Draft version November 25, 2026年 使用 LaTeX 双栏 样式，采用 AASTeX 7.0.1 排版

arXiv:2611.33092v1 [astro-ph.IM] 24 Nov 2026

Hesperos 与 Eosphoros：用于 SPHEREx 光谱测光的耦合式可微强制测光算子与前向模型

James R. Cheshire IV , [1] Spencer Everett , [1] Richard M. Feder , [2] James Bock , [1, 3] Tzu-Ching Chang , [3, 1] Yi-Kuan Chiang , Brendan P. Crill , [3, 1] Olivier Doré , [3, 1] Andreas L. Faisst , [5] Zhaoyu Huai , [1] Howard Hui , [1, 3] Gabriela Torrini [5] And Michael Zemcov [6, 3] The SPHEREx Science Team

[1] *Department of Physics, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA*
[2] *University of California at Berkeley, Berkeley, CA 94720, USA* [3] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA* [4] *Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan* [5] *IPAC, California Institute of Technology, MC 100-22, 1200 E California Blvd, Pasadena, CA 91125, USA* [6] *School of Physics and Astronomy, Rochester Institute of Technology, 1 Lomb Memorial Dr., Rochester, NY 14623, USA*

摘要

我们介绍 Hesperos & Eosphoros：一个强制测光算子和一个前向模型，分别用于执行和模拟宇宙历史、再电离时代与冰探索者光谱光度计（SPHEREx）的光谱测光测量。这两个程序以图形处理器加速的可微分 JAX 程序实现，可直接组合起来，测量下游产品（例如星系密度场）相对于建模的系统误差或扰动的响应，贯穿整个测量流水线。我们介绍其设计，并展示如何组合它们来计算天体测量扰动的响应；同时演示如何利用它们，为已知的图像级效应生成直接测量/流水线响应模板 ∂（重建密度）$/\partial$（系统振幅）。在生产规模下， Hesperos 每节点小时的吞吐量比先前基于中央处理器的 Tractor测量提高了 $\sim 700\times$ 倍；该测量此前是本实验数据处理的瓶颈。它还能在整个生产测量算子中传播自动导数。我们将测量算子 Hesperos与 The Tractor 的公开 SPHEREx 分支进行验证，并将前向模型 Eosphoros与 SPHEREx 天空模拟器进行验证，同时独立地将渲染结果与 GalSim总的来说，这为扩展、迭代和优化星系聚类测量流水线的能力带来了强有力的提升。

关键词： 宇宙学（343）、宇宙大尺度结构（902）、GPU计算（1969）、红外测光（792）、天文数据分析（1858）

1. 引言

当前及即将开展的宇宙大尺度结构（LSS）巡天（包括SPHEREx）的一个关键目标，是通过局域原初非高斯性 $f_{\rm NL}^{\rm loc}$在星系大尺度聚类中的印记来测量或约束N. Dalal 等， 2008；O. Dor´e 等， 2014）。对于 $f_{\rm NL}^{\rm loc}$ 而言，所关注的平滑大尺度特征在形式上与样本密度的大尺度调制所造成的污染特征极为相似由于非宇宙学效应，例如各向异性的恒星污染或银河系尘埃红化对河外源的影响。自最早的一些尝试（A. R. Pullen & C. M. Hirata 2013年）以来，人们就已清楚认识到，这些大角尺度系统误差及其清理会对恢复 $f_{\rm NL}^{\rm loc}$ 产生深远影响，而随着宇宙学样本统计保真度的不断提高，这种重要性也日益凸显。

具有已测量或建模空间剖面的系统误差，例如恒星污染和银河系尘埃消光，可能会在某种描述这些

电子邮箱：cheshire@caltech.edu

系统误差“模板”如何与该密度场耦合（D. Huterer 等，2013年；F. Elsner 等，2016年；N. Weaverdyck 与 D. Huterer，2021年）。尽管一些显而易见的系统误差污染源具有天体物理学性质，但有相当一类污染源来自观测者和巡天相关因素对样本选择的影响。对于 SPHEREx 而言，这两类因素存在一定重叠，例如，观测深度由黄道光辐射产生的光子噪声所决定。与观测相关的选择效应可能会给密度场引入更加微妙且非线性的异常，而这些异常很难通过模板去投影得到完全校正。对于 SPHEREx 这类强制测光实验，密度场的虚假调制可能同时继承自参考母星表自身的成像巡天以及实验本身的测量。本文聚焦于后者。此类观测选择效应高度依赖于测光测量本身及其相关不确定性，而后两者又取决于局部背景/密度、近距离明亮混合近邻源的存在、点扩散函数的变化（跨越天空以及焦平面），以及曝光特定瞬变异常或系统误差（例如，宇宙射线、图像边缘鬼影，见C. M. Fazar 等，2026年）。该实验的真实选择函数继承了所有这些属性，涵盖所有已测量像素、算法以及潜在的软件缺陷，因此需要仔细理解测量算子本身，以及它与这些效应之间的耦合关系。

表征这种响应的一种常见方法是合成源注入，即将具有已知属性的合成源直接插入真实的巡天图像中，并将其响应传递至完整的生产流水线。通过将响应与已知输入进行比较，可以直接测量选择函数中的重要组成部分，但计算成本是渲染注入的源并重新运行生产测光流程。注入过程对响应中存在的非线性几乎不作假设（并能捕捉这种非线性），因此大规模应用时是一种强大的工具。 BAL-ROG (E. Suchyta 等，2016年；S. Everett 等，2022年；D. Anbajagane 等，2025年) 对暗能量巡天开展了此类分析， SYNPIPE (S. Huang 等，2018年) 对 Hyper Suprime-Cam 开展了此类分析， OBIWAN (H. Kong 等，2020年) 则对扩展重子振荡光谱巡天和暗能量光谱仪开展了此类分析。H. Kong 等 (2025年) 利用图像模拟对暗能量光谱仪发光红星系（LRG）的密度调制进行了正向建模；基于此类图像模拟的估算与神经网络系统误差回归方法结合使用，见 OBIWAN；A. J. Rosado-Marín 等(2025年)，值得注意的是，他们还纳入了对系统误差红移依赖性的校正。

传统注入必须照单全收测量过程中的某些特征，尤其是点扩散函数（PSF）、图像级校准、天体测量学，以及银河系消光等天区效应，因此无法对这些因素进行表征。如果能够直接对图像级系统误差或测量属性的幅度求导，从而得到光度测量的响应，且该响应是平滑的，那么就可以直接计算选择函数对该系统误差的响应，计算成本仅取决于测量算子本身以及系统误差响应重建所需的分辨率。可微分处理流程也便于采用基于梯度的采样技术，这可能使其他类别的推断成为可能，尤其是在结合降维技术时。现代数组计算框架（如 JAX（J. Bradbury 等，2018年）提供了易于使用的接口，可用于构建采用自动微分（AD）的处理流程。自动微分的高效实现通常还能结合高效的即时编译（JIT）和图形处理器（GPU）加速。因此，构建快速、可微分的光度正向建模与测量处理流程，将有力拓展选择函数表征能力，并直接补充源注入。

利用导数表征流水线响应的方法已应用于弱引力透镜领域，在该领域中，检测会对测量剪切作出响应。最初采用数值方法实现这一点（E.S.Sheldon &E. M. Huff2017；E.S.Sheldonet al. 2020），后来也采用解析方法（X. Li & R. Mandelbaum 2023）以及自动微分（X.Li etal. 2024）。I. Mendoza et al. (2026年) 提出了一种快速、可微分的图形处理器流水线实现。这些方法都将微分和基于梯度的采样应用于感兴趣的信号，而非系统误差。可微分、自优化的流水线已在高能物理领域得到开发和广泛应用（P. de Castro &T. Dorigo 2019；N. Simpson&L.Heinrich 2023；G.C. Strong etal.2024；M.Hance&J. Robles 2025），但针对系统误差响应对流水线配置进行边缘化或优化，似乎仍是现代天文成像巡天中一个尚未充分探索的能力。

通过控制并组合图像级前向模型和强制测光测量管线，使其成为输入属性的单一可微函数，我们可以直接将图像级测光测量显著性表示为任何

这些输入所依赖的变量的函数，代价是每个求导维度需进行少量次前向计算以获得自动导数。这在很大程度上依赖于处理流程本身的性能，因此优化至关重要。

在这项工作中，我们介绍测量算子和前向模型（两者均以 JAX 实现），以及它们的组合；三者共同构成了这样一条处理流水线。HESPEROS[7] 是对 SPHEREx 窗口化强制测光测量的批处理、针对图形处理器优化的重新实现，并通过与 THE TRACTOR (D. Lang 等人, 2016a）的公开分支以及与 SPHEREx科学数据中心（SSDC）的生产 TRAC-TOR流程生成的星表进行对照验证（R. Akeson 等人, 2026）。 THE TRAC-TOR是一种通用的概率光度测量前向建模工具，也可以优化源位置、形状和点扩散函数；SSDC 使用其固定这些参数的模式，将问题简化为针对每个测光测量窗口的直接加权最小二乘问题。 HESPEROS专为直接处理这一特殊情况而设计，使图形处理器实现和可微性切实可行。

EOSPHO-ROS是一种加速且可微分的光谱测光前向模型，可在 SPHEREx探测器帧中渲染源；其渲染结果已通过与 GAL-SIM (B. T. P. Rowe 等人, 2015）的渲染结果进行对照验证。此外，它还包含经过验证的重新实现，用于将原始通量测量分箱到红移拟合所用的“基准”分箱（第6.1节），以及红移估计阶段。将它们组合起来，便可计算系统误差经过渲染及同一测量流水线（HESPEROS）时的导数，而该流水线也可应用于真实图像数据。

我们注意到，可微分且采用GPU加速的测光处理流程并非前所未有： ASTROPHOT (C. Stone 等, 2023年) 对未强制的测光测量参数（位置、形状、点扩散函数）采用了基于加速梯度的优化。相比之下，我们固定位置和形状，并对曝光中的 $\mathcal{O}(10^5)$ 拟合窗口所对应的线性系统进行求解，将导数通过测量传播至星表级量的响应。E. F. Schlafly 等人(2018年) 使用中央处理器对 DECam 图像执行了类似的批量线性求解，但不涉及导数。H. Bahk 与 H. S. Hwang (2026年) 提出了一种基于 JAX 的 THE TRACTOR 重新实现，并将其应用于公开的 SPHEREx 数据，在图形处理器上运行；不过，该工作侧重于深度恢复和测光的外部先验，而非系统误差、可微性和生产吞吐量。

论文中两者及其组成部分的示意图见图1。第2 节介绍 SPHEREx 的测量问题，以及其新型光谱测光巡天所特有的方面。第3节描述测量算子 HESPEROS，第4节描述 EOSPHOROS 前向模型，详细说明其设计、验证和吞吐量。第5 节描述二者的构成，并确立推断出的差分响应在何种条件下可信。第6 节描述从测得的测光目录到测得的三维密度场的传播过程（在EOSPHOROS中实现）。第7 节展示这一构成如何实现规模化应用，并在第8节讨论其局限性、新能力以及与源注入的关系。我们在第9节总结。

2. SPHEREx 测量

SPHEREx 是美国国家航空航天局的一项中型探索者级任务，旨在开展一项新颖的全天光谱巡天。借助线性可变滤光片（LVF），该任务可在 0.75 至 5.0 μm（J. J. Bock 等, 2026年），通过 102个基准波长通道绘制全天空图。线性可变滤光片的通带沿滤光片的一个轴线性变化（即安装在构成 SPHEREx焦平面的六个探测器阵列上时，图像平面的“y”方向），因此可在探测器视场（FoV）的不同位置对同一源进行一系列测量，从而重建低分辨率光谱（$R \sim 35$–130 个 $6\farcs15$ 像素）（P. M. Korngut 等, 2026年）。线性可变滤光片将光度测量过程与光谱采样直接耦合。这种模式对于光度巡天而言并不陌生，但此处由于分辨率的缘故，光谱耦合的重要性显著提高，因此采用“光谱光度测量”一词。观测波长因而也取决于几何因素，天体测量扰动会直接与光谱扰动耦合，从而可能直接影响红移估计。

SPHEREx 是一项强制测光巡天（D. Lang 等, 2016b；A. Dey 等, 2019），其源位置从一开始就根据外部参考成像目录 SPHEREx参考目录（Y. Yang 等, 2026年，编制中）确定。SPHEREx参考目录通过整合各种全天和大面积光学与红外巡天，并纳入经过筛选的数据，旨在支持 SPHEREx 的核心科学目标以及遗产目录的构建

[7] HESPEROS & EOSPHOROS 分别指代象征晚间和清晨星辰的古希腊神祇。二者都指金星，因此 HESPEROS& EOSPHOROS的名称也反映了强制测光算子和前向模型如何表现同一测量的两个方面。

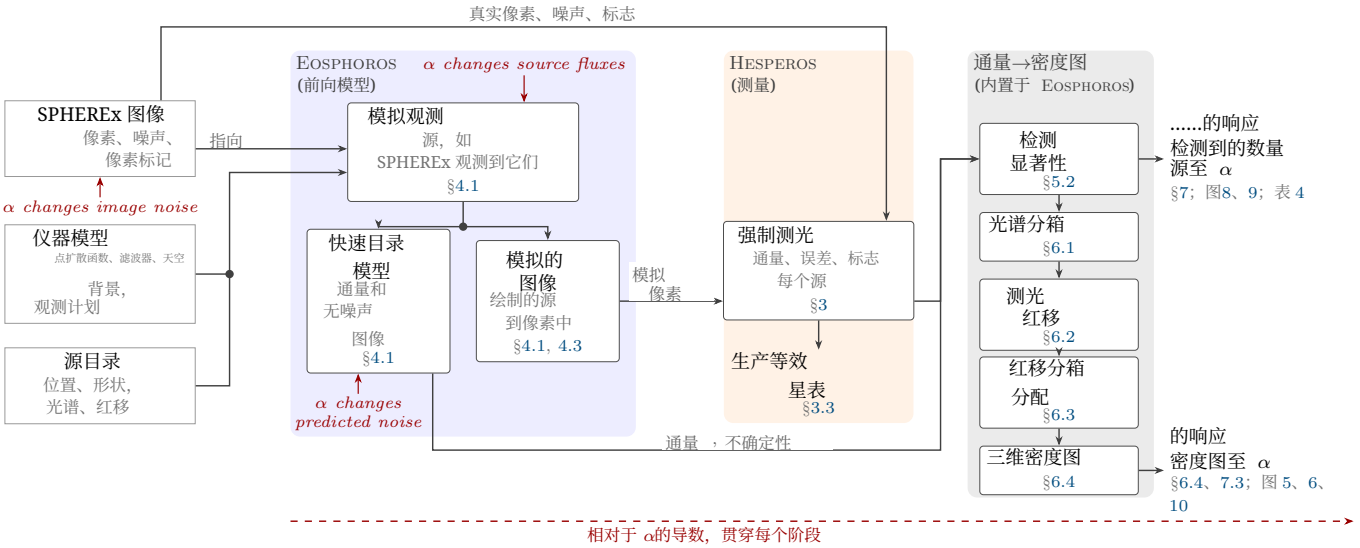


图1. HESPEROS & EOSPHOROS 测量的输入、输出和数据流概览。左侧从图像、仪器和源星表的输入产品开始，这些产品随后通过实线箭头所示的方式向下游提供数据。 EOSPHOROS前向模型根据帧几何、仪器属性和输入源组装场景，然后执行快速的目录直接评估，或将其图像级信息传递给生产测量算子 HESPEROS；真实数据则直接传递给 HESPEROS 。由此生成的星表可选择经过检测显著性步骤、光谱分箱、光度红移拟合和红移分箱分配，以生成三维密度图。系统振幅 α 可在红色标记的位置引入；它既可能直接影响测量通量，这种影响必须通过渲染源并使用 EOSPHOROS建模来估计；也可能影响真实图像中的方差或快速模型的噪声预测。关于系统误差 α 的导数会传播至每个阶段。

产品。位置和点扩散函数（对于扩展天体拟合，还包括形状）均为固定值，这意味着像素数据模型关于源通量是线性的，而这一点是第 3 节的关键基础。一个重要推论是，这些测量值以已知源及其位置、已知点扩散函数和图像定标为前提；因此，当前实现适用于 SPHEREx 这类强制测光实验，但未来计划将其扩展为使用类似算法进行源检测。

“级别”0–4用于指代不同的处理阶段，既指处理步骤，也指相关的输出产品；详情请参见R. Akeson 等人 (2026年）。经过初始星载缩减（0级，L0）以及定标和数据质量检查（L1）后，2级（L2）执行测光和天体测量定标以及像素标记，以生成定标后的线性可变滤光片图像或光谱数据立方体。3级（L3）是测光测量步骤，4级（L4）则指各种下游科学分析流程。作为测光测量的重新实现，并作为测量下游响应的工具， HESPEROS & EOSPHOROS 横跨3级和4级。我们通常会使用2级定标图像输出产品提供的图像帧结构、源位置、方差层和标志，

其中，3级测光目录是比较的目标，而非直接使用的产品。

3. HESPEROS：可微分强制光谱测光

HESPEROS是对 TRACTOR所描述的、基于 SPHEREx 3级窗口化强制测光流程的可微分 GPU批处理重实现，该流程基于 （2026年）。 HESPEROS 的目的并非相较于 TRAC-TOR所基于的处理流程优化测量本身，而是利用图形处理器大幅提高生产吞吐量，并在与 EOSPHOROS 结合时支持差分前向响应测量（第5节）。

3.1. 设计

在混叠场中，通过独立窗口执行点源强制测光，相当于求解大量规模较小且彼此独立的线性系统。测光测量无需遍历各个窗口，而可实现为单个批处理算子，用以分解数千个独立系统，并可轻松适配所使用的计算资源。

