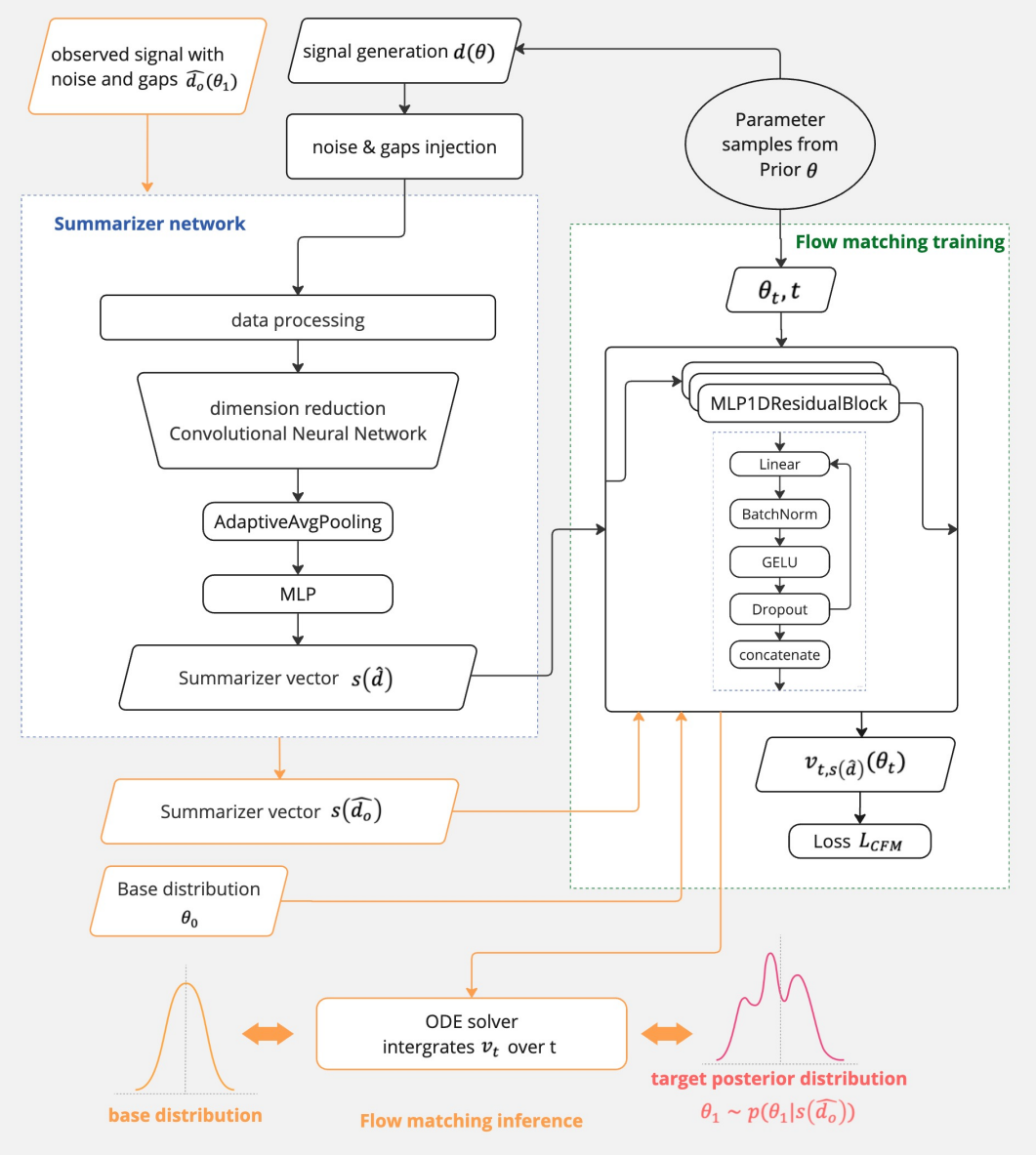


图 1：所提出的端到端模拟推断框架示意图。左上：前向过程生成注入噪声且带有缺口的合成信号。蓝色方框：摘要网络通过 CNN 和多层感知机（MLP）处理这些信号，生成 ˆ ˆ 低维摘要统计量 $s(d)$。绿色方框：以 $s(d)$为条件的流匹配网络使用残差块进行训练，以最小化流匹配损失 $\mathcal{L}_{CFM}$。橙色线条和底部橙色方框：在对观测到的 d_o进行推断时，通过常微分方程求解器，从基分布 $p(\theta_0)$ 出发，沿学习到的向量场 ˆ 积分，从而近似目标后验分布 $p(\theta_1|s(d_o))$。

如图2中的绿色方块所示，路径1的摘要器由一系列卷积块构成。每个卷积块包含一个步幅较大的单维卷积层，后接高斯误差线性单元（GELU）激活函数和批归一化层。在连续模块中逐步增大步幅，使网络能够快速缩短序列长度，并构建时间特征的层次化表示 [23]。末尾的自适应平均池化层确保输出具有固定大小，随后通过最终的线性投影生成摘要向量 $s(d)$，并将其用作 FM 中的条件。这种直接的时域方法无需任何中间数据变换，使模型能够直接从时间动态中学习最相关的特征。

B. 路径2：时频（小波）摘要器

当时间序列足够长时，直接使用一维卷积神经网络在计算上将变得不可行。为实现高效降维，同时减轻频谱泄漏，我们采用基于小波变换的时频表示。具体而言，遵循 [60], 中介绍的方法，我们使用 Wilson–Daubechies–Meyer (WDM)小波。它们构成完备的正交基，并且在时域和频域都具有良好的局部化特性。

时间序列 $d(t)$ 的变换会生成一个由小波振幅 a_{nm} 组成的二维网格，并以时间 t_n 和频率 f_m 为索引。每个基函数，即小波 $\chi_{nm}(t)$，都由加窗正弦波构成，其中窗函数（或“滤波器”）在时间上具有局部性。这种局部性的关键优势在于，时间 t 处的数据缺口所产生的影响会被限制在时频平面中的一个局部区域内。与傅里叶变换不同，傅里叶变换中的一个缺口会污染所有频率箱，而小波域中的缺口只会影响其滤波器与该缺口重叠的小波系数 [14]。这使得该表示对不完整数据具有更强的鲁棒性。

时域小波的形式为 [60]:

$$\chi_{nm}(t) = \begin{cases} \sqrt{2}(-1)^{nm}\cos(\pi m t/\Delta T)\eta(t-n\Delta T), & m+n=\text{even} \\ \sqrt{2}\sin(\pi m t/\Delta T)\eta(t-n\Delta T), & m+n=\text{odd} \end{cases} \tag{12}$$

其中，$\eta(t)$ 是时域中的迈耶窗函数，ΔT 是小波像素的时间范围，而该变换通过快速 $\mathcal{O}(N\log N)$ 算法 [60] 计算得到。通过将存在缺口的数据变换到该域中，我们生成了一个二维谱图，将缺口的影响隔离开来，从而为我们的摘要网络提供更清晰的输入。

接下来，对原始小波输出的模逐元素取对数。由此得到的实值谱图作为二维卷积神经网络摘要器的输入张量。如图中橙色方块所示2，该网络以标准的 3×3 2D 卷积层开始，对输入进行下采样，并将其投影到更高维的维度特征空间。核心特征提取由一系列残差块完成，这些残差块采用非对称卷积核（例如 3×9），专门设计用于捕捉相邻频带之间的相互作用，同时覆盖更长的时间跨度。重要的是，我们沿时间轴采用各向异性膨胀，并在后续阶段以几何级数递增膨胀率（$dr = 1$、2、4、8；在后续应用于90天信号时），如图 3所示。这种设计无需增加参数量，即可使感受野呈指数级增长 [61]。经过卷积阶段后，自适应平均池化层将空间维度压缩为单个特征向量。该

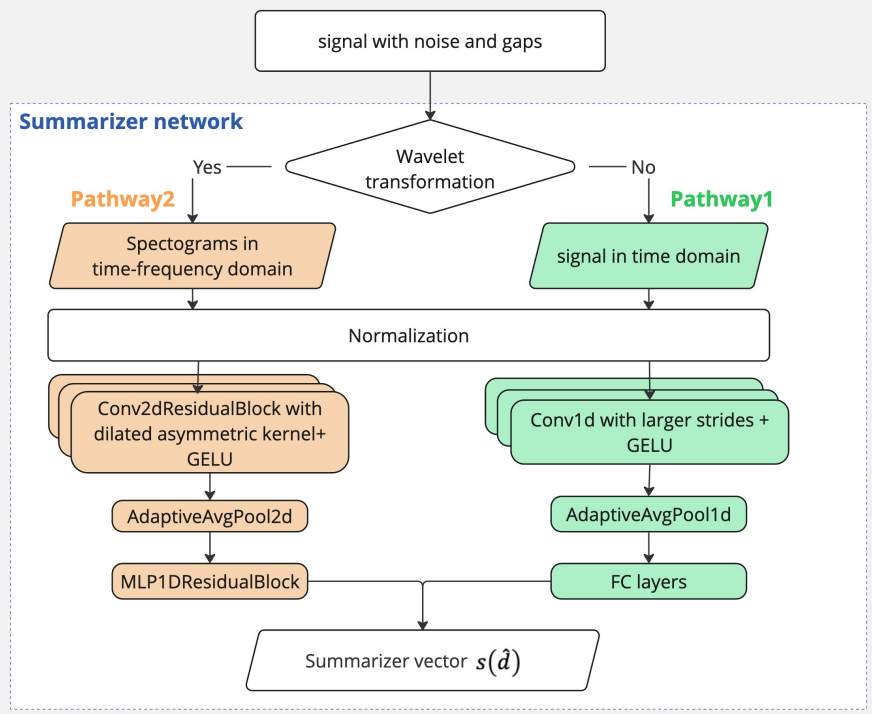


图 2：摘要　　器结构在两条路径中的示意图。右侧绿色部分表示路径 1，适用于数据；左侧橙色部分描述路径 2，适用于时频域数据。

向量最终通过一个稠密残差网络（残差多层感知机）进行投影，生成信息丰富的低维汇总统计量 $s(d)$。

C. 推断引擎：条件流匹配

无论采用哪种路径，提取出的摘要向量 s 都作为 FMPE 框架的条件输入。鉴于参数的数值相对较小，在将其纳入模型之前，我们会先进行对数变换。随后，我们使用复合神经网络架构对时变向量场 $v_\phi(\boldsymbol{\theta}_t, t \mid s)$进行参数化。首先，通过专用参数嵌入网络（残差多层感知机），将模拟时间 t与参数状态 $\boldsymbol{\theta}_t$ 共同投影到高维潜在空间。接着，通过门控线性单元（GLU）条件化，将该嵌入与摘要向量s拼接 [38, 39], ，从而有效融合当前流状态与观测数据上下文。组合后的表示作为主要流网络的输入——一个由密集层构成、采用 GELU 激活和丢弃法的深层残差多层感知机。该架构回归出所需的向量场，以将样本从简单的基分布传输至复杂的目标后验分布 $p(\boldsymbol{\theta} \mid s(d))$，从而通过数值常微分方程积分实现推断。

