## **Error Recognition and Image Analysis**



#### **Alison Peck**

Many thanks to Greg Taylor with additional contributions from Urvashi Rao, Sanjay Bhatnagar, Gustaaf van Moorsel, Justin Linford, Ed Fomalont



Atacama Large Millimeter/submillimeter Array
Karl G. Jansky Very Large Array
Robert C. Byrd Green Bank Telescope
Very Long Baseline Array



# This is mostly a review...

This week, we have learned about calibration, imaging, and about things that can go wrong with both.

#### To recap:

- Non-imaging analysis describes how to extract information directly from the (u,v) data
- Image analysis describes the many ways in which useful insight, information and parameters can be extracted from the image.
- Error recognition is used to determine defects in the visibility data and/or images during and after the 'best' calibration, editing, etc.



## **Issues: Radio Frequency Interference (RFI)**

#### **Culprits:**

Wireless devices at low frequencies Television/radio signals Aircraft, Radar Car radars

#### Problematic regions:

1215 – 1300 MHz mobile comm.

1675 – 1710 MHz mobile comm.

1755 – 1850 MHz mobile comm.

2155 – 2200 MHz mobile comm

4200 – 4220 MHz altimeters

4380 – 4440 MHz altimeters

5925 – 7250 level-probing-radar

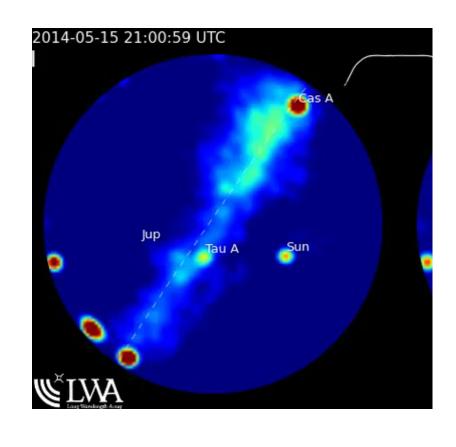
14000 – 14500 air to ground

15400 – 15700 radar

76000 − 77000 automobile radar ← Al MA

(and harmonics)

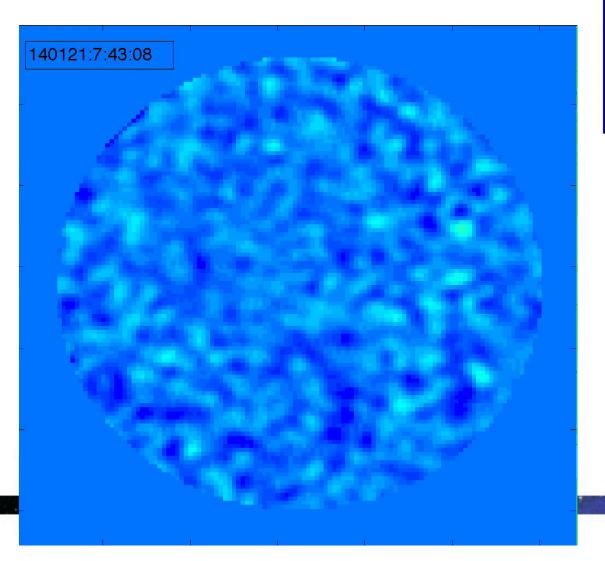
**VLA VLBA** 

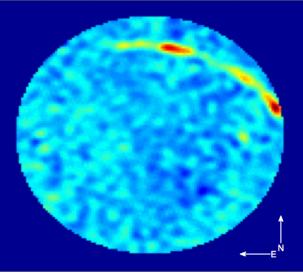




### **Fireballs**

Discovery of broad-band emission from meteors



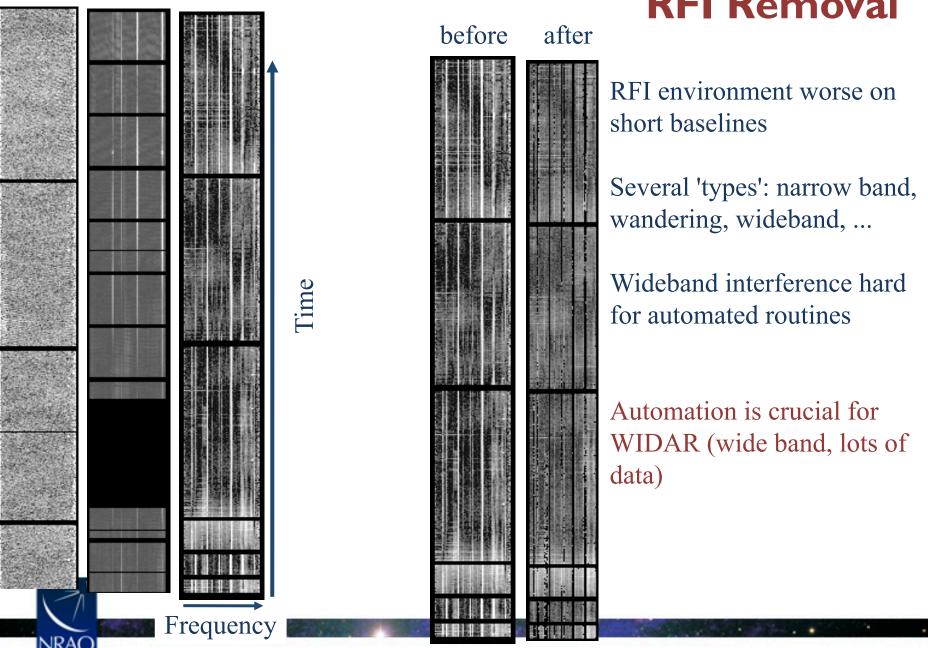


These happen about 1/week Peak is ~3000 Jy at 40 MHz

Obenberger et al

# 35 km 12 km 3 km baseline

### **RFI Removal**



**AIPS: SPFLG** 

### **General Procedure**

We assume that the data have been edited and calibrated reasonably successfully including self-calibration if necessary.

#### So, the first serious display of an image leads one—

- to inspect again and clean-up the data repeating some or all of the previous reduction steps.
  - removal of one type of problem can reveal next problem!
- once all is well, proceed to image-analysis and obtaining scientific results from the image.



### Image Plane or Data (u,v) Plane Inspection?

### Errors obey Fourier transform relationship

Narrow feature in (u,v) plane <-> wide feature in image plane

Wide feature in (u,v) plane <-> narrow feature in image plane

Note: often easier to spot narrow features

Data (u,v) amplitude errors <-> symmetric image features

Data (u,v) phase errors <-> asymmetric image features

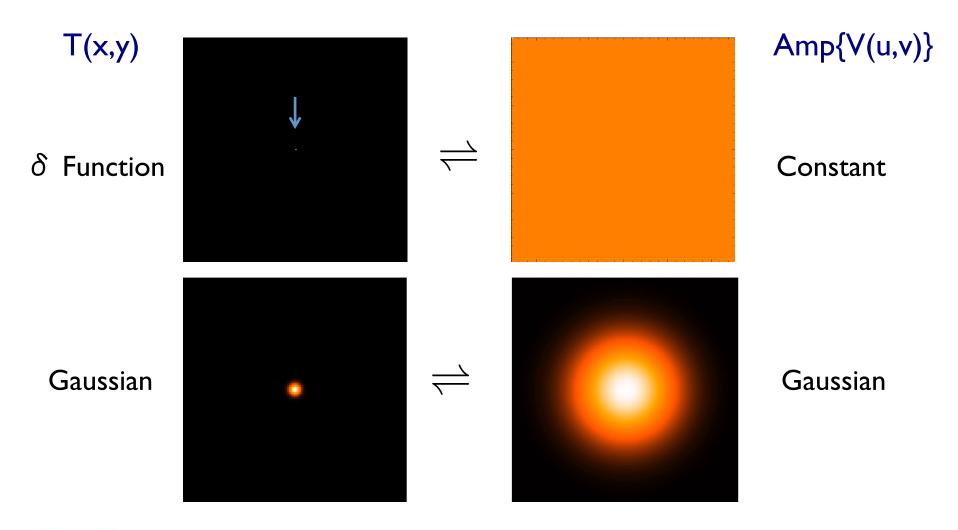
An obvious defect may be hardly visible in the transformed plane

A small, almost invisible defect may become very obvious in the transformed plane