单个 SPHEREx 2级探测器曝光的分辨率为 2048×2048 $6\farcs15$ 像素，其中

中央 2040×2040 对 Level 2 图像中的光线敏感。活动区域被划分为规则网格，每个网格单元边长为 5 像素，共有 $408 \times 408 = 166$ 个，总计 464 个窗口。每个窗口都由一个与相邻窗口重叠的 3 像素帧进行填充，从而得到一个$11 \times 11 = 121$ 总窗口大小（唯一裁剪图加上帧）。对于给定的强制测光源，仅当其属于某个窗口的唯一裁剪图成员时，才会记录其通量、拟合不确定度、背景和质量统计量；此外，它还可能在最多 8 个相邻窗口的帧中被测量，这些测量值将用于估计协方差，随后被丢弃。由于 3 像素帧的宽度小于 5 像素裁剪图，给定源只能延伸到 8 个相邻窗口，因此成员搜索是一个精确的 3×3 邻域搜索，但可能会截断那些通量尾部超出该范围的源的贡献。这种窗口化设置确保无论星表如何，测光都可以表示为一组固定线性系统；与逐源居中裁剪图方法相比，它还最大限度地减少了对同一源的重复观测。窗口化强制测光的几何结构如图 2所示。在探测器边界处，将活动区域之外的像素赋予零权重，以保持窗口的固定维度；这是一个关键的性能考量，将在第 3.4节中加以利用。

对于包含 n 个源的窗口，像素数据 $\mathbf{d}$ 与源通量呈线性关系，

$$\mathbf{d} = \mathrm{A}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \qquad \boldsymbol{\theta} \equiv (\mathbf{f}, b), \tag{1}$$

其中，n 设计矩阵 A 的前几列是归一化有效点扩散函数（ePSF；J.Anderson & I. R. King 2000），在每个源的亚像素位置处计算，并乘以参数向量 $\mathbf{f}$ 中的通量分量 $\boldsymbol{\theta}$；最后一列为常数，表示振幅为 b的局部平坦背景，$\boldsymbol{\epsilon}$ 则为像素噪声。背景振幅与通量共同拟合，不设先验。加权最小二乘（WLS）解为

$$\hat{\boldsymbol{\theta}} = \mathrm{N}^{-1}\mathrm{A}^{\mathrm{T}}\mathrm{W}\mathbf{d}, \quad \mathrm{N} \equiv \mathrm{A}^{\mathrm{T}}\mathrm{W}\mathrm{A}, \tag{2}$$

其中，W 是根据 2级方差平面构建的对角逆协方差权重矩阵，N 是正规矩阵，我们用 $\mathbf{f}$ 表示 $\boldsymbol{\theta}$ 中对应源的部分。

为方便起见，我们定义朴素不确定度 $\sigma_j^{\mathrm{naive}} = (\mathrm{N}_{jj})^{-1/2}$，即法矩阵的对角元素；它将源 j 视为单独处于其窗口中的源。我们还定义边缘不确定度 $\sigma_j^{\mathrm{marg}} = \left[\left(\mathrm{N}^{-1}\right)_{jj}\right]^{1/2}$，该不确定度对与混叠邻近源及背景之间的协方差进行边缘化处理。只有当某个源与其他源均不重叠时，这两者才相等；$\sigma^{\mathrm{marg}} \geq \sigma^{\mathrm{naive}}$ 且始终成立。仅对于退化解（即混叠严重和/或因图像质量标志而有许多像素被排除在拟合之外的解），才会在 N 的对角线上添加一个较小的岭项，以便在计算 σ^{marg}时对其进行正则化；但通量始终通过公式2 计算，不进行这种正则化，以避免产生偏差，并将这些测量值标记为退化。在计算像素立体角时，会考虑包含光学畸变在内的完整世界坐标系（WCS），但无论如何，这项修正都很小（$\sim$ 1在整个探测器阵列上为 %，$\sim 10^{-4}$ 在一个窗口内为 %）。

这种设置以及在每个窗口中单独拟合背景的做法，意味着完整曝光对应的式2 中的正规矩阵严格呈块对角形式，因此拟合可以完全分离，每个窗口只读取属于它的 121 个像素。一个源 j 只会记录在其位于中心裁剪区的那一次测量中，与其他窗口无关；即使其他窗口将源 j 作为邻源进行光度测量，也会完全独立地求解和处理。这使得一次曝光不再需要构建一个庞大的线性系统，而是 $\sim 1.7 \times 10^5$ 个小型、稠密且彼此独立的线性系统。通过窗口构造和零权重边缘填充，像素维度保持统一；而通过将窗口分组到“桶”中，并填充到一组较小的固定源数量（使用空的零权重源），参数维度（即成员源与背景项的总维数）也得以统一。因此，同一桶中的每个窗口都彼此独立且数组形状相同，从而显著提升数组性能。

针对设备内存的优化，每个存储桶会根据窗口槽位预算（源位置，实际位置或填充位置）进一步划分为“数据块”，从而能够扩展到所需的内存利用率。每个数据块都会作为一个即时编译程序一次性发送到设备（图形处理器），该程序会裁剪出每个窗口的像素，评估并归一化有效点扩散函数，同时使用一次批量乔列斯基分解构建并求解所有窗口的线性系统（C. Benoˆıt 1924年），并计算质量标志。只有结果会传回主机。通过存储桶/数据块的划分，可在构造时平衡单次曝光内的源密度；跨图像的负载平衡则通过按源总数对图像分组来实现。所引用的吞吐量数据是对具有代表性的源密度分布取平均后得出的。

$\sim 1.7 \times 10^5$ 线性系统随后只需向主机提交数量与数据块数量同量级的少量设备启动，而呈现给

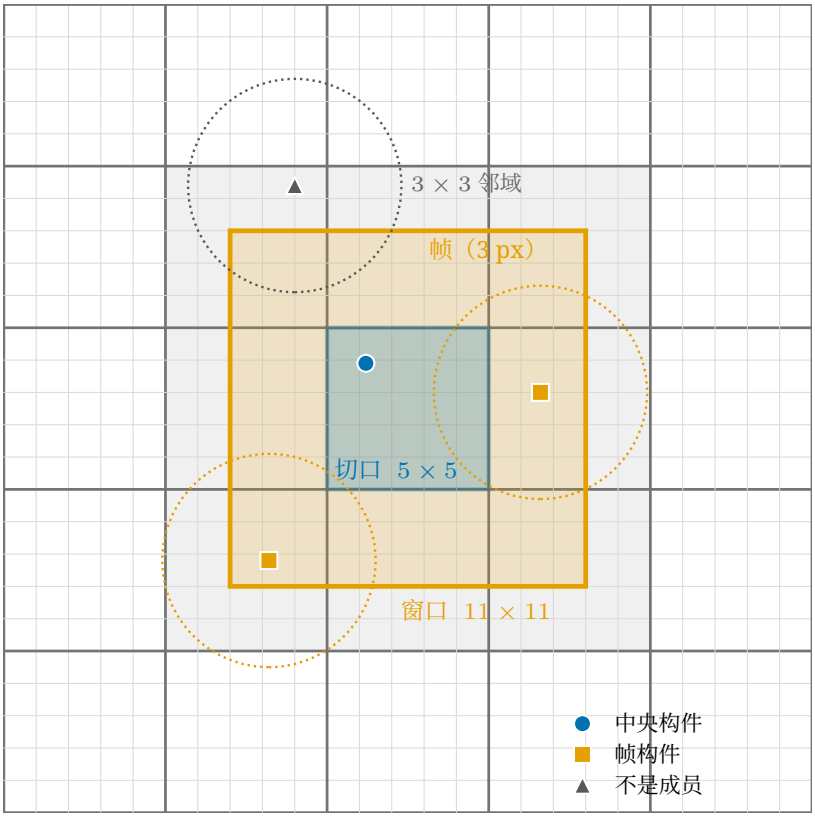


图 2. 拟合窗口几何示意图。对于图像视场网格中的每个唯一裁剪图，窗口由该唯一裁剪图中 5×5 像素的区域（蓝色）以及一个 3 像素的“帧”（橙色）构成，后者用于计入相邻裁剪图中邻居的贡献。裁剪图的 3×3 邻域（灰色）表示必须考虑是否纳入星表源的完整裁剪图集合。位于中央唯一裁剪图中的“中央成员”（蓝色圆圈）会记录其测量值；边缘区域中的“帧成员”（橙色方块）会进行测量，以了解它们对中央成员协方差的贡献，但其测量值会被丢弃；窗口外的源（灰色三角形）则不进行测量。虚线圆圈表示源的有效点扩散函数模型的支持边界（±3.3 像素）。

图形处理器的任务则是其特别擅长的一类任务：由静态形状矩阵上的简单矩阵运算组成的组合。这种架构通过支持可编译性，使该操作具备良好性能；同时也支持可微性，因为编译器和自动微分系统遍历的是同一个计算图。

由于填充步骤，最终结果的原始计算量略高于循环实现，但部署效率更高。在典型窗口中，乔列斯基求解大约 6× 比在相同中央处理器硬件上通过循环执行的奇异值求解快，速度可从1.2×对于包含 $n = 10$ 个成员的窗口到 $\sim 42\times$在 $n = 100$），并且求解器本身仅占拟合总时间的很小一部分，因此第 3.4节所述的性能主要得益于规模效应和打包效率，而非求解器本身。通过乔列斯基求解计算的导数在近乎退化的窗口中同样稳定，而这恰恰是通过奇异值求解计算导数不可靠的情况，因此乔列斯基方法适用于可微实现。

3.2. 与公开的 TRACTOR 分支进行验证

我们将结果与用于驱动SSDC L3测光的参考实现进行验证。该实现是一个公开分支[8]，属于 THE TRACTOR(D. Lang 等，2016年a)。关于此分支的详情，请参见R. Akeson 等(2026年) 和Z. Huai 等(2026年)。与上游 TRAC-TOR 版本相比，一个值得注意的区别是更新了混叠感知协方差（除 σ_j^{marg} 外，还求解 $\sigma_j^{\mathrm{naive}}$）。我们从三个不同尺度进行比较：对单次参考曝光进行详细比较；针对涵盖不同源密度的20次曝光的集合，依据一组预定义标准进行比较；以及与生产级L3星表的一个子集进行比较（第 3.3节）。

由于 11×11-像素窗口间隔为 5 像素，因此一个源属于 $(11/5)^2 \approx 4.8$

[8] https://www.github.com/Caltech-IPAC/tractor-spherex/

个窗口（平均而言），包括测量该源的窗口。本节中始终使用的参考曝光包含 744，529 个源，分布在 165，076 个窗口中（共 166，464 个窗口），每个窗口平均有 21。7 个成员。该曝光的 ID 为 2026年W12 2A 0530 2D1，拍摄于 2026年3月22日，视场中心位于右赤经 270.53°，赤纬 −20.87°，位于探测器 1（最短波长的探测器阵列）上。这是一次位于银河平面（$l = 8.8°$，$b = +0.9°$）中的高密度曝光，构成了一项有意设置的严格测试（参见附录 A，其中展示了此次曝光的窗口示例）。

TRACTOR 参考实现和 HESPEROS 求解都尝试通过不同的算法，以64位浮点精度求解公式2 中的窗口加权最小二乘问题，前者采用迭代法，后者采用直接法。我们以参考边缘不确定度$\frac{f_{\rm H}-f_{\rm T}}{\sigma^{\rm marg}_{\rm T}}$ 为单位对二者进行比较，其中 $f_{\rm H}$ 和 $f_{\rm T}$分别表示由 HESPEROS 和 TRACTOR分支估计得到的通量。

通量差异的主要原因在于用于连续计算表格化有效点扩散函数的插值函数。SSDC参考实现使用 FITPACK[9] 样条，而 HESPEROS 实现使用三次B样条。图3的上方面板展示了在参考曝光中744,529个源的基准 HESPEROS 配置下，两个实现的流量解之间的相对距离；同时还展示了一种配置：通过使用参考实现的样条对 HESPEROS的有效点扩散函数进行重采样，并匹配插值函数，使 HESPEROS 解保持不变，从而突出线性求解器本身的差异。表1总结了二者的一致性。使用默认插值函数时，超过 $1\,\sigma^{\rm marg}_{\rm T}$ 的测量值占1.4%；当插值函数匹配时，该比例为0.888%。

图3 的底部面板评估了 TRACTOR 参考解和 HESPEROS各自的求解结果，展示了两个求解器的存留比例如何随各自的边缘不确定性变化 $\sigma^{\rm marg}$，并与该代码自身拟合所得到的一个（精确但较慢的）64 位浮点数numpy （C. R. Harris 等，2020年）解进行比较。参考解采用迭代方法，因此在困难、标记严重且混合的窗口中无法完成求解。 HESPEROS则采用直接 Cholesky 解，结果与numpy 的实现非常接近，且没有任何源超出 $4 \times 10^{-3}\sigma$。

不确定性约定的中位数相差比率为 $\sigma^{\rm naive}$ 1.000，为 $\sigma^{\rm marg}$0.988，差异为——

[9] https://perazz.github.io/fitpack/

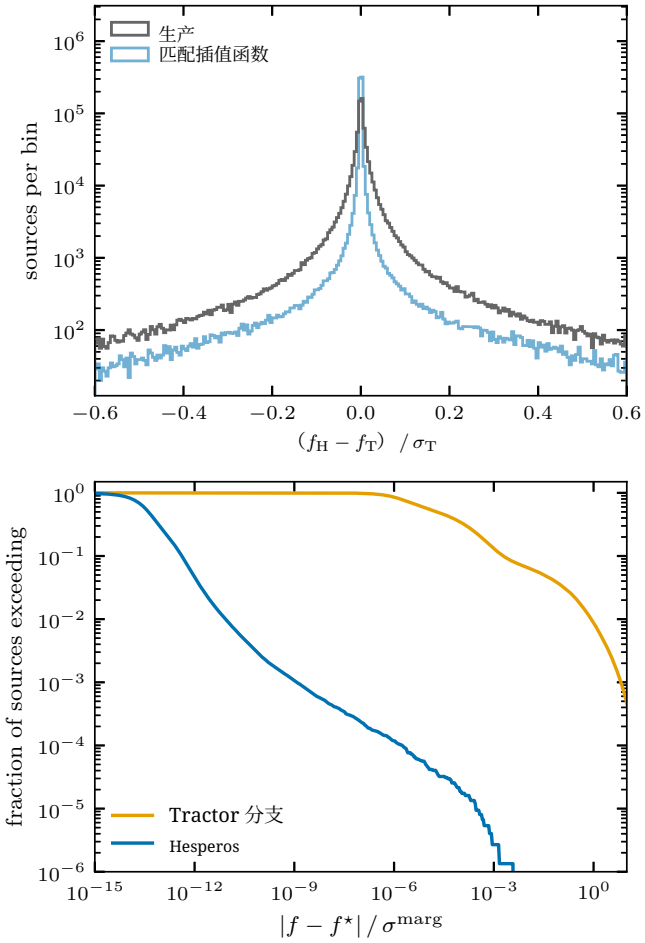


图3。 在参考曝光中，对于匹配的源集合， HESPEROS与参考 TRACTOR 实现之间的通量比较（禁用任一代码的拟合质量标志，以确保比较的是同一组对象且条件一致）。**上图**：参考曝光中 7.4×10^5 个匹配源的两种实现之间距离的直方图，距离以参考边缘不确定度为单位。黑色曲线表示 HESPEROS 生产插值函数；浅蓝色曲线表示与参考实现明确匹配的插值函数，以便单独考察求解器的差异。**下图**：超过与精确 numpy 加权最小二乘解 $f^\star$之间给定距离的源的累积分数，以求解器自身的边缘不确定度为自变量。 HESPEROS 解的分布始终更小，表明其与 numpy 解的一致性显著更好。

后者的差异源于 HESPEROS选择对背景项 b进行边缘化，而1式中的边缘不确定度计算采用了这一做法；参考实现则选择将其固定。20 次曝光检查中有 15 次通过，未通过的 5 次包括 3 次标记高度异常且曝光拥挤、会被分析排除的曝光；20 次中最密集、最拥挤的一次曝光；以及一次单个

表1. HESPEROS 与 TRACTOR分支及生产目录的通量一致性

比较	中位数	第99百分位。	第99.9百分位。
参考曝光			
H –T	0.007	1.4	7.4
H –T，匹配插值函数	6×10^{-4}	0.87	6.8
T – T⋆	2×10^{-5}	0.84	6.7
H – H⋆	4×10^{-14}	9×10^{-12}	1×10^{-9}
R7 生产目录			
H –L3	0.006	0.56	2.5

NOTE—绝对通量差异。H 和 T 是 HESPEROS 和TRACTOR 分支通量，而 H⋆ 和 T⋆ 是各代码自身拟合的精确双精度解，针对第 7.4×10^5 节参考曝光中的源3.2。L3 指 SSDC 生产星表，涵盖 7.2×10^8 相互未标记的测量值（第3.3节）。代码之间的差异以TRACTOR $\sigma^{\rm marg}$为单位，而与该代码自身 $\sigma^{\rm marg}$中的精确解之间的差异则以其为单位。

标记因边界效应而不同。附录A介绍了这项比较、不确定性约定，以及对两种实现之间测量值差异显著的情况所作的进一步考察。

3.3. 与生产目录进行验证

第 3.2 节的比较着重于求解器本身，并将其与单次曝光上的公开代码进行比较。这验证了 HESPEROS的流量解，但并未验证它能否复现生产目录及其弥散背景建模和定稿步骤。因此，我们还将其与内部发布的三级目录进行比较，该目录使用的是 SSDC L2 和 L3 处理软件的第一年（R7.0）版本[10] 。

