Proper motions of the Arches cluster with Keck LGS-Adaptive optics: the first kinematic mass measurement of the Arches

Will Clarkson

Division of Astronomy and Astrophysics, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, USA

E-mail: clarkson@astro.ucla.edu

Abstract.

We report the first detection of intrinsic velocity dispersion for the Arches - a young ($\sim$ 2 Myr), massive ($10^4 M_\odot$) starburst cluster located at only 26 pc in projection from the galactic center. This was accomplished using proper motion measurements within the central ($10'' \times 10''$) of the cluster, obtained with the laser guide star adaptive optics system at Keck Observatory over a 3 year time baseline (2006-2009). The expected velocity dispersion is sufficiently small that this is only possible with uniform data sets, which improve the proper motion measurements by a factor $\sim$ 5 over previous measurements from heterogeneous instruments. By careful, simultaneous accounting of the cluster and field contaminant distributions as well as the possible sources of measurement uncertainties, we estimate the internal velocity dispersion to be 0.15 ± 0.01 mas yr^{-1}, which corresponds to 5.7 ± 0.4 km s^{-1} at a distance of 8 kpc.

We employ two commonly used mass relations to convert the Arches velocity dispersion into a kinematic enclosed mass estimate. The standard core fitting relationship yields $(10 \pm 5) \times 10^3 M_\odot$ within the cluster half-mass radius of 0.4 pc, while the Leonard-Merritt moment-based estimator indicates $(11 \pm 2) \times 10^3 M_\odot$ within 0.2 pc. This provides the first test of mass estimates based on extrapolations of photometrically accessible stars using the three leading candidates for the present day mass function. While the mass predicted by a Salpeter mass-function without low-mass truncation is marginally above our kinematic estimate, all PDMF models still statistically consistent with the kinematic constraints. Extension of the radial coverage of the proper motion measurements should enable a more robust discrimination between PDMF possibilities.

22 July 2010 - *Sections 3-5 significantly revised (4 & 5 completely re-written as they seem to have confused most readers) . Abstract slightly updated; new figures and tables added. I have very good comments in-hand for Sections 1 & 2.*

1. Introduction

Star formation is a critical element in essentially all models of structure formation and evolution in the Universe. The best locations to study this process are the young (few My), massive ($M \sim 10^4 M_\odot$) clusters (hereafter YMCs). Since nearly all stars likely formed in clusters (e.g. Lada & Lada 2003), these clusters likely form the building blocks of the stellar populations in galaxies. The present-day characteristics of these clusters illustrate the interplay between their primordial conditions (particularly the initial mass function and the degree of primordial mass segregation that may be present), the effect of the tidal field in which the cluster lies, and dynamical mass segregation. Of particular interest is the present-day mass function (PDMF) of these clusters, which provides indirect constraints on the IMF and the degree of mass segregation, and therefore the physics of star formation.

Only seven YMCs younger than 10My are presently known in the Milky Way (Portegies Zwart et al. 2010). Photometric observations of the relatively nearby YMCs Westerland 1 and NGC 3603 indicate mass segregation (Harayama et al. 2008; Brandner et al. 2008) as well as a possibly top-heavy global IMF (Harayama et al. 2008). The YMC known as the Arches provides a more extreme test of cluster formation models due to its present location 26pc in projection from the galactic center (GC). In such a strong tidal field, a cluster like the Arches is expected to disrupt entirely in ~ 5 My (Kim et al. 1999; the present age of the Arches is likely 2-2.5 My; Najarro et al. 2004). Moreover, the high local gas density in the GC environment might produce an unusual IMF, with an overabundance of high-mass stars and a possible truncation at low-mass (Morris 1993: for the Arches, such a truncation must occur $\lesssim 1 M_\odot$; Kim et al. 2006).

Photometric attempts to determine the PDMF of the Arches suffer from crowding and variable extinction effects. Within the half-mass radius $r_{hm} \sim 0.4$pc (Stolte et al. 2002), several determinations of the mass function suggest a top-heavy PDMF with ($\Gamma \simeq -1.05 - -0.75$; Figer 1999; Stolte et al. 2005; 2002). However a recent estimate including corrections for differential extinction variations as well as radial variations in A_K itself, suggested that the global PDMF within r_{hm} is consistent with Salpeter ($\Gamma \simeq -1 \pm 0.2$; Espinoza et al. 2009).

The different PDMFs predict different total mass of stars M_{hm} within r_{hm}. Conditions in the inner Milky Way might lead to a truncation of the mass function at $\sim$ few $\times 0.1 M_\odot$ (Morris 1993); for truncation mass 1 $M_\odot$ and $0.1 M_\odot$, the top-heavy PDMF predicts $M_{hm} \simeq 1.08 \times 10^4 M_\odot$ and $1.2 \times 10^4 M_\odot$ respectively (Figer et al. 1999; Kim et al. 2000). In contrast, the Salpeter PDMF predicts $(2 \pm 0.6) \times 10^4 M_\odot$ and $(3.1 \pm 0.6) \times 10^4 M_\odot$ respectively (Espinoza et al. 2009). A *kinematic* measure of M_{hm} provides a direct test of the PDMF of the Arches cluster. In a pioneering proper motion study of the Arches, Stolte et al. (2008) used one epoch each of VLT/NACO and Keck/NIRC2 separated by 4.3 years to measure the motion of the cluster. However differential distortion between the cameras limited the proper motion precision to ~ 0.7 mas yr^{-1}, somewhat too coarse to measure the velocity dispersion, for which

the expected order of magnitude is about ~ 0.2 mas yr^{-1}‡.

We have observed the central $10'' \times 10''$ of the Arches across five epochs with a uniform observational setup (PIs Morris, Ghez). Using NIRC2 on Keck-2, behind the LGS Adaptive Optics facility (e.g. Wizinowich et al., Ghez et al. 2005), these cross-instrument systematics encountered by Stolte et al. (2008) are not present in our observations, and we are able to attain proper motion measurements with error lower than the expected velocity dispersion. We report here on our results, which provide the first kinematic mass estimate of the Arches cluster.

This communication is organised as follows. Section 2 describes the observations and positional measurement technique, while Section 3 describes the process of proper motion measurement and error assignment. Section 4 describes the techniques used to fit the cluster membership probabilities and kinematic parameters, the results of which are given in Section 5. Section 6 provides our mass measurement and new bulk motion measurement for the Arches, and briefly discusses the implications.

2. Observations & Measurements

Observations of the central ($10'' \times 10''$) of the Arches cluster were obtained between May 2006 and May 2009 with the Keck near-infrared camera (NIRC2: PI K. Matthews), behind the Laser Guide Star Adaptive Optics (LGSAO; Wizinowich et al. 2006; van Dam et al. 2006) system on the W. M. Keck II 10-meter telescope. All observations were obtained with the narrow-field mode of NIRC2 (field of view $10.2'' \times 10.2''$), which has a pixel scale of 9.952 $\pm$0.003 mas pix^{-1} (Yelda et al. 2010; hereafter Y10). Observations were taken in the K' filter ($\Delta\lambda = 0.35\mu m$, $\lambda_0 = 2.12\mu m$). Five epochs of the central field in K' have now been taken (Table 1), the second of which (2006 July 18) was reported in Stolte et al. (2008; 2010). Observations were designed to be as uniform as possible across the epochs, with detector-Y commanded to align with the S-N direction at each epoch, with the same pseudorandom dither pattern within a $0.7'' \times 0.7''$ box applied at each epoch, (Y10 and refs therein), and over as uniform a range of zenith angles as practical.

3. Analysis

Our goal is to extract relative proper motions of Arches stars against the field, which is mostly comprised of bulge stars. The analysis proceeds in the following stages: (1). Positions are estimated from a master-image at each epoch (Section 3.1). (2) The extracted positions are transformed into a common reference frame using likely cluster members, and proper motions extracted from the positional time-series in this frame (Section 3.2). (3) Possible sources of proper motion uncertainties are explored

‡ We use the term "velocity dispersion" to refer to both the dispersion in mas yr^{-1} and km s^{-1} throughout.

(Sections 3.3-3.6) and, when additional error sources are identified, motions re-extracted incorporating the updated errors.

3.1. Production of star-lists from each epoch

The procedures used to produce lists of stellar positions and fluxes from the image-sets at each epoch has been fully described elsewhere (Y10 and refs. therein); here we recapitulate briefly the aspects relevant for the present work. Images are calibrated and corrected for distortion and differential atmospheric refraction (Y10). Within an epoch, the corrected images are combined into a mean frame using positional shifts only, weighting by the Strehl ratio estimate for each image. Images are combined using the Drizzle algorithm (Fruchter & Hook 2002), and the mean frame is not supersampled. The shifts to use are estimated using cross-correlation of the scene between images. By combining using shifts only, we average through any rapid variations in effective distortion between images, and average over slow drifts in image orientation ($\sim 1'$ over the course of a night). A modified version of the IDL routine *Starfinder* (Diolaiti et al.) is then used to measure star positions in the mean frame (Y10 & refs. therein).

Within the epoch, the stack of images is also divided into three subsets spanning the same range of Strehl ratio and evolution, from which three submaps - mean images of each of the three subsets - are produced and positions measured. The rms of stellar position measurements across the three submaps is evaluated for each star and scaled by $1/\sqrt{3}$ to produce an estimate of the error on the positional measurement from the mean of the three submaps. This provides a measure of the random measurement error within an epoch (hereafter "centroiding error"; Figure 1).

3.2. Transformations to common reference frame and proper motion measurement

The Arches cluster moves rapidly with respect to the field (~ 5 mas yr^{-1}) and accounts for most of the stars in the field of view (Stolte et al. 2008). We therefore measure motions in the reference frame in which the cluster is at rest§. Details of this process are given in the Appendix; here we outline the important considerations.

First, stars are matched across epochs to produce a master catalogue containing all the original position measurements of each star. Each star-list is transformed to the frame of a single star-list at a chosen epoch t_0 . Motions are estimated from straight-line fits to the transformed measurements in t_0 . These measurements are then used to evaluate a refined cluster reference frame at a chosen time t_{ref} and the original star-lists are then re-mapped onto this refined frame and proper motions re-evaluated in this frame.

The choice of epoch t_0 is determined by the data quality and by the epochs of observation. Proper motions are determined from straight-line fits to positions in the chosen reference frame; choosing a reference frame near the pivot point of these straight

§ Throughout this discussion the term "frame" refers to reference frames of a given epoch or constructed from the positions, not to individual images.

line fits will minimize the error when mapping star-lists. Of the three deepest epochs (Table 1), epoch 2008.5 is closest to this pivot point, and is adopted as t_0 . However, once a first estimate of proper motions has been produced, the constructed reference frame is evaluated at a different time t_{ref} . This epoch is chosen to coincide with the pivot point of the straight-line fits to the greatest number of stars; we adopt $t_{ref} =$ 2008.0.

Because the field and cluster populations show significantly different motion, field stars are removed from consideration when evaluating the transformation parameters that map reference frames onto each other. This is achieved by clipping outliers in the vector point diagram; after a few passes the centroid of the cluster population is at zero motion in the vector point diagram.

When mapping star-lists between epochs, the transformation parameters are estimated by χ^2 minimization using the positional differences in each coordinate separately. The appropriate order of transformation was determined by evaluating the positional residuals as a function of order (Section 3.4). Measurements are inverse-variance weighted using the error estimates in each coordinate for each star. As part of the fitting, three passes of sigma-clipping (with 4σ bounds) are used to reduce sensitivity to measurement-outliers, mismatches or misidentification of cluster members among the reference stars. This typically removes a few stars from the reference list used for the mapping and can be regarded as a fine-tuning of the reference star list for a given mapping.

