



International  
Centre for  
Radio  
Astronomy  
Research

# Optimising the Processing and Storage of Radio Astronomy Data

Alexander Williamson, Pascal Elahi, Richard  
Dodson, Jonghwan Rhee, and Qian Gong



*ICRAR is a joint venture between Curtin University and The University of Western Australia and receives support from the Western Australian and Australian Governments.*





SKA-mid ~200 dishes

710 PB/year



512 stations of 256 antennas  
SKA-Low



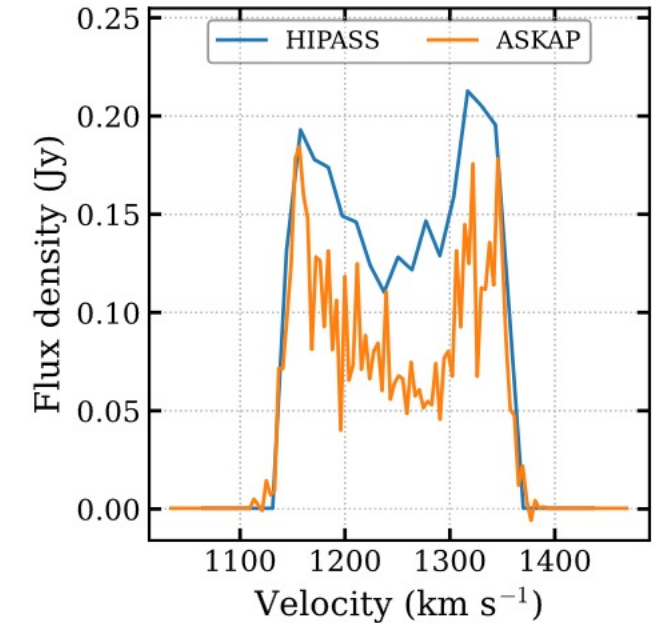
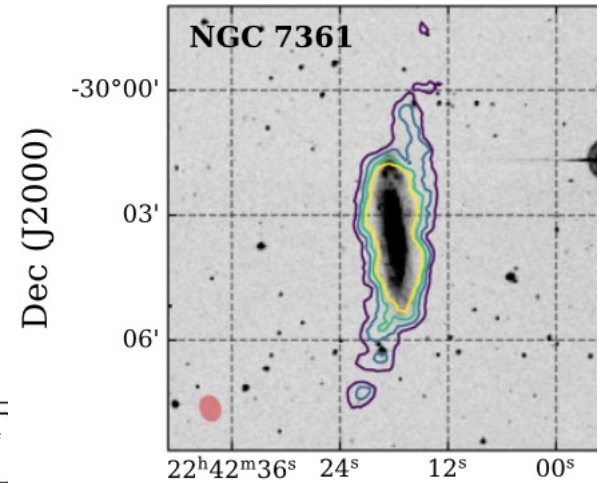
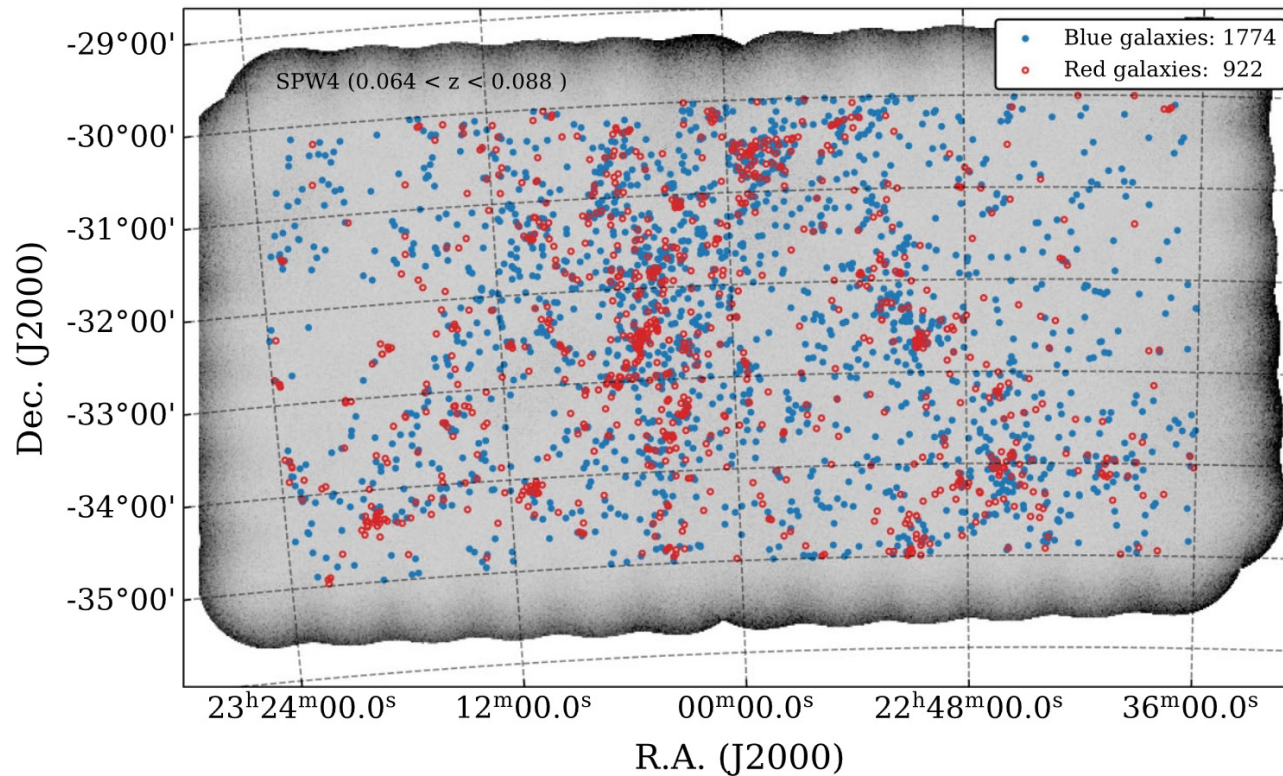


# ASKAP

(Australian SKA Pathfinder)

