



Vera C. Rubin Observatory
Systems Engineering

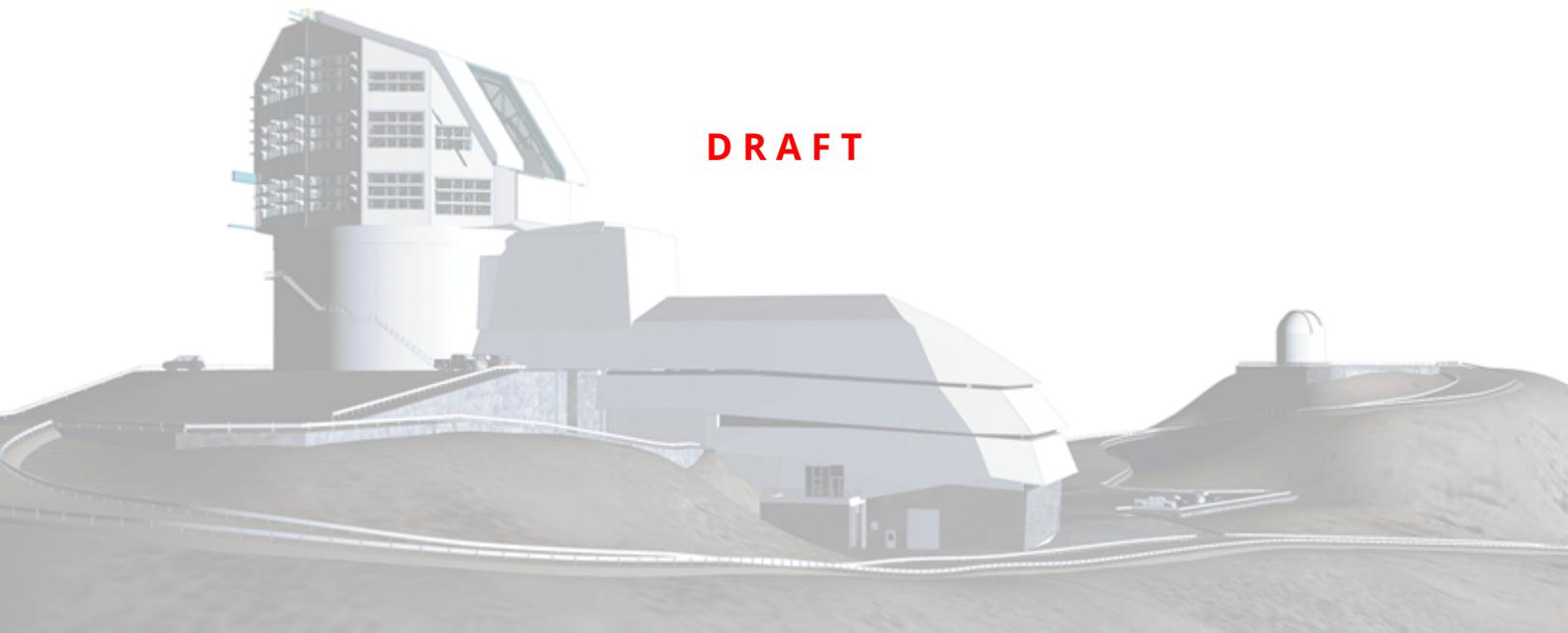
Determining the Accuracy and Precision of Astrometric Positions and Covariances from the LSST Science Pipelines Using Rubin's Operations Rehearsal 3 Data

Tom J. Wilson, and Tim Naylor

SITCOMTN-159

Latest Revision: 2025-07-04

DRAFT



Abstract

We carry out an initial investigation into the LSST Pipelines astrometric accuracy and precision using the Operations Rehearsal 3 dataset. For each of the available simulated tracts we create distributions of nearest-neighbour match on-sky separations to the simulation's truth table for narrow windows of magnitude, pipeline-derived astrometric precision, and (where appropriate) on-sky source density. For a series of such cross-matches – across a range of magnitudes and hence signal-to-noise ratios – for the source table, single-visit detections, we derive best-fit parameters such that we can derive astrometric uncertainties that match the separation distributions. Fitting for both a systematic uncertainty n , to be added in quadrature to the pipeline uncertainty, and a multiplicative scaling factor for the quoted precisions m , we find, globally, that scaling factors are unnecessary ($m \approx 1 \pm 0.1$), but that a systematic astrometric uncertainty of n in the range 0.003-0.007 arcsec (3-7 mas) should be included, with a weak decreasing relationship with increasing field density. This suggests that at fainter magnitudes (lower signal-to-noise ratios) the pipeline is correctly modelling all contributions to astrometric uncertainty, and that the deviations from “true” position are accurately reflected in the corresponding confidence in the measured position. For the very brightest objects a small – but significant relative to the statistical uncertainty – astrometric precision is not being recovered, from sources such as the global World Coordinate System solution, missing sources of noise not being propagated through the full pipeline, and so on. On the other hand, for object table, coadded image detections, this simple scaling relation does not hold. Instead we find a power-law fit between pipeline-derived and best-fit astrometric precisions, with a power-law slope of approximately 0.75, the origin of which we are unable to explain easily. Combining all simulated pointings we find a very tentative sub-arcsecond (0.6 mas) systematic plateauing astrometric precision for the Operations Rehearsal Object tables. We however caution that these results are relatively limited, due to the nature of the simulated fields chosen as part of the Operations Rehearsal, and that further investigations on “real” commissioning data – including fields with a higher density of (stellar) objects – will be necessary to determine the universality of these conclusions.

Change Record

Version	Date	Description	Owner name
1	YYYY-MM-DD	Unreleased.	First Last

Document source location: <https://github.com/lsst-sitcom/sitcomtn-159>

Draft

Contents

1 Introduction	1
2 Ensemble Astrometric Precisions	2
3 The LSST Science Pipelines and Operations Rehearsal 3	3
4 Assessing the Astrometric Precision of OR3	4
4.1 Pipeline Performance For The Source Table	7
4.1.1 Calculating Astrometric Precision Correction Factors	7
4.1.2 Source Table Astrometry	11
4.2 Pipeline Performance For The Object Table	11
4.2.1 Object Table Quadratic-Fit Precision Breakdown	14
4.2.2 Object Table Individual-Fit Precision Breakdown	14
4.2.3 Merged Object Table Analysis	15
5 Discussion	16
5.1 Magnitude-SNR Relationship	17
5.2 Magnitude-Precision Relationship	18
5.2.1 Simulated Seeing and The Effect of Seeing On Astrometric Uncertainty	20
5.3 Position Scatter vs. Centroid Uncertainty	20
5.4 Comparison with Ivezić et al. (2019) Predictions	22
6 Conclusion	22
A References	23
B Acronyms	24

Determining the Accuracy and Precision of Astrometric Positions and Covariances from the LSST Science Pipelines Using Rubin’s Operations Rehearsal 3 Data

1 Introduction

One of the most fundamental measurements that can be made of astrophysical objects in an image taken by a telescope is the position of the source. Initially measured in image coordinates (commonly referred to as the x - y plane), some transformation is made to convert these coordinates to celestial coordinates (most often in Right Ascension and Declination, RA and Dec respectively). However, we are also able to determine the precision with which we made that measurement, quantifying the belief we have in our measurement being the “true” location of the object in the face of e.g. noise in our image. The main task of any astronomical image pipeline is to extract, along with the brightnesses (and corresponding uncertainties) of the objects, these positions and position-measuring precisions of all detected objects, and collate them into a table (or catalogue, database, etc.) for use by astronomers.

For a large range of analyses in astronomy, but especially for the task of performing cross-matches from these catalogues of objects to other datasets, containing the positions and brightnesses of objects detected by other telescopes or surveys, it is important that the precisions of objects accurately reflect the “proper” certainty we have in the object location, and reflect the confidence we have in the true location of the object based on where we recorded it to be. Certain algorithms used in the determination of counterpart assignment (e.g. Budavári & Szalay, 2008, Wilson & Naylor, 2018a, Pineau et al., 2017) use the precision of the measurement of the object(s) as well as the location(s) of the object to compute relative likelihoods of two detections, one from each catalogue, being the same object detected twice and separate detections of two real astrophysical objects. It is therefore crucial that the astrometric measurements and their respective precisions are in alignment, and that the generated catalogues contain both accurate astrometry and reliable astrometric uncertainties.

In this work we perform a preliminary investigation into the accuracy of the astrometry as produced by the LSST Science Pipelines, working with precursor data for the Vera C. Rubin Observatory’s Legacy Survey of Space and Time (LSST; Ivezić et al., 2019).

2 Ensemble Astrometric Precisions

In isolation, it is quite a difficult task to verify an object’s astrometric precision is correct. We would ideally compare the separation between the object and the “ground truth” position, but we only have one measurement per object, which doesn’t give us much insight. In any case for catalogues of observations of sky sources we don’t know where the objects’ true locations are. However, we can at least solve the first problem by using the ensemble nature of similar groups of objects.

