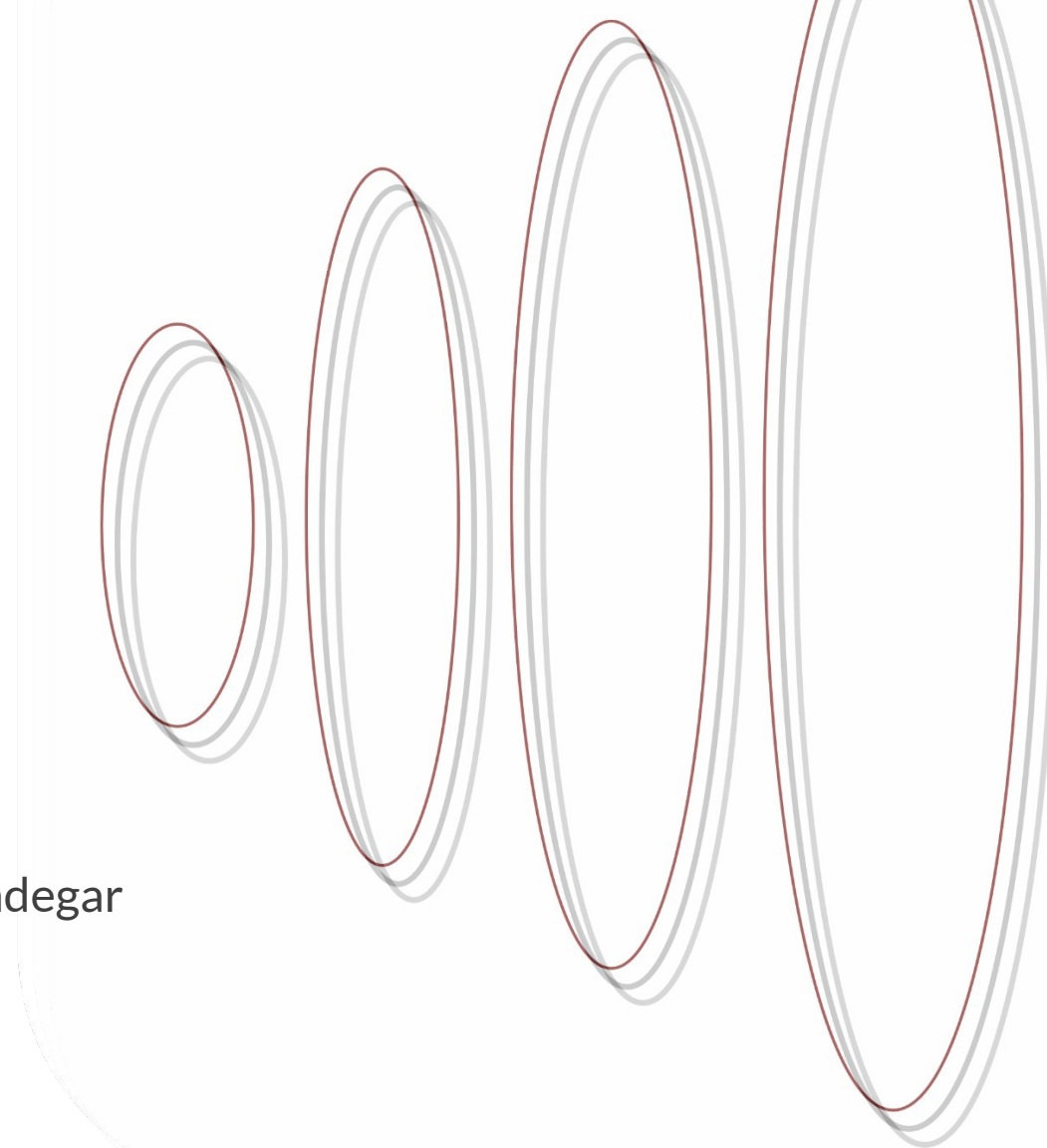
Simulation and Analysis Workshop Day 2

MAGIS - 100

Will Michaels, Brion Ye
Sanha Cheong, Sean Gasiorowski, Murtaza Safdari*, Maxime Vandegar
Joe Frisch, Michael Kagan, Ariel Schwartzman

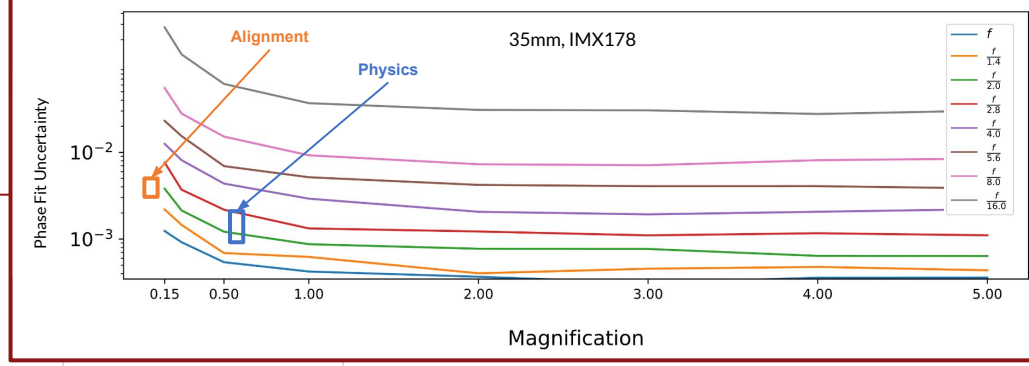
02/02/2022

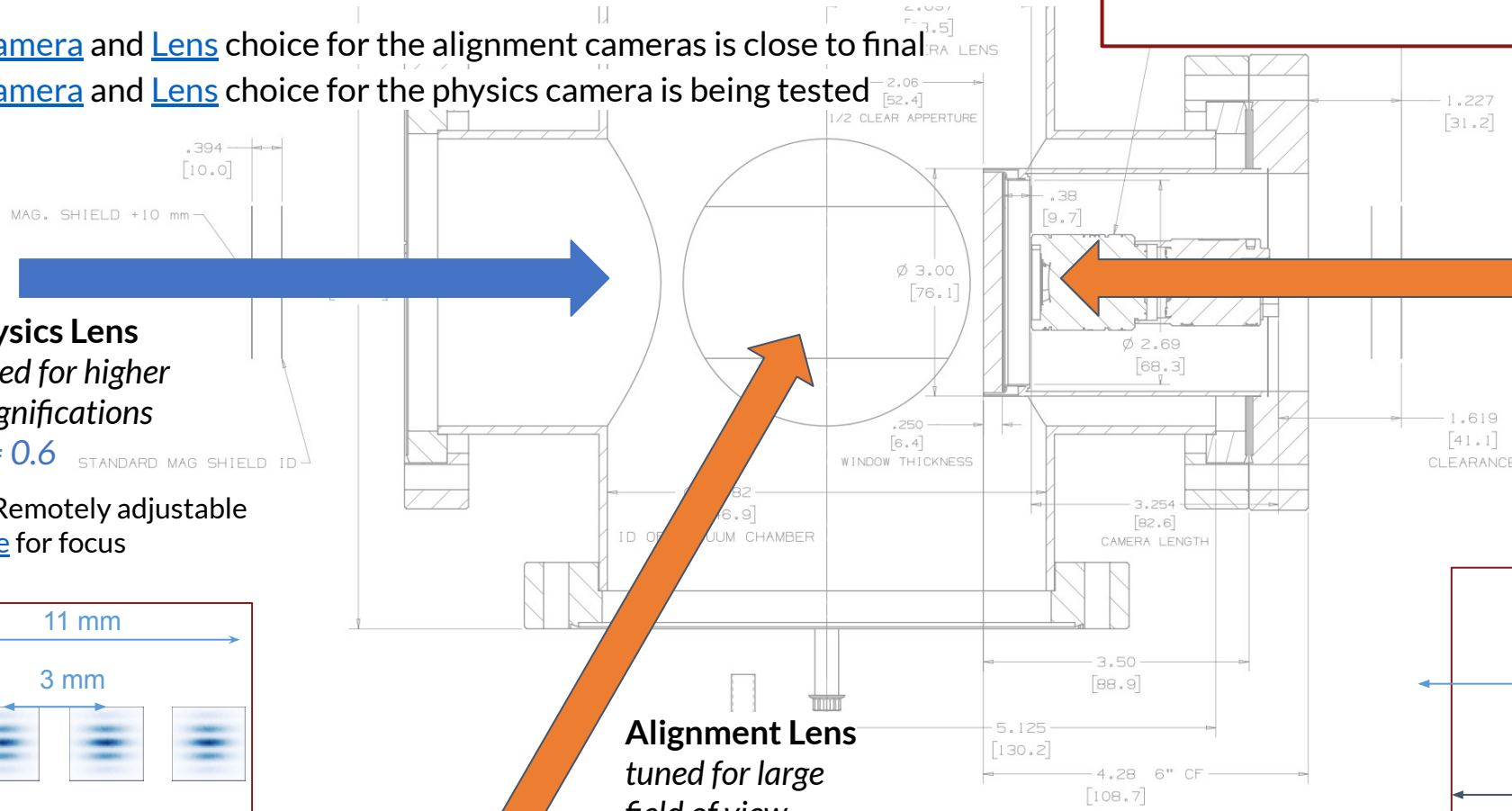


Diagnostic Imaging System

Design Choices

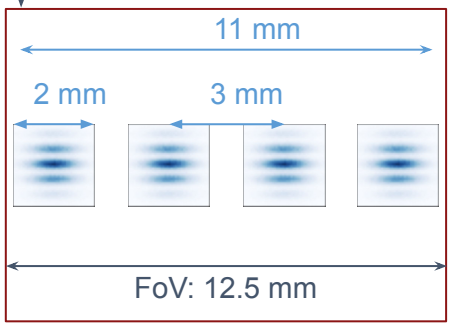
- Camera and Lens choice for the alignment cameras is close to final
- Camera and Lens choice for the physics camera is being tested





