



Photometry on Structured Backgrounds: Local Pixel-wise Infilling by Regression

Andrew K. Saydjari^{1,2} and Douglas P. Finkbeiner^{1,2} ¹ Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA; andrew.saydjari@cfa.harvard.edu² Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Received 2021 December 31; revised 2022 March 11; accepted 2022 April 18; published 2022 July 12

Abstract

Photometric pipelines struggle to estimate both the flux and flux uncertainty for stars in the presence of structured backgrounds such as filaments or clouds. However, it is exactly stars in these complex regions that are critical to understanding star formation and the structure of the interstellar medium. We develop a method, similar to Gaussian process regression, which we term local pixel-wise infilling (LPI). Using a local covariance estimate, we predict the background behind each star and the uncertainty of that prediction in order to improve estimates of flux and flux uncertainty. We show the validity of our model on synthetic data and real dust fields. We further demonstrate that the method is stable even in the crowded field limit. While we focus on optical-IR photometry, this method is not restricted to those wavelengths. We apply this technique to the 34 billion detections in the second data release of the Dark Energy Camera Plane Survey. In addition to removing many $>3\sigma$ outliers and improving uncertainty estimates by a factor of ~ 2 – 3 on nebulous fields, we also show that our method is well behaved on uncrowded fields. The entirely post-processing nature of our implementation of LPI photometry allows it to easily improve the flux and flux uncertainty estimates of past as well as future surveys.

Unified Astronomy Thesaurus concepts: Astronomy data analysis (1858); Photometry (1234); Sky surveys (1464); Catalogs (205); Gaussian Processes regression (1930)

Supporting material: animation

1. Introduction

Upcoming projects such as the Legacy Survey of Space and Time at the Vera C. Rubin Observatory (Jones et al. 2020) and observations from the Nancy Grace Roman Space Telescope (Akeson et al. 2019) will deliver measurements of stellar flux (brightness) for billions of stars, including those near the Galactic plane. Measurement of stellar flux in the Galactic plane is complicated by crowding (which makes it difficult to partition flux between nearby overlapping sources), and the presence of structured backgrounds³ (the flux of which should not be attributed to point sources).

The difficulties of crowding and structured backgrounds near the Galactic plane were explicitly noted in the Sloan Extension for Galactic Underpinnings and Evolution (Newberg & Sloan Digital Sky Survey Collaboration 2003), the Sloan Digital Sky Survey (SDSS) release at low galactic latitudes (Finkbeiner et al. 2004), and the Dark Energy Camera Plane Survey (DECaPS) (Schlafly et al. 2018). Some of these projects created modified photometric pipelines to better handle crowded fields and source deblending, but did not address structured backgrounds. Other works addressing photometry on structured backgrounds focus solely on methods for source identification (Molinari et al. 2011) or use polynomial fits of varying degrees to model the true background, both of which complicate estimating the uncertainty on the reported stellar flux (Traficante et al. 2015). A general technique able to handle both biases and uncertainties resulting from structured

backgrounds could not only improve future measurements in the galactic plane (Rubin and Roman), but also create value-added catalogs using images from any of the large-scale photometric surveys that target the Galactic plane such as SDSS, DECaPS, Wide-field Infrared Survey Explorer (WISE) (Schlafly et al. 2019), the Two Micron All Sky Survey (Skrutskie et al. 2006), Pan-STARRS1 (Magnier et al. 2020), Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE) (Benjamin et al. 2003), VISTA Variables in the Via Lactea (VVV) (Minniti et al. 2010), and the UKIRT Infrared Deep Sky Survey (UKIDSS) (Lawrence et al. 2007).

Instead of using classic background models that are smooth on a spatial averaging scale that is (and must be) larger than the point-spread function (PSF), we can instead perform an interpolation of the background behind the star. Using that local background model, we can separate the true flux of the star from background flux. In the machine-learning literature, reconstructing an unknown portion of an image is often called a *conditional infilling* or *inpainting* problem and many approaches exist, often using some version of generative adversarial networks (Pathak et al. 2016; Ulyanov et al. 2018; Van Veen et al. 2018; Wu et al. 2018). In addition to deriving corrections to the flux as a result of the true background, it is important to quantify the overall uncertainty in the reported flux. Without reliable uncertainty estimates, any downstream statistical inference, which is the main application of these astronomical catalogs, is polluted with overly tight constraints. This is a limitation for approaches based on neural networks because it is difficult to obtain reliable uncertainty estimates on their predictions.

We improve the estimates of stellar flux and error bars on structured fields by solving the infilling problem behind the star and obtaining an uncertainty on that infill. To do this, we perform a regression for the unknown pixels given the values of neighboring pixels that are known. The local pixel-wise infilling (LPI) used here is similar to Gaussian process

³ We refer to filaments of gas and dust as *background* regardless of whether those features are physically in front of or behind the star.



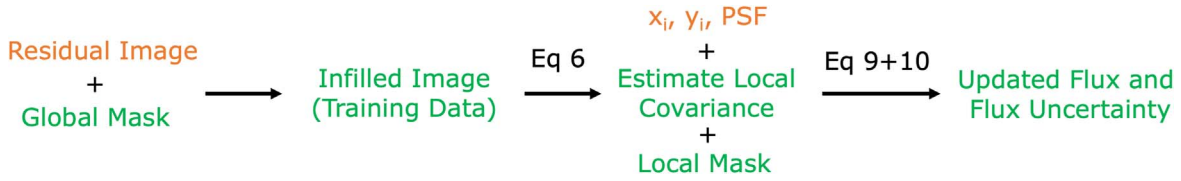


Figure 1. A scheme illustrating the steps used in this work to improve flux and flux uncertainty estimates for sources on structured backgrounds. Orange text represents data products from a previous photometric solution for the image of interest. Green text represents data products generated as part of this work.

regression (GPR), often called Gaussian process interpolation or Kriging. In GPR, the value of a function (the true background) is predicted at unobserved points conditioned on nearby observed points using a covariance function (called a kernel). Note that GPR assumes weak stationarity—that the kernel depends only on the (possibly angle-dependent) separation between points. Exact stationarity is often violated, but holds approximately in real applications. GPR is a form of nonparametric regression since each unobserved point is predicted without a prior specification on the functional form (linear, quadratic, etc.) of the interpolated image. However, GPR almost always uses a parametric kernel function. The benefit of parametric kernels is that the covariance can be sampled at any arbitrary separation. The downside is that an optimization over kernel parameters is required and one must consider whether the parametric kernel is a good model of the true covariance of the data.

In astronomy, GPR has already found numerous applications, including foreground removal (Mertens et al. 2018; Ghosh et al. 2020; Kern & Liu 2021; Soares et al. 2021), foreground modeling by point-source interpolation (Pinter et al. 2018), interpolating galactic-scale fields (Platen et al. 2011; Yu et al. 2017; Leike & Enßlin 2019; Williams et al. 2022), cleaning/interpolating missing data (Czekala et al. 2015; Baghi et al. 2016; Ndritu et al. 2021; Zhang & Brandt 2021), predicting photometric redshifts (Way et al. 2009; Almosallam et al. 2016), light-curve analysis (Evans et al. 2015; Littlefair et al. 2017; McAllister et al. 2017; Angus et al. 2018), radial velocity analysis (Haywood et al. 2014; Barclay et al. 2015; Rajpaul et al. 2015), removing instrumental systematics (Gibson et al. 2012; Junklewitz et al. 2016), interpolating atmospheric turbulence to improve astrometry (Fortino et al. 2021; Léget et al. 2021), modeling PSF variations (Gentile et al. 2013), or predicting object properties (Bu et al. 2020; Fielder et al. 2021).

The work of Fortino et al. (2021) notably improves astrometry (measurements of the position of stars) on images from the Dark Energy Camera (DECam), which is the same camera used for validating our photometric model in Section 5. The extraction of clean foreground models by interpolating to remove point sources or bad pixels shown in Pinter et al. (2018) and Zhang & Brandt (2021) serve as a proof of concept for this work. We further derive the correction to and uncertainties on the fluxes of those point sources as a result of the background model. Because we need to sample the covariance only on a grid (fixed by the pixelization of the camera) and need only a small number of points for the local background, we are able to use more flexible nonparametric models for the LPI covariance in contrast to the astronomical applications of GPR above.

2. Background Model

Ideally, we would like to be able to estimate the background underneath a star from a detailed analysis of the nearby

background. We find that we can successfully estimate the background and its uncertainty using an infilling technique, which will be the focus of Section 2.3. This technique dramatically improves flux and flux uncertainty estimates for stars in images of fields with significant nebulousity (Section 5.2). In the intervening sections, we provide detailed descriptions, simulations, and validations of this technique, but the eager reader may wish to skip to Section 5. Figure 1 broadly outlines the steps used in our technique.

2.1. Likelihood

Photometric pipelines process astronomical images in order to model the location and flux (brightness) of the stars in the image. Focusing on surveys at low galactic latitudes where stars far outnumber galaxies, we set aside modeling galaxies for now. In addition to stars which are effectively point sources, these images often contain diffuse components such as emission from interstellar gas or reflection off of interstellar dust (Magakian 2003). For example, the bright $H\alpha$ Balmer line of hydrogen and various emission lines for oxygen in different ionization states fall in the optical wavelength range (Lide 2004). Dense interstellar dust can also block background starlight, causing a decrease in detector counts over an extended region (Di Francesco et al. 2002). These diffuse components and background sky counts are often combined into a single background term b , which when added to the sum of N_{star} point sources yields the true image I . The contribution of a point source to the image is modeled as the flux of the star (a scalar) times a PSF, which is the spatial distribution of photons arriving at the focal plane from a point source. The PSF deviates from a single point because of atmospheric turbulence, telescope optics, and sensor imperfections. Thus, the true image (I) is

$$I(x, y) = b(x, y) + \sum_{i=1}^{N_{\text{star}}} f_i \times p(x - x_i, y - y_i) + \epsilon(x, y), \quad (1)$$

where b is the background, f_i is the flux of star i , $p(x - x_i, y - y_i)$ is the PSF centered at position (x_i, y_i) in the image, and $\epsilon(x, y)$ is the measurement noise of each pixel. Then, the residual between the model and the true image is $r = I - \hat{I}$, where we have dropped the explicit (x, y) dependence from here on.⁴

$$r = \epsilon + b - \hat{b} + \sum_{i=1}^{N_{\text{star}}} f_i p(x_i, y_i) - \sum_{i=1}^{\hat{N}_{\text{star}}} \hat{f}_i \hat{p}(\hat{x}_i, \hat{y}_i). \quad (2)$$

The noise in observed detector counts is from a mix of Poisson (photon events, shot noise) and Gaussian (read noise) processes. However, in optical surveys, we are generally in the

⁴ In the code released with this paper, we use the opposite convention for the residuals, $r = \hat{I} - I$.

large count limit and thus approximate the observed noise in each pixel as Gaussian. Then, we find the model parameters that minimize the negative log-likelihood, given some estimate of the per-pixel covariance matrix \hat{C} ,⁵

$$-2 \ln \mathcal{L}(\hat{f}_i, \hat{x}_i, \hat{y}_i, \hat{b}) = \ln(\det(2\pi\hat{C})) + r^T \hat{C}^{-1} r. \quad (3)$$

To solve this problem in full generality, one has to optimize the likelihood with respect to \hat{N}_{star} , a process known as transdimensional inference (Brewer et al. 2013; Portillo et al. 2017; Feder et al. 2020; Liu et al. 2021). Such an approach can be made Bayesian by including a prior on \hat{N}_{star} and the other parameters. In practice, a heuristic deblending code that finds peaks in the image, assigns sources to those locations, and estimates a parametric PSF model is often used. Then, the optimization is performed only over \hat{f}_i , \hat{x}_i , and \hat{y}_i with the other model parameters held fixed (Stetson 1987; Bertin & Arnouts 1996; Schlafly et al. 2018). We consider the latter approach. In what follows, we will also ignore uncertainties and corrections resulting from estimation of the stellar position (\hat{x}_i and \hat{y}_i) such as those outlined in Portillo et al. (2020) and any errors resulting from PSF mismodeling, focusing instead on the previously unaddressed uncertainties and corrections to \hat{f}_i in the presence of mismodeled backgrounds.

It is tempting to assume that r is mean zero (as assumed in Equation (3)) with \hat{C} diagonal (as is often done), which would be nearly true in the limit of perfect modeling and noisy data. Then, the diagonal of \hat{C} is simply the variance of each pixel (proportional to model counts assuming Poisson noise), and matrix inversion can be computed with little computational expense as element-wise division. However, the \hat{b} used in photometric pipelines is often not an accurate model of the background b and in this work we will consider how to deal with a non-diagonal C resulting from mismodeled backgrounds.

2.2. Literature Background Models

Popular photometric pipelines such as DAOPHOT (Stetson 1987), SExtractor (Bertin & Arnouts 1996), the Tractor (Lang et al. 2016), and *crowdsources* (Schlafly et al. 2018) use background models that are smooth over large (relative to the PSF) spatial scales. Prior works have explored different methods to obtain a robust and/or unbiased estimation of a constant background value for a subset of an image. Most of these methods operate on the histogram of pixel values in the subimage and estimate the mean, median, or mode, often with iterative outlier removal (“ σ -clipping”) (Da Costa 1992). The final background model applied at the pixel scale is then a smooth interpolation of individual subimage estimations. These approaches necessarily have an averaging scale (often 32–128 pixels) set by the size of the subimages considered (Bertin & Arnouts 1996). So \hat{b} can be a good model of background components that vary on a spatial scale larger than the averaging scale. However, high resolution imaging, from the Hubble Space Telescope, for example, indicates structures in the interstellar medium (ISM) at scales all the way down to $<1''$, which would not be well captured by \hat{b} .

In this case, the incorrect assumption that r is mean zero can bias the flux estimates \hat{f}_i for stars in the region. For example, if

the observed background is above the model ($(b - \hat{b}) > 0$), then the \hat{f}_i will overestimate the stellar flux so that r is closer to mean zero (Figure 2). The assumption that r has a diagonal C ignores nonzero (and as we will show in Section 4 often large) off-diagonal contributions to the flux uncertainty, which are a result of the spatially correlated residuals from the mismodeled background.

2.3. Local Pixel-wise Infilling

Since we cannot see behind the star, we describe the distribution of possible backgrounds by the set of images that are shifted horizontally/vertically by a number of pixels. We can view this as data augmentation of our single corrupted sample, where we have assumed local translation invariance of the statistics of the ISM (stationarity). This translation invariance should hold if the samples are *local enough*. We have *not* assumed rotational invariance (isotropy) because of the presence of highly directional filaments in the ISM (see Appendix D). Then, we perform a regression for the unknown pixels given the values of neighboring pixels that are known. Together, we term this approach LPI.