对于两个星表中均有效且通过拟合质量标志的 7.2×10^8 测量值，以参考值的边缘不确定性 $|\Delta f|/\sigma^{\rm marg}$为单位的流量距离中位数为 0.006，99 百分位数为 0.56σ，99.9 百分位数为 2.5σ。按单次曝光统计，带符号流量距离的中位数绝对值为$7 \times 10^{-5}\sigma$；在至少有 30 个有效测量值的 16,362 次曝光中，该值从未超过 0.18σ 。

输出目录的构成以及 L2 图像级标志的纳入方式完全相同。一个值得注意的不对称之处在于，由拟合步骤本身根据条件数标记的测量值数量，其中 HES-PEROS 标记了约 1.5×10^6（占相互未标记总体测量值的 0.2%）个未被生产目录标记的测量值，而反方向仅有5.5×10^3 。这些情况是由近退化混合引起的；尽管两种实现对于哪些视场条件较差的判断高度一致，但它们在构建星表时采用的阈值严格程度不同。标记比例同样是银纬的强函数，这是由于局部条件和密度存在差异；不过，这也提醒我们，不应过度解读汇总多个不同状态区域的统计数据。

3.4. 吞吐量

速率以长任务列表上的原始实际运行时间给出，采用各平台测得的打包最优点，并包括完整的启动开销[11]。一次曝光由在同一图形处理器设备上运行的多个独立进程测量，因此当一个进程运行其主机端阶段时，图形处理器仍可继续进行计算。达到图形处理器或多个图形处理器饱和状态所需的进程数，均按架构分别测量和优化。此处的基准测试在一组共 3,674 次探测器 1 曝光上运行，这些数据由两个各包含 1837 次曝光且不连续的观测周组成（2.5×10^9 通量测量）。

在点源模式下，单进程处理一次单独曝光，预热后在 NVIDIA GH200 上约耗时 0.180 秒（NVIDIACorporation 2023）。点源测光模式在同一 GH200 上可持续达到 2.97×10^4 次曝光/小时，而在配备四块图形处理器的 GB200（NVIDIA Corporation2024）节点上约为 8.5×10^4 次曝光/小时，不过不同运行之间的方差较大。在扩展模式下，每个源还需要进行一次 S´ersic 卷积，GH200 可持续达到 3.5×10^3次曝-光/小时。这约为 8.5× 点源处理路径的性能，不过具体表现取决于剖面的分布和密度。在同一批数据上，四块图形处理器的 GB200 节点可持续达到 1.24×10^4 次曝光/小时，其扩展模式的扩展性略优于点源处理路径（∼ 3.5× 相较于 ∼ 3× 单 GH200 的性能），这是因为扩展路径的打包更依赖内存，因此能更充分地受益于 Blackwell 架构显卡额外提供的高带宽内存。测试发现，点源模式下六个进程即可使 GH200 达到饱和，扩展模式下则为四个；GB200 显卡则在点源模式下每卡最多可从六个进程中受益，在扩展模式下最多可从八个进程中受益。

[10] https://irsa.ipac.caltech.edu/data/SPHEREx/docs/SPHEREx Expsupp.pdf

[11]每个进程都需承担即时编译带来的高昂初始化成本，因此需要运行更长时间的基准测试，才能了解与生产相关的吞吐量。

我们在相同的中央处理器硬件上运行 SSDC 参考实现和 HESPEROS（双路插槽、128 核的 AMD EPYC 7763 “Milan”，配备 256 GB 系统内存；SSDC 运行生产流水线所用的也是同一硬件），以便将算法差异与批处理/硬件加速的影响区分开来。结果发现，内存是性能瓶颈：参考 SSDC 处理流程在节点满载时每 CPU 节点小时可处理约 122 次曝光，而在同一硬件上运行 HESPEROS时约为每 CPU 节点小时 107 次曝光。我们并未投入大量精力来优化 EPYC 硬件上的中央处理器 HESPEROS 配置；这方面或许还能进一步优化，但关键在于， HESPEROS所带来的性能提升绝大部分源于该问题适合采用图形处理器硬件，而非拟合算法本身。四图形处理器 GB200 系统满载运行时， HESPEROS 的吞吐量相当于计算资源分配单位吞吐量提高了 $\sim 700\times$（$\sim 240\times$ 针对单 GPU GH200 节点）。各测试配置的性能比较见表 2。

为检验外推结果并在大规模条件下测试吞吐量，我们使用 HESPEROS，分别以点源模式和扩展模式运行，对首年“DR1”数据发布的全部数据进行了重新处理；输入数据为经校准的 DR1 2级图像[12], ，结果见表 2底部两行。这批数据略多于 SPHEREx 的首年数据，包含全天巡天的前两次巡天：1，312，429 次曝光，其中 1，116，233 次曝光至少包含一个被指定为扩展源的源。使用 DESI 遗留成像巡天（′）提供的 SA. Dey 等, 2019érsic 指数和类型，对扩展源进行强制测光；没有延展信息的源则作为点源处理。对全部图像和星表进行平均后，性能比时间较短的基准测试略有下降；这主要是由额外的文件加载以及其他受中央处理器或文件系统限制的操作造成的。包含 ~ 1.3M 次曝光的数据集以点源拟合模式处理，耗用 18.9 GB 节点小时（75.6 Blackwell GPU 小时）；而包含至少一个 Legacy巡天扩展源的 ~ 1.1M 次曝光则耗用 100.3 节点小时（401.6 GPU 小时）。使用默认点源模式处理整个巡天的测光，在四 GPU 节点上不到一天即可完成，使用单 GPU GH200 则需数天。处理一小部分观测或特定感兴趣区域只需数小时甚至数分钟。测量方式的变化（例如有效点扩散函数或其他校准因子）对下游数据的影响可以立即评估；可使用自动导数以较低成本探查响应；而注入等依赖重新运行测量的技术，则可大幅加速并扩展规模。

[12] SSDC 扩展源模式具有实验性质，未用于 DR1，因此我们暂缓详细验证，并提供该模式以对比 HESPEROS 点源与扩展源的测光吞吐量。

4. EOSPHOROS：前向模型

EOSPHOROS是一款光谱测光前向建模软件，可将源渲染到探测器帧中。通过与基于完整仪器的 SPHEREx 天空模拟器（B. P. Crill et al. 2025）进行验证后，EOSPHOROS 它代表了一种高度加速且更为近似的方法，但仍纳入了 SPHEREx 独特光谱测光测量中固有的完整空间-光谱 LVF-PSF耦合。

4.1. 光谱测光渲染

飞行前实验室测量（H.Hui et al. 2026）已刻画每个探测器阵列的仪器光谱响应，包括其线性可变滤光片。对于位于探测器位置（x， y）的源，我们从表中获取该像素处的 LVF响应 T（λ；x，y）。带内通量密度随后取带通加权平均值，

$$\bar{f}_\nu = \frac{\int f_\nu(\lambda) T(\lambda; x, y) \lambda^p \mathrm{d}\lambda}{\int T(\lambda; x, y) \lambda^p \mathrm{d}\lambda} \tag{3}$$

其中 $p = -1$ 体现了与表格按光子响应约定相对应的光子计数加权。用于求值该积分的标准求积法则会将数据插值到带通网格上；不过，对于采样比带通更精细的模板光谱（例如包含已分辨发射线的输入光谱模板），这可能会损失约 1% 的带内通量。因此，在这些情况下，我们选择采用累积积分形式，以保留单元平均通量密度。根据待渲染输入光谱的构造方式，这会体现为两种静态固定网格；也就是说，这属于一个不会影响可微性的配置步骤。

渲染步骤如下。首先将源投影到探测器坐标中，并剔除 PSF 图章与窗口区域不重叠的源。根据公式 3 计算带内通量，然后将每个源的 PSF 从视场位置网格插值到画布上，画布的过采样因子为 $18\times$，相对于原生 $6\farcs15$ 像素而言。使用经过缩放的可分离插值核，将图章放置在过采样网格上的亚像素位置处

表 2. HESPEROS 吞吐量与 SSDC 参考实现的比较

Code	每节点硬件	Mode	每节点小时曝光数	与 SSDC 比较
基准测试：3，674 次探测器 1 曝光				
SSDC（THE TRACTOR 分支）	2× EPYC 7763	点	122	1
HESPEROS	2×EPYC 7763	点	107	0.88
HESPEROS	1× GH200	点	2.97×10^4	240
HESPEROS	4× GB200	点	8.5×10^4	700
HESPEROS	1× GH200	扩展	3.5×10^3	···
HESPEROS	4× GB200	扩展	1.24×10^4	···
DR1 生产：全部六个阵列				
HESPEROS	4× GB200	点源	6.9×10^4	···
HESPEROS	4× GB200	扩展的	1.11×10^4	···

NOTE——原始实际运行时间吞吐量，包含启动开销。中央处理器各行对应双路、128 核的 AMD EPYC 7763 节点；图形处理器各行对应 NVIDIA GH200 和四图形处理器 GB200 节点。基准测试各行均采用各平台的最佳内存打包配置。GB200 点源速率在多次重复运行中为 8.1–9.2 $\times 10^4$ 。DR1 行对应在四个 GB200 节点上进行的完整生产运行：1,312,429 次曝光用时 18.9 节点小时（点源）；包含至少一个扩展源的 1,116,233 次曝光用时 100.3 节点小时（扩展源）。SSDC 扩展源模式为实验模式，未运行。

缩放因子为带内通量，并将每个 18×18 块重新求和至原生分辨率，从而精确保留通量。将源沿每个轴的亚像素位置 u 表示为整数亚像素单元部分 $c = \lfloor u \rfloor$ 和小数部分$r = u - c$，PSF $\mathcal{P}$ 的放置方式如下

$$\mathcal{I}(m) = \sum_k w_k(r)\, \mathcal{P}(m - c - k), \tag{4}$$

其中， m 索引像素内的过采样单元，且权重满足 $\sum_k w_k = 1$ 和 $\sum_k k\, w_k = r$，从而精确保留通量和质心。

渲染场景的最后一项成分是黄道光背景，它根据标准行星际尘埃模型的重新实现计算得出（T. Kelsall 等, 1998；SPHEREx 天空模拟器也使用同一模型），并从预计算表中读取。该模型关于天空位置、波长、历元和整体振幅均可微分，但关于底层尘埃模型的参数不可微分，因为此表是固定的。黄道光发射模型按像素逐波长计算，同时将近似视场位置固定；这隐含地假设黄道光辐射的整体强度在单个窗口所覆盖的 1.1 角分范围内不会剧烈变化，同时考虑探测器窗口中 y 维度上可能存在的显著光谱变化。黄道光辐射大部分会被式 b 1 的背景项消除，因此这里将其用于定义每次测量的光子噪声贡献的不确定性

而不是注入一种污染成分。黄道光辐射对 SPHEREx 测量的整体影响十分重要，值得在未来的一项专门研究中单独探讨（H. Hui 等，2026年b，准备中）。

EOSPHOROS 还提供了一种快速的近似直接星表生成模式，可在无需经过完整测量算子的情况下构建响应模板（第6节）。该路径使用基于背景和（由黄道光主导的）光子噪声水平的解析模型计算观测噪声，并采用近似迭代方法快速计算混叠效应。由于强制测光关于源通量是线性的（第5节所依赖的性质），混叠造成的通量污染是一个固定的稀疏串扰算子 K ，其元素为 $K_{ij} = g(r_{ij})$，适用于 $i \neq j$；其中 $g = k \circledast p$ 是匹配滤波核 k 与点扩散函数 PSF p卷积后，在两者探测器像素间距处求值并制成表格的结果。测得通量与真实通量之间的关系为 $\mathbf{f}_{\mathrm{meas}} = (\mathbf{I} + K)\mathbf{f}_{\mathrm{true}}$。去混叠后的通量可由诺伊曼级数得到：$(\mathbf{I} + K)^{-1} = \sum_m (-K)^m$；该级数截断于指定阶数，并收敛于可恢复源集合的精确线性强制测光解。当源 i 从其邻居泄漏的总分数通量小于 1（$|K| < 1$）时，该级数收敛。我们通过要求源的行和 $\sum_j K_{ij} < 0.9$，并单独处理超过该阈值的源来强制满足这一条件。这对于最密集、混合最严重的场景而言可能是一种稳健性较低的近似

方法，我们将其视为这种快速但近似方法的局限性。实际上，对于不受该判据影响的源，在典型场域中，该级数通常经过约十几次迭代即可令人满意地收敛（在高度混合场景中达到真实通量的$\lesssim$ 1%），而性能仅下降约 $\sim$ 5%。尽管这种方法是近似的，但它速度快且可微分，符合直接星表模式的总体目标；该模式可用于快速大规模探究各种效应，或生成合成星表实现。

被指定为扩展源的源，通过在傅里叶空间中将 S´ersic 轮廓与局部 PSF 进行卷积来渲染。要使这一过程具有良好的性能，关键在于利用仪器设计所提供的可分离性，将原本针对每个像素、关于轮廓、PSF、带通以及 LVF 上位置的四维积分拆分为可分离的分量。源的形状是天空的属性，而仪器 PSF 则属于光学系统；另一方面，光谱响应取决于光子在与 PSF 发生作用后落在探测器上的位置。中心位于 $(x_s,\ y_s)$ 的源随后对原生 SPHEREx 探测器像素 $(x,\ y)$ 的贡献为

$$I(x,y) = C(x,y)\,\bar{f}_\nu(x_s,y), \tag{5}$$

其中 C（$x,\ y$）是单位通量的 S´ersic 轮廓与 PSF 卷积后再对像素积分得到的结果，而 f_ν（$x_s,\ y$）是方程 3 中的带内通量密度，通过探测器位置（$x_s,\ y$）处的 LVF 响应进行评估。响应沿探测器 y 方向变化，即 LVF 梯度的方向；而沿探测器x 方向，我们近似认为响应保持在源所在位置不变。这一近似忽略了 LVF 通道横跨探测器x 方向时的轻微曲率或“微笑”效应，以及由光学畸变引起的变化，但在单个测量窗口的范围内，这些影响都很小。

我们在 18×倍过采样网格的傅里叶空间中组装空间因子 C，如下所示

$$\widehat{C}(\boldsymbol{\kappa}) = \widehat{\mathcal{P}}(\boldsymbol{\kappa})\ \widehat{S}_n(r_e\,|\mathsf{A}\boldsymbol{\kappa}|)\ \operatorname{sinc}\!\left(\frac{\kappa_x\delta}{2\pi}\right)\operatorname{sinc}\!\left(\frac{\kappa_y\delta}{2\pi}\right)\ \widehat{W}(\boldsymbol{\kappa};\mathbf{x}_s). \tag{6}$$

其中，$\boldsymbol{\kappa}$ 表示角波矢，$\mathcal{P}$ 表示在源位置处插值得到的 PSF图章的变换，r_e表示半光半径，$S_n(q)$表示单位通量 S´ersic 轮廓的径向零阶汉克尔变换，其索引为n，是无量纲波数 $q = r_e|\mathsf{A}\boldsymbol{\kappa}|$的函数。矩阵 A 表示源的椭率，该椭率已从天空帧旋转到探测器帧。sinc函数对每个边长为 $\delta = 0\farcs34$的过采样子像素单元内的光进行积分，而 $W(\boldsymbol{\kappa};\mathbf{x}_s)$ 是精确的相同子像素沉积权重的变换 w_k，即方程4 中用于点源版本计算的变换。每个源都在以其为中心的方形画布上进行渲染，单次逆快速傅里叶变换（FFT）即可得到该画布上的通量，对画布内的 18×18 过采样单元求和，从而得到 $C(x,\ y)$。S_n 没有闭式解，因此我们在 $0.3 \le n \le 6.2$和 $10^{-4} \le q \le 10^3$ 上构建表格，并使用其上的三次插值来计算 $\widehat{S}_n$。随后，该过程对于n、r_e 椭率、位置和通量均是可微分的，其计算成本基本与 n 无关（计算成本仅由画布大小决定；为了容纳延展性最大的轮廓，画布可能需要更大）。我们将这一设置及这些近似与 GALSIM 在第4.4节中针对各种轮廓生成的渲染结果进行了验证。

4.2. 与 *SPHEREx* 天空模拟器的关系

EOSPHOROS与 SPHEREx 天空模拟器一脉相承，并且在很大程度上共用其仪器描述，但二者的构造方式有所不同，目标也大相径庭。Sky Simulator 纳入了逼真的探测器读出、噪声采样和图像级伪影，以生成 SPHEREx 探测器实际观测到的完整曝光图像；而 EOSPHOROS生成的是理想化的渲染场景，并将其直接交给测量算子，以获得某种输入扰动的响应。尽管如此，二者仍共用相同的表列光谱响应、默认点扩散函数输入、黄道光背景模型和巡天计划。

Sky Simulator 包含一个“`quickcatalog`”直接星表模拟模块，它是 EOSPHOROS 最直接的模拟。两者都省去了图像和图章生成步骤，直接根据 PSF 和噪声模型生成星表通量；但 EOSPHOROS通过采用额外的近似并利用图形处理器的性能，将性能提升到了更高水平。EOSPHOROS 采用第 4.1 节所述的解析确定性噪声模型，而非进行大量蒙特卡洛抽样（默认情况下为 500 次，适用于 Sky Simulator/`quickcatalog`），并使用预先计算的黄道光发射表，同时充分利用仪限中央处理器的 `quickcatalog` 所不具备的图形处理器加速。星表级点源前向模型在通量和波长方面与 Sky Simulator 的相对差异均为 10^{-12}；对于 500 次噪声实现，解析不确定度对 Sky Simulator 抽样不确定度的复现精度达到其抽样误差下限以内。Sky Simulator 不包含自动微分，因此，EOSPHOROS所支持的微分响应模板也是一项新功能。

