

ASTROMETRY WITH THE WFIRST WIDE-FIELD IMAGER

THE WFIRST ASTROMETRY WORKING GROUP

ROBYN E. SANDERSON,¹ ANDREA BELLINI,² STEFANO CASERTANO,² JESSICA R. LU,³ PETER MELCHIOR,⁴ DAVID BENNETT,⁵
MICHAEL SHAO,⁶ JASON RHODES,⁶ SANGEETA MALHOTRA,⁵ SCOTT GAUDI,⁷ MICHAEL FALL,² ED NELAN,²
RAJA GUHATHAKURTA,⁸ JAY ANDERSON,² SHIRLEY HO,⁹ AND MATTIA LIBRALATO²

¹*TAPIR, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA*

²*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*

³*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*

⁴*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*

⁵*Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*

⁶*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*

⁷*Department of Astronomy, Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA*

⁸*University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95060, USA*

⁹*Departments of Physics and Astronomy, University of California, Berkeley, CA 94720, USA*

(Received XXX; Revised XXX; Accepted XXX)

ABSTRACT

The Wide-Field InfraRed Space Telescope (WFIRST) will be capable of delivering precise astrometry for faint sources over the enormous field of view of its main camera, the Wide-Field Imager (WFI). This unprecedented combination will be transformative for the many scientific questions that require precise positions, distances, and velocities of stars. We describe the expectations for the astrometric precision of the WFIRST WFI in different scenarios, illustrate how a broad range of science cases will see significant advances with such data, and identify aspects of WFIRST’s design where small adjustments could greatly improve its power as an astrometric instrument.

Keywords: Astrometry, space vehicles: instruments

1. INTRODUCTION

The wide field of view and stable, sharp images delivered by the Wide-Field Imager (WFI) planned for the Wide-Field InfraRed Space Telescope (WFIRST) make it an excellent instrument for astrometry, one of five major discovery areas identified in the 2010 Decadal Survey. WFIRST has two main advantages over other spacecraft missions: it can precisely measure very faint stars (complementary to *Gaia*); and it has a very wide field of view (complementary to the *Hubble Space Telescope*, *HST*, and the James Webb Space Telescope, *JWST*). Compared to *HST*, WFIRST’s wider field of view with similar image quality will provide many more astrometric targets in a single image, but also hundreds more anchors to the astrometric reference frame in any field, in-

cluding both background galaxies and stars with precise positions in the *Gaia* catalog ([Gaia Collaboration et al. 2016a,b](#)). In addition, WFIRST will operate in the infrared, a wavelength regime where the most precise relative astrometry has so far been achieved with adaptive optics images from large ground-based telescopes (e.g. $150\ \mu\text{as}$ from Keck, [Ghez et al. 2008](#)). WFIRST will provide at least a factor of three improvement in astrometry over the current state of the art in this wavelength range, while spanning a field of view thousands of times larger. WFIRST is thus poised to make major contributions to multiple science topics in which astrometry plays an important role. In most cases, these contributions can be achieved without major alterations to the planned mission or instrument. In this paper, we summarize a few of the many compelling science cases where WFIRST astrometry could prove transformational, and then outline the areas where a small investment of attention now will ensure that WFIRST’s impact on this science is significant.

Context	Estimated performance	§
Single-exposure precision	0.01 px; 1.1 mas	1.1
Typical guest-observer program (100 exposures of one field)	0.1 mas	1.1
Absolute astrometry accuracy	0.1 mas	3.3
Relative proper motions derived from High-Latitude Survey	$25 \mu\text{as yr}^{-1}$	4.1
Relative astrometry, Exoplanet MicroLensing Survey (per image)	1 mas	4.2
Relative astrometry, Exoplanet MicroLensing Survey (full survey)	3–10 μas	4.2
Spatial scanning, single scan	10 μas	2.4
Spatial scanning, multiple exposures	1 μas	2.4
Centering on diffraction spikes	10 μas	2.4

Table 1. Approximate expected astrometric performance of the WFIRST Wide-Field Imager for different types of observations. All estimates are for well-exposed point sources (refer to [Spergel et al. 2013](#) and [Table 3](#) for depths of the core survey programs). See the referred sections for details about the assumptions leading to each number. These estimates are order-of-magnitude only.

1.1. Expected astrometric performance of the WFIRST Wide-Field Imager

The astrometric performance of the Wide-Field Channel (WFC) in the WFI on WFIRST will depend on numerous elements, including hardware characteristics, the stability of the platform, characterization of the optics and of the detector (down to the individual pixel), the ability to design observations with the required properties for reference stars, and calibration of both the point-spread function (PSF) and of the geometric distortion of the focal plane. Here we summarize the assumptions used in this work; [Table 1](#) collects some performance estimates for different types of observations.

[Gould et al. \(2015\)](#) discuss astrometric science in the WFIRST Exoplanet MicroLensing (EML) survey field, but make very optimistic sensitivity estimates, and potential sources of systematic errors are ignored. For the purpose of this document, **we assume that the single-exposure precision for well-exposed point sources is 0.01 pixels, or about 1.1 mas**; this is consistent with current experience on space-based platforms such as *HST*, as long as a comparable level of calibration activities are carried out ([Hosek et al. 2015](#)). Depending on the platform stability and the quality and frequency of calibrations, this accuracy can be substantially improved by repeated observations as

$$\sigma_{\alpha,\delta} \propto \frac{\Delta\eta}{\sqrt{N}} \quad (1)$$

where Δ is the single-exposure point-source precision in pixels, η is the plate scale, and N is the number of observations. We assume that a gain by a factor ~ 10 (to 0.1 mas; i.e. 100 exposures) can be achieved for a typical Guest-Observer (GO) program. The EML survey, which will obtain many thousands of observations of each source, has more stringent requirements (§ 4.2). For proper motion (PM) measurements, the achievable accuracy depends on the single-image precision and the time baseline T between the first and last im-

ages. In the case of N evenly spaced images:

$$\sigma_{\mu} \propto \frac{\Delta\eta}{T\sqrt{N}}. \quad (2)$$

Improvements in astrometric measurements can also be obtained by special techniques, such as spatial scanning and centering on diffraction spikes, described in § 2.4, and by improvements in the pixel-level calibration, as discussed in § 5.2.

Thanks to its large field of view (FoV) and the availability of accurate reference sources from *Gaia*, each WFC exposure can achieve an *absolute* astrometric accuracy of 0.1 mas or better (see § 3.3 for details).

2. SCIENCE WITH WFIRST ASTROMETRY

The science enabled by astrometry with WFIRST spans size scales from the Local Group to exoplanetary systems, and provides important contributions to all three astrophysics goals in the NASA Science Plan. In this section, we survey the range of science topics to which this instrument can make important contributions. The astrometric precision needed for each of the following science cases is listed in [Table 2](#).

2.1. Motions of Local Group galaxies

The range and reach of WFIRST astrometry complement and extend *Gaia* and LSST astrometry. [Figure 1](#) compares the reach of current and planned PM surveys to the PMs corresponding to known velocities and distances of Local Group objects. Since the known orbital and internal velocities refer in almost every case to the radial component, these are intended only to represent the order of magnitude one might expect for the PMs (indeed, as in the case of M31, the orbital PMs may be significantly smaller than that inferred by radial velocity measurements alone). From this figure it is clear that to measure PMs of satellites beyond the Milky Way’s virial

§	Science case	Astrometric precision	
2.1	Motions of dwarf satellites in the Local Group	2.2×10^{-4} pixel yr ⁻¹	25 μ as yr ⁻¹
2.2	Motion of stars in the distant MW stellar halo	$\leq 2 \times 10^{-4}$ pixel yr ⁻¹	$\leq 25 \mu$ as yr ⁻¹
2.3	Low-mass end of the subhalo mass function	1.8×10^{-4} pixel yr ⁻¹	20 μ as yr ⁻¹
2.4	Detection & characterization of exoplanets	$\leq 9 \times 10^{-5}$ pixel	$\leq 10 \mu$ as
2.5	Structure of the MW bulge	$\leq 9 \times 10^{-5}$ pixel	$\leq 10 \mu$ as
2.6	Star formation in the MW	$\leq 4.5 \times 10^{-4}$ pixel yr ⁻¹	$\leq 50 \mu$ as yr ⁻¹
2.7	Isolated black holes & neutron stars	4.5×10^{-4} pixel	50 μ as
2.8	Internal kinematics in GCs	$\lesssim 1.8 \times 10^{-4}$ pixel yr ⁻¹	$\lesssim 20 \mu$ as yr ⁻¹

Table 2. Required astrometric precision (in units of both WFI pixels and μ as) for the different science cases discussed in §2.

radius will require better precision than LSST can achieve, at larger distances (and thus fainter magnitudes). This is the window of opportunity for WFIRST.

WFIRST astrometry is a crucial component of constraints on the nature of dark matter from the orbital and internal PMs of the satellite galaxies of the Milky Way (MW). The orbits of dwarf satellites can be used to help map the MW’s own dark matter halo out to its virial radius and beyond, placing our galaxy into a cosmological context. A complete knowledge of the MW halo’s properties enables tests of dark matter models through comparisons with predictions from simulations for its mass and shape, accretion history, and the mass and orbit distributions of its satellite galaxies. *HST* PMs of dwarf satellite galaxies, and *Gaia* astrometry in the inner Galaxy, will both make great strides towards this goal; however, *Gaia* has insufficient depth, and *HST* insufficient FoV, to reach to the edge of the MW halo (where the total mass is uncertain to a factor of 4; see [Eadie & Harris 2016](#)) or to obtain internal PM measurements for many dwarf galaxies (which are crucial to break velocity degeneracies and understand the small-scale distribution of dark matter¹).

The HSTPROMO campaign has used the *HST*, which has similar image quality to that expected for WFIRST, to measure PMs of both bound objects and stream stars in the MW (e.g. [Sohn et al. 2013, 2016](#)). In the coming years *HST* will set a PM baseline for many more distant satellites ([Kallivayalil et al. 2015](#)): by the time WFIRST is ready, this baseline will be roughly 8–10 years for these satellites, and far longer for dwarf galaxies with earlier observations. With its larger FoV and more sensitive detectors, WFIRST should be able to expand these measurements to more distant galaxies and achieve better accuracy thanks to the larger number of calibration objects available (and to the establishment of *Gaia*’s astrometric frame, see § 3). Figure 2 shows the estimated number of stars in each of the MW’s satellites that are accessible to WFIRST at the depth of the planned

high-latitude survey (HLS, $J < 26.7$, blue) and for a possible spatial-scanning mode on WFIRST ($H < 16$, red). The upper panel of the figure shows the expected tangential velocity error for each dwarf assuming PM precisions of 25 μ as yr⁻¹. Internal PM dispersions, as was recently done for the Large Magellanic Cloud with *HST* ([van der Marel & Kallivayalil 2014](#)), are reachable with WFIRST for the galaxies shown in cyan. These include three ultra-faint galaxies (Segue I, Draco, and Ursa Minor) that are sufficiently DM-dominated to distinguish between cold dark matter models (which predict a cuspy inner mass profile) and warm or “fuzzy” DM models (which predict cored inner profiles). Draco and Ursa Minor each have ~ 80 stars at $H < 16$ and their internal velocity dispersions are marginally resolved even at 25 μ as yr⁻¹, making them particularly good cases for spatial scanning to improve the internal PM accuracies by an order of magnitude.

2.2. Motions of stars in the distant MW halo

Besides dwarf galaxies, the MW’s halo contains the tidally disrupted remains of previously accreted galaxies, known as tidal streams. We expect that tidal debris should extend to at least the virial radius of the MW ([Sanderson et al. 2017](#)), but currently the most distant MW halo star known is an M giant at around 250 kpc ([Bochanski et al. 2014a](#)). The most distant known populations with statistical samples of stars (BHB, RR Lyr, and M giant stars) extend to around 150 kpc (half the virial radius) at the magnitude limit of current surveys ([Deason et al. 2012](#); [Bochanski et al. 2014b](#); [Sesar et al. 2017](#)), but the WFIRST HLS fiducial depth will reach to the MW’s virial radius down to the main-sequence turnoff. The orbits of distant stars probe the extent and total mass of the MW dark halo; they also represent a unique population of recently accreted small galaxies. With the HLS’s projected depth, the transition between the MW’s and Andromeda’s spheres of influence, and perhaps the splashback radius of the MW (e.g. [Diemer & Kravtsov 2014](#)), could also be detected. Proper motions from WFIRST are crucial to these endeavors, since complete phase-space information for these stars is the best way to confirm that stars associated in position at large dis-

¹ See work by the Gaia Challenge group at astrowiki.ph.surrey.ac.uk/dokuwiki/doku.php?id=tests:sphtri, summarized in Fig 2.2 of [The Theia Collaboration et al. \(2017\)](#).

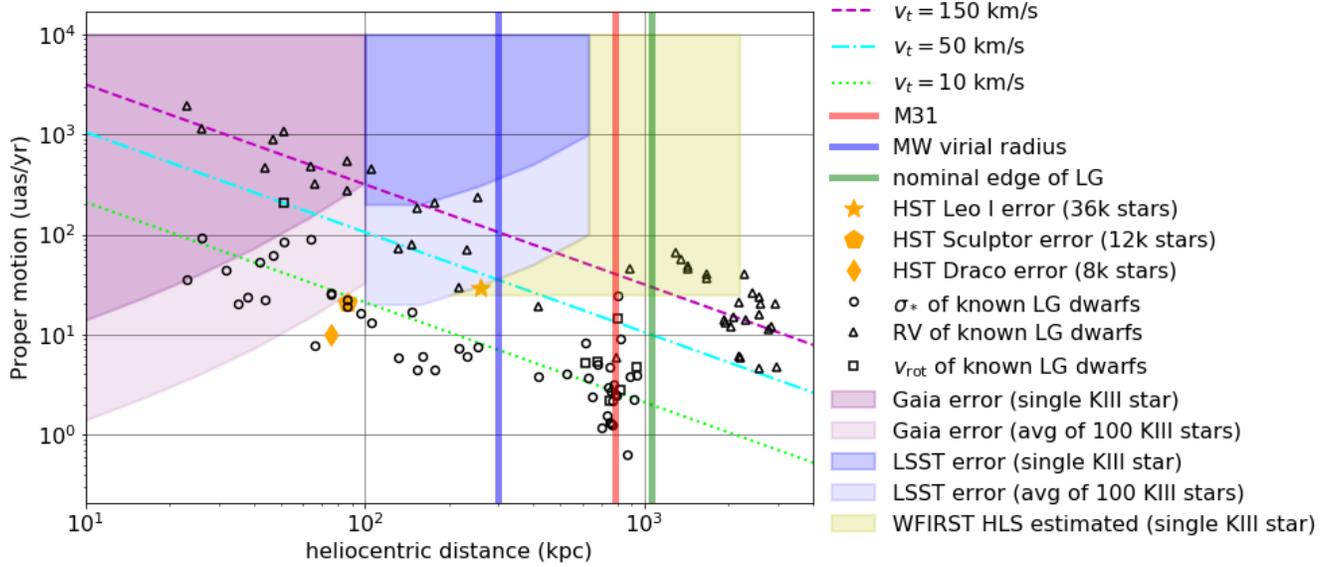


Figure 1. Proper-motion errors of various current and planned surveys and measurements, compared to the PMs corresponding to characteristic velocities and distances for Local Group objects. Shaded regions show the distances and PMs accessible to the *Gaia* and LSST surveys for single stars (darker) and naively averaging 100 stars (lighter); the orange star, diamond and pentagon are the quoted uncertainties on recent *HST* measurements of the PMs of Leo I, Draco and Sculptor respectively (Sohn et al. 2013, 2017). The approximate reach of the WFIRST HLS field for bright K giants (assuming 15 exposures over 5 years) is shown for comparison in yellow. Solid vertical lines mark a few important distances: the edge of the MW halo (blue), the distance to M31 (red), and the approximate edge of the LG (green). Small open symbols mark the PMs corresponding to the heliocentric orbital radial velocities (triangles), and internal velocity dispersions (circles) and rotational velocities (squares) of Local Group dwarf galaxies from McConnachie (2012). The diagonal lines show the PM associated with several characteristic transverse velocities as a function of distance: the Galactic circular velocity (magenta), the typical rotational velocity of a dIrr (cyan) and the typical internal velocity dispersion of a dSph (green).

tances are from the same progenitor and to connect groups on opposite sides of the galaxy through their orbits, leading to constraints on the mass profile and flattening of the Galactic dark halo at large distances (e.g. Belokurov et al. 2014).

High-velocity stars are another interesting target, whether for GO observations or as serendipitous objects in repeated survey fields. These stars’ orbits, which have an extremely wide radial range, can potentially also be used to constrain the overall shape and mass of the MW dark matter halo (Brown 2015).

