EVLA memo 86

# RFI excision in synthesis imaging without a reference signal

*T.J. Cornwell, R.A. Perley, K. Golap, S. Bhatnagar, NRAO*[1]

**Abstract:** *The excision of interfering signals is crucial to the continuation of radio astronomical observations into the future. Many algorithms for RFI excision require an estimate of the interference found by observations with a reference system. However, often the best measurements of the interference come from the scientific observations themselves – the sensitivity and sampling are guaranteed to be appropriate. This is similar to the logic of self-calibration whereby the best way to calibrate the telescope is to use the scientific observations. We develop and test an algorithm in which the interference is estimated by a least squares fit to the observations and removed by simple subtraction. We call this technique "RFI self-partitioning". Differentiation of interference from signal is dependent on the natural fringe rotation of celestial sources, and the lack of fringe rotation for ground-based interference. Our test is on VLA 333MHz observations of the closely circumpolar radio source NGC6251. The interference source is a radar transmitter at Albuquerque airport, some 200km from the VLA.*

## 1. Introduction

The best defense against RFI is clearly defense in depth – take every precaution at each point in the signal path. Bell *et al.* (2000) give an excellent summary of such approaches in a conference summary. Our topic in this paper is mitigation during the last stages of the data processing. We assume that all viable precautions have been taken upstream so that the system remains linear and unsaturated, and ask what can be done during the calibration and imaging to best protect the scientific content against the effect of interference. An array is after all a very sensitive machine for detecting radio signals, including radio interference. Inevitably there will be interference in the observations that has made it past all other lines of defense.

Calibration and imaging of radio synthesis observations has advanced as we have become more specific in defining the properties of the signal to be measured, and the measuring instrument itself. Briggs *et al.* 2000 noted that in some circumstances, RFI obeys closure, meaning that the RFI contribution to the observed visibility can be factorized into antenna dependent terms – a propagation effect and the complex gain of the antenna in the direction of the RFI.

---

*1. The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under cooperative agreement with the National Science Foundation.*

## 2. A model

Consider a collection of sources of narrow band RFI from stationary emitters. After fringe stopping, the visibility function measured will be:

$$V_{ij}^{\text{obs}}(-,t) = g_i(-,t)g_j^*(-,t)V^{\text{source}}\left(\underline{r}_i(t) - \underline{r}_j(t),-\right) + a_i\left(\underline{s}_i(t),-\right)a_j^*\left(\underline{s}_i(t),-\right)k_i\left(\underline{s}_i(t),-\right)k_j^*\left(\underline{s}_i(t),-\right)P(t,-)$$

**Equation 1**

From now on, we will drop the dependence on time and frequency:

$$V_{ij}^{\text{obs}} = g_i g_j^* V^{\text{source}} + a_i a_j^* k_i k_j^* P$$

**Equation 2**

This equation thus states that the RFI obeys closure relations. It is worth discussing briefly ways in which this equation (specifically the second term on the right hand side) can be violated:

- Excessive averaging in time and/or frequency
- Cross-talk between antennas

Given a model of the source, we may solve for the on and off axis gains in the least squares sense:

$$S = \sum_{ij} w_{ij} \left| V_{ij}^{\text{obs}} \sum g_i g_j^* V^{\text{model}} \sum a_i a_j^* k_i k_j^* \right|^2$$

**Equation 3**

Calibration of the observed data and removal of the estimated RFI can be performed thus:

$$V_{ij}^{\text{cal}} = \left(g_i g_j^*\right)^{-1}\left(V_{ij}^{\text{obs}} - a_i a_j^* k_i k_j^* P\right)$$

**Equation 4**

Under what circumstances is it possible to disentangle the source visibility and interference? In table 1, we show the characteristics we have assumed for the various effects.

| Term | Definition | Time variation | Frequency variation |
|---|---|---|---|
| $V_{ij}^{\text{obs}}$ | Visibility measured between antennas *i* and *j* | – | |

| | | | |
|---|---|---|---|
| $V^{\text{source}}$ | Source visibility function | *~ antenna crossing time* | *~ antenna crossing bandwidth* |
| $g_i$ | On axis gain for antenna *i* | *~ atmospheric coherence time* | *~ constant* |
| $a_i$ | Off axis gain for antenna *i* | *~ antenna beam crossing time* | *~ antenna beam crossing bandwidth* |
| $k_i$ | Propagation of interference to antenna *i* | *~ 1/(fringe frequency)* | *~ low order polynomial* |
| $P$ | Power of interfering source | *Intermittent or constant* | *Narrowband* |

**Table 1 Characteristics of various terms in the measured visibilities**

The strength of the interfering source can in principle be determined from observations with a wide-beam antenna or a narrow-beam antenna pointing at the source of interference (if known). Alternatively, we can treat the power as an unknown to be determined from the corrupted observations themselves. In fact, since we are mostly uninterested in the absolute value of the off axis gain, we can absorb the power into the gains:

$$V_{ij}^{\text{obs}} = g_i g_j^* V^{\text{source}} + a_i a_j^* k_i k_j^*$$

**Equation 5**

If the interference is broadband then it may help to search over a range of channels. For example, consider an interfering source over the horizon. A search in direction and range could be performed. In the more usual case, the interference will be narrowband and the propagation terms may be simplified to a fringe frequency[2].

$$V_{ij}^{\text{obs}} = g_i g_j^* V^{\text{source}} + a_i a_j^* e^{j\omega_e t}$$

**Equation 6**

It is the fringe frequency that enables solution for the on and off axis terms separately. By averaging over many cycles of the fringe frequency, the two terms will be split. In practice, we find the gains by a least squares fit for a given model.

$$S = \underset{\upsilon t}{\upsilon} \, \underset{ij}{\upsilon} \, w_{ij} \left| V_{ij}^{\text{obs}} \, \upsilon \, g_i g_j^* V^{\text{model}} \, \upsilon \, a_i a_j^* e^{j\upsilon_e t} \right|^2$$

**Equation 7**

[2] To derive this relationship, remember that in an interferometer, the fringes are stopped by some means, thus conferring a fringe rotation on the naturally constant interference.