Physics Lens
tuned for higher magnifications
m = 0.6

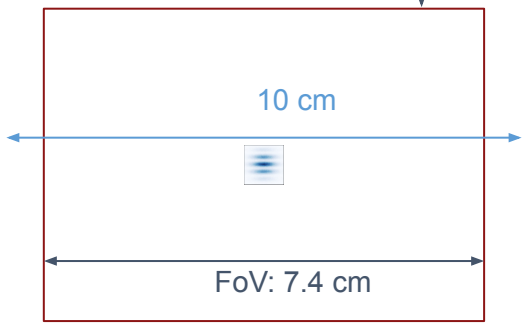
w/ Remotely adjustable base for focus



Alignment Lens
tuned for large field of view
m = 0.1

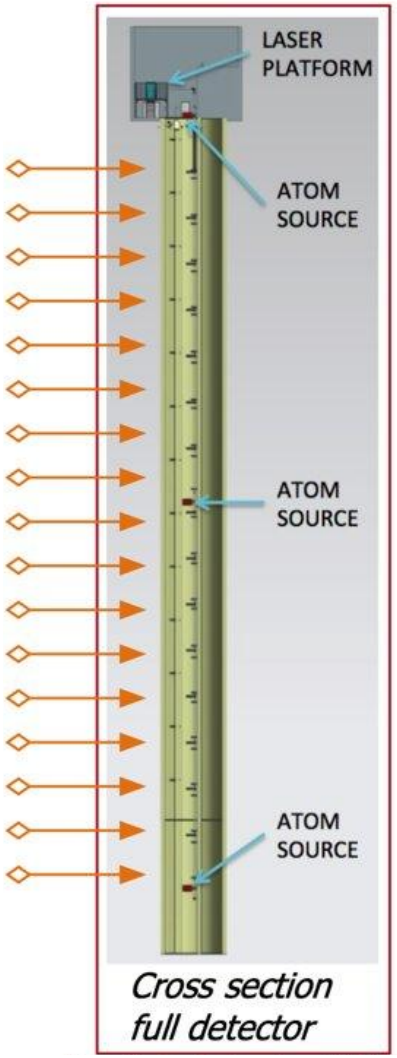
Alignment Lens
tuned for large field of view
m = 0.1

w/ Fixed mount



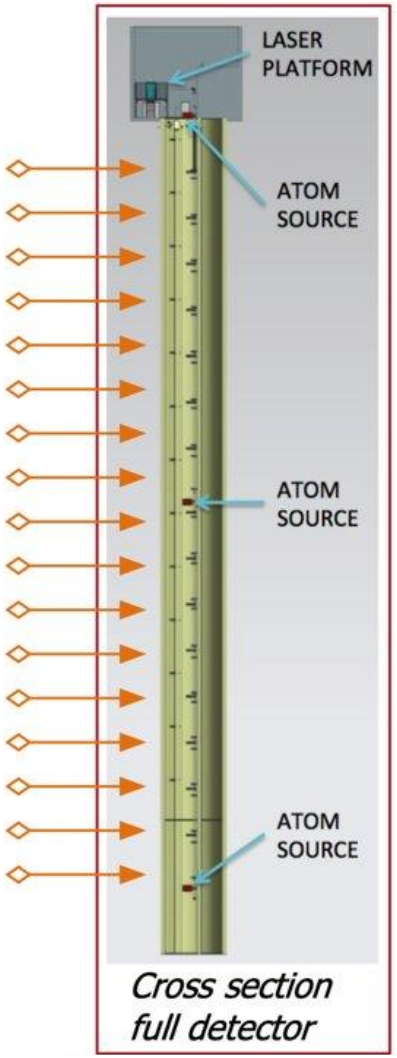
Resources for Diagnostic Imaging System - Alignment

Item	Value	Notes
# cameras	28	Two 6MP calib. cameras, 14 nodes
# images / drop	2	Reading 1 station / drop
Image size	10 MB	Could do image cropping if needed
Data / drop	O(20 MB)	Reading 1 station / drop
Drop rate	1 / minute (ave.)	Calib. slower than normal data taking
Run-time / day	24 hrs	
Calib. data rate	O(30 GB/day)	
Extracted data	O(10 MB/day)	Assuming O(10) floating point parameters extracted per drop



Resources for Diagnostic Imaging System - Physics

Item	Value	Notes
# cameras	14	One 6MP calib. cameras, 14 nodes
# images / drop	1	Reading 1 station / drop
Image size	10 MB	Could do image cropping if needed
Data / drop	O(10 MB)	Reading 1 station / drop
Drop rate	1 - 10 Hz	Assume 1Hz to start, 10 Hz stretch goal
Run-time / day	24 hrs	
Calib. data rate	O(1 TB - 10 TB/day)	
Extracted data	O(10 MB/day)	Assuming O(10) floating point parameters extracted per drop



Key Questions from the Diagnostic Imaging System

20MB/60s ; 30GB/day

Alignment:

Data-taking

Data storage and transmission
- Do we use the RPi's as buffers?
- Do we transmit the data up to the storage server right away?
  - Is it via ethernet?

How do we save the associated laser settings metadata?
- Pointers to database?

Cropped images to save on storage space?

Online processing

Calibration data analysis, feedback to drop system and live monitoring
- Calculate centroids, clouds widths online, O(n)

Does this happen on the Rpi's or directly on the servers?
- How easy is either option to develop?

Any other/detailed analysis software to run during online operations?
- Detect speckles or anomalies in the FoV?

Triggering off the RPi D/A or a centrally (delayed) Trigger Line?
- Follow-up in upcoming engineering meeting?

10-100MB/s ; 1-10TB/day

Physics:

Data-taking

Can the Rpi handle 10Hz (100MB/s) image transfer to the server either directly or as a buffer?

May need to use data compression for size reduction in addition to crop?

Online processing

Seems to set the data-processing rate for phase extraction

How fast can we extract a phase on a Rpi vs servers?

Data sizes are large, is all phase reconstruction online?

Offline

Will we plan yearly re-processing for new reconstruction algorithm use?

Simulations

Computing resources available for these simulations?
Test various alignment scenarios along
Simulate physics signals with images taken along the full length of the DIS

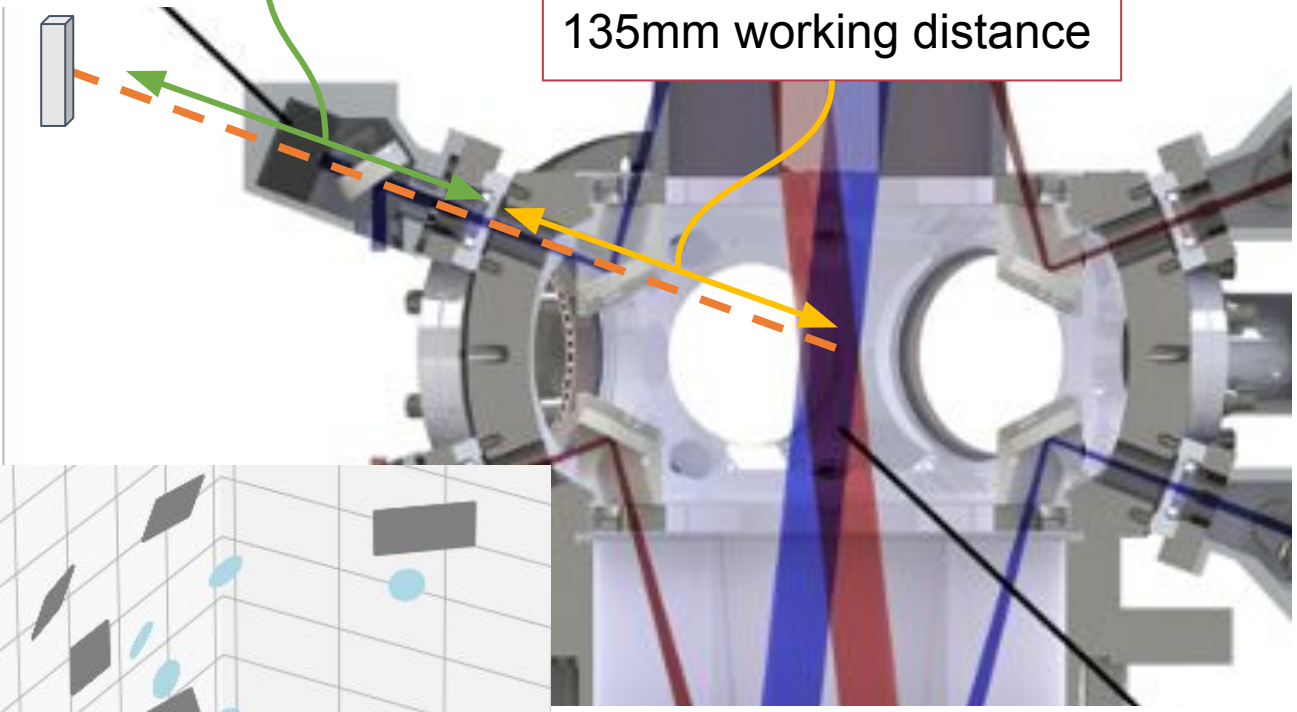
Off-Axis Views for Physics and Tomography