At distances of 100–300 kpc, preliminary work shows that PM precision of $25 \mu\text{as yr}^{-1}$ or better is required to eliminate outliers in groups and connect structures on opposite sides of the galaxy using their phase-space positions (Figure 3; Secunda et al. in prep). A similar precision would be needed to identify high-velocity stars at or near the Galactic escape velocity.

2.3. Constraining the low-mass end of the subhalo mass function

Astrometry from pointed GO projects could provide several routes to understanding the distribution of low-mass substructure in the MW. The abundance (or lack) of low-mass structure is a key to differentiating between cold and warm dark matter (DM) models, since cold DM predicts abundant substructure at small scales while in warm DM the mass

function is cut off at a characteristic scale related to the intrinsic temperature of the DM particle (and hence to its mass in the case of a thermal relic).

One route is to search for perturbations to tidal tails from distant globular clusters or dwarf galaxies. Current searches are focused on quantifying substructure in the spatial distribution of tidal debris but are limited by knowledge of the MW background/foreground in the region of the stream as well as by Poisson fluctuations in the star counts (e.g. Ibata et al. 2016) and by the uncertain dynamical ages of streams, which depend on modeling the orbit (e.g. Bovy et al. 2017). WFIRST’s capability to reach deep into the stellar main sequence at these distances will help mitigate the shot-noise issue. More importantly, obtaining astrometry of fields including tidal streams would allow superior selection of stream stars relative to the background/foreground, improving the sensitivity to density fluctuations and allowing better constraints on the time when material was first tidally stripped. Streams commonly stretch tens of degrees over the sky, so WFIRST’s large field of view is uniquely well suited to this application. A thin stream usually has a velocity dispersion of $1\text{--}10 \text{ km s}^{-1}$, so to provide a useful PM selection for a stream at 50 kpc would require relative PMs to a precision of about $20 \mu\text{as yr}^{-1}$.

Another possibility is to search for deviations in the apparent positions of quasars due to strong lensing by dark sub-

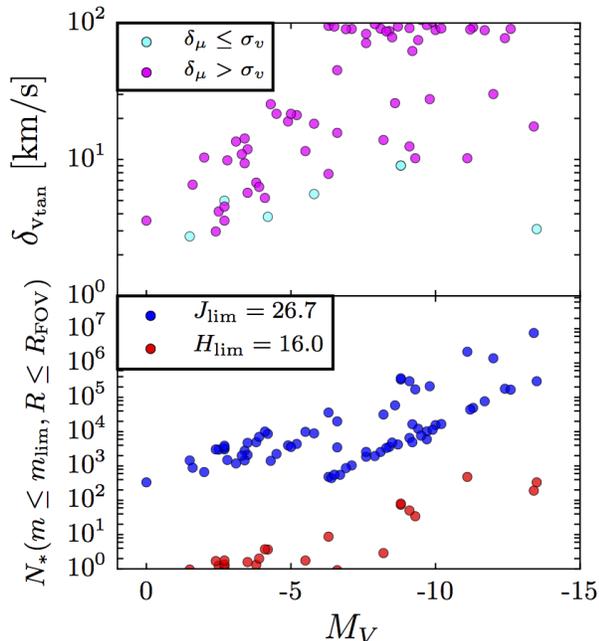


Figure 2. Top: estimated observational error in tangential velocity assuming PM precision of $25 \mu\text{as yr}^{-1}$. Galaxies in cyan have estimated velocity errors (δ_μ) comparable to or less than their intrinsic velocity dispersions (σ_v). **Bottom:** number of stars in Local Group dwarf satellite galaxies brighter than the limiting apparent magnitude of the High-Latitude Survey ($J < 26.7$, blue) and the limit for spatial scanning ($H < 16$, red). Plot courtesy Matthew Walker.

structures. For a distant quasar lensed by a $10^8 M_\odot$ subhalo at 50 kpc, the Einstein radius of the lens is roughly $20 \mu\text{as}$ (presuming a singular isothermal sphere). One route would be to look for the time-dependence of the lensing around a single quasar: for a subhalo moving at 200 km s^{-1} , the time to cross the Einstein radius is about 10 days. Alternatively, a wide field containing many quasars could be examined for statistical deviations between exposures taken at different times (separated by longer than 10 days). Either approach would require absolute astrometry (i.e. consistent between exposures) accurate to $20 \mu\text{as}$.

2.4. Detection and characterization of exoplanets

Very accurate PM estimates will make it feasible to search for the astrometric signature of exoplanets around nearby stars. For competitive constraints on exoplanet masses and orbital parameters, an instantaneous precision of better than $10 \mu\text{as}$ is required. The best constraints can be achieved for the most nearby stars ($d \lesssim 10 \text{ pc}$). A dedicated GO program that specifically observes those most promising targets with a flexible schedule is therefore complementary to the EML.

Because of their close proximity, the target stars are generally very bright. This makes high-precision astrometry pos-

sible by using one of two different methods: spatial scanning and diffraction spike modeling.

2.4.1. Spatial scanning

Spatial scanning involves intentionally slewing the spacecraft during integration to create extended tracks from bright target and reference stars in the field of interest. This spreads out the signal from each star over hundreds or thousands of pixels, thereby avoiding saturation while integrating orders of magnitude more photons, and averaging over pixel-level artifacts that may significantly affect pointed observations (Riess et al. 2014; Casertano et al. 2016). *HST* has attained precisions of $20\text{--}40 \mu\text{as}$ with this technique, limited in part by the small number of available reference stars and the variation of the focal plane geometry on the orbital time scale of the telescope ($\sim 1 \text{ hour}$). Because of its larger FoV and more stable orbit, we expect that the WFIRST WFC will be able to achieve precisions closer to the noise limit, about $10 \mu\text{as}$ per exposure. By combining multiple exposures it will then be possible to achieve a final relative astrometric accuracy of $\sim 1 \mu\text{as}$.

Desirable slew rates for spatial scanning are $0.5\text{--}10'' \text{ s}^{-1}$, roughly corresponding to $12\text{--}250$ pixels per read; this is the length of the region over which the light from each star will be spread within one readout frame. The fast, non-destructive reads of the WFIRST WFC will allow a clean separation of the signal accumulated within each pixel from different sources at different times, greatly reducing the confusion due to overlapping trails that has affected applications of this technique using WFC3/UVIS. Within the desired range of scanning speeds, it will be possible to observe unsaturated sources 7 mag brighter than the pointed-observation saturation limit, or $H_{\text{AB}} \sim 4 \text{ mag}$. The fastest available scan speed affects primarily the maximum brightness of the source that can be accommodated; slower scan speeds in the range $0.5\text{--}2'' \text{ s}^{-1}$ can achieve essentially the same benefits, but with a fainter saturation limit.

Both confusion effects and the signal-to-noise ratio for spatial scanning observations would benefit more than pointed observations from obtaining all independent reads for each exposure: unlike pointed observations, signal does not build up linearly over time in each pixel, but is deposited there during the narrow time interval in which a star passes over that pixel. Extending the interval between available reads increases both the background accumulated in each pixel without a corresponding increase in the signal, and the time interval over which signal from different stars in the same pixel cannot be cleanly separated. The availability of intermediate reads for download is of course subject to mission-level limits on science telemetry, so the number of reads to download may need to be determined on a scene- and project-dependent basis. Finally, spatial scanning observations will most likely need to be obtained under gyro control, as the required motion of the spacecraft will quickly exceed the size of the guiding window. More details, including error budgets, will be included in an upcoming white paper (Casertano et al., in preparation).

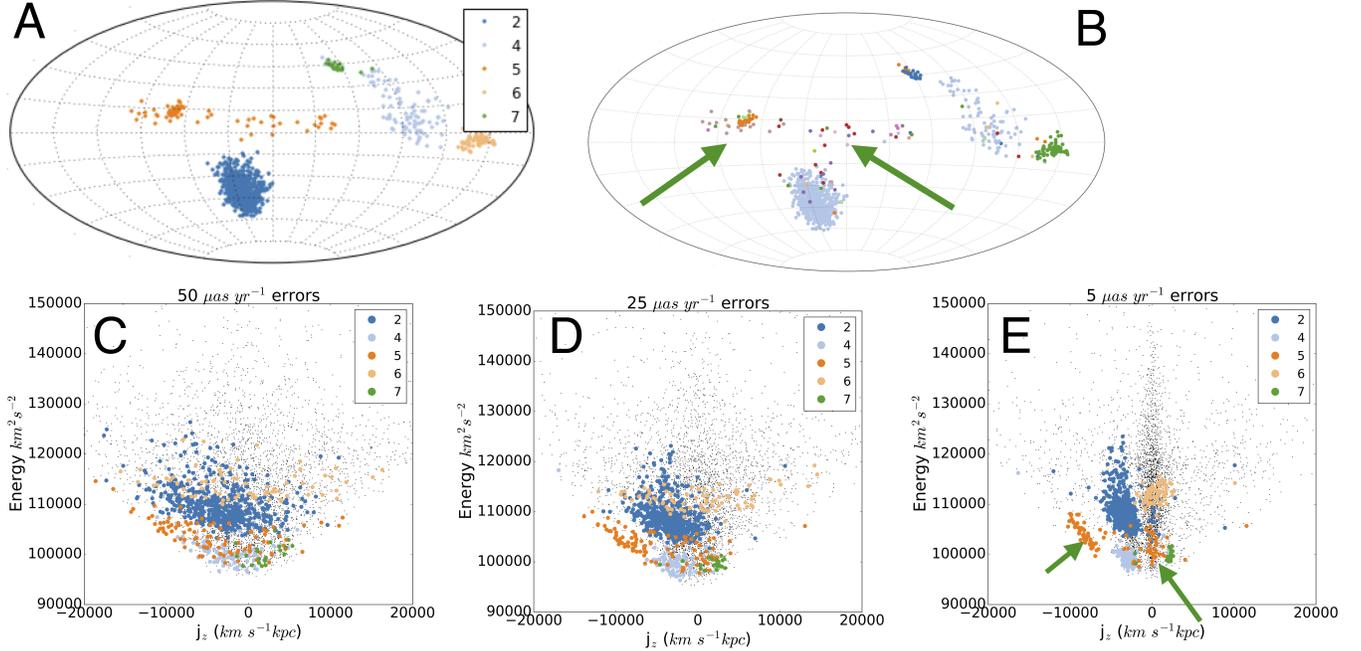


Figure 3. Identifying stellar structures in the distant Galactic halo. **Panel A:** Groups of stars identified in a mock stellar halo (Bullock & Johnston 2005) in the range 100–300 kpc, using sky positions (shown) and distances only. **Panel B:** Same stars colored by progenitor galaxy. Green arrows highlight the contribution of interlopers to group 5 (dark orange) in panel A. **Panels C–E:** view of the same groups in energy-angular momentum projection, which requires six-dimensional phase space information including PMs. Some outliers are already identifiable at $25 \mu\text{as yr}^{-1}$ (panel D) and structures are clearly distinguishable at $5 \mu\text{as yr}^{-1}$ (panel E; green arrows). Groups 2 and 4 (dark and light blue, respectively) in panel A are from the same tidally-disrupted progenitor galaxy but are found on opposite sides of the sky; with $\leq 25 \mu\text{as yr}^{-1}$ precision they can be associated through orbit integration, reflected in panels D and E by their similar values of angular momentum (j_z).

2.4.2. Centering on diffraction spikes

A second strategy for obtaining highly accurate astrometry of very bright stars involves centering on diffraction spikes. The approach is facilitated by the properties of the WFIRST H4RG detectors, which do not show “bleeding” of excess charges from saturated pixels to their neighbors (see §5.2). Astrometric precisions of $10 \mu\text{as}$ or better are achievable with this technique with integrations of 100 s for stars with $J = 5$ or, making use of the recently added optical R062 filter, $R = 6$ (Melchior et al. 2017).²

As for spatial scanning, measurement accuracy will likely be limited by systematic uncertainties, in particular the fidelity of corrections for optical distortion and pixel-level artifacts (cf. § 5). Thus, performing several exposures per visit is beneficial and should be able to yield precisions of $10 \mu\text{as}$ or better even in the presence of residual systematics.

2.4.3. Detection of Earth-mass exoplanets

These estimates indicate that a dedicated GO program with visits to target fields separated by months and spread out over the lifetime of WFIRST could detect Earth-mass exoplanets

² Diffraction-spike measurements are superior in the short-wavelength range because the diffraction spike is sharper. Unlike the core of the WFIRST PSF, the diffraction spike is well sampled even in the bluest WFIRST filter given a pixel scale of 0.11 mas.

around the most nearby stars, in some cases even in their respective habitable zones. In addition, it can probe Neptune-class planets around more distant stars and, by adding earlier measurements from *Gaia*, rocky planets with periods of >10 yr. Such measurements would be strongly synergistic with radial velocity campaigns, improving the mass constraints and breaking degeneracies in several orbital parameters (Pourbaix & Jorissen 2000), and enabling mass estimates of the direct-imaging exoplanets of the WFIRST coronagraph and possible starshade occulter programs (Melchior et al. 2017).

2.5. Detailed structure of the inner Milky Way

Gaia will revolutionize our understanding of Milky Way structure in the outer parts of the Milky Way, including the halo. However, *Gaia* has a very limited view of the inner Milky Way due to the limited sensitivity at optical wavelengths (Figure 4). WFIRST will probe significantly deeper into the inner Milky Way, allowing us to map the structure and kinematics and complement the *Gaia* view. As an example, the EML survey will obtain precise parallaxes and ultra-precise PMs for over 50 million stars in a small area of the Galactic bulge, enabling a detailed analysis of their kinematics and density distribution. Currently, studies of bulge stellar populations are limited by the quality of the PM and the need to remove foreground disk stars, typically achieved via kinematic or photometric filters (see, e.g., Clarkson et al. 2008).

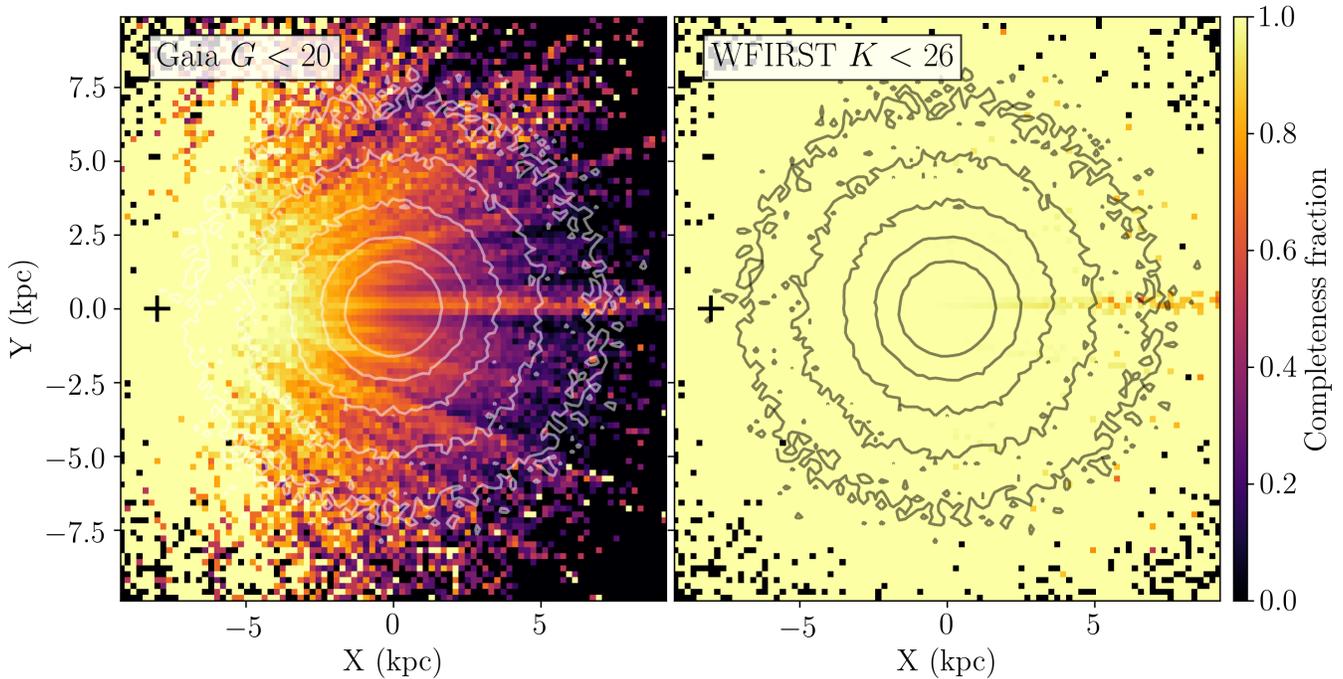


Figure 4. Simulated completeness of the distribution of red clump stars with distances $|z| < 500$ pc from the Galactic plane, based on the Milky-Way-like simulated galaxy in [Wetzel et al. \(2016\)](#). The left panel shows stars that *Gaia* would detect in the optical ($G < 20$); the right panel those seen by WFIRST in the infrared ($K < 26$). Grey contours in both panels show the density of the complete distribution on a logarithmic scale; the black cross marks the location of the Sun. The synthetic red-clump catalog was constructed by drawing stars in the range $-0.48 < M_I < -0.08$, $0.8 < V - I < 1.4$ ([Paczynski & Stanek 1998](#)) from [Marigo et al. \(2008\)](#) isochrones, distributed according to the age and stellar mass density of the simulated star particles (Sanderson et al. in prep). The [Schlegel et al. \(1998\)](#) three-dimensional extinction map was interpolated to determine apparent magnitudes and reddening, and *Gaia* G magnitudes were calculated using `pygaia` (github.com/agabrown/PyGaia). This simulated view ignores the effects of crowding and does not include a prominent Galactic bar (see [Romero-Gómez et al. 2015](#)). *Gaia* will largely be limited to heliocentric distances < 4 kpc in the plane, while WFIRST will measure parallaxes and velocities of stars well beyond the Galactic Center.

