Non-Imaging Data Analysis

Tim Pearson

Outline

Introduction
Inspecting visibility data
Model fitting
Some applications
- Superluminal motion
- Gravitational lenses
- The Sunyaev-Zeldovich effect
- The cosmic microwave background radiation

Introduction

Reasons for analyzing visibility data
- Insufficient (u,v)-plane coverage to make an image
- Inadequate calibration
- Quantitative analysis
- Direct comparison of two data sets
- Error estimation
  - Usually, visibility measurements are independent gaussian variates
  - Systematic errors are usually localized in the (u,v) plane

Statistical estimation of source parameters

Inspecting Visibility Data

Visibility Data

Fourier imaging

$$V(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(l,m)\, I(l,m) \exp[-2\pi i(ul+vm)]\, dl\, dm$$

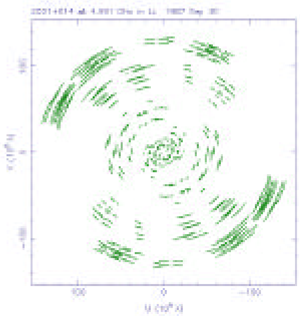
Problems with direct inversion
- Sampling
  - Poor (u,v) coverage
- Missing data
  - e.g., no phases (speckle imaging)
- Calibration
  - Closure quantities are independent of calibration
- Non-Fourier imaging
  - e.g., wide-field imaging; time-variable sources (SS433)
- Noise
  - Noise is uncorrelated in the (u,v) plane but correlated in the image

Inspecting visibility data

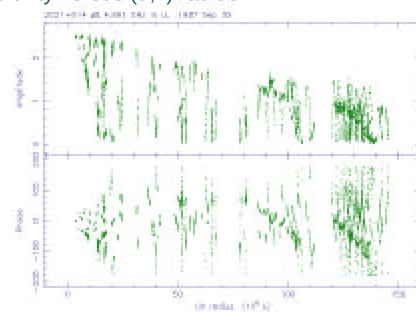
Useful displays
- Sampling of the (u,v) plane
- Amplitude and phase *v*. radius in the (u,v) plane
- Amplitude and phase *v*. time on each baseline
- Amplitude variation across the (u,v) plane
- Projection onto a particular orientation in the (u,v) plane

Example: 2021+614
- GHz-peaked spectrum radio galaxy at z=0.23
- A VLBI dataset with 11 antennas

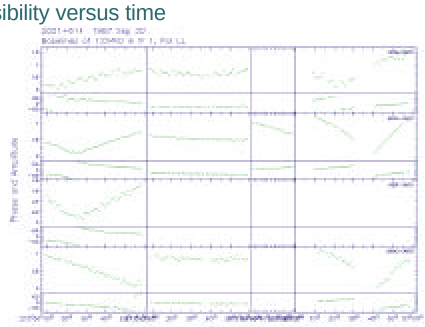
Sampling of the (u,v) plane



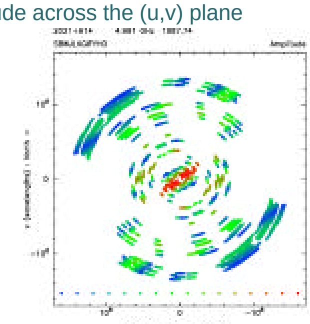
Visibility versus (u,v) radius



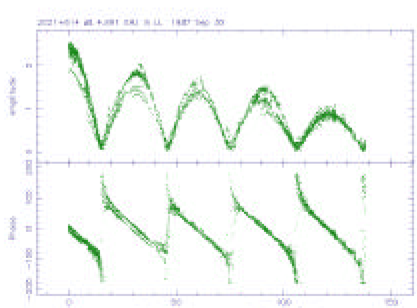
Visibility versus time



Amplitude across the (u,v) plane



Projection in the (u,v) plane



Properties of the Fourier transform

See, e.g., R. Bracewell, *The Fourier Transform and its Applications* (1965).

