VLA Scientific Memorandum No. 172:

A Test of the CS (Shortened C) Configuration

M.P. Rupen
National Radio Astronomy Observatory
Socorro, NM 87801
1 February 1997

Abstract

H I observations of NGC 1058, a face-on spiral galaxy with H I diameter $\sim$ 15 arcminutes, are used to compare the imaging capabilities of the current C with the proposed CS (shortened C) configuration of the VLA. The CS configuration gives much better images for all channels, corresponding to source sizes ranging from a few to over 12 arcminutes across, as measured by the recovered flux densities, the lack of negative bowls surrounding the emission, and detailed comparisons with images made by combining data from the C, CS, and D configurations. The flux densities derived from CS configuration agree with single-dish and D configuration measurements to within the thermal noise. CS configuration images of extended sources are comparable to those made by combining C and D configurations, and markedly superior to those made from C configuration data alone.

In the process of making these comparisons several important limitations to standard imaging techniques became apparent. CLEANing to a flux limit of $\sim 1\sigma$ is insufficient to remove the negative bowl (and to recover the true total flux density) for extended sources with a flux density in each beam only a few times the rms noise. Positive-definite maximum entropy algorithms like that implemented in the AIPS program VTESS perform very poorly on this sort of source, while maximum-emptiness algorithms like UTESS are better behaved but still sensitive to boxing. In any case the maximum emptiness/entropy algorithms also require many more iterations than normally supposed to reach the correct total flux and remove the negative bowl; for this sort of careful work the current implementations may not be significantly faster than CLEAN. Somewhat surprisingly, the deconvolved images produced by very deep CLEANs in IMAGR and those resulting from UTESS run to convergence agreed on a pixel-by-pixel basis to much better than the thermal noise, suggesting that the two algorithms are in practice less distinct than one might think (or wish).

1 Introduction

Many VLA projects require high-resolution imaging of large structures. High resolution requires the larger configurations, but this leads to a correspondingly larger central gap in the *uv*-coverage; if one wishes images which accurately recover both the small- and the large-scale structure, one must combine data from more than one VLA configuration. This is at best inconvenient, and Braun (1993) pointed out that one could at least partially 'plug' the central hole by taking one of the outer dishes and moving it to one of the central pads. Prompted by this work, Holdaway (1994) simulated noiseless observations of Cas A at 1.4 GHz, using the current C and D configurations, and Braun's proposed CS (shortened C) configuration. He

found that, especially for short ($\sim \pm 2$ hours) observations, images from the CS configuration were far superior to those from C configuration alone, and of comparable quality to those made by combining the C and D configurations, for sources up to 15arcmin across. This memo reports an observational test of this approach in the context of the most time-consuming observations encountered in practice, spectral line observations of neutral hydrogen in a nearby galaxy.

2 Observations and Deconvolution

A test CS configuration was set up at the end of C configuration in January 1995, by transferring the antennas on stations W12 and E12 to the stations W3 and E3. This differs somewhat from the configuration suggested by Braun (1993) and simulated by Holdaway (1994), who considered moving the antenna on the outermost pad of the north arm (N18) to station N1. In practice this creates problems because putting an antenna on N1 blocks the move to stations E1 and W1, both of which are occupied in D configuration (see Fig. 6b of the Green Book). The CS configuration discussed here, with antennas on W3 and E3, offers similar short-spacing advantages while avoiding this antenna-move problem. The resulting *uv*-coverage is shown in Fig. 1.

The test observations discussed here are of the neutral hydrogen in NGC 1058, a nearby ($D = 6.7h^{-1}$ Mpc) Sc II-III galaxy. This galaxy was chosen because it is large (H I diameter $\sim$ 15arcmin), fairly bright in H I (107 ± 7 Jy km/s, Allen and Shostak 1979), and has been observed exhaustively by MPR in both C and D configurations (AR296). At $02^h40^m + 37°08'$ it is also conveniently located for VLA studies. The data consist of

- **C configuration** observations of 14 and 15 June 1993, each day covering about 2100-0900 LST, for a total time on-source of 12.23 hours.
- **D configuration** observations of 7 and 8 November 1993, from 2100-0700 and 2200-0800 LST respectively, for a total time on-source of 2.67 hours (various pointings were observed during this run).
- **CS configuration** observations of 3 January 1995, from 2300-0800 LST, for a total time on-source of 5.42 hours.

All observations were taken in mode 2AC with on-line Hanning smoothing, giving 127 spectral channels with a channel width of 2.6 km/s. Reduction followed the normal AIPS procedures, with the same flux (3C 48) and phase (0234+285) calibrators throughout. No major or unusual problems were found in any data set. UVLIN was used to subtract the continuum, approximated as a linear fit to channels 11-31 and 95-115; where data sets were DBCON'd this subtraction was performed after that concatenation.

The *uv*-coverage for these observations is shown in Fig. 1, while the density of *uv*-data as a function of radius is compared for the various configurations in Fig. 2. The difference between the C and CS configurations is clear: while the C configuration coverage drops drastically below 400 wavelengths ($\sim$ 80m), the CS configuration extends inward to 200 wavelengths ($\sim$ 40m),

close to the shortest D configuration spacings. D configuration by itself has significantly more data on these short spacings, leading to much better low surface-brightness sensitivity; CS configuration is intended to fill in the short spacings for already-detected sources, not to replace D configuration for deep survey work.

To compare quantitatively the results of C, CS, and C+D configuration imaging, I considered two sets of images: one made with Briggs (1994) robust weighting at full ($\sim$ 15arcsec) resolution (the "full resolution" images), and the other made with natural weighting and a severe taper, yielding $\sim$ 55arcsec resolution (the "NALO" images). The former is intended to represent what a fair observer would actually do in practice, while the latter allows the short spacings full weight, and should give the deconvolution routines their best shot at modeling the true extended flux. Imaging and deconvolution in each case were done separately for (1) the full, C+CS+D configuration data, used as the reference data set; (2) the CS configuration alone; and (3) the C configuration alone. Table 1 gives some basic parameters of the images; and Table 2 gives details of six channels selected to span the range of extent of emission and total flux density (see also Fig. 16). The imaging and deconvolution steps brought up a number of troublesome details which might obscure the true differences between the configurations; these are discussed in the remainder of this section.

2.1 Deconvolution: CLEAN vs. MEM

While some imaging pundits passionately recommend maximum entropy, many observers have questioned the apparently mystic dependence of the resulting images on the choice of inputs, as well as the reliability of the resulting flux densities. Unfortunately the recovery of short-spacing information is precisely where the differences between the various deconvolvers might be expected to be greatest. We therefore tested four of those readily available and generally used within AIPS: IMAGR, which has an option for uv-based Clark CLEANing; APCLN, a venerable image-based Clark CLEAN; VTESS, the classic image-based maximum entropy (MEM) deconvolver; and UTESS, which is similar to VTESS but does not require positivity. The results are discussed in the following paragraphs, but the bottom line is that

(1) the results of IMAGR and APCLN are effectively indistinguishable, within the region properly deconvolved;

(2) VTESS is hopeless, as positivity leads to a biased solution (too high a flux density) when the signal-to-noise ratio (SNR) is low, as it is here on a beam-per-beam basis in all channels;

(3) the UTESS and CLEAN images with residuals restored agree very well (when run to convergence, to a fraction of the noise outside the area deconvolved), although UTESS never zeroes the residuals;

(4) the result of UTESS is not terribly sensitive to the choice of inputs (suggestions for which are given below);

(5) there is not a huge time benefit to using UTESS if one wishes to push both CLEAN and UTESS to the limit (to minimize the residuals);

(6) the flux densities derived from UTESS and CLEAN agree to within 0.1%, if proper allowance is made for the residuals.

These results were all obvious to imaging pundits (but apparently not to random observers like MPR) long ago.

- **IMAGR vs. APCLN**
  Both shallow and deep CLEANs of both the robustly and naturally weighted data show that the CLEANed images produced by IMAGR and APCLN are virtually indistinguishable over the inner quarter of the map. The tests employed were: differencing the two images at full resolution; differencing the two images at 60 arcsec resolution; blinking the images and their Fourier transforms; plotting azimuthal averages of both the difference images and their Fourier transforms. The most detailed tests used the full C+CS+D configuration data, although a quicker survey was done for the C and CS configurations independently. For purely practical reasons (disk space) I CLEANed the robustly weighted images with IMAGR, and the naturally weighted ones with APCLN.

- **Inputs to VTESS/UTESS**
  One common complaint from practicing astronomers trying to use UTESS/VTESS is that they don't know how to set the inputs. The following suggestions are based on the extensive discussion with M. Holdaway, with addenda resulting from conversations with Dan Briggs; I report on the effects of changing the various inputs in the sections on the individual tasks below. The important inputs then are as follows:

  (1) the region to be deconvolved: similar to boxing in CLEAN. The lore is that boxing may make a bigger difference for MEM than for CLEAN. For large objects of low SNR from beam to beam, our tests show that VTESS requires very careful boxing, while UTESS is more robust.

  (2) the maximum number of iterations: although the usual recommendation is for a few 10s of iterations, I found that some hundreds of iterations were required for convergence, and that those later iterations helped significantly in minimizing the residuals. Briggs (1994) also found that such "over-iteration" led to significantly better models in a very different context (super-resolution of slightly-resolved sources).

  (3) the required residual noise level: a noise level comparable to or slightly lower than the observed one seems reasonable, has worked well in past practice, and leads to convergence in a finite number (10s to 100s) of iterations.

  (4) the zero spacing flux: assuming one does not have true single-dish data, guessing a flux of about 10% of what one really believes seems to work well. One always wants to guess well below the true flux, because without true zero-spacing information

there is very little to push the flux back down, and (for VTESS) because some of the power in the algorithm derives from the positivity constraint, which means that overestimating the zero level is something of a disaster. In practice I found that guessing 10% of the flux seen in the *dirty* images is easy, agrees to first order with what the single-dish says for this galaxy, and gives good results.

(5) initial/bias image: a uniform brightness default image works very well in most cases.

We tested VTESS/UTESS primarily on the robustly weighted images, with a variety of inputs: with and without boxing; for tens to hundreds of iterations; varying the required noise level from roughly 100% to 50% of that seen in empty channels; and choosing as a guess of the true short spacing flux 10% of that found in the positive part of the original dirty map.

- **VTESS is hopeless**
  VTESS turned out to be something of a disaster, so I also ran a subset of these tests using SDE's maximum entropy task VM, with qualitatively similar results. The problems seem fundamental to the algorithm, and indeed most of these difficulties are to be expected when dealing with low SNR data, as discussed below.

  - *Floating flux scale and sensitivity to boxing:*
    Fig. 3 shows the flux density as a function of radius for the robustly-weighted images of the combined configuration (C+CS+D) data for channel 63 (520.6 km/sec). Each panel corresponds to a separate run of VTESS, with the required noise fit (in mJy) as indicated; in each case the algorithm converged in between 20 and 50 iterations. The entire inner quarter of the map was deconvolved. The results of a deep APCLN are shown for comparison. Depending on the choice of noise level (the actual value in channels with no signal is about 0.5 mJy/beam), the resulting images may show either a positive or a negative bowl surrounding the source, of virtually arbitrary strength. Here the total derived flux density varies from 3.4 to 17.1 Jy. Boxing helps matters considerably but does not eliminate the problem; this presumably is the origin of conventional wisdom that you can't trust MEM flux densities. Note that the "MEM bowl" is *not* some constant value that can be safely subtracted from the entire image; rather, the extraneous flux density is always higher in the inner part of the image. The reason for this is not clear. In any case, an algorithm which gives flux proportional to area deconvolved is rather frightening, particularly when one doesn't already know where the real flux is.
  - *Residuals large and correlated with image:*
    While CLEAN can (with enough iterations) effectively zero the residual image, maximum entropy algorithms cannot. Moreover, the residuals always trace the model image – in fact, one can derive the model from the residual image directly. Both statements follow trivially from the following argument. In MEM or any related maximization algorithm one is maximizing a function like

$$H - \lambda\chi^2$$

where H is the entropy, χ^2 is the usual sum of the squares of the residuals, and λ is a term balancing the importance of the entropy and the goodness-of-fit. The maximum corresponds to

$$\frac{dH}{dI} = \lambda \frac{d\chi^2}{dI}$$

where I is the image. For MEM, $\frac{dH}{dI} = -\log I$, while $\frac{d\chi^2}{dI}$ is basically the residuals, and the model image is roughly the exponential of the residual image[1].

Fig. 4 shows this effect in practice, for one of the more extended channels as imaged with the full C+CS+D configuration data. Even a boxed VTESS leaves 5% residuals over the source, not to mention 20 mJy sidelobes outside. APCLN by contrast has zeroed the residuals, at the expense of about a million CLEAN components (with a gain of 0.1; see below).

- *Instabilities when doing many iterations:*
  The VTESS results presented above refer to runs for 30-50 iterations. One might hope that further iteration would improve these results, but in fact both the goodness-of-fit and the entropy start to grow worse, and the flux in the model constantly decreases. VTESS actually "blows up" around iteration 50 (flux shoots up by an order of magnitude or more, then drops too low, while the goodness-of-fit jumps to much higher values and stays there). To test whether this was a bug or a feature I tried the parallel SDE task VM, which while not exploding does give the same basic behavior (model gets worse rather than better after a few 10s of iterations). The reason for this instability is not clear.

The first two effects – sensitivity to boxing and large and correlated residuals – make VTESS unsuitable for low-SNR imaging of sources of (initially) unknown structure. Holdaway's (1994) MEM deconvolutions of Cas A in his CS configuration tests were presumably successful because the simulated data were noiseless.

