The Astronomical Journal, 161:147 (24pp), 2021 March https://doi.org/10.3847/1538-3881/abd806

Estimating Distances from Parallaxes. V. Geometric and Photogeometric Distances to 1.47 Billion Stars in Gaia Early Data Release 3

C. A. L. Bailer-Jones[1], J. Rybizki[1], M. Fouesneau[1], M. Demleitner[2], and R. Andrae[1]
[1] Max Planck Institute for Astronomy, Heidelberg, Germany
[2] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Germany
Received 2020 December 9; revised 2020 December 30; accepted 2020 December 31; published 2021 February 25

Abstract

Stellar distances constitute a foundational pillar of astrophysics. The publication of 1.47 billion stellar parallaxes from Gaia is a major contribution to this. Despite Gaia's precision, the majority of these stars are so distant or faint that their fractional parallax uncertainties are large, thereby precluding a simple inversion of parallax to provide a distance. Here we take a probabilistic approach to estimating stellar distances that uses a prior constructed from a three-dimensional model of our Galaxy. This model includes interstellar extinction and Gaia's variable magnitude limit. We infer two types of distance. The first, geometric, uses the parallax with a direction-dependent prior on distance. The second, photogeometric, additionally uses the color and apparent magnitude of a star, by exploiting the fact that stars of a given color have a restricted range of probable absolute magnitudes (plus extinction). Tests on simulated data and external validations show that the photogeometric estimates generally have higher accuracy and precision for stars with poor parallaxes. We provide a catalog of 1.47 billion geometric and 1.35 billion photogeometric distances together with asymmetric uncertainty measures. Our estimates are quantiles of a posterior probability distribution, so they transform invariably and can therefore also be used directly in the distance modulus ($5\log_{10} r - 5$). The catalog may be downloaded or queried using ADQL at various sites (see http://www.mpia.de/~calj/gedr3_distances.html), where it can also be cross-matched with the Gaia catalog.

Unified Astronomy Thesaurus concepts: Catalogs (205); Galaxy structure (622); Bayesian statistics (1900); Parallax (1197); Stellar parallax (1618); Photometric parallax (1231); Distance indicators (394); Astrometry (80); Markov chain Monte Carlo (1889); Absolute magnitude (10)

1. Introduction

There are various ways to determine astrophysical distances. Near the base of the distance ladder on which almost all other distance measures are built are geometric parallaxes of stars. In recognition of this, the European Space Agency (ESA) implemented the Gaia mission to obtain parallaxes for over one billion stars in our Galaxy down to $G \simeq 20$ mag, with accuracies to tens of microarcseconds (Gaia Collaboration 2016a). The first two data releases (Gaia Collaboration 2016b, 2018) presented a significant leap forward in both the number and accuracy of stellar parallaxes. The recently published early third release (Gaia Collaboration 2020a, hereafter EDR3) reduces the random and systematic errors in the parallaxes by another 30%.

While parallaxes (ϖ) are the basis for a distance determination, they are not themselves distances (r). This is due to the nonlinear transformation between them ($\varpi \sim 1/r$) and the presence of significant noise for more distant stars. Small absolute uncertainties in parallax can translate into large uncertainties in distance, and while parallaxes can be negative, distances cannot be. Thus, for anything but the most precise parallaxes, the inverse parallax is a poor distance estimate. An explicit probabilistic approach to inferring distances may instead be taken. This has been discussed and applied to parallax data in various publications in recent years; a recent overview is given by Luri et al. (2018). The simplest approach uses just the parallax and parallax uncertainty together with a one-dimensional prior over distance. This yields a posterior probability distribution over distance to an individual star (Bailer-Jones 2015). A suitable prior ensures that the posterior converges to something sensible as the precision of the parallax degrades. This is important when working with Gaia data because, its truly revolutionary nature notwithstanding, in EDR3 only 33% of the sources have a parallax signal-to-noise ratio greater than 2 (13% greater than 5), and a further 24% have negative parallaxes. The shape and scale of the prior distribution should reflect the expected distribution of stars in the sample, including observational selection effects such as magnitude limits. The prior's characteristic length scale will typically need to vary with direction in the Galaxy (Bailer-Jones et al. 2018). More sophisticated approaches use other types of data, such as the star's magnitude and color (Astraatmadja & Bailer-Jones 2016a; McMillan 2018; Anders et al. 2019; Leung & Bovy 2019), velocity (Schönrich & Aumer 2017; Zucker et al. 2018), or spectroscopic (Sanders & Das 2018; Queiroz et al. 2020) or asteroseismic (Hall et al. 2019) parameters. In order to exploit such additional data, these methods must make deeper astrophysical assumptions than parallax-only approaches, and they may also have more complex priors. The benefit is that the inferred distances will usually be more precise (lower random errors), and hopefully also more accurate (lower systematic errors) if the extra assumptions are correct.

In the present paper, the fifth in a series, we determine distances for sources in EDR3 using data exclusively from EDR3. The resulting catalog should be more accurate and more useful than our earlier work, on account of both the more accurate parallaxes in EDR3 and improvements in our method. We determine two types of distance. The first, which we call "geometric," uses only the parallaxes and their uncertainties. We explored this approach in detail in the first two papers in this series (Bailer-Jones 2015; Astraatmadja & Bailer-Jones 2016a, hereafter Papers I and II) and applied it to estimate distances for 2 million stars in the first Gaia data

release (Astraatmadja & Bailer-Jones 2016b, Paper III) and 1.33 billion stars in the second Gaia data release (Bailer-Jones et al. 2018, Paper IV). Both papers used a (different) direction-dependent distance prior that reflected the Galaxy's stellar populations and Gaia's selection thereof.

Our second type of distance estimate uses, in addition to the parallax, the color and magnitude of the star. We call such distances "photogeometric." Along with the distance prior, this uses a model of the direction-dependent distribution of (extincted) stellar absolute magnitudes.

We construct our priors from the GeDR3 mock catalog of Rybizki et al. (2020). This lists, among other things, the (noise-free) positions, distances, magnitudes, colors, and extinctions of 1.5 billion individual stars in the Galaxy as a mock-up of what was expected to appear in EDR3. GeDR3mock is based on the Besançon Galactic model and PARSEC stellar evolutionary tracks. We exclude stars from GeDR3mock that simulate the Magellanic Clouds (`popid = 10`) and stellar open clusters (`popid = 11`). We divide the sky into the 12,288 equal-area (3.36 sq. deg.) regions defined by the HEALpixel scheme[3] at level 5, and we fit our prior models separately to each. In doing this, we only retain from GeDR3mock those stars that are brighter than the 90th percentile of the EDR3 magnitude distribution in that HEALpixel (Rybizki & Drimmel 2018; Gaia Collaboration 2020b). This is done to mimic the variable magnitude limit of Gaia over the sky, and it varies from 19.2 mag around the Galactic center to 20.7 mag over much of the rest of the sky (the median over HEALpixels is 20.5 mag).

We apply our inference to all sources in EDR3 that have parallaxes. As our prior only reflects single stars in the Galaxy, our distances will be incorrect for the small fraction of extragalactic sources in the Gaia catalog and may also be wrong for some unresolved binaries, depending on their luminosity ratios.

As some readers may be familiar with our previous catalog using GDR2 data (Paper IV), here is a summary of the main changes in the new method (which we describe fully in Section 2):

1. We update the source of our prior from a mock catalog of GDR2 (Rybizki et al. 2018) to one of EDR3 (Rybizki et al. 2020).
2. We replace the one-parameter exponential decreasing space density (EDSD) distance prior with a more flexible three-parameter distance prior (Section 2.3).
3. We again fit the distance prior to a mock catalog, but we no longer use spherical harmonics to smooth the length scale of the prior over the sky. We instead adopt a common distance prior for all stars within a small area (level 5 HEALpixels).
4. We introduce photogeometric distances (Section 2.4) using a model for the (extincted) color–absolute magnitude diagram, also defined per HEALpixel (Section 2.5).
5. In Paper IV we summarized each posterior with the mode and the highest density interval. The mode has the disadvantage that it is not invariant under nonlinear transformations. This means that if we inferred $r_{\rm mode}$ as the mode of the posterior in distance, then $5\log_{10} r_{\rm mode} - 5$ would not, in general, be the mode of the posterior in distance modulus. This is also the case for the mean. The quantiles of a distribution, in contrast, are invariant under (monotonic) nonlinear transformations. We therefore provide the median (the 50th percentile) of the posterior as our distance estimate. To characterize the uncertainty in this, we quote the 14th and 86th percentiles (an equal-tailed interval, ETI). These are therefore also the quantiles on the absolute magnitude inferred from the distance.

In the next section, we describe our method and the construction of the priors. In Section 3, we apply our method to the GeDR3mock catalog, giving some insights into how it performs. We present the results on EDR3 in Section 4, and we describe the resulting distance catalog in Section 5 along with its use and limitations. We summarize in Section 6. Auxiliary information, including additional plots for all HEALpixels—for both the prior and the results—can be found online.[4]

2. Method

For each source, we compute the following two posterior probability density functions (PDFs) over the distance r:

$$\text{Geometric:}\quad P_{\rm g}^{*}(r \mid \varpi, \sigma_\varpi, p)$$

$$\text{Photogeometric:}\quad P_{\rm pg}^{*}(r \mid \varpi, \sigma_\varpi, p, G, c)$$

where ϖ is the parallax, σ_ϖ is the uncertainty in the parallax, p is the HEALpixel number (which depends on Galactic latitude and longitude), G is the apparent magnitude, and c is the BP − RP color. The parallax and apparent magnitude will be adjusted to accommodate known issues with the EDR3 data, as detailed below. The star symbol indicates that we infer unnormalized posteriors. The geometric posterior uses just a distance prior. The photogeometric posterior uses this distance prior as well as a color–magnitude prior that we explain below. The posteriors are summarized using quantiles computed by Markov Chain Monte Carlo (MCMC) sampling.

2.1. Geometric Distance

The unnormalized posterior PDF is the product of the likelihood and prior:

$$P_{\rm g}^{*}(r \mid \varpi, \sigma_\varpi, p) = P(\varpi \mid r, \sigma_\varpi)\, P(r \mid p). \tag{1}$$

The likelihood is conditionally independent of p. We chose to make the second term, which we define in Section 2.3, independent of σ_ϖ.

2.2. Likelihood

Under the assumption of Gaussian parallax uncertainties, the likelihood is

$$P(\varpi \mid r, \sigma_\varpi) = \frac{1}{\sqrt{2\pi}\,\sigma_\varpi} \exp\left[-\frac{1}{2\sigma_\varpi^2}\left(\varpi - \varpi_{\rm zp} - \frac{1}{r}\right)^2\right] \tag{2}$$

where $\varpi_{\rm zp}$ is the parallax zero-point. In Paper IV we adopted a constant value of −0.029 mas for this zero-point, as recommended in the GDR2 release. For EDR3 the Gaia team has published a more sophisticated parallax zero-point based on analyses of quasars, binary stars, and the Large Magellanic

[3] https://healpix.sourceforge.io

[4] http://www.mpia.de/~calj/gedr3_distances.html

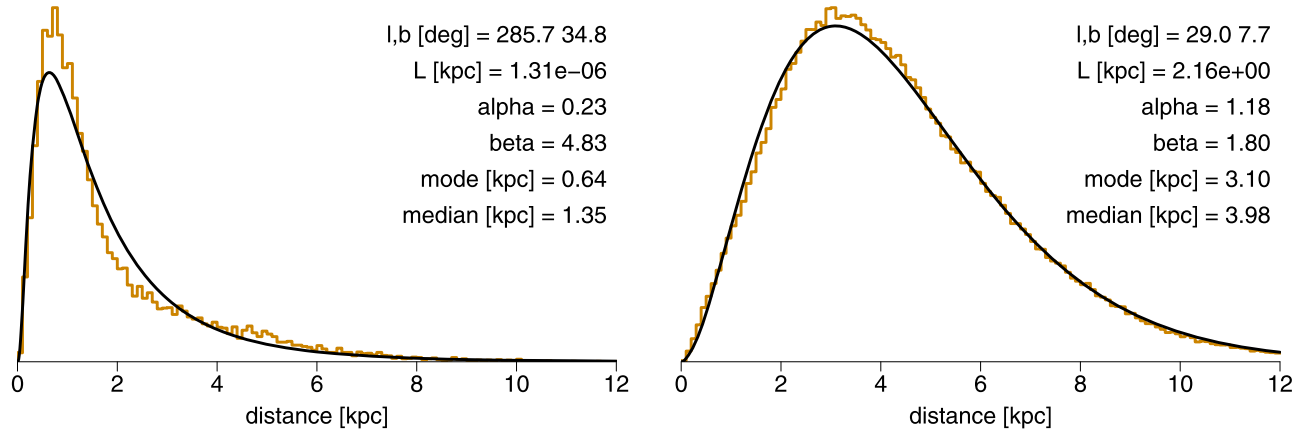

Figure 1. Distance priors for two HEALpixels, number 6200 at high latitude (left) and number 7593 at low latitude (right). The histograms show the distributions of the data in the mock catalog. The smooth curves are the fit of the GGD (Equation (3)) to these data, which defines the distance prior $P(r \mid p)$ with the parameters L, α, and β. Similar plots for all HEALpixels are available with the auxiliary information online.

Cloud (LMC; Lindegren et al. 2020b). This is a function of G, the ecliptic latitude, and the effective wavenumber used in the astrometric solution. Ideally, this last term was derived from the BP − RP color, and this is the case for the standard five-parameter (5p) astrometric solutions used for 585 million sources (Gaia Collaboration 2020a). But where BP − RP was unavailable or deemed of insufficient quality, the effective wavenumber was derived as a sixth parameter in the astrometric solution (6p solutions; Lindegren et al. 2020a), which is the case for 882 million sources. Overall, the zero-point ranges between about −0.150 and +0.130 mas (it is narrower for the 5p solutions), although the rms range is only 0.020 mas. We use this zero-point correction in Equation (2). Our geometric distances are therefore weakly conditioned also on G and c, but we omit this in the mathematical notation for brevity. For the 2.5 million sources that have parallaxes but no G (strictly, no `phot_g_mean_mag`), we use the EDR3 global zero-point of −0.017 mas (Lindegren et al. 2020a).

2.3. Distance Prior

In Paper IV we used the one-parameter EDSD distance prior, which models the space density of stars as dropping exponentially away from the Sun according to a (direction-dependent) length scale. Here we adopt the more flexible, three-parameter generalized gamma distribution (GGD), which can be written as

$$P(r \mid p) = \begin{cases} \dfrac{1}{\Gamma(\frac{\beta+1}{\alpha})} \dfrac{\alpha}{L^{\beta+1}} r^{\beta} e^{-(r/L)^{\alpha}} & \text{if } r \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

for $\alpha > 0$, $\beta > -1$, and $L > 0$. Here, $\Gamma()$ is the gamma function. This PDF is unimodal with an exponentially decreasing tail to larger distances. The mode is $L(\beta/\alpha)^{1/\alpha}$ for $\beta > 0$, and zero otherwise. The EDSD is a special case of the GGD with $\alpha = 1$ and $\beta = 2$. We fit the GGD prior for each HEALpixel separately via maximum likelihood using stars from the mock catalog. The HEALpixel (p) dependency on the left side of Equation (3) is equivalent to a dependency on α, β, and L.

Example fits for two HEALpixels, one at low Galactic latitude and one at high Galactic latitude, are shown in Figure 1. Although the GGD prior provides a better fit than the EDSD prior—which is why we use it—the parameter L may no longer be interpreted as a meaningful length scale, because it varies from 3e−7 to 1e4 pc over all HEALpixels. The appropriate characteristic scale of the GGD prior in this work is its median, for which there is no closed-form expression. The median varies between 745 and 7185 pc depending on HEALpixel (Figure 2). Fits for each HEALpixel can be found in the auxiliary information online.

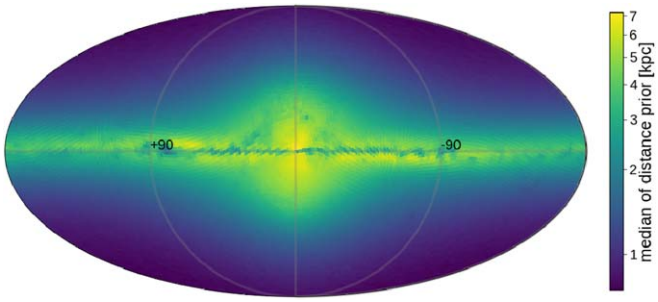

Figure 2. Variation of the median of the distance prior over the sky shown in Galactic coordinates on a Mollweide equal-area projection. The LMC and SMC are excluded from our prior.

In the limit of uninformative parallaxes, the geometric posterior converges on the GGD prior, so the median distance converges on the median of this prior. In Paper IV this convergence was on the mode of the EDSD prior. For the prior fits used in the present paper, the ratio of the GGD median to the EDSD mode ranges from 1.17 to 1.57. There are potential improvements one could make to the prior to give a better convergence in the limit of poor data. Some considerations are in Appendix A.

