This PDF version of the talk is missing animations, movies, etc. For a full version of this talk, go to www.joshdillon.net/talks

# Chasing the Cosmic Dawn with 21 cm Tomography

Josh Dillon, MIT

And, of course...

- Max Tegmark
- Jacqueline Hewitt
- Adrian Liu
- Aaron Ewall-Wice

- Jeff Zheng
- Chris Williams
- Jonathan Pober
- Andrei Mesinger

- Abraham Neben
- Miguel Morales
- Aaron Parsons

- **MWA**: Edward Morgan, Alan Levine, Steven Tingay, Adam Beardsley, Gianni Bernardi, Judd Bowman, Frank Briggs, Roger Cappallo, David Emrich, Bryna Hazelton, Daniel Mitchell, Divya Oberoi, Thiagaraj Prabu, Ian Sullivan, Randall Wayth, Rachel Webster

- **MITEoR**: Victor Buza, Hrant Gharibyan, Jack Hickish, Eben Kunz, Jon Losh, Andrew Lutomirski, Scott Morrison, Sruthi Narayanan, Ashley Perko, Devon Rosner, Nevada Sanchez, Katelin Schutz, Shana Tribiano, Matias Zaldarriaga, Kristian Zarb Adami, Ioana Zelko, Kevin Zheng, Richard Armstrong, Richard Bradley, Matthew Dexter, Alessio Magro, Michael Matejek, Edward Morgan, Qinxuan Pan, Courtney Peterson, Meng Su, Joel Villasenor, Hung-I Yang, Yan Zhu

- **HERA**: James Aguirre, Judd Bowman, Richard Bradley, Chris Carilli, David DeBoer, Daniel Jacobs, Matthew McQuinn, Dan Werthimer
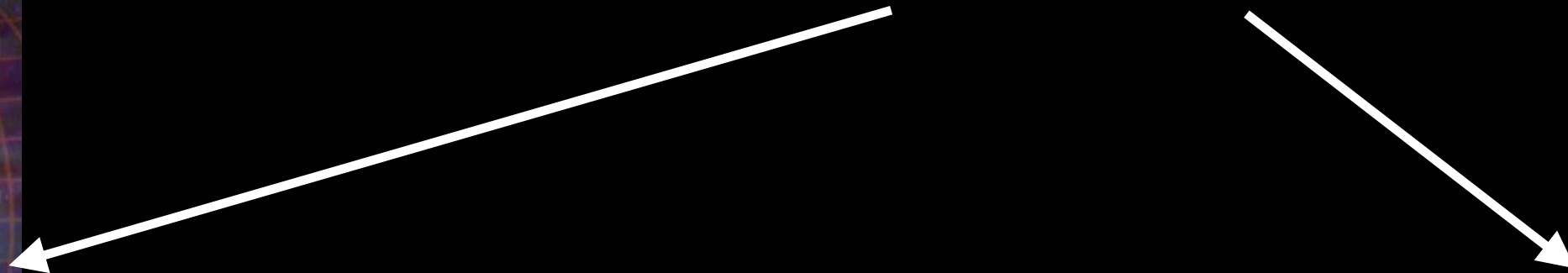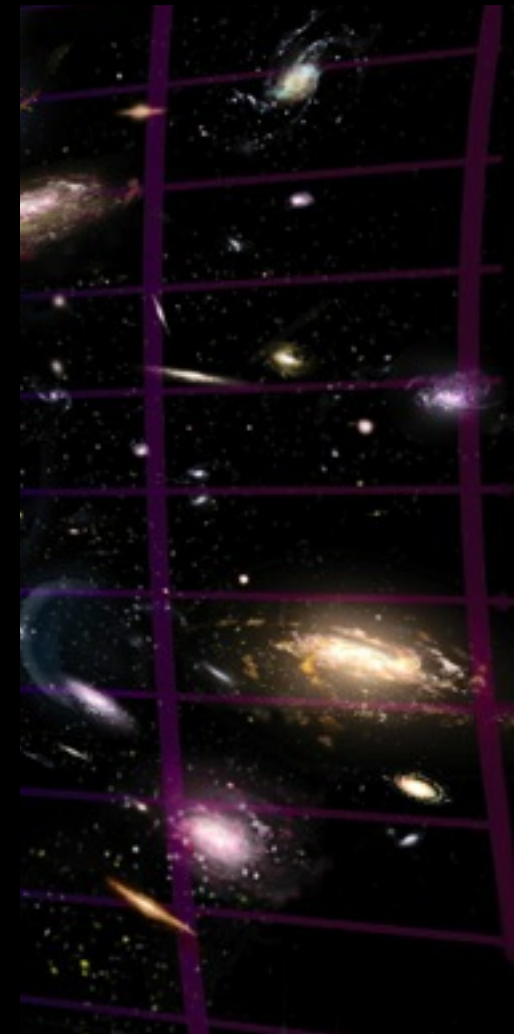
What is the "Cosmic Dawn"?

CMB
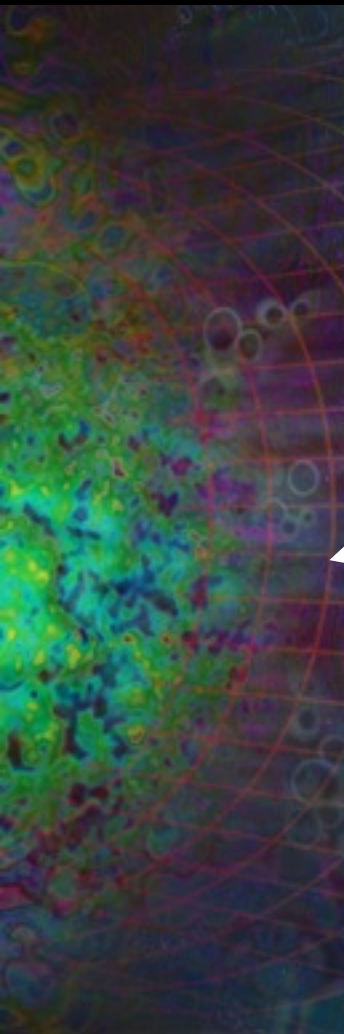
Observationally, we have few constraints on how we got from here to here.

Modern Galaxies

z = 1100

z < 6

Here's what we think...

Dark Ages   First Black Holes

The Cosmic Dawn

First Stars   The Epoch of Reionization

z =1100   z < 6

*Image: Avi Loeb & Scientific American*

And there's still a lot of open questions.

- What did the first stars, galaxies, and black holes look like and how did they form?

- What was the thermal and ionization history of the IGM and what determined it?

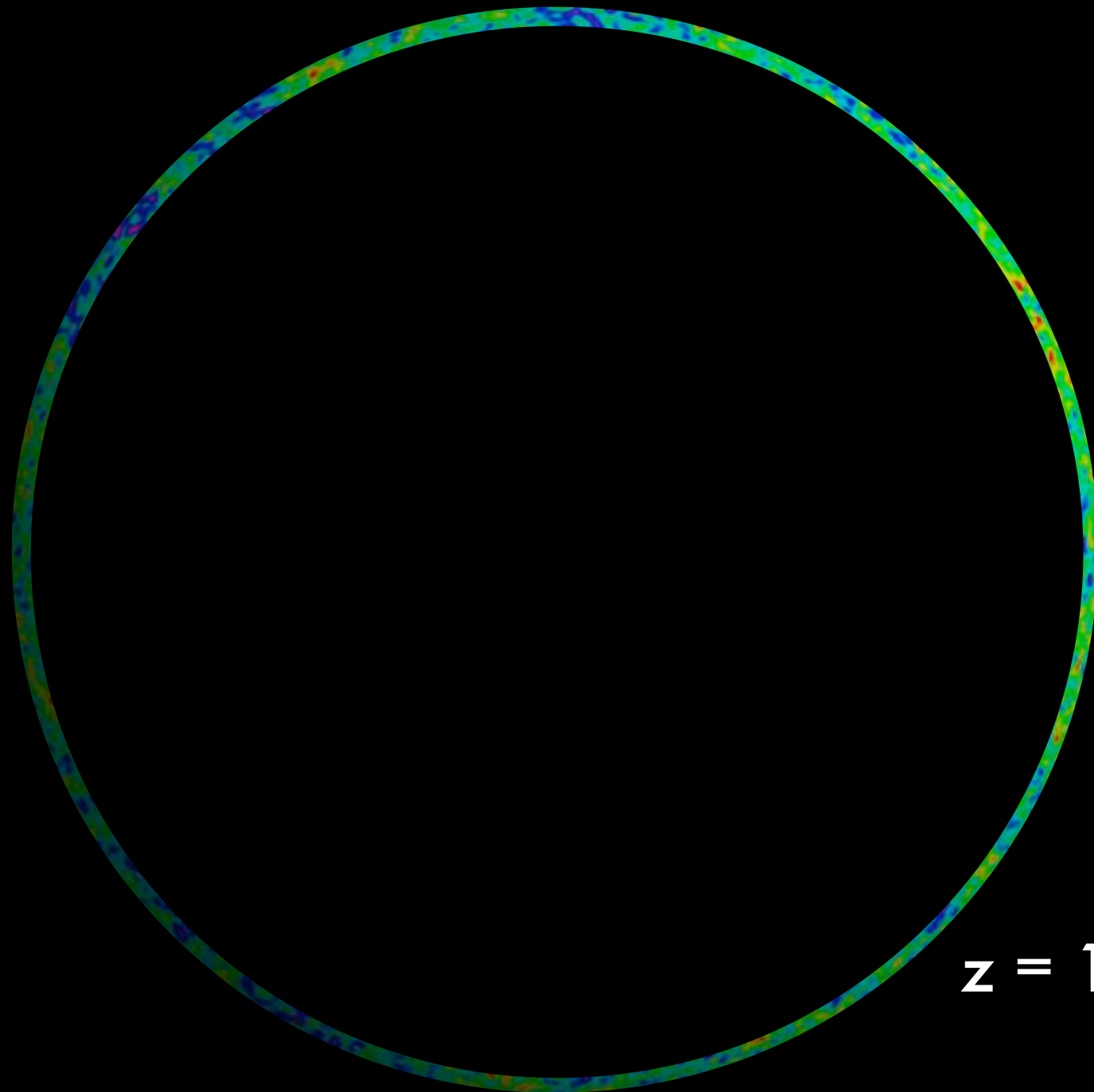- Can we measure the matter power spectrum during this epoch and test ΛCDM?

Image: Avi Loeb &

How can we observe the Cosmic Dawn?
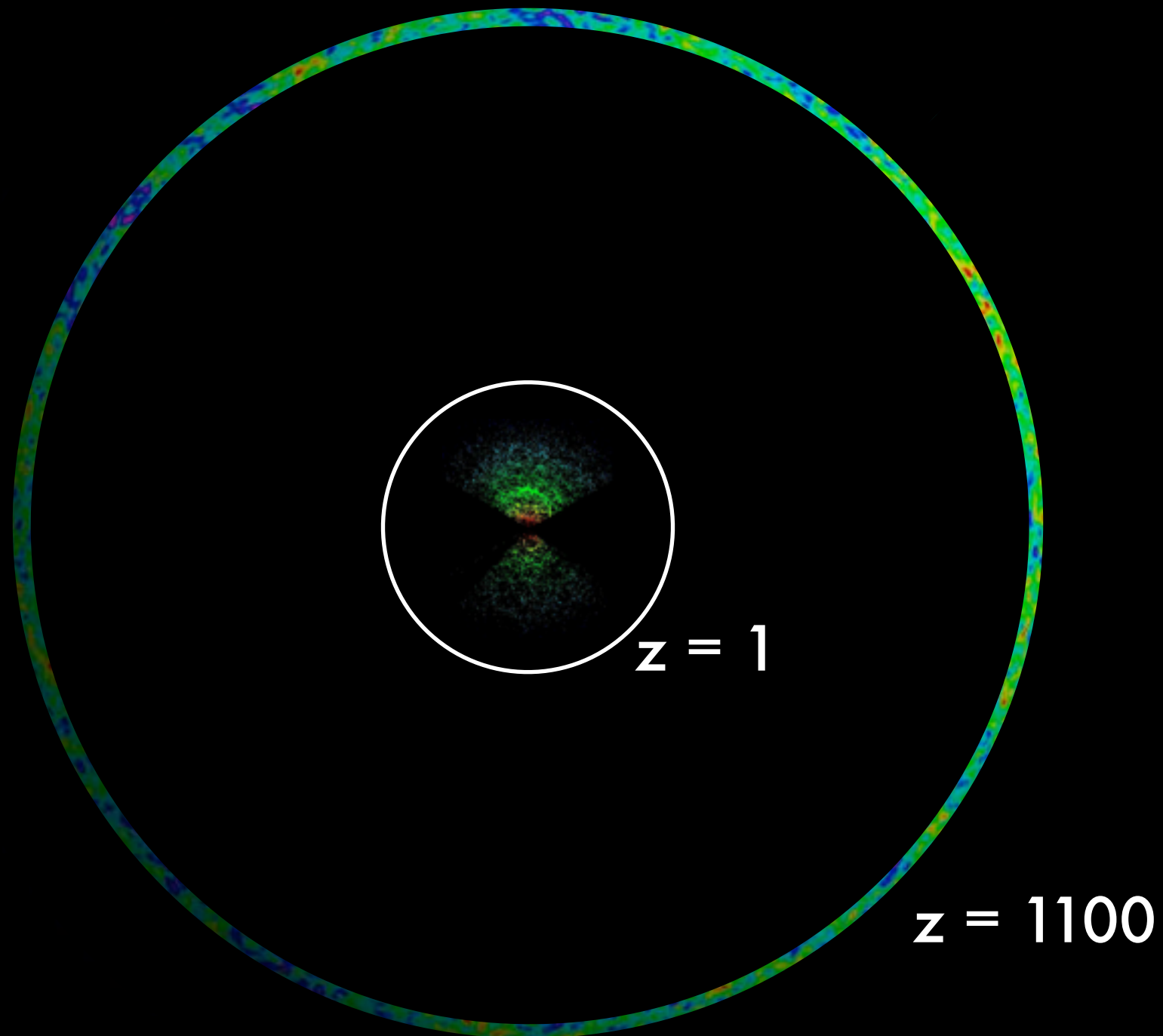
With the CMB...
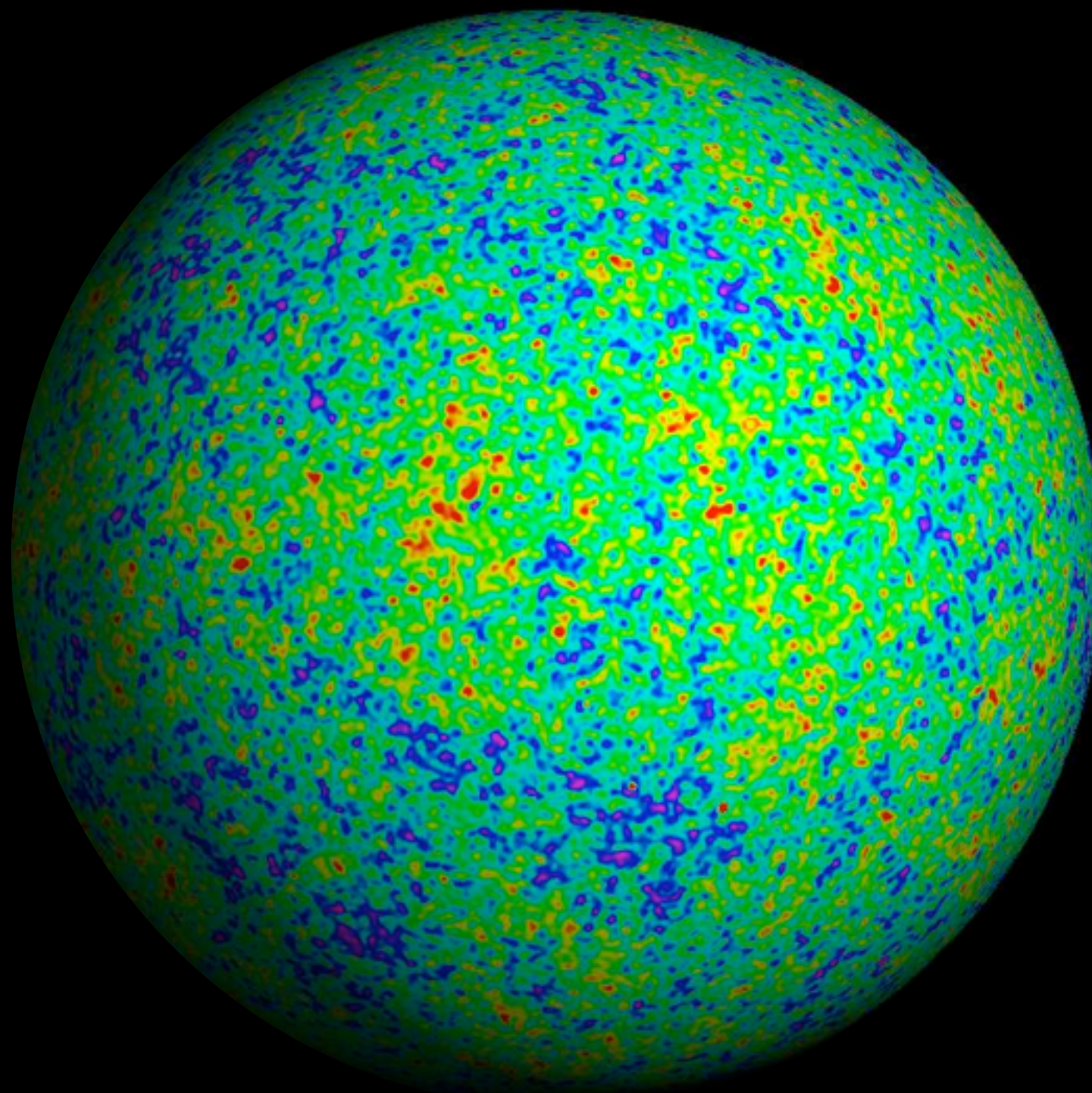
...we only get a thin shell at high redshift.

z = 1100

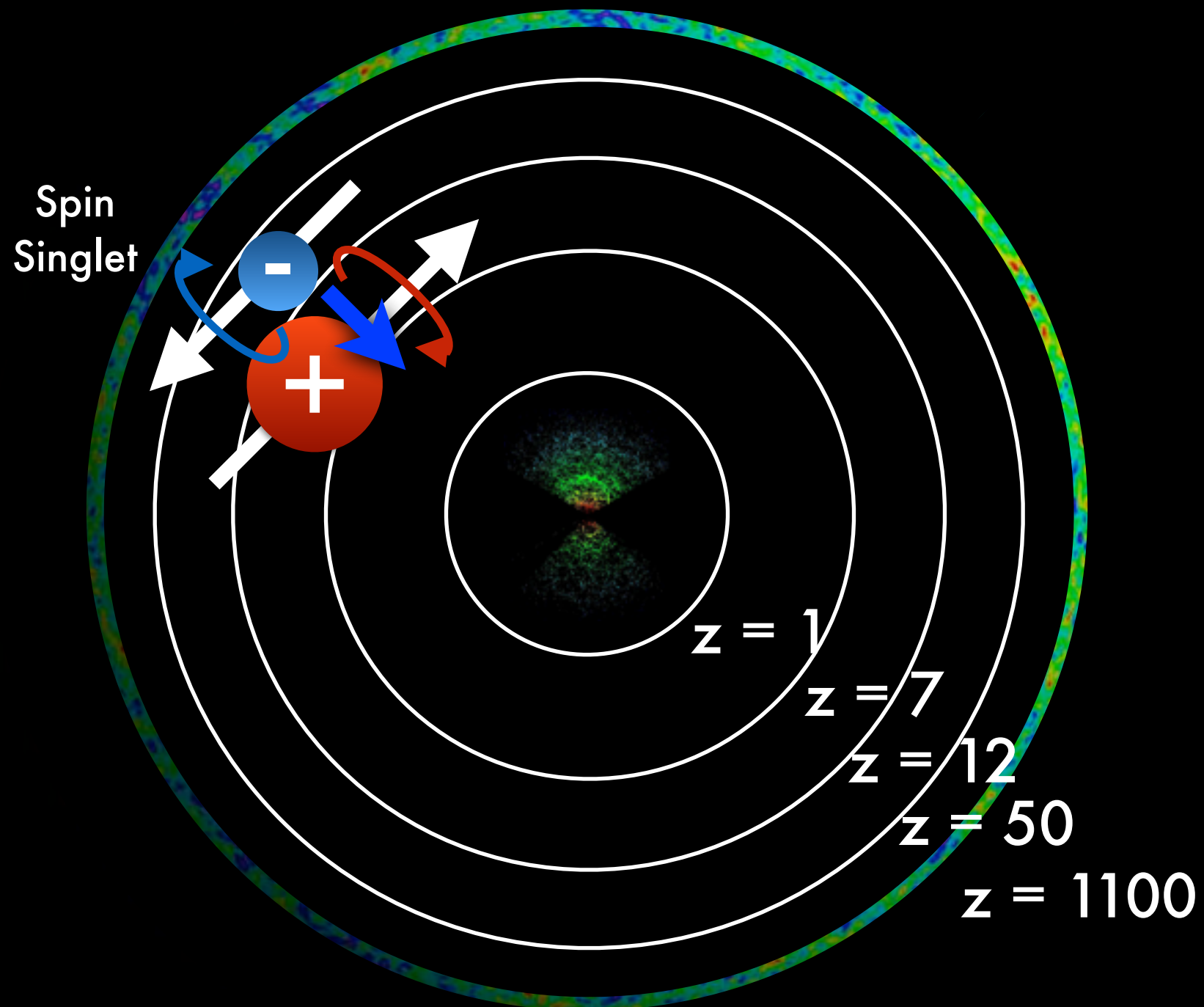# Galaxy surveys only tell us about the local universe.
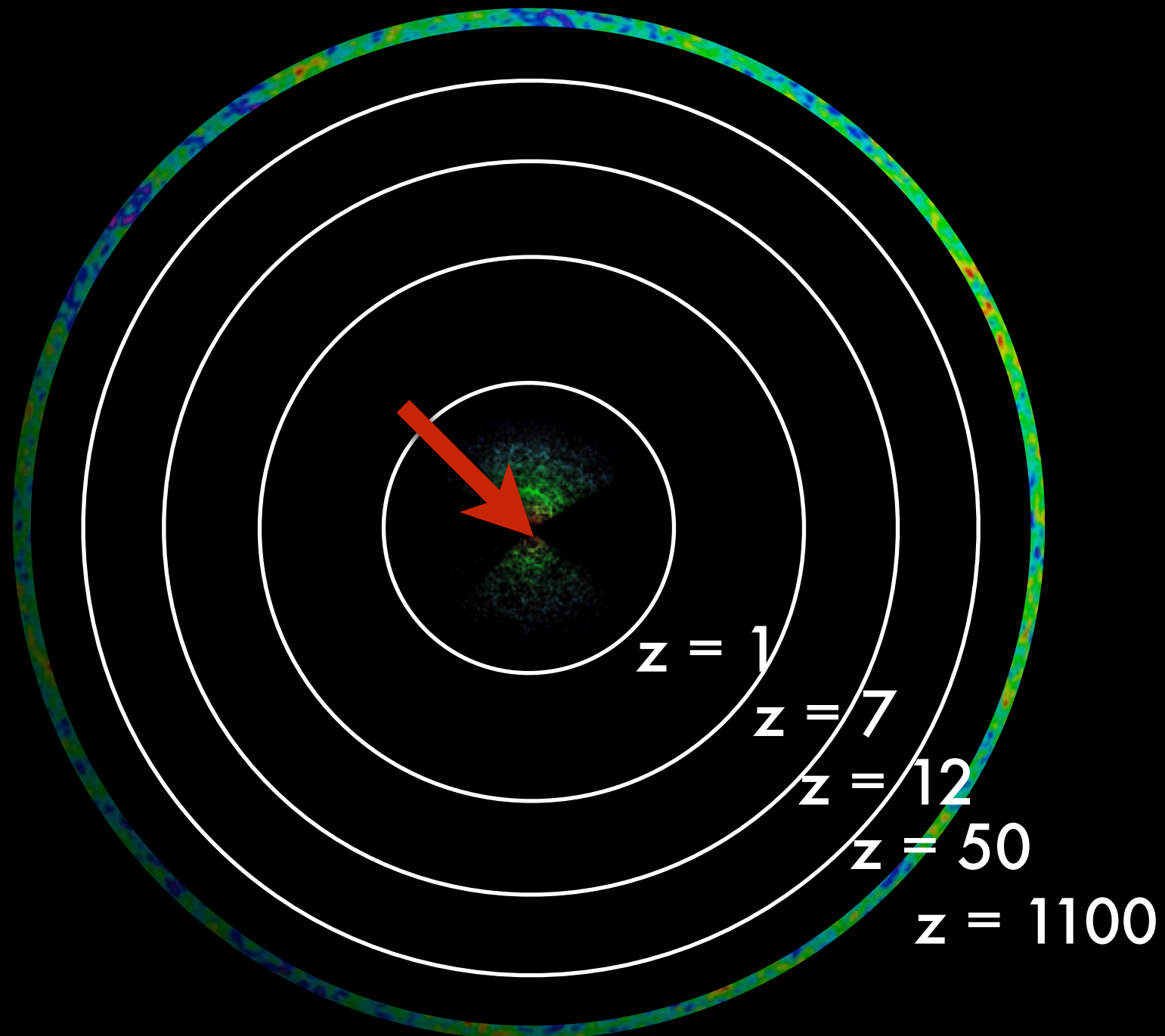

z = 1

z = 1100

So we turn to 21 cm Tomography.

Spin
Triplet

z = 1
z = 7
z = 12
z = 50
z = 1100
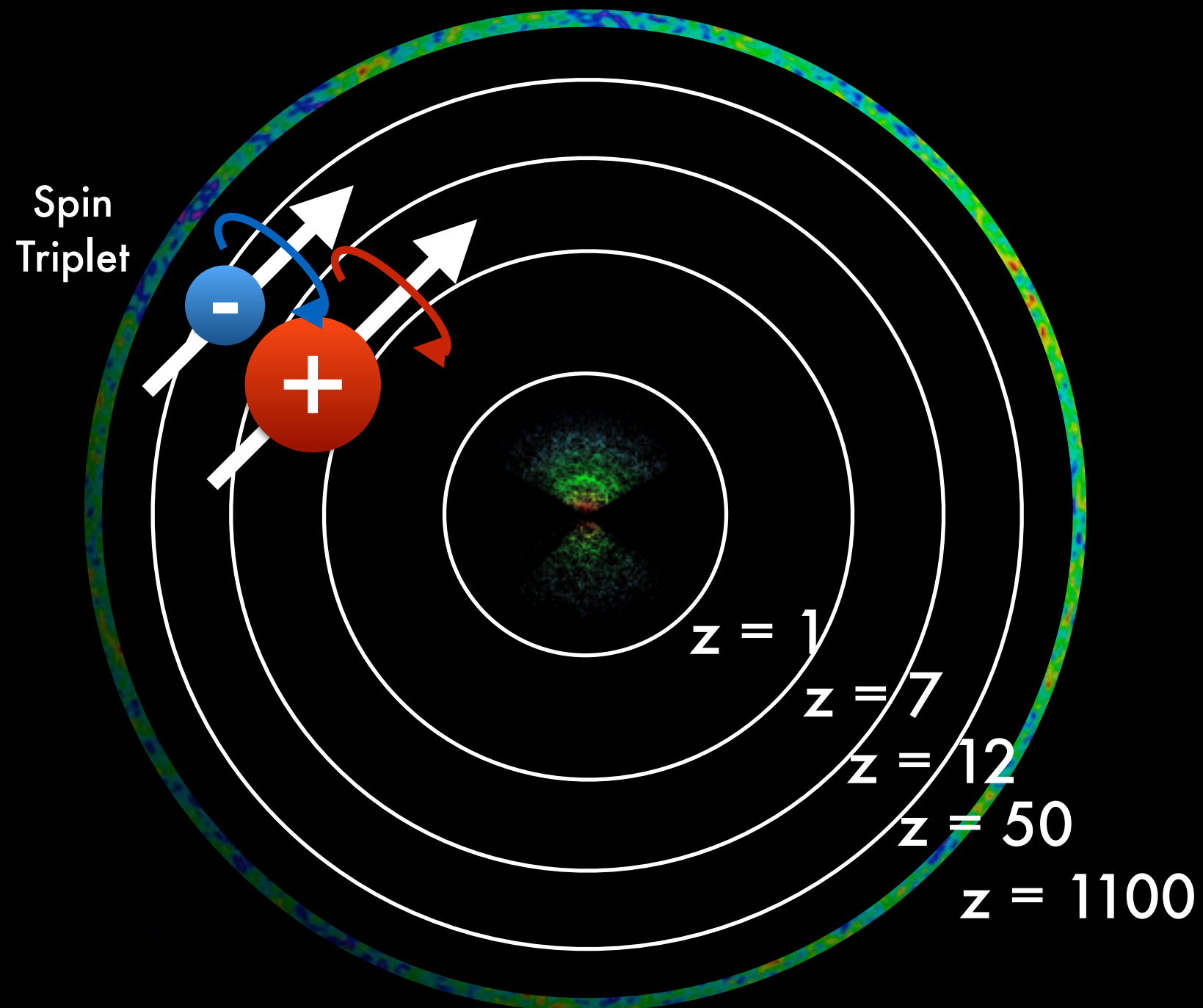
So we turn to 21 cm Tomography.
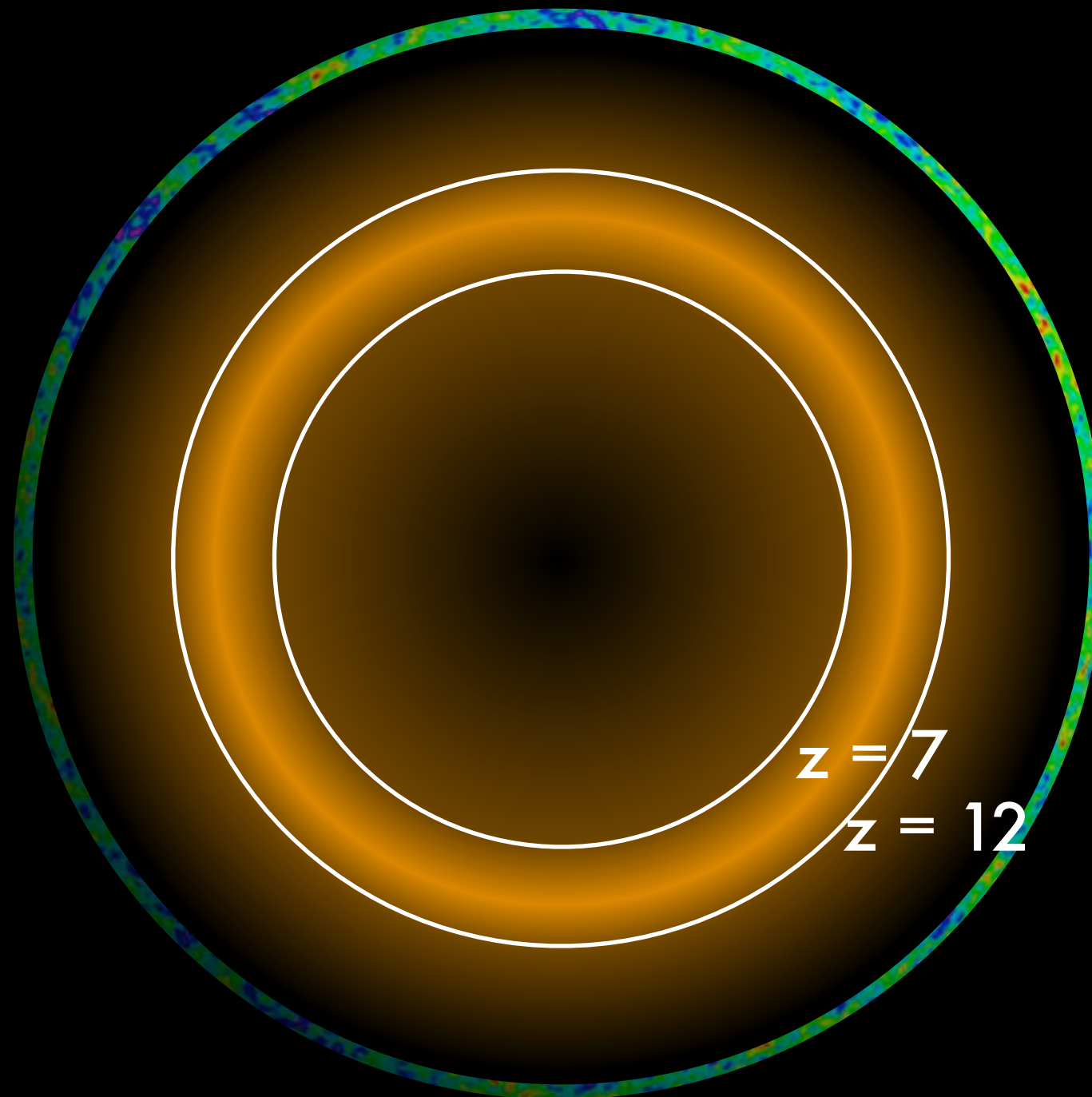
So we turn to 21 cm Tomography.