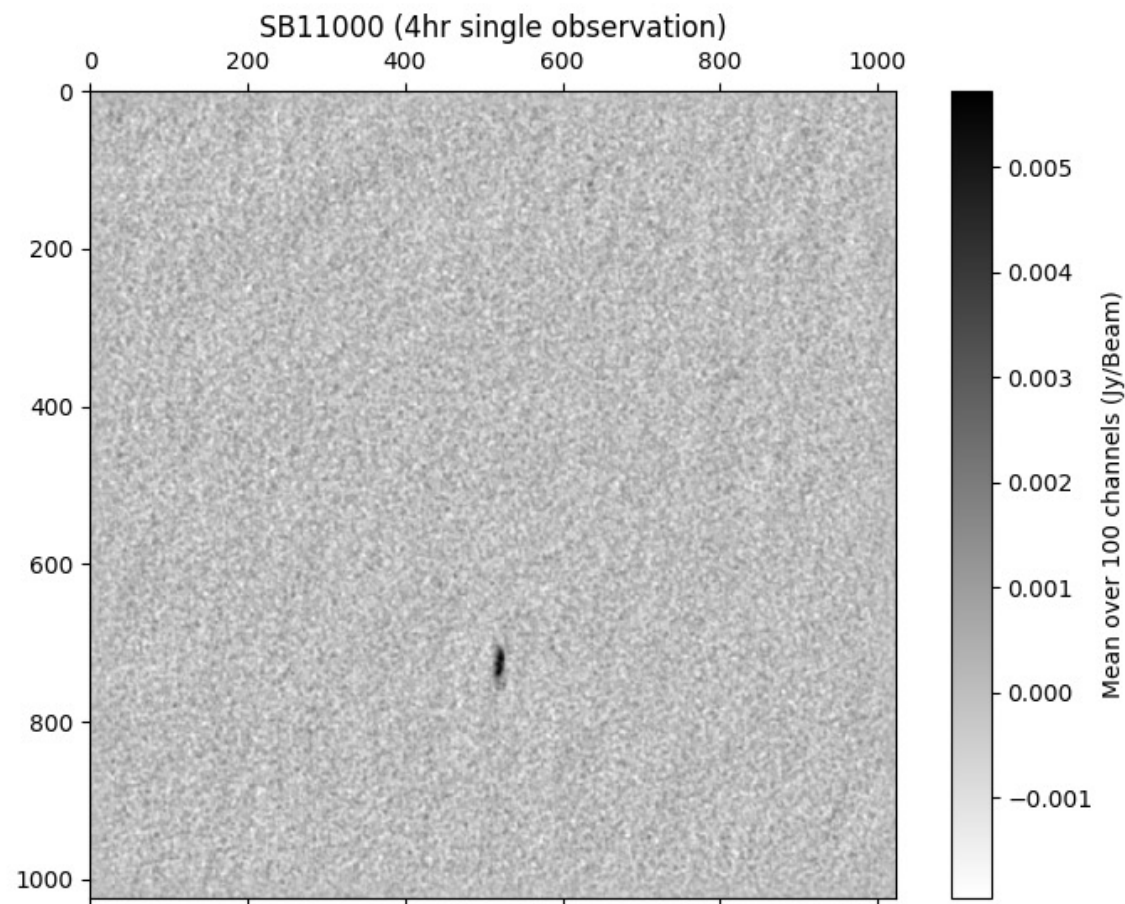
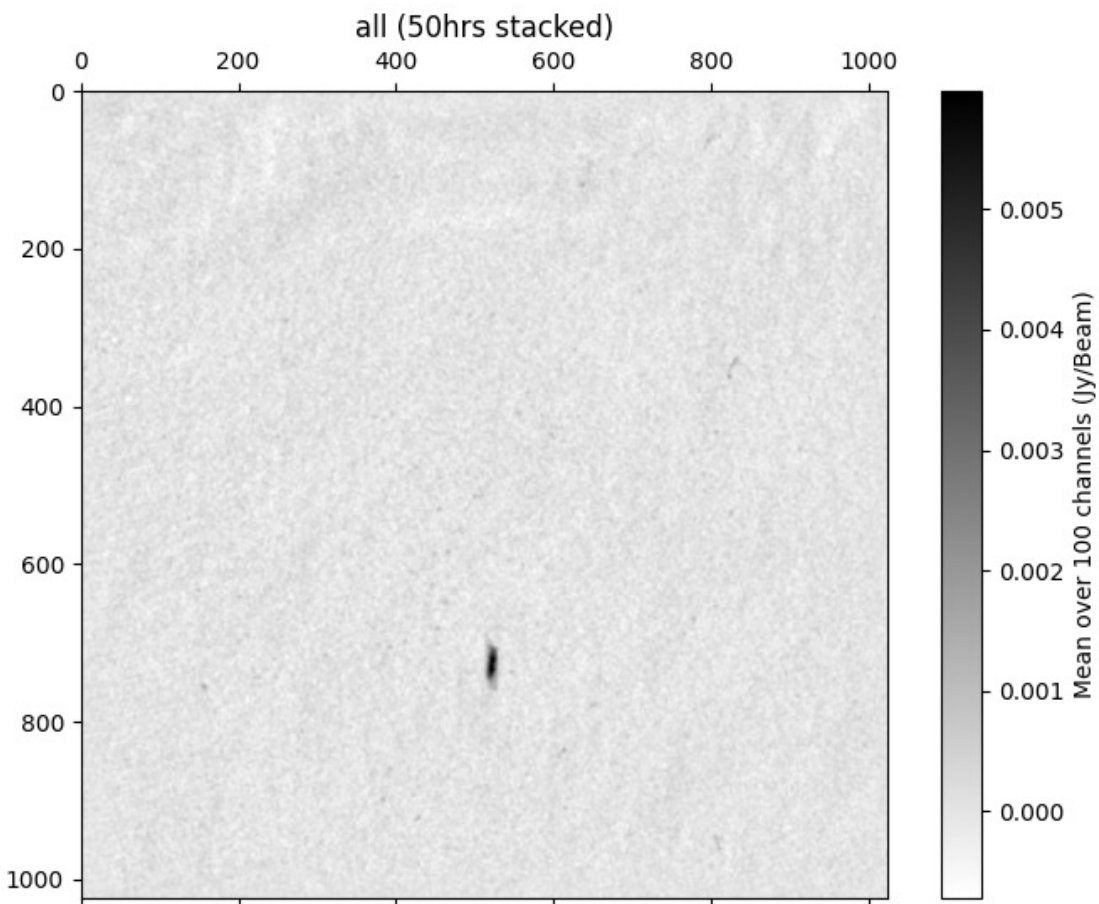


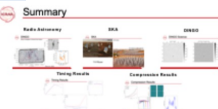
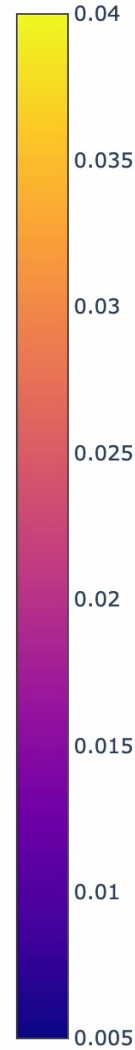
- 36 12m dishes
- Located at the Murchison Radio Observatory
- Provides a test bed for technology for the SKA and a full scientific instrument



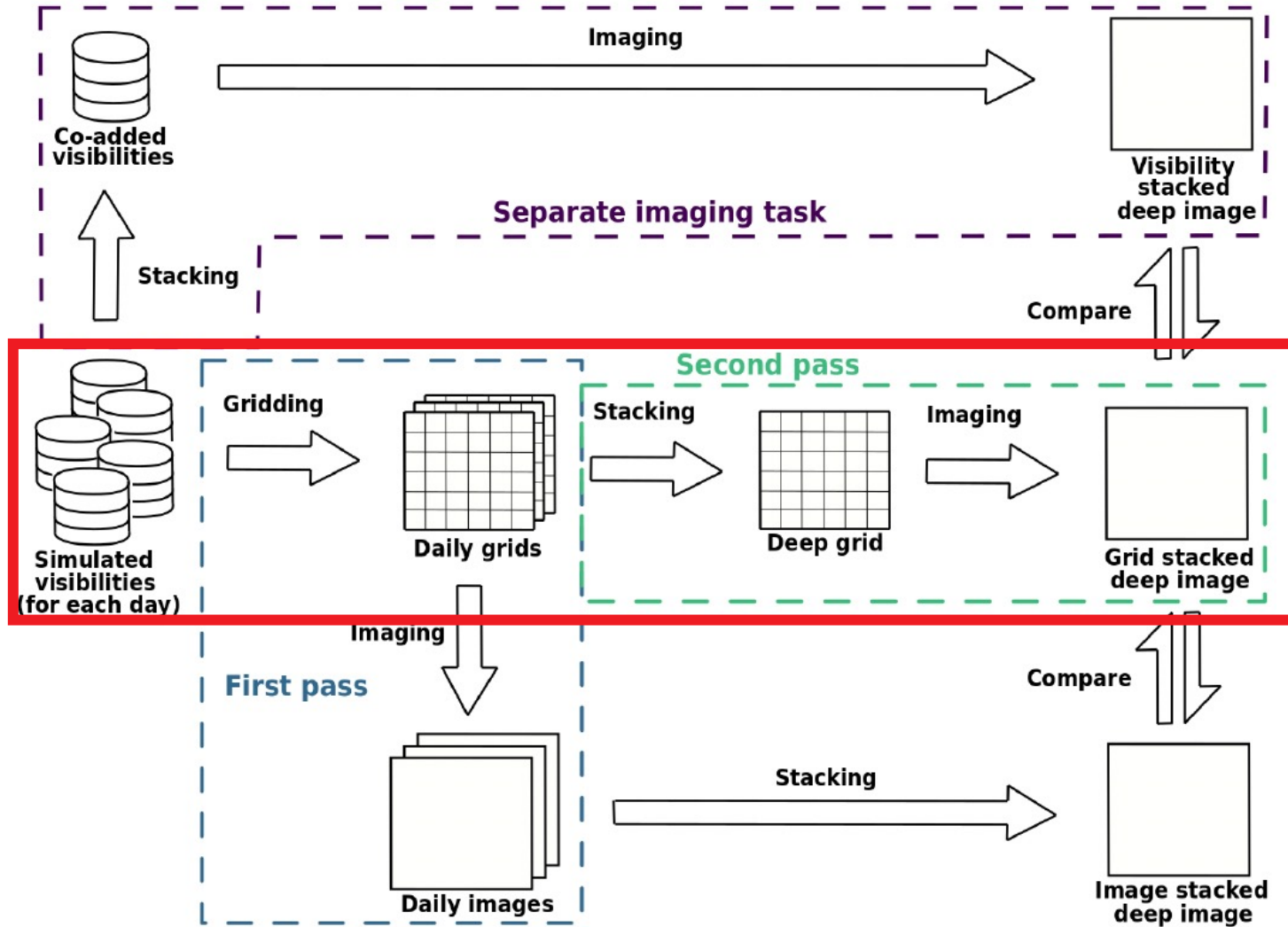
Rhee et al. (2023)







