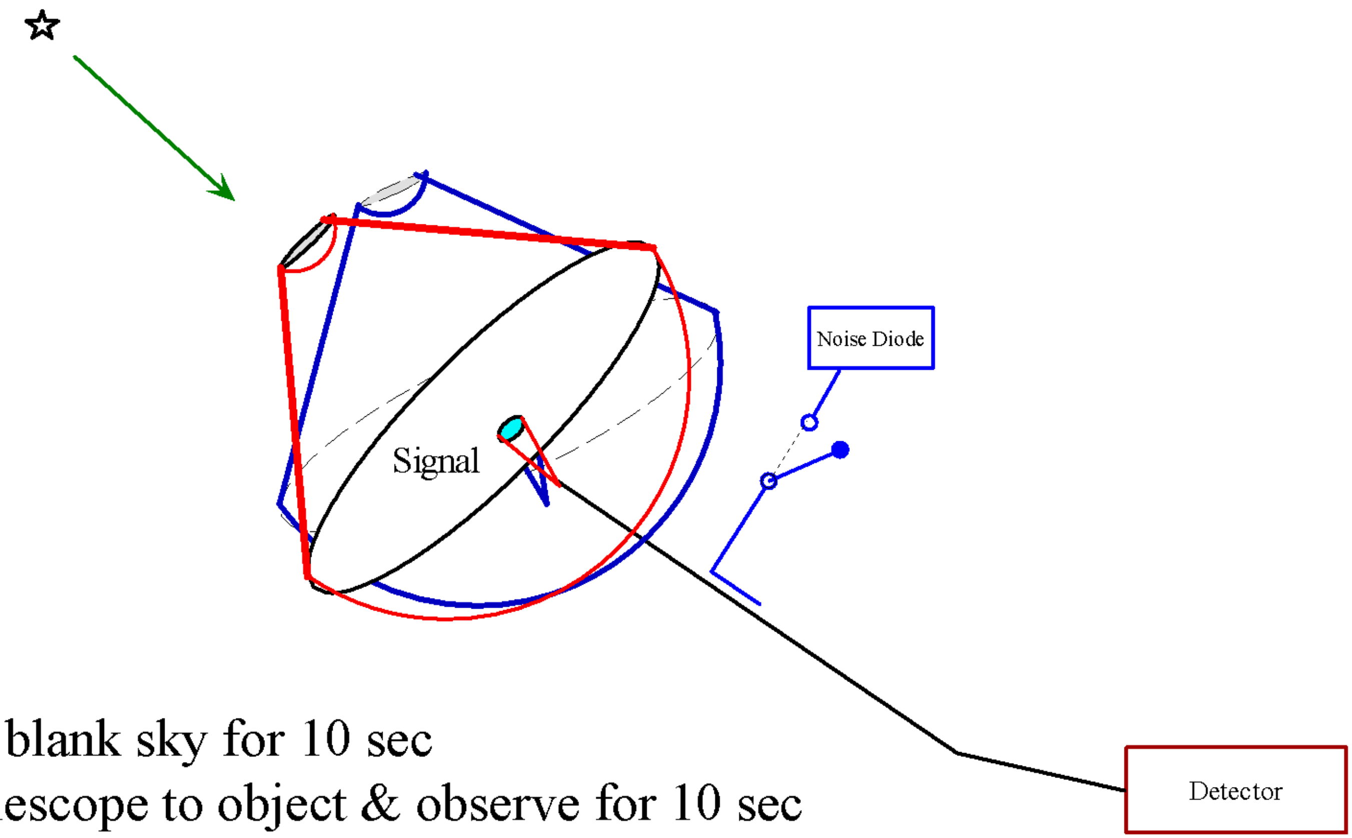
Data Reduction and Analysis Techniques

Ronald J. Maddalena

www.nrao.edu/~rmaddale/Education

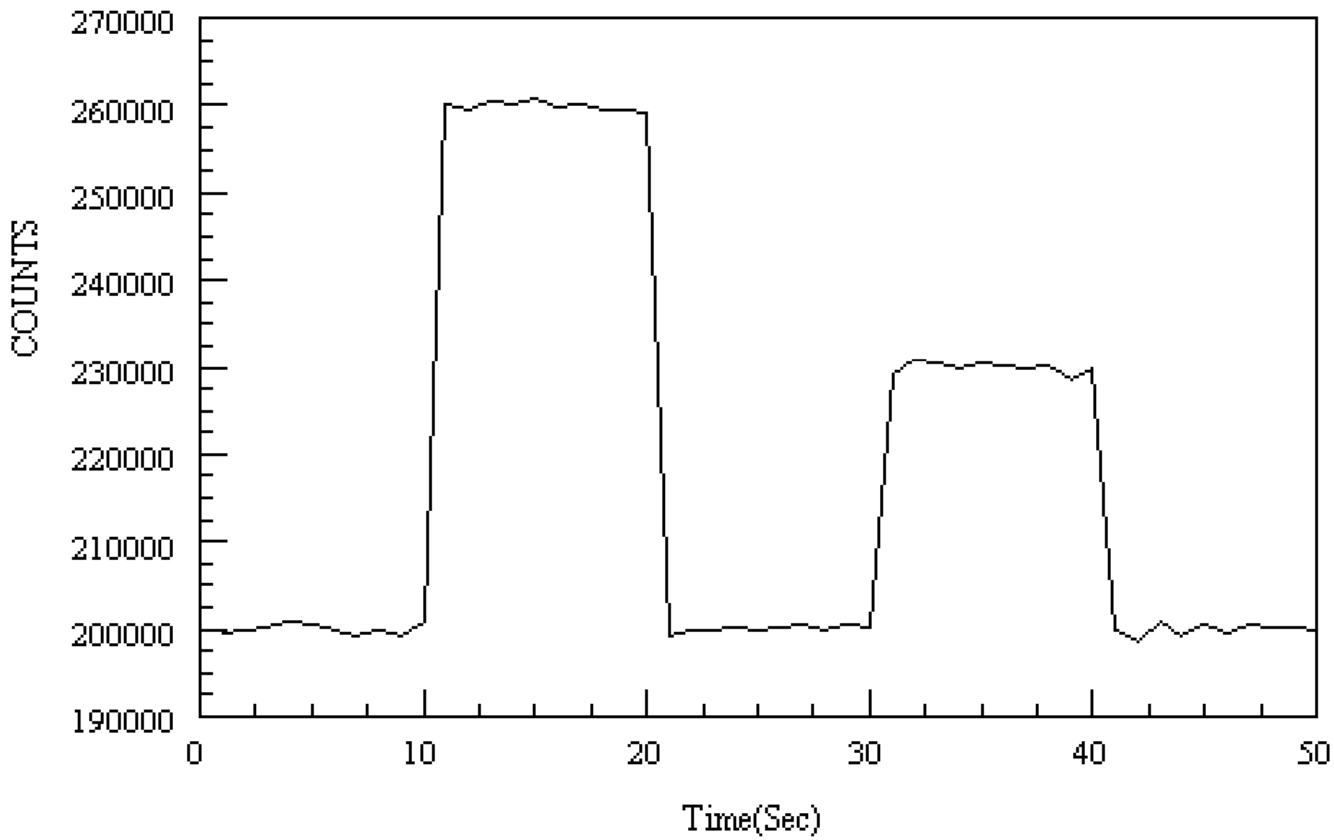
Continuum - Point Sources

On-Off Observing



- Observe blank sky for 10 sec
- Move telescope to object & observe for 10 sec
- Move to blank sky & observe for 10 sec
- Fire noise diode & observe for 10 sec
- Observe blank sky for 10 sec

Continuum - Point Sources
On-Off Observing



Continuum - Point Sources

On-Off Observing

- Known:
  - Equivalent temperature of noise diode or calibrator (T_{cal}) = 3 K
  - Bandwidth (Δv) = 10 MHz
  - Gain = 2 K / Jy
- Desired:
  - Antenna temperature of the source (T_A)
  - Flux density (S) of the source.
  - System Temperature(T_s) when OFF the source
  - Accuracy of antenna temperature (σ_{TA})

Continuum - Point Sources

On-Off Observing

$$T_S^{reference} = \frac{T_{cal} \cdot P_{cal\_off}^{reference}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}}$$

20 K

$$T_S^{signal} = \frac{T_{cal} \cdot P_{cal\_off}^{signal}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}}$$

26 K

$$T_A = T_S^{signal} - T_S^{reference} =$$

$$\frac{T_{cal}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}} \cdot \left(P_{cal\_off}^{signal} - P_{cal\_off}^{reference}\right)$$

6 K

$$\sigma_{T_A} = T_s \big/ \sqrt{\Delta\nu \cdot t}$$

0.002 K

SNR = 3000

Continuum - Point Sources

On-Off Observing –noise estimate

1. Write down data analysis equation:

$$T_A = \frac{T_{cal}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}} \cdot \left(P_{cal\_off}^{signal} - P_{cal\_off}^{reference} \right)$$

2. Use "propagation of errors":

$$\sigma_{T_A}{}^2 = \sum \left(\frac{\partial T_A}{\partial P_i} \right)^2 \sigma_{P_i}^2$$

3. Use the following substitutions :

$$\sigma_T = T \big/ \sqrt{\Delta\nu \cdot t} \qquad P = G \cdot k \cdot T$$

$$\rightarrow \sigma_P = P \big/ \sqrt{\Delta\nu \cdot t} \qquad \rightarrow \left(\frac{\sigma_P}{P} \right)^2 = \left(\frac{\sigma_T}{T} \right)^2 = \frac{1}{\Delta\nu \cdot t}$$

Continuum - Point Sources

On-Off Observing – noise estimate

$$Ts = \frac{T_{cal}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}} \cdot \left(P_{cal\_off}^{signal} - P_{cal\_off}^{reference}\right)$$

$$\sigma_{T_A}{}^2 = \sum \left(\frac{\partial T_A}{\partial P_i}\right)^2 \sigma_{P_i}^2 = \left(\frac{\partial T_A}{\partial P_{signal}}\right)^2 \sigma_{Psignal}^2 + \left(\frac{\partial T_A}{\partial P_{reference}}\right)^2 \sigma_{\mathrm{Pr}efrence}^2 + \left(\frac{\partial T_A}{\partial P_{cal\_on}}\right)^2 \sigma_{Pcal\_on}^2$$

$$\sigma_{T_A}{}^2 = \left(\frac{T_{Cal}}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}}\right)^2 \left(\sigma_{Psignal}^2 + \sigma_{\mathrm{Pr}efrence}^2\right) + \left(\frac{T_A}{P_{cal\_on}^{reference} - P_{cal\_off}^{reference}}\right)^2 \left(\sigma_{Pcal\_on}^2 + \sigma_{\mathrm{Pr}efrence}^2\right)$$

$$\left(\frac{1}{SNR}\right)^2 = \left(\frac{\sigma_{T_A}}{T_A}\right)^2 = \left[\frac{\left(P_{cal\_on}^{reference}\right)^2 + \left(P_{cal\_off}^{reference}\right)^2}{\left(P_{cal\_on}^{reference} - P_{cal\_off}^{reference}\right)^2} + \frac{\left(P_{cal\_offl}^{signal}\right)^2 + \left(P_{cal\_off}^{reference}\right)^2}{\left(P_{cal\_off}^{signal} - P_{cal\_off}^{reference}\right)^2}\right] \cdot \left(\frac{1}{\Delta\nu \cdot t}\right)$$

$$SNR = \frac{1}{\sqrt{103 + 30}} \cdot \left(10^4\right) \sim 900 \qquad (Not\,3000!)$$

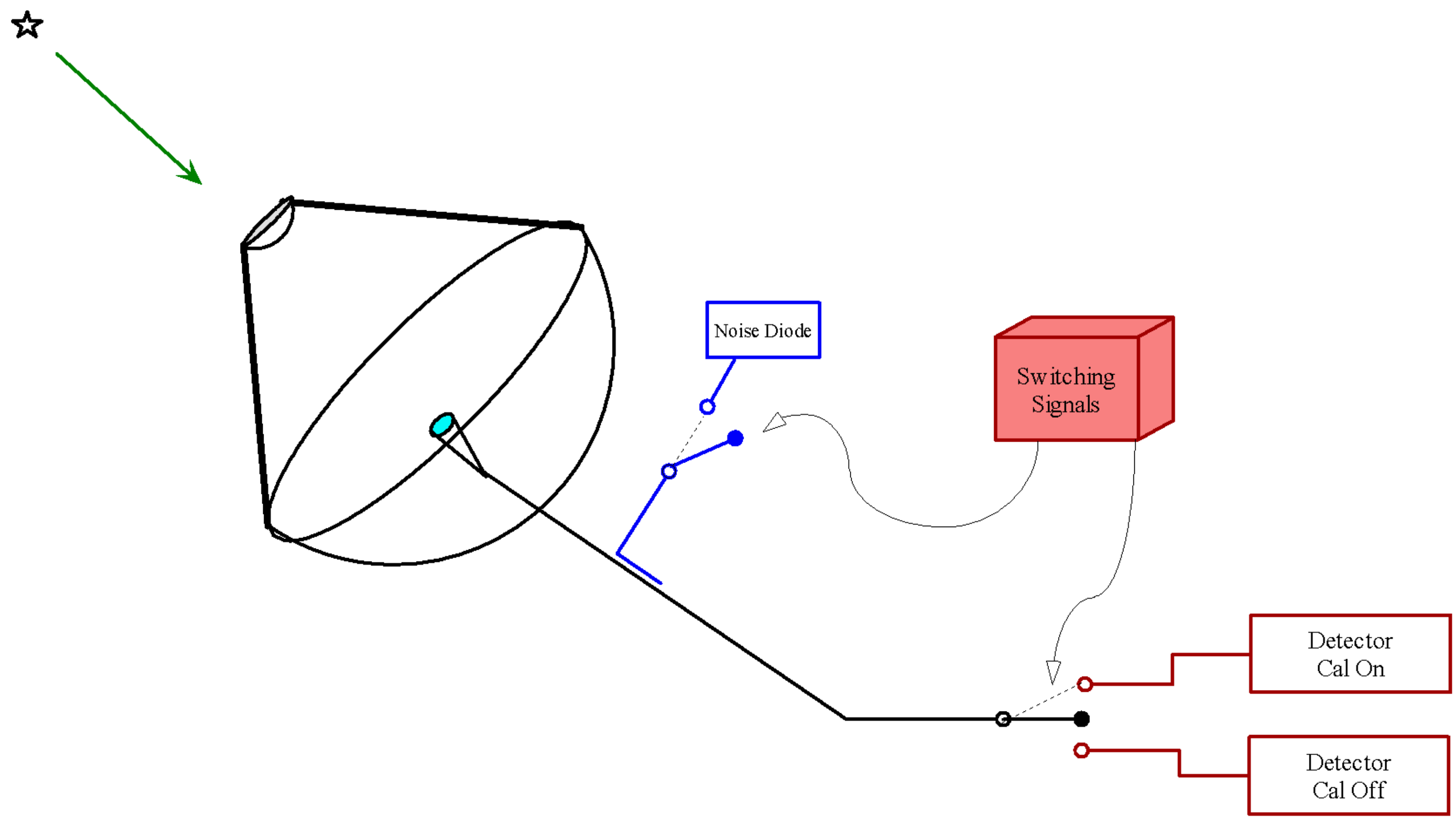
Continuum - Point Sources

Assumptions:

