Precision Interferometry Simulation For 21cm Cosmology

Danny Jacobs
October 2020
Arizona State University

Presenting work by the Radio Astronomy Software Group: Jonathan Pober, Bryna Hazelton, Adam Lanman, Daniya Seitova, Bharat Gehlot
And HERA: Steven Murray, Lily Whitler, James Aguirre, and many others

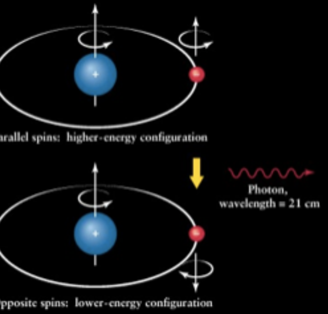
Outline

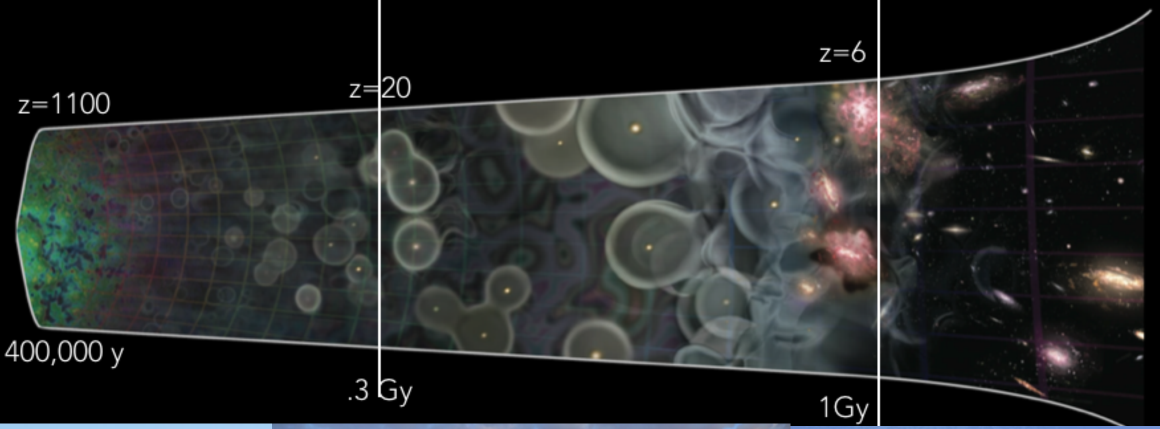
1. Short review: timeline and physics goal to power spectrum
2. Take some serious infrastructure
3. Low significance detections in the presence of systematics
    a. Paciga, Patil, Cheng.
    b. All fell victim to experimenter bias. Many lines of code, many parameters, simulators tightly coupled to analysis.
    c. How do we check pipelines producing 3 sigma results?
4. Instrument Designs Still differ
    a. Compare HERA and MWA.
    b. How do we design arrays to 1e-5 precision
    c. Not unique to interferometry, show Nivedita's plot
5. Solutions:
    a. Diversity of analysis methods. Enabled by data interchange standards. See pyuvdata and casacore ms
    b. Simulator test objects. Calibrated to first principles. Community-backed. Puvsim
6. Pyuvsim design goals
    a. Transparent and easy to read and use code. -> JOSS publication,
    b. Useful community product -> well defined sim parameters, use standard data interchanged, published reference products
    c. Well tested against analytic models -> Unit tests run analytic tests, reference sims lock it in, comparison to other simulators and data to keep it real.
    d. Accurate calculation of model, no approximations in the name of speed -> support for parallelization speedups
7. Design Details
    a. Test levels:
        i. Unittests of simple physics, What precision level?
        ii. reference simulations for external comparison and internal checkpointing. What precision level?
        iii. validation products for specific datasets Who's right?
    b. Unittests
        i. Many
        ii. Call out analytic diffuse tests as an open problem.
    c. Reference tests
        i. Ginned up arrays and samplings that cover physically relevant axes (time, baselines, frequencies, polarizations, sources)
    d. Testing against data
    e. Scaling.

Pre-Reionization
First stars activate 21cm line with Lyalpha
Visible in absorption

Reionization
Tau, As, $\sum m_\nu$,
first galaxies, supermassive blackholes

Post-Reionization
dark energy equation of state





LWA - Delillo et al 2020
RA - Deboer et al 2017
CHIME - Bandura et al 2014

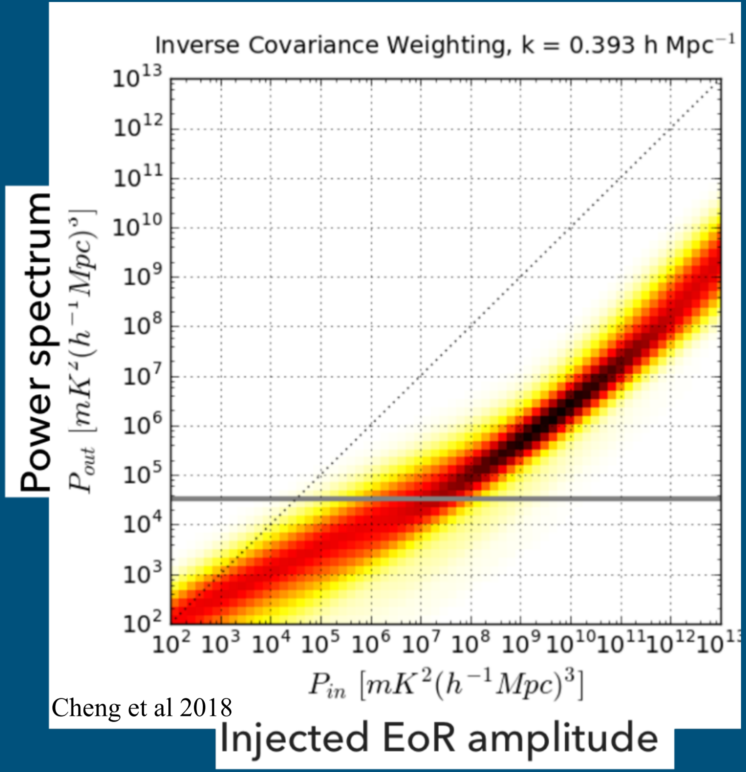
Low Significance detections in the presence of systematics not in the error model

Paciga et al 2013 - GMRT

Patil et al 2016 - Lofar

Cheng et al 2018 - PAPER

Unintentional experimenter bias is real and affects us all.