As the analysis proceeds, additional estimates of positional error become available to use as weights. When mapping star-lists onto each other, errors used in the weighting are the positional errors associated with each star-list. When mapping star-lists onto t_0 , these errors are just the centroiding error in each list; when mapping onto t_{ref} the errors associated with the target frame are the errors in the predicted position in t_{ref} based on the first pass of motion estimation. When evaluating velocities in a given frame, the positional error and error associated with the mapping into this frame (Section 3.3) are added in quadrature. When motions in t_{ref} have been produced, the distribution of velocity χ^2 values is examined for additional sources of random error not yet taken into account. Upon discovery of an additional error source, the entire analysis is repeated. Additional error determined from the velocity fits is associated with variations between epochs, and so is added in quadrature to the centroiding error in the frame mapping and carried through in subsequent frame mapping. The size of this additional error is examined in Section 3.5.

Positional time-series for a selection of objects along with the motions fitted to them, are provided in Figure 2. The proper motion precision adopted is shown in Figure 3. We are claiming sufficient proper motion precision to measure intrinsic dispersion. We now describe the error sources beyond centroiding error (Section 3.1) that are included in our analysis. Table 3 summarises the error budget of our astrometric measurements

3.3. Alignment error

Error in predicted positions due to the mapping between frames was estimated through Monte Carlo resampling: sets of half the reference stars were randomly drawn and the frame-mapping re-fit and re-evaluated for each trial set to produce a trial set of positions as transformed into the target frame. The rms of the differences between these predicted positions and those predicted from the full list is then adopted as the contribution to random positional error due to the fitting process. This error is always included when positional errors after transformation are needed (steps 4 and onward above).

3.4. Transformation order

At the level of a few percent of a pixel (comparable to the velocity dispersion we wish to measure), residual effective distortion (defined here as a spatially coherent pattern in the deviation between measured and true position at the detector) may be present between epochs. These variations might consist of both a spatially random and a spatially correlated part, and might consist of temporally random and/or correlated parts. To quantify spatially correlated time-variation, in step 6 above mappings between reference frames were re-fit separately across each (t-t_0) pair using polynomials of order $0 \leq M \leq 5$ using the same set of likely cluster members at each order. The rms of cluster members in X and Y as transformed to t_0 was evaluated for each order for each epoch (Figure 7), with errors on the rms evaluated from monte carlo resampling and re-fitting in a similar manner to the estimation of alignment errors (Section 3.3). Visual inspection suggests that for each epoch, a significant improvement is gained by using a second-order polynomial; order 3 is sometimes indicated along Y, and 4th or higher orders rarely bring about significant improvement.

The formal significance of the improvement of the fit when stepping up from order M-1 to M was estimated by evaluating the ratio $(\chi^2_\nu(M\text{-}1)\text{-}\chi^2_\nu(M))/\chi^2_\nu(M)$ for order $1 < M < 5$; assuming the residuals after mapping are indeed χ^2 distributed, this ratio should follow the F-distribution for the corresponding pairs of degrees of freedom for $M-1$ and M (e.g., Bevington & Robinson 2003, ch. 11). This produces a formal false-alarm probability that a difference in badness-of-fit of $\chi^2_\nu(M\text{-}1)\text{-}\chi^2_\nu(M)$ or greater could arise from random chance alone. This suggests that order $M > 3$ is not warranted for fits to either coordinate (Figure 5; Left).

The apparent improvement in fit significance at order $M = 5$ is probably an artefact of overfitting to the ~ 235 reference objects. A control test was conducted where stars at the observed positions were moved randomly under the expected velocity distribution of the cluster or (for 15% of objects) the field, perturbed by measurement error, and subjected to a 2nd order polynomial of similar magnitude to the parameters fitted to the real stellar positions. This indicated that the formal fit statistic is indeed sensitive to the polynomial order, provided the number of reference stars is sufficient. A 5th order polynomial (21 terms, or ~ 10 stars per term) is often spuriously indicated (Figure 5; Right). We therefore adopt a second-order polynomial for the frame mapping when

extracting motions. In principle, relative distortions between epochs might require a more complicated description, but this cannot be determined from the sample at hand.

3.5. Additional random error

When velocities have been extracted, the distribution of χ^2 values from the velocity fits is quite different from that expected if all random errors have correctly been included (Figure 7); clearly additional random positional variation is present between epochs that has not been accounted for by centroiding and alignment error. There are several candidate causes; for example the spatial pattern of PSF variation may differ between epochs, there may be residual relative distortion between reference frames that has not been fully accounted for by the frame mapping; or there may be a temporally random component owing to confusion by unrecognized sources (Section 3.6). We ask what size of additive error ϵ must be added in quadrature to the random error sources estimated thus far, such that the distribution of χ^2 values from the velocity fits most closely matches its expectation under purely random positional error along the time-series. Until this condition is met, the proper motion error reported from the fits to the positional time-series for each object underestimates the true size of the random component of the proper motion error.

Once stars are aligned into the t_{ref} reference frame, velocities are re-fit for trial values of ϵ_X, ϵ_Y, and the resulting distribution of χ^2 values from the velocity fits compared to expectations. Two tests were evaluated to make the comparison. First, the chi-squared test was evaluated for the difference between the χ^2 histogram and the theoretical expectation at each trial additive error. Because of the binning required, this statistic does not vary smoothly with the trial error; to find the minimum, a second-order polynomial was fit to the trough in the fit statistic. Second, the two-sided Kolmogorov-Smirnov test was used as a fit statistic to obviate the need for binning. This produces a smoother distribution of fit statistic, but at the expense of a broader range of allowed values. While the minima returned by the two measures are broadly consistent with each other, we adopt the chi-squared test since it appears to provide a more sensitive determination of the best-fit additive error (Figure 8). Errors on this determination of $\bar{\epsilon}$ are estimated by simulation; sets of positional time-series are constructed under gaussian noise with amplitudes as in the real data and perturbed by additional spatially uniform error ϵ_{in} (without changing the "measured" error). The rms of (input - recovered) values of $\bar{\epsilon}$ is then adopted as the error in the additive error. Only stars with 5 measurements are used to estimate ϵ.

As the balance of dominant error terms evolves with magnitude (e.g. Fritz et al. 2010), we might expect $\bar{\epsilon}$ to also vary with magnitude. In addition, there is some indication that the radial distribution of stars varies with magnitude (Figure 6); thus radially-dependent error-terms may correlate with magnitude. We therefore break the sample into three non-overlapping magnitude bins such that the number of stars with 5 good measurements is approximately uniform across the bins. The additive error and its

uncertainty is then determined for each magnitude bin following the above prescription in the previous paragraph (Figure 9).

Because they describe the mean additional statistical scatter required between epochs, the additive errors $\bar{\epsilon_X}, \bar{\epsilon_Y}$ are applied to the position lists at the stage of frame mapping. The distribution of χ^2 values from the velocity fits after re-mapping and re-fitting including additive error was evaluated for three cases: (1). no additive error; (2). a flat additive error (as determined from the $10 < K' < 16$ sample), and (3). additive error allowed to vary with magnitude. We find that a flat distribution of additive error with magnitude produces a velocity χ^2 distribution significantly more discrepant from statistical expectation than a magnitude-dependent additive error (Figures 7-10).

3.6. Source confusion

When a sufficiently bright star passes within ~ 1 FWHM of the PSF of a star of interest, the shape of its PSF can be sufficiently altered that its position measurement is biased, but not so altered that the measurement is rejected. In some cases this bias can be much larger than the positional measurement error ($\gtrsim 2$ mas for $\Delta K' < 2$; Ghez et al. 2008). The distribution of this confusion bias across the sample of stars depends on the spatial crowding and magnitude distribution of stars in the field of view. To estimate its order of magnitude for the Arches central field, we use the simulations of Fritz et al. (2010), which model the distribution of astrometric bias as a function of magnitude, for a K' distribution appropriate for the nuclear star cluster near the galactic center. The rms of the confusion bias σ_{bias} follows a power law whose normalization depends on the stellar density within the field. Of their three regions of interest, the stellar density within the Arches field matches most closely that of their $3.5''$ sample (Figure 6). This then predicts positional bias σ_{bias} of order 20% of the additive errors ϵ for $K' < 18$ and comparable to ϵ at $K' > 18$ (Table 2).

Relative motion across the PSF of the two components of a confused pair would imprint a spurious motion due to the resulting time-variation of the confusion bias. Inter-epoch variation in the PSF would thus cause varying positional bias between epochs even in the case of components that are perfectly stationary with respect to each other. Under the expectation that PSF variation between epochs is random, this error is subsumed within the additive error above (Section 3.5).

Linear trends in the relative separation of the star and its unrecognized confusing counterpart are in principle more problematic, as the spurious motion thus induced would be impossible to separate from the desired intrinsic motion. Indeed, for some of the rapidly-moving S-stars near the galactic center, apparent deviations from the orbital path on a timescale of up to a few years are clearly visible as the star of interest crosses the region of influence of the confusing source entirely during the timebase of the observations (Ghez et al. 2008; Gillessen et al. 2009); measurements confined to the time-interval of confusion would detect linear motion in the wrong direction entirely. However, for the Arches stars of interest here, relative motions of members of a confused

pair are too slow to have a significant impact on the motions we measure. We assume that the bias changes by $1 \times \sigma_{bias}$ in the time taken for the relative separation of confused components to change by the FWHM of the PSF. With expected velocity dispersion ~ 0.2 mas yr^{-1} (Stolte et al. 2008), confused pairs of cluster objects change their separation by $\lesssim 3\%$ of the FWHM per year, so that the astrometric bias will essentially be static for confused cluster-pairs. Cluster objects confused with field stars may be subject to relative motions ~ 5 mas yr^{-1}; in this case the proper motion bias may approach $\sim 0.1\sigma_{bias}$ yr^{-1}. We conclude that, for our measurements of motions in the Arches central field, proper motion bias due to confusion trends is a very small effect compared to other sources of error (Table 3) and can safely be ignored in our analysis.

4. Cluster Kinematics and Membership Probabilities

The cluster shares the field of view with a significant field population. To fit membership probabilities, we fit the kinematic parameters of the cluster and field components simultaneously with the relative contribution each component makes to the population in the image (Section 4.1). Armed with membership probabilities, we also estimate the velocity dispersion of the cluster and subsamples within the cluster by direct Maximum Likelihood fitting of the motions of likely cluster members (Section 4.2).

4.1. Kinematic fitting and membership probabilities

The field population is likely dominated by stars in the inner region of the Milky Way bulge and the outer regions of the central nuclear disk (Launhardt et al. 2003). Too few field objects are present in our sample to decompose the field population by distance based on our measurements. Both the bulge and the disk have a central concentration along our line of sight, however; for the Bulge we expect to preferentially sample field stars within a few hundred parsecs of the Arches population itself along the line of sight (e.g. Cabrera-Lavers et al. 2007). The velocity signature of the field component should thus be a sum of differential rotation along the line of sight and intrinsic velocity dispersion, sampled from the bulge and nuclear disk. Mass models fit to COBE photometry indicate that at ~ 30 pc from the galactic center, the CND and bulge contributions to the stellar population should be of comparable magnitude, although the uncertainty in the mass model is rather large (Launhardt et al. 2003). We find that the statistics of our field are still too poor to fit for multiple field components, so we parameterize the field population as a single kinematic component. Unlike many cluster studies (e.g. Platais et al. 2000), our field is expected to be significantly asymmetric in the vector point diagram (hereafter VPD), as the bulge velocity dispersion is larger along the galactic plane than perpendicular to it (Howard et al. 2010); at 350pc from the galactic center, for example, proper motion dispersions are of order $5, 3$ mas yr^{-1} along and perpendicular to the galactic plane, respectively (Clarkson et al. 2008).