This regression can be viewed as GPR with a nonparametric anisotropic kernel, so long as the covariance matrix we estimate is actually stationary. While our covariance matrix is approximately stationary (see Appendix D), we do not constrain it to be. We also do not require the covariance obey a functional form such that the covariance matrix is positive definite for all sample points, a common constraint in GPR. More simply, this regression can be viewed as a linear least-squares regression using the known pixel (k) values as the regressors to predict the unknown pixel (k_*) values in the high-dimensional pixel space (see notational correspondence in Appendix C). The local shifted samples are then the training data from which the parameters connecting the k - and k_* -pixels are estimated. Note that this model is not linear in the 2D spatial dimensions of the image. Even though we do not enforce stationarity, we find it convenient to use the language of GPR throughout.

We first estimate a local covariance matrix ($N_p^2 \times N_p^2$) for a subimage ($N_p \times N_p$ pixels) of the residuals centered on each star. We focus on applying LPI to the residuals because we want to predict nonzero mean deviations between the model and true background, assuming the star was not present in the image (Equation (3)). It is useful to decompose the N_p^2 pixels of the subimage into three classes: *good*, which we believe to be uncontaminated samples of the local background, *hidden*, which are obscured by the star of interest and should be interpolated, and *ignored*, which are contaminated either by other stars, cosmic-rays, pixel defects, etc. The *ignored* pixels should be excluded from background estimation and need not be interpolated. See Sections 4 and 5 for detailed examples of how pixels are partitioned between these classes.

The goal of LPI is to predict the value of missing data in a vector (i.e., the *hidden* pixels) given the data present (i.e., the *good* pixels) and a model for the covariance of the data vector. Consider a data vector X , of length equal to the number of pixels in a subimage around a star N_p^2 , and the corresponding mean vector μ and covariance matrix C . We denote the indices in X that are *good* k and those that are *hidden* k_* . Then, $\bar{X}(k_*|k)$, the mean of the missing values predicted conditionally using the known values, is given by Equation (4) and the covariance

⁵ In what follows, we will suppress a factor of 2 and normalization constants when writing likelihoods since they do not depend on model parameters and thus do not affect the optimization.

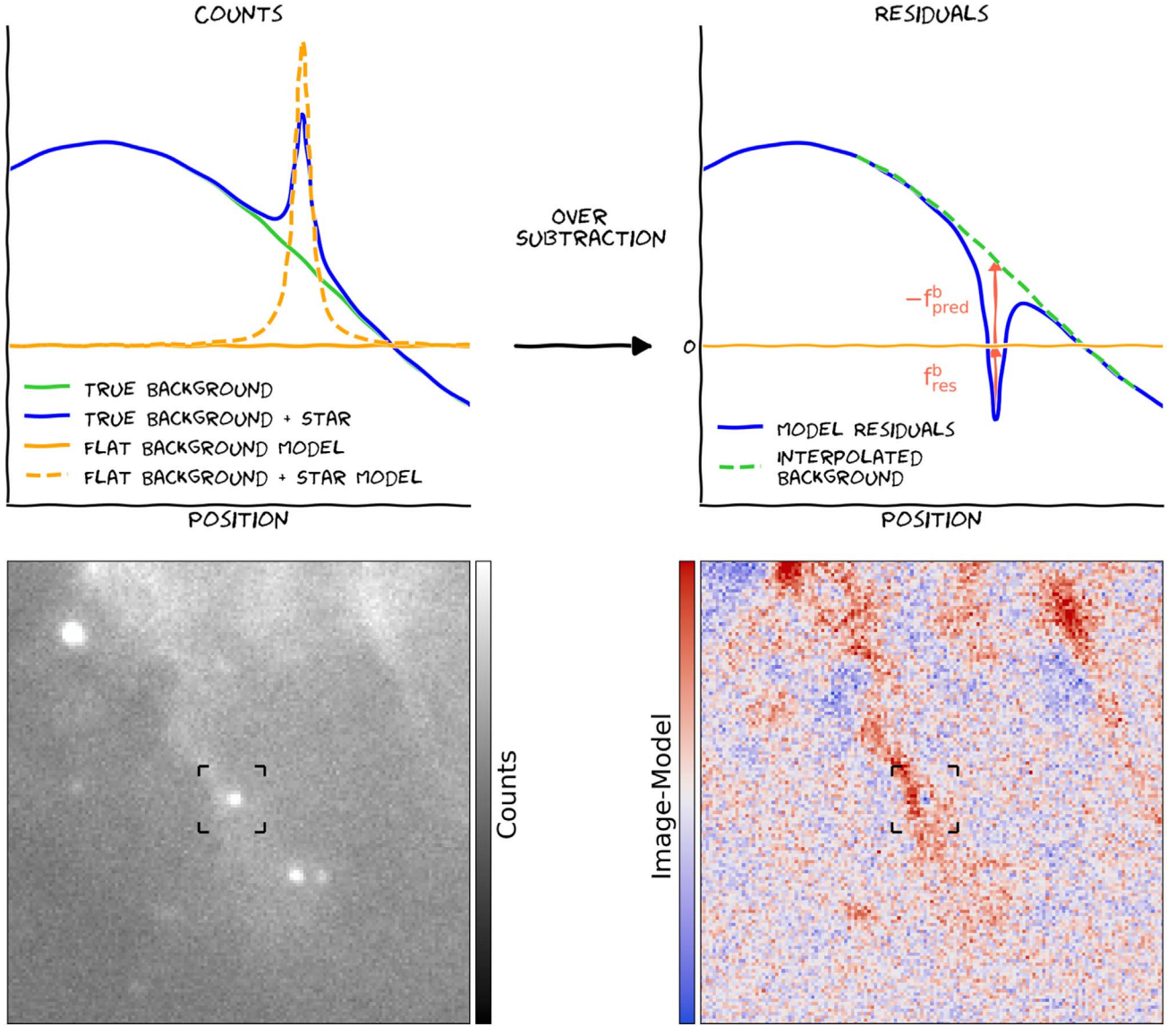


Figure 2. Top left panel: schematic showing overestimation of flux for a star (yellow dashed) given a flat background model (yellow solid) in the presence of a structured background (green solid). Top right panel: the residuals after subtraction, which illustrate that f_{res}^b accounts for the deviation of the residuals from zero and $-f_{\text{pred}}^b$ accounts for the deviation of the predicted background (green dashed) from zero. Bottom left panel: an example star imposed on a structured background in a r -band image from DECaPS2 (Section 5). Bottom right panel: residuals after photometric fit of image at left using `crowdsourc`. The fit drives the residual in a neighborhood around the highlighted source to be mean zero, relative to an incorrect background model, causing the center to be significantly oversubtracted.

of the prediction is given by Equation (5) (Rasmussen 2003),

$$\bar{X}(k_*|k) \equiv \bar{X}'_{k_*} = C_{k_*,k} C_{k,k}^{-1} [X_k - \mu_k] + \mu_{k_*}. \quad (4)$$

Here, $C_{k_*,k}$ denotes the C sub-block of rows with k_* -indices and columns with k -indices. The means of the pixels with k -indices and k_* -indices are denoted by μ_k and μ_{k_*} , respectively.

$$\text{cov}(X'_{k_*} - \bar{X}'_{k_*}) = C_{k_*,k_*} - C_{k_*,k} C_{k,k}^{-1} C_{k,k_*} \quad (5)$$

In GPR, C is often specified by a choice of parametric kernel (either chosen or with parameters learned from data). However, in LPI we learn \hat{C} directly from the data without any parameterization, as described next.

2.4. Covariance Construction

We learn \hat{C} (an $N_p^2 \times N_p^2$ matrix) for a subimage of $N_p \times N_p$ pixels (which we can view as a vector X of $N_p^2 \times 1$ numbers).

To obtain samples of X , we shift our $N_p \times N_p$ image up to N_x^{sample} and N_y^{sample} pixels horizontally and vertically from the stellar center. Assuming that all of the $N_p \times N_p$ subimages derived from a local region are drawn from the same distribution, we can use these samples to estimate C by $\hat{C} = V^T V / (N - 1) - \mu_X \mu_X^T$ where the rows of V are each a sample of N_p^2 pixels, N is the total number of samples, and μ_X^T is obtained by averaging over rows (samples). For sufficiently small spatial scales, the statistics of the dust/gas should be uniform and this stationarity assumption thus valid.

In practice, constructing V and computing $V^T V$ per star is too computationally expensive. For the application in Section 5, the reordering of the calculation we describe below is four orders of magnitude faster than a direct matrix multiplication per star for fields with the median number of stars (see Appendix A). Given that all of our samples are related simply by pixel shifts of the same image, we can perform the averaging over samples

as a spatial average using a computationally inexpensive boxcar mean. Consider an analogous problem in 1D, with samples $V_{ij} = X_{i+(j-1)}$, where X is 1 indexed.

$$\begin{aligned}\hat{C}_{ij} &= \frac{1}{N-1} \sum_{k=1}^N V_{ki} V_{kj} - \frac{1}{N^2} \left(\sum_{k=1}^N V_{ki} \right) \left(\sum_{k=1}^N V_{kj} \right) \\ &= \frac{1}{N-1} \sum_{k=1}^N X_{k+(i-1)} X_{k+(j-1)} \\ &\quad - \frac{1}{N^2} \left(\sum_{k=1}^N X_{k+(i-1)} \right) \left(\sum_{k=1}^N X_{k+(j-1)} \right)\end{aligned}\quad (6)$$

The elements of V^{TV} are simply the elements of a boxcar mean of width N on the product of X with itself shifted by $(i-j)$. This can be quickly computed for the entire image, and then referenced at the location of each star. Since we consider 2D image data, we implement Equation (6) with an additional spatial index.

2.5. Preliminary Infilling

The local covariance estimate must be computed over *good* pixels, excluding *hidden* and *ignored* pixels. Unfortunately the naive approach of considering the masked, shifted boxcar means fails because summing over only masked pixels in Equation (6) does not correspond to V^{TV} , and the resulting matrix is no longer guaranteed to be positive semidefinite (Baghi et al. 2016). Instead, we must make a reasonable guess to infill the masked pixels. As long as the masking fraction is low, the covariance matrix estimate is not strongly sensitive to the infill.

Our infilling procedure uses a boxcar mean of the *good* pixels to predict masked pixels using a variety of filter widths (both *hidden* and *ignored* pixels must be infilled in this initial step). The initial width is 19 pixels, and all pixels with a mean value determined by 10+ good pixels are infilled. The filter width is increased by a factor of 1.4 in successive infilling steps until all of the masked pixels are infilled. In heavily crowded/masked regions, it is useful to have a cutoff at a finite number of infilling steps (10), and infill the remaining pixels with the median image value. These numerical choices were qualitatively chosen for performance on the application in Section 5.

These infilled pixels should also have per-pixel uncertainties consistent with the rest of the image, so we replace each masked pixel by a Poisson draw. However, we must account for the counts already subtracted out in the smooth background model and detector gain, which pools g photoelectron counts per digital unit.⁶

$$g \times (\hat{b} + \hat{r}) \sim \text{Poisson}(g \times (\hat{b} + \bar{r})) \quad (7)$$

2.6. Deriving Corrections

One of the benefits of the LPI background model is the ease of deriving improved flux estimates. The maximum likelihood estimate for the flux correction to a star can be found by setting the first derivative of the log-likelihood $\ln \mathcal{L}(f^b)$ over the k_* -pixels with respect to the bias offset f^b equal to zero.

⁶ The background plus residuals should never be negative as such negativity would indicate that the model image included sources with negative flux or nonpositive portions of the PSF. Unfortunately, both of these situations can occur in *crowdsources* outputs on difficult fields, so we take $[\text{background} + \text{residuals}]$ in order to report a value that the user can decide to use based on quality flags.

Similarly, the variance of the new flux estimate $\hat{f} - f^b$ is obtained as the negative inverse of the second derivative of the log-likelihood with respect to the flux estimate,

$$-\ln \mathcal{L} \propto (f^b p + r - \bar{X}'_{k_*})^T \hat{C}^{-1} (f^b p + r - \bar{X}'_{k_*}), \quad (8)$$

$$f^b = -\frac{p^T \hat{C}^{-1} r}{p^T \hat{C}^{-1} p} + \frac{p^T \hat{C}^{-1} \bar{X}'_{k_*}}{p^T \hat{C}^{-1} p} = f_{\text{res}}^b - f_{\text{pred}}^b, \quad (9)$$

$$\sigma_{f+f^b}^2 = -(\partial_{f+f^b}^2 \ln \mathcal{L})^{-1} = (p^T \hat{C}^{-1} p)^{-1}. \quad (10)$$

We have suppressed the per-star PSF (p) position evaluation and i subscripts for clarity, highlighting the f^b dependence. Because \hat{C} learned in Section 2.4 was only estimated on pixels representative of uncontaminated background, we must adjust the per-pixel variance (diagonal) of \hat{C} to account for (Poisson) noise from modeled counts before computing Equations (9) and (10) (see Section 5.1).

We interpret the first term, which we denote f_{res}^b , as accounting for the fact that the initial photometric pipeline might not be able to achieve its target of zero residuals through modeling point sources. This often leads to overestimation of the flux from stars to soak up counts from bright structured backgrounds. In the worst case, photometric pipelines can add a large number of nonexistent sources to the model in order to account for the background flux. This suppresses evidence of background mismodeling in the residuals and is a failure mode beyond the scope of this work.

The second term, which we denote $-f_{\text{pred}}^b$, measures the expected deviation from zero residuals with respect to a flat background model as a result of the interpolated structured background. A pictorial scheme illustrating the role of these terms is depicted in Figure 2. We packaged all of the above steps required to compute LPI corrections to photometry into a code, CLOUDCOVER.JL, which we release with this work (see Section 7).

An alternative approach to the one pursued here would be to incorporate the predicted background and non-diagonal covariance matrix into Equation (3) as an improved \hat{b} and \hat{C} . In this way, the improved background model and uncertainties are incorporated into the original photometric solution. One difficulty with this approach is that we currently require an initial photometric solution in order to have access to residuals. Another complication is that the implementation will depend on whether the photometric pipeline solves for the flux of each source independently, or as a joint optimization problem. For these reasons, we present a pure afterburner as a proof of concept, which is conveniently compatible with any prior photometric solution.

3. Gaussian Approximation Validity

In the ISM, there are isolated dust clouds, filaments of gas, and shock fronts all of which are highly non-Gaussian features. This is in contrast to cosmic microwave background cosmology where the structures of interest are well described by Gaussian random fields and small deviations from Gaussianity are of interest. As a result of this strong heterogeneity, the non-Gaussian structure of the ISM cannot even be described as some functional transform of a Gaussian random field, such as by a log-normal random field. In order to draw the correspondence to Gaussian process methods and evaluate the validity of using LPI in the context of these non-Gaussian

fields, one can view the non-Gaussian field as being represented as a Gaussian mixture model (GMM)—a sum of many different Gaussian components.