In general, the precisions we wish to verify from our data reduction process are the “centroid” precisions – the measure of the ability we have to pinpoint exactly where on our observations the recorded objects are, independent of other factors that may influence those recorded positions (Wilson & Naylor, 2017, 2018b). Hence, if we are able to find a large number of objects that all have the “same” parameters, we can bypass this single-realisation problem. Once we have done so, we can use these multiple measurements to build an alternative method of determining the precision of these observations.

As an example, let us imagine we had a large number of objects, all with the same precision of position measurement, and we had a measurement of each of those object’s positions, subject to some noise (such as photon or shot noise in our exposures). If we were able to measure the deviation of each object’s measured position from its true position, and produce an array of deviations from true, then we expect that the distribution of these values follow a specific functional form. In this case, we should see that the number of objects in each small, unit-area region – say, some Δx by Δy rectangle in a 2-D plane around the origin, with deviations in e.g. RA and Dec – should follow a Gaussian distribution. In particular, the important parameter that controls the width of the distribution is its standard deviation, σ , which can be calculated by, in the case of a 1-D variable, $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$ with $\mu = \frac{1}{N} \sum_{i=1}^N x_i$. This can equivalently be seen in terms of the number of objects per unit-separation – thinking in terms of total distance from origin Δr , some sum-in-quadrature of x and y or RA and Dec – which, identically, follows a Rayleigh distribution and has exactly the same standard deviation σ calculation. This is simply an integration of the Gaussian function around 360 degrees, and hence has a scaling relation of $2\pi r dr$. The main visual difference between the two is that, due to this radial-dependence, the Rayleigh distribution vanishes at the origin, as there is simply no area for any objects to be placed in such a small annulus!

This then forms the crux of our validation: if the standard deviation of the pipeline-derived vs

true position differences agrees with the pipeline-derived precisions then we can have confidence in both the positions and positional uncertainties.

3 The LSST Science Pipelines and Operations Rehearsal 3

In practice, validating the astrometry of a photometric survey or a particular data release is hard, due to the aforementioned lack of ground-truth observation. For such cases we might turn to the incredible precision and accuracy of surveys such as *Gaia* (Gaia Collaboration et al., 2021) for what amounts, to all intents and purposes, to the “true” positions of objects, relative to our own catalogue’s precision. For the moment, however, we can do slightly better than that; Operations Rehearsal 3 (OR3) was a simulation of several nights of the Rubin Observatory’s operations, as it is anticipated it will operate during the early part of the LSST. Due to this, we, for once, *do* have ground truth observations against which to compare to! This allows us to test solely the data-reduction pipeline, removing from our analysis any potential issues with the physical telescope or with our comparison dataset.

We therefore extract all available tracts that were simulated in OR3 through the `Butler` using `lsst-scipipe-9.0.0` and the `intermittentcumulativeDRP/20240402_03_04/d_2024_03_29/DM-43865` collection. After filtering for tracts with sufficient numbers of sources we are left with 19: tracts 2661, 2662, 3200, 3346, 3384, 3385, 3534, 6914, 6915, 7149, 7683, 7684, 9348, 9590, 9591, 9638, 9812, 9880, and 9881. However, we will subsequently filter these down again and will, in the end, be left with six of these tracts (3346, 3384, 6914, 7149, 9590, and 9880) as our main targets of investigation. One simulation-specific filter to note is the dither pattern. Each simulated pointing is a densely dithered field, and hence there are a large, varying number of unique visits at each coordinate for each sky pointing in the dataset created from co-added images that forms the `Object` table. To avoid issues with our analysis where we could have introduced a disconnect between the magnitude of an observation and its signal-to-noise ratio (SNR), we limit analysis to sources in the `Object` tables that are in $N \pm 10$ visits, with N chosen to be the largest possible number of visits ($N \gtrsim 40$) while maximising the number of data points.

Furthermore, we will also combine tracts into “pointings,” each separate observing strategy, to investigate the single-visit image pipeline as well as the tract-level coadd pipeline performance. This results in eight combined-tract analyses of `Source` table reductions, all of which, by the larger number of objects available through no longer combining repeated observa-

tions, meet the post-filter cuts needed for analysis. In these cases we refer to sightlines by either coordinate or appropriate description assigned by the Project to the sky regions.

For each tract we load the pipeline-created catalogue and filter for NaN values in flux, flux uncertainty, and the three astrometric uncertainty values (RA, Dec, and covariance between the two). We then extract the position coordinates, astrometric uncertainty (circularised by computing the semi-major and semi-minor axes and taking the geometric mean, to reduce our problem to a one-dimensional one). We calculate photometric magnitudes and uncertainties by converting flux and flux uncertainty from nano-Janskies to AB magnitudes, propagating uncertainty through $m = -2.5 \log_{10}(f)$. The truth tables are loaded, for all `healpix` pixels the tract covers, for the `pointsource`, `sso`, and `galaxy` parquet files, which are then all combined into one table, from which the RA and Dec coordinates are extracted.

In the case of the `Source` table, we are not able to load a single catalogue in one go, since we now have the repeated observations at each pointing. This case will also aid in the avoidance of the “single realisation” issue, since we will observe the same astrophysical object repeatedly with each visit, providing increased number statistics for probing the ability of the pipeline to determine centroids and centroid covariances. For investigations in these cases, we load, at present, the first $N \approx 15$ single-visit reductions in each tract for a given pointing, filtering for repeated visit IDs; this provides us with at minimum 15 visits but upwards of roughly 60 visits – there being up to four tracts per pointing with no repeated visit IDs. These visits are merged into one large photometric catalogue, although the unique visit ID is kept to ensure that counterpart assignments can be performed on a per-visit basis during analysis, with truth-measured position separations then subsequently combined.

4 Assessing the Astrometric Precision of OR3

The first step in assessing the astrometric positions and precisions is to calculate the separations between our two datasets. Thus, the pipeline table is cross-matched to the truth table. For this, to avoid a cyclical process where we may bias results by using a cross-match algorithm to verify the astrometry that then gets input into our algorithm, we use a nearest-neighbour algorithm. Here each object has its on-sky separation computed to every source in the opposing catalogue – per unique visit ID for the `Source` table case – and has selected as its counterpart the source that is the smallest distance away. In crowded fields this is very sensitive to false-positive interlopers, so to reduce the effect this may have we calculate

nearest-neighbour matches for both datasets, and only select pairs which agree that each other are their closest neighbour. This removes from our sample any pairings for which there is disagreement over pairing solutions, but we are more concerned with purity than completeness in generating these samples. We use a five-arcsecond maximum cut-off radius, but in practice this is much larger than any separations we report, and as such we are insensitive to this particular parameter.

Next, we generate a series of sub-samples of our dataset, at a series of magnitudes. For this we generate samples at intervals of 0.25 magnitudes between 16th and 29th magnitude in the i -band – a lot of these, at the bright and faint ends, won't make our final sample cuts, due to low number counts, but we try them anyway, as higher-density parts of the sky may meet these criteria where other parts do not. Here we are largely using magnitude as a proxy for signal-to-noise ratio, to probe different regimes within the photometric images. For each magnitude (SNR) in turn, we select objects within 0.05 magnitude (i.e., $m \pm 0.05$), as well as selecting objects in a small range (0.02 arcseconds) of uncertainty around the mode of the pipeline astrometric precision distribution, determined on a case-by-case basis.

For these objects – all of which have the same magnitude and hence SNR, and same apparent astrometric precision – we calculate the histogram of cross-match separations to our truth table. To these we first compare the quoted astrometric precision, as well as including any systematic “deviations from true” that may be present in the data – chief of which for LSST will be an effect caused by unresolved contaminants, hidden with a brighter detection, causing perturbations to the maximum-likelihood measurement of its position. However, the OR3 data do not show any strong systematic terms, due to the areas of the sky simulated, and hence we do not discuss them any further here; the only important term in the separation between a measured position of an object and its true, simulated position is its pipeline-dependent centroiding. In addition, to the distribution of “correct” counterparts we further add a parameterisation of the false-positive matches that may not have been removed by our simple cross-match algorithm choice. These object separations generally follow a much broader distribution, with lengthscales a few times the typical separation between true cross-matches, and are, again, a nuisance parameter that we must account for to improve the robustness of the analysis but are otherwise not discussed further in this work.