扩展源渲染结果的匹配精度达到百分比量级，这主要受下文列出的已知常规差异影响；此外，实现的光学特征（扩展 PSF 晕和二向色镜鬼影，见 P. M. Ko-rngut 等，2026年）与 10^{-14} 的相对差异也在该量级。三项有意采用的不同约定分别是：带通通量加权（按每个光子而非单位能量评估）、PSF 图章的居中方式，以及光学畸变模型的符号约定。尤其是后两项约定，是为匹配飞行数据而制定的，而 Sky Simulator 创建时尚无法获取这些数据。整体差异小于绝对定标误差（M. Ashby 等，2026年，即将发表），且与相对定标误差的量级相当。

4.3. 可微性

EOSPHOROS 渲染器关于源通量、源位置以及其实现的系统误差算子均可微。通量的处理很直接，系统误差则取决于具体情况，但尤其是位置，需要特别谨慎。

我们将位置进入渲染的方式分为离散成分和连续成分。离散成分包括场景中包含哪些源、每个源落入哪个窗口、获取哪个已制表的有效点扩散函数位置、读取已制表滤波器映射的哪一行，以及落在哪个过采样单元中；连续成分则是单元内的亚像素偏移。我们固定离散成分，并对连续部分求导。该连续部分本身必须足够平滑，才能进行求导，这也对核的性质提出了限制。在每个过采样单元边界处设置节点的分段线性沉积，其位置导数将呈分段常数，并在跨越节点时发生跳变；而对于真实的有效点扩散函数，原生峰值像素的导数会在跨越此类边界时改变符号。连续可微的核消除了这一特征：在单元内部，其位置导数与二阶精度的中心有限差分相匹配，3.7×10^{-5} 在步长为 10^{-3} 个原生像素时，最大导数的 3.7×10^{-5}，并在步长为 10^{-4}时降至 3.7×10^{-7}。跨越单元边界时，它以一阶精度收敛，而分段线性核在一个数量级的步长范围内完全不收敛。

4.4. 外部验证

图 GALSIM （B. T. P. Rowe 等，2015）展示了在一系列 S´ersic 指数、轴比和位置角下渲染的八个代表性样本。图章峰值单位下的最大偏差用 1.5×10^{-7}到 2.3×10^{-6} 衡量，拟合通量通常仅有几分之一 10^{7} 的差异。关于.....的详细信息

更大数量的扩展剖面所进行的扩展剖面验证详见附录 B。

我们还通过一种恒等关系，按照核本身的定义对其进行验证。将同一个真实的表格化印章应用于两个相同的轮廓表示：一次通过渲染器使用的傅里叶域路径，另一次通过图像域中的显式循环求和。这代表同一操作的两种写法，两种方法所得通量和不确定度的相对差异仅为10^{15} 几个百万分之一。正如预期的那样，只要在其中一种方法中故意错配某项约定，就会使这种一致性大幅降低。

尽管这验证了轮廓、核和相关约定，但并未评估有效点扩散函数模型、仪器光谱响应或相对于真实探测器的定标。我们将在第7节重新讨论这一问题。

4.5. 吞吐量

在单个 GH200 上渲染全天区目录，EOSPHOROS 在宽巡天模式下可达到 1.01×10^{4} 个源每秒，使用的输入光谱取自R. M. Feder 等人（2024）。在图形处理器上处理一次观测平均需要几十分之一微秒，峰值设备内存占用为 18。5 GB。结果表明，在两个数量级的目录规模范围内，该性能与目录规模呈线性关系；处理 10^{8}个源的单次运行约需 2。7 个 GH200 节点小时。按照第6节所述，端到端运行完整的可微测量流程，包括渲染、重分箱、红移估计、层析分配以及雅可比-向量积，对 10^{8} 个源耗时 5。78 个节点小时，因此完整处理流程的耗时约为 $2.1\times$ 仅运行目录前向模式的耗时。在相同的中央处理器硬件上，测得 Sky Simulator`quickcatalog`在默认配置下每次观测的渲染耗时为 88。6 毫秒，而78 μ秒对应 EOSPHOROS，速度相差 $\sim5\times10^{4}$倍。需要指出的是，其中很大一部分耗时来自 Sky Simulator 执行的蒙特卡洛噪声抽样（默认配置下为 500 次，这是我们比较时采用的设置）；该参数可以调整。而 EOSPHOROS使用的是解析噪声。这些测量结果列于表3。

前向模型与测量（第 3.4 节）所需的计算成本处于相近水平，因此很容易将两者结合起来。在这一尺度上对测量进行前向建模，尤其是对系统误差的微分响应进行建模，也改变了系统误差模板的定义：我们不必再针对可能与密度场耦合的效应代理量进行回归，而是可以直接测量其响应；唯一的限制是系统误差如何与测量本身耦合的模型。对于那些对

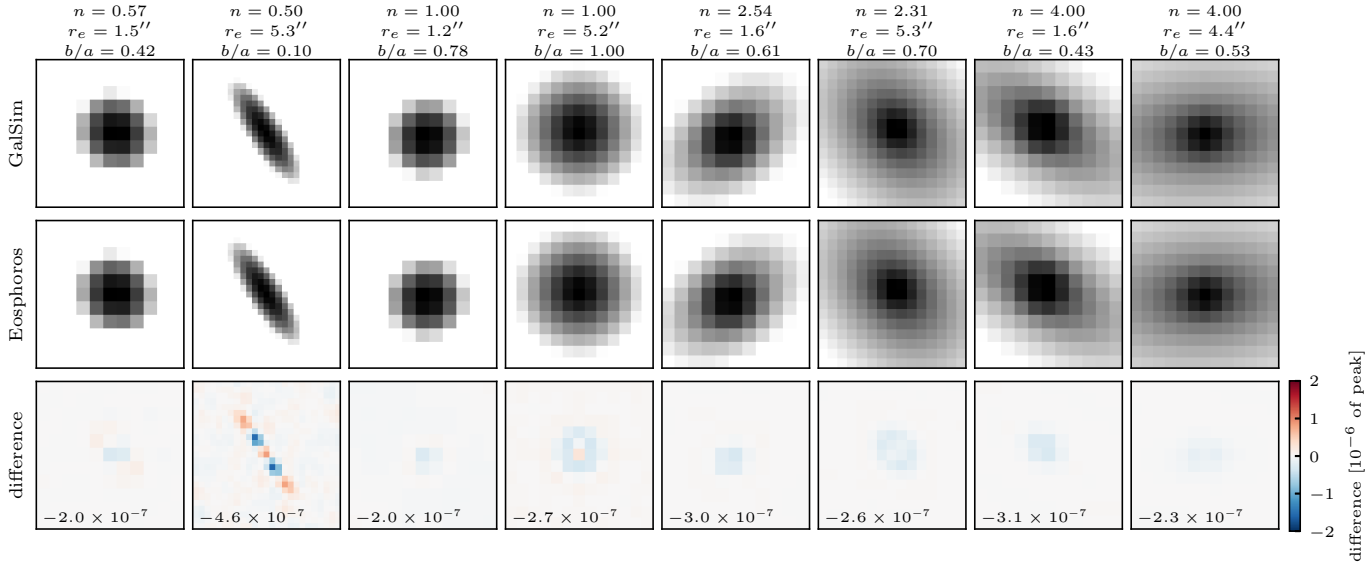


图4。 EOSPHOROS 附录 B 中八个源在一系列 S´ersic 指数和尺寸下的扩展渲染图。前两行展示了各代码在原生 SPHEREx 图像像素上的渲染结果，均以峰值归一化，并采用统一的对数色标。底行展示了以 10^{-6} 峰值的 GALSIM 为单位的差异，其中插图显示了通过相同的轮廓加权拟合从两幅图像中测得的分数通量差 $\delta f/f$ 。

表 3。 EOSPHOROS 星表吞吐量与 Sky Simulator 的比较

Code	硬件	每次观测所需时间	与 quickcatalog 的比较	10^8 源（小时）
星表前向模型				
quickcatalog，500次噪声抽样	16核CPU	88.6毫秒	1	1.2×10^6
quickcatalog，100次噪声抽样	16核CPU	19.8 毫秒	4.5	2.7×10^5
EOSPHOROS	16核CPU	1.78 μ秒	5.0×10^4	24
EOSPHOROS	1× GH200	0.20 μ秒	4.4×10^5	2.7
完整的可微测量流水线				
EOSPHOROS	1× GH200	...	...	5.8

NOTE—宽视场源，在整个巡天中每个源约有 490 次观测。中央处理器各行使用的是同一台搭载 16 核 Apple M4 Max 的笔记本电脑中央处理器（苹果公司 2024）。每源时间根据其单次观测速率外推，该速率与噪声抽样次数呈线性关系。完整的可微测量流水线见第6节，由源渲染、流量分箱、红移估计、层析分配和雅可比-向量积构成。

测量的影响易于建模的系统误差（例如，由黄道光辐射决定的观测深度），这在很大程度上缓解了系统模板去投影清理方法可能产生的偏差；这类偏差源于物理系统误差与密度场的耦合和回归量代理之间存在去相关。第 6 节通过将其与红移估计和样本选择相结合，对这一方面作了进一步阐述。

5. 构建模型与测量

前向模型与测量算子通过即时编译融合，将 EOSPHOROS 场景渲染组合到HESPEROS测量中，并对结果求导。本节介绍实现这一过程所需的条件、导数的含义，以及其有效的适用条件。

5.1. 测量几何

包含待测光的 SPHEREx参考目录源的二级图像源目录，或注入目录，会因强制测光而直接提供给测量引擎 HESPEROS 。我们使用这些相同的位置和 L2 提供的相同切平面畸变参数，借助 EOSPHOROS渲染这些源，确保渲染的源与其测量位置重合。回退模式使用飞行前制表式巡天计划，而非实际飞行指向和位置。整体偏移十八分之一个原生像素（0.34″）会导致通量误差达到百分之几十分之一，而整体偏移一个像素，则会使强制测光源的测得通量平均降低约 15%。经验证， HESPEROS 和 EOSPHOROS 的测量几何会产生完全相同的结果；世界坐标系/光学畸变、亚像素定位和位置导数的验证精度均达到或优于 1×10^{-8}像素。

5.2. 可微显著性阈值

我们将测量显著的概率定义为一项评价指标，也就是说，其拟合通量 f超过阈值 $\nu\sigma^{\rm marg}$ 的概率，

$$P_{\rm sig} = \Phi\left(\frac{\hat{f}}{\sigma^{\rm marg}} - \nu\right), \tag{7}$$

其中 Φ 为标准正态累积分布函数， f 为通量测量的无噪声期望。这一假设认为通量误差服从高斯分布；若存在未标记的图像伪影，这种近似可能不太适用，但对于此处的图像级演示已足够。自然引入的平滑尺度 $\sigma^{\rm marg}$ 恰好使原本不连续的阈值化问题变得可微分。我们采用 $\nu = 5$，并使用完整的边缘不确定度来计算混合源的适当显著性。请注意，这是针对每次曝光、每次测量的评价指标，并不一定反映某个源是否“被探测到”；后者取决于该源在所有采样波段中的全部测量值是否精度足够，从而能够纳入红移样本。第7 节报告的量相应为 $\mathrm{d}\ln N_{\rm sig}/\mathrm{d}\alpha$，反映的是图像级光谱测光，而非宇宙学样本的响应。

5.3. 微分

系统振幅 α 可能通过其通量及其不确定性在$P_{\rm sig}$ 中引发响应，而它进入其中的路径（或两者）对于自动微分具有机械上的相关性。一个 α 会改变场景通量，但不改变协方差（像素权重）或拟合所用的有效点扩散函数模型，从而改变输入数据 d。整个导数随后会通过 f流动，我们将其称为“通量路径”。如果 d 随 α线性变化，例如对于固定的零点偏移或加性背景，响应也呈线性，并且在任意振幅下导数都是精确的。如果 d 反之以非线性方式依赖于 α，那么响应也会继承相应的阶数；对于探测器位置的扰动，响应是渲染轮廓与拟合所假定轮廓之间重叠的导数，该响应在零偏移处达到最大值，并且对于连续轮廓在一阶时为零，通量损失为二阶（不过在像素网格上，由亚像素相位决定的残余一阶项仍可能出现，见第 7 节）。一个 α 影响噪声，而不是直接改变图像的通量，例如黄道背景水平，它通过逆方差权重 W 起作用，因而也通过 N 起作用；我们将其称为“方差路径”。影响通量的系统误差，例如银河系消光，相应地会通过“通量路径”直接作用于 d。逆方差权重的变化会使通量改变为 $\delta\hat{\boldsymbol{\theta}} = \mathrm{N}^{-1}\mathrm{A}^{\mathrm{T}}\,\delta\mathrm{W}\,(\mathrm{d} - \mathrm{A}\hat{\boldsymbol{\theta}})$，也就是说，与数据和最佳拟合模型之差成正比（其中像素在最佳拟合解上存在不一致的窗口，最容易因这些像素相对权重的变化而使该解受到扰动）。权重 $\mathrm{W} \to (1+\epsilon)\mathrm{W}$ 的均匀变化不会引起通量变化，但会将每个边缘不确定性按 $\sigma^{\rm marg} \to \sigma^{\rm marg}/\sqrt{1+\epsilon}$ 缩放，使得

$$P_{\rm sig} = \Phi\left(s\sqrt{1+\epsilon} - \nu\right), \qquad s \equiv \hat{f}/\sigma^{\rm marg} \text{ at } \alpha = 0, \tag{8}$$

与 s 源测得的显著性相符。权重发生非均匀变化（例如窗口内黄道光背景变化）可能会改变通量；无论如何，自动导数都会捕捉到这一响应。然而，只有当 s的偏移相对于阈值附近 Φ 的宽度较小时，这个导数才能充分描述该响应；如果每个像素的方差都按同一因子变化，则需要 $\epsilon \ll 2/\nu$。

位置扰动会通过通量路径起作用，因为在强制测光中，拟合所用的源轮廓固定在星表位置，只有渲染图像会受到影响。不过，SPHEREx 的这一特性并非纯粹由几何因素决定：线性可变滤光片沿探测器-y 方向的波长响应，将带通与探测器位置联系起来。沿 LVF 的色散轴发生位移，便会使其

通量相对于拟合轮廓发生移动，并影响该通量发射到场景中的波长。天体测量响应与光谱响应通过 LVF 耦合，因此在计算整体响应及其导数时都必须将二者纳入考虑。对于连续轮廓的匹配模型，这一光谱项是唯一在一阶上有贡献的项，而几何通量损失属于二阶项。然而，在真实像素网格上，一阶几何项仍然存在，并且在六个阵列中的五个阵列里是响应的主导项（第7.2节）。光谱项是

$$\left.\frac{\mathrm{d}\ln\hat{f}}{\mathrm{d}y}\right|_{\mathrm{spec}} = \frac{\mathrm{d}\ln\bar{f}}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}y} \tag{9}$$

沿着 LVF梯度方向移动 y。正如右侧第一项所示，被测光谱的陡峭程度也会影响这一光谱贡献对梯度振幅的影响，第二项所表示的 LVF响应也有同样的作用。位置依赖性与波长依赖性的耦合对无缝光谱学也很重要，通常被视为定标项（例如 Euclid Collaboration 等，2026年）。

以单个标量系统误差振幅为输入、以逐次测量概率 P_{sig} 为输出时，我们需要计算高维输出相对于低维输入的方向导数，因此前向模式微分是自然之选。主要成本在于一次性的初始步骤，即构建初始线性映射；此后便可低成本地将其应用于多个方向。计算单个导数方向的成本约为 $2.4\times$ 一次普通渲染，此外还需 $\sim 6\times$来构建一次线性映射。总体而言，对于一个源集合，单个方向会使运行时间比渲染和测量的一次典型前向遍历增加约 29%，之后每增加一个方向，运行时间还会增加 $\sim 8\%$。

请注意，计算出的导数仅在配置中被冻结的各项所构成的范围内有效，例如存在哪些源、每个源以哪个窗口为中心、各源所属的存储桶和数据块、分配的 ePSF 区域（定义单个有效点扩散函数模型的探测器子区域）等。实际操作中，窗口归属和源列表等因素由星表固定，而存储桶/数据块的分配只会影响浮点数排序；基于渲染位置的表列量（如 ePSF 区域）会在一致偏移 δ 个区域间距 $\approx \delta/L$ 的源时发生变化，其中 L 为区域间距，≈ 100 原生像素。对于 0.1 个探测器像素的一致偏移，大约有 0.1% 的源会进入不同区域。足以使某个源在这些状态之一之间发生变化的扰动，不会被自动导数捕获，在我们的实现中会引发误差。

5.4. 验证

由于我们的拟合对图像呈线性，因此可以针对这一组合构造精确测试：将源i自身的单位通量模型添加到图像中，应使源 i 的拟合通量恰好增加一个单位，同时使窗口中所有其他源的拟合通量恰好增加零。沿该方向渲染源 i 并使用 EOSPHOROS 对拟合HESPEROS 求导后，对于孤立源、共享同一窗口的混合源对，以及位于不同重叠窗口中的一对源，此测试通过时的相对误差均达到 10^{-9} 。我们也以类似方式测试反向模式梯度；相对于输入函数形式的拟合像素梯度，在包含混合窗口和背景的情况下，匹配误差达到 10^{8}, 分之一。这验证了算子沿所需方向的运作机制和内容；此外，我们还通过中心有限差分（FD）方法验证前述性质。尤其对于通量路径，由于拟合通量对场景振幅的响应呈线性，因此可以证明 FD 结果精确且不存在步长误差。对于这一情形（以及无法建立此类恒等关系的方差路径），AD 与 FD 导数的差异小于 10^{9}分之一。