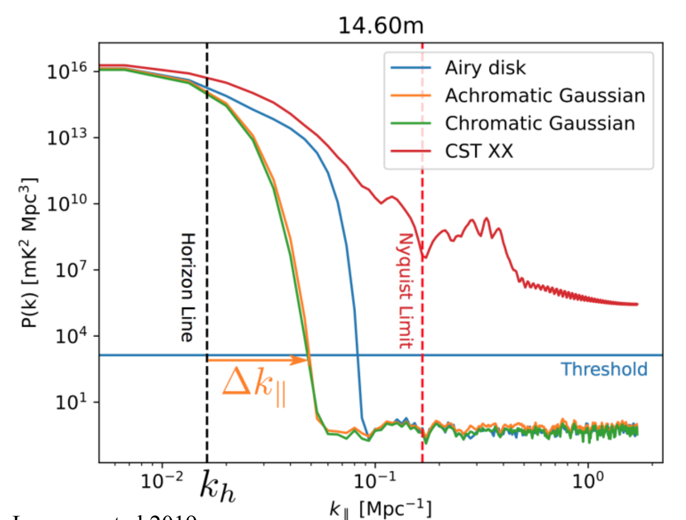


Cheng et al 2018

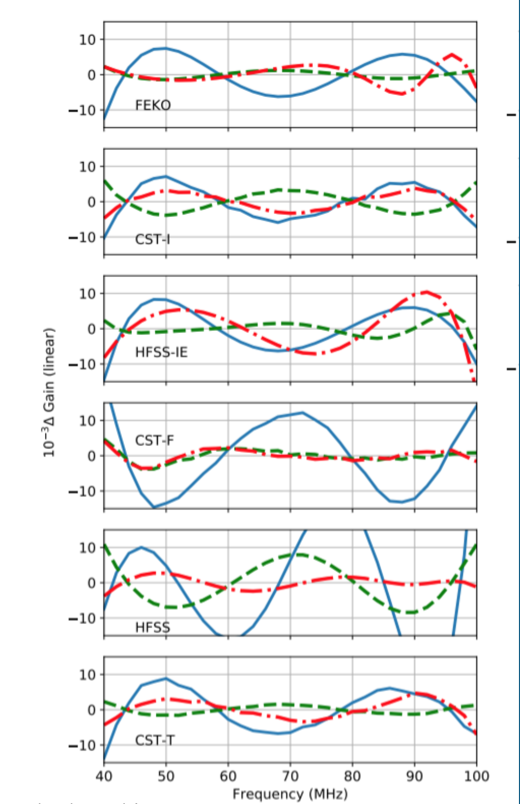
Instrument Design to 1:10,000

Small changes in instrument design make a big difference.

Ex Experiments vs SKA Observatory.



Lanman et al 2019

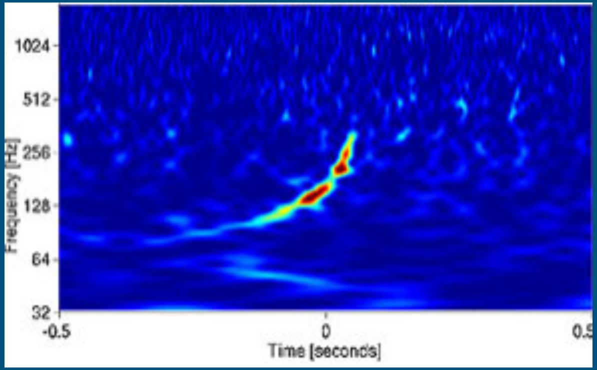


Mahesh et al in prep

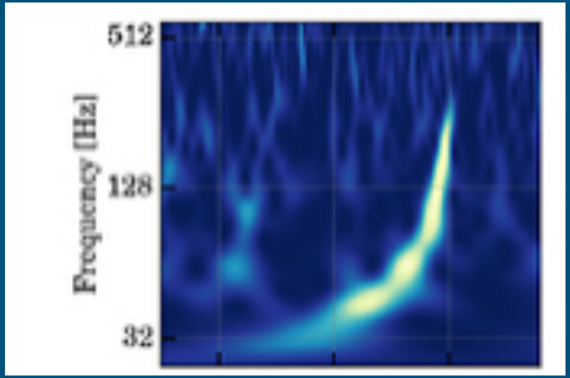
Beam	$A_{\rm eff}$ (m^2)	S/N (σ) (avoidance, subtraction)
Airy pattern	155	18.7, 90.8
Measured, feed at 5.3 m	93.0	12.7, 74.3
Measured, feed at 5 m	77.1	10.6, 67.9
Measured, feed at 4.5 m	68.5	10.0, 63.9

Neben et al 2016

How to distinguish reality from a false positive or negative



2010 - A surprise injection at LIGO



2018 - Routine detection in O2

The generalized interferometer model

Integral over entire sky

Beam and other propagation effects

Sky flux vs position s

Baseline Vector

$$V_{ij}(t,f) = \int\int_{4\pi} \mathrm{B}_i(\hat{s},t,f)\mathrm{B}_j^H(\hat{s},t,f)I(\hat{s},f)e^{\vec{b}\cdot\hat{s}}ds^2$$

