Title: Clean bias in ALMA Early Science images with a proposed 250m configuration

NAASC Memo # 2010-Oct-001

Mark Lacy & Remy Indebetouw

Date: October 26 2010

ABSTRACT

This memo describes results from simulations of clean bias in ALMA Early Science images made with a proposed 250m configuration, and some steps that may be taken to mitigate it.

1. Introduction

First formally recognized during survey projects with the VLA in the 1990s (White et al. 1997; Condon et al 1998), clean bias is an effect whereby cleaning of images with poor uv-plane coverage (typically "snapshots" - defined here as any observation too short for earth rotation to fill out the uv-coverage) can result in artificial changes to object fluxes and apparent image noise levels. It can be most easily understood in terms of inadvertent cleaning of bright image sidelobes, resulting in the subtraction of flux from the location of the real astronomical source (and this principle extends to the noise level in the image). Clean bias is strongly dependent on the dynamic range of the primary beam (peak to brightest sidelobes), and thus on the uv-plane coverage of the observations. Tests with simdata show that clean bias will be a very significant effect with ALMA Early Science (ES) data, particularly with snapshot observations. This memo describes the effects of clean bias on ES simulations, and outlines some potential mitigation strategies. Currently these simulations all use the currently proposed 250m ES configuration (see Appendix), but can easily be repeated for other ES configurations.

2. Simulations of clean bias

Figure 1 shows a series of simulations performed with short ES observations and default clean parameters, showing the clean bias effect on the noise becoming significant after only a few tens of clean iterations in the worst case of a 10min snapshot. Even the 1hr ES observation is significantly affected after about 100 iterations. It should be emphasized that clean bias also affects measured source fluxes, particularly if there are multiple sources in the field.

3. The ALMA ES beam and uv-plane coverage effects

Clean bias is a form of clean instability, and is dependent on the dynamic range of the beam. It is thus a strong function of both number and distribution of baselines, and of the degree to which earth rotation is able to fill out the uv-plane during an observation. Beam dynamic ranges for a range of exposure times using the ES 250m configuration and the closest matching FS configuration in size (number 6) are shown in Table 1, and two ES examples illustrated in Figure 2. In particular, we note that clean bias effects are likely to be negligible with the full science array in anything over a 1hr snapshot, or with the ES array in full synthesis (or near full-synthesis) observations, where the beam dynamic ranges are >10. Even so, it is clear from that for high fidelity imaging in full science a full synthesis is still worthwhile, the beam having a ten times higher dynamic range than a 10 minute snapshot. Note also the significant improvement in the ES beam dynamic range as the observation duration is increased from 30min to 1hr to 1.5hr, suggesting that increasing the maximum length of scheduling blocks even by a few tens of minutes is worthwhile in terms of the uv-coverage gained.