Atom Source Connection Node Cameras

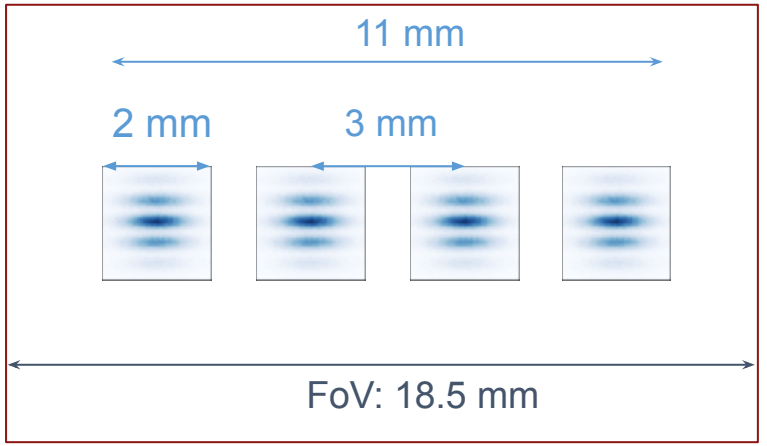
~254mm free behind the window

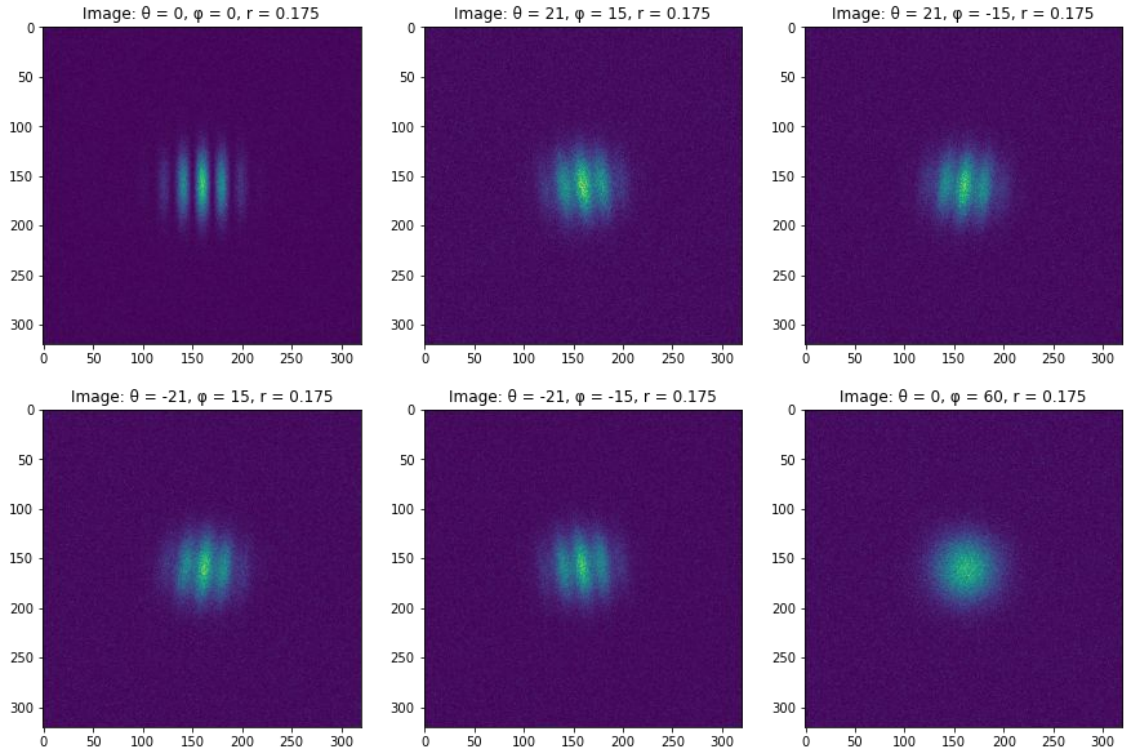
21°/15° head on view
16mm dia viewport, F/10

135mm working distance



4 port imaging with m = 0.4 lens using a 1/1.8” CMOS sensor





Primary Imaging System

Item	Value	Notes
# cameras / node	6	6 MP cameras, Off-axis imaging with 4 extra cameras
Image size	10 MB	Depending on # pixels
Data / drop	60 MB	Image + meta-data
Drop rate	1-10 Hz	Assume 1Hz to start, 10 Hz stretch goal
Run-time / day	24 hrs	
Raw data rate	O(6 TB – 60 TB) / day	Assuming 6 cameras
Extracted data	O(60 MB) / day	Assuming O(10) floating point parameters extracted per drop

Seems to set the data-processing rate for phase extraction

- How fast can we extract a phase? How many processors will we have?
- Data sizes are large, is all phase reconstruction online? Will we plan yearly re-processing for new reconstruction algorithm use?

3D reconstruction with off-axis imaging will require more computing resources

- Developing simulation and reconstruction algos now
- Will likely benefit from GPUs

Key Questions from Primary Imaging and Simulations

60-600MB/s ; 6-60TB/day

Physics:

Data-taking

Can we handle 10Hz (600MB/s) image transfer to the server?
Use data compression for size reduction in addition to crop?
How to save and associate appropriate metadata for each camera?

Online processing

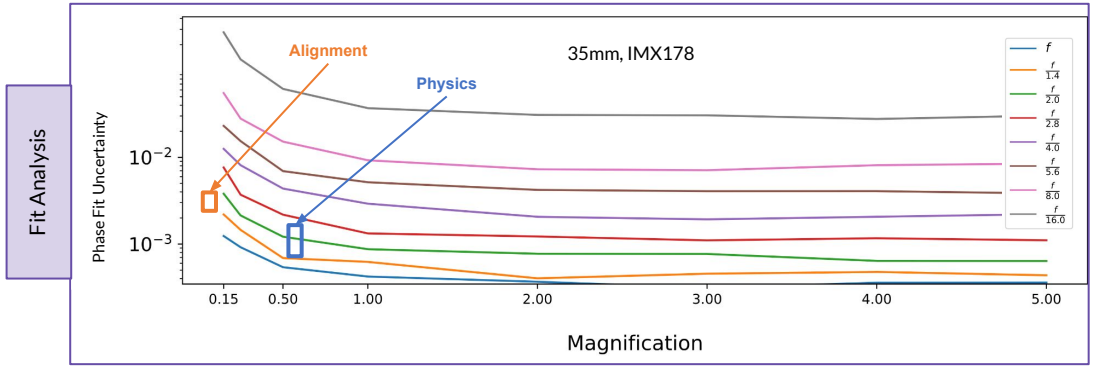
Seems to set the data-processing rate for phase extraction
How fast can we extract a phase from these images?
Data sizes are large, is all phase reconstruction online?
Resources needed for online reconstruction?

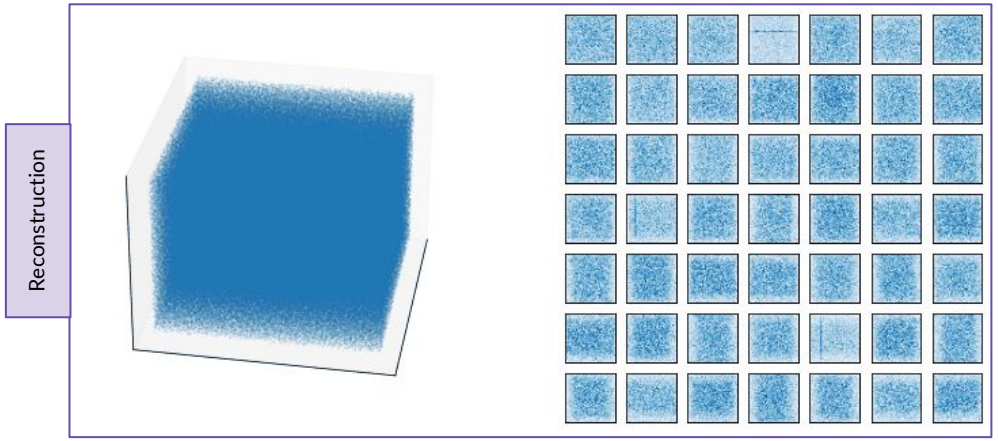
Offline

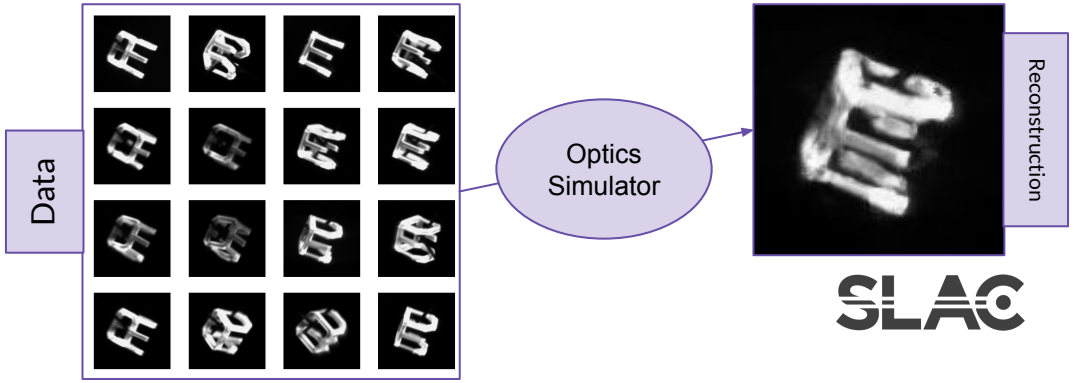
Will we plan yearly re-processing for new reconstruction algorithm use?
Resources needed for various reconstruction and fitting tasks?
Remote data access?