公式 9 中的光谱项也通过恒等式检查和经验中心有限差分检查进行了验证。对于具有幂律光谱能量分布的源，$f \sim \lambda^{\beta}$，$\mathrm{d}\ln f/\mathrm{d}y = \beta \cdot \mathrm{d}\ln \lambda_0/\mathrm{d}y$，我们可以立即将其与公式 9联系起来。我们使用了三个玩具幂律输入光谱，其 β 分别为 -1.2、0。4 和 -2.0。对于所有这些情况，测得的导数与恒等式的吻合度为 0.6%；剩余差异可归因于在有限网格上对光谱能量分布进行插值（该恒等式仅在连续积分时才精确成立）。有限差分验证所得结果与自动微分雅可比向量积的相对误差为 10^{9} ，步长为 10^{-3} 个原生像素。有限差分验证的只是连接方式，而非物理过程，因此两种验证都不可或缺。

这些结果验证了导数能正确地通过算子传播，但并未验证算子本身；这正是第 3.2 节和第 4.4 节讨论的主题。

6. 从通量到三维密度场

测光只是测量流水线中的一个步骤。从宇宙学家的角度来看，最终成果并非通量目录，而是对三维星系密度场的表征。

因此，我们关注的是扰动经过整个处理流程后对该密度场产生的响应。本节将介绍其余的测量步骤。

我们首先指出，通往下游密度场有两种可能的路径：一是将渲染器与第5节所述的完整测量算子组合起来；二是采用第 EOSPHOROS节所述的快速、近似、直接星表（仅 4.1）模型。前一种方法基于实际飞行帧，在逐像素层面运行，使用真实测得的背景和方差层；后一种方法则对巡天计划／指向和测量噪声均采用快速近似。后一种方法易于扩展至高密度，并能用于预测整个巡天的响应（包括尚未进行的观测），因此适合生成本节所述的响应模板。针对本节介绍的深度系统误差，我们在具有相同场景和相同曝光的情况下比较了这两种方法；此时拟合是精确的，因此通量得以一致地恢复，从而可单独考察解析噪声模型。在 125 次曝光和全部六个探测器阵列中，共有 422,492 次测量。直接星表方法的单次测量不确定度与图像级不确定度之比的中位数为 0.997，各阵列的中位数相差均在 1% 以内，第 16 至第 84 百分位数范围为 0.71–1.15。深度响应是通量和该不确定度的闭式函数，因此两者一致也就顺理成章。这正是我们期望的表现：第 5 节所述的完整组合，只有在刻画具有细致逐像素表现的效应，或探查解析模型失效的情形时才必不可少；第 8.3节将进一步讨论这一点。本节所述的响应采用快速的直接星表模式。

6.1. 基准测光通道

通量测量落在 LVF 上的位置是连续的，因此对应的波长也呈连续分布，这类测量称为“原生”通量测量。处理流程的初始步骤之一是对这些测量加权并分箱至 102 个“基准”通道，并在这些通道上进行固定网格红移拟合。对于某些应用，会使用数量更多的基准分箱，以获得更高的光谱分辨率，例如 306 个分箱。

每次测量都会分配给中心波长最近的基准滤光片，并根据观测波长及该滤光片归一化至峰值的透射曲线 $T_{\rm norm}$赋予相应权重。通量采用标准的逆方差加权 $T^2_{\rm norm}/\sigma^2$进行合并。未观测到的通道被赋予零权重，以补偿光谱覆盖不均匀的问题。基准分箱分配本身是一种静态几何特征，导数仅通过加权合并传播。

通道分箱的重新实现已通过验证：相对于现有的 L4 参考实现，其通量和不确定度的相对容差为 10^{-10}，且保留的测量列表（被丢弃、标记或设为 NaN 的测量模式）完全一致。

6.2. 红移估计

红移估计采用标准的模板拟合方法，这里仅作必要的概述。将分箱后的测光测量光谱与一组经过不同红移平移的光谱模板进行比较。对于源 i 和红移至网格节点 z 的模板 m，通过解析方式求得振幅轮廓，得到

$$\chi^2_{im} = C_i - B^2_{im}/D_{im} \tag{10}$$

采用 $C_i = \sum_f F_i^{f\ 2} w_i^f$、 $B_{im} = \sum_f F_i^f f_m^f w_i^f$和 $D_{im} = \sum_f (f_m^f)^2 w_i^f$在基准通道 f上使用权重 $w = 1/\sigma^2$。红移概率密度随后为 $P(z) \propto \sum_m P_{\rm tmpl}(m) \exp(-\chi^2_{im}/2)$，其中 $P_{\rm tmpl}$是光谱模板 m 的先验权重（对于 SPHEREx，这些先验取均匀分布）。 B 和 D 均为批量矩阵乘积，因此对星表进行拟合只需进行两次大型收缩运算，而非循环；基于与第 3.4 节所述类似的原因，这使得在图形处理器上的运算速度显著提升。

这种模板拟合估计器沿袭了代码 LEPHARE（S. Arnouts 等，1999；O. Ilbert 等，2006），SPHEREx 的实现则源自一种变体，该变体最初见于 N. R. Stickley 等 （2016年），并在R. M. Feder 等 （2024）中更新。生产用 SPHEREx 红移流水线、其开发状态及性能将发表于未来的一篇论文（D. Masters、Z. Huai 等，2026年，准备中）。作为光谱模板基底，我们使用 R. M. Feder 等 （2024）的库。这些光谱与第 4.1 节中用于渲染这些源的光谱相同，从而确保红移拟合具有自洽性[13]。

我们直接将红移拟合实现与 4 级生产 C++ 实现进行验证，两者在 1.5×10^{-5} 相对

[13]对于最终分析中将采用的实际宇宙学测量和模板，我们计划使用内部构建的经验星系SED模板；这些模板本身将成为即将发表的研究工作的主题（Z. Huai 等，2026年b，准备中）。

红移后验 $P(z)$最大偏差范围内一致，精度不超过参考实现自身的（32位浮点数）精度。

6.3. 层析分配

一个可能不可微分的步骤是红移估计以及向层析分箱的分配。红移估计通常取最大后验值；在当前实现中，则取红移概率密度函数（PDF）的期望值。即使期望值计算本身可能是可微分的，实际应用所需的分箱分配仍可能妨碍实际可微性。我们改为根据源的后验分布落入各分箱的质量，将通量测量向量为$\mathbf{F}_i$ 的源 i 分配至层析红移分箱 b ，

$$s_{ib} = \frac{\int_{z_b}^{z_{b+1}} P(z|\mathbf{F}_i)\mathrm{d}z}{\int P(z|\mathbf{F}_i)\mathrm{d}z}. \tag{11}$$

分箱边界仅通过静态掩膜参与计算，因此导数会沿归一化后的后验分布传播，分配过程仍然可微分。落在所考虑分箱之外的后验质量不参与选择，因此一个源在各分箱中的归属概率之和至多为 1，而不会重新归一化到红移网格上。

6.4. 密度场响应

每次源测量具有一定的显著概率 P_{sig}（公式 7），并且根据公式 11，将跨越一个光谱的测量集合分配至层析分箱。在此阶段应用基于红移拟合的最终选择，并将其视为静态处理（对于当前分析而言），随后将每个体素（像素-红移分箱组合）中的总体转化为分数密度调制 $\Delta n/\bar{n}$，其形式与角聚类系统误差模板分析所采用的形式相同。对于 SPHEREx，这些模板还包含测量经测量流水线处理后的完整光谱测光响应，直接将径向效应纳入响应图。

从像素数据到三维密度图的每个阶段，都实现为对形状固定的矩阵进行矩阵运算的组合。借助 JAX，我们可以通过处理流程计算一次方向导数，从而精确获得处理流程对该扰动的响应，无需进行蒙特卡洛采样，而计算成本仅略高于单次前向计算。图 5展示了巡天观测深度的此类三维响应测量，该深度通过参数 A 表示。 A 是一个无量纲振幅，用于调节每个通道中每个源的合并方差，其取值连续分布在总体均值 $\bar{v}_c$ 与其实际实现值之间，

$$v_{ic}(A) = \bar{v}_c + A\,(v_{ic}^{\mathrm{real}} - \bar{v}_c), \tag{12}$$

其中， v_{ic}^{real}是源 i在通道 c中的实际观测所得到的合并方差。 $A = 1$ 对应于实际巡天，而 $A = 0$ 则为每个通道赋予其天空平均方差。因此，模板 $\partial(\Delta n/\bar{n})/\partial A$ 在 $A = 1$ 处表示密度场对实际深度变化的线性响应。

该模板采用第4.1 节的解析噪声模型和目录直连模式构建，该模式将仪器与黄道光辐射平均水平的预测相耦合（该水平决定光子噪声预算），并与巡天扫描策略进行卷积；此处的巡天扫描策略通过飞行前制表代理巡天计划表示。这些图直接测量的是下游场对系统误差$\partial(\Delta n/\bar{n})/\partial A$ 经处理流程后的响应，而不是其可能与之去相关的代理量（其有效性取决于输入系统误差模型的可信度，以及标定产品/巡天计划的保真度）。例如，由于该模板以单次观测为基础构建，因此可以直接、完整地纳入黄道光辐射显著的时间变化。图6 展示了由单位深度变化产生的红移分箱迁移矩阵，说明空间上不均匀的红移恢复所引起的非对角耦合。通过完整的操作链（包括红移拟合和层析分配）计算得到的导数，在简化配置下与中心有限差分进行了验证，其 6×10^{-4} 相对误差在有限差分步长为0.01时为 A；该残差随步长的平方减小，符合预期。

7. 巡天尺度选择响应

本节展示了如何在巡天尺度上，利用真实帧元数据在合成场景中使用完整 EOSPHO-ROS+HESPEROS组合，以及可能测量到的效应类型。我们开展了一组测试观测：使用从竣工光学设计进行光线追迹得到的飞行前 PSF模型（P. M. Korngut 等，2026年）渲染源，但使用从飞行数据中的参考天体估计得到的真实有效点扩散函数进行测量。在对照组中，同一个基于飞行数据的有效点扩散函数既用于渲染，也用于求解。这些点扩散函数模型及其差异如图7所示。请注意，本文使用的有效点扩散函数模型本身仍处于初步阶段，未来可能发生变化；这里并非要说明 SPHEREx 实际的光学响应或 PSF模型，而是说明该处理流程对错误设定的模型（本例中已知为错误设定）的敏感性。

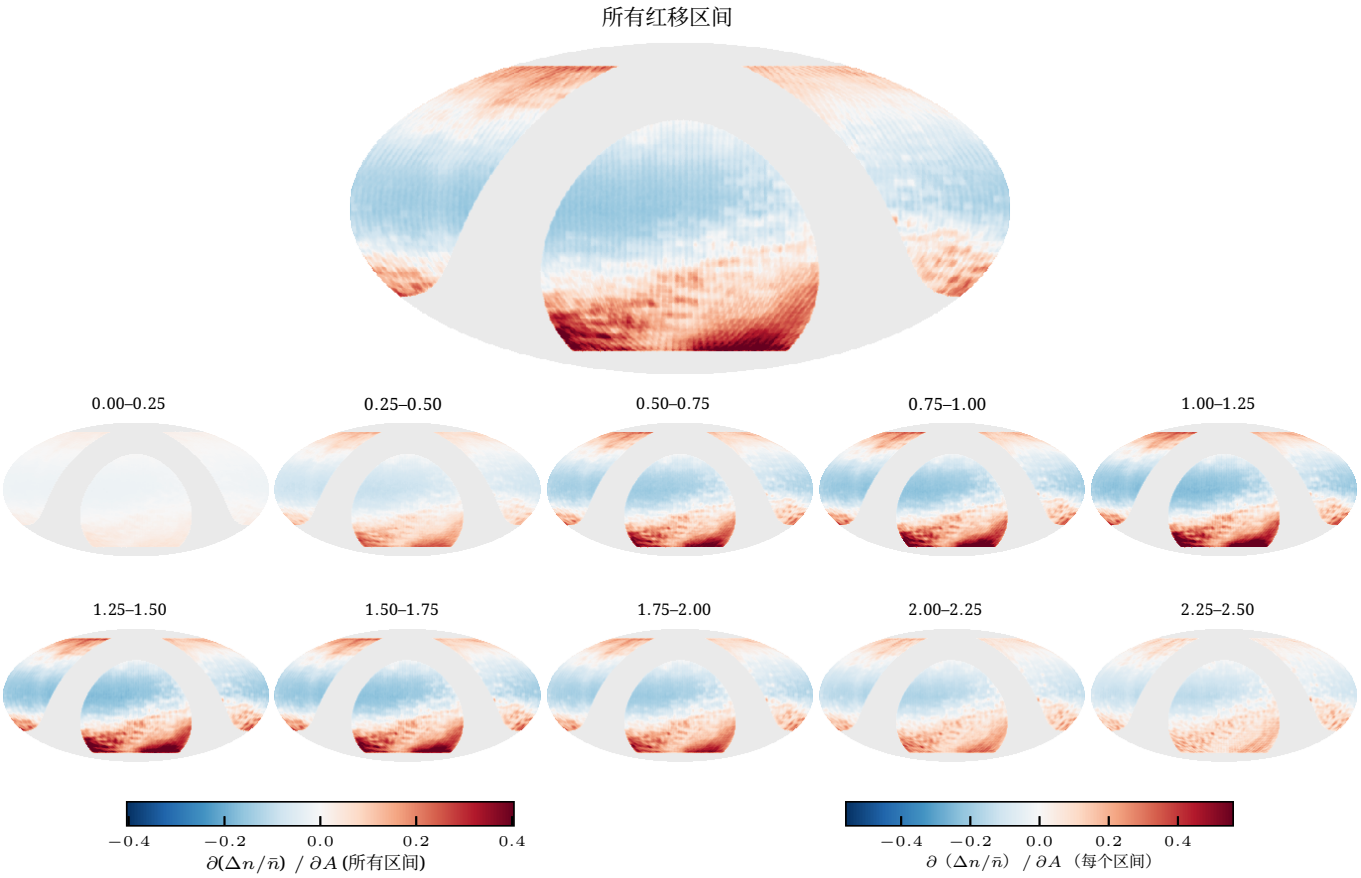


图5。 巡天深度系统误差发生单位变化时的密度场响应 $\partial(\Delta n/\bar{n})/\partial A$：基于模拟的两年巡天计划，采用 EOSPHOROS 的直接星表模式，对 SPHEREx 宽场巡天进行前向建模。 $A = 0$ 表示均匀覆盖， $A = 1$ 表示式 12所述的实际深度；模板显示的是在实际深度处计算的导数值，即完整组合分析流水线的输出产品。顶行显示径向分箱的汇总结果，并为便于可视化而以 1° 进行平滑；中间行和底行则显示从 0 到 2。5 的各个 $\Delta z = 0.25$ 分箱，并以 2°进行平滑。两者在整个天空上的均值均为零。地图采用黄道投影，并施加 $|b| < 20^\circ$银纬截断；此外，还对黄道极处黄道纬度高于 $|\beta| > 70^\circ$的 SPHEREx 深度巡天“深场”施加截断，以突出宇宙学分析的主要区域并避免其主导色标。源均匀分布于整个天空；图中的结构是巡天观测命中模式与特定于波长和视场上探测器位置的光学效应，与黄道光发射模型卷积后的结果；该模型决定了单次曝光中的光子噪声底限。梯度超出线性范围的体素中有 0.1% 被剔除。请注意，这表示源在共动体积中均匀分布时的影响，因此未必能代表实际 SPHEREx 样本选择的响应。此模板基于 10^8 个源总数，或 $\sim 6 \times 10^{10}$ 个测量值总数构建，通过快速目录直接（EOSPHOROS仅）模式在 GH200 上运行，耗时约 6 个节点小时。

7.1. 设置

合成源被渲染到对应于 $20\times$ ~ 3.5天“规划时段”的帧中，这些时段均匀分布在数据的最初 ~ 15 个月内；六个探测器总计包含 216,437 次曝光。这代表该时间段完整数据集的 16.6%。对于每幅图像，源均在真实的参考星表位置进行渲染，使用生产拟合进行测量，并通过 $P_{\rm sig}$ 计算导数。每次曝光包含 $\sim$ 2,500 个源；两个 PSF模型合计超过 $\sim 1.1 \times 10^9$ 个源，使用飞行前和实际飞行 PSF模型分别耗费 7.40 和 22.3 4×GB200 节点小时完成，这种差异源于 ePSF 在渲染步骤中所需的更大支持区域以及因此需要的更多像素。

系统振幅 α 是以探测器像素为单位的相干天体测量偏移；在拟合中固定源轮廓位于星表位置的情况下，将其应用于场景，具体如第5.3节所述。每次运行都会计算三组不同的梯度：在探测器x中相干地移动帧中的每个源；沿探测器y（LVF梯度方向）移动；以及沿一组随机方向移动。对于源自身位移所产生的局部线性响应，随机方向的

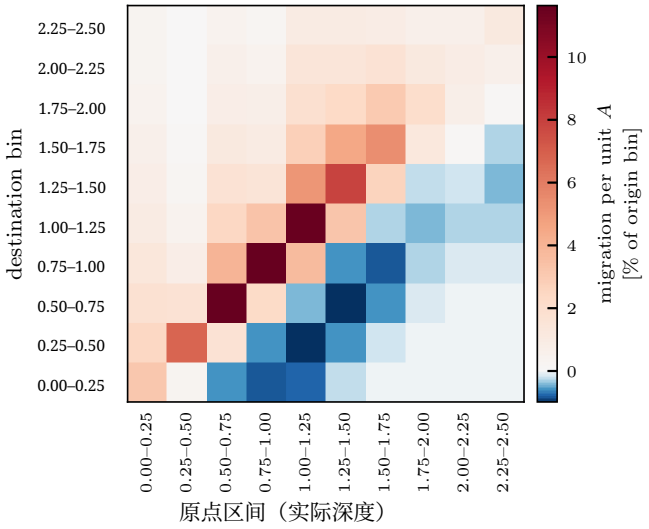


图 6. 红移分箱迁移矩阵$\partial(\Delta n/\bar{n})/\partial A$ 模板，显示每单位 A下各红移分箱的分数迁移，并以原点区间的计数归一化。明亮的对角线表示某个分箱自身被选中的源数量变化（$P_{\rm sig}$，以及其由公式11得到的平滑后验分布峰值），而非对角线则表示峰值分箱之外的后验质量变化，这种调制会将其从较低红移移向较高红移。请注意，源可能完全离开样本，因此图中表示的光度测量源总数不一定保持不变。此图汇总了物理性质不同的噪声机制，仅用于说明深度变化通过红移拟合引起的非对角耦合。