Fourier Transform theorems

Linearity

Visibilities of components add (complex)

Convolution

Shift

Shifting the source creates a phase gradient across the (u,v) plane

Similarity

Larger sources have more compact transforms

Fourier Transform theorems

$$F(u,v) = FT\{f(x,y)\}$$

i.e.,

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\exp[2\pi i(ux+vy)]\,dx\,dy$$

Linearity

$$FT\{f(x,y)+g(x,y)\} = F(u,v)+G(u,v)$$

Convolution

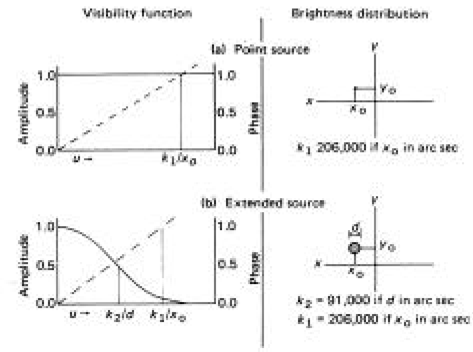
$$FT\{f(x,y)*g(x,y)\} = F(u,v)\cdot G(u,v)$$

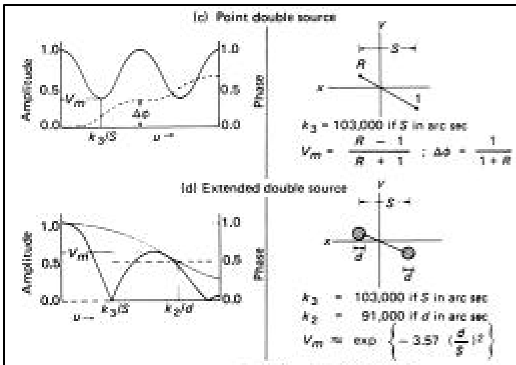
Shift

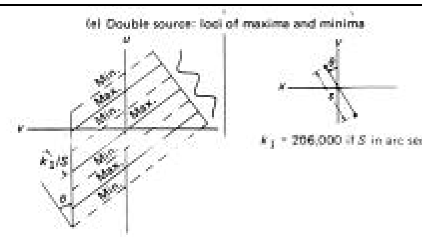
$$FT\{f(x-x_1,y-y_1)\} = F(u,v)\exp[2\pi i(ux_1+vy_1)]$$

Similarity

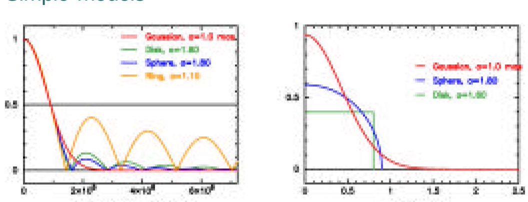
$$FT\{f(ax,by)\} = \frac{1}{|ab|}F\left(\frac{u}{a},\frac{v}{b}\right)$$







Simple models



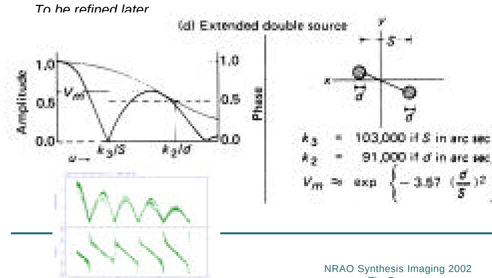
Visibility at short baselines contains little information about the profile of the source.

Trial model

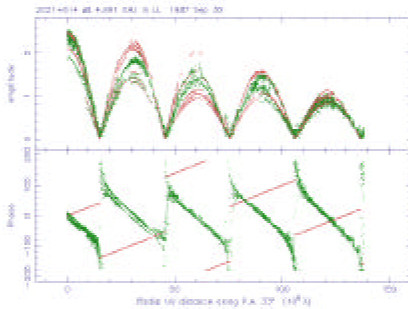
By inspection, we can derive a simple model:

Two equal components, each 1.25 Jy, separated by about 6.8 milliarcsec in p.a. 33º, each about 0.8 milliarcsec in diameter (gaussian FWHM)

To be refined later



Projection in the (u,v) plane



Closure Phase and Amplitude

The closure quantities

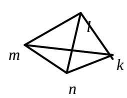
Antenna-based gain errors

$$V_{kl} \equiv |V_{kl}| \exp(i\phi_{kl}) = g_k g_l V_{kl}^{\rm true} \exp(i\phi_k) \exp(-i\phi_l)$$

Closure phase (bispectrum phase)