"Classical" Radiometer equation assumes:

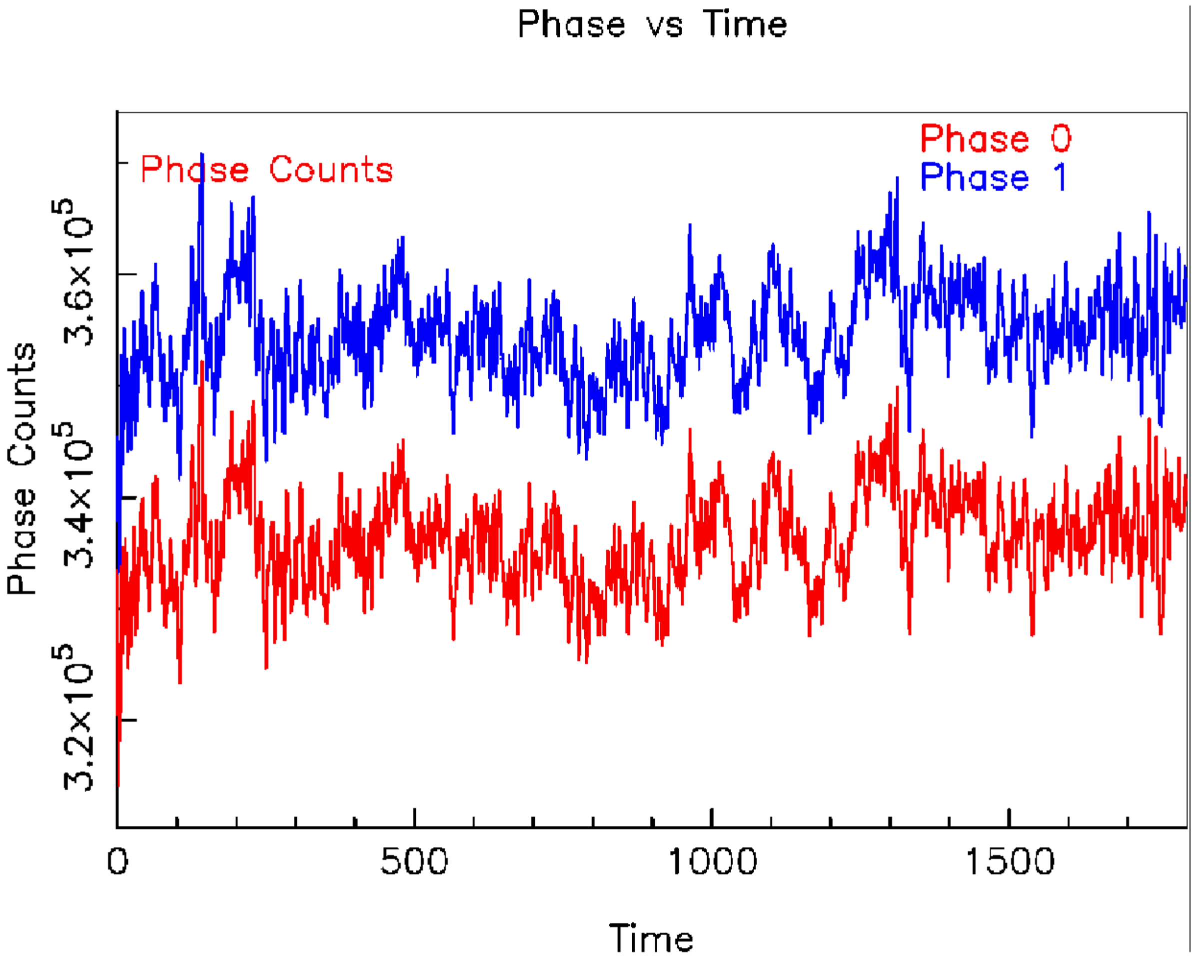
- Narrow bandwidths,
- Linear power detector,
- $T_A << T_s$,
- Noise diode temperature $<< T_s$,
- $t^{reference} = t^{signal}$
- $t_{cal\_on} = t_{cal\_off}$
- Blanking time $<< t^{signal}$
- No data reduction!