Simulations

Need to allocate significant resources (processors, GPUs, etc)
Vital in developing and tuning reconstruction techniques
Set upper bounds on accuracies and uncertainties
Fully integrated in the reconstruction algorithms
Used to generate datasets for training and analysis
Tuning simulations to match data

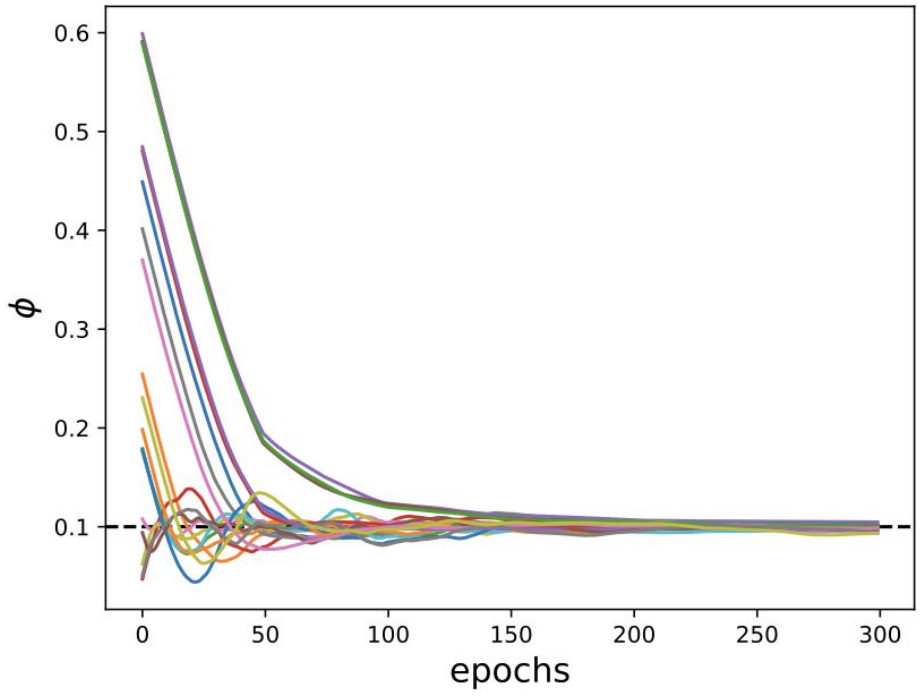


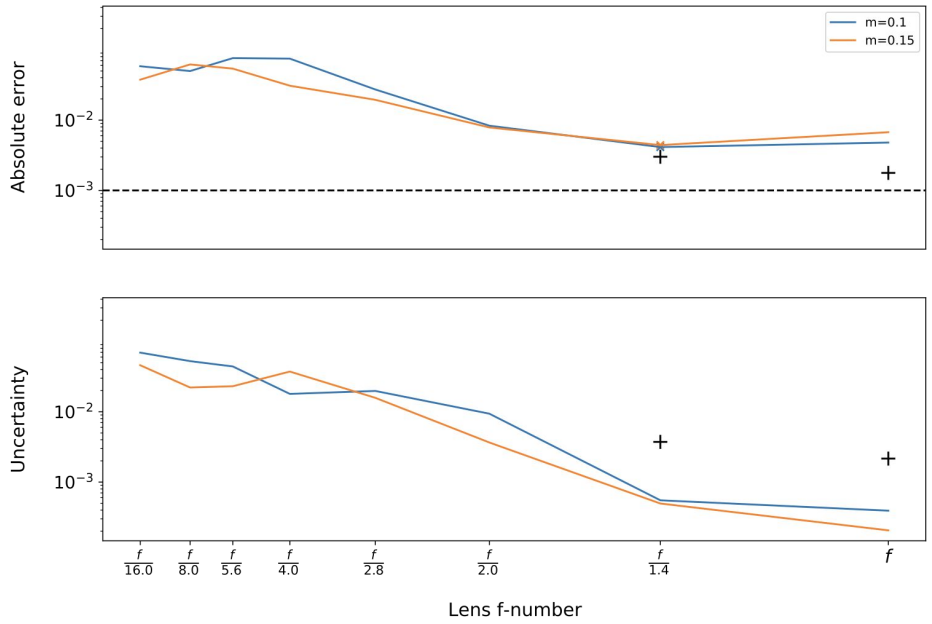




BACKUP

Function Fitting with 3D images

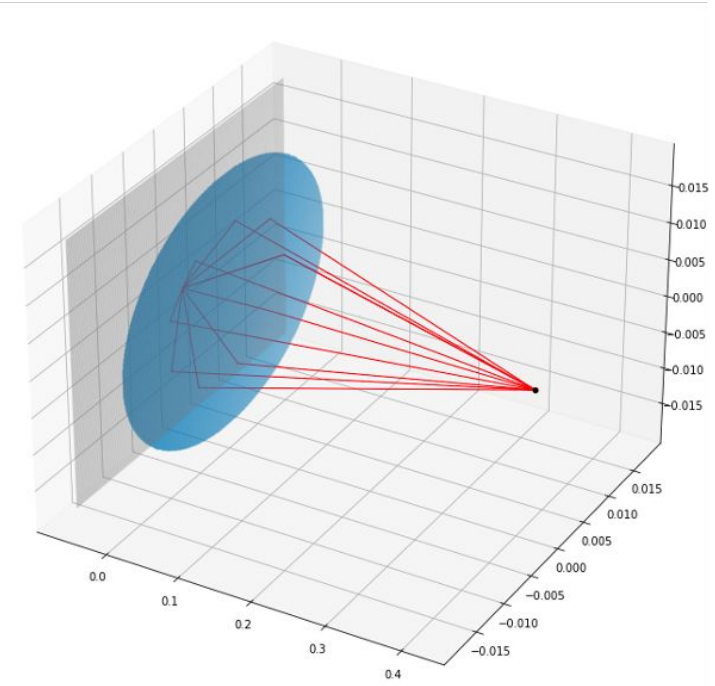




Differentiable-Optics (diffoptics)

diffoptics: Simulator of optical system and image generation

- Geometric Optics Approach: Ray tracing
- Building blocks: Optical Elements
  - Thin & thick lenses, sensors, viewing windows, mirrors, etc.
- Atom and photon shot noise included
- PSF implemented with convolutions



Python: written within PyTorch framework

- C++ and CUDA backend
- Vectorized, compile-able, and GPU compatible

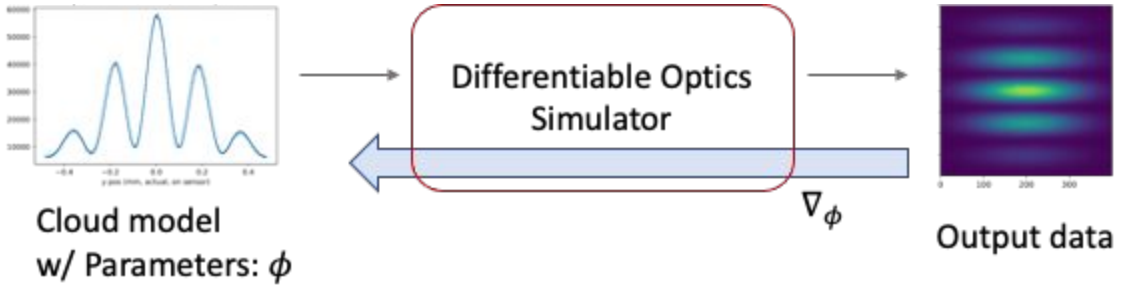
End-to-end differentiable (discussed in later slide)

Differentiable frameworks

Optics simulator is *differentiable*:

- Write function: f(x) {…}; ⇒ Get derivative function automatically: df(x) {…};

Why?