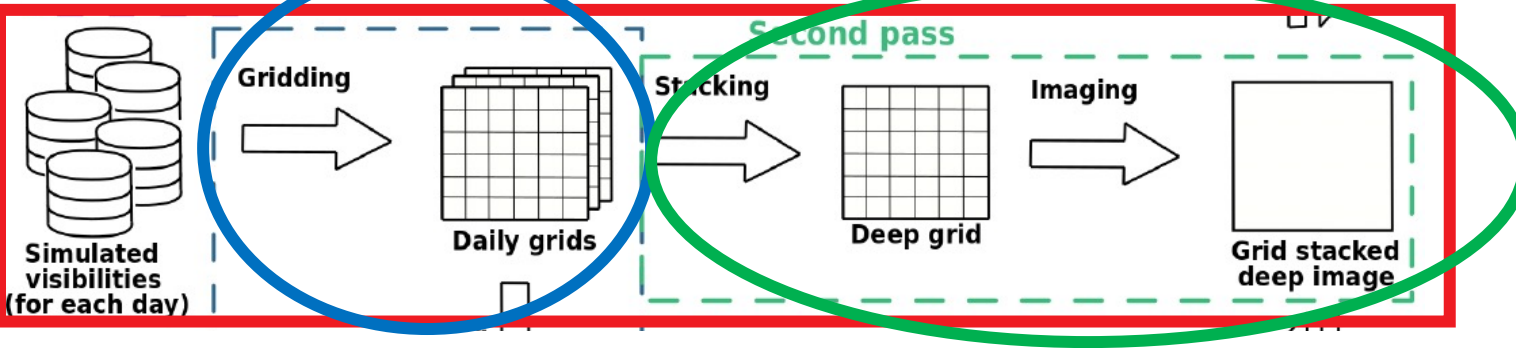
# DINGO Pipeline



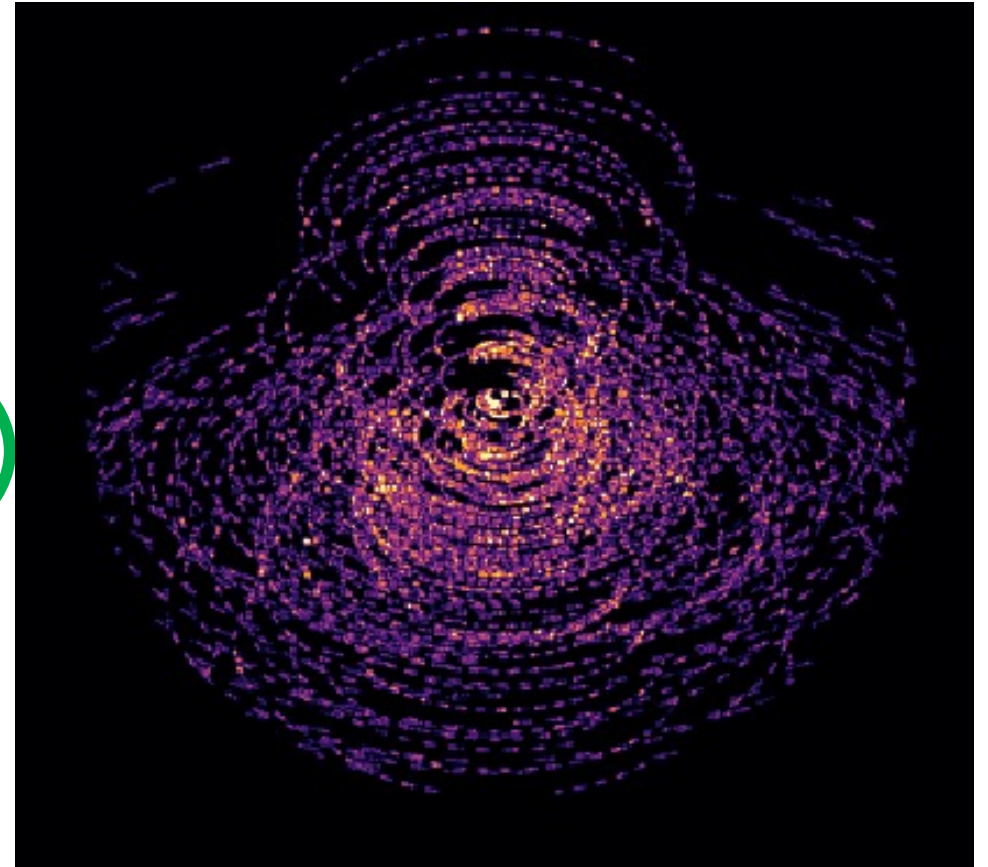
# DINGO Pipeline

Imager:  
gridding visibilities

cdeconvolver-mpi:  
stacking and imaging



Visibility grids  $\approx$  FFT







# Askapsoft with ADIOS2 and MGARD

---

Applications used for  
ASKAP data  
processing

Library that enables  
parallel I/O on HPC  
systems

Adaptive lossy compression  
algorithm

- Added a new interface mode for Askapsoft
- Uses ADIOS2 to perform I/O operations
- Facilitates in-memory compression via operator framework
- MGARD allows for us to set specific error bounds on compressed data, very good for scientific applications

