Response to the referee report of the SynPipe paper

We thank the referee for the careful reading and many valuable comments. Here we attached the updated draft where changes are highlighted in cyan color.

We believe that all the comments and questions from the referee have been addressed, and below are the detailed response to some of the referee's concerns.

---

0. About the general comment that this draft resembles a "technical supplement of the hscPipe paper"

We think:
* SynPipe paper presents important evaluations of the current HSC data. These results are useful to the general users of HSC instrument. Given that the hscPipe paper already contains lots of technical details and is quite long, we believe that a separate paper can better draw people's attention to these results.
* As we now presented in the new Discussion section, as an independent software, SynPipe is being used for any scientific goals, both within and outside the SSP community. For some of them (e.g. Murata et al. in prep. about the impact of blended objects in weak lensing analysis), SynPipe is the key technical component. We think it is useful to have a separate paper as reference to the readers of these works and the future SynPipe users. In addition, SynPipe has recently been adopted by the LSST community, and will be used to test the performance of the LSST pipeline. We think it is useful to summarize the development of SynPipe on the HSC side as an independent paper.
* We would like to draw the referee's attention to papers with the similar purpose, e.g. Balrog for the DES survey (http://adsabs.harvard.edu/abs/2016MNRAS.457..786S). These papers are proved to be useful and well-cited for similar reasons as we stated above.

---

1. Discussion:

We agree with the referee that this work can benefit from a more detailed discussion, which we have added before the Summary section. A few relevant points were included in the earlier Summary and Conclusion section, and have been moved to the new Discussion section.

We emphasize the importance of such synthetic object pipeline for modern cosmology survey, explain again the philosophy of our design, and compare it with other efforts for testing the data reduction pipeline.

Right now, SynPipe is specifically tailored for HSC images, but it will be included in the LSST pipeline in the near future. We demonstrate the broad scientific topics that can benefit from SynPipe, and describe the on-going efforts for further improvements.

About the current issues with the HSC pipeline and its results. They have been fully documented and discussed in both hscPipe technical paper (Bosch+2017; http://adsabs.harvard.edu/abs/2017arXiv170506766B) and the HSC SSP DR1 paper (Aihara+2017; http://adsabs.harvard.edu/abs/2017arXiv170208449A). We also point the reader there for more details.

---

2 Interpreter Language

The referee is correct that, in general, C++ should be more efficient in the heavy duty data reduction process. However, SynPipe here mainly works as a special user interface to hscPipe, which is also use Python for high-level algorithm. The reason for this choice is described in details in the Section 2 of the technical paper for hscPipe (Bosch et al. 2017). Below is the most relevant paragraph:

> Both the LSST Data Management codebase and the HSC Pipeline are written in a combination of C++ and Python. Flexibility and extensibility are core development goals at all levels of the code-base: we have attempted to make virtually every algorithm not just configurable but replaceable (even by external code). This flexibility is important for LSST in part because many algorithmic choices in LSST processing still require significant research and are under active development.....
>
> High-level algorithms are written in Python, primarily as subclasses of our Task class...

> Low-level algorithms that perform direct pixel-level processing or other computationally intensive operations are written in C++....

As the referee can see, flexibility and extensibility is the main reason that hscPipe choose to use Python as the high-level language. GalSim also uses the same logic. Since SynPipe 1) directly depends on hscPipe and GalSim; 2) will develop along with the hscPipe (and LSSTpipe in the future); 3) does not do any low-level computationally intensive operation, we think it is natural to follow this decision and use Python as our working language.

The efficiency of SynPipe, as the referee mentioned, certainly relates to the available cores of the CPU. What we meant to explain is that, under the same computational condition, SynPipe test costs similar amount time as the real data reduction. Most heavy duty is carried out by hscPipe, so using C or C++ for SynPipe will not increase the efficiency. The most time-consuming part of SynPipe test is the injection of synthetic objects to the individual CCD image. During this process, the most computationally intensive part is carried out by GalSim using C++, so using C or C++ for SynPipe here will not increase the efficiency much. At the
end, we would like to emphasize that the design of SynPipe currently focuses on testing data reduction in the most realistic way, in the future, we will gradually work on increase the efficiency of SynPipe.