2.4. Photogeometric Distance

We define the quantity Q_G as

$$Q_G \equiv M_G + A_G = G - 5\log_{10} r + 5. \quad (4)$$

The equality (=), which is a statement of flux conservation, holds only when all of the quantities are noise-free. If we knew Q_G for a star, then a measurement of G gives us an estimate of r. Given that the uncertainties on G in EDR3 are generally less than a few millimagnitudes (0.3–6 mmag for $G < 20$ mag; Gaia Collaboration 2020a), this would be a reasonably precise

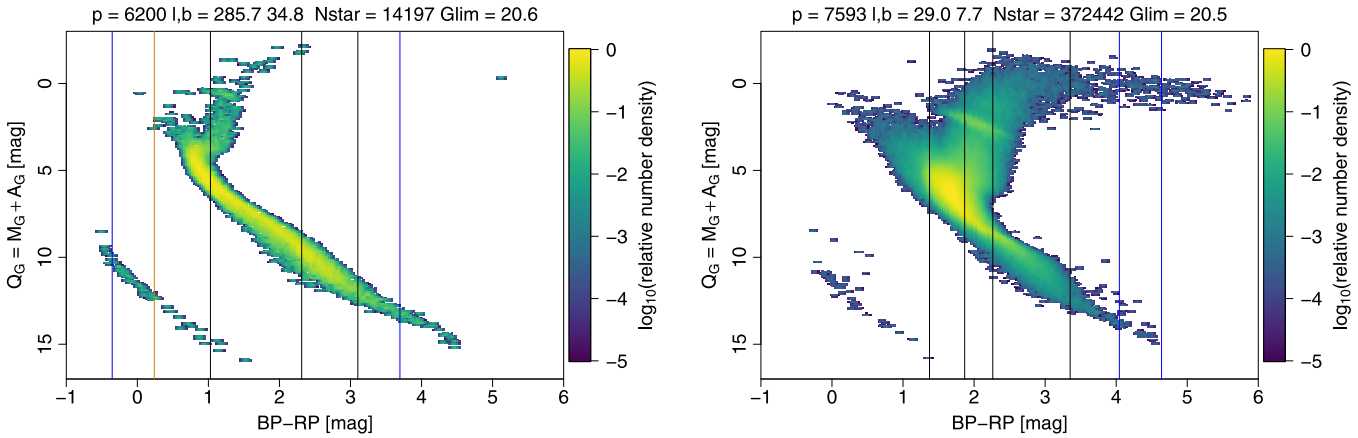

Figure 3. CQDs for HEALpixels 6200 (left) and 7593 (right) in the mock catalog. The density of stars is shown on a logarithmic color scale relative to the maximum density in each HEALpixel (so the zero-points of the density scales are not the same in the two panels). The text at the top of each panel gives the Galactic longitude and latitude (l, b) of the center of the HEALpixel in degrees, the number of stars, and the faintest magnitude. The vertical lines identify particular Q_G models that are shown in Figures 4 and 5. Similar plots for all HEALpixels are available with the auxiliary information online.

estimate. We do not know Q_G, but we can take advantage of the fact that the two-dimensional color–Q_G space for stars is not uniformly populated. This space (e.g., Figure 3), which we call the CQD, in analogy to the CMD (color–magnitude diagram), would be identical to the color–absolute magnitude diagram if there were no interstellar extinction. Thus, if we know the BP − RP color of the star, this diagram places limits on possible values of Q_G, and therefore on the distance to the star. We will use the mock catalog to model the CQD (per HEALpixel) and from this compute a prior over Q_G given the magnitude and color of the star.

The formal procedure is as follows, initially making no assumptions about G. We assume the color to be effectively noise-free. This is reasonable given the relatively low noise for most sources (13–120 mmag for $G < 20$ mag; Gaia Collaboration 2020a) and the fact that the prior is anyway imperfect (see Section 2.5). Using Bayes' theorem, the unnormalized posterior we want to estimate can be decomposed into a product of two terms:

$$P^*_{\rm pg}(r \mid \varpi, \sigma_\varpi, G, c, p) = P(\varpi \mid r, \sigma_\varpi)\, P(r \mid G, c, p). \quad (5)$$

The first term on the right side is the parallax likelihood (Section 2.2). It is independent of G, c, and p once it is conditioned on σ_ϖ, which is estimated in the Gaia astrometric solution using quantities that depend on the magnitude, color, scanning law, and so on (Lindegren et al. 2020a). The second term is independent of the parallax measurement process and thus of ϖ and σ_ϖ. We may write this second term as a marginalization over Q_G and then apply Bayes' theorem as follows:

$$\begin{aligned} & P(r \mid G, c, p) \\ &= \int P(r, Q_G \mid G, c, p)\, dQ_G \\ &= \int \frac{1}{P(G \mid c, p)}\, P(G \mid r, Q_G)\, P(r, Q_G \mid c, p)\, dQ_G \\ &= \frac{P(r \mid c, p)}{P(G \mid c, p)} \int P(G \mid r, Q_G)\, P(Q_G \mid r, c, p)\, dQ_G. \end{aligned} \quad (6)$$

In the last line, the first term under the integral is formally the likelihood for G (and is conditionally independent of c and p due to Equation (4)). However, as G is measured much more precisely than the intrinsic spread in Q_G—that is, the second term under the integral is a much broader function—we can consider G to be noise-free to a good approximation. This makes the first term a delta function, so the integral is nonzero only when Equation (4) is satisfied.

We make two further assumptions about the terms in the last line of Equation (6). The first is to make the distance prior independent of color, that is, $P(r \mid c, p) \to P(r \mid p)$. This is now the same distance prior as used in the geometric posterior (Equation (1)). The second is to assume that the CQD is independent of distance, that is, $P(Q_G \mid r, c, p) \to P(Q_G \mid c, p)$. This is not true in general, but we chose not to add this extra layer of dependence on GeDR3mock (see Section 2.5).

With these assumptions, the (unnormalized) posterior in Equation (5) can now be written as

$$\begin{aligned} P^*_{\rm pg}(r \mid \varpi, \sigma_\varpi, G, c, p) &\simeq P(\varpi \mid r, \sigma_\varpi)\, P(r \mid p) \\ &\times P(Q_G = G - 5\log_{10} r + 5 \mid c, p). \end{aligned} \quad (7)$$

The missing normalization constant, $1/P(\varpi, G \mid c, p)$, is not required. This posterior is simply the geometric posterior (Equation (1)) multiplied by an additional prior[5] over Q_G.

2.5. Q_G Prior

We construct the prior $P(Q_G \mid c, p)$ from the mock catalog. Given the complexity of the CQD and its variation over the sky, we do not attempt to fit the prior as a continuous 3D (position and color) parametric function. We instead compute a CQD for each HEALpixel, two examples of which are shown in Figure 3. Within each, we compute a series of one-dimensional functions at a series of colors in the following way. We divide the full color range of a given HEALpixel into strips of 0.1 mag width in color, and then for each strip we fit a model to the stellar number density as a function of Q_G (now ignoring the color variation in each strip). If there are more than 40 stars in a strip, we bin the data into bins of 0.1 mag and fit a smoothing spline with $\min(\lfloor N/4 \rfloor, 50)$ degrees of freedom

[5] We could have arrived at this without using the marginalization in Equation (6) if we assumed G to be noise-free from the outset, but the marginalization justifies how small the noise in G has to be for this to be valid.

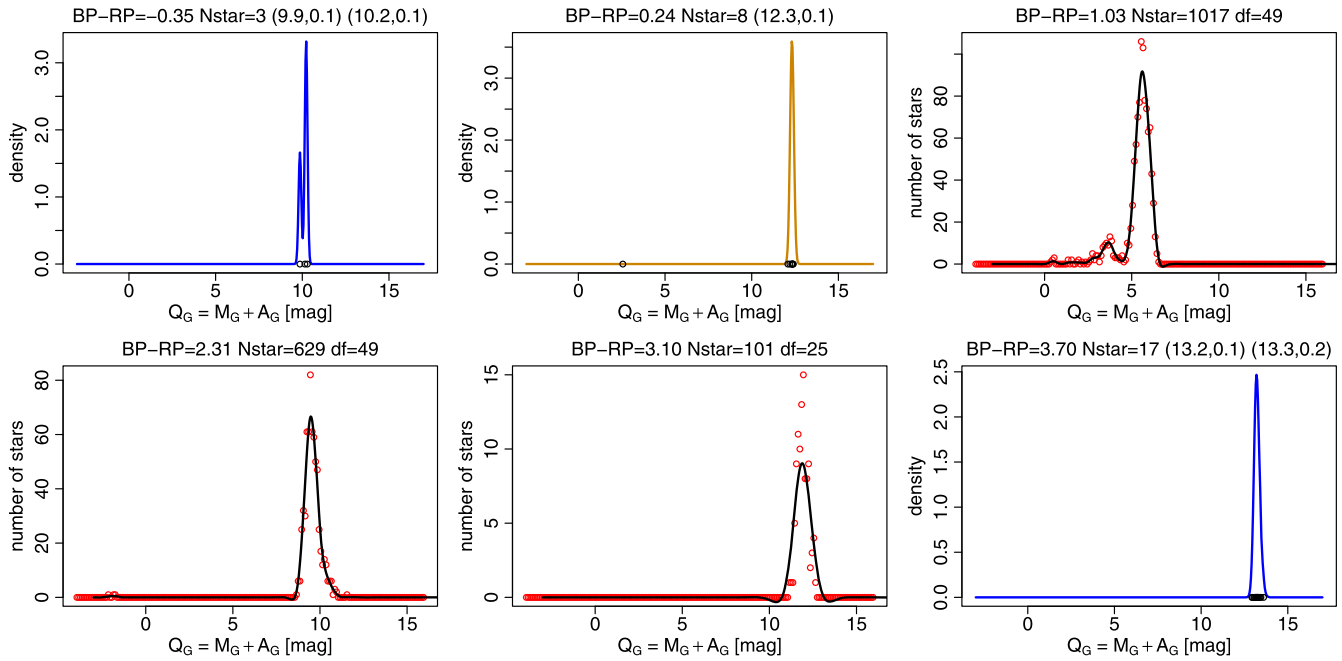

Figure 4. The Q_G prior models constructed from the CQD of HEALpixel 6200. Each of the six panels shows a fit to the mock data at a different BP − RP color, corresponding to the six vertical stripes shown in Figure 3 (left panel). Model fits using smoothing splines are plotted as black lines with the degrees of freedom as indicated and the (binned) data in the fit shown as red circles. Model fits using one or two Gaussian components are plotted as orange and blue lines, respectively, with the data in the fit shown as black circles and the mean and standard deviation of the fit components indicated in parentheses at the top of each panel. These density functions show the prior PDF $P(Q_G \mid c, p)$ at discrete colors before imposing the minimum threshold that ensures the prior density is always greater than zero. Similar plots for all color strips in HEALpixels are available with the auxiliary information online.

(df), where N is the number of stars in the strip (which can be many thousands). If there are fewer than 40 stars, we cannot fit a good spline. This generally occurs at the bluest and reddest ends of the CQD. Here the Q_G distribution is often characterized by two widely separated components, either the main sequence (MS) and white dwarf (WD) branches, or the MS and giant star branches (see Figure 3). Thus when $N < 40$, we instead fit a two-component Gaussian mixture model, with the constraint that the minimum and maximum standard deviation of each component be $\sigma_{\rm min} = 0.08$ mag and $\sigma_{\rm max} = 1.0$ mag, respectively. A full fit requires at least five stars, so if there are as few as two stars, we constrain the solution to first have equal standard deviations and then to have standard deviations of $\sigma_{\rm min}$. If $N = 1$, our model is a one-component Gaussian with mean equal to the Q_G of the star and standard deviation equal to $\sigma_{\rm min}$. If there are no stars, the model is null. Examples of the fits are shown in Figures 4 and 5.

As a smoothing spline can give a negative fit, and both these and the Gaussian models can yield very small values for the density, we impose that the minimum density is never less than 10^{-3} of the integrated density (computed prior to fitting the model). Thus our prior is nowhere near zero, meaning that even if the data indicate a Q_G in the regions where the mock catalog is empty, the posterior will not be zero. This allows sources to achieve distances that place them outside the occupied regions of the mock CQD.

For a given HEALpixel, each prior model refers to a specific color, namely the center of a 0.1 mag wide strip. This is larger than the uncertainty in the color for all but the faintest EDR3 sources. When evaluating the prior during the inference process, we compute Q_G from Equation (4), evaluate the densities of the two priors that bracket its color, then linearly interpolate. This ensures that our prior is continuous in color. If one of the models is null, we use the other model as is. If both models are null, or if the source is outside the color range of the mock CQD, we do not infer a photogeometric distance. The flag field in our catalog indicates what kinds of Q_G models were used (see Section 5).

The computation of Q_G in Equation (4) requires the G-band magnitude of the source. For this, we use the `phot_g_mean_mag` field in EDR3 corrected for the processing error described in Section 8.3 of Riello et al. (2020). This correction, which is a function of magnitude and color, can be as large as 25 mmag.

2.6. Posterior Sampling and Summary

The posteriors are formally the answer to our inference process. The geometric posterior has a simple parametric form that may be computed by the reader using the data in the EDR3 catalog and the parameters of our prior (available with the auxiliary information online). The photogeometric posterior is generally nonparametric. Both posteriors are asymmetric and not necessarily unimodal (Section 2.6.2).

There are a variety of statistics one could use to summarize these PDFs, such as the mean, median, or mode. There is no theoretically correct measure, and all have their drawbacks. We use quantiles, primarily because they are invariant under nonlinear transformations, and thus are simultaneously the quantiles of the posterior in distance modulus, $5\log_{10} r - 5$. We use the three quantiles at 0.159, 0.5, and 0.841, which we label $r_{\rm lo}$, $r_{\rm med}$, and $r_{\rm hi}$, respectively. The central quantile is the median. The outer two quantiles give a 68% confidence interval around the median. The difference between each quantile and the median is a Gaussian 1σ-like estimate of the uncertainty.

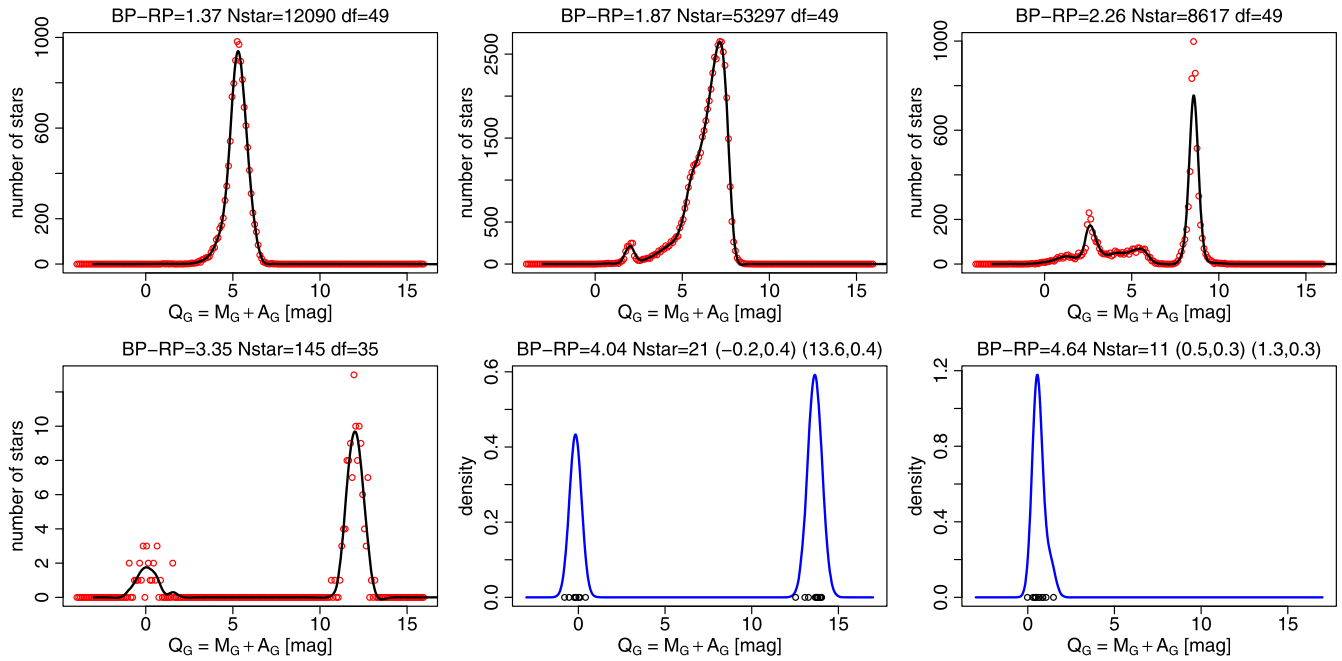

Figure 5. Same as Figure 4 but now for HEALpixel 7593, the CQD of which is in the right panel of Figure 3.

Due to the intrinsic asymmetry of the posteriors, we report the lower and upper values separately.

2.6.1. Markov Chain Monte Carlo

Neither the geometric nor photogeometric posteriors have closed-form expressions for their quantiles, so we must compute these numerically. We do this using MCMC, specifically the Metropolis algorithm.

We adopt the following scheme for the MCMC initialization and step size. We first compute the geometric distance posterior using the EDSD prior from Paper IV. The length scale of this prior is set to 0.374 $r_{\rm med}$, where $r_{\rm med}$ is the median distance of the stars in the mock catalog for that HEALpixel.[6] We use the mode of this posterior, $r_{\rm mode}^{\rm EDSD}$, which has a closed-form solution (Paper I), as the initialization for the geometric posterior. The initialization scheme for the photogeometric posterior is more complicated, in accordance with its more complicated shape, and depends on $r_{\rm mode}^{\rm EDSD}$, the fractional parallax uncertainty (fpu, σ_ϖ/ϖ), and the characteristic length scale of the Q_G prior model(s).

For both types of posterior, the step size needs to be adapted to the characteristic width of the posterior, which is generally wider the larger the fpu. We found a suitable step size to be $(3/4)r_{\rm init} \times \min(|\sigma_\varpi/\varpi|, 1/3)$, where $r_{\rm init}$ is the initialization value.

This scheme allows relatively short burn-ins: we use just 50. We experimented with chains of various chain lengths, employing various tests of convergence. Longer chains are always better, but as we need to sample around three billion posteriors, some parsimony is called for. We settled on 500 samples (after burn-in). Although the chains are not always settled, they are generally good enough to compute the required quantiles with reasonable precision. To quantify this, we obtained 20 different MCMC chains and computed the standard deviation of the median distance estimates and one-half the mean of the confidence intervals. The ratio of these is a measure of the convergence noise. Doing this for thousands of stars, we find this to be between 0.1 and 0.2 in general. For the geometric posteriors in particular, it can be larger for fractional parallax uncertainties larger than 0.3.

2.6.2. Multimodality

The posteriors can be multimodal. This is more likely to be the case for the photogeometric posterior at large fpu, as its prior can be multimodal. Multimodality is very rare for the geometric posterior.

Although multimodality is a challenge for MCMC sampling methods, we find that even widely separated modes can be sampled in our scheme. Our 68% confidence interval often encompasses the span of such multimodality. This is a blessing and a curse: the distance precision in a single mode may be quite good, yet a large confidence interval is obtained because of the presence of a second mode. To assist in identifying possible multimodality, we perform the Hartigan dip test (Hartigan & Hartigan 1985). This is a classical statistical test in which the null hypothesis is a unimodal posterior; that is, a small p-value suggests the distribution may not be unimodal. We select a threshold of 10^{-3} and set a flag to 1 if the p-value is lower than this, thereby suggesting possible multimodality. If the p-value is above this threshold or the test does not work for any reason, the flag is 0. The test is not particularly accurate and should not be overinterpreted. Furthermore, it is done on the MCMC samples, not on the true posterior, so it tends to be raised more often than expected because of the intrinsic noise of MCMC sampling.

