DMTN-068:
Lossy Compression Working Group

Goal

Determine whether some **additional** image products can be retained at a modest cost rather than requiring on-the-fly reconstitution from RAW.

Not meant to replace reprocessing where use-case demands the original image fidelity.

Methodology I

Prototype Lossy Compression implemented by separating Loss from Compression within *ci_hsc*

Loss: quantization applied to reduced single-epoch science and weight images recast as an integer representation (FITS with BSCALE)

Quantization (q): BSCALE = RMS / q

Compression:

evaluate lossless algorithms as these work well with integers

Methodology II

Single-epoch (equivalent to PVIs):

- Examine images to assess difference between quantized images and never-quantized versions
- Compare measurements (ForcedPhoto) made from never-quantized imaged and quantized single-epoch images

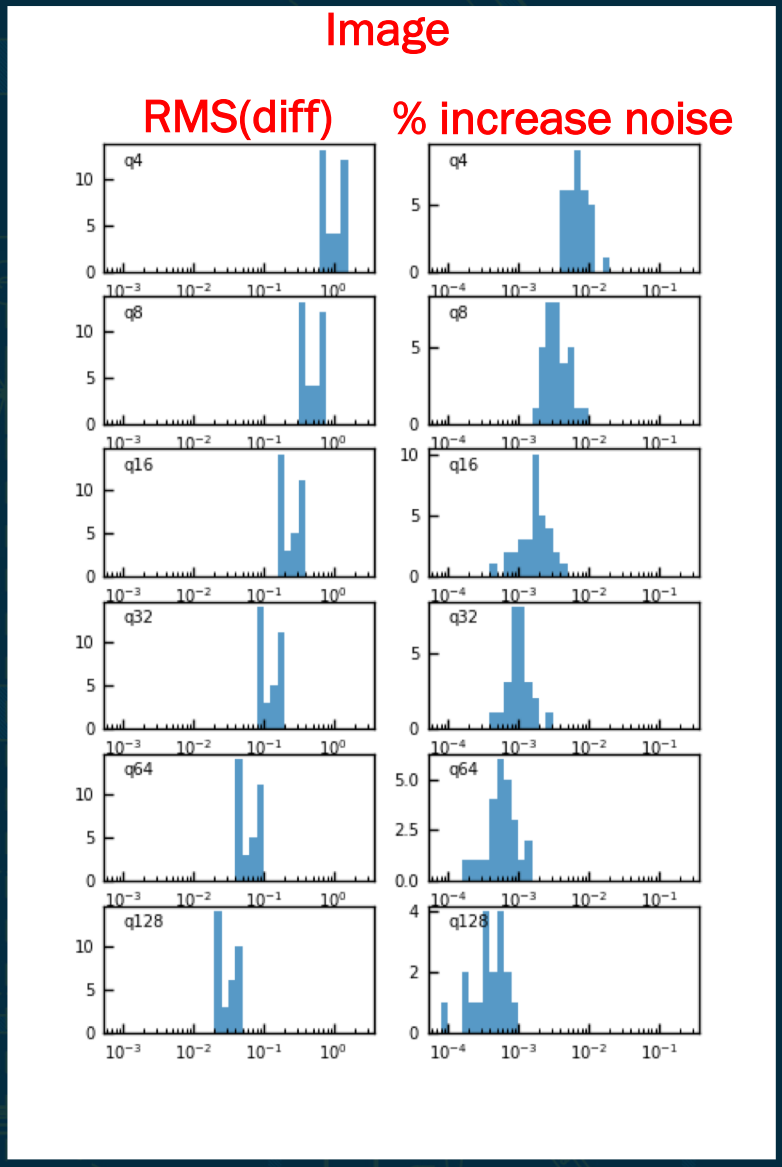
COADD: Construct COADDed patch from both never-quantized images and then from quantized images

