

**any questions?**

# Mapping Speed

- How long would GBT take to map half of sky to  $200 \mu\text{Jy}$  RMS, with  $T_{\text{sys}}=25\text{K}$ , 400 MHz of bandwidth at 600 MHz?
  - Beam size  $\sim 1.22\lambda/D$ , 50cm wavelength,  $\sim 20'$  beam.
  - need  $1.5\text{e}5$  beams.  $t_{\text{obs}}=(T/dt)^2/B$ ,  $dt=400\mu\text{K}$ ,  $t=10\text{s}/\text{beam}$ ,  $1.5\text{e}6 \text{ s}$  total  $\sim 20$  days.
- CHIME sensitivity after one day? Similar gain, 50K  $T_{\text{sys}}$ , 90 degree by 1 degree strip. 4 minutes to cross strip.  $dT=50/\text{sqrt}(240*400\text{e}6)=160 \mu\text{K}$ , or  $80 \mu\text{Jy}$ . Equivalent to  $\sim 100$  days of GBT(!).

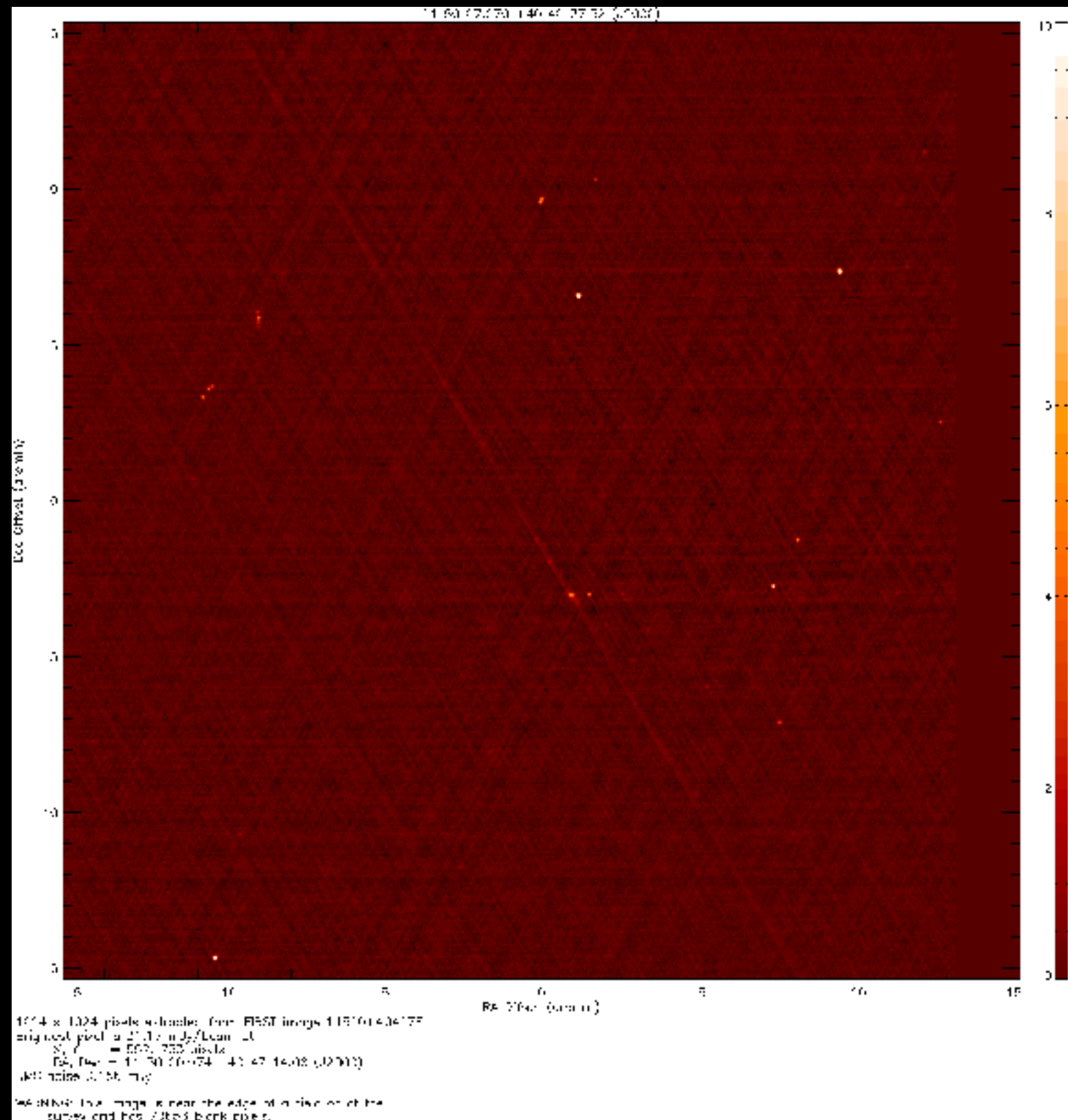
# Will This be Our Noise?

