The Astrophysical Journal, 889:126 (14pp), 2020 February 1 https://doi.org/10.3847/1538-4357/ab629b

Comparing Observed Stellar Kinematics and Surface Densities in a Low-latitude Bulge Field to Galactic Population Synthesis Models

Sean K. Terry[1,2], Richard K. Barry[2], David P. Bennett[2,3], Aparna Bhattacharya[2,3], Jay Anderson[4], and Matthew T. Penny[5]
[1] Department of Physics, Catholic University of America, 620 Michigan Avenue, N.E. Washington, DC 20064, USA; 41terry@cua.edu
[2] Code 667, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[3] Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[5] Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, USA
Received 2019 October 4; revised 2019 December 3; accepted 2019 December 15; published 2020 January 30

Abstract

We present an analysis of Galactic bulge stars from *Hubble Space Telescope* Wide Field Camera 3 observations of the Stanek window ($l, b = [0.25, -2.15]$) from two epochs approximately two years apart. This data set is adjacent to the provisional *Wide-field Infrared Survey Telescope* (*WFIRST*) microlensing field. Proper motions are measured for approximately 115,000 stars down to 28th mag in V band and 25th mag in I band, with accuracies of 0.5 mas yr^{-1} (20 km s^{-1}) at $I \approx 21$. A cut on the longitudinal proper motion μ_l allows us to separate disk and bulge populations and produce bulge-only star counts that are corrected for photometric completeness and efficiency of the proper-motion cut. The kinematic dispersions and surface density in the field are compared to the nearby SWEEPS sight line, finding a marginally larger-than-expected gradient in stellar density. The observed bulge star counts and kinematics are further compared to the Besançon, Galaxia, and GalMod Galactic population synthesis models. We find that most of the models underpredict low-mass bulge stars by ∼33% below the main-sequence turnoff, and upwards of ∼70% at redder J and H wavebands. While considering inaccuracies in the Galactic models, we give implications for the exoplanet yield from the *WFIRST* microlensing mission.

Unified Astronomy Thesaurus concepts: Galactic bulge (2041); Astronomical models (86); Gravitational microlensing (672); Stellar kinematics (1608); Luminosity function (942); HST photometry (756)

1. Introduction

The stellar environment toward the center of the Milky Way is of fundamental importance in our understanding of Galactic evolution, structure, and dynamics. Prior studies have focused on the important task of measuring the red clump (RC) giant stars toward the center of the galaxy (Stanek et al. 1994, 1997; Nataf et al. 2010) as these stars are very good standard candles to probe the Galactic bulge population. The central bulge in our Galaxy is kinematically and chemically distinct from other structures and is of great current interest as new observational techniques and technologies enable deeper and more complete assessment of its components. Until recently, the bulge was thought to be nearly homogeneously old and metal-rich (Ortolani et al. 1995 and references therein). Resolution and characterization of individual stars in the bulge into luminosity functions (LF) have shown this not to be the case (see Calamida et al. 2015 and references therein).

Surface densities are important in the study of stellar populations to disentangle different kinematic groups and to identify coeval subgroups. To be of value as an LF, stars in the sample must be identified as unambiguously point-like with unblended color. A precise, deep LF for stars in the Galactic bulge is difficult to produce, especially at low latitudes, due to several factors including crowding, extinction, and contamination of the sample with disk stars.

Gravitational microlensing requires the exceedingly precise alignment of a foreground mass (the lens) with a background star (the source) (see Gaudi 2012 for a review). A productive microlensing survey thus requires very dense stellar fields to supply a suitable microlensing event rate and optical depth (Mao 2008). Measurement of microlensing parallax, lens star color, and/or detection of finite source effects from a resolved source star are needed to yield the masses—critical for tests of planet formation theories (Suzuki et al. 2016). A thorough understanding of the underlying source-star distribution in the form of a deep LF is of principal importance in the prediction of the microlensing event rate and the exoplanet detection efficiency.

This detailed knowledge will be needed prior to the launch of *Wide-field Infrared Survey Telescope* (*WFIRST*) in order to optimize the microlensing mission's scientific yield (Spergel et al. 2015), and, indeed, forms the basis of the mission's exoplanet success criterion (Yee et al. 2014). Currently predicted detection rates have large uncertainties for several apparent reasons. First, the LFs that exist at present have mostly been conducted in visible light. The *WFIRST* mission will observe the bulge in the near-infrared (near-IR) J and H bands mostly, and the observational fields under consideration are heavily extincted in visible light. Consequently, stellar densities, microlensing event rates, and planet detection efficiencies are all extrapolated from visible observations. The lack of deep near-IR bulge stellar classification studies thus contributes to the uncertainty in expected yields for the mission. Second, while deep bulge luminosity functions do exist, they are focused on sight lines that are distant from projected *WFIRST* fields, and older studies are contaminated by foreground stars (Holtzman et al. 1998; Zoccali et al. 2000). Importantly, Stanek's field, the sight line that is closest to the proposed *WFIRST* exoplanet survey field has been observed using *Hubble Space Telescope* (*HST*) at several epochs and can now be analyzed in the context of the microlensing campaign itself.

Figure 1 shows the most up-to-date planned *WFIRST* microlensing survey fields from Penny et al. (2019, hereafter P19) overlaid

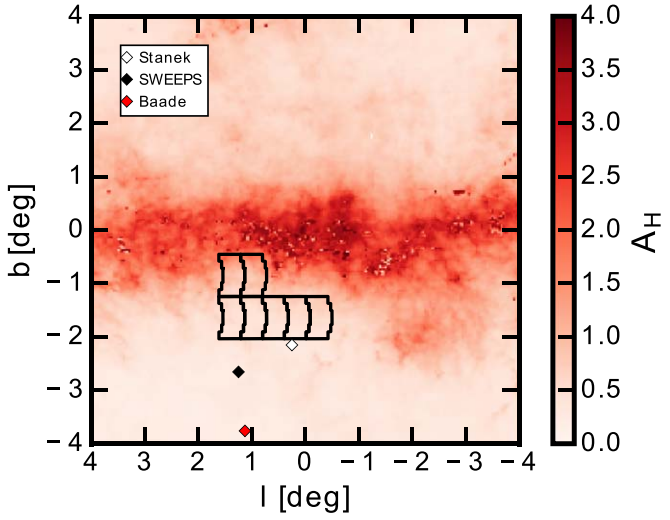

Figure 1. Current Cycle 7 *WFIRST* microlensing survey footprint (Penny et al. 2019), with *H*-band extinction map from Gonzalez et al. (2012) overplotted. The field locations for Stanek, SWEEPS, and Baade have been enlarged for visibility.

on an *H*-band extinction map near the Galactic center (Gonzalez et al. 2012). The white marker indicates the area of stars analyzed in this work and the approximate center of the Stanek field. The black and red markers indicate the approximate locations of the SWEEPS field and Baade's window, respectively.

Several studies have incorporated these *HST* Cycle 17 images in recent years. Age and metallicity estimates of globular clusters observed in the program were performed by Milone et al. (2012), Lagioia et al. (2014), Calamida et al. (2014a), and Baldwin et al. (2016). A broad study of the star formation rate (SFR) and initial mass function (IMF) was also conducted by the GO PI's (Gennaro et al. 2015) using predominantly the OGLE29 field images. The star formation history derived in the Stanek and SWEEPS fields using the Wide Field Camera 3 (WFC3) photometry was conducted by Bernard et al. (2018) and found to be quite similar between the two regions. Additionally, a detailed metallicity study of the fields were made by Renzini et al. (2018). They find that the most metal-poor and metal-rich components are essentially coeval and only a small fraction (~3%) of metal-rich bulge stars are 5 Gyr or younger. There has yet to be a substantial probe of the bulge stellar kinematics and surface densities (which are most important for microlensing) in the Stanek field with regard to the *WFIRST*.

The mission, launching in the mid-2020s, will discover thousands of exoplanets by microlensing (P19) and transits (Montet et al. 2017) in this area of the Galactic bulge. The mission is also expected to discover a small number of Mars-mass free-floating planets (FFPs; Sumi et al. 2011; Barclay et al. 2017; Mroz et al. 2019). A strong understanding of the population of disk (foreground and background) and bulge stars in this window is necessary.

This paper is organized as follows. We discuss the data, its reduction, and photometry in Section 2. In Section 3, we present our analysis of the astrometry and subsequent proper motion (PM) measurements, along with an analysis of the PM dispersion in the field and an estimation of the efficiency of the cut made on the the Galactic longitude μ_l direction. We include a brief Section 3.4 describing the Galactic rotation curve from the measured mean PM values along the sight line. In Section 4, we present corrected star counts in the field and compare these results with several empirical Galactic models. Section 4.2 includes the near-IR analysis of the bulge stars, again with star count results and model comparisons. We finish the paper in Section 5 with a discussion of the results and implications for the *WFIRST* microlensing survey.

Table 1
Bulge Field Properties

Field	l (deg)	b (deg)	A_I (mag)	R_{GC} (kpc)
Stanek	0.25	−2.15	1.284	0.32
SWEEPS	1.25	−2.65	1.004	0.43
Baade	1.06	−3.81	0.743	0.58

2. Observations and Data Reduction

Detailed descriptions of the observations are discussed in Brown et al. (2009, 2010), so only a brief overview is given here. The observations were conducted as part of the WFC3 Galactic Bulge Treasury Program[6] in 2010 June (GO-11664) and 2012 June (GO-12666) in various passbands; F390W, F555W, F814W in the WFC3/UVIS channel and F110W, F160W in the WFC3/IR channel. One of the main goals of the program was to gather deep images of four low-extinction fields toward the center of the Galaxy: OGLE29, Baade's Window, SWEEPS, and the Stanek Field. The program has recently published and subsequently updated (2018 June 5) their version 2 high-level science products (HLSP).[7] The *V* and *I* analysis in this paper began before the version 2 science products were released; therefore, the reduction, photometry, and astrometry presented stands on its own here. Both reductions use the "Kitchen Sink" (KS2) software package by Jay Anderson, as described later in this section. The outputs from both reductions show very similar results with regard to photometric and astrometric accuracy, as well as completeness measurements from artificial star tests (AST).

The planned *WFIRST* microlensing fields are centered at $[l, b] \approx (0.^\circ5, -1.^\circ7)$ with a total coverage area of 1.97 deg^2.[8] The OGLE29, Baade's Window, and SWEEPS fields are all outside of this area, with the Stanek Field ($\sim 2.'7 \times 2.'7$) covering a portion near the edge of the planned footprint at $[l, b] = (0.^\circ25, -2.^\circ15)$. The OGLE29 field $[l, b] = (-6.^\circ75, -4.^\circ72)$ lies substantially further away from the other three bulge fields, so it has been omitted from any comparisons in this paper. Table 1 shows a comparison of the extinction in these fields and their respective projected distances from the Galactic center (Reid et al. 2009). The first set of F555W and F814W observations were taken on 2010 June 27 with seven total exposures in each passband for a total exposure time of 2283 s (F555W) and 2143 s (F814W). Four of the seven exposures in each filter were sub-pixel dithered to allow for high-accuracy astrometry, and therefore, only the dithered frames were used in the subsequent analysis. The second epoch of observations were taken on 2012 June 27 with the F814W filter. Total exposure time for the 2012 images was 1983 s, and again, only the four sub-pixel dithered images were analyzed.