- **UTESS**
  UTESS is a "maximum emptiness" algorithm: it maximizes the entropy like VTESS, but does not require that the model be everywhere positive, thus avoiding most of the problems encountered with VTESS and VM for low SNR data.

  - *Less sensitive to inputs:*
    The inputs are similar to VTESS, but the result is far less sensitive to the details of

[1]Thanks to T. Cornwell for this explanation, which makes the result as obvious as it should have been to begin with.

those inputs. In particular, neither the derived flux nor the details of the final model depend significantly on the choice of noise level, and both are also less sensitive to the area used for deconvolution. For channel 63 (520.6 km/s), for instance, the flux in the restored image varied from 3.42 to 3.55 Jy for a factor two change in the input noise, and this variation can be ascribed entirely to the change in the level of the residuals (see the section below on scaling the residuals). Inadvertently setting the initial flux estimate to the negative of the true value gave a similar difference, compared to the correct inputs.

- *Convergence:*
With the noise level set to about 60% of the noise as measured in empty or CLEANed images, UTESS took 300-450 iterations to converge. (Setting the noise to the true level led to convergence in of order 100 fewer iterations, but with higher residuals and somewhat lower model fluxes; the noise levels used here were chosen as the smallest which would give convergence in any reasonable time.) The stopping criterion actually seems to be reaching an agreement of 0.05 times the requested noise. There is no sign of the VTESS/VM instability mentioned above: the model does improve at each step.

- *Much lower residuals:*
As argued above, all MEM-like algorithms must have residuals which are simply related to the source structure, and UTESS is no exception. However, the level of those residuals is an order of magnitude lower than given by VTESS: a peak of 0.02 mJy/beam for the worst full-resolution channel with UTESS, vs. the typical 0.1–0.2 mJy/beam from VTESS. This corresponds to 0.5% and 5% of the model intensity for UTESS and VTESS residuals, respectively. Minimizing the residuals is important because the restoring and dirty beams are different, in particular for the VLA configurations discussed here, and because a deconvolution which does not go very deep runs the risk of doing a poor job of recovering the missing short-spacing information.

- *Agreement with CLEAN:*
The final restored UTESS images, with residuals added, agree with the results of very deep CLEANs to within the level of the UTESS residuals. In general the disagreement is at a low, roughly-constant level across the part of the image that contains emission, as shown in Fig. 5 for the full-resolution version of the canonical channel 63. The best estimates of the total flux densities, correcting the residual flux scale as described below (JvM method), agreed with similar estimates from CLEAN to better than 0.3% (a few mJy) for six sample channels spanning a range of total flux density and size (520.6 to 559.4 km/s). As discussed above, UTESS cannot zero the residuals, so the flux densities before residual scaling are more discrepant (up to 20 mJy). Table 3 shows these results for the NALO (tapered, low-resolution) images. This agreement held for both the robustly- and the naturally-weighted images, and for CS, D, and C+CS+D configuration deconvolutions. Deconvolutions

of C configuration alone were less robust, with CLEAN generally recovering more flux, but this is hardly surprising since C configuration has so few short spacings to constrain the deconvolutions.

- *Timing:*
  UTESS, when run to convergence, does not seem to give a huge time advantage for the galaxy studied here, based on careful testing of a few images, and a rough (factor of few) impression from batch deconvolutions of entire cubes. For example, running UTESS to convergence (439 iterations) on channel 63 of the robustly-weighted C+CS+D configuration images took 4395s of CPU on navajo, an IBM/560 with 64 MB of memory, while APCLN to a similar level (residual flux 0.05mJy/beam; 260000 iterations with GAIN 0.1 and FACTOR −0.3) took 7034s. In each case the deconvolution region covered 43164 pixels. The relative speed of the two algorithms is however complicated by their detailed behavior; for instance, UTESS generally reaches a given flux level much more quickly than APCLN, but sometimes then 'loses' flux as iteration continues. This comparison requires more careful study.

In sum, while VTESS works only poorly on these data, UTESS does as good a job at deconvolution as CLEAN, with the caveat that it can never truly zero the residuals. This is not a major restriction except in the cases of very careful work, where one may wish to scour the map as clean as possible, and very non-Gaussian dirty beams, where the flux scale for the residuals is very different from that for the restored model. For the current study I chose to stick with CLEAN, because (1) the algorithm is widely used and trusted by H I observers; (2) CLEANing to convergence does zero the residuals, finessing the problem of residual scaling (discussed below); and (3) neither VTESS nor UTESS, as currently implemented in AIPS, restore the residuals outside the deconvolution region.

2.2 Deconvolution: How deep is deep enough?

The standard lore amongst H I pundits has been that CLEANing below 2-3 times the rms noise level is pointless, since the sidelobes of any remaining components are well below the noise, and so don't alter the general appearance of the map. Unfortunately, while the sidelobes of any single 1σ component may be negligible, the cumulative sidelobes of thousands of such components can be quite important; most obviously, they produce the well-known negative 'bowls' surrounding regions of emission. Since we are interested in quantitative comparisons, and are especially concerned with the recovery of large-scale structure, deeper CLEANing is required. Probably the simplest and safest procedure is to CLEAN to convergence, in the sense that the residuals are nearly zero and the flux density in the CLEAN model is stable. Figure 6 shows the flux CLEANed[2] as a function of the CLEAN flux limit and number of iterations for

[2]Note that the flux density and fractional errors in Fig. 6 refer to the flux density in the CLEAN model – no residuals are included. Tests involving a few channels and CLEAN flux limits between 0.01 and 2σ indicate that adding the residuals with no scaling would improve the fractional error in the total flux density by roughly a factor of two, independent of CLEAN depth or channel number (i.e., character and size of emitting region). This factor is of course related to the JvM scaling discussed in the next section.

a set of representative channels (see Table 2). This particular figure shows the results for the NALO images; the full-resolution images behave similarly, but require even deeper CLEANs (see Figure 7 for two examples).

Several results are apparent:

- Deep CLEANs are mandatory if one wishes to recover the total flux density accurately. For the NALO images CLEANing to 1σ left flux density errors of some 10s of per cent for the less extended (and lower SNR) channels, and the fractional errors for similarly deep CLEANs were several times worse for the full-resolution images. To model 90% of the total flux density for the NALO images requires CLEANing to about a quarter σ (10^4 iterations) for the weakest channels.

- Brighter (higher SNR) sources, even if more extended, need not be CLEANed as deeply to give accurate total flux densities. For instance, a several σ cutoff (4000 iterations) is adequate to recover 90% of the flux in the most extended channel (63) of the NALO cube. Similarly, the full resolution images, which necessarily have lower SNR, require a factor 5–10 deeper CLEANs than their NALO counterparts, as measured by flux limit or by NITER, to achieve the same total flux density. (The beam areas differ by roughly a factor 12.) Admittedly these statements are based on fractional errors, but it's hard to imagine what other measure to use, and the basic conclusion that one gets more bang per buck in CLEANing high-SNR sources seems reasonable. Note that each channel was boxed separately, so most of the CLEAN components do lie within the emitting region even for the channels with only faint emission.

- There are some odd features for the weakest channels – for instance, the "bump" in the CLEAN flux densities during light CLEANs. It is not clear what causes this; at any rate the flux density found by very deep CLEANing does agree with that found by the JvM scaling of residuals discussed below.

In the end I chose to CLEAN very deeply indeed, to avoid any confusion between incomplete deconvolution and real differences between the structure and flux recoverable using the different array configurations. Six representative channels (see Table 2) of the full-resolution images were CLEANed to about 1/25th of the rms noise, leaving only a few mJy in the residuals, and requiring of order a million CLEAN components (with a gain of 0.1)[3]. This is obviously prohibitive – each channel takes 2hrs of CPU on an IBM/580. The naturally weighted images, being much smaller and having higher SNR, could be CLEANed to an equivalent level in about a tenth the time, and all 37 NALO channels with significant emission were CLEANed to convergence. The flux densities measured in the six overlapping channels from the NALO and full-resolution images agreed to within a few mJy, for all configurations, demonstrating that conclusions drawn from the NALO flux densities are valid for the full-resolution data as well.

[3]The total number of visibilities in a single channel of the combined C+CS+D configuration data set is 598,010.

2.3 Deconvolution: Scaling the Residuals

One of the major difficulties in deriving flux densities from interferometric maps is the difference in units between the model (which is generally restored with a Gaussian beam) and the residuals (which are convolved instead with the dirty beam). Jörsäter and van Moorsel (1995; JvM) suggest that one should consider the "true" flux density as the sum of the flux density in the model and some constant times the apparent flux density in the residual image, as measured under the (false) assumption that the residuals are also characterized by the restoring beam. They argue that this constant should be a function of the distribution of the emission, the boxes used in CLEAN, and the shapes of the dirty and restoring beams; but it should *not* depend on the depth to which the map is CLEANed. This last implies that one may solve independently for the scaling factor for the residuals, and for the true flux density, by CLEANing to two different depths and requiring that both the scaling factor and the true flux density derived from the resulting images be the same. They further show that the flux density thus derived agrees with that found by CLEANing to convergence, even for relatively shallow CLEANs, to better than 10%. We applied this trick to both the CLEAN and the UTESS images discussed above.

The most obvious problem was with the C configuration data, which were simply not susceptible to this analysis. In many channels both the model and the residual flux densities grew as they were CLEANed more and more deeply, and in others the derived values of the "true" flux density and the multiplicative constant varied significantly (10s of per cent) from major cycle to major cycle of the CLEAN algorithm. Presumably this reflects the severity of the missing short spacing problem in this configuration, as even by "convergence" CLEAN leaves a significant negative bowl (see below). This problem recurred for both the robustly- and the naturally-weighted images. By contrast, the method worked well for the CS, D, and combined C+CS+D configuration data sets, yielding flux densities stable to a few per cent when CLEANing up to or deeper than a few times the rms noise. Some of these results are shown in Fig. 7.

In the cases in which it could be applied, the same technique worked equally well for images made by the maximum entropy task UTESS, with the final flux densities derived from that algorithm agreeing with those from CLEAN to the measurement accuracy of a mJy or 0.3% (Table 3). However, for the channels with little emission, boxing could make a big difference. NGC 1058 is a large, face-on galaxy, so it is quite possible that low level emission be spread across a large area even in weak channels; but the scaled-residuals approach works very poorly in that case, firstly because the SNR of the derived scaling factor and total flux is low, and secondly because that scaling factor depends on the distribution of the emission, which means that one should really employ a scaling factor which varies with the local structure. The weak channels, consisting of very low level extended emission plus fairly bright central 'blobs', appear to exacerbate this problem. This difficulty is not specific to this particular data set: trying the same approach on Mundell's B configuration observations of Seyfert galaxies gave results that differed by 20-30% (10-15 mJy) depending on the strictness of the boxing.

Even for CLEAN, scaling the residuals is not a panacea, because multiplying by a constant

is not the same as deconvolving an image. Consider the case of a low SNR 'woofly' structure with a large negative bowl surrounding it; clearly multiplication of the residuals cannot remove this bowl. But there are also more subtle effects, and while the total flux density may be derived very accurately, the details of the image are not, as is shown in the difference plots in the bottom panels of figure 7. These show the mean and rms difference between the partially-CLEANed images (with various scalings for the residuals) and the 'final', converged CLEAN image. While the CLEAN image with the usual unit scaling basically differs from the final image by a constant, the JvM scaling gives an image which agrees very well in the mean, but quite poorly in the dispersion. The curves labeled "best" in this figure represent an attempt to keep the best of both worlds, by scaling the residuals by the constant which minimizes the median difference between the CLEAN+scaled residual image and the converged CLEAN image. While this keeps the rms error fairly low, it doesn't yield very accurate total flux densities. We conclude that, while scaling the residuals can give very accurate total flux estimates, it cannot be used blindly to avoid lengthy deconvolutions when doing careful, quantitative analysis. Simulations would be very helpful to test what effect scaling the residuals has on the fidelity of the image – comparing with a converged CLEAN image is easy but hardly conclusive, given the known imperfections of that algorithm.

Given these difficulties I have chosen to CLEAN to ridiculously low levels, and to cite the standard flux densities as measured directly from the restored images. Since the residuals at the end of these deep CLEANs count for at most a few mJy, and since the derived scaling factors are typically about a factor 2, these standard flux densities are accurate enough for current purposes, and probably more accurate than the overall flux scale anyhow. For the six robustly-weighted channels studied in detail I also quote the "true" flux density as derived according to the Jörsäter and van Moorsel prescription (Table 3).

3 Results

The main question is how well the CS configuration measures, and subsequent deconvolution routines extrapolate, flux densities at the shortest spacings; and specifically, whether the CS configuration does sufficiently well that combination with D configuration would not be required even for large sources.

3.1 Integrated HI Spectra

Table 4 summarizes the available single-dish and interferometric H I flux determinations for NGC 1058. In general the single-dish measurements, *uncorrected* for the antenna response (primary beam), agree surprisingly well, given the observed size of the source (H I diameter about 14 arcmin). By contrast, earlier interferometric observations were low by 15% for WSRT (van der Kruit and Shostak 1984) and 35% for the VLA D configuration (Dickey, Hanson, and Helou 1990). Our observations do significantly better, missing only 8% of the flux in C configuration, and consistent with *no* missing flux in both the CS and combined configurations. The disagreement with the earlier D configuration observations presumably reflects their very

shallow CLEANs (500 components per channel, vs. of order 100,000 in our (tapered) images), as the authors remark on the large negative bowl surrounding the galaxy (Dickey, Hanson, and Helou 1990).