Naively, one might expect LPI to be valid only when the local pixels are drawn from a single component of the Gaussian mixture model and the local image is thus locally Gaussian. We will show in the context of a toy problem that LPI works well in a larger class of cases where we only require that the conditional distribution for the k_+ -pixels given the known k_- -pixels is Gaussian. Then, we will demonstrate that the LPI works well in practice on a real, filamentary dust image.

3.1. Toy Problem

Consider an image consisting of two pixels, pixel A and pixel B. Suppose the image is drawn from a population with a 50% chance of being drawn from $\mathcal{N}(\mu = [-1, 1], \sigma = [\epsilon, \epsilon])$ or $\mathcal{N}(\mu = [1, -1], \sigma = [\epsilon, \epsilon])$. The distribution of the pixels in this subimage is clearly non-Gaussian, composed of two separated Gaussian components (Figure 3). Further, the Gaussian approximation to this distribution is quite poor, placing maximal probability on the infinitesimally rare image $[0, 0]$.

However, when one pixel value is known (say pixel B without any loss of generality), the LPI prediction provides a good approximation of the distribution of the unknown pixel. Pictorially, the predicted conditional distribution on pixel A given the known value of pixel B is simply a renormalized horizontal line cut through Figure 3 at that value of pixel B. Numerically, we follow Equations (4) and (5) to obtain $\mathcal{N}(\mu = -1/(1 + \epsilon^2), \sigma = \sqrt{2}\epsilon)$, a close approximation to $\mathcal{N}(\mu = -1, \sigma = \epsilon)$ for pixel A when pixel B = 1, though with inflated variance.

The success of this model problem lends confidence to LPI being able to correctly interpolate non-Gaussian local subfields that may contain filaments. Two adjacent pixels transverse to the long filament direction will be highly anticorrelated on the positive or negative slope off of the peak of the filament. This will lead to two Gaussian components similar to those in Figure 3. When considering the slope changing along the transverse direction, more pairs of Gaussian components closer or farther from the origin will be added to the population. However, a GMM of components along the diagonal are all clearly still resolvable by LPI. It is only when there are components of the GMM that are not distinguished by the *good* conditioning pixels that LPI can be a poor approximation.

3.2. Interpolating in a Nebula

To demonstrate that the assumptions of LPI are compatible in practice with astronomical fields containing structured backgrounds, we mask and interpolate a real image from the WISE 12 μm dust map (Meisner & Finkbeiner 2014) corresponding to a portion of the Pipe Nebula (between Barnard 67 and 78) (Figure 4). We preprocess the initial image by first clamping outliers to be equal to the highest and lowest 0.01% percentile and then reducing the resolution from $6''$ to $45''$ pixel $^{-1}$.⁷ The covariance matrix for a $N_p = 49$ pixel subimage was estimated from samples on a larger local region (using a boxcar width of $N_x^{\text{sample}} = N_y^{\text{sample}} = 129$ pixels), shown in the left panel of Figure 4.

⁷ The WISE 12 μm dust map was smoothed during production to $15''$ so this is only a $3\times$ reduction in resolution. We downsample to $45''$ pixel $^{-1}$ only to put the features of interest at a pixel scale that is computationally inexpensive.

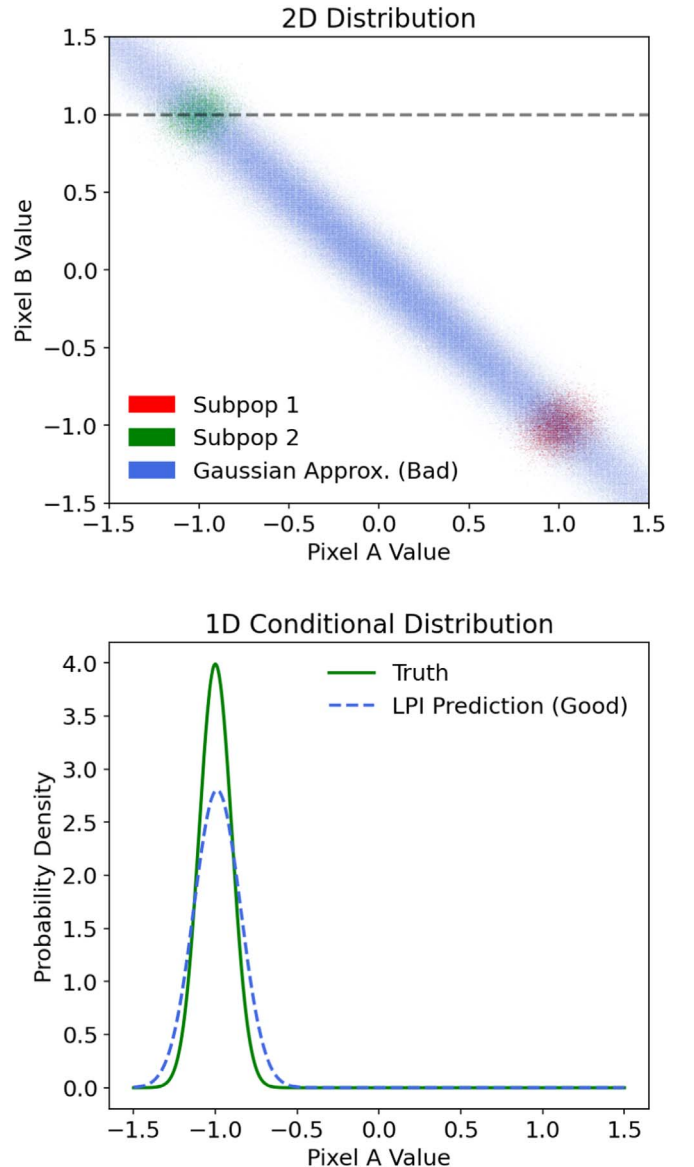


Figure 3. Top panel: probability distribution for the population in the toy 2D problem. The two Gaussian subpopulations are colored red and green. The Gaussian approximation to the distribution is shown in blue, and poorly describes the total population. A line cut along the gray dashed line yields the bottom panel. Bottom panel: the conditional probability distribution for pixel A given pixel B has a value of 1.0. While the true distribution (green) is more sharply peaked, the Gaussian process regression prediction (blue dashed) describes the conditional distribution well.

The true $N_p \times N_p$ subimage is shown in the “Truth” subpanel and features a core with filamentary connections toward dense regions at the top right and right of the subimage. In the masked image, the central mask with radius ~ 13 pixels defines the k_+ -pixels (*hidden*). The known true values of the *good* k_- -pixels are shown. The pixels far away from the interpolation region are *ignored* since they should have little impact on the interpolation and add to the computational expense. Videos animating Figure 4 as the sizes of the *ignored* region (https://faun.rc.fas.harvard.edu/saydjari/CloudCovErr/rseries_test.mp4) and *hidden* region (https://faun.rc.fas.harvard.edu/saydjari/CloudCovErr/thr_test.mp4) are changed demonstrate that the interpolation is robust to these choices.

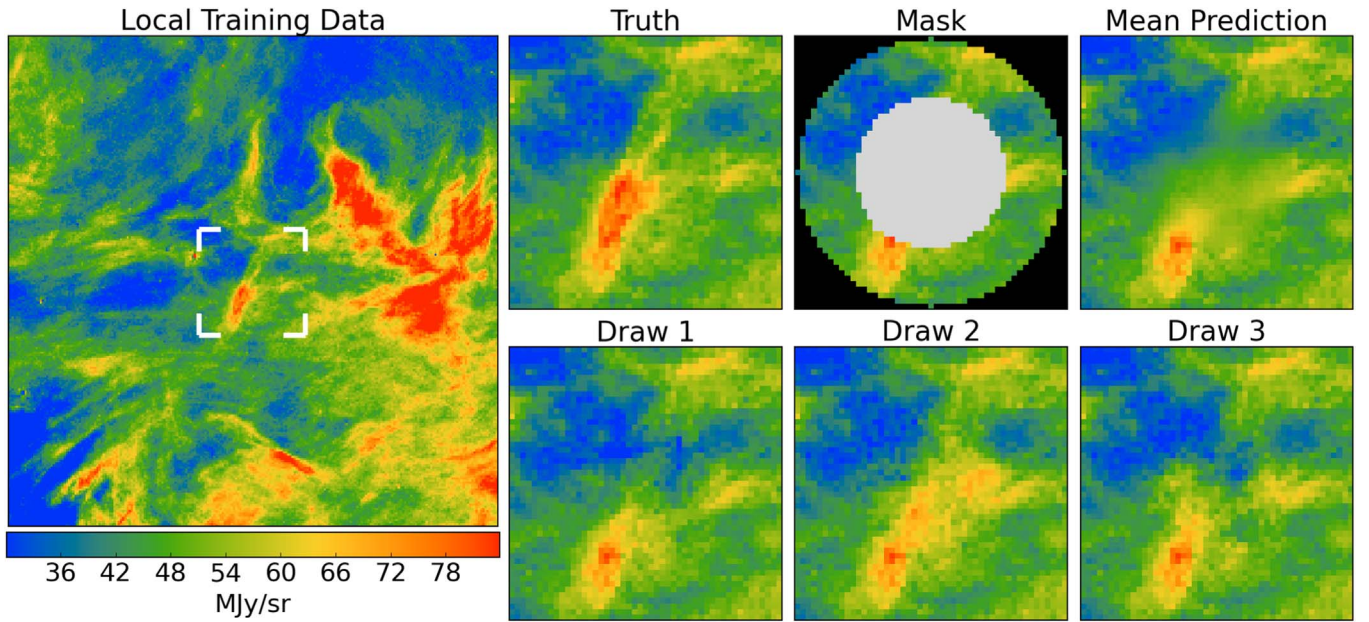


Figure 4. Left panel: image of dust emission from WISE 12 μm dust map containing a portion of the Pipe Nebula. The image scale shows all data used to train the LPI covariance, which is highly heterogeneous. Square marked by four white corners indicates subimage used in other panels. Top right panel: “Truth” shows the central $N_p = 49$ pixel square cutout of the WISE image. In “Mask,” the central gray pixels are those predicted (*hidden*) and the outer black pixels not used (*ignored*) during the prediction. In “Mean Prediction,” the *hidden* pixels are replaced with the mean LPI predictions and all other pixels display the true values. Bottom right panel: each image replaces the *hidden* pixels with a different random draw from the distribution of LPI predictions. All other pixels display the true values. Videos animating this plot as the size of the *ignored* region (https://faun.rc.fas.harvard.edu/saydjari/CloudCovErr/rseries_test.mp4) and *hidden* region (https://faun.rc.fas.harvard.edu/saydjari/CloudCovErr/thr_test.mp4) is changed demonstrate that these interpolations are robust, even for a narrow ring of *good* pixels. The animation includes the videos from both links. The first 17 s of the animation varies the size of the *ignored* region while the remaining 40 s varies the size of the *hidden* region.

(An animation of this figure is available.)

The mean prediction of the LPI is shown replacing the values of the k_* -pixels in the true image. Comparing this mean prediction to the “Truth,” we see that near the core, the interpolation correctly predicts a higher density, and on average predicts an overdensity extending toward the dense region at the right of the image. For the diagonal filament, the mean predicted density appears to have a much lower intensity between the two cores than the true image. However, the mean does not contain all the information provided by the LPI, and we must also consider the covariance.

One way to visualize the covariance of the LPI is to sample the distribution of interpolated images. Samples are easily obtained as $\sqrt{\hat{C}}y + \bar{X}'_{k_*}$, where $y \sim \mathcal{N}(0, 1)$, \bar{X}'_{k_*} is the predicted mean, and $\sqrt{\hat{C}}$ is the symmetric square root, which we obtain via singular value decomposition. Three different draws are shown in the bottom row of Figure 4. While none of these draws exactly reproduces the true image, the interpolated pixels have noise statistics which match the true image and some of the draws do have filaments connecting the central core to both overdense regions (i.e., Draw 2).

Thus, the LPI is consistent with the “Truth,” even though the LPI is not centered on the “Truth.” However, the masking scale here was specifically chosen to be much larger than the scale of the smallest resolved dust structures in the image. This mask thus represents a large information loss—a worst-case scenario for the infilling problem. The recovery of both a reasonable mean and substructure similar (but not identical) to the “Truth” in this limit is a testament to the capabilities of LPI.

4. Synthetic Validation

In order to show that our LPI bias offsets and error bars provide a substantial improvement over current photometric pipelines, we run our post-processing code on outputs from *crowdsourc*, a crowded-field photometric pipeline that was used to create catalogs from DECam and WISE (Schlafly et al. 2018, 2019). Throughout the rest of this work, we compare our flux and flux uncertainty estimates to the flux and flux uncertainty estimates from *crowdsourc*, which we use as the *conventional* benchmark. First we consider synthetic fields here, and then astrophysical images from the Dark Energy Camera Plane Survey (DECaPS2) survey in Section 5.

4.1. Treatment of Emission and Extinction

The first synthetic field we generate consists of two filaments, each having a Gaussian profile with an FWHM of 6.7 pixels (see Figure 5). One filament is positive and one negative relative to the background. These parallel filaments are separated spatially by 118 pixels and the background and Poisson noise are consistent with fields from DECaPS2.⁸ Stars, all of the same flux (~ 18 th g-band magnitude), are then injected on a regular grid with centers at integer pixel values. Each row in the grid is shifted by 1 pixel in the direction transverse to the filament in order to sample stars with different displacements relative to the filament. We then run *crowdsourc* on the synthetic image. We fix the PSF model to match the model used to inject sources and use less aggressive

⁸ For this reason, we quote flux in analog-to-digital units (ADUs) throughout synthetic tests.