Following the suggestion by the referee, we now try to explain this better in Section 2.2, Section 3.1, and Section 3.2.3.

---

3. Missing Definitions

The referee is correct that some of these basic definitions should be explained better.

Most of these concepts are explained and discussed in great details in the hscPipe paper (Bosch+2017) and the Data Release 1 Paper (Aihara+2017). We emphasize the basic definitions of these values at the end of the introduction section and point the authors to the Bosch+2017 paper for more details.

1. HSC pipeline performs multi-step photometric calibrations against the Pan-STARRS system, hence all the fluxes and magnitudes measured by hscPipe are already calibrated into the AB magnitude system. Therefore, the pixel values on the calibrated single-Visit and coadd images can be directly related to fluxes in physical unit, and can be converted into

magnitudes using a simple calibration zeropoint (27.0). Users of HSC survey data should not worry that the definition of magnitude is affected by filter or exposure time. Through out this work, we assume that all magnitudes of synthetic objects are also in AB system, and they are converted into fluxes using the same zeropoint.

2. The CCD processing provides us calibrated exposures that include both the detrended, background-subtracted image and an image containing per-pixel estimate of the variance. The following stacking process also takes the variance information into account, uses the inverse of the mean variance of the input images to set the weights for stacking, and calculates the per-pixel variance for the coadd images. Both the PSF and CModel photometry take these variance images into account, and provide meaningful estimations of flux error for the coadd images.

    - The HSC pipeline uses matched-filter method to estimate the PSF fluxes. Per-pixel variance information is taken into account when estimating the flux error. Please see Section 4.9.5 of Bosch+2017 for more technical details.
    - The HSC CModel photometry is modified and improved based on the SDSS version. It fits linear combination of two basic parametric models to the 2-D flux distribution of galaxy after taking PSF convolution into account. It has the advantage of providing consistent multiband photometry for extended objects and having reliable performance for very faint object. Please see Section 4.9.9 and Appendix 2 of Bosch+2017 for technical details of the CModel algorithm.
    - For CModel photometry, it is optional to use per-pixel inverse square root of the variance information as weights during the fitting (when usePixelWeights=True). By default, usePixelWeights=False, and a constant weight is used for an object. Again we will refer the reader to Bosch+2017 for more details about this.

3. We understand the referee's concern, and it is true that based on statistics of photons, we could have more clear definition of signal-to-noise ratio (S/N). However, in practice, it is not realistic to use "photon number" for statistics or separately account for the noise from readout, dark, and sky background. Each CCD on the camera has four amplifiers that have slightly different effective gain; individual CCD images need to be detrended, calibrated, and background subtracted before stacking; single exposures with different seeings are warped before they are coadded into a final image using appropriate weighting. All these processes make it difficult to trace back to the original photon statistics. For instance, on the coadd image, there are covariances between adjacent pixels due to the warping process. On the other hand, hscPipe calibrates the pixels values on both coadd and single-Visit images to

physically meaningful flux scale, and provides corresponding variance information that reflect the per-pixel uncertainties after all these complex reduction steps. Based on this design, scientific user of HSC data can use the flux value and its uncertainty from hscPipe to estimate the S/N of the object under certain photometric method.

  - Throughout this work, S/N is defined as the ration between a flux measurement and its uncertainty (flux / flux_err). Same definition is adopted by both Bosch+2017 (hscPipe paper) and Aihara+2017 (HSC DR1 paper), along with a series of science papers.
  - Under this definition, 5 sigma detection corresponds to a S/N=5 flux measurement under certain photometric method. The same concept is adopted by both Bosch+2017 (hscPipe paper) and Aihara+2017 (HSC DR1 paper), along with a series of science papers.
  - Both the PSF and CModel photometry algorithms convolve a model for the true object's morphology with the effective PSF model at the position of the object, which maximizes S/N (see Bosch+2017)

- The changes of text related to this comment are
  - At the end of Introduction section, we make it clear that we are using AB system, and explain the definiton of S/N and Nxsigma detection.
  - At the end of Section 2.2, we provide more details related to the flux measurements for PSF and CModel photometry. We also refer the reader to Bosch+2017 for more technical details.
  - Section 6.4: slightly more details about the usePixelWeights option and reference to hscPipe paper.