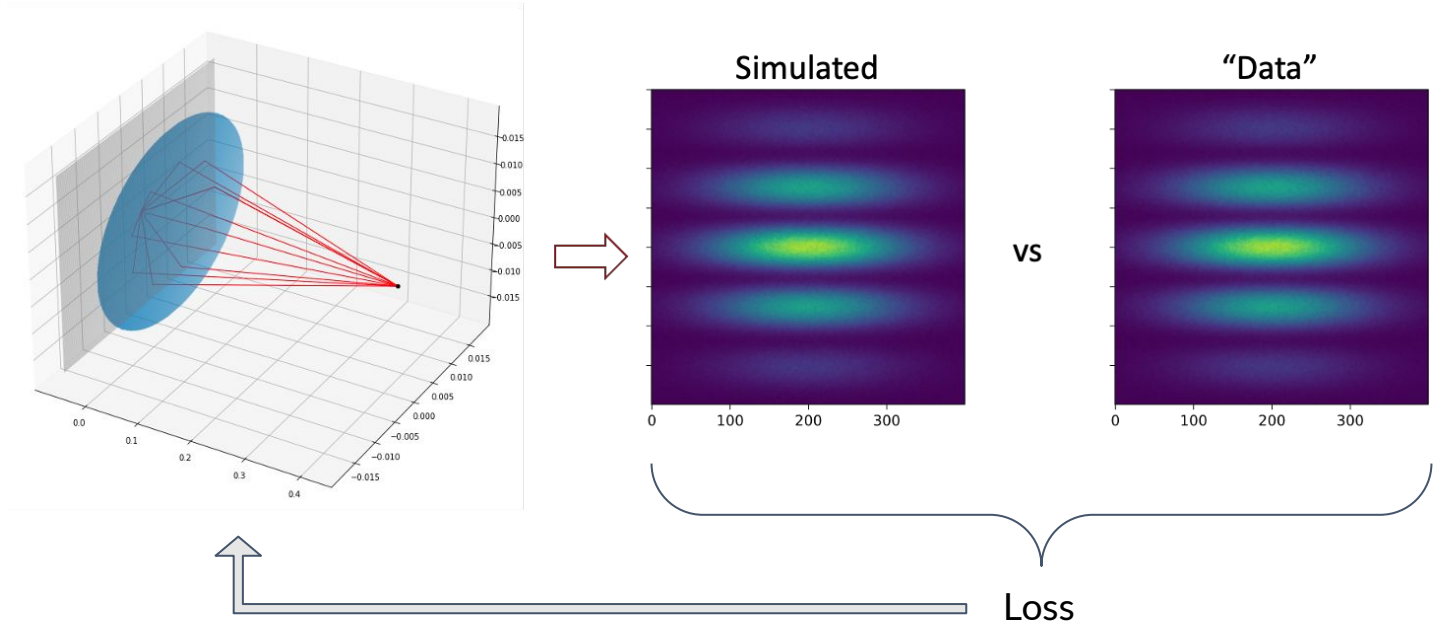


Derivatives give access to sensitivities of generated data w.r.t. parameters

- Enables *gradient-based optimization* approaches to calibration, optimization & reconstruction, etc.
  - E.g. **How should the cloud parameters ϕ be updated so simulations match data?**

$$\frac{\partial}{\partial \phi}$$

End-to-end differentiable simulator enables gradient-based optimization

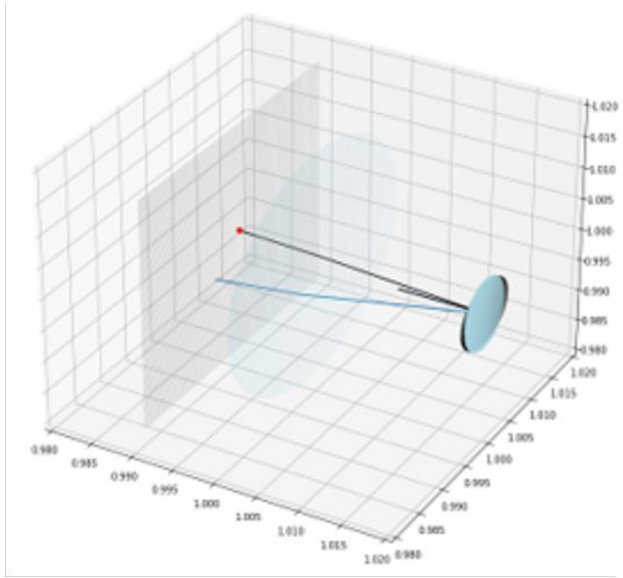


Reconstruction

- Update cloud parameters Θ with gradient descent

Calibration

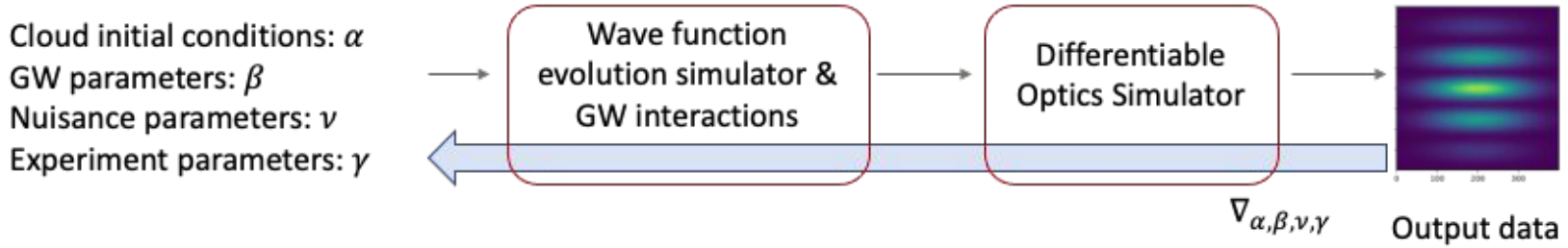
- Simulation and captured data do not match
- Fitting the simulator parameters to match the captured data



Simulation and Analysis

SLAC is interested in working on other simulators as well

- Long-term vision: fully differentiable inference pipeline (diffoptics, wave function evolution, phase aberrations, experimental settings, etc)



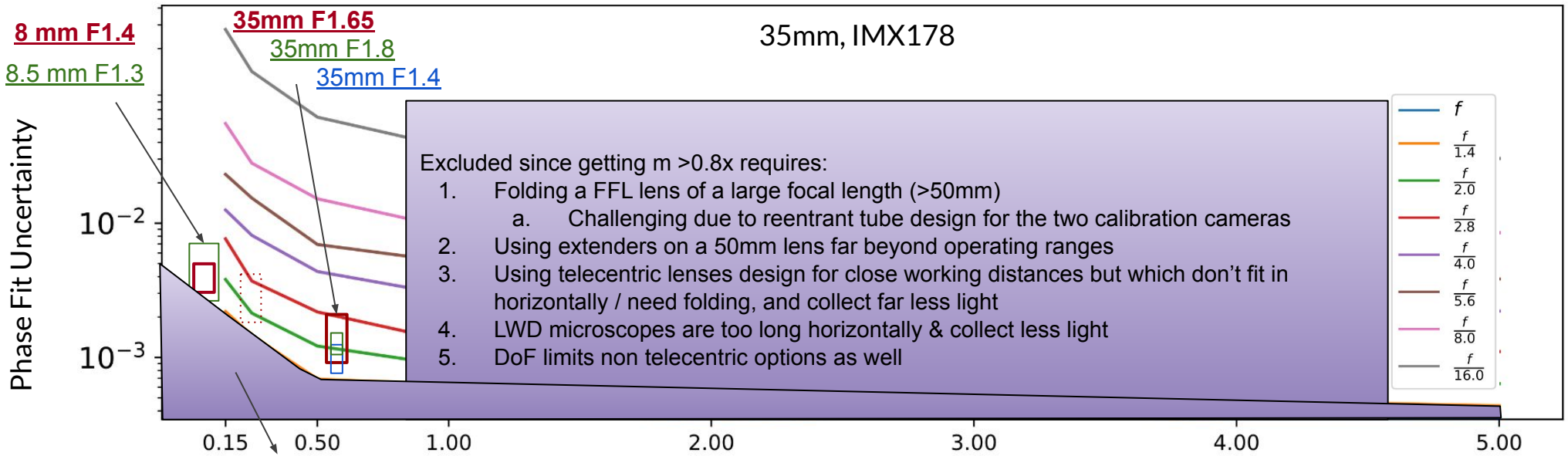
Broader Question: How will we use simulation in image and time-series analysis?

- Set upper bounds on accuracies and uncertainties?
- Fully integrated in the reconstruction algorithms?
- Used to generate datasets for training and analysis?