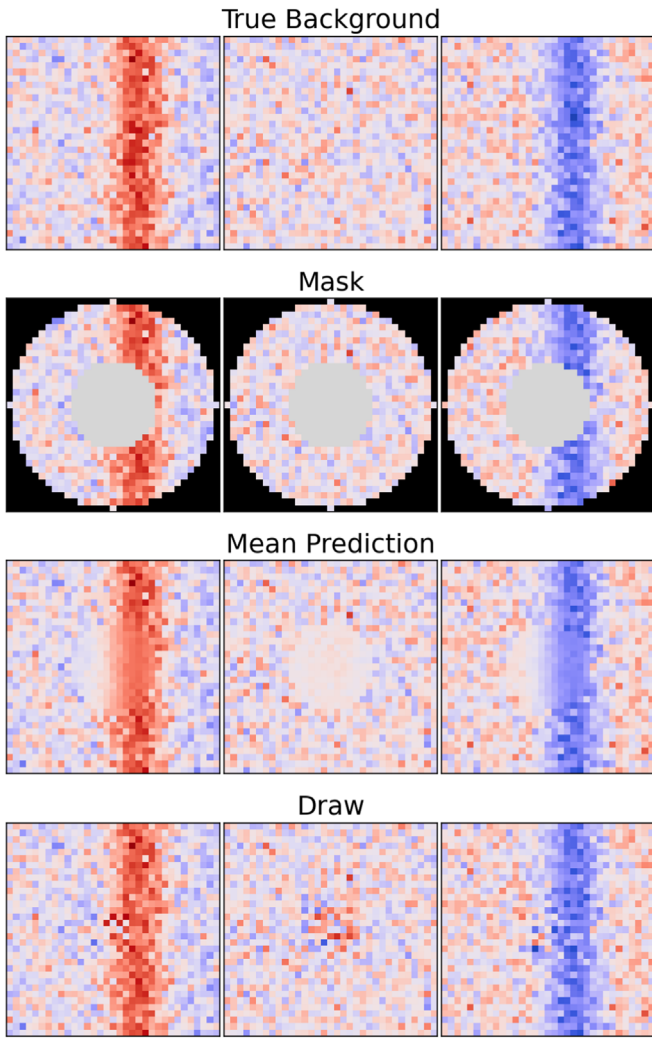


Figure 5. The columns show a stellar position near a positive filament (left), in between the filaments (middle), and near a negative filament (right) in the residuals of a synthetic image (red positive, blue negative). Top row: the true background before source injection is shown. Second row: the mask of *hidden* and *ignored* pixels is shown, as in Figure 4. Third row: the *hidden* pixels are replaced with the mean LPI prediction. Bottom row: the *hidden* pixels are replaced with a draw from the LPI distribution.

deblending than defaults to prevent breaking the filament into nonphysical point sources.⁹

We post-process the *crowdsources* outputs using the code base developed in this work, *CLOUDCOVER.JL*, with $N_p = 33$ and $N_x^{\text{sample}} = N_y^{\text{sample}} = 129$. The top panel of Figure 5 shows the residual image for the true background given the *crowdsources* background model. Each column shows a subimage centered on a star near the positive filament, negative filament, or between filaments. The k_* -pixel mask is obtained by a threshold on the PSF model used for source injection and masks a radius of ~ 10 pixels. Pixels beyond a radius of 16 were *ignored* (see second row of Figure 5). For the third and fourth rows, the k_* -pixels around each star are replaced by the mean prediction or by draws from the predicted distribution.

While the mean predictions reconstruct the filament position within the k_* -pixels correctly, the exact height of the filament

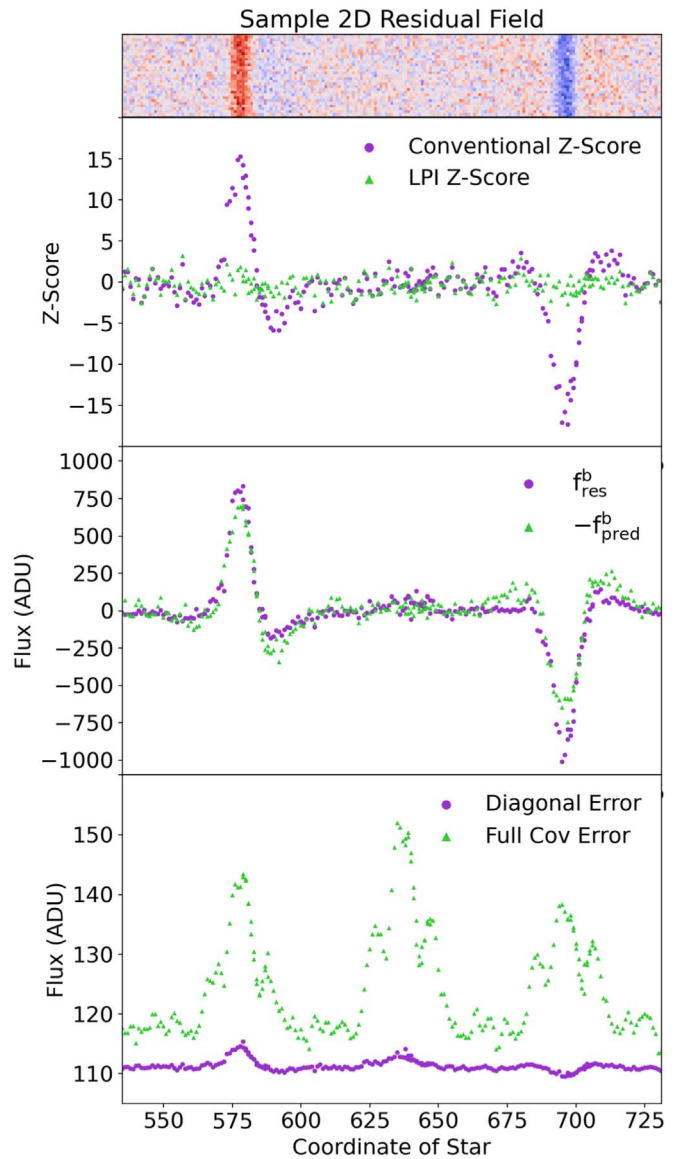


Figure 6. Top panel: residual field (2D) from synthetic image where the x-coordinate matches the 1D stellar coordinate used below. Second panel: Z-scores for injected stars as a function of position using conventional (purple circle) or LPI photometry (green triangle). The absence of positive and negative Z-scores correlated with the filament demonstrates that LPI corrects for the structured background. Third panel: the two debiasing terms, f_{res}^b (purple circle) and $-f_{\text{pred}}^b$ (green triangle), are shown to contribute similarly to correcting for the filaments. Flux is in ADU. Bottom: the predicted flux uncertainty is shown, assuming the covariance matrix is diagonal (purple circle) or using the full covariance matrix (green triangle).

peak is slightly underestimated. We explore the magnitude of this underestimation in Appendix B. The draw rows show interpolation that is almost indistinguishable from the background except for a slightly higher noise level, suggesting a slight overestimation of the uncertainty. However, that is expected. The averaging region for estimating the covariance matrix here is large and includes both filaments, a more heterogeneous environment than any of the regions reconstructed.

The second panel of Figure 6 shows that the Z-scores (observed flux—true flux)/(flux uncertainty) for the *crowdsources* solutions (purple circles) can be as large as $\pm 15\sigma$ and are correlated with the filament position. Note that these

⁹ We set a higher threshold for detection (10σ) and set the *nodeblend* and *sharp* mask bits. See Schlafly et al. (2018) for a detailed description of these parameters.

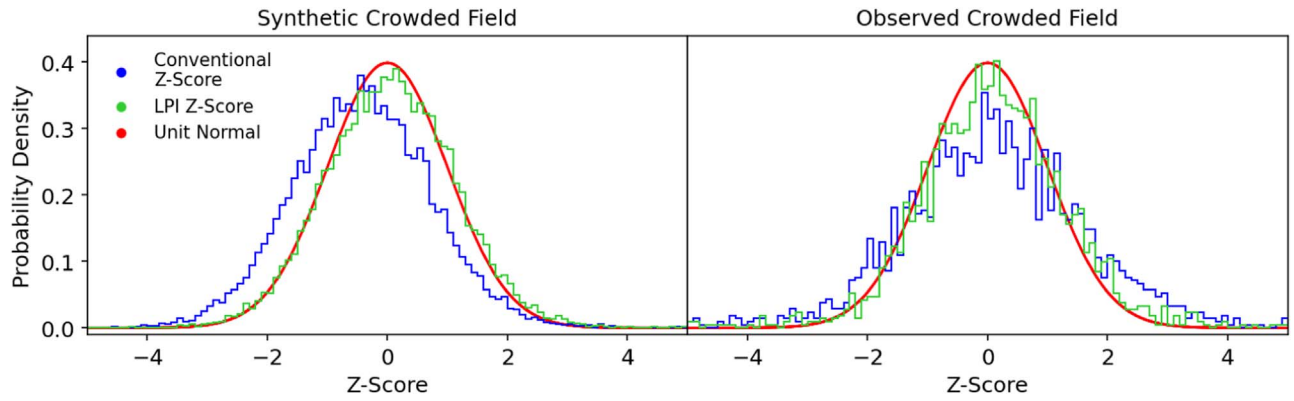


Figure 7. PDF of the conventional Z-scores (blue), LPI Z-scores (green), and a unit normal distribution (red), for reference. Left panel: there is a clear leftward shift in the conventional PDF, indicative of flux underestimation, which is not present in the LPI PDF. Thus, the major improvement of LPI photometry in the synthetic crowded field is correcting bias in the background model. Right panel: the major improvement for the real observed field is correcting the reported flux uncertainty (to be larger). This manifests as the LPI PDF being more peaked and closer to the normal distribution compared to the conventional PDF (which is too broad). In both cases, the improved match between the green and red curves indicates the improved performance of LPI photometry.

conventional Z-scores use uncertainties assuming the covariance matrix is diagonal, only accounting for the uncorrelated Poisson noise. In contrast, the Z-scores that include the bias offset we derive here and use error bars accounting for non-diagonal covariance (green triangles) are flat across the filament positions and are consistent with a standard normal distribution. In the third panel of Figure 6, we can see that the two bias offsets defined in Equation (9) have approximately equal contribution in this case.

One method of quantifying the normality of the Z-score distribution is evaluating the cumulative distribution function (CDF). However, when only a small fraction of the stars in the population is affected, a large local improvement such as the one shown in Figure 6 can appear as only a slight improvement in the CDF. In fact, one of the most common ways of *fixing* error bars in astronomy is to modify the error bars until the CDF matches a normal distribution. The modifications include adding an uncertainty floor, adding a fractional uncertainty (generally ascribed to absolute calibration errors), or simply rescaling the error bars by a scalar (Fitzpatrick & Massa 2009). While these modifications improve the distribution of the entire population, they ignore the detailed local environment of each star, in contrast to LPI photometry. Further, all of those previous distributional approaches will increase the uncertainty estimates for emission or reflection (positive filament) more than for extinction (negative filament).

In the fourth panel of Figure 6, we plot the predicted flux uncertainty assuming the covariance matrix is diagonal (purple circles) and using the full covariance matrix (including off-diagonal contributions, green triangles). For the diagonal error bars, there is a small peak for the positive filament and a small dip for the negative filament, which follows the independent Poisson pixel measurement noise. Thus, no method of rescaling the diagonal error bars can symmetrically treat both the positive and negative filament.

In contrast, the error bars derived from the full covariance matrix peak over both the positive and negative filaments. Away from the filaments on flat sky, the full error bars approach the diagonal approximation, agreeing with the intuition that away from coherent structure, the uncertainty is dominated by the independent per-pixel measurement noise. For stars on flat sky in between the filament, the covariance matrix estimate is trained on the local region containing both

filaments. The predicted uncertainty then displays large off-diagonal contributions. This is of course a limitation of the local covariance model; it can be influenced by local structure, even if that structure is not affecting the star of interest. This limitation drives us to make the number of samples (the size of the boxcar windows) as small as possible while maintaining numerical stability for matrix inversion.

4.2. Toward the Crowded Field Limit

To probe behavior of LPI photometry in the crowded field limit, we inject stars, all of the same flux (~ 18 th g -band magnitude), onto a flat sky background consistent with fields from DECaPS2. We inject 25,000 sources at random integer positions (excluding a buffer around the edge) into a single 2×4 k pixel CCD image. This source density is $10\times$ larger than that required to have isolated PSF footprints (given the 59×59 pixel footprint used by *crowdsources* for stars of this magnitude). We use only one magnitude of star (and with a large signal-to-noise ratio (S/N)) to maximize the probability of close sources being deblended.

Because there is a minimum separation below which two sources cannot be deblended, there is a population of detections that model two close stars as a single star with approximately twice the flux. The larger residuals at the source location resulting from this mismodeling are masked from the LPI covariance matrix estimation, and no large perturbation of the (*good*) k -pixels will indicate that a second star is acting as a *background* to the other. In other words, LPI works only on structured local backgrounds which are not entirely masked by the star. Thus both conventional and LPI Z-score for such sources are large and positive. To cleanly remove merged detections, we exclude stars with separation less than $4/3$ FWHM in the tests that follow (~ 6 pixels).

We display the PDF for sufficiently isolated detections in Figure 7. The PDF for the LPI Z-scores is centered at ~ 0 , while the conventional Z-scores are left shifted. This reflects one of the main biases in *crowdsources*; incompletely deblending sources from the sky background leads to flux underestimation (the left shift), especially in crowded fields.¹⁰ The PDFs having the same width agrees with the intuition that there is little off-

¹⁰ In contrast, other algorithms such as DAOPHOT were found to be biased high by this *confusion* noise (Ferrarese et al. 2000).

diagonal contribution to the flux uncertainties in this synthetic case. Thus, in this case, the main benefit of the LPI Z-scores is to provide a better background estimation, correcting the centering of the PDF.

To make a comparison with a real image, we use a *crowded* field *g*-band image from DECaPS2 (see Section 5) with $\sim 73,000$ detected sources (Figure 7, right panel). We inject 10% more synthetic sources using the same PSF model obtained by *crowdsources* for the original image, sampled from the flux distribution of sources in the original image, excluding sources below a minimum separation of $4/3$ FWHM. In this case, both the conventional and LPI PDFs are well centered at ~ 0 , but the conventional PDF has much larger tails. Thus, the main benefit of the LPI Z-scores on this real field is providing better uncertainty estimates. Compared to the synthetic field, the real field contains more stars ($73,000 > 25,000$), but only a small fraction of them are 18th magnitude or brighter (387) and we expect the leakage of flux from bright sources to more significantly affect the sky estimation. Therefore, both in the limit of poor sky estimation or correlated residuals from mismodeling, the LPI Z-scores improve the overall PDF in the case of crowded fields.

5. Validation on a Large Survey

We now turn to bias offsets and error bars computed on astronomical images from the Dark Energy Camera (DECam) as part of the release of the DECaPS2 (Saydjari et al. 2022). Briefly, DECaPS2 imaged $|b| < 10^\circ$, $5^\circ > \ell > -120^\circ$ in *grizY* filters, finding 34 billion detections (3.3 billion objects). We use DECaPS2 to demonstrate that underestimated error bars correlated with structure are present in most modern surveys of fields with structured backgrounds, our method is robust to the details of real data, and our method is scalable enough to run on all 34 billion detections from one of the largest photometric catalogs to date.¹¹

5.1. CCD Level Processing

Within CLOUDCOVER.JL, we also provide the specific processing choices made to run the DECam images that constitute DECaPS2. For each CCD in an exposure, we use the InstCal products from the DECam community pipeline (CP), which include the image, weight map, and data-quality mask. Since this is an afterburner code, we also take as inputs the *crowdsources* outputs for these images: the position and flux of sources, the background, and background+sources model images, and the position-dependent PSF model.