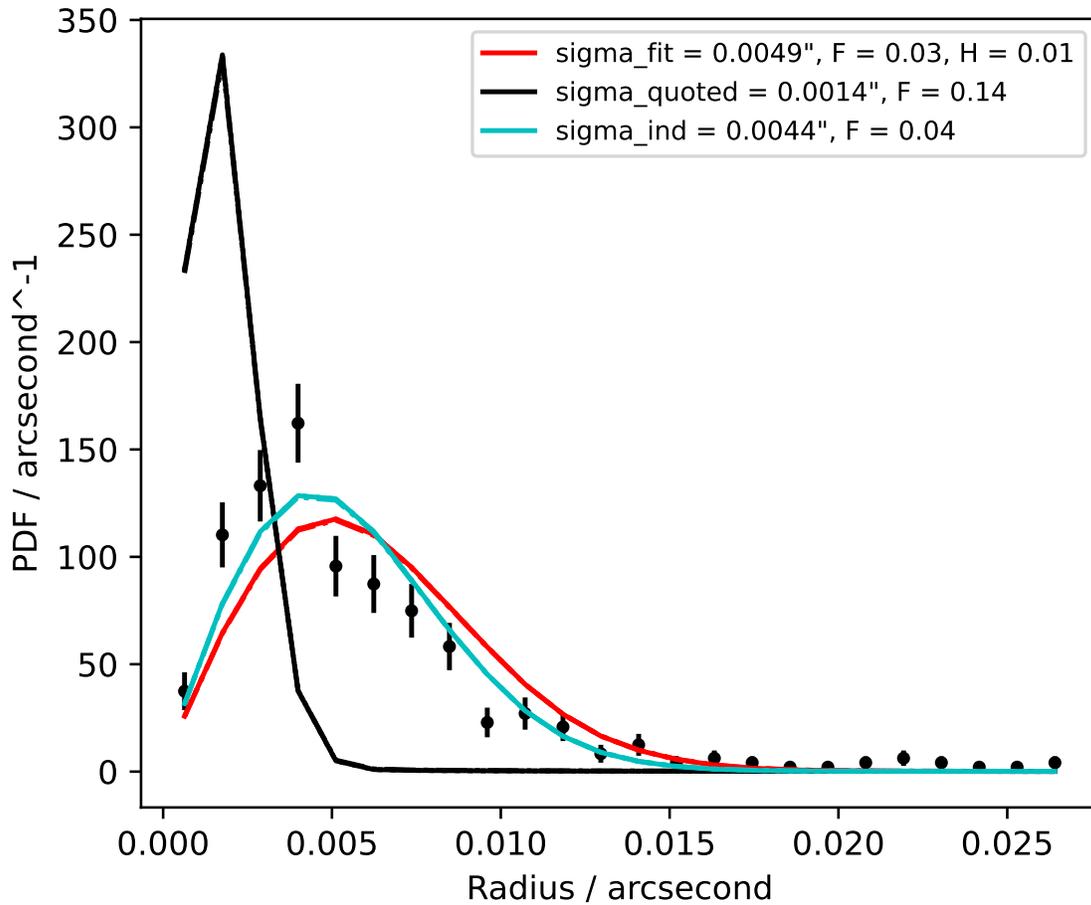


FIGURE 1: Comparison between astrometric precision model and distribution of measured-truth position separations. Black errorbars show the radial separation between pipeline-generated data and simulated positions for $i = 18.5$ Sources across the “Rubin SV” OR3 pointing at RA = 250 deg, Dec = 2 deg. Black lines show expected distribution of separations based on quoted astrometric precisions, while red and cyan lines show different best-fitting precisions. In this case the quoted astrometric uncertainties, as derived by the pipeline, are too small (black vs red/cyan solid lines).

4.1 Pipeline Performance For The Source Table

An example of the evaluation of astrometric uncertainty and position measurement, for the “Rubin SV 250 2” pointing, at $i = 18.5$ (ap03 fluxes, Source table), is given in Figure 1. Here the experiment becomes clear: we simply compare the ensemble of data-truth separations (black errorbars) with models for expected distributions of separations based on the pipeline-derived uncertainties (black line) and best-fitting uncertainties (cyan line). We can see in this case that the pipeline-derived precisions (expressed as a Gaussian σ) are too small (0.0014 arcsec, 1.4 mas) compared with the best-fit values (4.4 mas).

At fainter magnitudes, we see much better agreement. Figure 2 shows the case for $i = 22.75$, where we see excellent agreement between the quoted precision and the best-fit parameterisation of the position-position residual distribution. This trend continues for the “250, 2” pointing to its completeness limit of just below 24th magnitude, in agreement with the anticipated $5\text{-}\sigma$ depth (Ivezić et al., 2019).

Figure 3 summarises these results, plotting the individually derived astrometric precision that best fits the cross-match residuals against the pipeline-derived astrometric precision for the detections. Here, more obvious in the logarithmic-scale panel, we can see good agreement, and robust quoted astrometric precisions, across all pointings at faint magnitudes (larger uncertainties, top-right of the panel). However, we see systematic floors to the precisions for bright objects (towards the left of the panel).

4.1.1 Calculating Astrometric Precision Correction Factors

Although we quote the so-called “best-fit” astrometric precision for a single magnitude slice in Figures 1 and 2, we do not necessarily want to derive these on a per-tract-and-magnitude combination basis. We could instead generate the probability density functions (PDFs) for all magnitude (SNR) slices for a given tract or pointing, and then fit for the relationship between quoted precisions – with each magnitude slice having a single, shared precision – and the best-fit precisions. We choose to model this as $\sigma_{\text{fit}} = \sqrt{(m\sigma_{\text{quoted}})^2 + n^2}$, where n is a systematic astrometric precision that may be missing from the quoted precisions, which has its largest impact at bright magnitudes, and m is a multiplicative factor to correct for over- or under-estimation of the quoted precision, dominating the faint-end of the dynamic range probed. By inspection of Figure 3 this looks to be a good parameterisation of the relationship between best-fit and quoted astrometric precisions as a function of magnitude. However, you could

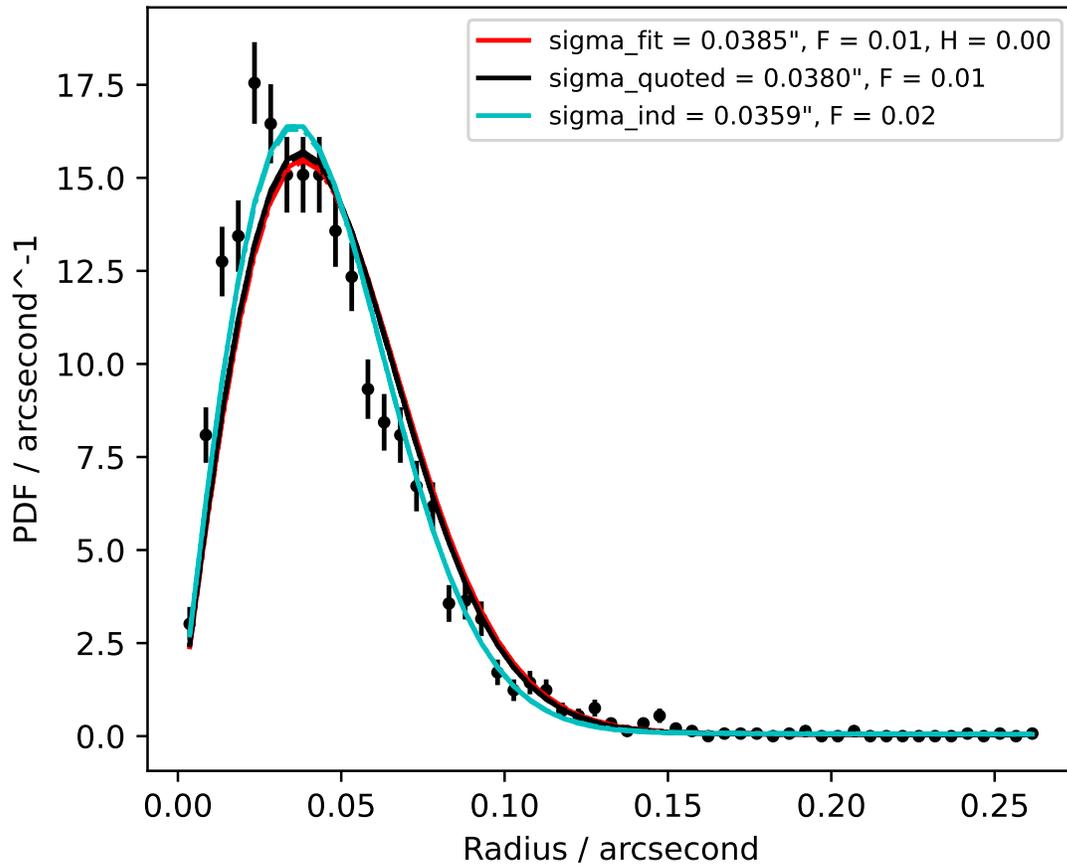


FIGURE 2: Comparison between astrometric precision model and distribution of measured-truth position separations. Figure has the same meanings as Figure 1, except data are $i = 22.75$.