Computing demands will depend on how we use simulations

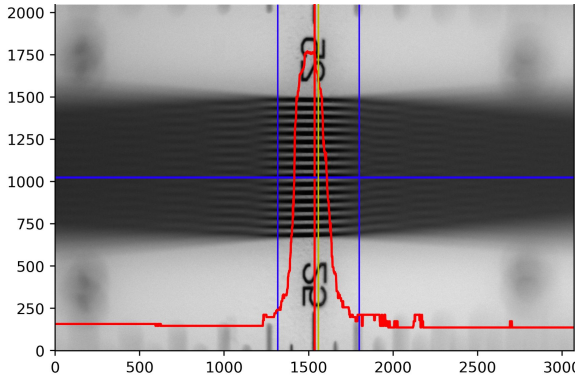
Diagnostic Imaging System

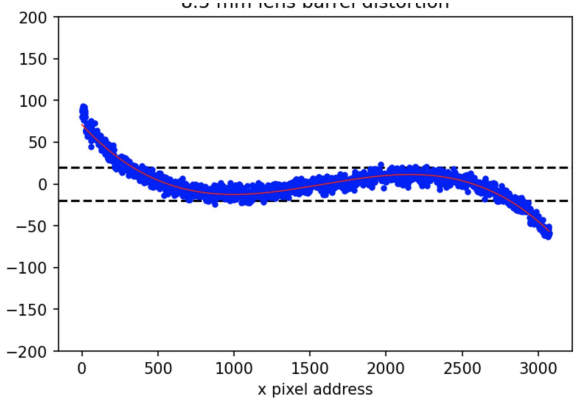
Extensive lab tests and system optimization (see backup)



Excluded because faster lenses call for longer working distances, even small focal lengths like 25mm (500mm); Need many extensions for the target magnifications

Magnification





Resources for Diagnostic Camera System - Physics

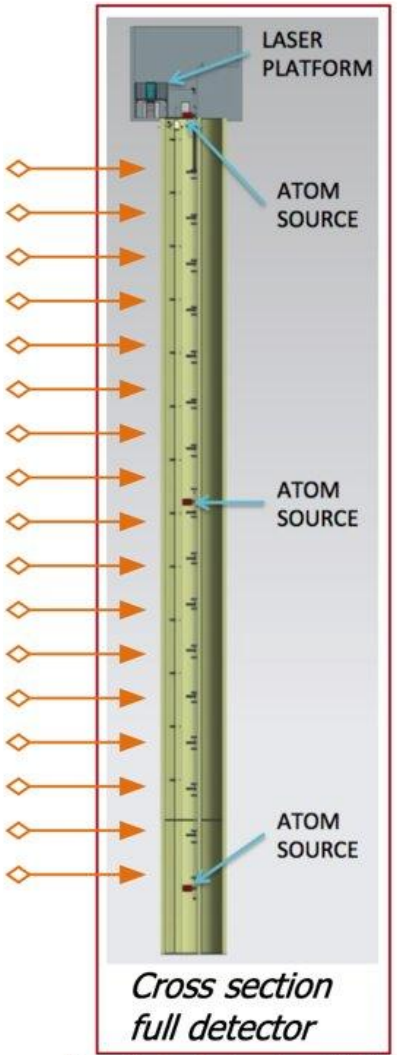
Item	Value	Notes
# cameras	14	One 6MP calib. cameras, 14 nodes
# images / drop	1	Reading 1 station / drop
Image size	10 MB	Could do image cropping if needed
Data / drop	O(10 MB)	Reading 1 station / drop
Drop rate	1 - 10 Hz	Assume 1Hz to start, 10 Hz stretch goal
Run-time / day	24 hrs	
Calib. data rate	O(1 TB - 10 TB/day)	
Extracted data	O(10 MB/day)	Assuming O(10) floating point parameters extracted per drop

Camera max FPS @ 12/16bit image w/ ISP off = 30 Hz
Running at 15 Hz -> 2 * min pixel exposures
Simulated Global Exposure min time of 1/30 s ~ 33ms

Can the Rpi handle 10Hz (100MB/s) image transfer to the server either directly or as a buffer?

Seems to set the data-processing rate for phase extraction

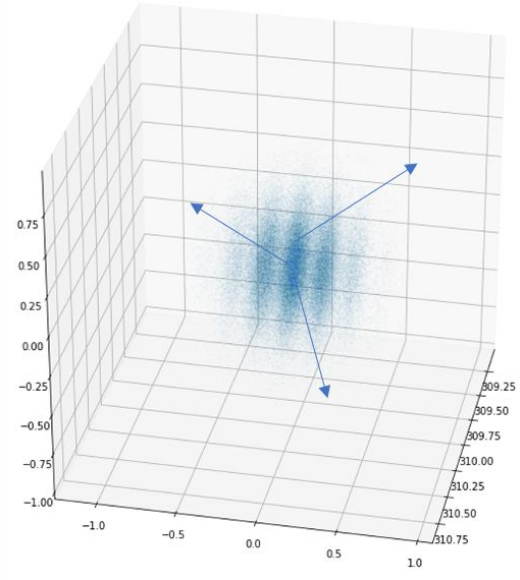
- How fast can we extract a phase on a Rpi?
- Data sizes are large, is all phase reconstruction online? Will we plan yearly re-processing for new reconstruction algorithm use?



Ray Tracing

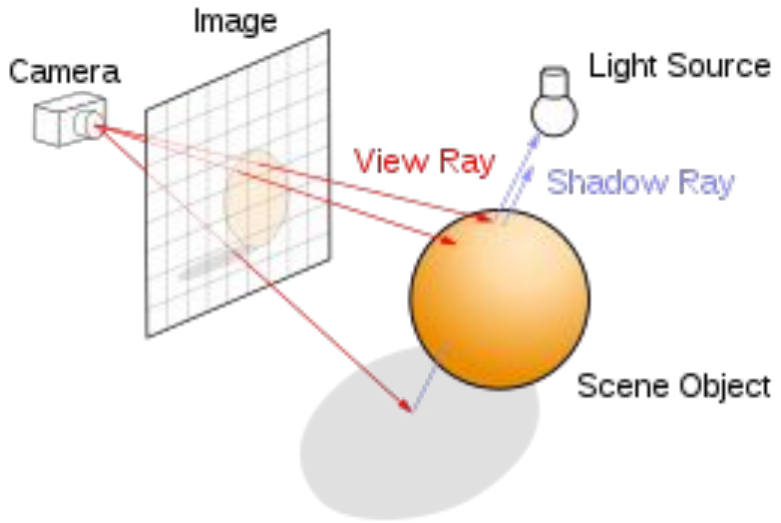
Forward Ray Tracing

- Simulates individual photons and their interactions with the system
- Photons can be sampled once and stored
- Vectorized
  - ~9B photons for a 10us image – O(1min)



Backward Ray Tracing

- Sampling rays from each pixels
- Integration of cloud density along rays
- **Computationally more efficient**



A little bit of 1D fourier space maths...

What does our density function look like?

$$g(x) = Ae^{-\frac{(x-b)^2}{2a^2}} \sin^2(kx + c)$$

In Fourier space this is a convolution of the FTs from *N* and $\sin^2$

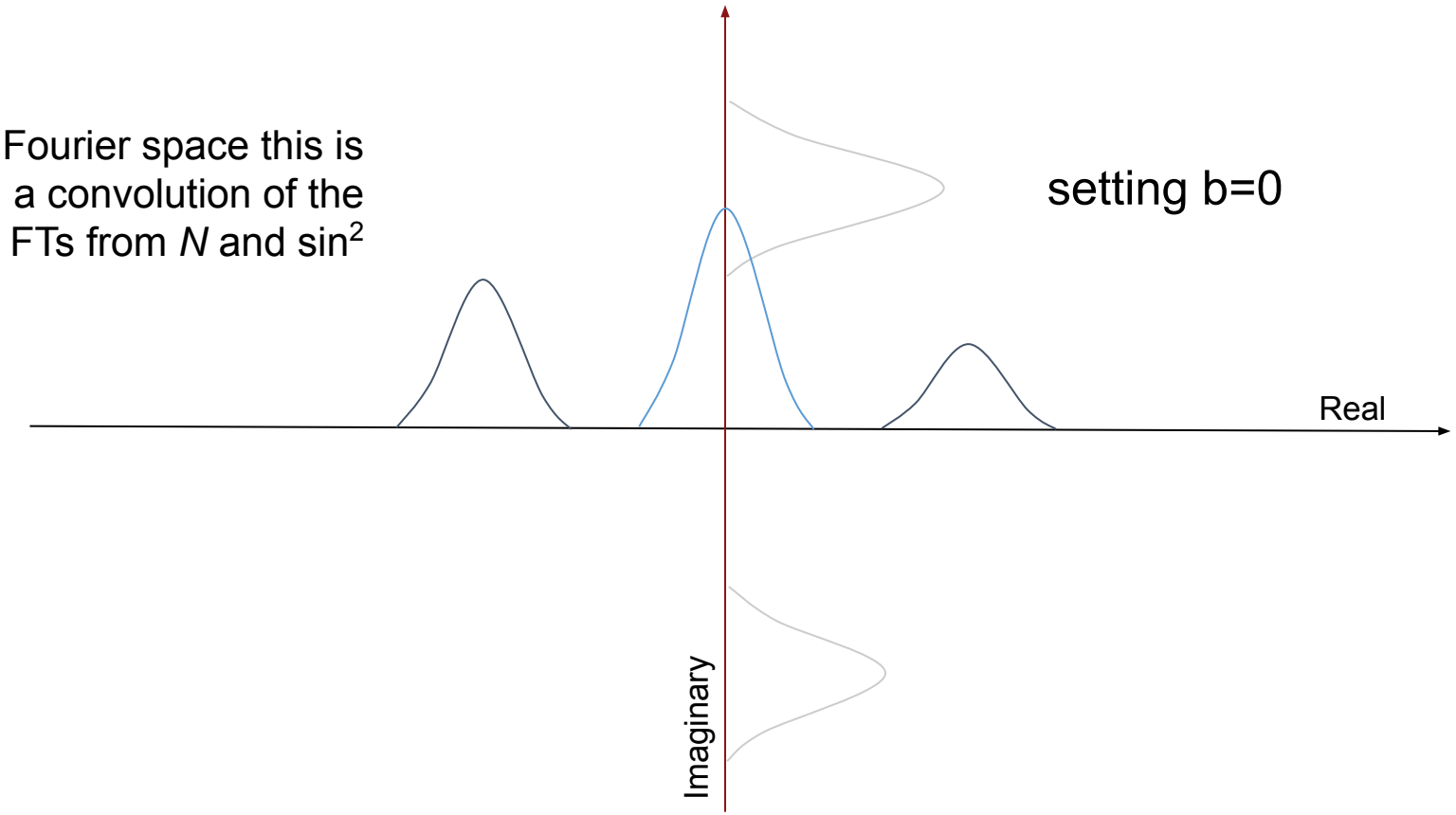
convolution	first function	$e^{i\,b\,\omega - 1/2\,(a^2\,\omega^2)}\,\lvert a\rvert$
	second function	$\frac{1}{2}\left(\sqrt{\frac{\pi}{2}}\right.$ $\left.\left(2\,\delta(\omega) - \frac{\delta(-2\,k+\omega)}{e^{(2\,i)\,c}} - e^{(2\,i)\,c}\,\delta(2\,k+\omega)\right)\right]$

A little bit of 1D fourier space maths...