[6] https://archive.stsci.edu/prepds/wfc3bulge/
[7] doi:10.17909/T90K6R
[8] https://wfirst.ipac.caltech.edu/sims/Param_db.html

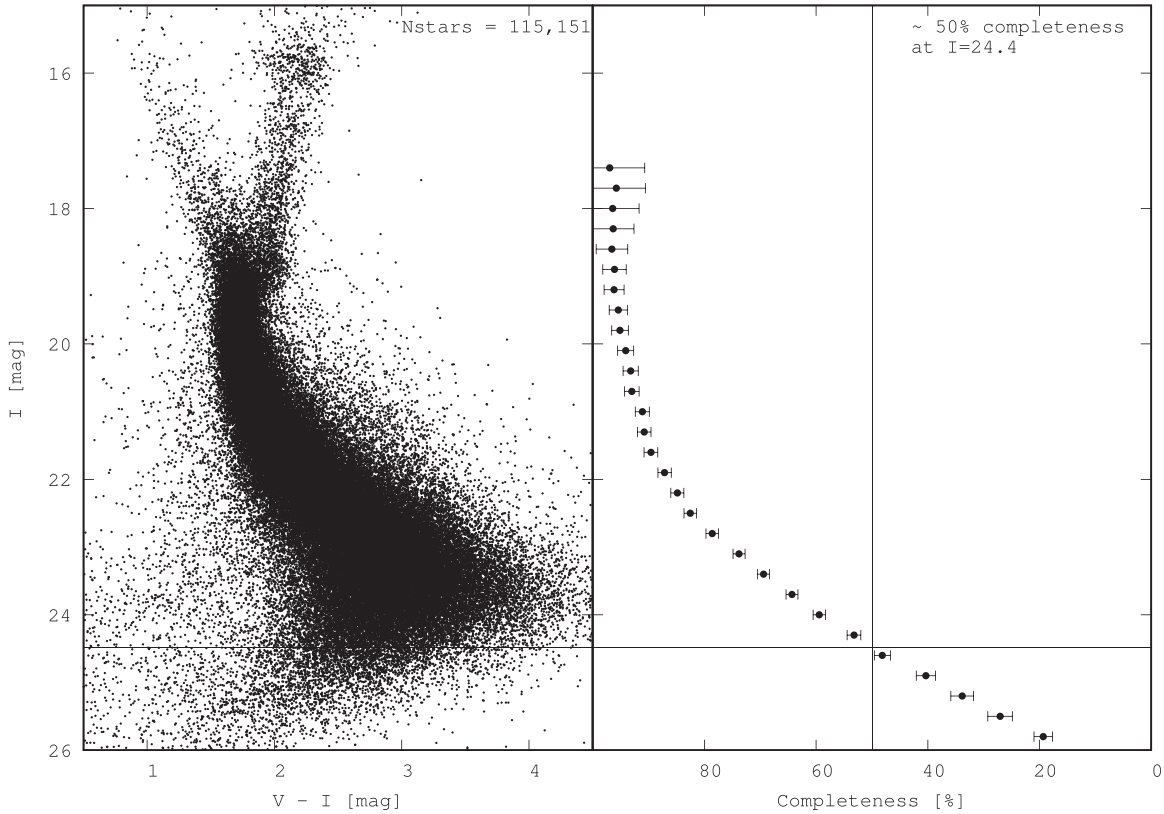

Figure 2. Left panel: CMD of Stanek's Field stars from the 2010–2012 data sets. Right panel: completeness as a function of *I* mag computed from artificial star tests.

The data reduction was performed using a combination of `img2xym` and `KS2` (Anderson & King 2006; Anderson et al. 2008); for the former, a standard 9×9 point-spread function (PSF) grid was used for each filter that accounts for spatial variation across the WFC3/UVIS detector. The routine interpolates the four closest standard PSF's near each star and uses this interpolated PSF for the individual local measurement. A perturbation to the standard PSF's was also included to account for variations in time and location of measurements. This method results in a much higher accuracy for position measurements than a library PSF from the images themselves. A distortion solution is then used to transform the position and flux measurements from the raw frame to a distortion-free sky frame. The distortion accuracy has been measured to be $\sim$1% of a pixel (Anderson & King 2006). In this manner, the positions and fluxes are measured to very high accuracy for both epochs. Additionally, we chose to cull the data set by the quality-of-fit parameter q, which is a measure of the difference in the PSF-fitted flux and the aperture flux in a 5×5 pixel radius centered on each source. We used a conservative rejection threshold of $q \geqslant 0.25$ because of the moderate crowding in these frames. This results in a majority of the remaining stars having photometric rms values $\leqslant$0.1 mag, and astrometric rms values of $\leqslant$1.0 mas in both *V* and *I* band.

The single-pass nature of the routine assumes that all stars are relatively isolated, which is not the case in this moderately crowded field. This places a limit on the fullness of the overall data set. Subsequently, the `KS2` routine was performed on the images. By running a multi-pass reduction routine, we can probe somewhat deeper and acquire well-measured flux and position values for the fainter stars. `KS2` takes as an input the catalog of well-measured bright stars produced from the initial single-pass reduction `img2xym`. The resulting photometry data set has a contribution of well-measured bright stars (F814W $\leqslant$ 17) from the single-pass routine and well-measured multi-pass routine stars (F814W $>$ 17). The instrumental photometry was then transformed to the VEGAMAG system using PySynphot (Astropy Collaboration et al. 2013, 2018). There are chip-based variations in the zero-point values for the WFC3/UVIS detector, which were accounted for during calibration. For the remainder of this text, F814W *I* band (simply "*I*") and F814W *V* band (simply "*V*") are reported throughout.

We combined the photometry and astrometry of the 2010 and 2012 data sets using a method analogous to Equations (1) and (2) to get a final data set of 115,151 stars. The left panel of Figure 2 shows the *I*, $(V-I)$ CMD of Stanek field stars, while the right panel shows the photometric completeness from artificial star tests conducted in Section 2.1. A clear population of foreground disk stars is seen as an un-evolved, blue branch at bright magnitudes. The stars in this branch as well as the older, red giant branch (RGB) above the main-sequence turnoff (MSTO) are used as tracer objects to map the PMs. A clear population of Red-Clump Giants (RC) is also apparent at magnitude $I \approx 15.8$. The redder "shoulder" following the main sequence (MS) at $2.0 < V-I < 4.0$ and I mag > 19 is comprised primarily of foreground (and background) disk stars. These disk stars overlap the bulge population substantially in color–magnitude space and become increasingly ambiguous as the distances to disk stars approach the "beginning" of the bulge population along the sight line. Further descriptions of this and the methodology for generating a "clean" bulge sample are given in Section 3 and subsections within.

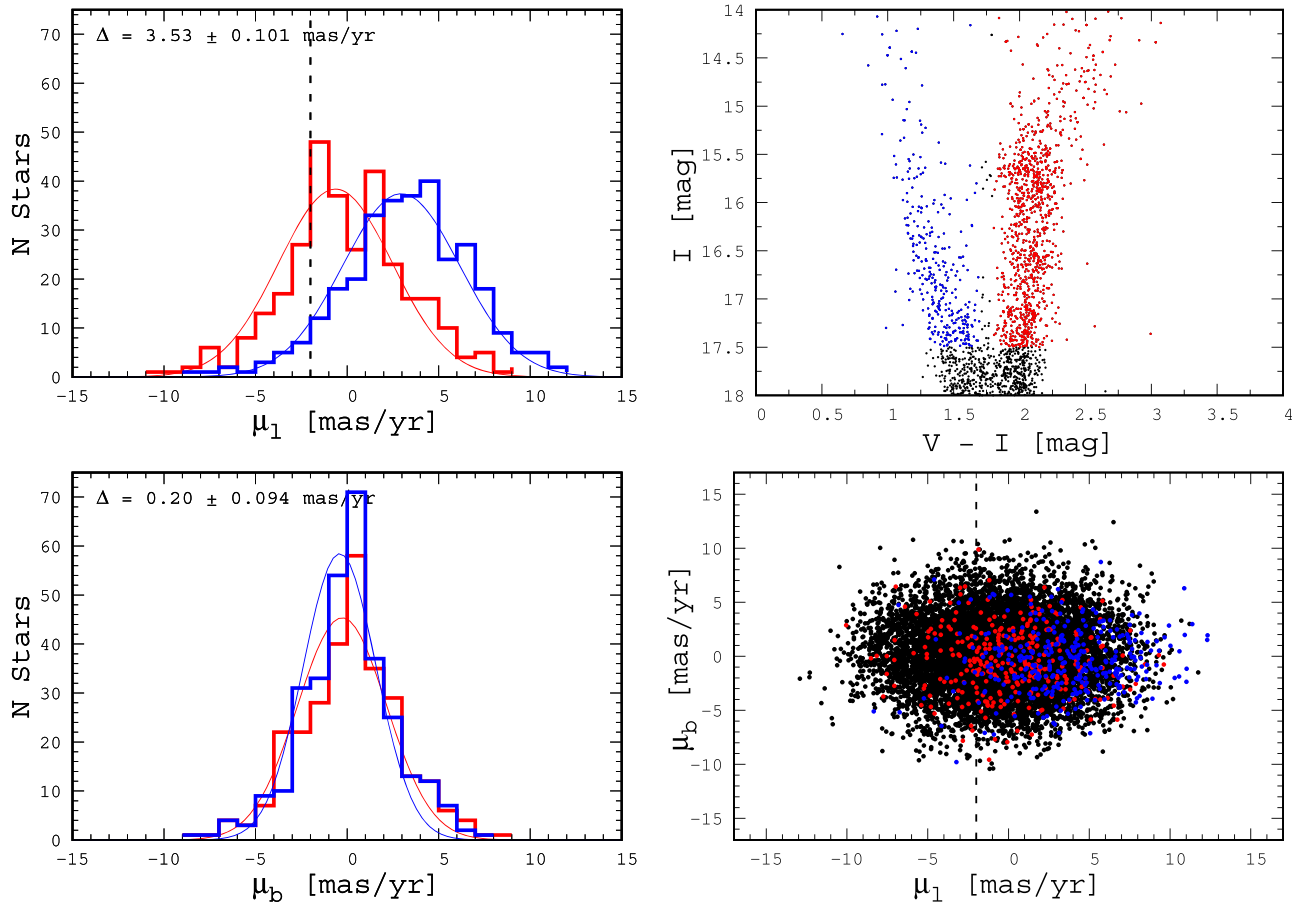

Figure 3. Top-left panel: galactic longitude PM histogram of the bulge (red) and disk (blue) populations as well as the calculated peak separation. The vertical dashed line corresponds to our PM-selection cut at -2 mas yr^{-1}. Bottom-left panel: galactic latitude PM histogram of the same populations. Top-right panel: bulge-dominated (red) and disk-dominated (blue) regions of the CMD chosen for the PM analysis. Bottom-right panel: PM vector point diagram of stars above the MSTO separated into the blue-plume branch (blue) and evolved bulge giant branch (red), all well-measured stars down to 21st mag are plotted in black.

2.1. Artificial Star Tests

Several AST were performed using approximately 200,000 fake stars in order to characterize the photometric completeness, PM accuracy, as well as errors in the reduction routine. The tests were conducted using an artificial star mode within the routine `KS2` to accept an input list of artificial star positions, magnitudes, and colors. The colors and magnitudes of fake stars were estimated by calculating the loci of each point along the MS of the real CMD including Gaussian noise around each source. Artificial stars were added to each image one by one and in a "tile by tile" pattern, adding and measuring several synthetic stars in each tile of size $\sim$120 $\times$ 120 pixels. This method is useful for avoiding major effects from crowding, which can be substantial in deep bulge field images.