This sounds very promising, but a channel-by-channel comparison is a much more stringent test. Fig. 8 compares the single-dish H I spectrum of Allen and Shostak (1978) with those from the WSRT (van der Kruit and Shostak 1984) and the combined VLA C+CS+D configurations.[4] We have removed a linear baseline from the single-dish data, and the VLA flux densities have been corrected for the primary beam response. The WSRT spectrum is as expected consistently low for the central channels, in which the H I is most extended; the asymmetry in the error, with a larger difference on the high-velocity than on the low-velocity wing, is a bit surprising, since the galaxy itself is very nearly symmetric. The VLA measurements on the other hand are in excellent agreement with the single dish, as shown in the difference spectrum in Fig. 8b. The relatively poor spectral resolution of the single dish (8.2 km/sec) is obvious in these residuals. Unfortunately Allen and Shostak do not report the thermal noise level of their observations, but the quoted system temperature (35 K), efficiency (0.51), integration time (12mins on, 12mins off), and bandwidth (8.2 km/sec) imply a thermal noise of $1\sigma \sim 62$mJy. Combined with the average VLA noise of ~ 9 mJy, the maximum residual between the two measurements, 196 mJy at 533.5 km/sec, corresponds to 3.1σ. Given this consistency, together with the lower noise level and higher spectral resolution of the VLA data, I take the C+CS+D configuration spectrum as the standard for all subsequent comparisons.

Before moving on to the comparison with the smaller configurations, it is worth noting that the above discussion shows the VLA can provide accurate H I spectra even of very large ($\sim$ 15 amin) and complex objects. It would be interesting to compare this spectrum with low-noise, multiple-pointing data from a large single dish, e.g. the NRAO 140-foot.

Figure 9 compares the total flux densities derived from the C, CS, and C+CS+D configuration data. In the interests of allowing simple noise calculations, these spectra have *not* been corrected for the VLA's primary beam. In each case the channels were CLEANed to 0.05σ. As with the WSRT, C configuration does a poor job of recovering the flux density in the central channels, where the source is largest. More surprising is the asymmetry of the residuals, shown in Fig. 9: the C configuration deconvolution consistently does a worst job of recovering flux on the low-velocity side, despite the apparent symmetry of the source. Further CLEANing, beyond the (constant) residual level required of all channels, did raise the flux in these channels a bit, but at an enormous cost: a factor 10 more CLEAN components for the most discrepant channel (510.2 km/sec, low by 500 mJy), for a total of two million components (vs. about 250,000 visibilities), raised the recovered flux density by only 200 mJy, leaving a $\sim$ 300mJy discrepancy. On the other hand, the WSRT data show a similarly-asymmetric disagreement, but at high rather than at low velocities. All this is important, because it shows that the accuracy of the recovered flux density is a complex function of the details of the flux distribu-

[4]The only other published single-dish *spectrum* is that from the Jodrell Mark I 250-ft. telescope in Lewis and Davies (1973). As their beam is significantly smaller than that of the Dwingeloo 25m (17x13 vs. 32 arcmin), I compare with the latter measurement. The two spectra in any case agree perfectly to within the noise, an astonishing result given the large extent of the H I in this galaxy (12–13 arcminutes) near the systemic velocity.

tion and of the interferometer. In particular, one *cannot* predict the magnitude of the error from the maximum extent of the source: the H I emission region is roughly the same size at velocities equally displaced from systemic, while the error in the C configuration flux density is very different depending on whether that displacement is a redshift or a blueshift. This implies for instance that velocity curves and velocity profiles derived for large galaxies using the larger configurations may be *systematically* in error, even if the flux integral appears to come out roughly right (as it does for our C configuration data).

Reaching at last the configuration we're actually interested in, flux densities derived from the CS configuration alone match those from the combined C+CS+D configurations to within the estimated noise level (Fig. 9b), in this case to better than 50 mJy. Since this source in different channels ranges from a few to 15 arcminutes across, we conclude that *CS configuration yields total flux densities indistinguishable from those derived from D or combined configuration observations.*

3.2 Images

Although total flux densities are an easy and fairly clear-cut indicator, the real question is whether the images resulting from CS configuration alone are similar, both qualitatively and quantitatively, to those from the combined C+D configurations; and in particular whether CS configuration does a much better imaging job than C configuration. The answer to both questions is yes, as shown in this section.

Fig. 10 presents images of one of the channels (520.6 km/sec) with the most extended H I roughly 12x13 arcmin. These are the robustly-weighted images, convolved to a common resolution of 30 arcsec for clarity. Note that these have been displayed with common contour levels, despite the rather different thermal noise. The C configuration image shows the expected negative bowl, creating hollows in the areas between the spiral arms, and making the emission appear sharper than the reality (here taken as the C+CS+D configuration image). CS configuration by contrast eliminates the negative bowl entirely, and reproduces (within the noise) all the salient features of the C+CS+D image, on all scales; note for instance the depression just south-east of the bar. The CS image does appear blotchier, but this simply reflects its higher noise level. This is shown in Fig. 11b: there are no systematic effects over the area of emission, and Fig. 12b shows quantitatively that the distribution of flux densities here is fully consistent with random thermal fluctuations. C configuration by contrast is dominated by a fairly constant offset, in addition to similar noise fluctuations (Figs. 11a, 12a). [5]

The azimuthally-averaged radial flux density profiles illustrate even more directly the advantages of the shortened C configuration (Fig. 13). The character of the C configuration bowl is obvious, with the largest (negative) deviations in the center, falling off with radius. CS configuration again gives results consistent with pure noise; note that the noise falls with radius,

[5]While these figures are all presented for 30 arcsec resolution, the results at full resolution are similar; the choice of 30 arcsec for plotting simply makes these results easier to see. In particular, there is no evidence that C configuration does a better job than CS configuration even at the highest resolution, although the noise level of the single-configuration images limits the usefulness of this statement.

as more and more pixels are brought into the azimuthal average.

One might hope that the maximum entropy algorithms would help the C configuration bowl significantly. The problem is that without the central spacings they don't know where to stop, and VTESS in particular gives a positive bowl of the same magnitude as the negative one found with CLEAN (Fig. 14). UTESS can be carefully massaged to give a better result but the only reliable way to choose the appropriate "golden"' parameters is on the basis of comparison plots like these, obtainable only through observations with the smaller configuration.

These difficulties are not confined to the channels with the most extended H I emission. Fig. 15 shows *dirty* maps from two edge channels, taken from the high-velocity side of the emission peak, where C configuration does a very good job of recovering the total flux densities. These images are the tapered, naturally-weighted ones; the robustly-weighted images show the same effect. Even for fairly small (4–5 arcmin) H I extents, C configuration has trouble with the low-level, extended emission; bear in mind that this is 12 hours of C configuration, compared with only 3.5 hours of CS. In both channels C configuration picks out only the peaks of the emission, and it is hard to imagine that one would even know how to box these maps to bring out more. That the total flux densities agree merely reflects the relatively low flux densities in these channels, which leads the integrated noise to mask the signal in the H I spectra.

Finally, Fig. 16 shows similar results for six channels chosen to span a range of intensities and H I distributions. These images were deeply CLEANed and convolved to a common resolution of 60 arcsec. The comparisons reported above hold for these channels as well, and I conclude that CS configuration represents a substantial improvement over C configuration in all cases.

4 Discussion and Conclusions

CS configuration represents a substantial improvement over C configuration in image fidelity on large size scales, and in particular for the measurement of total flux densities. The negative bowl seen in C configuration images is completely removed, even for sources with complex structure as large as 14 arcmin across. Further, CS configuration images agree, even at full resolution, with those made by combining C, CS, and D configurations; while C configuration by itself leads to images that are quantitatively in error and qualitatively mis-leading, even for sources only a few arcminutes in diameter. We conclude that the availability of CS configuration would obviate the need for follow-up D configuration observations of H I in large sources, and could substantially improve all H I observations currently made with C configuration alone.

This memo has focussed on large, woofly sources with fairly low signal-to-noise ratios in single beams. Holdaway (1994) gave similar results based on noiseless simulations of Cas A, scaled to various sizes, and also found that CS configuration was always an improvement over C configuration in terms of image fidelity, except in the case of full C configuration tracks (where the projected C configuration baselines became very short as the source set) when the two were roughly equivalent. Holdaway's results are appropriate for high SNR data, so CS configuration seems to be marvelous on all size scales, for both faint and bright sources. Clearly there is a strong case for allotting a substantial portion of the time normally reserved for C configuration

to the new, shortened, CS configuration.

Should CS configuration completely replace the current C configuration? In the scheme discussed here, one gains the short spacings of CS configuration at the cost of some intermediate C configuration baselines. This must lead to some degradation of image quality on intermediate size-scales, but this is almost impossible to quantify. Holdaway's simulations, using a known source structure, would seem ideal; but there the infinite signal-to-noise ratio probably leads to unnaturally good deconvolutions. The current memo, relying on real data, has no "truth image" to do the comparison, and measurement of the agreement between CS and C+CS+D configurations is limited by the CS noise level. One could re-do Holdaway's simulations with full-blown noise and calibration errors, but this would be both time-consuming and inconclusive, since the importance of those intermediate spacings must depend on the details of the source structure and the deconvolution scheme. Another argument is that projects which will want D configuration anyhow – for instance, mosaics at C configuration resolution, or surveys which need excellent sensitivity on both C and D configuration size scales – wouldn't gain by having the short CS spacings. There is also the question of spectral index or other comparative maps, where one wants as similar *uv*-coverage as possible at different wavelengths. Although these issues are almost impossible to check quantitatively, they seem very contrived; it is difficult to imagine a project which would not benefit from a few short spacings, or which would be harmed substantially by loss of a few intermediate ones.

In fact, I am led instead to propose going to shortened configurations in *all* of the larger configurations, A, B, and C, motivated only in part by the whimsical prospect of proposing for BS configuration. This would remove the only real argument, that of the need for scaled configurations, because AS, BS, and CS could be scaled versions of each other. D configuration would of course be different, but it always has been, both because the innermost antennas are placed in unique positions, and because shadowing removes some of the scaled short spacings. There would still be a need for D configuration for surface brightness sensitivity, but scaled configurations should remove many of the short follow-up observations currently required to give believable maps. For the VLA, this would mean some gain in available observing time; the main benefit however would be for the observer, who could do with a single configuration what now takes two, leading to simpler and quicker data reduction, and (for current single-configuration projects) to more accurate maps. The main scientific argument against this scheme is the loss of Fourier filtering in the largest configurations, which are frequently used to search for, or measure the fluxes of, small sources embedded in extended emission. This seems a reasonable argument for retaining the regular A and B configurations for some fraction of the time, but C configuration is hardly ever used thus, and could be abandoned.

5 References

Allen, R.J. and Shostak, G.S., 1979, *A& AS*, **35**, 163.

Braun, R., 1993, *VLA Scientific Memorandum No. 165: Telescope Placement at the VLA for Better Single Configuration Imaging.*

Dickel, J.R and Rood, H., 1978, *Ap. J.*, **223**, 391.

Dickey, J.M., Hanson, M.M., and Helou, G., 1990, *Ap.J.*, **352**, 522.

Holdaway, M.A., 1994, *VLA Scientific Memorandum No. 167: Evaluation of C Array Single Configuration Imaging.*

Jörsäter, S. and van Moorsel, G.A., 1995, *A.J.*, **110**, 2037.

van der Kruit, P.C. and Shostak, G.S., 1984, *A& A*, **134**, 258.

Lewis, B.M. and Davies, R.D., 1973, *MNRAS*, **165**, 213.

Table 1. NGC 1058 Images

Full-resolution images: full-resolution, robustly-weighted
(4 arcsec/pixel; 1024 x 1024 pixels; IMAGR robust=0)

Config.	Beam size [asec]	wtnoise[a]	rms noise [mJy/beam]
C	14.88 × 13.21 at 86.9°	1.18	0.675
CS	14.47 × 13.40 at 81.8°	1.17	0.74
D	47.57 × 43.13 at −86.4°	1.18	1.2
C+CS+D	15.44 × 13.52 at 87.0°	1.18	0.50

NALO images: low-resolution, naturally-weighted (with taper)
(8 arcsec/pixel; 512 x 512 pixels; IMAGR uvtaper= 4kλ)

Config.	Beam size [asec]	wtnoise[a]	rms noise [mJy/beam]
C	43.85 × 39.77 at −66.1°	1.50	0.88
CS	55.32 × 46.75 at 81.6°	1.44	0.95
D	72.10 × 64.35 at −85.3°	1.03	1.00
C+CS+D	53.66 × 46.87 at −83.2°	1.39	0.55

[a] *The expected noise level of images made with the requested weighting, compared to that which would be achieved with natural weighting, as reported by IMAGR.*

Table 2. Basic Characteristics of Selected Channels

Channel	Velocity [km/s]	Diameter [arcmin]	S_ν[a] [Jy]	Full-res.[b] [mJy/beam]	NALO[c] [mJy/beam]
48	559.4 km/s	3.5arcmin	0.0063 Jy	2.0	3 mJy/beam
50	554.2 km/s	4 arcmin	0.0565 Jy	3.2	4 mJy/beam
53	546.4 km/s	7 arcmin	0.200 Jy	2.8	12 mJy/beam
56	538.7 km/s	9 arcmin	0.599 Jy	6.7	40 mJy/beam
58	533.5 km/s	12 arcmin	1.23 Jy	11	60 mJy/beam
63	520.6 km/s	14 arcmin	3.56 Jy	12	80 mJy/beam

[a] *Total flux density in very deeply-CLEANed images; NOT corrected for primary beam attenuation.*