Let's start with the u,v plane



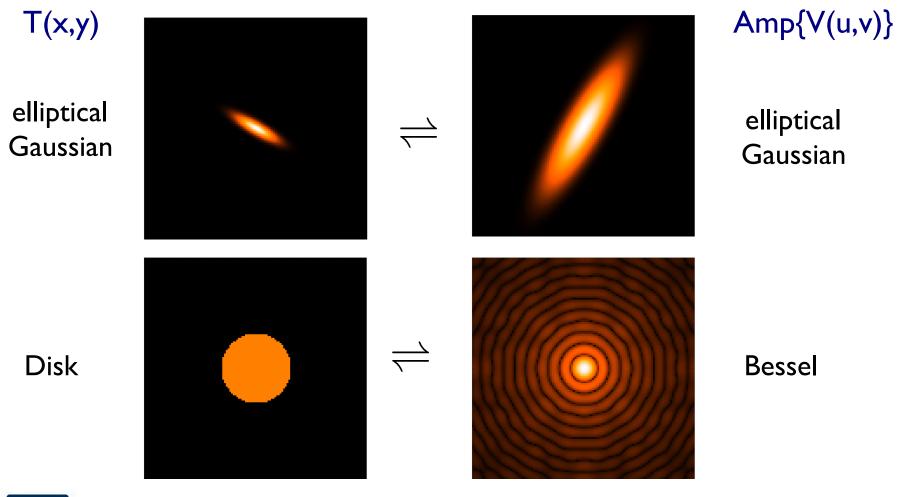
### **Some 2D Fourier Transform Pairs**





narrow features transform to wide features (and vice-versa)

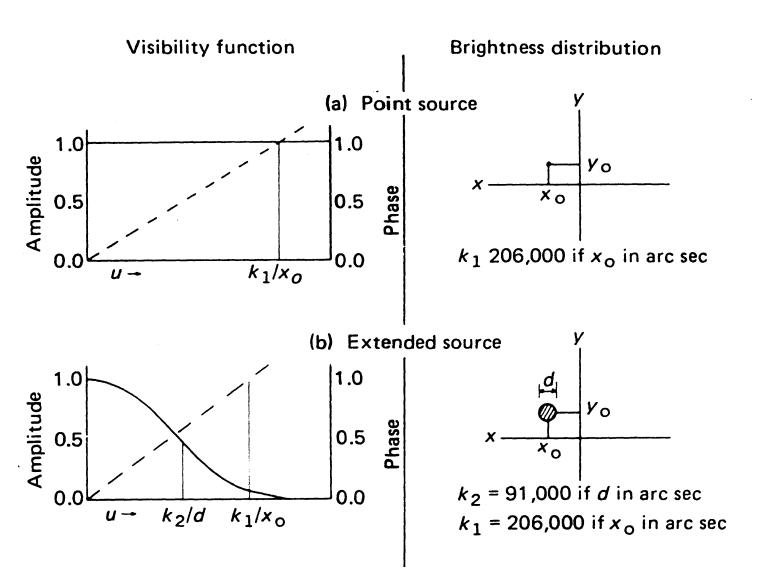
## **2D Fourier Transform Pairs**





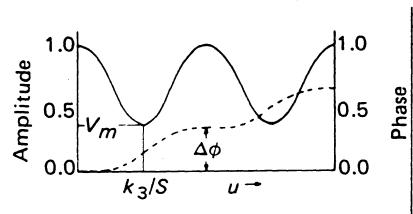
sharp edges result in many high spatial frequencies (sinc function, "ringing", Gibbs phenomenon)

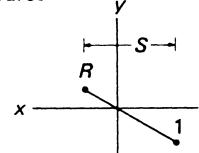
### **Visibility** ← F → **Brightness**



### **Visibility** ← F → Brightness

(c) Point double source

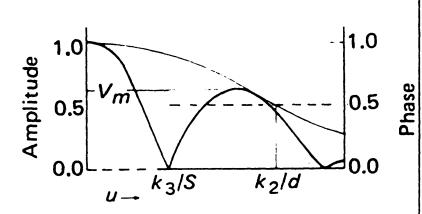


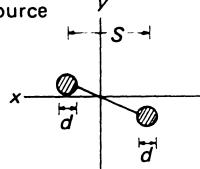


$$k_3 = 103,000 \text{ if } S \text{ in arc sec}$$

$$V_{m} = \frac{R-1}{R+1} ; \Delta \phi = \frac{1}{1+R}$$

(d) Extended double source



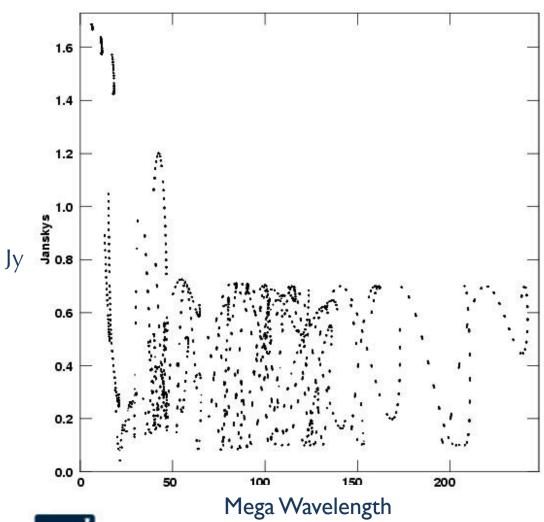


$$k_3 = 103,000 \text{ if } S \text{ in arc sec}$$

$$k_2 = 91,000 \text{ if } d \text{ in arc sec}$$

$$V_m \approx \exp \left\{-3.57 \left(\frac{d}{s}\right)^2\right\}$$





Visibility Amplitude versus Projected (u,v) spacing

General trend of data.

Useful for relatively strong sources.

Triple source model. Large components cause rise at short spacings.

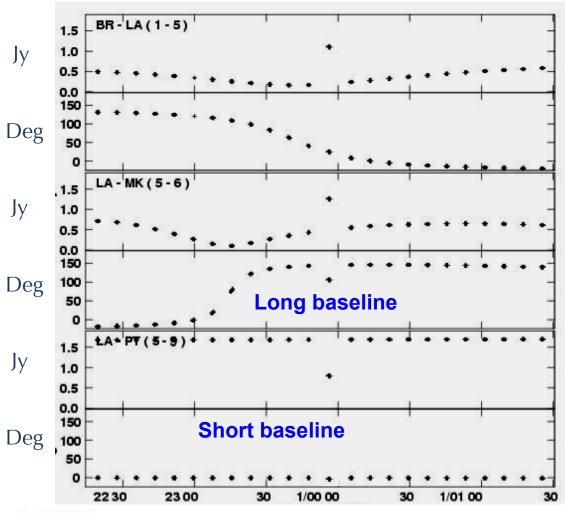
Oscillations at longer spacings suggest close double.



#### List of (u,v) Data

Source=	J0121+11	Freq=	8.434		1858511	Sort= TB		1	RR	
Vis #	IAT	Ant	Su	Fq	U(klam)	V(klam)	W(klam)	Amp	Phas	Wt
2191	0/22:35:08.2	2 5- 6	1	0	94220	23776	100371	0.614	-16	1.0000
3971	0/22:43:43.3	4 5- 6	1	0	97659	24517	96844	0.508	-13	1.0000
6431	0/23:07:05.1	5 5- 6	1	0	106307	26661	86632	0.154	17	1.0000
6611	0/23:07:14.9	8 5- 6	1	0	106364	26677	86557	0.152	17	1.0000
6791	0/23:07:24.8	1 5- 6	1	0	106421	26692	86483	0.150	18	1.0000
6971	0/23:07:34.6	4 5- 6	1	0	106477	26708	86408	0.148	19	1.0000
7151	0/23:07:44.4	7 5- 6	1	0	106534	26724	86333	0.146	19	1.0000
7331	0/23:07:54.3	0 5- 6	1	0	106591	26739	86259	0.144	20	1.0000
7511	0/23:15:06.8	4 5- 6	1	0	109027	27438	82930	0.101	74	1.0000
7691	0/23:15:16.6	7 5- 6	1	0	109081	27454	82854	0.101	75	1.0000
7871	0/23:15:26.5	0 5- 6	1	0	109135	27470	82777	0.102	77	1.0000
8051	0/23:15:36.3	3 5-6	1	0	109189	27486	82701	0.102	78	1.0000
8231	0/23:15:46.1	6 5- 6	1	0	109243	27502	82624	0.103	79	1.0000
8411	0/23:15:55.9	9 5- 6	1	0	109297	27518	82547	0.104	81	1.0000
9701	0/23:31:02.3	NO. 1001 P	1	0	114020	29035	75322	0.260		1.0000
9791	0/23:31:06.2	C 17 1900		0	114040	12607-201-7632		0.261		1.0000
10301	0/23:31:29.8	142. RESERVE 2010		0	114156	29082		0.266		1.0000
10861	0/23:39:02.0	THE RESERVE		0	116320	29863	71379	0.348		1.0000
10951	0/23:39:06.0			0	116339	100000000000000000000000000000000000000	71346	0.348		1.0000
11171	0/23:39:15.8	17 T. San	1	0	116384	29887	71264	0.350		1.0000

Old School, but sometimes worth-while: e.g., can search on e.g. Amp > 1.0, or large weight. Often need precise times in order to flag the data appropriately.