---

4. CModel algorithm

We agree with the referee that the details of CModel algorithm should be explained better. However, limited by the scope of this work, we can not afford to show all the technical details. As mentioned, we expanded the description in Section 2.2, and points the readers to the relevant sections in Bosch+2017.

- Related to the referee's detailed comments:

  - We should clearly point out that good performance of cModel photometry is **not** a

natural consequence of hscPipe, and hscPipe **does not** use GalSim model for galaxy photometry. These are two completely independent software, and SynPipe interacts with them at different stages of the test. In our tests, the synthetic galaxy is described using single-Sersic profile, and is generated using GalSim. For hscPipe, we test CModel photometry which uses very different model assumptions and algorithms for galaxy photometry. It first fits an exponential model to the image with both the shape and the amplitude free. Then it fits a de Vaucouleurs model in the same way. As last step, it fits both models simultaneously, keeping their ellipses fixed at the results from the previous two fits, and allowing only the two amplitudes to vary. Also, in CModel, both of these models are actually described by multi-Gaussian approximations to allow fast convolution (e.g. Hog & Lang 2013, Sheldon 2014). The CModel algorithm is not designed to model detailed flux distributions of galaxies accurately, but to provide robust and efficient photometry for large amount of galaxies, especially for the poorly-resolved and/or low S/N ones.
    - Although these tests are much more than just "internal consistency", we should mention that 1) single-Sersic is still oversimplified model for real galaxies; 2) we assume that there is no error is PSF modeling. These are the limitations of our current tests, and we mention them in the discussion section now.
    - About the central pixel value of Sersic model, please refer to Rowe et al. (2015; http://www.sciencedirect.com/science/article/pii/S221313371500013X) for details of this code. The Section 5.3 and Section 6 explain the integration over pixels and image rendering method in great details. We should point out that synthetic galaxies generated by GalSim have been used for various photometric and weak lensing tests. The algorithms involved should be considered well tested.

- We modify the text in the following way:

    - Section 3.2.1: We point the readers to the relevant sections in Rowe+2015 for more details.
    - At the end of Section 4.2, we clearly point out that CModel uses very different model assumptions, hence we are not only testing the internal consistency of CModel pipeline.
    - In the discussion section, we add discussion about the limitation of our single-Sersic model test.

---

5. Fig 4 and 9.

We understand the referee's point that random noises from photon statistics could be useful for comparison.

However, as we pointed out earlier, given the design of the camera, observational strategy, and the complex processes of data reduction, it is impossible to make statistics based on the "input photon number". Each pixel on the coadd image consists of information from CCDs with different characteristics and from separated exposures under different seeing and sky conditions. These individual exposures are first detrended, calibrated, background subtracted, and warped before they are coadded using weights that are related to their
variance information. Given these reasons, it is not realistic to conduct tests using photo number and random statistical errors as the referee requested.

Meanwhile, we should point out that, since hscPipe (and LSSTpipe in the future) is designed to provide high-level science-ready results for users, it carefully calibrates the fluxes into AB magnitude system (absolute flux level), which is physically more meaningful and convenient. In addition, the hscPipe tries to provide per-pixel variance information for both single-Visit and coadd images, which includes statistical uncertainties from sources mentioned and not mentioned by the referee (e.g. correction of brighter-fatter effect, crosstalk, cosmic-ray removal, number of single-visit images that go into the coadd etc.). For coadd image, the variance plane provided by hscPipe should be treated as a summary of all statistical errors, and all the flux error in hscPipe is estimated based on it. Notice that, for both PSF and CModel photometry, the estimation of statistical uncertainty is not as straightforward as for aperture photometry. And it is not realistic for us to provide independent estimations for each object.

As mentioned earlier, such definition of S/N has been adopted within the HSC SSP collaboration, and has been used for multiple technical and scientific purposes. Therefore, we think the S/N plotted in Fig 4 and Fig 9 are the most meaningful and useful ones for HSC data user.