Phases of an Observation
Total Power



Phases of an Observation

Total Power



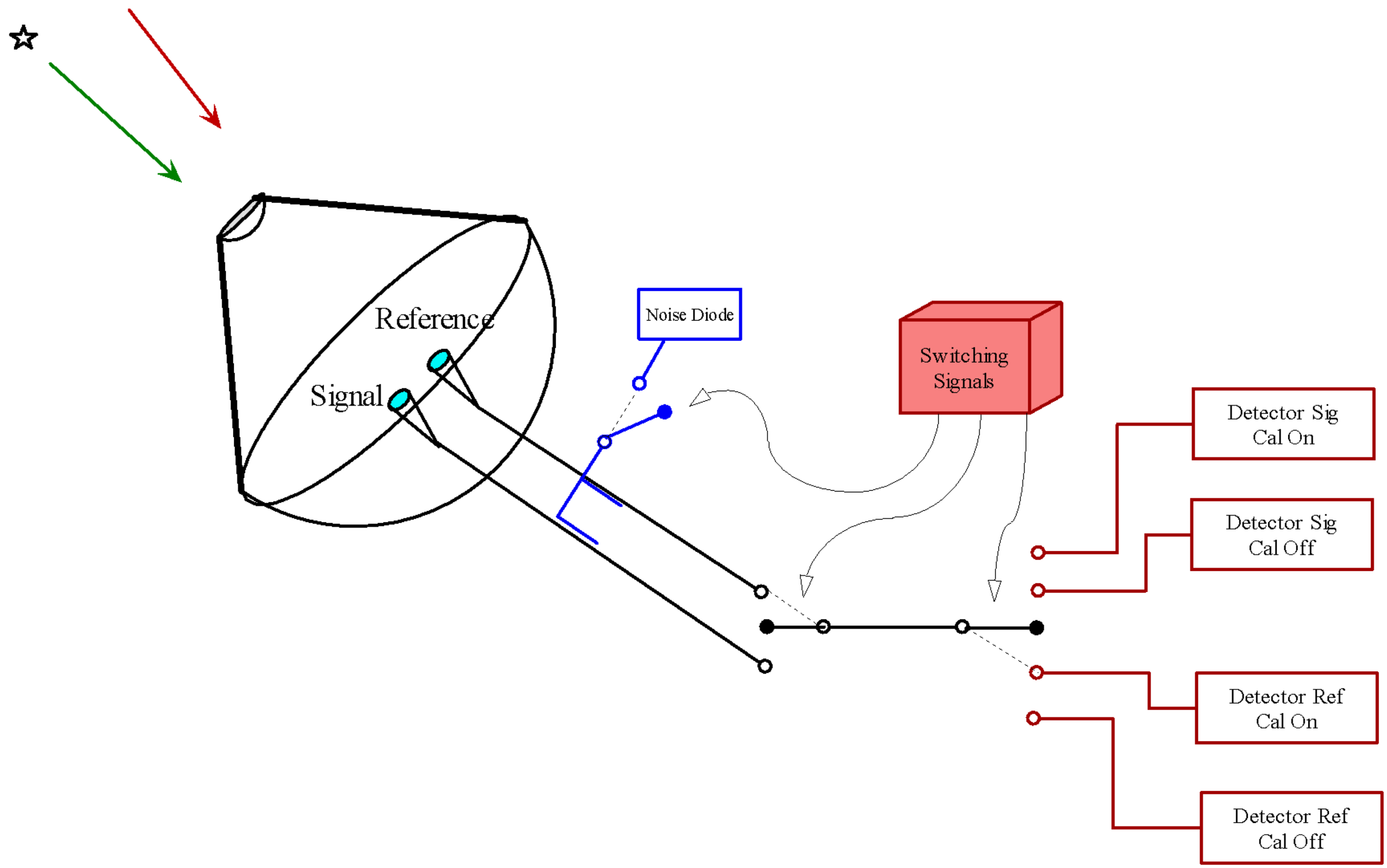
- Tcal = 4 K
- Ts = 100 K

- σ_{theor} =0.1 K
- σ_{meas} =1 K

- Shapes very similar
- Excess noise from atmospheric fluctuation

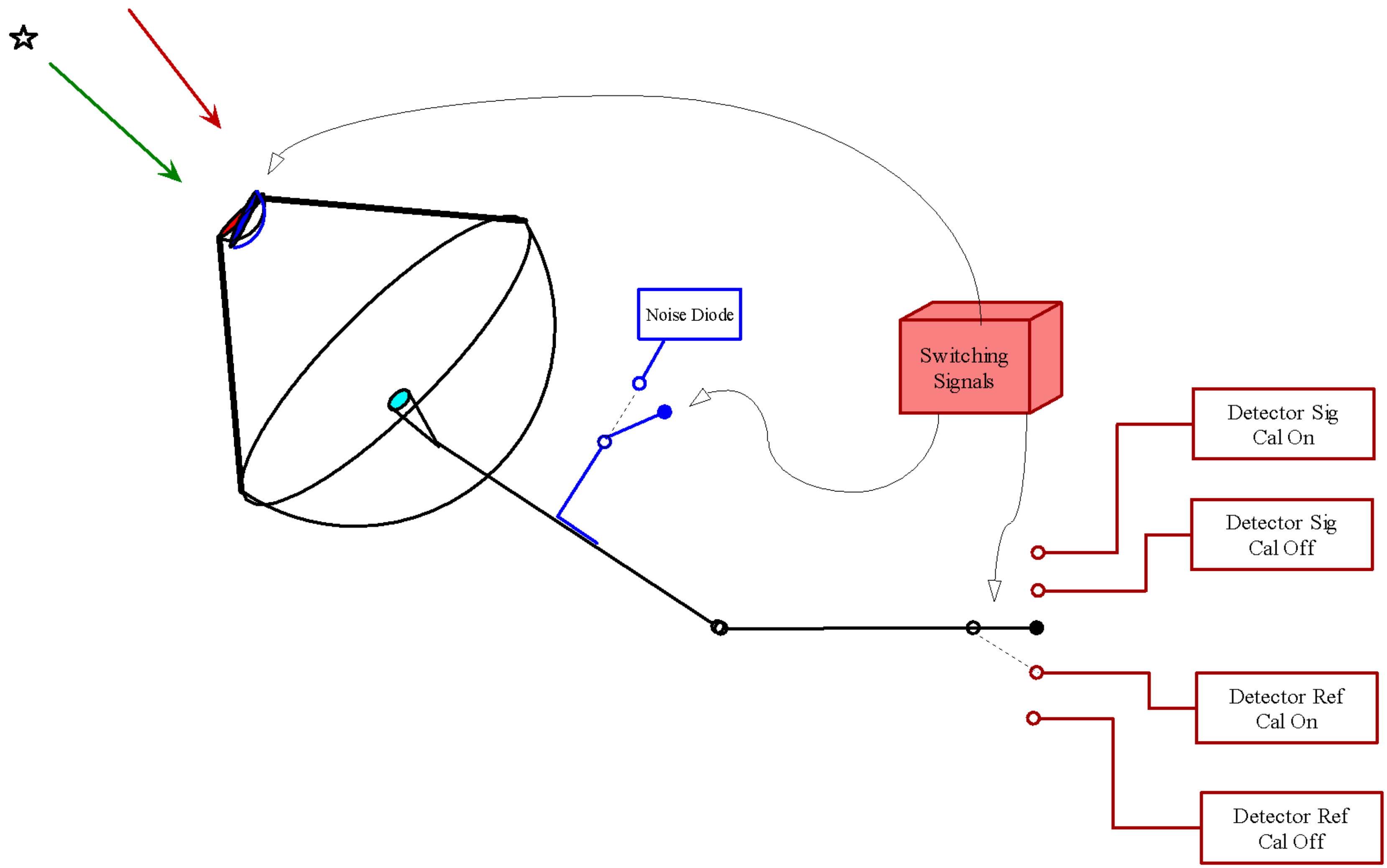
Phases of a Observation

Beam Switched Power



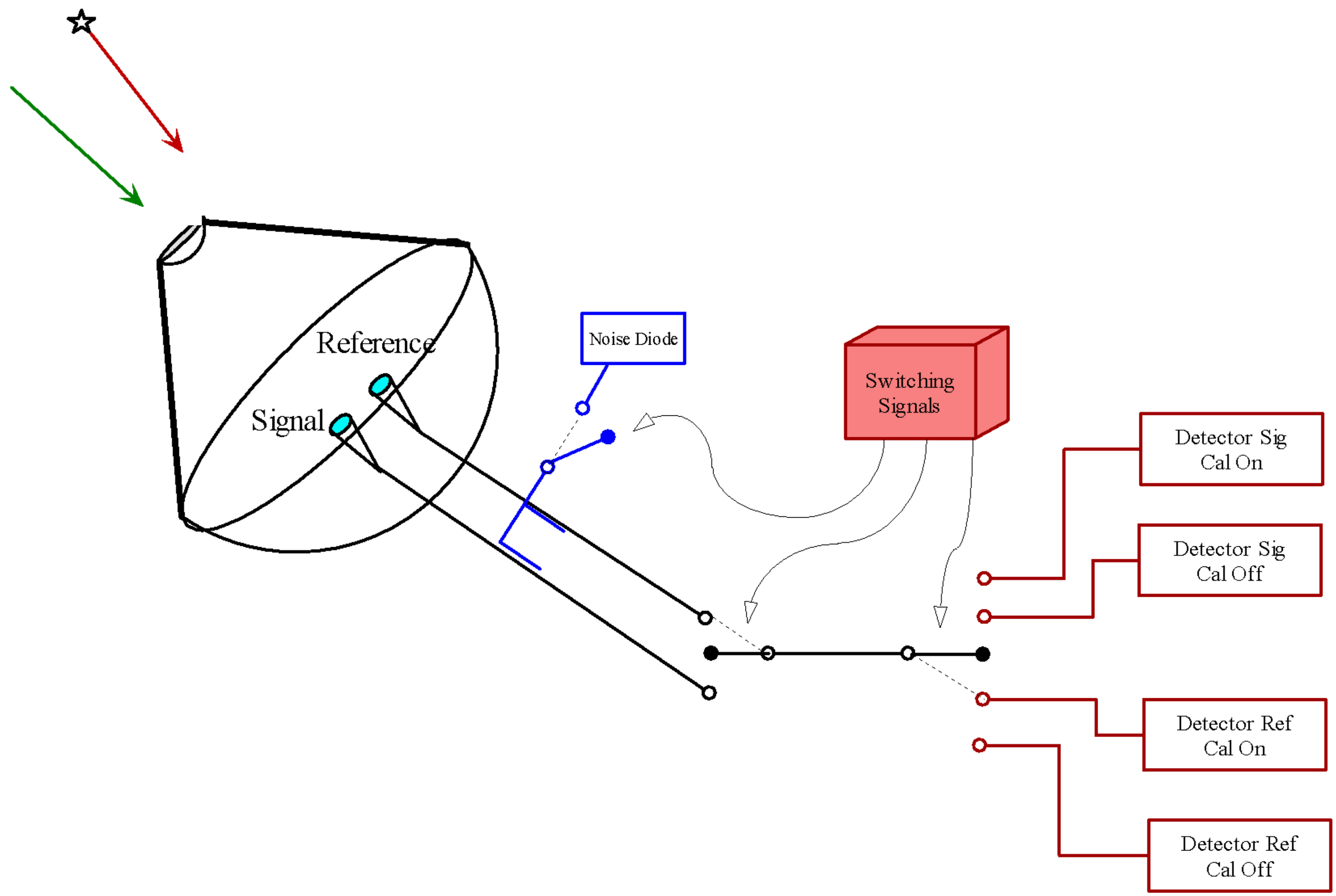
Phases of a Observation

Beam Switched Power



Phases of a Observation

Double Beam Switched Power



Continuum - Point Sources

Beam-Switched Observation

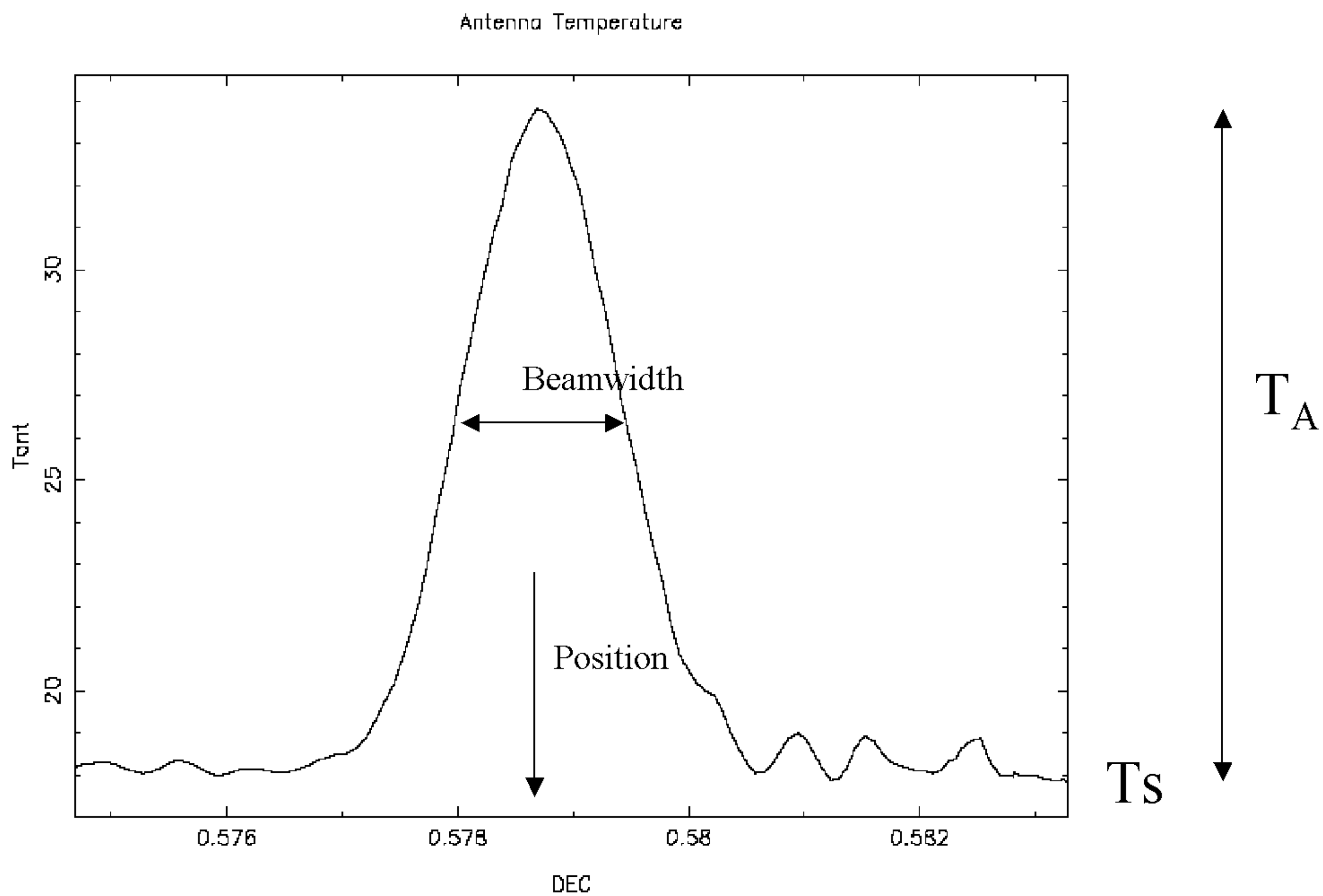
$$T_S^{reference}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{reference}(i) - P_{cal\_off}^{reference}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{reference}(i) + P_{cal\_off}^{reference}(i)\right)}{2}$$

$$T_S^{signal}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{signal}(i) - P_{cal\_off}^{signal}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{signal}(i) + P_{cal\_off}^{signal}(i)\right)}{2}$$

$$T_A = \left\langle T_S^{signal}(i) - T_S^{reference}(i) \right\rangle$$

Continuum - Point Sources

On-The-Fly Observation



Continuum - Point Sources

On-The-Fly Observation

If total power:

$$T_S(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}(i) - P_{cal\_off}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}(i) + P_{cal\_off}(i)\right)}{2}$$

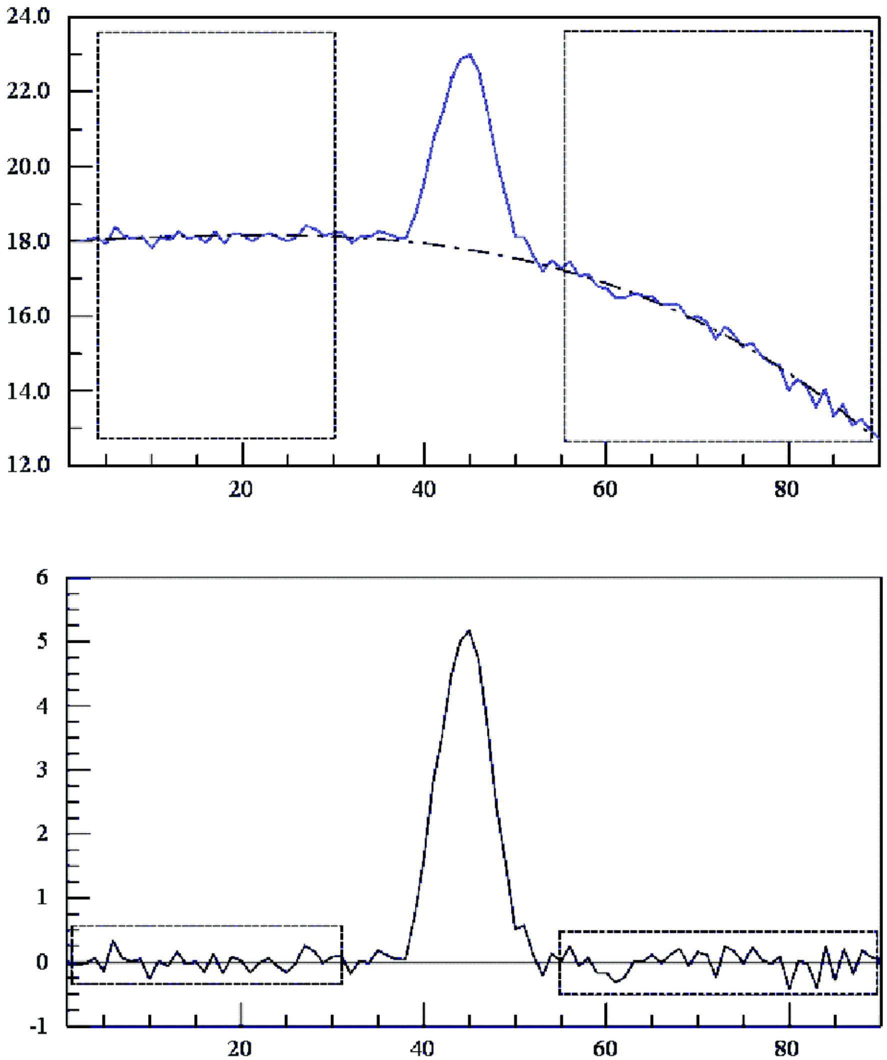
If beam-switching (switched power):

$$T_S^{reference}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{reference}(i) - P_{cal\_off}^{reference}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{reference}(i) + P_{cal\_off}^{reference}(i)\right)}{2}$$

$$T_S^{signal}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{signal}(i) - P_{cal\_off}^{signal}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{signal}(i) + P_{cal\_off}^{signal}(i)\right)}{2}$$

$$T_A(i) = T_S^{signal}(i) - T_S^{reference}(i)$$

Baseline Fitting
Polynomials



- Set order of polynomial
- Define areas devoid of emission.

- Creates false features
- Introduces a random error to an observation

$$\sigma^2_{Peak} = \sigma^2_{TA} + \sigma^2_{Polynomial}$$

Why Polynomials?

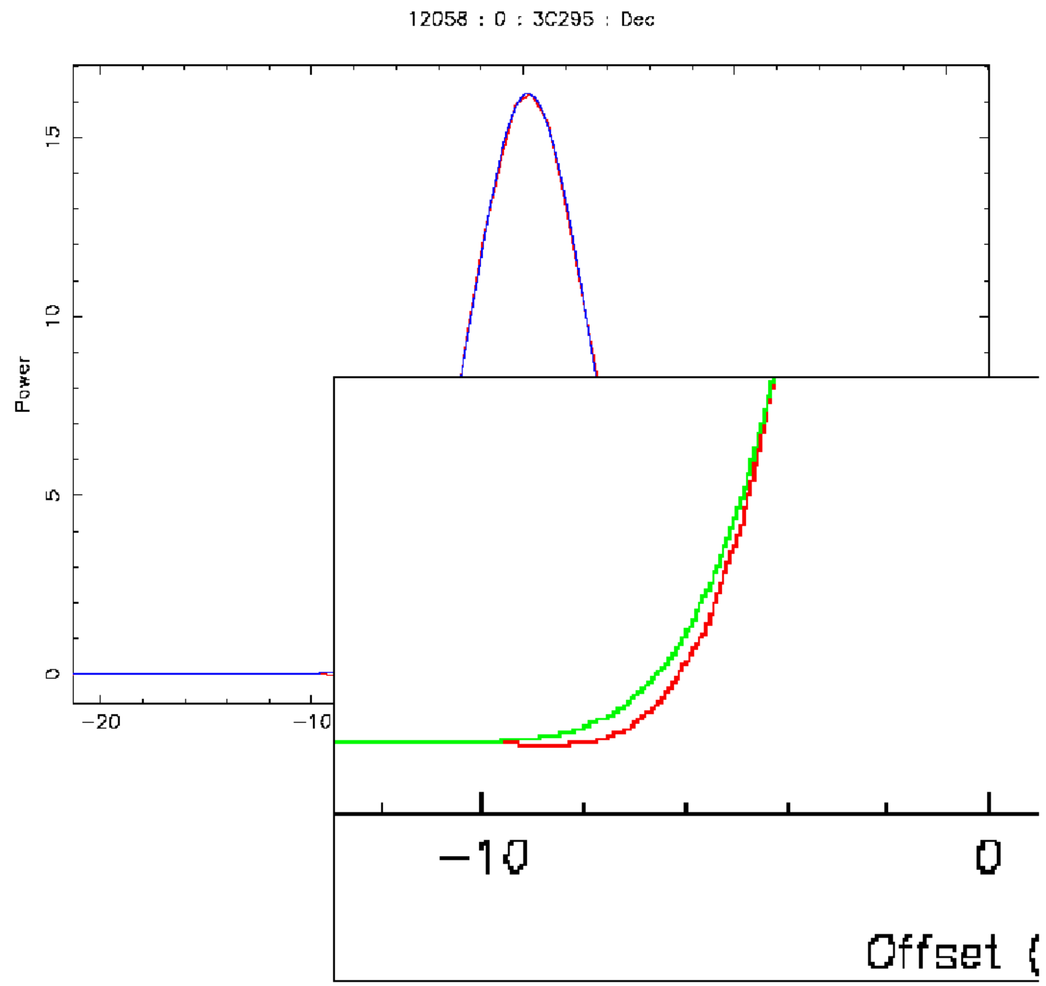
Continuum - Point Sources

Gaussian Fitting

- Define initial guesses
- Set flags to fit or hold constant each parameter
- Set number of iterations
- Set convergence criteria

- Fitted parameters
- Chi-square of the fit
- Parameter standard deviations.