To create the initial mask for pixels that need infilling (see procedure in Section 2.5), we use the flux of each source and a static PSF model by evaluating the PSF at the center of the CCD. Then, pixels around the stellar position are masked if the $\text{flux} \times \text{PSF} > \text{thr}$, where *thr* is a chosen threshold. This threshold was chosen to mask residuals in the core of bright stars resulting from mismodeling. For *crowdsources* models of DECaPS2 images, we found $\text{thr} = 20$ ADU acceptable for all bands and fixed this parameter for all exposures.

In addition, any pixel with a nonzero data-quality bit from the CP (cosmic-rays, bleed trails, satellite trails, etc.) were masked. Finally, we masked pixels with residuals larger than $\text{outthr} = 20,000$ ADU that were previously unmasked. These generally resulted from distorted PSFs near the edge of CCDs that were not captured in the per-star masking (see Appendix A). In the full DECaPS2 run, less than 1% of the images had to use the infill fallback of the image median for some subset of the pixels. After infilling, we add Poisson noise to each pixel in the mask as in Equation (7), where the gain used is that estimated from *crowdsources* as the median of the weight map \times image.¹²

To handle modeled sources near or beyond the edge, the image is padded using reflective boundary conditions. Then the image is multiplied with a shifted version of itself, averaged with a boxcar mean ($N_x^{\text{sample}} = N_y^{\text{sample}} = 129$), and stored as in Equation (6). For each source in the image, the covariance matrix for an $N_p = 33$ subimage is then constructed out of elements of those image products. The pixels in that subimage are then classified as *good* (k -pixels), *hidden* (k_* -pixels), or *ignored* based on the PSF for a source at that location and the flux. All pixels in the PSF above *thr* (same value as during infill) are selected as *hidden*. If the star has flux $< 10,000$ ADU, it is masked as if the flux were 10,000 ADU in order to ensure that the background behind every star is predicted.

Pixels masked by data-quality flags or outside a radius of 16 pixels that are not already *hidden* are *ignored*, and all remaining pixels are selected as *good*. By making the choice of $N_p = 33$, we implicitly impose a cutoff on the brightest star we can model well. For stars so bright that the whole N_p subimage is *hidden*, for example, there are no *good* pixels to use for the conditional prediction. However, the absolute uncertainty on flux calibration is usually the dominant term in the error budget for bright stars. So if the number of *good* pixels is too small, we compromise. The mask of all pixels beyond radius 16 is removed and the row of pixels around the outer edge of the subimage are forced to be *good* so there is at least a 1-pixel deep boundary for conditioning. This fallback changes the value of our quality flag *dnt* from 0 to 2.¹³ A table of *dnt* bit values is available in Section 7.

The PSF models from *crowdsources* can occasionally contain negative pixels, though usually only for non-photometric exposures.¹⁴ To handle these cases gracefully, the min/max ratio of the PSF is computed and if the PSF is negative, bit 3 of *dnt* is set. To quantify the negative portion of the PSF, bits 4 and 5 are set if the ratio is $< -1 \times 10^{-3}$ and -1×10^{-1} , respectively. In all cases, the absolute value of the PSF is used by CLOUDCOVER.JL.

The last step before computing Equations (9) and (10) is adjusting the per-pixel variance (the diagonal of the local covariance matrix) to account for the Poisson noise from the presence of the bright star. The local covariance matrix was originally estimated only on pixels that were representative of uncontaminated background and thus only have Poisson noise on the order of the background sky counts. Thus we add the background+sources model minus the background model to the diagonal. Since covariance matrices must be symmetric

¹¹ DECaPS2 with 3.32 B sources is comparable in size to Pan-STARRS1 with 2.9 B sources (Magnier et al. 2020), though DECaPS2 has fewer epochs and thus fewer detections. The second data release of the NOAO source catalog (NSC) (Nidever et al. 2021), which used SEXTRACTOR to uniformly reprocess public data including all of DECaPS1 and DECaPS2, contains 3.9 B. The Zwicky Transient Facility (Data Release 8) reports ~ 1.5 B objects (Masci et al. 2019; T. Price 2022, private communication).

¹² Random steps use exposure date-time seeds for reproducibility.

¹³ For heavily saturated stars, neither sky subtraction nor debrending noticeably change the predicted flux.

¹⁴ Before any photometric cuts, these detections comprise less than 0.02% of the total number of detections for DECaPS2.

Table 1
CLOUDCOVER.JL OUTPUT

Variable	Description
dflux	Flux uncertainty
dflux_diag	Diagonal flux uncertainty
cflux	LPI corrected flux
fdb_res	Bias offset from residuals
fdb_pred	Bias offset from predicted background
cchi2	Good pixel χ^2
kcond0	Initial number of <i>good</i> pixels
kcond	Number of <i>good</i> pixels
kpred	Number of <i>hidden</i> pixels
dnt	Quality flag on solution

positive semidefinite, we must never add a negative to the diagonal during this procedure. However, negative PSF models or sources modeled with negative flux can make this difference negative in rare cases, so we add the absolute value.

As a final quality-assurance check, we compute the χ^2 for the *good* pixels given the local covariance matrix estimate. This provides a self-consistency check on the quality of the local covariance matrix estimate. Since the *good* pixels were used for conditioning the prediction on the *hidden* pixels, assuming the local covariance matrix was a decent model, large χ^2 (relative to the number of degrees of freedom) act as a red flag.¹⁵ The output catalog columns are described in Table 1. The number of *ignored* pixels is obtained as $N_p - k_{\text{cond}} - k_{\text{pred}}$.

5.2. Nebulous Region

First we consider a *r*-band image of a highly structured H II region (CED 116 near Lambda Centauri) in the survey footprint (Figure 8). We then create many (2000) new images by injecting synthetic sources using the PSF model derived during the *crowdsourc* fit of the image. Each injection image consists of a number of synthetic sources that corresponds to 10% of the number of sources originally found so as to not significantly perturb the image. These sources are also sampled from the distribution of fluxes in the original solution to the image, excluding sources with *bad* flags set or brighter than 17th magnitude in the *g* band. The injected source-center positions are uniformly random, excluding $N_p = 33$ pixels around the edge of the CCD. The injection images are reprocessed by *crowdsourc* as if they were standard images so we can compute Z-scores for the photometric solutions relative to the injected fluxes.

By plotting the residuals (Figure 9, top) and the original Z-scores (Figure 9, middle), we visualize a clear spatial correlation between the sign and magnitude of the Z-scores and the residuals of the *crowdsourc* models. These correlations between structured backgrounds and biased photometry with underestimated error bars significantly complicate large-scale inference from photometry in the Galactic plane, such as for efforts mapping the interstellar dust distribution (Green et al. 2019). After adding our flux bias offset and using the error bars that include off-diagonal contributions, we obtain the bottom panel of Figure 9 where there is no clear correlation of the Z-scores with the residuals and the excessive population of

¹⁵ In Section 3.1, we observed that the Gaussian model need not be a good representation of the overall population. Thus, we cannot expect these values to follow a χ^2 distribution exactly. However, values that are orders of magnitude too high likely indicate numerical instability.



Figure 8. Color image (RGB = *zrg* bands) of a larger region where H α emission complicates photometry. The square marked by four white corners indicates the subimage used in Figure 9.

$>\pm 3\sigma$ outliers has been removed. We take this as strong evidence in support of using our post-processing code with other/future photometric pipelines.

5.3. Correction Correlations

We also want to confirm that using the LPI bias offset does not strongly affect regions without structured backgrounds. Since we have already shown that LPI improves over the conventional PDF (see Figure 7), we will examine the per-star percent changes. This allows us to confirm that LPI is not creating large flux errors in the process of debiasing. To do this we use injection tests as in Section 5.2, with an additional requirement that the *dnt* quality flag is zero.

The first panel of Figure 10 shows the per-star percent change versus the percent error in a field of typical stellar density ($\sim 14,000$ sources) without structured background.¹⁶ The maximum percent change is 20% due to the bias offset. If the bias introduced by structured backgrounds were the main source of error, we would expect the per-star corrections to cluster along the 1:1 line (black dashed). However, the distribution appears flat and parallel to the *x*-axis. This agrees with the intuition that in uncrowded regions, errors from structured backgrounds are subdominant (to deblending errors for example).

In the second panel, we compare to a crowded field ($\sim 195,000$ sources). The scatter introduced by the bias offset is larger, up to 50% in the most extreme cases, but the majority of stars are clustered near the origin. Only a slight positive trend is observed and might be attributed to the larger residuals in crowded fields. Finally, in a nebulous field (specifically the *g* band shown in the top panel of Figure 9), we see a strong positive correlation between the correction factors and the conventional percent differences, as expected. Thus, we correct

¹⁶ The classification of a region as having a structured background is somewhat subjective. In the context of DECaPS, this is automated using a neural network to flag regions as *nebulous* or not. (Schlafly et al. 2018).

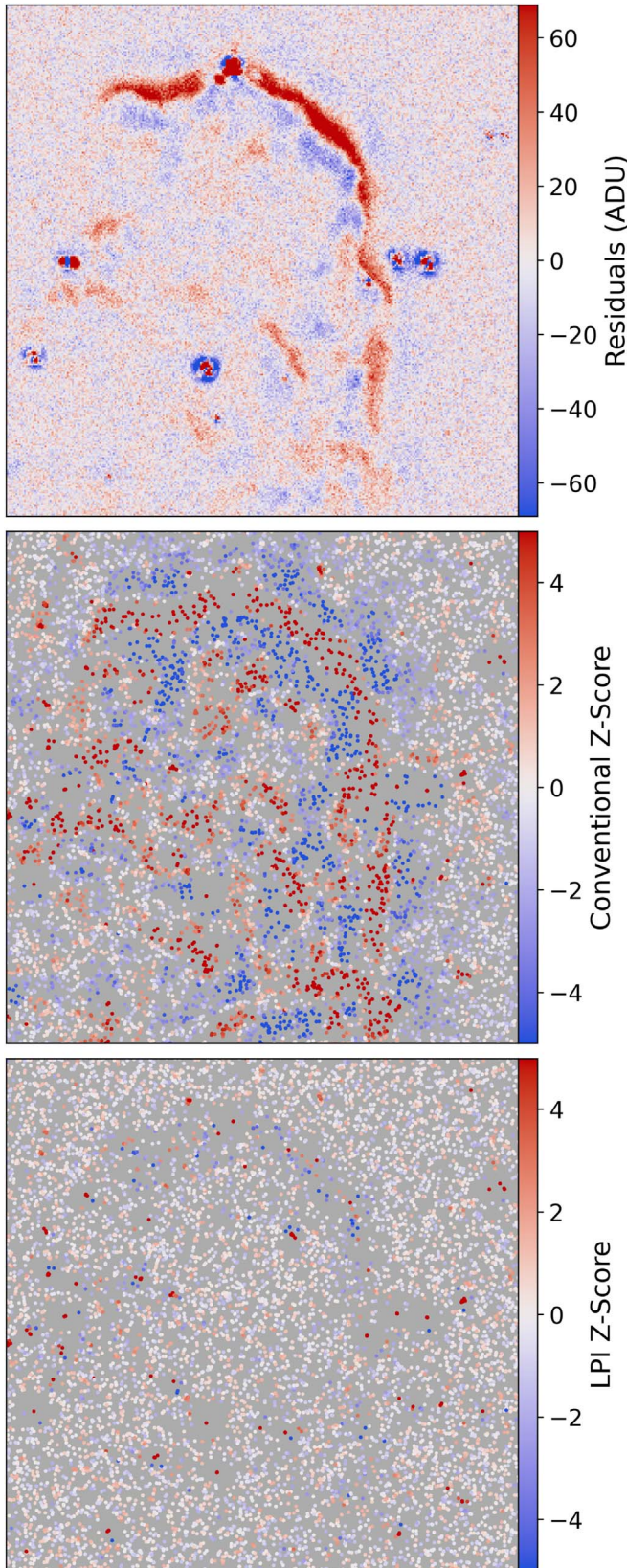


Figure 9. Top panel: residuals from `crowdsource` model of the central r -band region of Figure 8. Middle panel: Z-scores for injected sources as a function of position for the image above. Clear spatial correlation between the sign and magnitude of the Z-scores and residuals is observed (correlation between top and middle panels). Bottom panel: after LPI, the large population of $|Z\text{-scores}|$ larger than 3 are removed (the scatter points are whiter) and little spatial correlation is observed.

for bias resulting from structured background when present, but do not introduce bias into the catalog on either flat or crowded regions.

Another metric to examine is the following ratio:

$$\sigma\left(\frac{f^b}{\text{dcflux}}\right) / \sigma\left(\frac{\hat{f} - f}{\text{dcflux}}\right). \quad (11)$$

If the numerator is small, then the change in flux from the bias offsets is small relative to estimated uncertainty. We compare to the standard deviation of Z-scores to account for the fact that the distributions of uncorrected flux estimates for the three cases in Figure 10 have different widths. For the uncrowded, crowded, and nebulous regions, this metric is 0.23, 0.76, and 0.95, respectively. In the uncrowded field, this further demonstrates that the bias offsets do not significantly change the flux estimates relative to the flux uncertainties. And in the nebulous field, the bias offsets are large, as expected.

5.4. Full Survey Footprint

Zooming out, we evaluate the performance of the LPI error bars on the entire survey footprint (2700 deg², 6.5% of the sky). In the ideal case, we expect the Z-scores to be normally distributed when the uncertainties are correctly estimated. However, biases and failure modes in photometric pipelines introduce outliers. We measure the interquartile range (IQR), which is robust to outliers, for both the conventional and LPI Z-scores and compare this to the ideal value for a Gaussian distribution of 1.349. Then, we report by what scalar the error bars must be multiplied to achieve the IQR of a normal distribution, shown in Figure 11 ($N_{\text{SIDE}} = 32$). To do this, we use the injection tests (with the same criterion as before) and only analyze Healpix pixels with >30 stars.

In the conventional case, the rescaling required is over a factor of 5 toward the Galactic bulge. In contrast, the rescaling for the LPI error bars is only a factor of 2 toward the Galactic bulge. Comparing the median rescaling factors, we have 1.17 and 1.02 for the conventional and LPI error bars, respectively. To have a correct central value for the rescaling (1.0), there must be (and are) cases where the rescaling is less than 1. So the ideal case for Figure 11 is on average white with small red and blue fluctuations, as is seen for the majority of the LPI panel. Deviations from ideality toward the Galactic bulge likely result from the large masking fraction in this region of high stellar density, which impedes estimation of the residual covariance.