Visibility amplitude and phase versus time for various baselines

Good for determining the continuity of the data

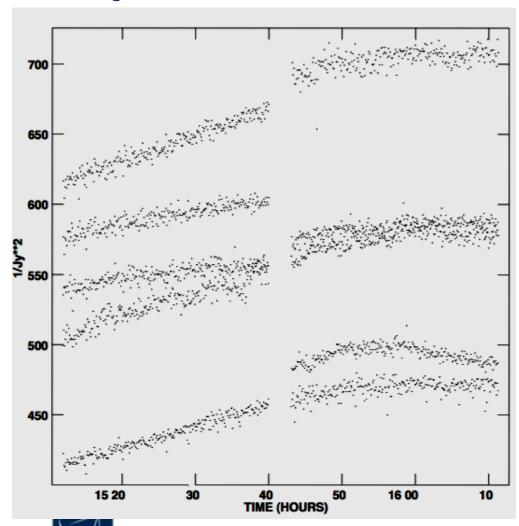
Should be relatively smooth with time

Outliers are obvious



Time in d/hh mm

Weights of antenna 4 with 5,6,7,8,9



All (u,v) data points have a **weight**. The weight depends on the antenna sensitivity, measured during the observations

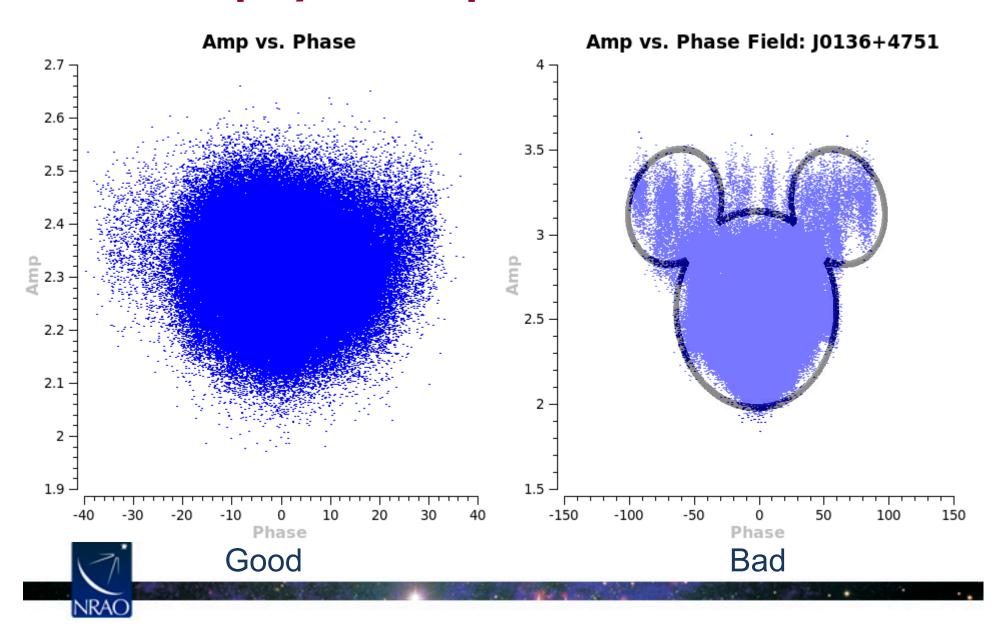
The amplitude calibration values also modify the weights.

Occasionally the weight of the points become very large, often caused by subtle software bugs.

A large discrepant weight causes the same image artifacts as a large discrepant visibility value.

Check weights to make sure they are reasonable.

# Data Displays - Amplitude vs Phase



# **Finding Errors**

### ---Obvious outlier data (u,v) points:

100 bad points in 100,000 data points gives an 0.1% image error (unless the bad data points are 1 million Jy)

LOOK at u,v DATA to find gross problem (you'd be hard pressed

to find it in the image plane other than a slight increase in noise)

#### ---Persistent small data errors:

e.g. a 5% antenna gain calibration error is difficult to see
in (u,v) data (not an obvious outlier), but will produce a
l% effect in image with specific characteristics (more later).
USE IMAGE to discover problem

#### ---Non-Data Problems:



Data ok, but algorithms chosen aren't up to the task.

## Error Recognition in the (u,v) Plane

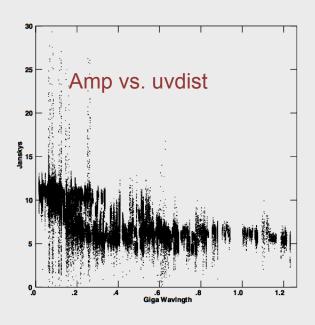
### Editing obvious errors in the (u,v) plane

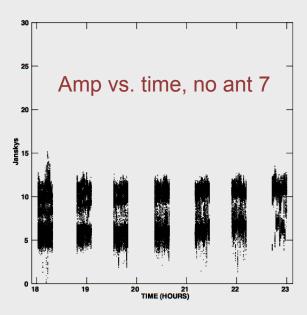
- --Mostly consistency checks assume that the visibility cannot change much over a small change in (u,v) spacing
- --Also, look at gain and phase solution tables from calibration processes. These values should be relatively stable.

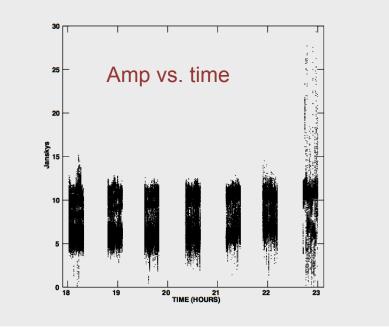
More details in the "white book" - Ekers, Lecture 15, p321



## **Visibility Amplitude Plots**







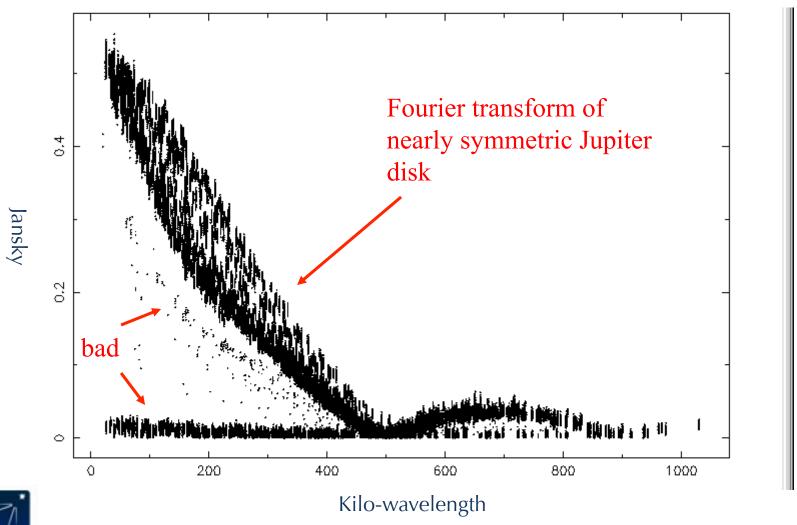
Amp vs. uvdist shows outliers

Amp vs. time shows outliers in last scan

Amp vs. time without ant 7 shows good data

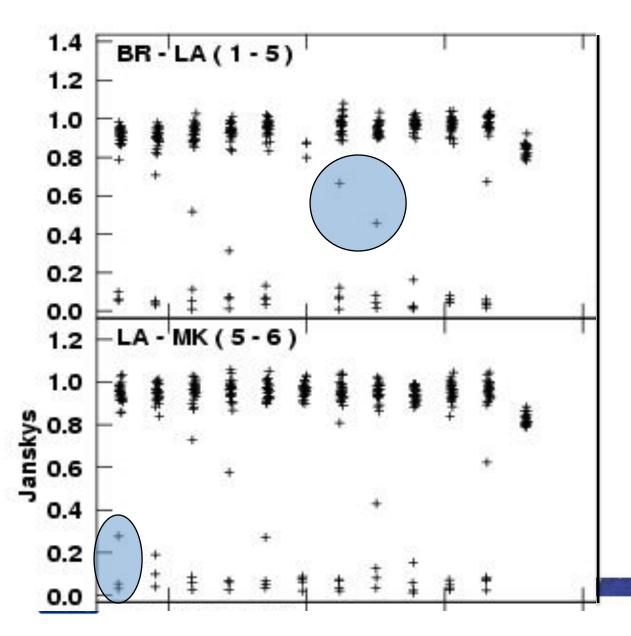
(3C279 VLBA data at 43 GHz)