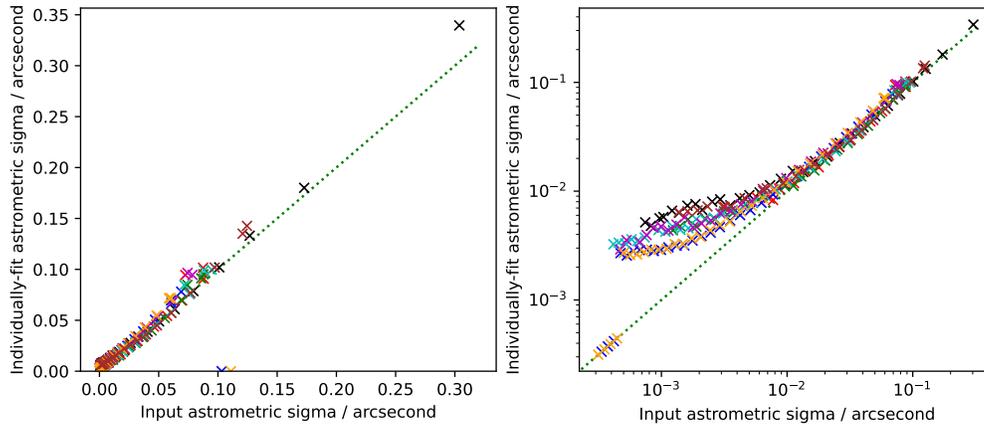


FIGURE 3: Comparison between best-fit individual and pipeline-derived precisions. All magnitude slices for each tract are shown, with decreasing brightness tracking towards the top-right of the figures, at lower SNRs. Each tract is overplotted in a separate colour, and the dotted green line shows the line $y = x$. Left-hand panel shows the relationship in a linear-linear plot, while the right-hand panel shows the scaling on a log-log scale.

choose any parameterisation that makes sense – an equivalent two-parameter model that we could investigate might be $\sigma_{\text{fit}} = a\sigma_{\text{quoted}} + b$, in which we parameterise a linear, instead of quadratic, fit. To limit degeneracies between m and n for small numbers of bins, we require any tract that we derive parameterisations for have five magnitude slices; ideally we should be able to robustly determine n from bright and m from faint objects in our dynamic range.

In Figures 1 and 2, in addition to the best-fit parameterisation for solely that magnitude slice within the given tract, we also show the best fit “hyper-parameter” result (red line). In this particular case, for the “SV 250 2” pointing, we find $m = 1.0052$ and $n = 0.0047$ arcsec, or a roughly constant quoted-fit astrometric precision at faint magnitudes (where n is swamped by lower SNRs and higher relative image noise), and a ≈ 5 mas systematic floor to the uncertainties at bright magnitudes. Figure 4 summarises these results for all tracts.

Overall, across all pointings simulated on the sky, we see a parameterisation that gives good agreement between lower-SNR precisions and the ensemble scatter of measured detections. m , the multiplicative factor in our functional relationship, varies from 0.96 to 1.19, while n , the systematic additive uncertainty, is measured to be between 3 and 7 mas. As also shown in Figure 4, there is tentative evidence for a relationship between these hyper-parameters and on-sky density, although this is driven by only a few high-density pointings, and thus more data will be needed to understand whether this is a real effect.

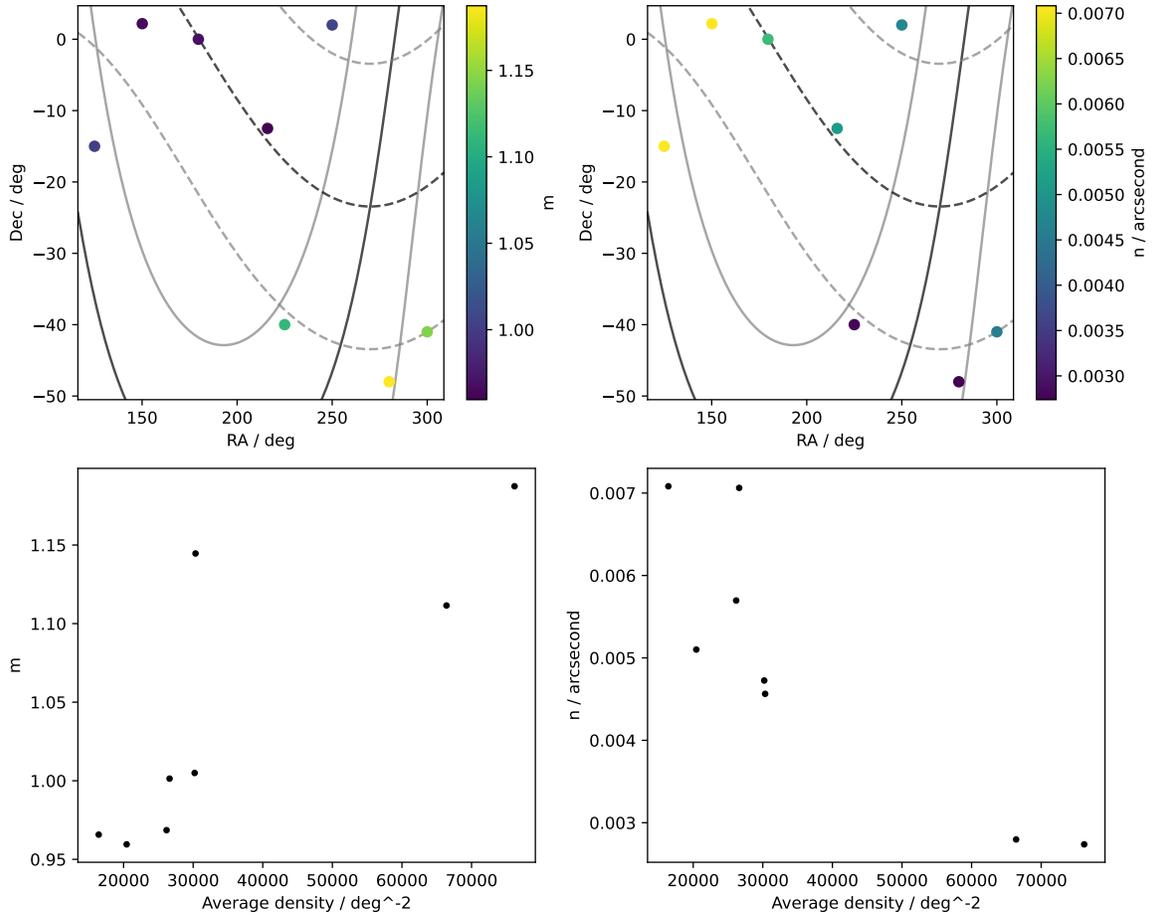


FIGURE 4: Summary of quadratic-fit parameterisation of astrometric precisions across all simulated pointing *Source* tables. Top row shows the m and n parameters from the relationship $\sigma_{\text{fit}} = \sqrt{(m\sigma_{\text{quoted}})^2 + n^2}$ respectively as a function of sky position of the pointing. For these top rows, solid and dotted lines show the Galactic and Ecliptic planes at $b = 0$ deg and $|b| = 20$ deg respectively. Bottom row shows m and n as a function of the average source density in each pointing showing tentative trends with the number of objects per unit area.

4.1.2 Source Table Astrometry

Preliminary evaluation of the pipeline’s ability to determine the positions and corresponding uncertainties – solely considering the precision with which the centroid was measured, dependent largely on the noise in the relevant image pixels – for the `Source` table is positive. The main conclusions are that a quadratic fit, systematic bright uncertainty combined with a scaled version of the quoted astrometric precision, produces good fits for all dynamic ranges of all sightlines. These parameterisations suggest that the main missing component of centroid uncertainty is a ≈ 5 mas systematic that should be accounted for. If the systematic is not included, the astrometric precisions of bright objects, above an SNR of approximately 100, would be over-estimated as compared to the scatter of their centroid measurements. On the other hand, at fainter magnitudes where this systematic is less dominant, we have good agreement between uncertainty and position-measurement scatter, suggesting that the pipeline is correctly determining the “statistical” uncertainty for these objects. However, we require further data, either observations or simulations, in more dense (stellar) fields to determine the validity of the ≈ 15 -20% underestimation of low-SNR uncertainties in the most-crowded fields simulated.

4.2 Pipeline Performance For The Object Table

As discussed previously (3), there are two different versions of each catalogue that the LSST pipeline will produce: the per-visit, time-series-enabling `Source` table, and the deep-stack, coadd `Object` table. It is important to verify the results of both catalogues, as there are key differences, and additional steps, that go into each reduction. `Source` table reductions are performed on each image separately, and hence there is no potential for the stacking of images to degrade the quality of the resultant catalogue, or create unforeseen issues. On the other hand, the `Object` table will probe much fainter – at 10-year depths, approximately 3.5 magnitudes for the “Wide-Fast-Deep”-cadence portions of the survey (Ivezić et al., 2019) – and therefore we can investigate the performance of the pipeline in much more crowded skies, relatively speaking.

The process, apart from the method by which we assemble our dataset for comparison to “truth,” is very similar. Instead of full pointings, we now have to analyse the sky as broken down into its tracts. Here, we often have a tract that contains a few edges of visits, and hence very few detections; this both affects our ability to analyse the output datasets, but also af-

ffects the robustness of the pipeline’s creation of the dataset, so we enforce minimum cuts as detailed in the previous section.