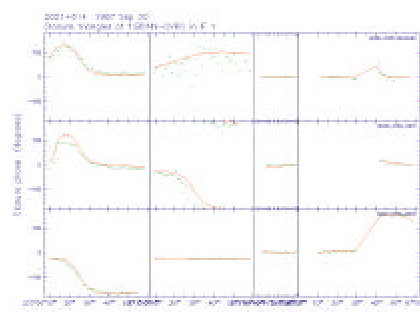
$$\Psi_{lmn}(t) = \phi_{lm}(t) + \phi_{mn}(t) + \phi_{nl}(t)$$

Closure amplitude $\dfrac{|V_{kl}| \cdot |V_{mn}|}{|V_{km}| \cdot |V_{ln}|}$



- Closure phase and closure amplitude are unaffected by antenna gain errors
- They are conserved during self-calibration
- Contain (*N*–2)/*N* of phase, (*N*–3)/(*N*–1) of amplitude info
- – Many non-independent quantities
- – They do not have gaussian errors
- – No position or flux info

Closure phase



Model Fitting

Model fitting

Imaging as an Inverse Problem

- In synthesis imaging, we can solve the **forward problem**: given a sky brightness distribution, and knowing the characteristics of the instrument, we can predict the measurements (visibilities), within the limitations imposed by the noise.
- The **inverse problem** is much harder, given limited data and noise: the solution is rarely unique.
- A general approach to inverse problems is **model fitting**. See, e.g., Press et al., *Numerical Recipes*.
  1. Design a model defined by a number of adjustable parameters.
  2. Solve the forward problem to predict the measurements.
  3. Choose **a figure-of-merit** function, e.g., rms deviation between model predictions and measurements.
  4. Adjust the parameters to **minimize the merit function**.
- Goals:
  1. Best-fit values for the parameters.
  2. A measure of the goodness-of-fit of the optimized model.
  3. Estimates of the uncertainty of the best-fit parameters.

Model fitting

Maximum Likelihood and Least Squares

- The model:

$$V(u,v) = F(u,v;a_1,\ldots,a_M) + \text{noise}$$

- The likelihood of the model (if noise is gaussian):

$$L \propto \prod_{i=1}^{N} \left\{ \exp\left[-\frac{1}{2} \left(\frac{V_i - F(u_i,v_i;a_1,\ldots,a_M)}{\sigma_i} \right)^2 \right] \right\}$$

- Maximizing the likelihood is equivalent to minimizing chi-square (for gaussian errors):

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{V_i - F(u_i,v_i;a_1,\ldots,a_M)}{\sigma_i} \right)^2$$

- Follows chi-square distribution with *N* – *M* degrees of freedom. Reduced chi-square has expected value 1.

Uses of model fitting

Model fitting is most useful when the brightness distribution is simple.

- Checking amplitude calibration
- Starting point for self-calibration
- Estimating parameters of the model (with error estimates)
- In conjunction with CLEAN or MEM
- In astrometry and geodesy

Programs

- AIPS UVFIT
- Difmap (Martin Shepherd)

Parameters

Example

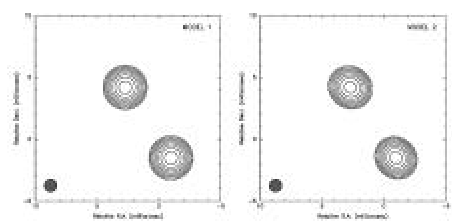
- Component position: (*x*,*y*) or polar coordinates
- Flux density
- Angular size (e.g., FWHM)
- Axial ratio and orientation (position angle)
    - *For a non-circular component*

6 parameters per component, plus a “shape”

This is a conventional choice: other choices of parameters may be better!
(Wavelets; shapelets* [Hermite functions])

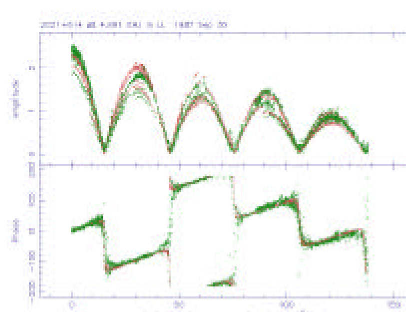
** Chang & Refregier 2002, ApJ, 570, 447*