Note that if the model is omitted, then the some part of the structure can be absorbed by the off-axis gains (Lesham and van der Veen, 2000). While this could be addressed in the deconvolution by suitable adjustments to the Fourier sampling, our approach is more straightforward and direct.

Once the on and off axis gains are estimated, calibration and excision requires the following straightforward calculation[3]:

$$V_{ij}^{\mathrm{cal}} = \left(g_i g_j^*\right)^{\omega l}\left(V^{\mathrm{obs}}\ \omega a_i a_j^* e^{j\omega_e t}\right)$$

**Equation 8**

The essence of our approach, which we call *"RFI self-partitioning,"* is therefore to use the RFI signal itself in the estimation and removal process. The antenna gains and propagation terms are estimated from the RFI. The data are partitioned into two parts – one with interference but no target, and one with target but no interference.

Although we have described the measurement process in terms of fringe frequencies, an alternate and equivalent description can be made in the visibilities from different patches of the sky. RFI self-partitioning is then a special case of a general algorithm for partitioning synthesis data according to region on the sky.

## 3. An algorithm

Our algorithm (Figure 1) solves for the various terms in Equation 8 in a round robin pattern similar to self-calibration, but with the addition of steps for interference estimation and removal.

To test this algorithm without an undue amount of software development, we wrote an AIPS++ Glish script (see Figure 2). All the necessary work can be done using existing AIPS++ tools such the imager and calibrater supplemented with Glish operations. Cross subtraction is performed using the AIPS++ table tool and Glish math. The antenna gains must be inverted before application; this too is done using the table tool and Glish math.

For production work, one would want to implement this algorithm in a more streamlined way. However for a test, this approach is adequate.

[3] This calculation is different from that proposed in Perley and Cornwell (2003) in which the data themselves are corrected by the estimated gains.

## Figure 1 RFI self-partitioning algorithm

1. Initialize on and off axis gains $\begin{matrix} g_i = 1 \\ a_i = 0 \end{matrix}$

2. Calibrate using current estimates of on and off axis gains

$$V_{ij}^{\mathrm{cal}} = \left(g_i g_j^*\right)^{\omega 1} \left(V^{\mathrm{obs}} \; \omega a_i a_j^* e^{j\omega_e t}\right)$$

3. Make Clean model from $V_{ij}^{\mathrm{cal}}$

4. Stop if Clean image is satisfactory

5. Predict model visibilities $V_{ij}^{\mathrm{model}}$

6. Solve for gains $g_i, a_i$ by minimizing

$$S = \mho_{vt} \mho_{ij} \, w_{ij} \left| V_{ij}^{\mathrm{obs}} \; \upsilon \; g_i g_j^* V^{\mathrm{model}} \; \upsilon \; a_i a_j^* e^{j\upsilon_e t} \right|^2$$

7. Return to step 2

**Figure 2 AIPS++ implementation of RFI self-partitioning**

a. Make two copies of MeasurementSet, one for the target *(Mt)* and one for the interference *(Mi)*.
b. Initialize interference source model to point source at the pole.
c. Predict model visibilities for *Mt* and *Mi* :
d. *Mt:*
    i. Solve for off-axis gains using antenna bandpass solution, B, in calibrater.
    ii. Apply off-axis gains to model visibility (contains Fourier transform of the target) to obtain predicted observed target visibility
e. *Mi:*
    i. Solve for on-axis gains using antenna gain solution, G, in calibrater.
    ii. Apply on-axis gains to model visibility (contains Fourier transform of the interference) to obtain predicted observed interference
f. Cross subtract:
    i. *Mt:* Subtract predicted observed interference visibilities to obtain estimate of observed visibilities in absence of interference
    ii. *Mi:* Subtract predicted observed target visibilities to obtain estimate of observed interference visibilities in absence of target
g. Update estimates of on axis gains and correct *Mt*.
h. Update model of target by clean deconvolution (or similar)
i. Stop if converged, else repeat from step c onwards.

## 4. A test

We tested RFI self-partitioning on VLA observations at 333MHz. The VLA sees strong, constant interference from the Albuquerque airport radar at ~333MHz. Our target source, NGC6251, at declination 86 degrees, was chosen carefully to keep the fringe rate relatively low so that the sampling time and data rate would be manageable with the current VLA correlator and computers.

**Table 2 Details of RFI excision observation**

| | |
|---|---|
| Configuration | D (up to 700m) with North arm in C (up to 2km) |
| Source | NGC6251 (declination ~ +86deg) |
| Observing date and time | 2004May21, 00:44UT-05:46UT |
| Integration time | 3.3s |
| Channelization | 3.1MHz total bandwidth, 127 channels for channel width of 24.4kHz |
| Polarization | RR |

In figures 3 and 4, we show the visibility amplitude spectra as a function of time and baseline. In addition to the strong lines shown in these plots, there are other weaker lines that occur for various antenna pairs. In figure 5, we show the visibility amplitudes for all baselines of channel 60 (the ABQ radar) as a function of time. The peak interference i very strong – 10,000 – 20,000Jy compared to an expected peak target flux of about 1Jy.