Because the cluster distribution in the vector point diagram is so much tighter than

the field distribution, fitting to the binned VPD (Stolte et al. 2008) is not appropriate for this dataset as there is no binning factor that allows simultaneous resolution of both cluster and field components. Instead we follow a more classical approach and maximize the likelihood L(data|component fractions, kinematic parameters, measurement errors) without recourse to binning. The component fraction π_k describes the mean surface density of stars in the image that belong to the k'th component; since our field fits entirely within the flat core of the surface density radial profile (Espinoza et al. 2009), we assume that π_k is uniform within our field of view. We do not know the true proper motion of each star, so we marginalize it out; the resulting likelihood L (data) is then the convolution of the likelihood due to the model and that due to measurement error. We follow standard practice (e.g. Jones & Walker, Kozhurina-Platais et al) in using bivariate gaussians to model the cluster in the VPD. Because the field population is sampled along a concentrated radius range, we also parameterize the field with a bivariate gaussian. The likelihood of finding a star at a given location in the VPD is thus given by the sum of K gaussian components:

$$\begin{aligned} L(\vec{v_i}) &= \sum_k^K \pi_k \frac{1}{2\pi|\mathbf{\Sigma}_{ki}|^{1/2}} \exp\left(-\frac{1}{2}(\vec{v_i}-\bar{\mu_k})^T \mathbf{\Sigma}_{ki}^{-1}(\vec{v_i}-\bar{\mu_k})\right) \\ &\equiv \sum_k^K \phi_{ki} \end{aligned} \tag{1}$$

where ϕ_{ki} describes the likelihood of finding a given star in a given component at its location in the VPD. The membership probability of a given star with the k'th component is then the usual (e.g. Kozhurina-Platais et al. 1995)

$$P(k)_i = \frac{\phi_{ki}}{\sum_j^K \phi_{ji}} \tag{2}$$

In (1), $\bar{\mu_k}$ is the centroid of the component in the VPD and

$$\begin{aligned} \mathbf{\Sigma}_{ki} &= \mathbf{Z}_k + \mathbf{S}_i \\ &= \mathbf{R}_k \mathbf{\Lambda}_k \mathbf{R}_k^T + \mathbf{S}_i \end{aligned} \tag{3}$$

is the covariance matrix of the K'th component for the i'th star. The orientation of the component in the VPD is described by the rotation matrix $\mathbf{R}_k$ and its velocity dispersions (squared) form the elements of the diagonal matrix $\mathbf{\Lambda}_k$. The (squared) measurement errors for the given star form the elements of $\mathbf{S}_i$, which is diagonal in detector coordinates. Once the best-fit observed components to the dispersion have been found, the intrinsic velocity dispersions and orientations are then found from the eigenvalues and eigenvectors of $(\mathbf{\Sigma}_{ki} - \mathbf{S}_i)$.

Best-fit parameters and component fractions are found by maximizing $\sum_i^N \ln L(\vec{v_i}) = \sum_i^N \ln\left(\sum_k^K \phi_{ik}\right)$ over the sample of reference stars of interest). To evaluate the fit, we employ the Expectation Maximization (hereafter EM) algorithm, in which the maximum-likelihood π_k and the kinematic parameters are evaluated sequentially and iteratively until convergence. This technique is well-established outside

astronomy (e.g. Chapter 9 of Bishop 2008) and is becoming more commonly employed to mixture problems in astronomy in which binning is undesirable and/or a low number of reference objects is available (e.g. Bovy et al. 2009).

Strongly-varying error is a significant complication, as convergence to best-fit parameters $(\bar{\mu_k}, \mathbf{\Sigma}_{ki})$ is slow and requires some supervision (Bovy et al. 2009). We are investigating the use of Markov Chain Monte Carlo to explore the probability distributions of the best-fit fractions and parameters; this type of approach will probably be essential when including the off-center fields, for which π_k is also expected to vary spatially. We have also investigated using EM with multiple components for the cluster; this technique requires some forcing to ensure the two cluster components have coincident centers in the VPD, likely due to low-number statistics. For the present investigation, we use a single cluster component and choose instead a variant of the technique of Khozurina-Platais et al. (1995 and refs therein) in which membership probabilities for each star are estimated using kinematic parameters fitted only from stars with roughly similar error (so $\mathbf{\Sigma}_{ki} \approx \mathbf{Z}_k + \bar{\mathbf{S}} = \mathbf{\Sigma}_k$ for the sample). When estimating membership probability, we break the sample into overlapping bins two magnitudes wide (so K'=(14.0-16.0), (15.0-17.0)...) and evaluate the membership probability to stars within 0.5 magnitudes of the center of each bin. In this way, the membership probability of every star is estimated from a sample consisting of at least 100 stars, with the star of interst always at least 1 magnitude from the edge of the sample.

To investigate the size of any misclassification biases that remain, synthetic datasets were simulated using the same parameters at all magnitudes and perturbed by the observed proper motion errors. The above fitting process was performed for a large number of trials and the recovered parameters observed as a function of magnitude. We see that a small bias does remain in both the component separation and the velocity dispersion; however the magnitude dependence of the recovered parameters is greatly reduced (Figure 7). We correct the fitted components for these biases.

Errors on the parameters thus fit are estimated by Monte Carlo. For the orientation of the components, the vector separation between components, and the component fractions we use full-sample bootstrap resampling.

4.2. Direct calculation

The above process fails when the sample size is $\lesssim 70$ stars, or contains a negligible field component (as occurs for $K \lesssim 14.5$). For these samples, dispersions are estimated directly from the motions of likely-cluster objects. Making the assumption that the cluster is intrinsically symmetric in the VPD, along each direction we maximize

$$L(\bar{v}, \sigma) = \prod_i^N (2\pi(\sigma^2 + e_i^2))^{-1/2} \exp\left\{-(v_i - \bar{v})^2 / 2(\sigma^2 + e_i^2)\right\} \quad (4)$$

for $\bar{v}, \sigma$ iteratively, where σ is the intrinsic velocity dispersion and e_i the measurement error. For each σ, $\bar{v}$ is obtained by weighted averaging while for each $\bar{v}$, σ is found

numerically by bisection. Errors are estimated by Monte Carlo using bootstrap resampling.

5. Results

5.1. Cluster Velocity Dispersion

The kinematic fitting and direct approaches produce complementary estimates for the intrinsic velocity dispersion of the Arches cluster. Both have been corrected for measurement error using the procedures described above. Tables 7 & 7 show the kinematic parameters fit to cluster and field, and the velocity dispersion estimate for the cluster respectively.

Both estimates yield cluster dispersions with at least 4σ significance for stars at $K < 18$, and usually closer to 8σ. The mean cluster velocity dispersion is 5.7 ± 0.4 km s^{-1} (kinematic fitting) and 5.3 ± 0.3 km s^{-1} (direct calculation). An increase in dispersion with magnitude is admitted by our data. For the brightest magnitude bin ($K < 14$), no field population is apparent so there the velocity dispersion is estimated from direct calculation alone. The kinematic fits suggest that the cluster major and minor axes in the VPD are not significantly discrepant, as expected since cluster stars brighter than $K = 17.5$ were used to define the reference frame.

5.2. Cluster bulk motion and field properties

The field component shows axis ratio that is roughly constant with magnitude, while its contribution to the sample in the field of view increases as fainter objects are probed. Its orientation in the VPD is consistent with the galactic plane, indicating differential rotation.

The bulk motion of the Arches with respect to the field population is 172 ± 15 km s^{-1} (the weighted average of the $14 < K < 16$ and $16 < K < 18$ bins). Including the $18 < K < 20$ bin revises this figure downward slightly (to 153 ± 11 km s^{-1}); however in this magnitude range the proper motion error curve rises steeply with magnitude (Figure 5), so objects this faint may be particularly vulnerable to misclassification bias.

5.3. Search for the cluster center

To decompose motions into radial and tangential directions, the cluster center must be located. We therefore attempt to do so from the spatial variation of cluster velocity dispersion, under the rationale that the peak in the velocity dispersion distribution should coincide with the true cluster center (Anderson & van der Marel 2010). However, the Arches cluster sample affords too few objects to accurately determine the cluster center in this way (Figure 13).

5.4. Limits on an IMBH in the Arches

In principle we can explore (x_c, y_c, M_{bh}) to constrain any central point-mass in the Arches. In practice the upper mass limit may not be very great... Worth a look...

6. Discussion

6.1. Bulk Motion

In their study of the Arches bulk motion, Stolte et al. (2008) found that, although the Arches is unlikely to be on a circular orbit, integration of its path through the potential of the inner Milky Way indicated the cluster was unlikely to pass sufficiently close to the GC to be a plausible source of the young stars in the very inner region of the GC. Our revised motion estimate makes the Arches orbit slightly more compatible with circular motion. Following the arguments of Stolte et al. (2008), if on a circular orbit, $\mathbf{v}.\mathbf{r} = 0$ then demands enclosed mass only 1.5σ above that measured photometrically (Launhardt et al. 2003). At first glance, our new bulk motion supports the conclusions of Stolte et al. (2008) that the Arches is unlikely to be a source of young stars for the GC star cluster. Integration of the Arches motion through the potential of the inner Milky Way using our new motion determination is required to draw further conclusions about the formation and subsequent motion of the cluster.

However, interpretation of the cluster bulk motion is complicated by two factors. Firstly, differential rotation of the field component(s) may make the true field motion a function of (unabsorbed) magnitude. In this scenario, the relative motion of the Arches against the field should therefore vary with magnitude. Secondly, extinction variations along the line of sight coupled with the low-number statistics ($\sim$ few tens of field objects in each magnitude bin; Table 7) may reduce the validity of a gaussian to represent the field component in the first place. Thus our quoted error of 15 km s^{-1} on the Arches bulk motion is likely an underestimate.

6.2. Velocity dispersion and mass

The present day kinematic structure of the Arches is likely to be somewhat discrepant from the predictions of a single-mass, spherical, isothermal cluster model. Full determination of the Arches mass probably requires simultaneous modeling of the mass and velocity profile using high-quality data at least out to the tidal radius (the approach in Schoedel et al. 2009), where models are projected onto the sky for comparison with data at each step. That investigation is beyond the scope of this report. Here we provide first-order estimates of the mass using two popular estimators, and briefly discuss where their shortcomings lie.

Probably the most frequently encountered estimator when only the velocity dispersion in the central region is available, is the core-fitting formula of King (1966); $\rho_0 \propto \sigma_0^2/r_{hm}^2$. For roughly isothermal, single-mass, spherical cluster models where the

external potential is unimportant, this relation links the central density to the velocity dispersion in the center. Adopting the mean velocity dispersion ($> 4\ M_\odot$) yields central density $\rho_0 \approx (3.7 \pm 0.1) \times 10^4 M_\odot$ pc^{-3}. Within a volume 0.4pc in radius, this implies mass $(10.0 \pm 0.2) \times 10^3 M_\odot$. We caution the reader that this estimate rests on several conditions which likely do not hold for the Arches cluster. In the strong tidal field of the GC region, the velocity dispersion outside the core region may be elevated, in which case the velocity field observed in projection towards the core may be elevated. This would cause ρ_0 to be artificially high. In his critique of the core-fitting formula, Merritt (1988) points out that the central velocity dispersion only relates uniquely to ρ_0 if the shape of the velocity ellipsoid is known at *all* radii; in the absence of this information, assuming isotropy can lead to significant errors in the derived velocity dispersion. This mass estimate M_{hm} may thus be incorrect by at least 50%, so we adopt $(10.0 \pm 5.0)\ \times 10^3 M_\odot$.