D. 数据增强

由于两条路径采用的输入数据结构各不相同，为了在存在噪声和缺口的情况下稳健地训练模型，训练过程中会采用不同的数据增强策略。

对于路径1中的时域数据，数据增强会在线进行。在每个训练步骤中，训练集中的一个干净信号都会与新生成的、随机实现的有色高斯噪声以及唯一的缺口掩码进行数据增强。

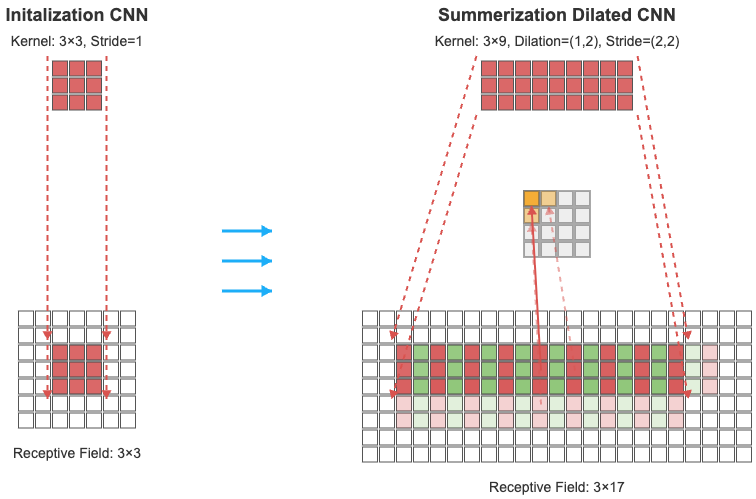


图 3：Pathway 2 中使用的非对称空洞卷积核示意图。该卷积核的时间维度（宽度）比频率维度（高度）更宽，旨在捕获时频谱图中的长程相关性。

这种动态增强是一种强大的正则化形式，可防止过拟合，并显著提升网络的鲁棒性和泛化能力 [62, 63]。通过持续向模型展示各种各样的噪声和缺口模式，迫使模型学习对这些随机扰动具有不变性的特征，从而防止对训练数据中特定实例的过拟合 [64–67]。此外，还针对噪声采用了课程学习策略 [68]，在训练初始阶段，噪声幅度会从零逐渐增加至最大水平。这种方法有助于模型首先按顺序从较简单的情景中学习特征，再逐步应对更具挑战性的高噪声情景。因此，它增强了模型的泛化能力，并加快了收敛速度 [69–71]。

对于 Pathway 2 中的时频域数据，实时生成小波所带来的计算开销使在线增强难以实现。因此，我们采用离线增强与随机加载相结合的方式。在这一通过并行化加速的预处理阶段，我们构建了一种一对多映射，使每个干净信号生成 N 种不同的谱图。每张谱图都会加入独立生成的噪声——这种方法已知能够提升模型的稳健性 [72]——并随机引入缺口，这与*SpecAugment* [73]提出的时频掩蔽策略类似。

训练流程专门针对这种特定的数据表示进行了设计。在每个训练轮次开始时，干净信号列表都会随机打乱。随后，在每次训练迭代中，数据加载器通过随机选择与批次中每个信号对应的一个 N 预计算增强谱图进行随机采样。这种分层（两级）采样策略确保模型在训练过程中反复接触同一底层信号的多种增强形式。因此，它保留了数据变化性 [73]，带来的关键正则化优势，同时受益于离线预处理所带来的计算效率。

E. 训练策略

我们采用联合的端到端训练策略，同时优化信号摘要网络 s_ψ 和流匹配网络 v_ϕ 。不同于依赖固定摘要统计量或预训练特征提取器的两步法，我们的架构能够学习最优的摘要统计量

ˆtic $s_\psi(d)$，其动态变化完全由下游推断任务引导。

a. 目标函数 整个架构使用单一的条件流匹配（CFM）损失函数进行训练。我们利用 Python 软件包 流匹配 [74] 来获得第 IIB 节和公式 13中提到的 OT 概率路径。模型参数 $\{\phi, \psi\}$ 通过最小化预测向量场与目标场之间的均方误差（MSE）进行更新：

$$\mathcal{L}_{CFM}(\phi,\psi) = \mathbb{E}_{t\sim\mathcal{U}(0,1),\boldsymbol{\theta}_1\sim p(\boldsymbol{\theta}),\boldsymbol{\theta}_t\sim p_t(\cdot|\boldsymbol{\theta}_1),d\sim p(d|\boldsymbol{\theta}_1)}\left[\left\|v_\phi(\boldsymbol{\theta}_t,t,s_\psi(\hat{d})) - (\boldsymbol{\theta}_1-\boldsymbol{\theta}_0)\right\|^2\right] \tag{13}$$

其中， θ_t 是 $\boldsymbol{\theta}_0$ 与时间 t 下目标 $\boldsymbol{\theta}_1$ 之间的线性插值。由于嵌入 ˆ$s_\psi(d)$ 是直接输入到 v_ϕ 的，因此该回归损失的梯度会反向传播 through 流网络并进入嵌入网络的卷积层，迫使 s_ψ 提取专门用于刻画后验分布几何结构的特征。

b. 优化计划 优化使用 Adam优化器 [75]执行。为稳定卷积特征提取器与残差流网络的联合训练，我们采用具有差分学习率 [76] 的两阶段优化计划：

1.**阶段 1（特征学习）：** 与流网络相比，我们为信号嵌入参数采用了更高的学习率（放大 5 倍）。这促使嵌入网络迅速调整其权重，从谱图中提取有意义的摘要向量。我们使用余弦退火调度器 [77]逐步降低学习率。

2.**阶段 2（微调）：** 特征提取器稳定后，我们会使用更低的全局学习率重新设置优化器，并降低嵌入层的乘数。此阶段将对提取的摘要与向量场预测之间的对齐关系进行微调，以确保高保真度的后验采样。

IV. 应用

A. 数据准备

我们通过模拟引力波信号来训练所提出的模型。正如训练策略中所述，我们的数据集由干净的（无噪声）波形 $h(t;\boldsymbol{\theta})$ 及其对应参数 $\boldsymbol{\theta}$ 组成。噪声 $n(t)$ 和数据缺口 $w(t)$ 会在训练期间即时应用，或在增强过程中应用，以便为每个训练步骤创建独特的数据实例$d(t) = w(t)[h(t;\boldsymbol{\theta}) + n(t)]$ 。考虑到生成信号的计算成本，我们通过聚焦于单个相对较强的银河系双星（GB）信号来简化问题，因为对多个双星进行总体推断将涉及完整的全局分析 [41]。

1. *Signal Model*

干净信号 $h(t;\boldsymbol{\theta})$ 使用 GPU 加速的 fastlisaresponse 软件包 [78],进行处理，该软件包可以将形式为 $h(t) = h_+ + ih_\times$ 的任意时域信号投影为 LISA 的探测器响应。银河系双星波形定义为

$$h_{+,\mathrm{src}} = -a(1+\cos^2\iota)\cos\Phi(t) \quad \text{and}$$

$$h_{\times,\mathrm{src}} = -2a\cos\iota\sin\Phi(t)\,,$$

其中 a 为振幅； ι为倾角；以及

$$\Phi(t) \approx -\phi_0 + 2\pi \left(f_0 t + \frac{1}{2}\dot{f}_0 t^2 + \frac{1}{6}\ddot{f}_0 t^3 \right) .$$

f_0 是初始引力波频率，上方的点表示其时间导数。初始相位为 ϕ_0。随后，使用偏振角 α [79] 将该波形变换到太阳系质心（SSB）参考系中。

$$\begin{bmatrix} h_{+,\mathrm{SSB}} \\ h_{\times,\mathrm{SSB}} \end{bmatrix} = \begin{bmatrix} \cos 2\alpha & -\sin 2\alpha \\ \sin 2\alpha & \cos 2\alpha \end{bmatrix} \begin{bmatrix} h_{+,\mathrm{src}} \\ h_{\times,\mathrm{src}} \end{bmatrix} .$$

我们的模型关注一组参数

. $\boldsymbol{\theta} = \{a, f, \dot{f}\}$。其余五个外禀参数在所有模拟中均保持不变：倾角（$\iota$）、初始相位（$\zeta_0$）、偏振角（$\alpha$）、黄纬（$\beta$）和黄经（$\lambda$）。我们假设等臂长解析轨道来模拟仪器响应，并采用 A 通道的第一代 TDI 响应。

2. *Noise Model*

噪声项 $n(t)$ 表示由 LISA 仪器扰动引起的随机波动，其来源既包括无法分辨的引力波源，也包括非引力波仪器干扰。在这项工作中，我们将对嵌入数据流d中的单个 波形进行推断，并忽略数据中存在多个重叠信号的可能性；这类情况可通过全局拟合框架 [12] 加以处理。我们假设每个通道中的噪声均为均值为零的弱平稳高斯随机过程，其颜色由相应 AET 通道的功率谱密度（PSD）表征。在此假设下，噪声在频域中不相关，因此噪声协方差矩阵严格为对角矩阵 Σ [53, 80, 81]:

$$\Sigma(f, f') = \langle \tilde{n}(f)(\tilde{n}(f'))^\star \rangle \tag{14}$$
$$= \frac{1}{2}\delta(f - f')S_n(f') . \tag{15}$$

对于 $f \in (0, \infty)$。其中， $\langle\cdot\rangle$ 表示对大量噪声实现取平均， δ 是狄拉克δ函数， S_n 是每个通道内噪声过程的PSD。

频域方程（14）在连续极限下离散化后，可得到两个频率箱之间噪声的协方差 f_i 和 f_j:

$$\tilde{\Sigma}_{ij} = \mathbb{E}_d[\tilde{n}(f_i)(\tilde{n}(f_j))^\star] \tag{16}$$
$$= S_n(f_i)\delta_{ij}/2\Delta f. \tag{17}$$

这里 $f_i \in [0, \Delta f, \ldots, (\frac{N}{2})\Delta f]$ 表示一个独立的频率箱， $\Delta f = 1/N\Delta t = 1/T_{\mathrm{obs}}$ 表示频率箱之间的间隔， N 表示时间序列的长度， Δt 表示采样间隔。

方程（17）表明，对于平稳高斯噪声， $i \neq j$ 的频率箱在统计上彼此不相关。只考察（17）的对角线元素，还可以进一步证明，噪声的实部和虚部各自服从高斯分布:

$$\mathrm{Re}(\tilde{n}(f_i)) = N\left(0, \frac{S_n(f_i)}{4\Delta f}\right) \tag{18a}$$
$$\mathrm{Im}(\tilde{n}(f_i)) = N\left(0, \frac{S_n(f_i)}{4\Delta f}\right) . \tag{18b}$$

在这一框架下模拟噪声时，噪声分量取样自方程（18a）和（18b），并使用适用于 A 通道的 S_n PSD，借助软件包`LISA Analysis Tools`[82]。随后，通过逆傅里叶变换合成一个精确信号，并将其叠加到这一特定的噪声实现上，最终形成方程（1）所述的数据流 d（t）。