Ideal Simulation Setup

The ideal case

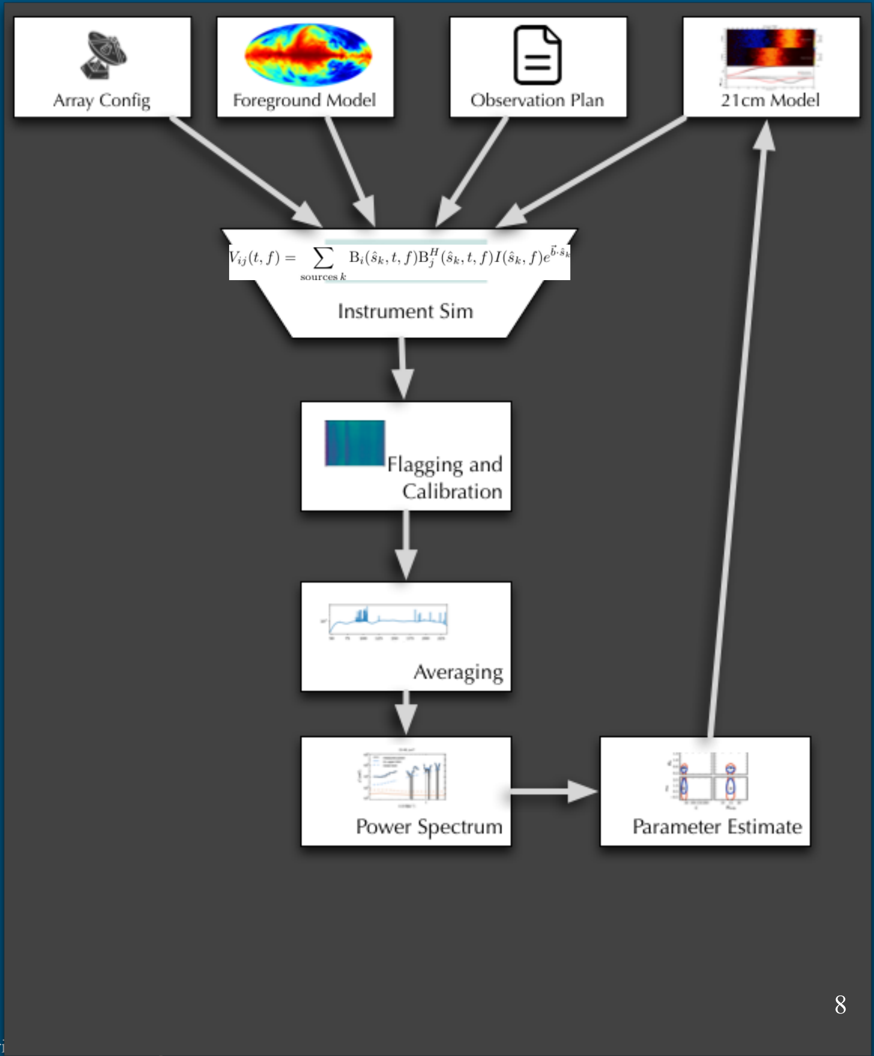
Add cosmology to sky model.

Calculate V_ij with independent simulator

Run through entire pipeline

Detect and fit parameters

Compare with injection.



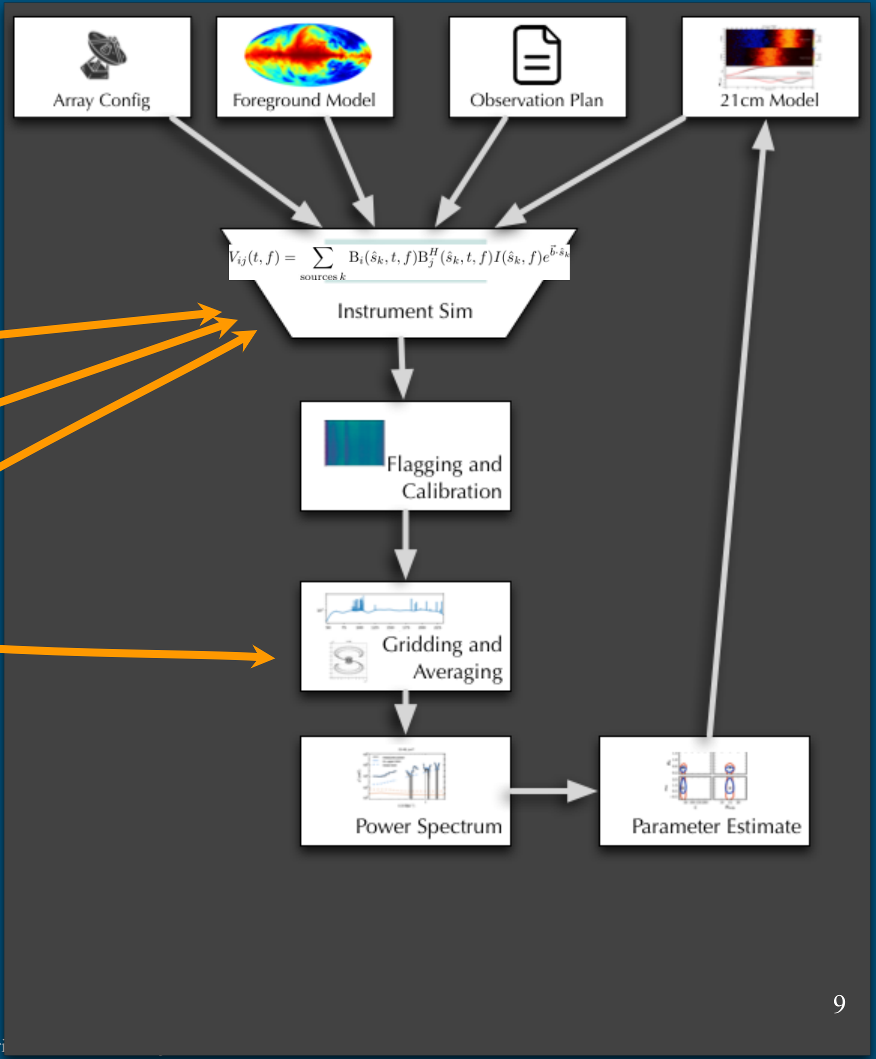
Typical Simulation Setup

Approximations in Vij (danger)

Sims checked mostly against data (danger!)

Sim/Analysis/validation codes by same person (more danger!)