[6] In Paper IV we used $(1/3)r_{\rm med}$ because the maximum-likelihood fit of the length scale is one-third of the mean. However, the median is a slightly biased estimator of the mean for the EDSD. For the typical length scales involved, we found empirically that the mean is about 12% (0.374/0.333) larger than the median.

3. Performance on the Mock Catalog

Before looking at the results on EDR3, we evaluate the performance of our method using the mock catalog, as here we know the true distances. In doing this, we add Gaussian random noise to the parallaxes using the `parallax_error` field in GeDR3mock, which is a model of the expected uncertainties in the EDR3 parallaxes. As the data are drawn from the same distance distribution and CQD from which the prior was constructed, this is a somewhat optimistic test, despite the noise. Unless noted otherwise, throughout this section the term "fpu" refers to the *true* fractional parallax uncertainty, computed using the true parallax.

3.1. Example Posteriors

Figure 6 shows examples of both types of posterior compared to their priors. At small fpu, for example, panels (a) to (c), the two posteriors are very similar, with a median (and mode) near the true distance, shown as the vertical line. As long as the fpu is not too large, the prior plays little role and the posterior can be quite different, for example, panel (d), although this can also occur at larger fpu, for example, panels (i) and (l). Panel (f) shows a multimodal photogeometric prior and posterior. The two types of prior sometimes disagree, as can the posteriors. In panel (h), which is for a 30% parallax uncertainty, the geometric posterior is more consistent with the true distance. Note that the parallax that the algorithm sees does not correspond to the vertical line, so for large fpu we cannot expect either posterior to peak near this. Panel (k) shows a multimodal posterior in which the true distance is close to a smaller mode. This happens here because the parallax has 50% noise, so the measured parallax corresponds to a smaller distance (where both geometric and photogeometric posteriors peak). At larger fpu—the bottom row is all for more than 1.0—the photogeometric prior is often more consistent with the true distance than the geometric one.

3.2. Comparison to Truth

3.2.1. Qualitative Analysis

Distance inference results for two HEALpixels are shown in Figures 7 and 8. We see a good correlation between the inferred and true distances out to several kiloparsecs (left columns). The degradation at larger distances is mostly due to stars with larger fpu, as can be seen in the middle columns of these figures. The fractional residual is defined as the estimated minus true distance, divided by the true distance. Note that these middle columns show the *true* fpu, as computed from the noise-free parallax, which is not the same as the *measured* (noisy) fpu that the inference algorithm encounters. (See Appendix B for a consequence of this difference.) At large fpu, the photogeometric distances perform better than the geometric ones, because even when the parallax is of limited use, there is still distance information from the color and magnitude via the Q_G model. For geometric distances, in contrast, as the measured fpu increases, the distance prior dominates the likelihood, so the median of the posterior is pushed toward the median of the prior. Hence at large fpu, the geometric distances to stars that are truly more distant than the median of the prior will generally be underestimated. Faraway stars tend to have larger fpu than nearby stars, because they have both smaller parallaxes and larger parallax uncertainties (as they are fainter). Thus as a whole, any *underestimation* of geometric distances to stars that are *beyond* the median of the prior will tend to be larger than the *overestimation* of the geometric distances to stars that are *closer* than the median of the prior. This explains why the distributions in the top left panels of Figures 7 and 8 flatten at larger distances. This feature is suppressed in the photogeometric distances (bottom left panels) because for large fpu, the Q_G prior can overrule the geometric prior. We also see more flattening for the low-latitude HEALpixel in Figure 8 than the high-latitude HEALpixel in Figure 7 because the low-latitude HEALpixel has larger fpu values on average.

The right columns of Figures 7 and 8 assess how well the estimated distance uncertainties explain the residuals, by plotting the distribution of residual/uncertainty. This is shown using three different representations of the uncertainty. The upper uncertainty, $r_{\rm hi} - r_{\rm med}$, and symmetrized uncertainty, $(r_{\rm hi} - r_{\rm lo})/2$, shown in blue and black, respectively, yield almost identical distributions. For the high-latitude HEALpixel 6200 (Figure 7), they are quite close to a unit Gaussian, in particular for the photogeometric estimates. The lower uncertainty, $r_{\rm med} - r_{\rm lo}$, shown in orange, is negatively skewed (larger tail to negative values), suggesting that the lower uncertainty measure, $r_{\rm lo}$, is slightly underestimated. This is more noticeable in the low-latitude HEALpixel 7593 (Figure 8), where we also see that the photogeometric estimates are slightly more skewed than the geometric ones.

3.2.2. Quantitative Analysis

To quantify the accuracy of our results, we use the median of the fractional distance residual, which we call the *bias*, and the median absolute of the fractional distance residual, which we call the *scatter*. These are robust versions of the mean and standard deviation, respectively. For normally distributed residuals, the mean equals the median, and the standard deviation is 1.48 times the median absolute deviation.

For HEALpixel 6200, the bias and scatter for the geometric distances over all stars are +0.29e−3 and 0.10, respectively. If we limit the computation of these metrics to the 50% of stars in this HEALpixel with $0 < \sigma_\varpi/\varpi < 0.20$, the bias is +5.3e−3 and the scatter is 0.037. The scatter in this subsample is smaller, as expected. The bias is larger because stars with small fpu tend to be nearer stars, whereas the distance prior is characteristic of all of the stars, which are more distant on average. Hence, the prior pulls up the distances for the small fpu subsample, leading to a more positive bias.

For the photogeometric distances, the bias and scatter over all stars are +5.7e−3 and 0.059, respectively, and for the $0 < \sigma_\varpi/\varpi < 0.20$ subsample are +2.5e−3 and 0.032, respectively. The scatter over the full sample is smaller for the photogeometric estimates than for the geometric ones, because the former benefit from the additional information in the stars' colors and magnitudes. The situation is particularly fortuitous here because of the near-perfect match between the Q_G models and the actual distribution of Q_G in the data. For the full sample, the bias is larger for the photogeometric distances than for the geometric ones, although still small on an absolute scale. For the small fpu subsample, the photogeometric distances are not much more accurate than the geometric ones, because the parallax dominates the distance estimate.

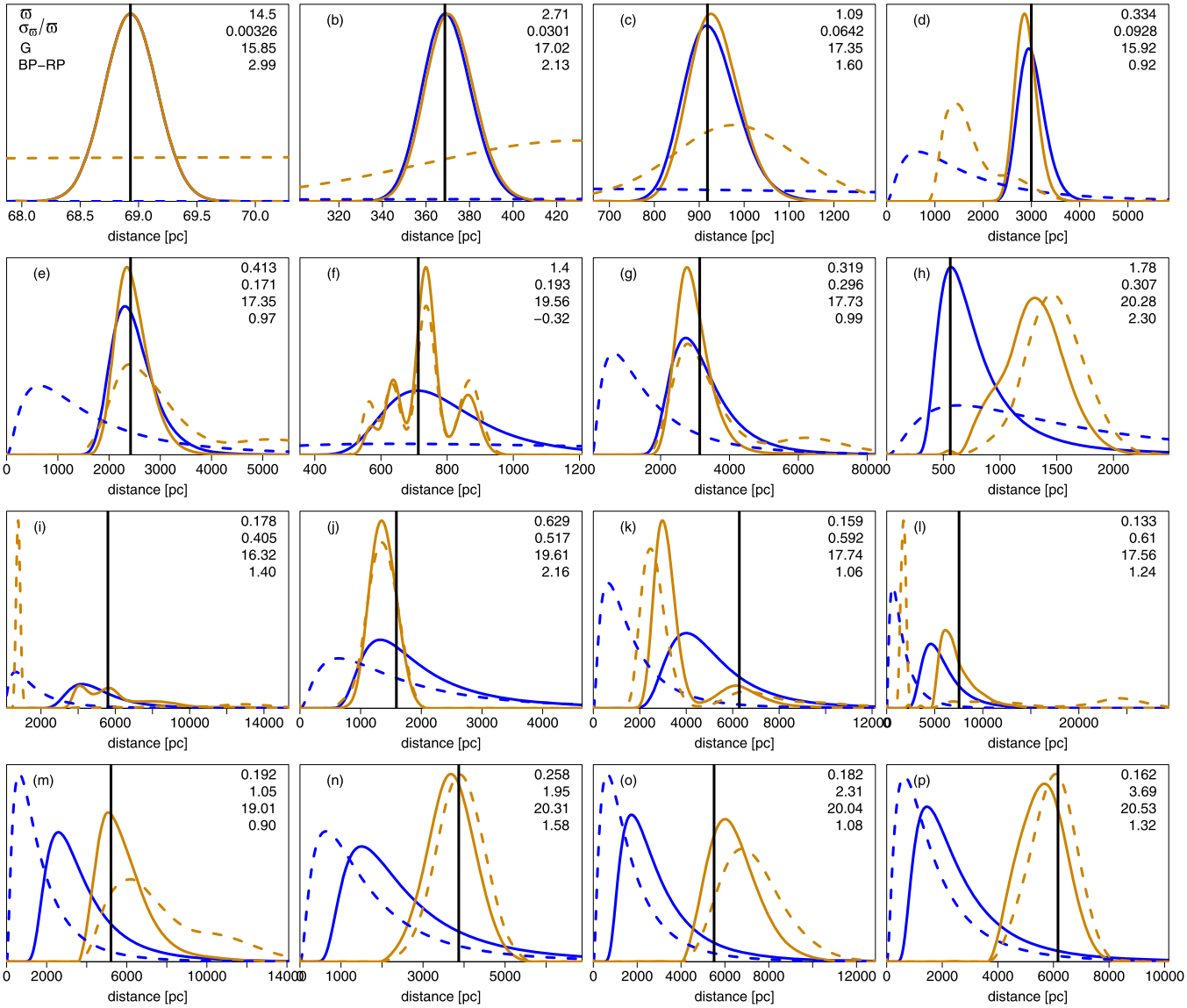

Figure 6. Example normalized posteriors (solid lines) and corresponding normalized priors (dashed lines) for geometric distances (blue) and photogeometric distances (orange) for various stars in the mock catalog (one per panel). These have been selected to show the variety; they are not a random subset. The vertical solid line is the true distance. The inverse of this is not the parallax seen by the inference because noise was added. All stars are from HEALpixel 6200, so the distance prior (blue dashed line) is the same in all panels. The four numbers in the top right corner of each panel are, from top to bottom, ϖ, true fpu, G, and BP − RP. Stars are ordered by increasing fpu. The two posteriors coincide in the top left panel.

Turning now to the low-latitude HEALpixel 7593 (Figure 8), we find that the bias and scatter in the geometric distances over all stars are −0.16e−3 and 0.27, respectively. There are two reasons for the larger scatter in this HEALpixel. The first is that the parallax uncertainties are larger: the median parallax uncertainty is 0.32 mas, as opposed to 0.15 mas in HEALpixel 6200. This in turn is because the stars are on average 0.9 mag fainter in HEALpixel 7593 (one reason for which is the larger extinction, as is apparent from Figure 3). The second reason is that the median true distance to stars is larger in this low-latitude HEALpixel than in the high-latitude one (4.0 kpc versus 1.2 kpc; see Figure 1). This may seem counterintuitive but is a consequence of distant disk (and bulge) stars at low latitudes that remain visible to larger distances despite the higher average extinction. At higher latitudes, in contrast, there are no distant disk stars and hardly any halo stars (which are scarce in Gaia anyway). Both of these facts contribute to the larger fpu in the low-latitude pixel (median of 1.18, central 90% range of 0.21–3.57) than in the high-latitude HEALpixel (median of 0.20, central 90% range of 0.03–1.08). Even if we look at just the 9% of stars in the low-latitude HEALpixel with $0 < \sigma_\varpi/\varpi < 0.20$, we get a bias and scatter of +25e−3 and 0.069, respectively, which are still significantly worse than the higher latitude HEALpixel for the same fpu range.

Concerning the photogeometric distances in HEALpixel 7593, the bias and scatter for all stars are −3.8e−3 and 0.17, respectively, and for the $0 < \sigma_\varpi/\varpi < 0.20$ subsample are +20e−3 and 0.062, respectively. For the full sample, we again see a

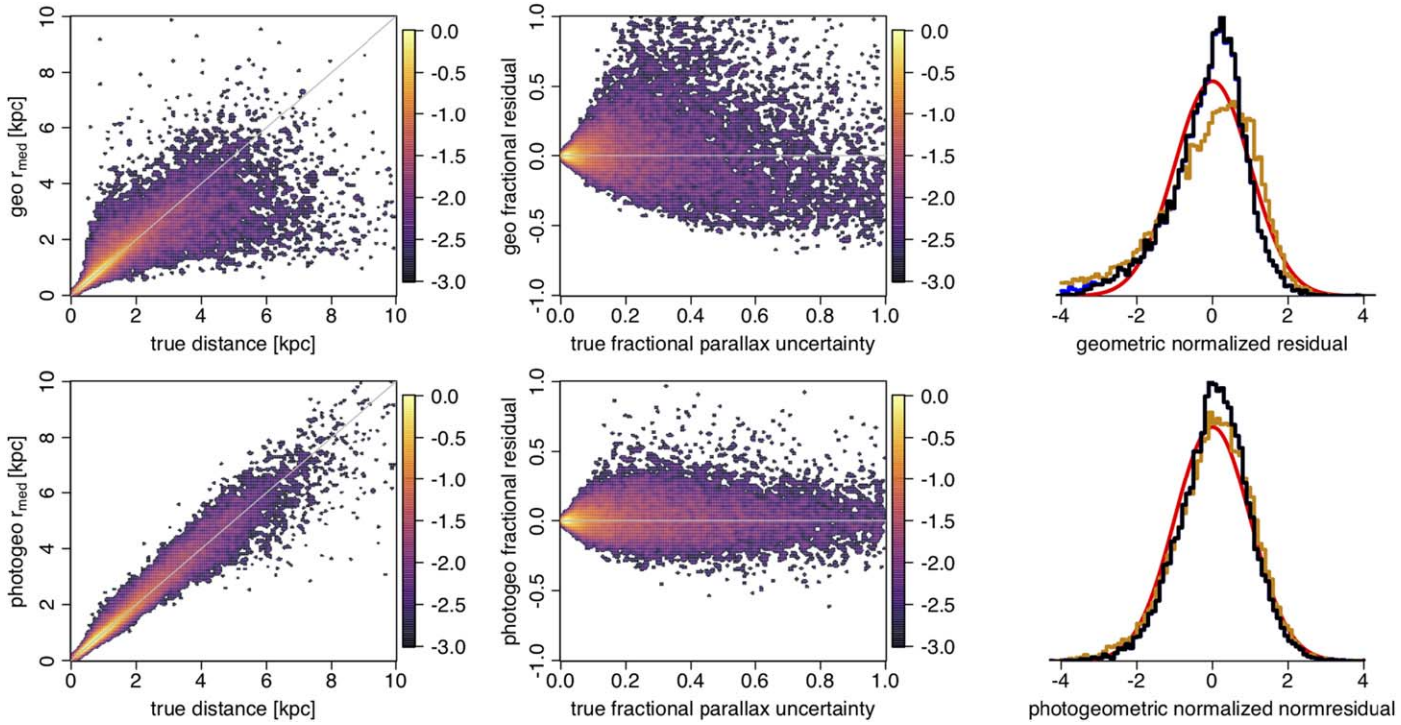

Figure 7. Results of the distance inference on mock catalog HEALpixel 6200. The top row shows geometric distances, the bottom row photogeometric ones. The left column compares the inferred distances (vertical axis) to the true distances for all sources. This covers the full range of fractional parallax uncertainties, which has a median of 0.20 and a central 90% range of 0.03–1.08. The middle column shows the fractional distance residuals as a function of the true fractional parallax uncertainty (fpu). In these first four panels, the color scale is a logarithmic density (base 10) scale relative to the highest density cell in each panel. The right column shows the normalized residuals: the difference between the inferred and the true value, divided by an uncertainty measure. The three colors refer to three uncertainty measures: orange is $r_{\rm med} - r_{\rm lo}$, blue is $r_{\rm hi} - r_{\rm med}$, and black is $(1/2)(r_{\rm hi} - r_{\rm lo})$. The blue and black lines virtually coincide. The smooth red curve is a unit Gaussian for comparison.

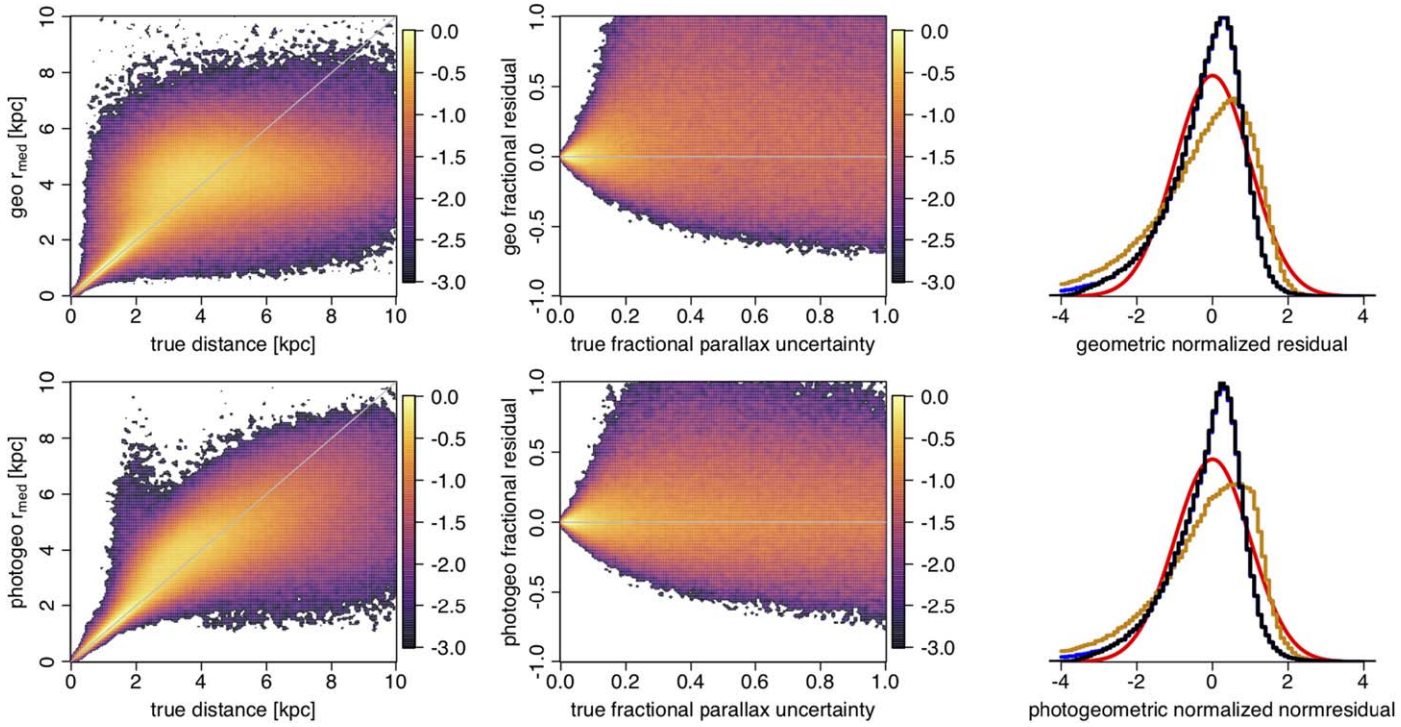

Figure 8. Same as Figure 7 but now for HEALpixel 7593. The median fpu is 1.18 and the 90% range is 0.21–3.57.

significant decrease in the scatter compared to the geometric distances. In a real application, we may get less benefit from the Q_G prior at low latitudes because our model CQD may differ from the true (unknown) CQD more than at high latitudes, on account of the increased complexity of the stellar populations and interstellar extinction near the Galactic plane.