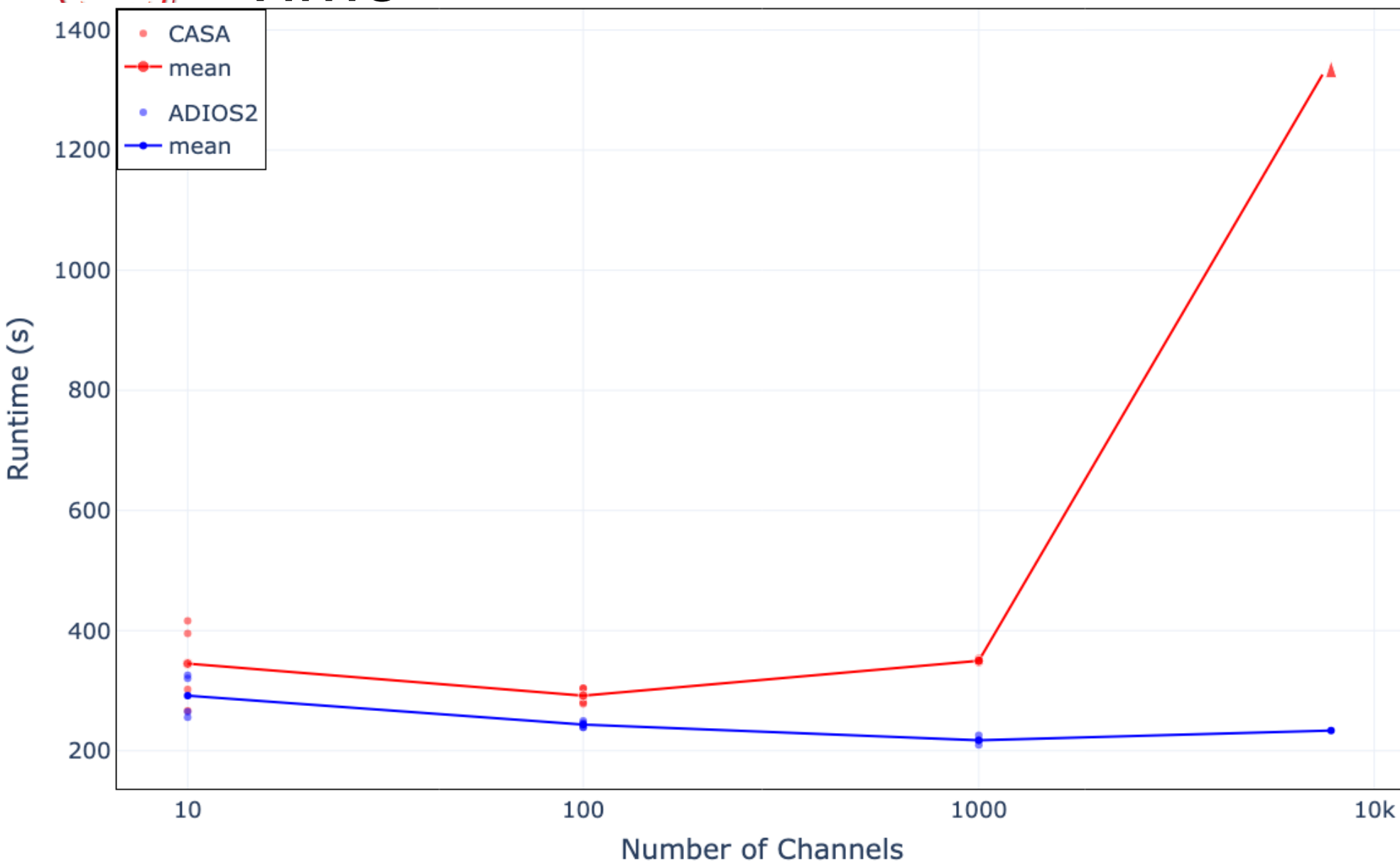


- Serial CASA
- Serial ADIOS2



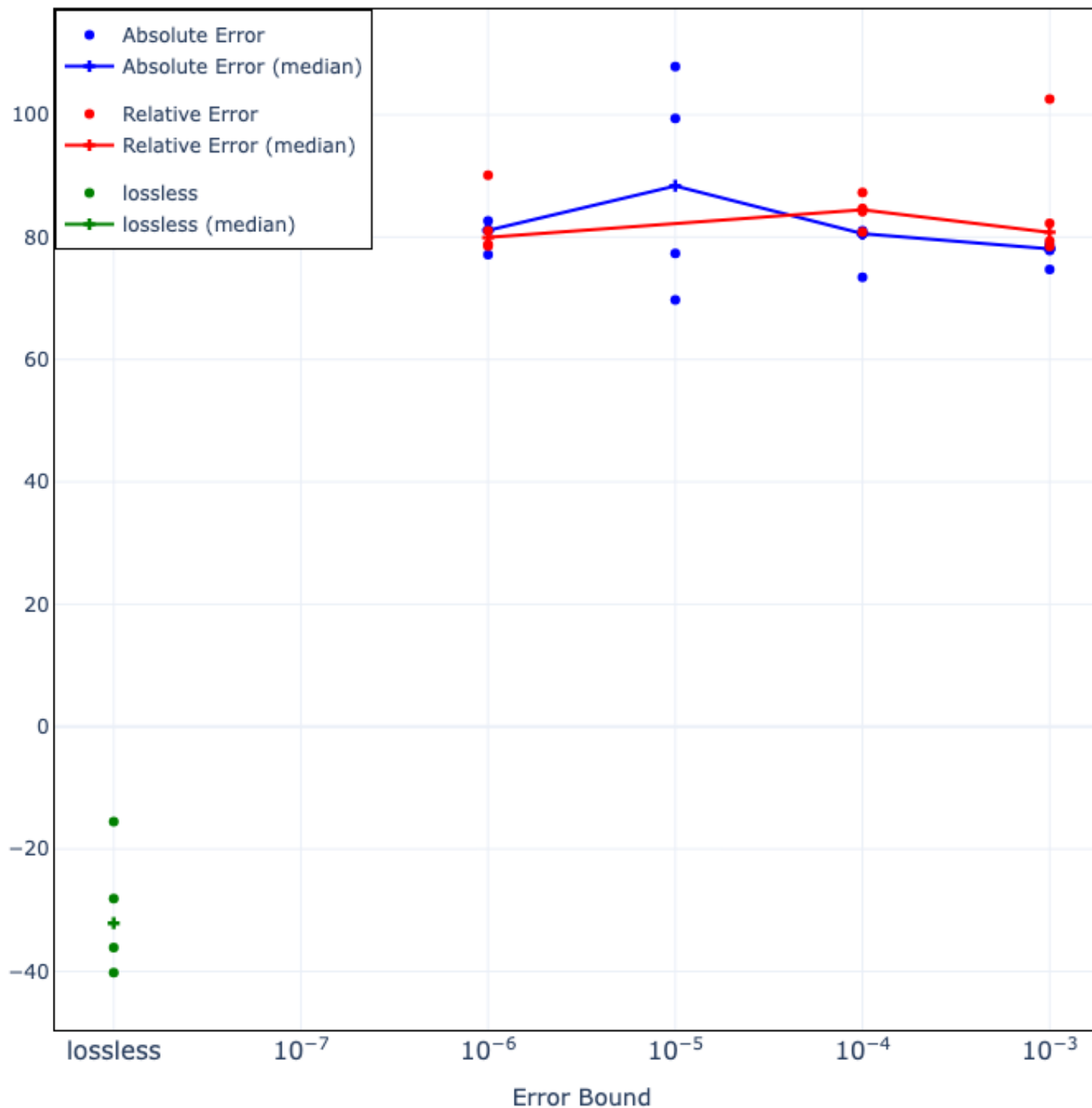


# Time



- Serial CASA
- Parallel ADIOS2

Time Difference (ADIOS2-CASA, s)

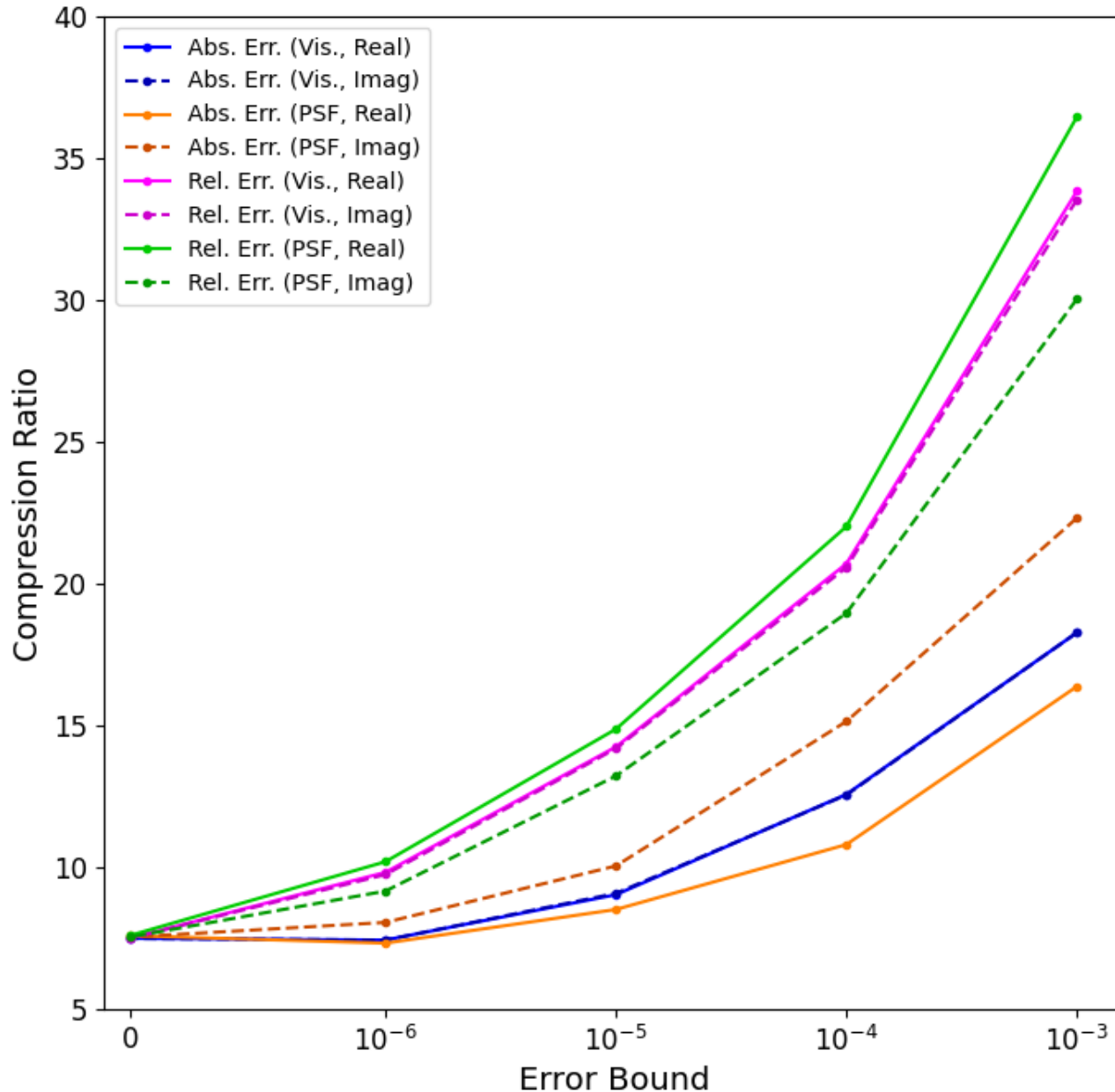


## Caveats:

- Small amount of data  
100 channels
- Purely CPU, no GPU compression
- The most basic of MGARD compression, doesn't use refactoring or anything special