An example, similar to Figure 2, is given in Figure 5 for the case of the Object table, tract 3384, for $i = 22.5$. Our best-fit individual astrometric precision (cyan line) is in good agreement with the scatter of cross-match separations. Similar to the Source table case, our quoted precision (black line) is slightly underestimated, suggesting a small missing systematic component to the astrometric precisions.

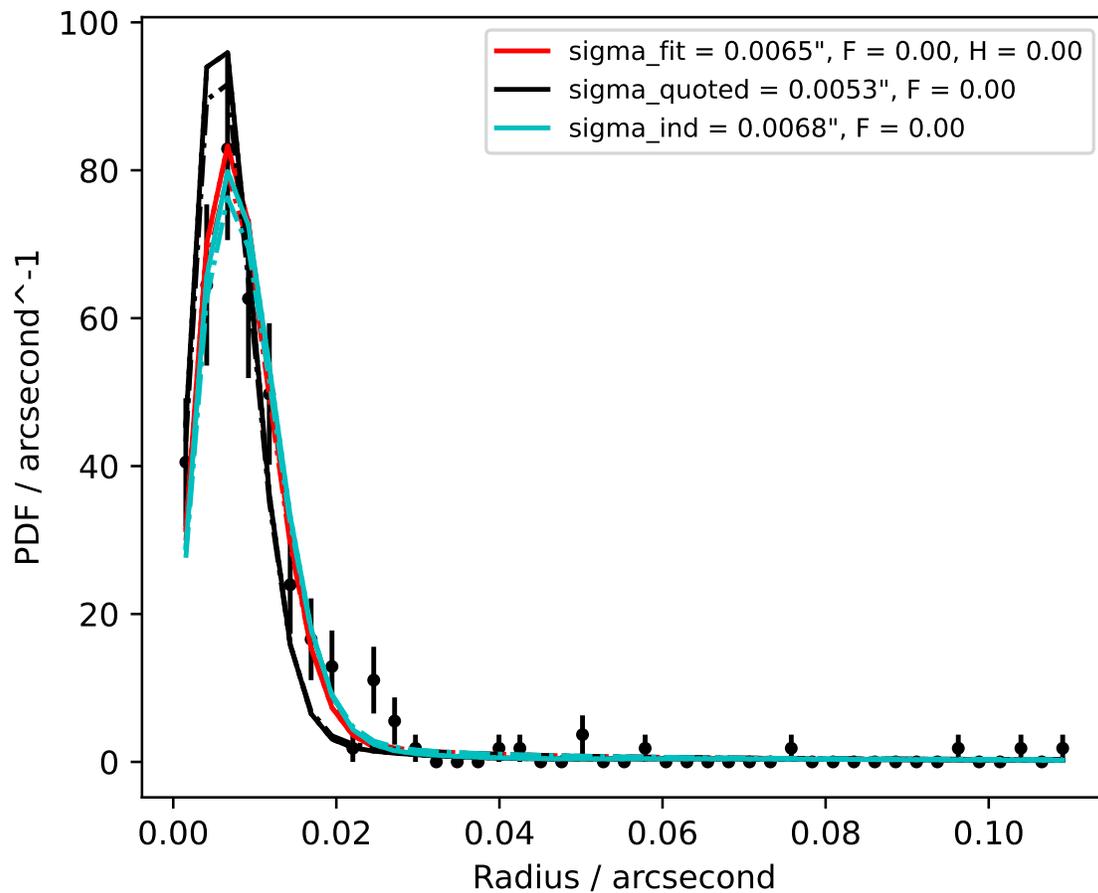


FIGURE 5: Comparison between astrometric precision model and distribution of measured-truth position separations. All symbols and lines have the same meaning as Figure 1. Object table analysis is now shown, however, for $i = 22.5$ detections in tract 3384.

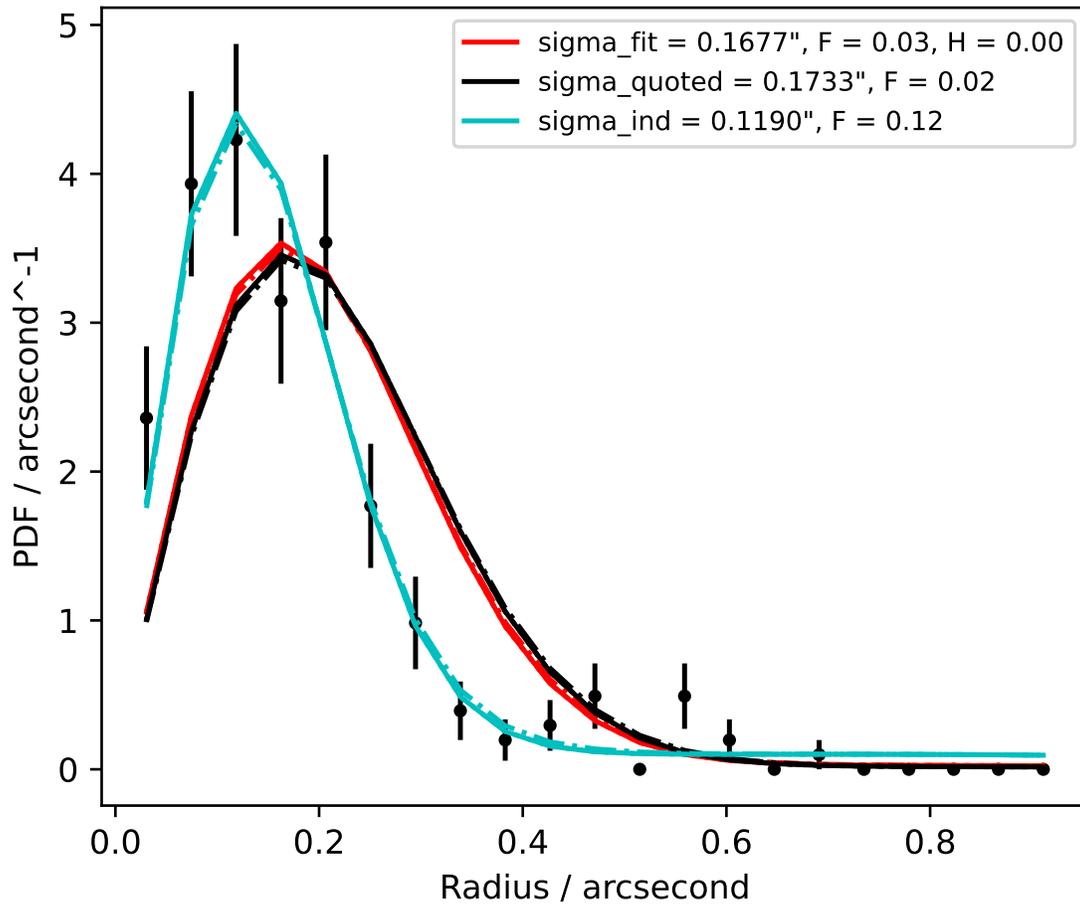


FIGURE 6: Comparison between astrometric precision model and distribution of measured-truth position separations. Figure is the same as Figure 5, except the data shown are $i = 25.75$.

4.2.1 Object Table Quadratic-Fit Precision Breakdown

However, as shown in Figure 6, we see a reversal of what we the previous Source table faint-magnitude results. Now the individual-slice best-fit precision is smaller than the quoted pipeline value. This is fainter than the single-visit completeness limit, so beyond the scope of comparison with the Source table results.

This results in a quadratic-fit scaling relation breakdown; however, this has a relatively simple explanation. Across all magnitude slices, the simultaneous maximum-likelihood fit will be driven by where there is more information, with higher number counts in the middle of our dynamic range, $23 \lesssim i \lesssim 24.5$. Here, the agreement between quoted astrometric precision and cross-match separations is good, and thus $m \approx 1$. However, at fainter magnitudes than $i = 24.5$, this overturn in standard deviation vs Gaussian- σ cannot be accounted for in our simple model, and thus the hyper-parameter best fit (red line) is simply unable to explain these 25th-magnitude position measurements. At this point we can see that the simplistic two-parameter model is unfit to describe the dependency between astrometric position scatter and pipeline-derived covariance, so we return to the individual-slice precisions.

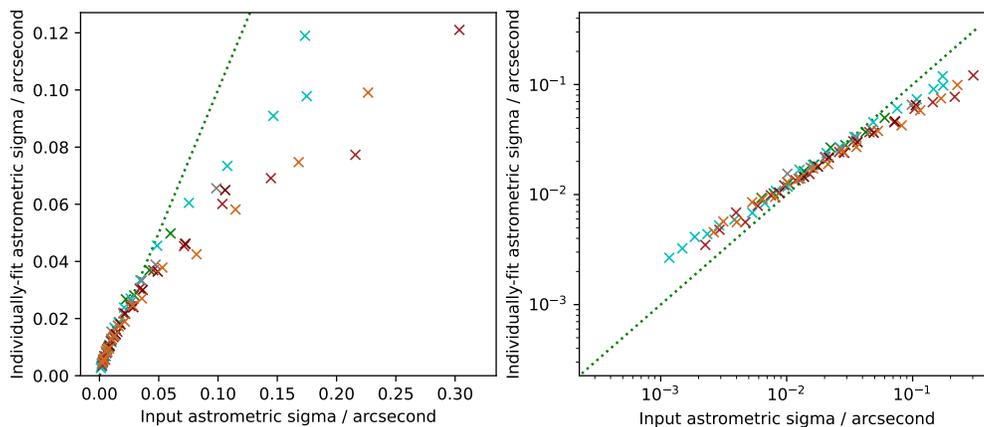


FIGURE 7: Comparison between best-fit individual and pipeline-derived precisions, for all OR3 tracts that pass the required quality and source-count cuts. All symbols and lines have the same meaning as Figure 3.