梯度应在足够大的源样本集合上平均为零，从而可作为零假设检验。

为源分配光谱，并对其重新采样，使其均匀分布于各红移区间，从而确保散粒噪声均匀。我们校准源的振幅，使其刻意分布在信噪比阈值两侧，以更好地了解这一更值得关注的区间内的行为。由于渲染得到的点扩散函数轮廓与拟合得到的点扩散函数轮廓不同，我们无法精确恢复输入的信噪比分布。两者在信噪比上的差异是加性的；对于飞行前渲染而言，这体现了本研究所探讨的模型设定错误。因此，我们报告每个源达到的信噪比，以及响应随该信噪比的变化。

经过所有选择后，一个分箱中的计数就是显著测量值的期望数量 $N_{\rm sig} = \sum_i P_{{\rm sig},\ i}$，因此我们报告

$$\frac{\mathrm{d}\ln N_{\rm sig}}{\mathrm{d}\alpha} = \frac{\sum_i \mathrm{d}P_{{\rm sig},i}/\mathrm{d}\alpha}{\sum_i P_{{\rm sig},i}}, \tag{13}$$

单位为原生探测器像素的倒数。这是总和之比，而非逐源比率的平均值，

对于聚集在显著性阈值附近的群体来说，这是一个重要的区别。

曝光内相干扰动的响应会因其帧和混合配置而相关，因此我们改为跨曝光估计不确定性，并针对每次曝光独立应用公式13 。这里的信噪比分布是实验设计的一部分，而非对任何现实人群的建模；公式13的归一化也继承了这一点。

7.2. 响应

图 8显示了各探测器阵列的飞行前渲染响应随达到的信噪比的变化。达到的信噪比相差 10 倍时，响应下降了四个数量级。接近阈值的低显著性源最有可能因给定扰动而越过阈值，因此这一结果符合预期。

图 9 展示了，表 4 列出了失配PSF和匹配PSF运行中沿LVF梯度方向的相干响应。当使用飞行前PSF渲染源时，沿LVF光谱维度的相干偏移会使显著测量值的预期数量每个原生像素变化 9–15%。这一变化在短波红外（SWIR）阵列上为负，在中波红外（MWIR）阵列上为正。使用拟合中所用的有效点扩散函数进行渲染，可使所有阵列上的响应降低至少 26 倍，降至每像素 0.4% 或更低。表 4 所示的不确定度是对每个阵列 $\sim 36,000$ 次曝光计算的标准误差；每次曝光的离散程度与平均响应相当，因此显著性反映的是样本对该平均值的分辨能力。使用飞行前PSF时，D1、D5 和 D6 上的探测器x 响应为 $+5.7$ 至 $+7.1 \times 10^{-2}$ 像素$^{-1}$，远高于对照组；由于LVF模型在这一范围内本身没有 x依赖性，这表明这三个阵列存在PSF形状失配。

探测器-x 响应在匹配情况下源于像素采样。第5.3节所述的一阶抵消对于连续轮廓严格成立，但有限像素化会使其失效；在离散像素网格上，和归一化剖面在像素中心附近更加尖锐，因此重叠的一阶导数不会消失。该项在每个阵列内均不随波长变化，在全部六个阵列上均为负值，平均后仅形成较小的净偏移。我们无法预测其总体幅度。探测器3飞行前响应中额外的散布，是由一个已知的光学特征造成的；该特征导致响应随位置快速且非单调地变化，

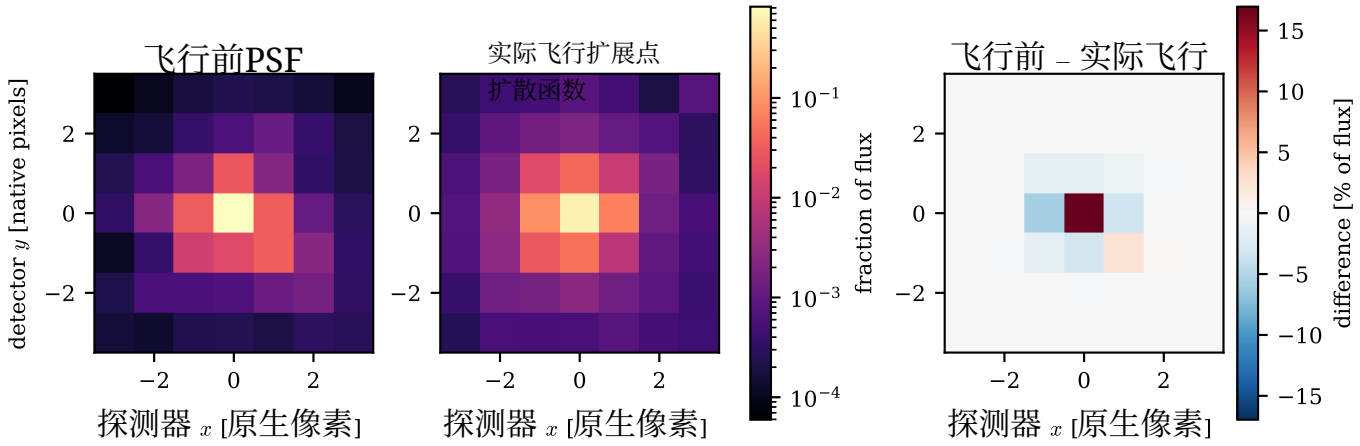


图7. 飞行前（左）与实际飞行中实测有效点扩散函数（中）模型的比较。模型均归一化为总通量为 1，并以对数色标显示，右侧面板展示二者之差。一般而言，PSF 高度依赖于焦平面位置；所示示例位于探测器 1 的像素 (877, 877) 处，该位置恰好对应一个表列的飞行前 PSF 值（而非在表列值之间插值得到的位置）。

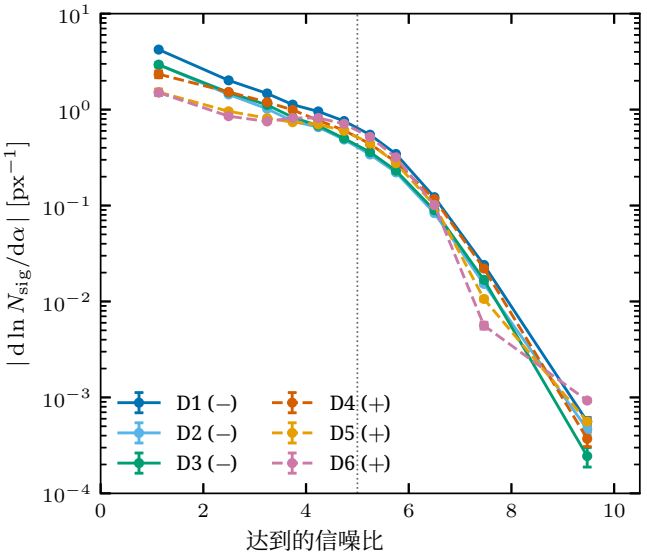


图8. 对于每个探测器，使用飞行前 PSF 渲染时，显著测量数量对沿探测器-y （线性可变滤光片梯度方向）的扰动的响应随达到的信噪比的变化。图例标出了响应的平均符号。虚线表示标称阈值 $\nu = 5$。误差条表示跨曝光计算的均值标准误差，且彼此独立。信噪比较低、接近阈值的源更可能因给定扰动而被推至阈值之上或之下；而明亮的源无论如何都显著，因此响应很小。

这一问题仅在后续的有效点扩散函数模型中才得到了恰当校正。

对于匹配模型对照，探测器-y 响应在 D1、D2 和 D6 上可分辨，在 D3–D5 上则与零一致。在可分辨处，这表示公式的光谱项的一种体现9，负值¯其中平均 $\mathrm{d}\ln f/\mathrm{d}\lambda$ 相对于 $\mathrm{d}\lambda/\mathrm{d}y$ 为负值，而在斜率为正时则为正值。对于所用的真实星系SED集合，这意味着在短波红外观测的短波长处响应为负，而在大多数星系SED已转折的中波红外观测处，响应为正。

在这一层面上，与匹配扩展点扩散函数的一致性说明处理流程和仪器模型在内部保持一致，因为我们尚未明确说明它们与天空实测数据之间的联系或以其为依据的情况。这种联系已通过实测扩展点扩散函数隐含地体现出来；该函数本身是直接针对校准数据的外部测光结果发现飞行前模型存在不足后，作为直接结果创建的（D. Masters 等，2026，准备中）。我们可以将这种不一致解读为需要更新渲染端的点扩散函数模型，而非反过来。光线追迹得到的点扩散函数与实际飞行中的点扩散函数之间的差异，早在测光验证初期就已发现，但更新后的模型尚未传递到 Sky Simulator，而 EOSPHOROS的渲染器正是基于该模拟器构建的。

7.3. HESPEROS 响应模板生成

的完整图像级测光流程也可以与第6 节中的分箱和红移赋值相结合，生成下游响应模板，其中包含所有像素级方差特征和标志；我们可以将其与 EOSPHOROS中的直接目录模板进行比较。图10 展示了观测等效的快速 EOSPHOROS 星表与使用 HES-的完整像素级方差层和标志生成的星表之间的比较

表4. 每个探测器阵列的相干天体测量响应

探测器	N_{exp}	相干 y	相干 x	随机方向
飞行前光线追踪 PSF 渲染				
D1	36, 131	-14.98 ± 0.05 (-300.3)	$+5.70 \pm 0.04$ ($+133.4$)	-0.02 ± 0.04 (-0.4)
D2	36, 126	-9.21 ± 0.05 (-203.3)	-0.22 ± 0.04 (-5.6)	-0.04 ± 0.04 (-1.0)
D3	36, 115	-10.25 ± 0.05 (-194.7)	-0.32 ± 0.04 (-7.6)	-0.03 ± 0.04 (-0.8)
D4	36, 021	$+9.53 \pm 0.04$ ($+220.3$)	-0.38 ± 0.04 (-10.3)	$+0.01 \pm 0.03$ ($+0.3$)
D5	36,033	$+11.70 \pm 0.05$ ($+232.5$)	$+6.84 \pm 0.05$ ($+148.4$)	$+0.02 \pm 0.04$ ($+0.6$)
D6	36,011	$+13.53 \pm 0.05$ ($+253.4$)	$+7.14 \pm 0.05$ ($+155.0$)	-0.02 ± 0.04 (-0.4)
匹配飞行 ePSF 渲染（对照）				
D1	36, 131	-0.23 ± 0.05 (-4.9)	-0.71 ± 0.04 (-16.1)	-0.04 ± 0.04 (-1.0)
D2	36, 126	-0.36 ± 0.05 (-7.5)	-0.65 ± 0.05 (-14.0)	$+0.01 \pm 0.04$ ($+0.1$)
D3	36, 115	-0.01 ± 0.05 (-0.3)	-0.41 ± 0.05 (-8.8)	$+0.05 \pm 0.04$ ($+1.3$)
D4	36, 022	$+0.10 \pm 0.05$ ($+2.1$)	-0.77 ± 0.05 (-17.0)	-0.11 ± 0.04 (-2.9)
D5	36,033	$+0.07 \pm 0.05$ ($+1.4$)	-0.31 ± 0.04 (-7.0)	-0.02 ± 0.04 (-0.5)
D6	36, 011	$+0.22 \pm 0.04$ ($+5.1$)	-0.15 ± 0.04 (-3.6)	-0.01 ± 0.04 (-0.2)

Note—扰动的相干响应 $\mathrm{d}\ln N_{\mathrm{sig}}/\mathrm{d}\alpha$，单位为 10^{-2} 原生探测器像素的倒数，对应第7.1节中的三种扰动方向。误差为该探测器各 N_{exp} 次曝光的标准误差，并根据公式13 对每次曝光独立计算；括号中的数值为相应的显著性。随机方向列应与零一致，并用作零假设检验。

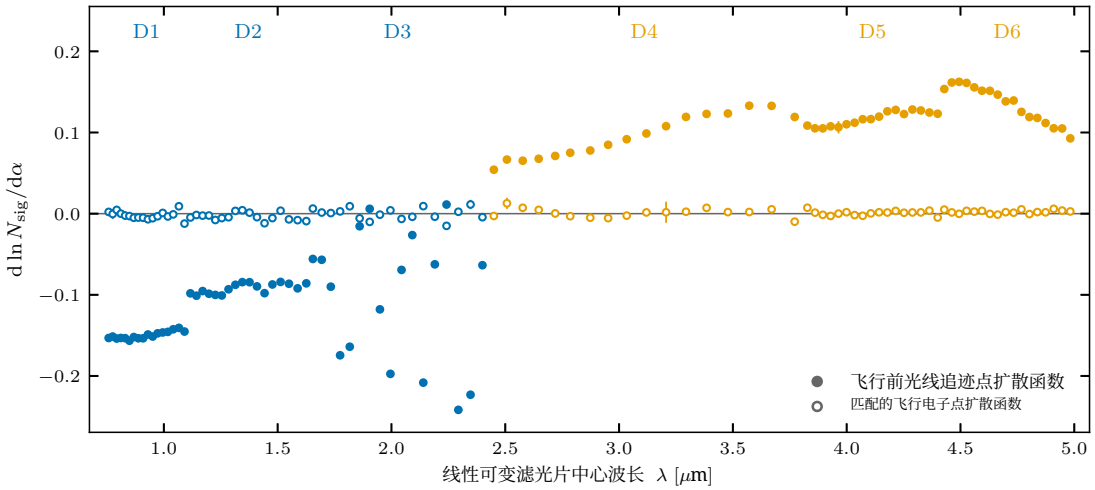


图9。 沿LVF梯度方向（探测器-y）的相干响应 $\mathrm{d}\ln N_{\mathrm{sig}}/\mathrm{d}\alpha$，分别针对飞行前PSF模型和在轨数据导出的ePSF，覆盖102个基准波长通道和六个物理探测器阵列。点表示分箱后的原生波长观测，误差条再次表示独立曝光之间均值的标准误差。匹配ePSF情形下的残余偏移很小，反映了LVF响应于这组输入光谱能量分布的结果。显著性是相对于曝光间散布而言的，并随所用曝光数量的平方根增加。

PEROS。主要差异体现在：银河系核球方向瞬变率较高，导致部分测量值缺失；此外，背景或环境特征各异，EOSPHO-ROS噪声模型未能将其纳入。HESPEROS版本包含所有回访观测、非计划观测或停机时段，并刻画局部背景条件的差异，这些差异可能未包含在解析模型中。两个模板均渲染在 $\sim 10^6$次曝光中观测到的 10^7 源；为简化处理并降低像素/体素采样要求，我们对径向分箱进行合并。这两个模板之间的相关性为$r = +0.940$，其差值的均方根约为任一模板均方根的三分之一。在这种密度下，快速 EOSPHOROS 模板的生成耗时为 0.53 个 GH200 图形处理器小时；相比之下，完整的 HESPEROS 模板耗时为 19 4× 个 GB200 节点小时（76 个 Blackwell 图形处理器小时）（对于 fp64 工作负载，代际差异仅为 [$\sim$ 10%] ）。

7.4. 天体测量误差边缘化

为了展示 HES-PEROS 梯度在实践中的另一种用途，我们评估了参考曝光中存在 0.25 像素的相干探测器x偏移的情况：在偏移位置使用 TRACTOR测量通量进行渲染，但在原始星表位置进行拟合。图 11 展示了通量恢复受影响的程度如何随源的原始信噪比而变化：若对偏移的影响不加处理，亮源会产生显著偏差；而即使采用均匀先验对偏移进行边缘化，也能够恢复无偏通量，且在所有信噪比下，第 16 至第 84 百分位数均落在零的 0.02σ 范围内。由于大量源共享这一相干偏移，拟合可以轻松地约束其影响，不过，其他与总体或信噪比相关的效应可能需要更有针对性的处理或采用不同方法。边缘化需要进行六次导数遍历和一次二阶导数计算，总工作量相当于在图像上进行一次测光拟合的约 11 倍。这个单参数案例刻意设计得较为简单，但增加方向时成本增长较低，加之测量算子整体速度很快，凸显了对更复杂或更多样的效应进行更精细处理的潜力。