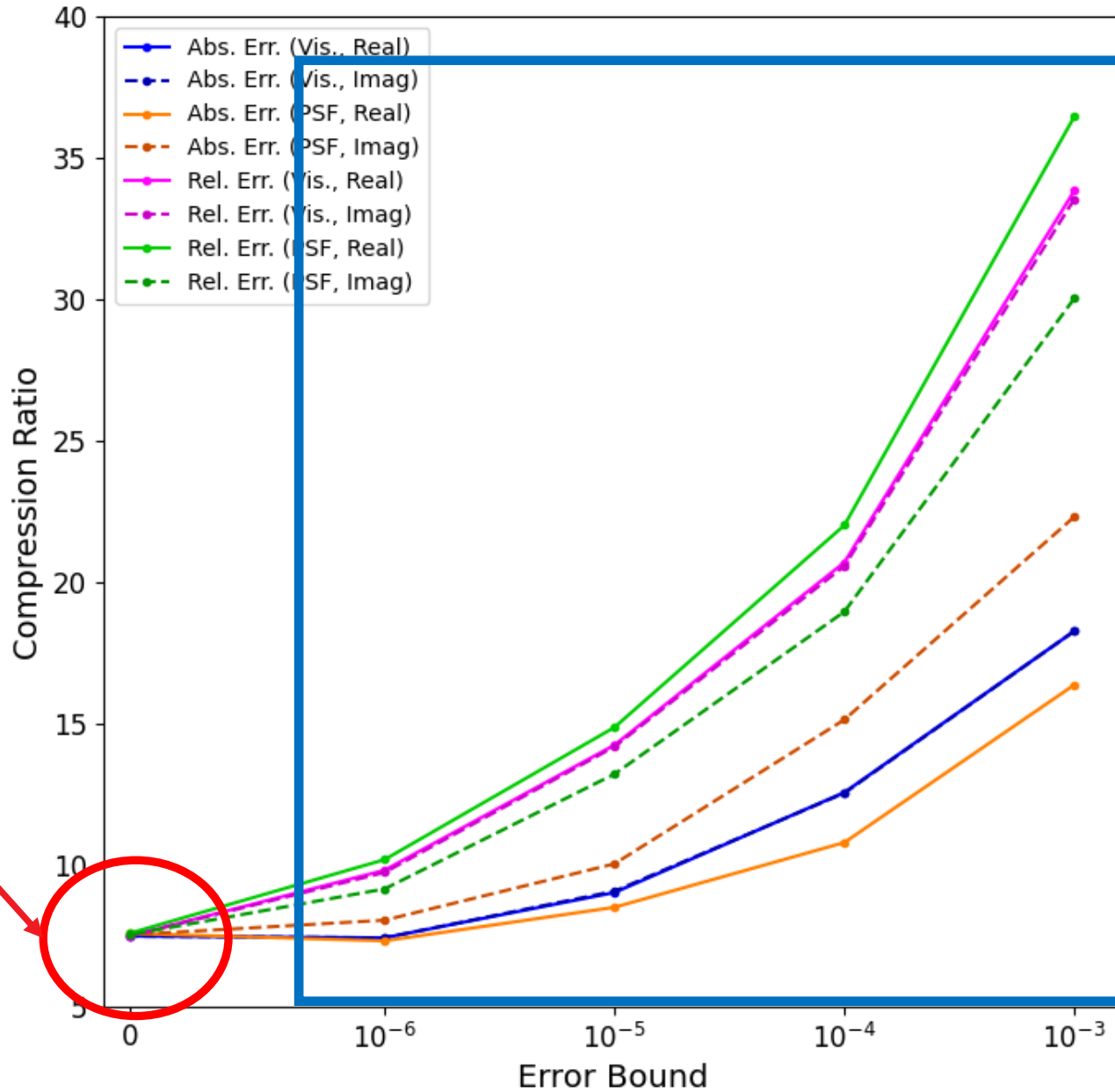


# Compression



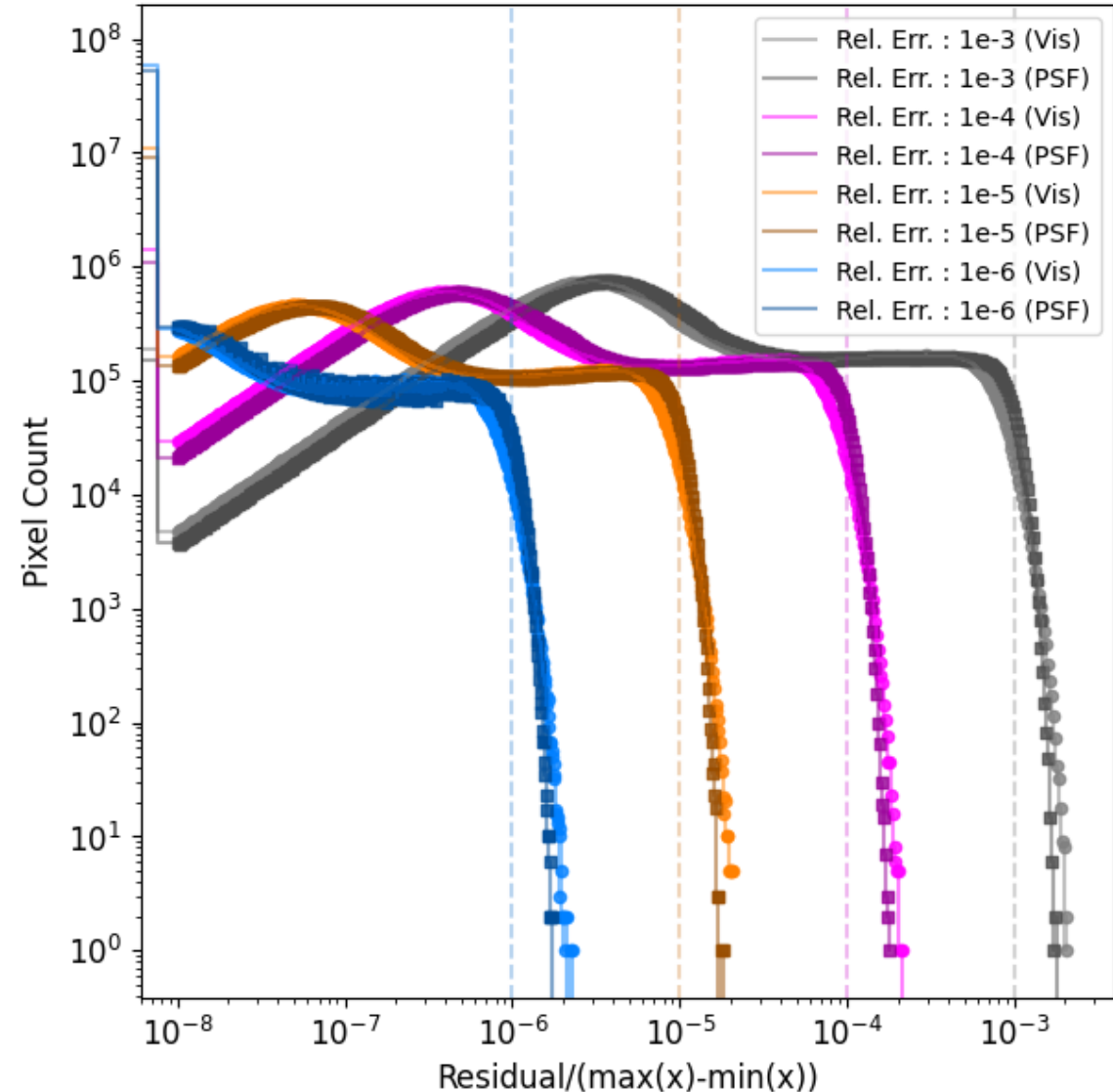
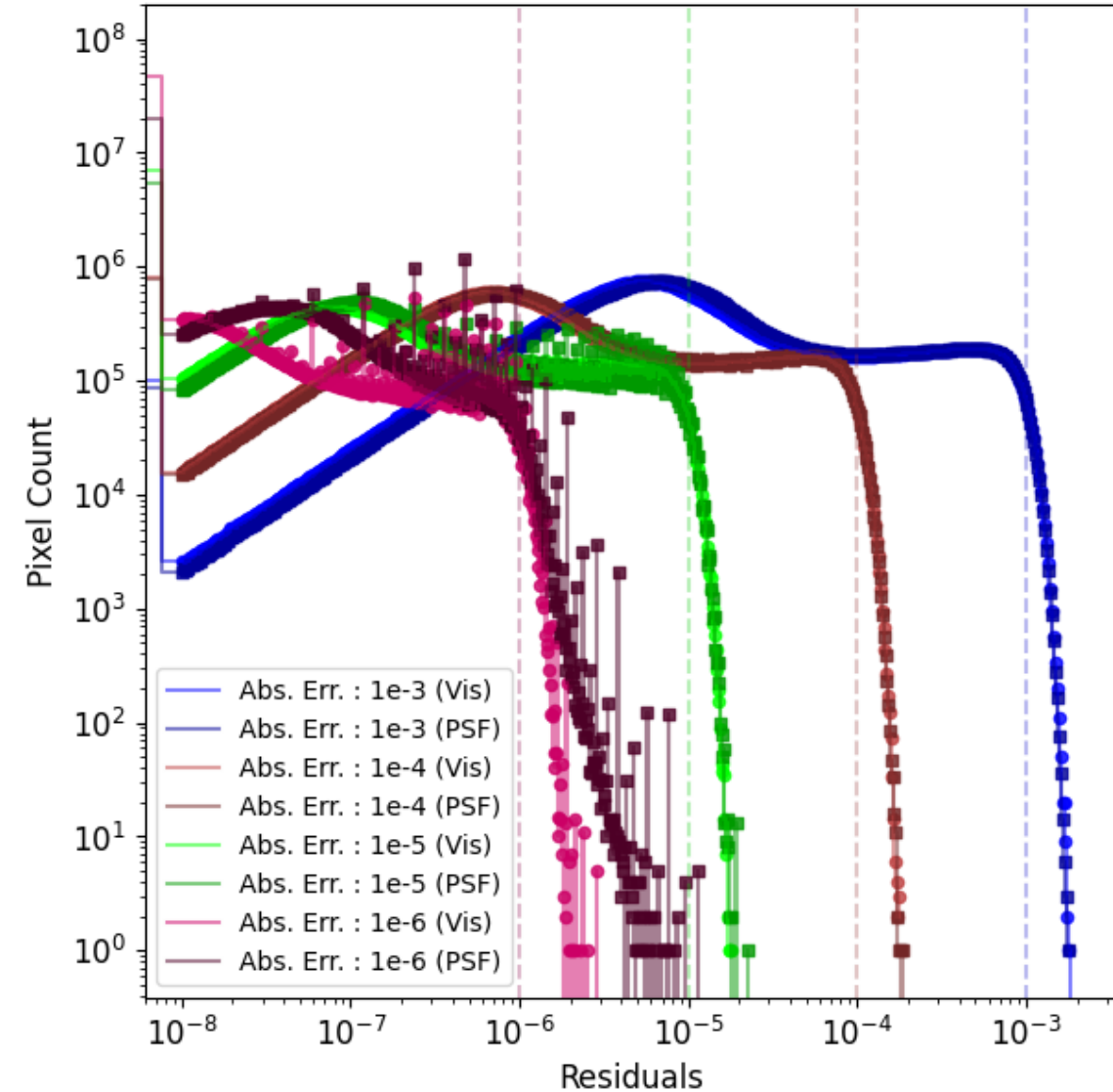
- Absolute and Relative Error bounds
- Complex Data
- Two different types of data: Visibilities and PSF