- Maybe. If we're lucky. But probably not...
- $1/f$  noise always exists in amplifiers.
- If integrate longer than  $1/f$  knee, noise gets *worse* with longer integrations.
- $1/f$  modulates  $T_{\text{sys}}$ , so often dominates after  $\sim\text{ms}$  timescales.
- Usual solution - switched receiver. Use two feedhorns going into one amplifier, switch back & forth between them. Can also switch against loads.
- How does this affect our noise?

# Will This be Our Noise?

- Maybe. If we're lucky. But probably not...
- $1/f$  noise always exists in amplifiers.
- If integrate longer than  $1/f$  knee, noise gets *worse* with longer integrations.
- $1/f$  modulates  $T_{\text{sys}}$ , so often dominates after  $\sim\text{ms}$  timescales.
- Usual solution - switched receiver. Use two feedhorns going into one amplifier, switch back & forth between them. Can also switch against loads.
- How does this affect our noise?
  - spend half time on-source, half time off-source. Then difference the two. Half as much time double the variance, subtracting two same-variance #'s doubles again, so  $\sigma$  goes up by 2,  $t_{\text{obs}}$  goes up by 4.

# Random Square Degree from FIRST



# Lots of Sources

- Our telescope will read out true sky convolved with telescope beam.
- If I point at source, but second source falls in beam, I will measure combined flux.
- Alternatively, convolve map of sky with telescope beam, what is median scatter?
- Unwanted sources contaminate signal, called *confusion noise*. Does not integrate down with time.
- How does importance of confusion noise scale with time?
- How does confusion noise scale with resolution?

# Lots of Sources

- Our telescope will read out true sky convolved with telescope beam.
- If I point at source, but second source falls in beam, I will measure combined flux.
- Alternatively, convolve map of sky with telescope beam, what is median scatter?
- Unwanted sources contaminate signal, called *confusion noise*. Does not integrate down with time.
- How does importance of confusion noise scale with time?
  - gets worse, since SNR on faint sources goes up.
- How does confusion noise scale with resolution?
  - gets better, since odds of second source in-beam go down.

# Quick GBT Confusion Limit

- FIRST survey has  $\sim 100$  sources/square degree to 1 mJy
- What is GBT beam size at 1.4 GHz?  $1.22\lambda/d \sim 10'$
- How many sources per beam?  $100 \cdot (10/60)^2 = 2-3$
- What is confusion limit?  $\sim \sqrt{2-3} \cdot (>1 \text{ mJy}) \sim \text{few mJy}$
- How long does it take to get there?  $3 \text{ mJy} = 6 \text{ mK}$ ,  $BW = 400$ ,  $T_{\text{sys}} = 25$ ,  $dT/T = 1/\sqrt{Bt}$  so  $t = (T/dT)^2/B = (25/0.006)^2/400 \times 10^6 = 0.04 \text{ s}$ .
- Say looking for 10 km/s spectral line -  $BW = 1.4 \times 10^9 \cdot (v/c) \sim 50 \text{ kHz}$ .  
 $t = (T/dT)^2/50 \times 10^3 = 6 \text{ min}$ .



# Receivers

- Radio waves come into detector.
- Something needs to measure them - usually an ADC (analog to digital converter)
- ADCs are noisy, and we may have signal loss en route through cables.
- So, usually amplify signal as soon as it comes in.
- Amplifiers usually quoted in logarithmic dB, 10 dB a factor of 10 in power. (sometimes electric field, so be careful)
- If I have 20K coming into 20 dB amplifier, how much comes out?
  - $20\text{dB} = 10^{(20/10)} = 100\times$  increase, so at 2000K.

# Receiver Noise

- If I add 10K noise before amplifier, what is my new power? If I add 10K after amplifier, what is new power?
- 3000K, 2010K respectively.
- If I had 1K signal, what is my output signal level?
- 100K, 100K.
- SNR?  $100/3000$  vs.  $100/2010$ . Adding noise after amplification didn't make much difference, but before makes huge difference.
- So, huge emphasis on noise of first amplifier in a system, and in reducing noise upstream of that. Downstream matters much less.