The output art-star positions and fluxes were then compared with the input art-star files. A star was considered found if it passed the following criteria:

$$\sqrt{(X_{\rm out} - X_{\rm in})^2 + (Y_{\rm out} - Y_{\rm in})^2} \leqslant 0.50 \text{ pix}, \quad (1)$$

$$|{\rm mag}_{\rm in} - {\rm mag}_{\rm out}| \leqslant 0.50 \text{ mag}, \quad (2)$$

where "in" and "out" denote input star and output star, respectively. Additional art-star tests were performed to estimate the PM accuracy by adopting the method of Calamida et al. (2015). Similar results were found, with somewhat lower accuracy due to the limited number of sub-pixel dithered images relative to SWEEPS. The dispersion of recovered PMs increases with dimmer magnitudes as expected, and with a measured accuracy better than 0.5 mas yr^{-1} (20 km s^{-1} when converting to transverse velocities at bulge distances) at magnitudes brighter than F814W $\approx$ 21 for most stars, with an accuracy $\sim$1.5 mas yr^{-1} (60 km s^{-1}) near the 50% photometric completeness in both V and I bands. This allows for an accurate determination of foreground disk or bulge population stars when culling the data set based on μ_l, down to faint magnitudes but clearly not as faint as SWEEPS, which is one of the deepest data sets to-date.

3. Estimating Proper Motions in the Field

The stars above the MSTO were used to calculate the mean PM of each population. The top-right panel of Figure 3 shows the two regions used; the blue-plume (BP) and RGB regions, respectively. The star counts in these two regions were kept similar to minimize the possible disk contaminants in the RGB sample following the procedure of Clarkson et al. (2008) and C14a. We then calculated the mean PM of each population, finding $[\bar{\mu}_\ell, \bar{\mu}_b] \approx (0.0, -0.3)$ mas yr^{-1} with a dispersion of $[\sigma_\ell, \sigma_b] \approx (3.10, 3.01)$ mas yr^{-1} for the RGB and $[\bar{\mu}_\ell, \bar{\mu}_b] \approx (3.5, -0.5)$ mas yr^{-1} with dispersion $[\sigma_l, \sigma_b] \approx (3.20, 2.14)$ mas yr^{-1} for the BP.

Following this, we adopt a cut at $\mu_l \leqslant -2.0$ mas yr^{-1} to exclude $\sim$85% of disk stars, while keeping $X \sim 35\%$ of total bulge population stars. A more accurate estimation of X can be

Table 2
Proper Motion Dispersions Along Sight Lines

Field Name	Bulge/Disk	l (deg)	b (deg)	σ_l (mas yr^{-1})	σ_b (mas yr^{-1})	σ_l/σ_b	Δt (yr)	N	[a]A_I	References
WFC3 Stanek	Both	0.25	−2.15	3.24 ± 0.02	2.97 ± 0.02	1.09 ± 0.01	2.00	10704	1.28	This work
⋯	Disk	⋯	⋯	3.20 ± 0.11	2.14 ± 0.09	1.50 ± 0.06	⋯	429	⋯	⋯
⋯	Bulge	⋯	⋯	3.10 ± 0.10	3.01 ± 0.09	1.03 ± 0.04	⋯	487	⋯	⋯
ACS SWEEPS	Disk	1.25	−2.65	2.92 ± 0.04	1.70 ± 0.03	1.72 ± 0.02	9.25	2500	1.00	C14b (Calamida et al. 2014b)
⋯	Bulge	⋯	⋯	3.05 ± 0.03	2.94 ± 0.04	1.04 ± 0.01	⋯	2500	⋯	⋯
WFPC2 BW	Both	1.13	−3.77	2.91 ± 0.06	2.51 ± 0.05	1.16 ± 0.03	5.24	1076	0.76	KR02 (Kuijken & Rich 2002)
WFPC2 Sgr-I	Both	1.27	−2.66	3.10 ± 0.06	2.73 ± 0.05	1.14 ± 0.03	⋯	1388	0.96	⋯
97-BLG-41	Both	1.32	−1.95	2.58 ± 0.07	2.13 ± 0.07	1.21 ± 0.04	5.15	612	2.03	K06 (Kozłowski et al. 2006)
98-BLG-6	Both	1.53	−2.13	3.26 ± 0.10	2.79 ± 0.12	1.17 ± 0.05	4.25	670	1.56	⋯
OGLE-II 3	Both	0.11	−1.93	3.40 ± 0.01	3.30 ± 0.02	1.03 ± 0.01	3.91	26763	1.63	R07 (Rattenbury et al. 2007)
OGLE-II 4	Both	0.43	−2.01	3.43 ± 0.02	3.26 ± 0.01	1.05 ± 0.01	⋯	26382	1.49	⋯
OGLE-II 39	Both	0.53	−2.21	3.21 ± 0.01	3.00 ± 0.01	1.07 ± 0.01	⋯	24820	1.48	⋯

Notes. Dispersions for SWEEPS given by Calamida via private communication.
[a] Extinction from Nataf et al. (2013)

made by integrating the Gaussian fit to μ_l (top-left panel of Figure 3), from negative infinity to the PM cut. The integral is of the form:

$$X = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{-2.0} \exp\left(\frac{-(x-\bar{\mu})^2}{2\sigma^2}\right)dx, \quad (3)$$

where $\bar{\mu}$ is the mean PM and σ is the dispersion of the distribution. A detailed estimation of X using this approach is described in Section 3.2. We chose the PM cut of −2.0 mas yr^{-1} by performing a simple optimization in which several thresholds were tested (i.e., −3.0, −2.5, −2.0, −1.5, −1.0 mas yr^{-1}, respectively) to find a resulting sample that best represents the cleaned distribution while maintaining statistical significance (∼20,000 stars passing the rejection criteria). This analysis also considered the effect of cutting based on the calculated PM dispersions for each population as described in Section 3.1. The cutting threshold value of −2.0 mas yr^{-1} was also adopted by the previous studies of Clarkson et al. (2008) and Calamida et al. (2015). Finally, the optimal cutting threshold is likely a function of latitude location near the Galactic plane, but a detailed analysis of this function is beyond the scope of this paper.

3.1. Proper Motion Dispersion

The PM dispersions σ_l, σ_b and their uncertainties were estimated using two techniques. First, each PM distribution was fit by a Gaussian using a χ^2 minimization routine, and then a dispersion and error on the dispersion were calculated from the resulting fit. This method accurately describes the distribution, with non-Gaussianity at most $\leqslant$10% integrated across the distribution. The second method takes a more direct approach, following that of Kozłowski et al. (2006; which is based on Spaenhauer et al. 1992 and references therein):

$$\sigma^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(\mu_i - \bar{\mu})^2 - \frac{1}{n}\sum_{i=1}^{n}\xi_i^2, \quad (4)$$

where μ_i are the individual PMs and ξ_i is the PM error (per coordinate) for the sample of n stars with mean PM $\bar{\mu}$. The error in PM dispersion is then:

$$\xi_\sigma = \left(\frac{\sigma^2}{2n} + \frac{1}{12n^2\sigma^2}\sum_{i=1}^{n}\xi_i^4\right)^{\frac{1}{2}}. \quad (5)$$

The error in PM dispersion relies on the finite size of the sample n and individual PM uncertainties. From these equations, we calculate PM dispersions ($\sigma_l \pm \xi_l$, $\sigma_b \pm \xi_b$) for the bulge-only population, disk-only population, and mixed bulge+disk population down to $I \approx 24$. Both methods of calculation are in agreement, and we adopt the former method due to 15% smaller PM errors on average across the entire magnitude range.

Table 2 and Figure 4 show a comparison of these results to those of Kuijken & Rich (2002) for Sgr-I and BW and Calamida et al. (2014b) for the SWEEPS field. We also include prior well-studied bulge fields from Rattenbury et al. (2007) and Kozłowski et al. (2006) that lie within ∼1.°5 of the Stanek field. The PM dispersion in the Stanek field is marginally larger in the longitudinal direction than most other bulge fields observed, which is to be expected for the nearest sight line to the Galactic plane. The increasing contamination of disk stars at lower latitudes leads to a further spread in the longitudinal dispersion. Additionally, Kozłowski et al. (2006) report a weak but measurable gradient in $\sigma_l(b)$ and $\sigma_b(l)$ that increases with decreasing Galactic latitude and longitude. We also find evidence of this weak gradient in our Figure 4 comparison (top panels). This weak gradient is also apparent in recent *Gaia* DR2 PM measurements (Brown et al. 2018).

Further, the OGLE 97-BLG-41 field is the only sight line with a lower latitude that has a significantly smaller dispersion in both σ_l, σ_b. This field is specifically pointed out by the authors as being intriguing in that it has the lowest measured dispersion, while being the lowest-latitude field of their study. It turns out that the next nearest field in the Kozłowski et al. (2006) study that meets our Table 2 selection criteria, OGLE 98-BLG-6, is also somewhat of an outlier. While both fields are relatively nearby one another, they have the largest difference in dispersion among the 35 sight lines in their sample. One explanation for this may be the increased extinction in these sight lines. From Nataf et al. (2013), the extinction in the 97-BLG-47 field is $A_I = 2.031$ and the extinction in the 98-BLG-6 field is $A_I = 1.560$. The circled data point in the upper-left

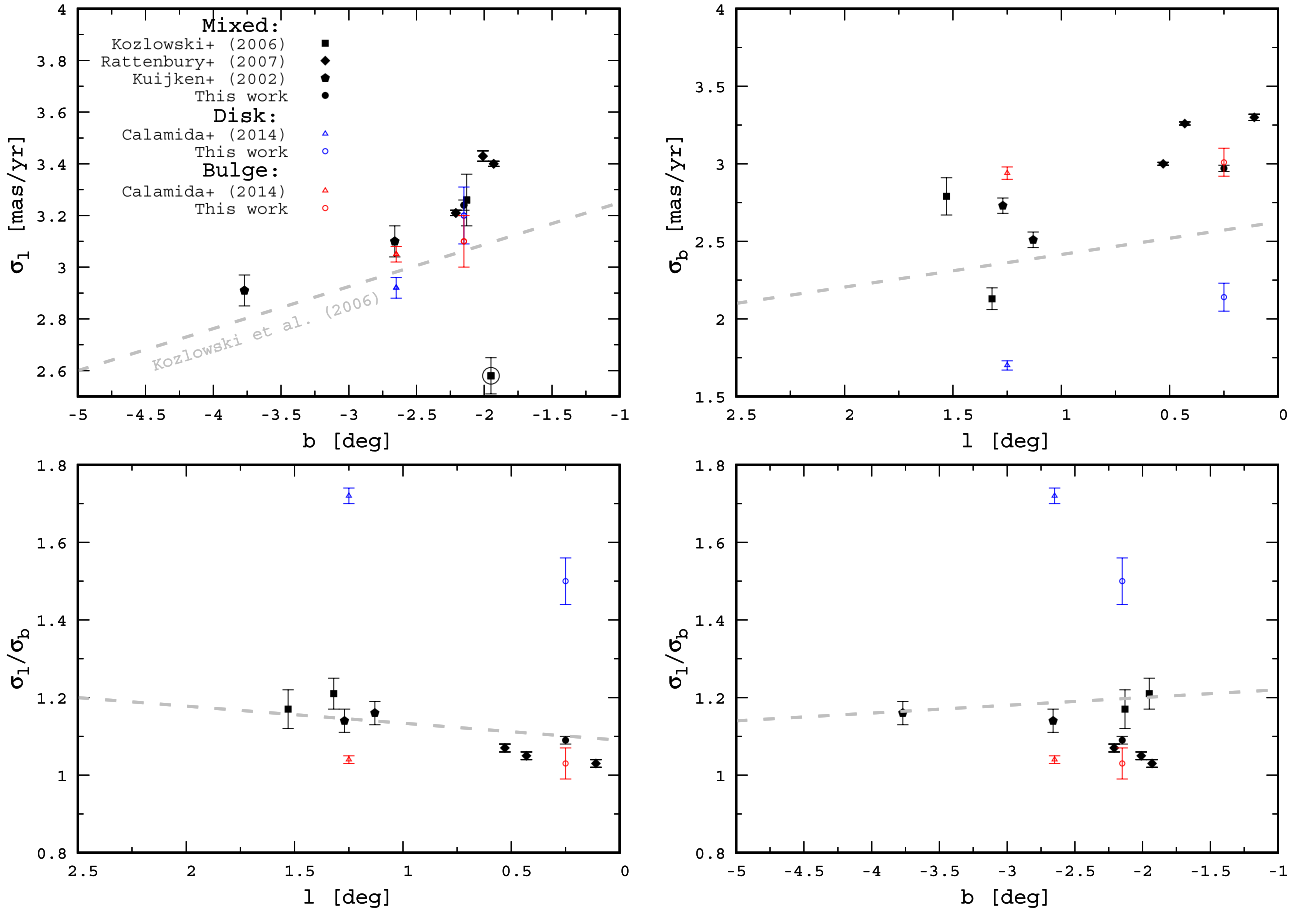

Figure 4. Proper motion dispersions and anisotropy ratios for the sight lines reported in Table 2. Black data points correspond to the "mixed" population. Blue and red data points correspond to the disk and bulge components as measured by Calamida et al. (2014b) and this work. The dashed line shows the linear regression fits from Kozłowski et al. (2006), and the circled data point in the top-left panel is the outlier, 97-BLG-41.