Both are statistical in nature and do not provide a direct determination of the distance to individual stars. According to the current requirements, a mission-long astrometric accuracy of $10 \mu\text{as}$ or better (with a stretch goal of $3 \mu\text{as}$) should be achieved at $H_{\text{AB}} = 21.6$ ([EML 20](#); see § 4.2).

At comparable accuracy in relative parallaxes, distances to individual stars can be measured to 9% at the bulge (3% if the stretch goal is achieved), and useful distance discrimination should be obtained to significantly fainter magnitudes. The two tangential components of the space velocity can be recovered to the same accuracy (in this regime, the distance error is dominant over the PM error in deriving the space velocity). This information will enable a much cleaner determination of the kinematics of the bulk of bulge stars in the EML survey field, and readily identify subgroups of stars—disk or bulge—with anomalous kinematics. If depth effects can be accounted for, the end-of-mission PM accuracy translates to a velocity precision of $\sim 1 \text{ km s}^{-1}$; together with the very large number of stars measured, this will permit a clear component separation of the spatially overlapping bulge and halo populations (see, e.g., [Minniti & Zoccali 2008](#)), and potentially identifying complex structures such as the anomalous

motions found in the X-shaped regions of the bulge (see, e.g., [Vásquez et al. 2013](#)). In principle, *Gaia* will achieve comparable precision over all of the bulge, but only for the bright red giants at $G_{\text{Gaia}} \sim 15$ or brighter; the uncertainties will be considerably larger (~ 2 orders of magnitude) at *Gaia*'s faint limit.

2.6. Star Formation in the Milky Way

With the advent of large infrared surveys of the Galactic Plane, many new young star clusters have been identified. The most massive of these young clusters are ideal laboratories for studies of star and cluster formation, stellar evolution, and cluster dynamics, but detailed studies of these regions are hampered by high and spatially variable extinction, high stellar densities, and confusion with foreground and background stars. Many of these limitations can be overcome with the addition of PMs observed in the infrared, to separate out the co-moving cluster members from the contaminating field population ([Stolte et al. 2008](#); [Hosek et al. 2015](#)). Furthermore, measurements of the internal velocity structure of star clusters provide constraints on the unseen stellar popula-

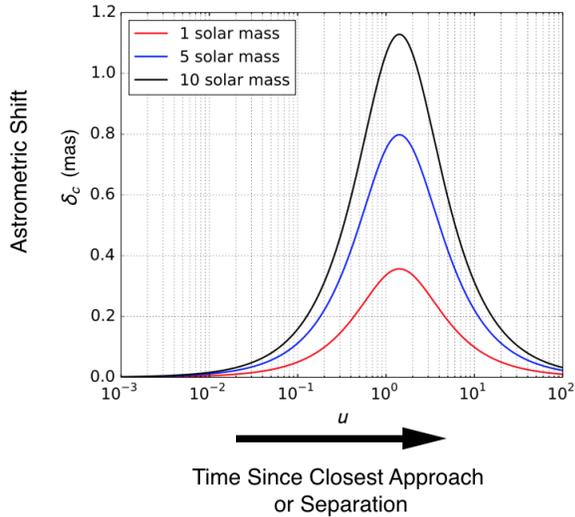


Figure 5. Astrometric shift of a background bulge star (source, $d=8$ kpc) lensed by a foreground compact object such as a black hole, neutron star, or white dwarf (lens, $d=4$ kpc). The astrometric shift changes as a function of the projected source-lens separation on the sky, u , in units of the Einstein radius. For the $10 M_{\odot}$ case, the Einstein radius is ~ 4 mas and the time for the source to cross the Einstein ring is typically >100 days.

tion from dynamical mass measurements, thereby informing cluster evolution models (Clarkson et al. 2012).

WFIRST is ideally suited for studies of massive young clusters and star forming regions in the Milky Way, given its wide field of view at infrared wavelengths, high spatial resolution, and potential for precise photometry and astrometry. For rapidly moving populations in the center of the Galaxy, a PM precision of ~ 0.5 mas yr $^{-1}$ per star is needed to separate cluster members from field stars. To obtain internal velocities or separate clusters in the disk, a PM precision of 0.05 mas yr $^{-1}$ or better is desired. Even higher astrometric precisions would enable searches for binaries and higher-order multiples.

An important factor to consider for this science case is that cluster members span a large range in brightness. The brightest and most massive cluster members in clusters beyond 4 kpc often have $J = 9$ or brighter. Careful calibration of persistence, shorter integration times or possibly narrow-band filters will be needed to reach both bright, massive members and faint, low-mass members of clusters.

2.7. Isolated Black Holes and Neutron Stars

Our Galaxy likely contains 10^7 – 10^8 stellar mass black holes and orders of magnitude more neutron stars (Agol & Kamionkowski 2002). Measuring the number and mass statistics of these stellar remnants will provide important constraints on the initial stellar mass function, the fate of massive stars and the initial-final mass relation, the star formation history of our Galaxy, and the fundamental physics of compact

objects. WFIRST has the ability to find such objects in large numbers through gravitational microlensing when a background star passes behind the compact object and is magnified photometrically. However, only the addition of WFIRST astrometry will enable us to measure the precise masses of these objects through astrometric microlensing. The apparent astrometric shift of the background star due to microlensing, which is proportional to $M^{1/2}$, is ~ 1 milli-arcsecond for a $10 M_{\odot}$ black hole at 4 kpc lensing a background star at 8 kpc (Figure 5). Thus, the necessary astrometric precision to detect isolated black holes is $<150 \mu\text{as}$; a factor of 2-3 better precision would also allow the detection of neutron stars.

While *Gaia* or ground-based adaptive optics systems may detect one or a few isolated black holes (Lu et al. 2016), WFIRST’s infrared capabilities and monitoring of the Galactic Center and bulge fields will yield the much larger samples needed to precisely measure the black hole and neutron star mass function and multiplicity. Microlensing by massive objects typically has long timescales, with Einstein crossing times >100 days for black holes, so WFIRST astrometry should be stable on these timescales; i.e., routinely calibrated on sky if possible.

2.8. Globular clusters

In the last decade, a wealth of revolutionary studies have dramatically changed the traditional view of globular clusters (GCs) as the best examples of “simple stellar populations:” stars with the same age and chemical composition. The presence of multiple stellar populations (MPs) in GCs has been widely established along all the stellar evolutionary phases (e.g., Carretta et al. 2009; Piotto et al. 2015 and references therein): spectroscopic studies have found significant star-to-star variation in light elements (e.g., Gratton et al. 2012 and references therein), while high-precision photometry, mostly from *HST* data, has clearly revealed the presence of distinct sequences in color-magnitude diagrams (CMDs) at all wavelengths (e.g., Milone et al. 2012; Piotto et al. 2015; Bellini et al. 2017c, see Fig. 6). Several GCs have also shown the presence of significantly He-enhanced subpopulations (e.g., Piotto et al. 2005, 2007) and even subpopulations with distinct iron content in a few cases like ω Cen, M22, Terzan 5, M54, NGC 5824, and M2 (Norris & Da Costa 1995; Johnson & Pilachowski 2010; Carretta et al. 2010; Milone et al. 2015). These observational findings present formidable challenges for theories of the formation and evolution of GCs, and have inaugurated a new era in GC research in which understanding how multiple stellar systems form and evolve is not just the curious study of an anomaly, but a fundamental key to understanding star formation.

Measuring the PMs of stars in GCs is the most effective way to constrain the structure, formation, and dynamical evolution of these ancient stellar systems and, in turn, that of the Milky Way itself. High-precision *HST* astrometry of GCs is now becoming available for a large number of objects (e.g., Bellini et al. 2014), but current PM catalogs are limited by the small field of view of *HST*, either to the innermost few arcminutes or to pencil-beam locations in the outskirts.

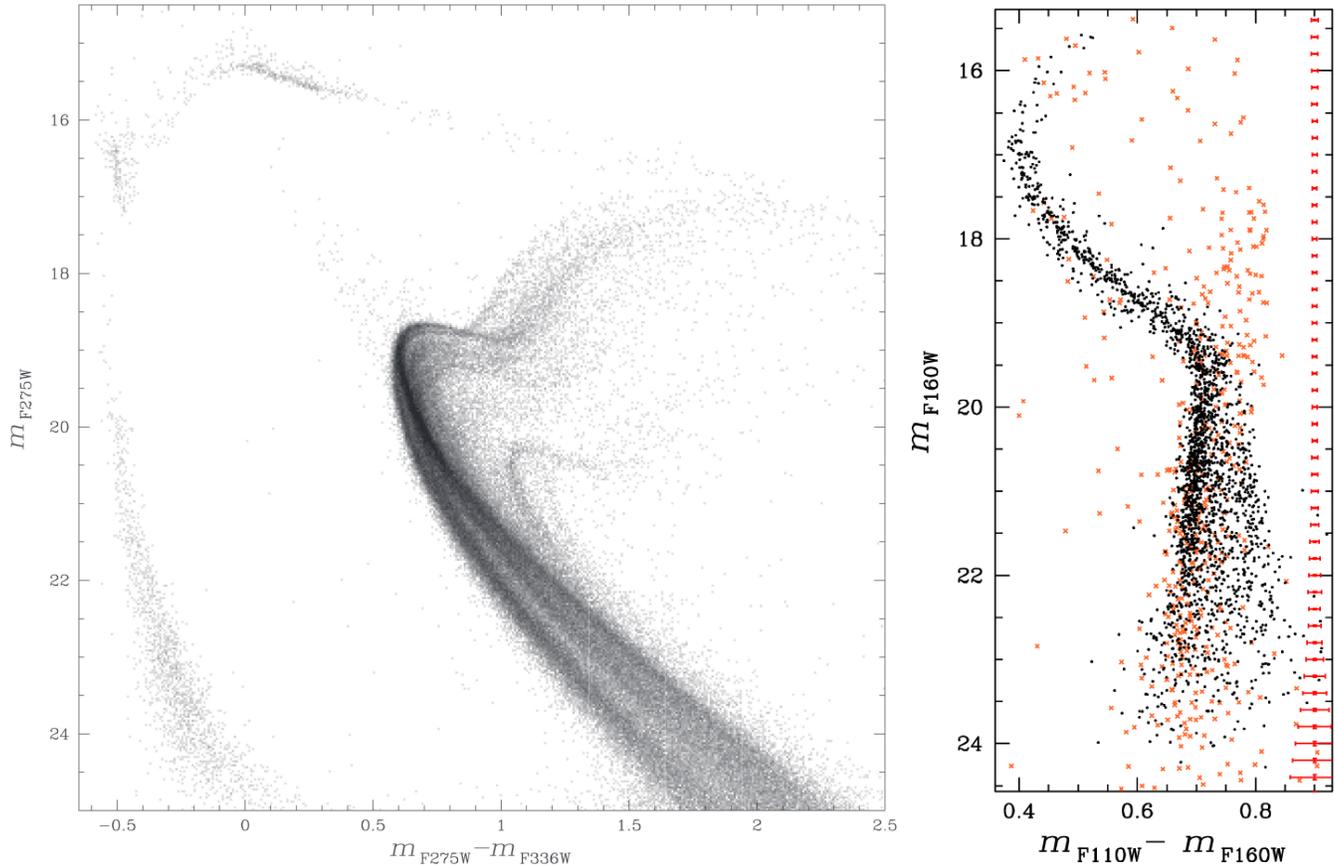


Figure 6. **Left:** The m_{F275W} vs. $m_{F275W} - m_{F336W}$ CMD of ω Cen, showing several subpopulations of stars in all evolutionary sequences (from Bellini et al. 2017a). MPs reveal themselves in high-precision photometry from the UV to the near IR. **Right:** m_{F160W} vs. $m_{F110W} - m_{F160W}$ CMD of an outer field of ω Cen, corrected for differential reddening. Field stars (orange dots) are identified using proper motions. From Milone et al. (2017).

While most dynamical interactions do happen in the center of GCs, answering many outstanding questions will require high-precision PMs of faint cluster stars over wide fields, for which WFIRST is by far the best tool. Here we discuss a few examples of such investigations.

2.8.1. Multiple-population internal kinematics

The PM-based kinematic properties of MPs have so far been characterized for only three GCs: 47 Tuc (Richer et al. 2013), ω Cen (Anderson & van der Marel 2010), and NGC 2808 (Bellini et al. 2015). The short two-body relaxation timescale in the inner regions of these clusters, where most observations have so far been focused, implies that any initial differences in the kinematic properties of different stellar populations have likely been erased. The cluster outskirts, however, have much longer relaxation timescales and could still retain fossil kinematic information about the early stages of cluster evolution (e.g., D’Ercole et al. 2008). The outer regions can thus provide a wealth of information and constraints on the formation and early dynamics of MP clusters, on the subsequent long-term dynamical evolution driven by two-body relaxation, and on the role of the Galactic tidal field in the outskirts of clusters. For instance, it has been shown

that second-generation stars in 47 Tuc (Richer et al. 2013) and NGC 2808 (Bellini et al. 2015) are characterized by an increasing radial anisotropy in the outer regions with respect to first-generation stars (see Fig. 7). Even further out, at distances approaching the tidal radius, the effects of the external tidal field are expected to lead to a more isotropic velocity distribution.

Both WFIRST’s wide field of view and its improved sensitivity will make revolutionary steps forward in understanding the initial differences in MPs if sufficient PM accuracy can be achieved. Due to mass segregation, the most abundant stars in the outskirts of GCs are low-mass, faint main-sequence objects. *Gaia* can only measure stars as faint as the turn-off region in most clusters, and therefore will not be able to provide enough statistics to properly characterize the kinematics of the outer cluster regions. The expected internal velocity dispersion of cluster stars near the tidal radius is of the order of $\lesssim 1\text{--}3 \text{ km s}^{-1}$, even for the most massive clusters. The PM error adds in quadrature, so it should be less than half the intrinsic velocity dispersion, i.e., $\lesssim 1 \text{ km s}^{-1}$, in order to measure dispersions in cluster outskirts. At the typical distance of Galactic GCs, $\sim 10 \text{ kpc}$, this translates into PM errors of the order of $\lesssim 20 \mu\text{as yr}^{-1}$.

