Doc No:	SC2/ANA/S210/004
Vers:	1.8
Category:	Report
Doc Type:	LaTeX
State:	Issued
Author:	Alex van Engelen and Douglas Scott
Date:	2005/05/25

A Correlation Study using SCUBA Data, with implications for SCUBA-2 data reduction

1 Introduction

This brief document presents follow-up work to the SCUBA-2 document SC2/S210/ANA/002.[3] There, the total-power atmospheric emission signal in SHARC-II data was compared with the model atmosphere used in SCUBA-2 simulations. One of the results was that SHARC-II timestream data contain a seemingly ubiquitous artifact: after removal of a common-mode signal from all (working) detectors, the detectors remain correlated with their nearest neighbours. This is apparent when using even small stretches of data (Fig. 13 of that document). Furthermore, when averaging longer stretches of data it becomes apparent that the region of correlation about a given detector is Gaussian in shape, with a full-width at half-maximum (FWHM) of approximately $35\,$arcsec (A. Kovacs, private communication), as can be seen in Fig. 1.

More precisely, if r_t^b is the data at detector b and timestep t, after a common-mode signal has been removed, then define a correlation statistic

$$C(b, b') = \frac{\langle r_t^b r_t^{b'} \rangle_t}{\sigma_b \sigma_{b'}}, \tag{1}$$

which ranges from -1 (completely anti-correlated detectors) to $+1$ (completely correlated detectors). On this scale, typical amplitudes of the observed effect are in the 0.1–0.2 range for b close to b' on the array.

The source of this effect is unknown. However, it seems to be common in SHARC-II data, and, curiously, its magnitude seems to decrease with increasing atmospheric opacity (Darren Dowell, private communication).

Depending on its origin, this effect could either be CSO-specific (e.g. some effect in the optics or on the detector array) or common to all sub-mm telescopes (e.g. an effect from the atmosphere). If it is present at the JCMT, there will be ramifications for SCUBA-2 data reduction strategies, in particular for the short-wavelength array. For this reason, it is worth searching for in existing data from the JCMT.

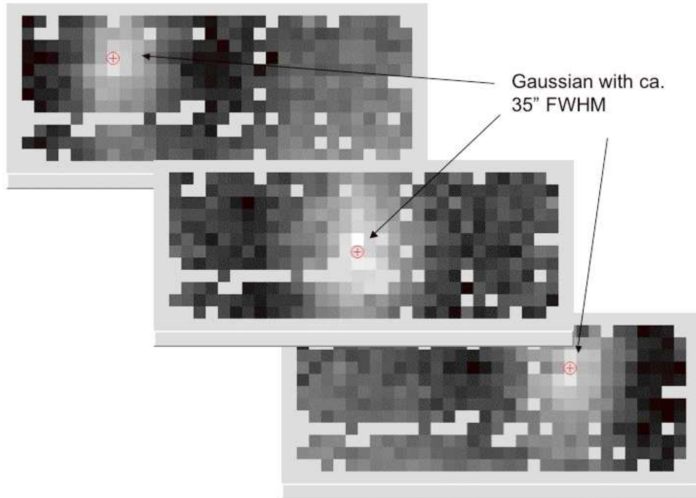

Figure 1: These images, courtesy of Attila Kovacs, show the correlation property that is the subject of this document. The correlation of each detector on the SHARC-II array with the one indicated in red is shown. Interestingly, this appears to be a circularly symmetric pattern, indicating that its origin may not be CSO-specific, and may be an issue for SCUBA-2. A similar pattern, using less data, was shown in van Engelen (2004).[3]

2 Description of procedure

Archived timeseries data from a 1999/2000 SCUBA scan-map of the Hubble Deep Field flanking region were obtained.[1] There is nothing particularly special about these data. They were chosen for convenience, but note that these data are also: simple, being taken in the same mode over many hours and on an essentially blank field; scan-map, so the data rate is approximately $8\,$Hz rather than $1\,$Hz; and with a chop sufficiently larger than the expected 35 arcsec correlation.