# Quantum Limit

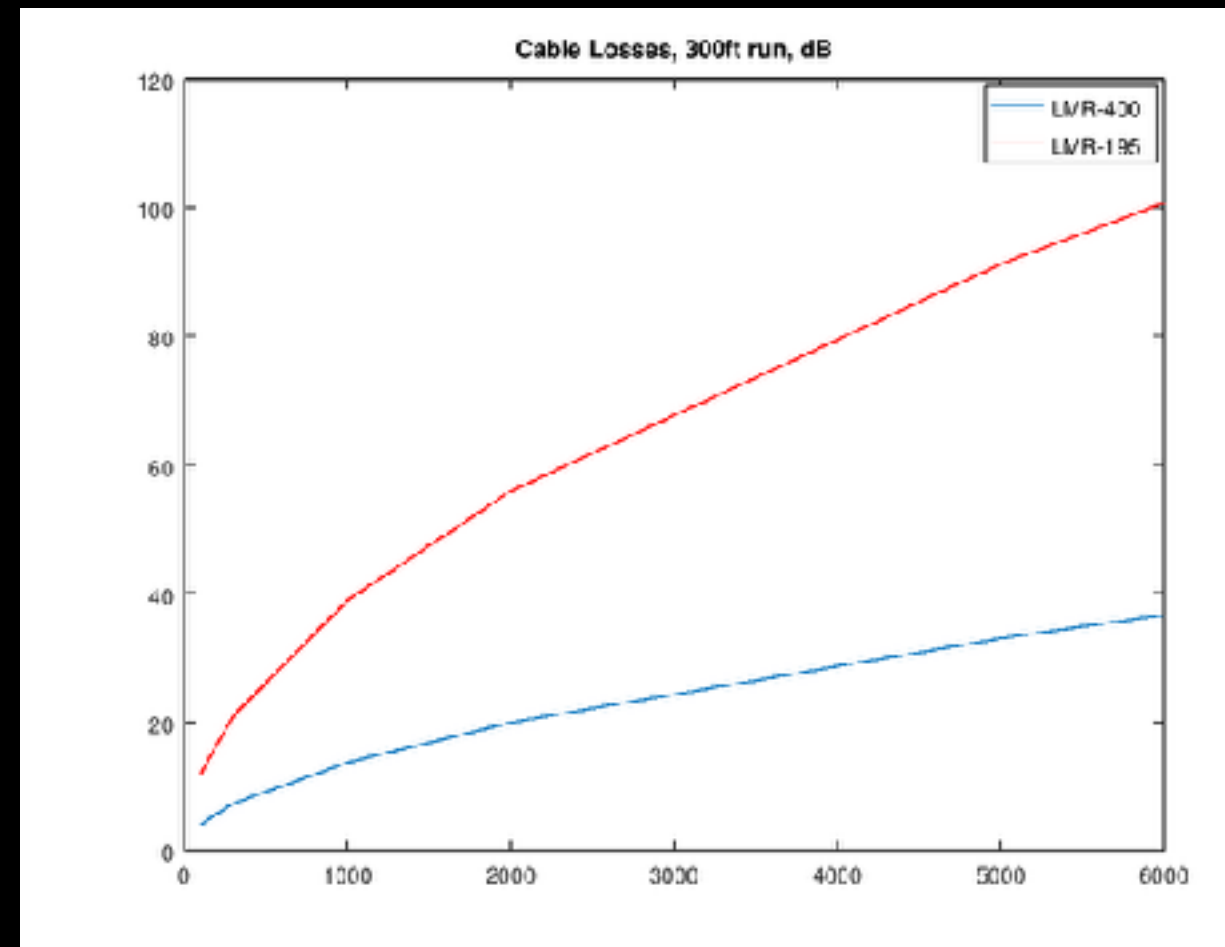
- Heisenberg uncertainty principle says  $\delta x \delta p > \hbar/2$
- But also,  $\delta E \delta t > \hbar/2$  - if I take a measurement over a short time, energy uncertainty grows (thankfully, since this is why structure exists in the universe).
- If I coherently amplify a wave (i.e. keep phase the same), time uncertainty should be less than a radian, or  $1/2\pi\nu$ .
- Means  $\delta E = k\delta T > \hbar\nu/2$ , or  $\delta T > \hbar\nu/2k$ . This is quantum limit.
- 30 GHz, limit is  $\sim 1$  K. Good amplifiers get within factor of few. Other noise wins at low frequency, but amplifier often sets noise at high frequency.
- Why do bolometers not suffer from this?