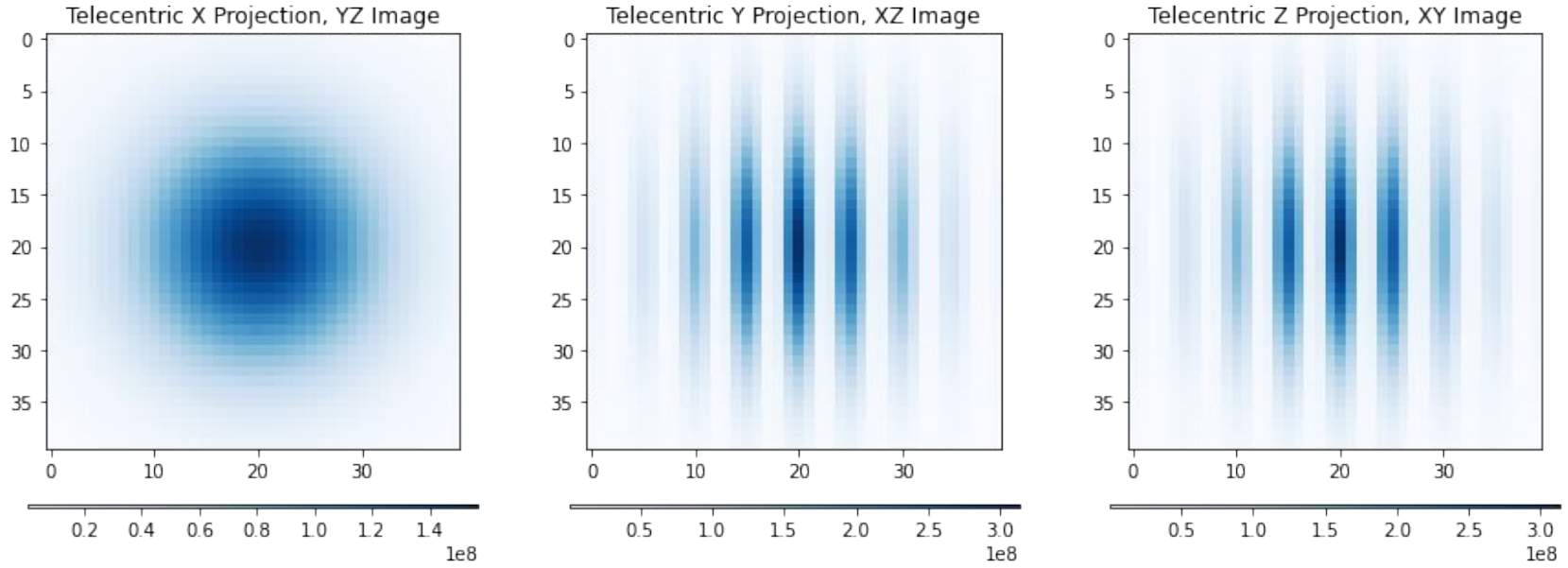
$$g(x) = Ae^{-\frac{(x-b)^2}{2a^2}} \sin^2(kx + c)$$

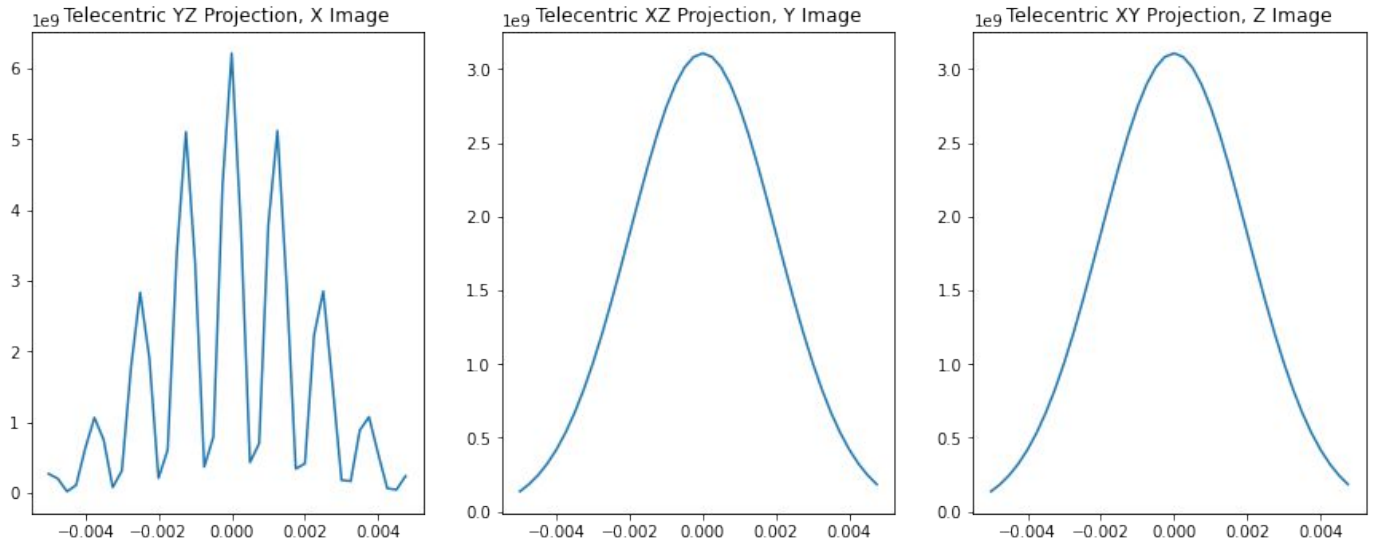
In Fourier space this is a convolution of the FTs from *N* and $\sin^2$