While it may be tempting to simply rescale the conventional estimates and then avoid the computational cost associated with LPI, a simple rescaling of error bars will never be able to remove the spatial correlations shown in Figure 9. Just as in the conventional case, these rescalings indicate a gap between the model complexity and reality that requires further method development to close. In the meantime, we leave it to users to decide whether to rescale the LPI uncertainty estimates.

We can evaluate relative metrics, such as the ratio between the stated flux uncertainty from LPI and `crowdsource`, with higher angular resolution (Figure 12, $N_{\text{SIDE}} = 512$).¹⁷ The

¹⁷ We report an outlier-clipped mean value of the ratio for stars within each Healpix pixel. The outliers are beyond $10 \times \text{IQR}/1.349$ and are computed per CCD. In addition to requiring $\text{dnt} = 0$, we impose $\text{cchi2}/\text{kcond} > 0.7$ from the `CLOUDCOVER.JL` flags. We also require no *bad flags* (which removes detections on CCD S7), $\text{rchi2} < 100$, and $\text{fracflux} > 0.5$ on the `crowdsource` outputs.

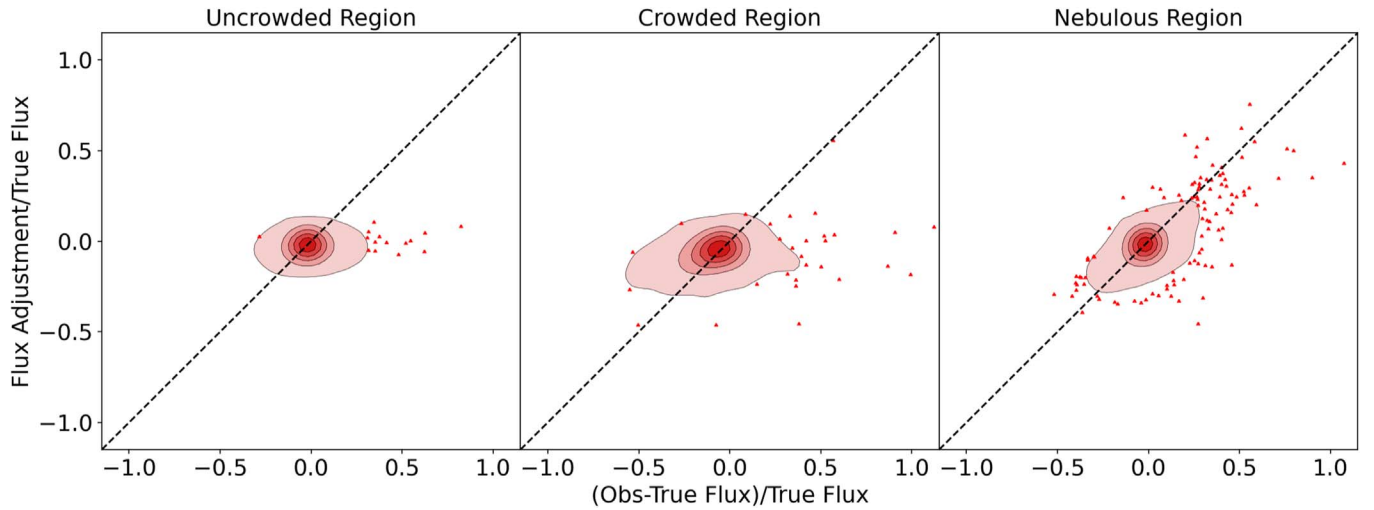


Figure 10. The ratio of f^b over true flux vs. the fractional error in flux estimation for injected sources using `crowdsourcing`. The colored contours represent density for the 2D histogram, and sources beyond the lowest contour are shown as scatter points. Sources that fall along the 1:1 line (black dashed) are correctly debiased by LPI, points that fall along the $y = 0$ line are misestimated by `crowdsourcing` and not corrected, and points that fall along the $x = 0$ line are incorrectly modified by LPI. Left panel: almost no correlation is observed for typical uncrowded fields and no outliers are introduced by the debiasing. Middle panel: for crowded regions, a slight correlation is observed and, while there is larger scatter, no outliers are introduced by debiasing. Right panel: for the nebulous region, there is a strong positive correlation (contours follow the 1:1 line), showing that the LPI analysis improves the flux estimate.

main feature is a large (relative) increase in the LPI error bars toward the Galactic bulge. This correlation with stellar density is consistent with intuition developed for the crowded-field test case (Figure 7). In addition, the peak associated with the Carina Nebula ($b = -0.8^\circ$, $\ell = -73^\circ$) and filaments associated with the Vela supernova remnant ($b = -2.8^\circ$, $\ell = -96^\circ$) shows that the LPI errors are sensitive to structured backgrounds as intended. More detailed analysis of these and other metrics over the survey footprint will be part of the DECaPS2 release (Saydjari et al. 2022).

6. Conclusion

In this work, we have derived a correction to both the estimated flux and flux uncertainties by using LPI to estimate the background behind a star. We showed that these corrections improve photometry on synthetic and real fields with structured backgrounds. In particular, we applied our method to all 34 billion detections in the DECaPS2 survey, demonstrating scalability and improvement across the entire survey footprint. In contrast to the commonly used method of rescaling error bars to obtain the expected chi-square, we showed that LPI adjusts uncertainty estimates according to the local background covariance, and correctly treats both positive and negative residuals. In the crowded limit, we found both improved estimation of contamination of the sky background by unmodeled stellar flux and uncertainty estimates from large residuals. We showed that LPI reduces to a conventional uncertainty (diagonal covariance) on fields without structured backgrounds, allowing for uniform processing of an entire photometric catalog. In addition, we point out that LPI applies to a broader class of images beyond Gaussian fields, even capturing highly non-Gaussian structures such as isolated filaments and clouds. We demonstrate this both analytically and numerically on a real dust image. Finally, we release our code base, which is a pure afterburner to any photometric code that outputs residual images and a PSF model. Thus, we suggest that this method can be easily used to create

value-added photometric catalogs for existing and future surveys in the Galactic plane.

While we restricted this work to point sources, the same technique should be extensible to correcting estimates of the flux from more extended sources, such as galaxies. In that case, the shape function for the extended source (i.e., the PSF-convolved Sérsic profile for a galaxy) takes the role of the PSF in the above discussion. Future directions include demonstrating that these corrections improve downstream inference, e.g., Bayesian fitting of photometry to jointly predict intrinsic stellar parameters (temperature, age, etc.) and foreground dust densities as a result of reddening. We expect that the improved flux uncertainties should expand the set of well-modeled sources. This is especially important for dust mapping in the Galactic plane, where there is significant extinction, and the stars in fields with structured backgrounds are precisely those most constraining for the dust clouds and filaments of interest.

7. Code and Data Availability

Data products and JUPYTER notebooks required to reproduce figures within the paper are publicly available on Zenodo at doi:10.5281/zenodo.5809521 (11.0 GB). We make the code base used here publicly available on GitHub CloudCovErr.jl¹⁸ and used v0.9.0 for this work. The data products associated with the full DECaPS2 run will be described in more detail and accompany the survey release paper.

A.S. gratefully acknowledges support by a National Science Foundation Graduate Research Fellowship (DGE-1745303). D.F. acknowledges support by NSF grant AST-1614941, “Exploring the Galaxy: 3-Dimensional Structure and Stellar Streams.” This work was supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions). We would like to thank Edward Schlafly and Lucas Janson for feedback during the development of this

¹⁸ <https://github.com/andrew-saydjari/CloudCovErr.jl>

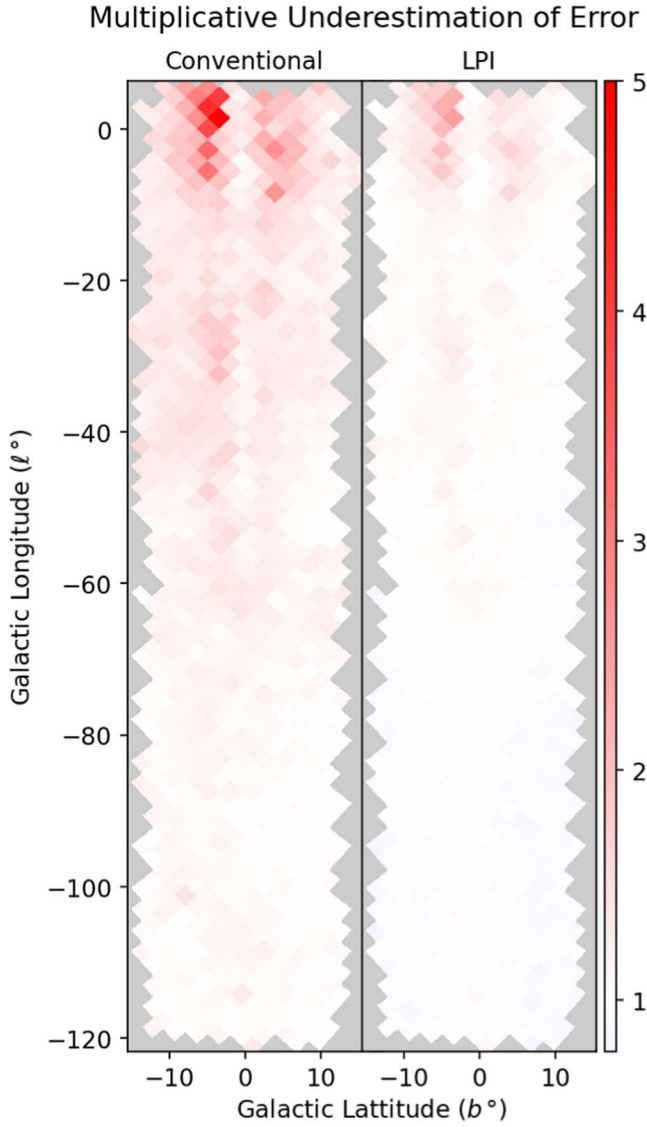


Figure 11. Factor by which the estimated flux uncertainty is too small for conventional (left panel) and LPI photometry (right panel). In the ideal case, this factor would be 1 (white) everywhere. The multiplicative factor is computed as $\text{IQR}/1.34896$ for the Z-scores of injected sources, where the denominator is the IQR for the standard normal distribution. The map is over the survey footprint of DECaPS2 (34B detections) at $N_{\text{SIDE}} = 32$. LPI photometry is a clear improvement, though correction toward the Galactic bulge is imperfect.

project and writing of this manuscript. We acknowledge Eugene Magnier, Josh Speagle, Catherine Zucker, Justina R. Yang, Nayantara Mudur, and Christopher Stubbs for helpful discussions. A.S. acknowledges Sophia Sánchez-Maes for helpful discussions and much support. Computations in this paper were run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University. We acknowledge the developers and maintainers of the Julia ecosystem, from whom we have learned and benefited greatly.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) Collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities

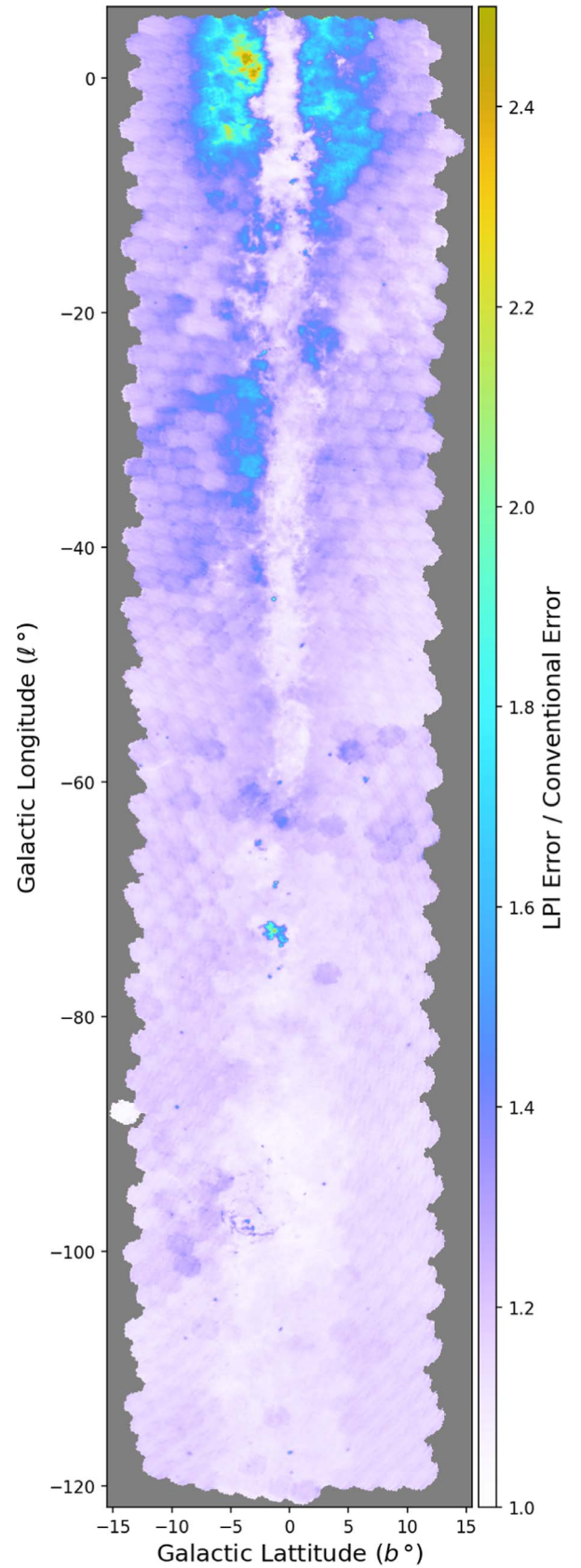


Figure 12. Ratio of uncertainty estimates for LPI relative to conventional photometry over the full survey footprint of DECaPS2 (34B detections). A value of 1 indicates LPI and conventional photometry agree on the uncertainty estimate. Features associated with the Galactic bulge, Carina Nebula ($b = -0.8^\circ$, $l = -73^\circ$), and Vela supernova remnant ($b = -2.8^\circ$, $l = -96^\circ$) stand out.

Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgicas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the NOIRLab (formerly known as the National Optical Astronomy Observatory), the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

Software: GNU Parallel (Tange 2021), ASTROPY (Astropy Collaboration et al. 2013), HSPY (Collette 2013), IPYTHON (Perez & Granger 2007), MATPLOTLIB (Hunter 2007), NUMPY (van der Walt et al. 2011), SCIPY (Virtanen et al. 2020).