# Quantum Limit

- Heisenberg uncertainty principle says  $\delta x \delta p > \hbar/2$
- But also,  $\delta E \delta t > \hbar/2$  - if I take a measurement over a short time, energy uncertainty grows (thankfully, since this is why structure exists in the universe).
- If I coherently amplify a wave (i.e. keep phase the same), time uncertainty should be less than a radian, or  $1/2\pi\nu$ .
- Means  $\delta E = k\delta T > \hbar\nu/2$ , or  $\delta T > \hbar\nu/2k$ . This is quantum limit.
- 30 GHz, limit is  $\sim 1$  K. Good amplifiers get within factor of few. Other noise wins at low frequency, but amplifier often sets noise at high frequency.
- Why do bolometer not suffer from this?
  - Not trying to amplify phase, so arrival time uncertainty can be much larger.

# Mixing

- We'd like to be able to observe at any frequency.
- ADCs will only work at some frequency. Usually (but not always) lower than frequencies we care about.
- Sending high frequencies along cables is also challenging. Cable loss grows quickly with frequency.
- Getting signal down from GBT at 5 GHz could lose 999,999,999 out of every billion photons(!)

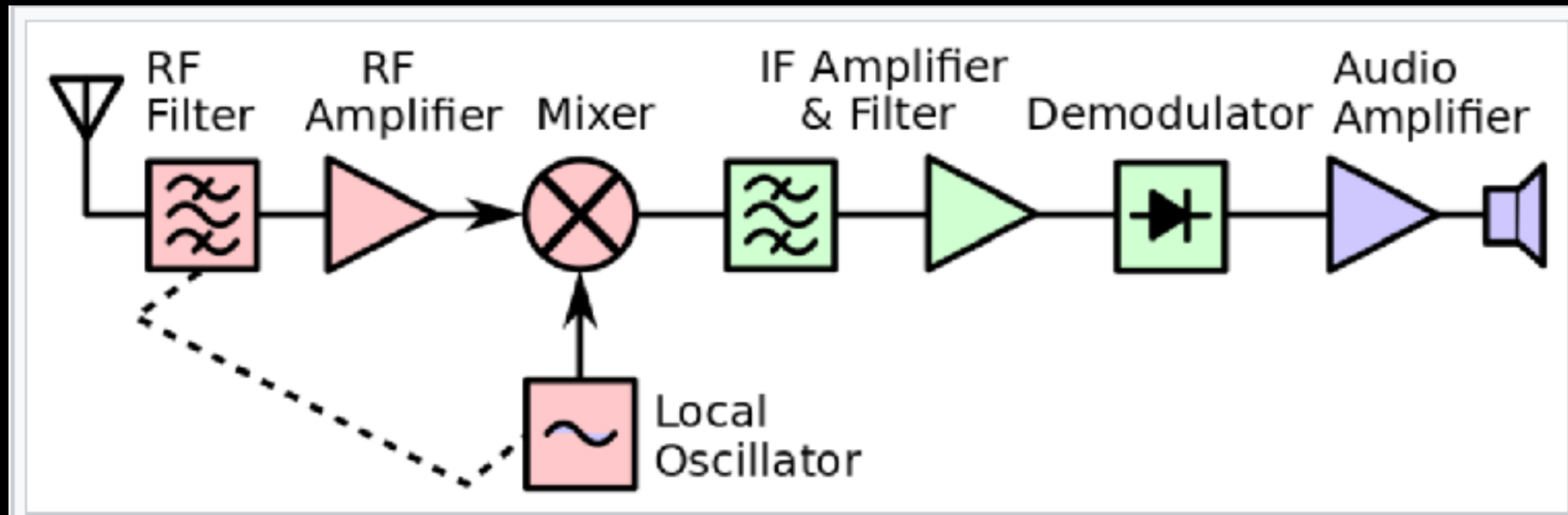


Cable losses from handy-dandy  
Times Microwave online calculator

# Mixing V2

- Can we do something about this?
- Yes! If I have a nonlinear component, and put in two signals  $V_1, V_2$ , output will be  $c_1(V_1+V_2)+c_2(V_1+V_2)^2+\dots$
- Second term gets a  $V_1V_2$  component. What does this look like if  $V_1$  and  $V_2$  are sine waves closely spaced in frequency?
- $V_1=a\sin(2\pi v_1t)$ ,  $V_2=b\sin(2\pi v_2t)$ . Angle summation formulas give  $\sin(2\pi(v_1-v_2)t)+\sin(2\pi(v_1+v_2)t)$
- We can filter out the  $v_1+v_2$  term with analog device, leaving  $v_1-v_2$ . We have shifted the signal to lower frequency.
- Process is called heterodyning (developed by Canadian Reginald Fessenden), and nonlinear device is called a mixer.
- Good mixers only put out one frequency combination, but others can put out various combinations of  $v_1, v_2$ , called intermodulation products. These are bad.