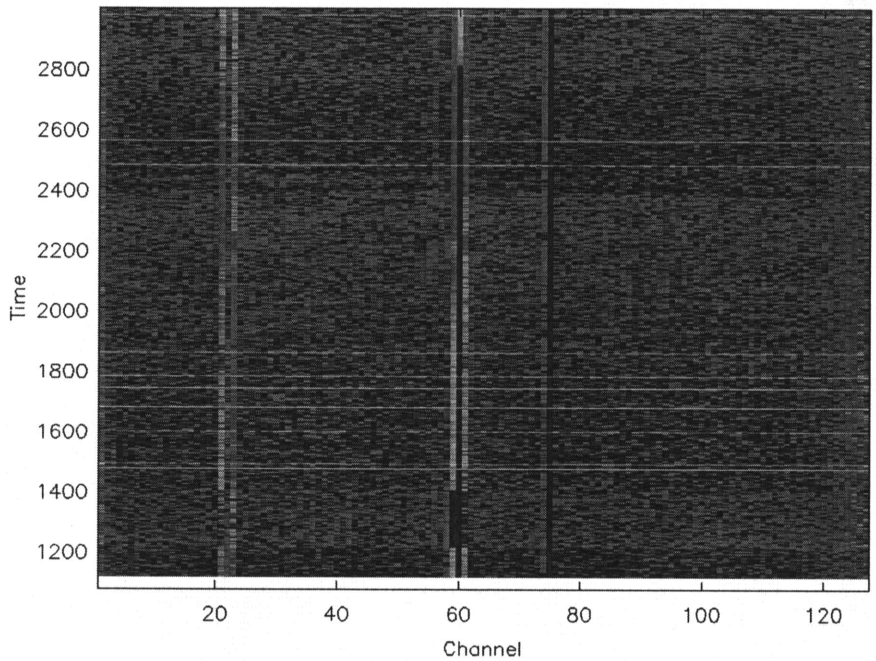



**Figure 3 Visibility amplitude as a function of channel and time for a given baseline. The ABQ radar is at channel 60 (332.9MHz).**