4.2.2 Object Table Individual-Fit Precision Breakdown

The relationship between the pipeline-derived astrometric precision and single-slice precision is less easy to explain. Figure 7 shows the summary of all individually fit precisions, for all

magnitude slices, across all tracts. Here, the breakdown of the connection between pipeline-derived uncertainties and the scatter in measured positions can clearly be seen, across all sky regions. More clearly visible in the right-hand figure panel, instead of our expected sum-in-quadrature model of systematic precision and scaled-pipeline precision, we see a power law fit, $y = Ax^k$, with $A \approx 0.38$, $k \approx 0.75$, across all tracts analysed.

We are currently unable to explain this power-law scaling relation between the coadd, Object table uncertainty as determined by the LSST pipeline and the precision that best accounts for the position scatter distribution.

4.2.3 Merged Object Table Analysis

We do not see the same evidence of a systematic “plateau” in Object table astrometric precision (Figure 7) that is demonstrated in Figure 3 for the case of the Source table. This is largely due to a lack of available data points at brighter magnitudes, due to the reduced number of observations without repeated measurements of the same bright source. Thus, to obtain better number statistics for these crucial magnitudes we performed a second analysis on the Object table data, in which we simply combined all eight previously analysed tracts together.

In our first attempt we continued to do exactly the same method as previously described, binning cross-match pair separations by small ranges in astrometric precision and magnitude. However, this gave poor results due to the varying magnitude-SNR parameterisation across the sky – with each tract having different sky backgrounds, for example. We therefore chose to perform a more rudimentary, but still elucidating, test on these tract-merged data.

Working purely in SNR, we iterated in even spacings of log-SNR, from 0.5 (SNR of ≈ 3) to 4 (SNR of 10,000). All objects, regardless of magnitude, astrometric precision, local crowding density, etc., were selected in the range from 0.9 to 1.1 times the particular SNR, to account for the logarithmic nature of the selection. The fully parameterised model for false matches was still used, although the functional form of the uncertainties was assumed to be Gaussian, as previously shown to be a valid assumption in this case. Once the false positives were accounted for, in a similar fashion to the previous analyses, the quoted and best-fit precisions were compared, as shown in Figure 8.

Here, with roughly eight times as many objects in our dynamic range, we extract good number statistics in our plots to much brighter magnitudes and higher SNRs, with an SNR of 10,000

roughly corresponding to an i -band magnitude of 17th. This reveals, brighter than an SNR of around 1,000 ($i \approx 20$), a plateauing of astrometric precision, as expected. However, we still do not obtain a linear fit to our astrometric precision dependency at faint magnitudes, and Figure 8 shows a combined-relationship parameterisation, $y = \sqrt{(mx^k)^2 + n^2}$, in which $m \approx 0.31$ and $k \approx 0.73$, similar to the results found in this regime for the individual-tract analyses. We find a “plateau” systematic of roughly $n = 0.6 \text{ mas}$, an order-of-magnitude decrease on the values obtained for the `Source` table results. These are in back-of-the-envelope agreement, however, as each tract has roughly 80-90 i -band visits, in addition to those from other bands that may have gone into the astrometric determinations, and for root-N scaling we need of order 100-150 visits to explain a roughly 10-12 times increase in precision.

This combined-tract analysis thus confirms further the power-law dependency of best-fit astrometric precisions (in the form of ensemble astrometric position scatter) on quoted astrometric precisions (covariances as determined by the pipeline on a per-object basis). It also provides a ballpark figure for the `Object` table systematic uncertainties of sub-milliarcsecond precisions for only a few years’ effective observation depths.

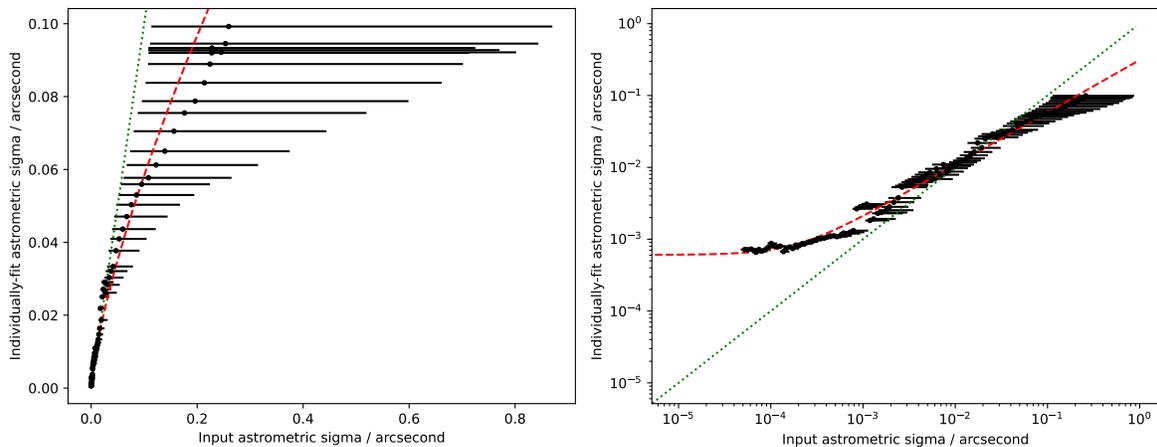


FIGURE 8: Comparison between best-fit individual and pipeline-derived precisions, for all OR3 tracts that pass the required quality and source-count cuts, combined into a single “pointing.” All symbols and lines have the same meaning as Figure 3. Additionally, x-axis errorbars show the 16th and 84th percentile ranges in the quoted astrometric precision for each analysed bin, and a best-fit relationship for $y = \sqrt{(mx^k)^2 + n^2}$ is plotted in a red dotted line.

5 Discussion

5.1 Magnitude-SNR Relationship

A separate diagnosis we can perform is a verification of the relationship between magnitude and SNR. SNR, defined as the ratio of the flux of the source to its photometric uncertainty, should be tightly correlated with brightness. The relations can be verified by investigation of the two extremes: faint objects, where the sky background dominates, and very bright objects where photon-counting statistics should be the majority component.

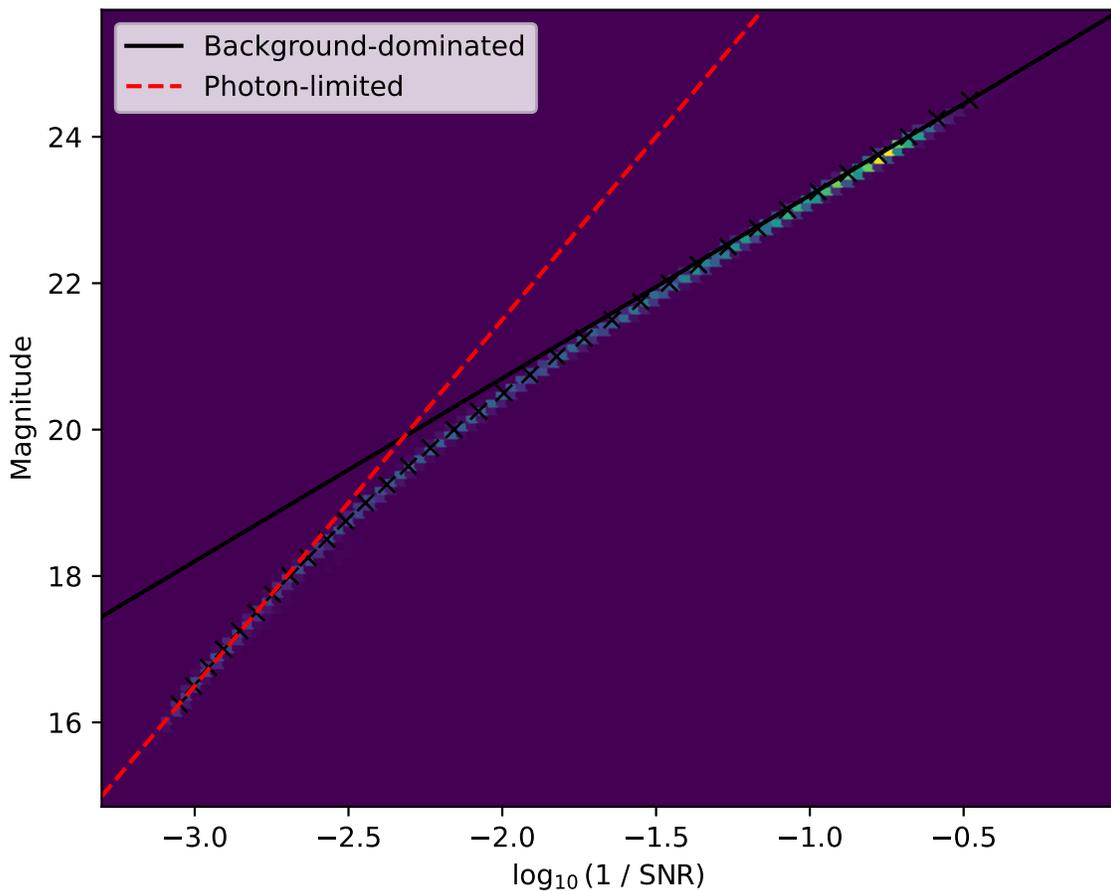


FIGURE 9: Verification of the SNR-magnitude relationship, for Source table objects in the “SV 225 -40” pointing. Background colours show histogram of sources in each magnitude-SNR bin, while black crosses show averages for each “magnitude slice” used in precision-verification analysis. Black and red lines respectively show the expected scaling in the bright and faint magnitude limits.