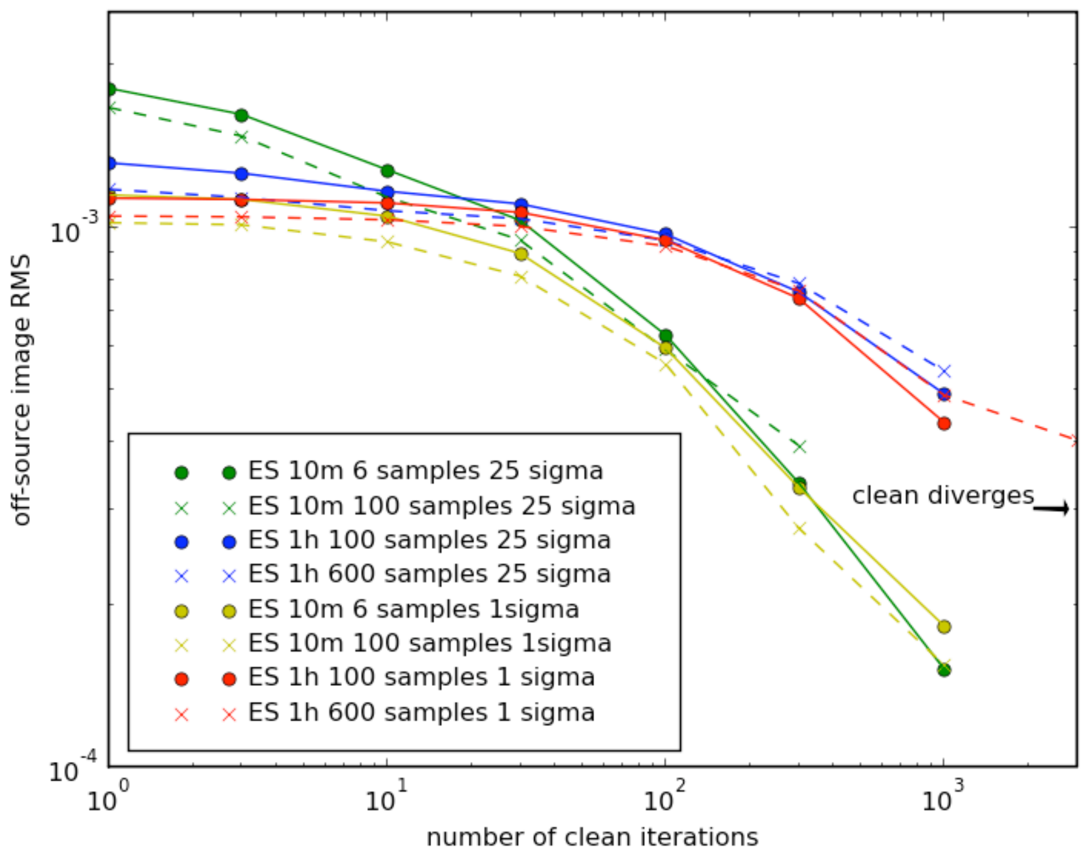

Figure 1: Clean bias simulations using the 250m ES configuration. Noise (image RMS, measured over 1/3 of the image, well within the PB, but away from the central point source) is plotted as a function of number of clean iterations (zero threshold). Green dots/solid green line: 10min ES observation with 6 samples (i.e. 100s integration time), green crosses/dashed green line: 10min ES observation sampled every 6s, blue dots/blue solid line: 1hr ES observation sampled 100 times, blue crosses/blue dashed line: 1hr ES observation sampled 600 times. All these simulations had a 25-sigma source at the center. Yellow/red symbols - as above but with a 1-sigma point source added. These simulations give the true noise level when niter is low (<10).

Table 1: Beam dynamic ranges (natural weighting, close to zenith)

Observation (345GHz)	Brightest sidelobe (beam normalized to unity)	Beam dynamic range (1/peak sidelobe)
10min ES250m (16 antennas)	0.45	2.2
30min ES250m	0.40	2.5
1hr ES250m	0.28	3.6
1.5 hr ES250m	0.20	5.0
4hr ES250m	0.12	8.3
8hr ES250m	0.092	10.9
10min FS (50 antennas, cfg06)	0.204	4.9
1hr FS	0.093	10.8
8hr FS	0.023	44

4. Mitigation strategies

Good practice when using clean will mitigate the effects of clean bias. Boxing the sources or placing a 3-sigma threshold on the clean will improve the results, with the best results coming from applying both a mask and a threshold (Figure 3). Adjusting the clean gain makes only a small improvement. Note also that although these effects are strongest for a snapshot observation, they are still noticeable in a 4hr track.