3. *Gaps Simulation*

LISA 数据流中的数据缺口大致分为计划性和非计划性两类。计划性数据间隙源于定期维护，例如前视角机构（PAAM）调整（∼100 秒，每天三次）、天线重新指向（每隔 ∼14天进行一次，每次持续∼3 小时）以及倾斜-长度耦合估计（2天，每年四次） [4]。相比之下，非计划性数据间隙由意外事件引起，例如短暂中断（∼10 秒），或更严重的情况，如微陨石碰撞。据估计，每年会发生 ∼30 次此类碰撞，每次会造成约一天的数据丢失 [4, 22, 83]。具体的模拟设置，例如 LISA Data Challenge 的“Spritz”，将这些情况模拟为每隔 10 至 15 天出现一次、持续 7 小时的固定缺口，以模拟每两周重新指向

。 [15]。

我们的研究旨在通过涵盖灵活多样的缺口情景，提升模拟的泛化能力。为避免过高的内存和计算成本，并考虑信号长度及路径2中采用的小波变换，每个缺口的持续时间均从3至12小时的均匀分布中抽取。相邻缺口之间的间隔从指数分布中抽取 [20, 23]。我们对该分布的速率参数进行校准，使占空比维持在约80%，与 [84]中引用的最新估计一致。

B. 30天信号

我们首先将路径1（一维时域摘要器）应用于一个30天信号，以此开始案例研究。该信号以 $\Delta t = 5$秒的采样间隔进行采样，得到长度为 $N = 518,400$的时间序列。

1. *Experimental Setup*

一个测试信号 h_0 使用真实参数生成 $\boldsymbol{\theta}_0 = \{a_0, f_0, \dot{f}_0\}$，其中 $a_0 = 5 \cdot 10^{-21}$，$f_0 = 2 \cdot 10^{-3}$赫兹，且 $\dot{f}_0 = 3 \cdot 10^{-10}$ 赫兹/秒。该信号被注入有色高斯噪声中，对应的信噪比约为45，该值根据 [85]中给出的定义计算得出。按照第IV A3 节所述的30天模型模拟缺口，从而生成最终观测到的测试数据流$\hat{}d_0$，如图4所示。

训练集由 20,000 个无噪声信号模拟组成 $h(t;\boldsymbol{\theta})$。这些信号的参数 $\boldsymbol{\theta}$ 是从以真实值 $\boldsymbol{\theta}_0$ 为中心的均匀先验分布 $p(\boldsymbol{\theta})$ 中抽取的。我们在参数的对数空间中进行均匀采样：

$$\begin{aligned}\ln a &\sim \mathrm{U}[\ln a_0 - 0.1, \ln a_0 + 0.1]\\ \ln f &\sim \mathrm{U}[\ln f_0 - 10^{-5}, \ln f_0 + 10^{-5}]\\ \ln \dot{f} &\sim \mathrm{U}[\ln \dot{f}_0 - 10^{-5}, \ln \dot{f}_0 + 10^{-5}]\end{aligned}$$

训练期间，每个干净信号都会即时叠加一组不同的有色噪声，并配以随机生成的缺口掩码。总计 5,000 个信号被留出，仅用于验证。

2. *Model and Training*

本实验的摘要网络是第 IIIA 节所述的一维卷积神经网络。如表I所示，该网络由 4 个卷积块构成，这些卷积块逐步缩短序列长度，同时增加通道深度（$1 \to 64,\ 64 \to 128,\ 128 \to 256,\ 256 \to 512$），步幅依次增大为 4、6、8 和 10。自适应平均池化层和最后的线性投影层生成一个 256 维摘要向量s。

摘要向量作为流匹配推断的条件化机制。该过程涉及使用由三层组成的残差网络对向量 s 进行门控线性单元条件化；这三层的隐藏维度分别为 64、128 和 256，用于参数嵌入和时间嵌入。随后，采用一个由 4 层构成的残差流网络，其隐藏维度分别为 512、256、128 和 64，以生成最终的参数估计。

完整框架采用两阶段优化计划训练了300个训练轮次。前250个训练轮次使用初始学习率 10^{-3}，随后在剩余50个训练轮次中进入学习率为 10^{-4} 的微调阶段，每个阶段均采用余弦退火计划。训练批量大小为64，整个过程约耗时3小时。作为对比，我们还训练了一个MAF模型，并使用相同的一维卷积神经网络汇总器，以确保公平比较。训练约耗时10小时。该MAF模型与近期关于LISA银河双星的研究所采用的方法相似 [41]。

3. *Results*

训练好的模型被用于推断测试数据的后验分布 $p(\boldsymbol{\theta} \mid s(d))$ d_0。图5 展示了使用 FM 和 MAF 得到的三个关注参数的近似后验分布。尽管这两种模型都成功定位了 $\dot{}$ 和 f 参数，f 所提出的 FM 框架还准确恢复了

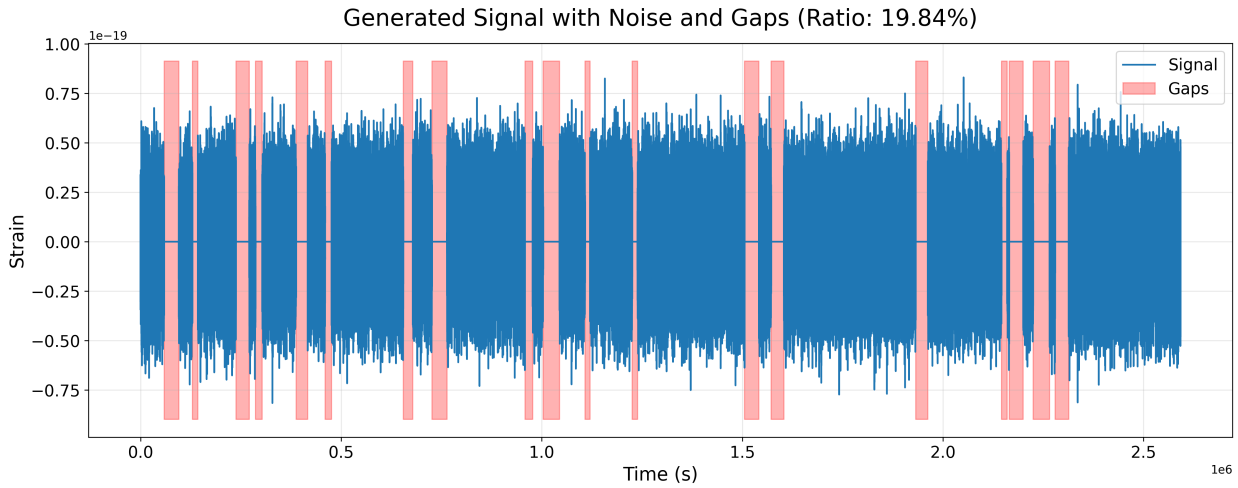


图4：时域中带有缺口的30天测试信号的一次实现

表1：路径1中摘要器的详细架构。

层类型	输入通道	输出通道	卷积核大小	步幅
一维卷积（区块1）	1	64	15	4
一维卷积（区块2）	64	128	15	6
一维卷积（区块 3）	128	256	17	8
一维卷积（区块 4）	256	512	21	10
自适应一维平均池化	512	512	-	-
线性投影	512	256	-	-

Note: Each Conv1d block includes Batch Normalization and GELU activation.

振幅 a。这表明，基于 FM 的推断引擎对数据缺口具有更强的鲁棒性，能够比 MAF 得到更准确的参数估计。此外，这一结果表明，Pathway 1 模型能够有效探索高维数据空间，学习能够抵御大量缺口和仪器噪声的特征，并可靠地恢复底层源参数。

为了进一步评估后验分布的统计可靠性，我们绘制了概率-概率图（PP 图），以展示经验覆盖率与名义可信水平之间的关系。我们注入了 100 组测试数据，每组数据都具有独特的有色高斯噪声实现和随机缺口模式，其参数 $\boldsymbol{\theta}$ 取自先验分布。测试详情可在线获取。[2]。随后，我们使用训练好的 FM 和 MAF 模型，分别推断每个信号的参数。结果 PP 图如图 6所示。FM 模型的图形紧贴对角线，表明其后验分布校准良好，在边缘地具有统计可靠性。相比之下，MAF 模型的 PP 图中，. a 存在显著偏差，而 f 和 f 存在轻微偏差，这证实了模型低估了参数的确定性。这与 [41], 中的发现一致：类似的校准失准使得后处理校准步骤成为必要。我们的

[2]详情请见 https://bpandamao.github.io/experiment_results/experiment_dashboard_fmvsmaf.html

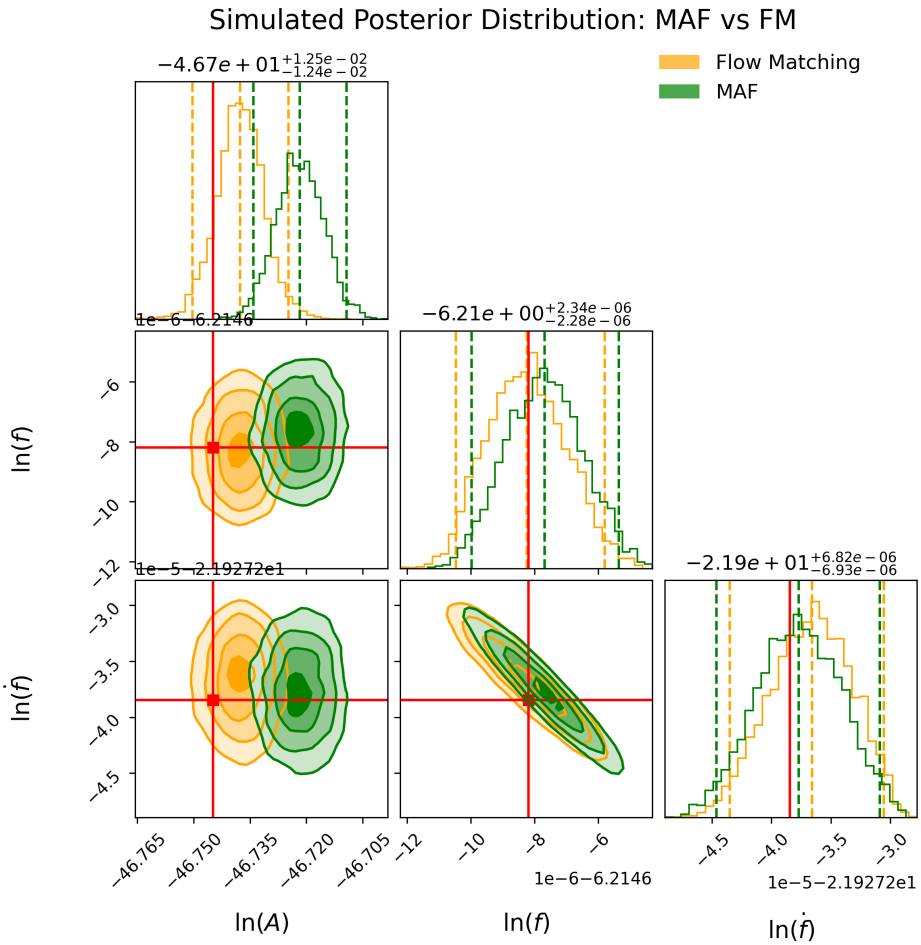


图 5：使用 Pathway 1 汇总器时，流匹配（FM）（橙色）与掩码自回归流（MAF）（绿色）对近似后验分布的推断结果对比。红线表示真实值。

然而，FM 框架可直接生成可靠的后验分布，展现出更强的鲁棒性，无需额外校准。

C. 单阶段与两阶段训练的比较

1. *Two-Stage training setup*

为定量探究端到端训练的优势，我们将所提出的方法与一种两阶段训练框架进行了比较，该框架明确地将特征提取与参数估计解耦。在第一阶段，我们沿用先前工作的策略 [23], ，训练去噪卷积自编码器（DCAE），从标准化、含噪且存在缺口的输入中重建干净、连续的波形。在训练期间，输入和目标信号都经过标准化，使其均值为零、方差为一；这一目标促使编码器学习潜在瓶颈表示，在编码信号时间特征的同时，不受其绝对振幅影响。

