Photometric Calibration of SuperBIT

Liam O'Shaughnessy

July 30, 2024

1 Introduction

SuperBIT is a balloon-bound telescope, which was lifted into the stratosphere for data collection in 2023 [8]. SuperBIT mapped dark matter distributions in clusters and imaged galaxies [9] [7]. However, the photometric data from SuperBIT bands has yet to be calibrated with a reference. By calibrating with a known source, not only can data units be physically quantified, but the behaviors of data in different regimes of the dynamic range can be better understood.

This project seeked to determine a photometric calibration of SuperBIT band data with respect to data from the Third Public Release of the Subaru Telescope Hyper Suprime-Cam, in the fields of a1689 and cosmosk.

2 Data

Data for the project included .fits files for band 0, band 1, and band 2 of SuperBIT pointed at a1689 and cosmosk regions, as well as .fits files for the g band and r band of HSC for those regions. SuperBIT .fits files had pixel values based on the ADU counts, based on photoelectron hits on the CCDs. HSC data was taken the hscMap feature of the PDR3 site [1]. From direct correspondence with HSC personnel, the HSC units were such that the sum of .fits pixel values in

a 12-pixel radius about a point gives the flux value for a conversion to AB magnitude. Therefore, for a given zero magnitude, a point source star with (0,0) being the object center, and S being the 12-pixel radius region about the center:

$$m_{AB} = -2.5 * \log(flux/m_0)$$

$$m_{AB} = -2.5 * \log(\frac{\sum_{i,j \in S} d_{ij}^{adj})}{m_0})$$

$$d_{ij}^{adj} = d_{ij}^{raw} - background = PSF_{ij} * (d_{0,0}^{raw} - background) = PSF_{ij} * d_{0,0}^{adj}$$

$$m_{AB} = -2.5 * \log(\frac{d_{0,0}^{adj}}{m_0} \sum_{i,j \in S} PSF_{ij})$$

$$f = \sum_{i,j \in S} PSF_{ij}$$

$$m_{AB} = -2.5 * \log(\frac{d_{0,0}^{adj}}{m_0} * f)$$

This allows the approximate magnitude of an HSC object to now be calculated, for the zero magnitude given in the .fits file . Background was estimated by, for some point (x,y), taking a 24-pixel radius disk about it and computing the median pixel value on the boundary. An $\overline{f}$ was computed by identifying a number (112) of star objects with a semi-major axis of 0.22 and .24 arcseconds (encompassing the stated angular resolution of the telescope [6]), then averaging their f values. A minimum-brightness cut was also applied to eliminate faint star outliers. This gives that $\overline{f} = 27.517, \sigma_f = 6.902$, computed from the a1689 data.

3 Method

The aim of the project was, for each galaxy observed in both SuperBIT data and HSC data, to compute the flux of the objects, then to find a correction that allowed a reasonable estimation of the reference flux from the *j*th HSC band via the flux data from the *i*th SuperBIT band. Due to the differences in bandpass distributions for the SuperBIT bands and HSC bands, this becomes

finding an α_{ij} in an approximation such that

$$Flux_{reference_j} = \int_0^\infty R_{HSC_j}(\lambda) * S_{obj}(\lambda)\, d\lambda \approx \alpha_{ij} * \int_0^\infty R_{BIT_i}(\lambda) * S_{obj}(\lambda)\, d\lambda + ...$$

where $R(\lambda)$ and $S(\lambda)$ are the band receptiveness and object spectrum respectively. There are higher-order corrections in this approximation, yet the α_{ij} term should provide a close estimate. For instance, if a SuperBIT band had a proportionally lower absorption rate than an HSC band for every wavelength of light, then this α_{ij} approximation would be extremely accurate and would simply scale between two identical distribution shapes. For differences in the bandpass distributions, however, such as in non-overlapping bands or different shapes, this approximation is less accurate. This α_{ij} would then serve to convert SuperBIT counts into a more understandable flux measurement.