Practical model fitting: 2021

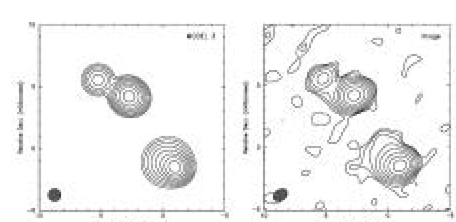


! Flux (Jy)	Radius (mas)	Theta (deg)	Major (mas)	Axial ratio	Phi (deg)	T
1.15566	4.99484	32.9118	0.867594	0.803463	54.4823	1
1.16520	1.79539	-147.037	0.825078	0.742822	45.2283	1

2021: model 2

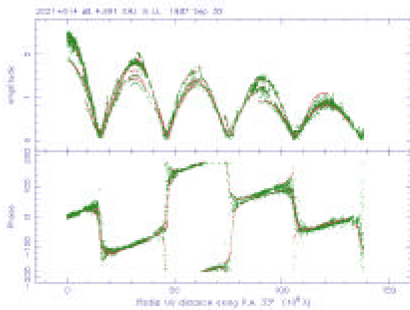


Model fitting 2021



! Flux (Jy)	Radius (mas)	Theta (deg)	Major (mas)	Axial ratio	Phi (deg)	T
1.10808	5.01177	32.9772	0.871643	0.790796	60.4327	1
0.823118	1.80865	-146.615	0.589278	0.585766	53.1916	1
0.131209	7.62679	43.3576	0.741253	0.933106	-82.4635	1
0.419373	1.18399	-160.136	1.62101	0.951732	84.9951	1

2021: model 3



Limitations of least squares

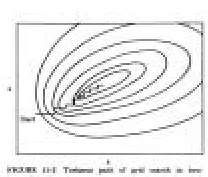
Assumptions that may be violated

- The model is a good representation of the data
  - *Check the fit*
- The errors are gaussian
  - *True for real and imaginary parts of visibility*
  - *Not true for amplitudes and phases (except at high SNR)*
- The variance of the errors is known
  - *Estimate from T_{sys}, rms, etc.*
- There are no systematic errors
  - *Calibration errors, baseline offsets, etc. must be removed before or during fitting*
- The errors are uncorrelated
  - *Not true for closure quantities*
  - *Can be handled with full covariance matrix*

Least-squares algorithms

At the minimum, the derivatives of chi-square with respect to the parameters are zero

$$\nabla\chi^2 = \frac{\partial\chi^2}{\partial a_k} = 0$$



Linear case: matrix inversion.

Exhaustive search: prohibitive with many parameters (~ 10^M)

Grid search: adjust each parameter by a small increment and step down hill in search for minimum.

Gradient search: follow downward gradient toward minimum, using numerical or analytic derivatives. Adjust step size according to second derivative

$$\nabla^2\chi^2 = \frac{\partial^2\chi^2}{\partial a_k \partial a_l}$$

For details, see *Numerical Recipes*.

Problems with least squares

Global versus local minimum

Slow convergence: poorly constrained model

Do not allow poorly-constrained parameters to vary

Constraints and prior information

Boundaries in parameter space

Transformation of variables

Choosing the right number of parameters: does adding a parameter significantly improve the fit?

Likelihood ratio or *F* test: use caution

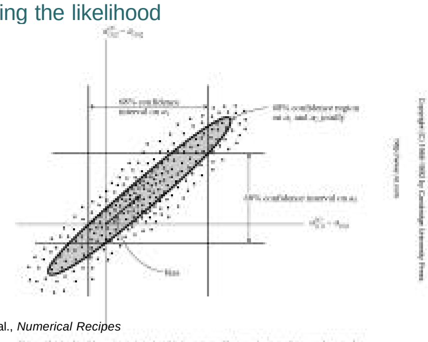
Protassov et al. 2002, ApJ, 571, 545

Monte Carlo methods

Error estimation

- Find a region of the *M*-dimensional parameter space around the best fit point in which there is, say, a 68% or 95% chance that the true parameter values lie.
- Constant chi-square boundary: select the region in which
  $$\chi^2 < \chi^2_{\min} + \Delta\chi^2$$
- The appropriate contour depends on the required confidence level and the number of parameters estimated.
- Approximate method: Fisher matrix.
  $$V_{ij} = \mathrm{cov}[\hat{a}_i, \hat{a}_j] \qquad (\hat{V}^{-1})_{ij} = -\frac{\partial^2 \ln L}{\partial a_i \partial a_j} \quad (a = \hat{a})$$
- Monte Carlo methods (simulated or mock data): relatively easy with fast computers
- Some parameters are strongly correlated, e.g., flux density and size of a gaussian component with limited (u,v) coverage.
- Confidence intervals for a single parameter must take into account variations in the other parameters ("marginalization").