z = 1
z = 7
z = 12
z = 50
z = 1100

A huge volume of the universe can be explored with 21 cm tomography.
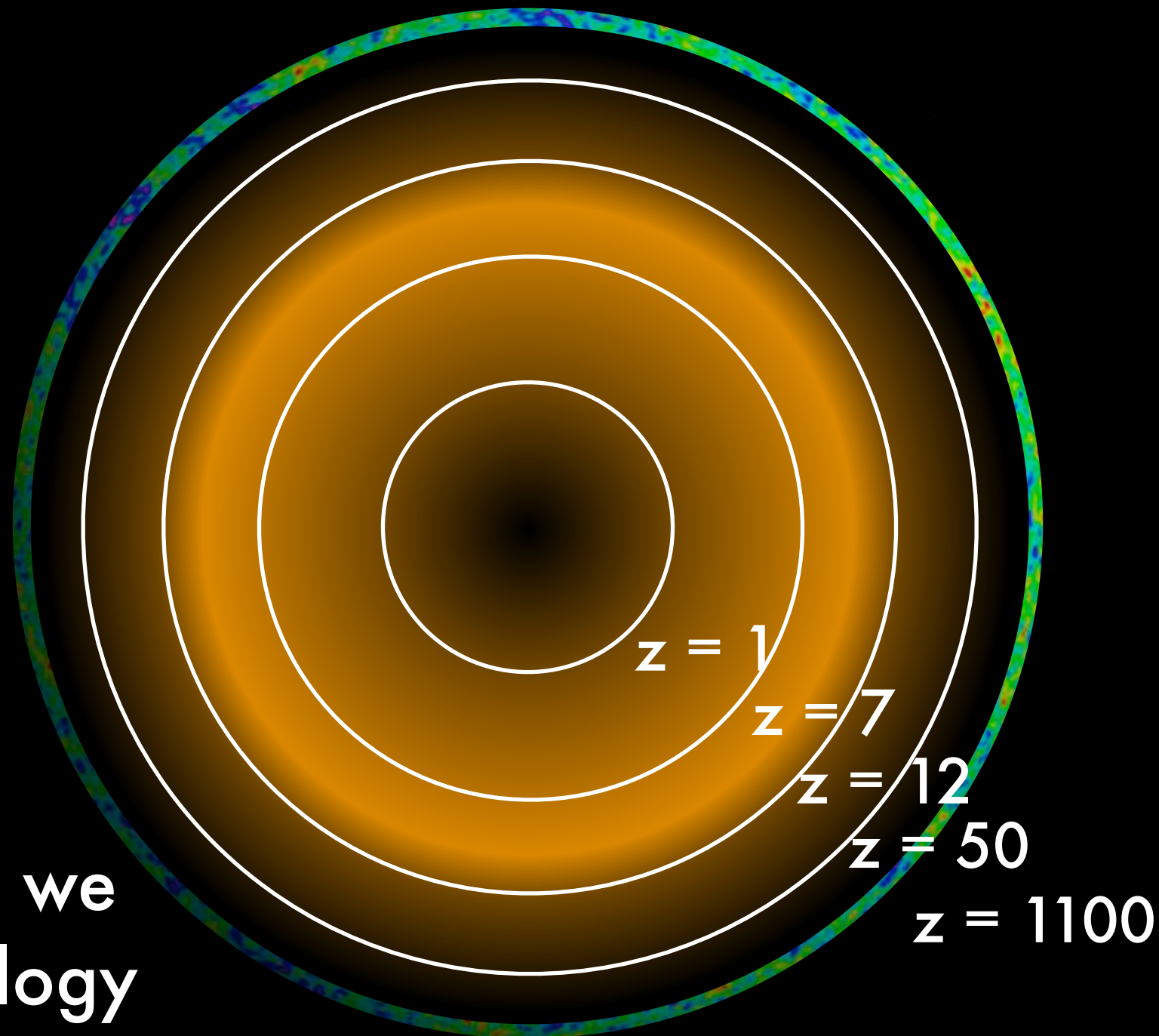
Which means we can do cosmology very precisely!

z = 1
z = 7
z = 12
z = 50
z = 1100

Our first target will be the "Epoch of Reionization"

z = 7
z = 12

Epoch of Reionization

Marcelo Alvarez, Ralf Kaehler, Tom Abel

The first detection will be statistical.

The 21 cm Power Spectrum

$\Delta_{21}(k)$ (mK)

$k(h\ \mathrm{Mpc}^{-1})$

10%
30%
50%
70%
90%
98%

Barkana (2009), Morales & Wyithe (2010)
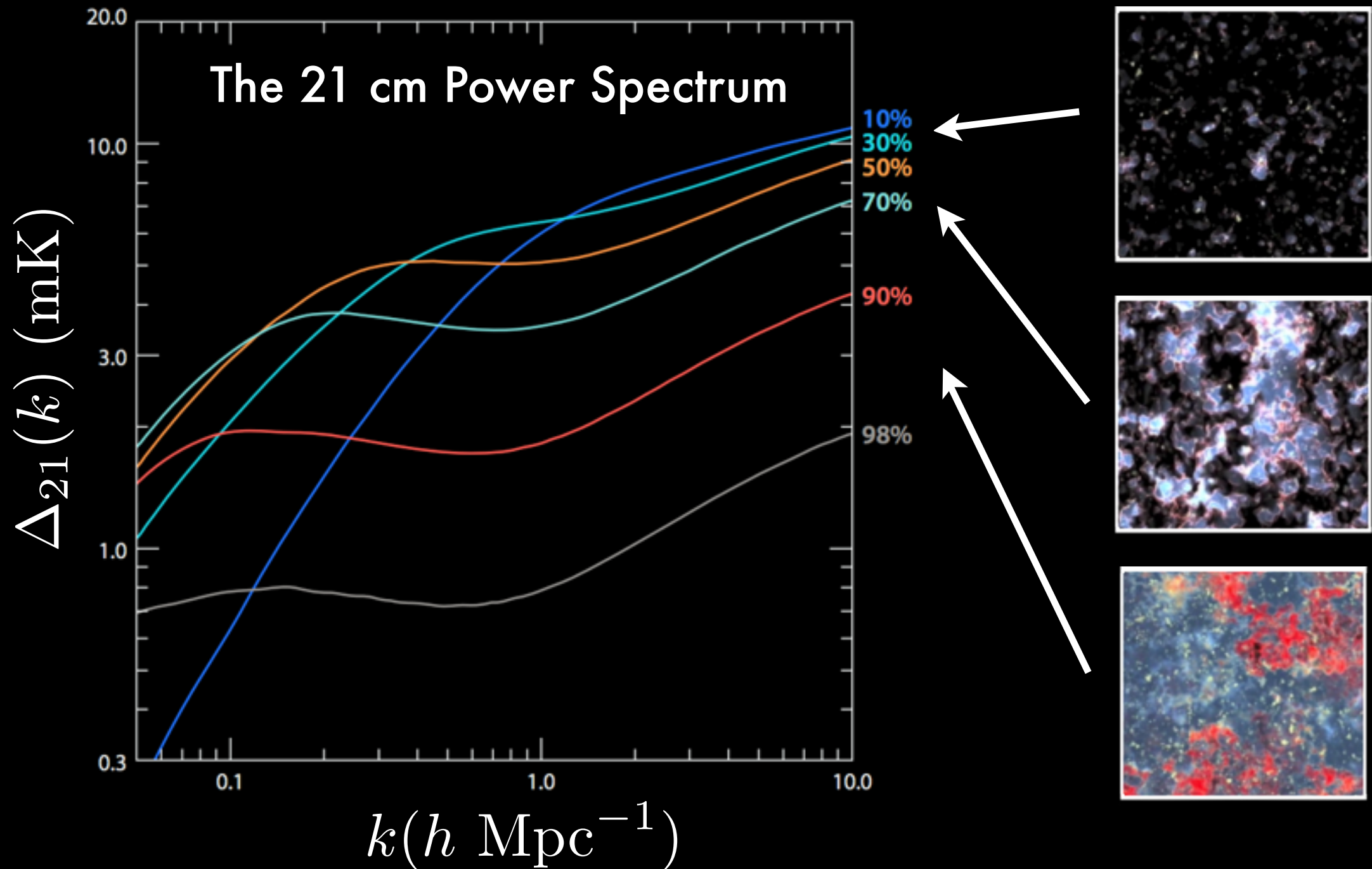
In Practice…

Cosmological Signal
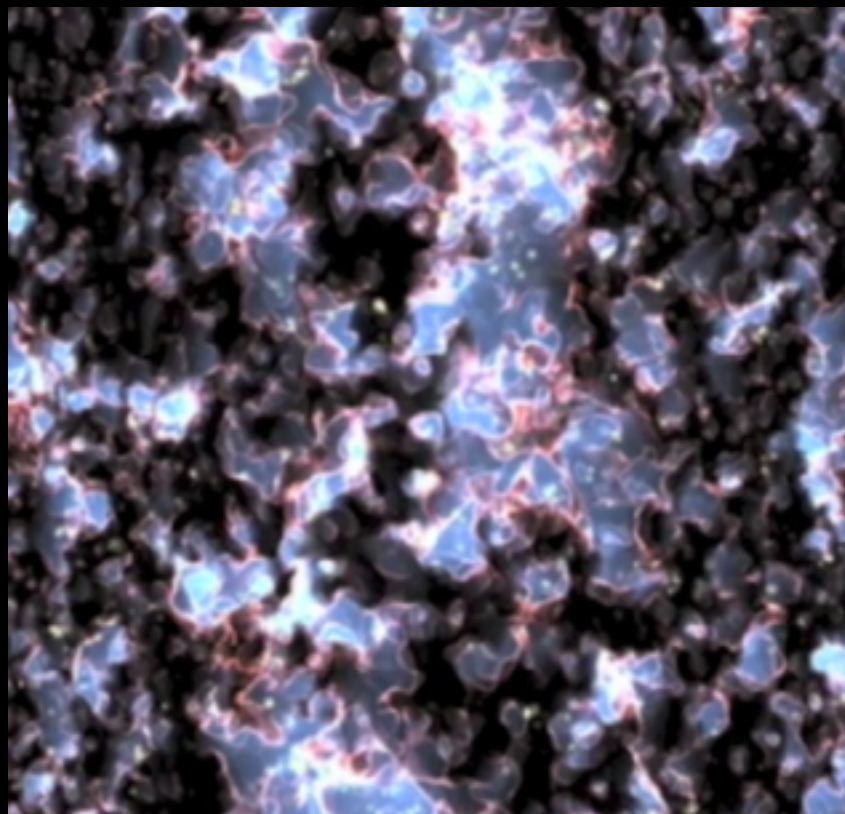
Frequency /
Line of Sight

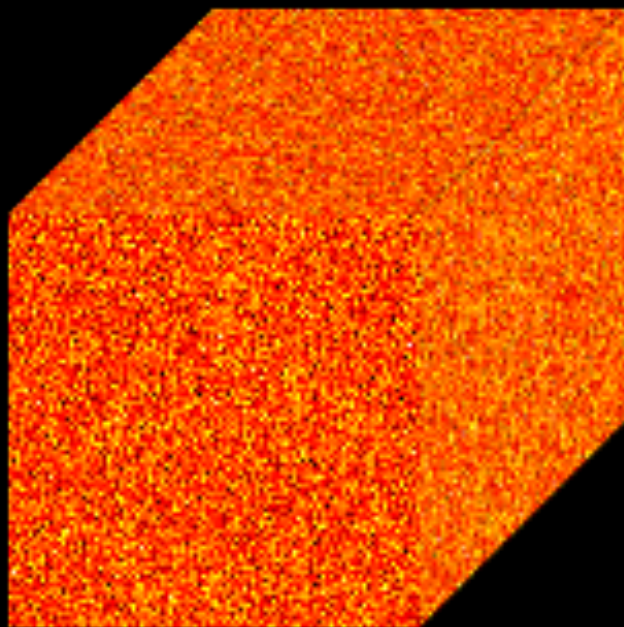# ...the cosmological signal is very dim.

Cosmological Signal

Frequency /
Line of Sight
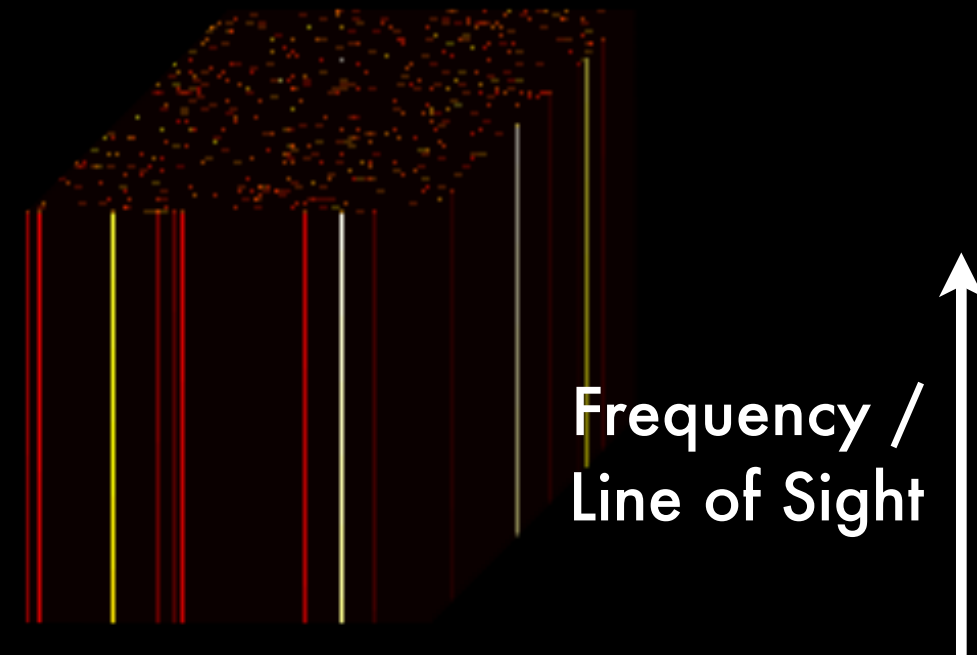
And the contaminants are roughly four orders of magnitude brighter.

Instrumental Noise

Bright Point Sources

Cosmological Signal

Frequency / Line of Sight

Galactic Synchrotron
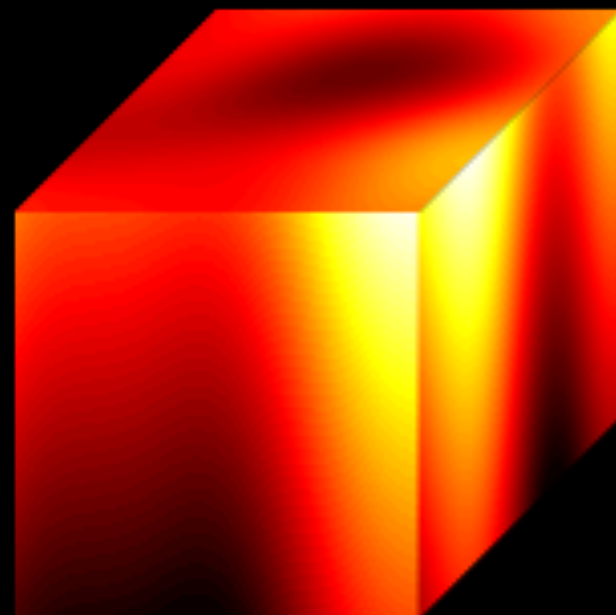
Unresolved Point Sources
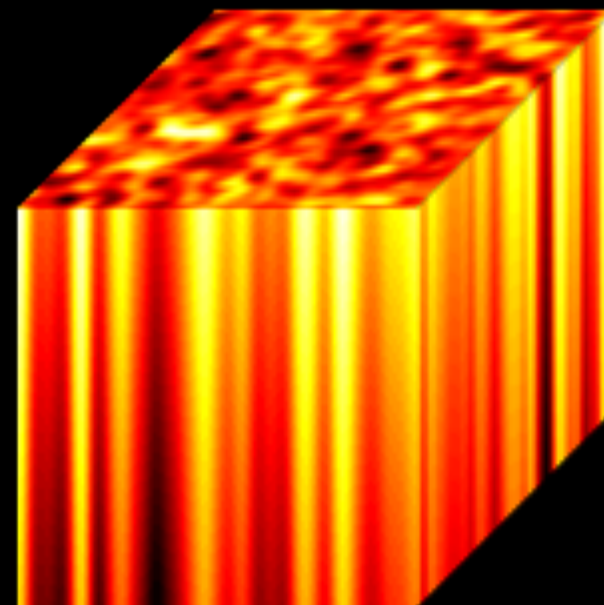
How can we separate the signal from bright foregrounds?

Using their spectral smoothness.

Synchrotron Foregrounds

Intensity

21cm Signal

4 orders of magnitude

Frequency

So instead of spherically averaged Fourier space…
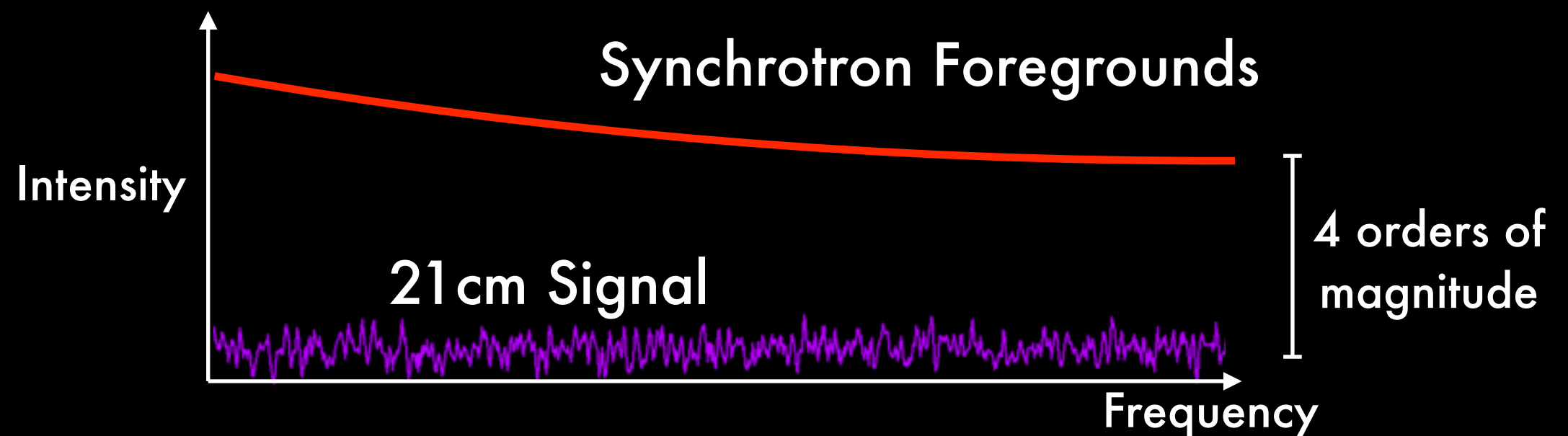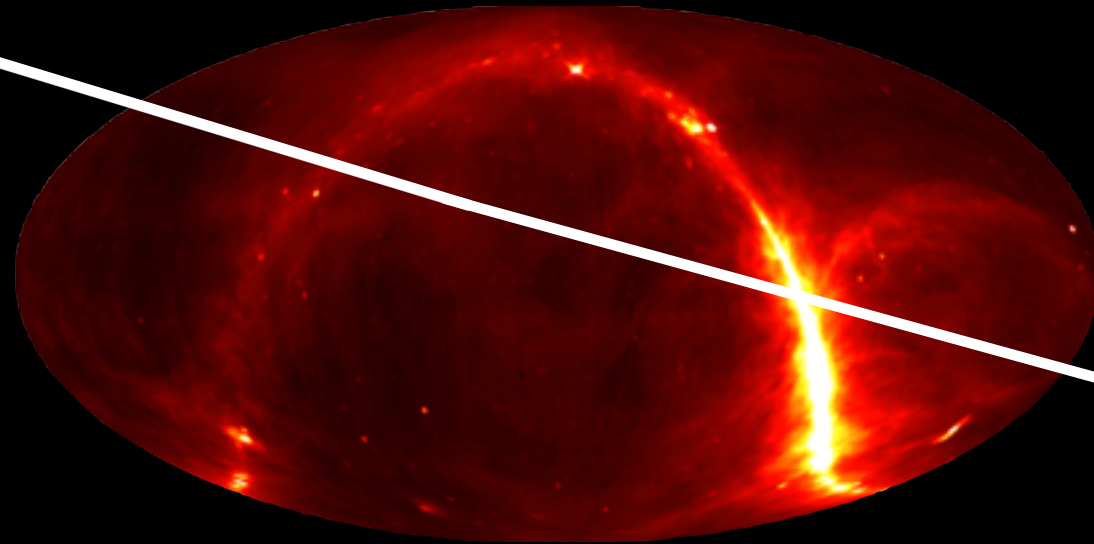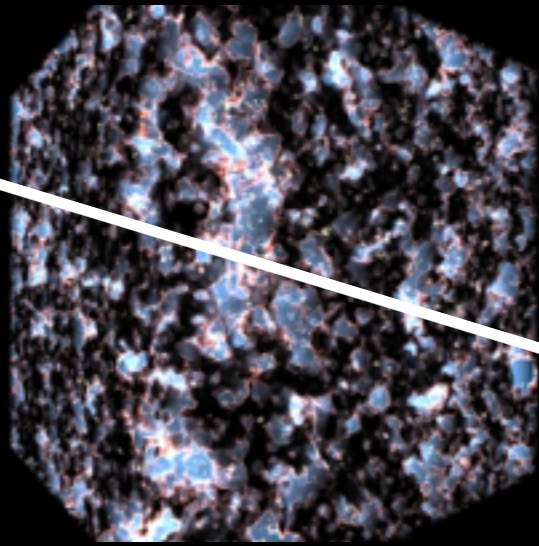
We separate out Fourier modes parallel and perpendicular to the line of sight.
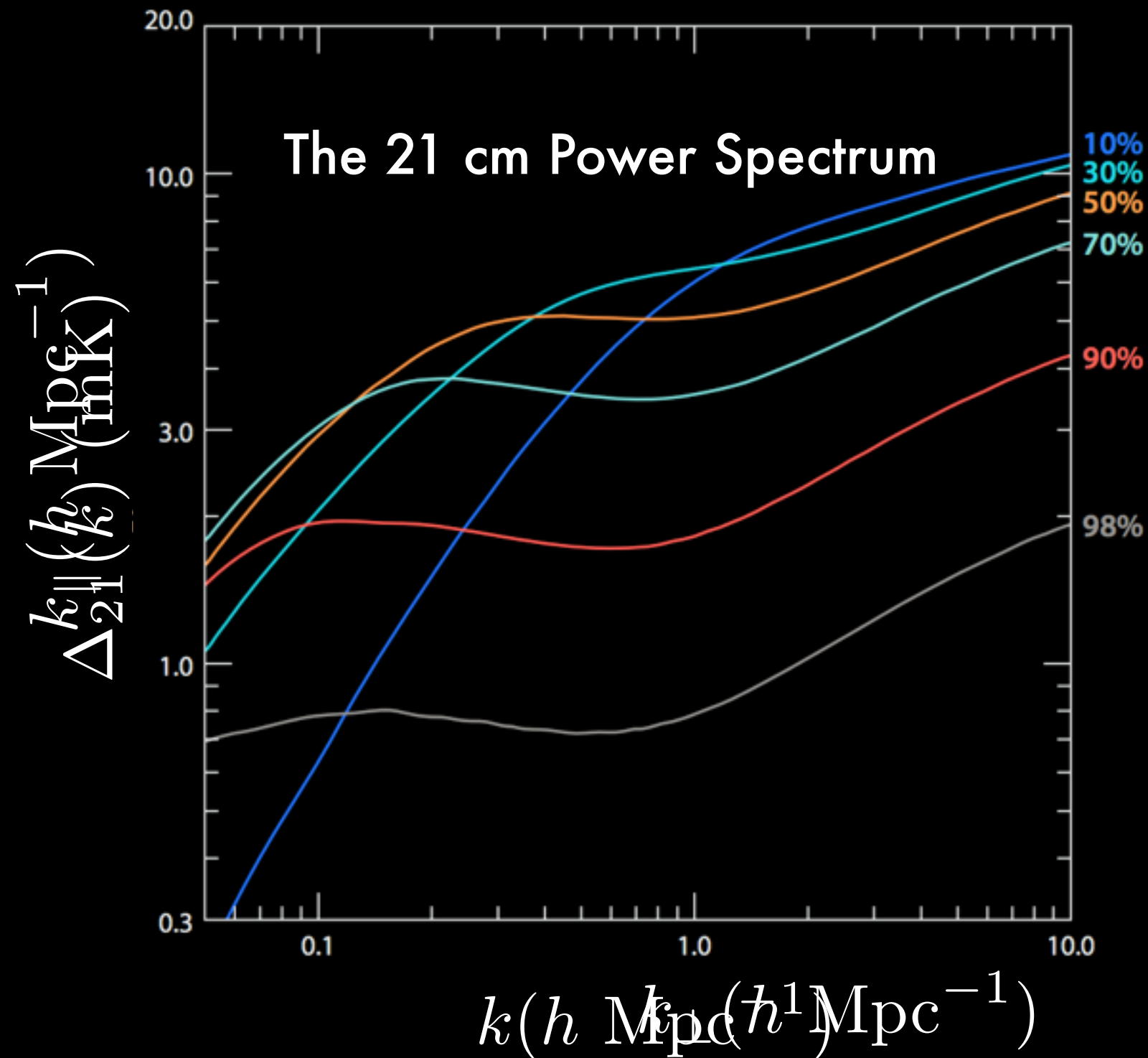
And we find an "EoR Window."
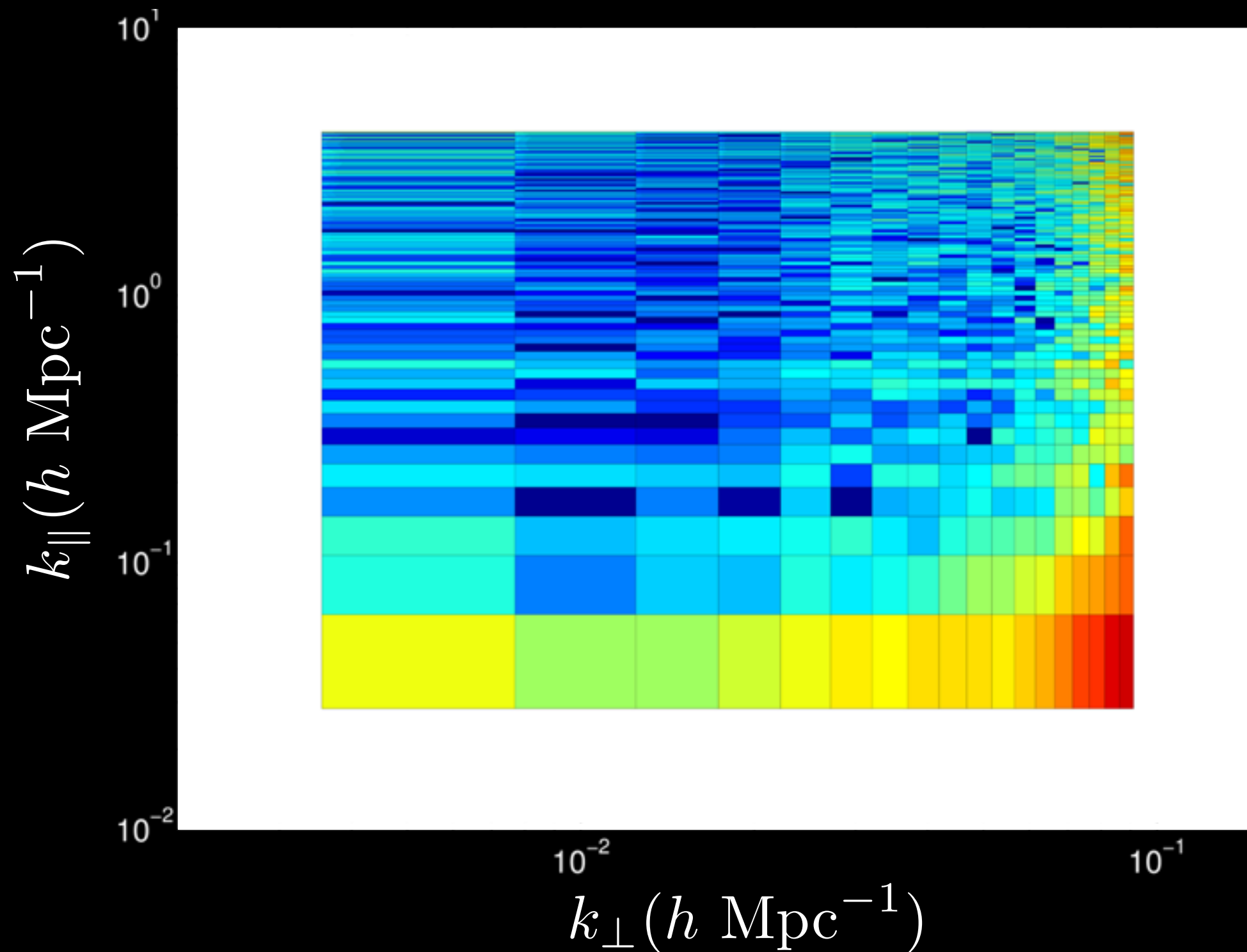
Dillon et al. (2014a)

And we find an "EoR Window."