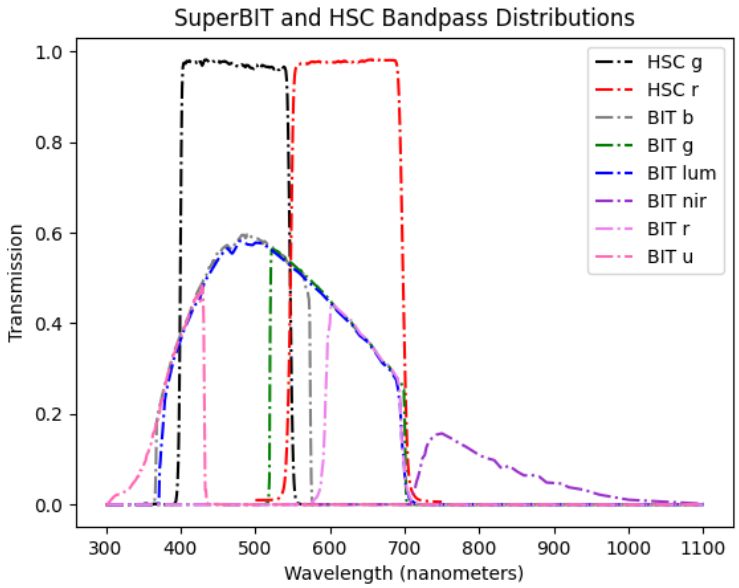

Figure 1: Bandpass Distributions for SuperBIT and HSC

The first step was to identify stellar objects in the .fits files and export relevant object parameters for analysis. This was performed using the Source Extractor software, which can take in the .fits files and return .cat files of chosen parameters per object [2]. For SuperBIT and HSC band data, the following parameters were returned:

- FLUXAUTO of object
- Error in FLUXAUTO
- Background at centroid
- Peak flux above background
- Threshold above background for analysis
- Barycenter positions along world x and y
- Barycenter right ascension and declination
- Semi-major and semi-minor axes of the object

The Source Extractor software identifies objects and parameters via the following method. For a given .fits file, where a linear detector is used, the pixel values are equal to the background plus the signal from the object. Source Extractor estimates the background contribution by creating a grid of rectangular meshes on the image, then, in each mesh, iteratively finding the mean/mode (depending on crowding) pixel value, clipping outliers, and repeating until pixel values lie within some range [5]. The mean/mode is then considered the background in the mesh (there is also an option to apply a median filter to smooth bright objects from interfering). Weighting can be applied at this point, yet this analysis did not employ this method. Once the background has been determined, candidate objects are selected via the following criteria: all pixel values are above a specified detection threshold (DETECTTHRESH), all pixels are adjacent, and there are more pixels than the minimum area threshold (here: 5) [2]. The filtering process then convolves the image with a chosen filter, and deblends to identify objects via essentially making a tree on the graph of pixel value vs. position, and identifying branches (relative maxima) as objects.

In terms of photometry, the auto flux measurement was employed, which used the Kron radius (I will hence call this Kron flux to be more specific). This is given by the sum of background-subtracted pixels within the Kron ellipse, which is determined via the moments-generated ellipse axes of the object, and a version of Kron's first moment algorithm[5]. The error in this measurement is given from calculated RMS noise. This measure of computing object flux appears to be less biased than the isophotal method. Object positions as well ellipse axes are calculated from moments of the pixel values, along with RA and Dec information given by WCS in the .fits files.

Once the object parameters were obtained in .cat files, a script was then run on each of the six SuperBIT band x HSC band pairs to create a catalog of common objects. The threshold was that the RA and Dec each had to be within $5 * 10^{-5}$ pixels for a1689 and $2 * 10^{-4}$ pixels for cosmosk to be considered the same object. Originally, this threshold was too generous, and objects would be double-counted, yet under the current thresholds, the number of the objects in the catalog and not in the catalog does indeed sum to the total. Additionally, only objects in the common region seen by all bands were included.