Mapping the likelihood



Press et al., *Numerical Recipes*

Applications

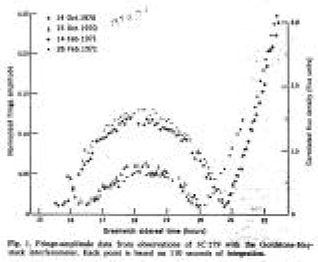
Application: Superluminal motion

Problem: to detect changes in component positions between observations and measure their speeds

- Direct comparison of images is *bad*: different (u,v) coverage, uncertain calibration, insufficient resolution
- Visibility analysis is a *good* method of detecting and measuring changes in a source: allows "controlled super-resolution"
- Calibration uncertainty can be avoided by looking at the closure quantities: have they changed?
- Problem of differing (u,v) coverage: compare the same (u,v) points whenever possible
- Model fitting as an interpolation method

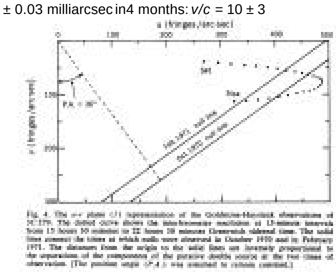
Superluminal motion

Example 1: Discovery of superluminal motion in 3C279 (Whitney et al., Science, 1971)

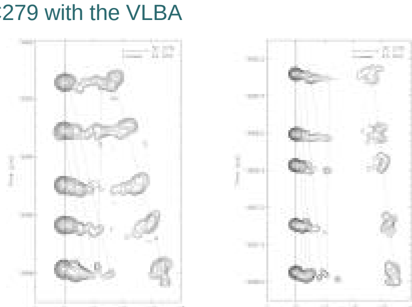


Superluminal motion

1.55 ± 0.03 milliarcsec in 4 months: v/c = 10 ± 3



3C279 with the VLBA



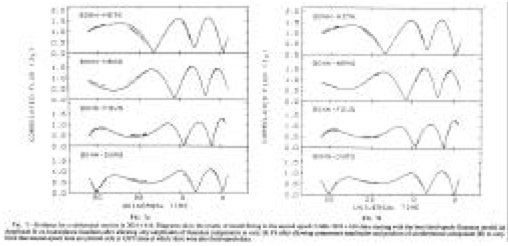
Wehrle at al. 2001, ApJS, 133, 297

Expanding sources

Example 2: changes in the radio galaxy 2021+614 between 1982 and 1987 (Conway et al. 1994, ApJ)

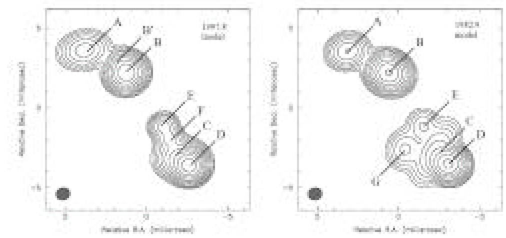
- Requires careful cross-calibration using a model: what changes to the model from one epoch are needed to fit the data from the other epoch, allowing for calibration errors and different (u,v) coverage?
- Closure phase shows *something* has changed.
- By careful combination of model-fitting and self-calibration, Conway et al. determined that the separation had changed by 69 ± 10 microarcsec between 1982 and 1987, for v/c = 0.13 h^{-1}

Expanding sources



Conway *et al.* 1994, ApJ

Expanding Sources

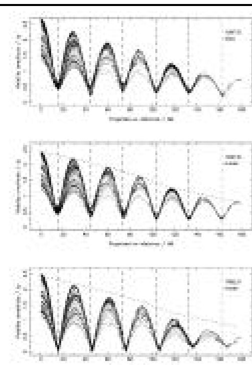


Tschager *et al.* 2000, A&A

Expanding Sources

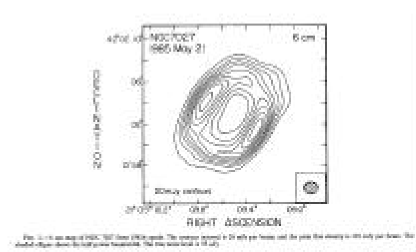
$v/c = 0.12 \pm 0.02\ h^{-1}$

Tschager *et al.* 2000, A&A

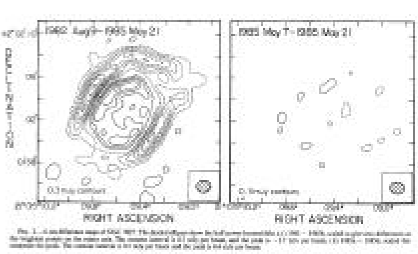


Expansion of Planetary Nebulae

Distance measurement (Masson, 1986, ApJ)