Appendix A Computational Details

In our LPI photometry implementation (see Section 7), for a field of average stellar density in DECaPS ($\sim 10,000$ sources/8 megapixels) the majority of the computational time is split evenly between element-wise image multiplication (Hadamard products) and per-star Cholesky decomposition. We choose to construct the local covariance matrix from elements of the boxcar mean of the initial image multiplied by a shifted version of that image (Equation (6)). For 2D, this costs $\sim 4N_p^2 N_x N_y$ FLOPs in contrast to constructing V^{TV} via matrix multiplication per star which is $\sim 2N_{\text{star}} N_p^4 N_x^{\text{sample}} N_y^{\text{sample}}$. Here N_x and N_y are the size of the original image and N_x^{sample} and N_y^{sample} are the size of the boxcar window in each direction.

There can be significant memory requirements to holding those image products in storage for use in the per-star covariance matrix construction. The alternative consists of grabbing the needed elements of the product images for every star at each shift, and throwing away the product image. However, this alternative procedure requires highly variable amounts of memory that is strongly dependent on the number of stars in the exposure. We believe in the context of large survey pipelines, stable memory requirements are highly desirable as they maximize the CPU utilization.

In order to not store $\sim 2N_p^2 N_x N_y$ Float32s during the processing of each CCD, ~ 70 GB for DECam with $N_p = 33$ as in Section 5, we break each image into tiles and process each tile and the stars within it separately. Doing so incurs an overhead cost since we have to include some padding beyond the furthest stellar position in a given tile. This tiling procedure allowed us to run the full DECaPS2 survey with 3.6–4.8 GB per exposure (median 3.7 GB) in 203 k core hours. The jobs were run on the FASRC Cannon cluster at Harvard University on an entire compute node with two water-cooled Intel 24-core Platinum 8268 Cascade Lake CPUs and 192 GB RAM running 64 bit CentOS 7. The runtime versus number of stars for the DECaPS2 run demonstrates high linearity with respect to the number of stars per exposure—approximately 15.3 core-ms/star in addition to a constant per exposure (~ 62 CCDs) of 3.4 core-h (see Saydjari et al. (2022), for more details).

The major failure mode for the code presented in this work occurs when numerical effects cause the estimated covariance matrix to not be positive definite, which must be true theoretically for any covariance matrix (and prevents Cholesky factorization). The main sources of instability we find are unmasked bright pixels which are not well captured by the `crowdsources` model. These outliers prevent a positive definite estimate of the covariance matrix and should be masked anyway as they do not represent the true background.

In DECaPS2, we reduced `outthr` to 5000 and reran the 980 CCDs (out of 1.32 million CCDs), which contained stars without positive definite covariance estimates.¹⁹ Most of these resulted from bright stars not fully masked by our procedure because `CROWDSOURCE` modeled them as multiple fainter sources. Afterward, only nine CCDs in the whole survey contain stars with nonpositive definite covariance estimates, and the majority of those result from large pupil ghost artifacts. In these cases where the source and sky model coming from `crowdsources` and the mask from the DECam NOAO community pipeline are all wrong, we simply leave all `CLOUDCOVER.JL` outputs as NaN as a warning to users. Using Float64 instead of Float32 (as we have done in all presented results) does enhance the numerical stability, but at too large a computational and memory cost.

Looking toward other large survey applications targeting the Galactic plane, such as Rubin and Roman, we comment that the code provided in Section 7 is not fully optimized. In several places, memory is re-allocated (which is inefficient) or the minimum memory footprint could be reduced by restructuring the code. Thus, the benchmarks above are not meant to demonstrate the optimality of the code, merely that estimating a covariance matrix and Cholesky factorizing it for 34 billion detections can be done faster than the original photometric solutions in crowded fields. For instance, the `crowdsources` solutions for DECaPS2 required 230 k core hours (13.1 kpix s^{-1}). Of course, a GPU-based implementation could see significant speedups on the limiting steps of element-wise image multiplication and Cholesky factorization (Naumov 2021), but large surveys will have to weigh the relative cost of GPU/CPU resources to the relative speedups.

¹⁹ Most of these cases were resolved during the reprocessing of DECaPS2 after rerunning the CP on images with saturation thresholds that were set too low (Saydjari et al. 2022).

Appendix B

Role of S/N and Aspect Ratio

It is important to understand how the quality of the LPI of the filament is a function of the S/N of the data, local environment used to estimate the covariance matrix, and number of training samples. To do this, we mask and interpolate 34 equally spaced positions centered along a synthetic filament with a masking radius of 5 pixels and $N_p = 33$. We report the average of the row of (11) pixels along the filament spine, pooling over all 34 reconstructions, as a fraction of the true filament height (Figure 13).

For what we call a sample aspect ratio (\log_2) of 0, we learn the covariance matrix over a symmetric region, using a boxcar width of 129 pixels in both the x and y directions ($N_x^{\text{sample}} N_y^{\text{sample}} = 2^{14}$) and obtain a reconstruction of $>80\%$ of the true peak value at an S/N ~ 1.4 . Here S/N is simply the peak height divided by the Poisson uncertainty for the total (peak and background) counts. As S/N increases, the predicted peak height converges to the true value, further validating the application of the LPI method to non-Gaussian structures. In the low S/N limit (S/N ~ 0.1), the noise dominates the

reconstruction and there is only a small positive bias due to the filament, as expected.

Instead of using a symmetric sampling of the x and y directions, we can consider more samples along the filament direction (positive sample aspect ratio) or transverse to the filament direction (negative sample aspect ratio). This amounts to using a boxcar width of (255, 65) for sample aspect ratio 2, given the filament along the x direction. The reconstruction converges more quickly the more samples are along the filament direction as opposed to the transverse direction. This is consistent with the intuition that in order to reconstruct the filament, the conditioning must overcome a large bias in the model for a flat background when a large fraction of the training sample is flat (as it is in the transverse direction relative to the filament). For numerical stability we must truncate the aspect ratios in the positive direction before there are $< 2N_p$ transverse samples because the image is translationally invariant (other than Poisson noise) along the filament direction. Increasing the overall number of samples (circle markers, Figure 13) improves numerical stability, but the main effect on convergence is the additional samples transverse to

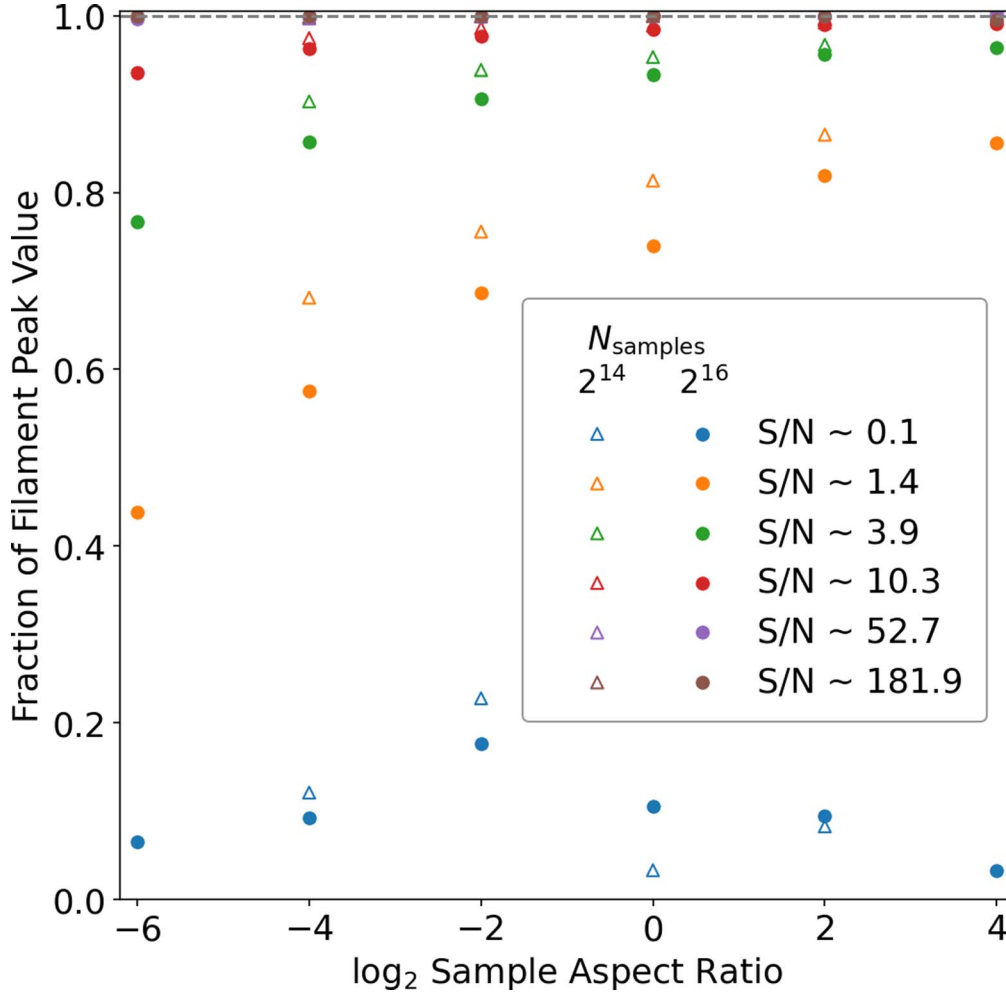


Figure 13. Convergence of the predicted filament height as a function of S/N. For S/N ≥ 1 , a large fraction of the filament height is reconstructed. To the right along the horizontal axis, the fraction of samples (used for estimating the covariance matrix) along the filament increases. The total number of samples is fixed to either 2^{14} (open triangle) or 2^{16} (closed circle) by changing the total area of the boxcar window. The main effect on convergence of increasing sample size is the additional samples transverse to the filament, causing an apparent shift of +2 along the aspect ratio axis with respect to the triangle markers. The ability of LPI to reconstruct the peak height thus agrees with the intuition that the more training data resembles the problem, the better the performance.

the filament (an apparent shift of +2 along the aspect ratio axis with respect to the triangle markers). In both cases, the ability of LPI to reconstruct the peak height agrees with the intuition that the more training data resembles the problem, the better the performance.

Appendix C

Pixel Space Linear Least-squares Notation

To reformulate our discussion in the language of a linear least-squares regression in pixel space, we view the *good* pixels (k) as the regressors for the value of the *hidden* pixels (k_*), the dependent variables. Let X and Y be matrices of the regressors and dependent variables, respectively, where each row represents a different sample (obtained by pixel shifts over the local region) out of a total of N samples. In what follows, we will ignore centering of the data for clarity. Then, the predicted linear parameters are given as usual,

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \quad (C1)$$

The prediction of the *hidden* pixels (\hat{y}) given the observed *good* pixels for a given stamp (x_0) is then

$$\hat{y} = \hat{\beta}^T x_0 = Y^T X (X^T X)^{-1} x_0. \quad (C2)$$

The covariance of this prediction is just

$$\begin{aligned} \mathbb{E}[(y - \hat{\beta}^T x)(y - \hat{\beta}^T x)^T] \\ = \frac{1}{N-1} (Y^T Y - Y^T X (X^T X)^{-1} X^T Y), \end{aligned} \quad (C3)$$

where the expectation value is estimated over the N samples. Then the correspondence to the Gaussian process regression notation used in the main text is clear after identifying $x_0 = X_k$, $\hat{y} = \bar{X}(k_*)$, and partitioning V into k and k_* sub-blocks so that we can see

$$\begin{aligned} C &= \frac{1}{N-1} V^T V = \frac{1}{N-1} \begin{bmatrix} X^T \\ Y^T \end{bmatrix} \begin{bmatrix} X & Y \end{bmatrix} \\ &= \frac{1}{N-1} \begin{bmatrix} X^T X & X^T Y \\ Y^T X & Y^T Y \end{bmatrix} = \begin{bmatrix} C_{k,k} & C_{k,k_*} \\ C_{k_*,k} & C_{k_*,k_*} \end{bmatrix}. \end{aligned} \quad (C4)$$

Appendix D

Approximate Stationarity of LPI Covariance

To show the approximate stationarity of the covariance matrix estimated in LPI, we consider the reconstruction of a portion of the WISE 12 μm dust map (Section 3.2). This test represents an extreme in heterogeneity of the local background used for estimating the local covariance matrix. In Figure 14 (left panel) we show the covariance of the central pixel with all other pixels in the 49×49 pixel subimage. In Figure 14 (middle panel), we show the mean value of the covariance as a function of pixel displacement for all pairs of pixels in the subimage. Note that the central pixel covariance strongly resembles the mean covariance as a function of displacement, though there are some small distortions. More quantitatively, in Figure 14 (right panel) we show the standard deviation of the covariance between pixel pairs with the same displacement normalized by the median value of the diagonal (pixel-wise variance). This demonstrates that the covariance function is stationary at the $\sim 10\%$ level outside the core, and $\sim 25\%$ level in the core.

The mean covariance as a function of displacement is highly anisotropic, agreeing with our intuition that anisotropy is required to handle filamentary structures. A more rigorous justification of stationarity could use a covariance matrix based only on the mean covariance as a function of displacement to perform the reconstructions as in Figure 4 and determine if there is any measurable difference in the reconstructions compared to our LPI procedure. Further, except for low amplitude structure near the edges of Figure 14 (middle panel), the covariance function appears to decrease rapidly as a function of radius. This suggests that the covariance function might be well approximated by a function such that the covariance matrix is positive definite when sampled at an arbitrary points (Bochner's theorem). Whether or not exact stationarity can be enforced or a well-behaved kernel function approximation found while maintaining sufficient flexibility to describe the diversity of structure in the ISM remains an exciting question for future work.

In Figure 15 we show a small sample of the diversity in covariance structure obtained in different regions of the DECaPS2 survey footprint (all r -band images): a nebulous region (left column), a dense field in the Galactic bulge (center column), and a low density field at high Galactic latitude (right

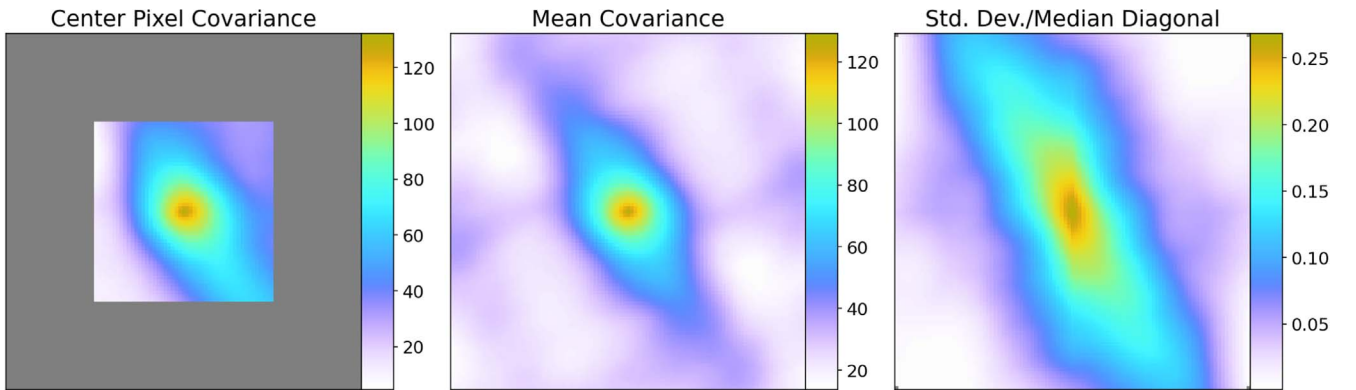


Figure 14. Covariance function for reconstruction of a portion of the WISE 12 μm dust map. Left panel: covariance of the central pixel with all other pixels in the 49×49 pixel subimage. Middle panel: mean value of the covariance as a function of pixel displacement for all pairs of pixels in the subimage. Right panel: standard deviation of the covariance between pixel pairs with the same displacement normalized by the median value of the diagonal (pixel-wise variance). This demonstrates that the LPI covariance function is stationary at the $\sim 10\%$ level outside the core, and $\sim 25\%$ level in the core.