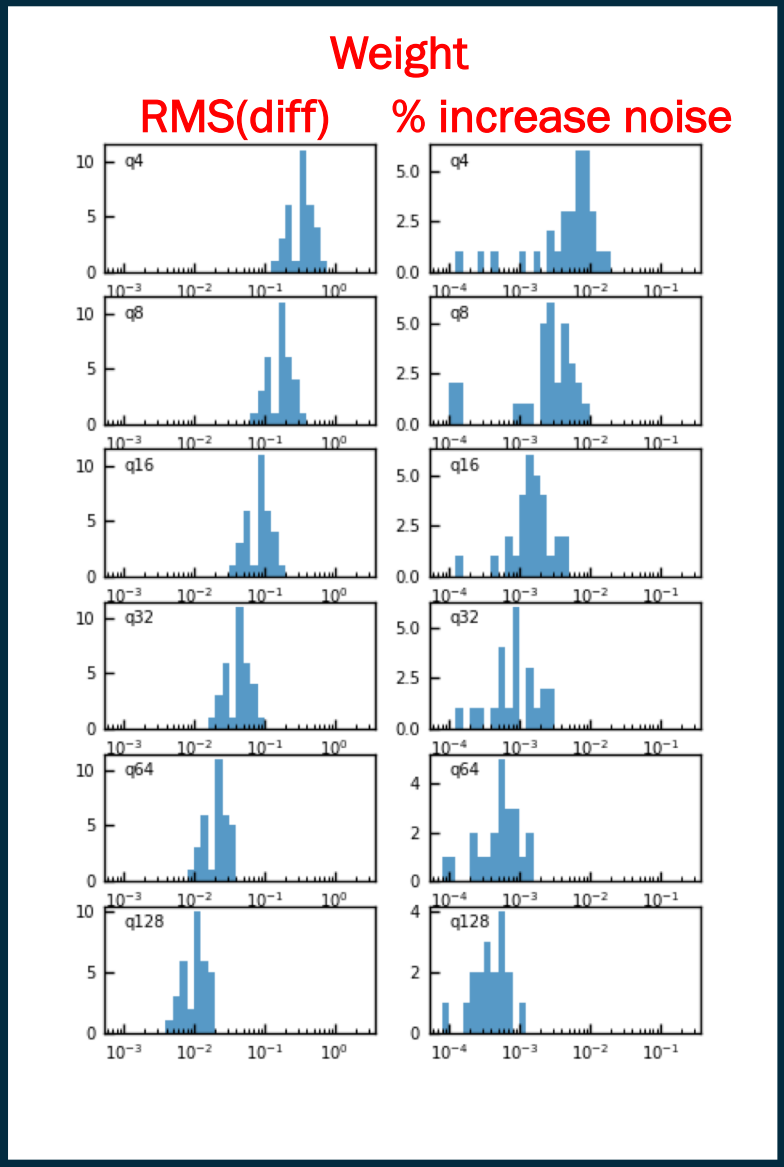
- Repeat image and measurement comparisons

Small set:

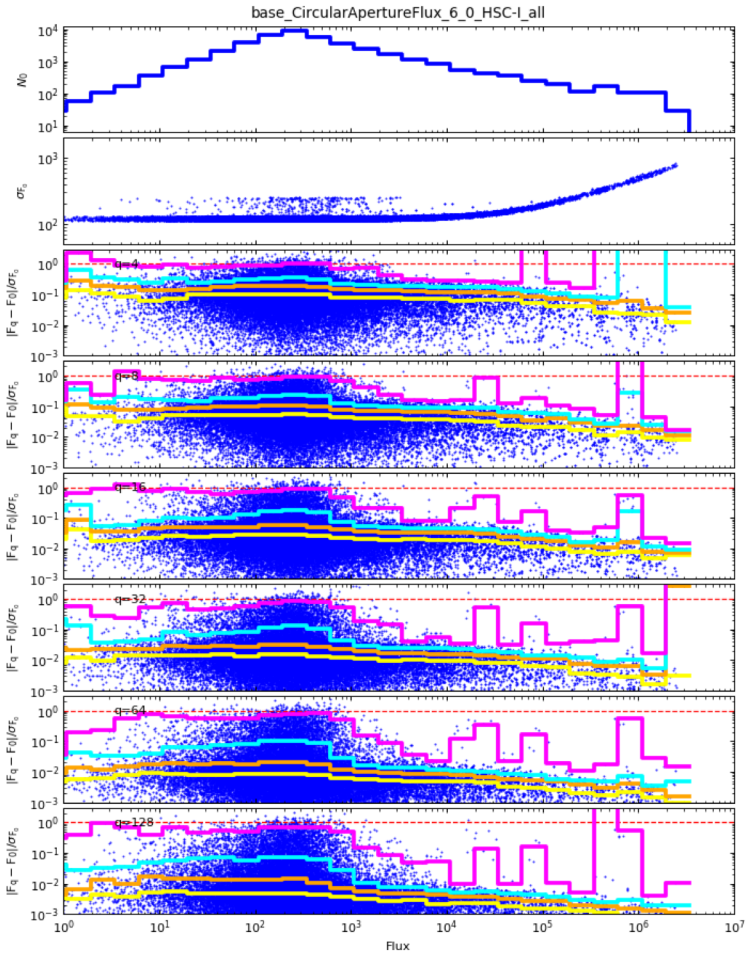
- 33 PVI CCDs
- COADD 1 patch @ 2-bands (typical depth ~5 images