# Compression

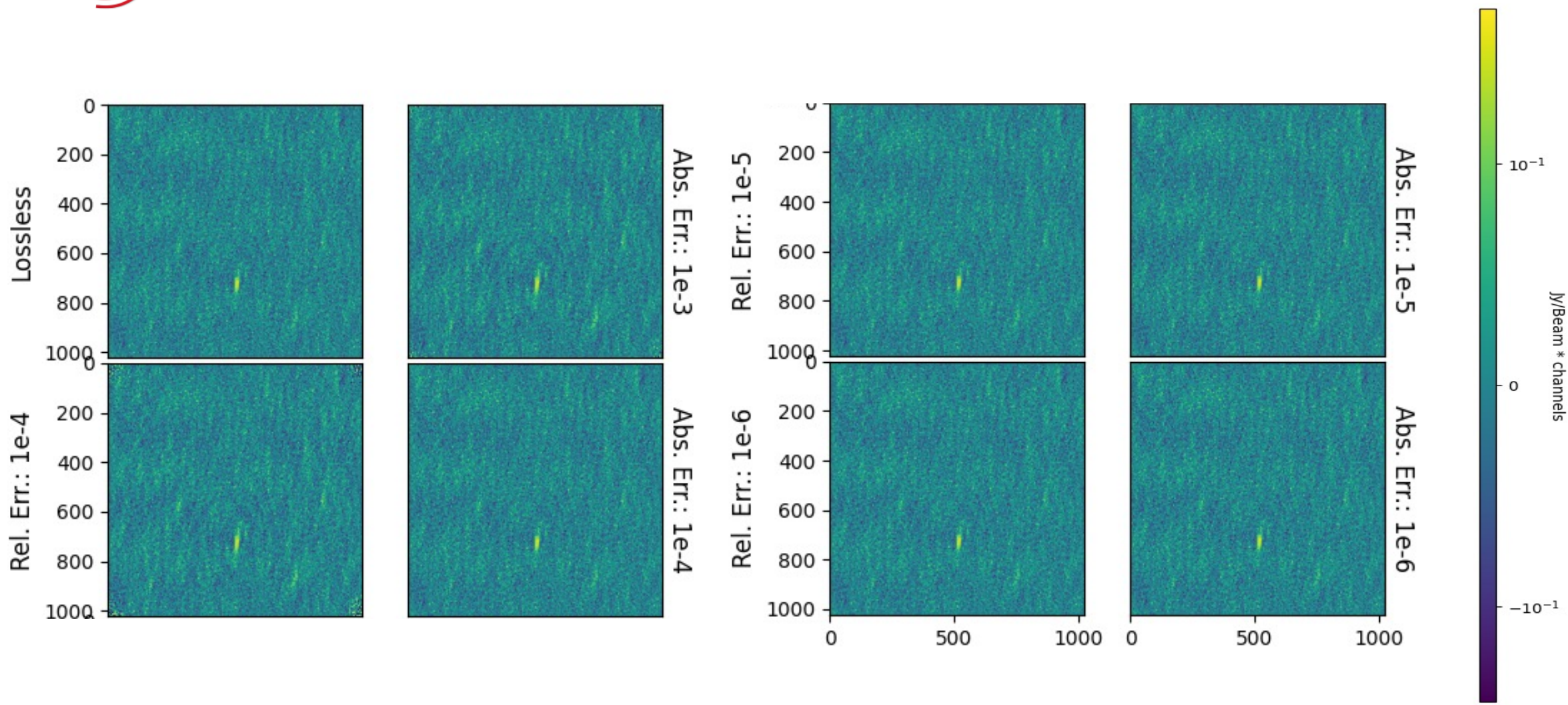




# Compression

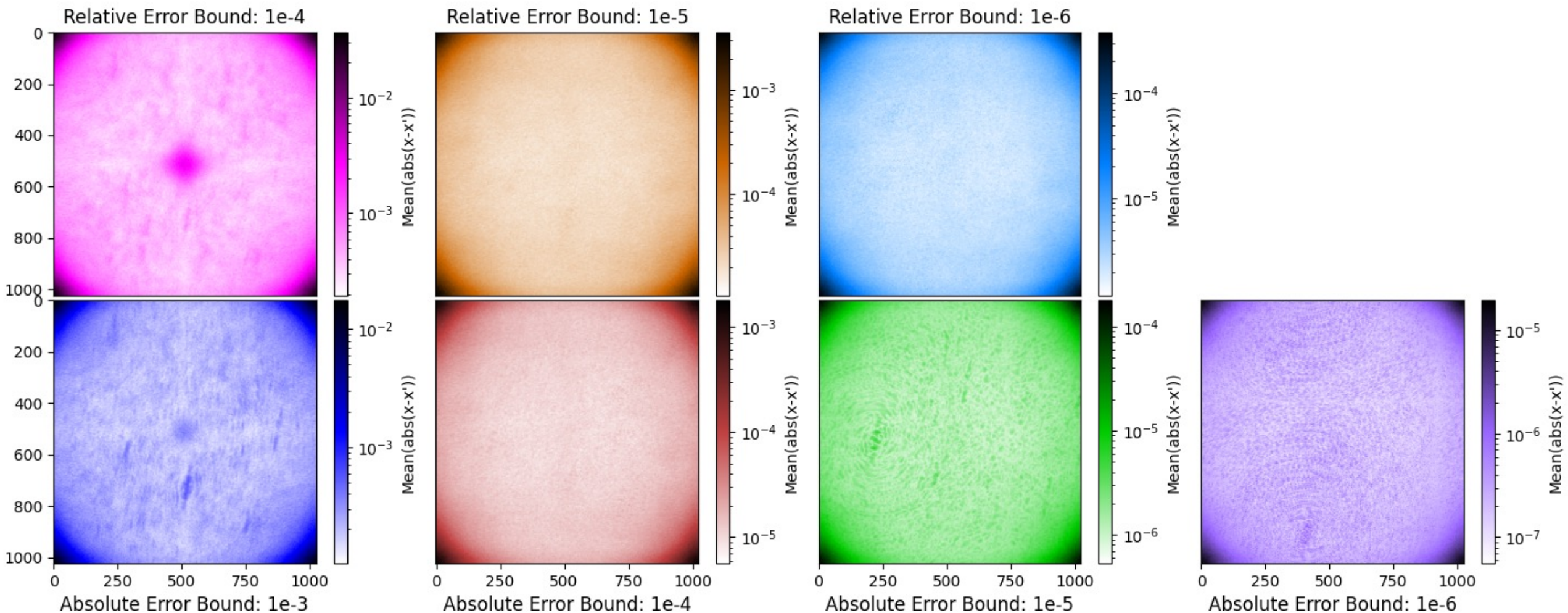


# Compression

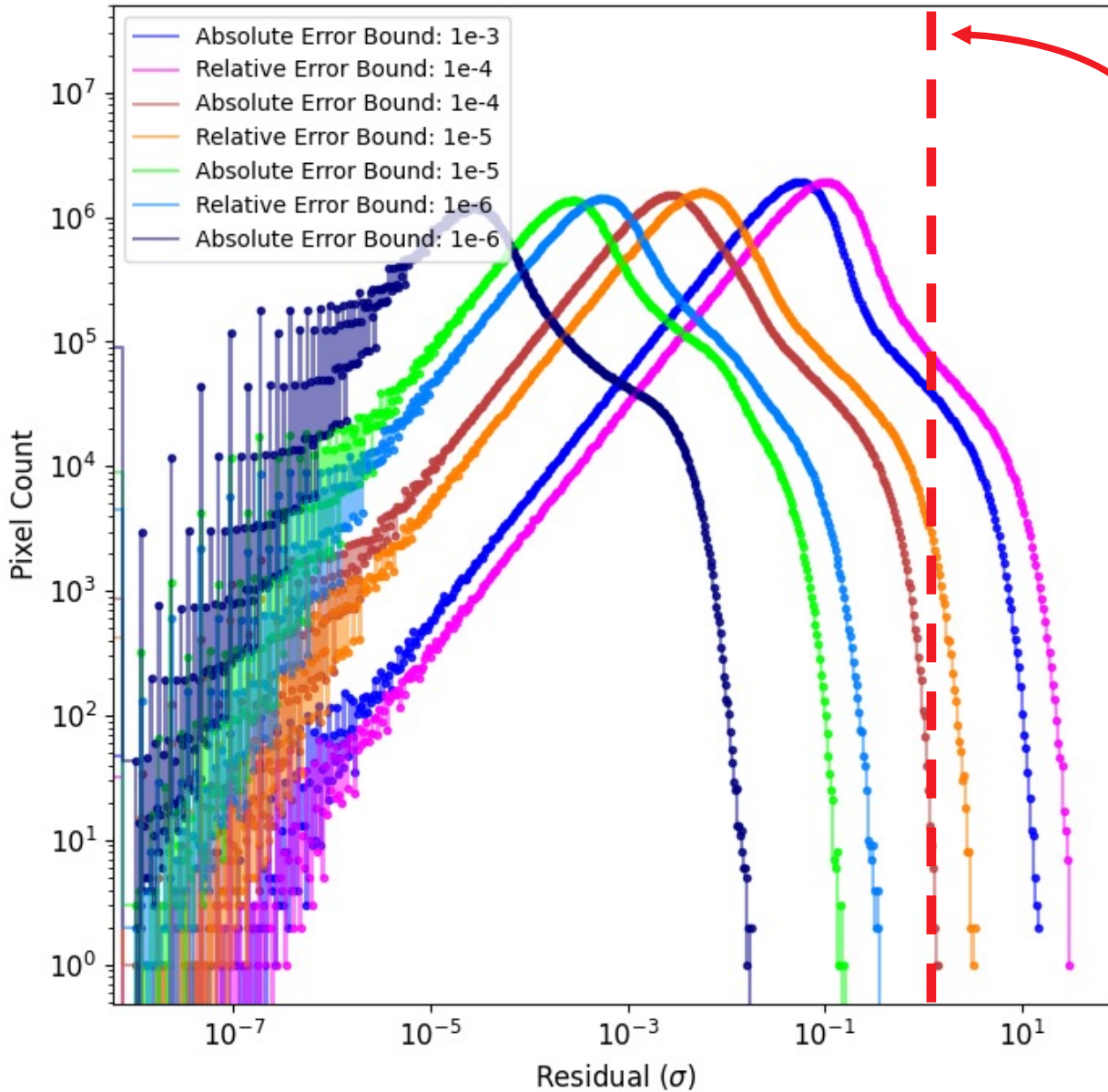