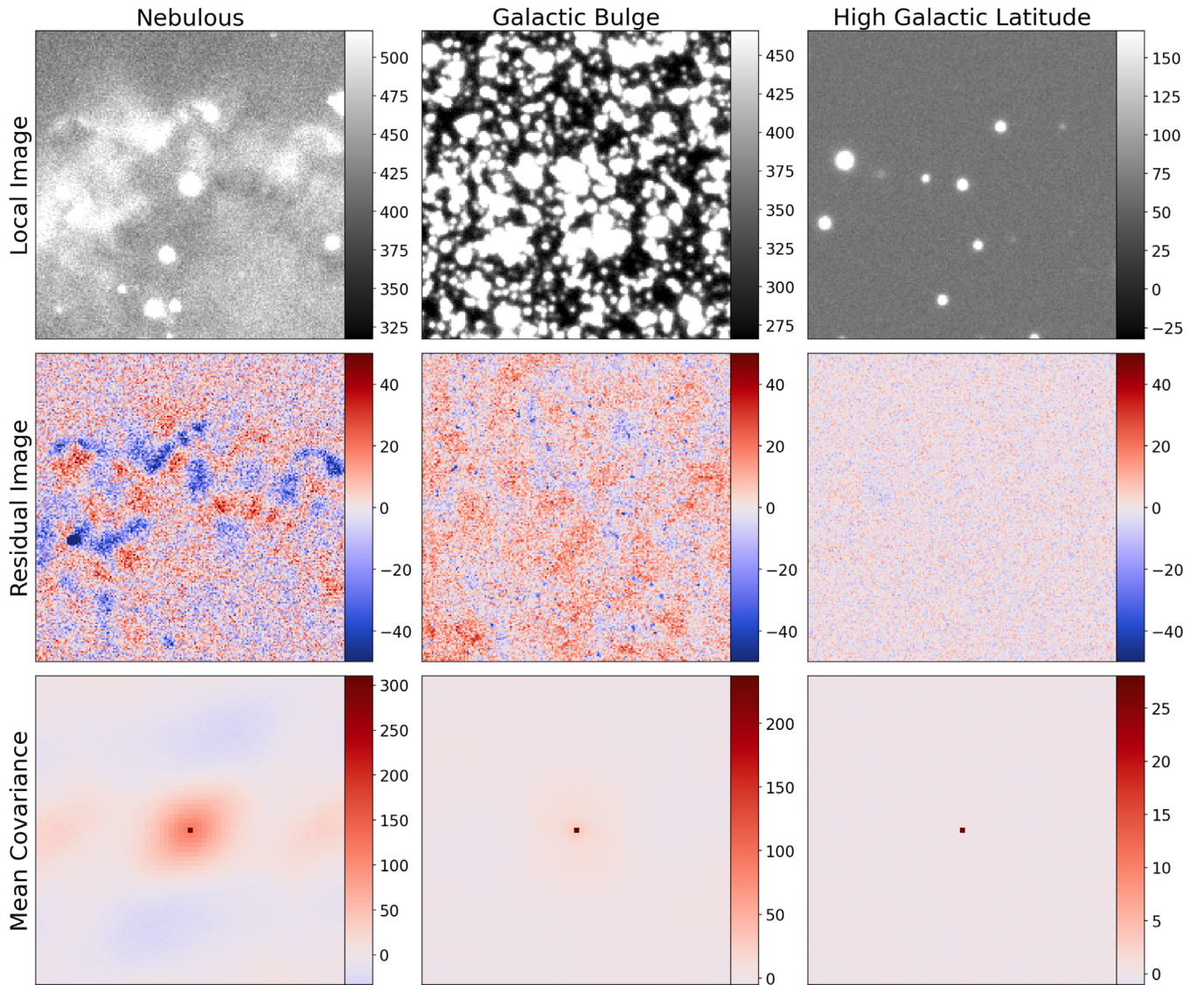




Figure 15. Diversity in covariance structure obtained in different regions of the DECaPS2 survey footprint (all r -band images): a nebulous region (left column), a dense field in the Galactic bulge (center column), and a low density field at high Galactic latitude (right column). In the top row, the raw image of the star and surrounding field are shown. Each image is stretched ± 50 around its median. In the center row, we display the `crowdsourced` residuals after infilling masked pixels as described in Section 2.5. The image spatial scale is chosen to include all training data used to estimate the local covariance matrix of the unmodeled background structure. In the bottom row, we display the mean stationary approximation (as in Figure 14, middle panel) to the covariance of the residuals used in our LPI procedure.

column). In the top row, the raw image of the star and surrounding field are shown. Each image is stretched ± 50 around its median. In the center row, we display the `crowdsourced` residuals after infilling masked pixels as described in Section 2.5. The image spatial scale is chosen to include all training data used to estimate the local covariance matrix of the unmodeled background structure. In the bottom row, we display the mean stationary approximation (as in Figure 14, middle panel) to the covariance of the residuals used in our LPI procedure.

For the nebulous region, the long-range correlations are comparable to the pixel-wise noise (magnitude of the central pixel) and are notably anisotropic. In the Galactic bulge, the pixel-wise noise is more dominant than long-range correlations. However, as evidenced by Figure 12, our covariance estimate in the bulge is imperfect. This is in part because our masking fraction in regions of high density is so large. Uncertainty due

to residuals from mismodeling neighboring stars (which are currently masked to access the covariance of the background) is likely important in this limit and incorporating this uncertainty into the covariance estimate is a route for future work. At high Galactic latitudes, the uncertainty is dominated by the pixel-wise uncertainty, as expected and as shown in Figure 6. While we have focused on the r band, the covariance functions for other bands are similar. The covariance function for the same region, but in different bands may differ as a result of varying S/N, seeing, artifacts from diffraction spikes, or the difference in backgrounds visible at those wavelengths.

ORCID iDs

Andrew K. Saydjari  <https://orcid.org/0000-0002-6561-9002>
 Douglas P. Finkbeiner  <https://orcid.org/0000-0003-2808-275X>

References

- Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv:1902.05569
- Almosallam, I. A., Lindsay, S. N., Jarvis, M. J., & Roberts, S. J. 2016, *MNRAS*, **455**, 2387
- Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, *MNRAS*, **474**, 2094
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Baghi, Q., Métris, G., Bergé, J., et al. 2016, *PhRvD*, **93**, 122007
- Barclay, T., Endl, M., Huber, D., et al. 2015, *ApJ*, **800**, 46
- Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, *PASP*, **115**, 953
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Brewer, B. J., Foreman-Mackey, D., & Hogg, D. W. 2013, *AJ*, **146**, 7
- Bu, Y., Kumar, Y. B., Xie, J., et al. 2020, *ApJS*, **249**, 7
- Collette, A. 2013, Python and HDF5 (Sebastopol, CA: O'Reilly)
- Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, *ApJ*, **812**, 128
- Da Costa, G. S. 1992, in ASP Conf. Ser. 23, Astronomical CCD Observing and Reduction Techniques, ed. S. B. Howell (San Francisco, CA: ASP), 90
- Di Francesco, J., Hogerheijde, M. R., Welch, W. J., & Bergin, E. A. 2002, *AJ*, **124**, 2749
- Evans, T. M., Aigrain, S., Gibson, N., et al. 2015, *MNRAS*, **451**, 680
- Feder, R. M., Portillo, S. K. N., Daylan, T., & Finkbeiner, D. 2020, *AJ*, **159**, 163
- Ferrarese, L., Silberman, N. A., Mould, J. R., et al. 2000, *PASP*, **112**, 177
- Fielder, C. E., Newman, J. A., Andrews, B. H., et al. 2021, *MNRAS*, **508**, 4459
- Finkbeiner, D. P., Padmanabhan, N., Schlegel, D. J., et al. 2004, *AJ*, **128**, 2577
- Fitzpatrick, E. L., & Massa, D. 2009, *ApJ*, **699**, 1209
- Fortino, W. F., Bernstein, G. M., Bernardinelli, P. H., et al. 2021, *AJ*, **162**, 106
- Gentile, M., Courbin, F., & Meylan, G. 2013, *A&A*, **549**, A1
- Ghosh, A., Mertens, F., Bernardi, G., et al. 2020, *MNRAS*, **495**, 2813
- Gibson, N., Aigrain, S., Roberts, S., et al. 2012, *MNRAS*, **419**, 2683
- Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, *ApJ*, **887**, 93
- Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, *MNRAS*, **443**, 2517
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jones, R. L., Bannister, M. T., Bolin, B. T., et al. 2020, arXiv:2009.07653
- Junklewitz, H., Bell, M., Selig, M., & Enßlin, T. 2016, *A&A*, **586**, A76
- Kern, N. S., & Liu, A. 2021, *MNRAS*, **501**, 1463
- Lang, D., Hogg, D. W., & Mykytyn, D. 2016, The Tractor: Probabilistic Astronomical Source Detection and Measurement, Astrophysics Source Code Library, ascl:1604.008
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, **379**, 1599
- Léget, P. F., Astier, P., Regnault, N., et al. 2021, *A&A*, **650**, A81
- Leike, R. H., & Enßlin, T. A. 2019, *A&A*, **631**, A32
- Lide, D. R. 2004, CRC Handbook of Chemistry and Physics, Vol. 85 (Boca Raton, FL: CRC Press)
- Littlefair, S., Burningham, B., & Helling, C. 2017, *MNRAS*, **466**, 4250
- Liu, R., McAuliffe, J. D., & Regier, J. 2021, arXiv:2102.02409
- Magakian, T. Y. 2003, *A&A*, **399**, 141
- Magnier, E. A., Schlafly, E. F., Finkbeiner, D. P., et al. 2020, *ApJS*, **251**, 6
- Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, *PASP*, **131**, 018003
- McAllister, M., Littlefair, S., Dhillon, V., et al. 2017, *MNRAS*, **464**, 1353
- Meisner, A. M., & Finkbeiner, D. P. 2014, *ApJ*, **781**, 5
- Mertens, F., Ghosh, A., & Koopmans, L. 2018, *MNRAS*, **478**, 3640
- Minniti, D., Lucas, P., Emerson, J., et al. 2010, *NewA*, **15**, 433
- Molinari, S., Schisano, E., Faustini, F., et al. 2011, *A&A*, **530**, A133
- Naumov, M. 2021, Parallel incomplete-LU and Cholesky factorization in the preconditioned iterative methods on the GPU, NVIDIA Tech. Rep. NVR-2012-003
- Ndiritu, S. W., Scaife, A. M. M., Tabb, D. L., Cárcamo, M., & Hanson, J. 2021, *MNRAS*, **502**, 5839
- Newberg, H. J. & Sloan Digital Sky Survey Collaboration 2003, AAS Meeting Abstracts, **203**, 112.11
- Nidever, D. L., Dey, A., Fasbender, K., et al. 2021, *AJ*, **161**, 192
- Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. 2016, in Proc. of the IEEE Conf. Computer Vision and Pattern Recognition (Piscataway, NJ: IEEE), 2536
- Perez, F., & Granger, B. E. 2007, *CSE*, **9**, 21
- Pinter, S., Bagoly, Z., Balázs, L. G., et al. 2018, in IAU Symp. 333, Peering Towards Cosmic Dawn, ed. V. Jelić & T. van der Hulst (Cambridge: Cambridge Univ. Press), 168
- Platen, E., van de Weygaert, R., Jones, B. J. T., Vegter, G., & Calvo, M. A. A. 2011, *MNRAS*, **416**, 2494
- Portillo, S. K. N., Lee, B. C. G., Daylan, T., & Finkbeiner, D. P. 2017, *AJ*, **154**, 132
- Portillo, S. K. N., Speagle, J. S., & Finkbeiner, D. P. 2020, *AJ*, **159**, 165
- Rajpaul, V., Aigrain, S., & Roberts, S. 2015, *MNRAS: Letters*, **456**, L6
- Rasmussen, C. E. 2003, Summer School on Machine Learning (Berlin: Springer), 63
- Saydjari, A. K., Schlafly, E. F., Lang, D., et al. 2022, arXiv:2206.11909
- Schlafly, E. F., Green, G. M., Lang, D., et al. 2018, *ApJS*, **234**, 39
- Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, *ApJS*, **240**, 30
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Soares, P. S., Watkinson, C. A., Cunningham, S., & Pourtsidou, A. 2021, *MNRAS*, **510**, 5872
- Stetson, P. B. 1987, *PASP*, **99**, 191
- Tange, O. 2021, GNU Parallel 20210622 (“Protasevich”), Zenodo, doi:10.5281/zenodo.5013933
- Traficante, A., Fuller, G. A., Pineda, J. E., & Pezzuto, S. 2015, *A&A*, **574**, A119
- Ulyanov, D., Vedaldi, A., & Lempitsky, V. 2018, in Proc. of the IEEE Conf. Computer Vision and Pattern Recognition (Piscataway, NJ: IEEE), 9446
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Van Veen, D., Jalal, A., Soltanolkotabi, M., et al. 2018, arXiv:1806.06438
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Way, M. J., Foster, L. V., Gazis, P. R., & Srivastava, A. N. 2009, *ApJ*, **706**, 623
- Williams, T. G., Kreckel, K., Belfiore, F., et al. 2022, *MNRAS*, **509**, 1303
- Wu, E., Wu, K., Cox, D., & Lotter, W. 2018, Image Analysis for Moving Organ, Breast, and Thoracic Images (Berlin: Springer), 98
- Yu, Y., Zhang, J., Jing, Y., & Zhang, P. 2017, *PhRvD*, **95**, 043536
- Zhang, H., & Brandt, T. D. 2021, *AJ*, **162**, 139