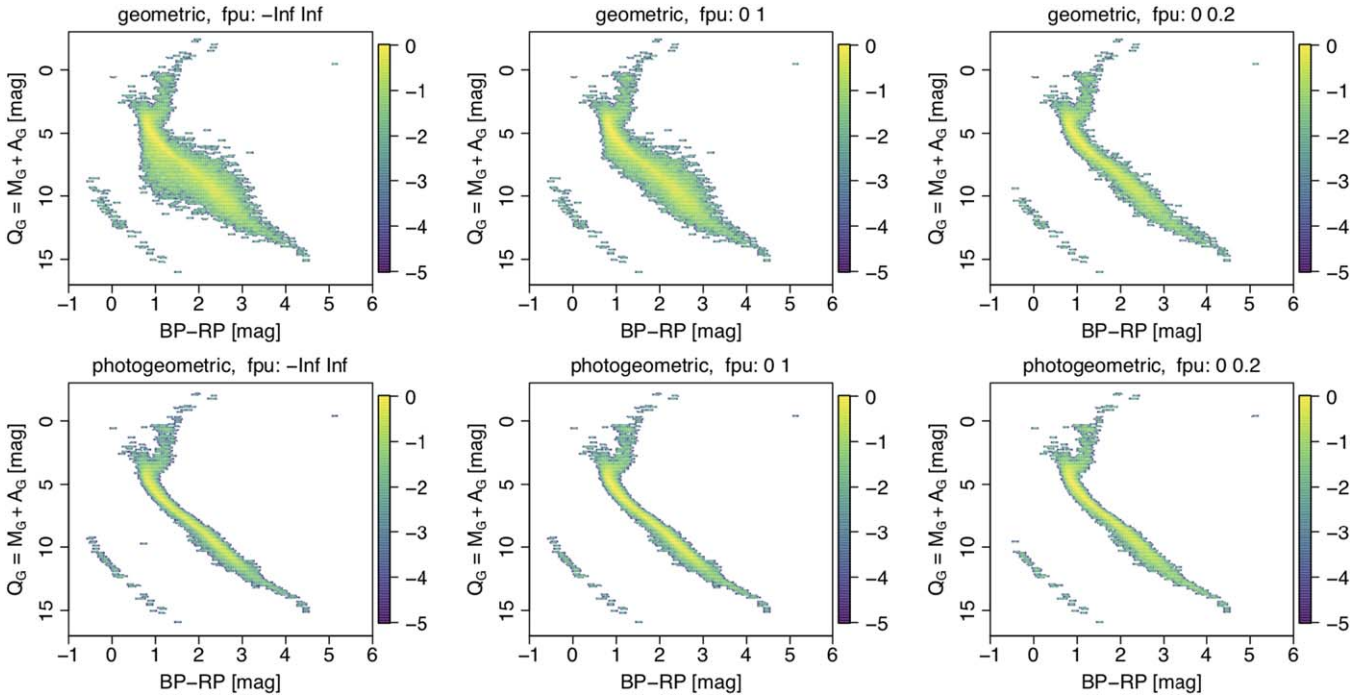

Figure 9. The CQD inferred for mock catalog HEALpixel 6200 using the median geometric distance (top row) and median photogeometric distance (bottom row) for three ranges of the true fractional parallax uncertainty (fpu): all (left), 0–1.0 (middle), and 0–0.2 (right). The color scale is a logarithmic (base 10) density scale relative to the highest density cell in each panel.

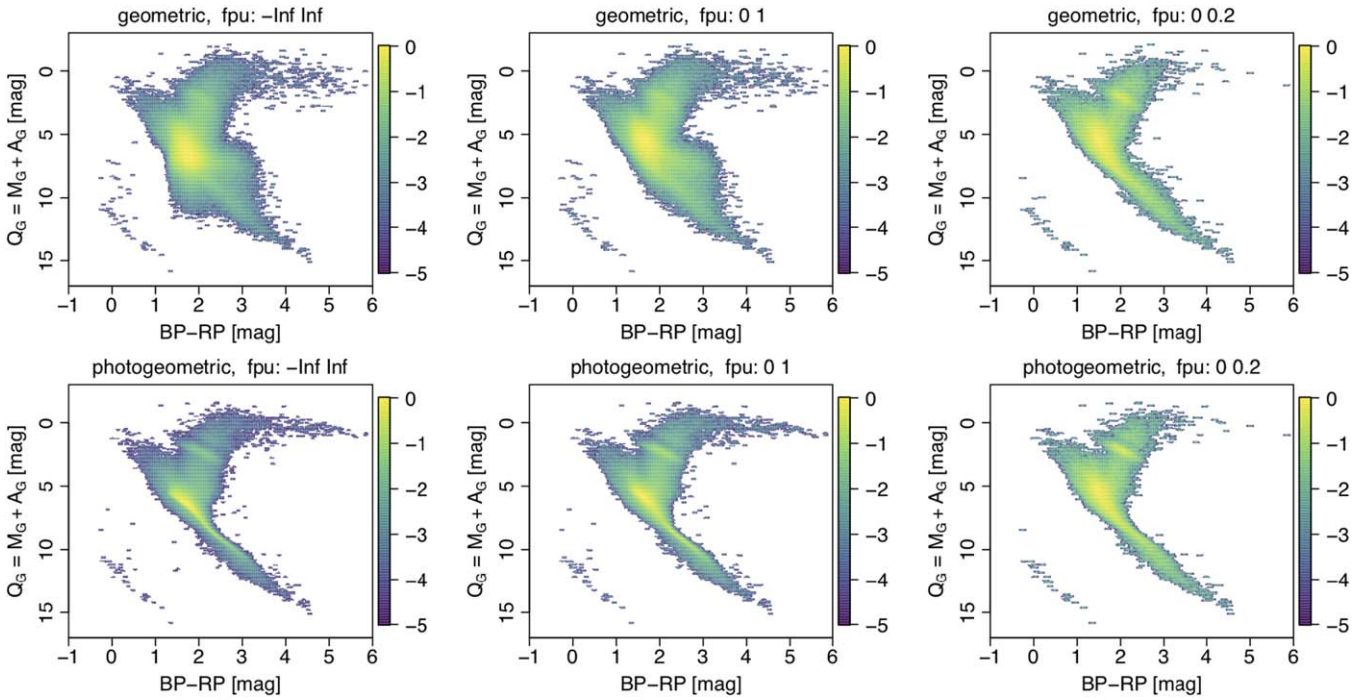

Figure 10. Same as Figure 9 but now for HEALpixel 7593.

3.3. Inferred CQDs

We can also assess the quality of our distance estimates by computing $Q_G = G - 5\log_{10} r_{\rm med} + 5$ and plotting the resulting CQD. We do this for both the geometric and photogeometric distances, for three ranges of fpu, for HEALpixel 6200 in Figure 9 and HEALpixel 7593 in Figure 10. These can be compared to the CQD for the same HEALpixels constructed using the true distances shown in Figure 3. Imperfect distance estimates can only move sources vertically in this diagram as the BP − RP colors are not changed. We see how the inferred main sequence is wider for the larger fpu samples for the geometric distances (left two columns in both plots), but much less so for the photogeometric distances. This is again due to the stabilizing influence of the Q_G prior. Both distance estimates are able to recover the primary structures: the main sequence, white dwarf sequence, giant branch, and horizontal

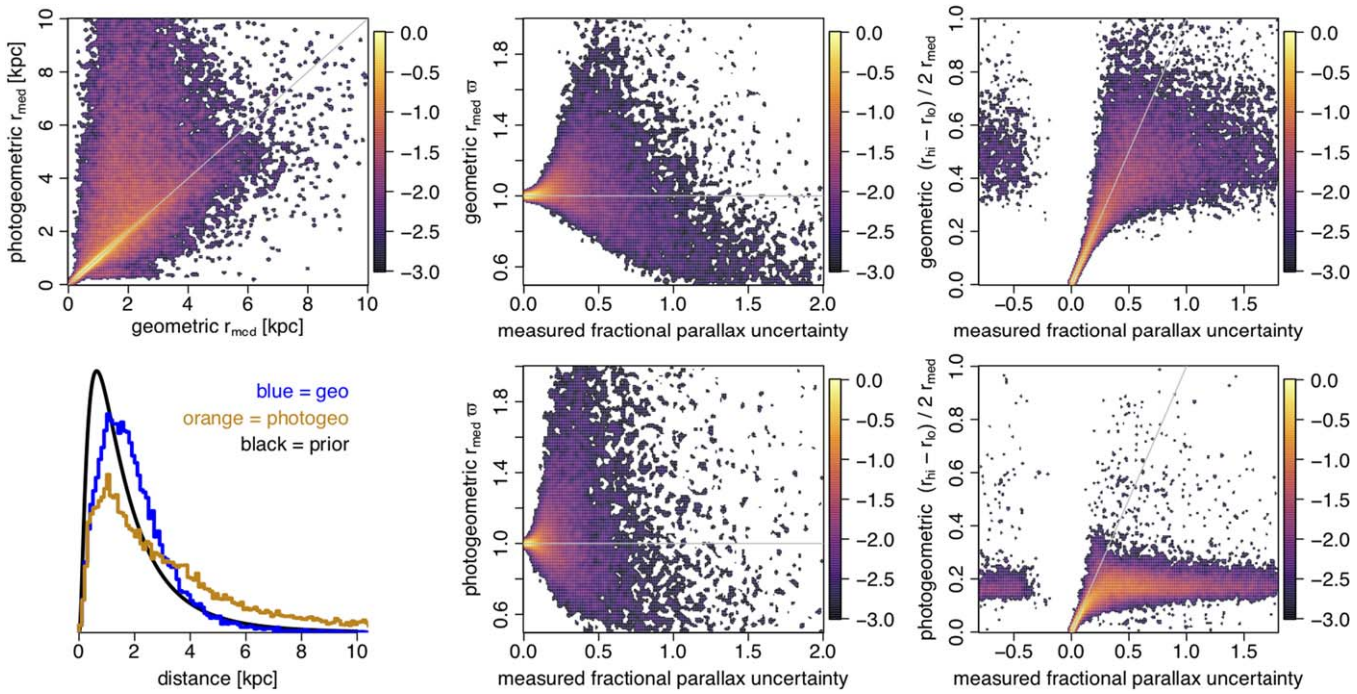

Figure 11. EDR3 distance results for HEALpixel number 6200 at $(l, b) = (285.^\circ7, 34.^\circ8)$. The color scale in the density plots is logarithmic (base 10) relative to the highest-density cell in each panel. The top left panel compares the median geometric and photogeometric distances. The bottom left panel shows normalized histograms on a linear scale of the median geometric (blue) and photogeometric (orange) distances, compared to the distance prior (black). The middle column shows the ratio of the inferred distance to the inverse parallax distance as a function of the measured fractional parallax uncertainty (fpu). Note that the apparent lack of sources in the lower panel at fpu values above about 1.0 is mostly a plotting artifact: regions too low in the density of sources are white. The two panels in the right column show the fractional symmetrized distance uncertainty also as a function of fpu (note the different scales). This plot is available for all HEALpixels with the auxiliary information online.

branch. These plots will be useful when it comes to analyzing the results on the real EDR3 data, because they do not involve the truth as a reference.

4. Analysis of Distance Results in EDR3

We applied our inference code (written in R) to the 1.47 billion sources in Gaia EDR3 that have parallaxes. This required 1.6×10^{12} evaluations of the posteriors and took 57,000 CPU core hours. Throughout this section, the term "fpu" of course refers to the *measured* fractional parallax uncertainty, as we do not know the true parallax.

4.1. Analysis of Two HEALpixels

4.1.1. Distance Distributions and Uncertainties

Results for our two example HEALpixels are shown in Figures 11 and 12. The two panels in the left column compare the two types of distance estimates. As expected, the photogeometric estimates extend to larger distances (see Section 3.2.1 for an explanation). The middle columns plot the ratio of the inferred distance to the inverse parallax distance (corrected for the zero-point). The latter is of course generally a poor measure of distance because it is not the true parallax, and this is the whole point of using an appropriate prior (see Section 1 and references therein). We see that both of our distance estimates converge to $1/\varpi$ in the limit of small fpu. Although the apparent lack of sources at large fpu in the lower middle panels is primarily a plotting artifact (due to the finite density scale), the two samples in the upper and lower panels are not identical because not all sources have photogeometric distances. For HEALpixel 6200, there are 24,007 sources with geometric distances and 23,829 with photogeometric distances. For HEALpixel 7592, these numbers are 385902 and 369608, respectively.

The panels in the right columns of Figures 11 and 12 show how the fractional symmetrized distance uncertainty varies with fpu. At small (positive) fpu, they are nearly equal for both geometric and photogeometric distances, because here the likelihood dominates the posterior. At larger fpu, the geometric distances become more uncertain, which is commensurate with their lower expected accuracy. For very large fpu ($\gg$1), the geometric distances and their uncertainties will be dominated by the prior, which for HEALpixel 7593 has a median of 3.98 kpc and lower (16th) and upper (84th) quantiles of 2.06 kpc and 6.74 kpc, respectively (corresponding to a fractional distance uncertainty of 0.59). The photogeometric fractional distance uncertainties tend to be smaller than the geometric ones. This is because the Q_G prior (Section 2.5) is usually more informative than the distance prior.

We extend the axes in the right panels of Figures 11 and 12 to negative fpu values, which occur when sources have negative parallaxes. One of the advantages of probabilistic inference is to provide meaningful distances for negative parallaxes (a quarter of all parallaxes in EDR3). Negative observed parallaxes usually correspond to sources with small true parallaxes, and although such measurements generally have reduced impact on the posterior, they do carry information. They do not yield precise distances, but insofar as the prior can be trusted, the posterior and resulting confidence intervals are meaningful. We see from the figures that the precisions are low for both types of distance, but sometimes more constrained for the photogeometric ones due to the

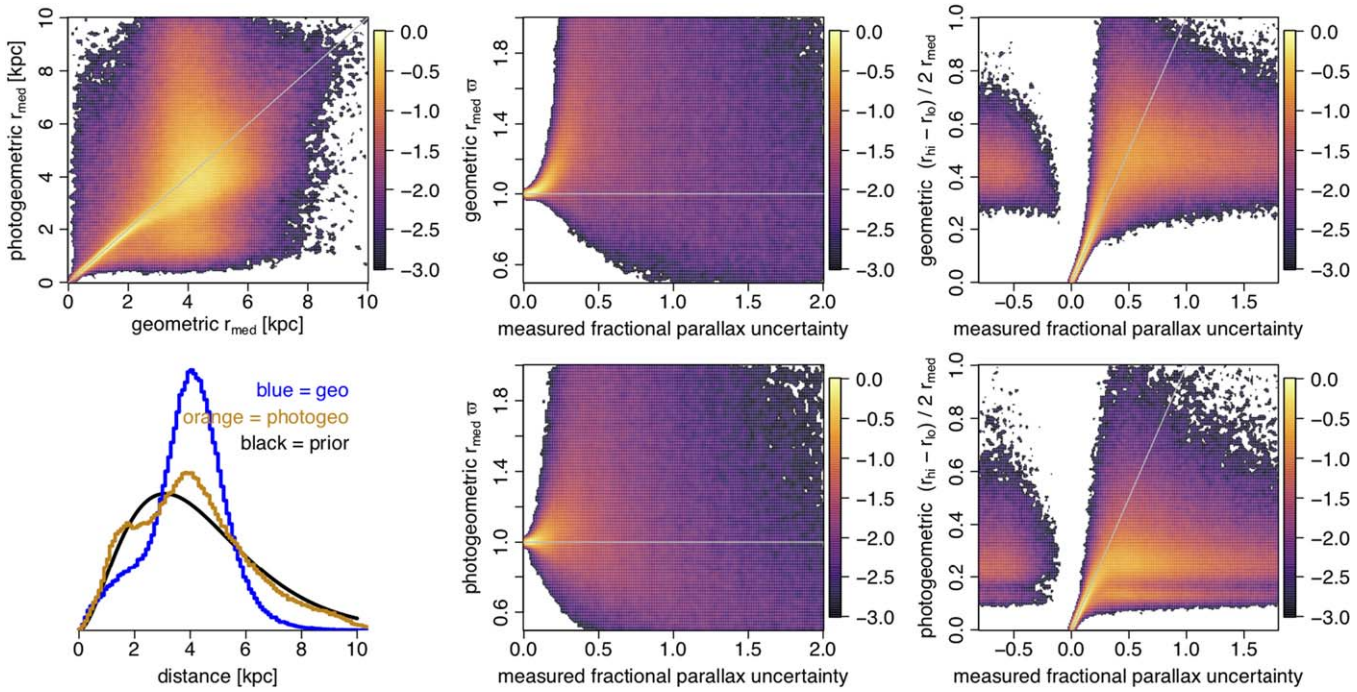

Figure 12. Same as Figure 11 but for HEALpixel number 7593 at $(l, b) = (29.^{\circ}0, 7.^{\circ}7)$.