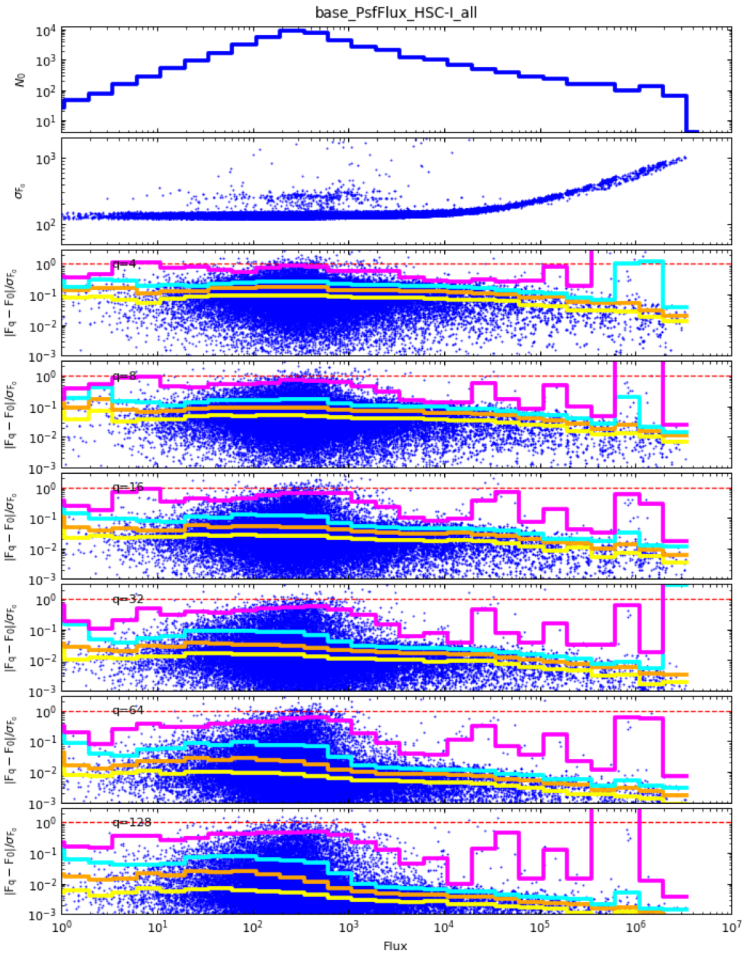
Image Difference



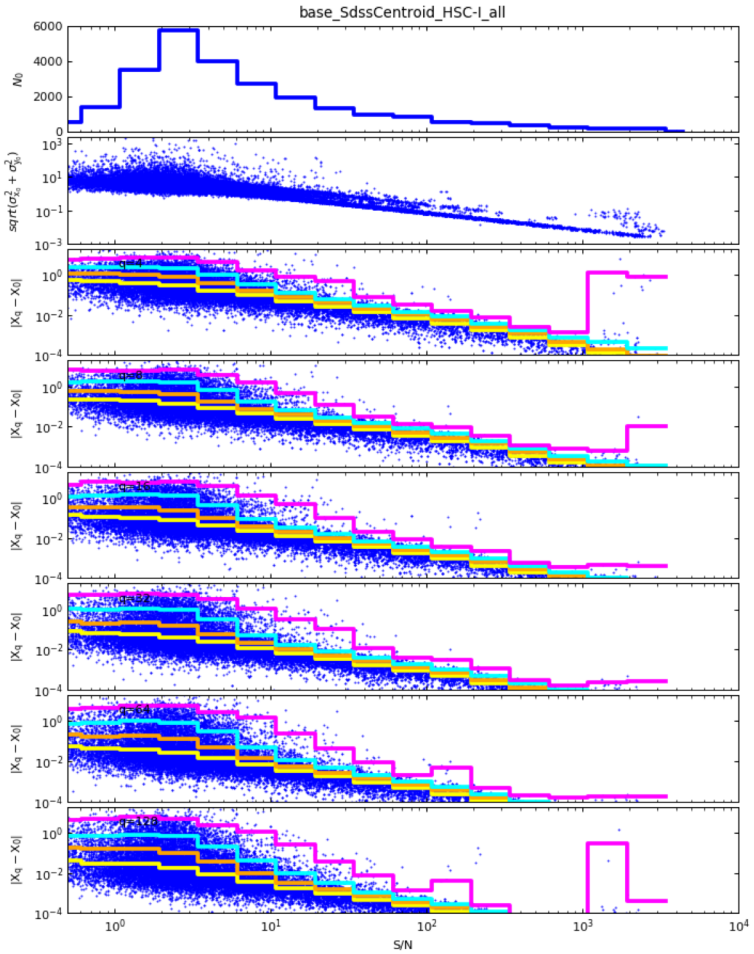


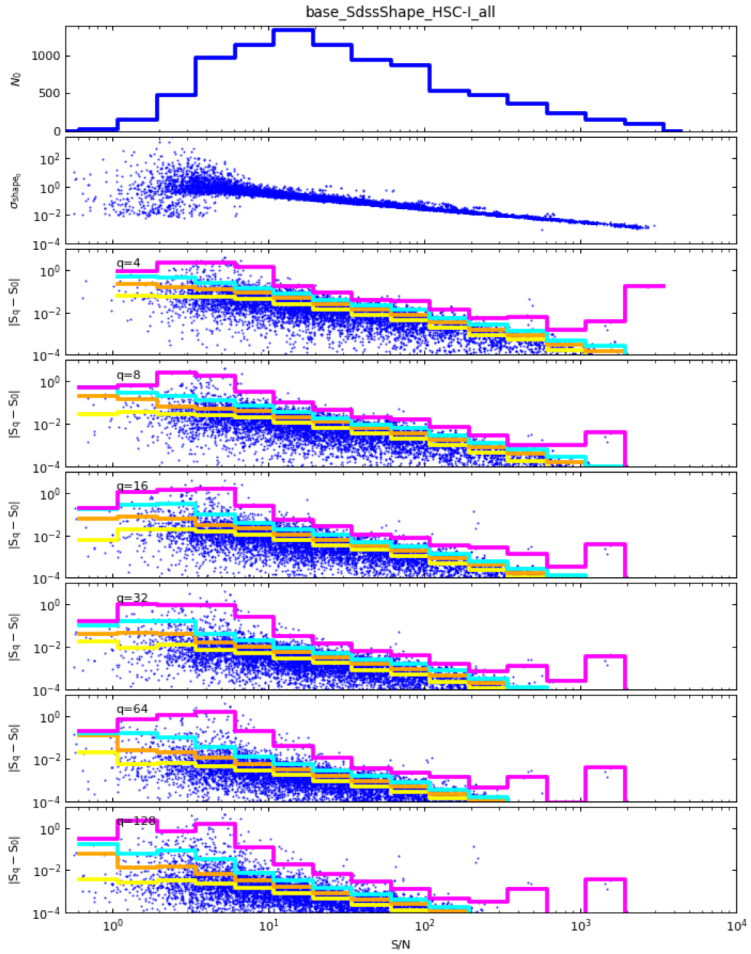
Flux Measurements (from PVIs)





Centroid & Shape from PVIs





Algorithmic Benchmarks

TABLE 7: Compression Factor Achieved

q	gzip	pigz	bzip2	pbzip2	lbzip2	lz4	lzop	zstd	zstd*	xz	xz′
4	6.73	6.73	9.96	9.95	9.96	3.69	3.11	6.29	6.29	10.06	8.19
8	5.54	5.53	8.20	8.20	8.21	3.34	2.96	5.42	5.42	8.20	6.79
16	4.69	4.69	5.41	7.01	7.03	3.11	2.82	4.82	4.82	6.96	6.01
32	4.04	4.03	6.14	6.14	6.14	2.93	2.66	4.35	4.35	6.00	5.44
64	3.62	3.62	5.47	5.47	5.48	2.82	2.47	3.94	3.94	5.29	4.95
128	3.38	3.37	4.88	4.88	4.88	2.66	2.32	3.56	3.57	4.75	2.51
vanilla	1.71	1.71	1.80	1.80	1.80	1.50	1.49	1.72	1.72	1.87	1.80

TABLE 8: Time to Compress per File

q	gzip	pigz	bzip2	pbzip2	lbzip2	lz4	lzop	zstd	zstd*	xz	xz′
4	4.45	1.18	5.00	1.42	0.85	0.21	0.24	0.36	0.12	55.27	3.33
8	6.06	1.64	4.91	1.39	0.82	0.21	0.24	0.42	0.15	50.94	4.12
16	8.27	2.24	4.33	1.39	0.82	0.27	0.27	0.55	0.18	50.09	4.27
32	10.30	2.76	5.27	1.42	0.79	0.24	0.27	0.58	0.21	47.36	4.64
64	11.79	3.00	5.39	1.52	0.88	0.24	0.30	0.61	0.24	47.48	5.09
128	12.76	3.21	5.91	1.61	0.94	0.27	0.30	0.67	0.21	54.21	2.52
vanilla	3.36	0.97	8.94	2.79	1.58	0.15	0.12	0.30	0.15	34.00	15.15