- We slightly modify the text to explain the definition of S/N and flux error better:
    - In Section 5.1.1, we explain the PSF flux error in more details.
    - In Section 5.2.1, we explain that for extended objects modeled by CModel, definition of S/N is not as straightforward as point sources. Here we choose to use the CModel flux and flux error to simply demonstrate the typical "significancy" of galaxies at fixed magnitude, and show that hscPipe can reliably detect and measure flux for galaxies down to certain magnitude. We also mention that uncertainty of CModel flux is dominated by systematics, and here the flux error only summarises the statistical ones. Therefore, our tests should provide more realistic estimation of average flux error

(including both systematic and statistical ones) for galaxy at certain magnitude.

6. Pure statistical errors in figures.

We agree with the referee that it is very useful to compare with the statistical uncertainties. However, as we explained for comment #5, "pure statistical random errors" based on photon statistics are not very practical to get from hscPipe.

Instead, the flux error from hscPipe for both PSF and CModel photometry are relied on the detailed variance information that summarize most of the statistical errors. These flux or magnitude error are also provided to the HSC data user as the most frequently used information on photometric uncertainties, it is useful to compare them with the results of our SynPipe tests.

On Fig 5, 6, 7, 10, 11, 12, and 15, we have overplot a pair of grey dashed lines that outline the running median of statistical uncertainties measured by the hscPipe (in positive and negative form). For magnitude, we simply use the magnitude error provided by hscPipe, for color, we add the magnitude uncertainties for both bands in quadrature.

Given the realistic situation, we think these comparisons cover the requirements by the referee, although are not exactly the same. We did not specifically mentioned these comparisions because the results are exactly as we expected based on the algorithms themselves: for PSF photometry, the statistical uncertainties from hscPipe are similar but still slightly lower compared to the median uncertainties from SynPipe test; For CModel, the statistical flux errors from hscPipe are smaller compared to the uncertainties from SynPipe test as systematic uncertainties (model-dependent errors) dominate the real error budget of CModel photometry.

We now briefly mention these comparisons in each section to draw readers attention to these information.

7. Precision and accuracy of PSF magnitude

We believe this comment is also covered by the previous one. For PSF magnitude, the statistical uncertainties obtained by hscPipe is similar but still slightly lower than the photometric uncertainties based on the SynPipe test. As mentioned, due to the correlated noise from the image warping and coadding process, the statistical uncertainties of PSF flux measured by hscPipe is likely to be over-optimistic.

Relevant text has been added to Section 5.1.2 and 5.1.3.

---

8. p. 7-8: "High level reduction pipeline": as stated more clearly in the hscPipe technical paper (Bosch et

al. 2017). Low-level reduction involves basic image detrending and calibration; high-level processes involves the ones that generate coadded images and science-ready catalogs. Now use the same explanation here to be consistent with the technical paper and provide more details.

---

9. "purity and completeness"

We agree with the referee that the purity and completeness tests are very important for survey like this, and SynPipe can be a useful tool for this purpose.

At the same time, careful tests for this purpose are beyond the scope of this paper, where we simply want to demonstrate the application of SynPipe and provide simple photometric benchmark tests for stars and galaxies.

- Following the suggestions from the referee, we modify the text with more discussions on this, and here are the main points:
    - Section 4.2: We explain that it is beyond the scope of this work to look into this carefully. Detection limits and completeness are normally defined for point sources. As we show using the PSF S/N, it depends on the seeing condition and show spatial changes which requires more careful designs and higher density of fake stars. Detection completeness

of point sources for HSC DR1 is investigated in Aihara+2017 (see Section 5.7) using a different method.
    - Misclassification is briefly discussed in Section 6.2. The design of our tests and the number density of synthetic galaxies are not perfect to thoroughly test the star-galaxy separation. So our results just demonstrate the capability and briefly discuss the current situation for star/galaxy separation. We find that, right now, hscPipe tends to provide a very complete sample for galaxy but with considerable fraction of star contamination at the faint end. In Aihara+2017, the star/galaxy separation issue is discussed in Section 5.7, where they draw the same conclusion using different method.
    - The current method adopted by hscPipe focuses on the magnitude different between PSF and CModel photometry and is known to have limitations (see Section 4.9.10 of Bosch+2017) and issues (e.g. the one related to the current CModel algorithm; see Appendix 2 in Bosch+2017). All these conclusions are reflected in our simple test. On the other hand, hscPipe team is working on replacing the star/galaxy separation algorithm with a much improved one (based on extreme deconvolution method, also see Section of 4.9.10 of Bosch+2017). We will test its performance after the next data release.
    - We expand the discussion in Section 6.2 to provide more information about star/galaxy separation tests and emphasize the importance of such tests as suggested by the referee.