在第二阶段，预训练的编码器被冻结，并仅用作特征提取

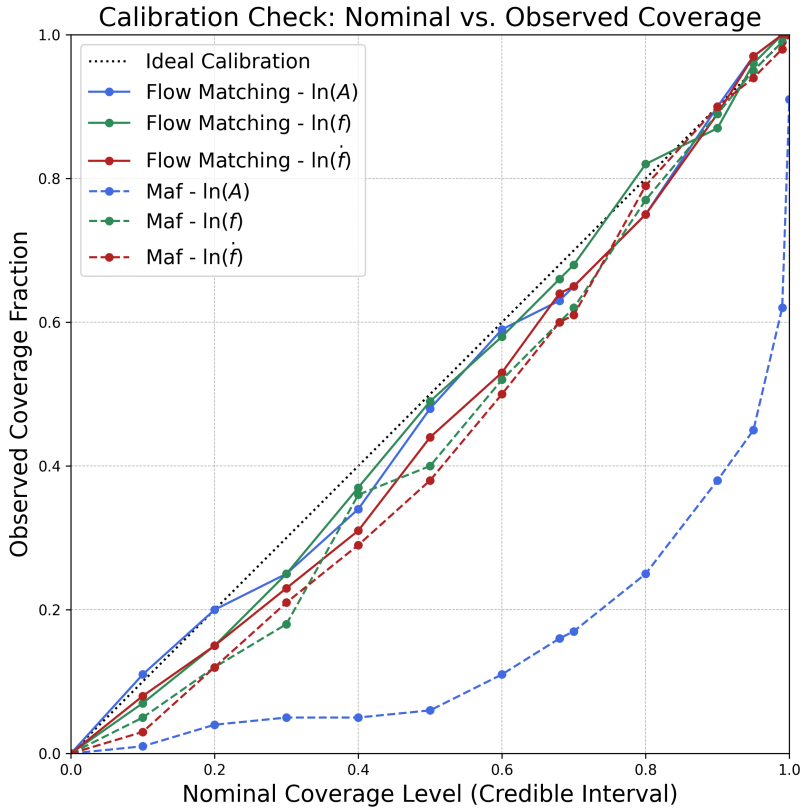


图 6：使用 100 个测试注入的 PP 图进行统计可靠性分析，比较关注参数的 FM（实线）和 MAF（虚线）。

器。为恢复第一阶段标准化过程中丢失的幅度信息，我们显式计算含噪输入信号标准差的对数。随后，将这个标量“尺度”特征与 256 维形态学瓶颈向量拼接，以构建完整的条件表示。最后，训练一个类似的 CFM网络，根据这一组合特征向量估计物理参数的后验分布。

2. *Results*

我们使用测试数据以及包含 100 个测试注入的总体研究评估了两个流程的性能。图7展示了估计的位置-测试信号的后验分布。总体研究的详细信息可在线获取。[3]。端到端（单阶段）模型生成的后验分布范围较窄、约束良好，并准确地以真实参数为中心。相比之下，尽管 DCAE 达到了其目标，如图8所示，两阶段流程的性能却显著下降。虽然它以中等精度恢复了 f 和 $\dot{f}$，但后验分布明显更宽，表明不确定性增加。至关重要的是，两阶段框架未能提取信号中关于振幅参数 a的信息。

[3]详情见 https://bpandamao.github.io/experiment_results/experiment_dashboard_decoupled_stage.html

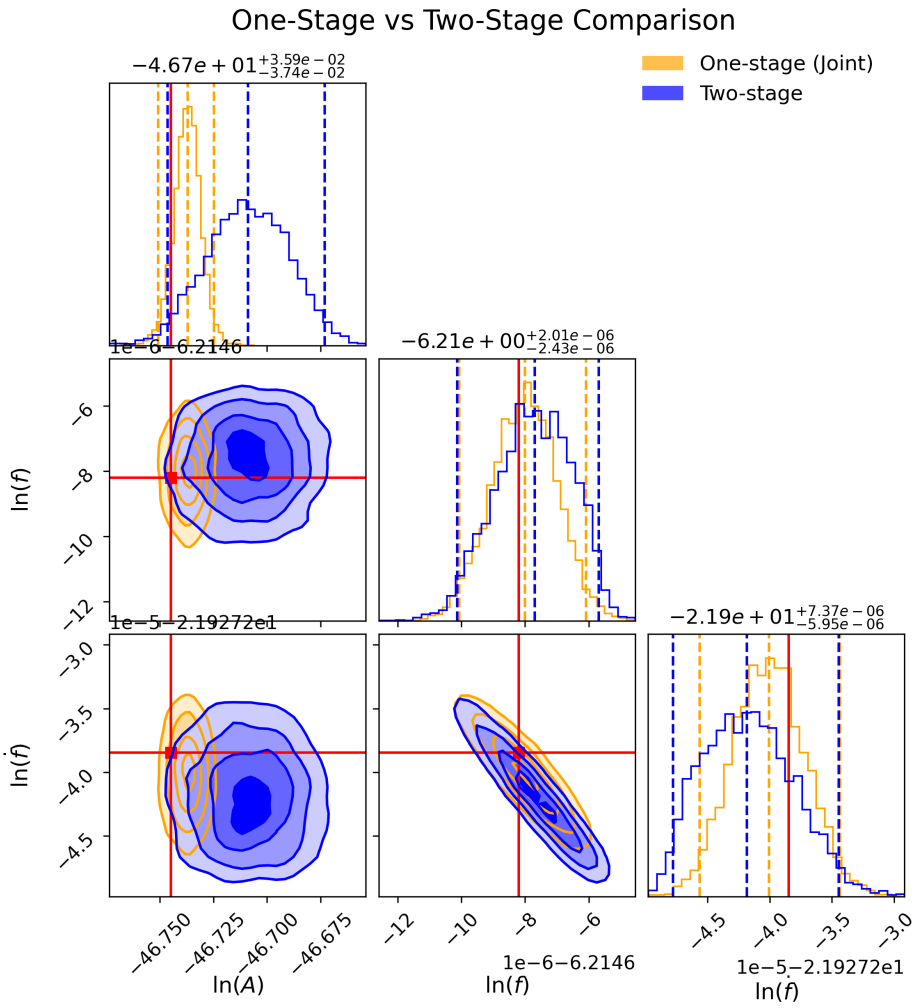


图 7：单阶段训练（橙色）与采用 DCAE 的两阶段训练（蓝色）所得到的近似后验分布推断结果对比。红线表示真实值。

图9所示的 PP 图从统计学角度印证了这一观察结果。端到端模型（实线）的覆盖率接近理想值，并紧贴对角线，表明其报告的不确定性在统计学上经过了良好校准。相比之下，两阶段模型（虚线）明显偏离对角线，尤其是振幅参数。

为探究两个框架的训练稳定性和收敛特性，我们监测了 FM 训练过程中 300 个训练轮次内的 MSE 损失。图 10 展示了两者的损失轨迹

单阶段端到端流程（蓝色）和解耦的两阶段流程（橙色）。尽管两者都表现出稳定的收敛，但可以观察到明显的性能差距。与联合训练不同，解耦流程在训练初期表现出严重的泛化不稳定性，其特点是验证集上的波动更大。在收敛水平方面，端到端流程最终的验证损失显著更低，约为 0.38，而两阶段训练约为 0.50。这一持续存在的差距表明，端到端训练能够有效利用

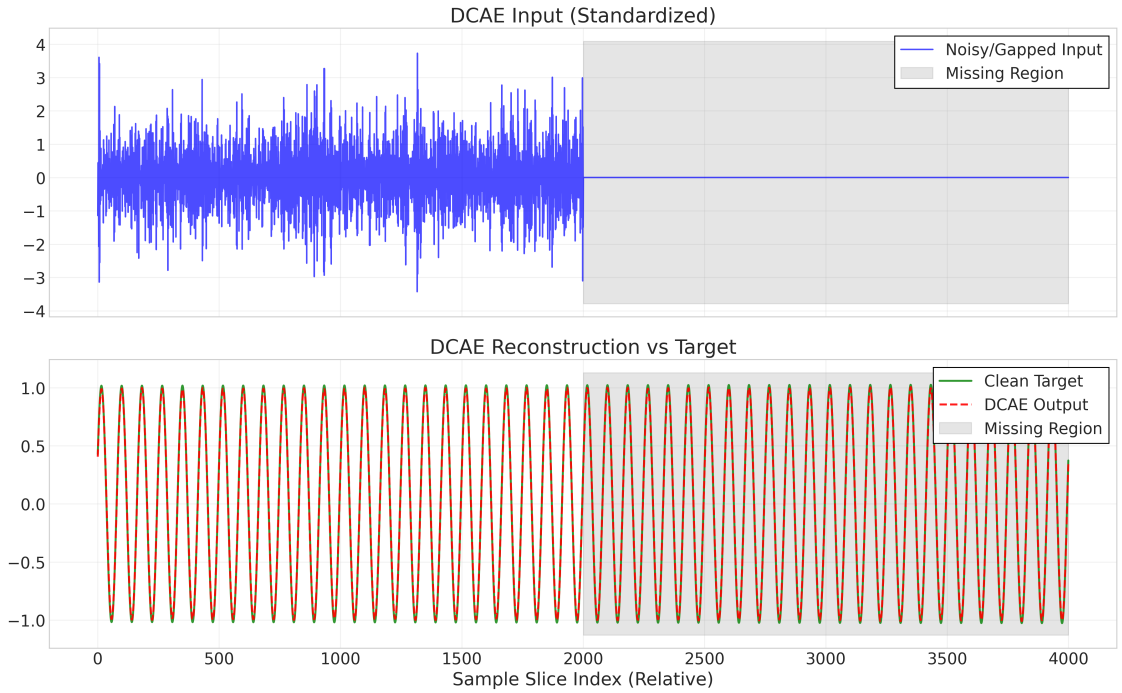


图 8：两阶段训练流程中 DCAE 在测试数据上的表现

完整的高维信号上下文，以最小化流匹配目标。相比之下，两阶段训练停滞在较高的损失水平，这表明存在不可消除的信息瓶颈；将信号压缩为潜向量（阶段 1）不可避免地会丢失细粒度相关性，而这些相关性对于阶段 2 所需的精确流匹配至关重要。损失上的这一量化差异与图7中参数估计结果所显示的更宽后验轮廓和更高不确定性直接相关。