Dillon et al. (2014a)
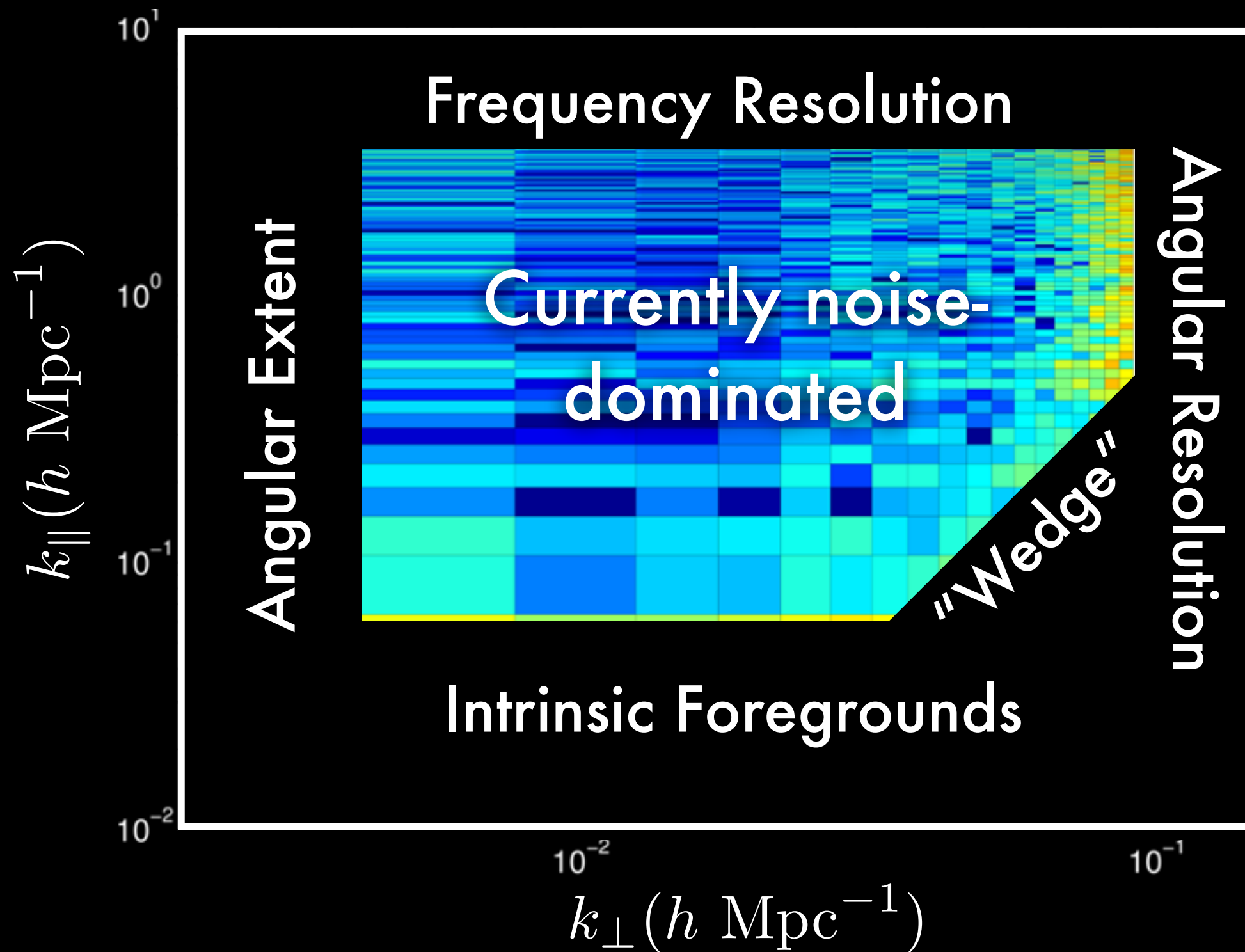
The "wedge" is the imprint of the chromaticity of the synthesized beam.
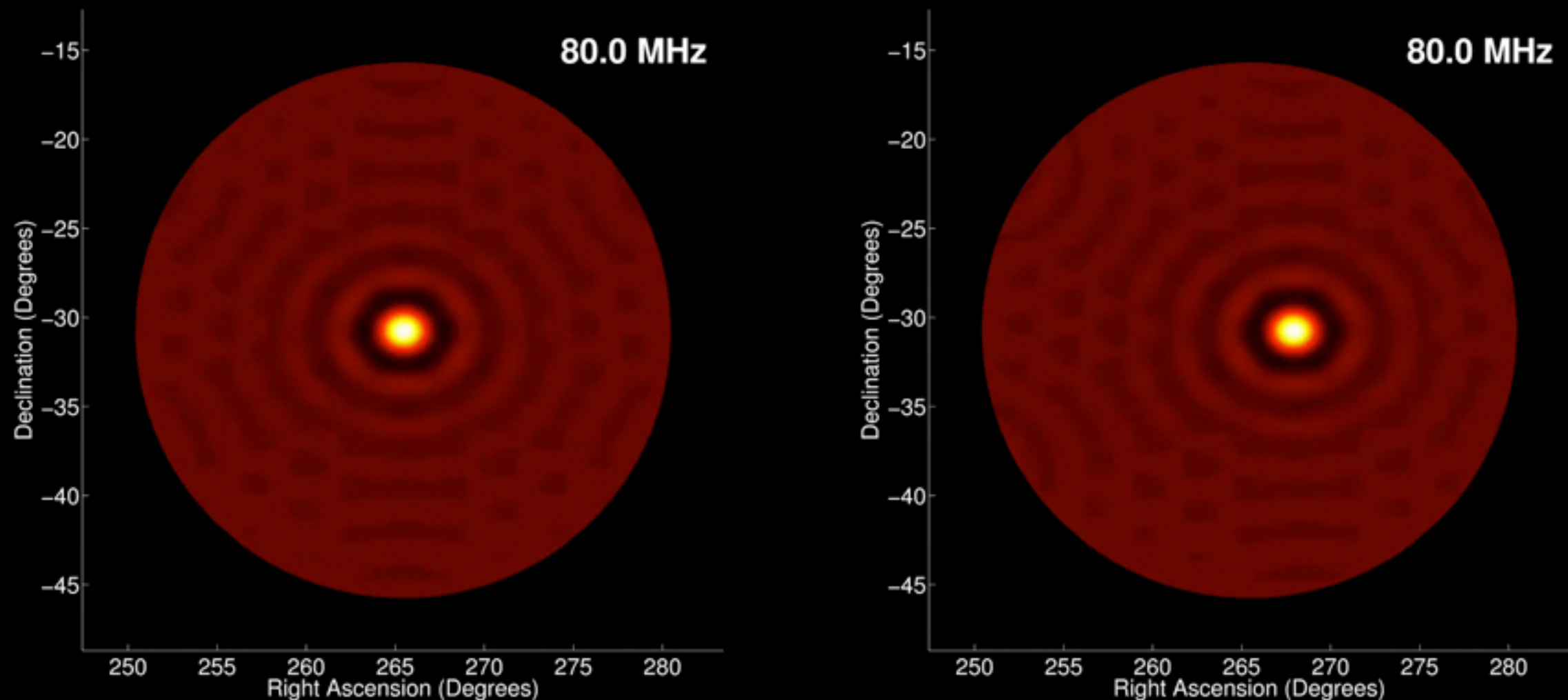
The point spread function has a complicated frequency dependence that introduces spectral structure to spectrally smooth foregrounds.

Dillon et al. (2014b)

But it's limited by geometry.
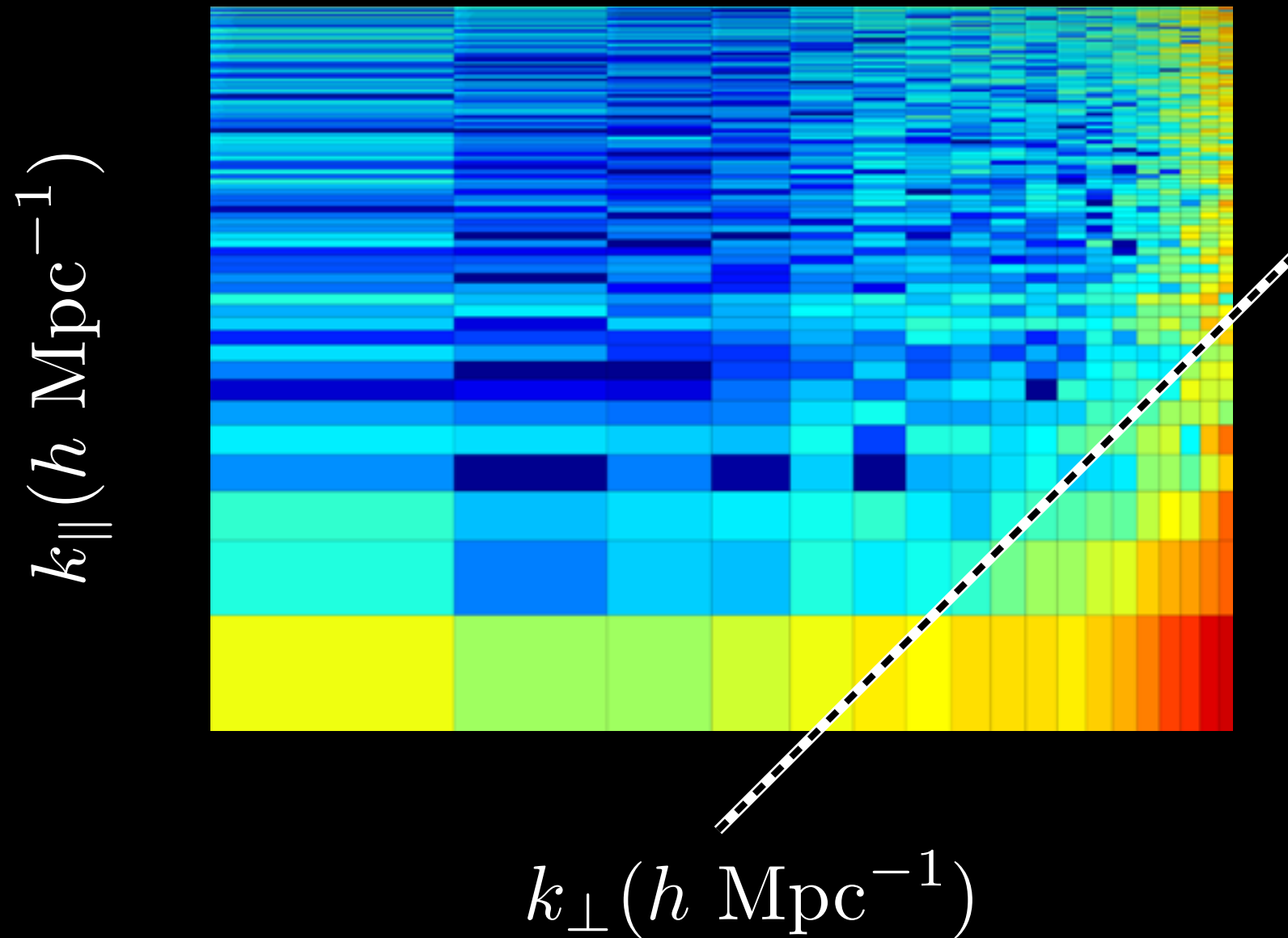
But it's limited by geometry.
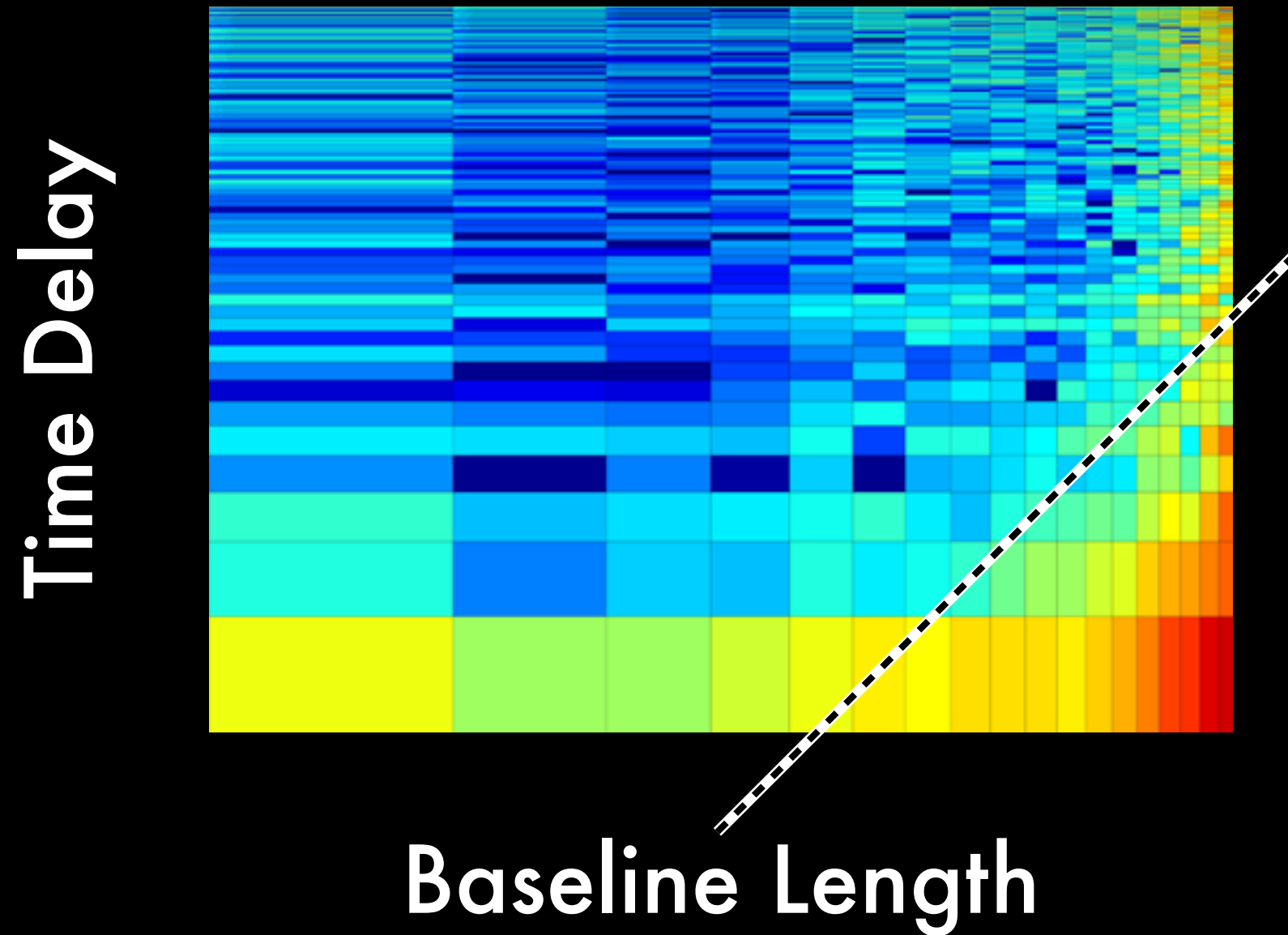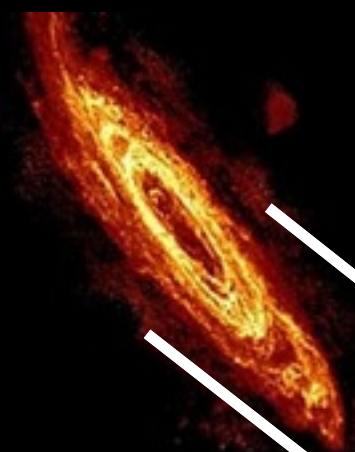
Time Delay

Horizon Delay
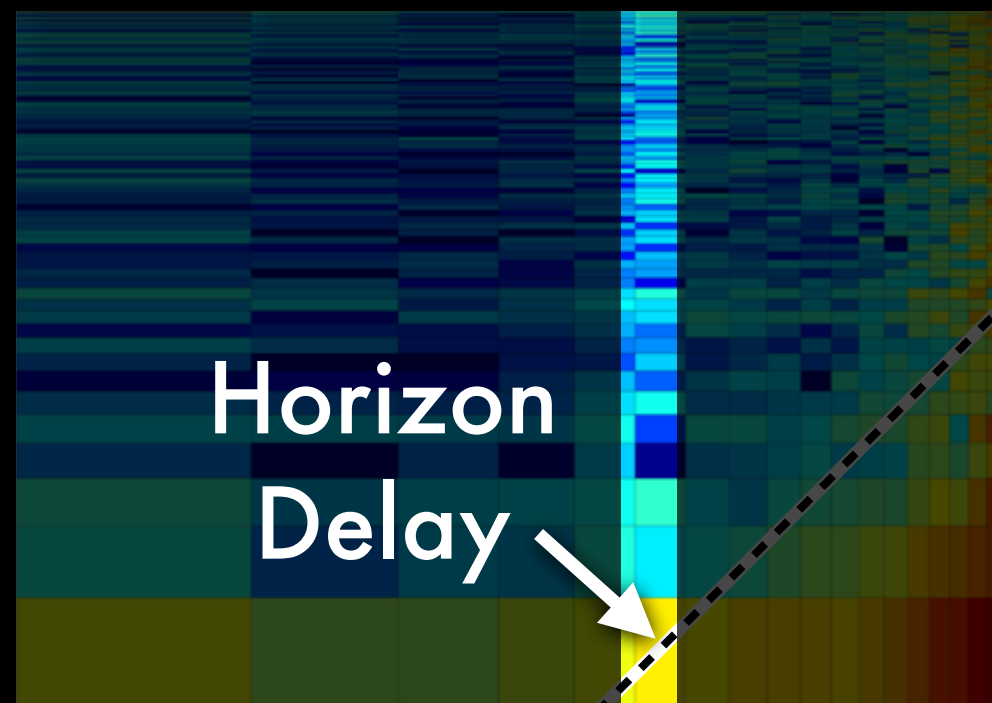
Baseline Length

Δt

The maximum delay of a foreground object is set by the horizon and the length of the baseline.

Parsons et al. (2012)

Time Delay

Horizon Delay

Baseline Length

Δt
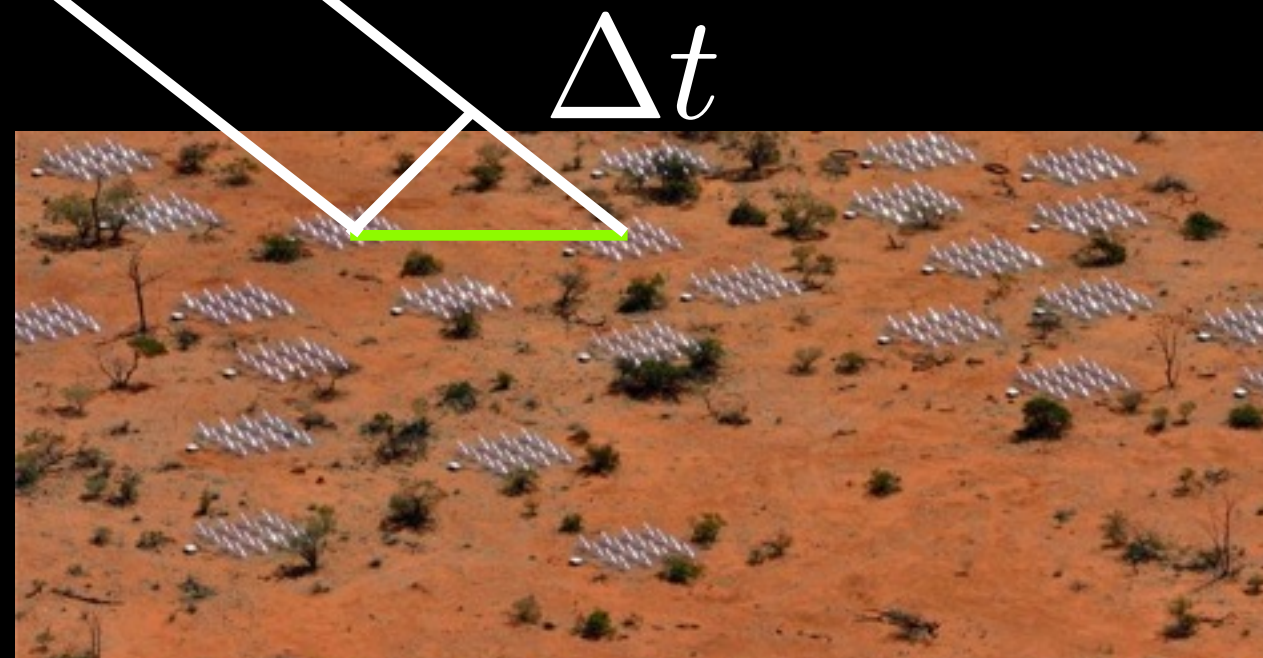
The maximum delay of a foreground object is set by the horizon and the length of the baseline.

Parsons et al. (2012)
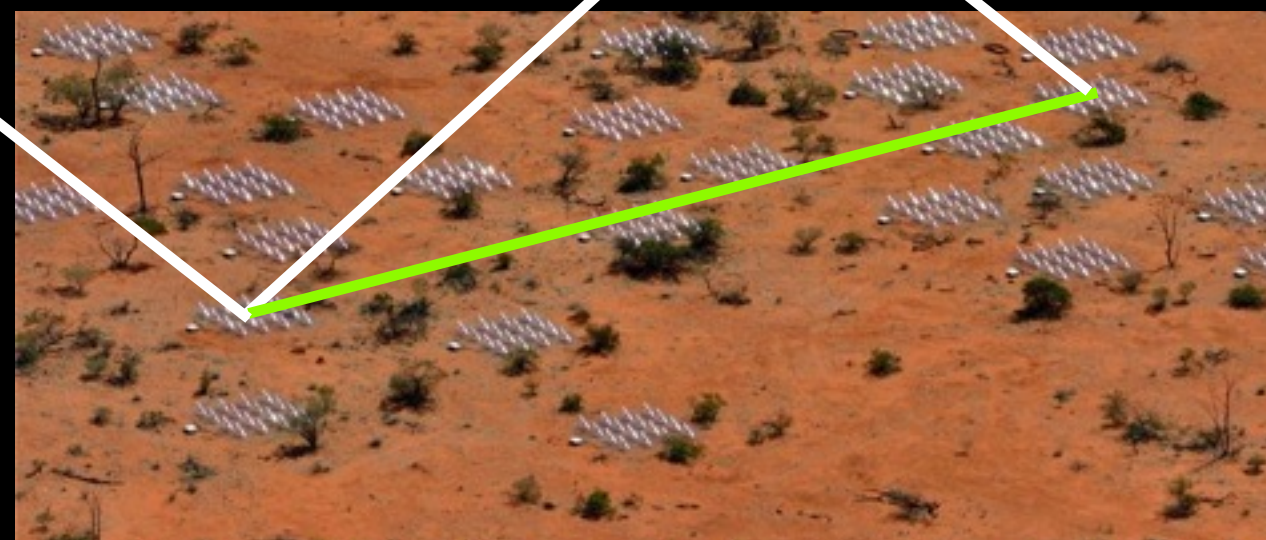
The wedge has been observed to be, as far as we can tell, foreground free.

$\log_{10}[\mathrm{mK}^2(h^{-1}\,\mathrm{Mpc})^3]$
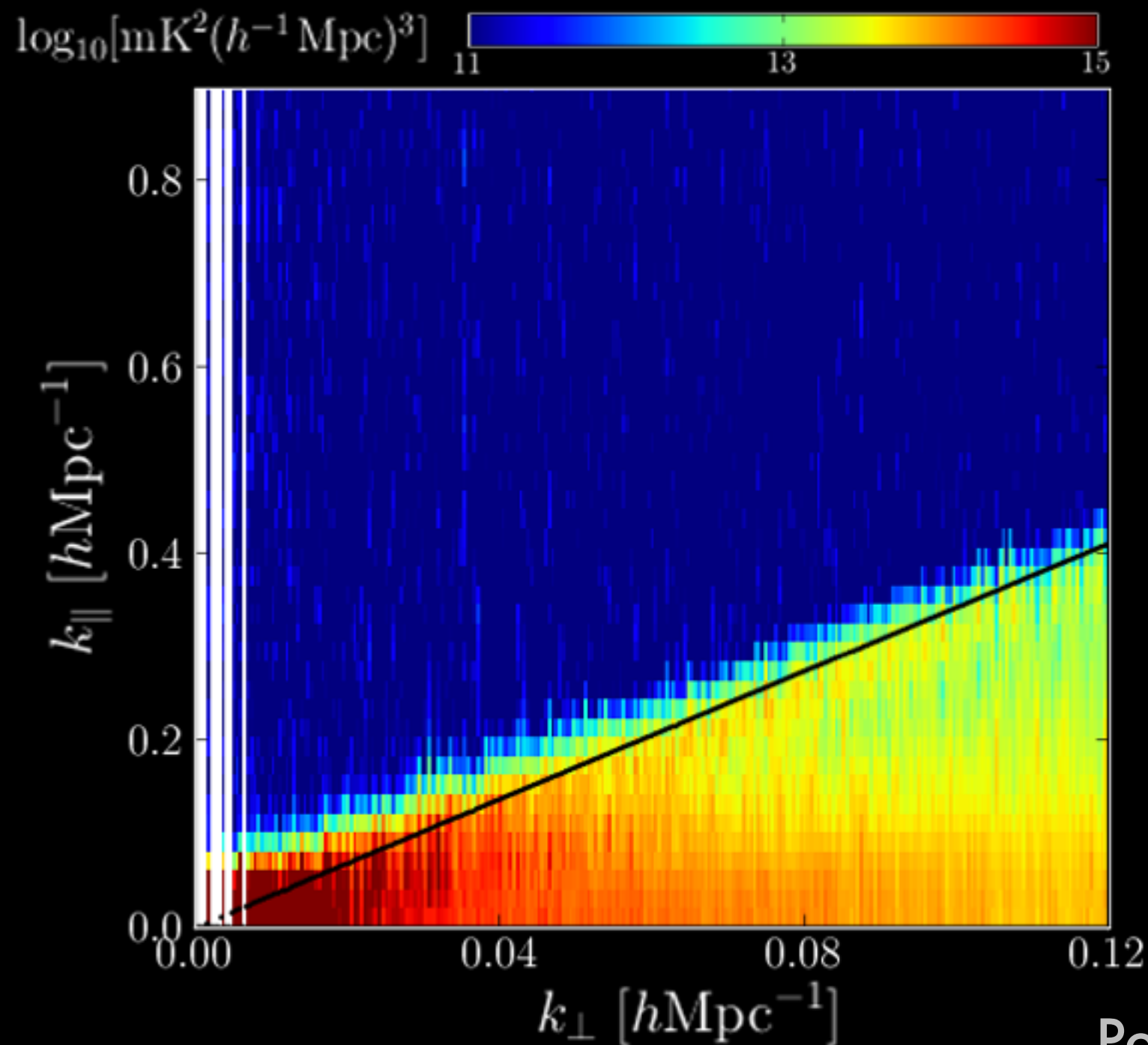
How do we keep the EoR window clean and understand the errors on our measurements?

In an ideal world,
there's an optimal estimator...

Invertible
Normalization
Matrix

Inverse
Covariance
Weighting

x

Data

Fourier
Transform
and Bin

Quadratic Power Spectrum Estimator
preserves all cosmological information
(adapted from CMB and galaxy survey work)

Liu & Tegmark (2011)

...with well-understood error properties.

$$\text{Cov}(\widehat{\mathbf{p}}) = \mathbf{MFM}^{\mathsf{T}}$$

Contains all the errors and error covariances

Fisher Information calculated from the covariance models:

$$F^{\alpha\beta} = \frac{1}{2}\text{tr}\left[\mathbf{C}^{-1}\mathbf{Q}^{\alpha}\mathbf{C}^{-1}\mathbf{Q}^{\beta}\right]$$

Liu & Tegmark (2011)
Dillon, Liu, & Tegmark (2013)

Sounds complicated. Why bother?

The Signal

The Signal

Just like the cosmological signal, there's power on many scales

The Noise

Blue noise represents noise for high angular frequency noise.

The Foregrounds

Low (spectral) frequency foregrounds dominate over signal and noise.