We also tried the Leonard & Merritt (1989; hereafter LM89) mass estimator, which relies on the observation that for a spherical, nonrotating cluster, the radial variation of two velocity dispersion components is sufficient to determine the mass distribution of the cluster. Assuming constant mass-to-light ratio with distance from the center, we decompose the motions into radial and tangential motions from the cluster center and the enclosed mass estimate M(R) produced from the LM89 relation. We have high-quality data for the inner 0.2pc $\times$ 0.2pc, with five epochs used over a three-year timebase. There is some uncertainty on the precise location of the center of the cluster. The LM89 estimate was therefore determined for several choices of the cluster center. Within the central 0.2pc, the M(R) profile rises steadily to about $(11 \pm 2)\ \times 10^3 M_\odot$. (**This estimate needs updating...**)

Although these estimates have a number of shortcomings, they represent the first kinematic mass measurements produced for this important YMC. Our two estimates suggest a range $(5 - 25)\ \times 10^3 M_\odot$ for the total mass contained within 0.4pc.

This range straddles the predictions from photometric studies of the PDMF, and we are unable *at this stage* to definitively judge whether a top-heavy or truncated PDMF really does obtain within the half-mass radius of the Arches. Of the three possibilities (Salpeter, Salpeter but truncated, top-heavy), the enclosed mass estimate $(31 \pm 6)\ \times 10^3 M_\odot$ predicted by extrapolating a Salpeter mass function all the way to the Hydrogen-burning limit is the least compatible with our data (Espinoza et al. 2009). Thus our motions admit the possibility that the IMF of the Arches may indeed be truncated below 1 $M_\odot$.

7. Conclusion

With uniform observational setup over a sufficient time baseline and careful accounting for a number of sources of proper motion error, we have measured the velocity dispersion of the Arches cluster for the first time. We have used this dispersion to test the photometric estimates of the present-day mass function (PDMF) within the half-mass radius $r_{hm} = 0.4$ pc. The total mass within r_{hm} is likely in the range $(5 - 25) \times 10^3 M_\odot$,

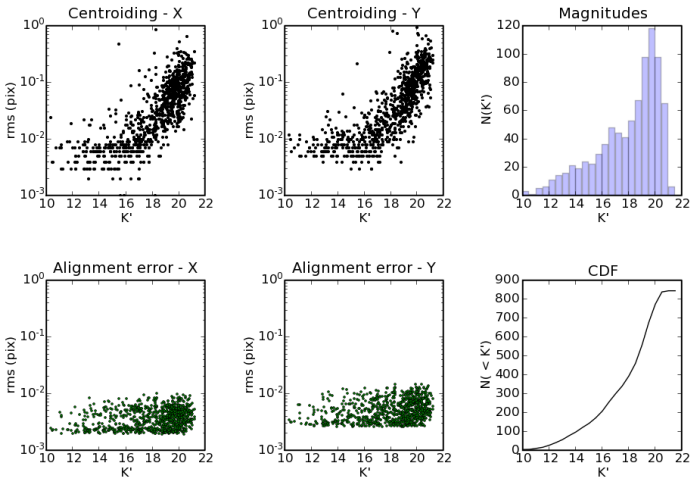

Figure 1. Errors and magnitudes for an example epoch (2008.50). Top left and middle: centroiding errors (i.e. intra-epoch repeatability; Section 3.1). Middle left and middle: alignment errors (Section 3.3). Right column: magnitude histogram and CDF.

which offers tentative support for either a top-heavy PDMF or a truncation in the PDMF at low mass.

However, this mass estimate suffers from significant limitations. On the modeling side, the Arches cluster may rotate, is subject to significant tidal shear which may elevate the observed velocity dispersion, and may not be spherical. On the observational side, we do not yet have sufficiently precise motions outside 0.2pc to constrain M(R) outside this region, and have had to resort to an extrapolation of the well-measured inner field. Observations of the outer fields due in summer 2010 should remedy this situation.

We have also revised the bulk motion of the Arches slightly downward. Our updated motion of 170 ± 18.9 km s^{-1} is only slightly lower than the 212 ± 29 km s^{-1} determined previously, due to two competing biases in the former work that nearly cancel each other. Taken at face value, this supports previous conclusions that the Arches is unlikely to pass within 10pc of the GC. Modeling of the path of the Arches through the inner bulge / GC potential will be the subject of a future communication.

Finally, we have provided the first estimate (to our knowledge) of the velocity dispersion of the Bulge along such a close sight-line to the galactic center; of (102, 65) $\pm$ (10,7) km s^{-1}.

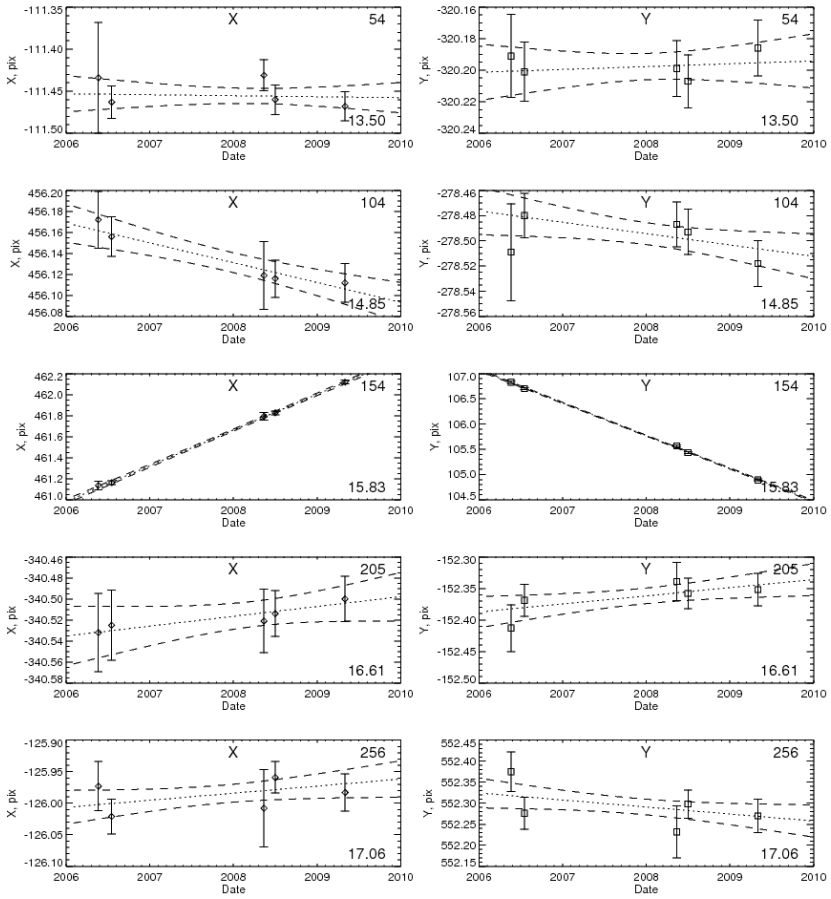

Figure 2. Five example positional time-series. Left columns show motion along X, right columns along Y. Object IDs and K' magnitudes are indicated in the left top and bottom corners respectively. The best-fit straight line to the motions are indicated in each case, as are 1σ positional error curves. Object 154 is likely a field object, as indicated by its large proper motion.

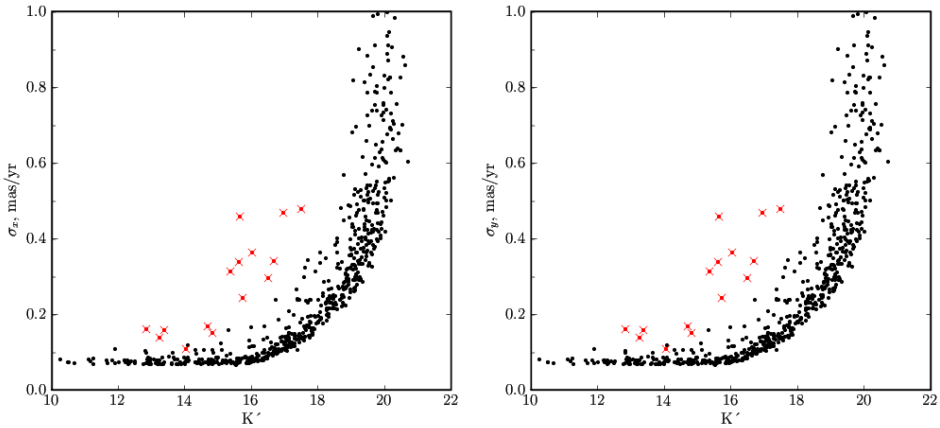

Figure 3. The distribution of adopted proper motion precision (Sections 3.5-3.6 and Table 3). Outliers due to likely mismatches are indicated and removed. *These panels should not be identical - to be updated.*

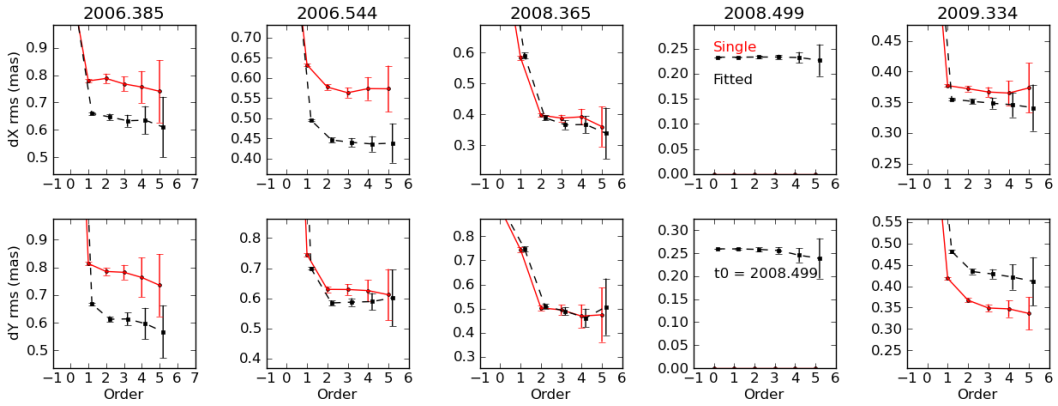

Figure 4. Positional residuals after mapping of each epoch onto the adopted reference frame, as a function of the polynomial transformation order adopted. Red circles: reference frame used is the star-list in 2008.5. Black squares: reference frame constructed by evaluating linear fits to the positional time-series for likely cluster members. Errorbars were estimated by monte carlo using random half-samples. Residuals are evaluated along detector X (*top row*) and detector Y (*bottom row*). We note that (1). Order 2 improves the mapping significantly over order 1, while higher orders are not supported by the data; (2). positions mapped onto the generated reference frame (black) show similar or lower residuals (for all but epoch 2008.37 x) than when mapped onto an individual star-list (red).