在两阶段方法中，DCAE 中的重构损失约束 256 维瓶颈层学习能够重构完整 518,400点信号的特征。因此，该表示倾向于编码能够解释最大方差比例的成分（例如主导频率成分及其导数），而不是捕捉所关注的联合参数分布。尽管第二阶段纳入了一个标量汇总统计量（对数标准差），但这一额外特征显然不足以恢复在初始重构阶段被丢弃的幅度信息，在初始重构阶段被丢弃的。

由此产生的校准失准表明，解耦架构——具体而言，中间去噪和标准化操作中的信息损失——削弱了模型捕捉准确参数推断所需细微相关性的能力。这些发现表明，与将重建视为完全独立的预训练步骤相比，在单一训练过程中联合优化中间目标（本例中为重建）和主要目标（参数估计），能够提升性能。

D. 更长的信号：90天信号

随着信号持续时间增加，Pathway 11D-CNN 模型的可扩展性成为一项重大的计算挑战。对于 90天信号，时间序列会变得过长，无法用一维卷积处理。因此，我们采用 Pathway 2，它利用小波变换作为一种强大的前端降维——

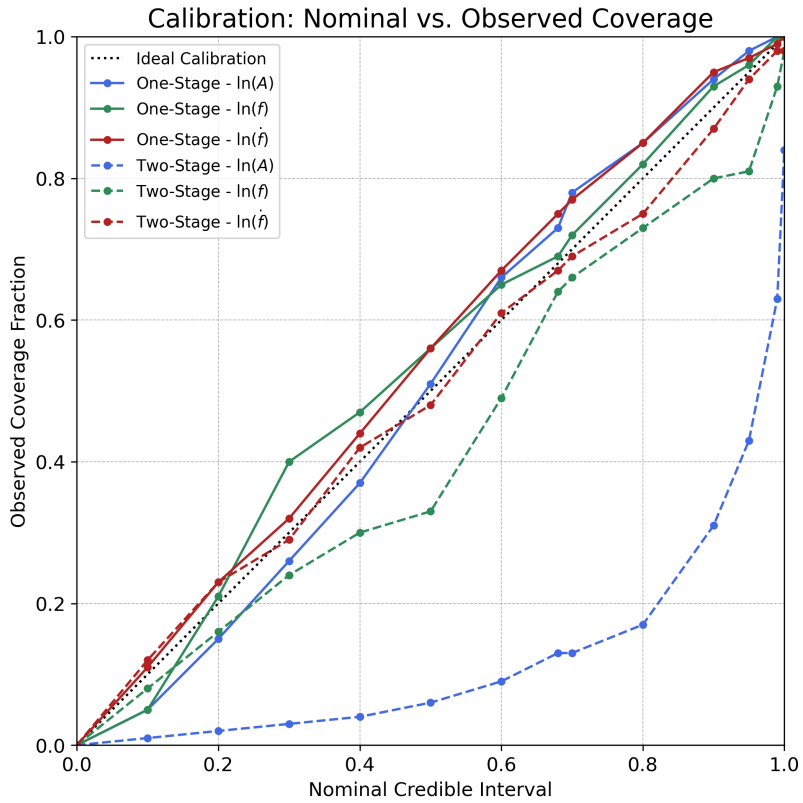


图 9：使用 PP 图对 100 次测试注入进行统计可靠性分析，比较单阶段训练（实线）和两阶段训练（虚线）在感兴趣参数上的表现。

过程。这种方法受到近期相关研究的启发，该研究成功地在类似情境下使用小波域分析进行噪声矩阵估计 [14]。

1. *Experimental Setup*

小波变换使用 Python 软件包将一维时间序列转换为二维时频谱图 `pywavelet` [60, 86, 87], ，所选 2^{11} 频率箱的示例如图 11所示。为优化特征提取并确保模型能够高效学习，我们对谱图进行切片，聚焦于可分辨银河系双星最敏感的频带，即 0.001 Hz 至 0.01 Hz [9, 88]。这种二维表示并非各向同性；其特征及相互依赖关系会在时间轴和频率轴之间存在显著差异。为此，我们的路径2 汇总器采用了由非对称卷积核构成的扩张二维卷积神经网络。

2. *Model and Training*

摘要网络是一个深度二维卷积神经网络，用于处理小波谱图对数绝对值。Pathway 2 中摘要器的核心是一系列残差块，这些残差块使用尺寸为 (3, 9) 的非对称二维卷积核——在频率方向较窄、在时间方向较宽，并在初始化时先使用对称卷积核。这种形状经过专门设计，可捕捉频率箱中的短程依赖，同时建模时间轴上的长程相关性。

此外，网络采用逐步增加的

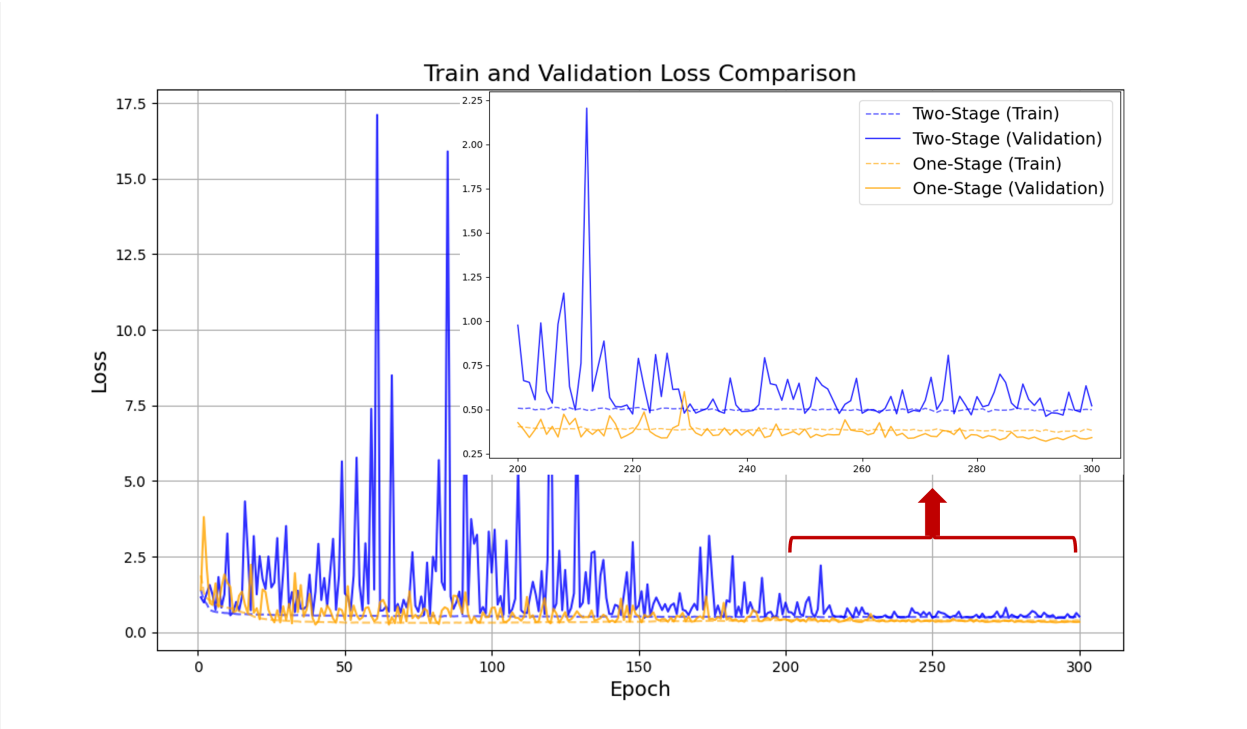


图10：单阶段训练（橙色）与采用 DCAE 的两阶段训练（蓝色）过程中每个训练轮次的损失值对比。训练集的损失为每个训练轮次的平均批次损失。

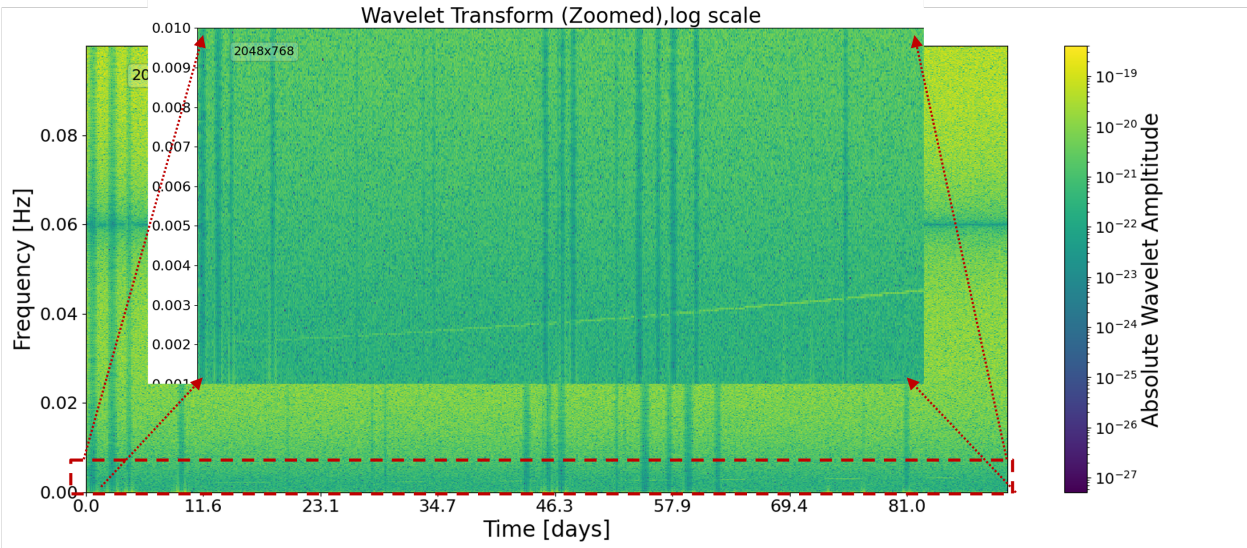


图11：时频域中的频谱图示例，包含 2^{11} 个频率箱，展示了存在间隙且带有噪声的 90天信号的对数频谱图，聚焦于 0.001 Hz 至 0.01 Hz 频段。

在连续模块中增加膨胀率（1、2、4 和 8）。这样，模型便能在时间维度上构建异常大的感受野，而不会导致计算成本相应地呈二次增长。网络在最终的 1-层“ResidualNet”将特征投影为 512 维摘要向量 s 之前，将其通道深度增加到 $64 \to 128 \to 256 \to 512$。

下面的 FM 部分与路径1类似，使用一个 4 层的‘ ResidualNet‘（隐藏维度为 64、128、256、512）来嵌入参数和时间，并使用一个 5 层的“ResidualNet” （隐藏维度为 1024、1024、512、256、128）作为主要的条件流。采用 0.4 的丢弃率进行正则化。

由于小波变换的计算成本高昂，实时 数据增强（如 30天模型中所采用的方式）并不可行。因此，我们采用离线增强策略。利用更宽的先验分布生成 10,000 条训练数据和 2500 条验证数据：

$$\log_{10}(a) \sim \mathrm{U}[\log_{10}(a_0) - 0.1, \log_{10}(a_0) + 0.1]$$
$$\ln(f) \sim \mathrm{U}[\ln(f_0) - 0.001, \ln(f_0) + 0.001]$$
$$\ln(\dot{f}) \sim \mathrm{U}[\ln(\dot{f}_0) - 0.001, \ln(\dot{f}_0) + 0.001]$$