- Restrict data to between the half power points for fitting to a telescope's beam
- Multi-component fits should be done simultaneously



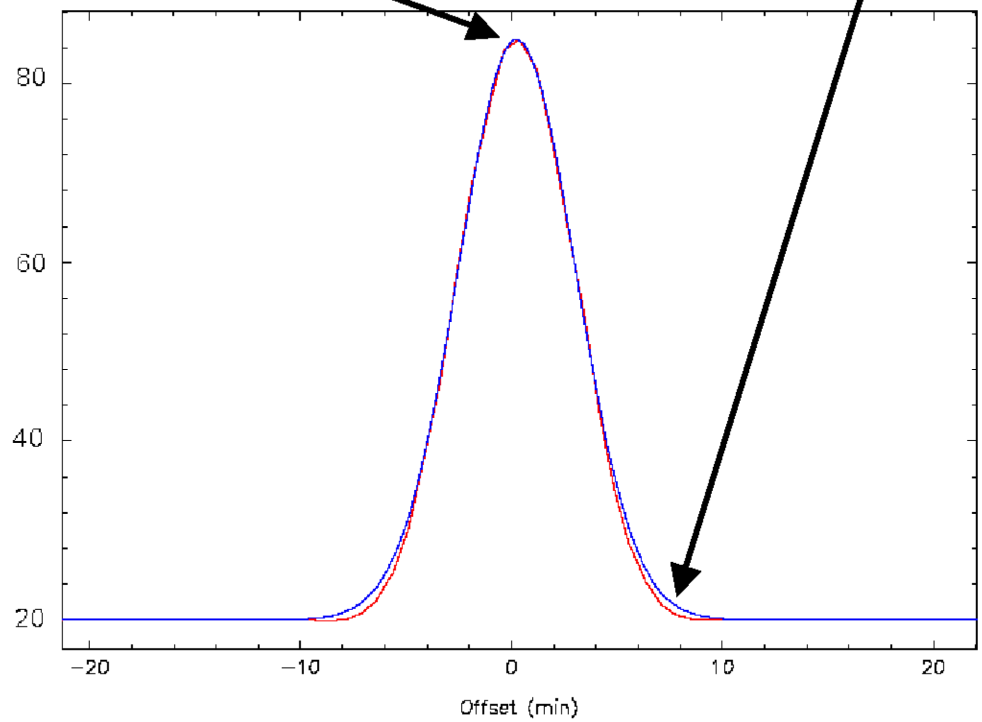
Continuum - Point Sources

Gaussian Fitting

Where is noise the highest?

Where is noise the lowest?

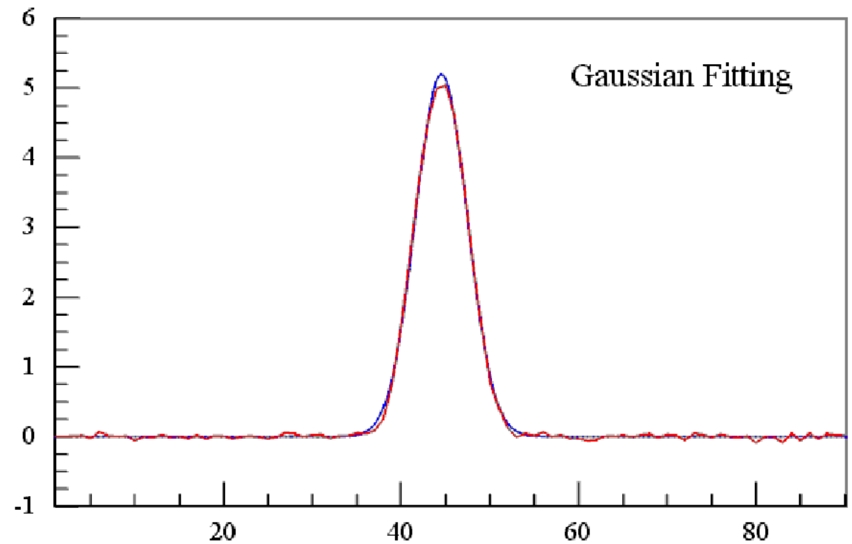
- σ changes across the observation.
- Weights ($1/\sigma^2$) for least-square-fit changes across the observation.
- For strong sources, should worry about using proper weights in data analysis.

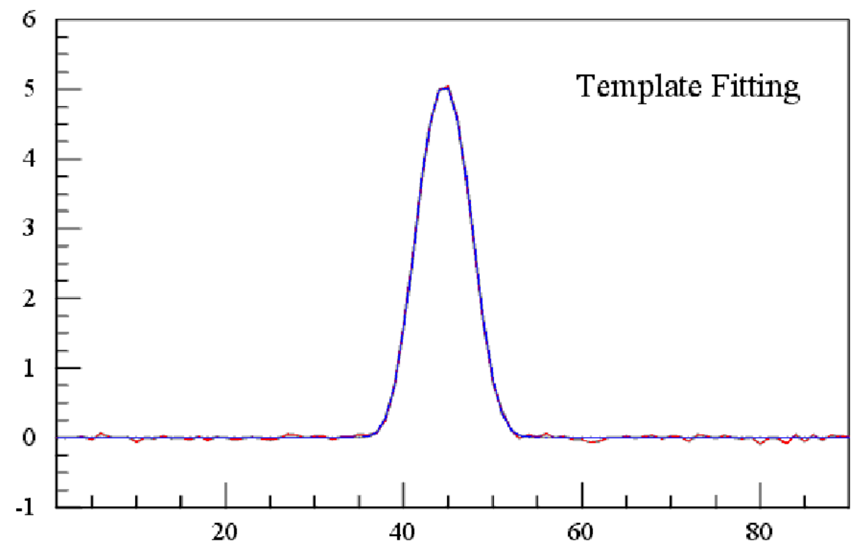


Template Fitting

- Create a template:
  - Sufficient knowledge of the telescope beam, or
  - Average of a large number o observations.
- ---------------
- Convolve the template with the data => x-offset.
- Shift by the x-offset.
- Perform a linear least-square fit of the template to the data:

Always try to fit physically-meaningful functions





Averaging Data / Atmosphere

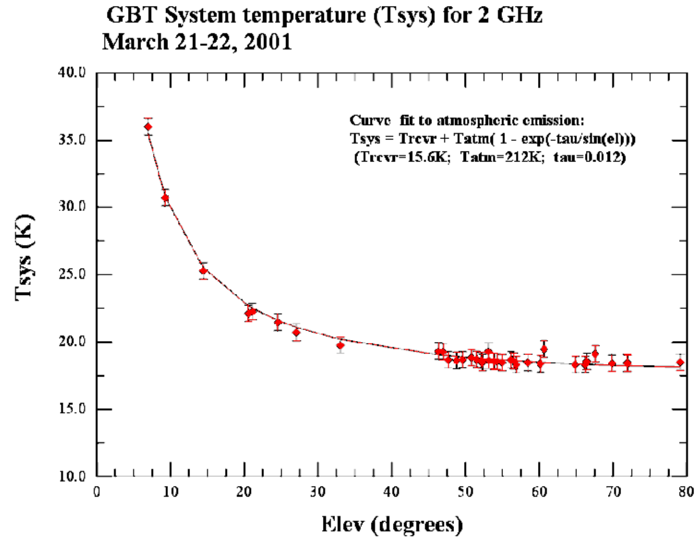
- T_s changes due to atmosphere emission.
- Use weighted average with weights = $1/\sigma^2$

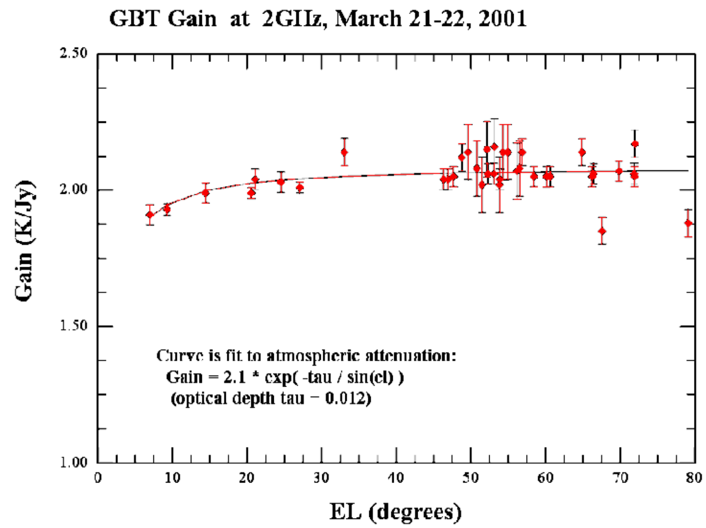
$$\langle T_A \rangle = \frac{\sum T_A \frac{1}{\sigma_j^2}}{\sum \frac{1}{\sigma_j^2}} \qquad \sigma_{avrg} = \frac{1}{\sqrt{\sum_j \frac{1}{\sigma_j^2}}}$$

- T_A changes due to atmosphere opacity.
- Opacity from the literature or theory, from a tipping radiometer, from atmospheric vertical water vapor profiles, or by "tipping" the antenna

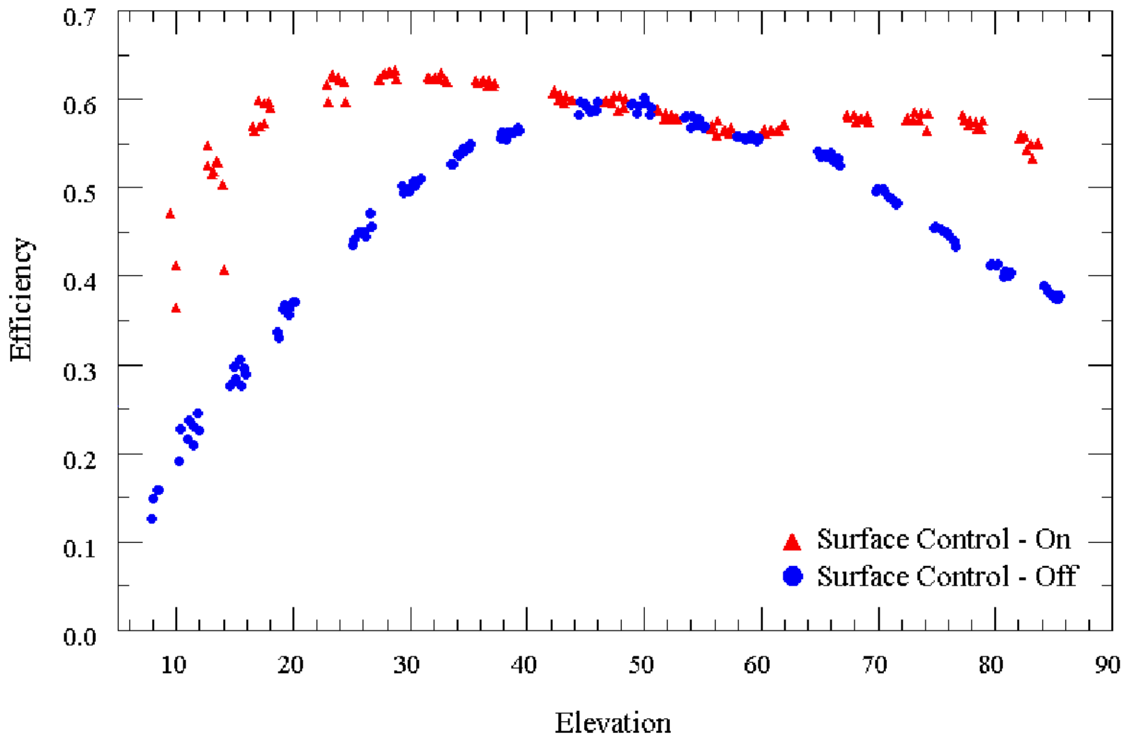
$$T_A' = T_A \cdot e^{tau/\sin(el)}$$

$$\sigma_{TA}' = \sigma_{TA} \cdot e^{tau/\sin(el)}$$





Gain Correction



$$T_A^* = T_A^{'} / \eta_A \qquad or \qquad T_B^* = T_A^{'} / \eta_M$$

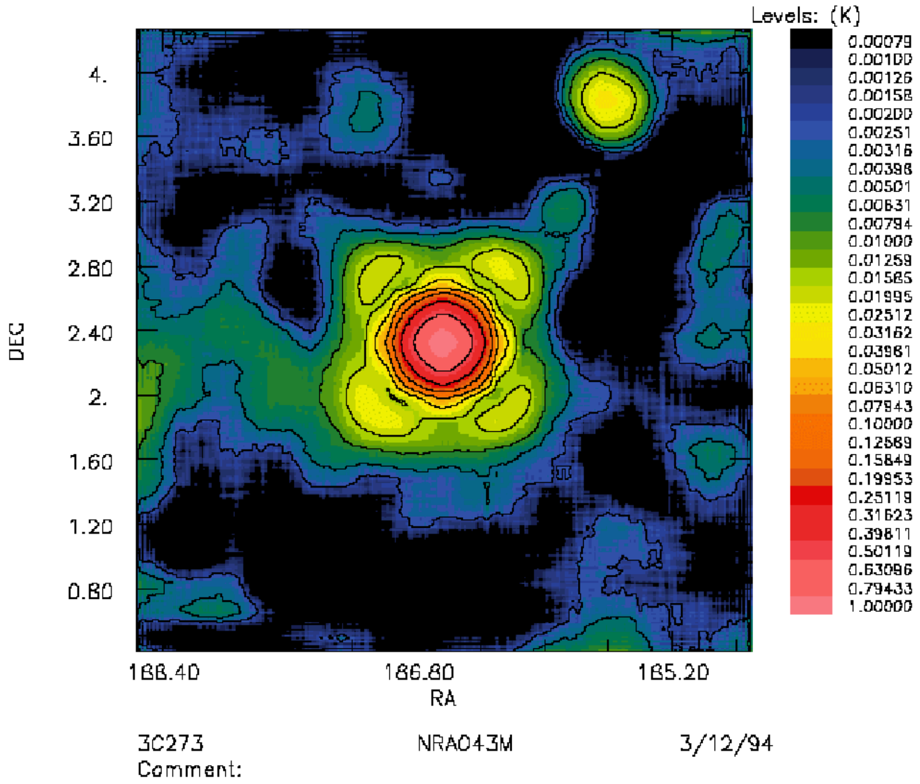
$$\sigma_{TA}^* = \sigma_{TA}^{'} / \eta_A \qquad\qquad \sigma_{TB}^* = \sigma_{TA}^{'} / \eta_M$$