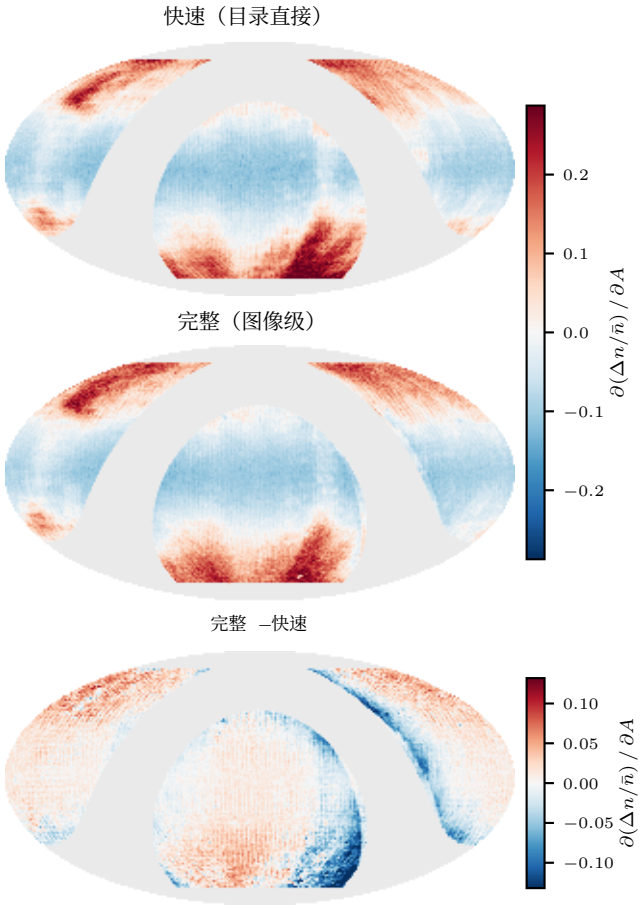


图10。 第 6 节中同一深度响应模板的快速目录版和完整目录版。顶部地图显示了图 5 中深度响应模板的一年巡天等效版本，该版本使用直接星表模式和解析噪声近似生成；中间地图显示了使用实际飞行噪声水平和 DR1 首年数据集中的图像级标记计算出的同一量，该计算通过完整的测光测量管线完成。两者均已测量 10^7 个源。底部地图显示了二者之间的差异，其主要由银河系核球附近的数据丢失，以及相对于模型变化的背景和环境特征所导致；这反映了快速目录对完整解的近似程度。为便于可视化，地图使用 1° 高斯进行平滑。

8. 讨论

8.1. 吞吐量提升的影响

HESPEROS&EOSPHOROS 的 GPU 加速吞吐量从根本上改变了关于大规模再处理任务，其中小区域的处理可在数小时内完成，而巡天规模的运行在单个节点上只需几天。系统—系统模板可以直接估计，并根据第 6.4 节所述的耦合测量级模型进行设计，然后在下游 批量处理。这些响应模板的生成不受蒙特卡洛抽样的限制，计算需求仅取决于所需的角分辨率和径向分辨率。在生产分辨率下，使用单个 GH200 节点生成符合实际的模板需要数小时。

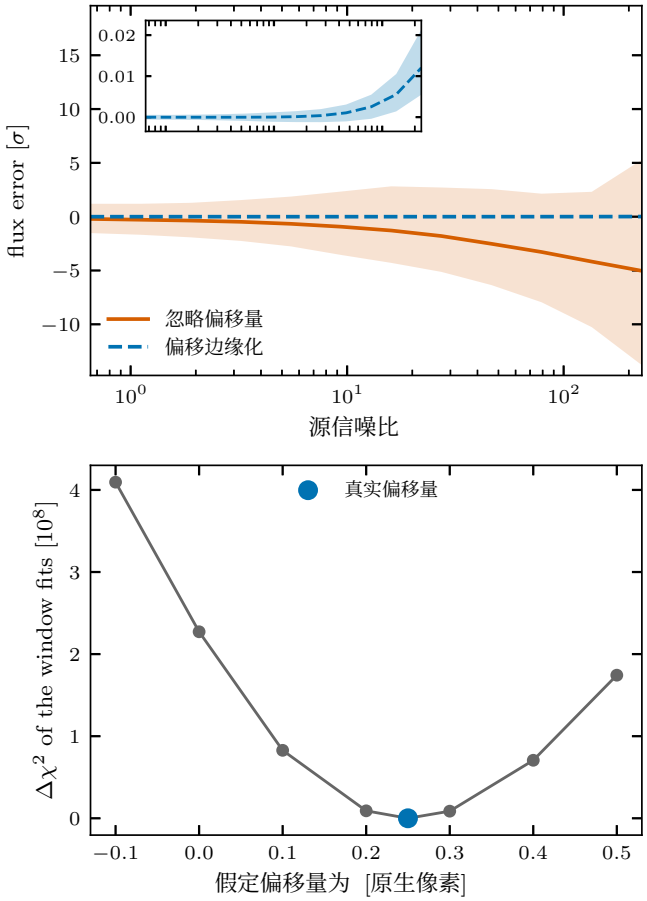


图11. 边缘化相干位置偏移的影响。在第 x 节的参考曝光中，对每个源沿探测器方向施加统一的 0.25 像素天体测量偏移3.2，并在假设不存在偏移的情况下进行拟合。此处特意选取了较大的偏移，以突出其影响。源以其 TRACTOR测得的真实图像通量进行渲染。 **上图**：相对于真实位置处拟合结果的误差（以西格玛为单位）随源输入信噪比变化的情况。橙色表示忽略偏移、让其未经处理地影响测量；蓝色虚线表示对偏移进行边缘化，使用 HESPEROS 。阴影带表示第 16 至第 84 百分位数；插图以扩展坐标范围展示边缘化的情况，以便看清这些百分位数。 **下图**：对于拟合中假设的每种偏移量，窗口拟合的 χ^2 变化；在真实的 0.25 像素偏移处达到最小值。此处使用网格扫描进行说明，但实际优化中使用的是梯度。

现在也可以通过处理流程大规模获取导数，从而能够对各个步骤进行基于梯度的优化，并实现新类别的推断。快速迭代测量配置（例如背景、点扩散函数或黄道光发射模型）的能力，以及随着巡天进展迅速响应不断到来的仪器数据的能力，使分析流水线得以快速迭代和优化，并以往在实践上（或经济上）不可能实现的对仪器响应的理解。

8.2. 与源注入的关系

通过导数获得的流水线响应天然地与源注入协同作用，实际上还为充分利用源注入提供了一个强大的框架。注入会对处理流程中所有（可能不可微分的）非线性和离散性方面进行采样，例如精细的去混叠拓扑、图像级瞬变和伪影、目录筛选，以及掩膜决策。相较于完整重新运行测量算子（第 5.3节），额外梯度的成本通常很低，因此每次注入都能以较低成本提供额外信息；而生产测量本身的速度如今也已进入能够在大范围天区内进行密集、多次实现注入的阶段。

利用该框架获得的响应，可以刻画或边缘化一类效应，而合成源注入不得不从一开始就继承这些效应，即由探测器增益校准、天体测量定标或点扩散函数等特征引起的响应。对这些效应进行仪器层面的详细刻画，对于获得尽可能准确的测量值仍然至关重要；但此处介绍的可微分方法提供了一种机制，可以评估其对下游的影响，或将知识不足吸收到误差预算中。

在响应平滑的区域及方向上，导数适合用于直接获得流水线响应；而对于高度非线性的效应，则更适合采用额外的注入。完整的注入处理流程还需要大量的记录管理，以及生产测量配置和输入真实值的记录，其中渲染和测量基础设施只是其组成部分。 EOSPHOROS渲染器可以很容易地替换为其他渲染器，例如 Sky Simulator 或 GALSIM （需适当考虑 LVF响应），同时仍可使用可微测量算子 HESPEROS，或者使用下游的分箱、红移估计和层析分配功能（尽管这些功能在EOSPHOROS中实现，但并不要求必须使用其中的渲染器）。HESPEROS和 EOSPHOROS 构成了可微分基础，配套论文（S. Everett 等，2026年，准备中）将在此基础上开发注入应用。

8.3. 局限性

我们指出这里展示的能力存在一些重要局限性。

此处展示的范围仅涉及微分，采用刻意不切实际、在红移上平坦的输入分布

仅用于测量处理流程的响应。正如第 5.3 节所述，在响应强烈非线性的情况下，导数可能不可靠；对于方差路径而言，响应还会强烈依赖所考虑的噪声机制。违反静态假设的大幅扰动（例如，扰动源的位置，使其落入不同的测量窗口、退化混合或掩膜中）也会使导数失效，不过在代码中遇到此类情况时，通常会引发硬错误。

我们还指出了第4.1 节的快速直接星表模式无法处理、必须借助 HESPEROS &EOSPHOROS完整组合的一些效应。第7 节本身就介绍了其中之一：天体测量 PSF/LVF模型保真度与响应的耦合。其他此类效应包括真实的背景变化和环境依赖性、ePSF 区域和窗口拓扑，以及图像级标记。第6节使用的巡天覆盖模式是基于完整任务计划的近似；在较大尺度上汇总较长时间段内的一组测量值时，这种近似效果良好，但在更精细的空间和时间尺度上，由于真实巡天中的标记观测或回访观测等特征，其准确性会越来越低。

第7 节的归一化取决于信噪比分布，可根据具体情况进行调整。

9. 结论

我们介绍了一种批处理、GPU加速且可微分的强制测光算子，可复现 SPHEREx 测量值 HESPEROS；还介绍了一种可微分的分光光度前向模型，可将源渲染到帧中 EOSPHOROS。我们展示了如何组合二者，以研究测量响应对天体测量扰动和所假设的PSF模型的响应。通过在完整的测量流水线中传播梯度，我们可以直接测量下游量对给定系统误差扰动的响应。

该 HESPEROS 测量流水线采用 Cholesky分解实现，将整个曝光视为一组相互独立的线性系统，并通过批处理高效地完成 JAX 编译及在图形处理器上的部署。通过逐源与公开的 TRAC-TOR 分支进行验证，并在星表层面与生产星表进行对比（在 7.2×10^8 个相互未标记的测量值上，中位数偏差为 0.006σ ），其速度大幅提升：在一台四图形处理器 GB200 节点上，每小时处理的曝光数为8.5×10^4 ，相较于基于中央处理器的 TRACTOR 处理流程，每节点小时的吞吐量 HESPEROS，从而实现了一系列新能力。目前正在评估将 HESPEROS采用为生产版 SPHEREx 测光管线，用于 DR1 之后的发布版本。

前向模型 EOSPHOROS 经验证，与GALSIM 相比，其误差小于图章峰值的几 10^6 分之一，适用于多种剖面，并且也通过了 SPHEREx 天空模拟器的验证。捕捉测量算子对系统误差响应的同一套下游导数传播，也可以直接由前向模型提供输入，从而直接估计建模的曝光级系统误差对下游的影响。该算子的运行速度使其能够大规模应用：处理整个巡天（$\mathcal{O}(10^{10})$ 个光度测量、 $\mathcal{O}(10^8)$ 个源），并进行红移估计和导数计算，耗时约一个 GPU日。前向模型和测量算子可以轻松组合，直接测量各种效应、标志或测量选项的选择响应，也可以单独使用。对此能力的初步演示有力地检测到（195–300σ 跨探测器，表 4）一种不匹配：使用飞行前光线追迹PSF模型渲染的场景，与使用基于飞行数据的有效PSF模型测得的场景之间存在差异。

加速的测光处理流水线能够快速迭代并优化测量，自然也有助于通过源注入测量巡天的选择函数，或训练注入模拟器。可微性自然可以进一步提升注入的效用：以摊销成本测量每次注入的许多额外扰动；或者通过反向模式微分，以不随参数数量变化的成本进行测量。HESPEROS本身速度很快，因此非常适合用于注入应用，因为这类应用通常依赖于重新运行生产流水线，若能大规模重复运行，便会受益匪浅；这将是配套论文的研究主题（S. Everett 等，2026年，准备中）。可微性也使哈密顿蒙特卡洛或其他基于梯度的采样技术得以应用，为系统误差的贝叶斯处理带来了新的可能性。对于诸如 $f_{\rm NL}^{\rm loc}$这类具有挑战性的大规模宇宙学目标，仔细刻画选择函数至关重要；本文所展示的可扩展性和能力提升，是实现这一目标道路上的一步。

数据和代码可用性

HESPEROS& EOSPHOROS作为 SPHEREx 4级科学任务的一部分，仍在持续进行内部开发，之后将与星表产品一同公开，具体时间遵循任务的公开时间表——

表，并在 SPHEREx 天空模拟器公开发布后推出。2级和3级数据产品来自 DR1 SPHEREx 数据处理流水线的内部版本，此后已通过 NASA/IPAC 红外科学档案馆（IRSA）提供[14]，，发布时间遵循任务的公开时间表（SPHEREx 团队 2026年）。第3.2节的比较是针对公开的`tractor-spherex` 分支，在提交`066680f`。

致谢

我们感谢美国国家航空航天局戈达德太空飞行中心通过合同向加州理工学院提供的 SPHEREx 项目支持。JRC 感谢 David & Ellen Lee 奖学金提供的支持。本研究所使用的高性能计算资源由喷气推进实验室信息与技术解决方案局提供资金支持。

计算在德克萨斯高级计算中心（TACC）提供的 STAMPEDE3、 LONESTAR6和 VISTA 集群上运行。部分工作在加州理工学院喷气推进实验室完成，依据的是与美国国家航空航天局签订的合同（80NM0018D0004）。作者还感谢使用 Anthropic 的 Claude OPUS 4.8、5.0 和 FABLE 5 模型协助进行文献综述、代码实现、稿件校对以及部分图表的完善。作者对所有生成内容进行了评估，并对本文所呈现的内容承担全部责任。

软件： THE TRACTOR（D. Lang 等，2016a），`tractor-spherex` （https://github.com/Caltech-IPAC/tractor-spherex），SPHEREx 天空模拟器（B. P. Crill 等，2025年），JAX（J. Bradbury 等，2018年），`numpy` （C. R. Harris 等，2020年），Claude OPUS 4.8、5.0，以及 FABLE 5 （https://www.anthropic.com/）

参考文献

Akeson, R., Dubois-Felsmann, G. P., Crill, B. P., 等，2026年，《天体物理学杂志增刊》，285，67，doi：10.3847/1538-4365/ae8843Anbajagane, D., Tabbutt, M., Beas-Gonzalez, J., 等，2025年，《开放天体物理学杂志》，8，65，doi：10.33232/001c.138627Anderson, J., & King, I. R. 2000，《太平洋天文学会出版物》，112，1360，doi：10.1086/316632苹果公司 2024，Apple M4 Max 芯片技术规格与新闻稿，https://www.apple.com/ca/newsroom/2024/10/apple-introduces-m4-pro-and-m4-max/Arnouts, S., Cristiani, S., Moscardini, L., 等，1999，《皇家天文学会月刊》，310，540。https://arxiv.org/abs/astro-ph/9902290Bahk, H., & Hwang, H. S. 2026年，arXiv 电子预印本，arXiv:2610.16007，doi：10.48550/arXiv.2610.16007Benoˆıt, C. 1924年，《法国地理学会通报》，第 G´eod´esique, 2，67，doi：10.1007/BF03031308Bock, J.J., Aboobaker, A. M., Adamo, J., 等，2026年，《天体物理学杂志》，999，139，doi：10.3847/1538-4357/ae2be2Bradbury, J., Frostig, R., Hawkins, P., 等，2018年，JAX：Python 的可组合变换+NumPy 程序，0.3.13 http://github.com/jax-ml/jax

Crill, B. P., Bach, Y. P., Bryan, S. A., 等. 2025年，《天体物理学杂志增刊》，281，10，doi：10.3847/1538-4365/ae04ccDalal, N.,Dor´e, O., Huterer, D., & Shirokov, A. 2008，物理评论D，77，123514，doi：10.1103/PhysRevD.77.123514de Castro, P., & Dorigo, T. 2019，《计算机物理通讯》，244，170，doi：10.1016/j.cpc.2019.06.007Dey, A., Schlegel, D. J., Lang, D., 等. 2019，《天文学杂志》，157，168，doi：10.3847/1538-3881/ab089dDor´e, O., Bock, J., Ashby, M., 等. 2014，SPHEREx 全天光谱巡天宇宙学，https://arxiv.org/abs/1412.4872Elsner, F., Leistedt, B., & Peiris, H. V. 2016，《皇家天文学会月刊》，456，2095，doi：10.1093/mnras/stv2777Euclid Collaboration, Gillard, W., Maciaszek, T., 等. 2026年，《天文学与天体物理学》，707，A227，doi：10.1051/0004-6361/202555859Everett, S., Yanny, B., Kuropatkin, N., 等. 2022，《天体物理学杂志增刊》，258，15，doi：10.3847/1538-4365/ac26c1Fazar, C. M., Zemcov, M., Dowell, C. D., 等. 2026年，arXiv 电子预印本，arXiv:2608.09862，doi：10.48550/arXiv.2608.09862Feder, R. M., Masters, D. C., Lee, B., 等. 2024，《天体物理学杂志》，972，68，doi：10.3847/1538-4357/ad596dHance, M., & Robles, J. 2025，《SciPost物理核心》，8，079，doi：10.21468/SciPostPhysCore.8.4.079

[14] https://irsa.ipac.caltech.edu/Missions/spherex.html[15] http://www.tacc.utexas.edu/