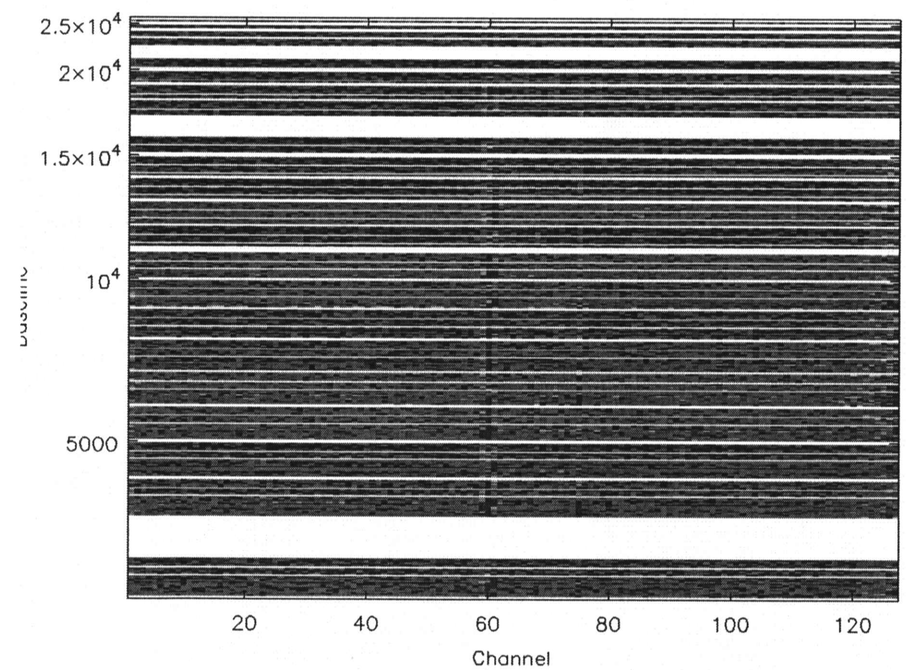



**Figure 4 Visibility amplitude as a function of channel and baseline for a given time.