panel of Figure 4 shows the 97-BLG-41 outlier. The authors decided not to use this data point in their linear regression fit for $\sigma_l(b)$.

Finally, it is worth noting the Kozlowski dispersion measurements reported were performed on the mixed bulge +disk population in all sight lines. Their sample was limited to the magnitude range $18.0 < I_{\rm F814W} < 21.5$, which is dominated by bulge MS stars near the turnoff, but will undoubtedly still be contaminated by disk stars. The contamination should become most severe in their lowest-latitude fields, which would seem at odds with their findings for the two fields described above. It is difficult to determine whether our dispersion results give further evidence for a subtle gradient in the rotational velocity of the bulge, as some of the previous studies have suggested.

3.2. Proper-motion Cutting Efficiency

In order to scale the cleaned bulge LF up to a "full" bulge LF, it is important to estimate a correction for the efficiency of the PM cut that is being made. Calamida et al. (2015) do not report any corrections for their SWEEPS PM-cut efficiency; however, P19 subsequently calculated an efficiency of the (bulge-only) PM-cut of 34 percent based on an estimate similar to that described in Equation (3) of this paper. The value of $X \sim 0.34$ is used to scale the entire cleaned SWEEPS LF up to the "full bulge" LF. For the Stanek LF, we calculated an efficiency factor in two ways. In the first method, we applied Equation (3) to the RGB population stars from the tracer region described earlier in this section. We find an efficiency factor $X \sim 0.32$. In the second method, we applied the same equation to the fully mixed (bulge + disk) PM-distribution as a function of I magnitude from approximately the red clump down to the 50% completeness threshold of $I \approx 24.4$. This results in an efficiency curve that covers each relevant magnitude bin. Figure 5 shows the PM-cutting efficiency curve for X, which increases from ∼0.23 at the brightest magnitudes to ∼0.36 at faint magnitudes. This curve flattens at $X \sim 0.31$ between intermediate magnitudes 19.5–23 corresponding to most of the observed MS, and is due to the two populations being fully mixed and indistinguishable from one another. The efficiency covering this flat region is consistent with the value calculated via method one. Method one will clearly under-correct the LF at brighter magnitudes and over-correct at the faint magnitudes, by upwards of ∼8%. To avoid introducing any additional inaccuracies, we adopt the second method for our PM-efficiency scaling.

Table 3 gives the details of each bin and the convolved Gaussian sigma from combining the PM distribution and the individual PM-error distributions. Again, the propagated errors are smaller by ∼15% using Equation (3). The fit to the distribution broadens as expected and increases at fainter bins, while the contribution from the PM-error distribution has a measurable but small effect on most of the full bulge LF. The residual disk star contamination increases at fainter magnitude bins, which is nearly

Table 3
Proper-motion Cutting Efficiency

Bin (mag)	$\langle \mu_l \rangle$ (mas yr^{-1})	σ_l (mas yr^{-1})	$\sigma_l + \xi_l$ (mas yr^{-1})	X
16–17	0.73 ± 0.12	3.749 ± 0.119	3.751 ± 0.119	0.233 ± 0.016
17–18	1.05 ± 0.11	3.875 ± 0.110	3.877 ± 0.110	0.216 ± 0.014
18–19	0.27 ± 0.06	3.726 ± 0.058	3.728 ± 0.058	0.271 ± 0.008
19–20	−0.19 ± 0.04	3.550 ± 0.042	3.555 ± 0.042	0.305 ± 0.006
20–21	−0.24 ± 0.04	3.517 ± 0.044	3.527 ± 0.044	0.309 ± 0.007
21–22	−0.30 ± 0.05	3.454 ± 0.050	3.479 ± 0.051	0.311 ± 0.008
22–23	−0.33 ± 0.05	3.482 ± 0.038	3.568 ± 0.043	0.316 ± 0.008
23–24	−0.56 ± 0.07	3.767 ± 0.064	3.932 ± 0.071	0.351 ± 0.010
24–25	−0.28 ± 0.26	5.123 ± 0.262	6.122 ± 0.402	0.358 ± 0.026

Note. $\sigma_l + \xi_l$ is the convolution of both Gaussian distributions, with associated uncertainty estimated by summing the Gaussian errors in quadrature. A larger X value corresponds to smaller relative scaling.

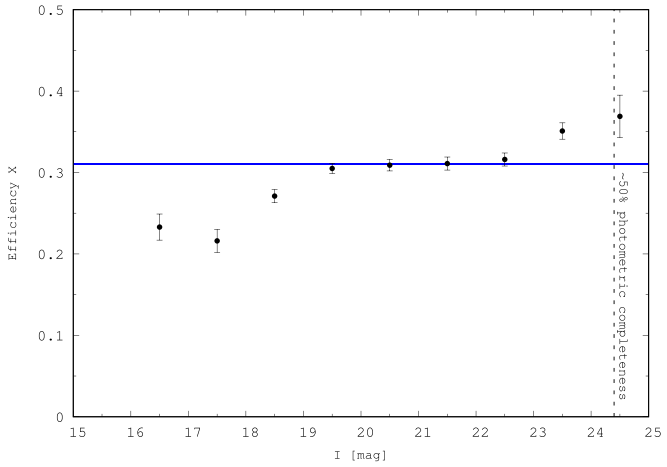

Figure 5. PM-cutting efficiency as a function of I magnitude. The blue horizontal line represents the efficiency measured from the tracer region above the MSTO (Figure 3).

impossible to accurately calculate. However, using estimates from Gennaro et al. (2015), we place a rough constraint of 0.5%–3.0% residual contamination that spans the range $I = 16$–25. Lastly, standard error propagation techniques are used for the values reported in Table 3, and associated error for the convolved Gaussian $\sigma_l + \xi_l$ is computed by summing the prior Gaussian errors in quadrature.

3.3. Anisotropy Ratio

The anisotropy ratio was also calculated for the Stanek field and found to be $\sigma_\ell/\sigma_b = 1.03$ for Bulge-only, $\sigma_\ell/\sigma_b = 1.50$ for disk-only, and $\sigma_\ell/\sigma_b = 1.09$ for the mixed populations, respectively. All are accurately measured despite the conservative number of proxy stars used. Spaenhauer et al. (1992), Kuijken & Rich (2002), Kozłowski et al. (2006), and others suggest an anisotropy ratio measurably larger than 1 due to the rotating bulge-bar component. Our results show a pure-bulge anisotropy ratio closer to 1, which is in contrast with the prior studies. However, our "mixed" population anisotropy measurement shows some evidence of rotation, although, this is likely due to the additional contribution of the foreground disk population. We caution that it may be difficult to draw any conclusive trends between the sight lines due to a majority of the prior fields only having measurements of the mixed bulge +disk populations rather than any pure component.

While this discrepancy may not seem severe, as Kuijken & Rich (2002) and Kozłowski et al. (2006) have stated, there are clearly measurable differences between our "Bulge" only dispersions and "Both" mixed dispersions in Table 2. Recall that at bulge distances, a PM of ∼3.0 mas yr^{-1} corresponds to a transverse velocity of ∼115 km s^{-1}. The smaller anisotropy measured for the Stanek field "Bulge" and/or "Both" populations is in agreement with the trend reported in Kozłowski et al. (2006), showing a decrease in anisotropy nearer to the minor axis (i.e., $\ell = 0$). Our measured dependence of σ_l/σ_b on b does not agree as well with the fields in these prior studies; however, it is marginally consistent with the observed σ_l/σ_b scatter.

We remind the reader that all of these fields reside in a relatively small area of the bulge, within ∼1.°5 of the minor axis and within ∼3° of the plane. Our measured dispersion gradient also agrees with the recent study of Clarke et al. (2019) who analyzed PM data from the Vista Variables in the Via Lactea (VVV) and the *Gaia* DR2 survey. It's important to emphasize the dependence of σ_ℓ on b and σ_b on ℓ, which is apparent when comparing dispersed fields against the lowest longitude field in the set by far (Stanek at $l = 0.25$). More detailed explanations of this dependence and other anisotropy descriptions are given in Clarkson et al. (2008) and references therein.

3.4. Rotation Curve

Following Kuijken & Rich (2002) and Kozłowski et al. (2006), a crude distance modulus can be calculated by removing the slope of the CMD and cutting a cross-section along the new de-colored MS. This results in a simple relative distance indicator for each star in the set, which follows the form:

$$M^* = m_I - 2(m_V - m_I). \quad (6)$$

Stars are binned by distance indicator and plotted against their mean PMs and PM dispersions, which are shown in Figure 6. As expected, the kinematics along the line of sight describe the near-side foreground disk stars, transitioning through the bulge population and very likely beyond the bulge to the far-side stars. This is detailing a rotation curve for the Milky Way along this sight line in terms of the relative PM and their dispersions, with limited resolution at the faintest magnitudes corresponding to backside disk star populations. The amount of contamination by disk star rotation is not large enough to significantly influence the mean velocity for all stars in a given distance indicator bin.

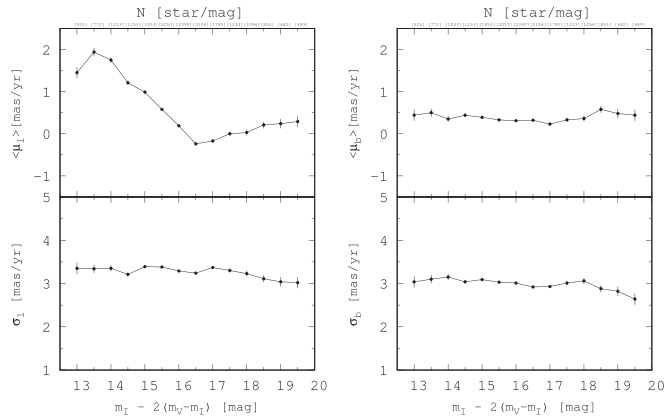

Figure 6. Left panel: galactic longitude mean PMs and dispersions of Stanek field stars as a function of the distance indicator. Right panel: galactic latitude component of the motions.