Naïve Filtering

Does an OK job,
but throws out lots
of information.
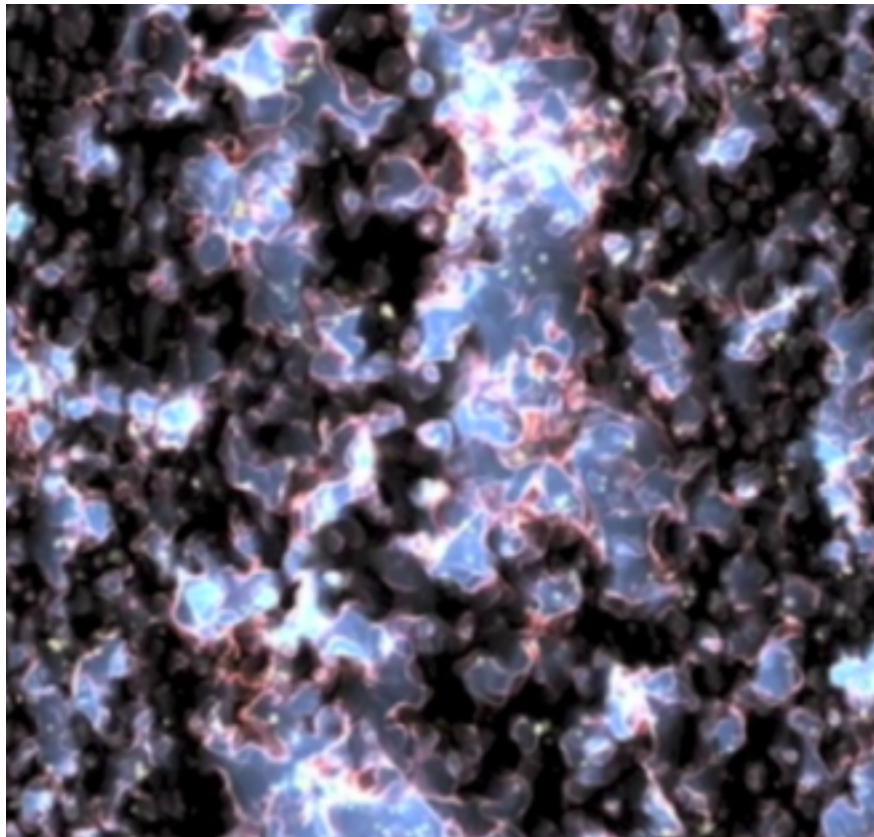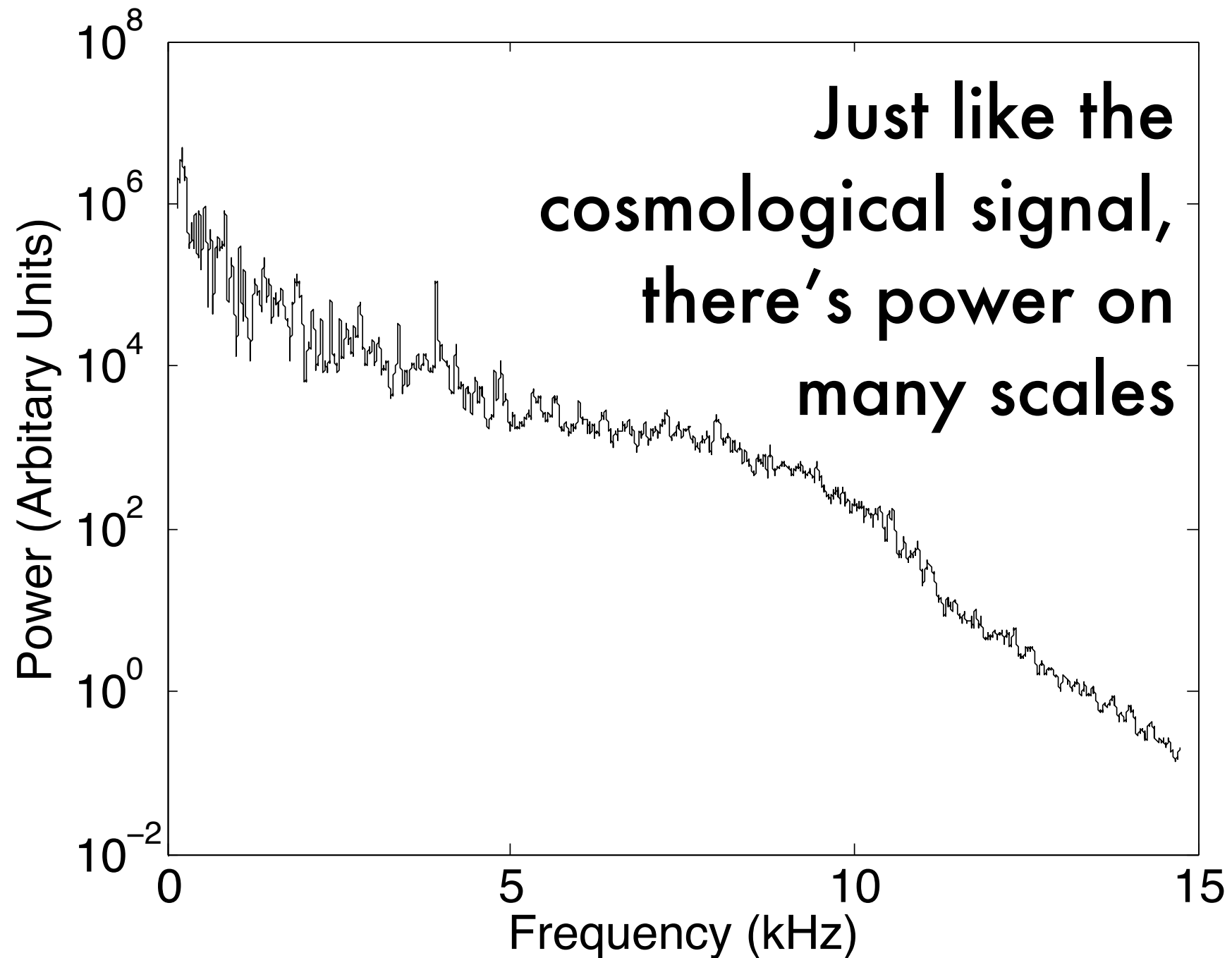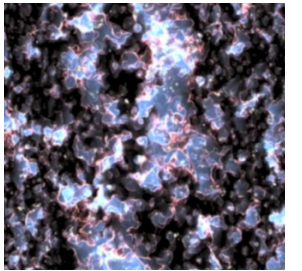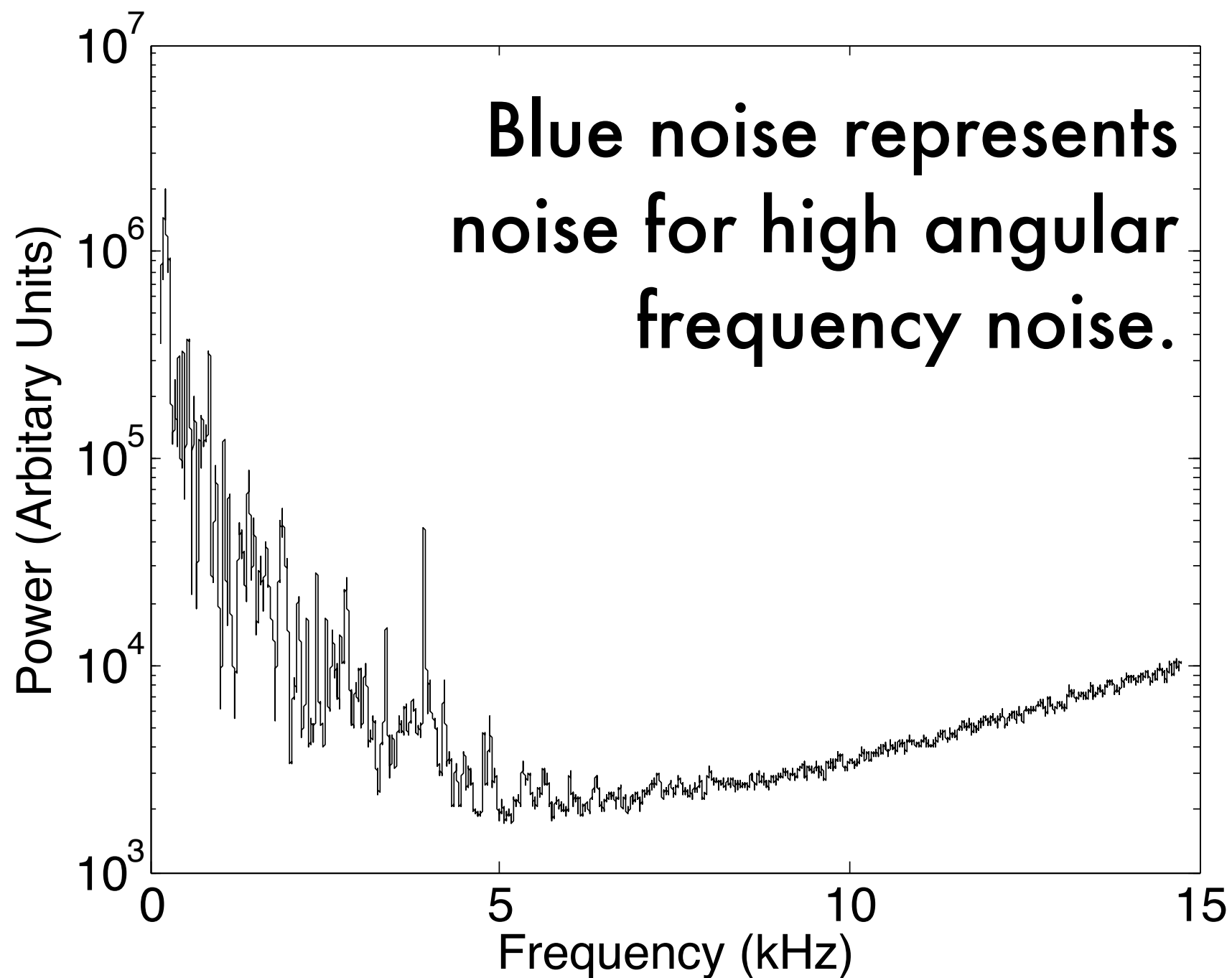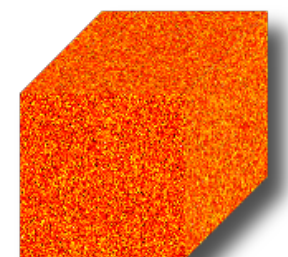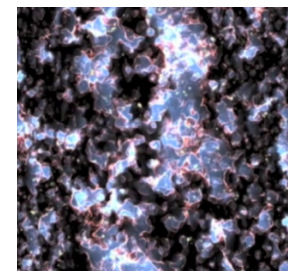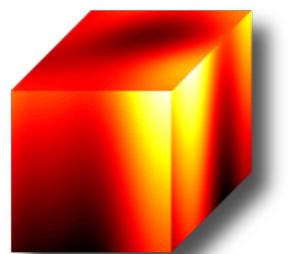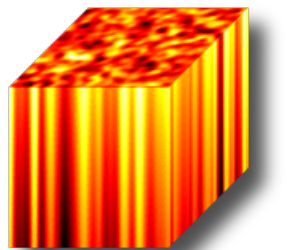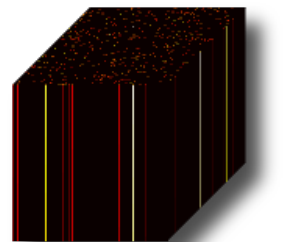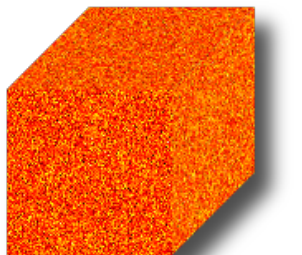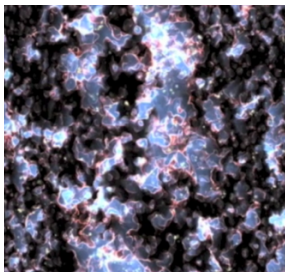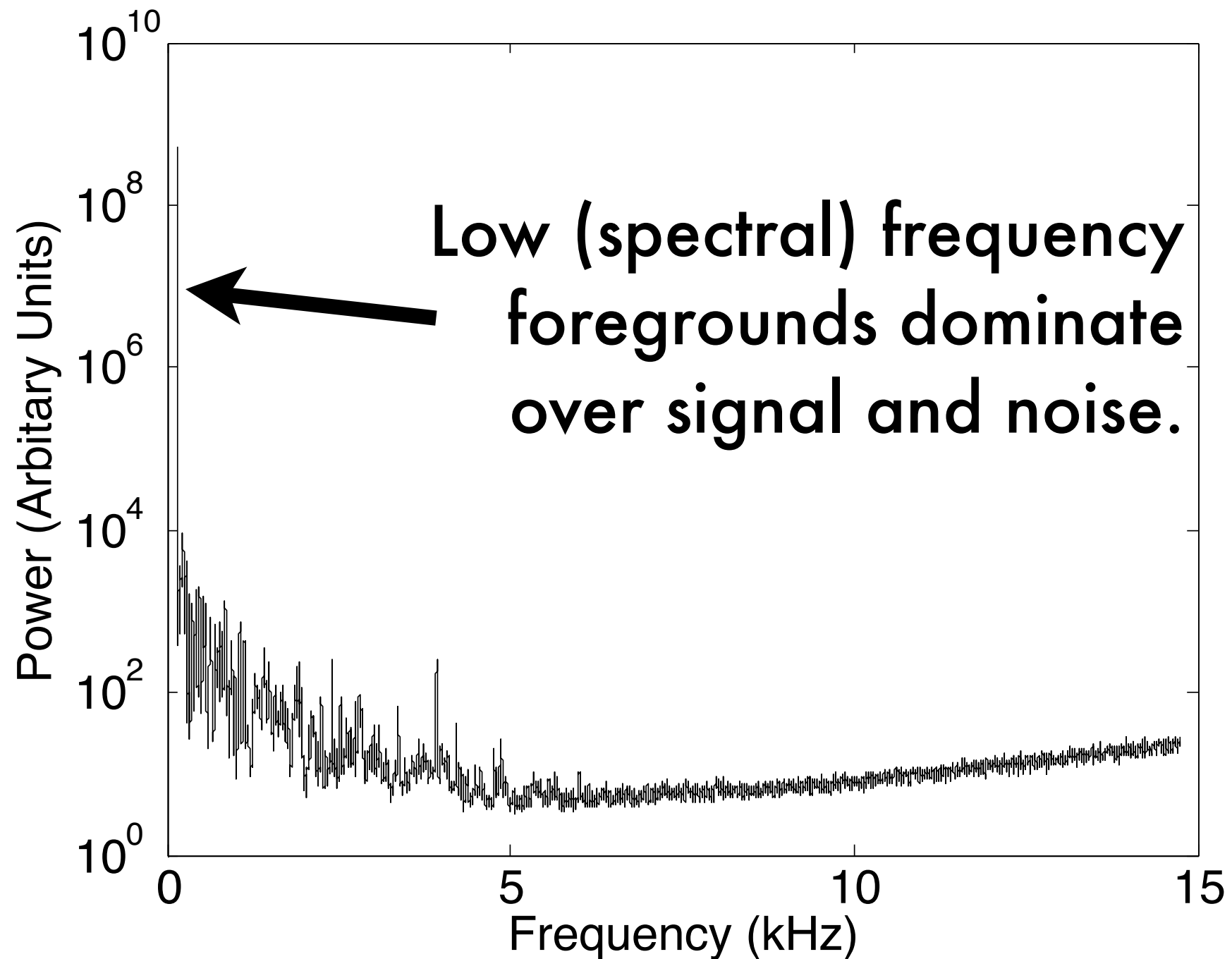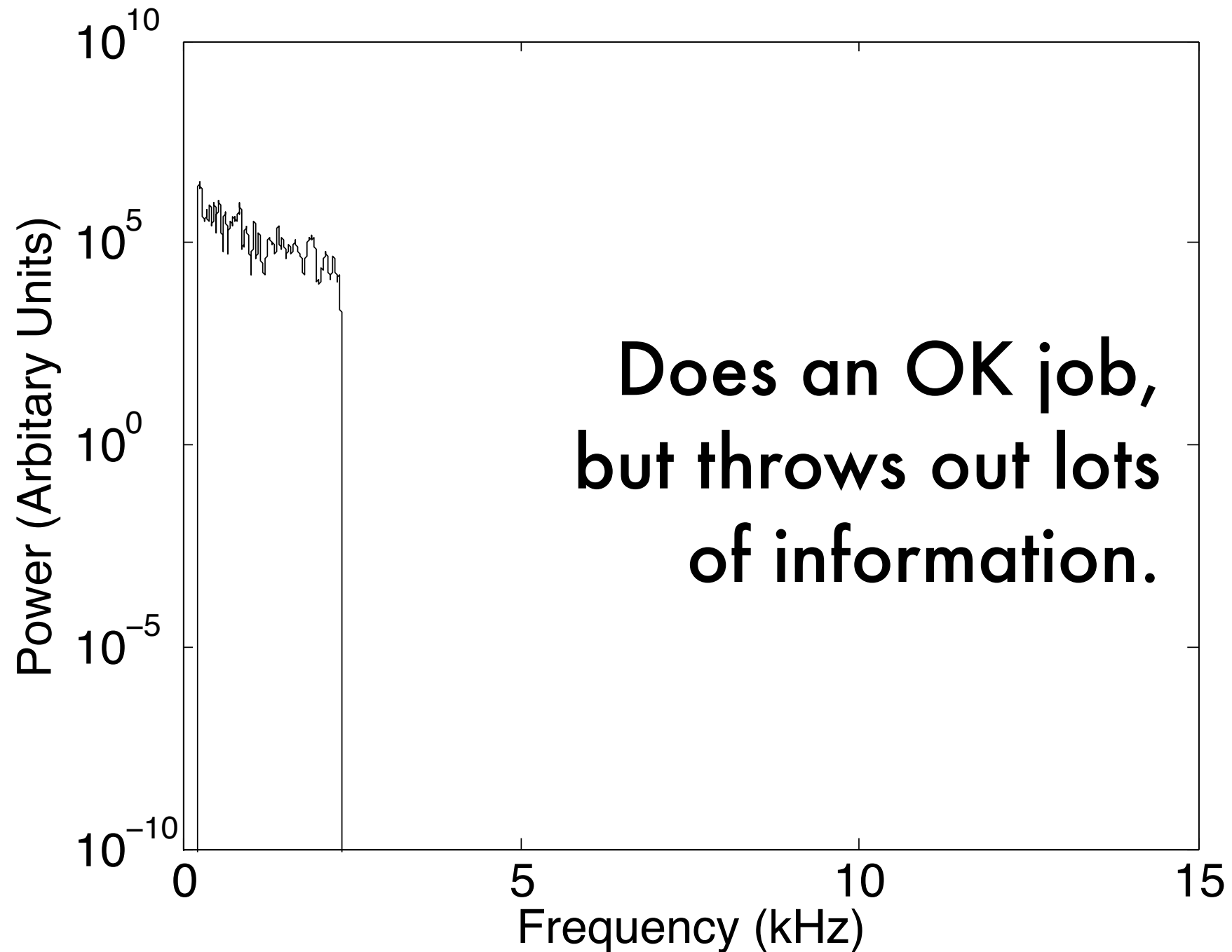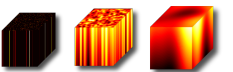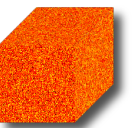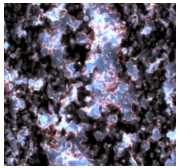
Inverse Variance Weighting

Preserves
as much
information
as possible.
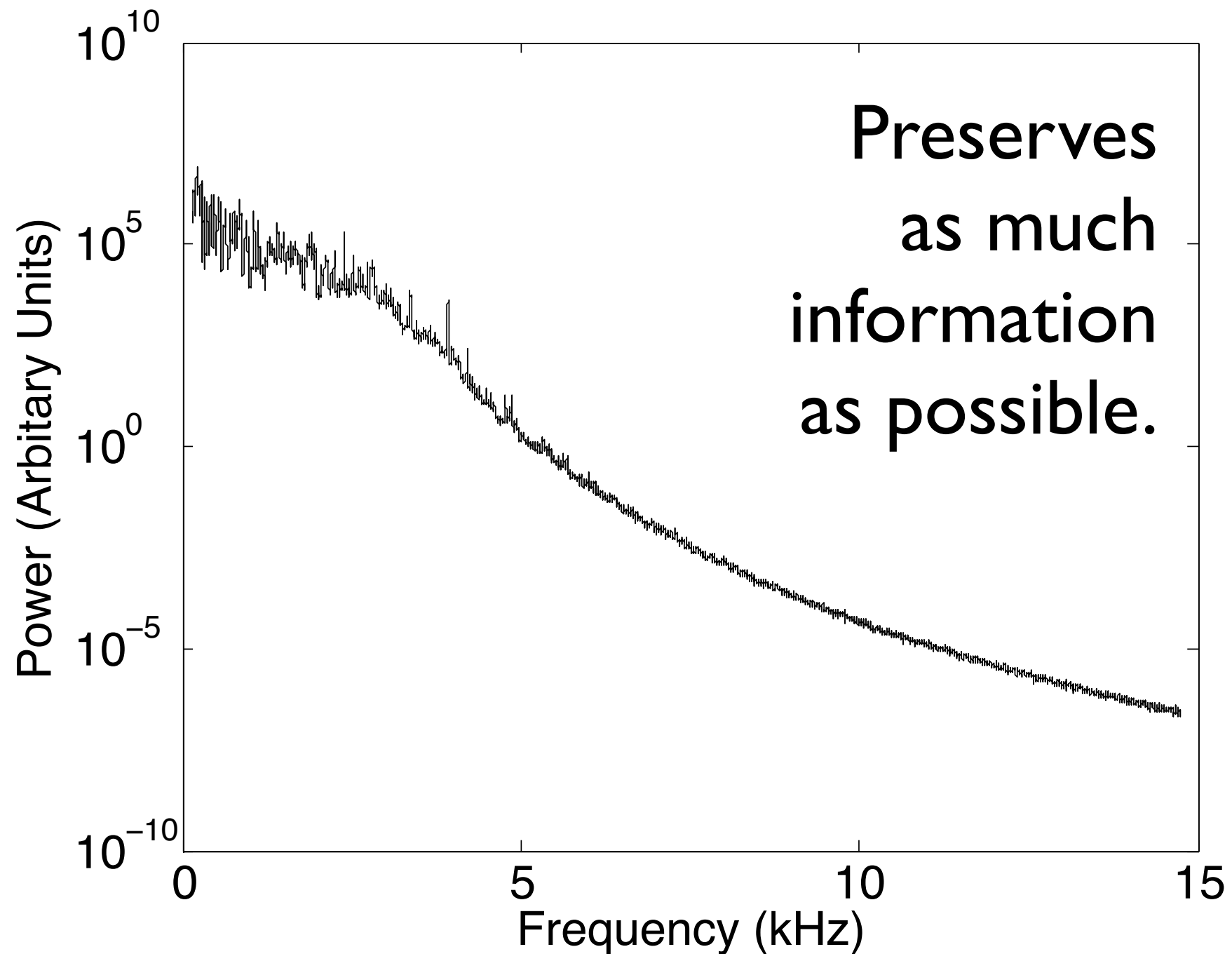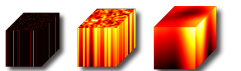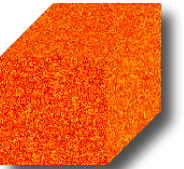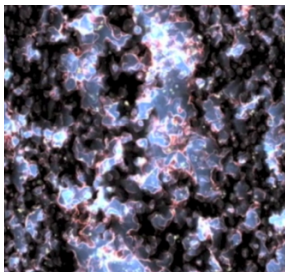
Careful statistics help isolate the signal from the foregrounds...

Recall...

$$\widehat{p}^{\beta} \equiv M^{\alpha\beta} \mathbf{x}^{\mathsf{T}} \mathbf{C}^{-1} \mathbf{Q}^{\alpha} \mathbf{C}^{-1} \mathbf{x}$$

and

$$\mathrm{Cov}(\widehat{\mathbf{p}}) = \mathbf{MFM}^{\mathsf{T}}$$

$$\mathbf{M} \sim \mathbf{I}$$

- Smallest errors, but errors are correlated

- Hard to cut out foregrounds

$$\mathbf{M} \sim \mathbf{F}^{-1/2}$$

- Decorrelated errors.

- Each band power represents a mutually exclusive yet collectively exhaustive piece of information.

A good estimator preserves the EoR Window.

$$\mathbf{M} \sim \mathbf{I} \qquad \mathbf{M} \sim \mathbf{F}^{-1/2}$$
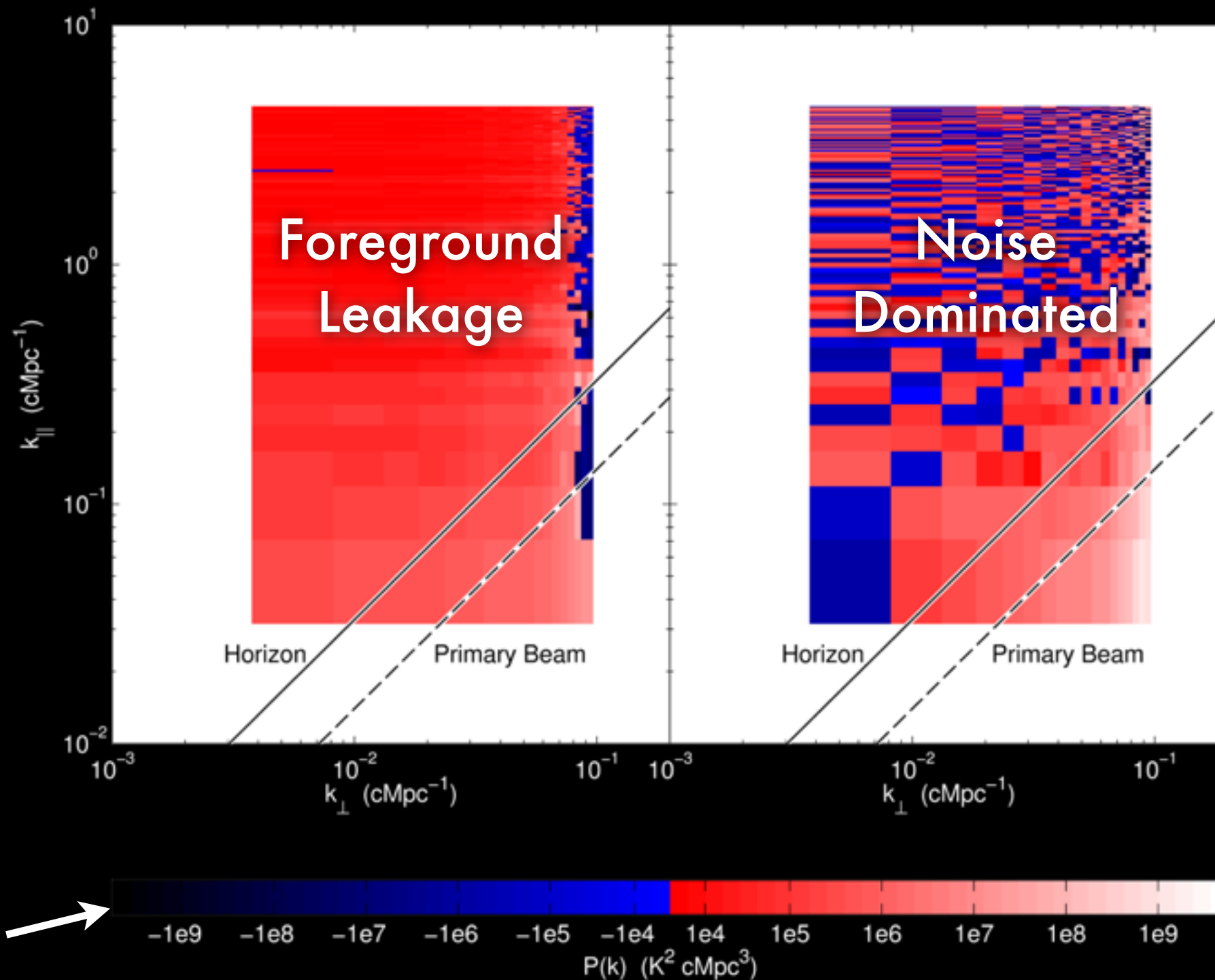
Cross Power Spectrum

But there's a catch.

$$C^{-1}$$

- Scales as $O(N^3)$

- Computationally infeasible with current data sets

Fast Power Spectrum Estimation

1. Generate lots of random data cubes from the model covariance, exploiting symmetries

All in O(NlogN)*

Dillon, Liu, & Tegmark (2013)

Because $\mathbf{N} = \mathbf{F}_\perp^\dagger \tilde{\mathbf{N}} \mathbf{F}_\perp$, and because $\tilde{\mathbf{N}}$ is a diagonal matrix, we define $\tilde{\mathbf{P}}_\mathbf{N}$ and $\mathbf{P}_\mathbf{N}^\dagger$ as follows:

$$\mathbf{P}_\mathbf{N} = \tilde{\mathbf{N}}^{-1/2}\mathbf{F}_\perp,$$
$$\mathbf{P}_\mathbf{N}^\dagger = \mathbf{F}_\perp^\dagger \tilde{\mathbf{N}}^{-1/2}. \quad (C2)$$

Since applying $\mathbf{P}_\mathbf{N}$ only requires multiplying by the inverse square root of a diagonal matrix and Fourier transforming in two dimensions, the complexity of applying $\mathbf{P}_\mathbf{N}$ to a vector is less than $\mathcal{O}(N \log N)$.

2. Constructing a Preconditioner for U

The matrix $\mathbf{U}$ (Equation 55) can be written as the tensor product of three Toeplitz matrices, one for each dimension, bookended by two diagonal matrices, $\mathbf{D}_\mathbf{U}$. Furthermore, since $\mathbf{D}_\mathbf{U}$ depends only on frequency (as we saw in Section III D 2), its effect can be folded into $\mathbf{U}_z$ such that

$$\mathbf{D}_\mathbf{U}[\mathbf{U}_x \otimes \mathbf{U}_y \otimes \mathbf{U}_z]\mathbf{D}_\mathbf{U} \equiv \mathbf{U}_x \otimes \mathbf{U}_y \otimes \mathbf{U}_z'. \quad (C3)$$

It is generally the ... proximated by the ... that the spatial c... is comparable with... ment. This assum... fairly compact arr... longest baseline— ... be optimal for 21... resolution on the c... rable to the fiducia... describe the cluste... sources. For the ... only, we can there...

where we have dro... ity. Looking bac... explains the stair... for every eigenvalu... values.

Since only a few... gogically useful to... preconditioning pr... rank 1 matrix by ... after the first eige... the other relevant ... follows:

where $\mathbf{v}_z$ is the n...

Let us now take

$\mathbf{U} + \mathbf{N}$:

$$\mathbf{P}_\mathbf{N}(\mathbf{U} + \mathbf{N})\mathbf{P}_\mathbf{N}^\dagger$$
$$= \mathbf{I} + \tilde{\mathbf{N}}^{-1}\dots$$
$$= \mathbf{I} + \tilde{\mathbf{N}}^{-1}\dots$$
$$\equiv \mathbf{I} + \overline{\mathbf{U}}.$$

Our next goal, ther... $\mathbf{P}_\mathbf{U}$ that, when app... to $\mathbf{I}$.

We now take a ... approximation to s... to the line of sight,

$$\overline{\mathbf{U}} \approx (\tilde{\mathbf{N}}_\perp^{-1/2} \otimes \dots$$
$$= (\tilde{\mathbf{N}}_\perp^{-1}) \otimes (\dots$$

where $\tilde{\mathbf{N}}_\perp$ is still a ... dimensions, genera... aged over frequency ... ditioning matrices, ... where $\mathbf{\Pi}$ has the ... $\overline{\mathbf{U}}\mathbf{\Pi}^\dagger = \overline{\mathbf{U}}$. The ma...

matrix that looks like $\mathbf{I} + \overline{\mathbf{U}}$ we can make it look like $\mathbf{I}$. So can we take $\mathbf{I} + \overline{\mathbf{U}} + \overline{\overline{\Gamma}}$, where $\overline{\overline{\Gamma}} \equiv \mathbf{P}_\mathbf{N}\Gamma\mathbf{P}_\mathbf{N}^\dagger$, and turn it into $\mathbf{I} + \overline{\mathbf{U}}$? Looking at $\overline{\overline{\Gamma}}$,

$$\overline{\overline{\Gamma}} = \tilde{\mathbf{N}}^{-1/2}\mathbf{F}_\perp\lambda_\Gamma\mathbf{v}\mathbf{v}^\dagger\mathbf{F}_\perp^\dagger\tilde{\mathbf{N}}^{-1/2}$$
$$= \lambda_\Gamma(\tilde{\mathbf{N}}^{-1/2}\tilde{\mathbf{v}}_\perp\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{N}}^{-1/2}) \otimes \mathbf{v}_z\mathbf{v}_z^\dagger, \quad (C20)$$

where λ_Γ is the sole eigenvalue we are considering and where $\tilde{\mathbf{v}}_\perp \equiv \mathbf{F}_\perp\mathbf{v}_\perp$.

Again, we will look at a preconditioner of the $\mathbf{P}_\Gamma = \mathbf{I} - \beta\mathbf{\Pi}$ where:

$$\mathbf{\Pi} \equiv \left(\tilde{\mathbf{N}}_\perp^{-1/2}\tilde{\mathbf{v}}_\perp\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{N}}_\perp^{1/2}\right) \otimes \mathbf{v}_z\mathbf{v}_z^\dagger. \quad (C21)$$

This time, the $\tilde{\mathbf{N}}_\perp^{\pm 1/2}$ matrices do not pass through the eigenvectors to cancel one another out. We now exploit the spectral similarity of foregrounds and the fact that $\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{v}}_\perp = \mathbf{v}_z^\dagger\mathbf{v}_z = 1$ to obtain

$$\mathbf{P}_\Gamma\overline{\mathbf{U}}\mathbf{P}_\Gamma^\dagger = \overline{\mathbf{U}} + \frac{\lambda_\mathbf{U}}{\lambda_\Gamma}(\beta^2 - 2\beta)\overline{\overline{\Gamma}}. \quad (C22)$$

This is very useful because it means that if we pick β properly, we can get the second term to cancel the $\overline{\overline{\Gamma}}$ terms we expect when we calculate the full effect of $\mathbf{P}_\Gamma$ and $\mathbf{P}_\mathbf{N}$ on $\mathbf{N} + \mathbf{U} + \Gamma$. Noting that the sole eigenvalue of $\overline{\overline{\Gamma}}$ is $\overline{\lambda}_\Gamma \equiv \lambda_\Gamma\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{N}}_\perp^{-1}\tilde{\mathbf{v}}_\perp$, we also define $\overline{\lambda}_\mathbf{U} \equiv \lambda_\mathbf{U}\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{N}}_\perp^{-1}\tilde{\mathbf{v}}_\perp$. Multiplying our preconditioner by our matrices, we see that the the equality of the single eigenvalues yields another quadratic equation for β:

$$1 + \overline{\lambda}_\mathbf{U} = 1 - 2\beta + \beta^2 + (\beta^2 - 2\beta + 1)\overline{\lambda}_\Gamma$$
$$+ \overline{\lambda}_\Gamma\frac{\lambda_\mathbf{U}}{\lambda_\Gamma}(\beta^2 - 2\beta). \quad (C23)$$

Solving the quadratic equation, we get

$$\mathbf{P}_\mathbf{U} \equiv \mathbf{I} - \left[\sum_l \left(1 - \sqrt{\frac{1}{1+\overline{\lambda}_l}}\right)\mathbf{v}_{z_l}\mathbf{v}_{z_l}^\dagger \otimes \overset{\leftrightarrow}{\delta}_{x,x_l} \otimes \overset{\leftrightarrow}{\delta}_{y,y_l}\right], \quad (C12)$$

where the pair of $\overset{\leftrightarrow}{\delta}$ matrices pick out a particular uv-cell. If we want to generalize to more eigenvectors of $\mathbf{U}_z$, we simply need to keep subtracting off sums of matrices on the right hand side of Equation (C12):

$$\mathbf{P}_\mathbf{U} \equiv \mathbf{I} - \sum_k \left[\sum_l \left(1 - \sqrt{\frac{1}{1+\overline{\lambda}_{l,k}}}\right)\mathbf{v}_{z_k}\mathbf{v}_{z_k}^\dagger \otimes \overset{\leftrightarrow}{\delta}_{x,x_l} \otimes \overset{\leftrightarrow}{\delta}_{y,y_l}\right], \quad (C13)$$

This works because every set of vectors corresponding to a value of k is orthogonal to every other set. Each term in the above sum acts on a different subspace of $\mathbf{C}$ independent of all the other terms in the sum.

If the relevant vectors $\mathbf{v}$ ... are precomputed, applying ...