Continuum - Extended Sources

On-The-Fly Mapping

- Telescope slews from row to row. Row spacing: ~0.9λ/2D
- A few samples /sec.
- Highly oversampled in direction of slew <0.3λ/2D
- Could be beam switching

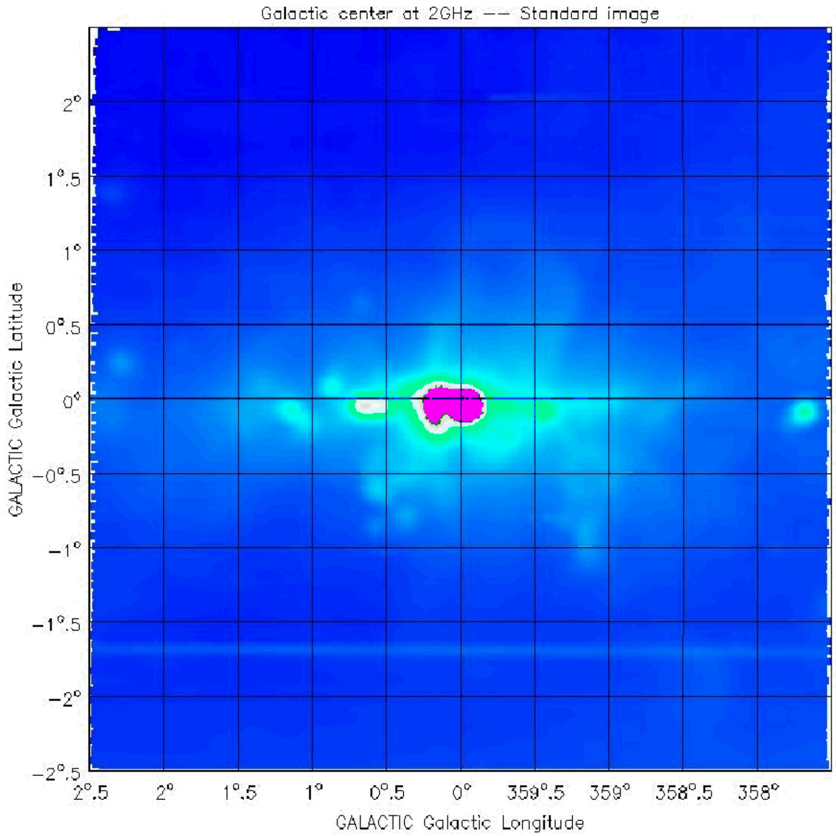
- Convert Power into T_S.
- Fit baseline to each row?
- Grid into a matrix



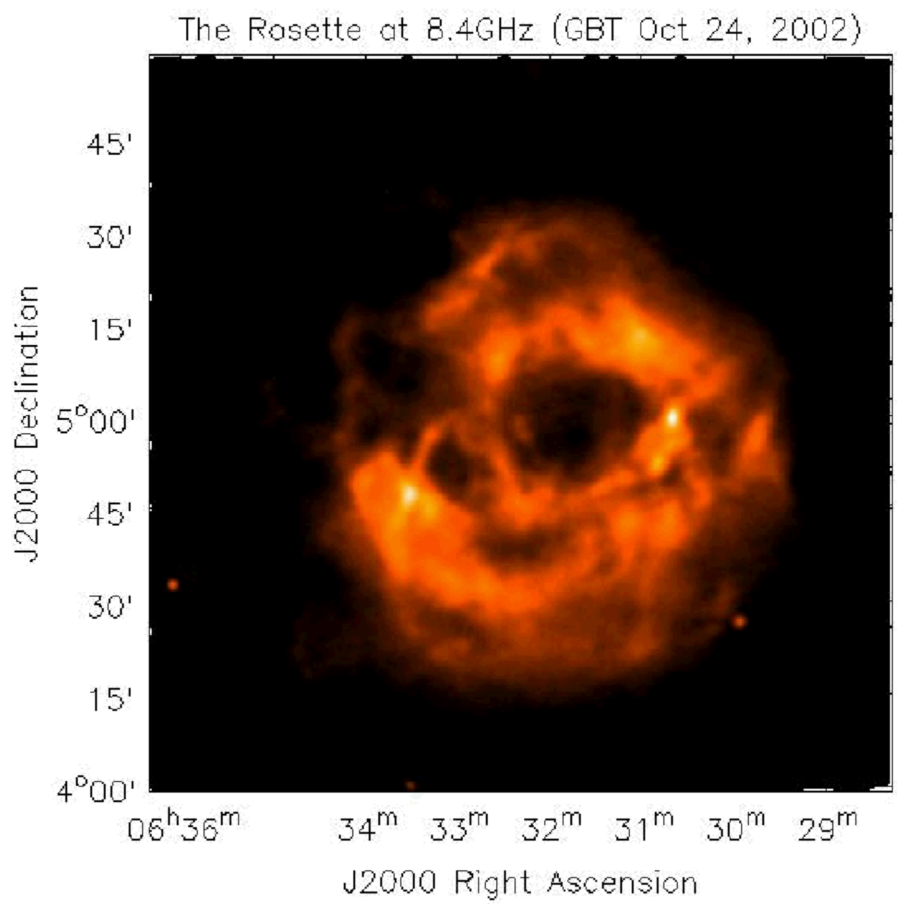
Continuum - Extended Sources

On-The-Fly Mapping - Common Problems

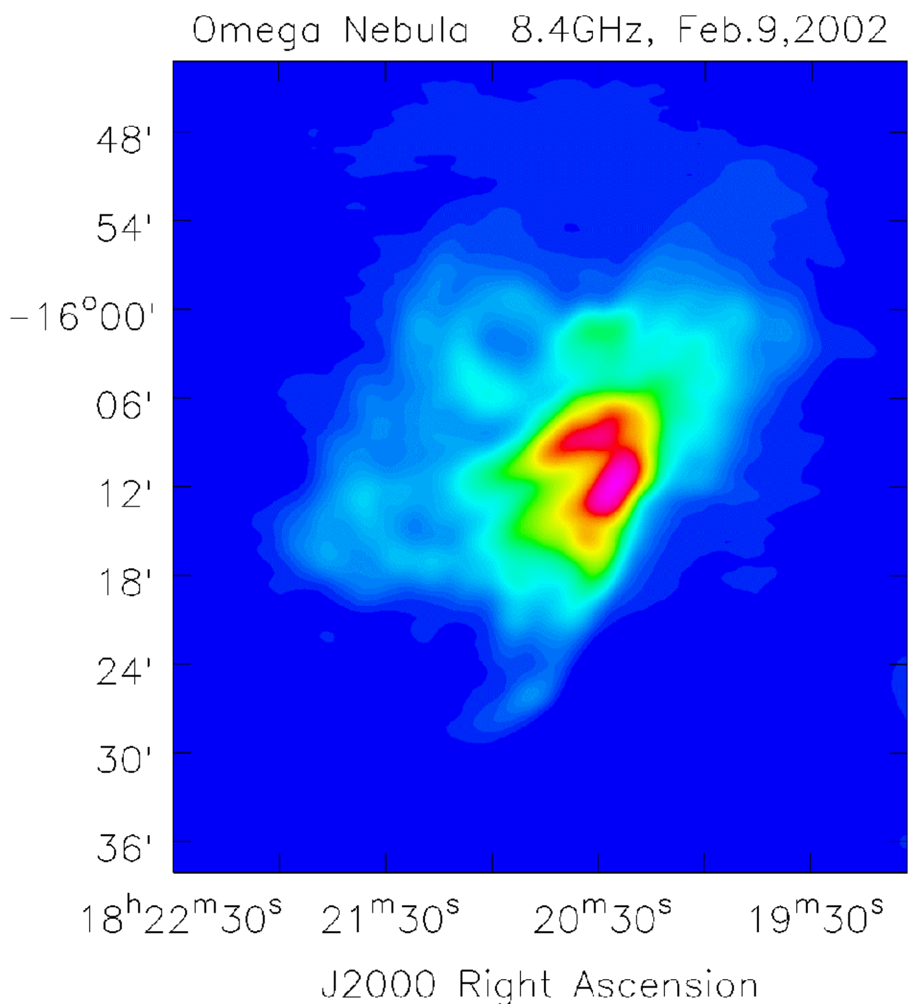
- Striping (Emerson 1995; Klein and Mack 1995).
- If beam-switched, Emerson, Klein, and Haslam (1979) to reconstruct the image.
- Make multiple maps with the slew in different direction.