For the sky-background case, we have $m = -2.5 \log_{10}(\text{flux})$ – for a given definition of “flux” and associated zeropoint that we don’t need to worry about here – and $\text{SNR} \propto \text{flux}/B$, where B is some constant background uncertainty. Rearranging for flux and plugging into the equation

for the magnitude, we get $m = -2.5 \log_{10}(B) - 2.5 \log_{10}(\text{SNR})$. If, as we show in Figure 9, we define $x \equiv -\log_{10}(\text{SNR})$ and $y \equiv m$, then we have, effectively, $y = 2.5x + C$.

On the other hand, the extremely bright objects will have $\text{SNR} \propto \sqrt{\text{flux}}$, since the uncertainty in flux will scale as the square-root of the flux. Now, $m = -2.5 \log_{10}(\text{flux}) = -2.5 \log_{10}(\text{SNR}^2) = -5 \log_{10}(\text{SNR}) = 5x$. Although we did not end up with some scaling parameter D as we did in the sky-background case, we explicitly need to add one, since we don't have any knowledge at this stage of the zeropointing of the magnitudes, and hence simply re-scale our photon-limited case as $y = 5x + D$, such that our brightest data point defines D .

Figure 9 shows the results for a typical case, the Source table for the "SV 225 -40" pointing. Here we can see that our scaling relations hold at both bright and faint fluxes, with a transition between the two at intermediate SNRs. The relationship is also very singular: there is a tight range of recorded SNRs for a given magnitude, or magnitudes measured for a given SNR. This both confirms, in a small way, the photometric portion of the LSST Science Pipelines, but also validates our visit cuts made previously.

5.2 Magnitude-Precision Relationship

A similar verification that can be done on the ensemble data, without binning data up into magnitude-astrometric precision slices to perform cross-match scatter vs uncertainty analysis, is the scaling relation between astrometric precision and SNR. The theoretical relationship between the SNR of a source and the precision with which you can determine its centroid on a detector is derived by King (1983). The precision scales as $\sigma \propto a/\text{SNR}$, with a the "sigma" of the Gaussian approximation to the Airy disk seeing profile, related to its full-width at half-maximum (FWHM), the so-called "seeing" parameter.

This relationship, again for the "SV 225 -40" pointing, is shown in Figure 10, for Source table detections. Here we can, essentially, verify the image quality of the observations, as the intercept of the line should allow the seeing of the image to be determined. A few representative seeings are overlaid on Figure 10, 0.8 to 1.2 arcsec – roughly "good" to "below average" seeing for the Rubin Observatory site – from which we can tentatively suggest that these particular visits were simulated with seeing on the better end of the range.

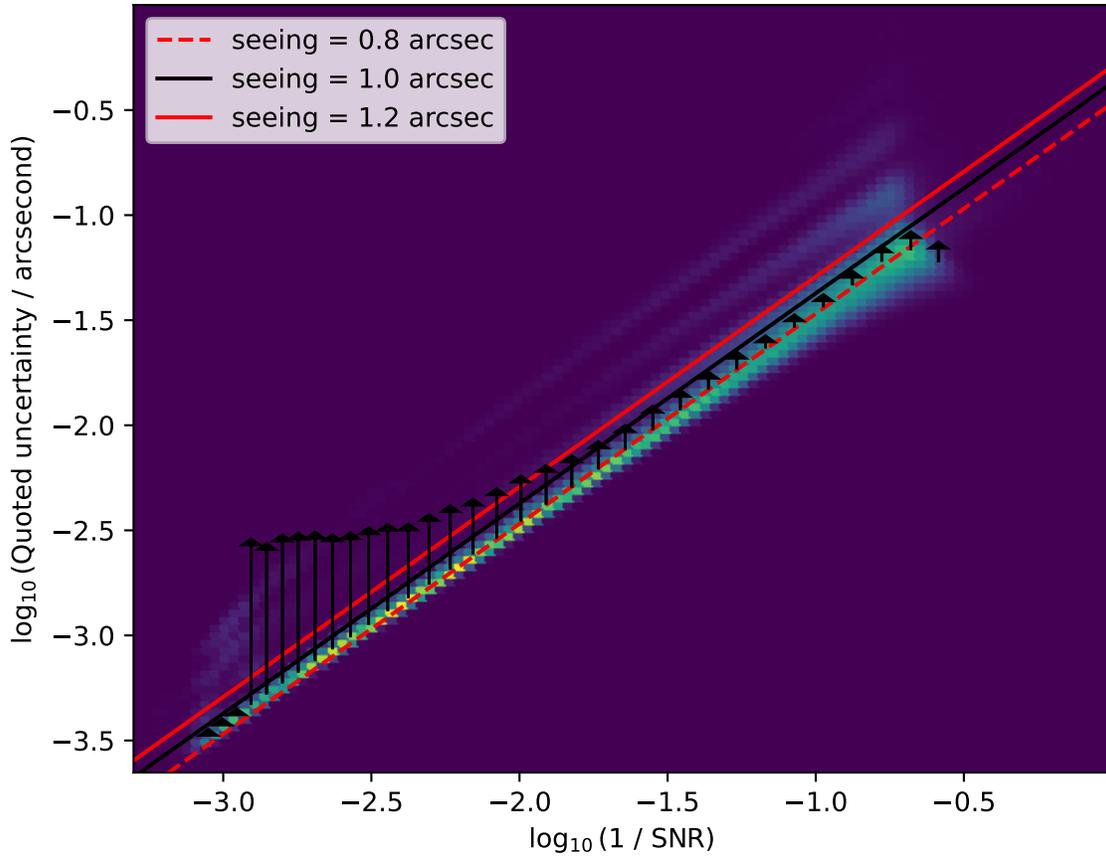


FIGURE 10: Astrometric precision-magnitude relationship, for Source table objects in the “SV 225 -40” pointing. Background colours show histogram of sources in each magnitude-uncertainty bin. Lines show the scaling relations for a few typical “seeing” values. Arrows show the change in astrometric precision required to explain cross-match separation scatter for each magnitude slice bin, with arrows going from “quoted” pipeline precision to “best-fit” precision.

5.2.1 Simulated Seeing and The Effect of Seeing On Astrometric Uncertainty

However, also visible in Figure 10, are a few stripes of unexplainably poor quality uncertainties for a given SNR. Seen at SNRs of approximately 10, but with astrometric uncertainties of around 0.1 arcsec, we are unable to explain these artefacts in the ensemble data. For the Source table analyses, in which a set of individual visits were combined, a plausible if perhaps unlikely explanation is that one or two visits had simulated an unusually large seeing, perhaps to emulate poor weather conditions. In these cases, it is likely that detections would be pushed to lower SNRs as well as correspondingly higher position uncertainties, resulting in a few of these “ghost” relationships in our ensemble data.

However, we also see these stripes in the Object table, coadd catalogues, at SNRs of greater than 100 with astrometric precisions of order 0.01 arcsec, for a seeing of approximately 2.5 arcsec. Since these data are derived from the deep-stack images, it seems less likely that we would obtain detections solely from such poor-quality imaging, if such a high seeing was even simulated. Given the ghost-relationships appear in both coadd and single-visit images, we simply highlight them here for further potential investigation, and conclude that the *overall* relationship between SNR and statistical astrometric precision is robust in the OR3 data.

5.3 Position Scatter vs. Centroid Uncertainty

We reported in the previous section the changes required to update the pipeline-quoted astrometric precisions to bring them into alignment with the scatter in repeat measurement of the positions of detections. However, especially in the case of the Object table results, we should emphasise that the conclusions may be made backwards. While the conclusion can certainly be made that there is *tension* between the positions and corresponding uncertainties in these first-look data, we have implicitly trusted the maximum-likelihood positions more than the second-order information in the covariance matrix. It could in fact be the case that the quoted coadd astrometric uncertainties are good, and there is some issue affecting the pipeline-derived coordinates of those objects.