Independent simulation is the key

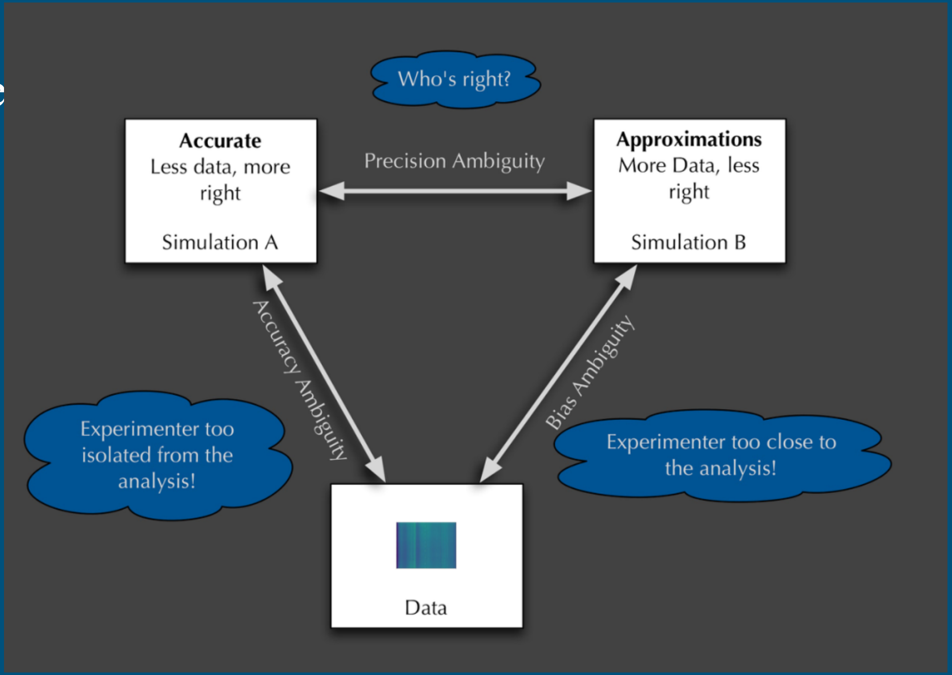


How do we know if our simulation is right?

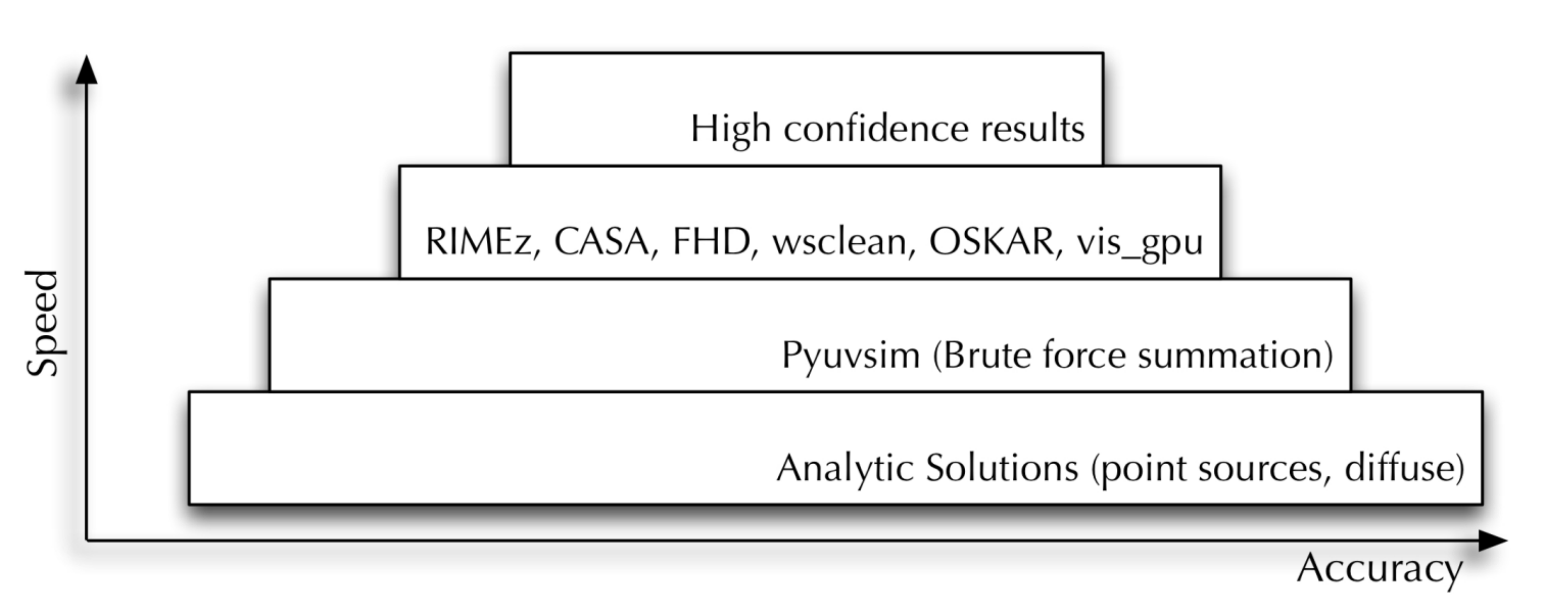
Recognize different classes of ambiguitie

Validation from first principles

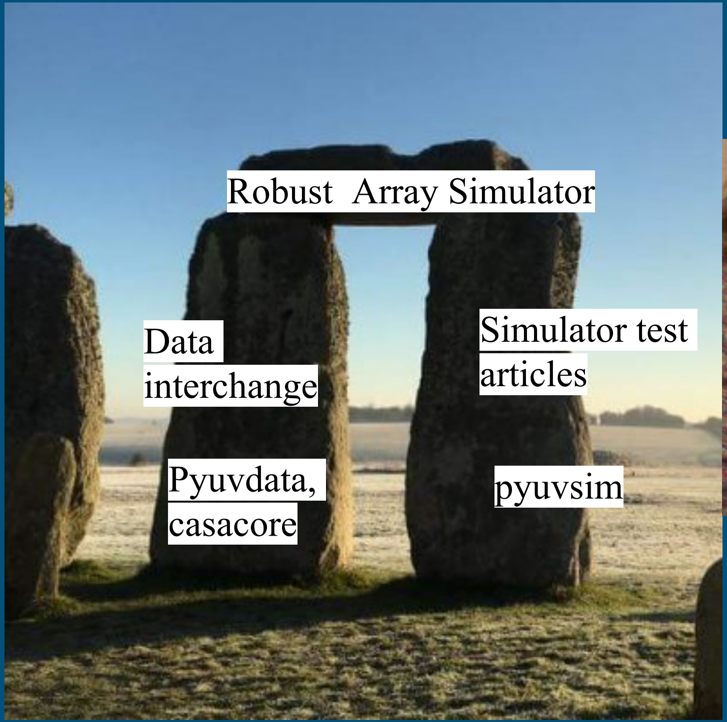
Multiple comparison options



Building high confidence validation



An awful awful Managementy slide.