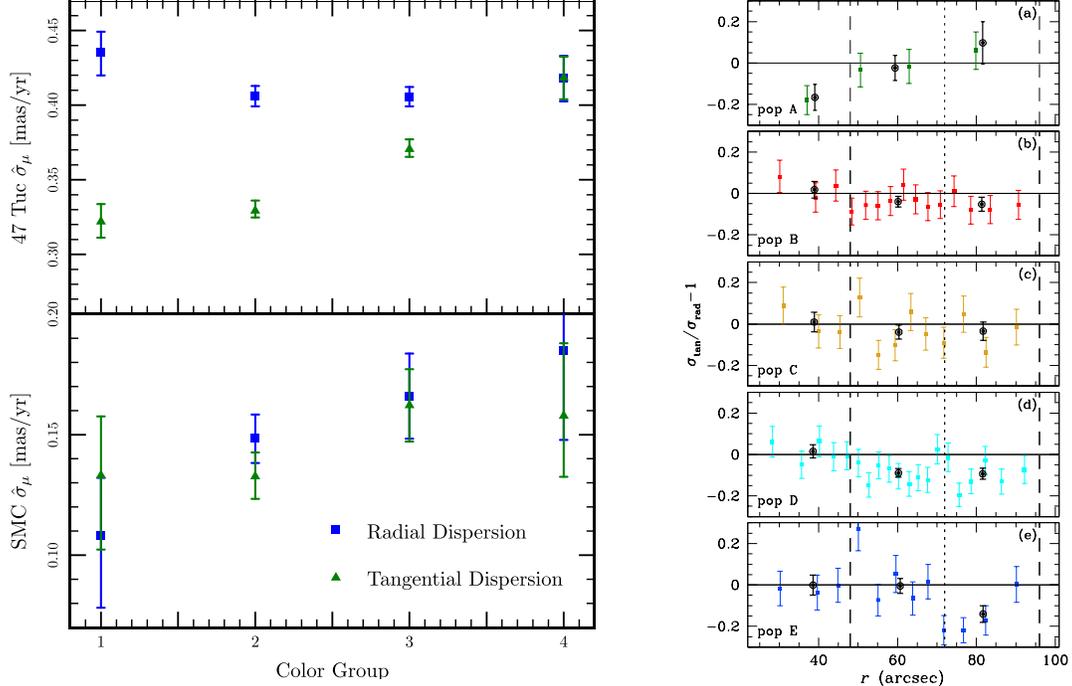


Figure 7. **Left:** Radial (blue) and tangential (green) PM dispersions as a function of color for the bluest (left) and reddest (right) parts of the MS of 47 Tuc (top) and of the SMC (Bottom). (From Richer et al. 2013). **Right:** Deviation from tangential-to-radial isotropy (horizontal line) for the 5 subpopulations in NGC 2808. Vertical lines mark the locations of r_h , $1.5 \times r_h$, and $2 \times r_h$ (From Bellini et al. 2015).

2.8.2. Energy equipartition

It is widely assumed that GCs evolve towards a state of energy equipartition over many two-body relaxation times, so that the velocity dispersion of an evolved cluster should scale with stellar mass as $\sigma \propto m^{-\eta}$, with $\eta = 0.5$. Recently, Trenti & van der Marel (2013) used direct N-body simulations with a variety of realistic initial mass functions and initial conditions to show that this scenario is not correct (see also Bianchini et al. 2016). None of these simulated systems reaches a state close to equipartition: instead, over sufficiently long timescales the mass-velocity dispersion relation converges to the value $\eta_\infty \sim 0.08$ as a consequence of the Spritzer instability (see Fig. 8). These intriguing results have yet to be observationally tested.

To measure η , a wide range of stellar masses must be probed. Again, this task is out of reach for *Gaia* because of its relatively bright magnitude limit, but WFIRST will easily measure high-precision PMs down to the hydrogen burning limit (HBL; $\sim 0.08 M_\odot$) and out to the tidal radius, thus constraining both the current state of energy equipartition in a cluster and its past dynamical evolution. As in the previous case, PM errors of the order of $\lesssim 20 \mu\text{as yr}^{-1}$ are needed.

2.8.3. The hydrogen-burning limit and the brown-dwarf regime

WFIRST will also make it possible to study the luminosity functions of GCs beyond the HBL and into the brown-dwarf regime. Close to the HBL, old stars show a huge difference in luminosity for a small difference in mass, resulting in a plunge of the luminosity function toward zero for stars with masses just above this limit. Stars in GCs are homogeneous

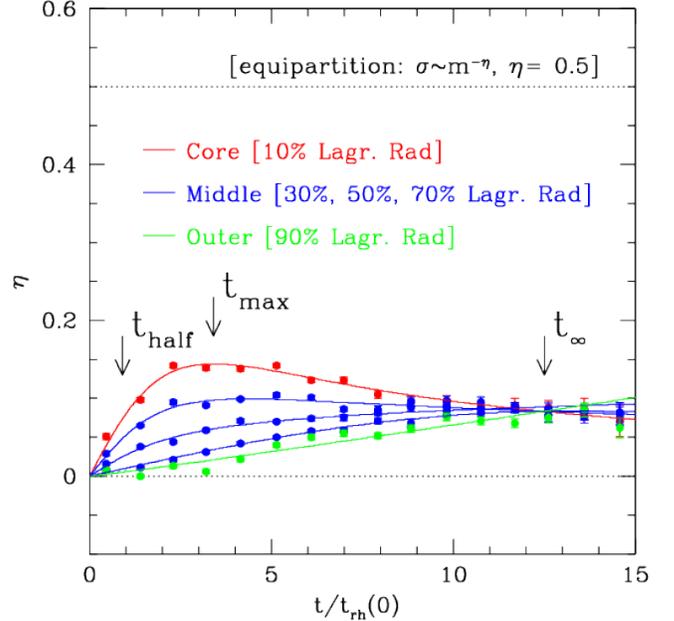


Figure 8. Time evolution of the energy equipartition indicator η for single main-sequence stars in N-body simulations, from Trenti & van der Marel (2013). The time along the abscissa is expressed in units of the initial half-mass relaxation time $t_{rh}(0)$. Complete energy equipartition ($\eta=0.5$; dotted line) is never attained, confirming previous investigations based on stability analysis.

in age, distance, and chemical composition (within the same

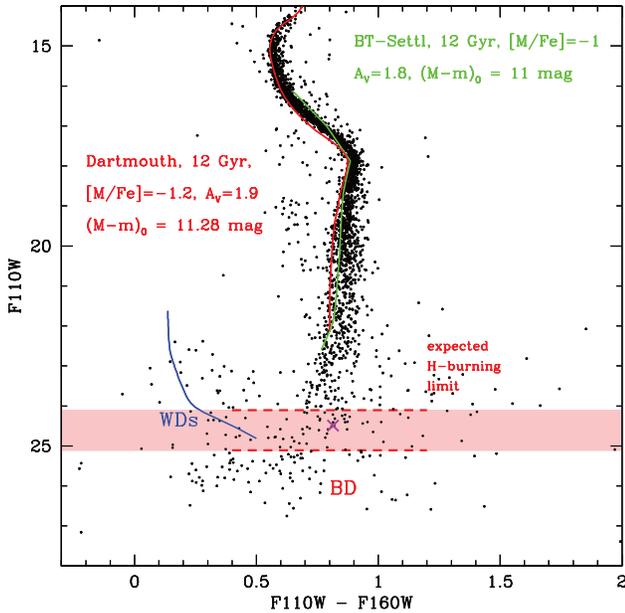


Figure 9. Deep NIR CMD of the GC M4, from Dieball et al. (2016). The white-dwarf and brown-dwarf regions are labeled, and low-mass stellar models are over-plotted in green and red. The expected end of the H-burning sequence is marked with red dashed lines and a shaded area.

subpopulation), so at the typical GC age of > 10 Gyr, stars with masses below the HBL will have faded by several magnitudes relative to those above it, creating a virtual cutoff in the luminosity function (e.g., Bedin et al. 2001; Dieball et al. 2016; see also Fig. 9).

The best place to observe the properties of stars approaching the HBL is once again outside a cluster’s core region, where contamination by light from much brighter red-giant-branch stars is negligible. The brown-dwarf regime in GCs is unexplored ground, so many new and intriguing discoveries may be waiting for WFIRST. Due to the relatively low number density of low-mass MS and brown-dwarf stars in the cluster outskirts, WFIRST is the perfect astronomical tool for these investigations as well. Proper-motion-based cluster-field separation is needed to create clean samples of cluster members; PM errors of the order of a few tenths of mas yr^{-1} might be sufficient to separate cluster stars from field stars, while errors one order of magnitude smaller would also enable studies of the internal kinematics of cluster stars in these lowest-mass regimes.

3. WFIRST ABSOLUTE ASTROMETRIC PERFORMANCE

High-precision absolute astrometric measurements with *HST* have typically been based on the positions of well-measured background galaxies within the field of view (FoV), but a new major improvement in absolute astrometry measurements is imminent. When WFIRST begins observing the universe, the ESA mission *Gaia* will already

be complete, providing absolute astrometric positions of unprecedented accuracy everywhere on the sky.

In broad terms, there are two methods to determine the absolute astrometric accuracy of a scientific observation. The first method is based on the information available before an observation is made, essentially the same information that is used to point the telescope (hereafter the “a priori” method). This information includes the celestial coordinates of the guide stars (GSs) and the locations of the scientific instruments relative to the GSs in the focal plane of the telescope. The second method is based on the information available after an observation is made, namely the positions of all sources with accurate coordinates in external catalogs (hereafter, the “a posteriori” method). Although typically up to four stars will be used to guide WFIRST observations, many more, fainter stars within each exposure can be used to improve the absolute astrometric precision with the a posteriori method (Bellini et al. 2017b).

3.1. The a priori method

3.1.1. Imaging mode

According to WFIRST’s current design, GSs can be placed on any four of the 18 WFI detectors, but at most one GS will be assigned to a given detector. Following a slew to a new position on the sky, the attitude control system will determine when the attitude and angular rate errors have fallen below a suitable threshold, indicating that conditions are acceptable for acquiring GSs. Once a GS is selected on a particular detector, a pre-guide window of 64×64 pixels is centered around the guide star to lock the telescope on its position. The size of this window is set large enough to accommodate $3\text{-}\sigma$ uncertainties in the attitude. GSs thus have to be relatively isolated, to avoid confusing the guider into locking on a different nearby star and introducing additional astrometric error. This GS isolation requirement is met when no other star of magnitude $m_{\text{guide}} + 2.5$ or brighter is present within the pre-guide window.

After the post-slew attitude error has been reduced by the attitude control system, the size of the pre-guide window is reduced to 16×16 pixels, which are then read with an update rate of 5.86 Hz (about 6 times per second). This rate sets both a minimum brightness (at which a star has sufficient signal-to-noise to be detected) and a maximum brightness (above which a star will saturate on this timescale) for guide stars. The minimum brightness a star must have to qualify as a GS is $H_{\text{AB}} = 17$ (or $H_{2\text{MASS}} \sim 15.5$), while the saturation limit is set to $H_{2\text{MASS}} = 9.5$.³

³ Based on Monte Carlo simulations, about 22% of stellar flux lands within the central pixel in the H158 filter. The Galaxy-Survey exposure time calculator (wfirst.ipac.caltech.edu/sims/tools/wfDepc/wfDepc.html) predicts that about 227 electrons land within the central pixel in 1/6 second for the H158 filter. Assuming 16-bit detectors and electronic saturation occurring before physical saturation, 60 000 counts are collected in 1/6 second for stars with a magnitude $H_{2\text{MASS}} = 9.5$, defining the saturation limit.

GUIDING: IMAGING MODE $9.5 < H_{2\text{MASS}} < 15.5$ (no bright neighbors within $5''$)

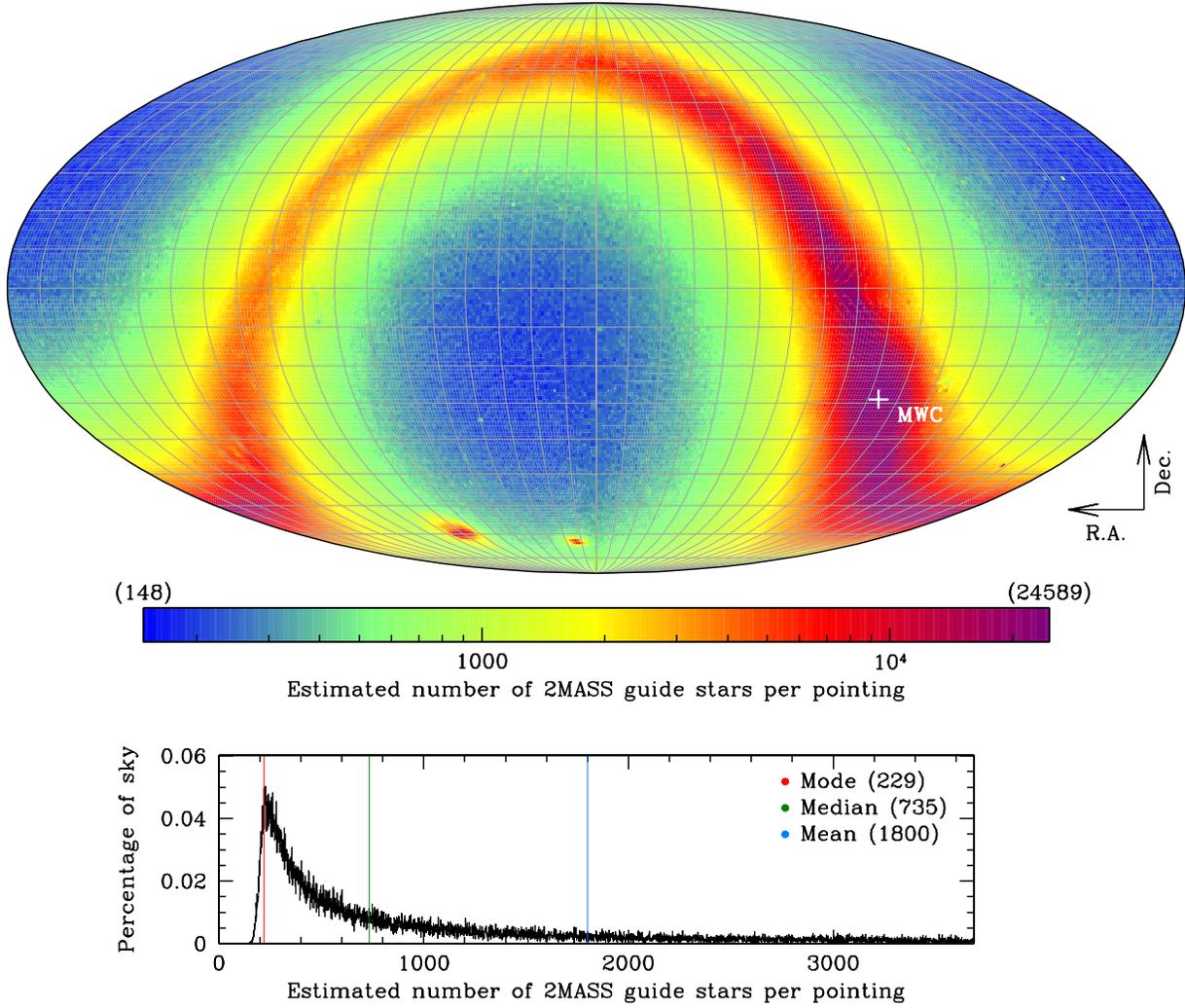


Figure 10. Top: Map of the number of available guide stars within a WFIRST WFI FoV as a function of sky position, based on the 2MASS catalog. The Galactic center is marked with a white cross. The minimum and maximum numbers of available guide stars are reported in parentheses at the edges of the color bar. **Bottom:** distribution of the number of guide stars per WFIRST pointing. We expect at least 700 available guide stars per pointing (about 40 GSs per chip; green line) over half the sky.

Fig. 10 shows a density map of the sky based on the number of expected GSs within the WFIRST FoV, based on the 2MASS catalog, which is fairly complete in the magnitude range $9.5 \leq H_{2\text{MASS}} \leq 15.5$. A single WFIRST pointing will contain at least 700 GSs satisfying this criterion over half the sky, and there are on average at least ~ 8 GSs per chip to choose from for the imaging mode even in the sparsest regions of the sky.

