Evaluation of coadding and single-frame photometry methods

Dustin Lang

2017–09–08

This report describes some simple experiments to evaluate the signal-to-noise performance of coadding and single-frame photometry methods. The script for these experiments is publicly available[1] and uses the *Tractor* code[2] to generate synthetic images and photometer them.

The main question to be answered in this experiment is how much signal-to-noise is lost (how much extra photometric error is introduced) when individual exposures are coadded before performing photometry, as compared to performing photometry on single exposures and averaging the results at "catalog level".

This experiment assumes we are doing forced photometry on two ground-based images. That is, it assumes that we have a high-quality catalog (eg, from Euclid VIS) in which we have detected a single point source. It assumes that we have correctly calibrated the astrometry of the ground-based images, so we know exactly where in pixel coordinates the point source will be found; and it also assumes that we have a correct point-spread function (PSF) model for each image. It assumes the two image were taken with the same bandpass filter, and asks how well we can measure the flux of the point source in the images. We will vary the PSFs and the per-pixel noise. In typical ground-based images, the per-pixel noise is due primarily to the Poisson distribution of the sky background, which is well approximated by Gaussian noise; readout noise adds to this, but is usually fairly small in comparison. We will assume that the images have been photometrically calibrated so that they are all in the same units; thus longer exposures will result in the same measured flux values but will have smaller per-pixel noise and smaller errors in the measured fluxes.

We examine four cases, building up to the most realistic:

- **Case 1**: the two images have the same per-pixel noise, and the same PSF
- **Case 2**: the two images have different levels of per-pixel noise, but the same PSF
- **Case 3**: the two images have the same per-pixel noise but different PSFs
- **Case 4**: the two images have different per-pixel noise and PSF (but the same per-image signal-to-noise).

For each of these four cases, we show results for three different photometry methods. In all cases, we are doing model-based forced photometry: we are fitting for the flux where our pixel-space image model best matches the observed pixels, given the per-pixel noise in the images.

The methods are:

- **Simultaneous fitting**: We do a single fit for the flux that produces the best match to both images at the same time. That is, we are finding the flux that produces the

[1] *https://github.com/dstndstn/euclid/blob/master/coadd.py*
[2] *https://github.com/dstndstn/tractor*

smallest sum of chi-squared differences between the model and the images, summing the chi-squared of both images.

- **Single-frame averaging**: We fit for the flux in each image independently, and then compute a weighted average of the fluxes. That is, we compute a best-fit flux for each image, and then average the flux measurements in "catalog space".
- **Coadding**: We add together the pixels of the two images (with weights), and then fit for the flux that best matches the coadded image. When doing the fit, we compute the correct PSF model for the coadded images. We test different *coadd weights* α, where the coadd is $C = \alpha A + (1 - \alpha)B$ for images A and B, and the sum is pixelwise.

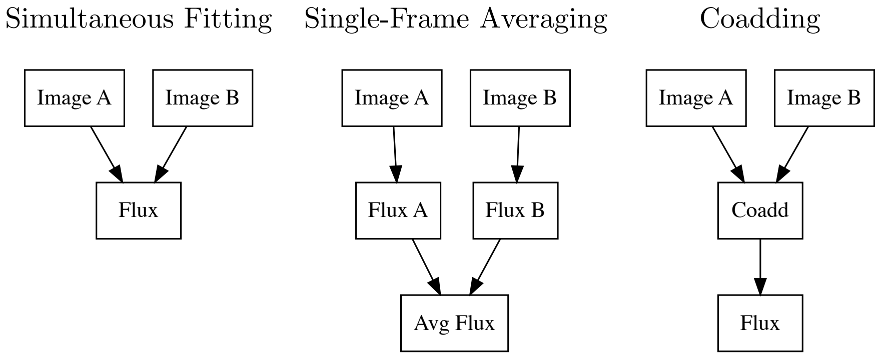

Figure 1: Schematics of the three photometry methods presented in this report. In *Simultaneous Fitting*, the flux is optimized to minimize the sum of chi-squared values between the model and data, summed over both images. In *Single-Frame Averaging*, we measure a flux for each image (by minimizing chi-squared in the single image), then weighted-average the two flux measurements. In *Coadding*, we compute a coadd (average the images), then fit a flux to the coadd.