# Example Edit – plotms





## Drop-outs at Scan Beginnings



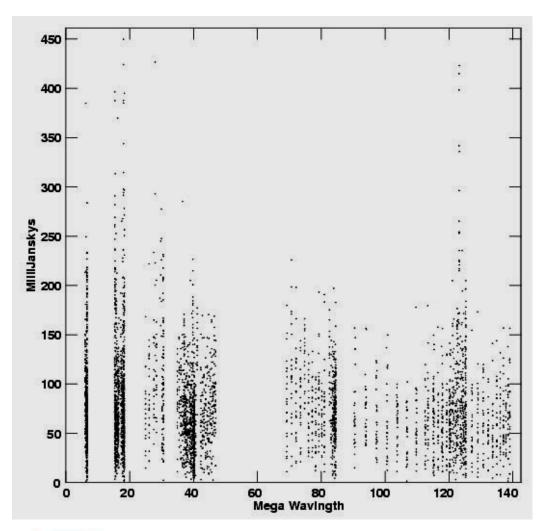
Often the first few points of a scan are low. E.g. antenna not on source.

Software can remove these points (aips,casa 'quack')

#### Flag extension:

Should flag all sources in the same manner even though you cannot see dropout for weak sources

### **Editing Noise-dominated Sources**



No source structure information is detected. Noise dominated.

All you can do is quack and remove outlier points above ~3sigma. Precise level not important as long as large outliers are removed.



# **Image Displays**

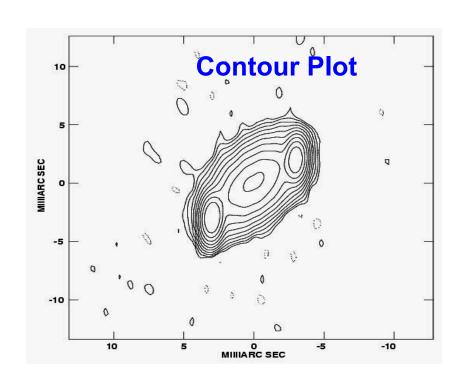
Pixel values 8 15 21 28 35 36 37 33 43126217132 28 8 15 25 34 44 54 54 53 48 61173288168 34 8 14 25 40 52 67 79 77 74 63 78199316177 34 7 14 24 40 60 77 97109102 93 74 79191289158 5 11 22 37 58 86108130137123105 79 73154220113 8 17 33 54 81116139156156133107 75 61106140 69 12 24 45 72105143162170161131 99 66 47 64 75 4 8 18 32 58 88124160171169152118 86 55 36 36 36 16 7 16 27 42 70101135162164156134100 71 44 27 15 34 43 51 77105133150146135112 81 56 34 19 34 73 70 59 79100120130122110 88 61 2 14 69141112 65 73 87100106 96 83 64 3 23121238167 69 62 69 77 81 70 58 42 3 34180338217 69 48 52 56 57 47 4 42222402242 65 36 37 38 37 29 21 4 44229398228 56 26 25 25 22 16 

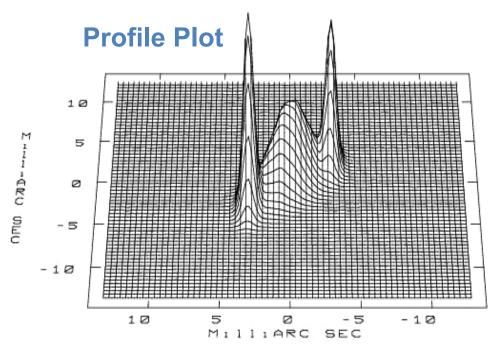
### Digital image

Numbers are proportional to the intensity

Old School

# **Image Displays**





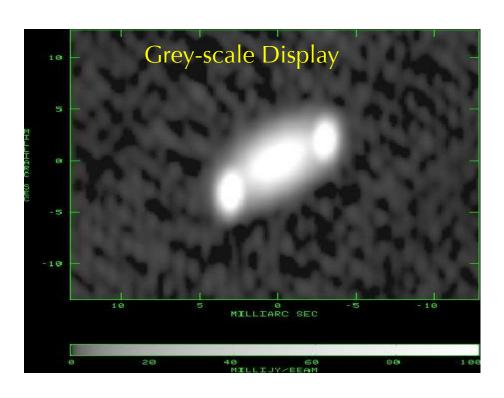
### These plots are easy to reproduce and print

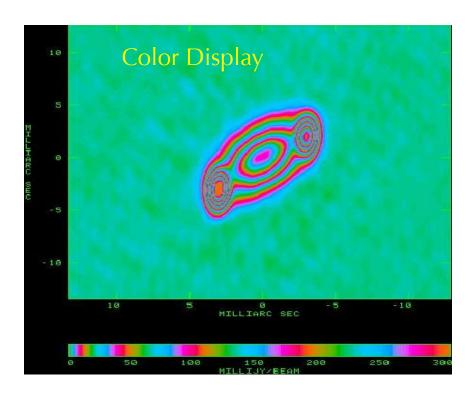
Contour plots give good representation of faint emission.

Profile plots give a good representation of the bright emission



# **Image Displays**





### TV-based displays are most useful and interactive:

Grey-scale shows faint structure, but not good for high dynamic range and somewhat unbiased view of source
Color displays more flexible; e.g. pseudo contours



### **Error Recognition in the Image Plane**

#### Some Questions to ask:

#### Noise properties of image:

Is the rms noise about that expected from integration time?
Is the rms noise much larger near bright sources?
Are there non-random noise components (faint waves and ripples

#### Funny looking structure:

Non-physical features; stripes, rings, symmetric or anti-symmetric Negative features well-below 4xrms noise Does the image have characteristics that look like the dirty beam?

#### Image-making parameters:

Is the image big enough to cover all significant emission? Is cell size too large or too small? ~4 points per beam okay Is the resolution too high to detect most of the emission?



# **Obvious Image Problems**

VLBA observations of SgrA\* at 43 GHz

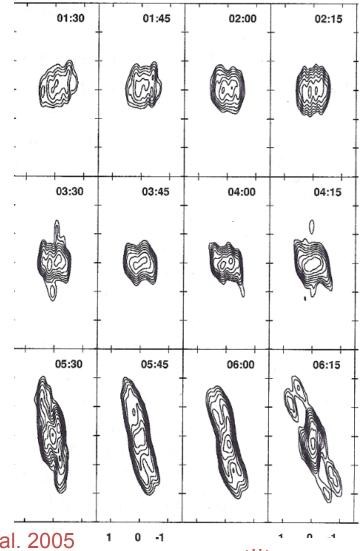
This can't be right. Either SgrA\* has bidirectional jets that nobody else has ever seen or...

Clear signs of problems:

Image rms > expected rms

Unnatural features in the image

How can the problems be found and corrected?





Miyoshi et al. 2005

milliarcsec

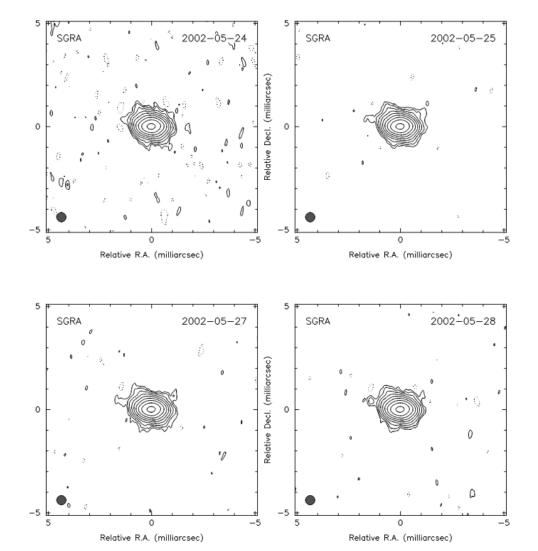
# **High Quality Images**

#### **Reality**

With care we can obtain good images.

#### What were defects?

Two antennas had ~30% calibration errors at low elevations.