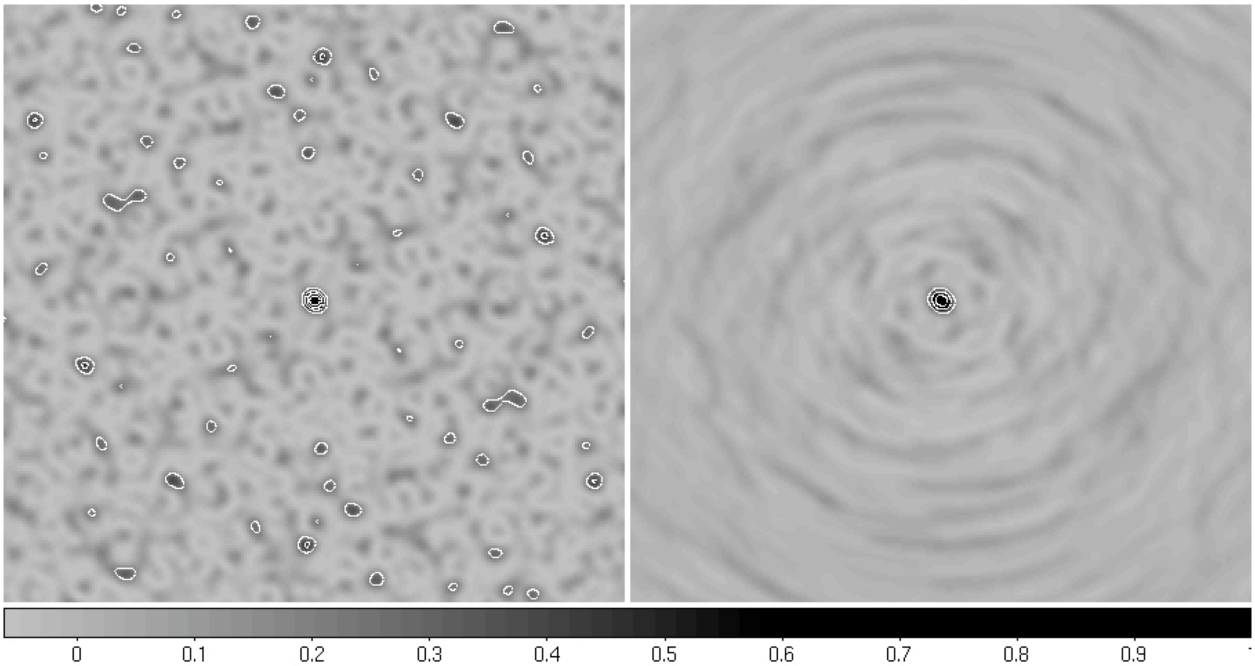

Figure 2: 10min ES beam on the left, 4hr ES beam on the right. Contours are at 20,40,60 and 80% of the beam. Note the dramatically improved sidelobe response in the rotation synthesized beam.

5. Conclusions

Without prior knowledge of the source structure, any ES snapshot will be vulnerable to flux and noise distortions in the image. This applies both to images of multiple point sources (for example, an extragalactic survey, or any extragalactic observation deep enough to detect field sources), images of a single point source against a complex background (most observations in the galactic plane), as well as imaging observations of all types.

The ALMA pipeline in particular performs an initial clean, searches for sources in the resulting image to define clean boxes, and then cleans again. Early science sidelobes are 40% of the main beam, so if the initial clean is shallow, we expect sidelobes to have a good chance of being chosen for boxing by the heuristics. On the other hand, we have shown here that an unmasked clean of even a point source can diverge very quickly, so it is also possible for the pipeline's initial shallow unmasked clean to diverge, with negative effects on boxing heuristics. We intend to explore these effects and verify these hypotheses using the actual ALMA pipeline scripts in the future. This instability results from the inherent nonlinearity of the clean algorithm, and has nothing to do with its implementation in CASA or any script such as the Pipeline in which the source location is unknown a priori.

We conclude that even for relatively isolated point sources, the combined effects of strong sidelobes and the inherent nonlinearity of the clean algorithm, uv coverage by earth synthesis should be encouraged during Early Science, explicitly in software and implicitly in documentation.