The pyuvsim interferometer model

Integral over entire sky

Beam and other propagation effects

Sky flux vs position s

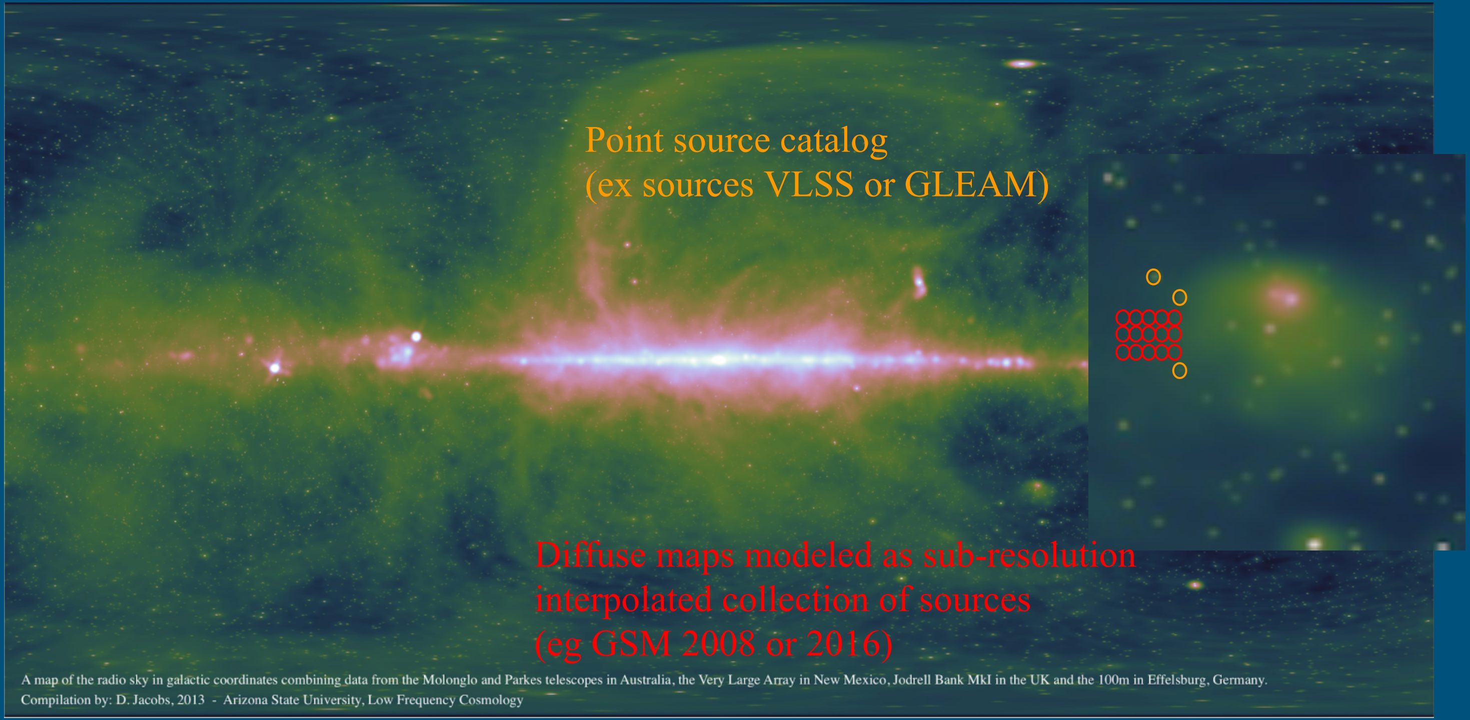
Baseline Vector

$$V_{ij}(t,f) = \int\int_{4\pi} \mathrm{B}_i(\hat{s},t,f)\mathrm{B}_j^H(\hat{s},t,f)I(\hat{s},f)e^{\vec{b}\cdot\hat{s}}ds^2$$

Quantize sky into arbitrarily dense collection of sources

$$\hat{s} \longrightarrow \hat{s}_k$$

$$V_{ij}(t,f) = \sum_{\text{sources } k} \mathrm{B}_i(\hat{s}_k,t,f)\mathrm{B}_j^H(\hat{s}_k,t,f)I(\hat{s}_k,f)e^{\vec{b}\cdot\hat{s}_k}$$



$$V_{ij}(t,f) = \sum_{\text{sources } k} \mathrm{B}_i(\hat{s}_k, t, f)\mathrm{B}_j^H(\hat{s}_k, t, f) I(\hat{s}_k, f) e^{\vec{b}\cdot\hat{s}_k}$$

Gaussian Component Model (Standard Clean Component decomposition)

Pyuvsim design goals

Goal	Done	To Do
Transparent and easy to read code.	1. Publication in JOSS*, which has high community standards 2. pyuvsim.readthedocs.io	Developers guide Set up more users
Useful community product	1. Well defined reference sim parameters 2. Standard data interchange 3. Published comparison analysis	Paper detailing test protocol
Well tested against analytic models	1. Analytic comparisons run in unittests 2. Reference sims lock in results 3. Comparisons to other sims and data to keep it real	Paper detailing analytic diffuse (Lanman in prep)
Accurate calculation of model	1. No approximations in the name of speed 2. Support for parallelization speedups 3. HPC scaling tests	HPC time on XSEDE next quarter HTCondor version under test. GPU version

*Lanman et al 2019, Journal of Open Source Software

Versioning Approach

We use a `generation.major.minor` format.

- Generation - Release combining multiple new physical effects and or major computational improvements. Testing: Backed by unittests, internal model validation, and significant external comparison.
- Major - Adds new physical effect or major computational improvement. Small number of improvements with each release. Testing: Backed by unittests, internal model validation and limited external comparison.
- Minor - Bug fixes and small improvements not expected to change physical model. Testing: Backed by unittests

Pyuvsim Validation Layers

Level	Run Time	Setup	Version
Unittests	Run on every git push in minutes	Most Tolerances 1e-8	`minor`
Reference Simulations	Run every major version in hours	Routine Checkpointing and detailed validation	`major`
Validation products	HPC jobs 10k cpu-hours or more	Tuned to match observation or simulator needs	`generation`