Bad detectors were determined by using SURF: for the $850\,\mu$m array there were 7 of 37 detectors flagged; at $450\,\mu$m there were 13 of 91 detectors flagged. SURF was also used to flat-field and extinction-correct the data. A locally-written procedure was then used to iteratively solve for, and remove, the DC offsets, the linear drifts, the anomalous spikes and the overall sky-noise signal contained in the data. This procedure is a slightly modified version of that described in van Engelen (2001)[2]).

The statistic $C(b, b')$ was then calculated for given stretches of data, for all pairs of (working) detectors, with the results plotted as a function of the detector separation distance.

It can be shown that the average value of $C(b, b')$ over all pairs of detectors is 0; this is because a DC term is subtracted from each detector's output, and a common-mode (array mean) signal is also subtracted. Given that we must have $C(b, b') = 1$ for $b = b'$, this means that in the case of completely de-correlated detectors, $C(b, b')$ for $b \neq b'$ is actually negative, on average. This statistic is still useful, however, because we are looking for an increase in C as the detector separation becomes small.

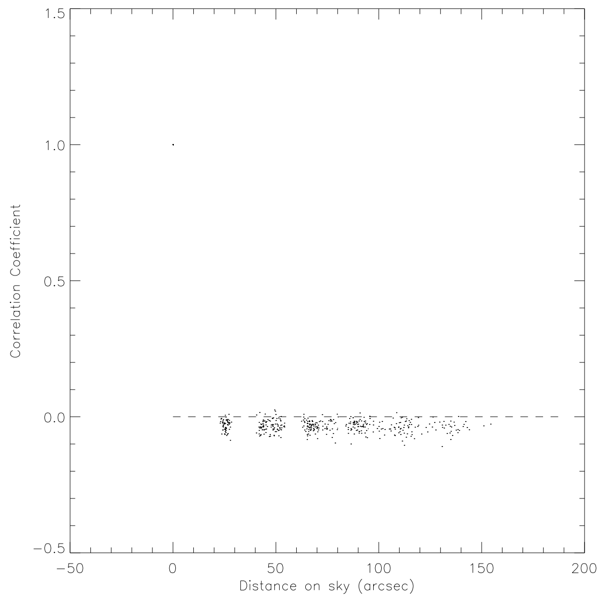

Figure 2: Correlation statistic from equation (1) as a function of detector separation distance for approximately 10 minutes of simulated independent white-noise data. In this simulation, each detector's data has the same variance. The main point here is that for uncorrelated data, the estimate for non-zero detector separation is negative but constant. This is because the mean value of the correlation must be zero, while the zero separation correlation is unity.

3 Results

The entire process was tested on simple simulated data, mainly for debugging purposes, but also for comparison with results from the real data. Figure 2 shows results for Gaussian white noise which are independently drawn for each detector. Each pair of detectors is plotted as a point.

Figure 3 shows results of this procedure for both $450\,\mu\mathrm{m}$ and $850\,\mu\mathrm{m}$ SCUBA data, in which the time average in equation (1) is taken over 62.5 minutes of time-series data. In both cases, there is clearly no increase in the correlation coefficient as the separation distance decreases. This is in stark contrast to the SHARC-II data, in which the correlation is noticeable after integrating for only tens of seconds.[3] Integrating the SCUBA data for longer times gives similar results. [1]Due to the difficulty in finding independent calibration information, as well as the difference in observing modes (chopped vs. not chopped), it is difficult to make a quantitative comparison of the 'expected' level of correlation to be seen at JCMT; however we are integrating over *much* larger timescales here!]