setting b=0

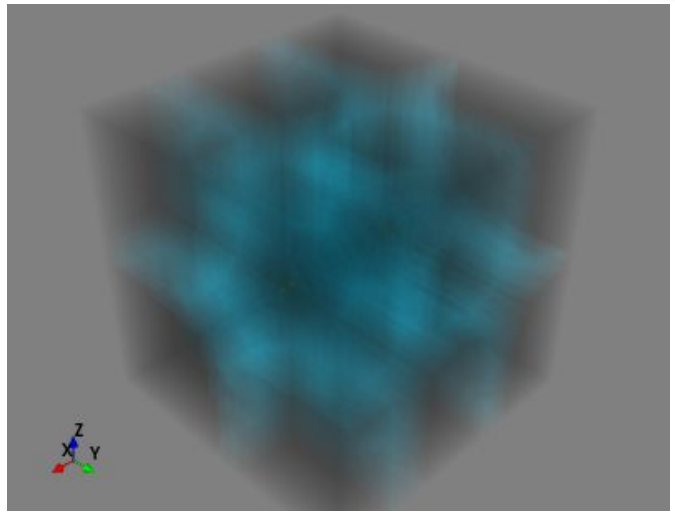


Density used



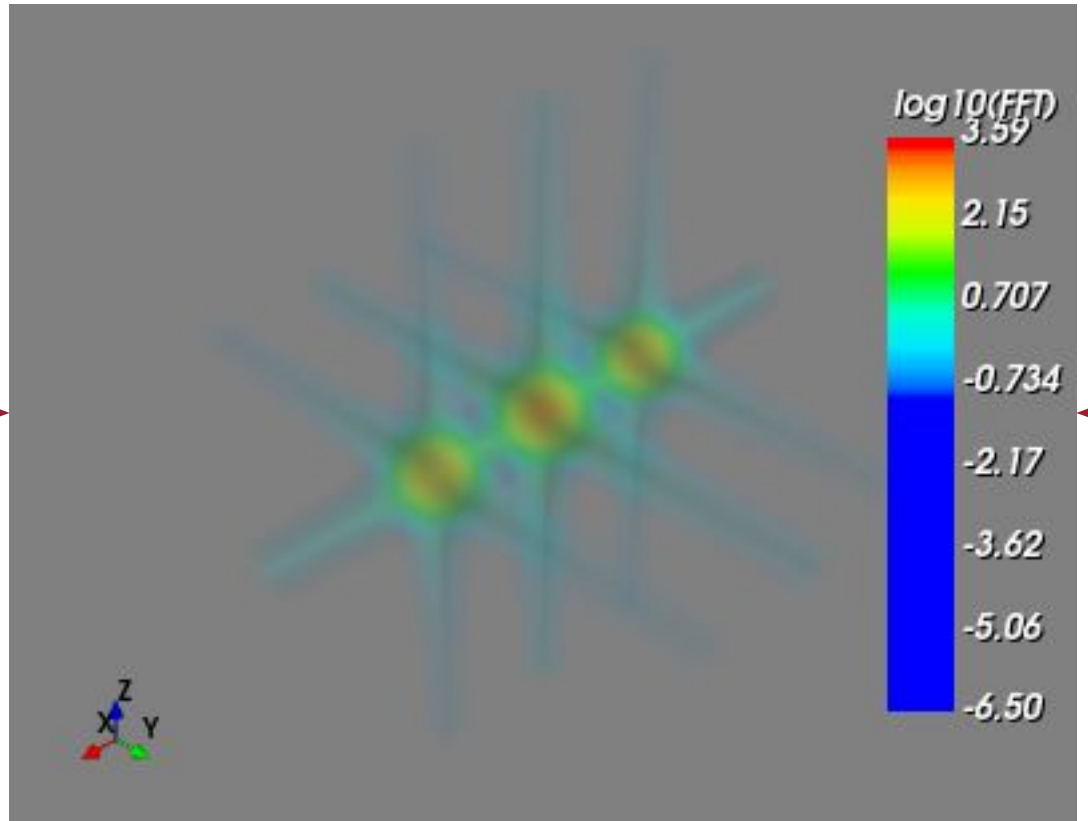


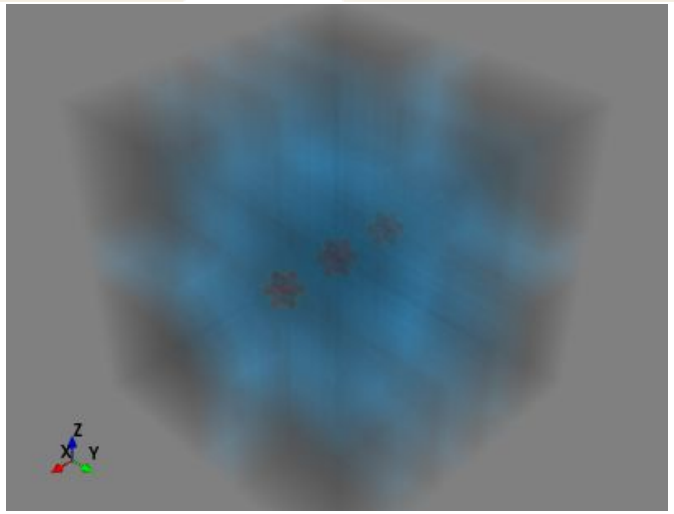
3D fourier picture of our cloud



Imaginary

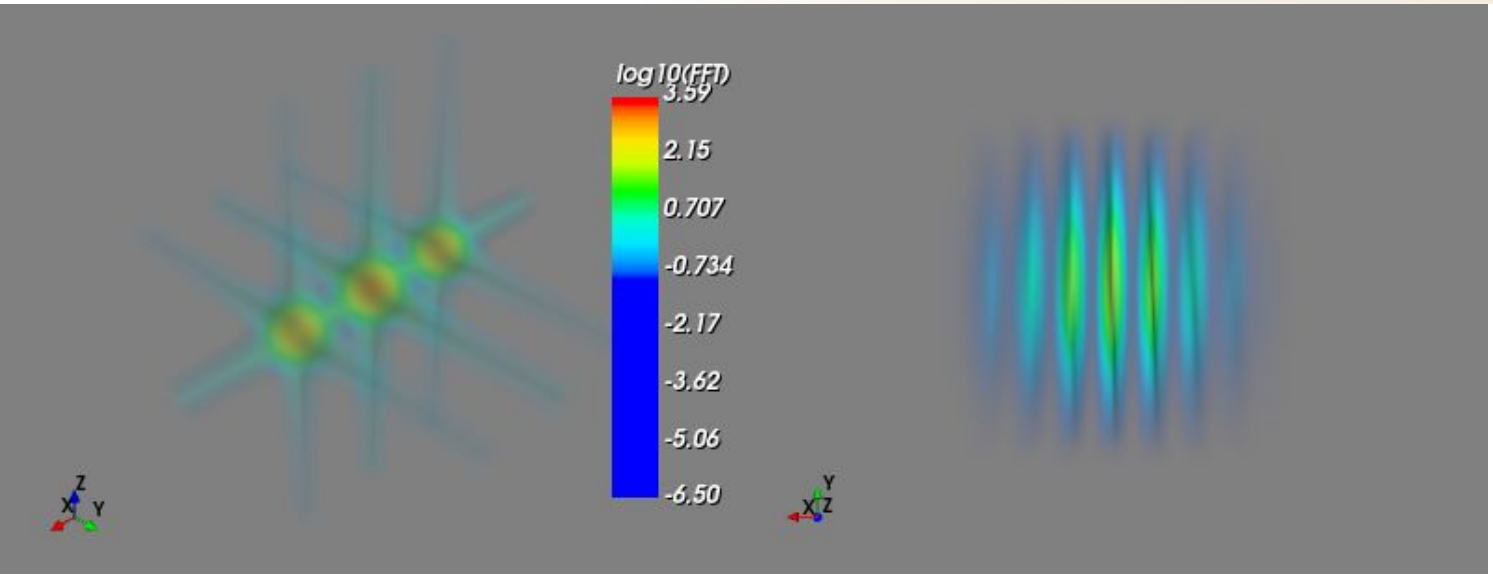
Absolute



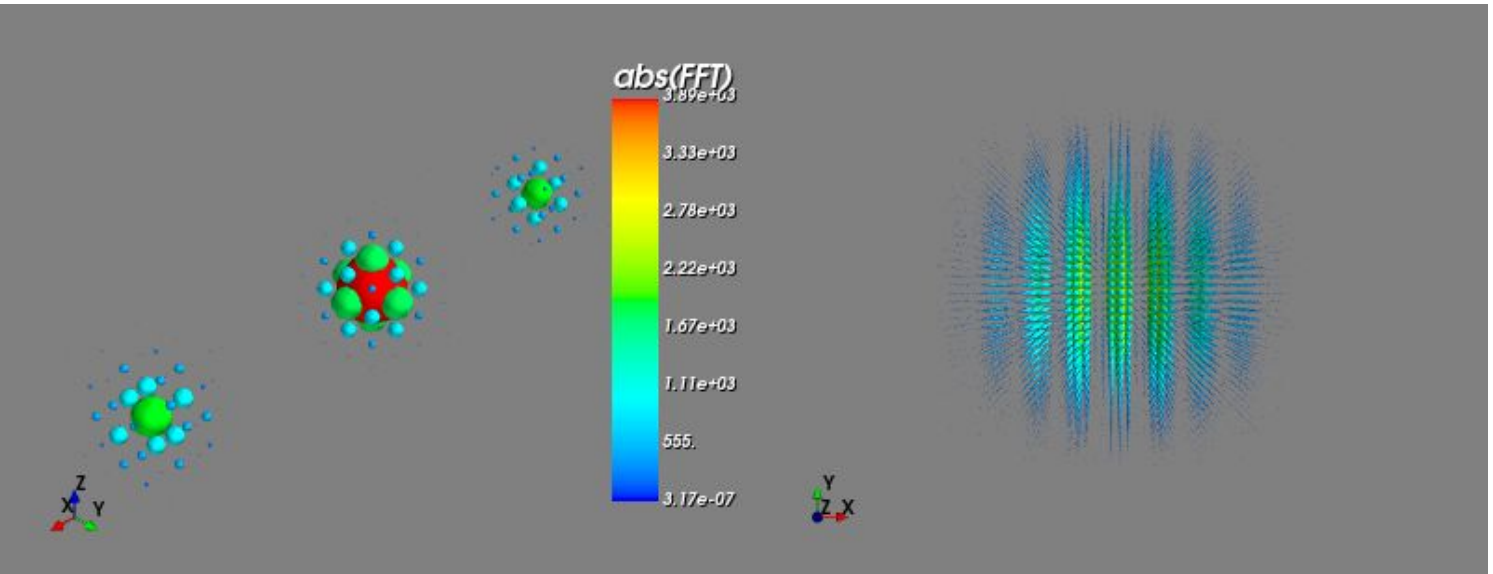


Real

How to best represent clouds?

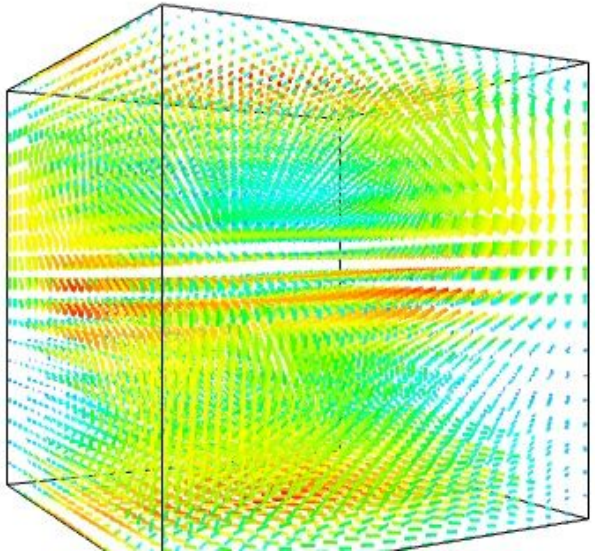


Colors - Interpolation



Glyphs - Points

How to best represent gradients?



Vector flow?

https://docs.enthought.com/mayavi/mayavi/mlab_case_studies.html#mlab-case-studies