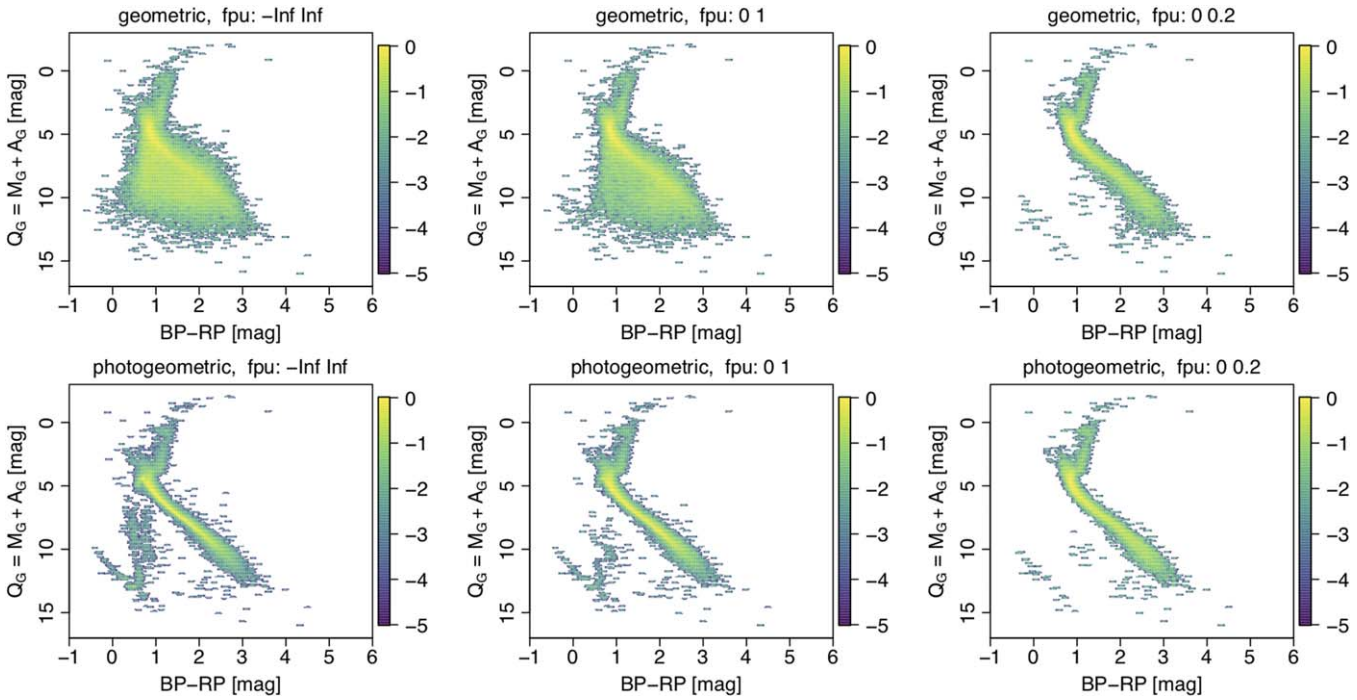

Figure 13. CQD inferred for EDR3 HEALpixel 6200 using the median geometric distance (top row) and median photogeometric distance (bottom row) for three ranges of the measured fpu: all (left), 0–1.0 (middle), and 0–0.2 (right). In total, there are 24,007 sources with geometric distances and 23,829 with photogeometric distances. No other filtering has been applied. The color scale is a logarithmic (base 10) density scale relative to the highest density cell in each panel, so it is not comparable across panels. This plot (including also a comparison with the prior CQD) is available for all HEALpixels with the auxiliary information online.

additional use of color and magnitude. In some senses, the negative fpu regime is a continuation of the $\sigma_{\varpi}/\varpi \gg 1$ regime (see Figures 3 and 6 of Paper II).

4.1.2. *Color–Q_G Diagrams*

From the inferred median distances we can compute the median Q_G via Equation (4) and then plot the CQD. This is shown in Figure 13 for HEALpixel 6200 for the geometric distance (top row) and photogeometric distance (bottom row) for three different ranges of the fpu. Interstellar extinction should be low toward this high-latitude field (around 0.15 mag in GeDR3mock), $Q_G \simeq M_G$, so this CQD is similar to the color–absolute magnitude diagram. In all of the panels, we see a well-defined main sequence and giant branch, as well as a white dwarf sequence in some of the panels. Comparing the

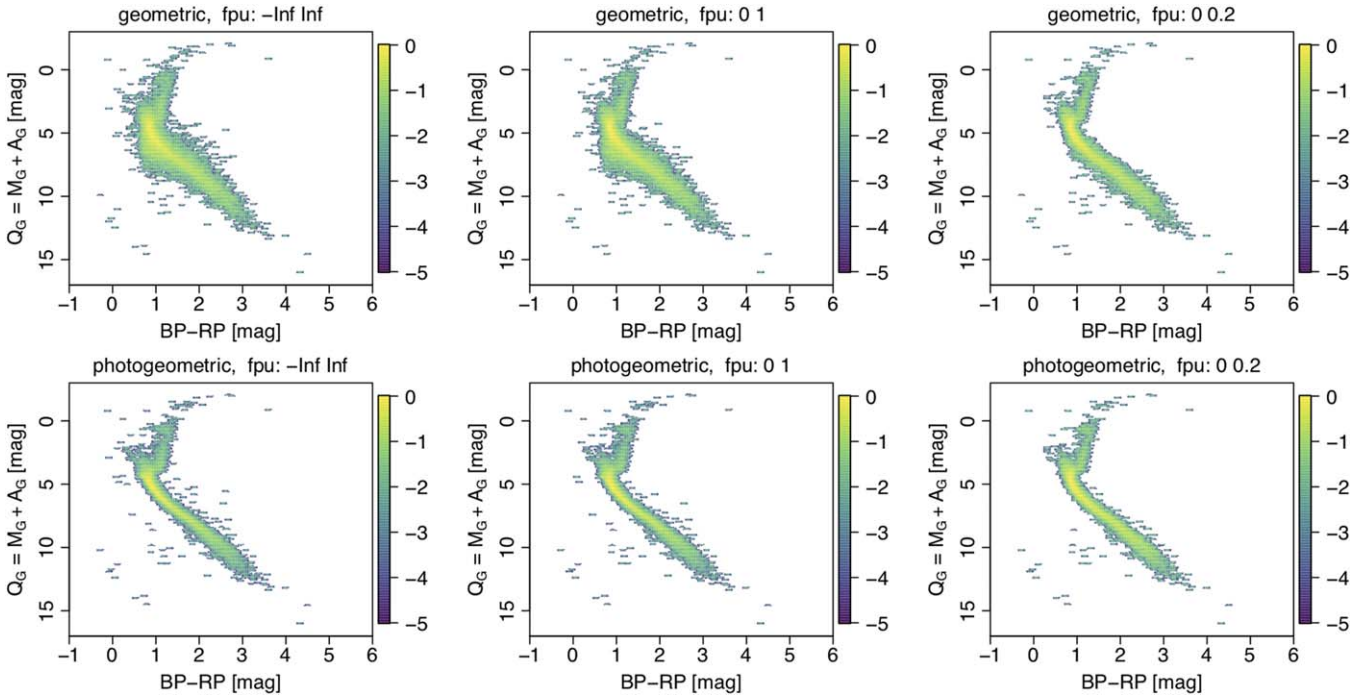

Figure 14. Same as Figure 13 but now excluding the 54% of sources in this HEALpixel with $G > 19.0$ mag.

upper and lower panels, we see how the photogeometric distances constrain the Q_G distribution more than the geometric distances do. The puffing-up of the geometric CQD is due to sources with large fpu: their distances tend to be underestimated (see Section 3.2.1) so Q_G becomes larger—intrinsically fainter—for a given G (see Equation (4)). This puffing-up diminishes as we successively reduce the range of fpu, as shown in the middle and right columns of Figure 13.

The photogeometric CQD for the full fpu range (bottom left panel of Figure 13) shows a conspicuous blob of sources at $\mathrm{BP} - \mathrm{RP} \simeq 0.5$ mag between the MS and WD sequences. These are sources with spuriously large parallaxes, well known from GDR2 (Arenou et al. 2018) and still present, if less so, in EDR3 (Fabricius et al. 2020; Gaia Collaboration 2020b). They are usually close pairs of sources that receive incorrect astrometric solutions, as the EDR3 astrometric model is only suitable for single stars (Lindegren et al. 2020a). Figure 13 shows that spurious parallaxes are less common among the smaller fpu subsample. The Q_G prior will often help to constrain the distance of these spurious solutions and thus place them on the correct part of the CQD. This is only partially successful at around $\mathrm{BP} - \mathrm{RP} \simeq 0.5$ mag in this HEALpixel, however, because the distance prior may still be pulling truly very distant sources with larger fpu toward us.

Sources with spurious parallaxes are preferentially faint. To quote from Gaia Collaboration (2020a): "For faint sources ($G > 17$ for 6p astrometric solutions and $G > 19$ for 5p solutions) and in crowded regions the fractions of spurious solutions can reach 10% or more." This can be seen in Figure 14, where we replot the CQD only for sources with $G < 19.0$ mag. This also reduces the puffing-up of the geometric CQD, although some of this reduction occurs simply because magnitude is correlated with fpu, so a magnitude cut also lowers the fpu.

These effects can be seen more prominently in the low-latitude HEALpixel 7593, shown in Figures 15 and 16. Due to the larger mean distance of stars at low latitudes (see Section 3.2.2), as well as the more complex stellar populations and larger mean extinction (up to 3.5 mag), the CQD is more complex. For the full fpu range, the geometric CQD in Figure 15 is quite washed out, due in part to large fpu values and spurious parallaxes, although an extincted red clump is visible. The photogeometric CQDs are cleaner, with a better-defined main sequence. The CQD for the $G < 19.0$ mag subsample (Figure 16) again shows the removal of spurious sources. Section 3.2 of Fabricius et al. (2020) analyzes spurious astrometric solutions and offers more sophisticated ways of identifying them than a simple magnitude cut.

4.2. All Sources

We now look at a representative sample of the entire catalog. All plots and analyses in this section use a random selection of 0.5% of all sources from each HEALpixel. This has 7,344,896 geometric and 6,739,764 photogeometric distances.

Figure 17 shows the distribution of distances. As expected, the photogeometric distances extend to larger distances than the geometric one. The fractional symmetrized distance uncertainties as a function of distance are shown in Figure 18 for three different magnitude ranges. As noted earlier, the photogeometric distance uncertainties are generally smaller than the geometric ones, at least for fainter sources. This plot also shows again that photogeometric estimates extend to larger distances.

4.2.1. Color–Q_G Diagrams

Figure 19 shows the CQD over the whole sky. Because the sample is a constant random fraction per HEALpixel, it is numerically dominated by sources at low-latitude Galactic latitudes where there can be significant interstellar extinction. This is apparent from the upper diagonal feature—especially clear in the photogeometric panel—which is the red clump stretched by extinction/reddening. The white dwarf sequence

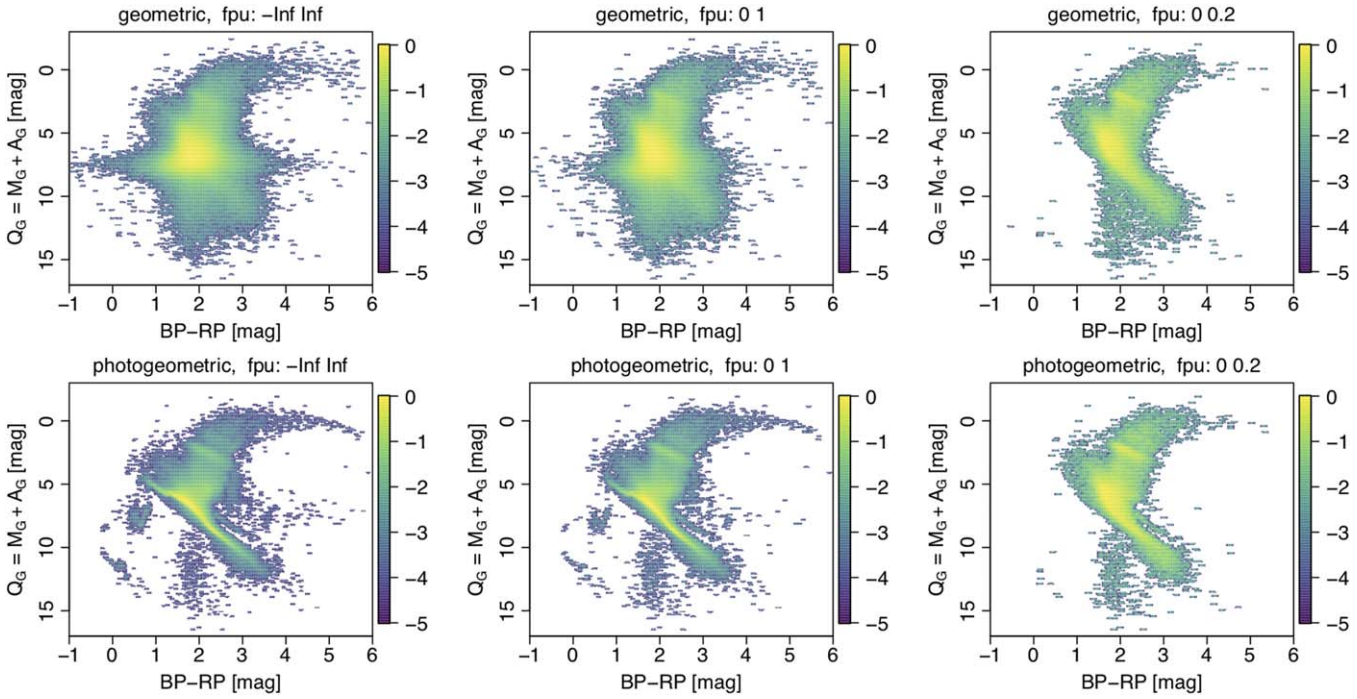

Figure 15. Same as Figure 13 but now for HEALpixel 7593. All sources are shown (no magnitude cut). In total, there are 385,902 sources with geometric distances and 369,608 with photogeometric distances.

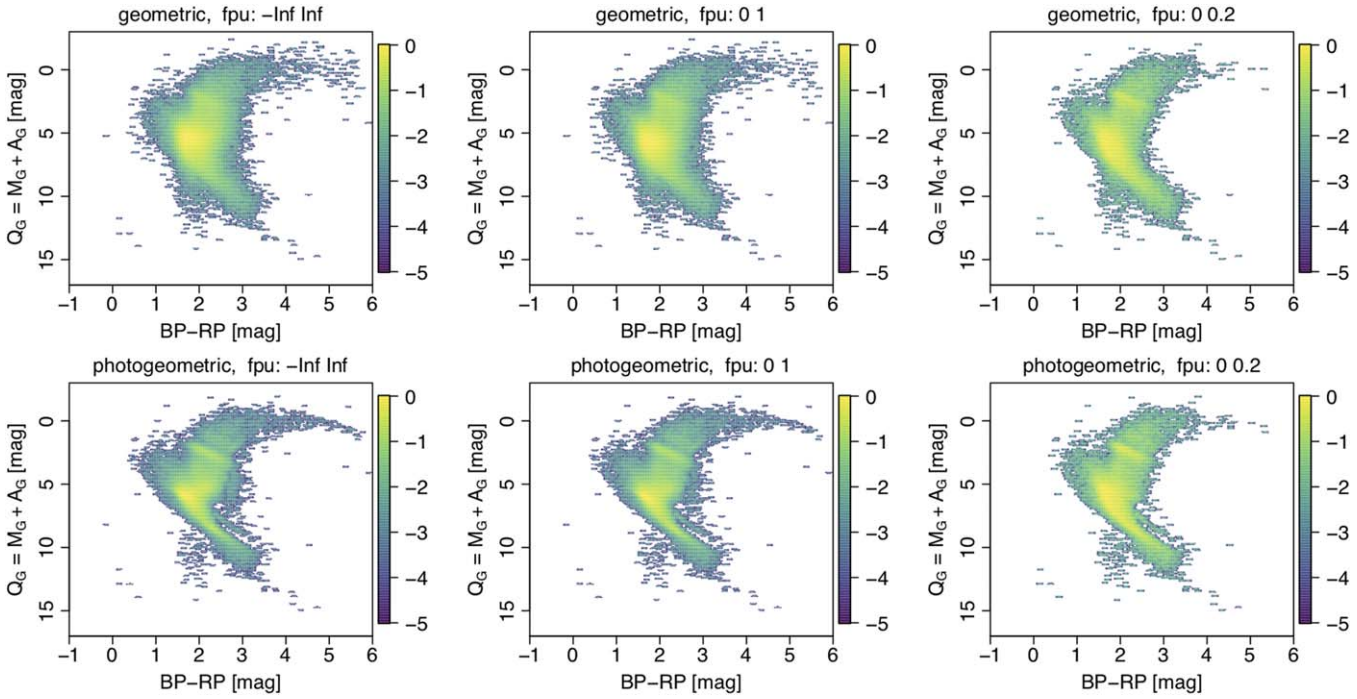

Figure 16. Same as Figure 15 but now excluding the 70% of sources in this HEALpixel with $G > 19.0$ mag to remove spurious sources.

appears clearly in the photogeometric CQD. Although some white dwarfs are correctly placed in the CQD by the geometric distances, they are not visible here, due to the finite dynamic range of the plotted density scale. Furthermore, for reasons explained in Section 3.2.1, faint nearby sources with large fpu tend to have their geometric distances overestimated and therefore their Q_G underestimated, thereby pushing them up from the true white dwarf sequence. These plots have not filtered out spurious sources, some of which are clearly visible in the photogeometric CQD as the blob between the upper MS and the white dwarf sequence. Other broad differences between the geometric and photogeometric CQDs were explained in Section 3.3.