4 Conclusion

The correlation property described above which exits in SHARC-II data at $350\,\mu\mathrm{m}$ is not detected in SCUBA data at either $450\,\mu\mathrm{m}$ or $850\,\mu\mathrm{m}$. This suggests that the effect is related to the CSO and/or the SHARC-II instrument, and is unlikely to arise in the atmosphere. This study closes one of the loose ends left from the previous study,[3] namely the only large-

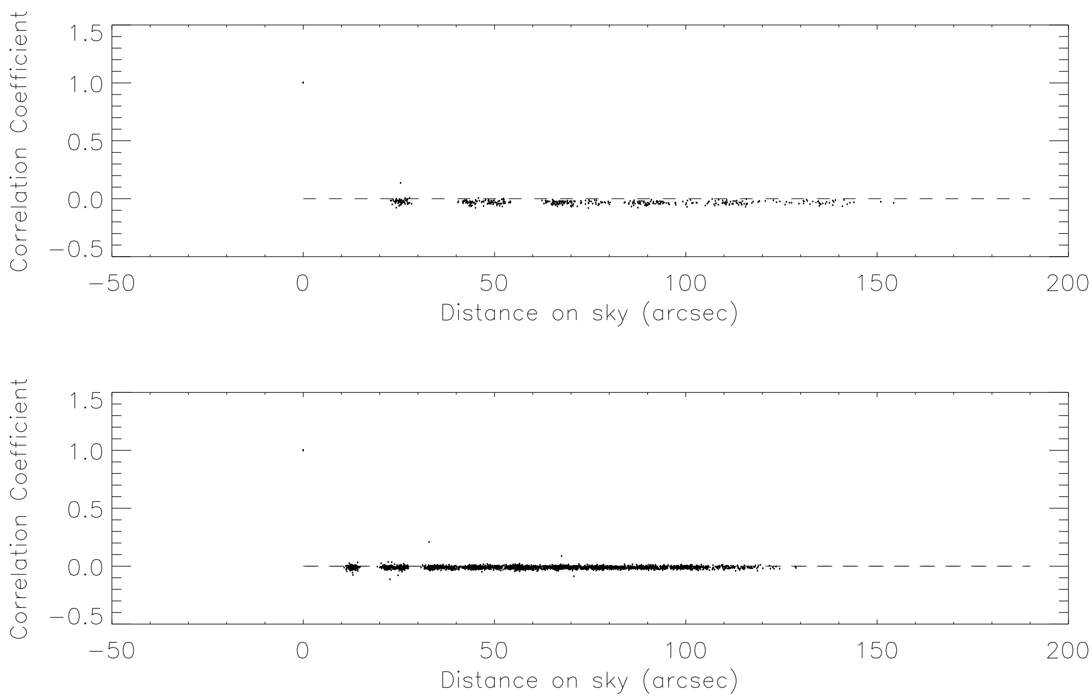

Figure 3: Correlation statistic from real SCUBA data for all pairs of detectors, as a function of detector separation distance on the sky. Top image: $850\,\mu\mathrm{m}$ array; bottom: $450\,\mu\mathrm{m}$ array. Note that there is no evidence for an increase at separation distances less than 35 arcsec. There are several data points that are well-away from zero; these are due to incomplete masking of bad detectors and/or spikes, and form a very small percentage of the whole data-set.

scale, non-instrument specific effect measured in blank-field SHARC-II data after removing a common-mode signal from each detector's timestream.

From the perspective of SCUBA-2 data reduction, this is good news, since it implies that one of the complications in SHARC-II data will not be present. The more general lesson to be learned, however, is that we should expect different, although equally subtle and complicated, effects to be present in SCUBA-2 data. The DR software needs to be prepared to test for and remove whatever systematic effects turn out to exist in real SCUBA-2 data.

References

[1] C. Borys, S. C. Chapman, M. Halpern, and D. Scott. A SCUBA scan-map of the Hubble Deep Field : measuring the bright end of the submillimetre source counts. *Monthly Notices of the Royal Astronomical Society*, 330:L63–L67, March 2002.

[2] Alexander van Engelen. Undergraduate Honours Thesis: New map-making ideas for submillimetre experiments. Simon Fraser University, 2001.

[3] Alexander van Engelen. Analysis of atmospheric emission using simulations and SHARC-II data. SCUBA-2 Project SC2/ANA/S210/002, 2004.