- NB - both  $v+dv$  and  $v-dv$  will generally get mixed in - these are the *sidebands*. Can make mixers that separate them out, can filter one out beforehand, or can just eat them.

# Super Heterodyne Receiver



- Signal comes in. Gets amplified/filtered (often amplified before the filter. Why? half dB loss typical, equals 30K if warm)
- Separate signal from local oscillator gets piped into mixer, to shift to lower intermediate frequency (IF).
- IF much easier to move around. Usually goes into another thing that does the detecting, often gets mixed again.

# CHIME

- Instead, we could sample directly if ADC fast enough.
- CHIME works 400-800 MHz. Normally need to run at 1600 Msamp/sec (why? Nyquist...)
- BUT - if we analog filter everything outside of band, then we could run at 800 Msamp/s. We would alias low-frequency power, but that is gone.
- This is called working in the second Nyquist zone.
- Other low-frequency telescopes (e.g. PAPER, HERA) directly sample electric field w/out any tricks at all.



# Interferometry

- Resolution from single dishes not great. But we can do better!
- If I have two telescopes separated by distance  $d$ , observing source at angle  $\theta$  at wavelength  $\lambda$ , what is phase difference in arriving electric field?
- Source in far field (probably) so path length is  $d\sin(\theta)$ . Phase difference is  $2\pi d/\lambda \sin(\theta)$ .
- If I measure electric fields and multiply together  $\langle E_1 E_2^* \rangle$ , I will get  $E_0^2 \exp(2\pi i d/\lambda \sin(\theta))$
- Call  $d/\lambda$   $b$ , the dish separation in wavelengths, and assume small angles on the sky, and  $I = E_0^2$ . Then we get  $I \exp(2\pi i b \theta)$

# Signal over Sky

- If I have many sources, I have to integrate over the (2D) sky brightness. Gives:  $\langle E_1 E_2^* \rangle = \int I(\theta) \exp(2\pi i b \cdot \theta) d^2\theta$
- Have you seen this before?
- Yes! This is just one component of the FT of the sky, measured at wavevector  $b$ .
- This time-average is the fundamental output of a radio array, called a *visibility*.
- Resolution set by *baseline length*, not by dish size.