3.1.2. Grism mode

The availability of GSs all across the sky for the grism mode is estimated following the same procedures as for the

imaging mode. The GS magnitude range for the grism mode is expected to be $8 < H_{\text{AB}} < 14$, but Hirata (2016, private communication) has suggested using a faint limit of $H_{\text{AB}} = 13$ (or even as $H_{\text{AB}} = 12$) to account for the different spectral types of GSs. If the faint limit is $H_{\text{AB}} = 14$, then at least 12 GSs are available per WFIRST pointing. Assuming a random distribution of GSs across the FoV, the probability of having at least 4 guiding detectors is 100% (see Nelan et al. 2015). If the faint limit is $H_{\text{AB}} = 13$, then a few regions (about 0.001% of the sky) will have ≤ 4 GSs

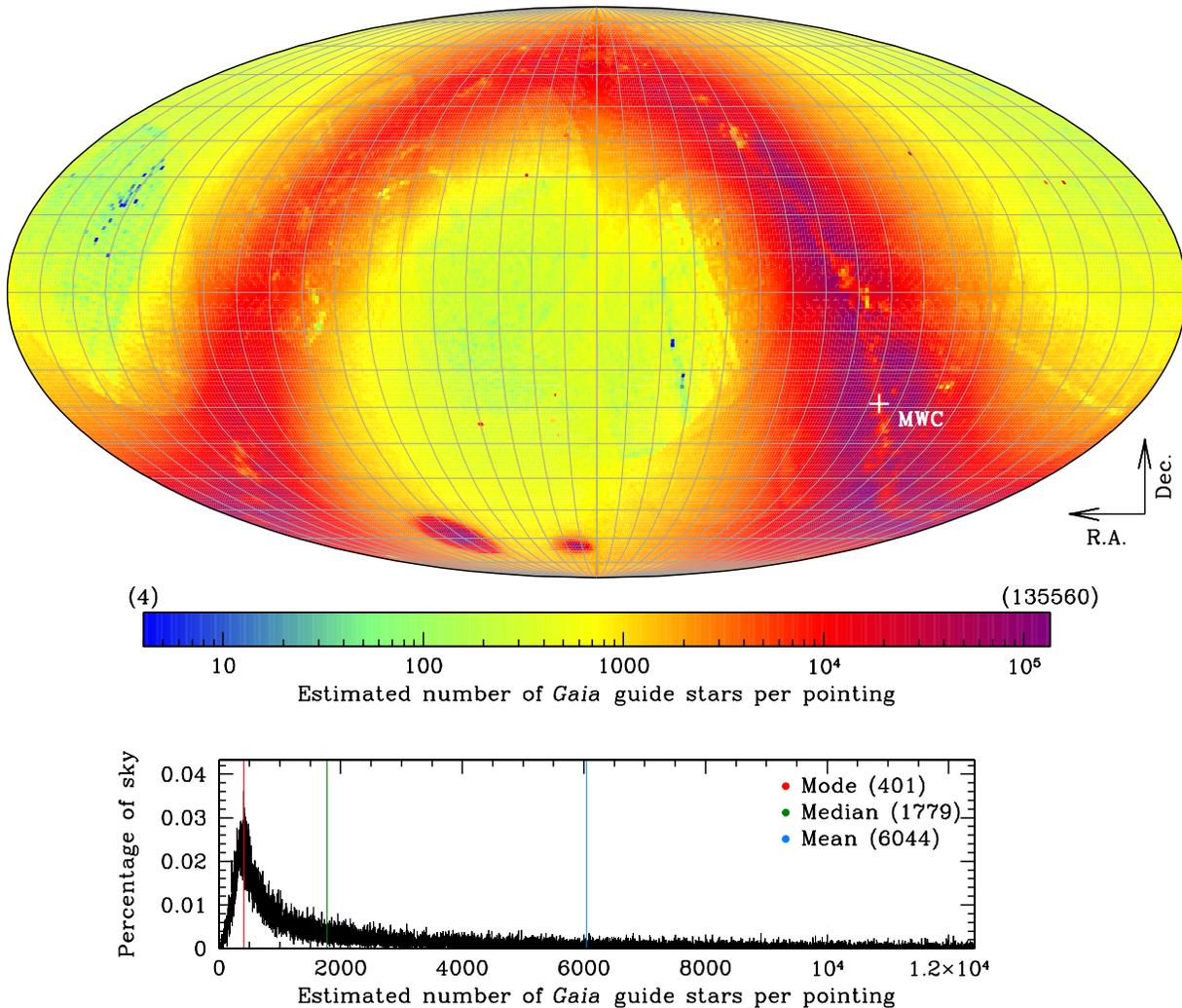
A-POSTERIORI METHOD: ALL *Gaia* DR1 ($G > 17.3$)

Figure 11. As Figure 10, but for the a posteriori method using the *Gaia* DR1 catalog.

available for the entire WFI FoV.⁴ If the faint limit is further decreased to $H_{AB} = 12$, then about 3.5% of the sky will have four or fewer GSs and $\sim 0.13\%$ of the sky (~ 58 sq. deg.) will have less than two GSs. The affected regions are mainly around the Galactic poles.

3.2. The a posteriori method

The a posteriori method uses significantly fainter (but more plentiful) stars than the a priori method. This is due to the significantly different saturation time scale in the guide-star window relative to the rest of the detector. The 16×16 pixel

guide window is read every $\sim 1/6$ seconds, so GSs must not saturate within this time interval. On the other hand, the science images are read 16 times slower than the guide windows, and pixels will be irrecoverably saturated if they reach the full well before the fourth such read (i.e., before ~ 8.19 seconds). The precise location of such a “hard-saturated” source is difficult to obtain using aperture-based centroids or PSF fitting. The bright limit is thus set by this criterion to $H_{2\text{MASS}} = 13.7$ (approximately $G_{\text{Gaia}} = 17.3$), a factor of 48 lower in flux (or a 4.2 mag difference) than the saturation criterion in the GS window. Assuming the same S/N limit for GSs applies to the a posteriori method, the faint limit in the *shortest* possible science exposure with WFIRST, 8.19 sec, is $H_{2\text{MASS}} = 22$ (approximately $G_{\text{Gaia}} = 25.6$): about 2 mag fainter than the 2MASS catalog itself. Most WFIRST expo-

⁴ Even when 4 GSs are available, they could of course land on fewer than four chips.

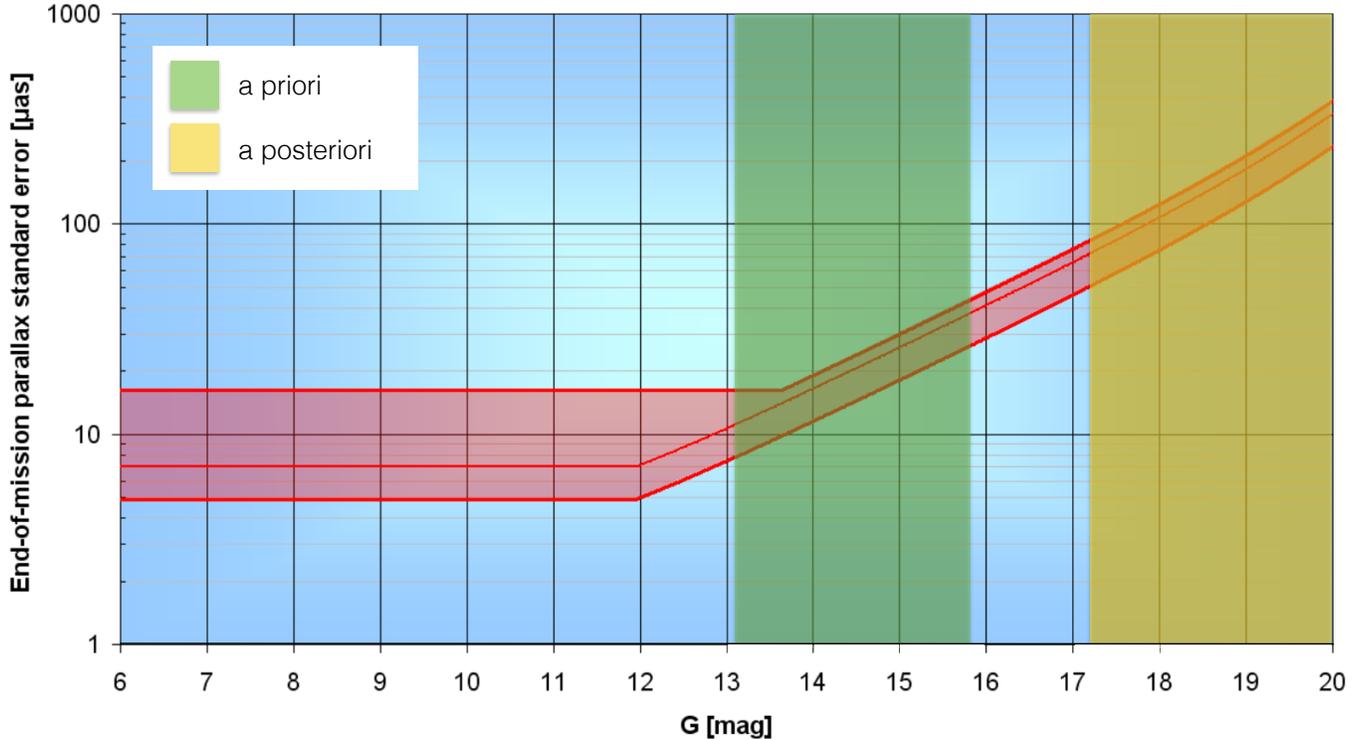


Figure 12. Expected *Gaia* end-of-mission astrometric errors as a function of magnitude, adapted from [de Bruijne 2012](#). *Gaia* stars that can be used for the a priori (a posteriori) method fall within the green (yellow) region.

tures will of course be much longer than 8.19 seconds, thus pushing far beyond the *Gaia* faint limit of $G_{\text{Gaia}} \sim 20.5$.

Figure 11 shows the all-sky density map of the number of expected *Gaia* sources that can be used for the a posteriori method. There are at least 1780 such sources per WFIRST pointing in half the sky. The minimum number of sources predicted by this map is currently unreliable due to the incompleteness of *Gaia* DR1 ([Arenou et al. 2017](#)); these regions will contain many more sources in future *Gaia* data releases, so this map can be considered a lower limit on the number of suitable a posteriori targets per field. Based on the density of *Gaia* stars around the North Galactic pole, where DR1 is fairly complete, a more reliable minimum is ~ 450 stars per WFIRST FoV. Simulations of the *Gaia* catalog based on the Besançon Milky Way model (which includes disk, bulge and spheroid but no accreted halo) also estimate the minimum density of stars to be at least 300 per WFIRST FoV at the Galactic poles (Fig. 3 of [Robin et al. 2012](#)).

3.3. Expected absolute astrometric performance

The WFIRST mission will begin operation in the second half of the 2020s, several years after *Gaia* has completed its 5-year mission in 2019.⁵ Near the faint *Gaia* limit ($19 < G_{\text{Gaia}} < 20$), PMs in the *Gaia* catalog will have an end-of-mission error of about $0.2\text{--}0.3 \text{ mas yr}^{-1}$ ([de Bruijne 2012](#); see Fig. 12). This translates into a position uncertainty

of about $1.4\text{--}2.1 \text{ mas}$ (or $\sim 0.01\text{--}0.02$ WFIRST WFI pixels) at the start of WFIRST operations, and about twice as much by year 5. Position errors of this size will have a significant impact if the goal is to achieve high-precision (to better than 0.01 pixels) absolute astrometric measurements. In the following, we provide expected estimates based only on catalog errors. All other sources of errors (geometric-distortion residuals, centroiding errors, etc.) are ignored (more in [Bellini et al. 2017b](#)).

Single-epoch GO and Guest-Investigator (GI) observations of a random location on the sky may have to rely solely on the information contained in prior astrometric catalogs (in particular *Gaia*'s catalog) to determine the absolute position of their sources. Assuming an average per-star positional error of 2 mas (corresponding *Gaia*'s expected end-of-mission astrometric error for $G_{\text{Gaia}} = 19$, extrapolated to the late 2020s), and ignoring all other sources of errors (e.g., geometric-distortion or source centroiding errors), it will be possible to obtain absolute a posteriori positions to better than $\sim 0.05 \text{ mas}$ (or about 5×10^{-4} WFI pixels) over half the sky. In regions with the lowest stellar densities, the expected absolute position error increases to $\sim 0.1 \text{ mas}$ ($\sim 10^{-3}$ WFI pixels).

For the planned WFIRST mission surveys (e.g., HLS and EML survey), repeated WFIRST observations spanning several years can be used to improve *Gaia*'s PMs, especially at the faint end, and to derive absolute positions and PMs for many fainter sources. The absolute astrometric precision for the planned surveys is expected to be significantly better than

⁵ There is the possibility that the *Gaia* mission will be extended five extra years.

what can be done with *Gaia* alone, but is difficult to quantify at this time.

For all stars suitable for the a priori method, *Gaia*'s expected end-of-mission astrometric error ranges between 10 and 80 $\mu\text{as yr}^{-1}$ (de Bruijne 2012), but WFIRST will likely choose the four GSs among the brightest available sources. We estimate that at least 7–8 GSs in the range $10.0 < H_{2\text{MASS}} < 10.2$ will be available over half the sky. If these stars land on at least four different WFI chips (a near 100% chance; Nelan et al. 2015), assuming their average magnitude is $H_{2\text{MASS}} \sim 10.1$ (corresponding to $G_{\text{Gaia}} \sim 13.7$), and assuming they have a *Gaia*-extrapolated position error of ~ 0.15 mas in late 2020s, then the a priori method is expected to offer absolute position measurements at the 0.075 mas level or better (7×10^{-4} WFI pixels) for half the sky. For the entire sky, on the other hand, we always expect at least 7–8 GSs within any given WFI FoV if the faint limit is relaxed to $H_{2\text{MASS}} = 12.4$ (or roughly $G_{\text{Gaia}} = 15.8$). This translates into an upper limit for the expected a priori astrometric error of 0.2 mas (or about 2×10^{-3} pixels).

The proposed 5-year extension to the *Gaia* mission would improve WFIRST astrometry substantially. In this case each *Gaia* source will have twice as many measurements over twice the time baseline, providing an increase in precision by a factor of $2\sqrt{2}$; moreover, WFIRST would only need to extrapolate *Gaia*'s positions for 5 years instead of 10, resulting in an additional factor of 2 improvement. Together, this means a 5-year extension to *Gaia* would result in a factor of 5.6 improvement in absolute astrometry for WFIRST, to an expected accuracy of about 20 μas for the a posteriori method described above.

Finally, the possibility to guide the telescope using more than four guide stars, perhaps even one per detector, would significantly improve the determination of the temporal stability of the imager and the global distortion solution. If only four guide stars are used, we recommend choosing well-separated stars spread across the WFI, preferably near each of the four corners, in order to minimize guiding errors.

4. ASTROMETRY WITH CORE SCIENCE PROGRAMS

Many of the features that make WFIRST an excellent astrometric instrument, notably the requirements for excellent PSF modeling and thermal stability, are already addressed by the core science programs such as the High-Latitude Survey and Exoplanet MicroLensing survey. Given their use of repeated visits to the same fields, both of these programs also offer an opportunity to produce excellent astrometry for all observed stars in their footprints with little extra analysis. *We recommend that derived proper motions should be provided as part of the object catalogs for both surveys once multiple epochs have been observed.* In the HLS particularly, a further cross-match of stars to the LSST catalog would extend the LSST survey into the IR regime for the region covered by HLS, identify variable stars that can be used as standard candles (particularly for RR Lyrae stars, as the period-luminosity relation is much tighter in the IR), and allow for cross-validation of PMs.

Here we summarize the current state of relevant requirements for these two surveys, broken down into three categories: Basic Science Requirements (BSRs), requirements for the High-Latitude Imaging Survey (HLIS), and requirements for the EML survey.

4.1. Astrometry with the High-Latitude Survey

The current requirements related to astrometry, as of July 2017, include the following:

- : **BSR 2:** WFIRST WFI shall measure shapes of galaxies at $z=0-2$ in at least 2 bands and fluxes in at least 4 bands for photometric redshifts, at a depth equivalent to a 5-sigma point source detection at AB magnitude $J < 26.9$ or $H < 26.7$, with photometric accuracy of 1% and with rms uncertainties (in the shape measurement filters only) below 10^{-3} in the PSF second moment and below 5×10^{-4} in the PSF ellipticity, in the HLS imaging survey.
- : **HLIS 7:** Obtain photometry, position, and shape measurements of galaxies in 3 filters (J , H , and F184), and photometry and position measurements in one additional color filter (Y ; only for photo- z).
- : **HLIS 8:** Obtain $S/N \geq 18$ (matched filter detection significance) per shape/color filter for galaxy $r_{\text{eff}} = 180$ mas and AB mag = 24.7/24.6/24.1 ($J/H/F184$).
- : **HLIS 9:** Determine PSF second moment to a relative error of $\leq 9.3 \times 10^{-4}$ rms (shape/color filters only).
- : **HLIS 10:** Determine PSF ellipticity to $\leq 4.7 \times 10^{-4}$ rms (shape/color filters only).
- : **HLIS 11:** System PSF EE50 radius ≤ 0.12 (Y band), 0.12 (J), 0.14 (H), or 0.13 (F184) arcsec.

The reference dither pattern for the HLS is a set of 3–4 dithers of a size intended to cover the chip gaps, repeated to tile the field and in each of 4 filters: Y , J , H , and F184. A second pass over each field follows six months later at a different roll angle.

The projected, approximate bright and faint point-source limits of the HLS are summarized in Table 3. The faint limits are as stated in the requirements above. The bright limits were estimated based on the statement in the reference mission that pixels in the HLS will be read non-destructively every 5.4 seconds, and assuming that any pixels that saturate before the fourth such read will be hard-saturated (see the discussion in § 3). Using the GS ETC (see § 3), assuming 25% of light in the central pixel for all filters and 65,000 electrons as the saturation level, the values listed in the table give the approximate bright limit (probably accurate to within 0.3–0.4 magnitudes).