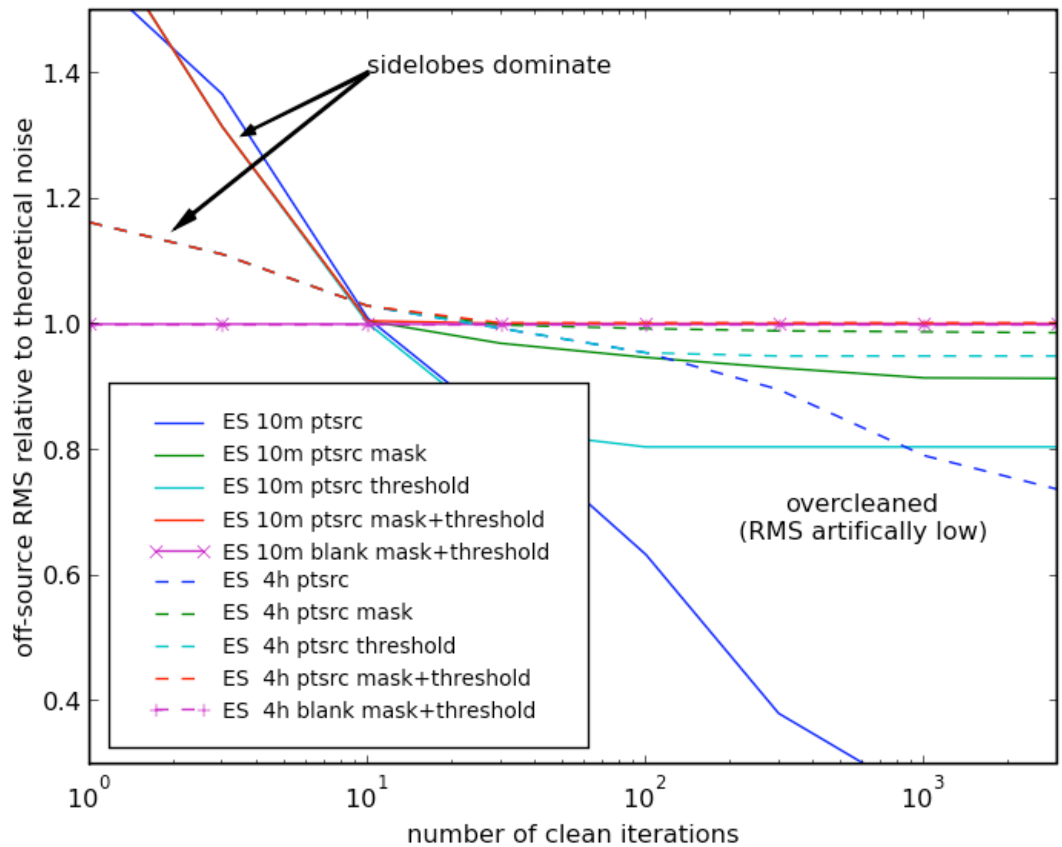

Figure 2: Image noise as a function of clean iterations in 10min and 4hr ES simulations. All simulations have a 25-sigma point source at the phase center, except for the magenta lines, which are for blank fields to indicate the true noise level.

References

Condon, J.J. et al. 1998, AJ, 115, 1693
White, R.L. et al. 1997, ApJ, 475, 479

Appendix

The ES configuration used was a strawman design by R. Hills dated 2nd September 2010, with the following antenna positions. Note that it was not optimized, and small improvements in the uv coverage may be possible.

```
# observatory=ALMA
# coordsys=LOC (local tangent plane)
# x y z diam pad#Pad,X,Y
27.93  -2.60  28.8   12.0  A004
54.29  -0.50  28.8   12.0  A006
```

49.46	20.31	28.8	12.0	A007
24.03	57.34	28.8	12.0	A009
-10.10	29.31	28.8	12.0	A021
-25.52	54.02	28.8	12.0	A023
17.77	-25.06	28.8	12.0	A036
-3.61	-40.67	28.8	12.0	A046
-47.01	-107.01	28.8	12.0	A062
60.01	-106.38	28.8	12.0	A064
117.80	22.30	28.8	12.0	A067
-106.40	71.10	28.8	12.0	A068
-126.60	-66.90	28.8	12.0	A069
141.00	-42.60	28.8	12.0	A070
-84.20	139.40	28.8	12.0	A071
47.77	161.05	28.8	12.0	A137