Solving, we finally have our $\mathbf{P}_\Gamma$ that acts on $\mathbf{I} + \overline{\mathbf{U}} + \overline{\Gamma}$ and yields $\mathbf{I} + \overline{\mathbf{U}}$:

$$\mathbf{P}_\Gamma = \mathbf{I} - \left(1 - \sqrt{\frac{1+\overline{\lambda}_\mathbf{U}}{1+\overline{\lambda}_\mathbf{U}+\overline{\lambda}_\Gamma}}\right) \times \left[\left(\tilde{\mathbf{N}}_\perp^{-1/2}\tilde{\mathbf{v}}_\perp\tilde{\mathbf{v}}_\perp^\dagger\tilde{\mathbf{N}}_\perp^{1/2}\right) \otimes \mathbf{v}_z\mathbf{v}_z^\dagger\right]. \quad (C24)$$

Finally, generalizing to multiple eigenvalues and taking advantage of the orthonormality of the eigenvectors, we have

$$\mathbf{P}_\Gamma = \mathbf{I} - \sum_{k,m}\left[\left(1 - \sqrt{\frac{1+\overline{\lambda}_{\mathbf{U}_k}}{1+\overline{\lambda}_{\mathbf{U}_k}+\overline{\lambda}_{\Gamma_{k,m}}}}\right) \times \left(\left(\tilde{\mathbf{N}}_\perp^{-1/2}\tilde{\mathbf{v}}_{\perp_m}\tilde{\mathbf{v}}_{\perp_m}^\dagger\tilde{\mathbf{N}}_\perp^{1/2}\right) \otimes \mathbf{v}_{z_k}\mathbf{v}_{z_k}^\dagger\right)\right]. \quad (C25)$$

The result of this somewhat complicated preconditioner is a reduction of the condition number of the matrix to be inverted by many orders of magnitude (see Figure 5).

Lastly, we include Fourier transforms at the front and the back of the preconditioner, so that the result, when multiplied by a real vector, returns a real vector. Therefore, the total preconditioner we use for $\mathbf{C}$ is:

$$\mathbf{F}_\perp^\dagger \mathbf{P}_\mathbf{U}\mathbf{P}_\Gamma\mathbf{P}_\mathbf{N}(\mathbf{R}+\mathbf{U}+\mathbf{N}+\mathbf{G})\mathbf{P}_\mathbf{N}^\dagger\mathbf{P}_\Gamma^\dagger\mathbf{P}_\mathbf{U}^\dagger\mathbf{F}_\perp. \quad (C26)$$

---

which can be interpreted as a set of matrices describing spectral coherence, each localized to one point source, and all of which are spatially uncorrelated. And likewise, we can write down $\mathbf{G}$ as:

$$\mathbf{G} = \sum_{i,j,k}\left[\lambda_{x_i}\lambda_{y_j}\lambda_{z_k}\mathbf{v}_{x_i}\mathbf{v}_{x_i}^\dagger \otimes \mathbf{v}_{y_j}\mathbf{v}_{y_j}^\dagger \otimes \mathbf{v}_{z_k}\mathbf{v}_{z_k}^\dagger\right]. \quad (C15)$$

We now make two key approximations for the purposes of preconditioning. First, we assume that all the z_k eigenvectors are the same, so $\mathbf{v}_{z_k} \approx \mathbf{v}_{z_{n,k}}$ for all n, all of which are also taken to be the same as the eigenvectors that appear in the preconditioner for $\mathbf{U}$ in Equation C13. Second, as in Section C 2, we are only interested in acting upon the largest eigenvalues of $\mathbf{R}$ and $\mathbf{G}$. To this end, we will ultimately only consider the largest values of $\lambda_{n,k}$ and $\lambda_{i,j,k} \equiv \lambda_{x_i}\lambda_{y_j}\lambda_{z_k}$, which will vastly reduce the computational complexity of the preconditioner.

Our strategy for overcoming the difficulty of the different bases is to simply add the two perpendicular parts of the matrices and then decompose the sum into its eigenvalues and eigenvectors. We therefore define

$$\Gamma \equiv \mathbf{R} + \mathbf{G} \quad (C16)$$

(choosing the symbol Γ because it looks like $\mathbf{R}$ and sounds like $\mathbf{G}$). Given the above approximations, we can reexpress Γ as follows:

$$\Gamma \approx \sum_k \left(\Gamma_{\perp,k} \otimes \mathbf{v}_{z_k}\mathbf{v}_{z_k}^\dagger\right), \quad (C17)$$

where we have defined each $\Gamma_{\perp,k}$ as:

$$\Gamma_{\perp,k} \equiv \left(\sum_n \lambda_{n,k}\overset{\leftrightarrow}{\delta}_{x,x_n} \otimes \overset{\leftrightarrow}{\delta}_{y,y_n}\right) + \left(\sum_{i,j}\lambda_{i,j,k}\mathbf{v}_{x_i}\mathbf{v}_{x_i}^\dagger \otimes \mathbf{v}_{y_j}\mathbf{v}_{y_j}^\dagger\right). \quad (C18)$$

Due to the high spectral coherence of the foregrounds, only a few values of k need to be included to precondition for Γ. Considering the limit on angular box size imposed by the flat sky approximation and the limit on angular resolution imposed by the array size, this should require at most a few eigenvalue determinations of matrices no bigger than about 10^4 entries on a side. Moreover, those eigenvalue decompositions need only be computed once and then only partially stored for future use. In practice, this is not a rate limiting step, as we see in Section III E 2.

We now write down the eigenvalue decomposition of Γ:

$$\Gamma = \sum_k \left(\sum_l \lambda_{l,k}\mathbf{v}_{\perp,l}\mathbf{v}_{\perp,l}^\dagger\right)\mathbf{v}_{z_k}\mathbf{v}_{z_k}^\dagger. \quad (C19)$$

Before we attack the general case, we assume that only one value of $\lambda_{l,k}$ is worth preconditioning—we generalize to the full $\mathbf{P}_\Gamma$ later. We now know that if we have a

Fast Power Spectrum Estimation

1. Generate lots of random data cubes from the
   model covariance, exploiting symmetries

2. Calculate our quadratic estimator

3. Monte Carlo many quadratic estimators to
   get error bars and window functions.

$$\text{Use } \mathrm{Cov}(\mathbf{q}) = \mathbf{F}$$

$$\text{To avoid } F^{\alpha\beta} = \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\mathbf{Q}^{\alpha}\mathbf{C}^{-1}\mathbf{Q}^{\beta}\right]$$
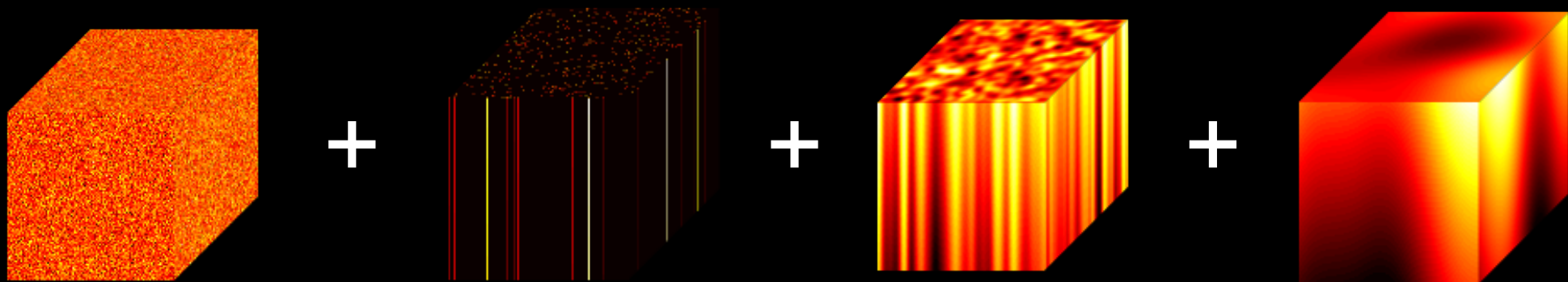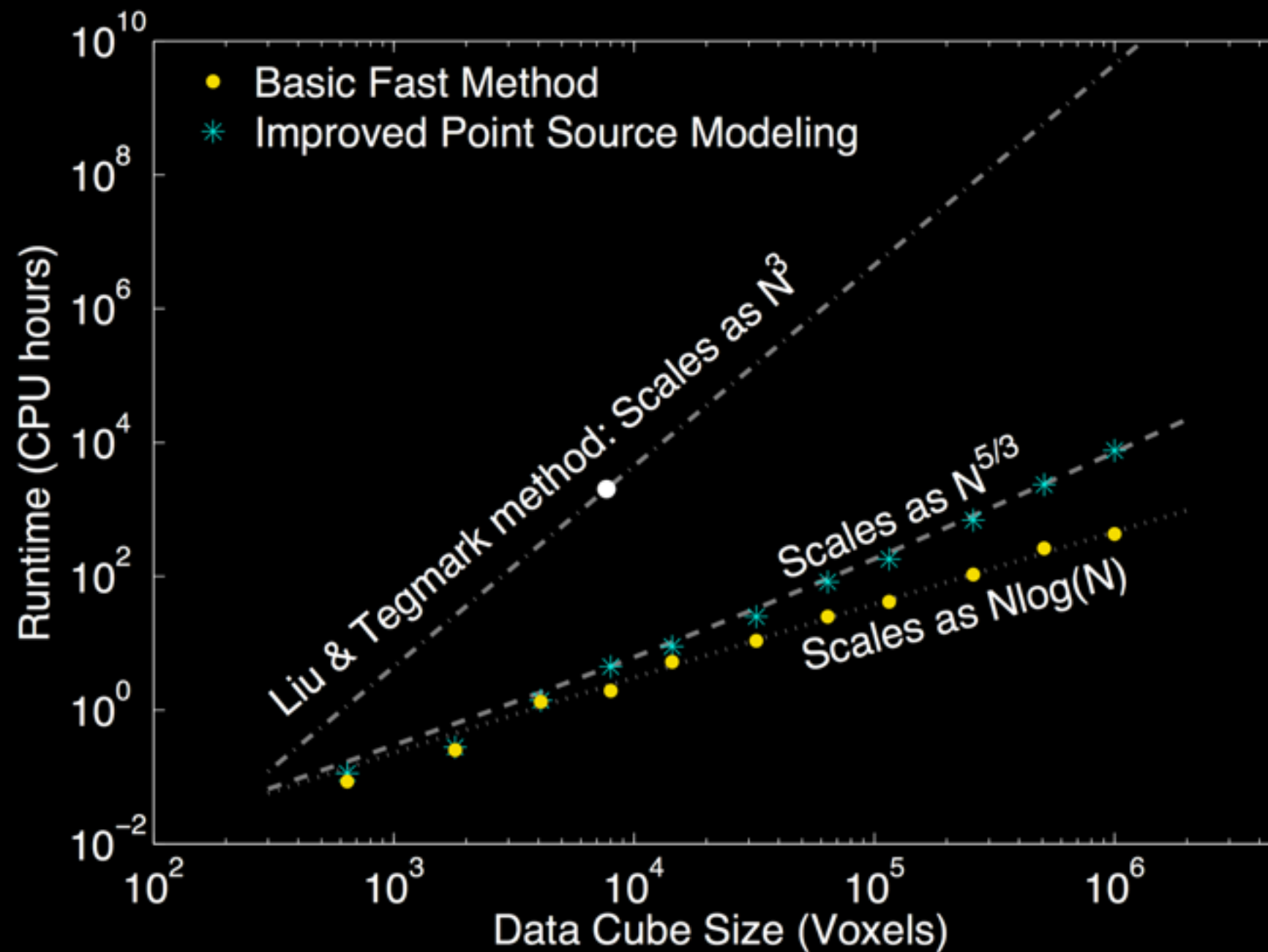
All in O(NlogN)*

*Dillon, Liu, &
Tegmark (2013)*

It works as fast as advertised.

Dillon, Liu, & Tegmark (2013)

This is timely because there's there's a lot of 21 cm interferometers now up and running.

MWA

PAPER

GMRT

LOFAR

21CMA

MITEoR

And lots of related experiments in 21 cm Cosmology:
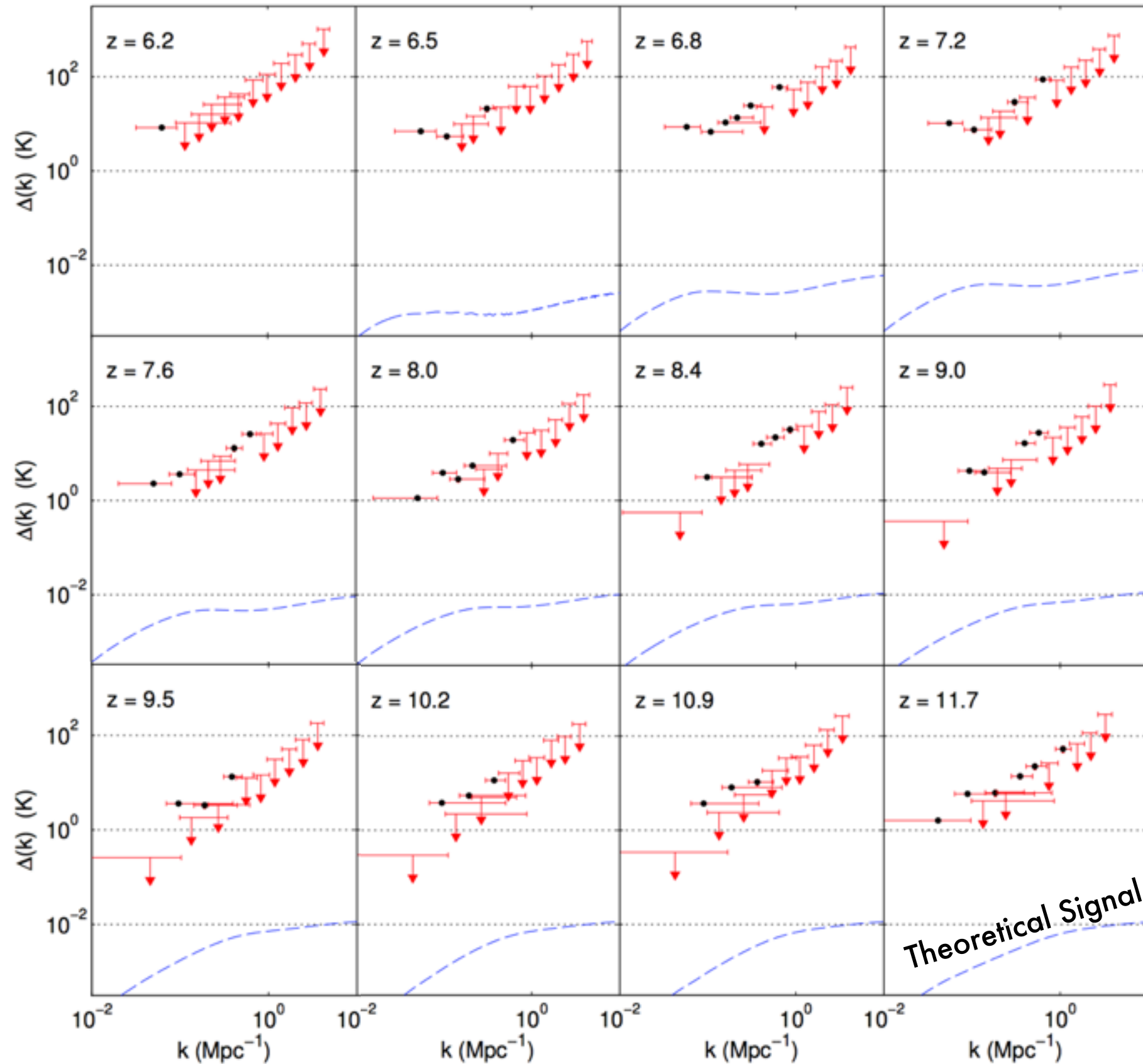ASKAP, BAOBAB, BINGO, CHIME, CRT, DARE, EDGES, EMBRACE, GBT, KAT-7, LEDA, LWA, MeerKAT, SKA...and more

The first application of our power spectrum estimation technique was to the MWA.

- Data taken in March 2010 with the Murchison Widefield Array 32 tile prototype array in Western Australia

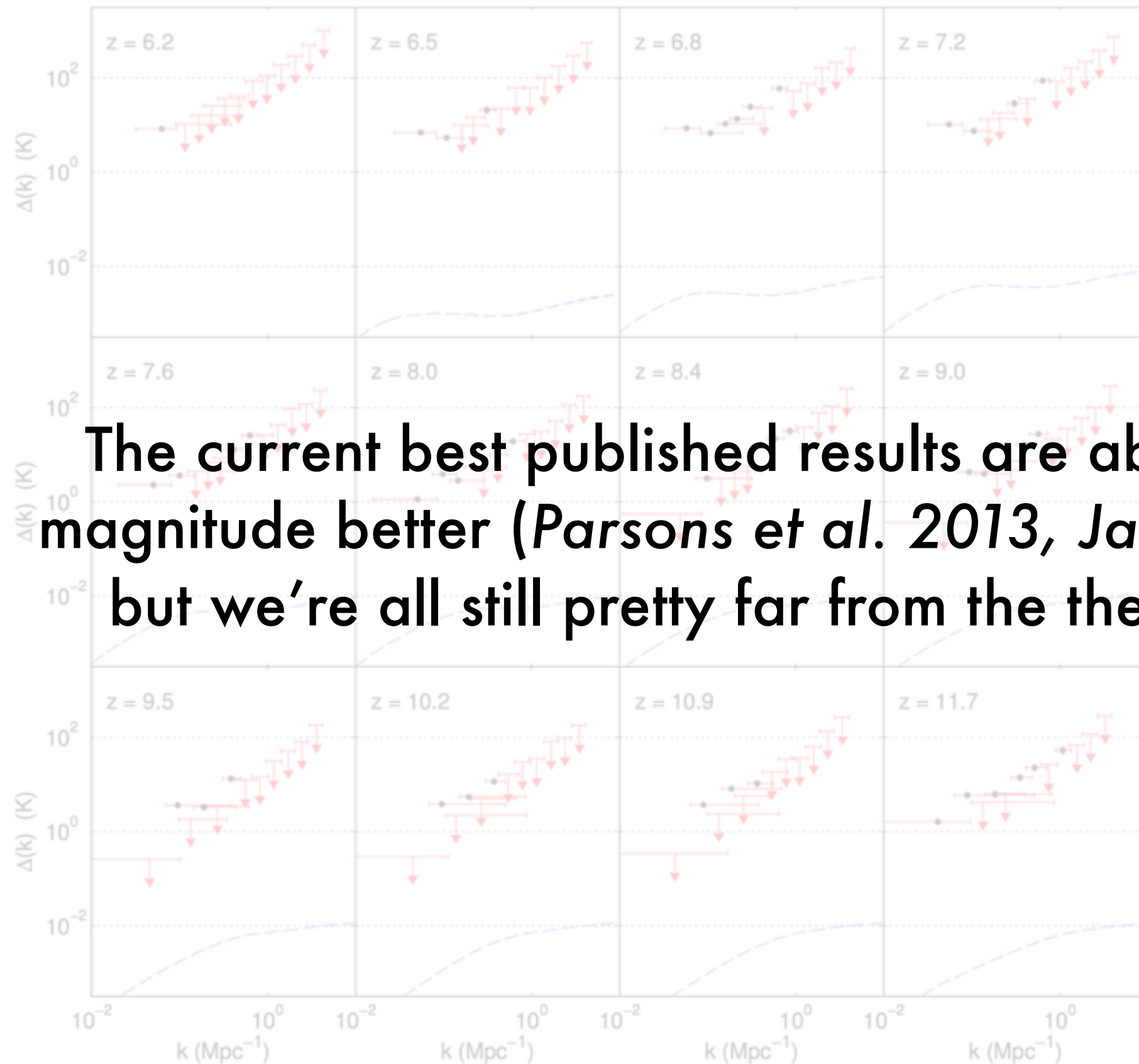- Approximately 3-5 hours of observation per frequency band

Image:

Results:

We set power spectrum limits across a wide range of scales and redshifts.

Dillon et al. (2014a)

The current best published results are about an order of magnitude better (*Parsons et al. 2013, Jacobs et al. 2014*), but we're all still pretty far from the theoretical signal.
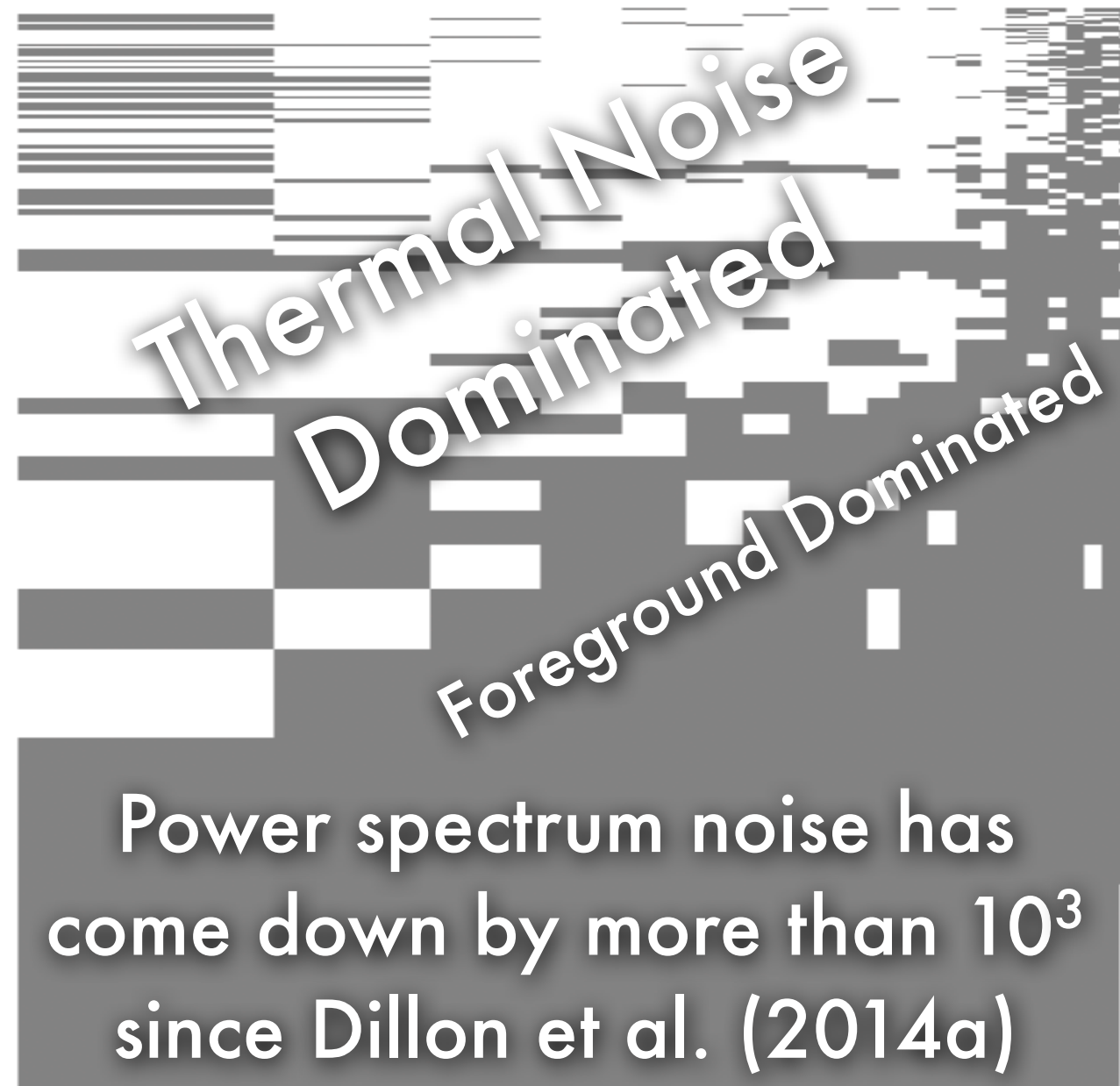
Dillon et al. (2014a)

MWA 128-Tile Preliminary Data

Buffer

Horizon Wedge

Primary Beam Wedge

[Preliminary Data Removed From Online Talk]

MWA 128-Tile Preliminary Data

Thermal Noise Dominated

Foreground Dominated

Power spectrum noise has come down by more than 10^3 since Dillon et al. (2014a)

[Preliminary Data Removed From Online Talk]

Work is still ongoing.

- Further refining calibration

- Integrating down with more data, while performing jackknife tests of data quality

- Improving our foreground and foreground uncertainty models

We also need to understand our maps.

The Instrument   True Sky

Noise where:

Measurements $\longrightarrow$ $$\mathbf{y} = \mathbf{Ax} + \mathbf{n}$$ $\langle \mathbf{nn}^\dagger \rangle \equiv \mathbf{N}$

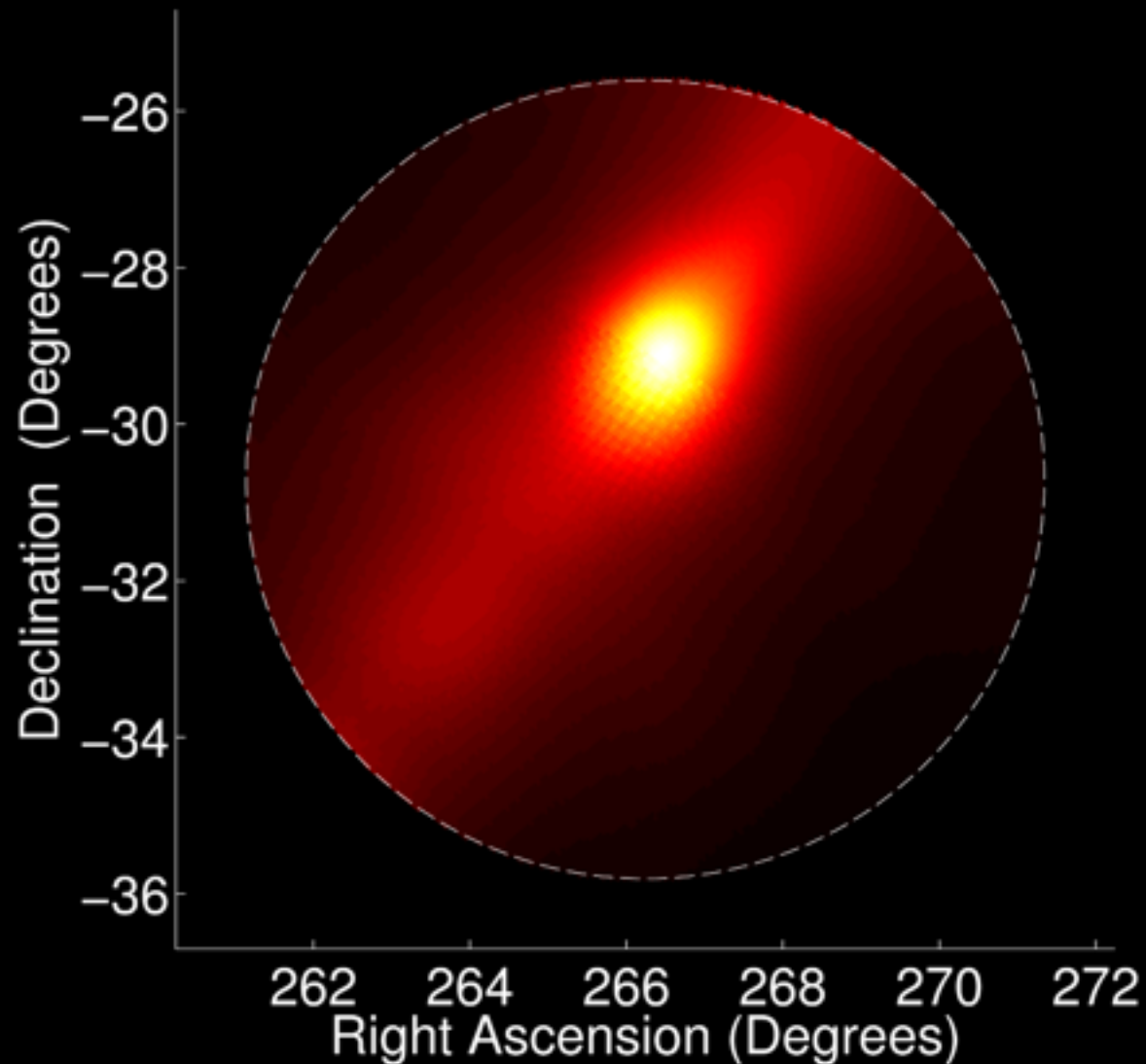Optimal Map Estimator $\longrightarrow$ $$\widehat{\mathbf{x}} = \mathbf{DA}^\dagger \mathbf{N}^{-1} \mathbf{y}$$

Normalization

We also need to understand our maps.

The Instrument   True Sky

Noise where:

Measurements $\longrightarrow$ $$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \langle \mathbf{n}\mathbf{n}^\dagger \rangle \equiv \mathbf{N}$$

Optimal Map
Estimator $\longrightarrow$ $$\widehat{\mathbf{x}} = \mathbf{D}\mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{y}$$

Normalization

$$\langle \widehat{\mathbf{x}} \rangle = \mathbf{D}\mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A}\mathbf{x} \equiv \mathbf{P}\mathbf{x}$$

Matrix of PSFs

Dillon et al. (2014b)

Our maps have different statistics than the true sky.

True Sky: $\mathbf{x}$

"Dirty" Map: $\langle \widehat{\mathbf{x}} \rangle = \mathbf{Px}$

Dillon et al. (2014b)
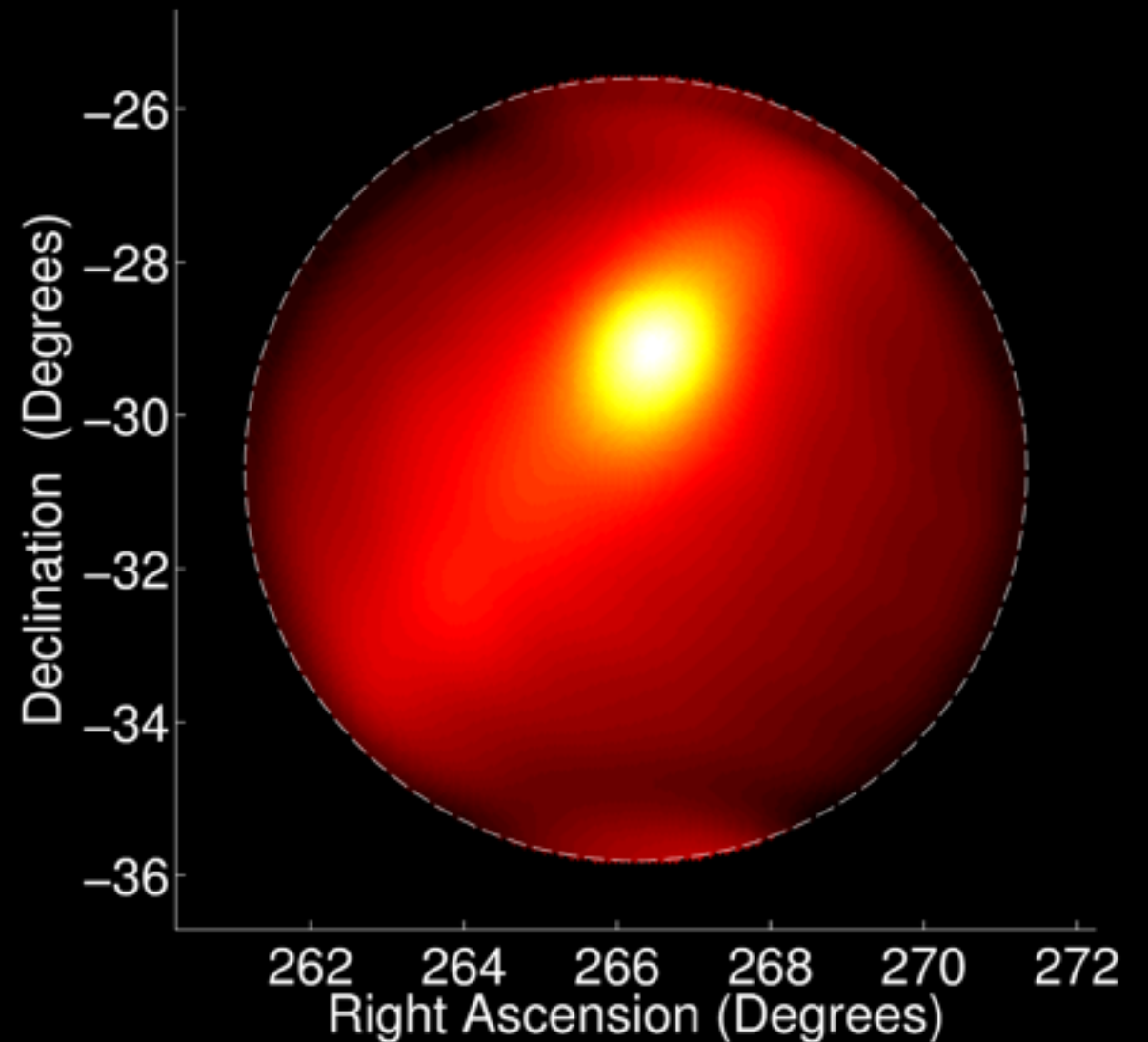
Our maps have different statistics than the true sky.