Epoch	($t_{int} \times N_{coadd}$) (s)	N_{images}	N_{use}	FWHM (mas)	Strehl	N_*	$N_{*,\mathrm{uncrowd}}$	K'_{max} (mag)
2006 May 21	3.00 × 10	15	15	61.05	0.261	660	649	20.2
2006 Jul 18	3.00 × 10	52	38	56.95	0.349	657	642	20.6
2008 May 13	3.00 × 10	146	72	66.66	0.219	556	536	20.3
2008 Jun 01	3.00 × 10	89	83	54.96	0.373	845	810	21.2
2009 May 02	2.80 × 10	119	108	51.47	0.442	968	917	21.3

Table 1. Summary of observations. Reading left-right, the columns are: Epoch of observation, the total integration time for each image, the number of images observed, the number of images used, the median FWHM and Strehl ratio over the set of accepted images N_{use}, the number of stars measured within the mean image stack in each epoch; the number surviving the cut on proximity to a known neighbour, and finally the faintest magnitude measurement within the epoch.

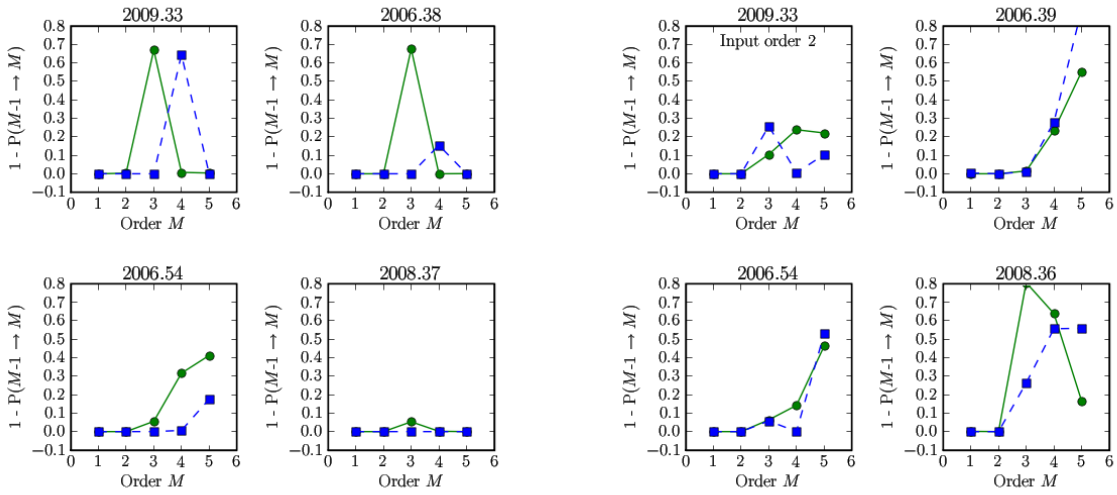

Figure 5. Formal significance of the fit improvement when transformations of increasing order are used to map starlists between epochs. For each step up in order M-1 to M, the false-alarm probability is shown that corresponds to random chance producing a decrease in badness-of-fit at least as great as that observed (Section 3.4). This statistic is evaluated separately for residuals in X (green circles, solid line) and Y (blue squares, dashed line). *Left* 2×2 *panels:* measured positions. *Right* 2×2 *panels:* the same test applied to a control experiment where the observed positions are perturbed under the cluster and field motion distributions, and a second-order polynomial of similar magnitude to that fit from the real data is added to simulate epoch-to-epoch distortion variations. A polynomial of order 2 produces a formally significant improvement in the fitting.

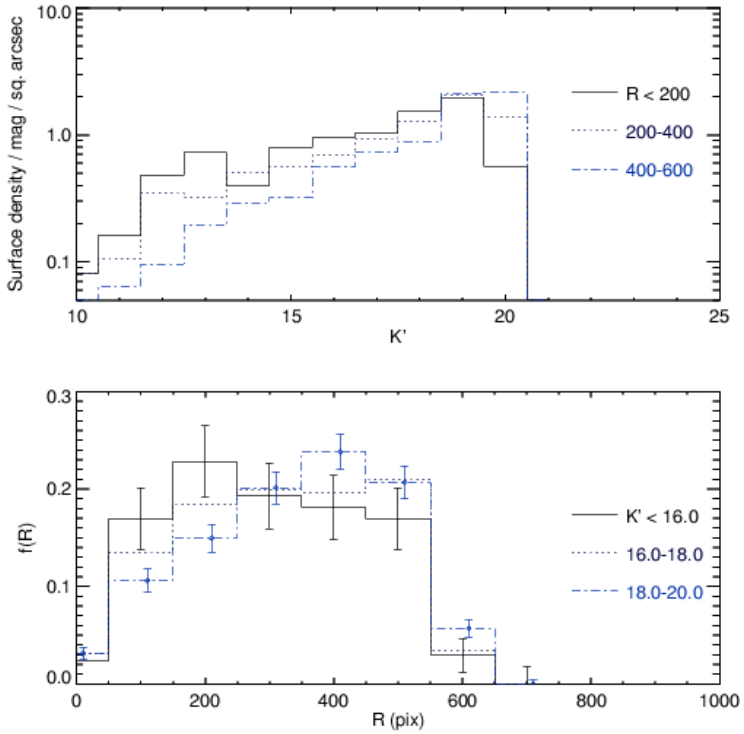

Figure 6. Measured distribution of stars within the field of view, from the 2009.33 epoch. Completeness corrections have not been applied. *Top:* Surface density per magnitude as a function of K', for three bins in radius (as pixels) from the center of the frame. *Bottom:* Radial distribution of stars in three magnitude ranges; Poisson error bars are indicated for the brightest and faintest magnitude ranges.

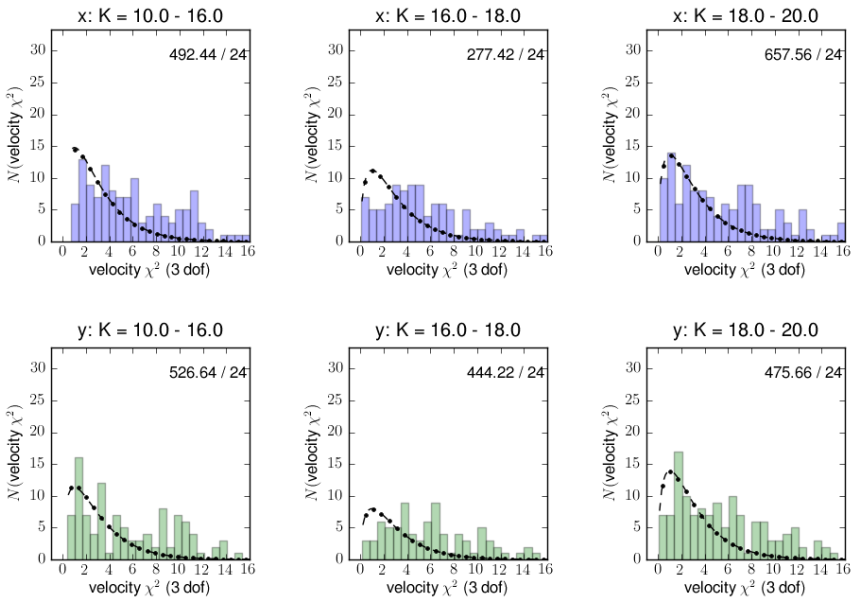

Figure 7. Assessment of the distribution of χ^2 from the velocity fits when motions are fit using only centroiding and alignment errors. Columns break the sample of stars into bright, medium and faint magnitiude bins. Histograms show the distribution of velocity-fit χ^2 values in X (blue; top-row) and Y (green;bottom row). The numbers inset give the values of the chi-square test statistic per degree of freedom for the comparison of the observed histogram to the predicted distribution (dashed curve).

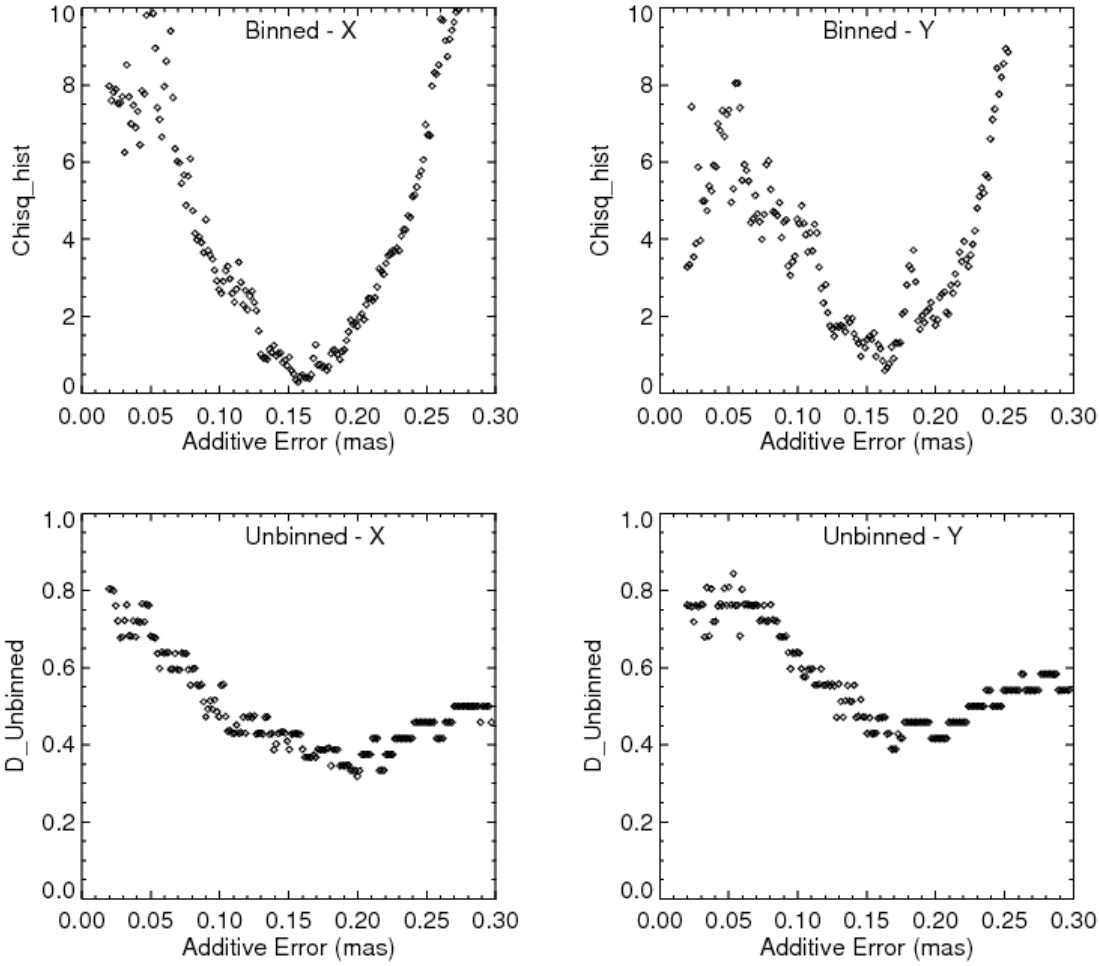

Figure 8. Example determination of the additive errors ϵ_x (left), ϵ_y (right), for objects in the brightness range ($10.0 \leq K' \leq 16.$). For each trial additive error, the chi-square test (top) and two-sided K-S test (bottom) are evaluated, with the minimum found by fitting a polynomial to the trough.