[b] *Peak flux density in C+CS+D configuration full-resolution image.*

[c] *Peak flux density in C+CS+D configuration NALO image.*

Table 3. NALO Flux Density Comparisons: UTESS vs. APCLN

Channel	Config.	UTESS [Jy]	APCLN [Jy]	UTESS (JvM) [Jy]	APCLN (JvM) [Jy]	rms noise [Jy]
48	C	0.137	0.160	–	–	0.017
48	CS	0.031	0.0316	0.029	0.030	0.014
48	D	–	–	–	–	–
48	C+CS+D	0.0089	0.0092	0.0063	0.0063	0.0088
50	C	−0.044	−0.044	–	–	0.017
50	CS	0.0542	0.0549	0.0534	0.054	0.015
50	D	–	–	–	–	–
50	C+CS+D	0.0546	0.0563	0.0563	0.0565	0.0089
53	C	0.124	0.1399	–	–	0.016
53	CS	0.200	0.2014	0.1976	0.197	0.014
53	D	–	0.1909	–	0.1895	0.019
53	C+CS+D	0.204	0.2032	0.2005	0.200	0.0088
56	C	0.4982	0.609	–	–	0.017
56	CS	0.5551	0.5692	0.5669	0.566	0.014
56	D	0.603	0.606	0.6039	0.604	0.020
56	C+CS+D	0.5965	0.6013	0.5998	0.599	0.0088
58	C	0.952	1.096	–	–	0.017
58	CS	1.213	1.24	1.239	1.238	0.014
58	D	1.21	1.224	1.226	1.225	0.020
58	C+CS+D	1.223	1.231	1.23	1.23	0.009
63	C	3.009	3.275	–	–	0.016
63	CS	3.526	3.562	3.561	3.558	0.016
63	D	3.533	3.552	3.556	3.552	0.012
63	C+CS+D	3.548	3.560	3.560	3.558	0.009

Comparison of H I flux densities derived from deep CLEAN (10^6-iteration APCLN) and maximum emptiness (300-500 iteration UTESS) deconvolutions of the NALO (naturally-weighted, low-resolution) images. Flux densities were measured by summing over the area used in the deconvolutions (which was the same for UTESS and APCLN). **UTESS/APCLN:** residuals restored with unit scaling. **JvM:** residual scaling per JvM, derived from CLEANs to 1σ and 0.5σ; residuals were multiplied by 2.00 (CS), 1.30 (D), 1.73 (C+CS+D). No consistent scaling could be derived for C configuration. **rms noise:** statistical error estimate based on the dispersion measured in blank regions, and the size of the region summed to give the quoted flux densities.

Table 4. Flux Integrals: Single-dish vs. Interferometers

Telescope	Raw Flux Integral [Jy km/sec]	Corrected F.I.[a] [Jy km/sec]	Beam size [arcmin]	Reference
Jodrell Mk I	114 ± 17 [b]	—	17 x 13	Lewis and Davies 1973
Jodrell Mk II	81.3 ± 8.6 [c]	—	31 x 33	Lewis and Davies 1973
NRAO 91m	59 [d]	—	10 x 10	Dickel and Rood 1978
Dwingeloo 25m	106 ± 7 [e]	—	36 x 36	Allen and Shostak 1979
	102.9 ± 2.1 [f]	—	36 x 36	(new baseline)
WSRT	—	87.8 [g]	?? x ??	van der Kruit and Shostak 1984
VLA D	67 ± 3 [h]	—	30 x 30	Dickey, Hanson, and Helou 1990
VLA C	92.8 ± 0.3 [i]	—	30 x 30	this work
VLA CS	99.3 ± 0.3 [i]	—	30 x 30	this work
VLA C+CS+D	99.7 ± 0.2 [i]	102.1 ± 0.2 [j]	30 x 30	this work

[a] Flux integral corrected for the antennas response (primary beam). For single-dish measurements we preferentially report only the observed flux density, as this correction is a very dicey matter, depending for instance on the galaxy's size as a function of velocity.

[b] 1σ error including baseline uncertainties, but otherwise unspecified.

[c] 1σ error including baseline uncertainties, but otherwise unspecified.

[d] No error quoted.

[e] Error is dominated by 7% error in absolute flux scale.

[f] This is the flux integral found by fitting a linear baseline to the regions the VLA spectrum suggests are free of emission. The 1σ error reflects our estimate of the thermal noise alone (see text), presumed to add in quadrature.

[g] No error estimate was given.

[h] It is not clear whether this flux has been corrected for primary beam attenuation, though our best guess is not. Similarly the nature of the quoted error is unclear.

[i] 1σ error reflecting map noise alone (presumed to add in quadrature within the images and from channel to channel).

[j] Uncorrected 1σ error scaled by the ratio of corrected to raw flux. This is probably an underestimate but the primary beam correction leads to map noise which depends on position, making correct noise computation difficult. The rms noise quoted here is consistent with the distribution of flux densities measured in the primary-beam-corrected, line-free channels, with those again assumed to add in quadrature when summing over velocity; this gives $1\sigma \sim 0.14$ Jy km/sec.

Table 4: Comparison of HI flux integrals for NGC 1058 measured by single dishes and interferometers.

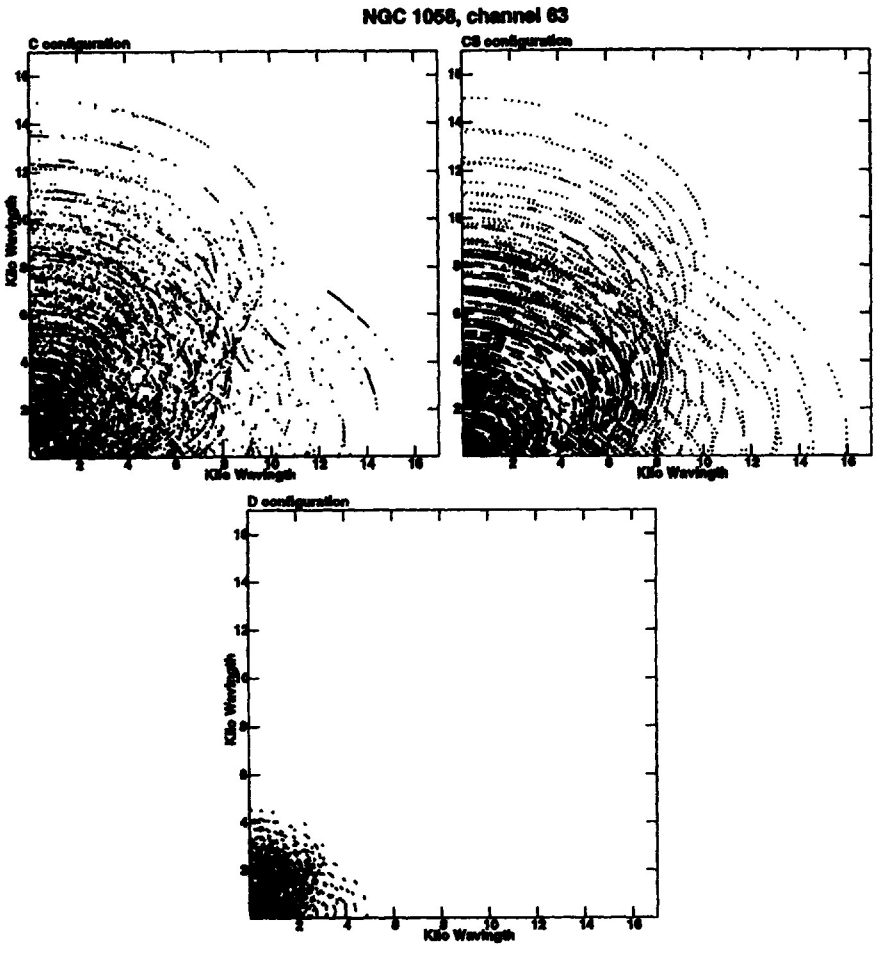


Figure 1: Comparison of the *uv*-coverage obtained from the C, CS, and D configuration observations. These have been scaled to have roughly the same total number of points plotted; the observations covered ±6, ±4.5, and ±5.5 hours about the meridian in the C, CS, and D configurations, respectively. Note the high concentration of short spacings in the CS compared to the C configuration.

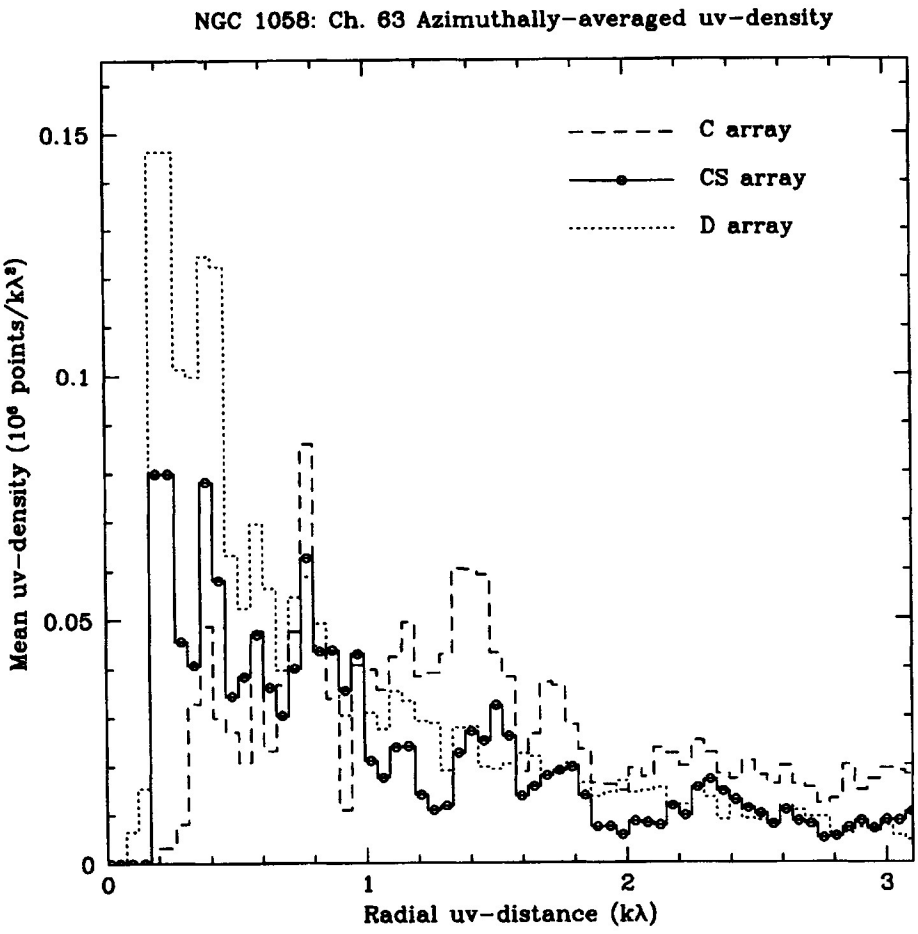


Figure 2: The density of data in the *uv*-plane as a function of radius for the C, CS, and D configurations, obtained as the Fourier transform of the naturally-weighted dirty beam. The resolution is 48.4 wavelengths.

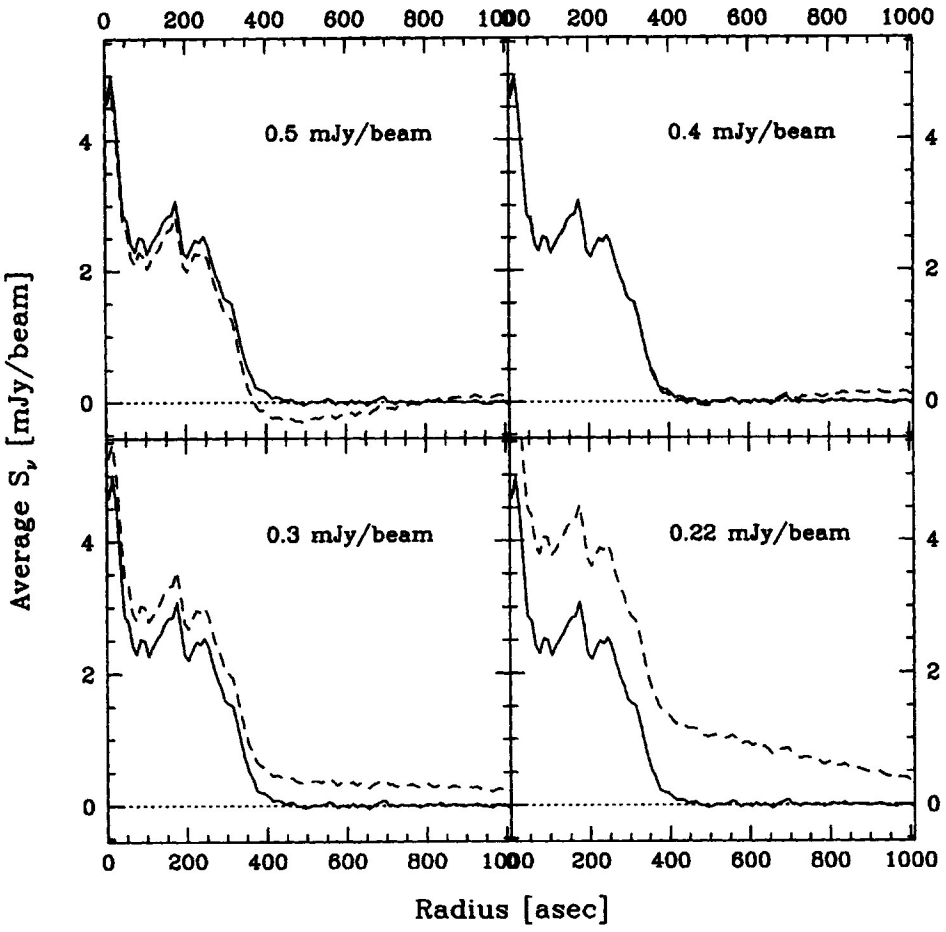


Figure 3: Azimuthal averages of the flux densities after deconvolution of the combined C+CS+D configuration robustly-weighted images of the channel centered on 520.6 km/sec. The noise level is about 0.5 mJy/beam. The solid line, identical in each panel, shows the results of a deep APCLN (to 0.01 mJy/beam); the dashed line shows the results of an unboxed VTESS, with the labels indicating the required noise fits. All models were convolved with the beam given by APCLN (15.44x13.52 arcsec), and residuals were restored with no scaling. VTESS was run to completion, taking between 20 and 50 iterations; further iteration did not qualitatively change the resulting images. VTESS deconvolved the entire inner quarter of the maps, which included the entire radial range illustrated here. This particular APCLN used a single box, but the results for an unboxed APCLN were indistinguishable in these plots.