- We need to know $\mathbf{P}$ to estimate power spectra and model foregrounds.

- Nominally, $\mathbf{P}$ maps every point on the true sky to every point in the dirty map at every frequency and knows about every observation…so it's hard to calculate.
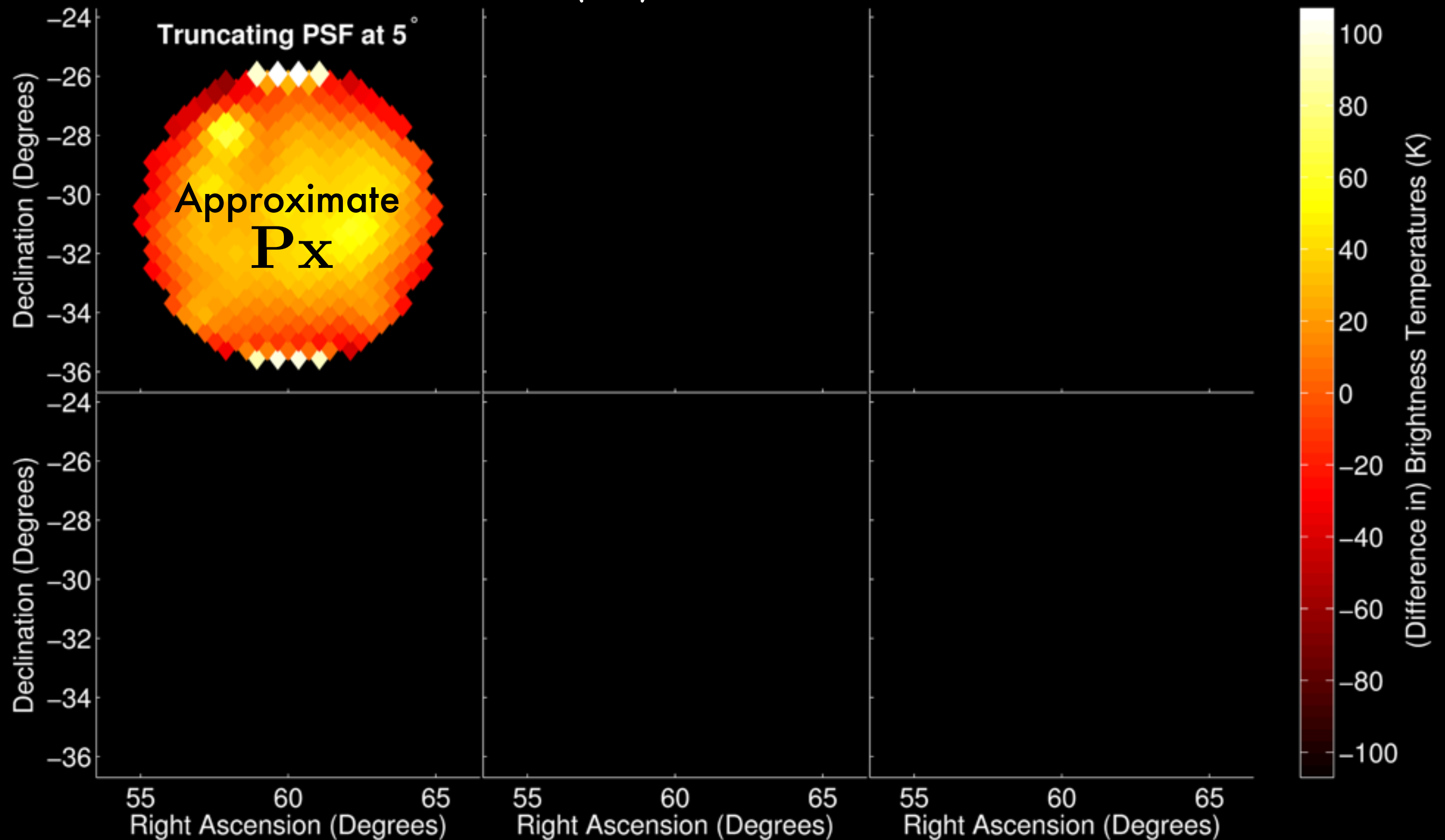
Dillon et al. (2014b)

Three ways to make it faster...

1. Truncating the PSF.

2. Combing together multiple sequential observations.

3. Fitting the PSF's translational variations with low-order polynomials.

All have speed vs. accuracy tradeoffs.

Dillon et al. (2014b)

Truncating the PSF trades speed for accuracy.

$$\langle \widehat{\mathrm{x}} \rangle = \mathbf{P}\mathbf{x}$$

Dillon et al. (2014b)

Truncating the PSF trades speed for accuracy.
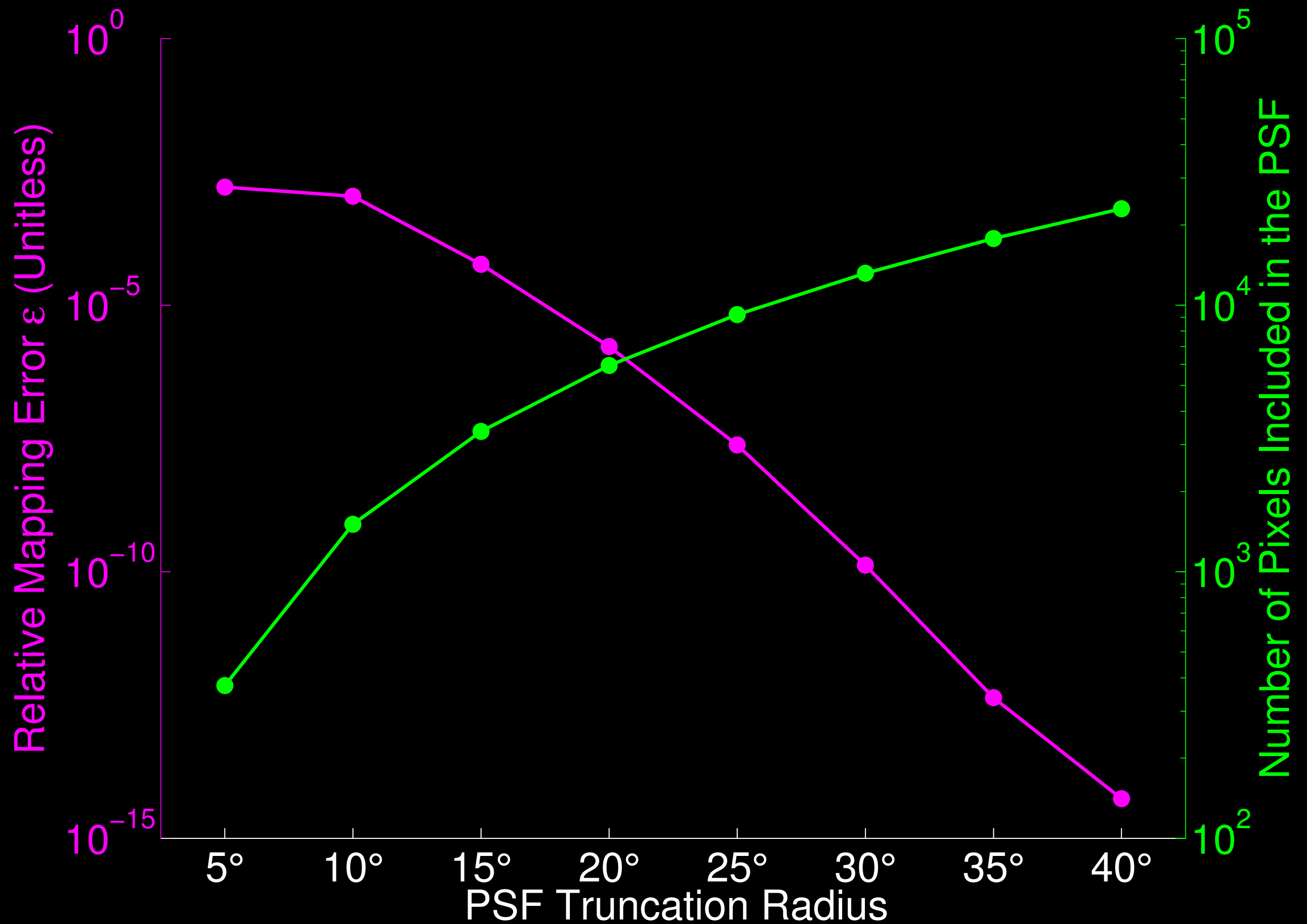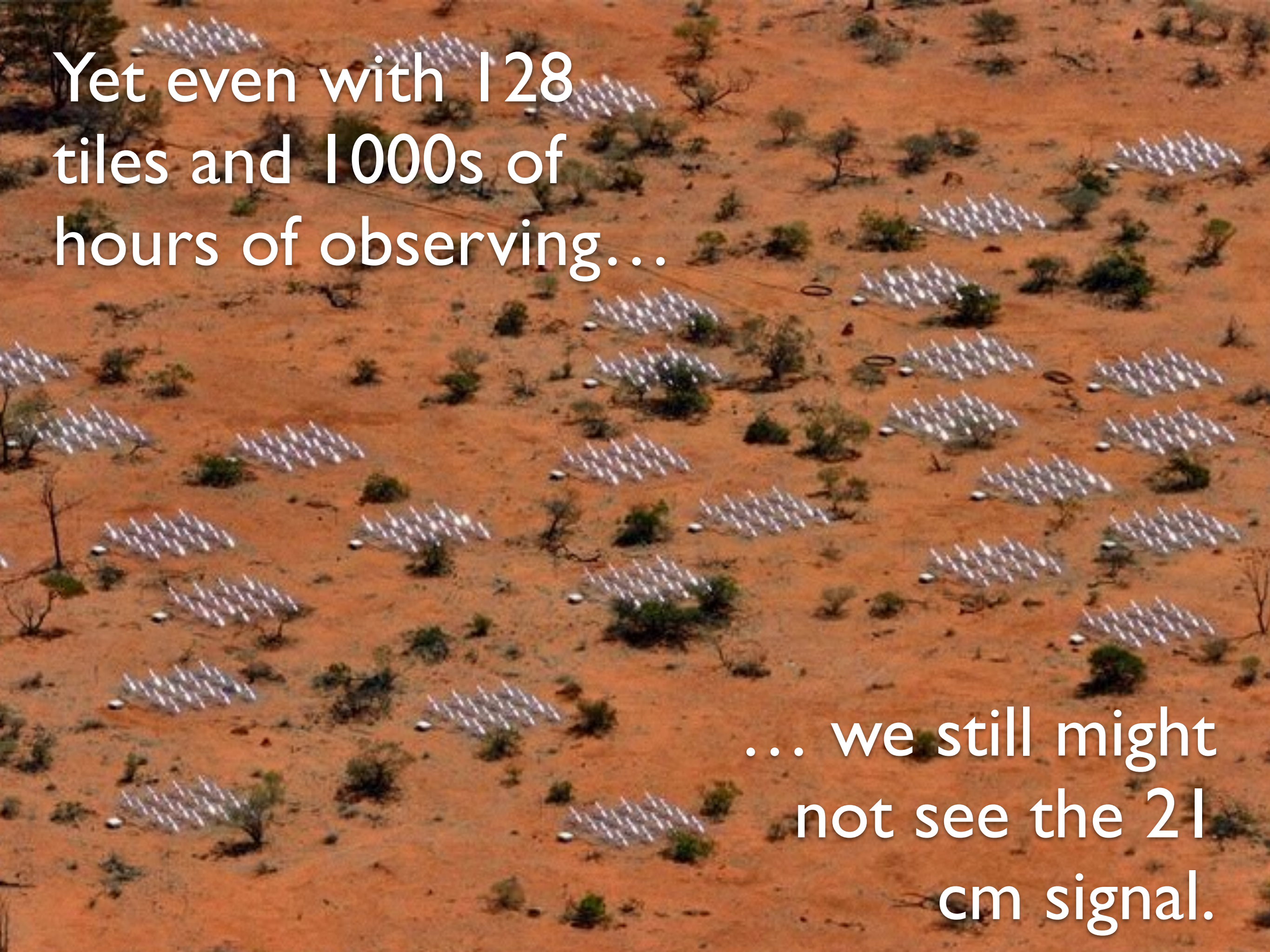
Dillon et al. (2014b)

What's next?

- End-to-end simulation incorporating optimal mapmaking with quadratic power spectrum estimators.

- Apply an integrated mapmaking and power spectrum pipeline to real data to better keep the EoR window clean and to try to subtract foregrounds.

Yet even with 128 tiles and 1000s of hours of observing…

… we still might not see the 21 cm signal.

How do we build a more sensitive telescope?

As is usually the answer in astronomy:

Go bigger.
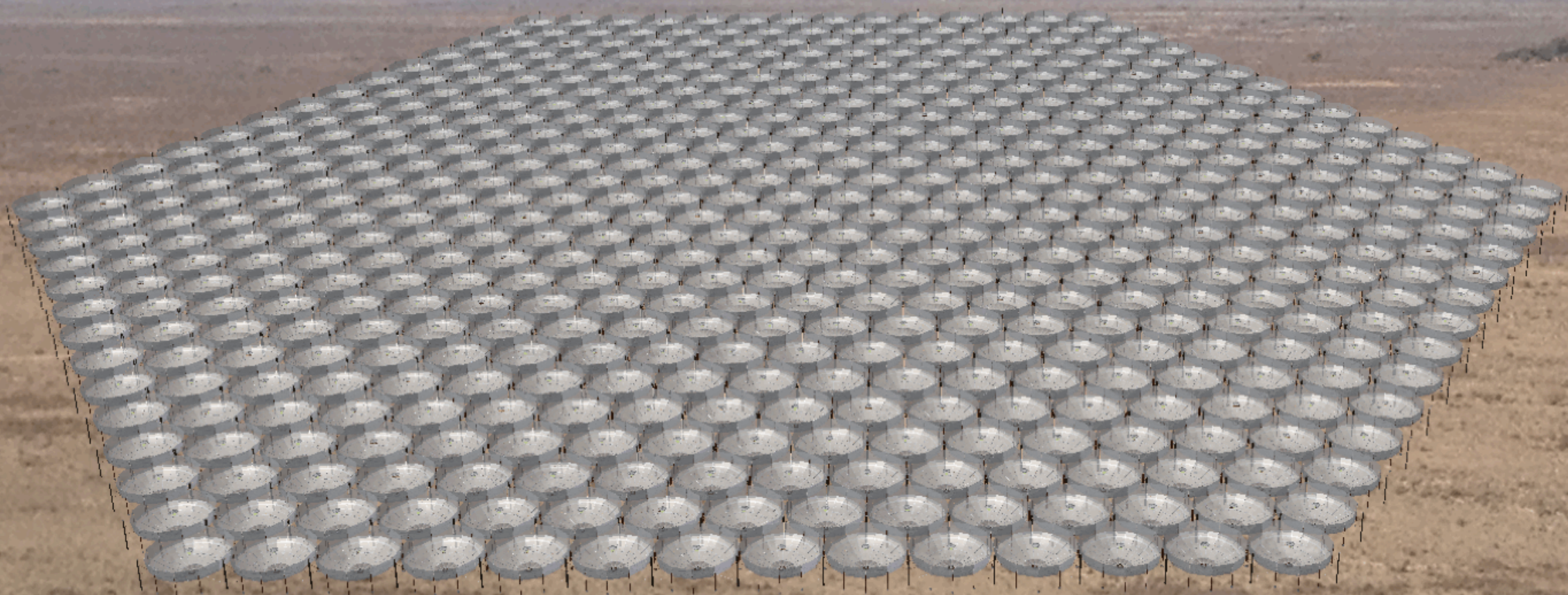
HERA
The Hydrogen Epoch of Reionization Array
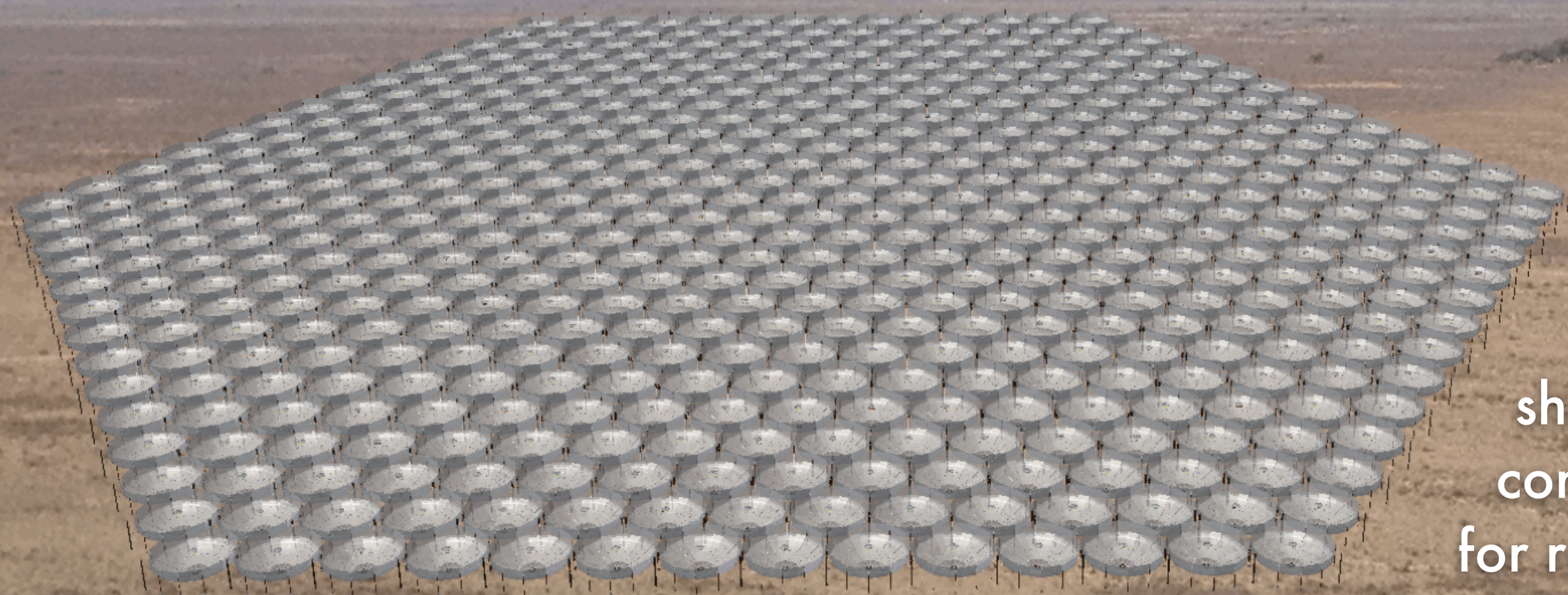
shipping container, for reference

HERA

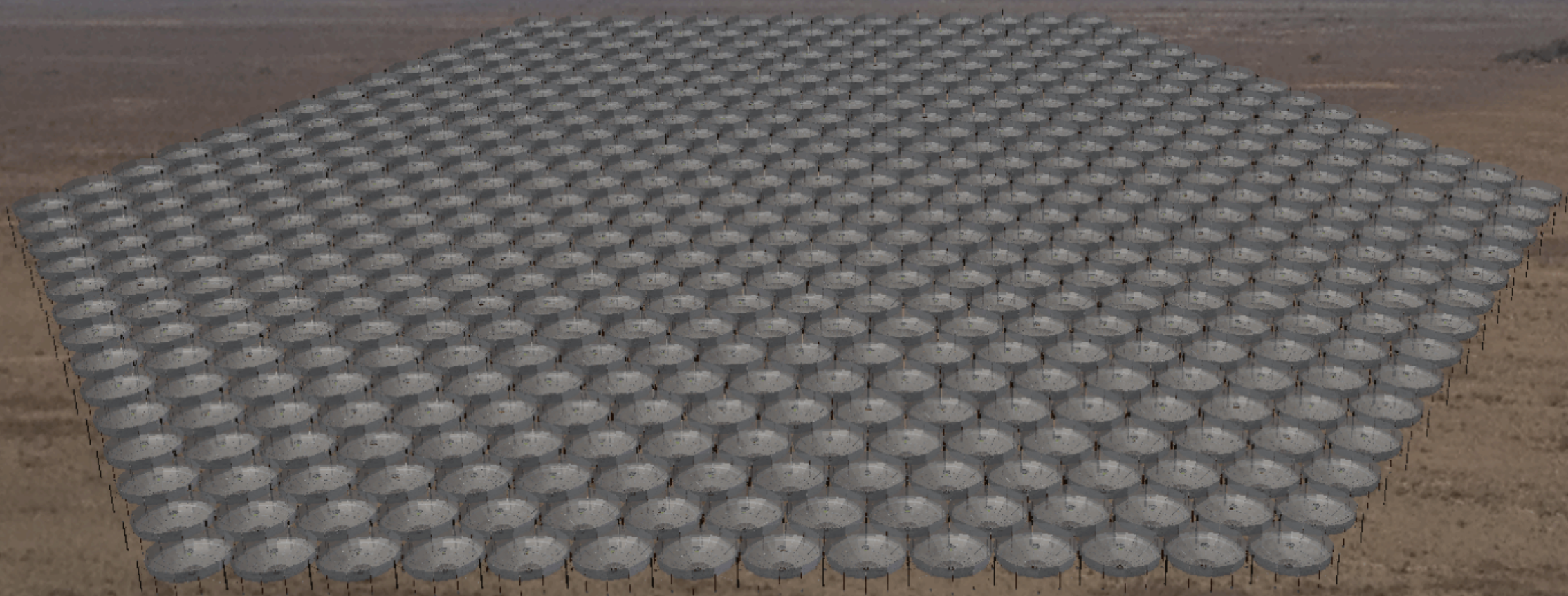brings together teams and technical lessons from PAPER, MWA, and MITEoR.
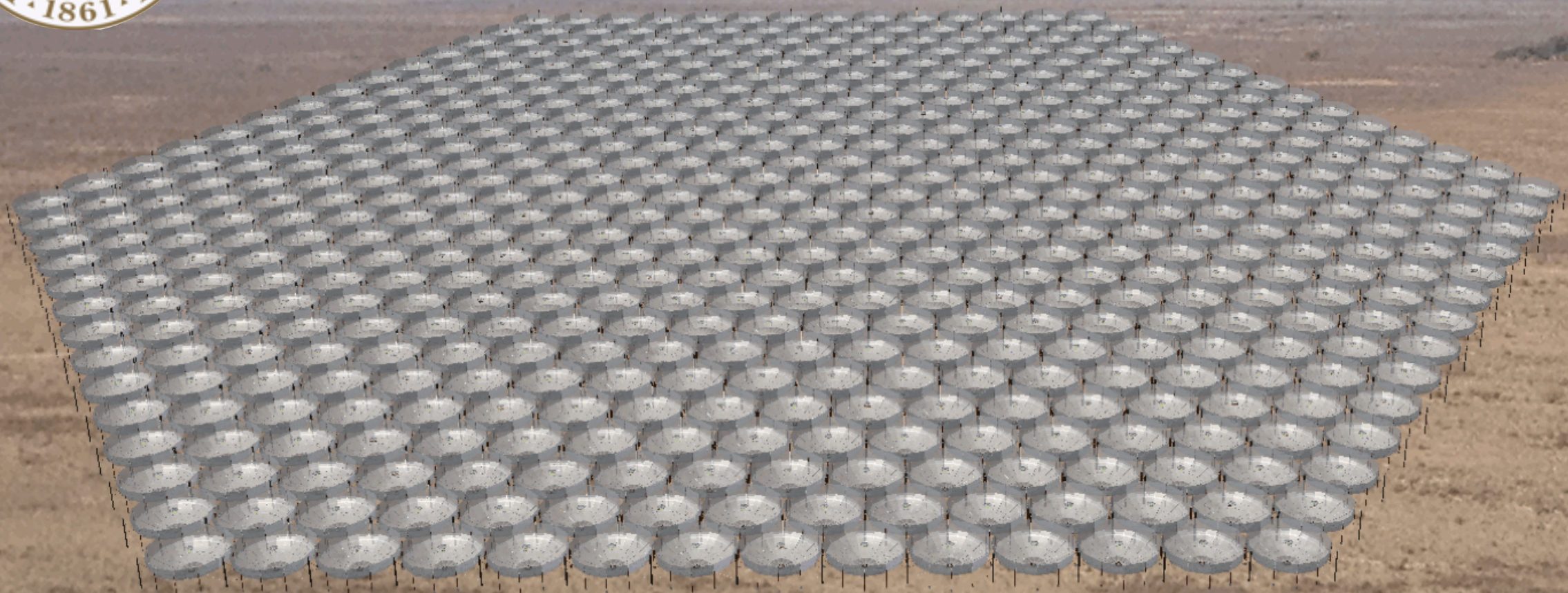
HERA

brings together teams and technical
lessons from PAPER, MWA, and MITEoR.

HERA will have:

- At least 331 stationary dishes that vastly increase sensitivity (at the cost of field of view).

HERA prototype element in the foothills behind Berkeley.

14 m diameter dishes

The HERA Stripe

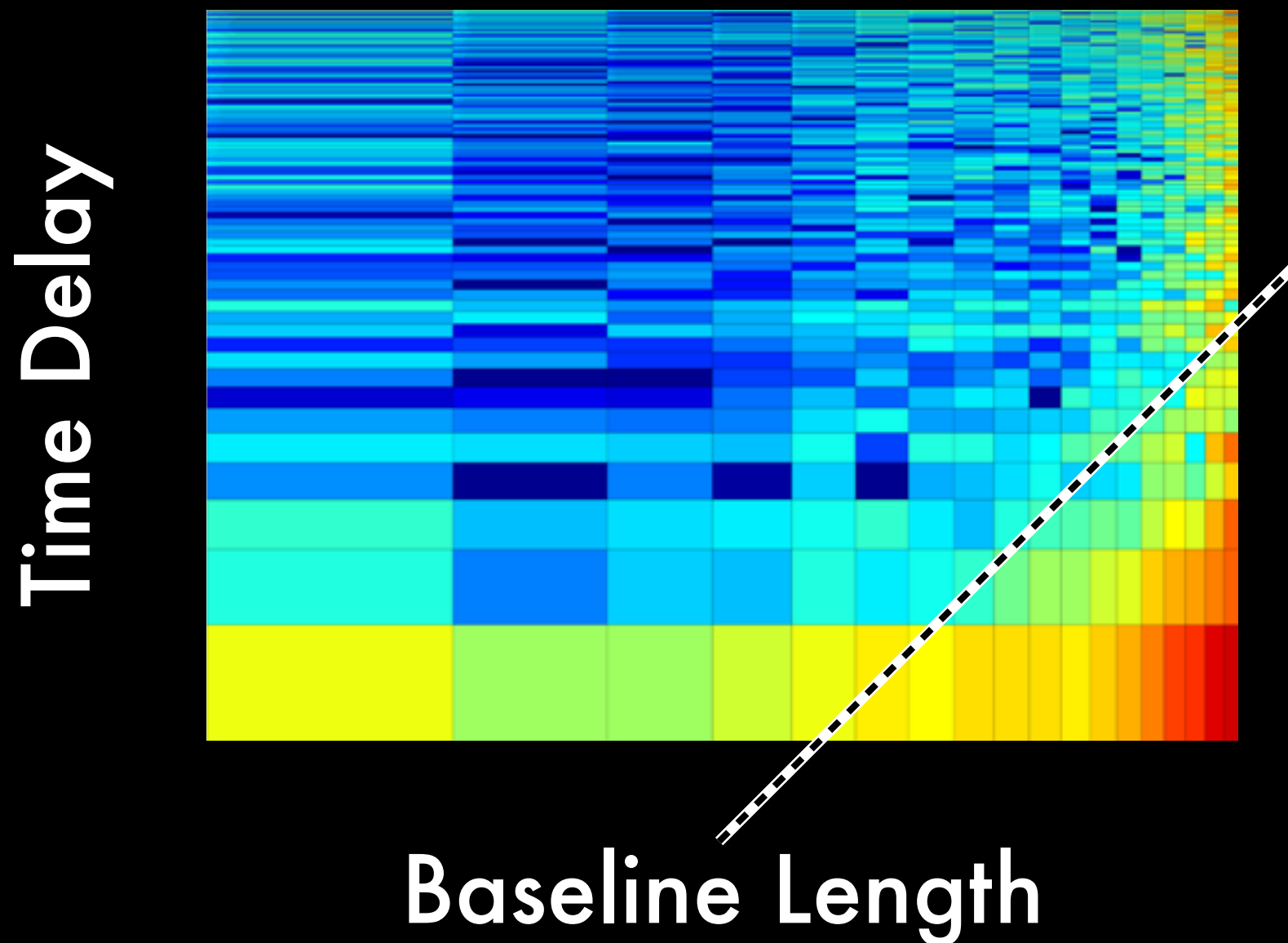HERA is a drift scan instrument that maps out a stripe of constant declination.

HERA will have:

- At least 331 stationary dishes that vastly increase sensitivity (at the cost of field of view).