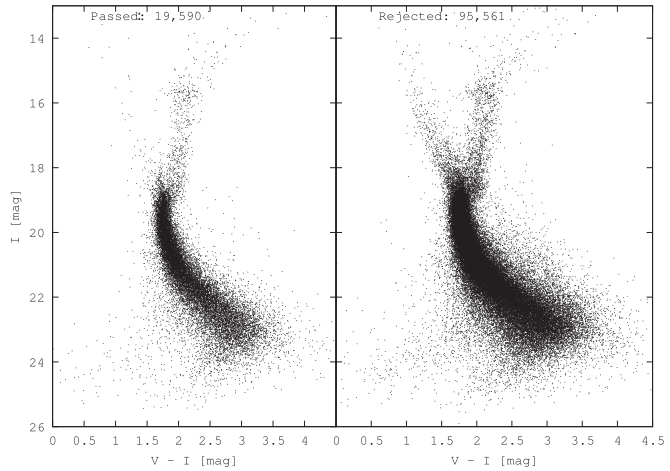

Figure 7. Left panel: CMD of stars that passed the PM cut. Right panel: CMD of rejected stars from the cut.

For detailed descriptions of the velocity profile and kinematics, we refer the reader to Zhao et al. (1994), Izumiura et al. (1995), Zhao et al. (1995), and, more recently, Clarkson et al. (2018).

4. Bulge Star Counts

The completeness corrections and PM-cutting efficiency leads to a proper accounting of all bulge sources along this sight line. Figure 7 shows the cleaned CMD after PM cuts have been applied (left panel), along with stars that are rejected as a result of the cut (right panel). Such a large rejection threshold is needed to ensure the final data set is sufficiently cleaned of foreground star contamination. The cleaned CMD of the Stanek field shows some evidence of a population of younger bulge BSS just brighter and bluer than the MSTO. This would seem to verify the results of Clarkson et al. (2011), although we do not conduct a further analysis in the current paper. There is also a small amount of residual contamination from foreground stars that remains in the cleaned data set. These sources have high enough (e.g., <-2 mas yr^{-1}) PM to pass the cut; they are likely bright foreground disk stars with sporadically high PM, or large bright stars counter-rotating on the far side of the bulge/disk. Additionally, there is evidence of the WD cooling

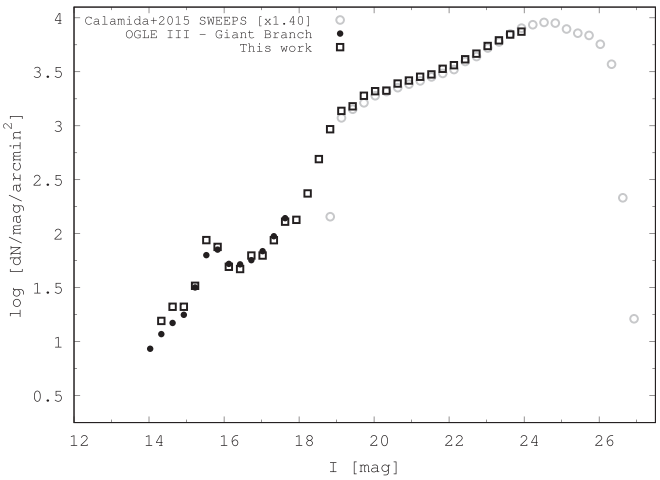

Figure 8. Stanek field bulge star counts compared to Calamida et al. (2015) SWEEPS field bulge star counts (gray). RGB stars from OGLE-III (Szymański et al. 2011) in the Stanek field are plotted as solid black points. The SWEEPS counts are scaled by ∼40% to match the larger surface density of the Stanek field.

sequence in the bulge previously detected by C14a that reside in a similar location on the cleaned CMD ($0 < V - I < 1.5$ and $23 < I < 25$).

The bulge-only star counts are presented in Figure 8, after being corrected for photometric completeness and PM-cutting efficiency. Data with low completeness ($\lesssim$50% at $I \gtrsim 24$) are not plotted. Star counts from the OGLE-III (Szymański et al. 2011) fields overlapping the *HST* Stanek coverage are shown covering primarily the RC at $I \sim 15.7$. The OGLE-III stars have not been completeness-corrected. The very deep SWEEPS field bulge star counts of Calamida et al. (2015) are shown in Figure 8 as well. The SWEEPS star counts are found to be approximately 40% less than Stanek integrated over the magnitude range $I = 19.5$–23.5. The increased surface density between these fields is expected but slightly larger than values estimated by Wegg & Gerhard (2013), who show projections of the fiducial density measurements of RC stars identified in DR1 of the VVV survey (Saito et al. 2012). The ∼7% larger-than-expected surface density we find is due to several factors. A higher residual disk star contamination in our pure-bulge sample is likely caused by a marginally lower cleaning efficiency, which can be seen as a difference between the disk population σ_l reported on Table 2. As described earlier, the nearer location of the Stanek field to $\ell = 0$, $b = 0$ cause a more severe contamination. Recall the center of the Stanek field is $\sim 1.^{\circ}0$ closer to the minor axis and $\sim 0.^{\circ}5$ closer to the major axis relative to the center of SWEEPS.

4.1. RC/MS Ratio

To confirm there is no intrinsic difference in the bulge-only population of evolved RGB stars and MS stars between Stanek and SWEEPS, we measured the ratio of the surface densities for the RC and MS range in the two fields. For the RC measurement in the Stanek field, we used the `BLG101.3.map` catalog from OGLE-III. This star list is nearest to the Stanek *HST* field with significant overlap. We measure the centroid of the RC in this field to be $I_{\rm rc} = 15.721$ and $(V-I)_{\rm rc} = 2.099$, and from here we measure the RC star count in the window by integrating over the range $I = 15.571$–15.871. The MS star

Table 4
Bulge Parameters for Each Population Synthesis Model

Reference	Age (Gyr)	Fe/H (dex)	IMF ($dn/dm \propto m^{-\alpha}$)	SFR	Bar Angle (°)
Besançon 1612	10.0	0.00 ± 0.40	$\alpha = 2.3$, $m \geqslant 0.7M_{\odot}$ $\alpha = 0.5$, $0.7M_{\odot} > m \geqslant 0.15M_{\odot}$	single burst	12.9
Galaxia 0.7.2	10.0	0.00 ± 0.40	$\alpha = 2.3$, $m \geqslant 0.7M_{\odot}$ $\alpha = 0.5$, $0.7M_{\odot} > m \geqslant 0.15M_{\odot}$ $\alpha = 1.3$, $0.15M_{\odot} > m \geqslant 0.07M_{\odot}$	single burst	12.9
GalMod 18.19	6.0−12.0	$0.00^{+0.3}_{-0.4}$	$\alpha = 2.3$, $m \geqslant 1.0M_{\odot}$ $\alpha = 2.7$, $1.0M_{\odot} > m \geqslant 0.5M_{\odot}$ $\alpha = 1.8$, $0.5M_{\odot} > m \geqslant 0.16M_{\odot}$	Rosin (1933)	27.9

count is measured by integrating the *HST* WFC3 star counts over $I = 21$–23. Finally, we measure $\mathrm{RC/MS_{Stanek}} = 0.020$ for the raw counts.

Similarly, for the SWEEPS field, we used `BLG104.5.map` from OGLE-III for the RC measurements. The RC centroid was measured to be $I_{\mathrm{rc}} = 15.376$ and $(V - I)_{\mathrm{rc}} = 1.843$. To account for the difference in the I magnitude location of the RC between the fields ($\Delta I = 0.345$), the MS magnitude range integrated over for SWEEPS was $I = 20.65$–22.65. We measure $\mathrm{RC/MS_{Sweeps}} = 0.021$ for the raw counts. Our results show that the ratios for both fields are consistent, and they confirm that there is indeed no significant intrinsic variation in population types among these two bulge sight lines.

4.2. Near-IR Star Counts

The version 2 science products from the WFC3 Bulge Treasury Program (Brown et al. 2009) were used to analyze the J(F110W) and H(F160W) photometry, astrometry, completeness, and surface density. The methods described in Section 2 were used for these redder data, with the exception of the photometric calibration. The photometry given by the version 2 data products is in the STMAG system; thus, we converted these magnitudes to VEGAMAG using the zero-points from Deustua et al. (2017). The field of view for the WFC3/IR camera is smaller than UVIS, at a scale of $136'' \times 123''$, so additional offsets were performed to provide the IR camera full coverage of the UVIS area ($162'' \times 162''$). The total exposure time for the J-band data was 1255 s, and 1638 s for the H-band data.

The right panel of Figure 10 shows the H-band LF for bulge stars in the Stanek field along with three simulated LF's (normalized to dN mag^{-1} arcmin^{-2} and transformed to VEGAMAG) from the models. The slope of the LF below the MSTO is steeper than the I-band results as expected for a bulge with a significant amount of low-mass dwarfs. Of the three population synthesis models, GalMod most closely predicts the surface density in this near-IR band; however, the higher low-mass cutoff of the IMF precludes the model from probing the dimmest population stars. From this, we are unable to deduce whether the modeled LF "turns-up" (like Galaxia), or "turns-down" (like Besançon) at the faintest magnitudes. A future update to the GalMod software will include the MIST stellar tracks (Choi et al. 2016; Dotter 2016), which will allow for much deeper CMD's and LF's (S. Pasetto 2019, private communication). Lastly, Table 4 shows comparisons between relevant parameters for each population synthesis model. The major difference again is the low-mass contribution in the IMF's and the bar angle.

4.3. Comparison with Bulge Population Synthesis Models

As a basis for estimating *WFIRST* exoplanet yields, population synthesis models are used to generate stellar surface densities and microlensing optical depths, which in turn can be used to estimate microlensing event rates. Based on the event rate, a detection efficiency for microlensing planets can then be derived while placing limits (via the models) on expected planet yields. The better a model can accurately describe surface densities in very crowded and highly extincted fields, the more accurate the estimations of planet detection efficiencies can be (P19). The P19 simulation study is a very detailed project that has estimated the planet detection efficiency and expected exoplanet yields for the current Cycle 7 *WFIRST* design. P19 compared only one Galactic model in their study and were subsequently required to make adjustments to the yields in order to match observations after discrepancies in the model were considered. We expand on this part of the P19 study by comparing several popular Galactic models to the Stanek field observations.

First, the models analyzed here are all publicly available either by direct access via web interface or download and compile. Second, the models all generally describe the central bulge as a boxy triaxial bar shape (Dwek et al. 1995) with two models using a bulge-bar angle of ∼12°–13° with respect to the Sun-Galactic line and the third model assumes an angle of ∼28°. They all use relatively similar scale lengths, each incorporates disk and bulge kinematics, and they include some form of prescription for the warp and flare of the thin disk and bulge, respectively. The built-in photometric system chosen for each model is the Johnson Cousins (*UBVRIJHKL*) system. As we show in the following subsections, each model has its strengths and weaknesses; therefore, it is a worthy exercise to perform the comparison of these sophisticated models to a set of real data. Lastly, it is important to keep in mind the following caveat: the Galactic bulge-bar is arguably the most complex MW component, and it is very likely impossible for any model of our galaxy to perfectly simulate all of the complexities that lie within such a complicated dynamical system.