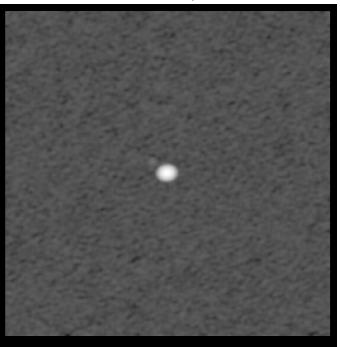


milliarcsec

## **EXAMPLE I: Data bad over a short period of time**

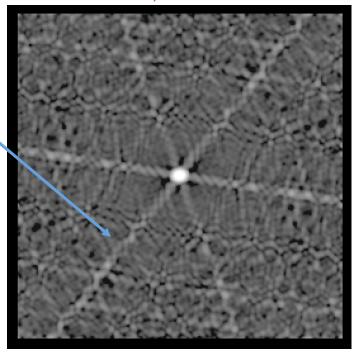
Results for a point source using VLA. 13 x 5min observation over 10 hr. Images shown after editing, calibration and deconvolution.

no errors: max 3.24 Jy rms 0.11 mJy



6-fold symmetric pattern due to VLA "Y". Image has properties of dirty beam.

10% amp error for all antennas for 1 time period rms 2.0 mJy



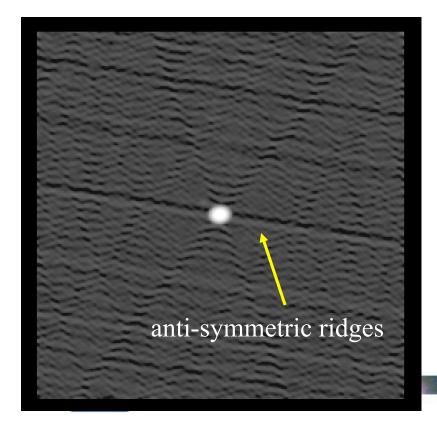


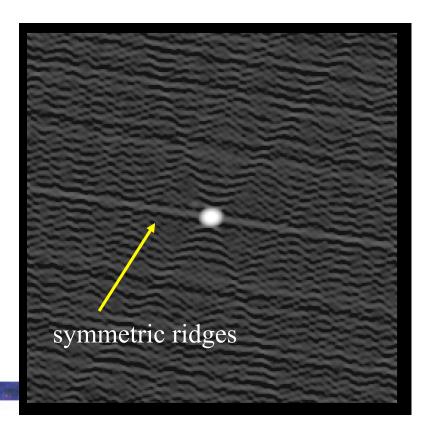
### **EXAMPLE 2: Short burst of bad data**

#### Typical effect from one bad antenna

10 deg phase error for one antenna at one time rms 0.49 mJy

20% amplitude error for one antenna at one time rms 0.56 mJy (self-

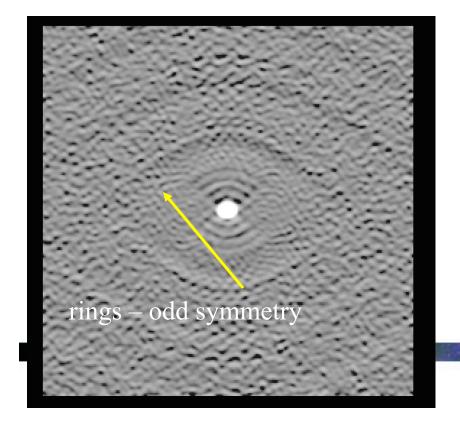


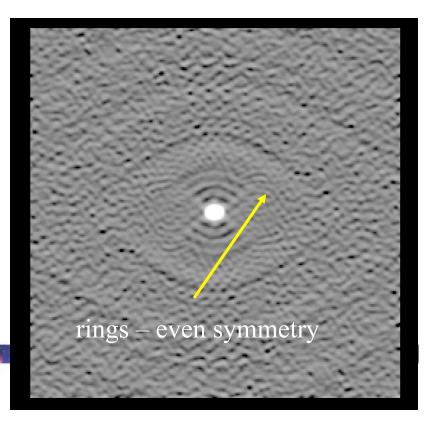


#### **EXAMPLE 3: Persistent errors over most of observations**

NOTE: 10 deg phase error to 20% amplitude error cause similar sized artifacts

10 deg phase error for one antenna all times rms 2.0 mJy 20% amp error for one antenna all times rms 2.3 mJy



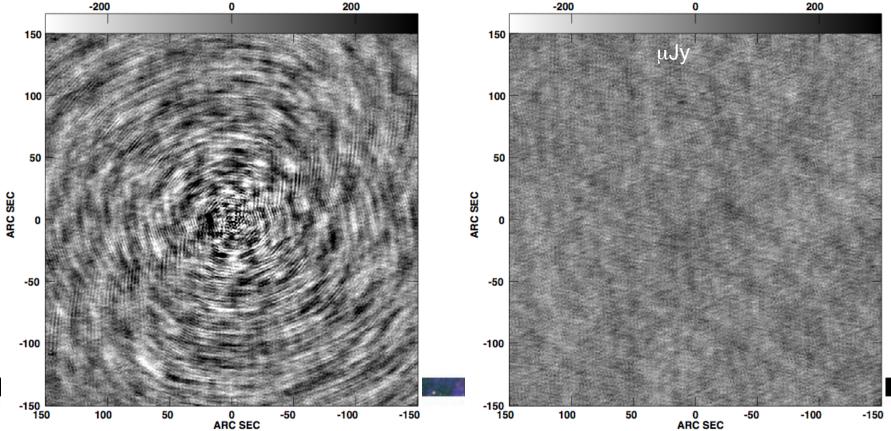


## **EXAMPLE 4: Spurious Correlator Offset Signals**

Occasionally correlators produce ghost signals or cross talk signals Occurred during change-over from VLA to EVLA system

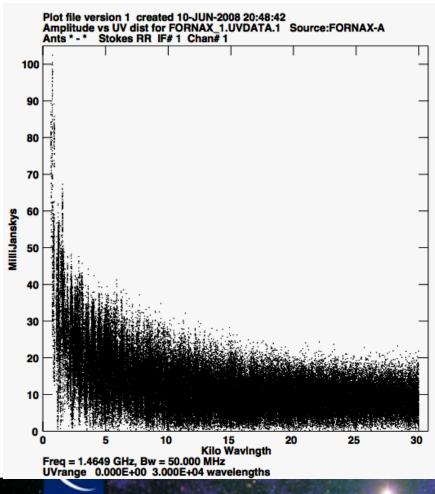
Symptom: Garbage near phase center, dribbling out into image





### **Deconvolution Errors**

Even if the data are perfect, image errors and uncertainties will occur because the (u,v) coverage is not adequate to map the source structure.

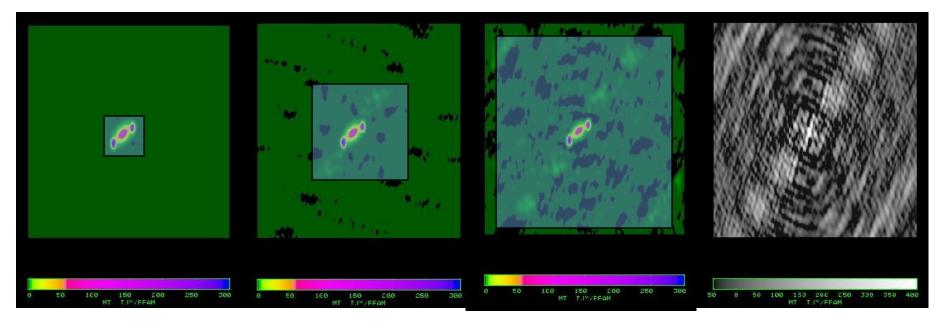


The extreme rise of visibility at the short spacings makes it impossible to image the extended structure. You are better off imaging the source with a cutoff below about 2 kilo-wavelengths

Get shorter spacing or single-dish data



# **Cleaning Window Sensitivity**



Tight Box

One small clean box

Middle Box

One clean box around all emission

Big Box

Clean entire inner map quarter

**Dirty Beam** 

Make box as small as possible to avoid cleaning noise interacting with sidelobes



## **How Deep to Clean?**

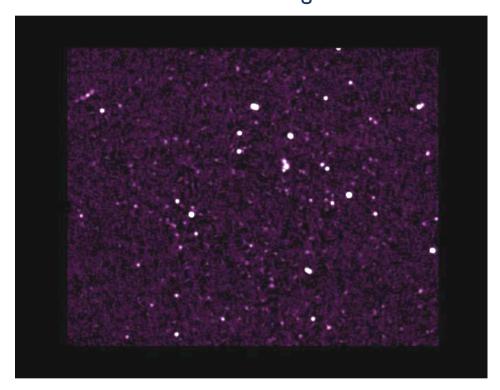
**Under-cleaned** Properly cleaned Over-cleaned Residual sidelobes Background is thermal dominate the noise **Emission from** noise-dominated; second source sits Regions within no "bowls" around atop a negative "bowl" clean boxes sources. appear "mottled"