4.2.2. Distribution on the Sky

Figure 20 shows the mean distance of sources (i.e., mean of $r_{\rm med}$) in each HEALpixel in our catalog, as well as the ratio of these in log base 2. Over all HEALpixels, the 5th, 50th, and

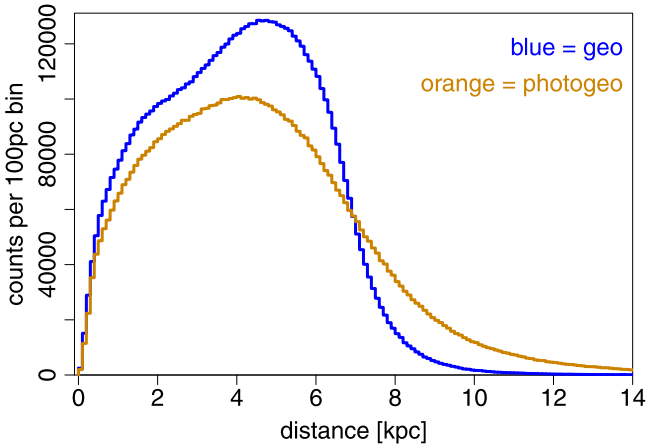

Figure 17. Distribution of inferred geometric and photogeometric median distances, $r_{\rm med}$, in EDR3. This plot uses a random sample of 0.5% of all sources in each HEALpixel.

95th percentiles of the mean of the geometric distances are 1.3, 2.1, and 4.4 kpc, respectively. The percentiles for the mean of the photogeometric distances are 2.2, 3.3, and 5.0 kpc. These translate into low ratios of geometric to photogeometric distances in general. Only in the Galactic plane and the bulge are the two mean distances comparable. At high Galactic latitudes, the photogeometric average is easily twice as large as the geometric average.

4.2.3. Galactic Spatial Distribution

Figure 21 shows the projected distribution of stars in EDR3 in the Galaxy using our distance estimates. The Sun is at the origin, and we see the expected larger density of sources in the first and fourth Galactic quadrants. Finer asymmetries in the distribution projected onto the Galactic plane (upper panels) are presumably due to both a genuine asymmetry in the Galactic population and Gaia's scanning law. These, as well as nearby dust clouds, also explain the various radial lines pointing out from the origin. The lack of sources in the fan around the positive x-axis in the lower panels is due to extinction in the Galactic plane. The overdensity in the same direction in the upper panels is the projection of the bulge. The lower panels demonstrate the point made earlier (Section 3.2.2) about being able to see sources to larger mean distances at lower Galactic latitudes.

The high-density rays extending below the Galactic plane (lower panels of Figure 21) are in the directions of the Magellanic Clouds. Many stars in these satellite galaxies are in EDR3—they are some of the densest HEALpixels—yet they are so far away (50–60 kpc) that most have poor (and often negative) parallaxes, such that the inferred geometric distances are dominated by the prior (see Appendix B for further discussion). Our photogeometric distances are similarly poor, because we excluded the Magellanic Clouds from the mock CQD out of which our Q_G priors are built. This was intentional: anyone interested in estimating distances to sources in the Magellanic clouds can do better than just use Gaia parallaxes and photometry.

Figure 22 shows the fractional distance uncertainties also in Galactic projection. As expected, the uncertainties generally increase with distance from the Sun, but there are exceptions due to bright, distant stars having more precise distances than faint, nearby ones. The rays toward the Magellanic clouds also stand out as having larger uncertainties on the whole.

4.3. Validation Using Clusters

Figures 23 and 24 show our geometric and photogeometric distances and their uncertainties for members of various star clusters. The membership lists have been drawn from Paper IV. NCG6254 (=M10) and NGC 6626 (=M28) are globular clusters; the rest are open clusters. Recall that our prior does not include star clusters. The horizontal dashed line in each panel shows the inverse of the variance-weighted mean parallax of the members, that is, a pure parallax distance for the cluster. Both of our distance estimates congregate around this for small, positive fpu, but deviate for large or negative fpu, as one would expect. We generally see a larger deviation or scatter for the geometric distance: compare in particular the panels for NGC 2437 (=M46) and NGC 6254. Despite this, the weighted mean of our distances is often quite close to the pure parallax distance, even for clusters up to several kiloparsecs away.

We nevertheless emphasize that the inverse of the variance-weighted mean parallax will usually be a better estimate for the distance to a cluster than the mean of our distances. This is because any combination of our individual distances will reuse the same prior many times. If stars have large fpu values, this product of priors will dominate and introduce a strong bias into the combined distance. This would particularly affect clusters beyond a few kiloparsecs.

4.4. Comparison to Other Distance Estimates

Figure 25 compares our distance estimates for 36,858 red clump (RC) stars with those estimated by Bovy et al. (2014) using high-resolution Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017) DR16 spectra. This method selects sources using color, effective temperature, metallicity, and surface gravity and is calibrated via stellar evolution models and high-quality asteroseismology data. Given the narrowness of the red clump locus in the parameter space, their distances are expected to be precise to 5% with a bias of no more than 2%.

The 5th, 50th, and 95th percentiles of fpu for this sample are 0.01, 0.05, and 0.27, respectively, and of G are 10.4, 13.4, and 16.2 mag, respectively. The fractional bias and rms of the deviations of our estimates relative to those of Bovy et al. are +0.05 and 0.31, respectively, for the geometric distances and +0.03 and 0.29, respectively, for the photogeometric distances. For reference, the fractional bias and rms of the deviations of the APOGEE red clump estimates relative to the StarHorse (Queiroz et al. 2020) estimates (see next paragraph) for the same sample are +0.05 and 0.21, respectively. The parallaxes for this sample are mostly of such high quality that the prior does not strongly affect our posteriors, although we still see a slight improvement in the photogeometric distances over the geometric ones. When counting the percentage of sources where the Bovy et al. estimate is within our upper and lower bounds (+7% error margin from Bovy et al.), we find that 65% are compatible with the geometric distances and 69% with photogeometric (we expect 68% to be within 1σ). If we do the same for the StarHorse estimates (which also have upper and lower percentiles) for the red clump sample, we see that 84% of the StarHorse estimates are within 1σ of the Bovy et al. estimates.

Figure 26 compares our distance estimates for 307,105 stars with those estimated by Queiroz et al. (2020) using their StarHorse method, which uses APOGEE DR16 spectra, multiband photometry, and GDR2 parallaxes. This sample

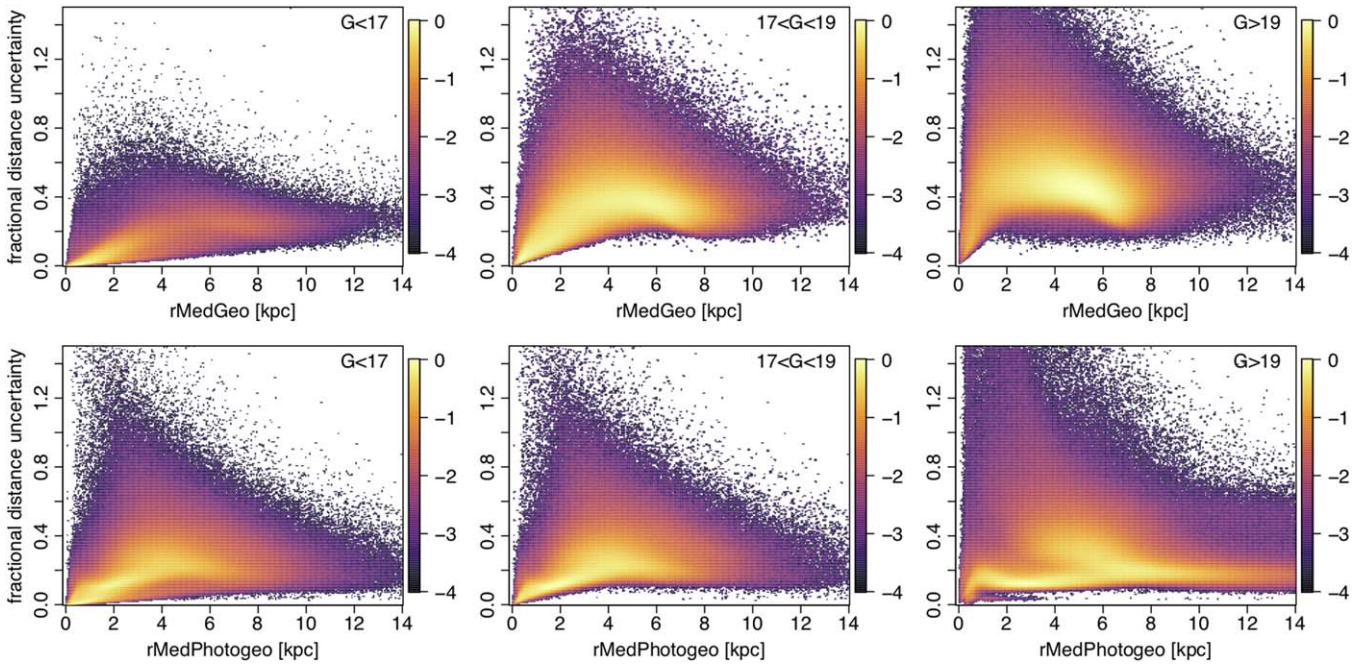

Figure 18. Fractional symmetrized distance uncertainty, $(r_{\rm hi} - r_{\rm lo})/2r_{\rm med}$, vs. distance for the geometric distance estimates (top) and photogeometric distance estimates (bottom) for the three different G ranges. The color scale is a logarithmic density (base 10) scale relative to the highest-density cell in each panel. This plot uses a random sample of 0.5% of all sources in each HEALpixel.

comprises around one-third main-sequence stars; the rest are turnoff stars and giants, excluding the red clump stars used in the previous comparison. StarHorse estimates a posterior probability distribution that the authors likewise summarize with a median, so our distance estimates are directly comparable. They report achieving typical distance uncertainties of 11% for giants and 5% for dwarfs.

The 5th, 50th, and 95th percentiles of fpu for this sample are 0.002, 0.02, and 0.46, respectively, and of G are 10.2, 13.3, and 16.6 mag, respectively. The fractional bias and rms of the deviations of our distance estimates relative to the StarHorse estimates are 0.00 and 0.30, respectively, for the geometric distances and −0.01 and 0.23, respectively, for the photogeometric distances. As this sample extends to larger distances (and larger fpu) than the sample in Figure 25, we begin to see that our geometric distances (and to a lesser extent our photogeometric distances) are smaller than the StarHorse distances beyond about 6 kpc, which is where some of the large fpu sources will have true distances beyond the median of the distance prior.

5. Distance Catalog

5.1. Content

The distance catalog includes an entry for all 1,467,744,818 sources in EDR3 that have a parallax. All of these have geometric distances, and 1,346,621,631 have photogeometric distances. In comparison, there are 1,347,293,721 sources in EDR3 that have defined G-band magnitudes,[7] BP − RP colors, and parallaxes and so could in principle have received a photogeometric distance estimate, but did not due to missing Q_G prior models.

[7] By this we mean the `phot_g_mean_mag` field is defined. We do not make use of the other estimates of G from the Gaia catalog if this field is null.

The fields in our catalog are defined in Table 1. Note that 3% of the sources have changed their `source_id` identifier from GDR2 to EDR3 (Fabricius et al. 2020), so the `source_id` cross-match table `dr2_neighbourhood` provided with EDR3 should be used to find the best match before doing source-by-source comparisons between the two releases. `r_med_geo` in Table 1 is the median ($r_{\rm med}$) of the geometric distance posterior and should be taken as the geometric distance estimate. `r_lo_geo` ($r_{\rm lo}$) and `r_hi_geo` ($r_{\rm hi}$) are the 16th and 84th percentiles of the posterior and so together form a 68% confidence interval around the median. $r_{\rm hi} - r_{\rm med}$ and $r_{\rm med} - r_{\rm lo}$ are therefore both 1σ-like uncertainties on the distance estimate and are generally unequal due to asymmetry of the posterior. The fields `r_med_photogeo`, `r_lo_photogeo`, and `r_hi_photogeo` are defined in the same way for the photogeometric distance posterior.

We cannot overstate the importance of the uncertainties provided. They reflect the genuine uncertainty in the distance estimate provided by the median. As $r_{\rm hi} - r_{\rm lo}$ is a 68% confidence interval, we expect the true distance to lie outside of this range for one-third of the sources. This is the nature of statistical uncertainty and should never be ignored.

The field `flag` is a string of five decimal digits defined in Table 2. Flag `A` is set to 2 if the source is fainter than the faintest mock source used to make the prior for that HEALpixel. The estimated distances can still be used. Faint stars tend to have poor parallaxes, so the distance uncertainties will generally be larger in these cases. The two digits of flag `B` refer to the Hartigan dip test, as explained in Section 2.6.2. We find that 2% of geometric posteriors and 3% of photogeometric posteriors may not be unimodal according to this test, although this test is not particularly accurate, so this is only a rough guide. Even when the sampled posterior shows a true, significant bimodality (or even multimodality), the 68% confidence interval sometimes spans all modes.

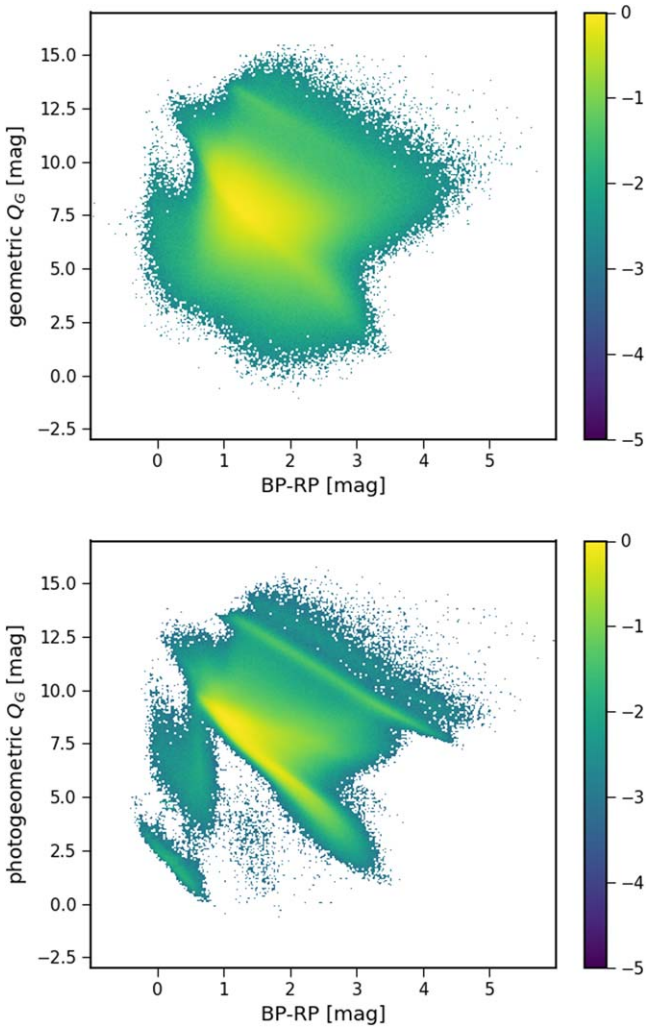

Figure 19. The EDR3 CQD over the whole sky using the geometric distances (top) and photogeometric distances (bottom). This plot uses a random sample of 0.5% of all sources in each HEALpixel. These plots include sources of all magnitude and fpu and so include sources with spurious parallaxes.

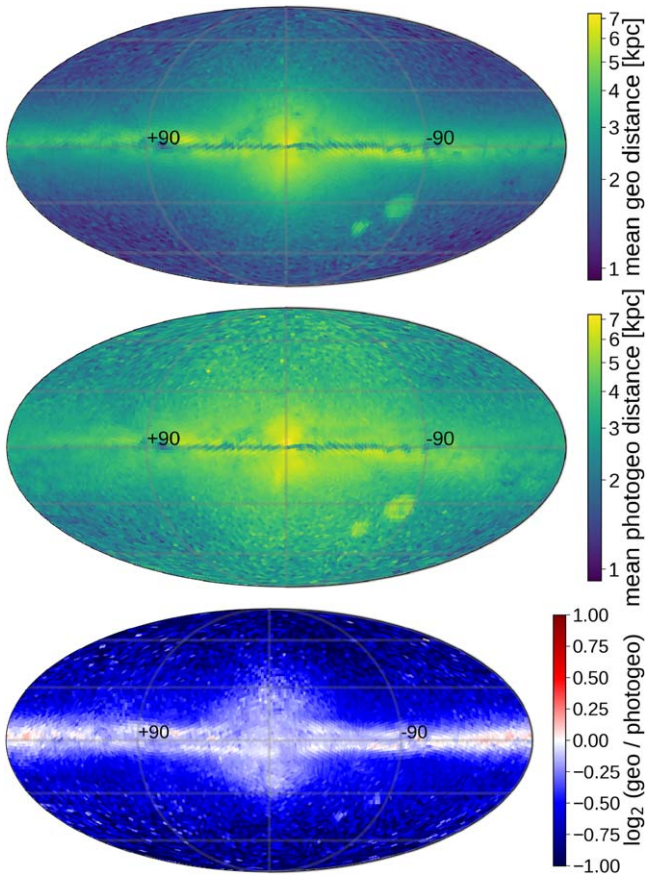

Figure 20. Mean distance of sources per HEALpixel (level 5) for our median geometric distances (top) and median photogeometric distances (middle), and the $\log_2$ ratio of these (bottom), that is, $\log_2$(geo/photogeo). This plot uses a random sample of 0.5% of all sources in each HEALpixel.

The two digits of flag `C` indicate the nature of the two Q_G models that were used to construct the Q_G prior. If both numbers are between 1 and 3, then two models bracket the source's color and were combined by linear interpolation, as explained in Section 2.5. If only one of them is zero, then only a single model was used. If both flags are zero, then there is no nonnull model within 0.1 mag color of the source, so the photogeometric posterior is not computed. There is one special value of this flag: 99 means the star lacked the necessary data to compute the photogeometric distance.

We provide additional information on the prior for each HEALpixel in the auxiliary information online, including plots like Figures 1, 3, and 4 and a table with the three parameters of the geometric prior (Equation (3)).

5.2. Filtering

We have not filtered out any results from our catalog. Parallaxes with spurious parallaxes remain, as do sources with negative parallaxes (the latter is no barrier to inferring a sensible distance; Bailer-Jones 2015). Any filtering should be done with care, as it often introduces sample biases. The flag field we provide is for information purposes; we do not recommend using it for filtering. Lower quality distances will arise from lower quality input data. These can be identified using the various quality fields in the main Gaia catalog of EDR3, which is easily cross-matched to our catalog using the `source_id` field, as shown in the example in Section 5.4. Useful quality metrics may be `ruwe`, `parallax_over_error`, and `astrometric_excess_noise`, as defined in the EDR3 documentation, where users will also find advice on their use. See in particular Section 3.2 of Fabricius et al. (2020) for suggestions for filtering spurious parallaxes.