Since α would be the ratio of object fluxes, confounding sources of flux were eliminated via various methods. First, a cut on peak flux (the mean of the SuperBIT peak flux data) was imposed to ensure possibly-saturated SuperBIT pixels did not interfere with flux calculations. Second, another cut was imposed on object elliptical semi-major axis to attempt to cut out stars (time-variable). This cut was generated by first constructing a functional form for the second moment (proxy for semi-major axis of object) as a function of an azimuthally-averaged SuperBIT PSF's integral upper limit (proxy for flux). This created a model that replicated that, as signal-to-noise increases, more flux from the object falls above Source Extractor's threshold, and hence the apparent angular size increases. This form was then scaled and fitted to data that appears to observe this pattern. Additionally, a minimum SuperBIT flux value of 0.01 was

required. Finally, for objects appearing to be outliers, if there appeared to be morphological oddities (e.g. blending in the HSC data), these objects were excluded.

4 Results

Some band combinations, as mentioned above, are more informative than others, as the estimated accuracy of the lower-order model is better - this arises from the "closeness" of the bandpass distributions. In this case, between SuperBIT and HSC, BIT band 1 resembled the HSC g band (symmetrically as well), and BIT band 2 resembled the HSC r band. This analysis, hence, primarily focused on the calibration values for these band pairings. Calibrations for BIT band 0 would likely require finding a reference telescope with a similar bandpass that has data from a common region of the sky.

As for linear regressions, this analysis employed orthogonal regression for data with two-dimensional uncertainties [4]. Uncertainties in object flux were taken from Source Extractor's flux error measurement, and these errors were assumed to be uncorrelated between SuperBIT and HSC. Assuming that these uncertainties are Gaussian, the linear regression $y = mx + b$ is given by maximizing the likelihood of Gaussians, with parameters being the data covariance $\mathbf{S}_i$ and model-predicted values on the line with minimal noise $\mathbf{Z} = (x, y)$, for the N observed data points $\mathbf{Z}_i = (x_i, y_i)$. This can be derived using

$$p(x_i, y_i|\mathbf{S}_i, x, y) \propto exp(-\frac{1}{2} * [\mathbf{Z}_i - \mathbf{Z}]^T \mathbf{S}_i^{-1} [\mathbf{Z}_i - \mathbf{Z}])$$

$$\hat{\mathbf{v}} = (-\sin\theta, \cos\theta), \theta = \arctan m$$

$$ln(\mathscr{L}) = K - \sum_{i=1}^{N} \frac{(\hat{\mathbf{v}}^T \mathbf{Z}_i - b\cos\theta)^2}{2 * \hat{\mathbf{v}}^T \mathbf{S}_i \hat{\mathbf{v}}}$$

where $\hat{\mathbf{v}}$ is the unit vector orthogonal to the best-fit line, and K is an constant from the coefficient on the product of the Gaussian likelihoods [4].

The log-likelihood was maximized via Markov Chain Monte Carlo using the emcee Python package to find the best-fit line for each band combination observing both a1689 and cosmosk. The emcee package uses an affine-invariant ensemble sampler algorithm to efficiently sample distributions [3] [10]. Using the so-called "uninformative prior" distribution, set over a range centered on previous best-fit results, a Gaussian initial distribution also centered on previous results, and 50 chains ran for 3,000 steps, best-fit parameters were estimated from the sampled likelihood after burn-in and thinning were applied. The results from the MCMC are shown below.