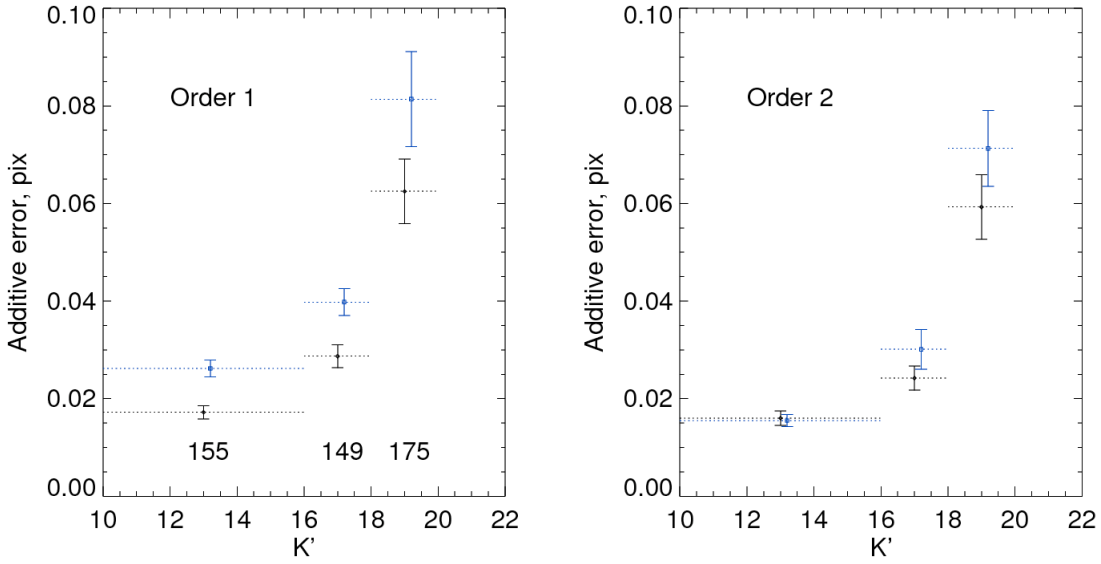

Figure 9. Additive error as a function of magnitude, for two orders of frame:frame transformation. Black diamonds: X; blue squares: Y. Numbers inset in the left panel give the sample size in each magnitude bin. When frames are mapped with a second-order polynomial, the additive errors in each direction become more consistent with each other.

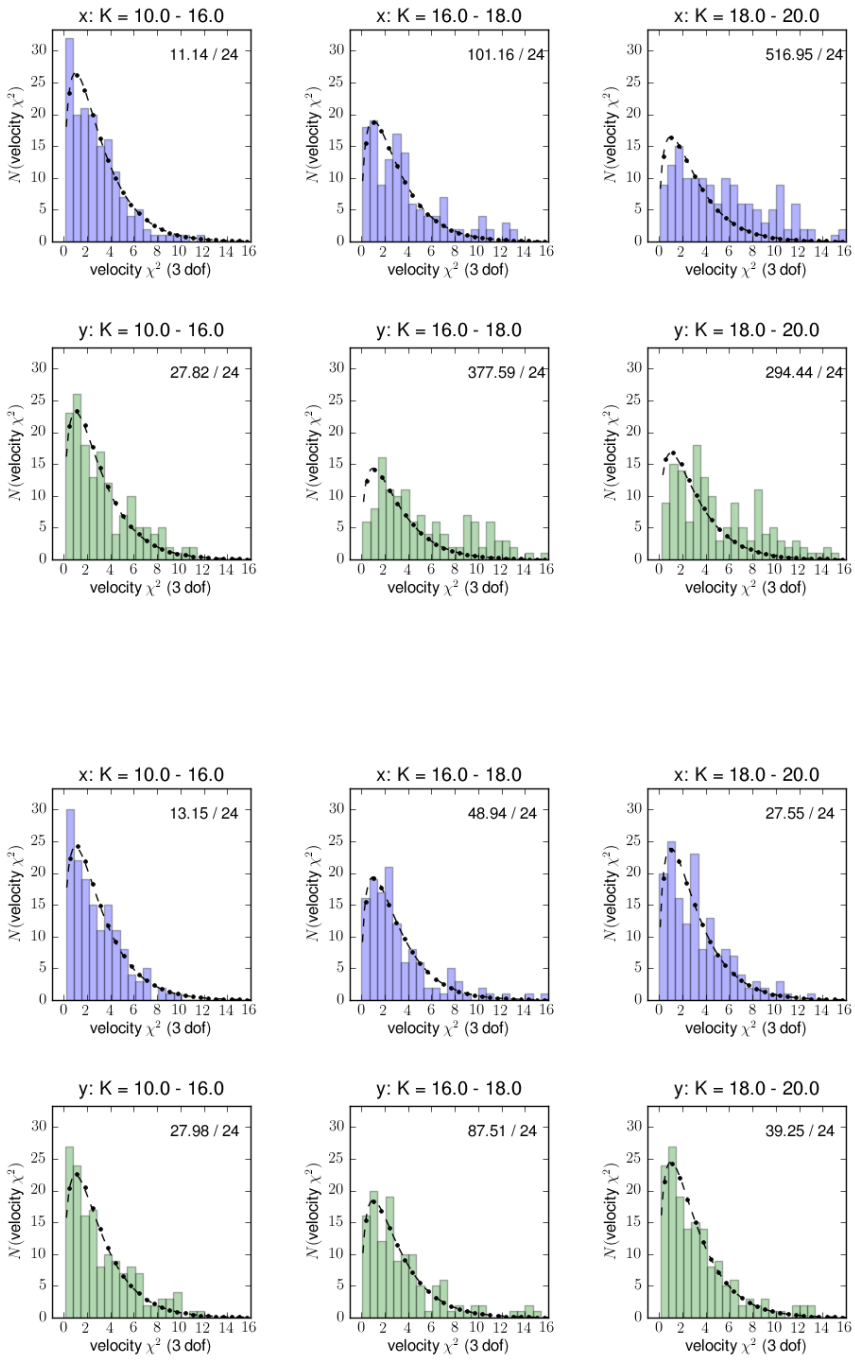

Figure 10. As Figure 7, for constant additive error (Top) and additive error computed from a fit to separate determinations for each magnitude-range (Figure 9)

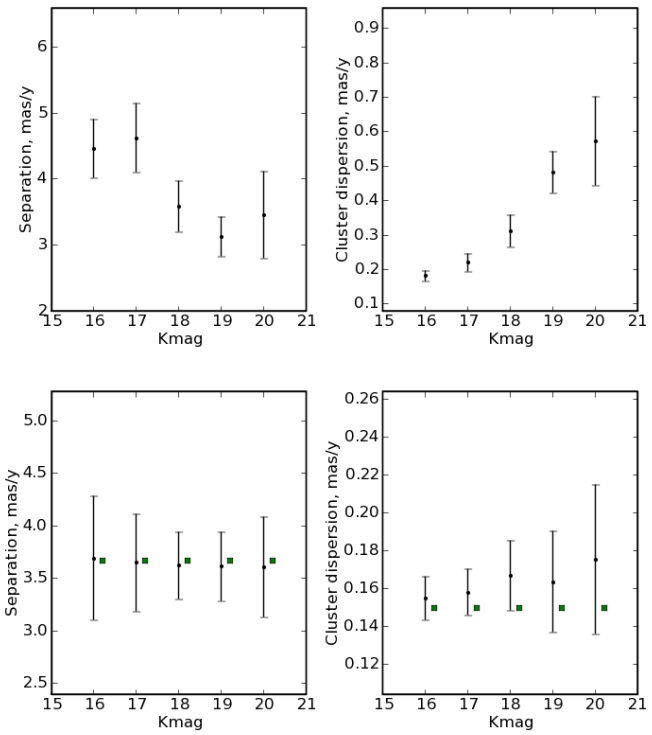

Figure 11. Cluster-to-Field separation and the fitted velocity dispersion. *Top row:* Fitted parameters with 1σ errorbars from Bootstrap Monte Carlo resampling of the motions. *Bottom row:* component separation and error-corrected velocity dispersion from simulated sets of motions, with a constant intrinsic dispersion and errors sampled from the observed dataset, and with constant input separation (input parameters denoted with green squares).

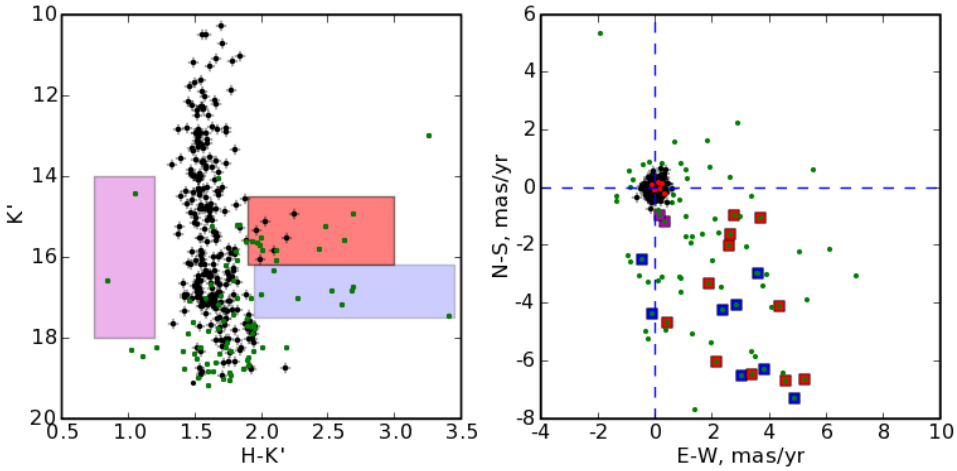

Figure 12. CMD of the field of view and VPD for all objects with proper motion error < 0.5 mas yr^{-1}. (The CMD is limited by the depth of the H-band imaging and thus spans a smaller magnitude range than the K' astrometric observations.) Objects with $P_{\rm cluster} > 0.9$ are shown in black. Likely-field objects falling in the shaded regions in the CMD are denoted with squares in the VPD. Red circles in the VPD indicate the six objects in the red shaded region in the CMD that show $\rm P_{cluster} > 0.9$.