NGC7027



Application: Gravitational Lenses

Gravitational Lenses

- Single source, multiple images formed by intervening galaxy.
- Can be used to map mass distribution in lens.
- Can be used to measure distance of lens and H_0: need redshift of lens and background source, model of mass distribution, and a **time delay**.

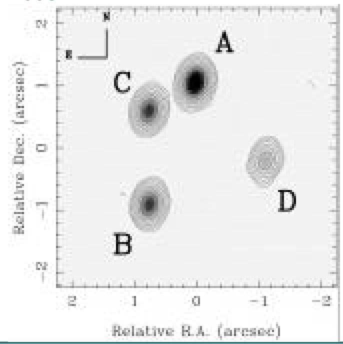
Application of model fitting

- Lens monitoring to measure flux densities of components as a function of time.
- Small number of components, usually point sources.
- Need error estimates.

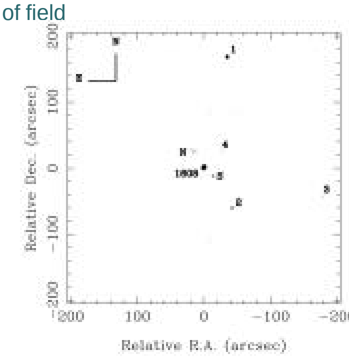
Example: VLA monitoring of B1608+656 (Fassnacht et al. 1999, ApJ)

- VLA configuration changes: different HA on each day
- Other sources in the field

VLA image of 1608

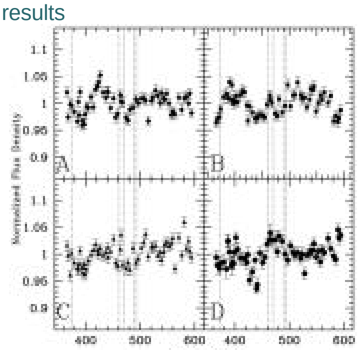


VLA image of field



1608 monitoring results

B – A = 31 days
B – C = 36 days
H_0 = 59 ± 8 km/s/Mpc



Application: Sunyaev-Zeldovich effect

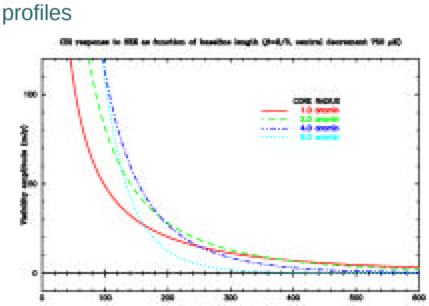
The Sunyaev-Zeldovich effect

- Photons of the CMB are scattered to higher frequencies by hot electrons in galaxy clusters, causing a negative brightness **decrement**.
- Decrement is proportional to integral of electron pressure through the cluster, or electron density if cluster is isothermal.
- Electron density and temperature can be estimated from X-ray observations, so the linear scale of the cluster is determined.
- This can be used to measure the cluster distance and H_0.

Application of model fitting

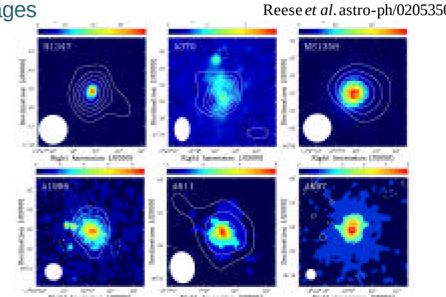
- The profile of the decrement can be estimated from X-ray observations (beta model).
- The Fourier transform of this profile increases exponentially as the interferometer baseline decreases.
- The central decrement in a synthesis image is thus highly dependent on the (u,v) coverage.
- Model fitting is the best way to estimate the true central decrement.