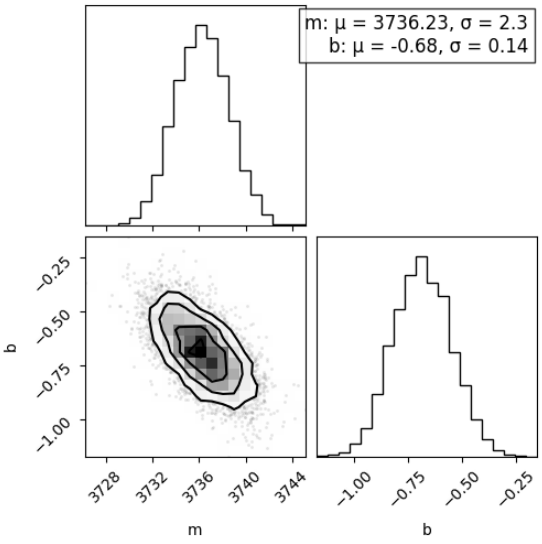

Figure 2: a1689 Data

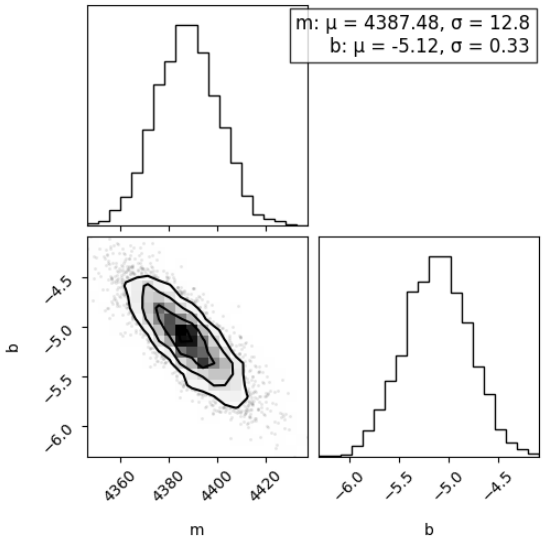

Figure 3: cosmosk Data

Figure 4: Triangle Plots for BIT 1 x HSC g

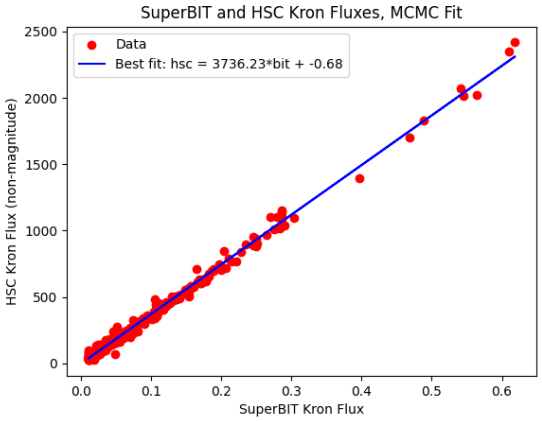

Figure 5: a1689 Data

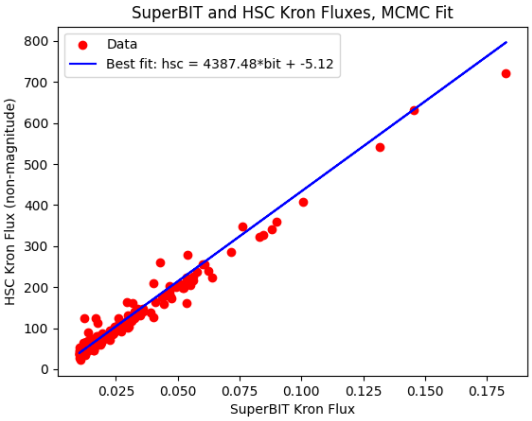

Figure 6: cosmosk Data

Figure 7: Regressions for BIT 1 x HSC g

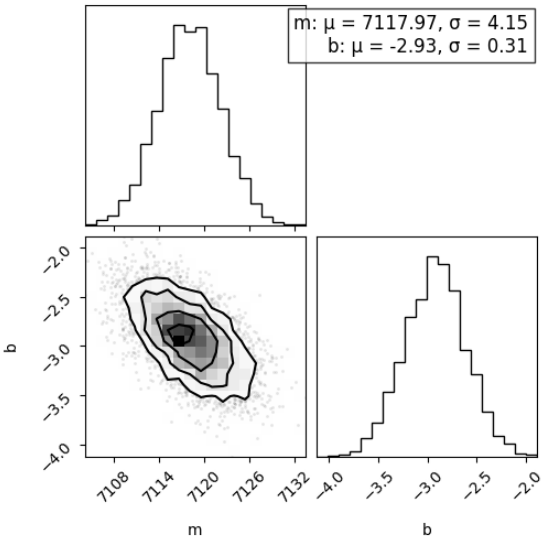

Figure 8: a1689 Data

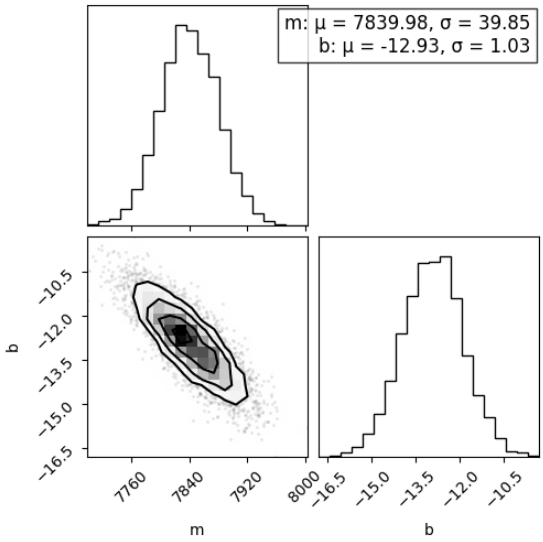

Figure 9: cosmosk Data

Figure 10: Triangle Plots for BIT 2 x HSC r

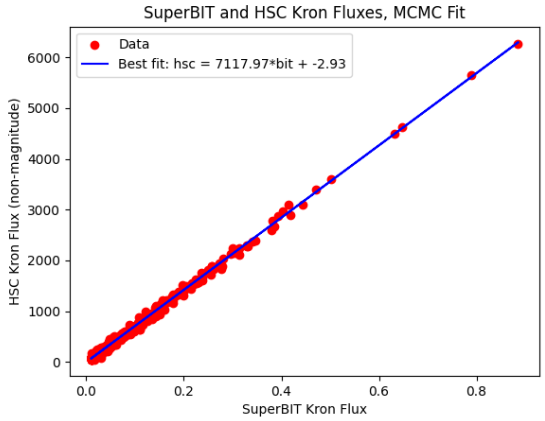

Figure 11: a1689 Data

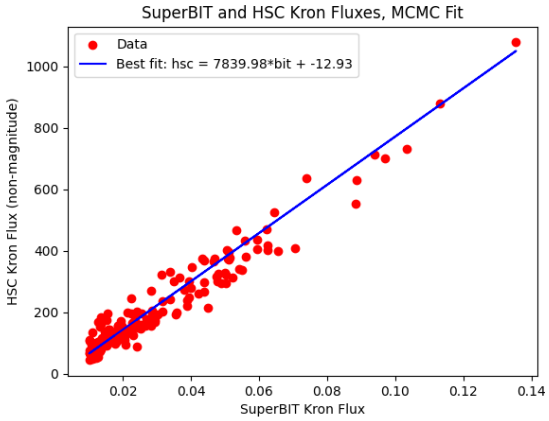

Figure 12: cosmosk Data

Figure 13: Regressions for BIT 2 x HSC r

5 Discussion

As seen in the MCMC regression plots, the regressions yield neighboring, yet apparently statistically different, flux ratio values for calibration. These differences could be the result of a number of effects, such as differences in the emission spectra of objects in the data, redshifts of objects, the relative depths of SuperBIT and HSC, etc. For instance, the cosmosk data that remains has, on average, dimmer flux values than a1689. Additionally, the HSC measurements for the cosmosk field, coming from one of their ultra-deep regions, are particularly dense. Calculating the average number of objects around each object in a 0.005 degree radius within the common windows (a "neighbor density"), the SuperBIT and HSC densities for a1689 (and the SuperBIT densities for cosmosk) are about 35-50 objects per square arcminute - the neighbor density of HSC on cosmosk appears to be around 150 objects per square arcminute. It is possible that the increased occurrence and likelihood of blending is producing a negative effect, such as, for instance, rendering invalid data from high-flux SuperBIT objects due to a lack of unblended counterparts in HSC.