接下来，我们预先生成一个大型频谱图数据集，其中每个干净信号都会通过不同的噪声实现和随机缺口模式多次扩增。随后，按照第IIIE节所述，在这个经过预处理的多样化数据集上训练模型。

3. Results

我们将训练好的 Pathway 2 模型应用于一组90天测试数据，其中注入了真实参数 $a_0 = 1.5 \cdot 10^{-21}$ 、 $f_0 = 2 \cdot 10^{-3}$ Hz 和 $\dot{f}_0 = 10^{-10}$ Hz s，并叠加有色高斯噪声，对应的信噪比约为 40。模型训练了 1,000 个训练轮次，用时 5.5 小时。如图12所示，从推断得到的后验分布可见，模型成功恢复了注入的参数。这表明模型能够扩展到更长时长的信号，同时保持对数据缺口的鲁棒性。

为验证我们的架构选择，我们对联合多维后验分布进行了统计覆盖率分析。虽然边缘后验分布看起来可能正确，但校准是否可靠，真正取决于经验覆盖率（真实参数落入所得可信区间的概率）是否与名义覆盖率（预期可信水平）相符。我们使用高斯核密度估计，为一组模拟信号计算最高后验密度（HPD）可信区域，并据此绘制了这些覆盖率曲线。

我们在图13中比较了三个模型的后验覆盖率：1）我们提出的模型，采用由 (3,3) 卷积核初始化的非对称 (3,9) 卷积核，以及 512 维汇总向量。2）采用对称 (3, 3) 卷积核和 512 维汇总向量的对照模型。3）更简单的对照模型，采用对称 (3, 3) 卷积核和 256 维汇总向量。

结果表明，非对称模型表现更优，其覆盖曲线紧贴对角线，说明其后验分布在统计上准确且校准良好。相比之下，两种对称核模型都明显偏离对角线，证实无论网络深度如何，它们都会产生校准不良的后验分布。这表明，非对称膨胀核是正确学习时频谱图中与参数相关的各向异性特征的关键组成部分，而不仅仅是摘要器的深度。

V. 讨论

在本研究中，我们提出了一种稳健且灵活的基于模拟的推断框架，专为从含噪且不完整的时间序列数据中准确进行参数估计而设计。我们提出了一种联合训练的深度学习降维网络，使 FMPE 方法能够应用于高维 LISA 数据流，从而应对 LISA 数据分析中的一项核心挑战：在不依赖计算成本高昂的似然评估的情况下，处理数据缺口引起的非平稳性。

本研究的一项主要贡献是展示了一种可扩展的架构，能够处理全规模的 LISA 信号。通过联合训练摘要网络与推断引擎，我们确保降维过程专门针对保留物理参数的高密度信息而优化。

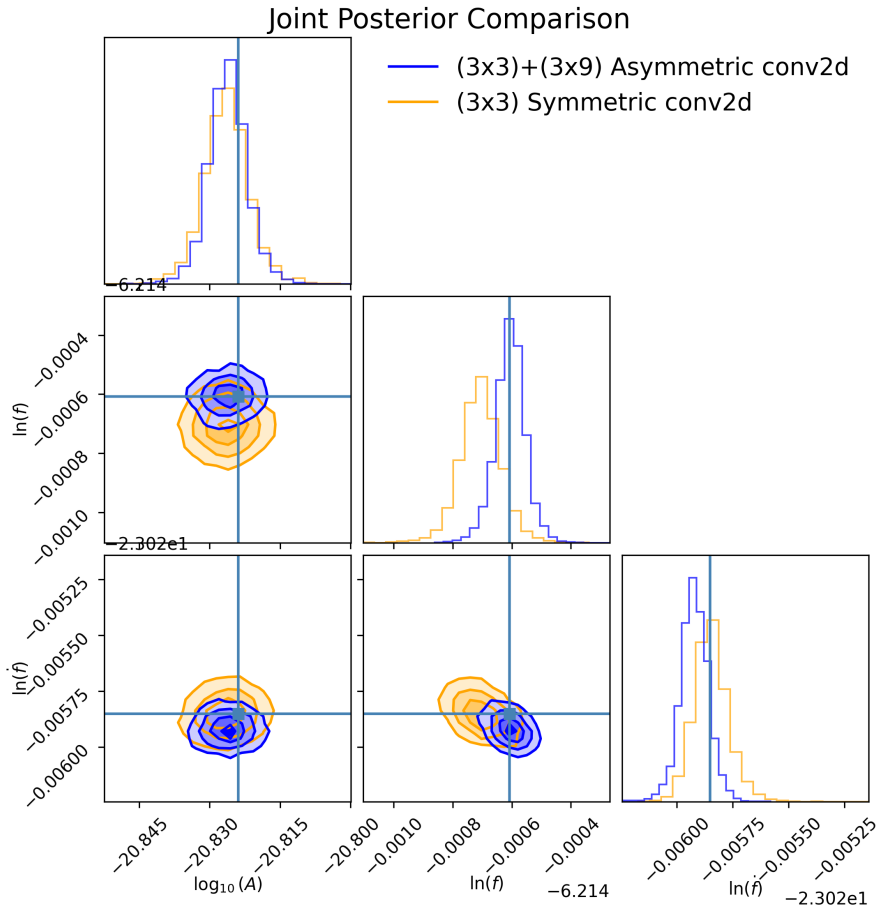


图 12：使用 Pathway 2 模型推断得到的 90天测试数据参数后验分布。蓝线表示真实注入的参数值。

我们的数值实验表明，这种端到端方法显著优于两阶段流程，例如依赖中间 DCAE 的流程；研究显示，这类流程会在重建阶段丢失关键的幅度信息。此外，我们证实，与传统 MAF 相比，FMPE 具有更高的稳定性和更好的覆盖率校准，尤其是在数据受到严重仪器噪声污染且存在缺口时。

我们提出了一种双路径策略，以平衡精度和可扩展性。对于计算上可处理的时长，我们的时域一维 CNN（路径1）通过直接从原始时间动态中学习，得出可靠的可信区间。

对于长持续时间信号（例如，90 天），我们引入了小波域摘要器（Pathway 2）。该路径利用小波的局部化特性来分离缺口，并采用一种新颖的、使用非对称卷积核的扩张二维卷积神经网络，大幅降低时频谱图的维度。这项创新使得在基于模拟的推断框架内，对长达一年的信号进行推断在计算上切实可行。

展望未来，有多个方向可以用于完善和拓展这一框架。首先，有必要对单阶段与两阶段训练范式进行更系统的分析。尽管我们的初步结果更支持端到端方法，但未来的研究应

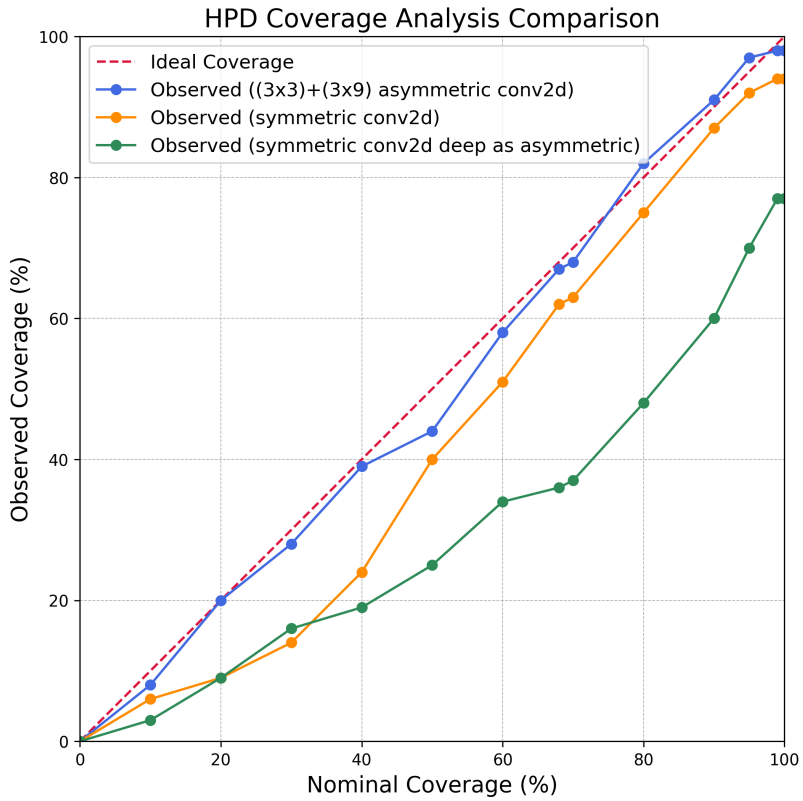


图 13：多维覆盖曲线图，对比了三个模型的统计校准检验结果：我们提出的非对称 (3, 9) 核模型（蓝色）、采用相同摘要器结构的对称 (3, 3) 核模型（绿色），以及更简单的对称 (3, 3) 核模型（橙色）。非对称模型贴合对角线，表明其可信体积准确；而两个对称模型均表现出明显的校准失准。

严格量化更广泛的非线性摘要统计量和嵌入架构中的信息损失。将我们提出的摘要器与其他成熟的压缩网络进行比较，有助于区分哪些特定的归纳偏置有助于在存在缺口时实现稳健的参数恢复。

其次，小波变换目前的计算成本使 Pathway 2 必须离线生成数据，从而限制了训练过程中所见噪声实现的多样性。为解决这一问题，我们提议将 FMPE 的实现从 PyTorch 迁移到 JAX 框架。JAX 的即时（JIT）编译和自动向量化（“vmap”）功能非常适合加速信号处理操作。这将实现高效的实时小波变换和训练，大幅减少 I/O 瓶颈，并使模型能够在每个训练步骤中看到独特的噪声和间隙实现，从而更好地泛化。

第三，尽管 WDM 小波为我们的谱图提供了稳健的基函数，但未来的工作应系统地评估其他小波族。对于特定类别的引力波信号或毛刺形态，不同的基函数或许能提供更优的稀疏性或特征隔离。此外，我们的非对称卷积核实现是通过经验确定的；要优化这些卷积核的维度，还需要更严谨的理论框架，或可依据时频不确定性原理或啁啾

目标信号的特征。

最后，小波变换能够分离时频平面中的特征，这表明该框架非常适合解决重叠源的“鸡尾酒会”问题。尽管本研究聚焦于单源推断，该模型也可扩展用于全局拟合分析。通过在多源实现上进行训练，网络能够学习重叠银河系双星的群体特征，从而更接近为完整的 LISA目录提供全面的解决方案。