- A maximally packed configuration with short baselines that are less foreground-contaminated.
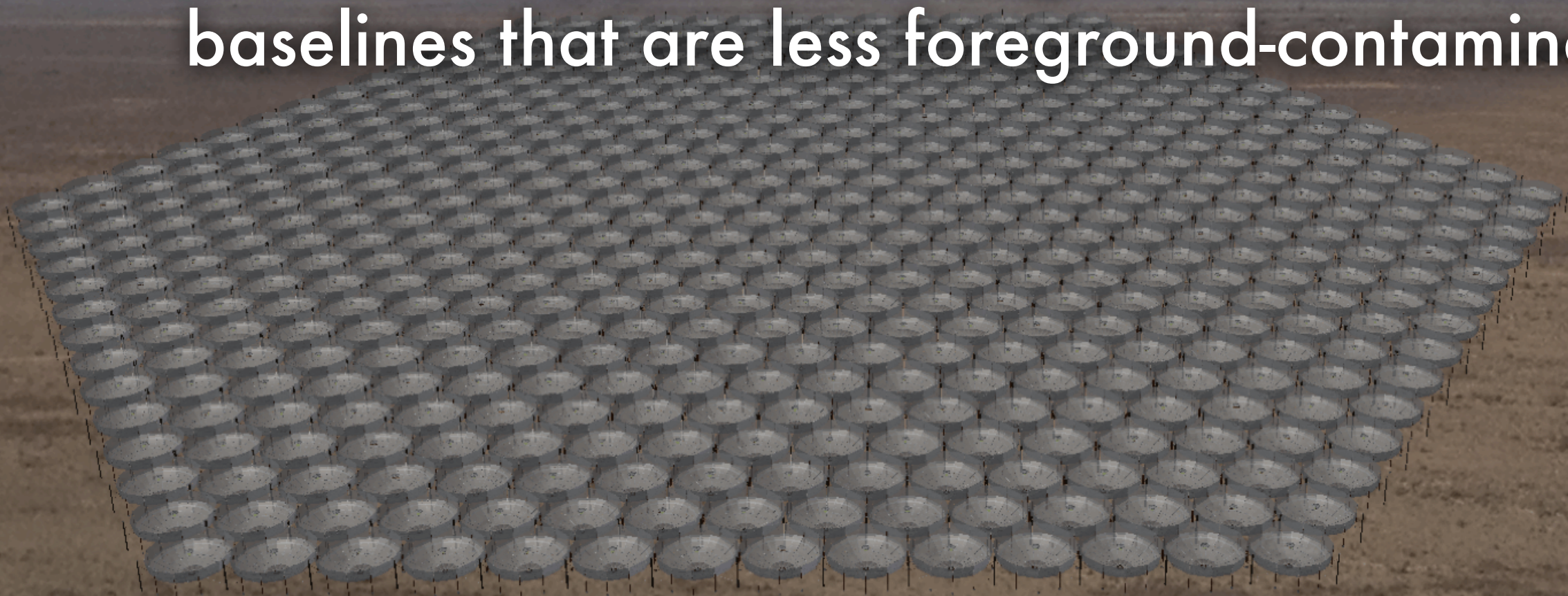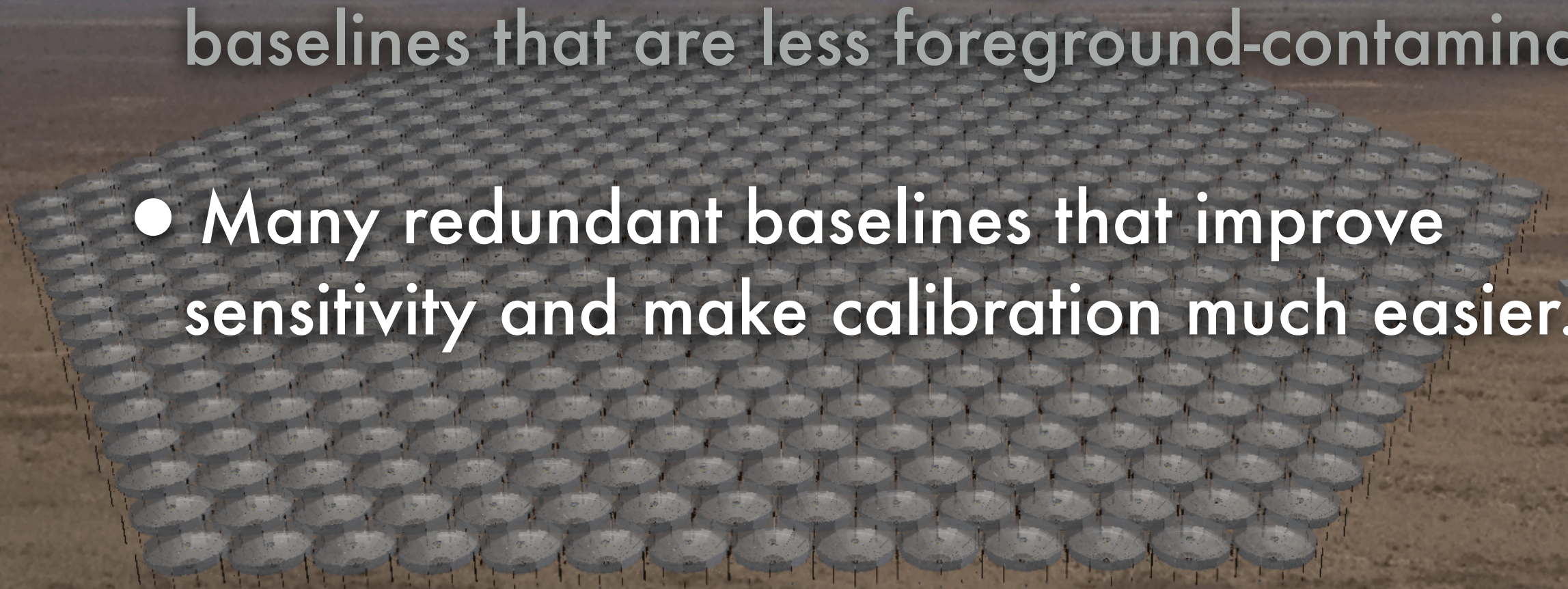
Recall, shorter baselines have "less wedge" in them.

HERA will have:

- At least 331 stationary dishes that vastly increase sensitivity (at the cost of field of view).

- A maximally packed configuration with short baselines that are less foreground-contaminated.

- Many redundant baselines that improve sensitivity and make calibration much easier.
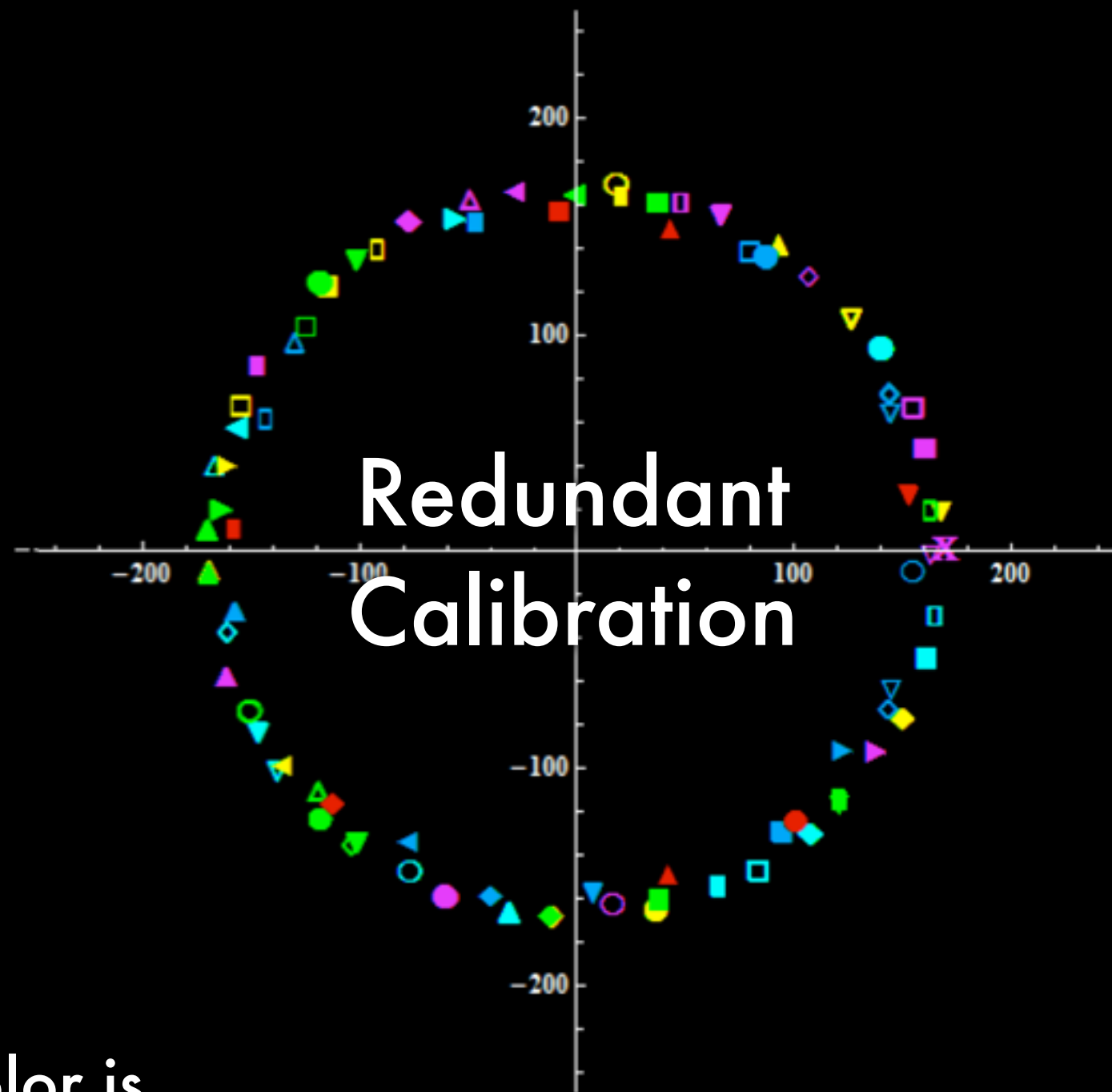
Redundant baselines make the precise calibration necessary for 21 cm tomography much easier.

Liu et al. (2010)
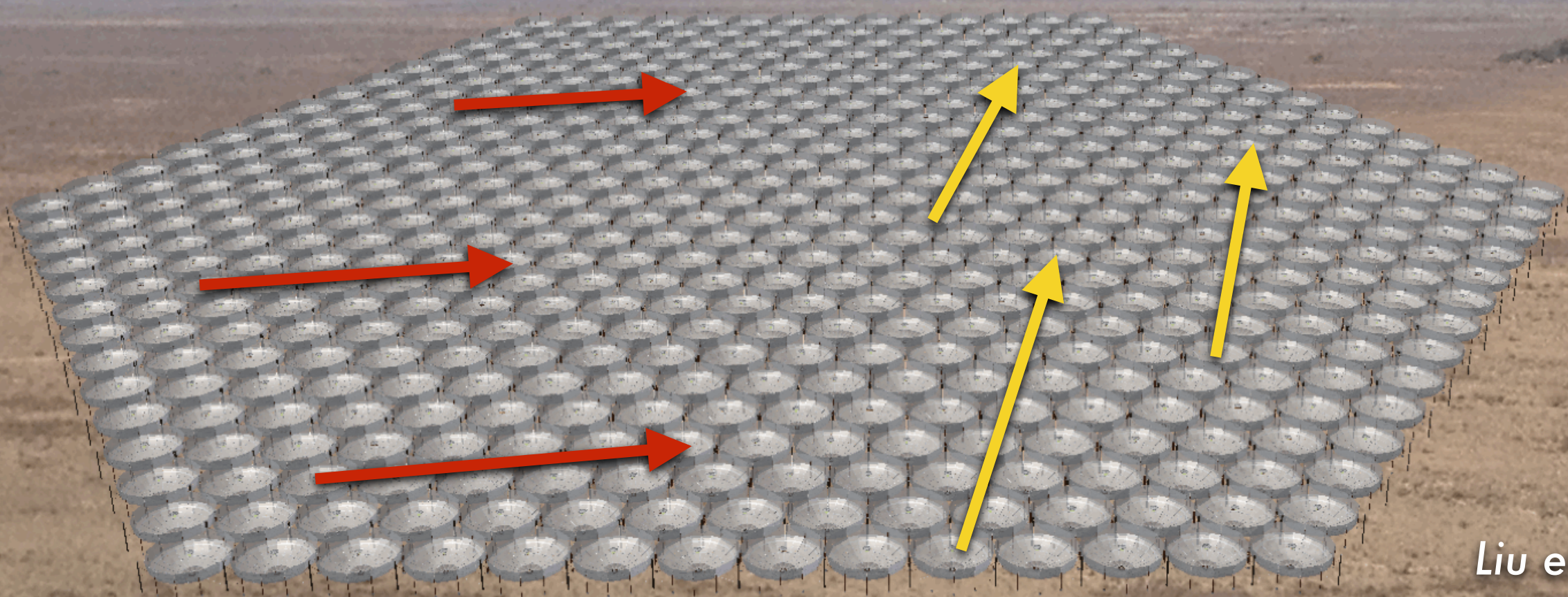
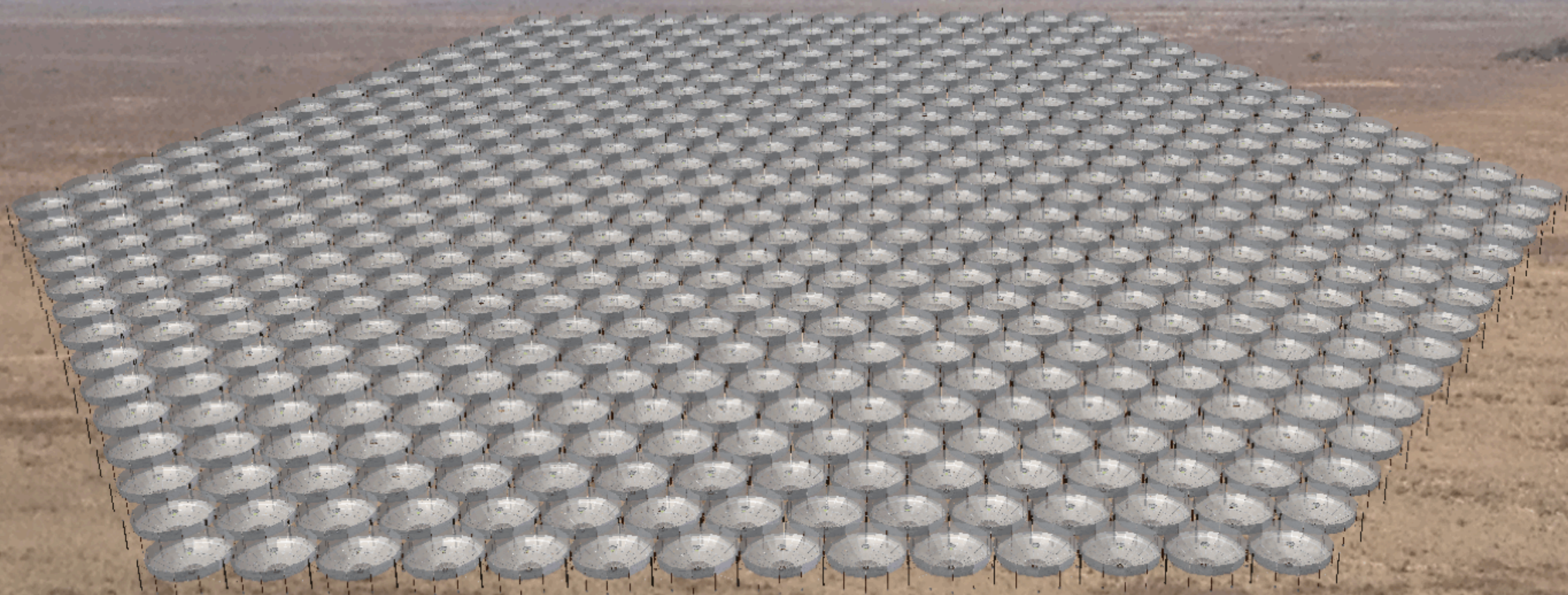# MITEoR: a prototype highly-scalable interferometer for 21 cm cosmology.

Redundant baselines allow us to quickly and precisely calibrate the amplitudes and phases of every antenna.

Redundant Calibration

Each shape/color is
a unique baseline.

Zheng et al. (2014)

So, what can we expect
to see with HERA?

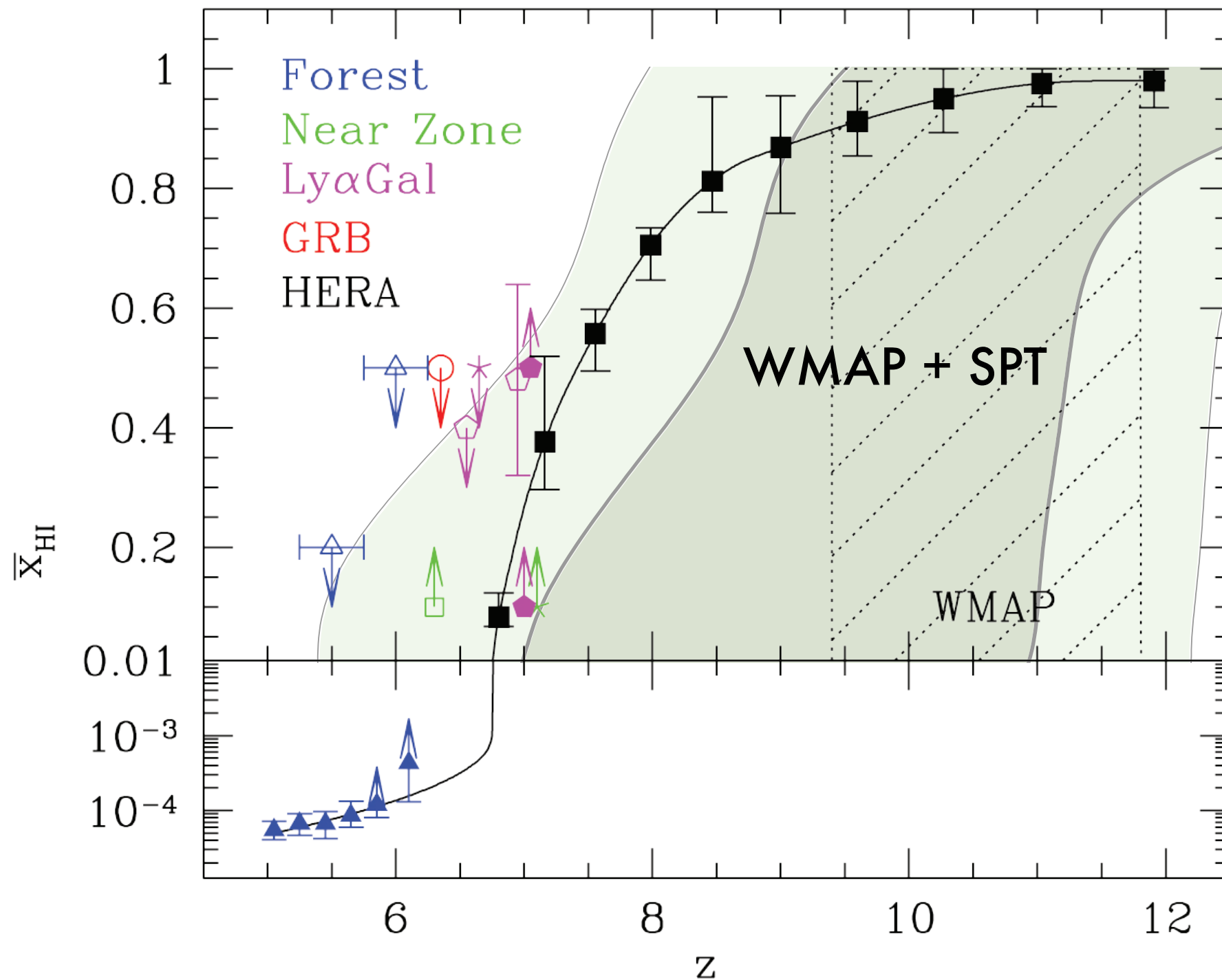# We'll constrain the ionization history of the universe...

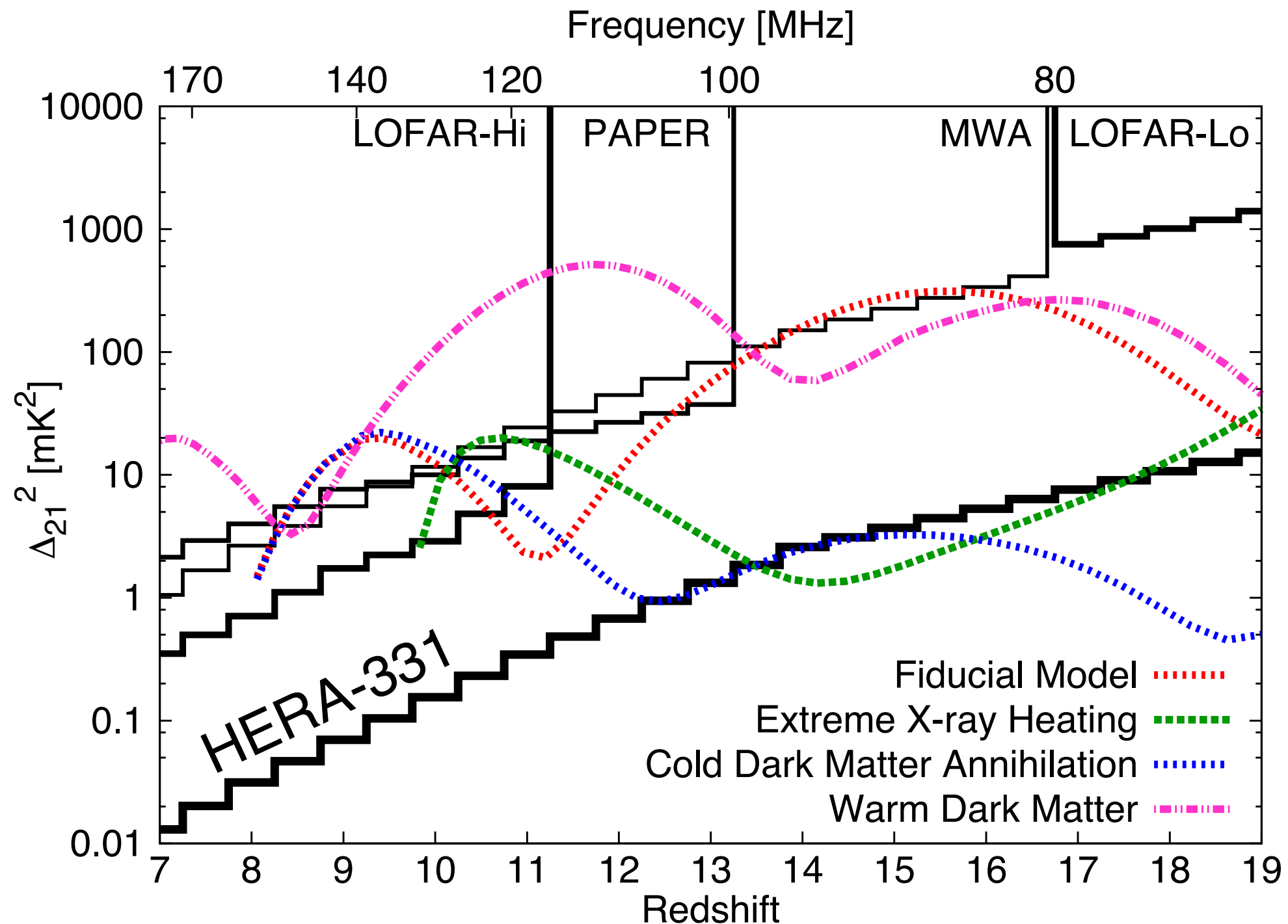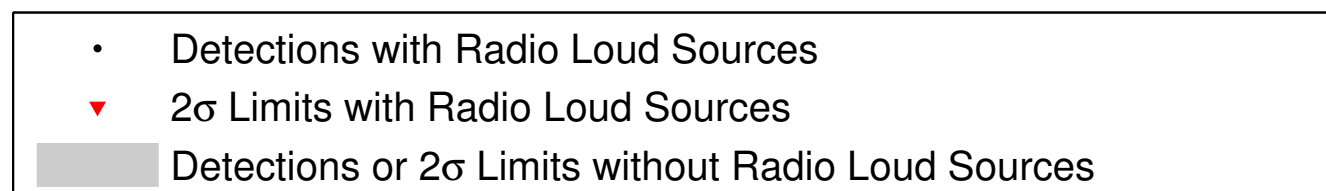Figure: Judd Bowman + Zahn et al. (2012)
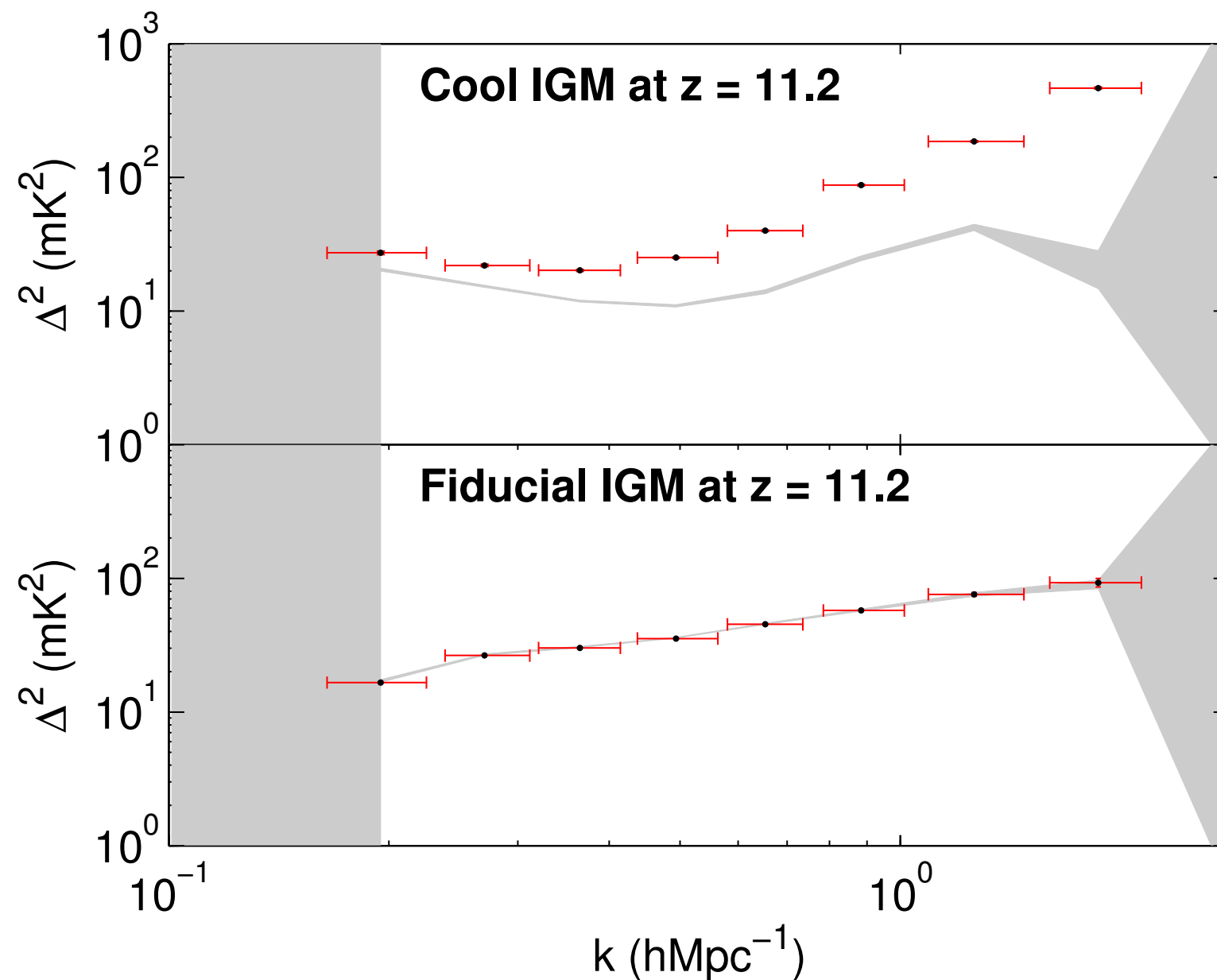
...and its thermal history


Figure: Aaron Ewall-Wice

We'll constrain X-ray heating and the population of high redshift quasars via the 21 cm forest.

*Ewall-Wice,*
Dillon, et al.
(2013)

And we'll also provide the first tight constraints on the astrophysics underlying reionization.

- ζ: Ionizing efficiency

- R_{mfp}: Mean free path of ionizing photons

- T_{vir}: Minimum virial temperature (and thus mass) of ionizing galaxies

21cmFAST

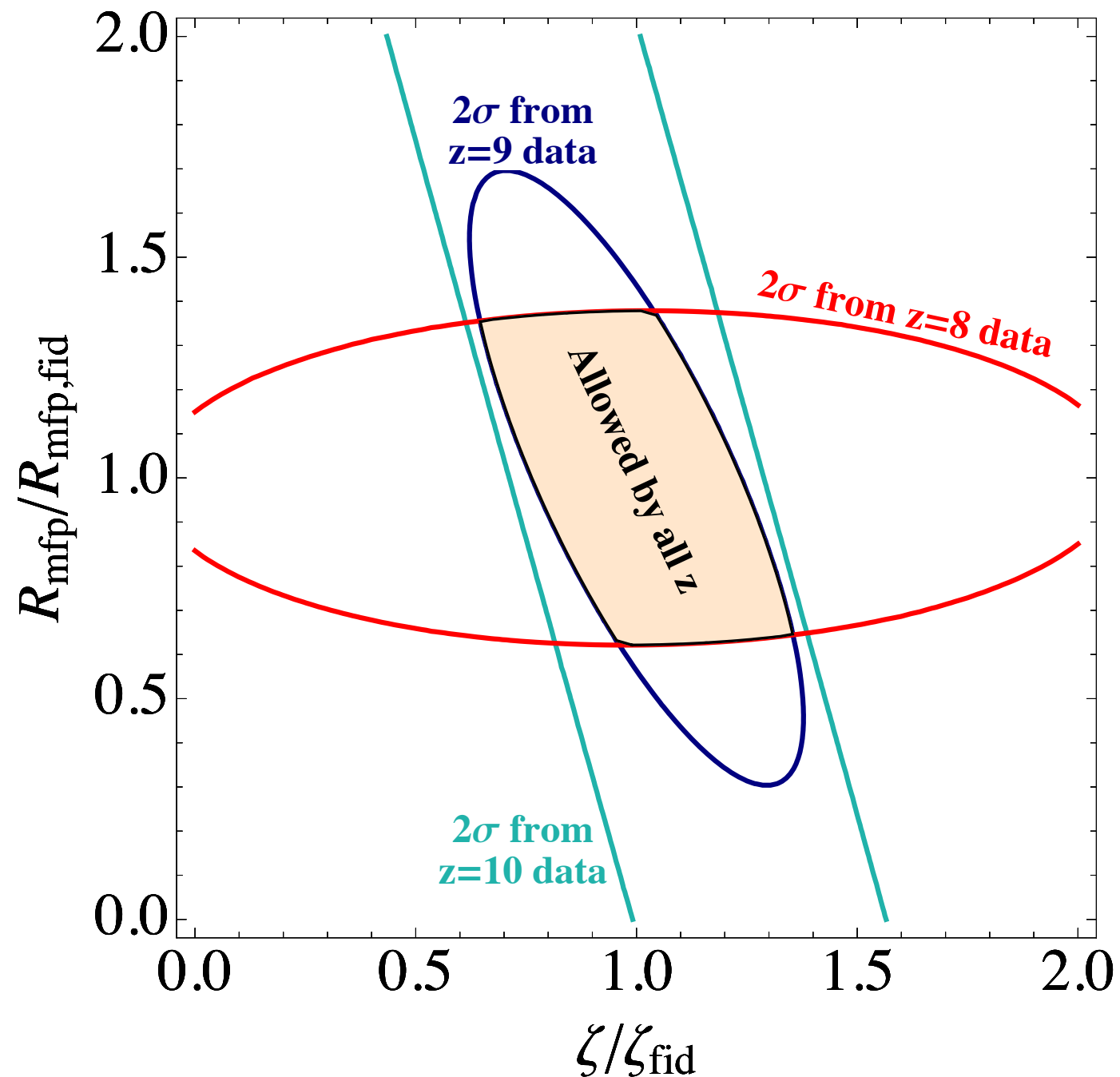*Messinger et al. (2010)*
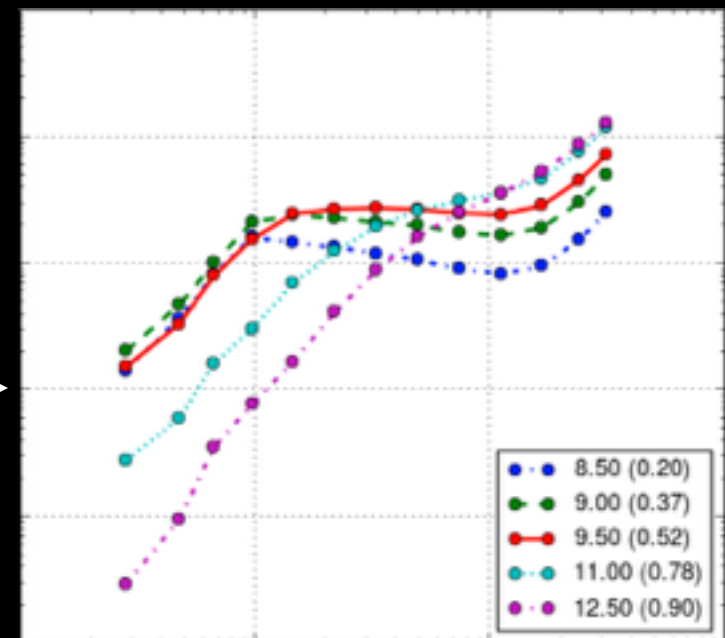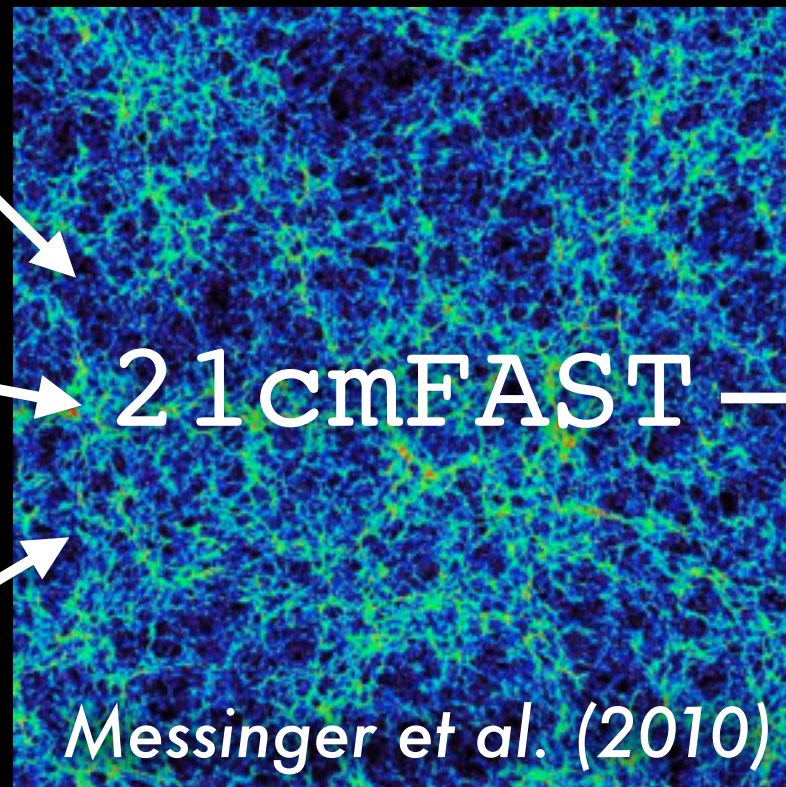
Qualitatively different 1D Power Spectra as a function of k and z

Pober, Liu, Dillon, et al. (2013)

Using a Fisher matrix analysis, we can jointly constrain all three parameters...