# Improvement of Image

Removal of low level ripple improves detectability of faint sources

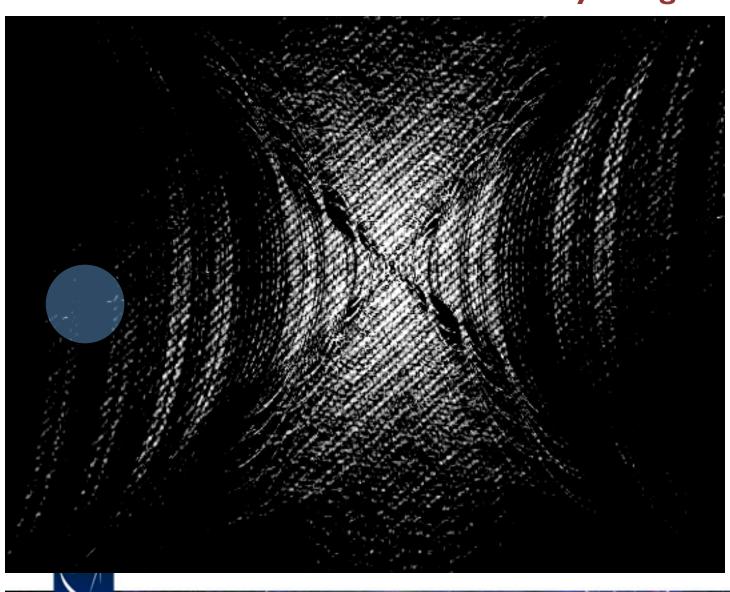
Before editing

After editing





### **Fourier Transform Dirty Image**

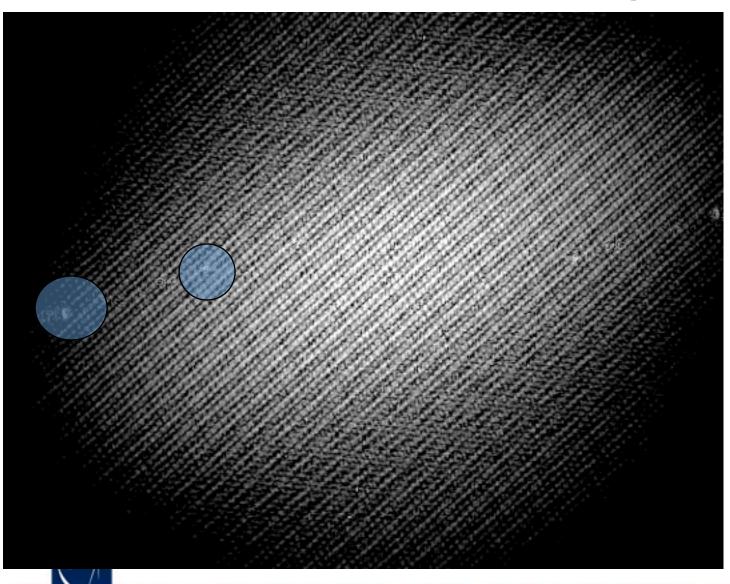


Shows the (u,v) data as gridded just before imaging

Diagonal lines caused by structure in field

A few odd points are not very noticeable

### Fourier Transform Clean Image

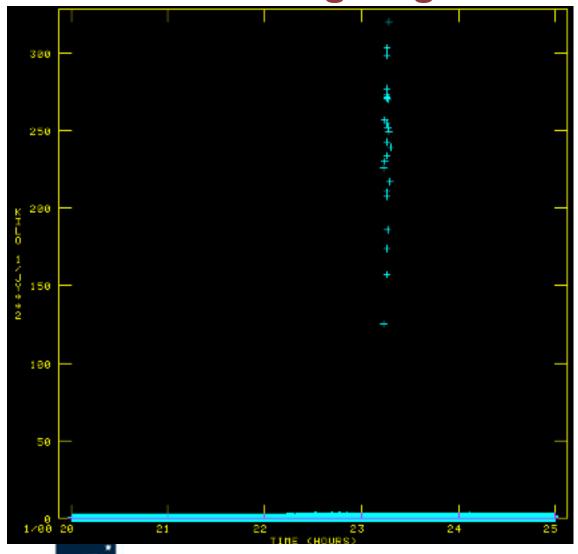


Shows the (u,v) data from clean image.

Diagonal lines still present. Notice that clean does an interpolation in the u,v plane between u,v tracks.

The odd points are smeared, but still present. These produce the low level ripples.

## Bad weighting of a few (u,v) points



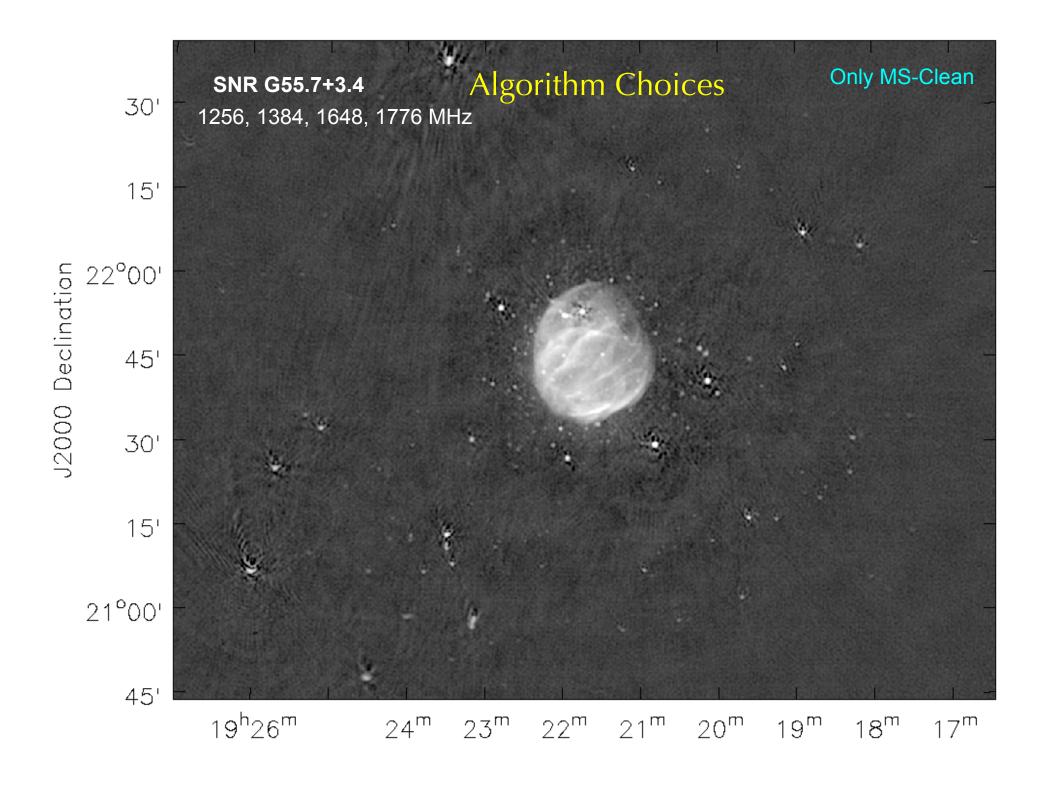
After a long search through the data, about 30 points out of 300,000 points were found to have too high of a weight by a factor of 100.