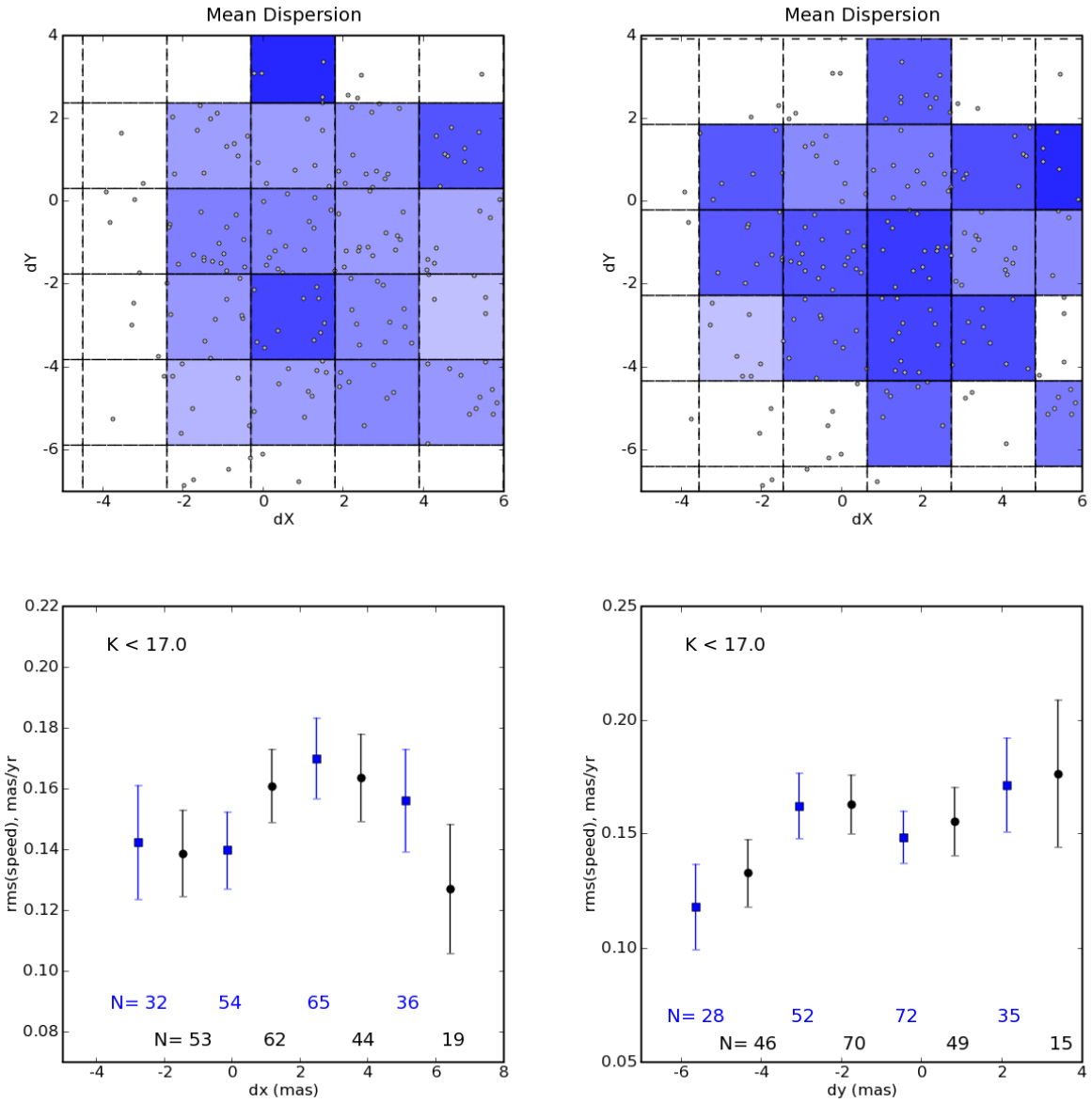

Figure 13. Search for the cluster center using dispersions. Top Row: mean dispersion $\sqrt{(\sigma_x^2 + \sigma_y^2)/2}$ from independent bins of cluster members at $K < 17.0$. Bins are color coded by mean dispersion. The right panel is similar to the left panel, but with the bin centers dithered with respect to the left panel. Bottom: mean dispersion for strips in X (left) and Y (right); here the black circles and blue squares indicate strip centers dithered by half a strip-width with respect to each other. Inset numbers give the sample size in each strip.

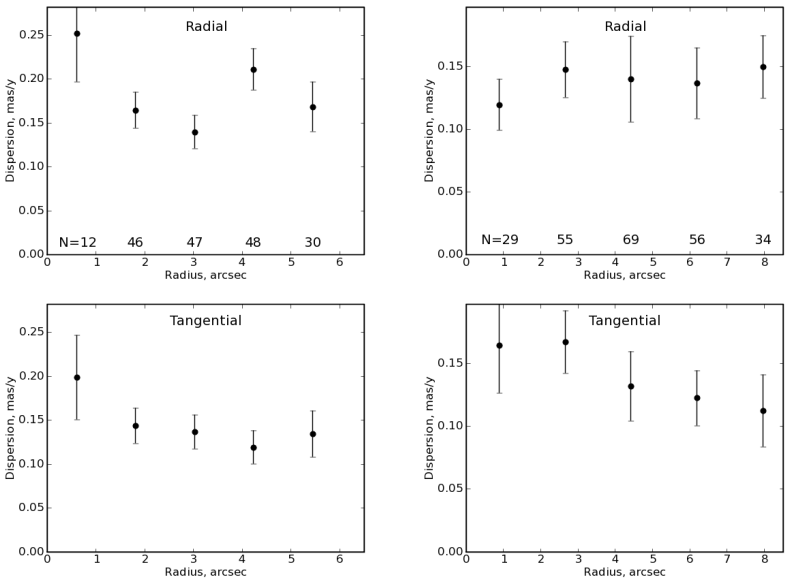

Figure 14. Radial (top) and tangential (bottom) velocity dispersions for two choices of cluster center, from cluster stars with $K < 17$. Left: cluster center coincident with the frame center. Right: cluster center offset from the field center in (X,Y) by (-0.5, 1.5) mas. Inset numbers give the number of cluster members in each radial bin.

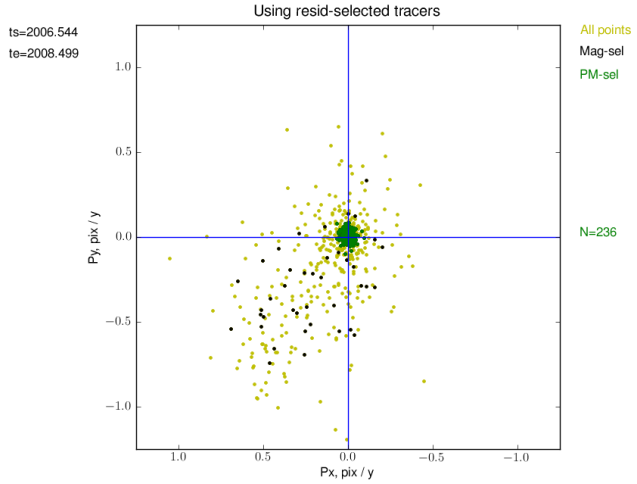

Figure 15. Culling of reference stars during the frame mapping process, in this case mapping epoch 2006.54 onto the frame of epoch t_0 =2008.50. *Yellow points:* all matches. *Black points:* selected for magnitude ($K' \leq 17.5$). *Green points:* reference stars selected by position residual from the center of mass of the magnitude-selected sample and with outlyers clipped during the mapping (in this plot the motions have been shifted to the center of mass of the selected objects).

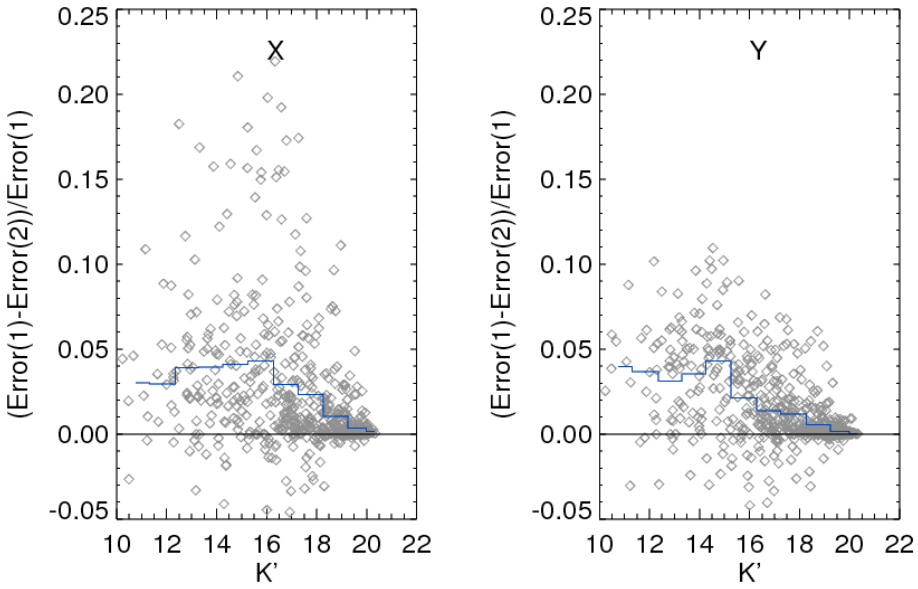

Figure 16. Reduction in velocity error when frames are mapped onto a reference list constructed from a first pass at proper motion fitting (2), over frame mapping onto a single starlist at epoch t_0 (1). The change in error is expressed as the ratio of the difference between (1) and (2) to the original error (1). Reading left-right,panels indicate errors along X and Y.

Parameter		2006.38	2006.54	2008.37	2008.50	2009.33
N_{ref}		238	239	235	241	233
$N_{<4}$		9	10	21	11	6
Δ	x'	-83.94 ± 6.03 × 10^{-3}	9.26 ± 4.01 × 10^{-3}	4.16 ± 3.71 × 10^{-3}	(-4.06 ± 17.716)×10^{-4}	6.53 ± 3.52 × 10^{-3}
(pix)	y'	-16.88 ± 4.74 × 10^{-3}	5.69 ± 4.39 × 10^{-3}	3.35 ± 5.64 × 10^{-3}	(3.66 ± 25.421)×10^{-4}	3.51 ± 3.66 × 10^{-3}
x	x'	1.00005 ± 2.11 × 10^{-5}	0.99994 ± 9.12 × 10^{-6}	1.00025 ± 1.34 × 10^{-5}	0.99999 ± 6.18 × 10^{-6}	1.00002 ± 8.24 × 10^{-6}
()	y'	(6.64 ± 0.185)×10^{-4}	(7.73 ± 0.079)×10^{-4}	(1.33 ± 1.425)×10^{-5}	(2.32 ± 5.858)×10^{-6}	(2.75 ± 0.090)×10^{-4}
y	x'	(-3.15 ± 0.177)×10^{-4}	(-3.53 ± 0.089)×10^{-4}	(1.23 ± 0.105)×10^{-4}	(-2.40 ± 4.766)×10^{-6}	(-2.54 ± 0.072)×10^{-4}
()	y'	1.00046 ± 1.97 × 10^{-5}	1.00028 ± 9.62 × 10^{-6}	1.00014 ± 1.34 × 10^{-5}	1.00000 ± 6.66 × 10^{-6}	1.00009 ± 7.99 × 10^{-6}
x^2	x'	28.90 ± 5.946	20.61 ± 3.463	-18.24 ± 3.989	-0.27 ± 1.986	4.63 ± 2.818
(×10^{-8} pix^{-1})	y'	11.78 ± 4.751	15.66 ± 3.266	23.10 ± 4.404	0.48 ± 1.937	20.65 ± 2.938
xy	x'	1.80 ± 5.818	20.39 ± 3.772	58.96 ± 4.594	-0.28 ± 2.040	-0.58 ± 2.973
(×10^{-8} pix^{-1})	y'	-5.05 ± 6.935	-10.19 ± 3.985	-28.61 ± 5.203	-1.52 ± 2.680	-20.03 ± 3.335
y^2	x'	-3.04 ± 5.917	5.55 ± 3.048	1.09 ± 3.540	1.71 ± 1.686	-0.59 ± 2.959
(×10^{-8} pix^{-1})	y'	33.22 ± 6.228	52.07 ± 3.664	75.11 ± 5.451	-1.13 ± 2.488	-0.47 ± 2.633
1.0-M (×10^{-4})		-2.55 ± 0.150	-1.14 ± 0.064	-1.95 ± 0.095	(1.97 ± 4.303)×10^{-2}	-0.58 ± 0.062
1.0-M_y/M_x (×10^{-4})		-4.08 ± 0.275	-3.34 ± 0.135	1.10 ± 0.189	(-5.72 ± 9.564)×10^{-2}	-0.71 ± 0.104
θ_{rot} ($''$)		-10.11 ± 0.250	-11.62 ± 0.129	1.13 ± 0.182	(-4.88 ± 7.925)×10^{-2}	-5.46 ± 0.116
θ_{skew} ($''$)		7.20 ± 0.553	8.66 ± 0.234	2.82 ± 0.364	-0.002 ± 0.153	0.43 ± 0.245

Table 2. Transformation parameters taking the starlist in each epoch into the reference frame t_{ref}. *Top two rows*: Number of reference stars N_{ref} used in the mapping, along with the number of reference stars $N_{<4}$ that appear in fewer than four epochs. *Next six rows:* Coefficients of the polynomial fits $x' = f(x, y)$ and $y' = g(x, y)$ (top and bottom rows respectively in each pair). For reference, a quadratic term of size 10.0×10^{-8} pix^{-1} would introduce displacement 0.25 mas at the edges of the detector, comparable to the centroiding error for bright ($K' < 16$) objects (Table 3). *Bottom four rows:* The linear parts of the transformations written in terms of global scaling M, nonuniform magnification M_y/M_x, rotation θ_{rot} and departure from perpendicular axes θ_{skew}. (Global shifts Δ appear in the polynomial fits and are not repeated.)