Parallaxes from the 6p astrometric solutions (identified by `astrometric_params_solved = 95`) are not as accurate as those from the 5p solutions (Lindegren et al. 2020a) because they were normally used in more problematic situations, such as crowded fields, and are also fainter on average than the 5p solutions. Sources with 6p solutions should not be automatically removed, however. Their larger parallax uncertainties reflect their lower quality. In some applications, users may want to filter out sources with large absolute or relative distance uncertainties. One must exercise caution here, however, because uncertainty generally correlates with distance or magnitude (among other things), so filtering on these quantities will introduce sample biases.

5.3. Use Cases

For stars with positive parallaxes and $\sigma_\varpi/\varpi < 0.1$, the inverse parallax is often a reasonably good distance estimate

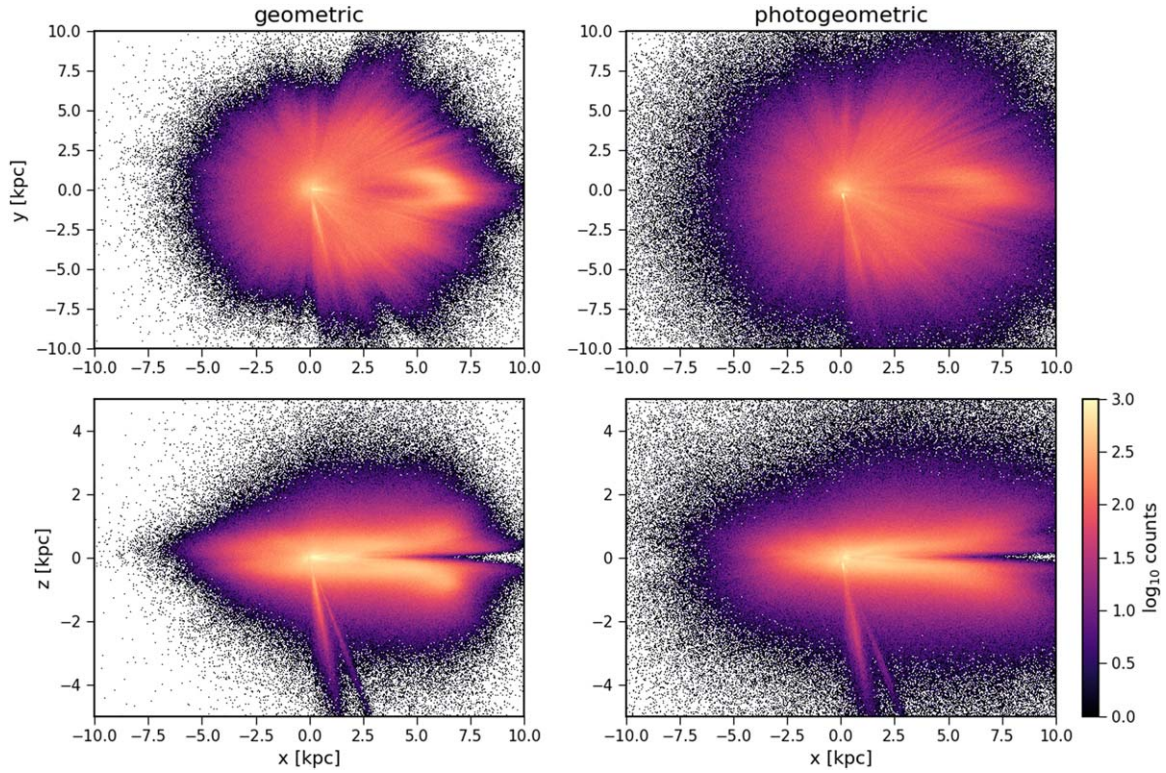

Figure 21. Projected distribution of EDR3 stars in the Galaxy using our geometric distances (left) and photogeometric distances (right). The projections are in Galactic Cartesian coordinates with the Sun at the origin. The Galactic North Pole is in the positive z direction, and the Galactic center is at around (+8, 0, 0) kpc. Galactic longitude increases counterclockwise from the positive x-axis. The top plots are the view from the Galactic North Pole. The bottom plots are a side view. This plot uses a random sample of 0.5% of all sources in each HEALpixel.

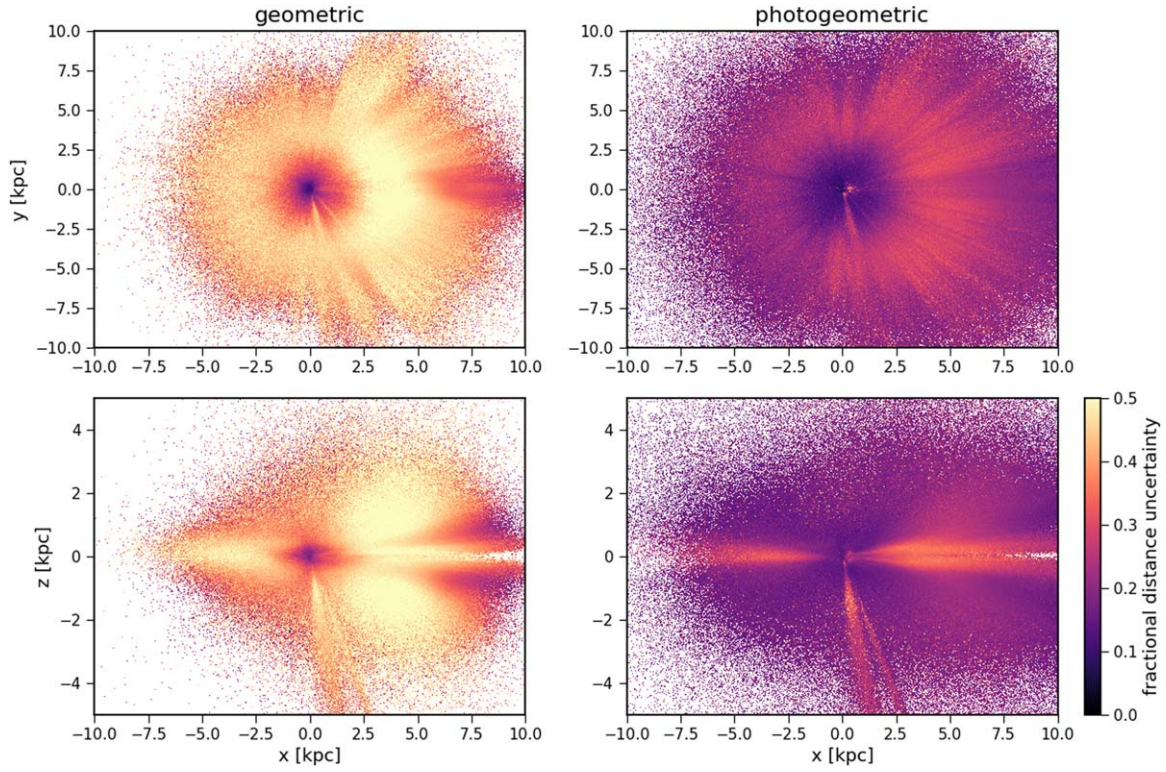

Figure 22. Same as Figure 21 but now showing the fractional symmetrized distance uncertainties, that is, $(r_{\rm hi} - r_{\rm lo})/2r_{\rm med}$.

for many purposes (when using a suitable parallax zero-point). This applies to 98 million sources in EDR3. For sources with negative parallaxes or $\sigma_\varpi/\varpi > 1$ (704 million sources), our distances will generally be prior dominated, and while the photogeometric distances could still be useful, the geometric ones are probably less so. The sweet spot where our catalog

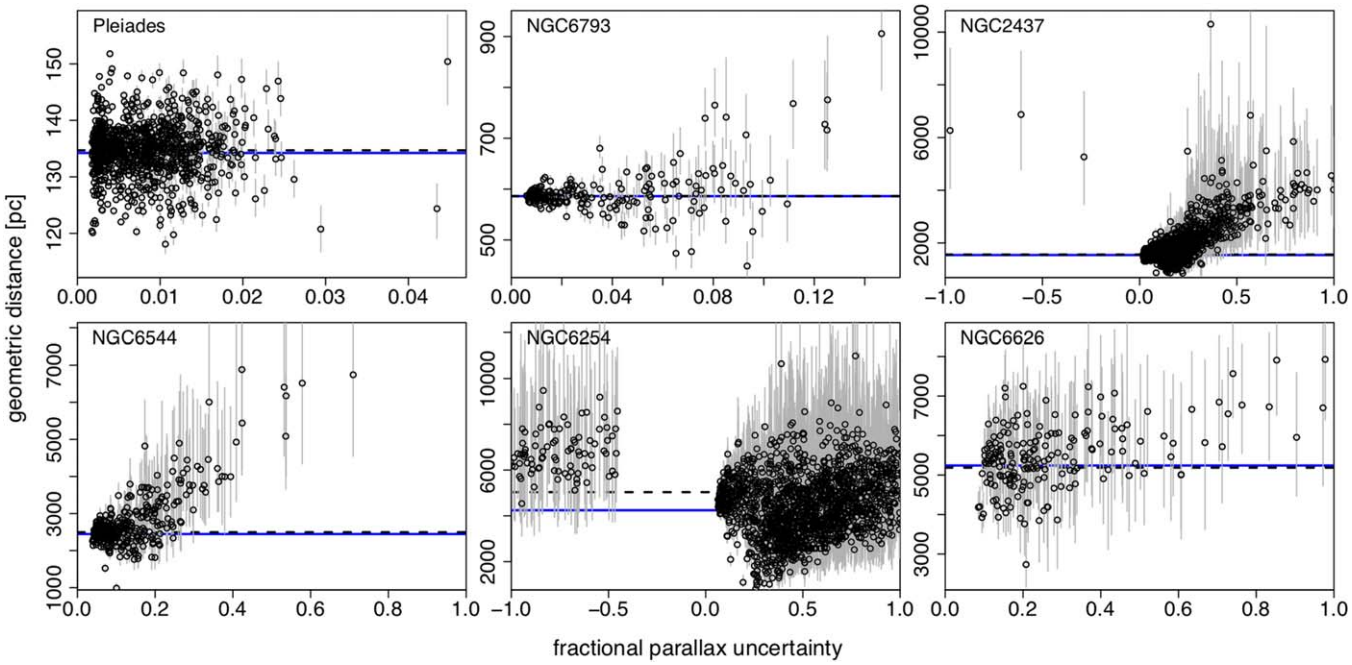

Figure 23. Validation of the geometric distance estimates using star clusters (one per panel). Each panel shows the estimated distance, $r_{\rm med}$, of the cluster members as open circles, as a function of the fractional parallax uncertainty σ_ϖ/ϖ. The error bars show the lower ($r_{\rm lo}$) and upper ($r_{\rm hi}$) bounds of the confidence intervals. The distance range spans everything in the plotted fpu range, but a few stars lie outside of the plotted fpu range for some clusters. The dashed horizontal line is the inverse of the variance-weighted mean parallax for all cluster members (including any beyond the fpu limits plotted). The solid horizontal (blue) line is the weighted-mean geometric distance for the same stars, where the weight is the inverse square of the symmetrized distance uncertainty. The clusters are ordered by increasing parallax distance.

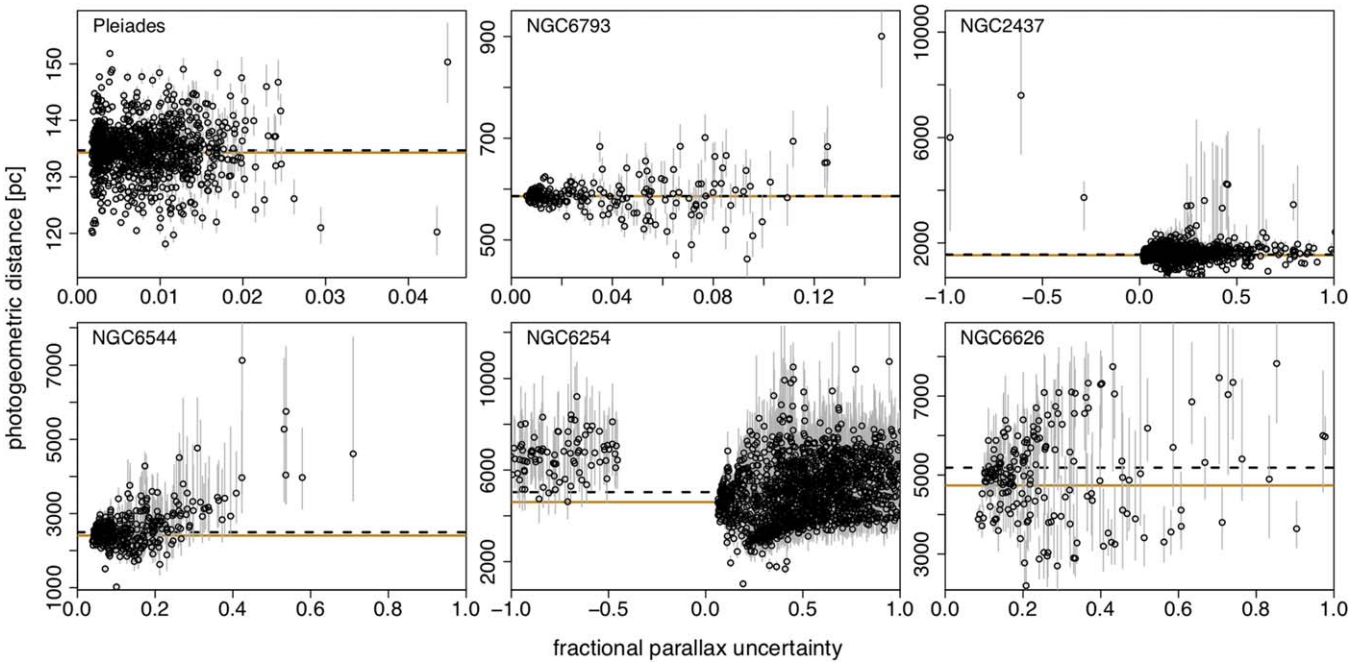

Figure 24. Same as Figure 23 but now for photogeometric distances. The solid horizontal (orange) line is the weighted-mean photogeometric distance.

adds the most value is for the remaining 665 million sources with $0.1 < \sigma_\varpi/\varpi < 1$.

The choice of whether to use our geometric or photogeometric distance depends on the specific situation and what assumptions you are willing to accept. In the limit of negligible parallax uncertainties, they will agree. At large fractional parallax uncertainties, our photogeometric distances will generally be more precise than geometric ones, because they use more information and have a stronger prior (see Figures 11 and 12). Whether they are also more accurate depends on how

Table 1
Format of the Distance Catalog Showing Results on Five Fictitious Sources

source_id	r_med_geo (pc)	r_lo_geo (pc)	r_hi_geo (pc)	r_med_photogeo (pc)	r_lo_photogeo (pc)	r_hi_photogeo (pc)	flag
4295806720	3547.973	2478.490	4741.725	2705.790	2307.170	3357.151	10033
34361129088	291.709	275.786	306.577	290.840	277.130	304.291	10033
38655544960	318.105	312.888	323.334	318.807	313.264	323.045	10033
5835726683934945280	7547.806	4509.953	11817.191	5299.187	4060.932	7178.086	10033
5835726688222520960	6316.000	3860.044	10591.593	NA	NA	NA	10099

Note. The source_id is the same as in EDR3. r_med_geo, r_lo_geo, and r_hi_geo are the median, 16th percentile, and 84th percentile of the geometric distance posterior in parsecs. r_med_photogeo, r_lo_photogeo, and r_hi_photogeo are the median, 16th percentile, and 84th percentile of the photogeometric distance posterior in parsecs. The flag is defined in Table 2. The distances are shown here rounded to three decimal places, but are provided in the catalog with 32-bit floating-point precision, which guarantees a precision of at least one part in 2^{24} (17 million). The photogeometric fields can be missing, indicated here with NA.

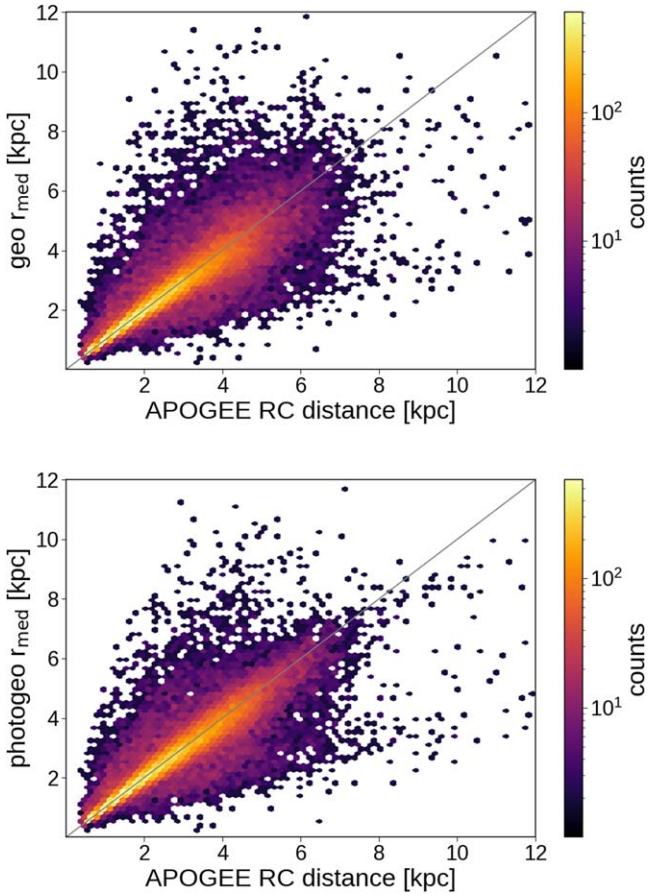

Figure 25. Comparison of APOGEE DR16 red clump star distance estimates from Bovy et al. (2014) to our geometric estimates (top panel) and to our photogeometric estimates (bottom panel) for a common sample of 36,858 sources.