Previously, imposing certain cuts/restrictions was tried - this was a mixed bag. One, for instance, was restricting a1689's flux range to [0,0.07], so that it was in the same regime as the cosmosk flux range. This did not necessarily improve the MCMC. Imposing a signal-to-noise cut (of about 10) was also motivated yet didn't necessarily improve the agreement; going higher than 20 runs into problems due to BIT 2 x HSC r having so much outlier elimination already that the sample size was low. Some of these techniques also cause the intercept values for the regression to become increasingly negative.

In summary, it appears that the first-order calibration factors are somewhat known: around 3700 to 4400 for BIT band 1 to HSC g (non-magnitude), and around 7100 to 7900 for BIT band 2 to HSC r (non-magnitude). Further investigation into this would be very helpful in determining a value with uncertainty, such as attempting to quantify the effect of object spectra or attempting to calibrate on different reference telescopes and different regions of sky. Additionally, there is the need to calculate a calibration for BIT band 0, as HSC does not have a comparable band.

References

[1] Aihara et al. "Third data release of the Hyper Suprime-Cam Subaru Strategic Program." In: *Publications of the Astronomical Society of Japan* 74.2 (Feb. 2022), pp. 247–272. ISSN: 2053-051X. DOI: 10.1093/pasj/psab122. URL: http://dx.doi.org/10.1093/pasj/psab122.