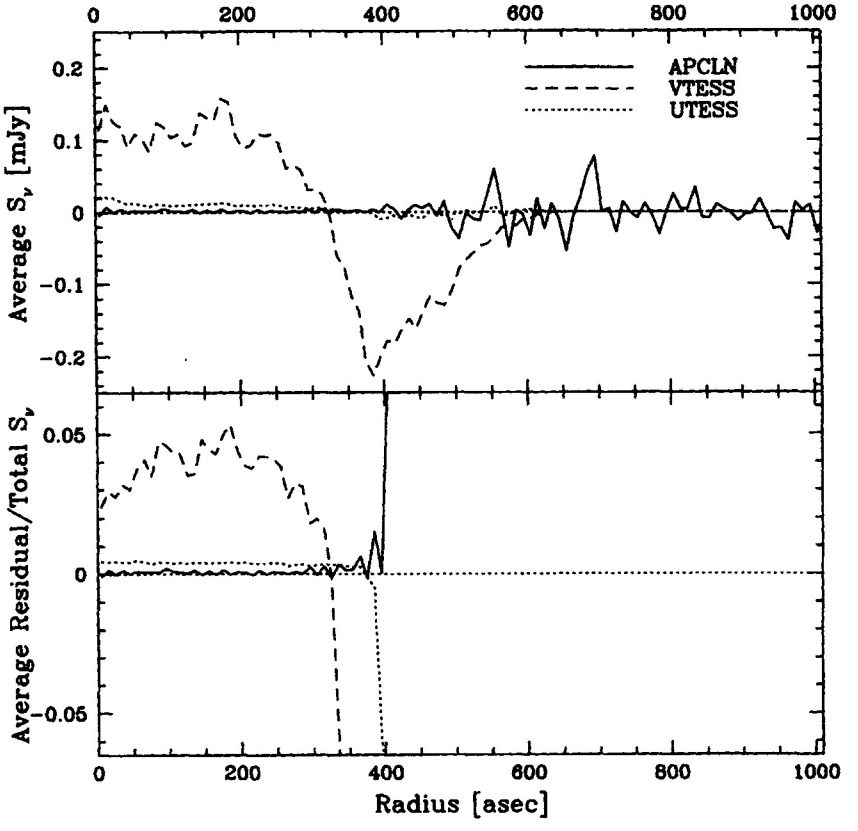


Figure 4: Residuals of boxed deconvolutions of the robustly-weighted, C+CS+D configuration images of channel 63, centered on 520.6 km/sec. The noise level is about 0.5 mJy/beam. The solid line corresponds to a deep APCLN (to 0.01 mJy/beam); the dashed line represents a VTESS run (past convergence) to 30 iterations, with a required noise level of 0.3 mJy/beam; while the dotted line shows a UTESS run with the same noise level, to convergence at 439 iterations. The top panel shows the azimuthally-averaged flux densities, while the bottom panel shows those flux densities divided by the correspondingly averaged total (model plus residuals) flux densities. The zeroing of all but the APCLN line in the top panel beyond 620 arcsec corresponds to the edge of the single box used, while the bottom panel is artificially limited to radii less than 420 arcsec to make the figure easier to read. The top panel shows the systematic residuals left by both VTESS and UTESS, with the latter doing a much better job but still unable to zero the residuals; the panel below illustrates the fractional level of these residuals within the emission region, about 5% for VTESS and a factor 10 lower for UTESS. The total residual flux density is about 5.5 mJy for UTESS and around 210 mJy for VTESS, compared with a total flux density of about 3600 mJy.

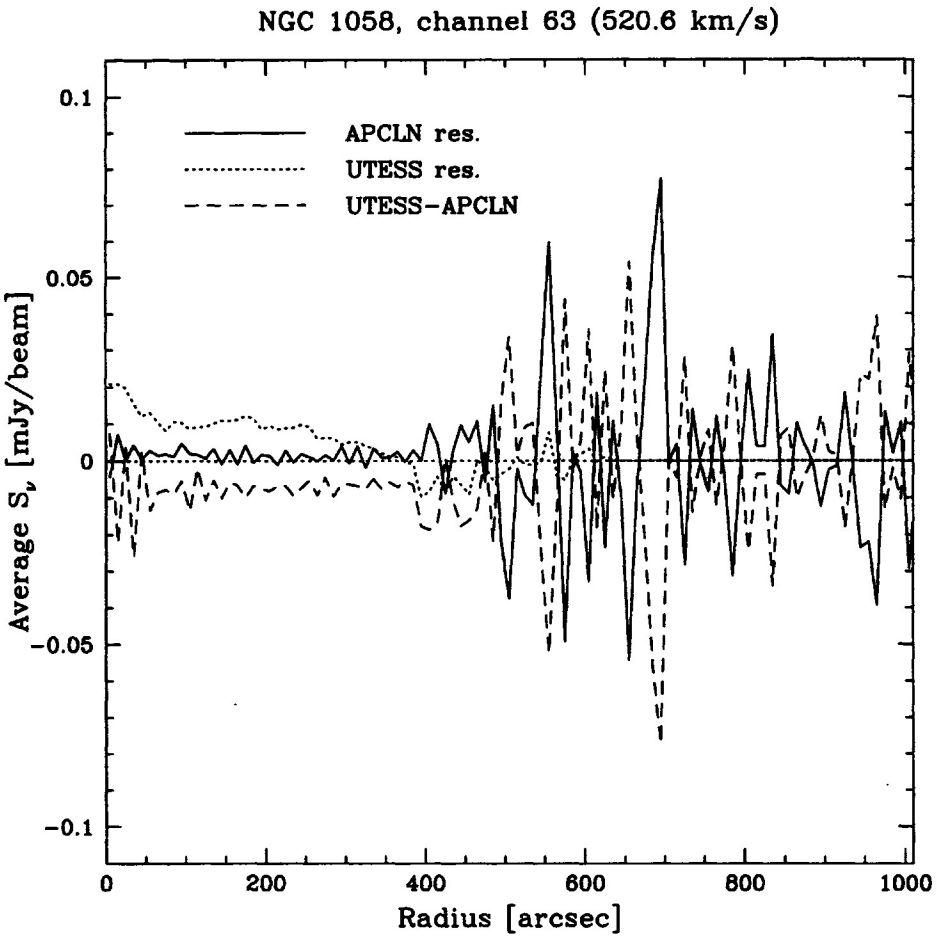


Figure 5: Comparison of boxed APCLN and UTESS for the robustly-weighted, C+CS+D configuration images of channel 63, centered on 520.6 km/sec. The noise level is about 0.5 mJy/beam. The solid line shows the residuals from a deep APCLN (to 0.01 mJy/beam); the dashed line shows the residuals of UTESS with the noise set to 0.3 mJy/beam, run to convergence at 439 iterations; while the dotted line shows the difference between the final (model plus residuals) UTESS and APCLN images. This figure again shows azimuthal averages of the flux densities. On a beam-per-beam basis the results are effectively identical.

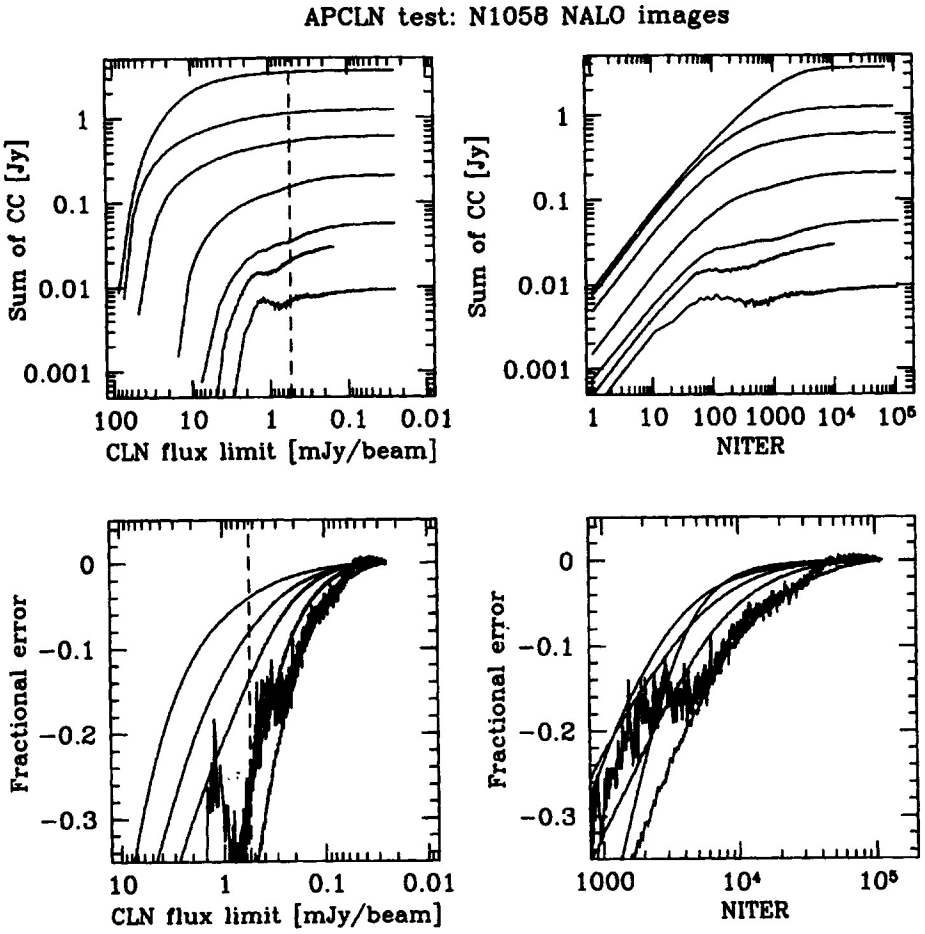


Figure 6: Flux density in CLEAN components, and fractional error. as a function of the APCLN flux limit and number of iterations (NITER). The plots shown here are for the NALO (naturally-weighted, tapered, low-resolution) ($\sim$ 55 arcsec) images made from the combined C+CS+D data set; see Fig. 7 for a comparison with the larger high-resolution images. The fractional error is calculated assuming the deepest CLEAN recovers the correct total flux density. The vertical dashed line indicates the noise level outside the region with emission ($1\sigma \approx 0.5$ mJy/beam), which is consistent for all channels. The channels shown are (in order of increasing HI size and total flux density) channels 48, 49, 50, 53, 56, 58, and 63; channel 49 is excluded from the fractional error plots because it was not CLEANed as deeply as the others. The same order (top-to-bottom $\rightarrow$ high-to-low channels) is preserved in the fractional error plots, with two exceptions: (1) channel 48 wanders up and down, reflecting the decrease in CLEAN flux between $\sim$ 100 and a few thousand components; (2) channel 63 shows a much sharper improvement in fractional error as a function of the number of iterations than do the other channels.