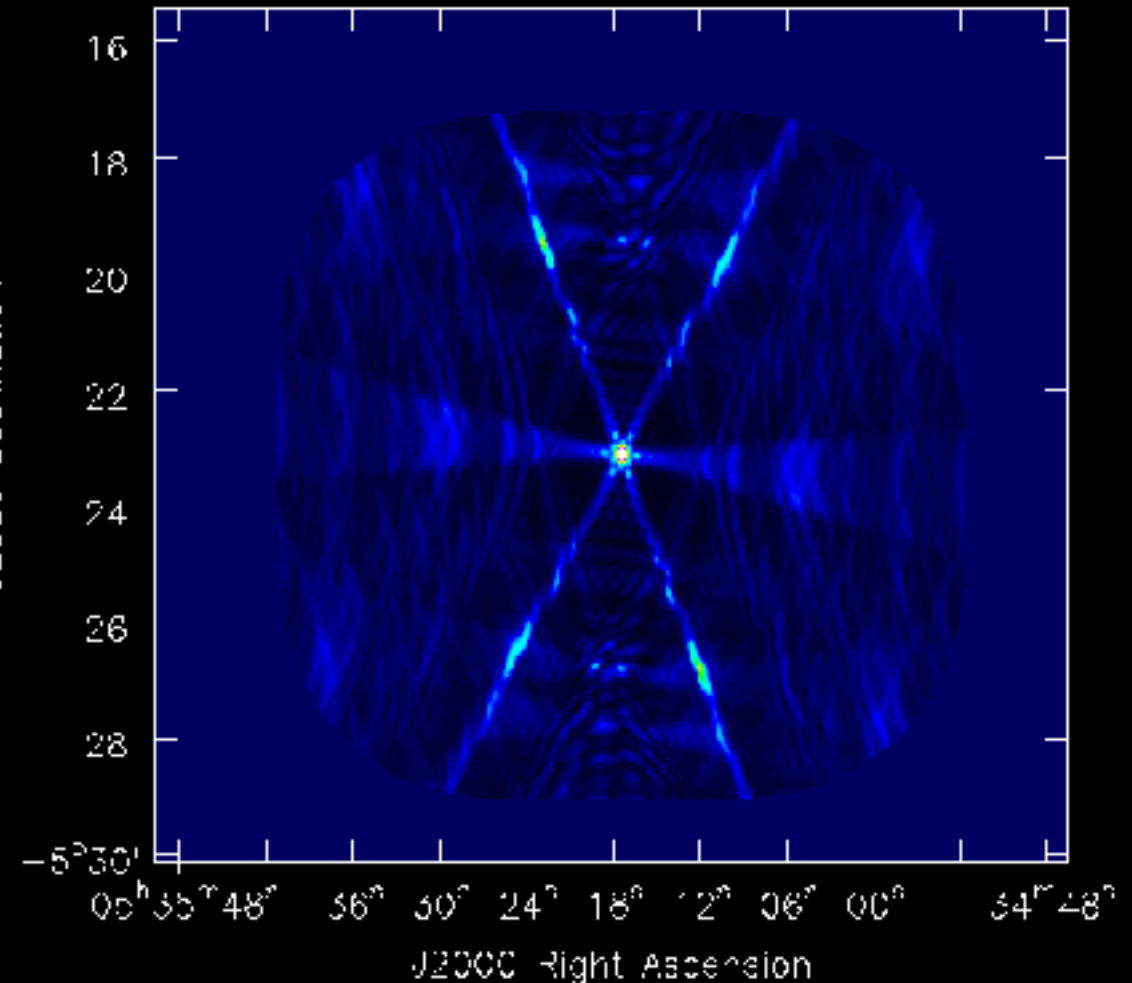
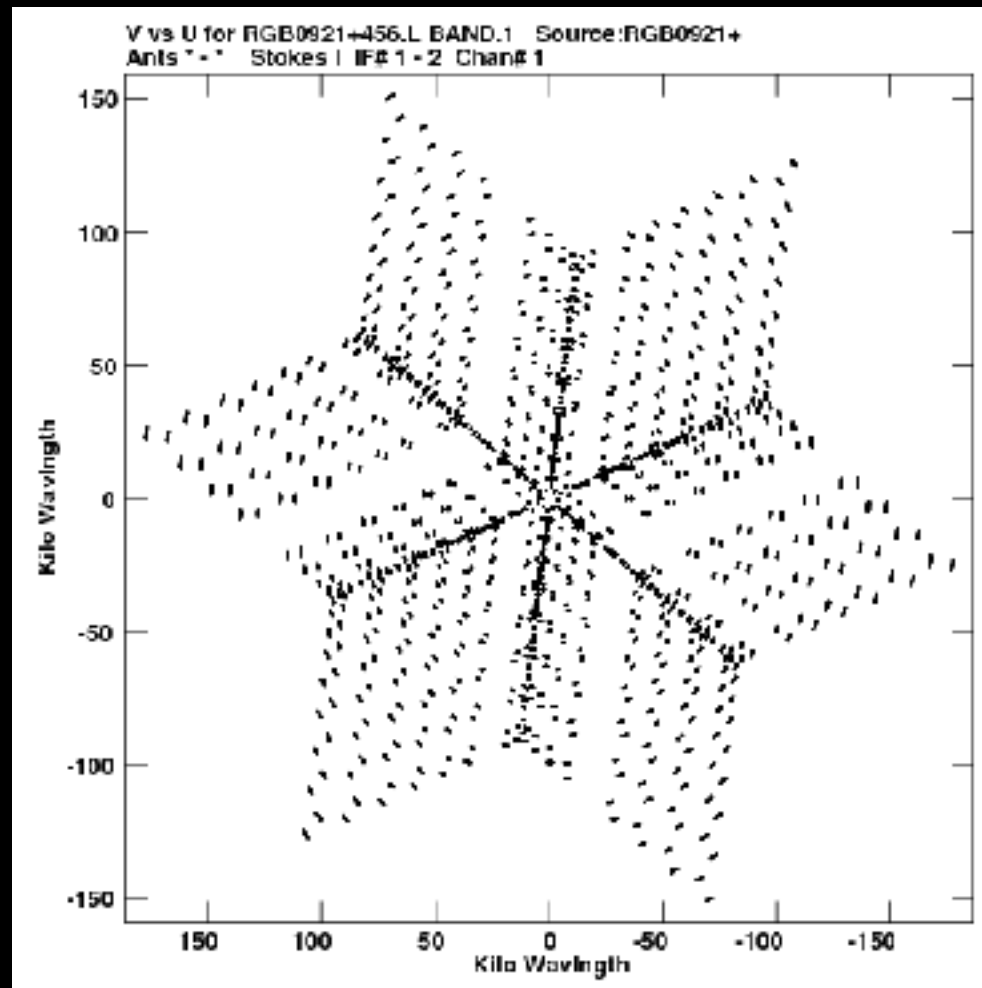
# Interferometer