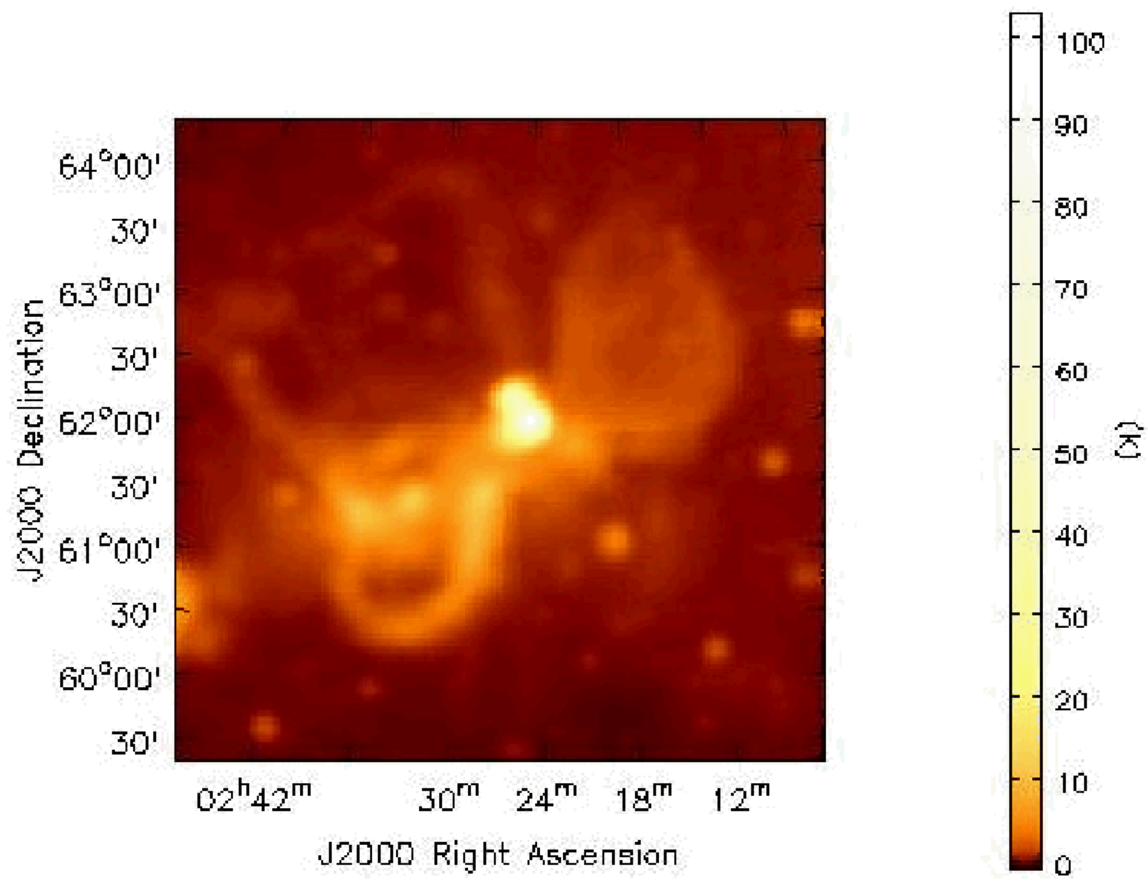
GBT Continuum Images – Rosette



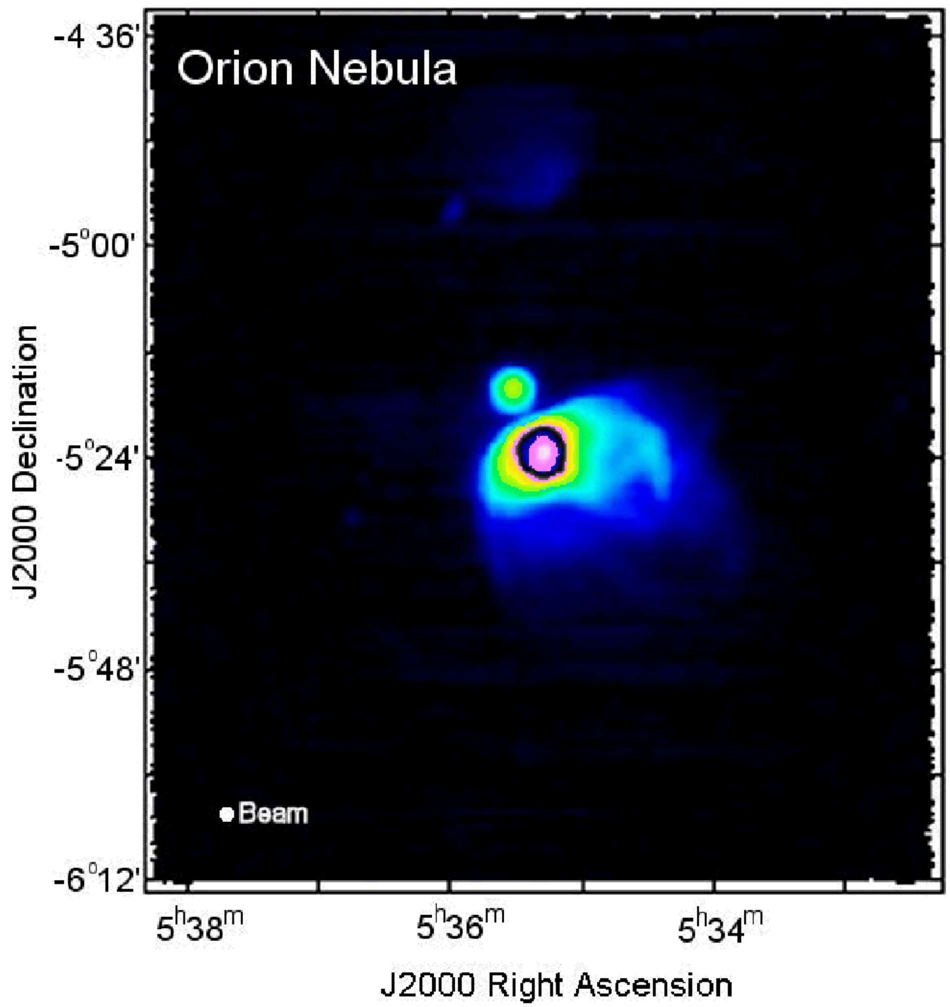
GBT Continuum Images – M17



GBT Continuum Images – W3

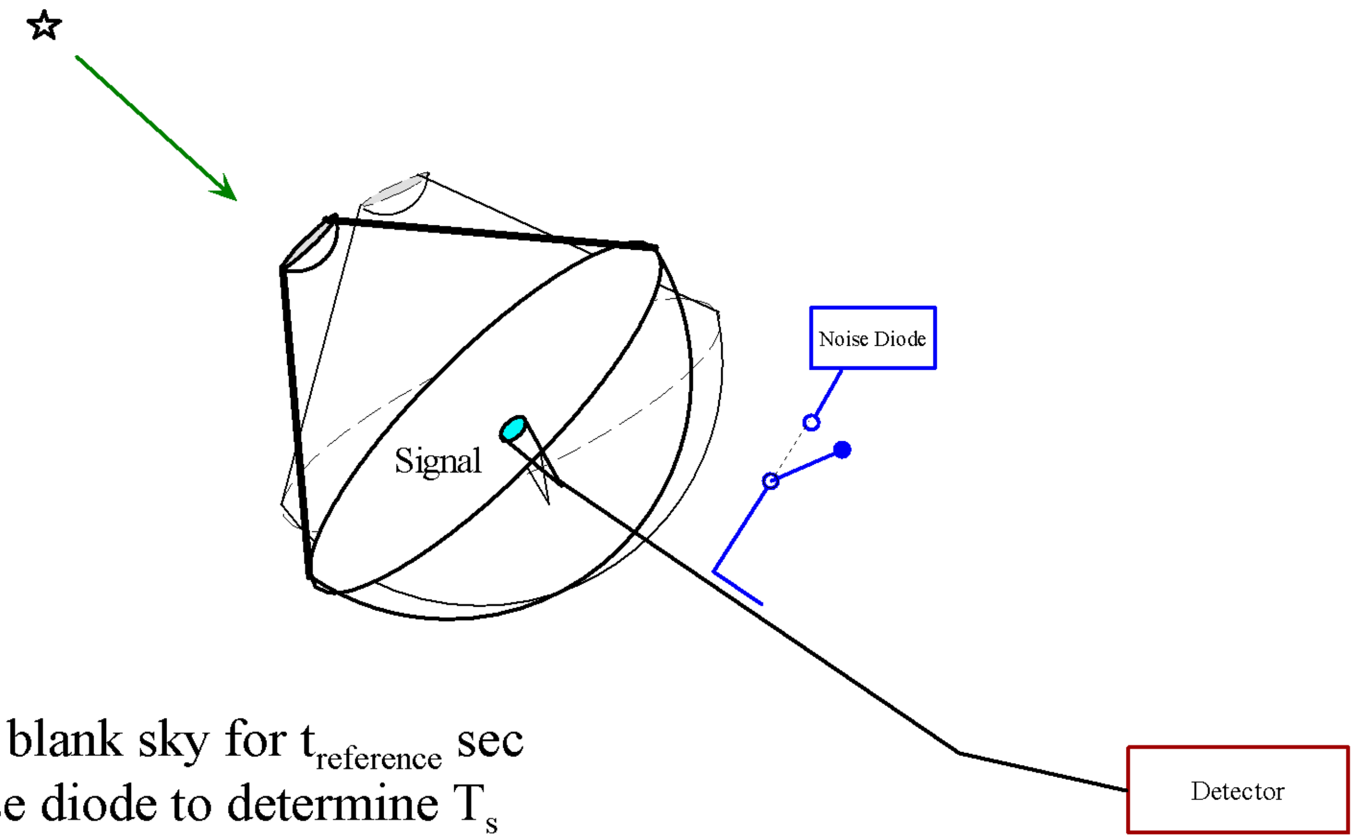


GBT Continuum Images - Orion



Spectral-line - Point Sources

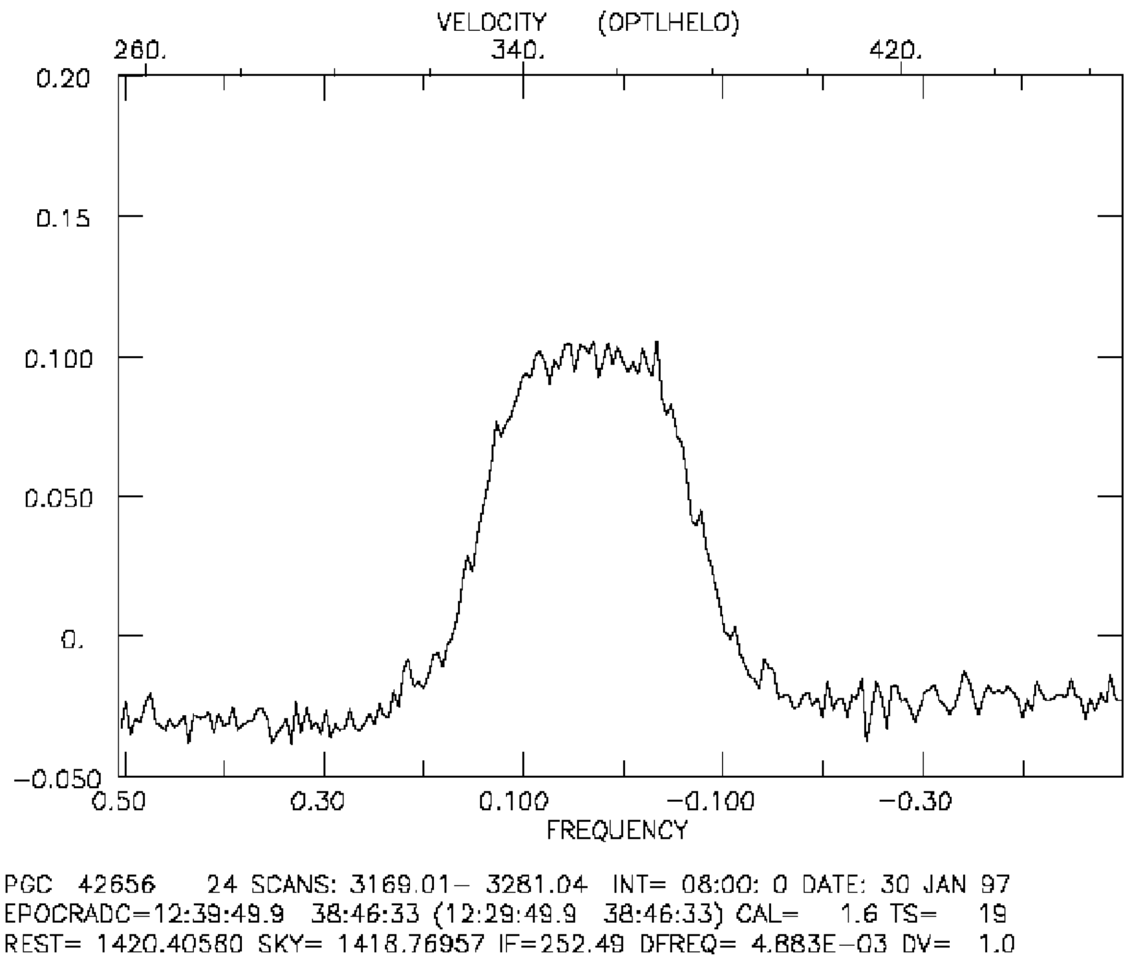
On-Off Observing



- Observe blank sky for $t_{reference}$ sec
- Fire noise diode to determine T_s
- Move telescope to object & observe for t_{signal} sec
- Can observe an extended source using this technique -- 'signal' observations arranged in a "grid" map.

Spectral-Line - Point Sources

Position-Switched Observing



Spectral-Line - Point Sources

Position-Switched Observing

$$T_A(f) = T_s^{reference}(f) \cdot \left[\frac{P^{signal} - P^{reference}}{P^{reference}} \right]$$

Smoothed/Averaged T_s of Denominator

Signal (line expected)

Reference (No line expected)

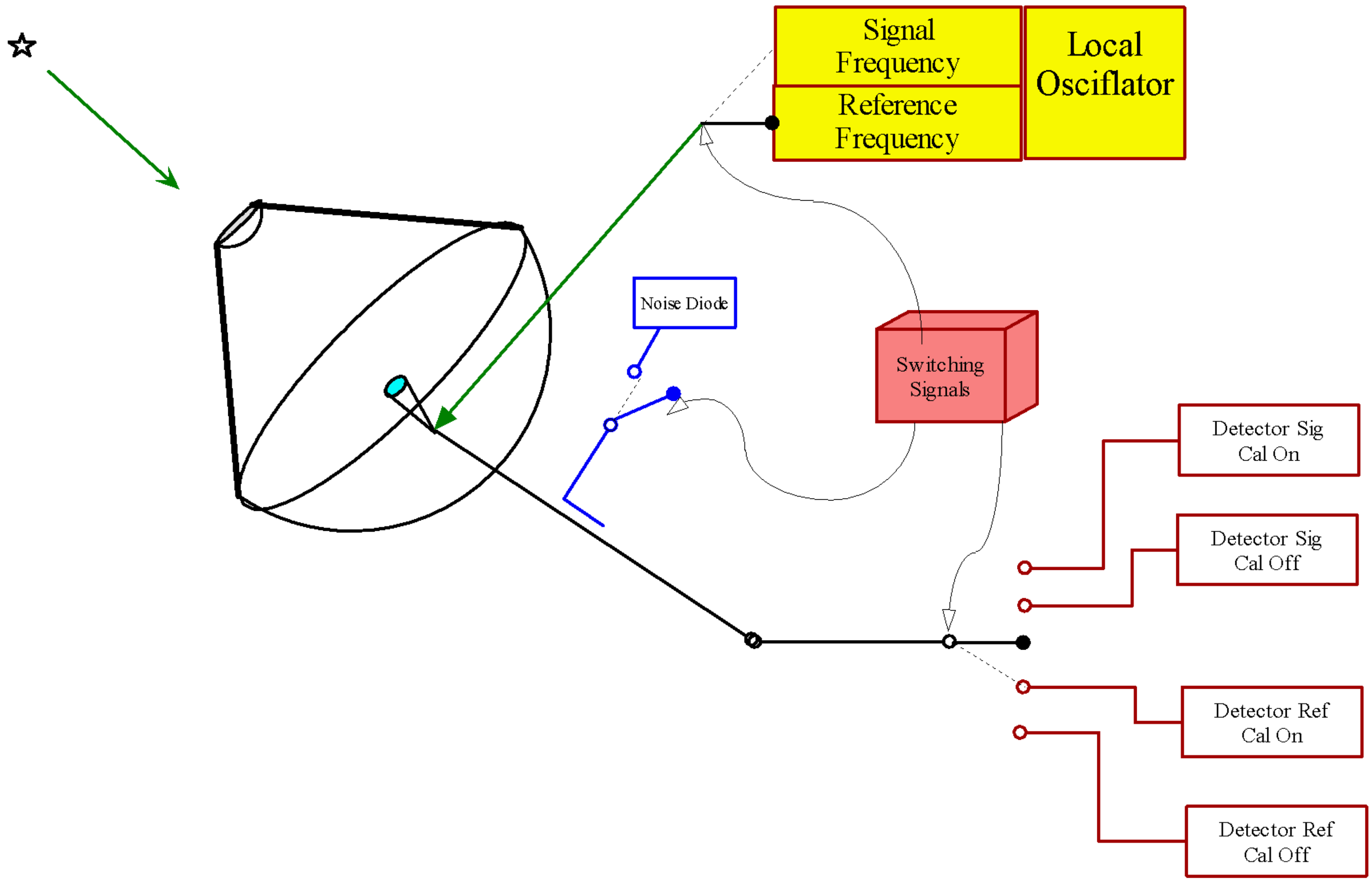
$$T_s^{reference}(f) = \left\langle \left(T_{cal}/2 \right) \cdot \left(\frac{P_{cal\_on}^{reference}(f) + P_{cal\_off}^{reference}(f)}{P_{cal\_on}^{reference}(f) - P_{cal\_off}^{reference}(f)} \right) \right\rangle_{M\_Channels}$$

$$\left(\frac{\sigma_{TA}}{T_A} \right)^2 \sim \frac{K}{\Delta\nu / N_{channels}} \left(\frac{1}{t^{reference}} + \frac{1}{t^{signal}} \right) + \left(\frac{\sigma_{Ts}}{T_S} \right)^2$$

- But only for weak lines and no strong continuum!
- Constant depends upon details of the detecting backend