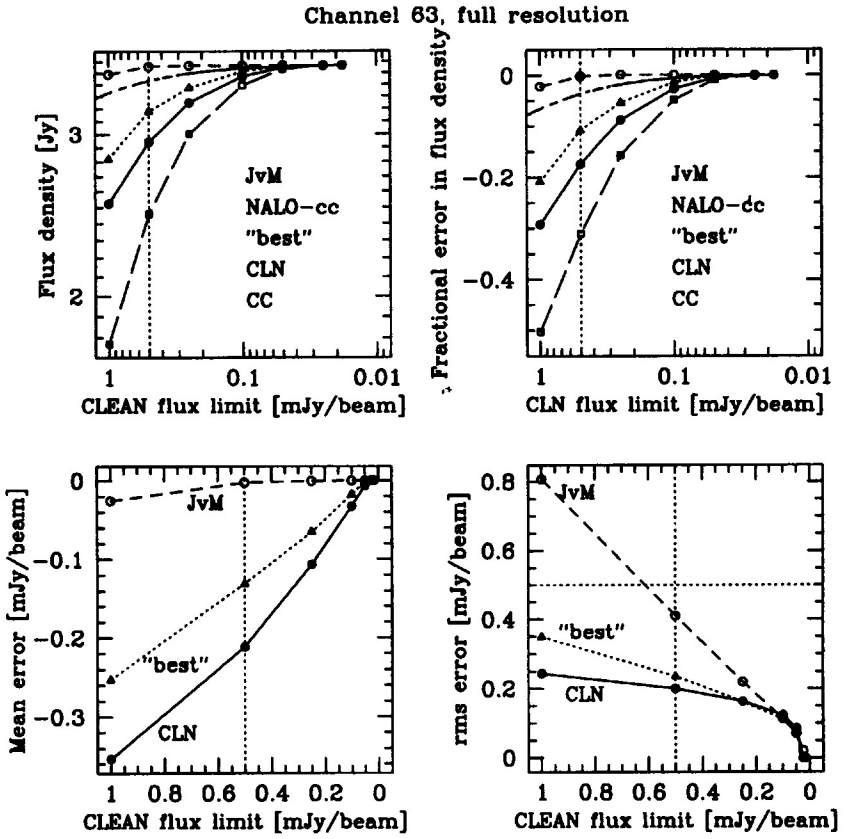
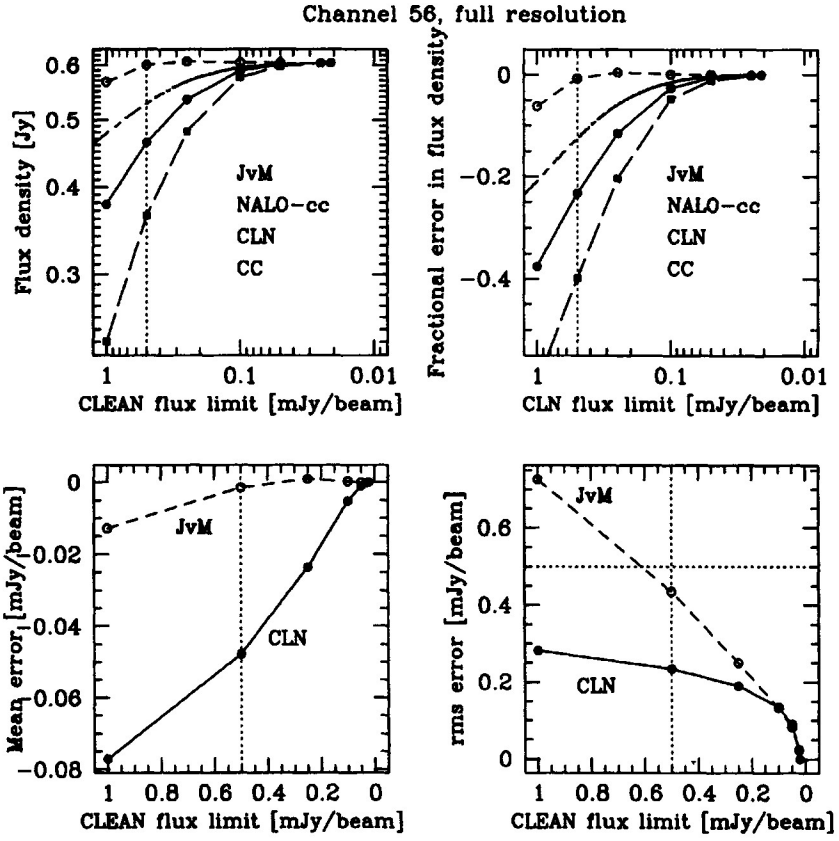


Figure 7: The usefulness of scaling the residuals, for the full-resolution, robustly-weighted images. The upper plots show the total flux density in the deconvolved region of the final image, and the corresponding fractional error (assuming the deepest CLEAN to be perfect). The lower plots show the mean and the rms dispersion in the "error" maps, calculated as the difference between each image and the deepest CLEAN map with standard (no scaling) residuals restored. All except the NALO-cc curves refer to the full-resolution, robustly-weighted images. The CC curves refer to CLEAN components only (no residuals added); the CLN curves refer to the standard CLEAN images (residuals added, no special scaling); "best" curves correspond to images made with the residuals scaled to minimize the median of the difference between the image and the "truth" (deepest CLEAN) image; JvM refers to images with residuals scaled according to the JvM algorithm, intended to give the correct total flux density; and NALO-cc refers to the CLEAN components from the NALO (tapered, low-resolution) images, and is included mostly for comparison with Fig. 6. Note that the CC flux is consistently a factor ~ 2 worse than the CLN estimate, in terms of the fractional error; this corresponds (as of course it must) with the JvM scaling factor.

(a) Channel 63.


Channel 56, full resolution
Flux density [Jy]
0.6
0.5
0.4
0.3
JvM
NALO-cc
CLN
CC
1
0.1
0.01
CLEAN flux limit [mJy/beam]
Fractional error in flux density
0
−0.2
−0.4
JvM
NALO-cc
CLN
CC
1
0.1
0.01
CLN flux limit [mJy/beam]
Mean error [mJy/beam]
0
0.02
0.04
0.06
−0.08
JvM
CLN
1
0.8
0.6
0.4
0.2
0
CLEAN flux limit [mJy/beam]
rms error [mJy/beam]
0.6
0.4
0.2
0
JvM
CLN
1
0.8
0.6
0.4
0.2
0
CLEAN flux limit [mJy/beam]

(b) Channel 56.

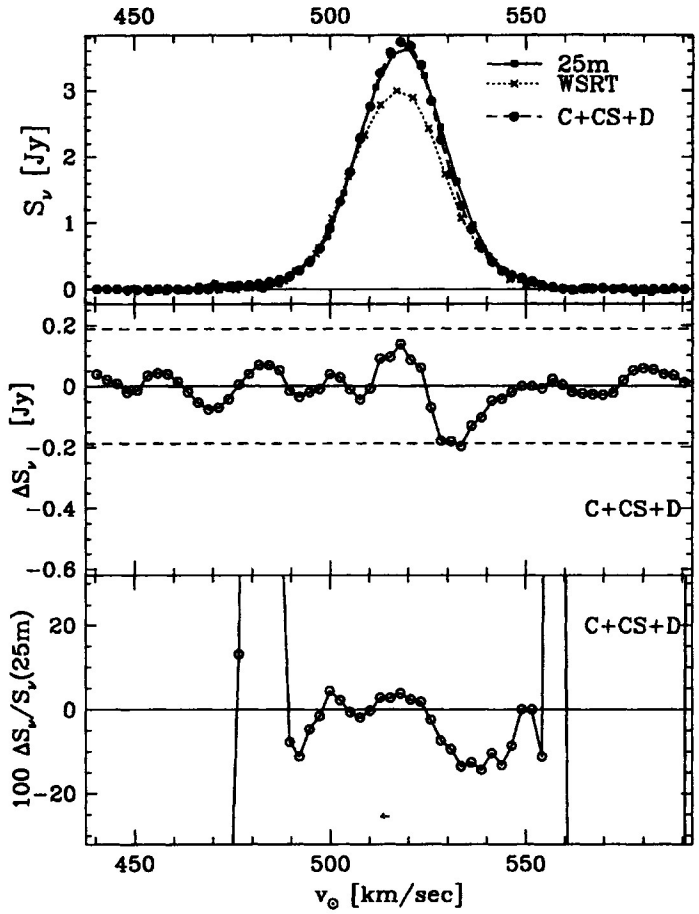


Figure 8: Comparison of single-dish and interferometric HI spectra. Velocities are heliocentric and calculated using the optical convention. VLA flux densities were measured off deeply CLEANed (0.05σ), NALO images with (unscaled) residuals restored, and corrected for the VLA's primary beam response.

(a) Top panel: HI spectra from the Dwingeloo 25m (solid line and squares; Allen and Shostak 1978), the WSRT (dotted line and crosses; van der Kruit and Shostak 1984), and the combined VLA C+CS+D configurations (dashed line and circles). The velocity resolutions are 8.2, 8.25, and 2.59 km/sec, while the data are sampled every 4.13, 4.12, and 2.59 km/sec, respectively. A linear baseline has been removed from the 25m data. Typical 1σ errors are ~62 mJy for the 25m and ~9 mJy for the VLA, while no error was quoted for the WSRT. Zero flux density is indicated by the solid horizontal line.

(b) Middle panel: The difference between the VLA and 25m spectra, with the latter interpolated linearly to the VLA resolution. The solid horizontal line shows zero flux density, while the dashed lines represent $\pm 3\sigma$ ($\pm$188 mJy). Note that the errors are expected to be correlated over the resolution of the single-dish data, 8.2 km/sec = 3.2 VLA channels.

(c) Bottom panel: The difference spectrum from the middle panel, expressed as a fraction of the interpolated single-dish flux density. The solid horizontal line again corresponds to 0.

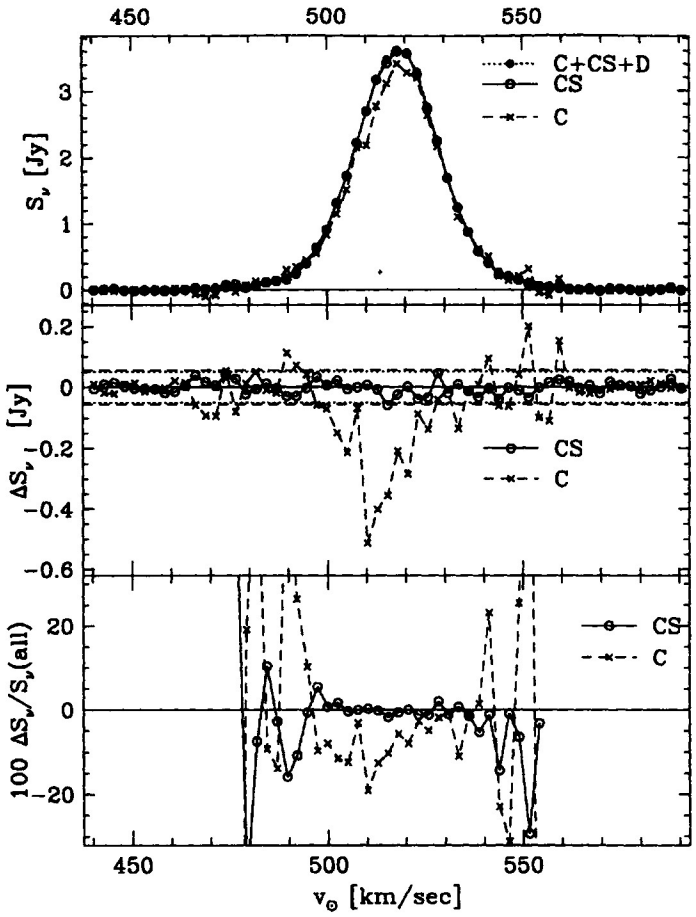


Figure 9: HI spectra from C, CS, and C+CS+D configurations. Velocities are heliocentric and calculated using the optical convention; flux densities have *not* been corrected for the VLA's primary beam response. All flux densities are taken from deeply CLEANed (0.05σ) NALO images with (unscaled) residuals restored. The velocity resolution and sampling are both 2.59 km/sec.

(a) Top panel: The HI spectra, with the configurations as indicated. 1σ noise levels are roughly 16.6, 14.5, and 8.8 mJy for the C, CS, and C+CS+D configurations respectively. The solid horizontal line indicates 0 Jy.

(b) Middle panel: Difference between the C and CS configuration spectra and the C+CS+D configuration spectrum. The solid horizontal line indicates 0 Jy, while the dashed and dotted horizontal lines show $\pm 3\sigma$ for the C and CS configuration difference spectra respectively ($\pm$56.4 mJy for C, $\pm$50.9 mJy for CS configuration).

(c) Bottom panel: The difference spectrum from the middle panel, expressed as a fraction of the C+CS+D configuration spectrum. These spectra are truncated outside the velocity range 475 to 555 km/sec for clarity. The solid horizontal line again corresponds to 0.

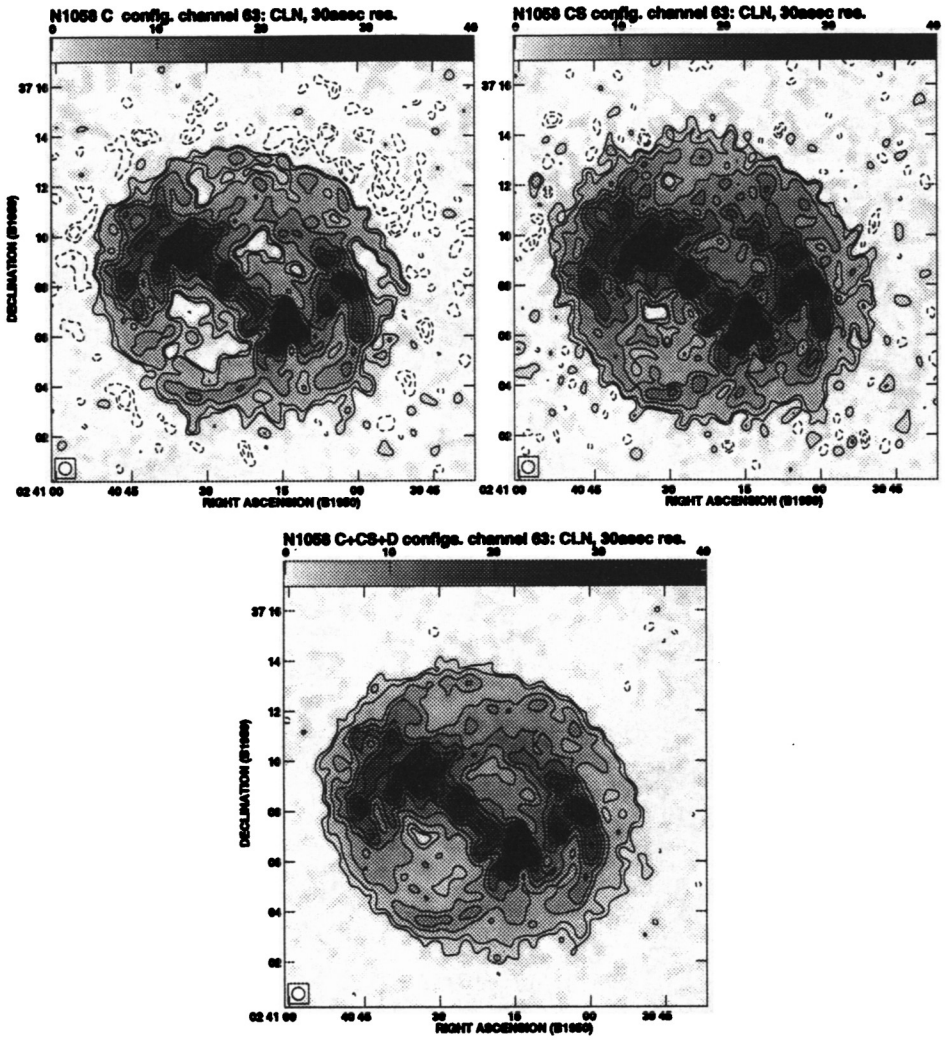


Figure 10: Comparing the configurations: grey-scale/contour plots of one of the channels with the most extended HI emission, that centered on 520.6 km/sec. These are the robustly-weighted images, CLEANed to 0.05σ, the CLEAN components restored with the beam given by APCLN with (unscaled) residuals added, and the result convolved to a common 30 arc-sec resolution. The contouring levels in each case are -5, -3, 3, 5, 10, 15, ..., 65 times 0.6 mJy/beam, while all greyscales extend from 0 to 40 mJy/beam. The noise levels are 0.9, 1.0, and 0.6 mJy/beam for the C, CS, and combined C+CS+D configurations, respectively. (a) C configuration; (b) CS configuration; (c) C+CS+D configuration.