4.2. Astrometry with the Exoplanet MicroLensing Survey

The current astrometry-related requirements being discussed (as of 29 June 2017) for the EML survey are:

HLS	Y	J	H	F184
Bright limit	15.5	15.6	15.7	15.3
Faint limit	–	26.9	26.7	–

Table 3. Bright and faint point-source limits for the WFIRST High-Latitude Survey (AB magnitudes), based on the requirements described in § 4.1.

- : **EML 8:** Relative photometric measurements in the primary microlensing filter that have a statistical S/N of ≥ 100 per exposure for a $H_{AB} = 21.6$ star.
- : **EML 14:** The EE50⁶ radius of the PSF in the wide filter shall be $< 0''.15$.
- : **EML 19:** The relative astrometric measurements shall have a statistical precision of 1 mas per measurement for $H_{AB} = 21.6$.
- : **EML 20:** Relative astrometric measurements will have systematic precision of 10 μ as over the full survey (stretch goal of 3 μ as).

Currently, both narrow (2-pixel wide) and large (10'' wide) possibilities for dithering are being explored for this core project. From the perspective of astrometric calibration, large dithers are crucial to accurately measure skew and kurtosis in the wings of the PSF.

Parallaxes and PMs over the bulge field are part of the mission of this program to characterize the masses of the star-planet pairs that will be discovered. The survey is therefore requesting the first two (spring/fall) and last two bulge observing seasons over the full time-baseline of the mission. This is also optimal for general astrometry in the bulge fields but may pose problems for understanding long-term variations in the PSF (on timescales of a year or so) prior to the end of the mission.

4.3. Guest Observer/Investigator Astrometry

Much of the science described in § 2 will likely be carried out through Guest Observer and Guest Investigator programs. It is therefore crucial that the benefit of the astrometric requirements for the two core science programs listed above be made available to GO/GI programs as well (by providing the calibration information for these programs to GOs/GIs and allowing for multiple astrometric calibrations within the archive) in order to achieve the promised precision. This need is discussed in depth in § 5.8.

5. RECOMMENDATIONS

Here we consider what is most likely to have an effect on the astrometric performance of the Wide-Field Imager. We

highlight areas where astrometry-specific considerations are especially important and can add significant extra science capability with little to no extra cost. Our recommendations are summarized in Table 4.

5.1. Geometric Distortion

Geometric distortion (GD) is the most significant systematic contributor for astrometry that is not currently covered by explicit requirements for either the HLS or EML survey. A dedicated set of observations to autocalibrate the GD of the WFI is currently being considered. There are two main ways to solve for the GD: using previous knowledge of the stellar positions in the field from existing astrometric catalogs (the ‘‘catalog’’ method), or via autocalibration, in which stellar positions themselves are iteratively solved for together with the GD. Each of these approaches has different advantages and disadvantages, but both depend strongly on the precision with which the position of stars can be measured using appropriate point-spread functions (e.g., Anderson & King 2003, 2004; Anderson et al. 2006; Anderson & King 2006; Bellini & Bedin 2009, 2010; Bellini et al. 2011; Libralato et al. 2014, 2015; Massari et al. 2016; Dalessandro et al. 2016).

The catalog method is less demanding of telescope time, since it requires fewer images to calibrate the GD and monitor temporal variations, but it strongly depends on the quality of the astrometric catalog used as a reference, since all the residuals and the systematic errors present in the reference catalog will also be present in the GD solution. In addition, unless the reference catalog is based itself on images taken very close in time to the WFIRST calibration images, proper motions (and their errors) can introduce significant residuals in the GD solution. The autocalibration approach requires more images, and therefore more telescope time, but offers a self-consistent calibration solution and can be designed to be formally insensitive to proper-motion-related errors.

On-sky GD calibration has historically been performed using large dithered exposures (as wide as the FoV in some cases) of a homogeneously-distributed, moderately dense stellar field. Stars in the EML survey fields are homogeneously distributed, and while their overall stellar density can be too high, this can be mitigated by using only the brightest stars in each field to calibrate the GD. If the exposure time has been carefully chosen, the bright, unsaturated stars will still be reasonably far apart from each other, and their surrounding neighbors typically a few magnitudes fainter, so that the bright stars can still be considered fairly isolated. The Baade window was successfully used by Anderson et al. (2006) and Libralato et al. (2014) to calibrate the GD of the ground-based ESO WFI@MPG and HAWK-I@VLT detectors, respectively. The globular cluster ω Cen, another possible target field for WFIRST calibration, was used by Libralato et al. (2015) to calibrate the GD of the infrared WFI VIRCAM@VISTA. The same technique of using only the brightest stars could be applied to the crowded regions in the core of ω Cen. On the other hand, the stellar density near the tidal radius ($\sim 48'$, Harris 1996, 2010 edition) may be too low. Other globular clusters could also be used for calibra-

⁶ 50% Encircled Energy.

§	Topic	Recommendation
5.1	Geometric distortion	This significant systematic for astrometry, not currently covered by core science requirements, should be considered in calibration plans for the WFI.
5.2	Pixel-level effects	Ground-based calibration should be considered, based on results of current tests by several labs. A spatial scanning mode would mitigate these effects for bright stars.
5.3	Filters & Chromaticity	Likely straightforward to calibrate, but should be aware of systematic effects.
5.4	Readout Hysteresis	Straightforward to minimize based on experience with current generation of HXRG detector/amplifier combinations.
5.5	Scheduling	
	HLS	Optimal to evenly space observations over full time of survey, to extent permitted by other requirements. Current example schedules vary in PM outcome by factor of $\gtrsim 2$.
	EML survey	Programming an occasional larger dither will significantly help calibrate for general astrometry. Largest possible time-spacing between first and last exposures is optimal; regular intermediate observations will increase understanding of long-term PSF variations.
	GO	The TAC process should allow for multi-year GO proposals to optimize PM baselines. For proposals covering large sky areas, time between field revisits should be maximized.
5.6	Jitter	This may be an issue for WFIRST where it was not for <i>HST</i> , given large requirement (14 mas). Requirements of the HLS for galaxy shape determination should help.
5.7	Data Management	Downloading every read with no coadds for at least part of the FoV is highly desirable for spatial scanning observations (see Section 2.4). Downloads of guide star postage stamps are crucial and inexpensive for PSF jitter correction.
5.8	High-level data products & Archive	Astrometry (linked to the <i>Gaia</i> frame) and astrometric uncertainties (including PSF centroiding error estimates) should be part of the high-level products. A requirement should be set on the astrometric uncertainty. The archive should allow for multiple upgradable astrometric solutions.

Table 4. Summary of main recommendations for astrometry.

tion, but because of its overall high number of members and its large tidal radius, ω Cen is the best target.

5.1.1. Example of an autocalibration strategy

A possible autocalibration strategy for the WFIRST WFI could be modeled on the work of Libralato et al. (2015) for the VIRCAM@VISTA detectors. The VIRCAM WFI comprises 16 $2k \times 2k$ VIRGO detectors, for a total FoV of about 1.3×1.0 sq. deg, but the very large gaps between the chips bring down the effective FoV to 0.59 sq. deg. The calibra-

tion program⁷ used a combination of small and large 5×5 dithers. Large dithers were used to cover the gaps between chips, monitor low-frequency distortions, and construct a single common reference system for all observations; the small dithers were included to enable independent modeling of the high-frequency residuals of the GD within each chip (more in Libralato et al. 2015). The choice of a 5×5 dither pattern was a compromise to obtain a sufficiently high-precision GD correction in a reasonable amount of telescope time. Ex-

⁷ ESO proposal 488.L-0500(A), PI: Bellini.

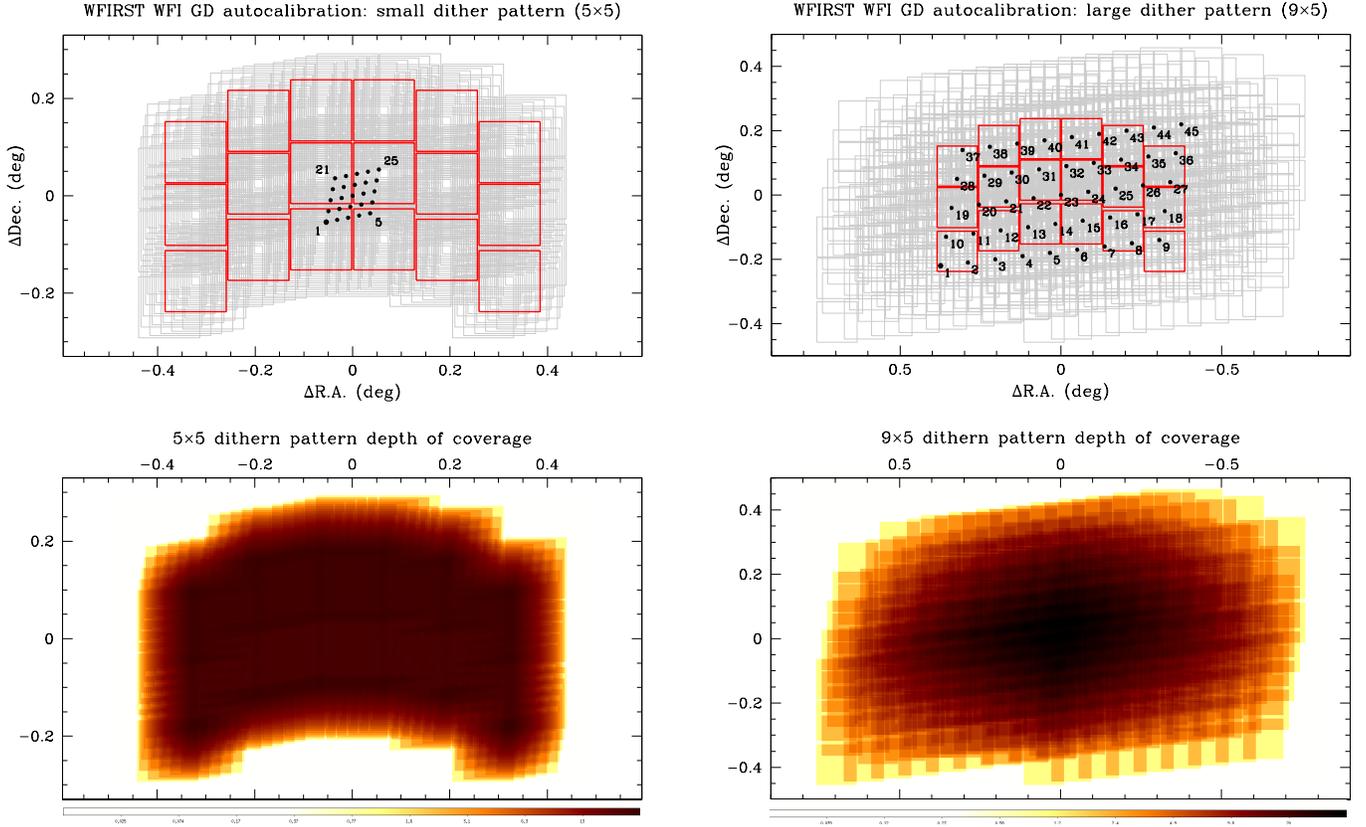


Figure 13. Left column: Example of a small, 5×5 dither pattern that covers each WFI detector from corner to corner. **Right column:** Example of a large, 9×5 dither pattern that covers the entire FoV from corner to corner. **Top row:** Dither pattern layout on-sky, with the center of each dither marked by a black dot, the WFI outline of the central dither shown in red, and outlines of the other dithers in grey). **Bottom row:** Depth-of-coverage map (number of repeat observations as a function of position) for the assumed dither strategies, on a logarithmic scale. See §5.1 for details.

tremely small dithers (from a few subpixels to a few pixels apart) are not strictly needed to characterize the PSF in well-populated star fields, since nature distributes stars randomly with respect to the pixel boundaries (see also [Anderson & King 2000](#)).

Figure 13 shows an example of a possible dither strategy for WFIRST following this plan. In the left column is a plan for a small 5×5 dither pattern that covers each WFI detector from corner to corner. In the top left panel, the black dots mark the center of each of the 25 dithers, with the detector layout of the central dither shown in red and other dithers in grey. The bottom left panel shows the resulting depth-of-coverage map on a logarithmic scale, with a maximum of 25 different images covering the same patch of sky. Most of the map is covered by at least 22 images, but only 12–15 of these come from the same chip, so that the same star is typically imaged in 12–15 different chip locations. The 5×5 pattern never repeats the same shift along the X or Y direction, thus guaranteeing that the same stars will never fall on the same column or row, to minimize the impact of possible detector defects or degeneracies in the distortion solution.

The right column of Figure 13 shows a plan for a 9×5 pattern of large dithers that covers the entire WFI FoV from

corner to corner. Because of the rectangular shape of the WFI FoV, 9 dithers on the X axis are needed to cover the FoV with similar spacing to the 5 dithers along the Y axis. As for the small dithers, the layouts and centers of each pointing are shown on top and the resulting depth-of-coverage map is shown on the bottom. In this case, the layout results in a maximum of 39 different images at the center of the pattern.

The dither patterns shown in Figure 13 all have the same telescope rotation angle, but in order to properly calibrate the skew terms of the distortion, a few observations of the same field at different roll angles would be highly beneficial. It is not obvious to suggest exactly how many of these rotated exposures should be taken, but sampling the full circle every $45\text{--}60^\circ$ should suffice. The total FoV covered by the large dither pattern in Figure 13 allows for the central pointing to be rotated by any angle and still be fully within the covered region.

The proposed dither strategy makes use of 25 small dithers and 45 large dithers, plus 6 or 8 additional pointings (60° or 45° sampling, respectively) to constrain the skew terms, for a total of 76–78 distinct pointings. Experience calibrating the *HST* GD shows that convergence in the GD solution is achieved when stellar positions transformed from one image

to another taken with a different pointing have rms residuals comparable to the stellar centroiding errors. Simulations to assess the precision of the GD correction as a function of the adopted dither strategy are ongoing (Bellini, in preparation) to determine, among other things, whether fewer pointings than the example shown here could be sufficient.

An additional complication introduced by WFIRST’s large FoV is due to the use of tangent-plane projections: adopting the same projection point for images taken more than a few arcminutes apart results in significant positional transformation residuals. The GD calibration therefore has to be carried out on the celestial sphere rather than on any given tangent plane, adding an extra layer of complexity (see also Libralato et al. 2015).

5.1.2. Long-term monitoring of the GD solution

The EML survey is intended to characterize the PSF and fine-tune the GD solution of the WFIRST WFI, possibly including other sources of systematic effects such as intra-pixel sensitivity variations. The current design of the EML survey employs small dithers and fixed rotation angle, suggesting that a satisfactory autocalibration of the GD using only EML survey images will be very hard to achieve. The catalog method could instead be used to fine-tune and monitor the GD solution, probably using *Gaia* as the reference catalog, but the lack of different telescope rotation angles may result in poorly constrained skew-term variations if these are present, as is the case for the WFC of the Advanced Camera for Surveys on *HST*. A preliminary investigation into the possibility of using EML-like simulated images of the Bulge, *Gaia* absolute stellar positions, and WebbPSF-based WFIRST PSF models (Bellini, in preparation) showed that: (i) stellar positions measured by PSF models that do not take into account jitter variations are significantly affected by pixel-phase-like errors (of the order of a few to several hundredths of a pixel); and (ii) the density of *Gaia* stars in the simulated EML survey field (about 5000 stars per chip) is adequate to solve for the GD.

Improved PSF models, either derived independently for each individual exposure or as a function of jitter rms, will address the pixel-phase issues and allow time-monitoring and fine-tuning of the GD solution. Jitter will vary with the reaction wheel speed, particularly at speeds that excite a structural vibration mode. Thus, the level of jitter is expected to change with time even on short time scales, but should have an rms well below 14 mas most of the time. Regardless, because of the time variability, and because excitation of the telescope structure can cause line-of-sight jitter not sensed by the gyroscopes, the guide-star data will be extremely valuable for characterizing the jitter. It would therefore be very useful to downlink the reads of the guide windows together with each exposure, especially given the small amount of additional data involved. Furthermore, the jitter-dependent PSF models obtained for the filters employed in the EML survey will not necessarily apply equally well to other WFIRST filters. If this turns out to be the case, it would be helpful to include settling criteria that allow more stringent jitter rms

constraints when images are taken for the purpose of calibrating and/or monitoring the GD. The current settling criteria include constraints on both position and angular rates, but additional criteria aimed at achieving better stability prior to calibrations would be helpful and should be investigated, since a smaller jitter rms implies smaller pixel-phase errors, helps in removing the degeneracy between centroiding accuracy and GD residuals, and would thereby make calibration more efficient.