# Compression



# Compression

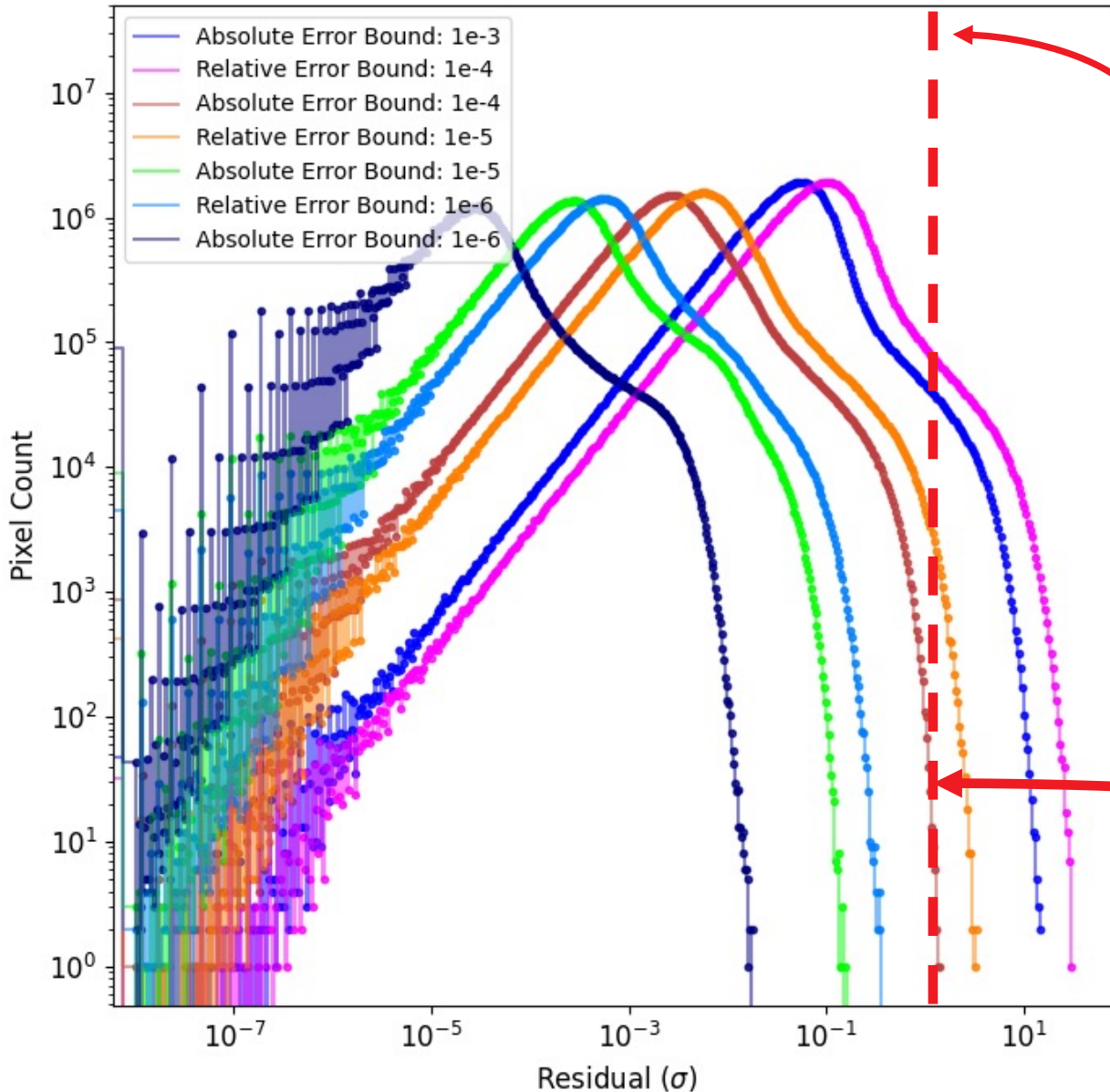


$1\sigma$ :

Errors above here are detectable  
(for general purposes)