Pyuvsim physics unittests

1. Beam Jones matrix modeling and interpolation
2. Sky catalog interpolation and spectral modeling
3. Source position rotation, baseline rotation, vector product, baseline redundancy
4. Simulation configuration setup and file generation
5. Single source analytic model, small angle approximation, horizon cutoff.
6. Heterogeneous beams

$$\underset{4}{V_{ij}(t,f)} = \underset{5}{\sum_{\text{sources }k}} \underset{1}{\mathrm{B}_i(\hat{s}_k,t,f)}\underset{6}{\mathrm{B}_j^H(\hat{s}_k,t,f)}\underset{2}{I(\hat{s}_k,f)}\underset{3}{e^{b\cdot\hat{s}_k}}$$

Pyuvsim reference simulations

First Gen - 6 sims testing times, frequencies, sources and beams

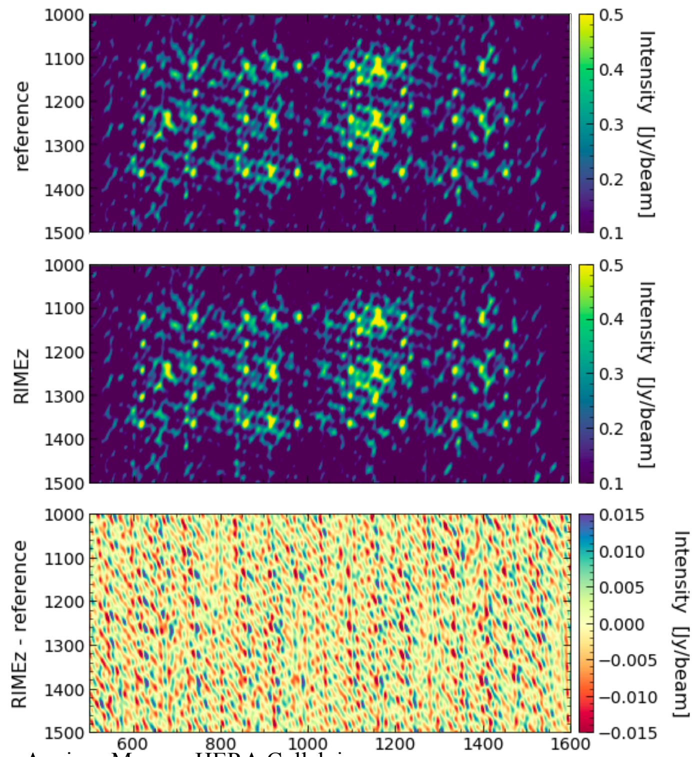
Name (beam)	Purpose
1.1	Test imaging and source orientation.
1.2 (gauss)	Check that sources move appropriate and rise/set, and pass through the beam properly.
1.2 (uniform)	Check that sources move appropriate and rise/set (stay visible near horizon).
1.3 (gauss)	Check that visibilities have sensible frequency evolution. Get observable fringes. Realistic primary beam.
1.3 (uniform)	Check that visibilities have sensible frequency evolution. Get observable fringes.
1.4	Check phasing precision and simulate realistic data.

github.com/RadioAstronomySoftwareGroup/pyuvsim/tree/master/reference_simulations

Reference sims validating HERA pipeline

RIMEz: fast m-mode sim by Zac Martinot (UPenn grad)

Validation comparison by Lily Whitler (ASU Undergrad)



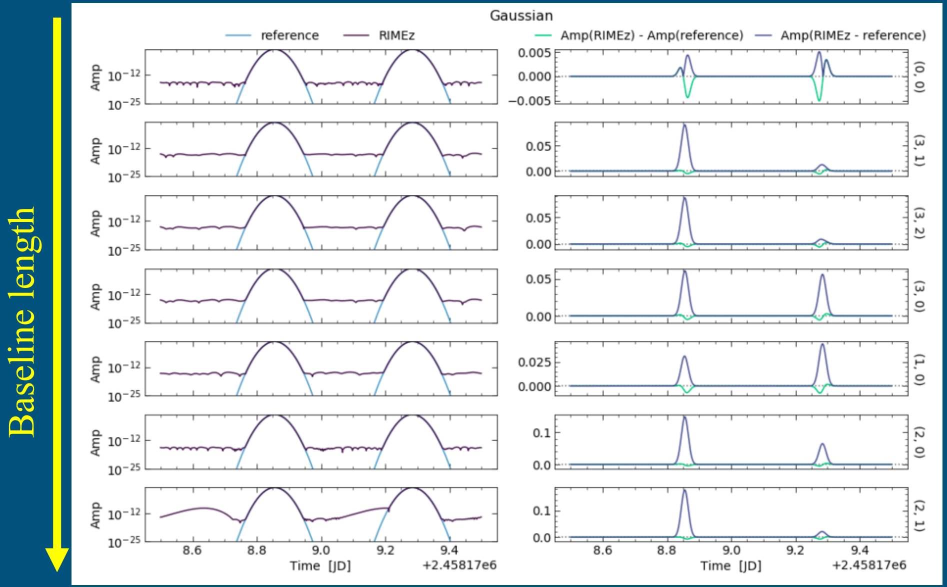
Aguirre, Murray, HERA Collab in prep

Reference sims validating HERA pipeline

RIMEz: fast m-mode sim by Zac Martinot (UPenn grad)

Validation: comparison by Lily Whitler (ASU Undergrad)

Conclusion: RIMEz floor at 1e-16 due to limited precision in FFT. Position errors at arcseconds to minutes due to astrometry engine imprecision



Aguirre, Murray, HERA Collab in prep

Pyuvsim reference simulations

First Gen - 6 sims testing times, frequencies, sources and beams

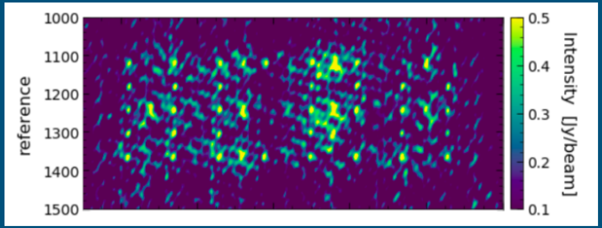
Second Gen - 5 sims, larger more physical beams, polarization

Name (beam)	Purpose
2.1 (airy)	Replace the first generation reference simulations by covering multiple axes.
2.2 (UVBeam)	Check that visibilities are sensible with UVBeam.
2.3 (airy)	Check that visibilities are sensible with a known power spectrum diffuse map.
2.3 (airy)	Check that visibilities are sensible with the healpix map.
2.4	Check phasing polarized response(not done yet).

github.com/RadioAstronomySoftwareGroup/pyuvsim/tree/master/reference_simulations

Accurate calculation of model	1. No approximations in the name of speed 2. Support for parallelization speedups 3. HPC scaling tests	HPC time on XSEDE next quarter HTCondor version under test. GPU version

Naive Scaling



?