TABLE 9: Time to Decompress per File

q	gzip	pigz	bzip2	pbzip2	lbzip2	lz4	lzop	zstd	zstd*	xz	xz′
4	0.21	0.24	2.30	1.21	1.27	0.15	0.18	0.27	0.24	0.82	1.06
8	0.24	0.27	2.33	1.12	1.24	0.18	0.18	0.27	0.27	0.97	1.27
16	0.27	0.27	2.02	1.12	1.21	0.18	0.18	0.27	0.24	1.09	1.33
32	0.30	0.30	2.42	1.24	1.24	0.15	0.18	0.27	0.24	1.27	1.42
64	0.30	0.30	2.42	1.27	1.09	0.18	0.21	0.27	0.27	1.42	1.48
128	0.30	0.33	2.82	1.30	1.24	0.24	0.21	0.30	0.27	1.52	0.67
vanilla	0.39	0.36	4.36	1.48	1.27	0.15	0.12	0.24	0.24	3.58	3.52

Conclusions/Recommendations

Single-epoch (equivalent to PVIs) q=16-32

- Less than 1/10 sigma in flux for S/N>5
- Less than 1/10 pixel centroid and shape S/N>10

COADD: not yet a realistic assessment compared to LSST Survey w/ 100-1000 image depth.

Do not recommend that LSST DRP catalogs reflect measurments made from quantized/compressed images.

Can achieve compression factor of 5-7 using BZIP or possibly XZ.

If speed becomes important each of these have multi-threaded variants.

Extras

TABLE 3: Maximum Flux Difference (in units of σ_{F_0}) for 90% of Objects

	PSF Flux				Aperture Flux			
q	S/N=3	S/N=5	S/N=10	S/N=100	S/N=3	S/N=5	S/N=10	S/N=100
4	0.266	0.234	0.208	0.171	0.332	0.269	0.208	0.171
8	0.161	0.136	0.110	0.087	0.220	0.170	0.113	0.089
16	0.106	0.078	0.056	0.044	0.163	0.113	0.058	0.044
32	0.074	0.050	0.031	0.022	0.117	0.068	0.033	0.022
64	0.067	0.043	0.018	0.011	0.097	0.064	0.019	0.011
128	0.049	0.027	0.011	0.006	0.067	0.042	0.014	0.006

TABLE 4: Maximum Centroid and Shape Difference for 90% of Objects

	Centroid				Shape			
q	S/N=3	S/N=5	S/N=10	S/N=100	S/N=3	S/N=5	S/N=10	S/N=100
4	1.791	0.898	0.269	0.013	0.370	0.231	0.112	0.010
8	1.364	0.590	0.129	0.007	0.183	0.118	0.058	0.005
16	1.063	0.375	0.066	0.003	0.255	0.083	0.030	0.002
32	0.778	0.276	0.039	0.002	0.136	0.039	0.016	0.001
64	0.704	0.258	0.036	0.001	0.080	0.030	0.009	0.001
128	0.540	0.169	0.029	<0.001	0.066	0.029	0.006	<0.001

TABLE 5: Maximum Flux Difference (in units of σ_{F_0}) for 90% of COADD Objects

	PSF Flux				Aperture Flux			
q	S/N=3	S/N=5	S/N=10	S/N=100	S/N=3	S/N=5	S/N=10	S/N=100
4	0.399	0.296	0.264	0.225	0.443	0.392	0.291	0.208
8	0.296	0.193	0.181	0.152	0.281	0.255	0.184	0.140
16	0.234	0.142	0.128	0.107	0.237	0.186	0.120	0.089
32	0.182	0.114	0.107	0.086	0.174	0.147	0.103	0.077
64	0.193	0.107	0.102	0.083	0.156	0.126	0.096	0.071
128	0.114	0.034	0.019	0.024	0.080	0.067	0.030	0.010

TABLE 6: Maximum Centroid and Shape Difference for 90% of COADD Objects

	Centroid				Shape			
q	S/N=3	S/N=5	S/N=10	S/N=100	S/N=3	S/N=5	S/N=10	S/N=100
4	1.857	1.077	0.414	0.020	0.155	0.181	0.120	0.008
8	1.185	0.677	0.281	0.015	0.144	0.113	0.098	0.006
16	1.094	0.506	0.191	0.011	0.116	0.071	0.056	0.004
32	0.856	0.353	0.142	0.009	0.075	0.057	0.049	0.003
64	0.827	0.284	0.114	0.008	0.070	0.053	0.045	0.004
128	0.652	0.156	0.073	0.006	0.029	0.010	0.005	>0.001