**

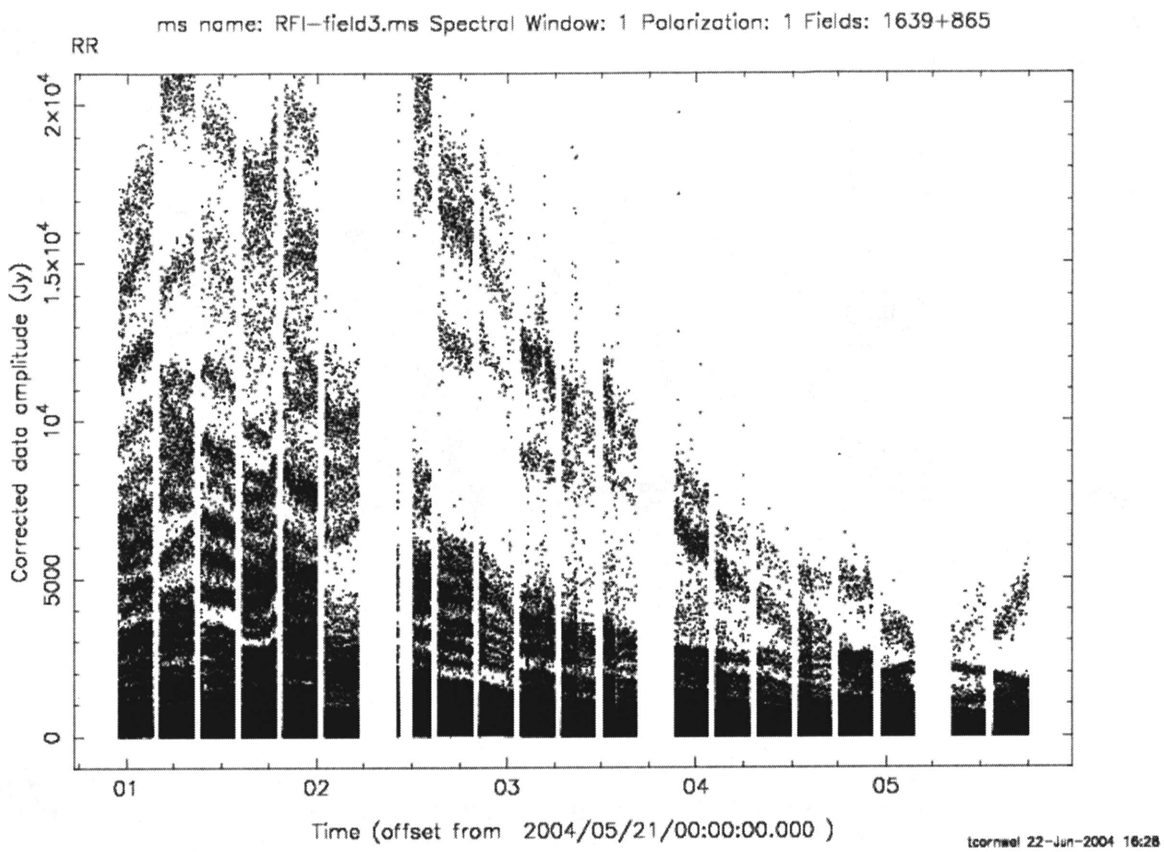
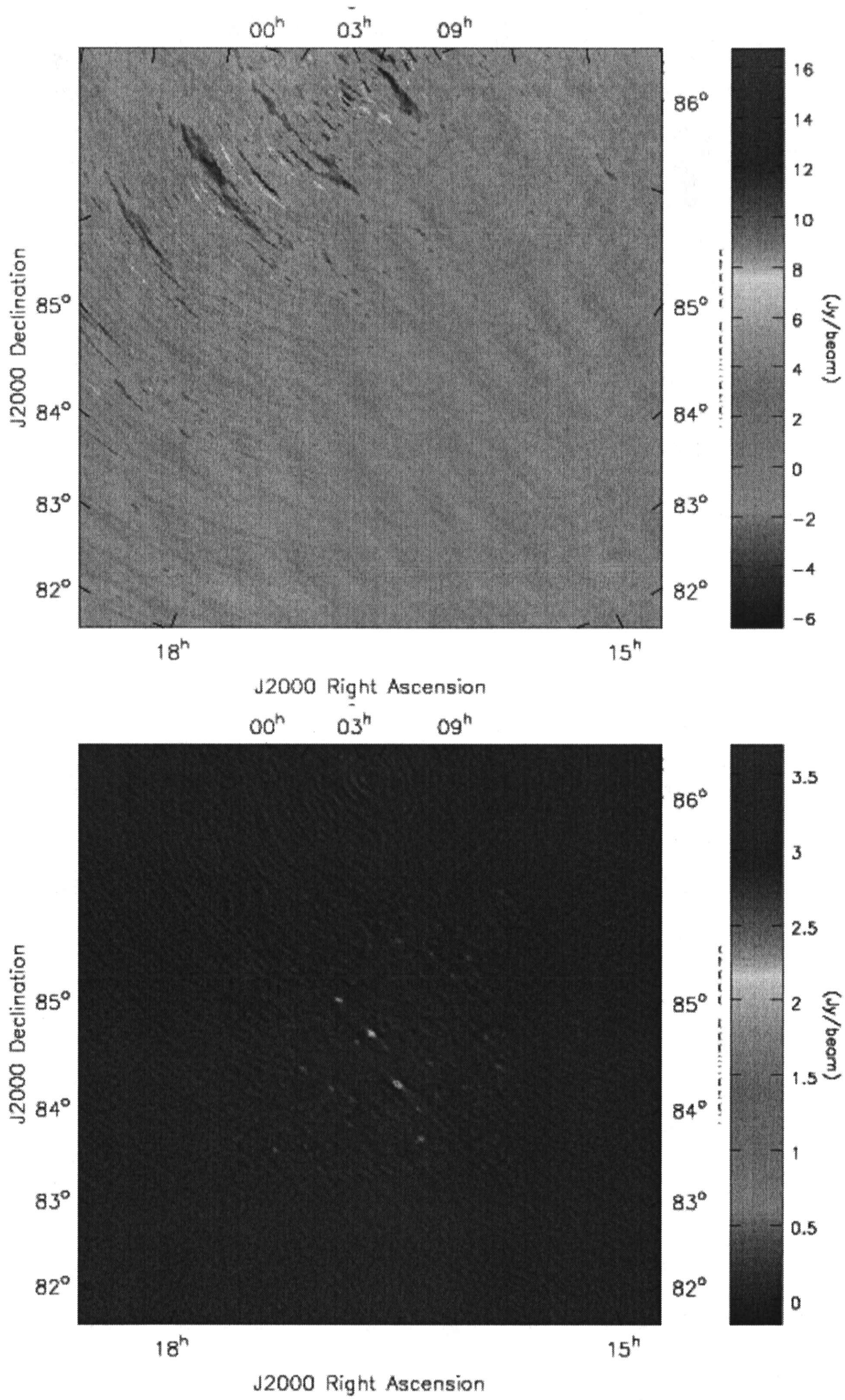
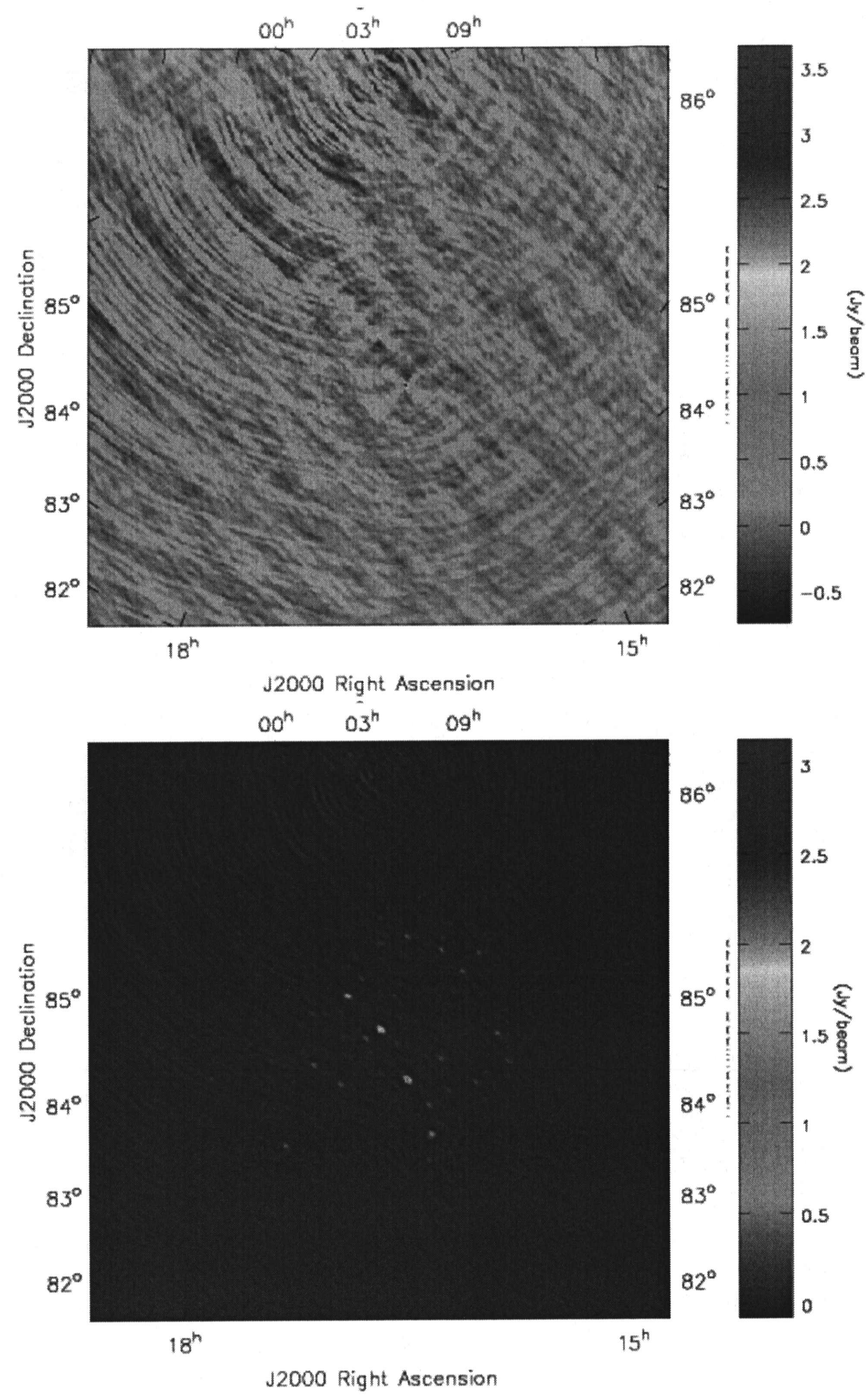