On the other hand, while the King (1983) scaling law is good (cf. Figure 10), with all Source table pointings producing strong agreement, Object table tract analyses only show good agreement for their brightest detections. The analysis performed on the merged-tract data suggests that the relationship between pipeline-quoted and “scatter-based” uncertainties holds across the

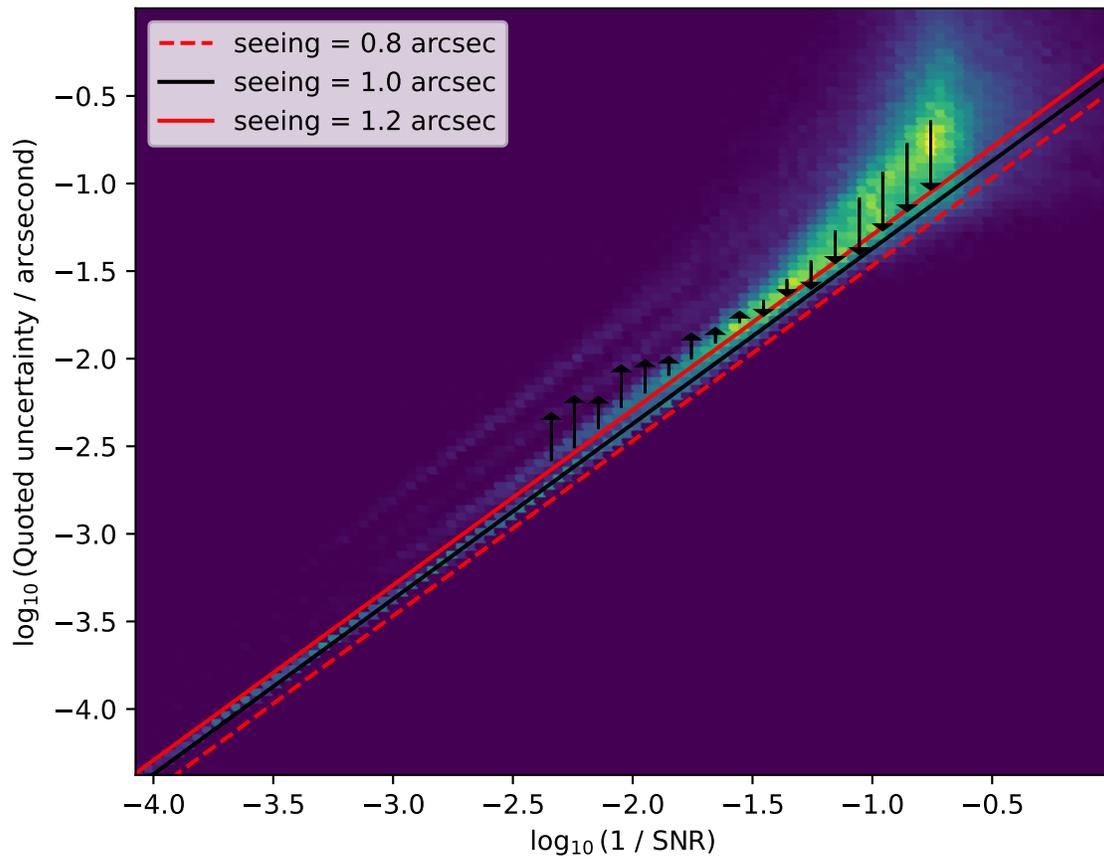


FIGURE 11: Astrometric precision-magnitude relationship, for Object table objects in tract 9980. All symbols and lines have the same meaning as Figure 10.

whole sky, if one takes SNR to be the “true” parameter of pipeline performance. Here, we then see a breakdown between quoted precision and SNR – cf. the increasingly large inter-centile ranges in Figure 8, additionally demonstrated in Figure 11. Unlike the `Source` table relationship (Figure 10), the `Object` table covariances deviate from the theoretical scaling law lower than an SNR of roughly 50 ($i \approx 23.5$). So while we caution against blind faith in the conclusions drawn, and suggest independent verification of the two potential causes of `Object` table astrometry tension, we tentatively suggest that SNR may be the “tiebreaker” which confirms the covariances as being the parameter that is incorrectly determined.

5.4 Comparison with Ivezić et al. (2019) Predictions

Ivezić et al. (2019) describe, in section 3.2.3, the predicted precision of LSST observations. They cite the precision relationship as being $700\text{mas} / \text{SNR}$ with a 10mas independent systematic. Our verification of the seeing being approximately 0.8 arcsec confirms the validity, in a cyclical fashion, of this assumption, although we note that if one follows the King (1983) methodology one must use the Gaussian *sigma* rather than the FWHM, and assume this correction factor was included in the Ivezić et al. results. Ivezić et al.’s table 3 lists predicted typical single-visit *r*-band precisions for 21st, 22nd, 23rd, and 24th magnitude: 11, 15, 31, and 74 mas respectively. In comparison, in the *i*-band – and thus an ever-so-slightly apples-to-oranges comparison – we obtain best-fit precisions of 7, 14, 31, and 78 mas for, e.g., our “SV 225 -40” pointing. However, Ivezić et al. (2019) note that a more accurate systematic astrometric uncertainty could be slightly lower, more like 7mas , decreasing in higher-density fields. Our preliminary single-visit analysis confirms these assumptions, both the faint-magnitude statistical uncertainties, but also the expected systematic uncertainties, with our values for n ranging from 3mas to 7mas , decreasing with increasing field density. However, this confirmation should be taken lightly, as it uses simulated data to confirm the theoretical framework that potentially went into those simulations!

6 Conclusion

We have performed a first-pass analysis of the validity of astrometric positions and their uncertainties as produced by the LSST pipelines. The positions of ensembles of objects with similar properties (magnitude, SNR, local on-sky source density, etc.) are compared to recorded positions from the Operations Rehearsal 3 simulation truth table – which can be substituted with any other much-higher-precision dataset, as will be the case for future work. From this we

obtain astrometric precisions as determined by the scatter of measured positions, to which we can compare the precisions of the objects as given by the pipeline-produced catalogue. We investigate both isolated best-fit precisions and a parameterisation for the scaling relation that involves a missing systematic uncertainty, n , and a multiplicative scaling factor for the quoted uncertainty, m .

For the Source table, we find that $m \approx 1$ across most of the sky, but suggest that for the brightest objects an approximately 0.005 arcsec (5 mas) systematic precision be added in quadrature, such that the precisions of these objects then match the scatter in measured positions. However, we are currently unable to report confident results for the Object table as our best-fit precisions are in disagreement with both quoted precisions and any reasonable model for scaling relation. Instead we find a power-law dependency between scatter-based precision and pipeline-derived covariance, with an exponential of $b \approx 0.75$. We find a tentative sub-milliarcsecond systematic precision needed for the brightest detections in coadded images.

Finally, we investigate a few other relationships the data can verify, such as between SNR and magnitude or SNR and statistical astrometric precision, all of which appear to be robust and produce characteristic values, such as the quality of the seeing in the images, that are sensible.

A References

Budavári, T., Szalay, A.S., 2008, ApJ, 679, 301 (arXiv:0707.1611), doi:10.1086/587156, ADS Link

Gaia Collaboration, Brown, A.G.A., Vallenari, A., et al., 2021, A&A, 649, A1, URL <https://ui.adsabs.harvard.edu/abs/2021A&A...649A...1G> (arXiv:2012.01533), doi:10.1051/0004-6361/202039657

Ivezić, Ž., Kahn, S.M., Tyson, J.A., et al., 2019, ApJ, 873, 111, URL <https://ui.adsabs.harvard.edu/abs/2019ApJ...873..111I> (arXiv:0805.2366), doi:10.3847/1538-4357/ab042c

King, I.R., 1983, PASP, 95, 163, doi:10.1086/131139, ADS Link

Pineau, F.X., Derriere, S., Motch, C., et al., 2017, A&A, 597, A89 (arXiv:1609.00818), doi:10.1051/0004-6361/201629219, ADS Link

Wilson, T.J., Naylor, T., 2017, MNRAS, 468, 2517 (arXiv:1703.03750), doi:10.1093/mnras/stx629

Wilson, T.J., Naylor, T., 2018a, MNRAS, 473, 5570 (arXiv:1710.04896v1), doi:10.1093/mnras/stx2692

Wilson, T.J., Naylor, T., 2018b, MNRAS, 481, 2148, URL <http://adsabs.harvard.edu/abs/2018MNRAS.481.2148W> (arXiv:1809.00018), doi:10.1093/mnras/sty2395

B Acronyms

Acronym	Description
DM	Data Management
Dec	Declination
FWHM	Full Width at Half-Maximum
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
OR3	Operations Rehearsal 3
RA	Right Ascension
SE	System Engineering
SNR	Signal to Noise Ratio
SV	Science Validation