Data Points:
60 antennas, 1 time, 100 frequencies, 43 sources (spelling HERA) = 177k voxels

Sky Model:
43 source (spelling HERA)

Compute: ~0.4 CPU-hours

Data Points:
128 antennas, 60 times, 600 frequencies = 292M voxels

Sky Model:
1M GLEAM sources +100k diffuse NSIDE 128 = 1.1M sources

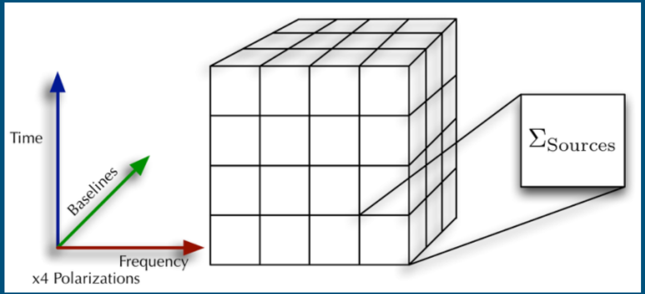
Naive scaling: 1.4M CPU-hours?!?

Simplistic Parallelization

Very Parallel*
Good for spreading out

Vector rotations and summations
Good for multicore/LAPACK optimization

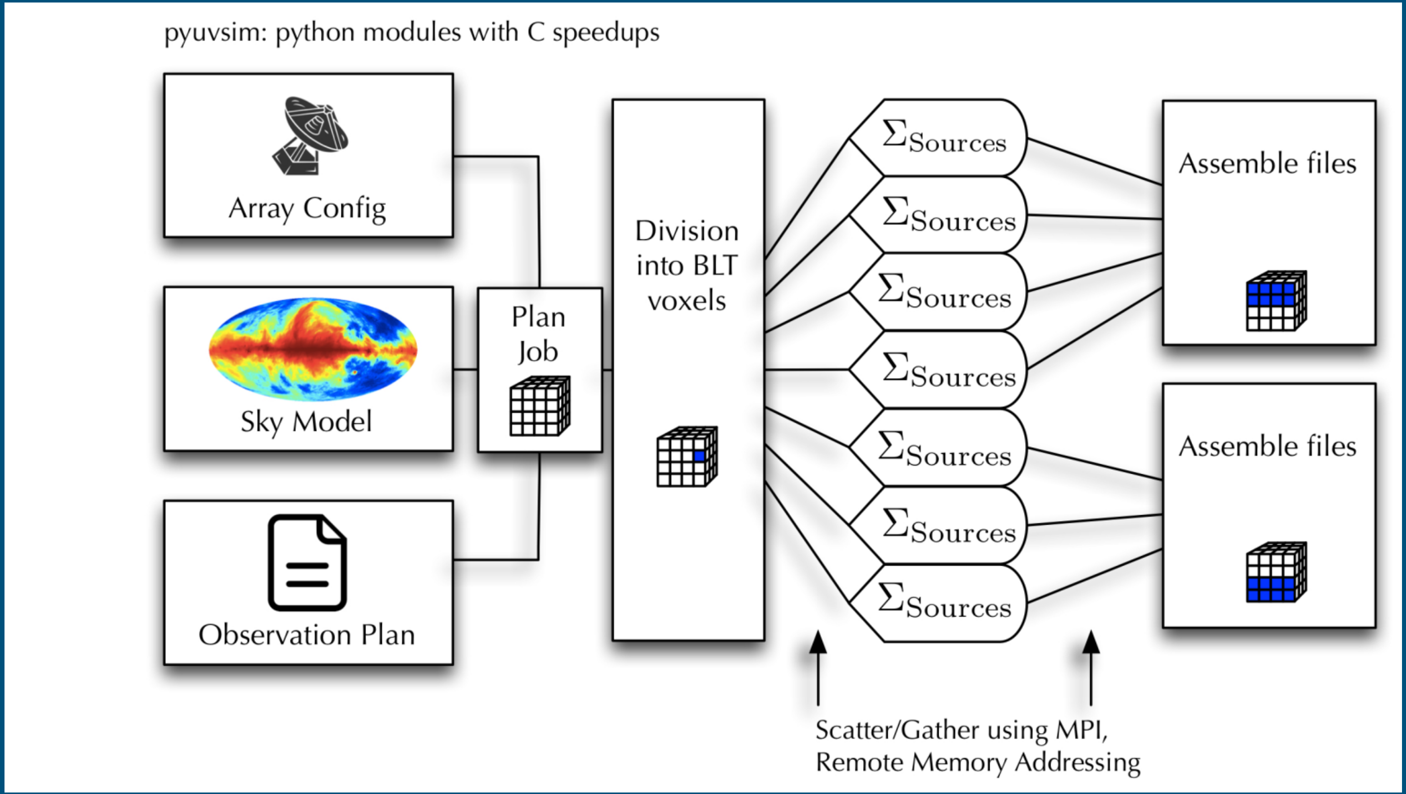
$$V_{ij}(t,f) = \sum_{\text{sources } k} \mathrm{B}_i(\hat{s}_k,t,f)\mathrm{B}_j^H(\hat{s}_k,t,f)I(\hat{s}_k,f)e^{\vec{b}\cdot\hat{s}_k}$$



*There are other ways to break down the Vij matrix, see eg gpu_vis.