[2] E. Bertin. *SExtractor Documentation - Release 2.24.2*. Source Extractor Documentation. 2017. URL: https://sextractor.readthedocs.io/_/downloads/en/latest/pdf/.

[3] Daniel Foreman-Mackey et al. “emcee: The MCMC Hammer.” In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pp. 306–312. ISSN: 1538-3873. DOI: 10.1086/670067. URL: http://dx.doi.org/10.1086/670067.

[4] David W. Hogg, Jo Bovy, and Dustin Lang. *Data analysis recipes: Fitting a model to data*. Data Analysis Guide. 2010. arXiv: 1008.4686 [astro-ph.IM]. URL: https://arxiv.org/abs/1008.4686.

[5] B. W. Holwerda. *Source Extractor for Dummies v5*. Source Extractor Manual. 2005. arXiv: astro-ph/0512139 [astro-ph]. URL: https://arxiv.org/abs/astro-ph/0512139.

[6] M. Iye et al. “Current Performance and On-Going Improvements of the 8.2 m Subaru Telescope.” In: *Astronomical Society of Japan* 56 (2 2004). DOI: https://doi.org/10.1093/pasj/56.2.381.

[7] Veome Kapil. *Analyzing the Gravitational Lensing performance of SuperBIT to constrain Dark Matter Self Interactions*. Princeton Senior Thesis. 2021.

[8] Richard Massey, C. Barth Netterfield, and William C. Jones. “Forty days and forty-five nights at space’s edge.” In: *Nature Astronomy* 8 (2024), p. 264. DOI: https://doi.org/10.1038/s41550-024-02202-1.

[9] Mohamed M. Shaabad et al. “Weak Lensing in the Blue: A Counter-intuitive Strategy for Stratospheric Observations.” In: *The Astronomical Journal* 164.6 (2022). DOI: https://doi.org/10.3847/1538-3881/ac9b1c.

[10] A. Sokal. “Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms.” In: Springer US, 1997.