**Figure 5 Visibility amplitude for channel 60 (the ABQ radar).**

In Figures 6 and 7, we show the result of applying RFI self-partitioning to channel 75 only and to channels 70-80. These results took 6 iterations. Even though the RFI in the first completely swamps the true source, the partitioning works well. With some good channels to help in the second case, the performance improves. The excision from the original data is shown in Figure 8. The peak has decreased by about a factor of 1000 and so the closure is good to about 0.1%.




**re 2 Image of NGC6251 in channel 75 only, (top) without partitioning, (bottom) with.**




**re 3 Image of NGC6251 in channels 70 – 80, (top) without partitioning,**
**(bottom) with.**

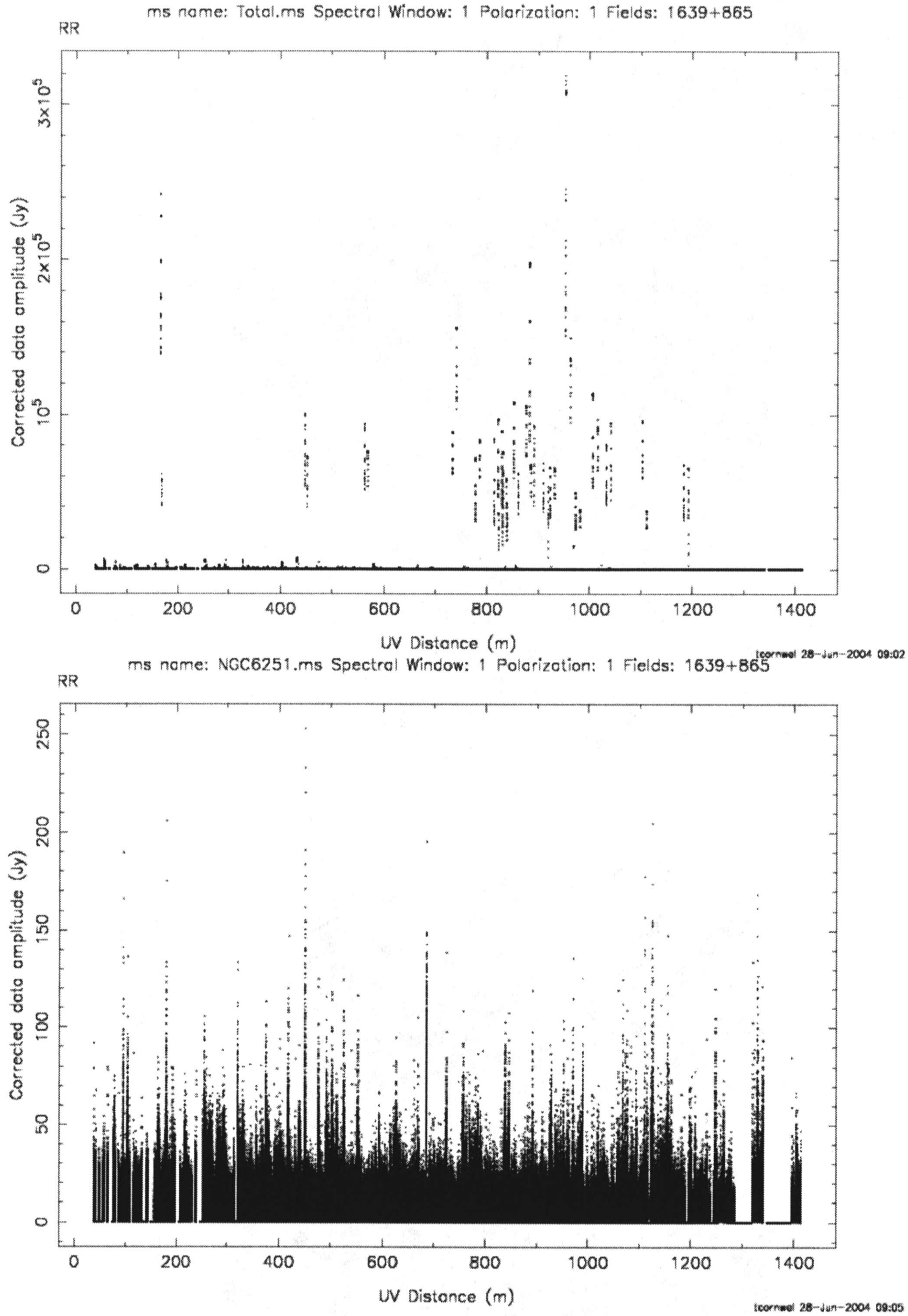



**Figure 8 Visibilities for channels 70 - 80 (top) without RFI partitioning and (bottom) with. No other editing has been applied.**

In our test, excision fails on the strongest channel. In figure 7, we show the visibility spectrum before and after RFI self-partitioning using 100 channels centered on channel 60. The line at channel 75 (333.27MHz) is well removed but channel 60 is not, and the wider line at 13 (332.0MHz) changes little. In figure 8, we show the dirty image of channel 60 before excision and the residual image after removing the best model. The RFI is removed to about 1% accuracy but the residual pattern is still sufficiently strong to swamp NGC6251. It is not clear why this line is removed less successfully than the line at channel 75. Further observational tests are required.