该Python第IVA 节中 Lisa 类信号应用的代码可在 GitHub 上获取。[4]。

致谢

我们感谢 Avi Vajpeyi 提出了采用小波变换这一富有见地的建议，并感谢他协助使用 pywavelet 软件包进行实现。所有计算均在单张 NVIDIA L40S-48Q GPU 和 Ubuntu Linux 操作系统上完成。模型训练使用 Python 软件包 PyTorch 实现；MAF模型使用 nflows 实现 [89]；我们感谢奥克兰大学电子研究中心（CeR）提供 Nectar研究云的使用权限并给予协助。Ruiting Mao 感谢奥克兰大学提供的博士奖学金。MCE 和 JEL 感谢新西兰政府资助的 Marsden资助项目 MFP-UOA2131，该项目由新西兰皇家学会 Te Ap¯arangi 管理。

[1] P. Amaro-Seoane、H. Audley、S. Babak、J. Baker、E. Barausse、P. Bender、E. Berti、P. Binetruy、M. Born、D. Bortoluzzi、J. Camp、C. Caprini、V. Cardoso、M. Colpi、J. Conklin、N. Cornish、C. Cutler、K. Danzmann、R. Dolesi、L. Ferraioli、V. Ferroni、E. Fitzsimons、J. Gair、L. G. Bote、D. Giardini、F. Gibert、C. Grimani、H. Halloin、G. Heinzel、T. Hertog、M. Hewitson、K. Holley- Bockelmann、D. Hollington、M. Hueller、H. In- chauspe、P. Jetzer、N. Karnesis、C. Killow、A. Klein、B. Klipstein、N. Korsakova、S. L. Larson、J. Livas、I. Lloro、N. Man、D. Mance、J. Martino、I. Ma- teos、K. McKenzie、S. T. McWilliams、C. Miller、G. Mueller、G. Nardini、G. Nelemans、M. Nofrarias、A. Petiteau、P. Pivato、E. Plagnol、E. Porter、J. Re- iche、D. Robertson、N. Robertson、E. Rossi、G. Rus- sano、B. Schutz、A. Sesana、D. Shoemaker、J. Slut- sky、C. F. Sopuerta、T. Sumner、N. Tamanini、I. Thorpe、M. Troebs、M. Vallisneri、A. Vecchio、D. Vetrugno、S. Vitale、M. Volonteri、G. Wanner、H. Ward、P. Wass、W. Weber、J. Ziemer和P. Zweifel，《激光干涉仪空间天线》，（2017），arXiv:1702.00786 [astro-ph.IM]。[2] M. Colpi、K. Danzmann、M. Hewitson、K. Holley- Bockelmann、P. Jetzer、G. Nelemans、A. Pe- titeau、D. Shoemaker、C. Sopuerta、R. Stebbins、N. Tanvir、H. Ward、W. J. Weber、I. Thorpe, A. Daurskikh, A. Deep, I. F. N´u˜nez, C. G. Marirrodriga, M. Gehler, J.-P. Halain, O. Jen- nrich, U. Lammers, J. Larra˜naga, M. Lieser, N. L¨utzgendorf, W. Martens, L. Mondin, A. P. Ni˜no, P. Amaro-Seoane, M. A. Sedda, P. Auclair, S. Babak, Q. Baghi, V. Baibhav, T. Baker, J.-B. Bayle, C. Berry, E. Berti, G. Boileau, M. Bonetti, R. Brito, R. Buscicchio, G. Calcagni, P. R. Capelo, C. Caprini, A. Caputo, E. Castelli, H.-Y. Chen, X. Chen, A. Chua, G. Davies, A. Derdzinski, V. F. Domcke, D. Doneva, I. Dvorkin, J. M. Ezquiaga, J. Gair, Z. Haiman, I. Harry, O. Hartwig, A. Hees, A. Heffernan, S. Husa, D. Izquierdo, N. Karne- sis, A. Klein, V. Korol, N. Korsakova, T. Kupfer, D. Laghi, A. Lamberts, S. Larson, M. L. Jeune, M. Lewicki, T. Littenberg, E. Madge, A. Mangiagli, S. Marsat, I. M. Vilchez, A. Maselli, J. Mathews, M. van de Meent, M. Muratore, G. Nardini, P. Pani, M. Peloso, M. Pieroni, A. Pound, H. Quelquejay- Leclere, A. Ricciardone, E. M. Rossi, A. Sartirana, E. Savalle, L. Sberna, A. Sesana, D. Shoemaker, J. Slutsky, T. Sotiriou, L. Speri, M. Staab, D. Steer, N. Tamanini, G. Tasinato, J. Torrado, A. Torres-Orjuela, A. Toubiana, M. Vallisneri, A. Vecchio, M. Volonteri, K. Yagi, and L. Zwick, “LISA 定义研究报告”, (2024),arXiv:2402.07571 [astro- ph.CO].

[3]J. Baker, J. Bellovary, P. L. Bender, E. Berti, R. Caldwell, J. Camp, J. W. Conklin, N. Cor- nish, C. Cutler, R. DeRosa, *et al.*, arXiv预印本 arXiv:1907.06482（2019）。

[4] https://github.com/bpandamao/SBI_gaps

[4] O. Burke, S. Marsat, J. R. Gair, 和 M. L. Katz, 《物理评论D》111, 124053（2025）。[5] E. Berti、V. Cardoso 和 C. M. Will, 《物理评论D》73, 064030（2006）。[6] A. Sesana、F. Haardt、P. Madau 和 M. Volonteri, 《天体物理学杂志》623, 23（2005）。[7] J. R. Gair、S. Babak、A. Sesana、P. Amaro-Seoane、E. Barausse、C. P. L. Berry、E. Berti 和 C. Sop- uerta, 《物理学杂志：会议系列》840, 012021 2017( )。[8] S. Babak、J. Gair、A. Sesana、E. Barausse、C. F. Sopuerta、C. P. L. Berry、E. Berti、P. Amaro-Seoane、A. Petiteau 和 A. Klein, 《物理评论D》95（2017）, 10.1103/physrevd.95.103012。[9] N. Cornish 和 T. Robson, 《物理学杂志：会-议系列》840, 012024（2017）。[10] T. B. Littenberg 和 N. Yunes, Classical and Quan-tum Gravity 36, 095017（2019）。[11] B. Willems、A. Vecchio 和 V. Kalogera, Phys-ical Review Letters 100（2008）, 10.1103/phys-revlett.100.041102。[12] T. B. Littenberg 和 N. J. Cornish, 《物理评论D》107, 063004（2023）。[13] Q. Baghi、J. I. Thorpe、J. Slutsky、J. Baker、T. Dal Canton、N. Korsakova 和 N. Karnesis, 《物理评论D》100, 022003（2019）。[14] N. Pearson 和 N. J. Cornish, arXiv预印本 arXiv:2509.05479（2025）。[15] Q. Baghi, arXiv预印本 arXiv:2204.12142 （ 2022 ）。[16] M. Armano 、 H. Audley、G. Auger、J. T. Baird、M. Bassan、P. Binetruy、M. Born、D. Bortoluzzi、N. Brandt、M. Caleno、L. Carbone、A. Cavalleri、A. Cesarini、G. Ciani、G. Congedo、A. M. Cruise、K. Danzmann、M. de Deus Silva、R. De Rosa、M. Diaz-Aguil´o、L. Di Fiore、I. Diepholz、G. Dixon、R. Dolesi、N. Dunbar、L. Ferraioli、V. Ferroni、W. Fichter、E. D. Fitzsimons、R. Flatscher、M. Fres- chi、A. F. Garc´ıa Mar´ın、C. Garc´ıa Marirro- driga、R. Gerndt、L. Gesa、F. Gibert、D. Giar- dini、R. Giusteri、F. Guzm´an、A. Grado、C. Gri- mani、A. Grynagier、J. Grzymisch、I. Harrison、G. Heinzel、M. Hewitson、D. Hollington、D. Hoy- land、M. Hueller、H. Inchausp´e、O. Jennrich、P. Jetzer、U. Johann、B. Johlander、N. Karne- sis、B. Kaune、N. Korsakova、C. J. Killow、J. A. Lobo、I. Lloro、L. Liu、J. P. L´opez-Zaragoza、R. Maarschalkerweerd、D. Mance、V. Mart´ın、L. Martin-Polo、J. Martino、F. Martin-Porqueras、S. Madden、I. Mateos、P. W. McNamara、J. Mendes、L. Mendes、A. Monsky、D. Nicolodi、M. Nofrarias、S. Paczkowski、M. Perreur-Lloyd、A. Petiteau、P. Pi- vato、E. Plagnol、P. Prat、U. Ragnit、B. Ra¨ıs、J. Ramos-Castro、J. Reiche、D. I. Robertson、H. Rozemeijer、F. Rivas、G. Russano、J. Sanju´an、P. Sarra、A. Schleicher、D. Shaul、J. Slutsky、C. F. Sopuerta、R. Stanga、F. Steier、T. Sumner、D. Tex- ier、J. I. Thorpe、C. Trenkel、M. Tr¨obs、H. B. Tu、D. Vetrugno、S. Vitale、V. Wand、G. Wan- ner、H. Ward、C. Warren、P. J. Wass、D. Wealthy、W. J. Weber、L. Wissel、A. Wittchen、A. Zambotti、C. Zanoni、T. Ziegler 和 P. Zweifel, 《物理评论快报》116, 231101（2016）。[17] M. Armano、H. Audley、J. Baird、P. Binetruy、M. Born、D. Bortoluzzi、E. Castelli、A. Cav- alleri、A. Cesarini、A. M. Cruise、K. Danz- mann、M. de Deus Silva、I. Diepholz、G. Dixon、R. Dolesi、L. Ferraioli、V. Ferroni、E. D. Fitzsi- mons、M. Freschi、L. Gesa、F. Gibert、D. Giar- dini、R. Giusteri、C. Grimani、J. Grzymisch、I. Har- rison、G. Heinzel、M. Hewitson、D. Hollington、D. Hoyland、M. Hueller、H. Inchausp´e、O. Jennrich、P. Jetzer、N. Karnesis、B. Kaune、N. Korsakova、C. J. Killow、J. A. Lobo、I. Lloro、L. Liu、J. P. L´opez-Zaragoza、R. Maarschalkerweerd、D. Mance、N. Meshksar、V. Mart´ın、L. Martin-Polo、J. Mar- tino、F. Martin-Porqueras、I. Mateos、P. W. Mc- Namara、J. Mendes、L. Mendes、M. Nofrarias、S. Paczkowski、M. Perreur-Lloyd、A. Petiteau、P. Pi- vato、E. Plagnol、J. Ramos-Castro、J. Reiche、D. I. Robertson、F. Rivas、G. Russano、J. Slutsky、C. F. Sopuerta、T. Sumner、D. Texier、J. I. Thorpe、D. Vetrugno、S. Vitale、G. Wanner、H. Ward、P. J. Wass、W. J. Weber、L. Wissel、A. Wittchen 和 P. Zweifel, 《物理评论快报》120, 061101（2018）。[18] E. Castelli、Q. Baghi、J. G. Baker、J. Slutsky、J. Bobin、N. Karnesis、A. Petiteau、O. Sauter、P. Wass 和 W. J. Weber, 《经典与量子引力》42, 065018（2025）。[19] L. Wang、H.-Y. Chen、X. Lyu、E.-K. Li 和 Y.- M. Hu, “加窗与修复：处理天琴数据缺口”（2024）, arXiv:2405.14274 [gr-qc]。[20] J. R. Gair、S. Deng 和 S. Babak, “数据缺口和毛刺对 LISA 数据中瞬态引力波的污染”（2025）, arXiv:2510.16770 [gr-qc]。[21] J. Carr´e 和 E. K. Porter, “数据缺口对 LISA 银河系双星参数估计的影响”（2010）, arXiv:1010.1641。[22] K. Dey、N. Karnesis、A. Toubiana、E. Barausse、N. Korsakova、Q. Baghi 和 S. Basak, 《物理评论D》104（2021）, 10.1103/physrevd.104.044035。[23] R. Mao、J. E. Lee 和 M. C. Edwards, 《物理评论D》111, 024067（2025）。[24] J. Blackman、B. Szilagyi、C. R. Galley 和 M. Tiglio, 《物理评论快报》113（2014）, 10.1103/physrevlett.113.021101。[25] K. Cranmer、J. Brehmer 和 G. Louppe, 《美国国家科学院院刊》117, 30055（2020）。[26] H. Gabbard、C. Messenger、I. S. Heng、F. Tono- lini 和 R. Murray-Smith, 《自然·物理》18, 112