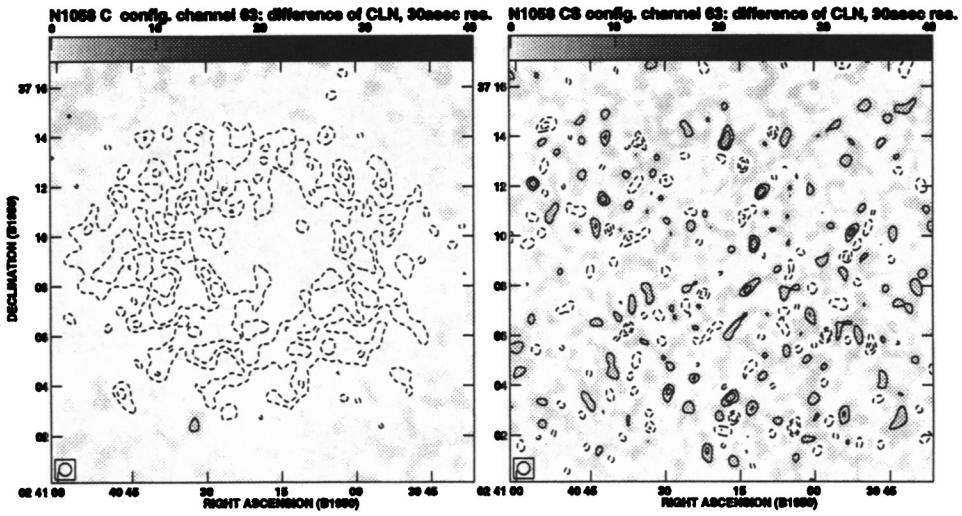


Figure 11: Comparing the configurations: grey-scale/contour plots of the differences between the deconvolved C and CS configuration images and those from the combined C+CS+D configurations for the HI in a single channel centered on 520.6 km/sec. Images, contours, and grey-scales as in Fig. 10.

(a) C configuration difference image. The mean flux difference over the galaxy is -1.7 mJy/beam, with an rms of 1.3 mJy/beam; in regions free of emission and well away from the negative bowl, the mean difference is consistent with zero, with an rms of 0.5 mJy/beam.

(b) CS configuration difference image. The mean flux difference over the galaxy is 0.02 mJy/beam, with an rms of 1.1 mJy/beam; well away from the emission the mean difference is 0.04 mJy/beam over a similar area, with an rms of 0.94 mJy/beam.

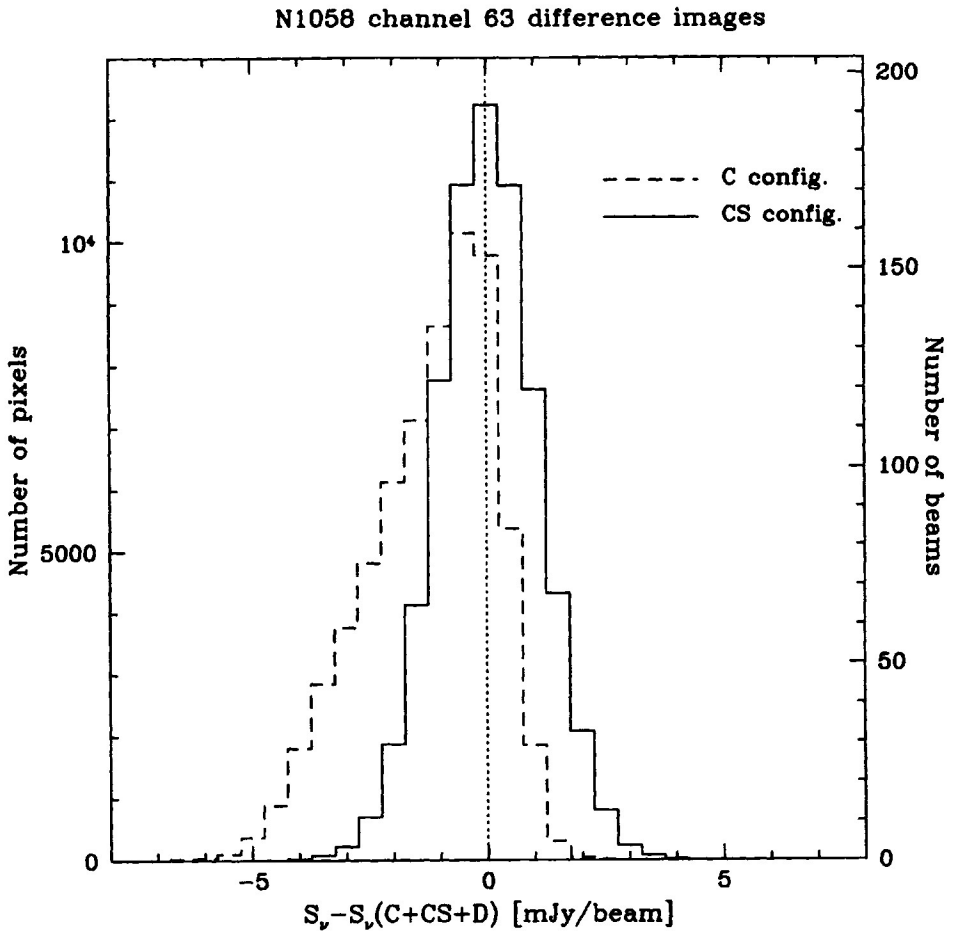


Figure 12: Comparing the configurations: histograms of the differences between the C/CS and C+CS+D configuration images for the HI in the channel centered on 520.6 km/sec. Data are as described in Fig. 10. The C configuration data are shown with a dashed line, the CS configuration data with a solid line, and the dotted vertical line represents 0 mJy/beam. The histograms are over the areas shown in Fig. 10, giving a total of 64009 pixels (1004.2 30-arcsec beams). The mean of the C configuration histogram is −1.6 mJy/beam, and the dispersion is 1.3 mJy/beam; both are significantly different from the pure-noise case. The corresponding figures for the CS configuration data are 0.02 and 1.1 mJy/beam, consistent with thermal noise centered on 0 mJy/beam.

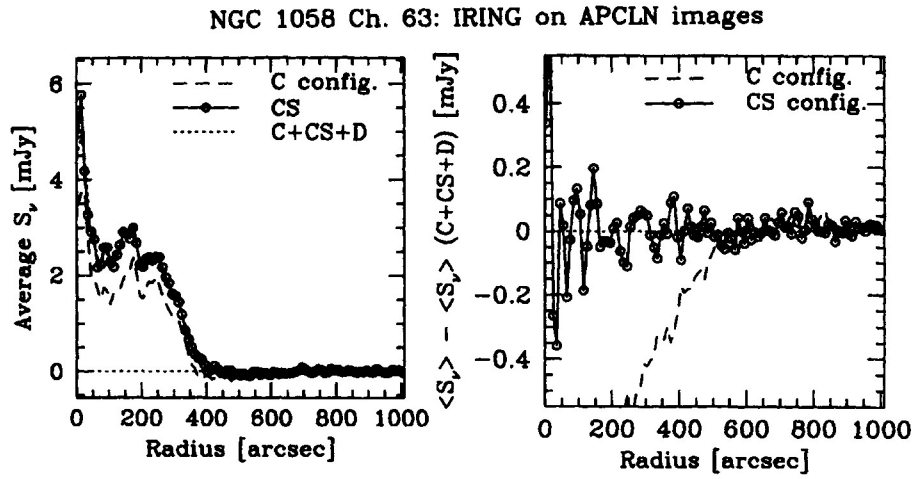


Figure 13: Comparing the configurations: azimuthal averages over rings centered on the centroid of the C+CS+D configuration image (which corresponds fairly well to the geometric center of the emission). As usual we plot the results for the channel centered on 520.6 km/sec. These are taken again from the robustly-weighted images, CLEANed to 0.05σ, the CLEAN components restored with the beam given by APCLN with (unscaled) residuals added, and the result convolved this time to a common 15.44x13.52 arcsec resolution (set by the C+CS+D configuration beam).

(a) Azimuthally-averaged flux densities as a function of radius. The dashed, solid, and dotted lines show the results for the C, CS, and C+CS+D configuration images, respectively; the open circles also represent CS configuration data. Once again note the agreement between the CS and combined configurations, while C configuration is low throughout the emission region. The noise level increases towards small radii because the number of pixels entering into the averages is reduced.

(b) The difference between the azimuthally-averaged flux densities for the single and for the combined configurations. The dashed line represents the difference between C and C+CS+D configuration data; the solid line and open circles shows the difference between CS and C+CS+D. The noise increases towards smaller radii as in (a). The CS configuration are consistent with the combined C+CS+D configuration data within the thermal noise, while the C configuration shows the classic negative bowl due to missing short spacings.

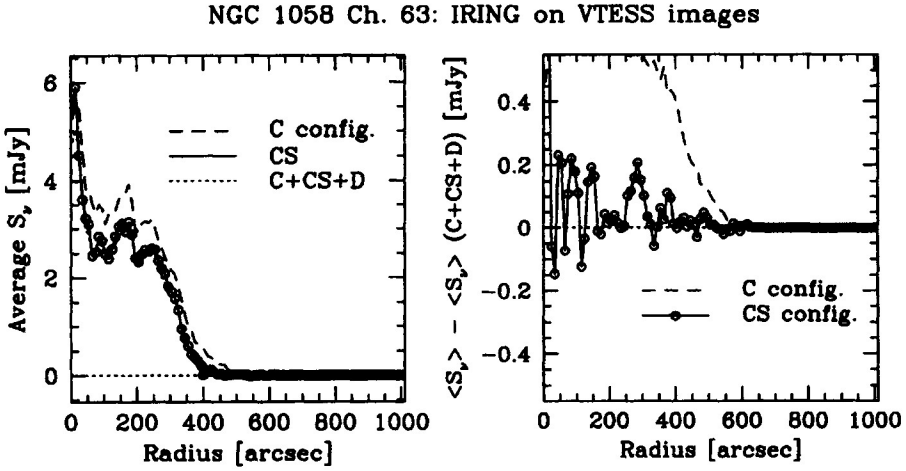


Figure 14: Comparing the configurations: VTESS images. As fig. 13 but for (boxed) VTESS rather than APCLN images. (a) Azimuthal averages, (b) azimuthal averages of difference images. VTESS has inverted the C configuration bowl but has not changed its magnitude.