Note that the first method, Simultaneous Fitting, saturates the Cramér–Rao bound, meaning that it extracts all the available information in the images. It is the most expensive in terms of computation time and memory; it requires that all the images to be measured are in memory at once. We include it in this experiment only for comparison; one of the results of this report is that Single-Frame Averaging always performs exactly the same. This is because the flux measurements are sufficient statistics and are distributed as Gaussians, so they can be combined at "catalog level" with simple equations and no loss of information. Single-Frame Averaging has the strong advantage that each exposure is photometered independently, so it is computationally cheap, can be distributed, and is embarrassingly parallel.

1 Results

For each of the four cases, we perform two analyses. In the first analysis, we compute the expected signal-to-noise performance of the different methods. Since the Coadding method has a parameter α, we evaluate the performance at all different parameter values. In the second analysis, we produce simulated images, and then run $10,000$ trials where we add random Gaussian noise to the two images and run each of the three photometry methods. We then show the distributions of measured fluxes, to confirm that they indeed match the analytically predicted performance.

Case 1: Same noise, same PSF

Figure 2 shows results for Case 1. In this case, all three methods produce exactly the same results. Since the two images have the same PSF and per-pixel noise, the best coadd weight $\alpha = 0.5$.

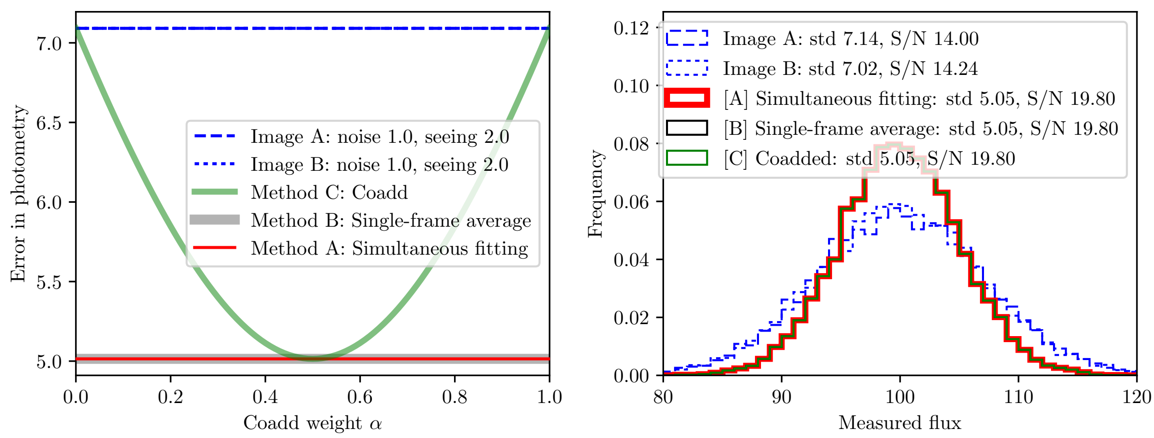

Figure 2: Results for Case 1 (same PSF, same per-pixel noise). **Left**: the expected photometric uncertainty (error) for the different methods, as a function of the coadd weighting. (Smaller is better.) At the top of the plot, the two individual exposures, which have the same noise and PSF, give the same photometric error. When the coadd weight is $\alpha = 0$ or 1, the error in the coadd photometry equals that of the individual images. With $\alpha = 0.5$, the coadd method performs as well as the other two methods, which produce errors that are a factor of $\sqrt{2}$ smaller than the individual exposures.
Right: Results of simulations adding noise to synthetic images and running each of the photometry methods. The measured fluxes are shown for $10,000$ trials. In this Case, the flux measurements on the individual exposures have a standard deviation of about 7, while all three of the other methods produce exactly the same results, with a standard deviation $1/\sqrt{2}$ as large; roughly 5. Note that these results are entirely consistent with the analytically-computed expected performance values.

Case 2: Different noise, same PSF

In Case 2, one image has half the per-pixel noise as the other. Results are shown in Figure 3. In this case, again, each of the three methods produce the same results. The best coadd weight α is found to correspond to inverse-variance weighting of the images: Image A, with half the noise, is given a weight four times greater than that of Image B.