Minor comments:

1. p.7: should be 'mas' not 'max', corrected.

2. p.10, about 'rectangular cutout':

    - the images are stored in a 2-D array, hence has a rectangular nature to it. And, it is much easier to convolve the model and shift the 2-D rectangular array before we add them to the HSC coadd images. When we generate the model image array using GalSim, we try to make sure that all fluxes of the model is contained in the image. For galaxies, we cutout the model at 10 x Re of the Sersic model. This is fine for most galaxies, but still miss small fraction of light for galaxy with very high Sersic index. First, we try to avoid model with very high Sersic index in the input catalog, as they are often problematic models for real galaxies; Second, given the magnitude distribution of the model galaxies (most have i>23.0 mag), the results we are showing here are not sensitive to the flux

distributions outside 10 x Re. The referee is correct that estimation of R_e could be sensitive to the very outskirt when Sersic index is high, but we'd like to point out that the CModel method used in hscPipe does not use any pixel with S/N < 5, and only provides n=4 model fits (dev) to the "footprint" of the galaxy. We do not expect accurate R_e estimates for high Sersic index galaxies in the beginning.
  - We add a footnote to briefly explain this choice.

3. p.11, about "noise":

   - It is Poisson noise, and we state that clearly in the text.

4. p.13:

   - We think the location and context of this figure makes it very clear that it is for the input parameters. Now we state that clearly in the caption.

5. Fig. 4 and 5: has been corrected.

6. p.22:

   - For stars, first, fraction of stars with such extreme colors are very low. The discrepancy can barely be noticed on the comparisons of 1-D color distributions. We think the likely reasons for this are:
     - There are still scattered light on the y-band images that are not accounted for. This known problem can affect the accuracy of y-band photometry and is discussed in Section 5.8.14 in Aihara+2017.
     - Background level is much higher in z and y-band, and background subtraction could suffer from higher uncertainties. Right now, SynPipe can still not test the impact of background subtraction on photometry, and it will be investigated in the future.
     - We modify text in Section 5.1.3 to briefly explain this.
   - As for galaxies, the same background and y-band scatter light issues also exist. However, the dynamical ranges for i-z and z-y colors are narrower. Also the CModel color uncertainties are higher than the PSF ones. Both of these factors make it harder to spot discrepancies of color distributions. However, we still see discrepancies between the input and recovered z-y colors.
   - We expand the discussion in Section 5.2.3 a little to address this issue.

7. p.22

- The referee is correct that, for highly blended stars, the biases on their PSF colors are still quite big. What we meant here is that the biases for PSF magnitude are much worse as the total fluxes of blended stars are highly underestimated. As for PSF colors, although there are still long tails that indicate the biases, at least the peak of the distributions are around zero now. We adjust this sentence to make it more clear.
- As for general comments for PSF photometry of highly blended stars. Unfortunately we are not perfectly clear what causes this issue. It is clearly related to the performance of deblender in hscPipe, which is know to still have issue and is being improved now (please see Section 4.8 of Bosch+2017). Given the depth of HSC survey, the situation for object deblending becomes much more tricky. At this point, we have to leave this as a caveat. PSF magnitude is still considered the best photometry for point sources, but we will remind the users to avoid using it for highly blended objects. The text has been modified to reflect this point.

8. P.29

    - Actually it should be (z-y) color, and has been corrected.

9. Astrometric Calibration Fig 14

    - HSC image has pixel scale of 0.168 arcsec/pixel. We use arcsec as unit since when evaluating the astrometric calibration for HSC DR1, arcsec or mas are used in Aihara+2017. Information about pixel scale is added to the caption of the figure and the text.

10. the definition of highly blended

    - The referee is correct, that we should use consistent definition for highly blended objects, which is b>0.05. We have make it consistent across the draft.

11. Table 1, 2, 3, and 4

    - The statistical uncertainty is actually the standard deviation of (mag_output - mag_input) within each input magnitude bin. This is the what we used to describe "precision". For accuracy, we use the mean value of magnitude difference in each input magnitude bin.
    - All the captions of tables have been updated to make it more clear.