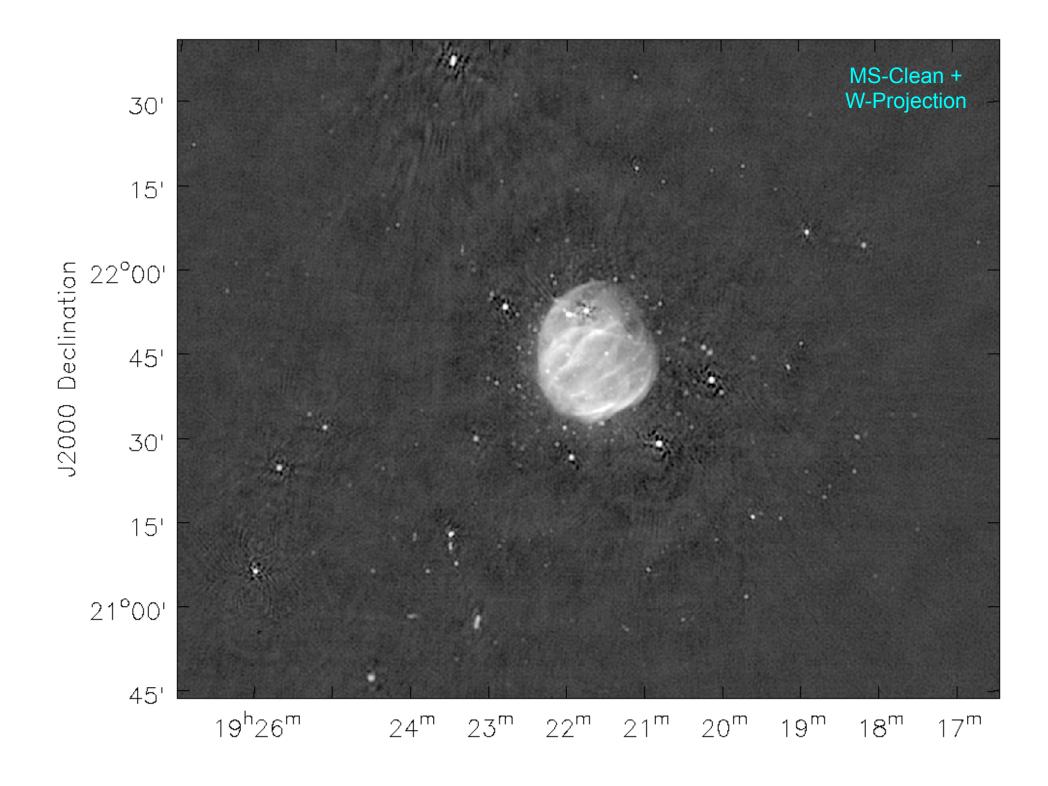
Effect is <1% in image.

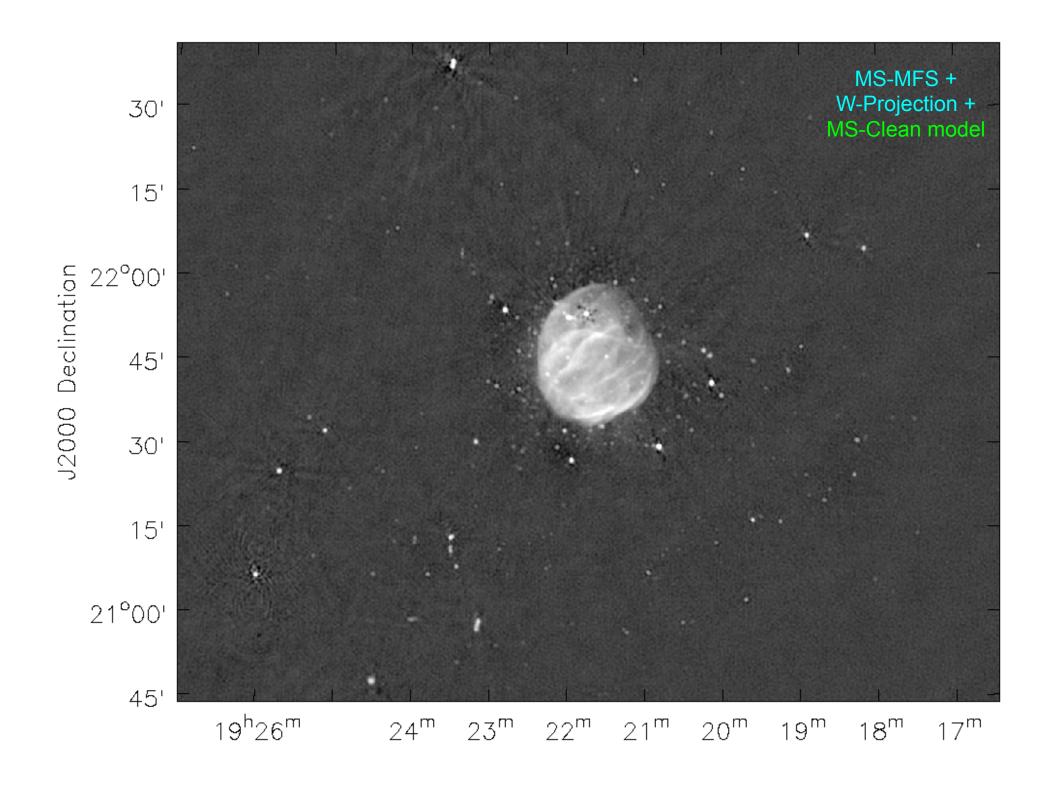
#### Cause??

Sometimes in applying calibration produced an incorrect weight in the data. Not present in the original data.

These problems can sneak up on you. Beware.







#### SUMMARY OF ERROR RECOGNITION

Source structure should be 'reasonable', the rms image noise as expected, and the background featureless. If not,

#### (u,v) data

Look for outliers in (u,v) data using several plotting methods.

Check calibration gains and phases for instabilities.

Look at residual data (u,v data - clean components)

#### **IMAGE** plane

Do defects resemble the dirty beam?

Are defect properties related to possible data errors?

Are defects related to possible deconvolution problems?

Are other corrections/calibrations needed?

Does the field-of-view encompass all emission?



#### **IMAGE ANALYSIS**

- Input: Well-calibrated datasets producing high quality images
- Output: Parameterization and interpretation of image or a set of images

This is very open-ended

Depends on source emission complexity

Depends on the scientific goals

Examples and ideas are given.

Many software packages besides CASA

(e.g.. AIPS, IDL, DS-9) are available, depending on needs.

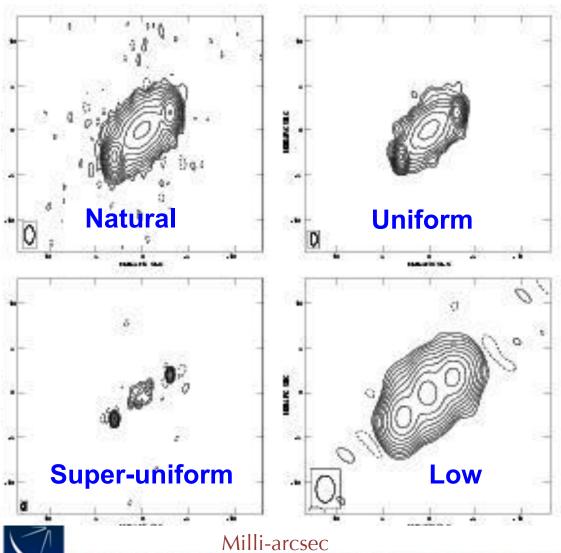


#### **IMAGE ANALYSIS OUTLINE**

- Desired Resolution of radio source.
- Parameter Estimation of Discrete Components
- Image Comparisons
- Positional Registration



# Image at Several Resolutions



Different aspect of source structure can be see at various resolutions, shown by the ellipse in the lower left corner of each box.

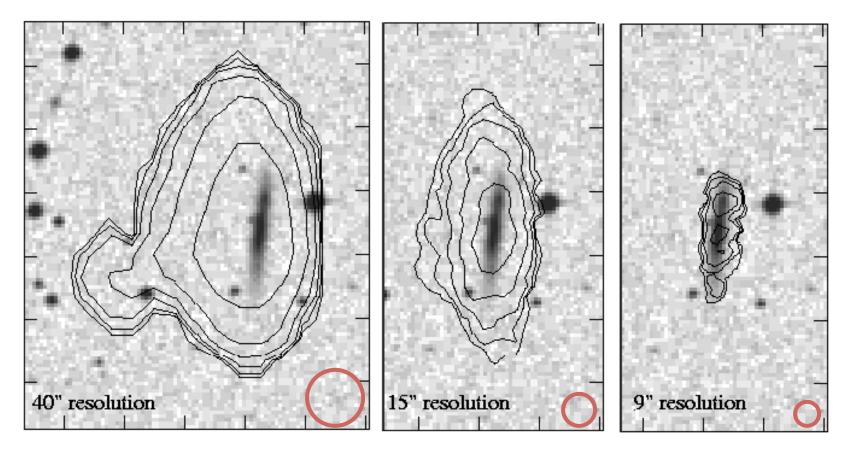
# SAME DATA USED FOR ALL IMAGES

For example,
Outer components are small from SU resolution
There is no extended emission from low resolution