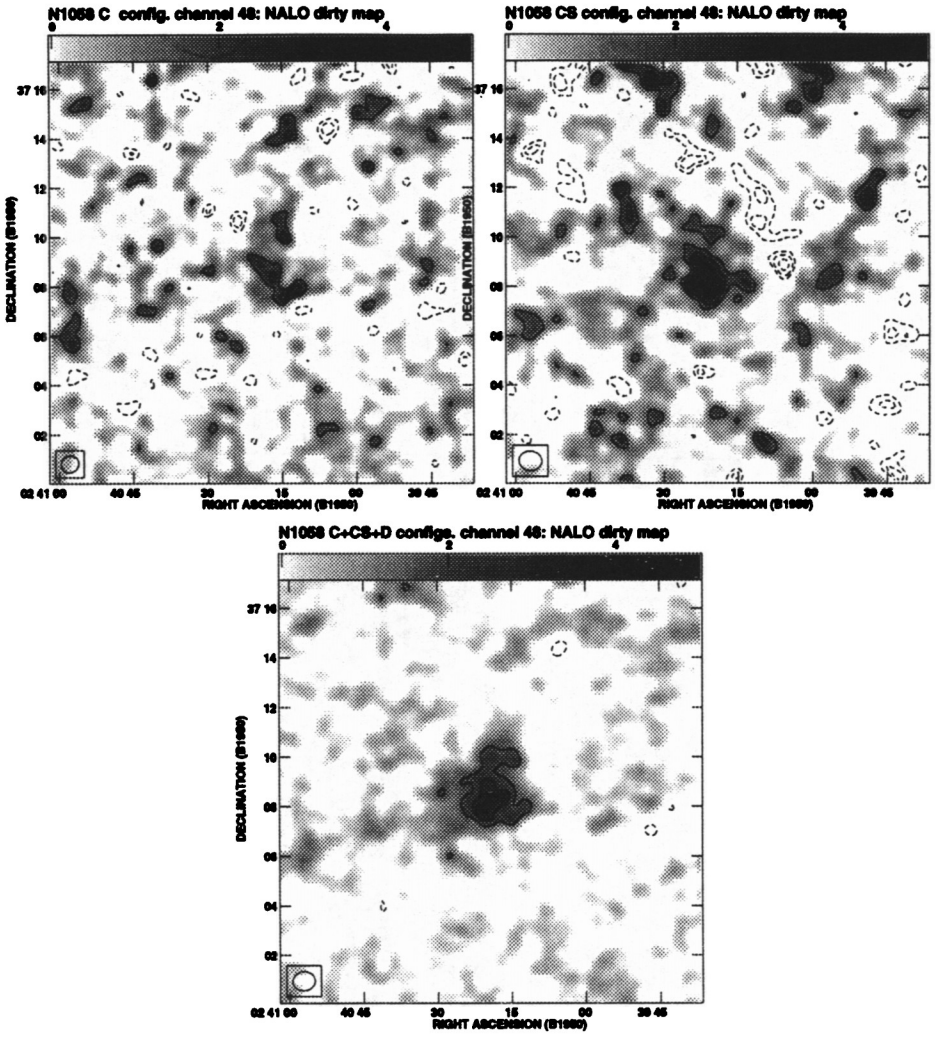
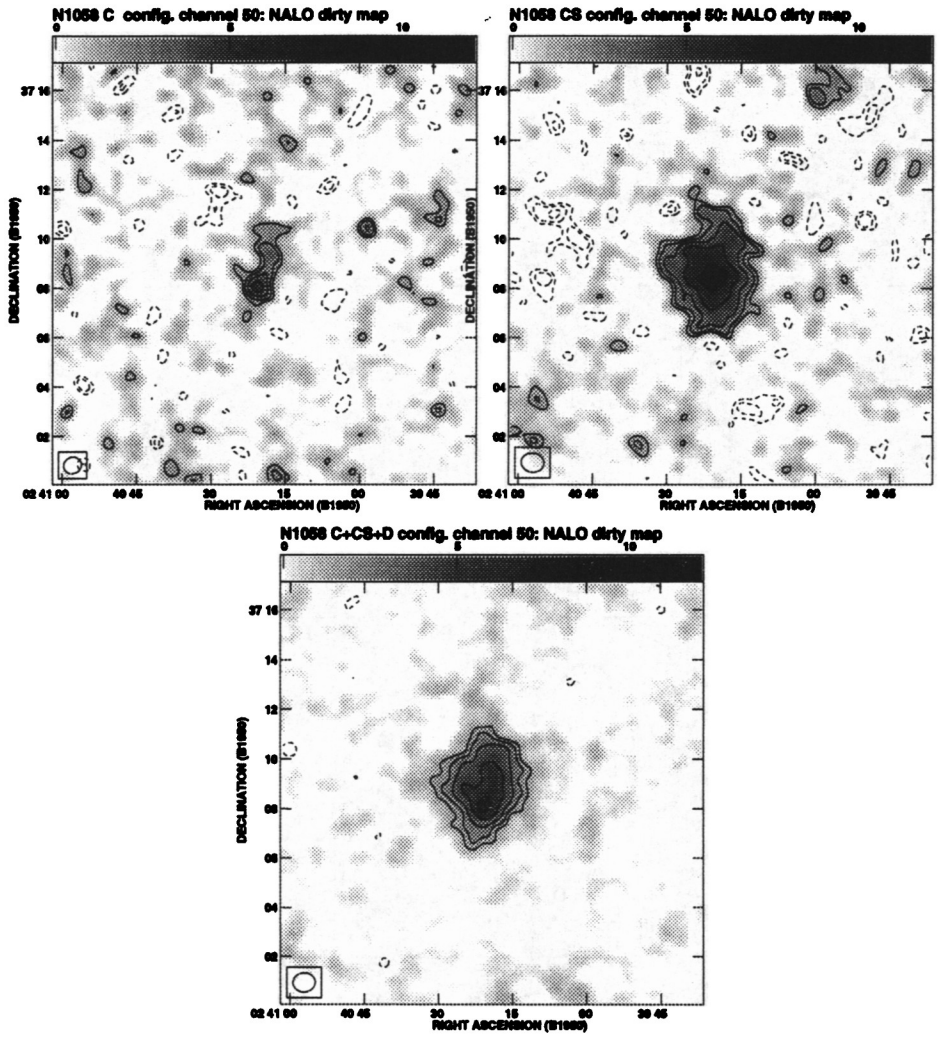


Figure 15: Comparing the configurations: dirty maps of edge channels. These images are based on the tapered, naturally-weighted data, which should give the best sensitivity to large-scale structures by giving maximum weight to the shortest spacings. All maps are contoured logarithmically at $\pm 2.828^{n/2}$ times 0.6 mJy/beam, for n = 0, 1, 2, ... Gaussian fits to the dirty beams are shown in the lower left corner of each plot, and are 43.7x39.8 arcsec at $-66.1°$, 55.3x46.7 arcsec at 81.6°, and 53.7x46.9 arcsec at $-83.2°$ for the C, CS, and C+CS+D configurations, respectively.

(a) Channel 48, centered on 559.4 km/sec. Here the greyscale ranges from 0 to 5 mJy/beam. C configuration loses the larger-scale emission, distorting the character of the emission region.



(b) Channel 50, centered on 554.2 km/sec. Here the greyscale ranges from 0 to 12 mJy/beam. The "skirt" surrounding the outermost C+CS+D configuration contour is due to the wings of the naturally-weighted beam. Once again C configuration qualitatively misrepresents the emission, while CS configuration while noisy retains the overall size of the source and the presence of a central bright spot.

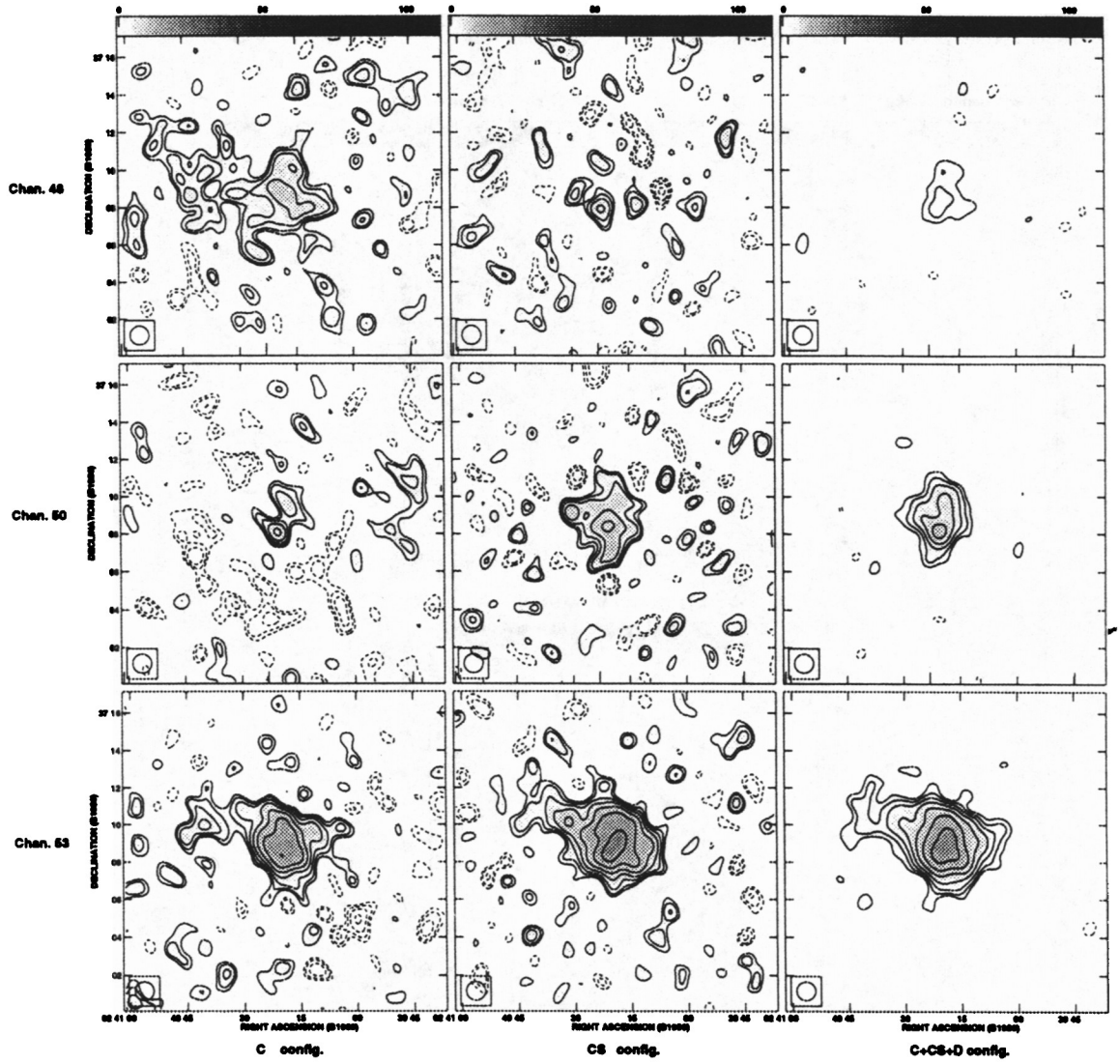
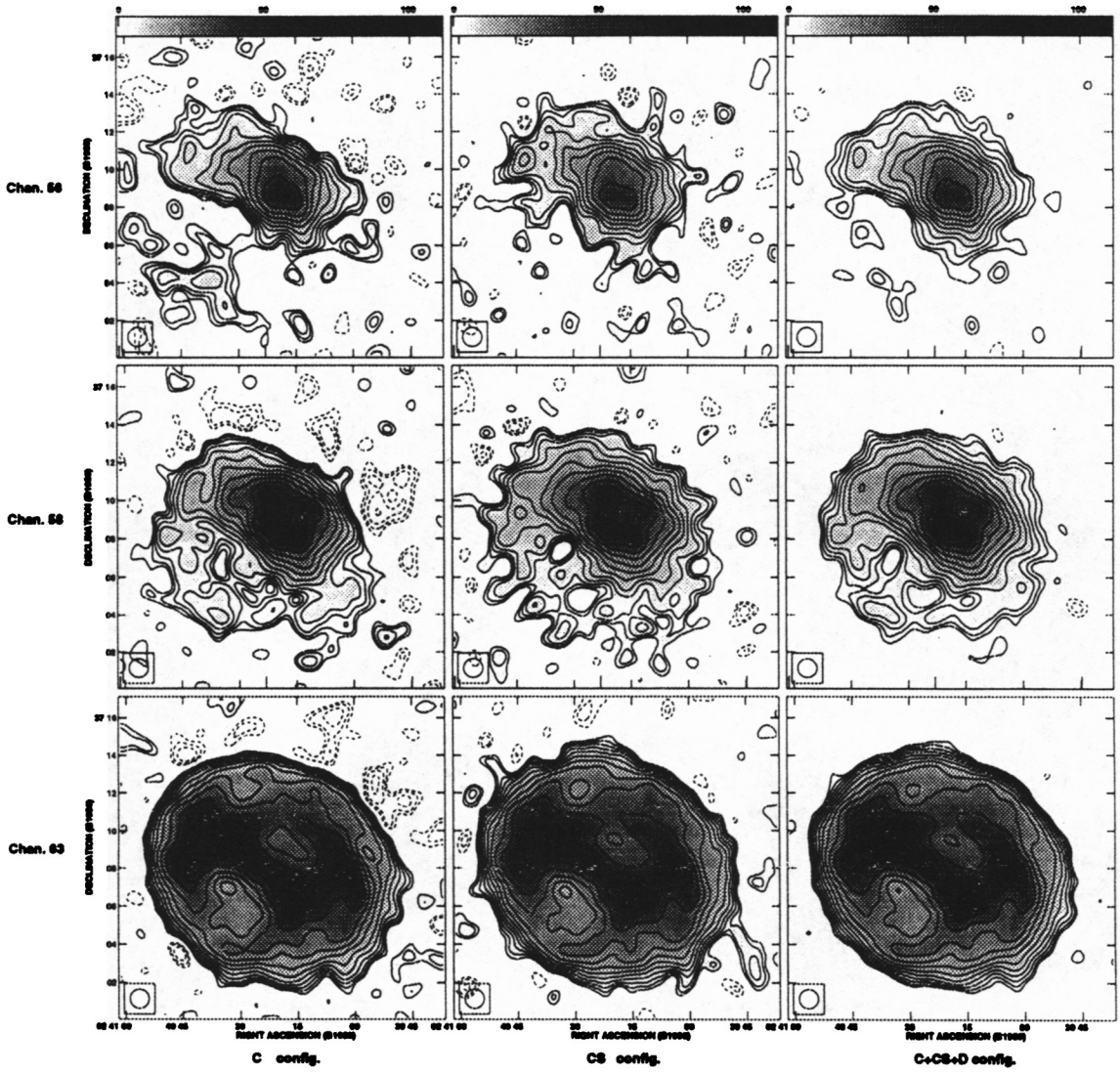


Figure 16: Comparing the configurations: six sample channels of the CLEANed NALO images, convolved to 60arcsec resolution. Left-hand column is C, middle column CS, and right column C+CS+D configuration. The channels are labeled on the left-hand side; see Table 2 for further information. All maps are contoured logarithmically at $\pm 2.828^{n/2}$ times 0.6 mJy/beam, for $n = 0, 1, 2, \ldots$ The gray scale ranges from 0 to 110 mJy/beam. The rms noise in emission-free regions is $\sim$ 1.1, 1.0, and 0.6 mJy/beam for the C, CS, and C+CS+D configuration images, respectively.

(a) Channels 48, 50, and 53. The blotchiness of the C/CS images is due to the deep CLEAN.



(b) Channels 56, 58, and 63. Note the negative bowl in the C configuration images.