- An array of radio telescopes working together is called an interferometer. Every pair of antennas (called a *baseline*) gets multiplied together in the *correlator*, producing a visibility.
- Vector separation labelled by  $(u,v,w)$  where  $u$  is EW,  $v$  is NS (and  $w$  is vertical separation, generally much less important).
- UV coverage is sampling of points in UV plane - denser sampling means better understanding of sky.
- Telescope PSF given by FT of UV coverage - which may not surprise you given FT optics result from Monday. Every antenna acts as a radiator.

# How many operations to correlate?

- Correlation used to be done by analog circuits. Almost uniformly no longer the case.
- VLA - 27 dual pol receivers, 8 GHz bandwidth (total - they split up). 54 inputs,  $n_{\text{pair}} = 54 \times 55 / 2 \sim 1500$  \*  $8 \times 10^9 \times 8$  ops/complex multiply+add = 100 TFlops.
- CHIME - 1024 dual pol, 400 MHz - 6.7 PFlops. Lots of crunching needed for large arrays.

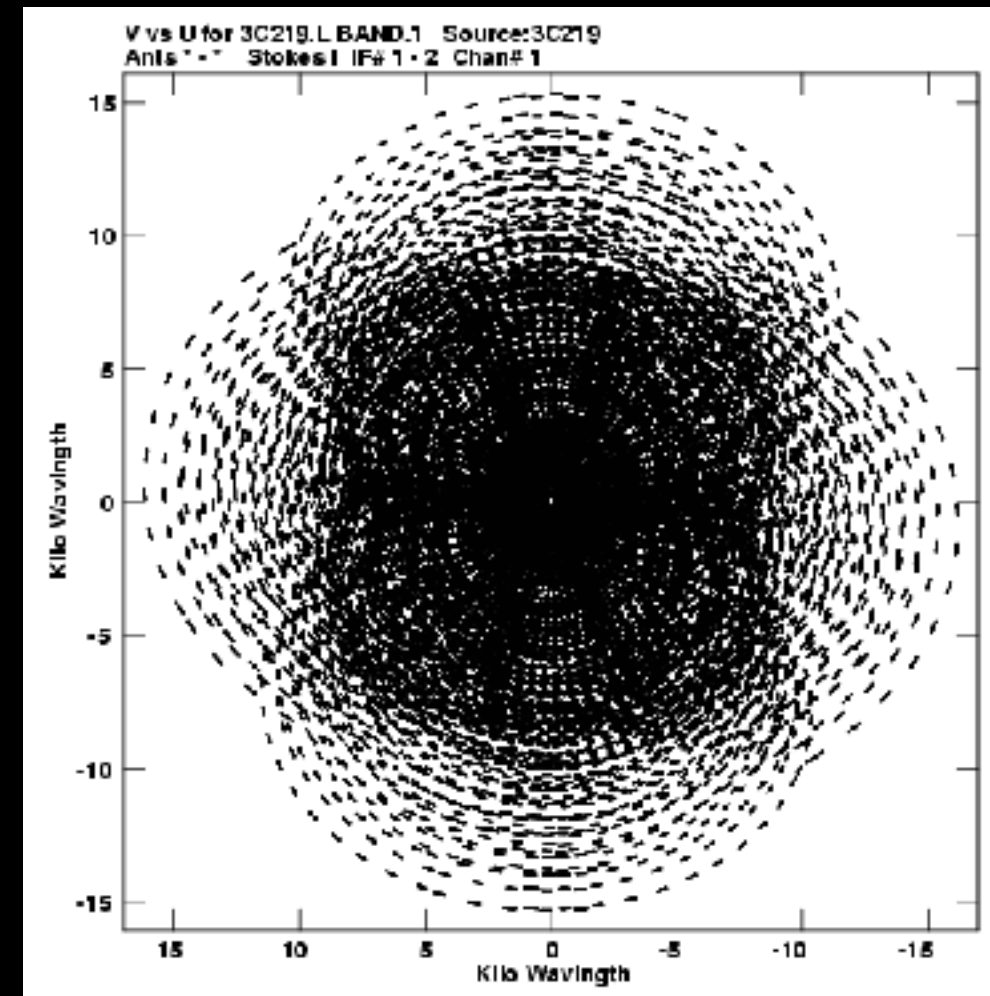
# VLA UV Coverage



Left: Instantaneous UV coverage from VLA  
Right:

# Rotation

- We want as much UV coverage as possible, right?
- Instantaneous snapshots usually not great.
- However, we are on a rotating platform. Wait long enough, and baseline moves relative to source.
- UV coordinate set by baseline separation perpendicular to field center. Fills things out enormously.
- Multiple channels also samples UV plane radially.



# Primary Beam

- Dishes have a response on the sky, usually limits field of view
- In interferometer, each dish sees primary *beam times* sky, so visibility is  $\int A(\theta) I(\theta) \exp(2\pi i \mathbf{u} \cdot \theta) d^2\theta$  where  $A$  is the primary beam (in power)
- In UV space, we measure not the sky transform, but the sky transform convolved with the primary beam transform. PB sets UV-space resolution
- PB is electric field response squared, so in UV space, that's electric field response convolved with itself. But, electric field response is just aperture illumination, so PB transform is just dish+feed autocorrelation with itself in wavelengths.

# Imaging

- If we build telescopes, we usually want to take pictures of things...
- If I measure FT of sky, I can just IFT to get map, right?
- Well... There are usually gaps in UV coverage. If baselines sampled more densely than dish diameter, this might work.
- Not possible for high-resolution. To make a map of sky, we must fill in UV plane with some guess.