4.3.1. Besançon

Version 1106 of the Besançon Galactic Model (Robin et al. 2003, 2012, hereafter BGM1106) was used by P19 for their

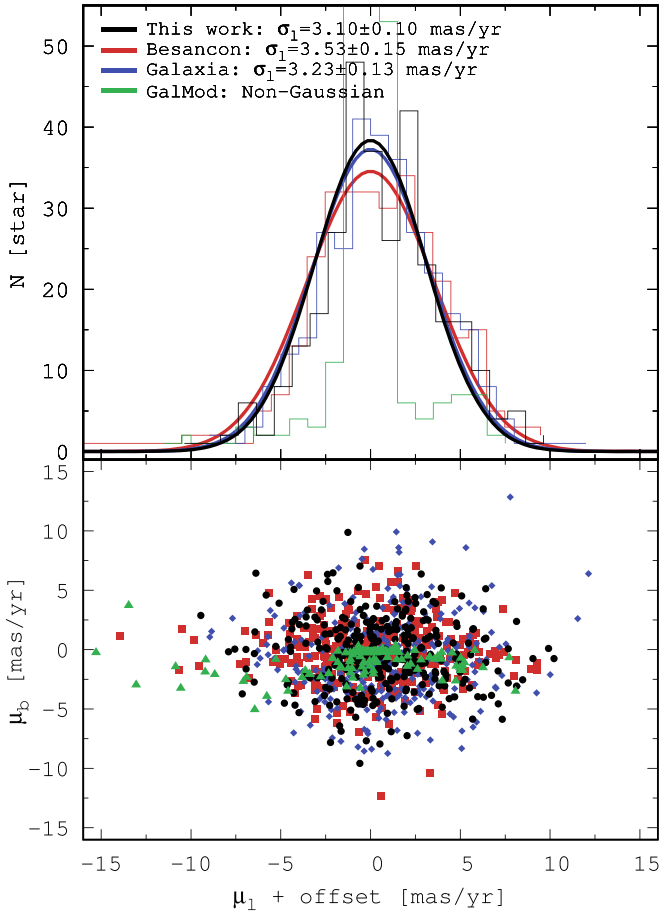

Figure 9. Top panel: PM histogram for the observed Stanek field (black curve) mean longitudinal component, μ_l, compared to the three population synthesis models. Bottom panel: vector point diagram of bulge proxy stars selected from the RGB in each sample. Selection size is ∼400 stars for each sample.

analysis and subsequent microlensing event rate and detection efficiency estimation. They give a detailed description of BGM1106 in Penny et al. (2013), so the details given here will be limited to the differences between the P19 model and the current model used in this paper (version 1612, hereafter BGM1612[9]). A full description of the uncertainties in BGM1106 can be found in Section 6.2 of P19.

BGM1612 uses a slightly larger bulge-bar angle of 12.°9, a thin-disk density law (Einasto 1979) with an 8% smaller scale length of 2.17 kpc for the old stars, and the same scale length of 5 kpc for the young stars. There is a central hole in the disk with a scale length of 1.33 kpc for old stars and 3 kpc for young stars, which is virtually unchanged between versions. P19 measure the BGM1106 kinematics toward the SWEEPS field, and find PM dispersions roughly similar to the detailed PM study of Clarkson et al. (2008), with the exception of the Galactic longitude dispersion σ_l for the red-bulge population stars. The authors find the dispersion from the simulation to be larger than the observed value by a factor of 1.73 ± 0.12, which leads to microlensing event timescales that are too short. The too-fast kinematics in the model have since been corrected in the current version, and we find only a marginal increase in the longitudinal PM dispersion between the current model and observations. As shown in Figure 9, BGM1612 has a larger dispersion by a factor of 1.14 ± 0.05. Note the circular velocity (at the distance of the Sun) that the current version uses is $V_{\rm LSR} = 244.6$ km s^{-1}. Further, the PM's reported by BGM1612 and the subsequent models are in an absolute frame, whereas the Stanek *HST* measurements are made in an arbitrary frame since there is no absolute reference to measure the motions against. We apply a simple offset to the Besançon model's bulge and disk proxy populations to correct for this.

The bulge-bar angle of 12.°9 that the model assumes is significantly smaller than some prior observational studies (Stanek et al. 1994; Rattenbury et al. 2007; Cao et al. 2013) that measure a bar angle ∼25°–45°. P19 point out that a smaller bar angle leads to a larger spread of bulge stars along the line of sight. This ultimately results in larger Einstein radii and event timescales for bulge–bulge lensing and, to a lesser extent, bulge-disk lensing. The Einstein radius depends on the relative distance to the source and lens by:

$$r_{\rm E} = \sqrt{\frac{4G}{c^2} M_l D_l \left(1 - \frac{D_l}{D_s}\right)}, \quad (7)$$

where G and c are the fundamental constants for gravity and speed of light, respectively, M_l is the mass of the lens, and D_l and D_s are the distance to the lens and source, respectively. It is also clear with a smaller bar angle, the ratio of bulge lenses to disk lenses will be larger. Robin et al. (2012) attempt to reconcile this discrepancy by comparing different versions of their model with varying bar angles to determine if they get a more favored fit. They conclude that a larger bar angle gives a higher likelihood at locations farther from the minor axis and close to the plane, but a lower likelihood nearer to the minor axis and higher latitude. Their results agree with the study of Cabrera-Lavers et al. (2007, 2008) from an analysis of 2MASS data; however, several authors have pointed out that the 2MASS data does not probe the RC population. The survey is much shallower and only measures the upper red giant population with a broader range in luminosity and much larger fraction of disk contamination. Additionally, Simion et al. (2017) presented a 3D description of the bar/bulge from the VVV survey and found a strong degeneracy between the bar angle and the RC absolute magnitude dispersion. This degeneracy may be what is causing the Besançon Galactic Model (BGM) to underpredict the bar angle.

One final discrepancy between BGM and observations is the choice of bulge IMF the model uses. BGM1612 uses the Padova Isochrones (Marigo et al. 2008; Bressan et al. 2012), with a broken power law, $dn/dm \propto m^{-\alpha}$ for the bulge population. An IMF slope of $\alpha = 0.5$ is used for the low-mass range $0.15M_\odot < m < 0.7M_\odot$ and $\alpha = 2.3$ for $m > 0.7M_\odot$. The low-mass cutoff of $0.15M_\odot$ for the IMF is higher than that needed for the bulge population, as a major fraction of bulge stars will be very low-mass (VLM) dwarfs (Calamida et al. 2015). The low-mass IMF slope of $\alpha = 0.5$ is shallower than the BGM1106 model, which used a slope of $\alpha = 1.0$, both of which are shallower than low-mass slopes obtained from observations. P19 fit a more reasonable slope from Sumi et al. (2011) to the model and found better agreement with the shape of the SWEEPS LF, but they ultimately decided to keep the BGM1106 low-mass slope and correct for the inaccuracy in their further analysis. Adding VLM stars (e.g., $m < 0.15M_\odot$) to the model IMF will result in

[9] https://model.obs-besancon.fr/index.php

better agreement in surface densities at the dimmest magnitudes, especially in redder wavebands, which *WFIRST* will utilize for its microlensing survey. It's worth noting that adding brown dwarfs (BD) will also clearly increase the surface density; however, the mass function of BD in the bulge is quite uncertain (Sumi et al. 2011; Mróz et al. 2017; Wegg et al. 2017). Adding these VLM stars and some fraction of BD to the bulge IMF will result in increased optical depth and event rates per star for bulge–bulge lensing.

4.3.2. Galaxia

The Galaxia[10] version `0.7.2` code (Sharma et al. 2011) largely implements the Besançon model, but includes a wide variety of input parameters with more flexibility than that of BGM1612. Some adjustable parameters include the choice of an analytical or *N*-body seeded model from Bullock & Johnston (2005), the ability to subsample the simulated data by a given fraction in order to reduce runtimes and file size, and no restriction to the size of a given catalog (Besançon has set a new maximum of two million stars per simulation as of 2019 May). The model is run by a `C++` compilation. As Galaxia is essentially an altered version of BGM with extended capabilities, we proceed further with an analysis of the three discrepancies that were described in the previous subsection; the bulge IMF, the μ_ℓ kinematics, and the bulge-bar angle, in order to investigate any differences in their significance within this modified software:

1. Like Besançon, Galaxia uses the Padova isochrones for the bulge population but importantly also includes isochrones from Chabrier et al. (2000) for the lower-mass regime $0.07M_\odot < m < 0.15M_\odot$, below the BGM1612 cutoff. This results in a better agreement with the surface densities at the faintest magnitudes in both the *I*-band and the *H*-band LFs. Although Besançon can model WDs, this version of Galaxia does not account for them. This ultimately has little to no effect on our interpretation of star counts as this population is very faint and does not overlap our observation range. Lastly, BDs are not being modeled in this code, which may affect the resulting surface densities at the faintest levels but likely much fainter than the magnitude range of interest.
2. The analysis of the Galaxia kinematics follows the same procedure described in Section 4.3.1, with the focus of comparing the modeled σ_l values with the Stanek field measurements. Figure 9 shows the PM diagram and histogram comparison. The Galaxia longitudinal dispersion $\sigma_l = 3.23 \pm 0.13$ is consistent with both the observed value from the Stanek *HST* measurements and the BGM1612 dispersion. The circular velocity that Galaxia uses is marginally lower than BGM1612, at $V_c = 224.8$ km s^{-1}.
3. Galaxia adopts the same bulge-bar angle of $\sim$13° that BGM uses, which again is smaller than the results found by numerous observational studies in the past. Both models incorporate warp and flare of the thin/thick disk, derived from Robin et al. (2003). As stated above, the Galaxia input form allows the warp/flare to be turned on or off, while BGM1612 is hard-coded to always be on.

Overall, the Galaxia simulation results are tightly correlated with BGM1612 as expected. The bulge and disk kinematics are consistent with observed measurements from *HST*, and the addition of VLM stars in the bulge IMF is an advantage over Besançon for surface density calculations at dimmer (and redder) magnitudes. However, the too-low bar angle still persists in the current version.

Finally, the *N*-body model that Galaxia implements is a self-consistent realization of the formation of the stellar halo in Milky Way-type galaxies via the formulation presented in Bullock & Johnston (2005). The approach follows the evolution of accreted satellite galaxies in the halo formation process, and makes important distinctions between the evolution of light and dark matter within the host galaxies. The density of the accreted halo follows a varying power-law distribution, which changes radial slope from -1 within 10 kpc to -4 beyond 50 kpc. The distribution of stars is more centrally located in the halo compared to the dark-matter distribution. This is expected for the stars building the stellar halo to be more tightly bound than the dark-matter material that builds up the dark-matter halo. The model is largely successful in reproducing the observed properties of surviving Milky Way satellites and the stellar halo. The characteristics of the inner bulge region are not incorporated in the *N*-body spawning of particles and is generally not suited for studying the most central regions of the Milky Way. For these reasons, we do not conduct an analysis of the *N*-body model with regard to the Stanek field.

4.3.3. GalMod

GalMod[11] (Pasetto et al. 2016, 2018, 2019) version `18.19` is the most recently published population synthesis model considered in this work. The simulation offers significantly more adjustable parameters over the previous models, which include the choice of 24 different photometric bands, fine-tuning of the density normalization factor, ρ, for each composite stellar population (CSP; which also includes the ISM), 16 different combinations of SFR/IMF for each CSP, and the dark matter (DM) circular speed factor, scale radius, and flattening. Further, there are some distinct differences between GalMod and the previous models; the model incorporates a more realistic bulge-bar angle of $\sim$28°. GalMod generates convolved PDFs for star counts and then populates them with synthetic stars to obtain the CMDs and star counts, whereas Besançon and Galaxia work in the opposite manner. GalMod's PDF-generating technique is particularly useful in the era of very large surveys that we are currently entering. Although this process differs from the prior models, quantitatively, the GalMod approach to generating CMDs is identical to the other models. GalMod uses the geometry-independent ray-tracing extinction model of Natale et al. (2017), which is based on Draine & Li (2007), and can realize a collisional or collisionless model generator for an *N*-body integrator.