（2022）。[27] S. R. Green、C. Simpson 和 J. Gair，《物理评论D》102，104057（2020）。[28] M. J. Williams、J. Veitch 和 C. Messenger，《物理评论D》103，103006（2021）。[29] F. Stachurski、C. Messenger 和 M. Hendry，《物理评论D》109，123547（2024）。[30] R. Srinivasan、E. Barausse、N. Korsakova 和 R. Trotta，arXiv预印本 arXiv:2506.22543（2025）。[31] M. Dax、S. R. Green、J. Gair、J. H. Macke、A. Buo- nanno 和 B. Sch ¨olkopf，《物理评论快报》（）。[32] B. Liang 和 H. Wang，《天文技术与仪器》（2025）。[33] J. Wildberger、M. Dax、S. R. Green、J. Gair、M. P ¨urrer、J. H. Macke、A. Buonanno 和 B. Sch ¨olkopf，《物理评论D》107，084046（2023）。[34] V. Raymond、S. Al-Shammari 和 A. G ¨ottel，《皇家天文学会月刊》542，1103（2025），https://academic.oup.com/mnras/article-pdf/542/2/1103/63957040/staf1282.pdf。[35] R. Legin、M. Isi、K. W. Wong、Y. Hezaveh 和 L. Perreault-Levasseur，《天体物理学杂志快报》985，L46（2025）。[36] C.-Y. Xiong、T.-Y. Sun、J.-F. Zhang 和 X. Zhang，《物理评论D》111，024019（2025）。[37] Y. Lipman、R. T. Chen、H. Ben-Hamu、M. Nickel 和 M. Le，arXiv预印本 arXiv:2210.02747（2022）。[38] J. Wildberger、M. Dax、S. Buchholz、S. Green、J. H. Macke 和 B. Sch ¨olkopf，《神经信息处理系统进展》36，16837（2023）。[39] B. Liang、M. Du、H. Wang、Y. Xu、C. Liu、X. Wei、P. Xu、L.-e. Qiang 和 Z. Luo，《机器学习：科学与技术》5，045040（2024）。[40] B. Liang、H. Guo、T. Zhao、H. Wang、H. Evange- linelis、Y. Xu、C. Liu、M. Liang、X. Wei、Y. Yuan、M. Du、P. Xu、W. Qian 和 Z. Luo，《中国物理快报》42，081101（2025）。[41] R. Srinivasan、E. Barausse、N. Korsakova 和 R. Trotta，“LISA 银河系双星的基于模拟的总体推断：绕过全局拟合”（2025），arXiv:2506.22543 [astro-ph.GA]。[42] M. Park、M. Gatti 和 B. Jain，《物理评论D》111，063523 2025（）。[43] H.-Y. Liu 和 Y.-T. Wang，“用于引力波信号识别的自监督学习”（2023），arXiv:2302.00295 [gr-qc]。[44] D. George 和 E. Huerta，《物理学快报B》778，64 70 2018–（）。[45] Y. X. Fu 和 W. H. Li，《分子生物学与进化》14，195（1997），https://academic.oup.com/mbe/article-pdf/14/2/195/11165471/9fu.pdf。[46] G. Weiss 和 A. von Haeseler，《遗传学》**149**，1539（1998）。[47] P. Fearnhead 和 D. Prangle，《皇家统计学会会刊：B辑（统计方法学）》**74**，419（2012），https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2011.01010.x。[48] B. Jiang、T.-Y. Wu、C. Zheng 和 W. H. Wong，《辛尼卡统计学》27，1595（2017）。[49] D. Lanchares、O. G. Freitas、J. Gonz´alez-Nuevo 和 J. A. Font，“评估归一化流在引力波数据时频表示参数估计中的表现”，（2025），arXiv:2505.08089 [gr-qc]。[50] L. S. Finn，《物理评论D》46，5236（1992）。[51] E. E. Flanagan 和 S. A. Hughes，《物理评论D》57，4566 1998 9710129（），arXiv:gr-qc/。[52] R. Meyer、M. C. Edwards、P. Maturana-Russel 和 N. Christensen，《威利跨学科评论：计算统计学》14，e1532（2022）。[53] A. I. Burke 和 O. Burke，（2021）。[54] G. Papamakarios、T. Pavlakou 和 I. Murray，《神经信息处理系统进展》30（2017）。[55] R. T. Chen、Y. Rubanova、J. Bettencourt 和 D. K. Duvenaud，《神经信息处理系统进展》31（2018）。[56] W. Grathwohl、R. T. Chen、J. Bettencourt、I. Sutskever 和 D. Duvenaud，arXiv预印本 arXiv:1810.01367（2018）。[57] C. Finlay、J.-H. Jacobsen、L. Nurbekyan 和 A. Oberman，载于 *International conference on machine learning* 2020 3154 3164 （PMLR，）第 – 页。[58] X. Liu、C. Gong 和 Q. Liu，“流动直接而快速：利用整流流学习生成和迁移数据”，（2022），arXiv:2209.03003 [cs.LG]。[59] N. Kornilov、P. Mokrov、A. Gasnikov 和 A. Korotin，《神经信息处理系统进展》37，104180（2024）。[60] N. J. Cornish，《物理评论D》102，124038（2020），arXiv:2009.00043 [gr-qc]。[61] F. Yu 和 V. Koltun，arXiv预印本 arXiv:1511.07122（2015）。[62] Z. Zhong、L. Zheng、G. Kang、S. Li 和 Y. Yang，“随机擦除数据增强”，（2017），arXiv:1708.04896 [cs.CV]。[63] S.-A. Rebuffi、S. Gowal、D. A. Calian、F. Stimberg、O. Wiles 和 T. A. Mann，《神经信息处理系统进展》34，29935（2021）。[64] L. Perez 和 J. Wang，arXiv预印本 arXiv:1712.04621（2017）。[65] R. Takahashi、T. Matsubara 和 K. Uehara，《IEEE电路与系统视频技术汇刊》30，2917（2019）。

[66] K. Maharana、S. Mondal 和 B. Nemade，Global Transitions Proceedings 3，91（2022）。[67] C. Shorten 和 T. M. Khoshgoftaar，Journal of big data 6，1（2019）。[68] Y. Bengio、J. Louradour、R. Collobert 和 J. Weston，载于*Proceedings of the 26th annual international conference on machine learning* 2009 41 48 （）第 – 页。[69] H. Einsle 、 M. A. Bizouard 、 T. Regimbau 和 M. Sakellariadou，《物理评论D》112，063056（2025）。[70] P. Soviany、R. T. Ionescu、P. Rota 和 N. Sebe，International Journal of Computer Vision 130，1526（2022）。[71] F. Liu、T. Zhang、C. Zhang、L. Liu、L. Wang 和 B. Liu，Electronics 12，1676（2023）。[72] J. Salamon 和 J. P. Bello，IEEE Signal Processing Letters 24，279（2017）。[73] D. S. Park、W. Chan、Y. Zhang、C.-C. Chiu、B. Zoph、E. D. Cubuk 和 Q. V. Le，arXiv预印本 arXiv:1904.08779（2019）。[74] Y. Lipman、M. Havasi、P. Holderrieth、N. Shaul、M. Le、B. Karrer、R. T. Q. Chen、D. Lopez-Paz、H. Ben-Hamu 和 I. Gat，“流匹配指南与代码”，（2024），arXiv:2412.06264 [cs.LG]。[75] D. P. Kingma 和 J. Ba，“Adam：一种随机优化方法”，（ 2017 1412 6980 ），[ ] arXiv:。cs.LG。[76] J. Howard 和 S. Ruder，载于 *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*（2018），第328–339页。[77] I. Loshchilov 和 F. Hutter，arXiv预印本 arXiv:1608.03983（2016）。[78] M. Katz，“fastlisaresponse”，https://github.com/mikekatz04/lisa-on-gpu（2021）。[79] M. L. Katz、J.-B. Bayle、A. J. K. Chua 和 M. Vallisneri，《物理评论D》106，103001（2022）。[80] N. Wiener *et al.*，《数学学报》55，117（1930）。[81] A. Khintchine，《数学年刊》109，604（1934）。[82] M. Katz 和 L. A. T. Collaboration，“mikekatz04/LISAanalysistools：v1.1.0”（2025）。[83] P. Amaro Seoane、M. Arca Sedda、S. Babak、C. P. L. Berry、E. Berti、G. Bertone、D. Blas、T. Bogdanovi´c、M. Bonetti、K. Breivik、R. Brito、R. Caldwell、P. R. Capelo、C. Caprini、V. Cardoso、Z. Carson、H.- Y. Chen、A. J. K. Chua、I. Dvorkin、Z. Haiman、L. Heisenberg、M. Isi、N. Karnesis、B. J. Ka- vanagh、T. B. Littenberg、A. Mangiagli、P. Mar- coccia、A. Maselli、G. Nardini、P. Pani、M. Peloso、M. Pieroni、A. Ricciardone、A. Sesana、N. Tamanini、A. Toubiana、R. Valiante、S. Vretinaris、D. J. Weir、K. Yagi 和 A. Zimmerman，《广义相对论与引力》54（2021），10.1007/s10714-021-02889-x。[84] M. Colpi、K. Danzmann、M. Hewitson、K. Holley- Bockelmann、P. Jetzer、G. Nelemans、A. Petiteau、D. Shoemaker、C. Sopuerta、R. Stebbins、 *et al.*，arXiv预印本 arXiv:2402.07571（2024）。[85] R. Mao、J. E. Lee、O. Burke、A. J. K. Chua、M. C. Edwards 和 R. Meyer，《物理评论D》109，083002（2024）。[86] N. Cornish，“WDM变换”，https://github. com/eXtremeGravityInstitute/WDM_Transform（2020）。[87]M. Digman，“WDMWaveletTransforms”，https: // github.com/XGI-MSU/WDMWaveletTransforms（2023）。[88] T. B. Littenberg、N. J. Cornish、K. Lackeos 和 T. Robson，《物理评论D》101（2020），10.1103/physrevd.101.123021。[89] C. Durkan、A. Bekasov、I. Murray 和 G. Papa- makarios，“nflows：PyTorch 中的归一化流”（2020）。