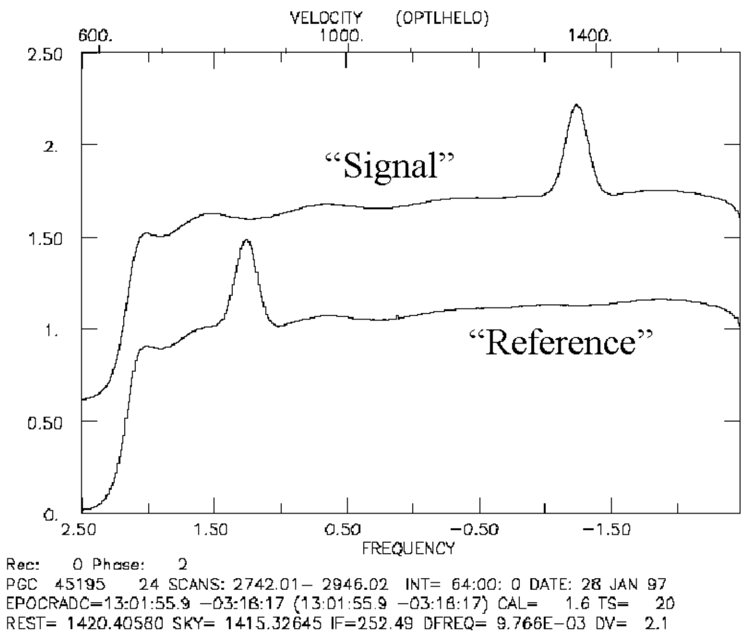
Phases of a Observation

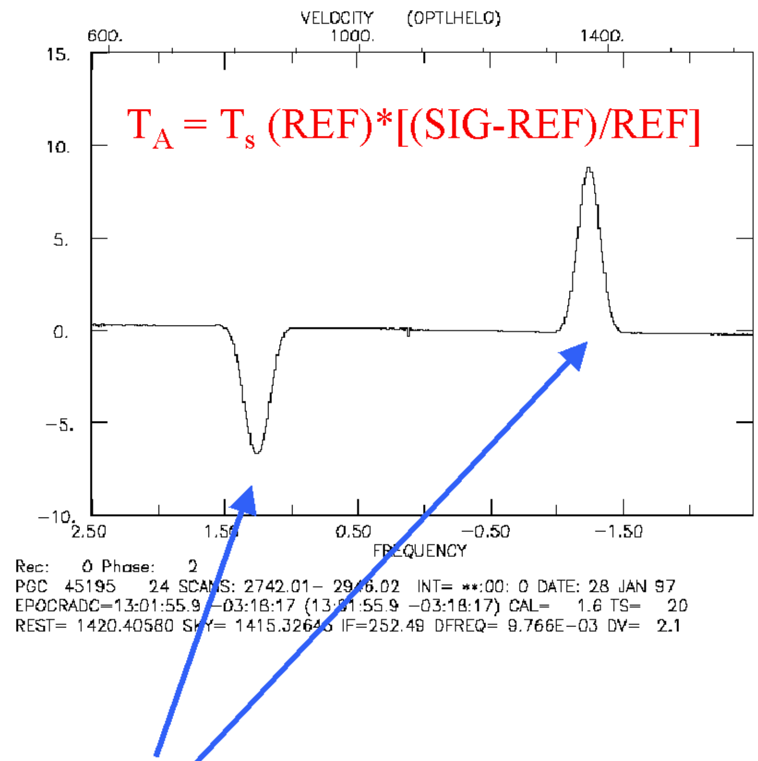
Switched Power – Frequency Switching



Spectral-Line - Point Sources

Frequency-Switched Observing - In band





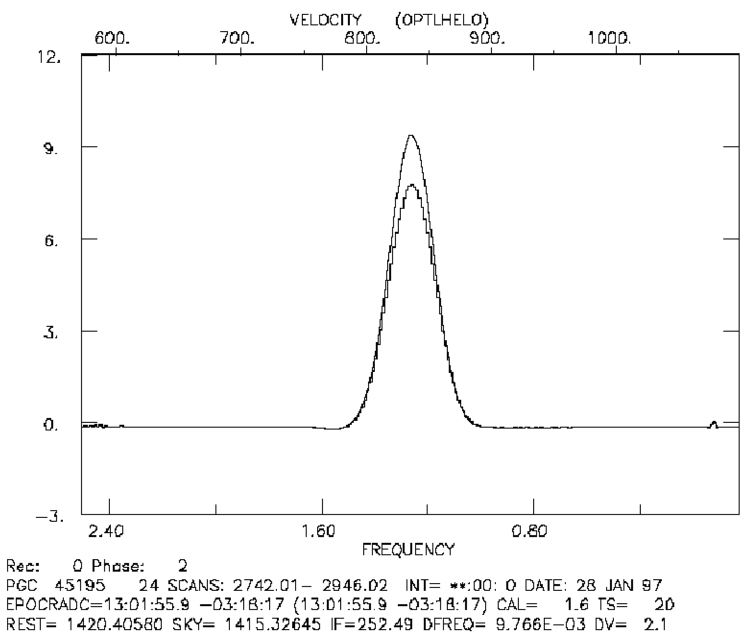
Line appears twice – should be able to 'fold' the spectra to increase SNR

Spectral-Line - Point Sources

Frequency-Switched – “Folding” In Band

$$T_A(f) = T_S^{reference}(f) \cdot \left[\frac{P^{signal}(f) - P^{reference}(f)}{P^{reference}(f)} \right] +$$

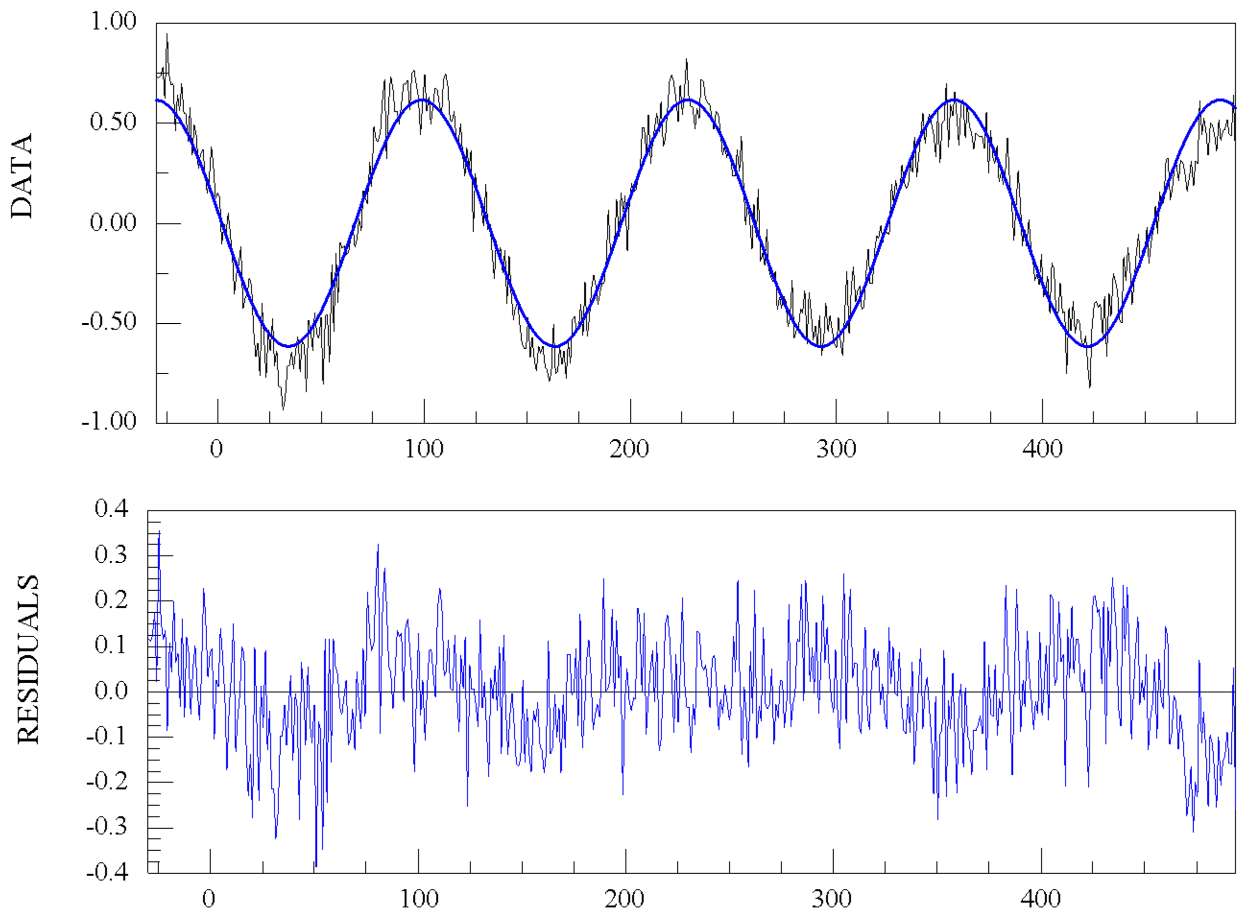
$$T_S^{signal}(f) \cdot \left[\frac{P^{reference}(f+\Delta f) - P^{signal}(f+\Delta f)}{P^{signal}(f+\Delta f)} \right]$$



$$T_{ant}(f) = T_S^{reference}(f) \cdot \left\{ \frac{P^{signal}(f) - P^{reference}(f)}{P^{reference}(f)} - \frac{P^{signal}(f+\Delta f) - P^{reference}(f+\Delta f)}{P^{reference}(f+\Delta f)} \right\}$$

Spectral-Line
Baseline Fitting

- Polynomial: same as before
- Sinusoid

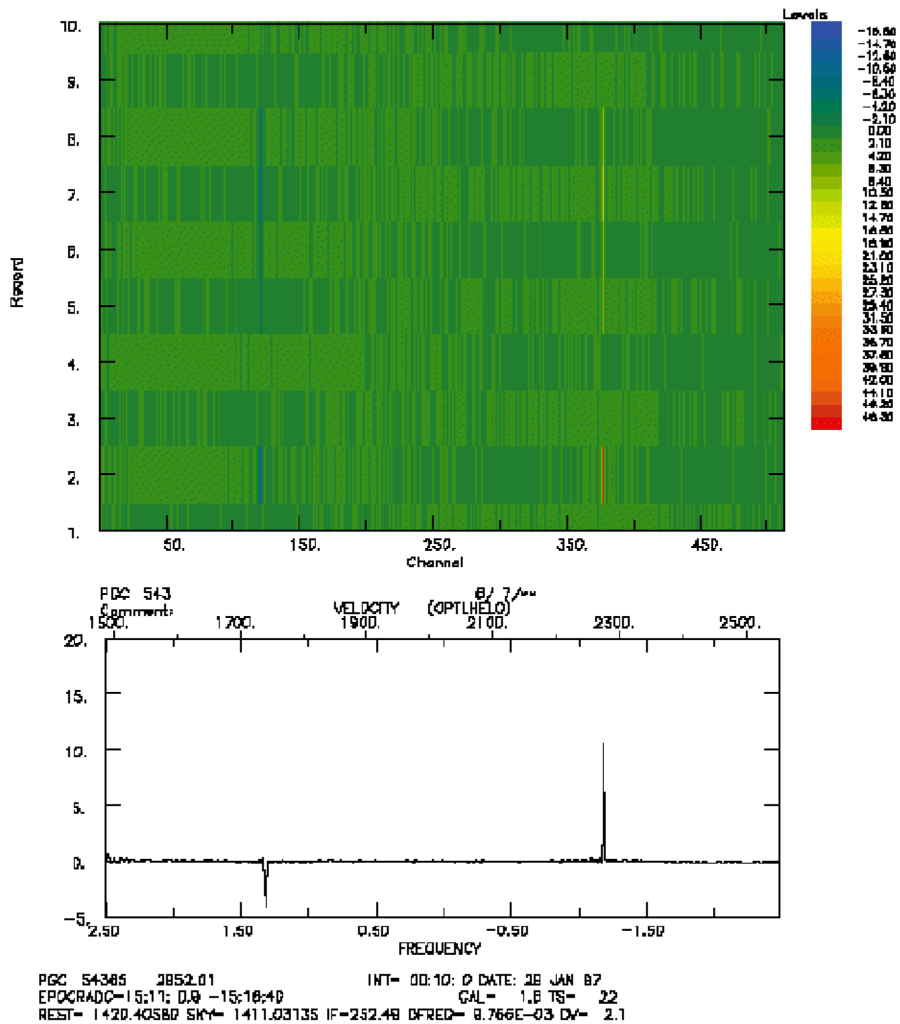


Spectral-Line
Other Algorithms

- Velocity Calibration
- Velocity/Frequency Shifting & Regriding
  - Doppler tracking limitations
- Smoothing – Hanning, Boxcar, Gaussian
  - Decimating vs. non-decimating routines
  - For "Optimal Filtering", match smoothing to expected line width
- Filtering – low pass, high pass, median, …
- Moments for Integrated Intensities; Velocity centroids, ….