Additionally, there is a private version of the code that has tools to implement machine-learning techniques for data-fitting convergence, that uses dynamical estimators connected to the global Galactic potential, and has other features. For an extensive comparison between GalMod and Besançon, we refer the reader to Pasetto et al. (2016). The remainder of this section

[10] http://galaxia.sourceforge.net/

[11] https://www.galmod.org/gal/

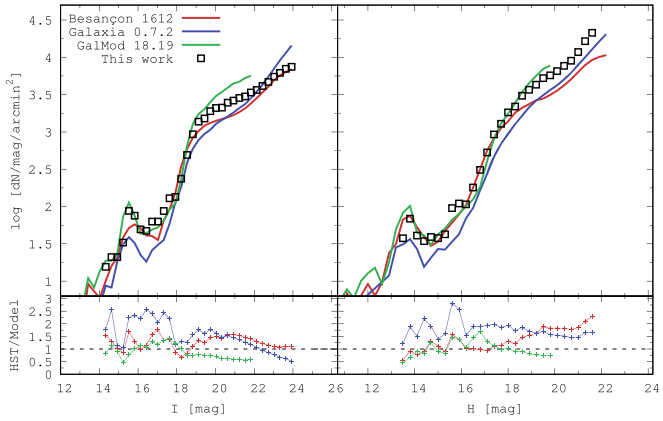

Figure 10. Top-left panel: *I*-band LF of Stanek compared to the models. Lower-left panel: ratio of raw Stanek counts to the models raw counts for the corresponding surface densities. Top-right and lower-right panels: similar comparison for the *H*-band data. The extinction is identical for each sample; however, extinction-corrections are not applied here.

will address the three discrepancies outlined in the previous subsections:

1. There are four different IMFs that can be used to simulate the Bulge/Bar CSP within GalMod; Salpeter (1955), Scalo (1986), Kroupa (2001), and Chabrier (2003). If no IMF is specified by the user, the model implements Kroupa (2001) by default. This is a broken power law with varying slopes; $\alpha = 2.3$ for $1.0M_{\odot} \leqslant m < \infty$, $\alpha = 2.7$ for $0.5M_{\odot} \leqslant m < 1.0M_{\odot}$, and $\alpha = 1.8$ for $0.16M_{\odot} \leqslant m < 0.5M_{\odot}$. Figure 10 shows the GalMod predicted star counts compared to the Stanek field observations and other models. The overall shape of the LF agrees well in both *I* and *H* bands, while the model overpredicts star counts in the *I*-band range $I > 19$. The normalized *H*-band counts agree quite well with observations and are certainly the most consistently integrated across a majority of the magnitude range.
2. The GalMod kinematics exhibit a mean PM value consistent with prior results; however, they show significant non-Gaussianity for the distributions in both μ_l and μ_b directions. The green data points in the bottom panel of Figure 9 show the PM distribution for the bulge RGB proxy stars from GalMod, which are clearly not well fit by a Gaussian. Finally, the circular velocity, V_c, at the Sun location that GalMod uses, is 220.8 km s^{-1}, which is marginally smaller than those of the previous two models.
3. The 28° bulge-bar angle that Galmod uses is in better agreement with some prior observations. This angle will lead to proper descriptions of bulge–bulge and bulge-disk lensing probabilities, and a more realistic ratio of bulge-to-disk source stars for microlensing.

As stated above, the SFR in each CSP can take one of four different forms; constant, exponential, linear, or Rosin (1933). The Rosin (1933) SFR is used by default and describes a rapid increase of the SFR up to a given time (free parameter), followed by a shallow decrease to the present day. This SFR has the functional form:

$$\psi(t) = \tau_{(t_1,t_2)}\psi_0 t^{\beta} e^{-\frac{t}{h_\tau}}, \quad (8)$$

Table 5
Observed Bulge Luminosity Function

V	log N_V	I	log N_I	J	log N_J	H	log N_H
14.629	0.484	14.329	1.192	14.1	1.284	13.5	1.575
14.929	1.262	14.629	1.323	14.4	1.363	13.8	1.837
15.229	1.086	14.929	1.322	14.7	1.626	14.1	1.612
15.529	1.505	15.229	1.517	15.0	1.359	14.4	1.537
15.829	1.630	15.529	1.941	15.3	1.474	14.7	1.592
16.129	1.907	15.829	1.878	15.6	1.481	15.0	1.576
16.429	1.982	16.129	1.694	15.9	1.688	15.3	1.628
16.729	1.774	16.429	1.673	16.2	1.930	15.6	1.979
17.029	1.614	16.729	1.798	16.5	2.003	15.9	2.041
17.329	1.820	17.029	1.799	16.8	2.132	16.2	2.029
17.629	1.679	17.329	1.940	17.1	2.394	16.5	2.256
17.929	2.025	17.629	2.111	17.4	2.564	16.8	2.491
18.229	2.097	17.929	2.128	17.7	2.629	17.1	2.724
18.529	2.255	18.229	2.373	18.0	2.856	17.4	2.966
19.129	2.495	18.529	2.691	18.3	3.072	17.7	3.110
19.429	2.849	18.829	2.968	18.6	3.195	18.0	3.261
19.729	3.042	19.129	3.138	18.9	3.294	18.3	3.340
20.029	3.093	19.429	3.181	19.2	3.394	18.6	3.484
20.329	3.185	19.729	3.279	19.5	3.484	18.9	3.564
20.629	3.239	20.029	3.321	19.8	3.564	19.2	3.633
20.929	3.247	20.329	3.326	20.1	3.618	19.5	3.721
21.229	3.254	20.629	3.392	20.4	3.695	19.8	3.758
21.529	3.260	20.929	3.420	20.7	3.747	20.1	3.816
21.829	3.280	21.229	3.455	21.0	3.817	20.4	3.886
22.129	3.283	21.529	3.476	21.3	3.888	20.7	3.954
22.429	3.287	21.829	3.528	21.6	3.976	21.0	4.071
22.729	3.286	22.129	3.562	21.9	4.049	21.3	4.216
23.029	3.290	22.429	3.616	22.2	4.159	21.6	4.328
23.329	3.305	22.729	3.669				
23.629	3.342	23.029	3.740				
23.929	3.357	23.329	3.792				
24.229	3.388	23.629	3.846				
24.529	3.402	23.929	3.873				

Note. Magnitude cutoff is at the 50% completeness threshold. Units for star counts N are consistent with that of Figure 10, i.e., [dN mag^{-1} arcmin^{-2}].

where β is the power-law exponent ($1 \neq \beta > 0$) and h_τ is the timescale ($h_\tau > 1$). The prior models both use a single burst for their bulge SFR.

Table 5 shows the binned LF for each of the four observed wavelengths (*V*, *I*, *J*, *H*) and Figure 10 shows the Stanek field LF and the Galactic models in *I*-band and *H*-band. BGM1612 underpredicts star counts by ∼33% at the RC location and integrated over the magnitude range 19.5–23.5 in *I*-band. Galaxia undercounts bulge stars more severely at brighter magnitudes. However, because of the accounting for VLM stars in the IMF, Galaxia counts more faint stars than BGM1612. This more closely matches the observed Stanek field numbers and even over-counts at the faintest *I*-band end. In both plots, Galaxia overtakes BGM1612 in counts dimmer than the $0.15M_{\odot}$ low-mass cutoff (at ∼21 mag in *I*-band and ∼19 mag in *H*-band). The lower panels of Figure 10 show the raw surface density ratio between observation and the models. GalMod most closely predicts the star counts in both wavebands, with an accuracy of ∼10% at $I < 20$ and ∼5% at $H < 19$.

With the larger bulge-bar angle, and higher predicted surface densities, GalMod stands to be a promising tool for further simulations and accurately predicting microlensing observables for *WFIRST*. The low-mass cutoff presents an issue for predicted

surface densities at the dimmest magnitudes, as well as bulge star kinematics that are not in solid agreement with observations or other models. The purpose of this Section 4.3 is not to fully simulate the microlensing mission itself, but to conduct a detailed comparison of popular Galactic models to observations.

5. Conclusion

We have accurately measured PMs of the disk, bulge, and mixed stellar populations to within $\sim$20 km s^{-1} around $I = 21$ and $H = 20$ and within $\sim$60 km s^{-1} near the 50% completeness limits of $I = 24$ and $H = 22$ and have likely probed the far-side disk population beyond the Galactic center. Our measured PM dispersions for each population are largely in agreement with prior studies; however, we measure a bulge component anisotropy ratio of $\sigma_l/\sigma_b = 1.03 \pm 0.04$, which is significantly smaller than prior results from studies in nearby bulge fields. We note that most of the past studies analyzed the "mixed" populations and are contaminated by disk rotation. The exception is the study of Calamida et al. (2015), who measure the pure-bulge anisotropy ratio in the SWEEPS field to be in agreement with our result. We applied magnitude-dependent scaling factors to the cleaned star counts in both I and H bands in order to properly account for bulge members that are excluded by our PM cut.

The resulting bulge surface densities were compared to several Galactic population synthesis models. We find that the Besançon and Galaxia models generally underpredict star counts in both wavebands from 33% to upwards of 75% in the most severe case (i.e., $H > 18$). The bulge-bar angle that is smaller than previously measured values may be due to a degeneracy between this angle and the RC absolute magnitude dispersion, as discussed in Simion et al. (2017).

On the other hand, the newest population synthesis model, GalMod, produces bulge star counts that are in best agreement with the observed values along the Stanek sight line. Although the model overpredicts bulge star counts by $\sim$35% below the MSTO in I-band, it only overpredicts the counts by $\sim$5%–10% in the H-band data. These redder data are more closely aligned with the wavebands that the *WFIRST* microlensing mission will utilize during its observations.

There are two drawbacks of note with GalMod; the bulge kinematics show a non-Gaussian PM distribution in both μ_l and μ_b directions. This disagrees with the measured kinematics of Stanek field observations presented in this paper as well as prior observational results in nearby bulge fields (Clarkson et al. 2008; Calamida et al. 2015). The second drawback is the low-mass cutoff of the empirical IMF; they do not allow for the inclusion of the dimmest stars, which need to be factored in when attempting to simulate the lens and source-star distribution and characterization in bulge sight lines for *WFIRST* observations.

The previous study of P19 worked to adjust the microlensing event rate based on a similar comparison to ours, between the SWEEPS field (Calamida et al. 2015) and an older version of the Besançon model that used a shallow IMF and unrealistic bulge kinematics. The authors subsequently corrected for these inaccuracies so that their simulations correctly predict microlensing event rates that match observations. Our star count results and comparison with the Galactic models support P19 with regard to the kinematics and low-mass IMF corrections that are required. As P19 point out, there is an urgent need to advance simulation capabilities, particularly the correcting of stellar surface densities in highly crowded bulge fields, in order to optimize *WFIRST*'s observing strategy with regard to direct mass measurements. Lastly, an important precursor advancement we have also shown in this paper is an accurate description of the source magnitude distribution very nearby to the provisional *WFIRST* microlensing fields and to a (near-IR) depth overlapping what is achievable by *WFIRST*.

5.1. *Implications for* WFIRST

Over the four design reference mission studies, beginning with the mission proposed by the 2010 decadal survey, there has been significant variance in estimates of the expected microlensing event rate, optical depth, and exoplanet yield. Calculations of these mission success criteria were based on disparate population synthesis models, measurements along sight lines distant from the proposed *WFIRST* microlensing fields, and shallow LFs that do not overlap the *WFIRST* wavebands. In this work, we have measured the stellar populations directly adjacent to the *WFIRST* microlensing fields to near-IR wavebands J and H. These measurements take the form of kinematic distributions and dispersions in the field, as well as accurately measure stellar surface densities, which are compared to several population synthesis models.