- Unobserved parts of UV could be anything! The art of imaging is sensibly filling in those unobserved areas.
- Remember - we understand very well how to predict data/measure  $\chi^2$  given a map of the sky. It's only inverse problem that is ill define.



# CLEAN

- One standard way to do this is CLEAN. Pretend the sky is full of point sources.
- Make a dirty map - direct FT of visibilities. Requires putting visibilities on a grid - how fine does that grid need to be?
- Look for brightest peak (or peaks).
- Subtract from data. Repeat.
- Experience has shown that subtracting fraction of peak brightness works better in practice - say 30-50% of peak.
- When should we stop?

# CLEAN

- One standard way to do this is CLEAN. Pretend the sky is full of point sources.
- Make a dirty map - direct FT of visibilities. Requires putting visibilities on a grid - how fine does that grid need to be?
  - UV resolution set by dish diameter in wavelengths. Need to have grid be significantly finer than that.
- Look for brightest peak (or peaks).
- Subtract from data. Repeat.
- Experience has shown that subtracting fraction of peak brightness works better in practice - say 30-50% of peak.
- When should we stop?
  - We know what noise in dirty map should be - weight is just sum of visibility weights, so stop when noise is (roughly) equal to that.

# CLEAN CTD.

- Process of cleaning gives us a list of source fluxes/positions, plus dirty map that may be noise.
- Usual reported thing is to put sources into map, convolve with Gaussian with same size as main dirty beam, add in noise map.
- Significant freedom available in how we weight data. Often density in UV space higher at short baselines.
- Least-squares (“natural”) weighting may not give “best” images since weight across UV plane uneven.
- For pretty pictures, you may want to use “uniform” weight, where each *area* of UV plane gets same weight.

# Bayesian

- General imaging problem is usually underconstrained. We have to make up data to make an image.
- Amongst all the possible maps that agree with the data, *you* need to decide which one you think makes sense.
- Which brings us to the Reverend Bayes.  $P(a|b)P(b)=P(b|a)P(a)$ .  $P(m|d)=P(d|m)P(m)/P(d)$  for data  $d$  and map  $m$ .
- Drop  $P(d)$  because we already have that.  $P(d|m)$  is straightforward to calculate. Effectively, mapping problem is deciding on  $P(m|d)$
- There are many ways to do this - multiscale clean where you fit Gaussians instead of sources, maximum entropy... Do think about what you're looking at as it will guide choice of imaging.