Given that the EML survey will make use of only two of the WFI filters, it is important to note that filter elements are known to add significant contributions to the geometric distortion (e.g., Bellini & Bedin 2010; Bellini et al. 2011; Libralato et al. 2014). Exposures taken with the other filters must be used to monitor the time dependency of the GD solution in those filters. In principle *all* WFI exposures can, and probably will, be used for this purpose. This would typically be done with the catalog method, but autocalibration can be applied when properly-dithered exposures are available, so that any variation of the GD solution in all filters can be mitigated to the fullest extent possible.

The possibility of ground calibration of both the distortion and intra-pixel sensitivity variations (see §5.2) should also be considered. As was found with *HST*, a successful astrometric calibration could reasonably be expected to improve WFIRST’s point-source localization, and therefore all astrometry-related measurements, by an order of magnitude. Such an improvement would multiply WFIRST’s reach in distance or velocity sensitivity for astrometry, thereby unlocking an entirely new space for discoveries.

5.2. Pixel-level effects

5.2.1. Quantum efficiency variations

Variations in the quantum efficiency *within* a single pixel can affect the accuracy of localization and therefore the astrometric precision. Hardy et al. (2014) measured the intra-pixel response function for the H2RG detectors to be flown on JWST, which are direct precursors of those planned for WFIRST. They found that the variations in the response per pixel (shown in the left-hand panel of Figure 14) appear to be mainly caused by redistribution of charges from pixel to pixel rather than by variations in pixel sensitivity. The most important effect in redistributing charge between pixels was the diffusion of charge to neighboring pixels, followed by interpixel capacitance (measured between 0 and 4 percent). They also find that the type of small defect visible in Figure 14 occurs in roughly ten percent of pixels. Additional testing of next-generation detectors more closely resembling those to be used for WFIRST is ongoing, but we expect that they will exhibit lower levels of variation.

5.2.2. Placement error

To translate pixel-level effects into predictions for the precision of localization, Michael Shao’s group has made some preliminary measurements of the pixel placement error in H2RG detectors. The “effective” pixel position, which is

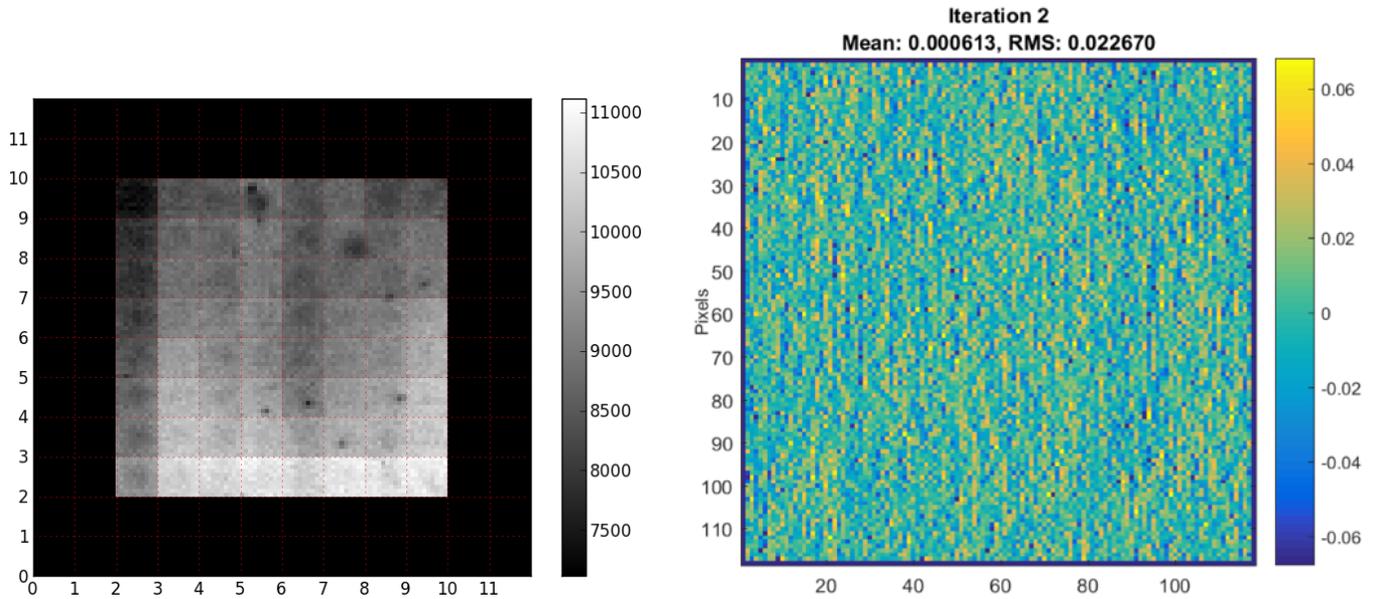


Figure 14. **Left:** Figure 4 of [Hardy et al. \(2014\)](#), showing the pixel response at 650 nm for an 8-pixel-square region of an H2RG detector. **Right:** Pixel offsets in a 128×128 region of an H2RG detector (Figure courtesy Michael Shao).

defined as the location of the centroid of the quantum efficiency (QE) within each pixel, was measured in these tests. Pixel offsets can have multiple causes, including the QE variations within a pixel discussed in [Hardy et al. \(2014\)](#) and in §5.2.1. These tests considered the pixel offset of a 128×128 pixel section of a H2RG detector and measured a rms ~ 0.02 pixel offset error, twice the value assumed in this work for single-exposure precision (see §1.1). The right-hand panel of Figure 14 shows the pixel offset in the X direction for the portion of the H2RG detector that was tested. By eye, the pixel placement errors appear to be random, but the tested region is not large enough to detect larger-scale systematic errors. Currently, the accuracy of pixel position measurement is estimated at $\lesssim 0.5\%$, but the group is in the process of more thoroughly verifying this as well as pushing towards 0.1% accuracy level.⁸

A 0.02 pixel position error corresponds to a single-image astrometric error of about 2 mas, so calibration of this effect will be important if WFIRST wants to deliver astrometry at the level of *Gaia* (10–100 times better) or even LSST precision (5–10 times better). Multiple dithered images and spatial scanning can be used to improve accuracy over the “raw” single-image error, depending on the brightness of the targets.

5.2.3. Ground- versus space-based calibration of subpixel effects

A limitation of calibrating subpixel effects once the telescope is in space is the issue of telescope jitter, which can change the PSF on a time scale of hours. The use of images

of crowded fields to solve for subpixel errors in the detector relies on a stable PSF over a period of time long enough to collect sufficient photons to calibrate subpixel effects. The presence of time-variable telescope jitter prevents this from happening by many orders of magnitude. It is almost certain that the combination of jitter and photon noise will not allow on orbit calibration better than just assuming a perfect detector, given the measured 0.02 pixel errors in H2RG detectors.

It may not be as time-consuming to scan and calibrate this type of variation on the ground as [Hardy et al. \(2014\)](#) discuss in their work. They measured intrapixel QE variations by scanning a spot image across ~ 100 pixels using an extremely time-consuming process. For larger regions containing $\sim 10^4$ pixels, the accuracy of measurements of pixel spacing using this approach will be limited by the positional accuracy of the translation stage used to perform the scan, which will likely be less accurate than the micrometer stage used by [Hardy et al.](#) Therefore, while this is a good approach to measure intrapixel QE, it’s not sufficient for calibrating the dimensional accuracy of a large focal plane array for astrometry. The approach of scanning a spot across a pixel individually would be prohibitively slow for the WFIRST focal plane, which will contain 300 million pixels. Instead, the tests described in § 5.2.2, which consider all pixels simultaneously, could potentially be scaled up to calibrate all the detectors before launch. Current estimates are that this scanning process can calibrate the focal plane array roughly 10^4 – 10^5 times faster than the technique in [Hardy et al.](#) For a ~ 300 megapixel camera like the WFIRST WFI, such a calibration is estimated to take about 1–2 weeks, not including setup time.

⁸ The main sources of error now are spurious fringes due to ghost reflections from, e.g., the dewar window, which have been minimized by tuning the laser over a broad range of wavelengths.

5.2.4. Persistence

Persistence of brightly illuminated regions is known to affect H2RG devices, especially in areas that have been saturated beyond the full-well depth (see the Wide Field Camera 3 Instrument Handbook⁹, section 7.9.4). Characterizations of the persistence for both H2RGs and H4RGs are currently ongoing in several laboratories, which should establish a model for the persistence amplitude and decay time. Such a model can then be implemented to test the biases arising from persistence, which are relevant not only for precision astrometry but also for weak-lensing measurements in the HLS. It remains to be confirmed whether the model from ground-based testing is consistent with the persistence experienced in flight, for which several exposures of suitably bright stars should be sufficient.

If persistence is found to be problematic, a dark filter could be employed to block the incoming light during slews, or slew trajectories could be chosen to avoid bright stars. Experience from *HST* indicates that without a dark filter, persistence during slewing/tracking will be significant for stars brighter than 4th magnitude at the maximum slew rate. Given that there is ~ 1 star brighter than 6th magnitude per 10 square degrees of sky, avoiding these sparsely distributed sources during slewing should be fairly straightforward.

5.2.5. Mitigation strategies

Spatial-scanning and diffraction-spike measurements (§ 2.4) distribute the photons over hundreds or thousands of pixels, and are therefore more robust against pixel-level effects. Spatial scans are also robust against jitter (§ 5.6). A complete summary of requirements for an astrometric spatial scanning mode will be presented in an accompanying report (Casertano et al., in preparation).

5.3. Filters & Chromaticity

Chromatic-induced systematic residuals in the geometric distortion may have a significant impact on WFIRST. The bluest filters of the Wide-Field Camera 3 (WFC3) Ultraviolet-VISible (UVIS) detector on *HST* have chromatic-induced residuals of a few hundredths of a pixel (see Fig. 6 of Bellini et al. 2011). In that case, it is likely that the problem is due to a chromatic effect induced by fused-silica CCD windows within the optical system, which refract blue and red photons differently and have a sharp increase of the refractive index below 4000Å. As a consequence, the F225W, F275W, and F336W filters are the most affected. WFIRST detectors are not based on silica CCDs, but the possibility of similar chromatic-induced systematic effects should be taken into account.

In addition, especially for wide-band filters, the geometric distortion affecting redder and bluer photons is likely to be slightly different due to diffraction. A test using WFC3/UVIS observations in the F606W filter of

blue-horizontal-branch and red-giant-branch stars in ω Cen showed that the measured positions of blue and red stars are off by ~ 0.002 UVIS pixels on average, with respect to their true positions. Filter-dependent residuals could therefore introduce small but still significant color-dependent systematic effects in wide-band WFIRST filters. These filter-dependent systematics are expected to be stable over time, so their calibration should be straightforward.

5.4. Hysteresis in readout electronics

WFIRST should be aware that, on at least one existing instrument with HXRG detectors, the readout electronics are affected by significant hysteresis (Libralato et al. 2014). The effect was first observed by J. Anderson in data taken by the HAWK-I detector on the VLT while observing a standard field to help calibrate JWST and the HAWK-I detector itself, and later by Libralato et al. (2014) in all the fields observed with HAWK-I during the commissioning of the detector (see their Table 2). Figure 15 shows the results of Anderson’s investigation on the top, and of Libralato et al. (2014) on the bottom. In both cases, a periodic signal in the astrometric residuals is observed in the calibration data, believed to be caused by the alternating readout directions for the different amplifiers. A detailed description of the effects of hysteresis on astrometry and how to minimize them is presented in Section 5 of Libralato et al. (2014). For the H2RG chips of the HAWK-I detectors, the hysteresis effect produces ± 0.035 pixels (or ± 3.7 mas) of systematic error, that can be easily modeled and corrected for. The WFIRST detectors will have 32 amplifiers, and scientific full-frame images will make use of all of them. If similar hysteresis effects to those of the H2RG of the HAWK-I camera are present as well in the WFIRST detectors, they will likely be easily minimized as was done for HAWK-I.

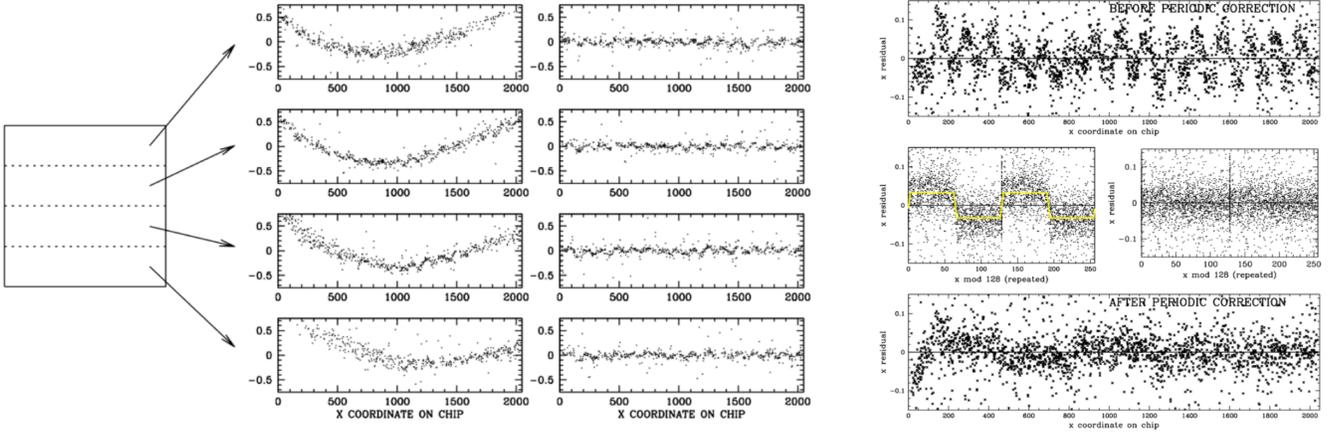
5.5. Scheduling

A small amount of extra attention to scheduling can give great payout for astrometric measurements. As indicated by Equation 2, the ideal scheme for measuring proper motions is to space revisits to the same field as evenly as possible over the longest possible time baseline. Here we discuss a few considerations for the core science and guest observing components of the WFIRST mission.

5.5.1. High-Latitude Survey

The observation of HLS fields is planned to take place over a 5-year baseline, but the detailed schedule is not yet finalized. Starting the HLS with an initial exposure of each field in Year 1 and re-observing at least once in each field as late as Year 5 would produce an astrometric survey that extends *Gaia*’s PM precision ($\sim 25\text{--}50 \mu\text{as yr}^{-1}$) to stars six magnitudes fainter than *Gaia* can reach. However, the current range of schedules considered for the HLS can affect this projected precision by factors of a few. We recommend breaking ties between otherwise equivalent programs by considering the time-distribution of revisits.

⁹ www.stsci.edu/hst/wfc3/documents/handbooks/currentIHB/wfc3_cover.html



Periodic trend before S correction (a)

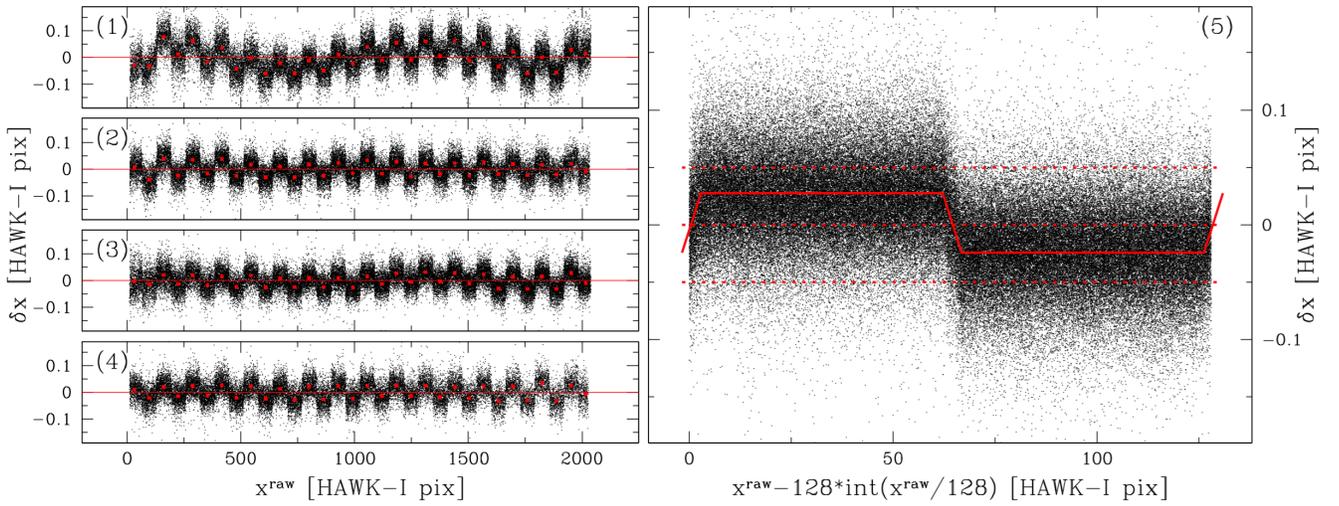


Figure 15. Hysteresis in the HAWK-I detectors on the VLT, precursors to those planned for WFIRST. **Top:** Results from an investigation by Anderson. Four horizontal strips of the detector were analysed, as shown on the far left. The central columns show the distortion as a function of x position in each strip; the vertical scale is in pixels (1 pix = 100 mas). The raw distortion is shown to the left, while to the right are shown the residuals after subtraction of a smooth global polynomial, revealing a high-frequency periodic signal. On the far right are close-up views of the astrometric effect. The top row shows one polynomial-subtracted stripe; while the central row shows the residuals phase-wrapped by 128 columns. The resulting step function (left center row) has a half-amplitude of 0.035 pixel; the right center row shows residuals after subtraction. The unwrapped, corrected residuals are shown in the bottom row. **Bottom:** Results from Libralato et al. (2014). The left-hand panel shows δx residuals for each of the four HAWK-I detectors after the polynomial correction is applied. The right-hand panel shows a periodogram, with a period of 128 pixels, containing all the points plotted on the left, with the median shown as a solid red line. The dashed red lines are at 0 and ± 0.05 HAWK-I pixels (about ± 5.3 mas).