These results can be used to directly answer mission-critical scientific needs stated in Spergel et al. (2013) and Yee et al. (2014). Particularly, Yee et al. Section 1.2: "*Improve characterization of the WFIRST fields*," as well as Spergel et al., the "*scientific need to measure the: source star luminosity function, near-IR event rate, and relative bulge-to-disk planet frequency*." With regard to the Galactic models used in this work, we re-iterate P19 Section 6.4: there is still room for improvement to Galactic models; we can place tighter constraints on microlensing observables and source/lens properties with updated models. Particularly, new PM results from *Gaia* DR2 (Brown et al. 2018) can be included in the models to better characterize the kinematics within the *WFIRST* microlensing fields. There are plans to include these new PM measurements in a future GalMod release (S. Pasetto 2019, private communication). Additionally, the future *Gaia* DR3 and DR4 releases offer prospects of high-accuracy positions and kinematics for the farthest stars yet toward the Galactic center at $\sim$8 kpc. The stellar properties within the models as well as future *WFIRST* preparatory work will benefit greatly from these future releases.

Our PM analysis presented here shows that the stars along the sight line to Stanek's Window in the bulge and disk exhibit a PM (longitude and latitude, combined) of about 4.2 mas yr^{-1}. While *WFIRST* will observe the Galactic bulge for six seasons over five years during its primary mission, it will not visit the same field with the same orientation for about 3.5 yr, corresponding to $\sim$15 mas total PM per detected event. This is the case, however, only for bulge-disk lensing. When compared to *WFIRST*'s pixel scale of 110 mas, we may expect a typical star to exhibit 0.14 pix of motion from one visit to the next. Using the color-dependent centroid shift method (Bennett et al. 2006, 2015; Bhattacharya et al. 2018) as well as the image elongation method (Bennett et al. 2007), we may expect a precision of 11% on the lens-source separation with 3.5 yr of baseline.

Finally, the near-IR source magnitude distribution and other results presented here can be be combined with future studies to further simulate the scientific yields of the *WFIRST* microlensing

survey. For instance, the previous extrapolation errors in the GULLS simulation software (Penny et al. 2013) can now be mitigated with newer, corrected models.

The authors would like to thank Stefano Pasetto for GalMod technical assistance. We also thank Annalisa Calamida, Kailash Sahu, and Noé Kains for practical help and assistance. Finally, we thank the anonymous referee for constructive comments that led to a stronger manuscript.

This work was performed in part under contract with the Center for Research and Exploration in Space Sciences & Technologies (CRESST-II), which is a collaboration between NASA GSFC and surrounding universities. This work makes use of data products from GO-proposals 11664 and 12666 of the NASA/ESA *Hubble Space Telescope*, obtained by STScI. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. D.P.B. and A.B. were supported by NASA through grant NASA-80NSSC18K0274.

This publication makes use of the software packages: Astropy (Astropy Collaboration et al. 2018), gnuplot, KS2 (Anderson et al. 2008), Matplotlib (Hunter 2007), and Numpy (Oliphant 2006).

ORCID iDs

Sean K. Terry https://orcid.org/0000-0002-5029-3257
Richard K. Barry https://orcid.org/0000-0003-4916-0892
David P. Bennett https://orcid.org/0000-0001-8043-8413
Jay Anderson https://orcid.org/0000-0003-2861-3995
Matthew T. Penny https://orcid.org/0000-0001-7506-5640

References

Anderson, J., & King, I. R. 2006, ACS Instrument Science Report 2006-01
Anderson, J., Sarajedini, A., Bedin, L. R., et al. 2008, AJ, 135, 2055
Astropy Collaboration, Price-Whelan, A., Sipőcz, B., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Baldwin, A. T., Watkins, L. L., van der Marel, R. P., et al. 2016, ApJ, 827, 12
Barclay, T., Quintana, E. V., Raymond, S. N., & Penny, M. T. 2017, ApJ, 841, 86
Bennett, D., Bhattacharya, A., Anderson, J., et al. 2015, ApJ, 808, 169
Bennett, D. P., Anderson, J., Bond, I. A., Udalski, A., & Gould, A. 2006, ApJL, 647, L171
Bennett, D. P., Anderson, J., & Gaudi, B. S. 2007, ApJ, 660, 781
Bernard, E. J., Schultheis, M., Di Matteo, P., et al. 2018, MNRAS, 477, 3507
Bhattacharya, A., Beaulieu, J.-P., Bennett, D., et al. 2018, AJ, 156, 289
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Brown, A., Vallenari, A., Prusti, T., et al. 2018, A&A, 616, A1
Brown, T. M., Sahu, K., Anderson, J., et al. 2010, ApJL, 725, L19
Brown, T. M., Sahu, K., Zoccali, M., et al. 2009, AJ, 137, 3172
Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931
Cabrera-Lavers, A., González-Fernández, C., Garzón, F., Hammersley, P., & López-Corredoira, M. 2008, A&A, 491, 781
Cabrera-Lavers, A., Hammersley, P., González-Fernández, C., et al. 2007, A&A, 465, 825
Calamida, A., Bono, G., Lagioia, E. P., et al. 2014a, A&A, 565, A8
Calamida, A., Sahu, K. C., Anderson, J., et al. 2014b, ApJ, 790, 164
Calamida, A., Sahu, K. C., Casertano, S., et al. 2015, ApJ, 810, 8
Cao, L., Mao, S., Nataf, D., Rattenbury, N. J., & Gould, A. 2013, MNRAS, 434, 595
Chabrier, G. 2003, PASP, 115, 763
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Clarke, J. P., Wegg, C., Gerhard, O., et al. 2019, arXiv:1903.02003
Clarkson, W., Sahu, K., Anderson, J., et al. 2008, ApJ, 684, 1110
Clarkson, W. I., Calamida, A., Sahu, K. C., et al. 2018, ApJ, 858, 46
Clarkson, W. I., Sahu, K. C., Anderson, J., et al. 2011, ApJ, 735, 37
Deustua, S., Mack, J., Bajaj, V., & Khandrika, H. 2017, Instrument Science Report WFC3 2017-14, 19
Dotter, A. 2016, ApJS, 222, 8
Draine, B. T., & Li, A. 2007, ApJ, 657, 810
Dwek, E., Arendt, R. G., Hauser, M. G., et al. 1995, ApJ, 445, 716
Einasto, J. 1979, in IAU Symp. 84, The Large-scale Characteristics of the Galaxy, ed. W. B. Burton (Dordrecht: Reidel), 451
Gaudi, B. S. 2012, ARA&A, 50, 411
Gennaro, M., Brown, T., Anderson, J., et al. 2015, in ASP Conf. Ser. 491, Fifty Years of Wide Field Studies in the Southern Hemisphere: Resolved Stellar Populations of the Galactic Bulge and Magellanic Clouds, ed. S. Points & A. Kunder (San Francisco, CA: ASP), 182
Gonzalez, O., Rejkuba, M., Zoccali, M., et al. 2012, A&A, 543, A13
Holtzman, J. A., Watson, A. M., Baum, W. A., et al. 1998, AJ, 115, 1946
Hunter, J. D. 2007, CSE, 9, 90
Izumiura, H., Catchpole, R., Deguchi, S., et al. 1995, ApJS, 98, 271
Kozłowski, S., Woźniak, P. R., Mao, S., et al. 2006, MNRAS, 370, 435
Kroupa, P. 2001, MNRAS, 322, 231
Kuijken, K., & Rich, R. M. 2002, AJ, 124, 2054
Lagioia, E. P., Milone, A. P., Stetson, P. B., et al. 2014, ApJ, 782, 50
Mao, S. 2008, arXiv:0811.0441
Marigo, P., Girardi, L., Bressan, A., et al. 2008, A&A, 482, 883
Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012, A&A, 540, A16
Montet, B. T., Yee, J. C., & Penny, M. T. 2017, PASP, 129, 044401
Mróz, P., Udalski, A., Skowron, J., et al. 2017, Natur, 548, 183
Mroz, P., Udalski, A., Skowron, J., et al. 2019, arXiv:1906.02210
Nataf, D., Udalski, A., Gould, A., Fouqué, P., & Stanek, K. 2010, ApJL, 721, L28
Nataf, D. M., Gould, A., Fouqué, P., et al. 2013, ApJ, 769, 88
Natale, G., Popescu, C. C., Tuffs, R. J., et al. 2017, A&A, 607, A125
Oliphant, T. E. 2006, A Guide to NumPy, Vol. 1 (USA: Trelgol Publishing)
Ortolani, S., Renzini, A., Gilmozzi, R., et al. 1995, Natur, 377, 701
Pasetto, S., Crnojević, D., Busso, G., Chiosi, C., & Cassarà, L. 2019, NewA, 66, 20
Pasetto, S., Grebel, E. K., Chiosi, C., et al. 2018, ApJ, 860, 120
Pasetto, S., Natale, G., Kawata, D., et al. 2016, MNRAS, 461, 2383
Penny, M., Kerins, E., Rattenbury, N., et al. 2013, MNRAS, 434, 2
Penny, M. T., Gaudi, B. S., Kerins, E., et al. 2019, ApJS, 241, 3
Rattenbury, N. J., Mao, S., Debattista, V. P., et al. 2007, MNRAS, 378, 1165
Reid, M., Menten, K., Zheng, X., et al. 2009, ApJ, 700, 137
Renzini, A., Gennaro, M., Zoccali, M., et al. 2018, ApJ, 863, 16
Robin, A. C., Marshall, D. J., Schultheis, M., & Reylé, C. 2012, A&A, 538, A106
Robin, A. C., Reylé, C., Derriere, S., & Picaud, S. 2003, A&A, 409, 523
Rosin, P. 1933, Journal of Institute of Fuel, 7, 29
Saito, R. K., Hempel, M., Minniti, D., et al. 2012, A&A, 537, A107
Salpeter, E. E. 1955, ApJ, 121, 161
Scalo, J. M. 1986, FCPh, 11, 1
Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3
Simion, I. T., Belokurov, V., Irwin, M., et al. 2017, MNRAS, 471, 4323
Spaenhauer, A., Jones, B. F., & Whitford, A. E. 1992, AJ, 103, 297
Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757
Spergel, D., Gehrels, N., Breckinridge, J., et al. 2013, arXiv:1305.5422
Stanek, K., Mateo, M., Udalski, A., et al. 1994, ApJL, 429, L73
Stanek, K., Udalski, A., Szymański, M., et al. 1997, ApJ, 477, 163
Sumi, T., Kamiya, K., Bennett, D. P., et al. 2011, Natur, 473, 349
Suzuki, D., Bennett, D., Sumi, T., et al. 2016, ApJ, 833, 145
Szymański, M. K., Udalski, A., Soszyński, I., et al. 2011, AcA, 61, 83
Wegg, C., & Gerhard, O. 2013, MNRAS, 435, 1874
Wegg, C., Gerhard, O., & Portail, M. 2017, ApJL, 843, L5
Yee, J. C., Albrow, M., Barry, R. K., et al. 2014, arXiv:1409.2759
Zhao, H., Spergel, D. N., & Rich, R. M. 1994, AJ, 108, 2154
Zhao, H., Spergel, D. N., & Rich, R. M. 1995, ApJL, 440, L13
Zoccali, M., Cassisi, S., Frogel, J. A., et al. 2000, ApJ, 530, 418