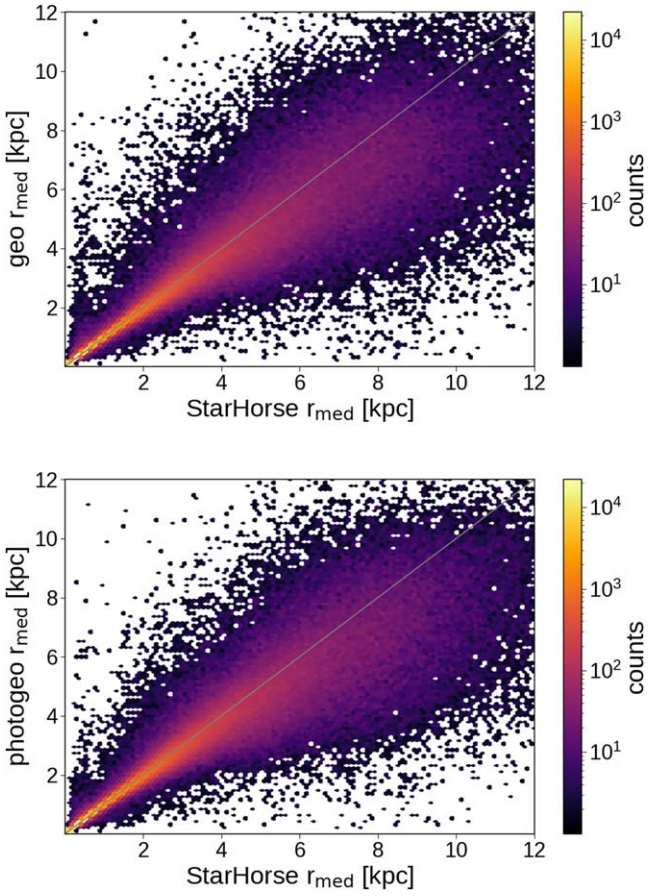

Figure 26. Comparison of StarHorse distance estimates from Queiroz et al. (2020) to our geometric estimates (top panel) and to our photogeometric estimates (bottom panel) for a common sample of 307,105 sources.

well the Q_G prior matches the true (but unknown) Q_G distribution. The Q_G model reflects the stellar population and interstellar extinction in a small patch of sky (HEALpixel of area 3.36 sq. deg.). The GeDR3mock catalog and our prior should model these reasonably well at higher Galactic latitudes, but may be less accurate at lower latitudes where extinction is higher and the stellar populations along the line of sight are more complicated. If you do not want to rely on color and magnitude information in the distance inference, use the geometric distance, as the distance prior is less sensitive to the exact stellar population in GeDR3mock.

Some example use cases are as follows:

1. Look-up of distance (or distance modulus) for particular sources of interest using their source_id or other identifier matched to this. EDR3 includes a cross-match to many existing catalogs. Positional cross-matches can also be done on the EDR3 data site or using TAP uploads, and at other sites that host our catalog.

Table 2
Flag Field in the Catalog as a String of Five Decimal Digits ABBCC

A		Source magnitude compared to the limit used to make the prior
	0	Source has no G-band magnitude
	1	$G \leqslant G_{\rm lim}$
	2	$G > G_{\rm lim}$
B		Hartigan dip test for unimodality. Left digit, geometric; right digit, photogeometric
	0	Unimodal hypothesis okay
	1	*Possible* evidence for multimodality
C		Q_G models used in prior. Left digit, bluer model; right digit, redder model
	0	Null (no model)
	1	One-component Gaussian
	2	Two-component Gaussian
	3	Smoothing spline
	Special setting:	
	99	Source lacks G and/or BP − RP

2. Identification of sources within a given distance (or distance modulus) range. The confidence intervals should be used to find all sources with a distance r satisfying $k(r_{\rm med} - r_{\rm lo}) < r < k(r_{\rm hi} - r_{\rm med})$, where the size of k will depend on the desired balance between completeness and purity of the resulting sample. A better approach would be to use the actual posterior to get a probability-weighted sample. For the geometric distances, our posterior can be reconstructed using the geometric distance prior provided for each HEALpixel in the auxiliary information online. Readers interested in using our photogeometric priors should contact the authors.
3. Construction of absolute-color–magnitude diagrams. One of the reasons that we provide quantiles for our distance estimates is that $5\log_{10}(r_{\rm med}) - 5$ is the median of the distance modulus posterior. (This would not be the case if we provided the mean or mode, for example.) Using G from EDR3, one can then compute Q_G, and from this the absolute magnitude M_G, if the extinction is zero or otherwise known. The same can be done for any photometric band from any other catalog. When computing Q_G in this way with Equation (4), the user should remember to apply the correction to the EDR3 G-band magnitude as described in Section 8.3 of Riello et al. (2020).
4. Construction of the three-dimensional spatial distribution of stars in some region of space. This may also assist in selection of candidates in targeted follow-up surveys.
5. As a baseline for comparison of distance or absolute magnitude estimates obtained by other means.
6. Our distances could be used for another layer of inference, such as computing transverse velocities using also the EDR3 proper motions, although users will need to consider the appropriate error propagation. In particular, if the error budget is not dominated by a single source (e.g., not just the distance), users are advised to infer their desired quantities directly from the original parallaxes, perhaps using the priors provided here.

Users should realize that uncertainties in the parallaxes in EDR3 are correlated between different sources to a greater or lesser degree depending on their angular separations (Fabricius et al. 2020; Lindegren et al. 2020a). Caution must therefore be exercised when combining either the parallaxes or our distances, such as averaging them to determine the distance to a star cluster. In such a case, the simple "standard error in the mean" may underestimate the true uncertainty, and the same prior would be used multiple times. One should instead set up a joint likelihood for the sources that accommodates the between-source correlations and solve for the cluster distance directly.

5.4. Access

Our distance catalog is available from the German Astrophysical Virtual Observatory at http://dc.g-vo.org/tableinfo/gedr3dist.main, where it can be queried via TAP and ADQL. This server also hosts a reduced version of the main Gaia EDR3 catalog (and GeDR3mock). Typical queries are likely to involve a combination of the two catalogs. By way of example, the following query returns coordinates, our distances, BP − RP, and the two Q_G values using the median distances, for all stars with a low `ruwe` in a one-degree cone in the center of the Pleiades. This should run in about one second and return 22,959 sources:

```
SELECT
  source_id, ra, dec,
  r_med_geo, r_lo_geo, r_hi_geo,
  r_med_photogeo, r_lo_photogeo, r_hi_photogeo,
  phot_bp_mean_mag − phot_rp_mean_mag AS bp_rp,
  phot_g_mean_mag − 5 ∗ LOG10(r_med_geo) + 5AS qg_geo,
  phot_g_mean_mag − 5 ∗ LOG10(r_med_photogeo) + 5
    AS gq_photogeo
FROM gedr3dist.main
  JOIN gaia.edr3lite USING(source_id)
WHERE ruwe < 1.4
  AND DISTANCE(ra, dec, 56.75, 24.12) < 1
```

A bulk download for the catalog is also available at the URL given above. Our catalog will also become available soon, together with the full EDR3 catalog hosted at https://gea.esac.esa.int/archive/ and its partner data centers. At these sites, the table names `gedr3dist.main` and `gaia.edr3lite` may well be different.

5.5. Limitations

When using our catalog, users should be aware of its assumptions and limitations:

1. We summarize the posteriors using only three numbers (quantiles), which cannot capture the full complexity of these distributions. This is more of a limitation for the photogeometric posteriors. The confidence intervals should not be ignored.
2. Most sources in EDR3 have large fractional parallax uncertainties, and our distances correspondingly have large fractional uncertainties, especially for the geometric distances.
3. The poorer the data, the more our prior dominates the distance estimates. Our prior is built using a sophisticated model of the Galaxy that includes 3D extinction, but it will not be perfect. If the true stellar population, extinction, or reddening law is very different in reality, our distances will be affected. In Section 3.2.1, we explained, using results on simulated data, what biases can occur and why.
4. Sources with very large parallax uncertainties will have a posterior dominated by the prior. The median of this varies between 745 and 7185 pc depending on HEALpixel (Figure 2). Stars with large fpu values that truly lie well beyond the prior's median will have their geometric distances underestimated; stars with large fpu values that lie closer than the prior's median will have their geometric distances overestimated. As distant stars generally have larger fpu values than nearby stars, and distant stars are more numerous, the former characteristic will dominate among poor-quality data. This leads to a bias in distance estimates, one that is probably unavoidable (see Appendix A). Poor data remain poor data.
5. Our prior is spatially discretized at HEALpixel level 5, that is, in patches of 3.36 sq. deg. on the sky. The distance prior and CQD change discontinuously between HEALpixels, and this may be visible in sky maps of posterior distances. The Q_G priors (constructed from the CQD) are formed by a linear interpolation over color whenever possible, so in these cases there should be no discontinuity of distance with color within a HEALpixel.
6. Our inferred distances retain all of the issues affecting the parallaxes, some of which have been explored in the EDR3 release papers (Fabricius et al. 2020; Lindegren et al. 2020a). We applied the parallax zero-point correction derived by Lindegren et al. (2020b), which is better than no correction or a single global correction but is not perfect. Any error in this will propagate into our distance estimates. The published parallax uncertainties are also probably underestimated to some degree (Fabricius et al. 2020). Gaia Collaboration (2020a) and Riello et al. (2020) report some issues with the EDR3 photometry, such as biased BP photometry and therefore BP − RP colors for very faint sources, which could affect our photogeometric distances. These distance estimates additionally suffer from any mismatch between the published EDR3 photometry and the modeling of this—in particular the passbands—used in the GeDR3mock catalog, which forms the basis for our Q_G priors.[8] Note that we applied the G-band magnitude correction to the EDR3 photometry as described by Riello et al. (2020).
7. We implicitly assume that all sources are single stars in the Galaxy. Our distances will be incorrect for extragalactic sources. The geometric distances will be wrong for unresolved binaries if the parallax for the composite source is affected by the orbital motion. Even when this is not the case, the photogeometric distance may still be wrong, because the G-band magnitude will be brighter than the Q_G prior expects (binaries were not included in the prior).
8. By design, we infer distances for each source independently. If a set of stars is known to be in a cluster, and thus have a similar distance, this could be exploited to infer the distances to the individual stars more accurately than we have done here. In its most general form, this involves a joint inference over multiple sources. Various methods exist in the literature for doing this, such as Palmer et al. (2014), Cantat-Gaudin et al. (2018), and Olivares et al. (2020). Likewise, in order to estimate the distance to the cluster as a whole, one should be aware that averaging our individual distances will compound the prior. If the fpu of the individual sources is large, this product of priors would dominate the distance estimate more than desired. A joint inference can easily be set up to overcome this.

6. Summary

We have produced a catalog of geometric distances for 1.47 billion stars and photogeometric distances for 92% of these. These estimates, and their uncertainties, can also be used as estimates of the distance modulus. Geometric distances use only the EDR3 parallaxes. Photogeometric distances additionally use the G magnitude and BP − RP color from EDR3. Both types of estimate involve direction-dependent priors constructed from a sophisticated model of the 3D distribution, colors, and magnitudes of stars in the Galaxy as seen by Gaia, accommodating both interstellar extinction and a Gaia selection function. Tests on mock data, but moreover validation against independent estimates and open clusters, suggest our estimates are reliable out to several kiloparsecs. For faint or more distant stars, the prior will often dominate the estimates. We have identified various use cases and limitations of our catalog.

Our goal has been one of inclusion: to provide distances to as many stars in the EDR3 catalog as possible. This has required us to make broad, general assumptions. If one focuses on a restricted set of stars with some approximately known

[8] We compared simulations of the G-band magnitude and the BP–RP color between the GeDR3mock passbands and those published for EDR3, using isochrones at 4 Myr and 1 Gyr. The differences in the G magnitudes are below 6 mmag, except for sources bluer than −0.15, where it can be as high as 700 mmag. For BP − RP using the BP bright band (in GeDR3mock), the difference is around 10 mmag, but up to 25 mmag for sources with BP − RP > 1.2 mag and up to 100 mmag for sources with BP − RP < −0.2 mag. For BP faint, the BP − RP difference is around 20 mmag, but up to 60 mmag for sources with BP − RP > 0.5 mag and up to 100 mmag for sources with BP − RP < −0.15 mag.

properties, it will be possible to construct more specific priors and to use these to infer more precise and more accurate distances. Better distances may also be achievable by using additional data, such as spectroscopy or additional photometry.

We thank the IT departments at the Max Planck Institute for Astronomy and the Astronomisches Rechen-Institut for computing support. This work was funded in part by the DLR (German space agency) via grant 50 QG 1403. It has made use of data from the ESA mission Gaia (http://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC; http://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. This research made use of TOPCAT, an interactive graphical viewer and editor for tabular data (Taylor 2005); Vaex, a tool to visualize and explore big tabular data (Breddels & Veljanoski 2018); matplotlib, a Python graphics library (Hunter 2007); `HEALpix` (Górski et al. 2005) and `healpy` (Zonca et al. 2019); the NASA Astrophysics Data System; and the VizieR catalog access tool, CDS, Strasbourg.

Facility: Gaia.

Appendix A
Thoughts on a Better Distance Prior

The strong dependence of the geometric posterior on the distance prior in the limit of large parallax uncertainties is an unavoidable consequence of inference with noisy data. We saw something similar in Paper IV. This leads to a distance bias mostly for distant stars with large fpu. Could this be avoided? Conceptually one would like a distance prior that depends on the true fpu, but this is impossible because the true parallax is not known. One may be tempted to use the measured fpu instead, but this is not what we want: a star with a large true fpu could have a small measured fpu due to noise, and thereby be treated incorrectly. Its use is also theoretically dubious because it places the parallax—a measurable—in the prior and in the likelihood. We experimented with using a prior conditioned on σ_ϖ, but we found that this did not help (see the technical note GAIA-C8-TN-MPIA-CBJ-089 with the auxiliary information online). One may achieve something close to what is desired by simply shifting the distance prior to greater distances so that it better represents stars with a larger true fpu, which is where the prior is needed more. Yet this would detrimentally affect the distance estimates for nearby stars. It seems a poor trade-off to sacrifice accuracy on high-quality data for a better prior on low-quality data. Conditioning the prior on the star's magnitude may help, and this is what our photogeometric distances do (Section 2.4).

Appendix B
The Limit of Poor Parallaxes

We tend to think that a large fpu means that the likelihood is uninformative and that the posterior converges toward the prior. Consider a red clump star in the LMC with a true parallax of 0.02 mas and a typical parallax uncertainty of 0.2 mas for a star with $G = 19$ mag. The true fpu is $0.2/0.02 = 10$. Let us assume initially that we actually measure a parallax of 0.02 mas; that is, we have a measured fpu of 10. (Of course in this lucky case the inverse parallax would be the correct distance, but it is very rare in practice.) In the LMC HEALpixel 8275, our distance prior has a median of 1.2 kpc because we exclude the LMC from our prior, so we might expect to see many sources with this inferred distance. In fact, we see many sources with larger inferred distances (see the plot with the auxiliary online information). The reason is that the likelihood of a measurement of 1 mas (corresponding to a distance of 1 kpc) is still at 4.9σ and therefore quite unlikely. This shows that even when the fpu is large the parallax can be quite informative.

One should remember, however, that our inference never sees the true parallax but only the measured parallax, which is normally distributed around the unknown true value (with a standard deviation that is also only estimated). So it is quite likely that our measurement of the above red clump star gives us a parallax measurement of, say, 0.4 mas. In that case, the measured fpu is 0.5, and the likelihood of 1 mas, that is, a 1 kpc distance, is only 3σ away from this measurement. Taking the parallax *measurement* into account essentially redistributes probability mass into the wings of the likelihood and therefore to higher and lower (also negative) parallax values. Given the truncation of negative parallaxes when calculating the posterior, this implies that the median distance estimate is lower for the true measurements, compared to the idealized inference using the true parallax. Similarly, one should be careful not to interpret plots involving the measured fpu as though it were the true fpu.

ORCID iDs

M. Fouesneau https://orcid.org/0000-0001-9256-5516

References

Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, A&A, 628, A94
Arenou, F., Luri, X., Babusiaux, C., et al. 2018, A&A, 616, A17
Astraatmadja, T. L., & Bailer-Jones, C. A. L. 2016a, ApJ, 832, 137
Astraatmadja, T. L., & Bailer-Jones, C. A. L. 2016b, ApJ, 833, 119
Bailer-Jones, C. A. L. 2015, PASP, 127, 994, (Paper I)
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58, (Paper IV)
Bovy, J., Nidever, D. L., Rix, H.-W., et al. 2014, ApJ, 790, 127
Breddels, M. A., & Veljanoski, J. 2018, A&A, 618, A13
Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
Fabricius, C., Luri, X., Arenou, F., et al. 2020, A&A, in press
Gaia Collaboration, Brown, A. G. A., & Vallenari, A. 2016b, A&A, 595, A2
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2020a, A&A, in press
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016a, A&A, 595, A1
Gaia Collaboration, Smart, R. L., Sarro, L. M., et al. 2020b, A&A, in press
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hall, O. J., Davies, G. R., Elsworth, Y. P., et al. 2019, MNRAS, 486, 3569
Hartigan, J. A., & Hartigan, P. M. 1985, AnSta, 13, 70
Hunter, J. D. 2007, CSE, 9, 90
Leung, H. W., & Bovy, J. 2019, MNRAS, 489, 2079
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2020a, A&A, in press
Lindegren, L., Bastian, U., Biermann, M., et al. 2020b, A&A, in press
Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
McMillan, P. J. 2018, RNAAS, 2, 51
Olivares, J., Sarro, L. M., Bouy, H., et al. 2020, A&A, 644, A7
Palmer, M., Arenou, F., Luri, X., & Masana, E. 2014, A&A, 564, A49
Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2020, A&A, 638, A76
Riello, M., De Angeli, D., Evans, D. W., et al. 2020, A&A, in press
Rybizki, J., Demleitner, M., Bailer-Jones, C. A. L., et al. 2020, PASP, 132, 074501

Rybizki, J., Demleitner, M., Fouesneau, M., et al. 2018, PASP, 130, 074101
Rybizki, J., & Drimmel, R. 2018, gdr2_completeness: GaiaDR2 Data Retrieval and Manipulation, Astrophysics Source Code Library, ascl:1811.018
Sanders, J. L., & Das, P. 2018, MNRAS, 481, 4093
Schönrich, R., & Aumer, M. 2017, MNRAS, 472, 3979
Taylor, M. B. 2005, in ASP Conf. Ser. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29
Zonca, A., Singer, L., Lenz, D., et al. 2019, JOSS, 4, 1298
Zucker, C., Schlafly, E. F., Speagle, J. S., et al. 2018, ApJ, 869, 83