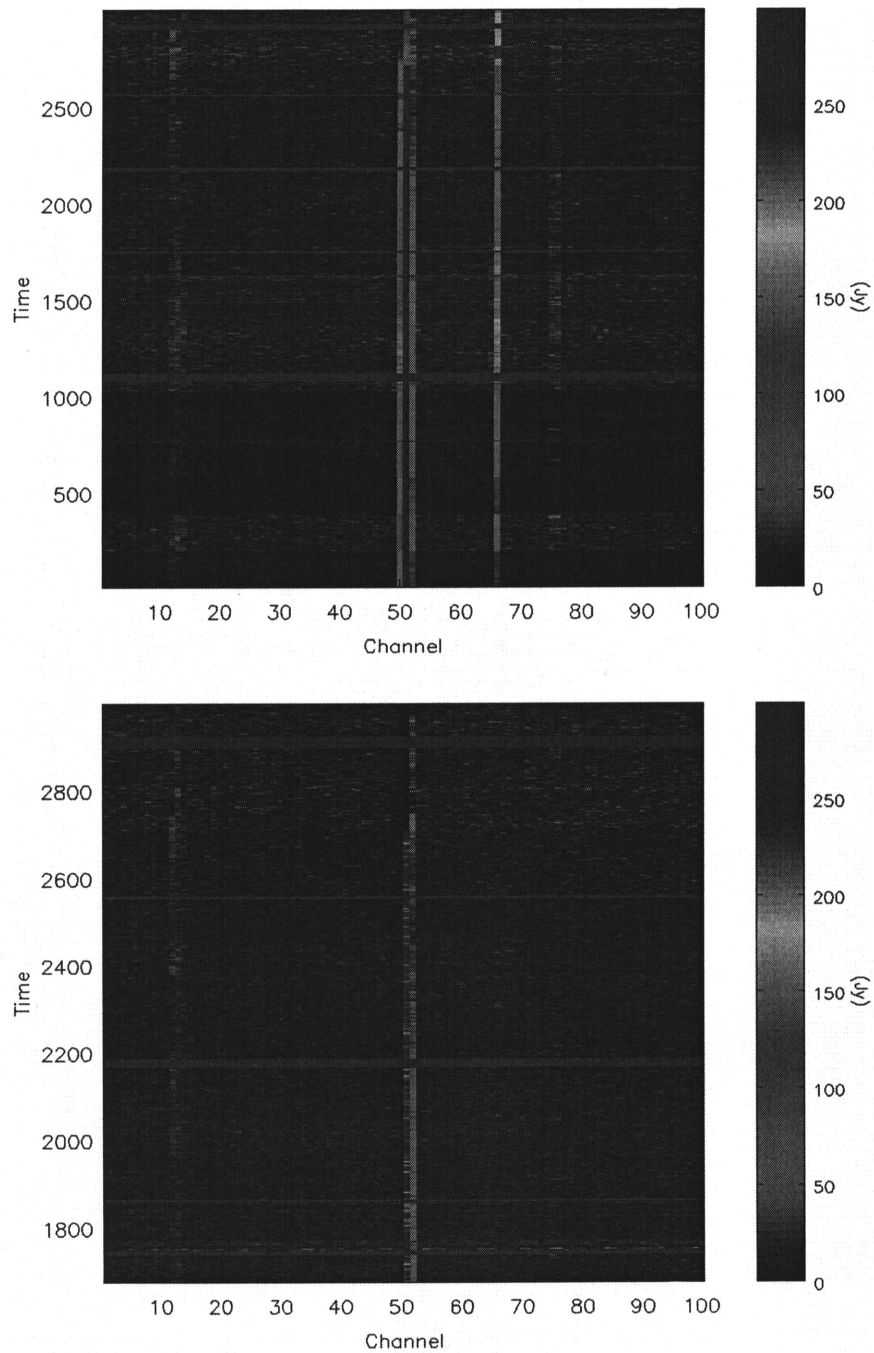



**Figure 4 Visibility spectrum (top) before RFI self-partitioning, and (bottom) after. One hundred channels centered on channel 60 were used. These are displayed via the same range. The peak on the top image is about 100 times larger than the**

**maximum displayed. The bottom image fits within the displayed range. Note tha the first ten channels have been excluded from these plots.**

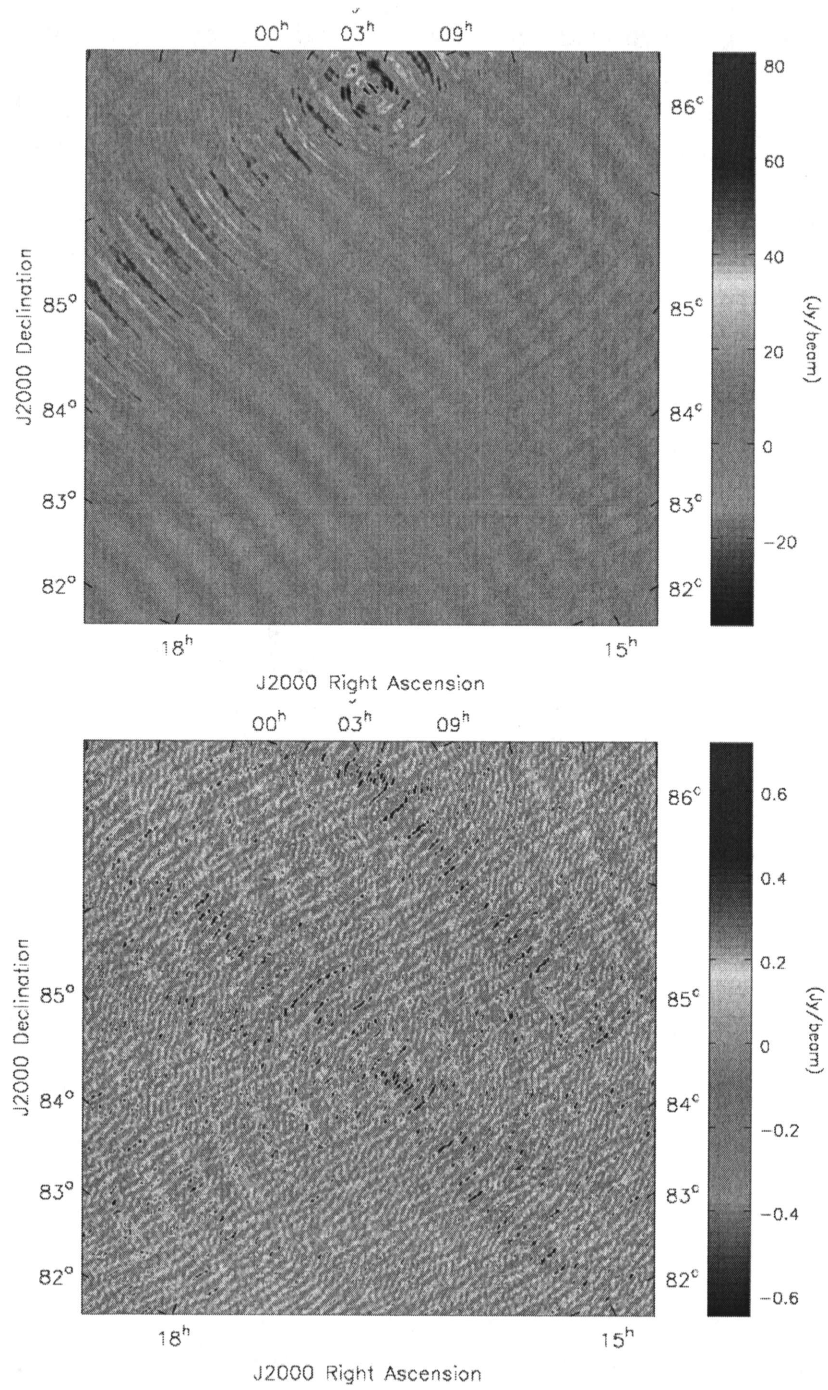



**Figure 5 Image of NGC6251 in channel 60 (top) without any partitioning, and (bottom) with. The partitioning has clearly failed.**