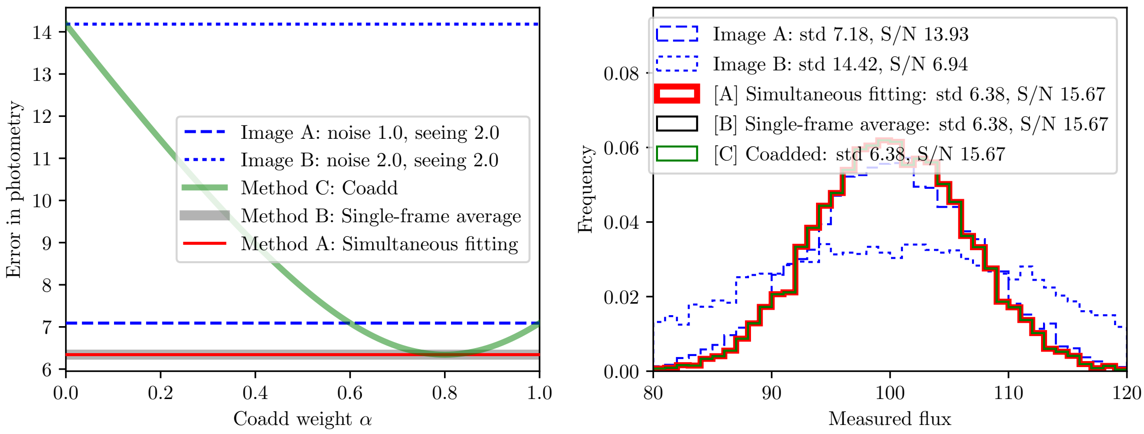

Figure 3: Results for Case 2 (same PSF, different per-pixel noise). **Left**: the expected photometric uncertainty (error) for the different methods, as a function of the coadd weighting. Image A has half the per-pixel noise as Image B, so provides photometric measurements with half the uncertainty. The coadd weight that results in the best performance is $\alpha = 0.8$; this corresponds to inverse-variance weighting the images.
Right: Results of simulations where we add noise to synthetic images and run each of the photometry methods $10,000$ times. The measured fluxes for the three methods are all identical, as in Case 1.

Case 3: Same noise, different PSF

In Case 3, one image has half the seeing size (full-width at half-max, FWHM) as the other. Results are shown in Figure 4. Here, the coadding method performs slightly worse than the other methods. This can be explained by the fact that optimal source detection requires the use of a “matched filter”—convolution by the correct PSF—but in the coadded image, the two images with different PSFs have been mixed together. In other words, the matched filters for the two images would place different relative weights on the pixels; the PSF of the coadd places intermediate weights on the pixels, so is less than optimal for both input images.

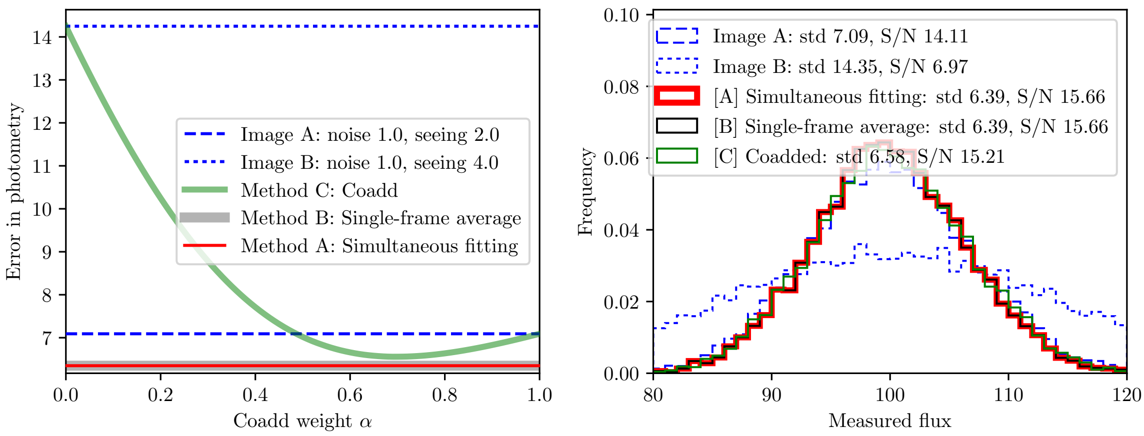

Figure 4: Results for Case 3 (same per-pixel noise, different PSF). **Left**: the expected photometric uncertainty (error) for the different methods, as a function of the coadd weighting. Image A has half the seeing size as Image B, and provides photometric measurements with half the uncertainty. The coadd weight that results in the best performance is $\alpha \sim 0.7$. Even with the best coadd weight, the coadd method perform slightly worse than the other methods. Since the majority of the signal-to-noise is carried by Image A, the difference is quite small.
Right: Results of simulations where we add noise to synthetic images and run each of the photometry methods $10,000$ times. The measured fluxes for the three methods are as predicted: Simultaneous Fitting and Single-Frame Averaging still perform exactly the same, but Coadding performs slightly worse.