Accurate calculation of model	1. No approximations in the name of speed 2. Support for parallelization speedups 3. HPC scaling tests	HPC time on XSEDE next quarter HTCondor version under test. GPU version

Optimizations



Source Scaling

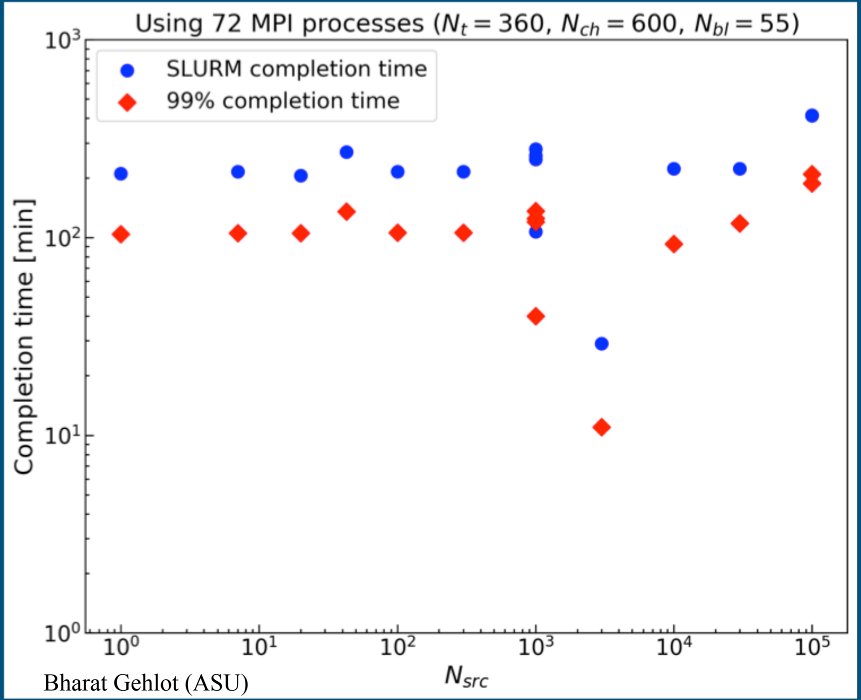
Constant Ncores

Same Data volume

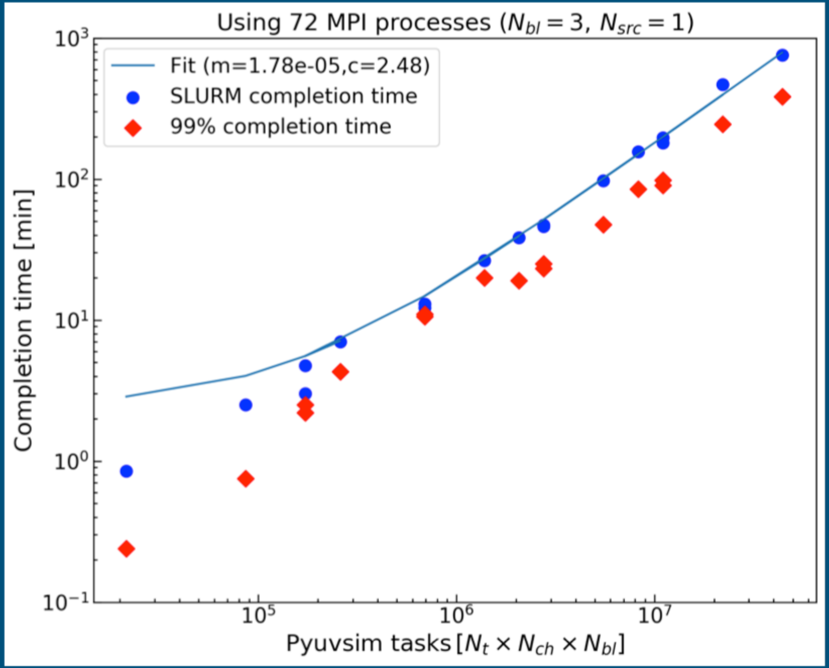
Increasing Source counts

Completion time is flat up to 2e4 sources

Appears to be linear thereafter. (Not shown)

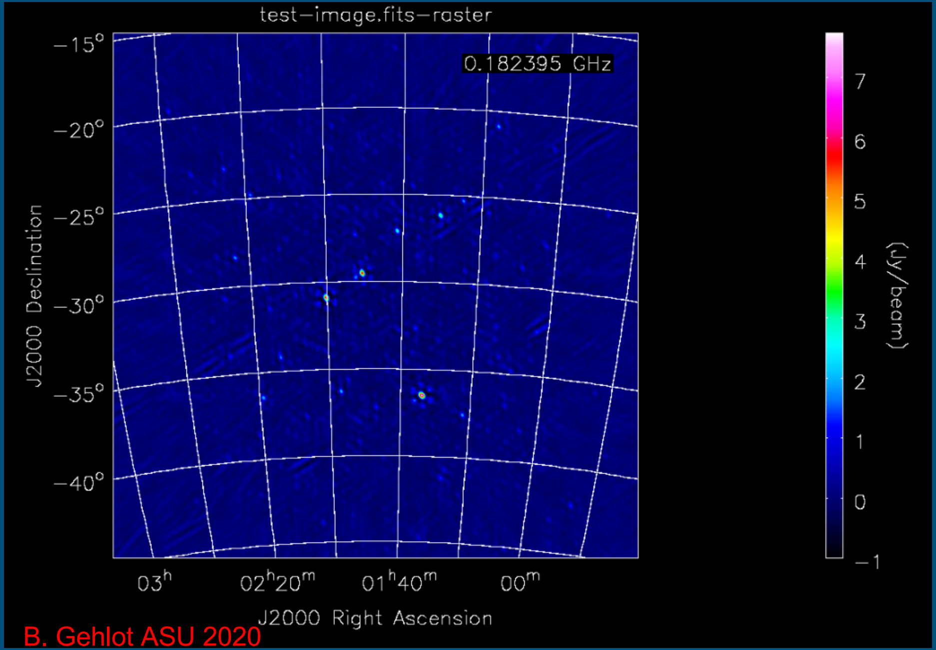


Linear with data volume



Big Run

The MWA "Golden Minute" reference sample. The reference snapshot for multiple pipelines. One of the most studied 21cm data sets in the world.



Antennas	128 MWA Layout
Beam Model	Latest EM Sim
Times	60
Frequencies	768
Polarizations	1
Total Data Volume	374M data points
Sources	~1M GLEAM
Cores	1000
Cores Per Task	4
RAM per Core	4.5GB
Projected run time	23 hours
Actual run time	12 hours
Queue time	28 days (!)

Next steps

Goal	To Do
Transparent and easy to read code.	Developers guide Set up more users
Useful community product	Paper detailing test protocol
Well tested against analytic models	Paper detailing analytic diffuse (Lanman in prep)
Accurate calculation of model	HPC time on XSEDE next quarter HTCondor version under test. GPU version