Pober, Liu, Dillon, et al. (2013)
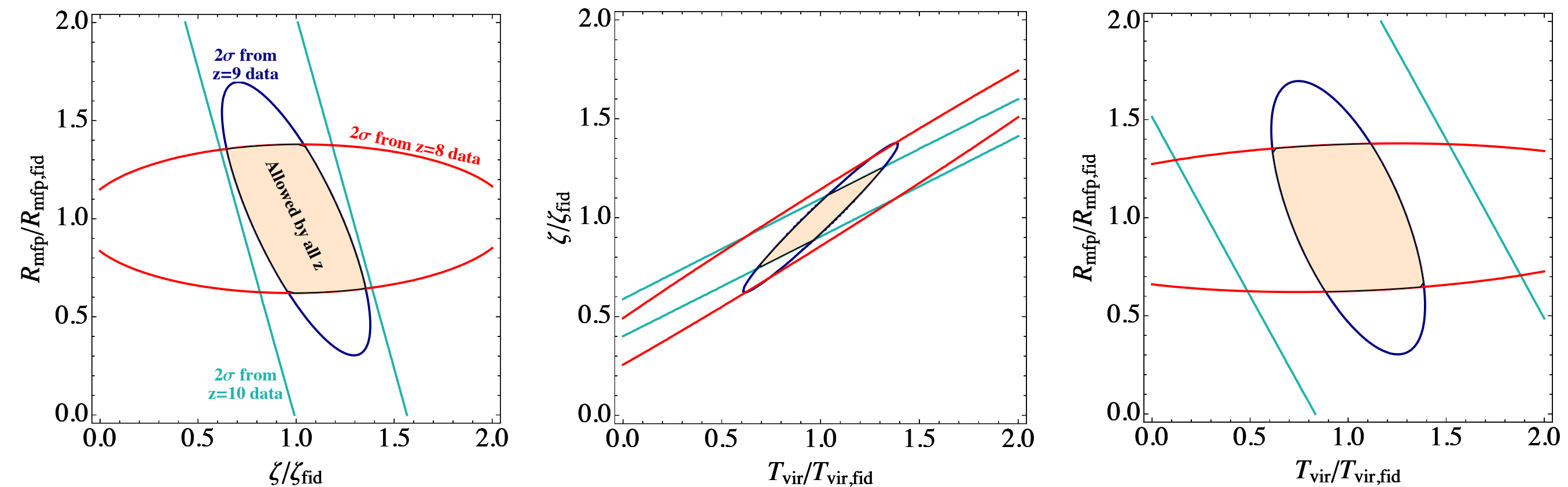
...and break degeneracies using information from multiple redshifts.
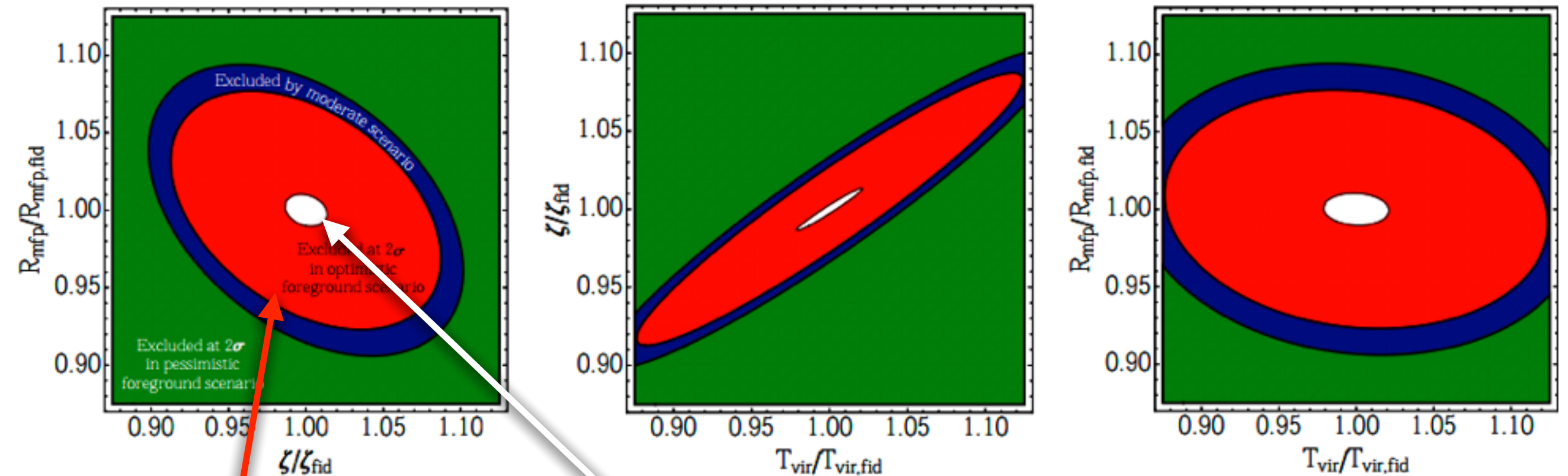

Pober, Liu, Dillon, et al. (2013)

And if we can get better and foreground subtraction and work within the wedge...
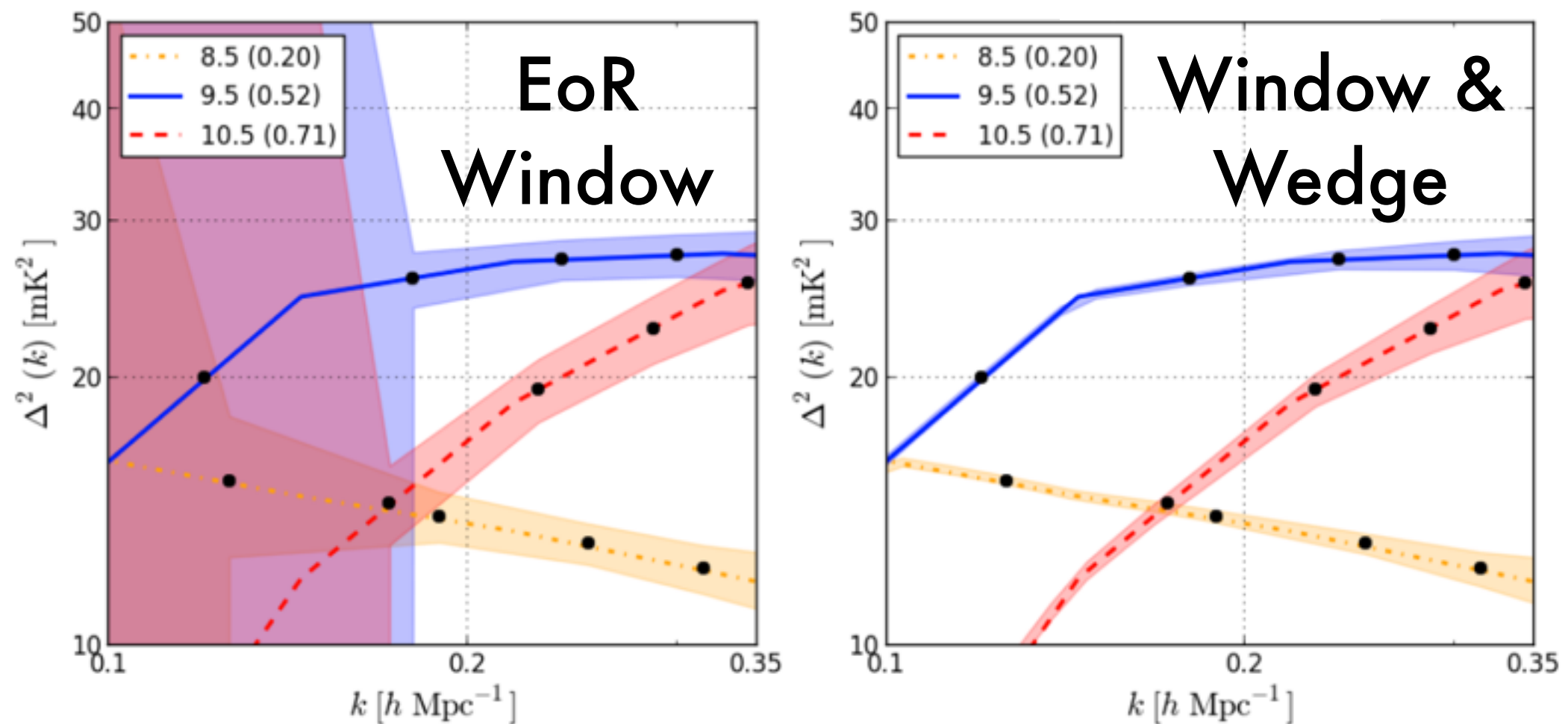
This is no small task! We'll need even better statistical algorithms and a precise understanding of both foregrounds and our instrument.

Pober, Liu, Dillon, et al. (2013)

…we can improve the parameter constraints from ~5% to ~1%.

Working inside the EoR window.

Working inside the wedge too.

(These are parameters still unconstrained by an order of magnitude or more.)

Pober, Liu, Dillon, et al. (2013)

Next steps…

- What degeneracies exist between cosmological parameters and reionization parameters using 21 cm tomography?

- How can other cosmological probes complement and be complemented by 21 cm?

 - This hasn't been investigated in the context of the EoR Window.

In Conclusion

- 21 cm Tomography will open up a huge volume of the universe during the unexplored "Cosmic Dawn."

- Maps and power spectrum measurements require careful, rigorous statistics and new, fast algorithms.

- We've already made great progress with the MWA, setting upper limits over many redshifts.

- HERA will draw on the lessons of MWA, PAPER, and MITEoR with vastly increased sensitivity and can convincingly detect the EoR and tightly constrain the physics behind reionization and the Cosmic Dawn.
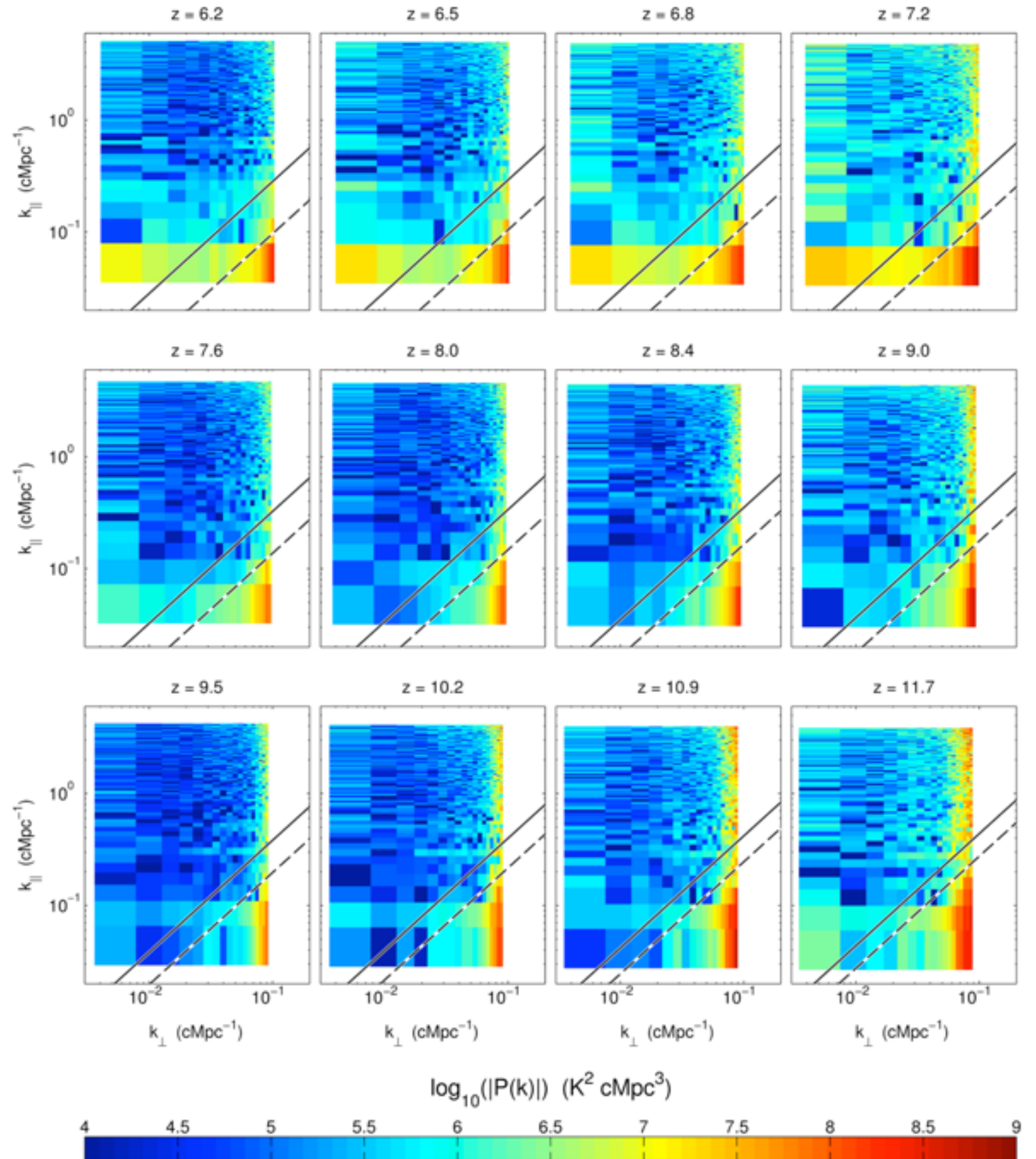
Backup Slides

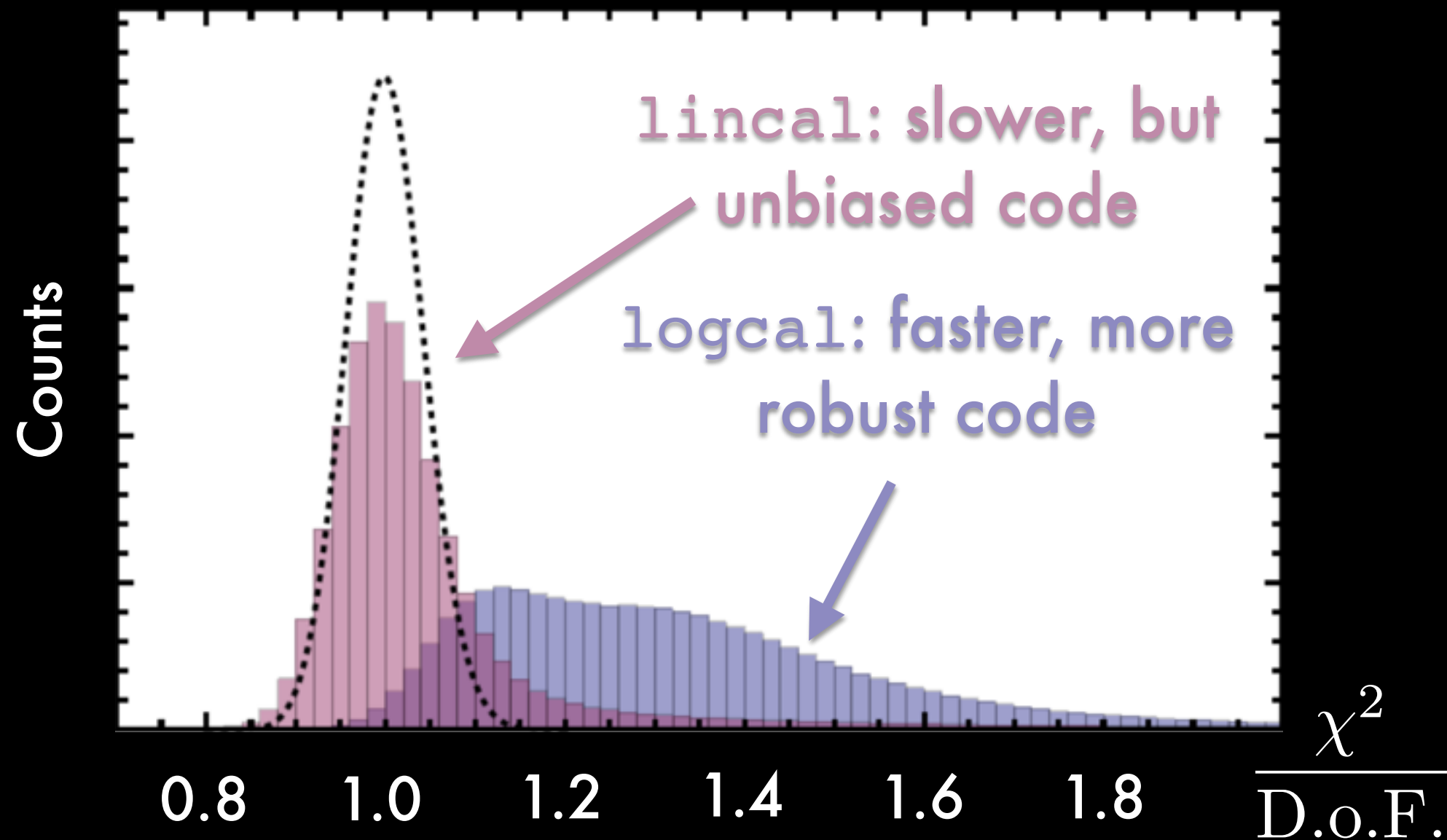# The "wedge" is the imprint of the chromaticity of the synthesized beam.
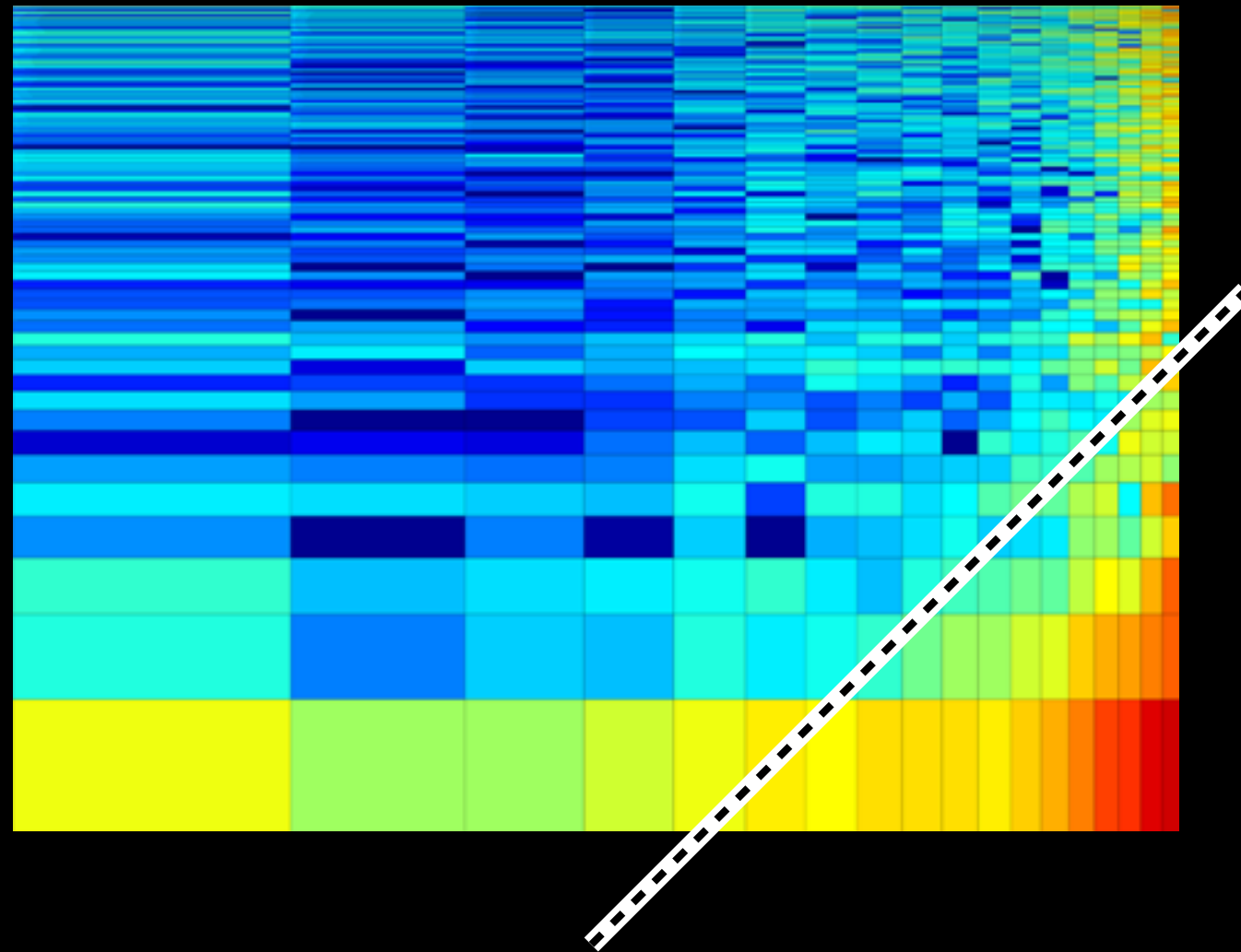
$$k_\parallel = \left[\theta_{\text{field}} \frac{D_M(z)E(z)}{(1+z)D_H}\right] k_\perp$$

Results:

The wedge evolves
with frequency in
just the way we
expected.

Dillon et al. (2014a)

Redundant baselines allow for a quantitative test of calibration and the real-time identification of problems.

lincal: slower, but unbiased code

logcal: faster, more robust code

$$\chi^2 = \sum_{\text{all baselines}} \left[(g_i g_j^*) V_{i,j}^{\text{fit}} - V_{i,j}^{\text{measured}} \right] / \sigma^2$$

$\dfrac{\chi^2}{\text{D.o.F.}}$

0.8   1.0   1.2   1.4   1.6   1.8

Counts

Zheng et al. (2014)

We'll begin imaging the EoR directly.

100Mpc

size of GOODS survey area

z = 8

Representative HI shells and HII holes
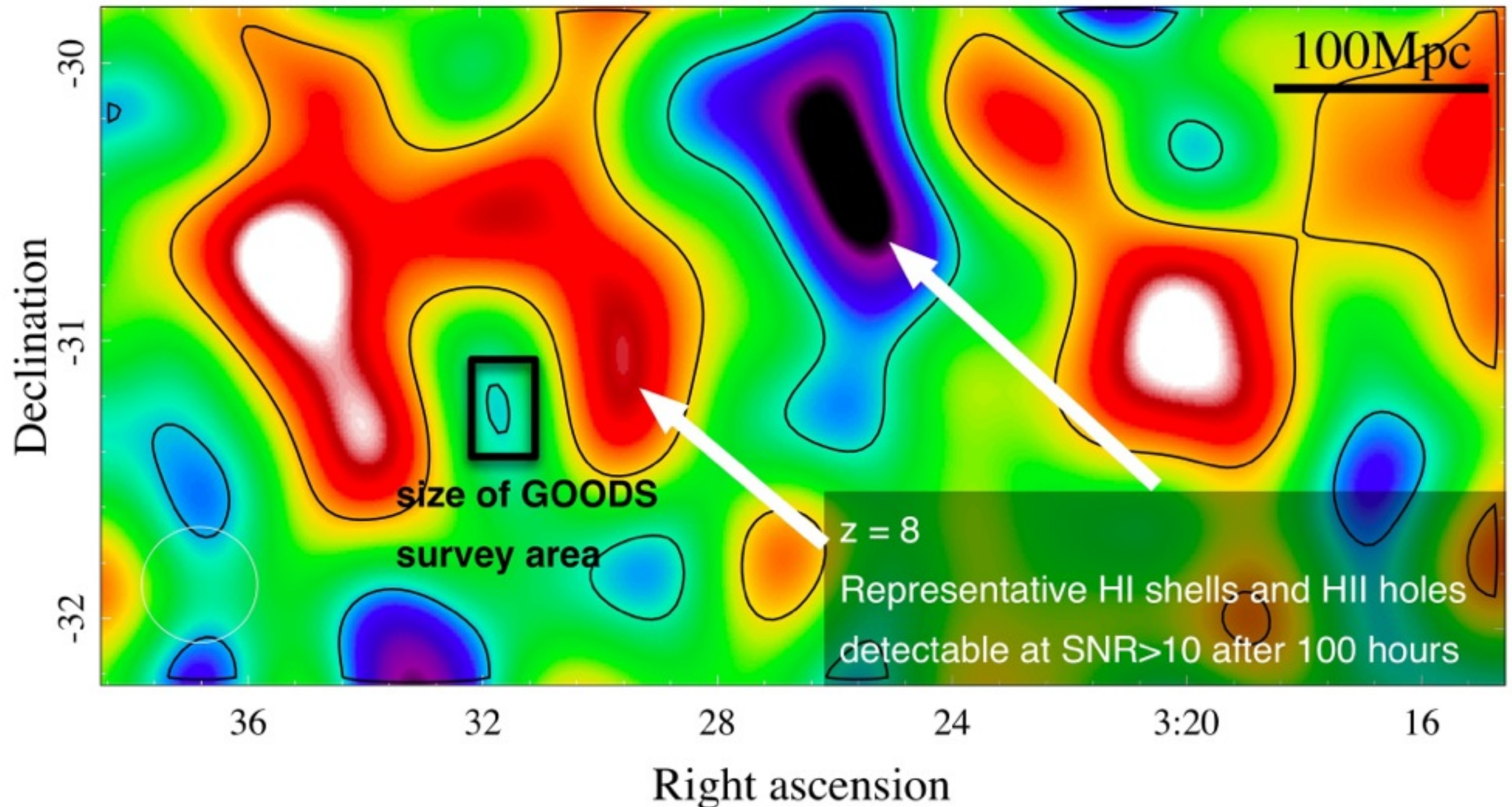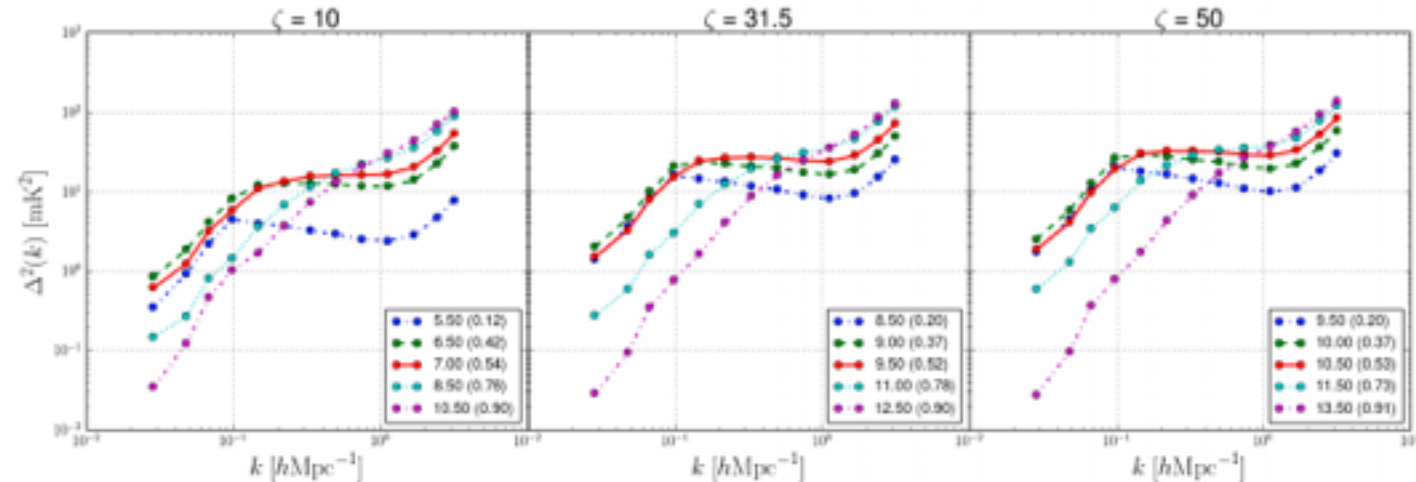
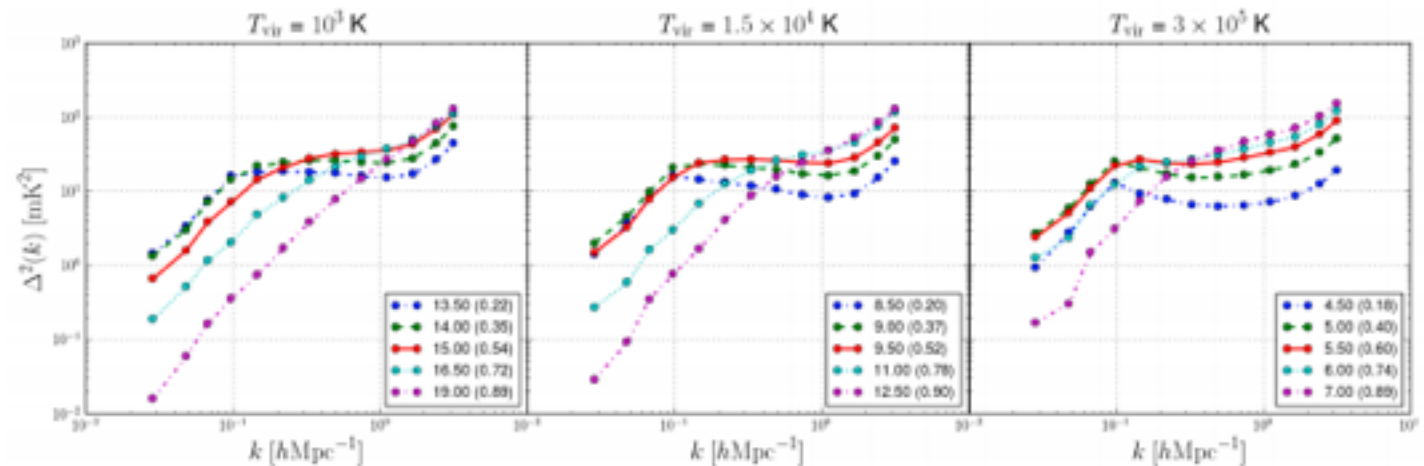detectable at SNR>10 after 100 hours

Figure: Danny Jacobs

Varying the reionization parameters yields qualitatively different power spectra.

Varying ζ:

Varying $T_{\rm vir}$:

Varying $R_{\rm mfp}$:

Pober, Liu, Dillon, et al. (2013)