Case 4: Different noise, different PSF

In Case 4, one image has half the seeing size (FWHM) as the other, but the second image has half as much noise. As such, the images are equally sensitive. This is a realistic case because many ground-based surveys will take longer exposures (leading to less per-pixel noise) when the seeing is worse, in order to produce a survey of uniform depth. The results for Case 4 are shown in Figure 5. In this case, the coadd method produces significantly worse results. Perhaps surprisingly, the best coadd weight α is not 0.5—equal weighting for the equally-sensitive images—but is roughly 1/3. The loss of signal-to-noise from using the Coadding method is roughly 5%, which may not seem like much, but observing time goes with signal-to-noise squared, so this corresponds to more than a 10% increase in required observing time.

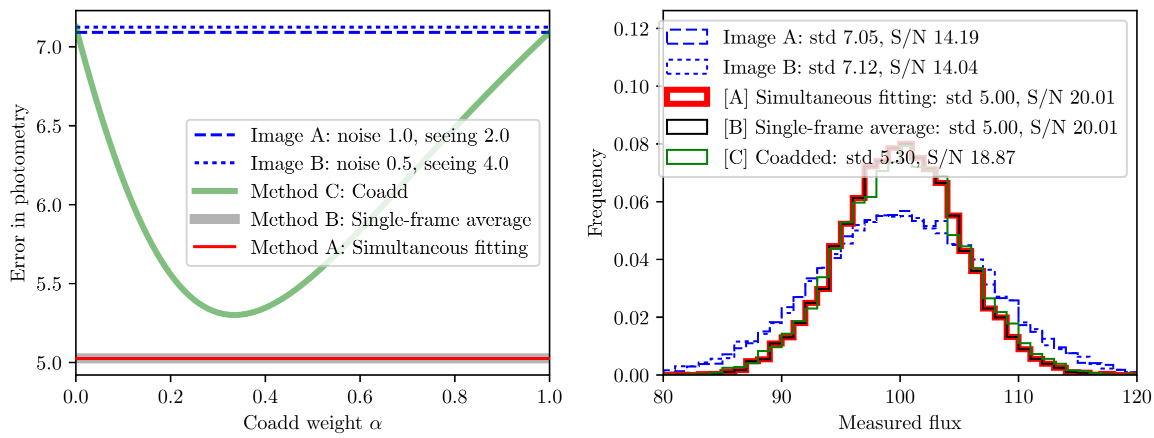

Figure 5: Results for Case 4 (different per-pixel noise, different PSF, same signal-to-noise). **Left**: the expected photometric uncertainty (error) for the different methods, as a function of the coadd weighting. Image A has half the seeing size as Image B, but more per-pixel noise, and thus provides photometric measurements with approximately equal uncertainty. (The signal-to-noise values are not exactly equal due to the effects of detector pixelization.) The coadd weight that results in the best performance is $\alpha \sim 0.33$. Even with the best coadd weight, the coadd method perform significantly worse than the other methods.

Right: Results of simulations where we add noise to synthetic images and run each of the photometry methods $10,000$ times. The measured fluxes for the three methods are as predicted: Simultaneous Fitting and Single-Frame Averaging still perform exactly the same, but Coadding performs worse.

Conclusions

These experiments show that for the realistic case (Case 4) of different PSFs and different per-pixel noise levels in ground-based images, creating a coadd and then photometering it results in a significant (5%) loss of signal-to-noise, or more than 10% increase in required telescope time. On the other hand, performing photometry on individual exposures and then computing weighted averages at "catalog level" is shown to exactly match the signal-to-noise performance of simultaneous fitting, and saturates the Cramér–Rao bound (that is, extracts all available information).

Photometering individual exposures also has the programmatic advantage that it is highly distributed: each input image is processed independently, and updating the final photometric measurement when a new images is taken is simple and fast. A broader advantage is that per-exposure information (such as spatially-varying filter response) can be tagged on to single-exposure measurements and incorporated into more advanced downstream processing, in which one would not summarize the single-exposure measurements with a catalog-level average, but instead forward-model the single-exposure measurements.

The code for this experiment is freely available.[3]

[3] *https://github.com/dstndstn/euclid/*