## 5. Implications for EVLA and SKA

As discussed by many authors (see *e.g.* Perley and Cornwell, 2003), observing to allow RFI excision must be performed with little time and frequency averaging. Our test was carefully arranged so that the interference fringe rate could be accommodated with the current VLA correlator. In the general case of a source anywhere on the sky, the fringe rate will be much higher, and the maximum allowed averaging time much smaller. Since the effects of RFI are worse for the more compact configurations, the processing requirements may be no worse than those for the larger configurations without RFI excision. However, since after excision, the data may be averaged to a rate commensurate with the source size rather than the entire sky, it may be attractive to perform excision must be real time. We see two ways to do this. First, the role of the model in our algorithm may be simply ignored. In that case, the deconvolution algorithm must be modified to account for the excised spatial frequencies (Lesham and van der Veen, 2000). Second, a model may be specified *a priori*, or built up over the course of observing, requiring rapid real time calibration and imaging that is challenging but feasible. The array would auto-calibrate and auto-edit in real time.

The impact on the computing requirements warrants in depth study and is beyond the scope of this paper. However, we can make some general scaling arguments here. The necessary data rate is set by the requirement that no significant time or frequency smearing occurs over the entire sky. This is equivalent to observing with omni-directional antennas. For antennas of diameter $D$ observing at wavelength $\lambda$, the data rate expansion is therefore:

$$\left|\frac{D}{\lambda}\right|^2$$

Cornwell (2004) has shown that the data rate for imaging the full primary beam goes as $B^2D^{-6}$, and so the data rate for RFI excision goes as $B^2D^{-4}$, meaning that small antennas are not as bad as for the pure imaging case. If the maximum baseline affected by RFI is $B_{\mathrm{RFI}}$ and the maximum baseline for full field imaging is $B_{\mathrm{Imaging}}$ then the RFI processing is less than the worse case imaging if:

$$\left(\frac{D}{\lambda}\right)^2 \le \left(\frac{B_{\mathrm{Imaging}}}{B_{\mathrm{RFI}}}\right)^3$$

For EVLA Phase 1, we will have to be able to image in A configuration at 20cm, and hence the same computing hardware would handle RFI excision in D configuration. If for EVLA Phase 1, we can image at 350km resolution at 20cm, then RFI excision up to B configuration is possible.

The feasibility of our self-partitioning approach has a considerable impact on the design of SKA. A major advantage of post-correlation RFI excision is that the interference is estimated and removed where it is most easily detected and where it does the most damage – in the imaging step. To the extent that this strategy is successful, other pre-correlation approaches may not needed. In particular, if stations of antennas were to be used instead of single antennas then instead of adaptively steering the station beams to null out interference, the station beams should be held as constant as possible and our technique used at the back end. This would largely avoid the nasty problem of wide-field imaging with unstable station beams. Our method is equivalent to adaptively nulling towards the interfering sources using the entire synthesis array, rather than just the individual elements. The disadvantage is that the data stream post-receiver must have sufficient headroom to accommodate large interference signals.

Finally, we are accustomed to designing telescopes with high levels of in-beam closure. Clearly we must now design for all-sky closure.

## Summary

We have described and tested an adaptive, post-correlation technique for RFI excision, which we call *"RFI self-partitioning."* This amounts to an extension of the usual self-calibration technique to include estimation and removal of RFI sources seen through the antenna sidelobes. The necessary closure is seen in our tests at levels close to 0.1-1%.

A reasonable question is what RFI self-partitioning gains over just excising completely the affected channels. In the case of sparse RFI (few channels have RFI), there is very little gain. However, for denser RFI (a significant number of channels have RFI), an upper limit is the square root of the fractional filling factor. In principle, every channel could suffer from RFI and our method would still work. The practical limits have yet to be explored.

The technique can still be improved in a number of ways:

- In principle, a fringe search over the entire sky (or horizon) would constrain the solutions and improve stability and robustness.
- Some simple thresholding on the antenna gains may improve noise performance.
- Our algorithm provides an excellent way to find interference. The channels affected could simply be flagged instead of excised.

## References

Bell, J.F., Ekers, R.D., and Bunton, J.D., (2000), *"Summary of the Elizabeth and Frederick White Conference on Radio Frequency Interference Mitigation Strategies,"* PASA **17**, 3

F. H. Briggs, J. F. Bell & M. J. Kesteven, (2000), *"Removing Radio Interference from Contaminated Astronomical Spectra Using an Independent Reference Signal and Closure Relations"*, AJ, **320**, 3351-3361

Cornwell, T.J., (2004), SKA memo 49.

Lesham, A. & van der Veen, A. J., (2000), *"Radio astronomical imaging in the of strong radio interference,"* IEEE Tr. Information Th., **46**, no. 5, pp. 1730-17

Perley, R.A., and Cornwell, T.J., (2003), EVLA memo 63.