K'		Centroiding, Alignment (mas) 2006.39	2006.54	2008.37	2008.50	2009.33	Additive (mas)	Confusion bias (mas)	Motion (mas/y)
$10-16$	x	0.25, 0.09	0.10, 0.05	0.25, 0.06	0.06, 0.04	0.08, 0.03	0.16 ± 0.02	0.0 - 0.03	**0.076**
	y	0.23, 0.08	0.07, 0.05	0.19, 0.08	0.11, 0.04	0.09, 0.05	0.15 ± 0.02		**0.074**
$16-18$	x	0.41, 0.11	0.23, 0.05	0.40, 0.08	0.14, 0.05	0.14, 0.04	0.24 ± 0.02	0.03 - 0.10	**0.130**
	y	0.42, 0.10	0.20, 0.05	0.43, 0.09	0.17, 0.05	0.18, 0.05	0.30 ± 0.03		**0.153**
$18-20$	x	1.10, 0.11	0.92, 0.05	1.03, 0.08	0.60, 0.05	0.59, 0.04	0.59 ± 0.06	0.1- 1.0	**0.378**
	y	1.38, 0.10	1.05, 0.05	1.35, 0.09	0.81, 0.05	0.77, 0.06	0.71 ± 0.08		**0.478**

Table 3. Astrometric error budget. For each magnitude bin, the top (bottom) row gives errors in X (Y). For each star, centroiding, alignment and additive error describe random variation between epochs. The effect of confusion bias on motions depends on its variation between epochs; random variation is already included in the additive error, while linear trends masquerading as spurious motions are expected to be $\lesssim 10\%$ of the confusion bias across the epochs for all objects (Section 3.6).

K	14.0-16.0	15.0-17.0	16.0-18.0	17.0-19.0
N	87	107	142	167
π_{cl}	0.81 ± 0.044	0.72 ± 0.043	0.72 ± 0.038	0.59 ± 0.044
$\Delta\mu$ (mas yr^{-1})	3.95 ± 0.592	4.58 ± 0.495	4.96 ± 0.523	3.63 ± 0.449
$\sigma_{a,f}$ (mas yr^{-1})	2.32 ± 0.294	2.52 ± 0.236	2.92 ± 0.278	2.89 ± 0.275
$\sigma_{b,f}$ (mas yr^{-1})	1.50 ± 0.259	1.66 ± 0.243	1.81 ± 0.271	1.80 ± 0.158
$\sigma_{a,cl}$ (mas yr^{-1})	0.16 ± 0.022	0.17 ± 0.040	0.17 ± 0.031	0.23 ± 0.073
$\sigma_{b,cl}$ (mas yr^{-1})	0.14 ± 0.014	0.15 ± 0.014	0.16 ± 0.021	0.16 ± 0.051
θ_f (^{o})	-58.87 ± 14.766	-61.34 ± 12.702	-54.02 ± 10.269	-61.97 ± 6.442
$\sigma_{b,cl}/\sigma_{a,cl}$	0.83 ± 0.106	0.89 ± 0.095	0.98 ± 0.121	0.69 ± 0.185
$\sigma_{b,f}/\sigma_{a,f}$	0.65 ± 0.109	0.66 ± 0.109	0.62 ± 0.113	0.62 ± 0.085

Table 4. Fitted kinematic parameters of cluster and field. For each magnitude range, rows give: the number of stars, the cluster fraction, the separation between cluster and field centers in the VPD, the semimajor and minor axes of the field, the semimajor and minor axes of the cluster, the orientation of the field component, and the semiminor/major axes of the cluster and field. Errors are estimated from Monte Carlo simulations.

K	N	σ_x (mas yr^{-1})	σ_y (mas yr^{-1})	σ_x (km s^{-1})	σ_y (km s^{-1})
10.0-14.0	67	0.130 ± 0.017	0.123 ± 0.016	4.912 ± 0.639	4.680 ± 0.593
14.0-16.0	72	0.161 ± 0.019	0.129 ± 0.016	6.088 ± 0.739	4.878 ± 0.606
16.0-18.0	107	0.177 ± 0.027	0.180 ± 0.030	6.721 ± 1.034	6.839 ± 1.142
18.0-20.0	97	0.224 ± 0.039	0.148 ± 0.046	8.508 ± 1.498	5.629 ± 1.753

Table 5. Arches velocity dispersion in each co-ordinate. Reading left-right, columns are: Magnitude range of interest, number of cluster stars in this magnitude range, intrinsic velocity dispersion and error in each coordinate, first in mas yr^{-1} and then km s^{-1} assuming the Arches is at 8 kpc.

Appendix A. Production of motions from star-lists

Here we provide details of the procedures used to produce proper motions from the star-lists. The steps are:

1. Choice of initial reference stars, and initial mapping onto t_0 : An initial list was constructed of eleven bright stars that were well-measured in all epochs and cover the full area of the detector, and used for the initial registration of each epoch onto $t_0 =$ 2008.5. This epoch was chosen because its starlist is of high quality (Tables 1 & 3) and this time will be near the pivot point of the straight line fits to most of the positional time-series. A 6-term linear transformation was used for this mapping for each epoch, accounting for positional shift, global scaling, rotation, a difference in scale factors in X and Y ("on-axis skew") and differences in the angle between axes ("off-axis skew").

2. Matching of all stars within t_0 : Using this initial mapping, all stars were matched to their counterparts in t_0 (where present) by proximity in t_0 and magnitude. Matching radius 5 pix (approximately the PSF core FWHM) and a broad magnitude tolerance 3 mag were used. This yields positional differences between predicted and observed positions (hereafter "deltas") in t_0 for matched pairs using the first-guess transformation.

3. Fitting of reference-frame mapping for matching: The previous step typically produces deltas for ~ 300 stars at $K' < 17.5$ across each pair of epochs (t-t_0). The field population displays significant motion dispersion in a preferential direction close to the galactic plane. Field objects must therefore be removed from the sample of reference stars to avoid biasing the offsets and magnification factors when mapping the reference frames. From the positional deltas of reference stars, the center of the distribution in the vector point diagram is estimated and the standard deviation of motions in each direction from this center of mass estimated. Objects farther than 2σ from this estimated center of mass are removed. This process is repeated twice to produce a cleaned list of reference stars; typically 260 objects survive this process. These objects are used to re-map the epochs onto t_0 using a full 6-term linear transformation. Clipping of outliers in this epoch mapping typically removes a further 30 stars in each epoch (Figure 15).

4: Trim coincident close pairs of stars: At this stage we have the master-list of measurements of each object, in the frame in which the object was originally measured. To mitigate confusion by known objects as much as possible, all coincident pairs with separation < 75 mas are removed from consideration for each epoch. This typically removes 20-40 objects from the position-lists at each epoch (column $N_{*,kept}$ in Table 1). The result is a matched catalogue of 1114 objects present in at least two epoochs.

5. Reference frame-mapping for motions: Armed with the matched list of objects and their measurements at each epoch, likely-cluster members (at $K < 17.5$) are used as reference stars to map each epoch onto t_0 , using the same weighting and clipping as step 3. We find (Section 3.4) that a second-order transformation in X and Y is sufficient to capture most of the residual higher-order effective distortions between epochs without falling prey to overfitting of few stars with a high-order transformation.

6. Motion extraction in t_0 : A first pass at stellar motions is estimated by fitting a linear trend to the positional time-series $x(t), y(t)$ of each star in the reference frame t_0 . For each star, the weighted mean time $\bar{t} = \sum_i t_i w_i / \sum_i w_i$ is evaluated so that the fit becomes $x(t) = a + b(t - \bar{t})$. Weights $w_i = 1/\sigma_i^2$ are the inverse of the variance of each measurement due to positional uncertainty. This removes correlation between errors in the parameters (e.g. Press et al. 1992); the center of mass of the data is first determined then the slope pivoted about this point to find the best-fit positional gradient. (The covariance matrix of the fit for each star is reported in Table X). The proper motion error is the formal error on the best-fit slope: $\sigma_b^2 = 1/\left(\sum_{i=1}^{N} \frac{(t-\bar{t})^2}{\sigma_i^2}\right)$. To mitigate sensitivity to short-term excursions in position, for objects measured in ≥ 4 epochs, two passes of sigma-clipping at 3σ are applied. Note that $\bar{t}$ is a property of $\sigma_i(t)$ and thus is estimated separately for each star and for each co-ordinate. Motions are estimated for all 805 stars with ≥ 3 epochs of measurement.

7. Refinement of the reference position-list: When choosing a reference frame in which to evaluate motions, our goal is a reference-list onto which cluster members can be mapped with as little scatter as possible due to measurement and fitting error. We generate a reference frame by evaluating at some time t_{ref} the straight-line fits to the positional time-series of cluster reference stars. Positional errors in this predicted frame ("predictive errors") are evaluated by propagating the errors on the fit coefficients a, b for each star. By choosing t_{ref} to be near the pivot point $\bar{t}$ of the greatest number of reference stars, we aim to minimize the error of the predicted positions in the constructed reference frame. The distribution of $\bar{t}$ is nearly Gaussian with $\bar{t} = 2008.0 \pm 0.4$ (1σ); we therefore adopt $t_{ref} = 2008.0$ to evaluate the reference epoch. To evaluate the degree to which this mean reference frame improves the mapping, motions and their errors for each star were evaluated using the quad sum (centroiding + alignment) errors when mapped to t_0 and (centroiding + alignment + predictive) when mapped to t_{ref} . Motion errors are improved by up to 20% for some bright objects, with median improvement up to 4% for well-measured objects (Figure 16).

8. Re-mapping and re-extraction of motions: Finally, the star-list from each epoch is mapped onto the constructed reference-frame t_{ref} and motions in this frame evaluated in the manner of Step 7. Table 2 gives the fitted parameters and the number of reference stars used in the mapping from each epoch to the t_{ref} frame. Provided the motions of cluster stars in the field of view do not themselves describe a second-order or lower transformation (e.g. rotation or contraction of the cluster) to within our ability to determine, then the parameters taking reference frame 2008.50 to $\bar{t}$ should be consistent with zero, as is observed (Table 2). We find that, when applied to stars near the edges of the detector, the size of the positional shifts due to the quadratic terms in the mapping are 1-few times the centroiding error for bright ($K' < 16$) objects (Table 2).

9. Evaluation and incorporation of additional error sources: The distribution of fits to the velocities thus produced were examined for additional sources of random error. It became readily apparent that a significant source of error along the time-series was not taken into account by the steps above. When characterised (Section 3.5), steps 6-9

were repeated with this error term included.

References

Acknowledgments

Support for this work was provided by NSF grant AST 04-06816 and the NSF Science and Technology Center for Adaptive Optics, managed by the University of California, Santa Cruz (AST 98-76783), and the Levine- Leichtman Family Foundation. The W. M. Keck Observatory is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.