### Imaging and Deconvolution of Spectral Line Data:

Type of weighting in imaging



**Smoothed** 

Robust = +1

Robust=-1



HI contours overlaid on optical images of an edge-on galaxy

#### **Parameter Estimation**

#### Parameters associated with discrete components

- Fitting in the image
  - Assume source components are Gaussian-shaped
  - Deep cleaning restores image intensity with Gaussian-beam
  - True size \* Beam size = Image size, if Gaussian-shaped. Hence, estimate of true size is relatively simple.
- Fitting in (u,v) plane (aka model-fitting)
  - Better estimates of parameters for simple sources
  - May be possible even when imaging is not
  - Can fit to more source models (e.g. Gaussian, ring, disk)
- Error estimates of parameters
  - Simple ad-hoc error estimates
  - Estimates from fitting programs
  - Monte Carlo simulations if model-fitting



#### **IMAGE FITTING**

3.6

+/-

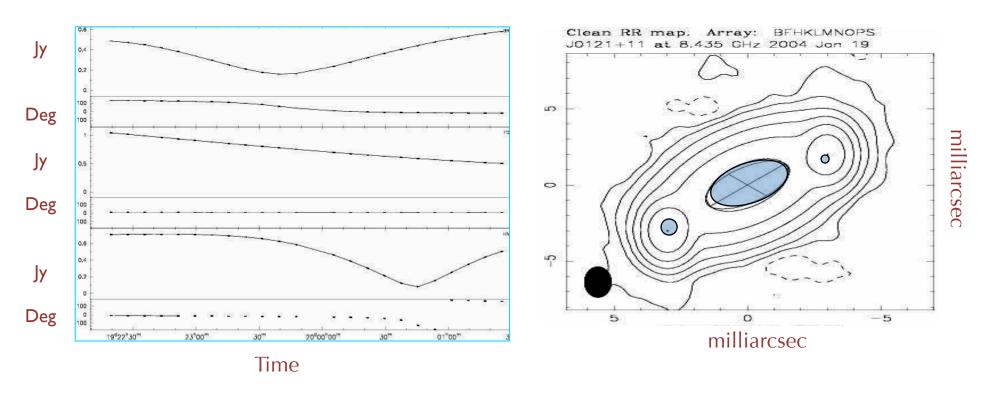
0.2

deg

Pos ang

```
Component
         2-Gaussian
 Peak intensity
                  = 0.104 +/- 0.005 JY/BEAM
 Integral intensity= 0.998 +/- 9,47 JANSKYS
 X-position
                      255.986 +/- 0.0029 pixels
                                   0.0032 pixels
 Y-position
                      257.033 +/-
                                   0.02
   Major ax
                     19.99
                              +/-
                                        pixels
   Minor ax
                      9.98
                                   0.03
                                        pixels
                              +/-
                                   0.1
                     135.3
                                         deg
   Pos ang
                                                        1-Gaussian
                                              Component
                                                Peak intensity
                                                                = 0.300 +/- 0.005 JY/BEAM
                                                Integral intensity= 0.302 +/- 0.008 JANSKYS
                                                X-position
                                                                    270.991 +/-
                                                                                 0.001 pixels
                                                Y-position
                                                                    267.018 +/-
                                                                                 0.001 pixels
                                                                    0.53
                                                                                 0.01 pixels
                                                 Major ax
                                                                    0.00
                                                 Minor ax
                                                                                 0.05 pixels
                                                 Pos ang
                                                                           +/-
                                                                                 1.1
                                                                                      deg
                                                                   21.6
                                                                     AIPS task: JMFIT
 Component
           3-Gaussian
                     = 0.393 +/- 0.004 JY/BEAM
   Peak intensity
                                                                     Casa tool: imfit
   Integral intensity= 0.403 +/- 0.008 JANSKYS
   X-position
                           241.007 +/-
                                         0.001 pixels
   Y-position
                          241.988 +/-
                                         0.001 pixels
                                        0.01 pixels
     Major ax
                           1.54
     Minor ax
                           0.21
                                   +/-
                                         0.01 pixels
```

# (u,v) DATA FITTING



DIFMAP has particular clear (u,v) fitting algorithm

Fit model directly to (u,v) data Compare model to data Contour display of image Ellipses show true component size. (SNR dependent resolution)



#### **COMPONENT ERROR ESTIMATES**

```
P = \text{Component Peak Flux Density}
\sigma = \text{Image rms noise}
P/\sigma = \text{signal/noise} = S
P/\sigma = \text{signal/noise}
```



### **Comparison and Combination of Images of Many Types**

### FORNAX-A Radio/Optical field

Radio is red Faint radio core in center of NGC1316

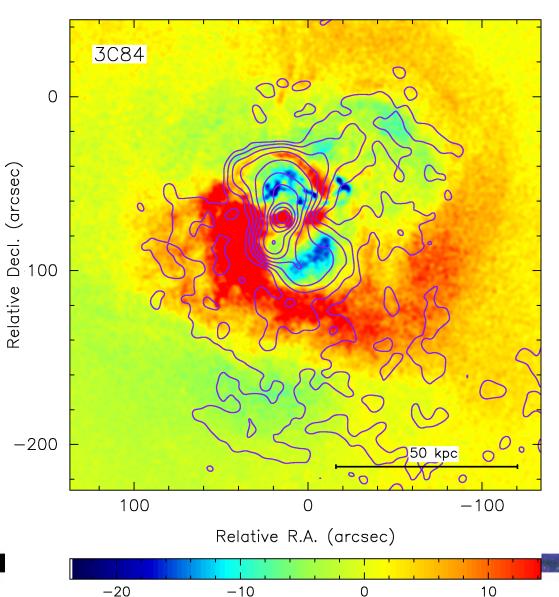
Optical in blue-white

Frame size is  $60' \times 40'$ 





## **COMPARISON OF RADIO/X-RAY IMAGES**



Contours of radio intensity at I.4 GHz

Color intensity represents X-ray intensity smoothed to radio resolution

#### IMAGE REGISTRATION AND ACCURACY

• Separation Accuracy of Components on One Image due to residual phase errors, regardless of signal/noise:

Limited to 1% of resolution

Position errors of I:10000 for wide fields, i.e. 0.1" over 1.4 GHz PB

• Images at Different Frequencies:

Multi-frequency. Use same calibrator for all frequencies.

Watch out at frequencies < 2 GHz when ionosphere can produce displacement. Minimize calibrator-target separation

• Images at Different Times (different configuration):

Use same calibrator for all observations. Daily troposphere changes can produce position changes up to 25% of the resolution.

Radio versus non-Radio Images:

Header-information of non-radio images often much less accurate than for radio. For accuracy <1", often have to align using coincident objects.

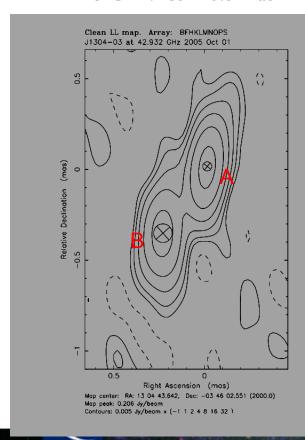
### Radio Source Alignment at Different Frequencies

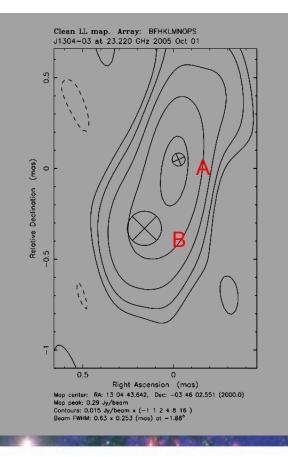
Self-calibration at each frequency aligns maximum at (0,0) point Frequency-dependent structure causes relative position of maximum to change Fitting of image with components can often lead to proper registration

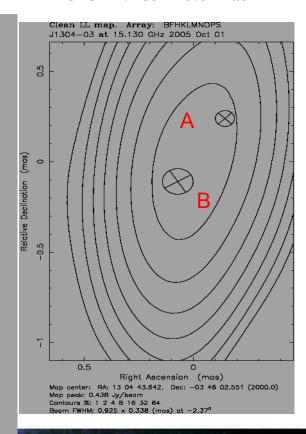
43 GHz: res = 0.3 mas

23 GHz: res = 0.6 mas

15 GHz: res = 0.8 mas









### **ANALYSIS: Summary**

- Analyze and display data in several ways
  - Adjust resolution to illuminate desired interpretation, analysis
- Parameter fitting useful, but be careful of error estimates
  - Fitting in u,v plane and/or image plane
- Registration of a field at different frequencies or wave-bands can be subtle.
  - Whenever possible use the same calibrator
  - May be able to align using 'known' counterparts



# **Further Reading**

- http://www.nrao.edu/whatisra/
- www.nrao.edu
- 2010 Lecture on Non-Imaging Analysis
- Synthesis Imaging in Radio Astronomy, ASP Vol 180, eds Taylor, Carilli & Perley
- Numerical Recipes, Press et al. 1992