Harris, C. R., Millman, K. J., van der Walt, S. J., 等. 2020年，《自然》，585，357，doi：10.1038/s41586-020-2649-2Huai, Z., Bock, J. J., Cheng, Y.-T., 等. 2026年，《天体物理学杂志》，1000，56，doi：10.3847/1538-4357/ae472bHuang, S., Leauthaud, A., Murata, R., 等. 2018年，《日本天文学会刊物》，70，S6，doi：10.1093/pasj/psx126Hui, H., Bock, J., Condon, S., 等. 2026年，arXiv 电子预印本，arXiv:2602.09139，doi：10.48550/arXiv.2602.09139Huterer, D., Cunha, C. E., & Fang, W. 2013年，《皇家天文学会月刊》，432，2945，doi：10.1093/mnras/stt653Ilbert, O., Arnouts, S., McCracken, H. J., 等. 2006，《天文学与天体物理学》，457，841。https://arxiv.org/abs/astro-ph/0603217Kelsall, T., Weiland, J. L., Franz, B. A., 等. 1998，《天体物理学杂志》，508，44，doi：10.1086/306380Kong, H., Burleigh, K. J., Ross, A., 等. 2020年，《皇家天文学会月刊》，499，3943，doi：10.1093/mnras/staa2742Kong, H., Ross, A. J., Honscheid, K., 等. 2025年，《宇宙学与天体粒子物理学杂志》，2025，146，doi：10.1088/1475-7516/2025/01/146Korngut, P. M., Bock, J. J., Condon, S., 等. 2026年，arXiv 电子预印本，arXiv:2603.29835，doi：10.48550/arXiv.2603.29835Lang, D.,Hogg, D.W., & Mykytyn,D. 2016a，The Tractor：概率天文源探测与测量，天体物理学源代码库，ascl:1604.008 http://ascl.net/1604.008Lang, D., Hogg, D. W., & Schlegel, D. J. 2016b，《天文学杂志》，151，36，doi：10.3847/0004-6256/151/2/36Li, X., & Mandelbaum, R. 2023年，《皇家天文学会月刊》，521，4904，doi：10.1093/mnras/stad890Li, X., Mandelbaum, R., Jarvis, M., 等. 2024，《皇家天文学会月刊》，527，10388，doi：10.1093/mnras/stad3895Mendoza, I., Guinot, A., Becker, M. R., 等. 2026年，arXiv 电子预印本。https://arxiv.org/abs/2604.22048NVIDIA Corporation. 2023，NVIDIA GH200 Grace Hopper超级芯片数据手册，NVIDIA Corporation。
h
t
t
p
s://www.nvidia.co
m/en-us/data-center/grace-hopper-superchip/
NVIDIA Corporation. 2024年，NVIDIA Blackwell 架构数据手册，NVIDIA Corporation。
h
t
t
p
s://
d
a
m-
c
d
n.
n
v
d.orangelogic.co
m/AssetLink/01dyp27s6aj4e0483cs5goo15354m438.pdfPullen, A. R., & Hirata, C. M. 2013年，《太平洋天文学会出版物》，125, 705，doi：10.1086/671189Rosado-Marín, A. J., Ross, A. J., Seo, H., 等. 2025年，《宇宙学与天体粒子物理学杂志》，2025年，018，doi：10.1088/1475-7516/2025/07/018Rowe, B. T. P., Jarvis, M., Mandelbaum, R., 等. 2015年，天文学与计算，10, 121，doi：10.1016/j.ascom.2015.02.002Schlafly, E. F., Green, G. M., Lang, D., 等. 2018年，《天体物理学杂志增刊》，234, 39，doi：10.3847/1538-4365/aaa3e2Sheldon, E. S., Becker, M. R., MacCrann, N., & Jarvis, M. 2020年，《天体物理学杂志》，902, 138，doi：10.3847/1538-4357/abb595Sheldon, E. S., & Huff, E. M. 2017年，《天体物理学杂志》，841, 24，doi：10.3847/1538-4357/aa704bSimpson, N., & Heinrich, L. 2023年，载于《物理学会议系列》，第2438卷，《物理学会议系列》（英国物理学会），012105，doi：10.1088/1742-6596/2438/1/012105SPHEREx 团队. 2026年，SPHEREx 第一版数据发布（DR1）光谱图像，IPAC，doi：10.26131/IRSA678Stickley, N. R., Capak, P., Masters, D., 等. 2016年，arXiv 电子预印本。https://arxiv.org/abs/1606.06374Stone, C., Courteau, S., Cuillandre, J.-C., 等. 2023年，《皇家天文学会月刊》，525, 6377，doi：10.1093/mnras/stad2477Strong, G. C., Lagrange, M., Orio, A., 等. 2024年，《机器学习：科学与技术》，5, 035002，doi：10.1088/2632-2153/ad52e7Suchyta, E., Huff, E.M., Aleksić, J., 等. 2016年，《皇家天文学会月刊》，457, 786，doi：10.1093/mnras/stv2953Weaverdyck, N., & Huterer, D. 2021年，《皇家天文学会月刊》，503, 5061，doi：10.1093/mnras/stab709

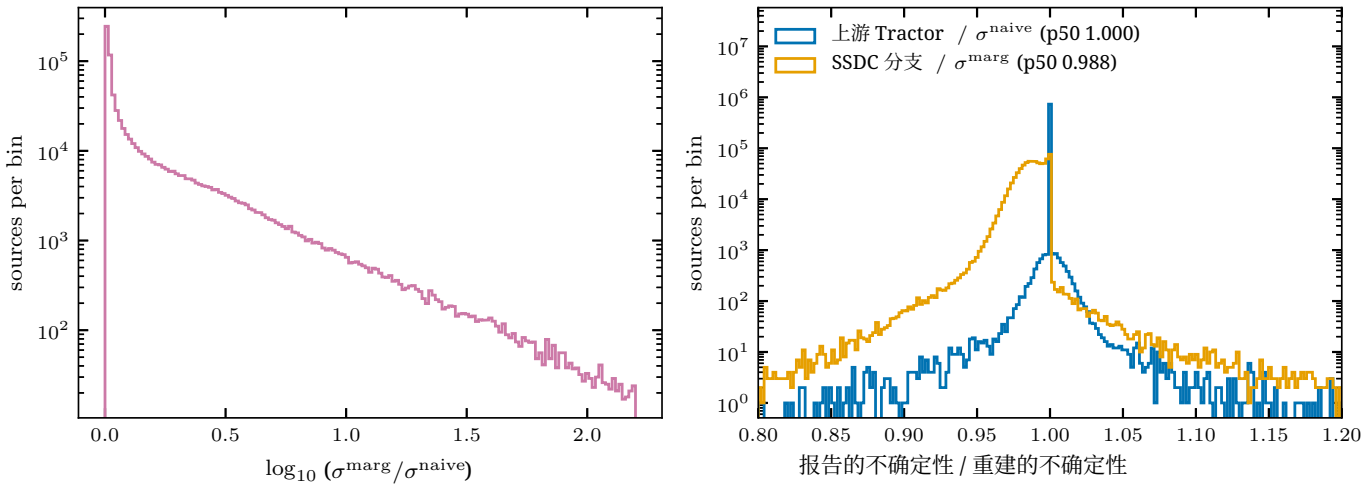


图 12. 参考曝光的不确定性约定与比较。**左**：HESPEROS 边缘误差与朴素误差之比，展示协方差（混合）贡献的影响。**右**：比较 HESPEROS 的不确定性约定与上游 THE TRACTOR 代码（蓝色）的结果；上游代码提供朴素误差 $\sigma^{\rm naive}$。同时也与公开的 SSDC 分支（橙色）进行比较，该分支将混合项纳入边缘不确定度 $\sigma^{\rm marg}$。SSDC 分支看似存在的 1% 缺失／堆积，是由于其在计算不确定性时选择冻结背景项 b 所致；我们则选择对其进行边缘化。在 HESPEROS 中冻结该项后，与上游分支的朴素不确定性相比，也能达到相近程度的一致性。

附录

A. HESPEROS 验证详情

对于通量测量本身，SSDC 分支与上游 TRACTOR 代码得出的结果相同，差异在于不确定性。图12 展示了参考曝光下 $\sigma^{\rm naive}$ 和 $\sigma^{\rm marg}$ 分别与上游TRACTOR 以及 SSDC 分支的比较，并比较了 HESPEROS派生的 $\sigma^{\rm naive}$和 $\sigma^{\rm marg}$。$\sigma^{\rm naive}$ 在第1至第99百分位数范围内，将上游构建版本的不确定性重现至 1。000；而 $\sigma^{\rm marg}$ 与 THE TRACTOR 分支吻合，其中位数为 0。988（范围为第1至第99百分位数；差异源于参考版本的 Fisher 矩阵固定背景项的方式）。

图 13 展示了这个密集银河系参考曝光场中亮源的一些吻合良好的测量值。13。另一方面，图 14 展示了两个实现版本在参考曝光中差异最大的一些案例。这些案例是由退化混合以及大量从拟合中排除的标记像素造成的，导致拟合条件不佳（在实际分析中，这些测量值，甚至整个曝光，很可能会在后续处理阶段被标记）。 HESPEROS 即使在最差案例中，解也始终更接近其 WLS 最优解。

使用了观测密度范围内的 20 次曝光来比较各实现，并依据一组保守的预定义标准进行检查：恢复的标志必须完全准确；位于参考解（精确版本） 1σ 之外的源必须达到 $< 0.75\%$；并且求解器之间的距离必须与参考解自身在 $\rho \geq 0.9$ 处的残差相关。20 次曝光中有 15 次通过了所有这些标准。在未通过的 5 次中，有 3 次是严重受掩膜影响的帧，应当将其剔除并标记后提交给巡天项目重新尝试（通常是由于瞬变事件）；其中 1 次是比较中观测密度最高的视场，包含 39% 的退化窗口；最后 1 次则是由于边界效应，单行中存在一个标志差异。对于一组匹配的未标记观测，HESPEROS 在 20/20 次曝光中都比精确解更接近最优解，这由距最优解超过 1σ 的源所占比例评估得出；例如，对于其中一次高度退化的曝光，参考解的比例为 42.2%，而 HESPEROS 为 12.7%。在具有相同源集和相同剔除条件的情况下， HESPEROS 直接解在所有被考虑的情形中，都以这一指标优于迭代参考解。

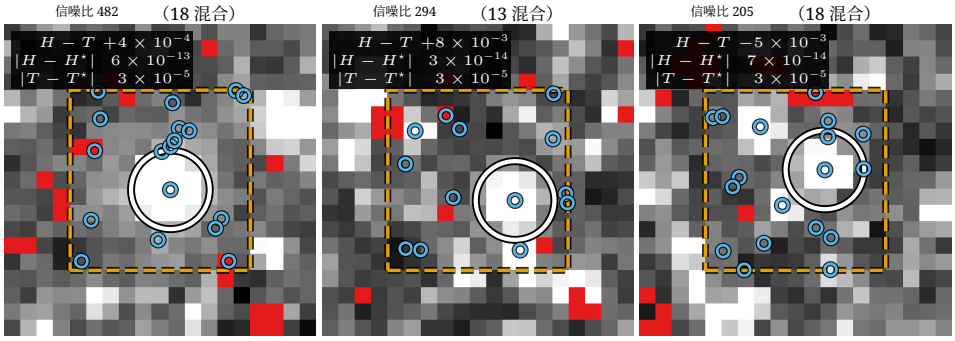


图13. 参考曝光中，两种实现对一些低标记率高信噪比（$S/N > 200$）源的测量结果相当一致，以下展示其中一部分。测量的源以白色圆圈突出显示，而唯一裁剪图（橙色虚线方框）中的所有光度测量源均以蓝色圆圈显示。插图展示了该源的 HES-PEROS-TRACTOR 差值（“H − T”），单位为 TRACTOR 测量值的 $\sigma^{\rm marg}$；同时还展示了各通量与其自身 WLS 最优解的绝对距离（分别以 HESPEROS 和 THE TRACTOR的 $H^\star$ 和 $T^\star$ 表示），单位为各自的 $\sigma^{\rm marg}$。L2图像中的红色像素为标记像素，拟合时已将其排除。

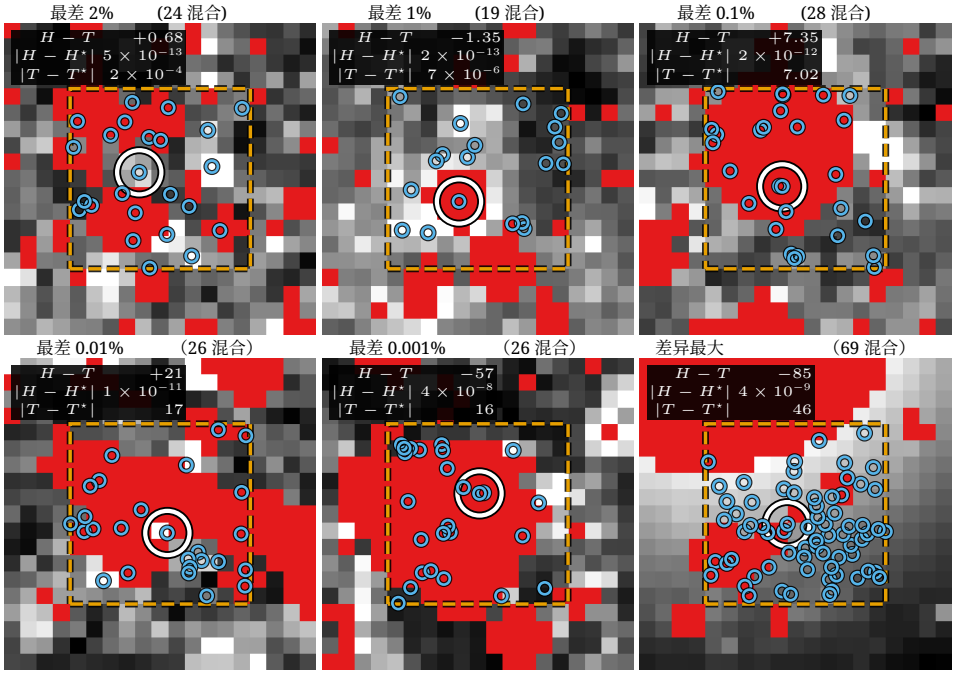


图14. 在参考曝光中， HESPEROS与 TRACTOR 参考值之间差异特别大的部分测量结果，取自参考曝光中刻意选取的相距较远的位置。图中展示了通量差异最接近若干分级百分位数的单个通量测量值，以白色圆圈突出显示；同时还显示了其 11×11像素测量窗口，以及其他同时进行光度测量的源（以蓝色圆圈表示）。图中展示了从最差的 2% 到最差的 0.001% 的百分位数，以及曝光中差异最大的单个测量值。插图中的数值与图13相同；如前所述，红色像素表示拟合时排除的标记像素。差异最大的案例通常是由过多的标记像素造成的，但即使在这种情况下， HESPEROS 与其自身的参考最优值（$H^\star$）仍始终更接近。差异最大的单个案例，是由图像裁剪区域左上边缘之外、亮星光晕附近大量可能为虚假或被遮挡的源造成的。

B. S´埃尔西克渲染验证

为了评估 EOSPHOROS 的扩展 S´ersic 渲染与 GALSIM 的渲染效果，我们使用两种代码，对随机选取的 9,799 个已渲染 Legacy巡天形状的测量通量差异进行比较。我们使用 GALSIM 在同一周期画布上、以与 EOSPHOROS 相同的方式，在 18×18 个过采样亚像素单元的中心放置 S´ersic 轮廓，并将其与高斯点扩散函数进行卷积，采用相同的亚像素权重，然后求和至 SPHEREx 原生

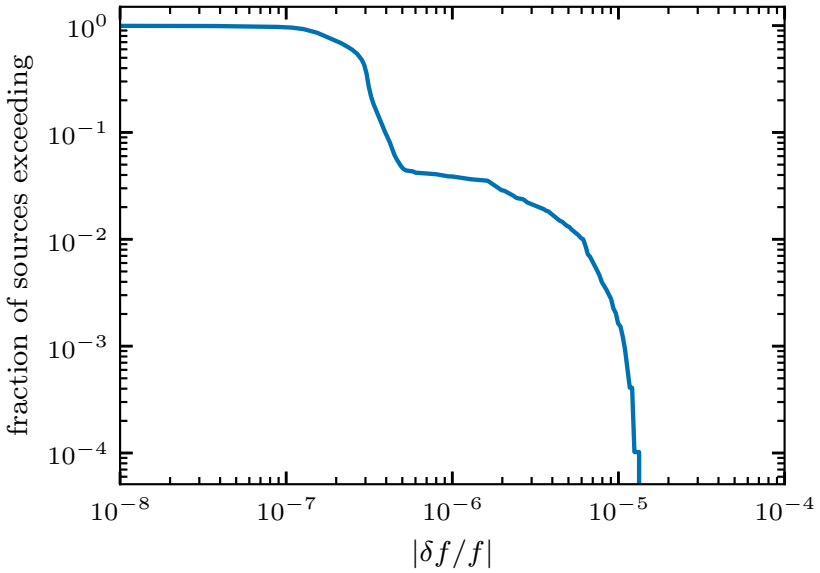


图15。 对于从参考星表（Legacy巡天）形状总体中随机选取的9,799个星系样本，在随机亚像素位置、图像旋转和奇偶性下，对比 EOSPHOROS 与 GALSIM扩展渲染时，分数通量差超过 $|\delta f/f|$的源所占比例。对于每幅图像，均使用相同的（GALSIM）轮廓测量通量。超过 5×10^{-7} 的星系中，有90%过于扩展，无法完整地放入画布中，因此预计其精度只能达到大约这一水平。

像素。 EOSPHOROS 使用默认设置渲染源。由于点扩散函数是共享的，这项测试涵盖了轮廓、剪切和方向、单元积分以及放置。形状从 1.45×10^{6} 扩展源（$r_e \geq 1''$）的十二个参考星表视场中，在（索引、半光半径、轴比和位置角）上联合绘制。分数通量差 $|\delta f/f|$ 通过对 GALSIM 轮廓分别拟合 GALSIM 和 EOSPHOROS图像的最小二乘拟合计算得出，并采用均匀像素噪声。

图 15 显示了分数通量差超过 $|\delta f/f|$ 的源所占比例，整体数值都很小；在渲染的星系中，差异超过 5×10^{-7} 的部分源里，90% 是非常大的源，其延伸范围超出了可用画布边缘的最大值。预计这些源的准确度也只能达到这一水平左右。通量差的中位数为3×10^{-7}，最大的分数差为 1.4×10^{-5}。与测量噪声相比，这些差异小得多；即使将最大的分数误差应用于第 3.2 节参考曝光中测得的最亮未饱和源（信噪比 514），也只会导致通量差 0.007σ，中位差为 $1.5 \times 10^{-4}\sigma$。