5.5.2. Exoplanet MicroLensing Survey

The EML survey cycles through its 10 target fields at 52-second intervals using the wide filter, with one set in the blue filter every 12 hours. Currently there are six total seasons planned, half in spring and half in fall, in order to measure relative parallaxes with a similar level of accuracy to that of PMs. Some observations will be front-weighted at the start of the survey, but at least one season should be planned for the end of the mission to obtain the longest possible time baselines. Pointing and solar-panel orientation requirements are likely to separate seasons by almost exactly 6 months, which

may cause some complications in calibrating time-dependent PSF effects (see §5.1).

5.5.3. Guest Observing

Since WFIRST is an infrared telescope, it can provide astrometry for regions of the Galaxy, such as the disk plane and bulge, that are completely inaccessible to the current generation of optical astrometric instruments. The EML survey will cover one such region, but to take full advantage of WFIRST's astrometric capabilities, it will be crucial to allow

multi-year proposals from guest observers in order to optimize for PM baselines, as it is currently done for *HST*.

5.6. Jitter

The current WFIRST requirements impose a maximum jitter rms of ≤ 14 mas, far larger than *HST*'s jitter rms of 2–5 mas. Jitter of the size allowed by WFIRST's requirements could have a significant impact on the shape of the PSF. Preliminary simulations of WFIRST's geometric-distortion corrections based on EML-like Bulge images (Bellini, in preparation), which make use of time-constant, spatially-variable, WebbPSF-based library PSF models¹⁰, show that there is significant degeneracy (at the 0.02–0.05 pixel level) between the achievable geometric-distortion correction and pixel-phase errors in stellar positions.

One way to break the degeneracy is to spatially perturb the library PSF in each individual exposure, so as to tailor the library PSF to the particular jitter status of each image. Because of the geometric distortion, jitter-induced PSF variations are expected to affect the PSF of different WFI detectors in a different way. Using this workaround to calibrate the geometric distortion would require images with roughly 20–40 thousand bright and isolated sources, homogeneously distributed across the WFI FoV, in order to map local PSF variations on scales of 500 pixels or so.

Another possibility would be to exploit the enormous number of images in the EML survey to map PSF variations at different jitter rms values, thereby creating a jitter-sampled set of spatially-variable model PSFs. This technique will be limited by the use of only 2 filters for the EML survey, since it is unclear whether the models generated by these two filters will apply equally well to the rest of the filterset.

5.7. Data management

Both pointed and spatially scanned astrometric observations will be constrained by the data downlink rate. For the WFIRST reference mission at L2, the downlink rate is estimated to be about 1.3 TB day⁻¹, barring addition of extra ground stations (Spergel et al. 2015). Particularly for pointed astrometric observations, this limits the number of reads per exposure that can be downloaded. The current plan is to allow configuration of the options for averaging and saving reads, similar to JWST. For astrometry, two options are particularly important:

- Downloads of guide star postage stamps are crucial and inexpensive (0.1% overhead!) for PSF jitter correction.
- The ability to download every read with no coaddition for at least part of the field is especially desirable for spatial scanning observations, as discussed in Section 2.4.

¹⁰ WebbPSF has been adapted to compute accurate WFIRST PSF models. See pythonhosted.org/webbpsf/wfirst.html.

5.8. High-level data products and archive

The combination of a wide field of view, the plentiful availability of reference stars with extremely accurate astrometry from Gaia, and the resolution and stability of a space-based platform makes it possible for WFIRST to achieve extremely accurate absolute astrometry: better than 100 μ as for essentially all WFC imaging products (see §3 and Bellini et al. 2017b). Any field observed more than once by the WFI is thus a potential astrometric field, providing the community with a wealth of opportunities for high-precision astrometry studies in many different domains, including many of the science cases described in §2 of this report. However, for this potential to be realized for all users throughout the mission, *we strongly recommend that the data processing pipeline and the archive incorporate from the outset the necessary elements to obtain, propagate, and maintain in practice the astrometric accuracy that the mission characteristics make possible in principle.*

Specifically, we recommend that the mission consider the possibility of achieving the following goals:

1. The initial (a priori) astrometric information for each image should be based on guide star positions known to Gaia accuracy, together with an accurate WFC geometric distortion model and the analysis of guide-star window data to extract accurate instantaneous GS positions.
2. The WFC geometric distortion model should be verified, and updated if needed, with sufficient frequency to maintain no worse than 100 μ as precision, on the basis of on-orbit data on the geometric stability of the WFC focal plane.
3. The pipeline to generate Level 2 products should include an a posteriori alignment step based on cross-matching sources found in each image with the Gaia catalog; this information should be incorporated as an alternate (preferred) astrometric solution in the image metadata.
4. The accurate astrometry thus determined should be propagated to Level 3 and Level 4 data products.
5. The Archive should have the ability to retain and distribute multiple astrometric solutions for each data product, together with their pedigree and uncertainty. The Archive could also incorporate community-provided astrometric solutions if deemed useful.
6. The Data Management System should have the ability to update the astrometric information for higher-level products when the astrometric solution for the contributing data products is updated.

These recommendations are based in part on our experience with *HST*. The astrometric accuracy available for *HST* data early in the mission was originally limited by the quality of the guide star positions, therefore modest effort was placed

into improving other components of the astrometric fidelity, such as the knowledge and time evolution of the relative positions of instruments and guiders. Now that substantially better positional accuracy is available for guide and reference stars, retrofitting the *HST* pipeline and archive to improve the final absolute astrometric accuracy of *HST* processed data has proven complex and resource-intensive. We recognize that several of these recommendations go beyond the current science requirements and may exceed the baseline capabilities of the mission as currently planned. However, incorporating these considerations, to the extent possible, into the design of the WFIRST Data Processing and Archive systems will greatly improve the quality and accessibility of mission data for astrometry, improving science outcomes in this area and ultimately reducing total development costs when compared with adding similar capabilities at a later time.

5.8.1. Typical astrometry-related queries to object catalog

These queries were submitted to the Archive Working Group as part of their “20–questions” use case development.

AWG-1: Give me positions, IR magnitudes, PMs of all stars in the HLS or EML survey within a color-magnitude box/isochrone cutout.

AWG-1a: Also return LSST optical magnitudes, PMs, etc. for the selected stars.

AWG-2: Return positions, magnitudes, PMs, distances of stars in a specified field that LSST identifies as standard-candle variable stars (e.g. RR Lyrae).

AWG-3: Run a group finder [or other analysis software] I provide on the above data.

AWG-4: Give me all frames from any observing program, and any associated calibration information or data flags, that intersect a defined region on the sky. (I.e. it would be great to be able to re-reduce data from different observations to measure PMs with more frames/longer time baseline)

AWG-5: Give me positions, parallaxes, PMs, plus respective errors of all *Gaia* stars within a WFIRST pointing above a mag threshold.

AWG-6: Extension of above: Give me predicted positions and uncertainties of *Gaia* stars within a WFIRST pointing above a mag threshold at some observation date+time.

AWG-7: Give me positions, PMs, magnitudes, etc. of all XX-type stars within YY pc from the Sun in this ZZ WFIRST pointing.

Acknowledgments. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

RES is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under grant AST-1400989.

REFERENCES

- Agol, E., & Kamionkowski, M. 2002, MNRAS, 334, 553
- Anderson, J., Bedin, L. R., Piotto, G., Yadav, R. S., & Bellini, A. 2006, A&A, 454, 1029
- Anderson, J., & King, I. R. 2000, PASP, 112, 1360
- . 2003, PASP, 115, 113
- . 2004, Multi-filter PSFs and Distortion Corrections for the HRC, Tech. rep.
- . 2006, PSFs, Photometry, and Astronomy for the ACS/WFC, Tech. rep.
- Anderson, J., & van der Marel, R. P. 2010, ApJ, 710, 1032
- Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., et al. 2017, A&A, 599, A50
- Bedin, L. R., Anderson, J., King, I. R., & Piotto, G. 2001, ApJL, 560, L75
- Bellini, A., Anderson, J., & Bedin, L. R. 2011, PASP, 123, 622
- Bellini, A., Anderson, J., Bedin, L. R., et al. 2017a, ApJ, 842, 6
- Bellini, A., & Bedin, L. R. 2009, PASP, 121, 1419
- . 2010, A&A, 517, A34
- Bellini, A., Casertano, S., Nelan, E., & Kruk, J. 2017b, WFIRST Technical Report
- Bellini, A., Milone, A. P., Anderson, J., et al. 2017c, ApJ, 844, 164
- Bellini, A., Anderson, J., van der Marel, R. P., et al. 2014, ApJ, 797, 115
- Bellini, A., Vesperini, E., Piotto, G., et al. 2015, ApJL, 810, L13
- Belokurov, V., Koposov, S. E., Evans, N. W., et al. 2014, MNRAS, 437, 116
- Bianchini, P., van de Ven, G., Norris, M. A., Schinnerer, E., & Varri, A. L. 2016, MNRAS, 458, 3644

- Bochanski, J. J., Willman, B., Caldwell, N., et al. 2014a, *ApJ*, 790, L5
- Bochanski, J. J., Willman, B., West, A. A., Strader, J., & Chomiuk, L. 2014b, *AJ*, 147, 76
- Bovy, J., Erkal, D., & Sanders, J. L. 2017, *MNRAS*, 466, 628
- Brown, W. R. 2015, *ARA&A*, 53, 15
- Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, 635, 931
- Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009, *A&A*, 505, 117
- . 2010, *A&A*, 520, A95
- Casertano, S., Riess, A. G., Anderson, J., et al. 2016, *ApJ*, 825, 11
- Clarkson, W., Sahu, K., Anderson, J., et al. 2008, *ApJ*, 684, 1110
- Clarkson, W. I., Ghez, A. M., Morris, M. R., et al. 2012, *ApJ*, 751, 132
- Dalessandro, E., Saracino, S., Origlia, L., et al. 2016, *ApJ*, 833, 111
- de Bruijne, J. H. J. 2012, *Ap&SS*, 341, 31
- Deason, A. J., Belokurov, V., Evans, N. W., et al. 2012, *MNRAS*, 425, 2840
- D’Ercole, A., Vesperini, E., D’Antona, F., McMillan, S. L. W., & Recchi, S. 2008, *MNRAS*, 391, 825
- Dieball, A., Bedin, L. R., Knigge, C., et al. 2016, *ApJ*, 817, 48
- Diemer, B., & Kravtsov, A. V. 2014, *ApJ*, 789, 1
- Eadie, G. M., & Harris, W. E. 2016, *ApJ*, 829, 108
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016a, *A&A*, 595, A1
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016b, *A&A*, 595, A2
- Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, *ApJ*, 689, 1044
- Gould, A., Huber, D., Penny, M., & Stello, D. 2015, *Journal of Korean Astronomical Society*, 48, 93
- Gratton, R. G., Carretta, E., & Bragaglia, A. 2012, *A&A Rv*, 20, 50
- Hardy, T., Willot, C., & Pazder, J. 2014, *Proc. SPIE*, 9154, 91542D
- Harris, W. E. 1996, *AJ*, 112, 1487
- Hosek, Jr., M. W., Lu, J. R., Anderson, J., et al. 2015, *ApJ*, 813, 27
- Ibata, R. A., Lewis, G. F., & Martin, N. F. 2016, *ApJ*, 819, 1
- Johnson, C. I., & Pilachowski, C. A. 2010, *ApJ*, 722, 1373
- Kallivayalil, N., Wetzell, A. R., Simon, J. D., et al. 2015, *ArXiv e-prints*
- Libralato, M., Bellini, A., Bedin, L. R., et al. 2014, *A&A*, 563, A80
- . 2015, *MNRAS*, 450, 1664
- Lu, J. R., Sinukoff, E., Ofek, E. O., Udalski, A., & Kozłowski, S. 2016, *ApJ*, 830, 41
- Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, 482, 883
- Massari, D., Fiorentino, G., McConnachie, A., et al. 2016, *A&A*, 595, L2
- McConnachie, A. W. 2012, *AJ*, 144, 4
- Melchior, P., Spergel, D., & Lanz, A. 2017, *ArXiv e-prints*
- Milone, A. P., Marino, A. F., Cassisi, S., et al. 2012, *ApJL*, 754, L34
- Milone, A. P., Marino, A. F., Piotto, G., et al. 2015, *MNRAS*, 447, 927
- Milone, A. P., Marino, A. F., Bedin, L. R., et al. 2017, *MNRAS*, 469, 800
- Minniti, D., & Zoccali, M. 2008, in *IAU Symposium*, Vol. 245, *Formation and Evolution of Galaxy Bulges*, ed. M. Bureau, E. Athanassoula, & B. Barbuy, 323–332
- Nelan, E., Holfeltz, S., Chayer, P., & McLean, B., e. a. 2015, *WFIRST Technical Report*
- Norris, J. E., & Da Costa, G. S. 1995, *ApJ*, 447, 680
- Paczyński, B., & Stanek, K. Z. 1998, *ApJL*, 494, L219
- Piotto, G., Villanova, S., Bedin, L. R., et al. 2005, *ApJ*, 621, 777
- Piotto, G., Bedin, L. R., Anderson, J., et al. 2007, *ApJL*, 661, L53
- Piotto, G., Milone, A. P., Bedin, L. R., et al. 2015, *AJ*, 149, 91
- Pourbaix, D., & Jorissen, A. 2000, *A&AS*, 145, 161
- Richer, H. B., Heyl, J., Anderson, J., et al. 2013, *ApJL*, 771, L15
- Riess, A. G., Casertano, S., Anderson, J., MacKenty, J., & Filippenko, A. V. 2014, *ApJ*, 785, 161
- Robin, A. C., Luri, X., Reylé, C., et al. 2012, *A&A*, 543, A100
- Romero-Gómez, M., Figueras, F., Antoja, T., Abedi, H., & Aguilar, L. 2015, *MNRAS*, 447, 218
- Sanderson, R. E., Secunda, A., Johnston, K. V., & Bochanski, J. J. 2017, *MNRAS*, 470, 5014
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Sesar, B., Hernitschek, N., Dierickx, M. I. P., Fardal, M. A., & Rix, H.-W. 2017, *ApJL*, 844, L4
- Sohn, S. T., Besla, G., van der Marel, R. P., et al. 2013, *ApJ*, 768, 139
- Sohn, S. T., van der Marel, R. P., Kallivayalil, N., et al. 2016, *ArXiv e-prints*
- Sohn, S. T., Patel, E., Besla, G., et al. 2017, *ApJ*, 849, 93
- Spergel, D., Gehrels, N., Baltay, C., Bennett, D., et al. 2015, *ArXiv e-prints*
- Spergel, D., Gehrels, N., Breckinridge, J., Donahue, M., et al. 2013, *ArXiv e-prints*
- Stolte, A., Ghez, A. M., Morris, M., et al. 2008, *ApJ*, 675, 1278
- The Theia Collaboration, et al. 2017, *ArXiv e-prints*, arXiv:1707.01348
- Trenti, M., & van der Marel, R. 2013, *MNRAS*, 435, 3272
- van der Marel, R. P., & Kallivayalil, N. 2014, *ApJ*, 781, 121
- Vásquez, S., Zoccali, M., Hill, V., et al. 2013, *A&A*, 555, A91
- Wetzell, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, 827, L23