SZ profiles



SZ images

Reese *et al*. astro-ph/0205350



Application: Cosmic Background Radiation

The Cosmic Background Radiation

- The CMB shows fluctuations in intensity at a level of a few μK on scales from a few minutes of arc to degrees. Short-baseline interferometers can detect these fluctuations.
- Inflation models predict that CMB intensity is a **gaussian random process** with a **power spectrum** that is very sensitive to cosmological parameters.
- The power spectrum is the expectation of the square of the Fourier transform of the sky intensity distribution: i.e., closely related to the **square of the visibility** VV^*.

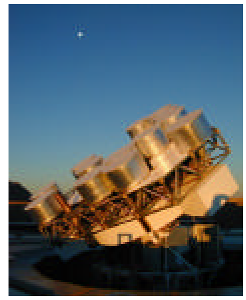
CMB interferometers

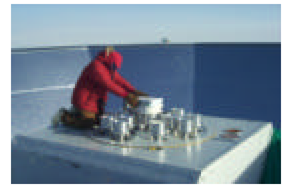
- CAT, DASI, CBI, VSA

Primary beam

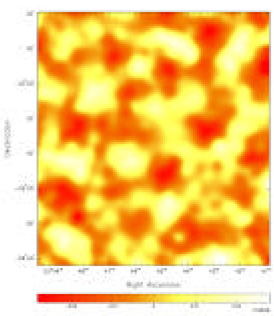
- The observed sky is multiplied by the primary beam, corresponding to convolution (smoothing) in the (u,v) plane: so the interferometer measures a smoothed version of the power spectrum.
- Smoothing reduced by mosaicing.

CBI and DASI





CBI mosaic image



An interferometer image is dominated by the effects of (u,v) sampling.

The image is only sensitive to spatial frequencies actually sampled.

Optimum analysis is most straightforward in the (u,v) plane.

CMB analysis

- A **parameter estimation** problem. The parameters are
  - Either: band powers
  - Or: cosmological parameters

$$\Omega, \Omega_{\rm cdm}, \Omega_b, \Lambda, H_0, n_s, C_{10}$$

- This can be approached as a **Maximum Likelihood** problem: compute the probability of obtaining the data, given the parameters, and maximize wrt the parameters.
- As the noise is gaussian and uncorrelated from sample to sample, this is best approached in the visibility domain.

CMB analysis

Likelihood

$$L = \frac{1}{\pi^n \det \mathbf{C}} \exp\left(-V^\dagger \mathbf{C}^{-1} V\right)$$

Covariance matrix

$$\mathbf{C} = \langle V^\dagger V \rangle = \mathbf{C}_N + \mathbf{C}_{\rm sky}$$

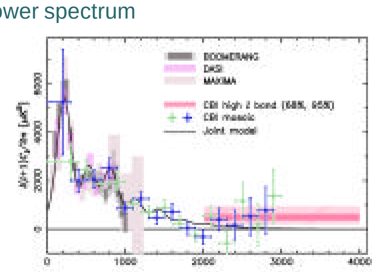
Noise (diagonal)

$$(\mathbf{C}_{\rm sky})_{kk'} \propto \int\!\!\int d^2\mathbf{v}\, C(v) \tilde{A}(\mathbf{v}_k - \mathbf{v}) \tilde{A}^*(\mathbf{v}_{k'} - \mathbf{v})$$

FT of primary beam

Power spectrum

CBI power spectrum



3 fields, each 42 pointings, 78 baselines, 10 frequency channels: ~ 600,000 measurements. Covariance matrix 5000 x 5000.
astro-ph/0205384–0205388

Summary

- For simple sources observed with high SNR, much can be learned about the source (and observational errors) by inspection of the visibilities.
- Even if the data cannot be calibrated, the **closure quantities** are good observables, but they can be difficult to interpret.
- Quantitative data analysis is best regarded as an exercise in **statistical inference**, for which the maximum likelihood method is a general approach.
- For gaussian errors, the ML method is the **method of least squares**.
- Visibility data (usually) have uncorrelated gaussian errors, so analysis is most straightforward in the (*u,v*) plane.
- Consider visibility analysis when you want a quantitative answer (with error estimates) to a simple question about a source.
- Visibility analysis is inappropriate for large problems (many data points, many parameters, correlated errors); standard imaging methods can be much faster.

[redacted]