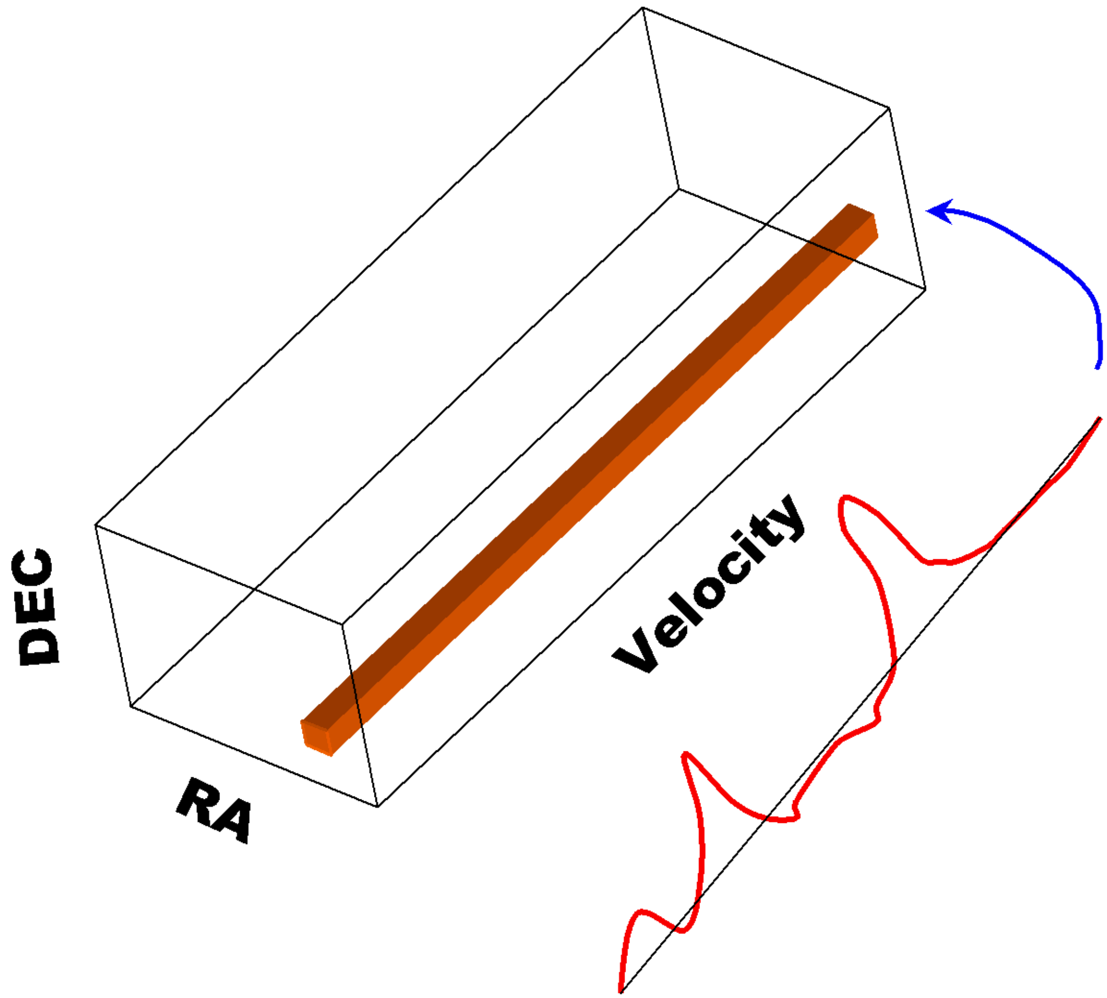
Spectral-Line

RFI Excision



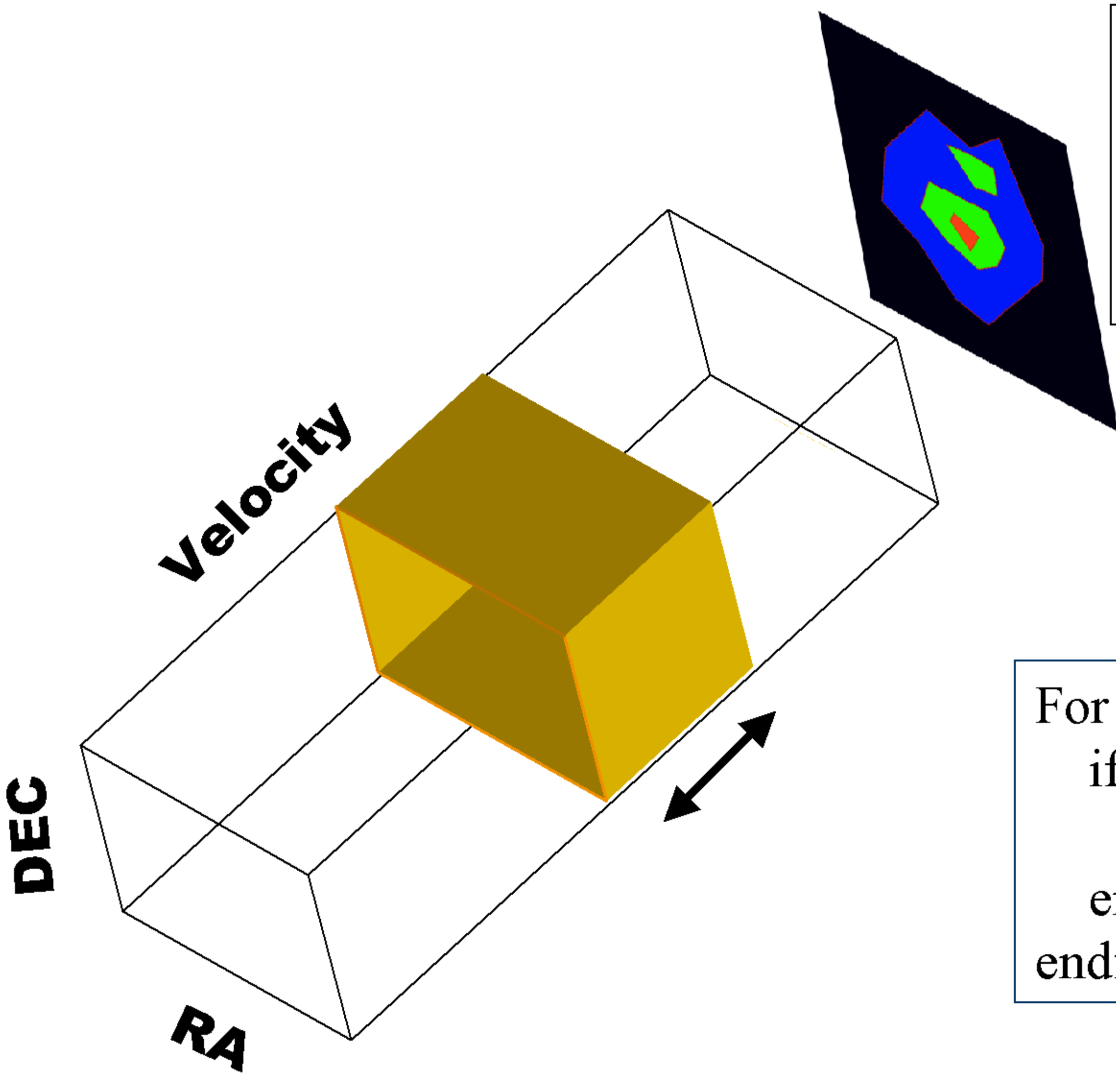
Spectral-Line Mapping

Grid or On-the-Fly



Spectral-Line Mapping

Grid and On-the-Fly

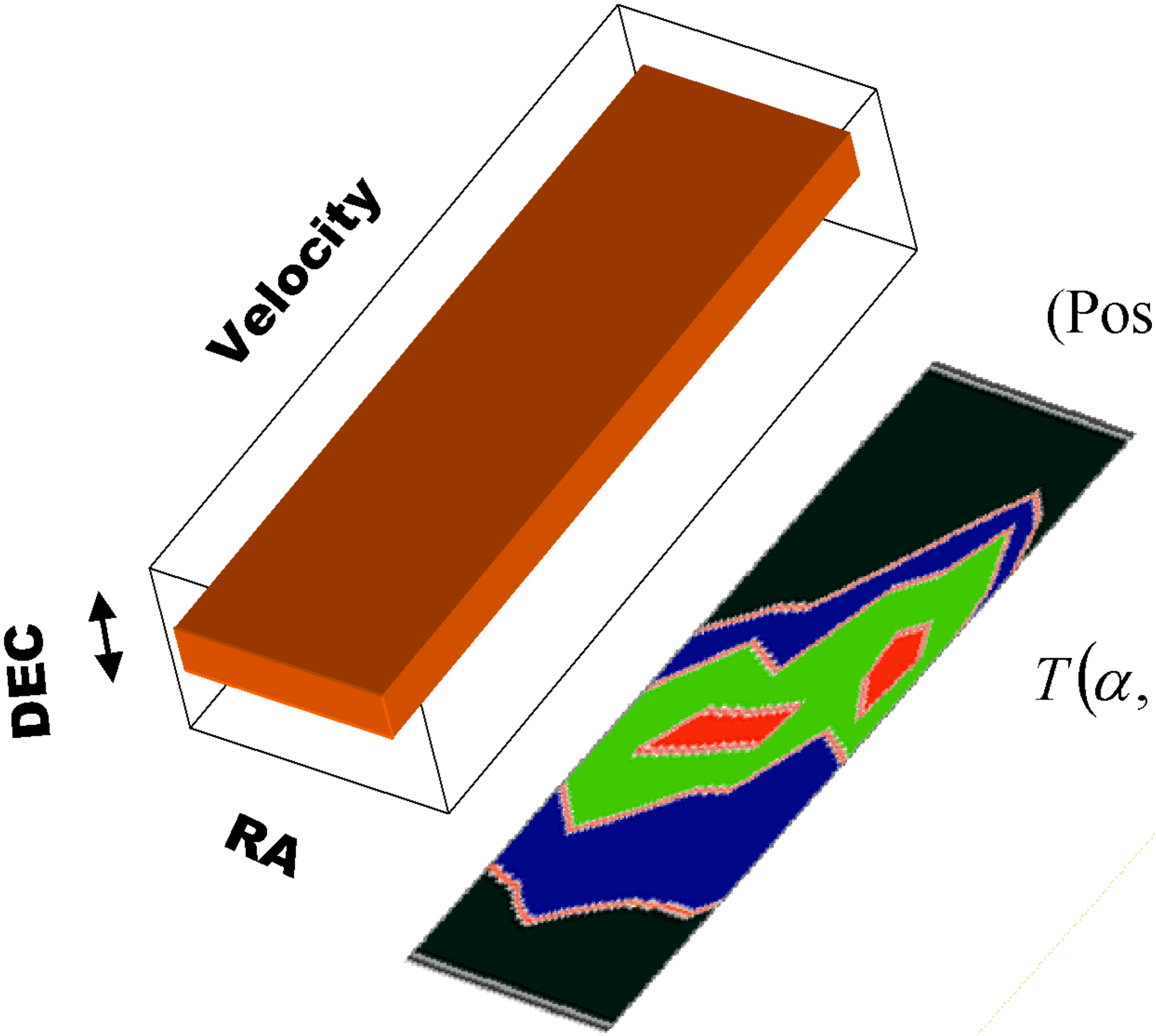


$$W(\alpha,\delta) = \sum_{V_i=V_{min}}^{V_{max}} T(\alpha,\delta,V_i)\cdot \Delta V_i$$

(If V1=V2 => Channel Map)

```
For {v=vmin} {v <=vmax} {v++} {
    if T(α,δ,v) > Tmin then
        W(α,δ)=W(α,δ)+ T(α,δ,v)
    endif
endfor
```

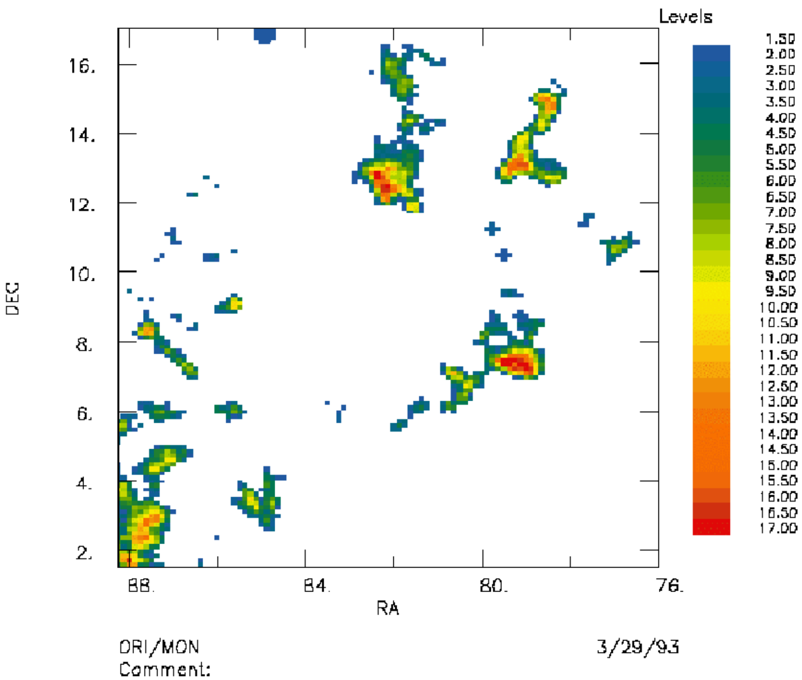
Spectral-Line Mapping
Grid and On-the-Fly

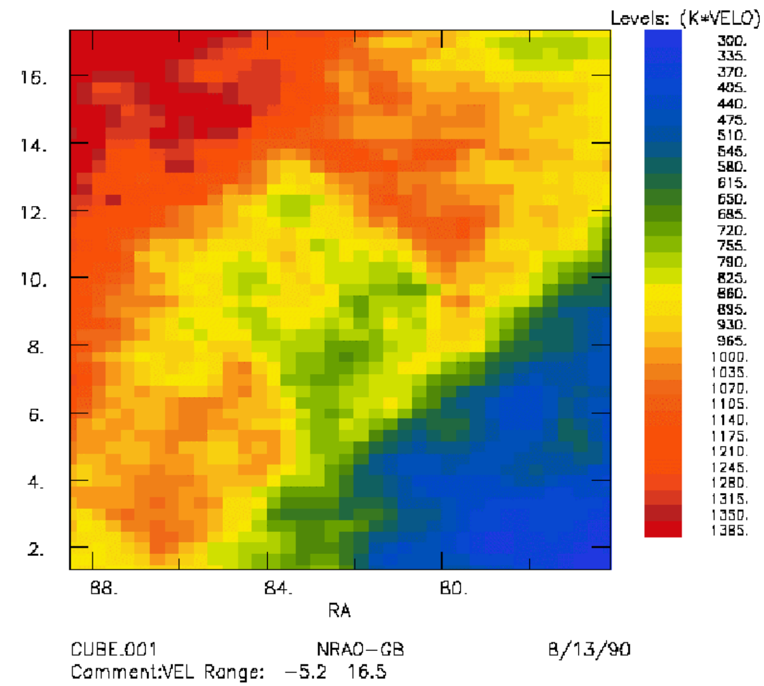


(Position-velocity map)

$$T(\alpha,V)=\sum_{\delta=\delta_{\min}}^{\delta_{\max}} T(\alpha,\delta,V)$$

Spectral-Line Mapping





The Future of Single-Dish Data Analysis

- Increase in the use of RDBMS.
- Support the analysis of archived data.
- Sophisticated visualization tools.
- Sophisticated, robust algorithms (mapping).
- Data pipelining for the general user.
- Automatic data calibration using models of the telescope.
- Algorithms that deal with data sets.
- Analysis systems supported by cross-observatory groups
- More will be done with commercial software packages