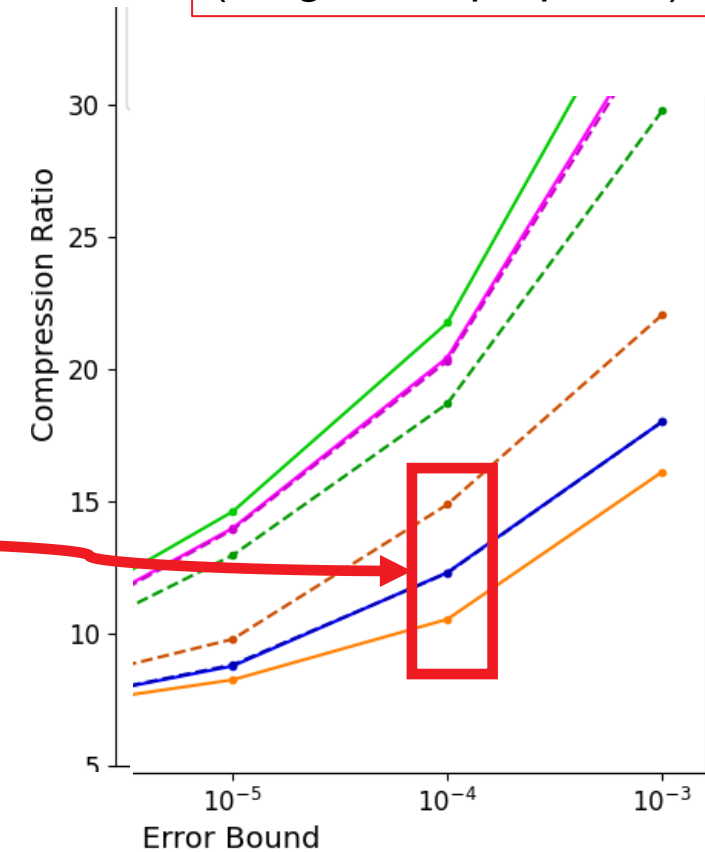


# Compression

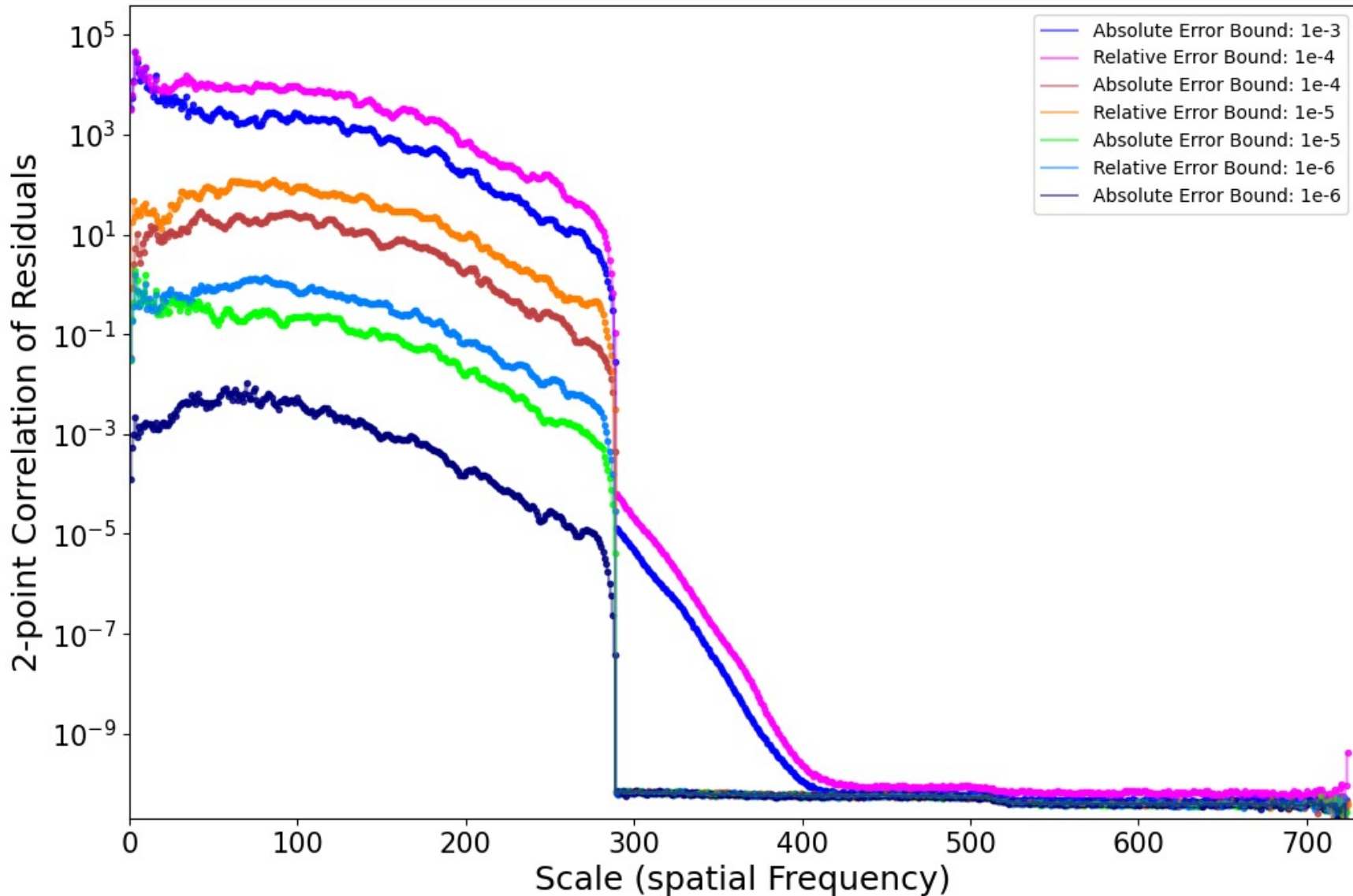


$1\sigma$ :

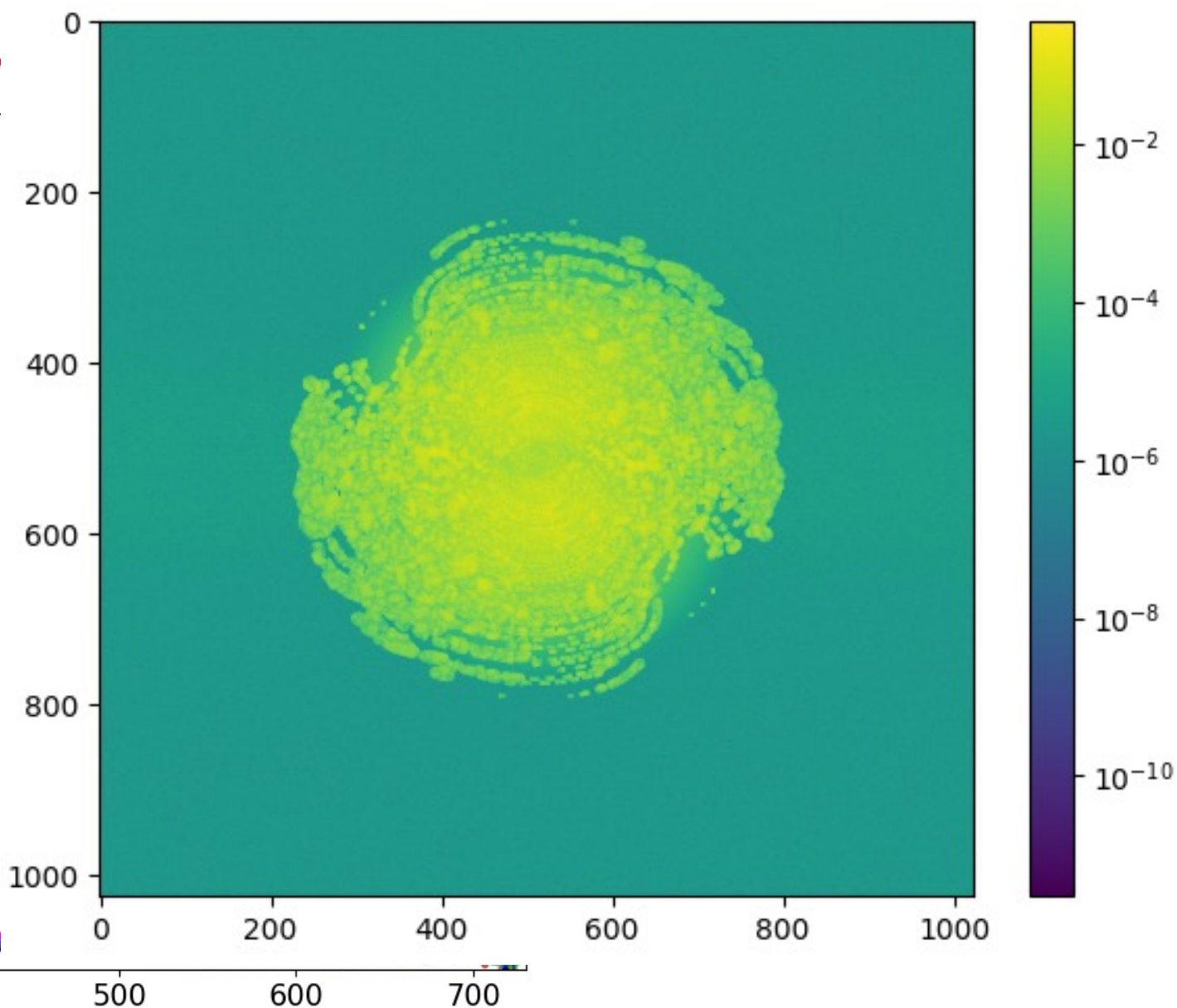
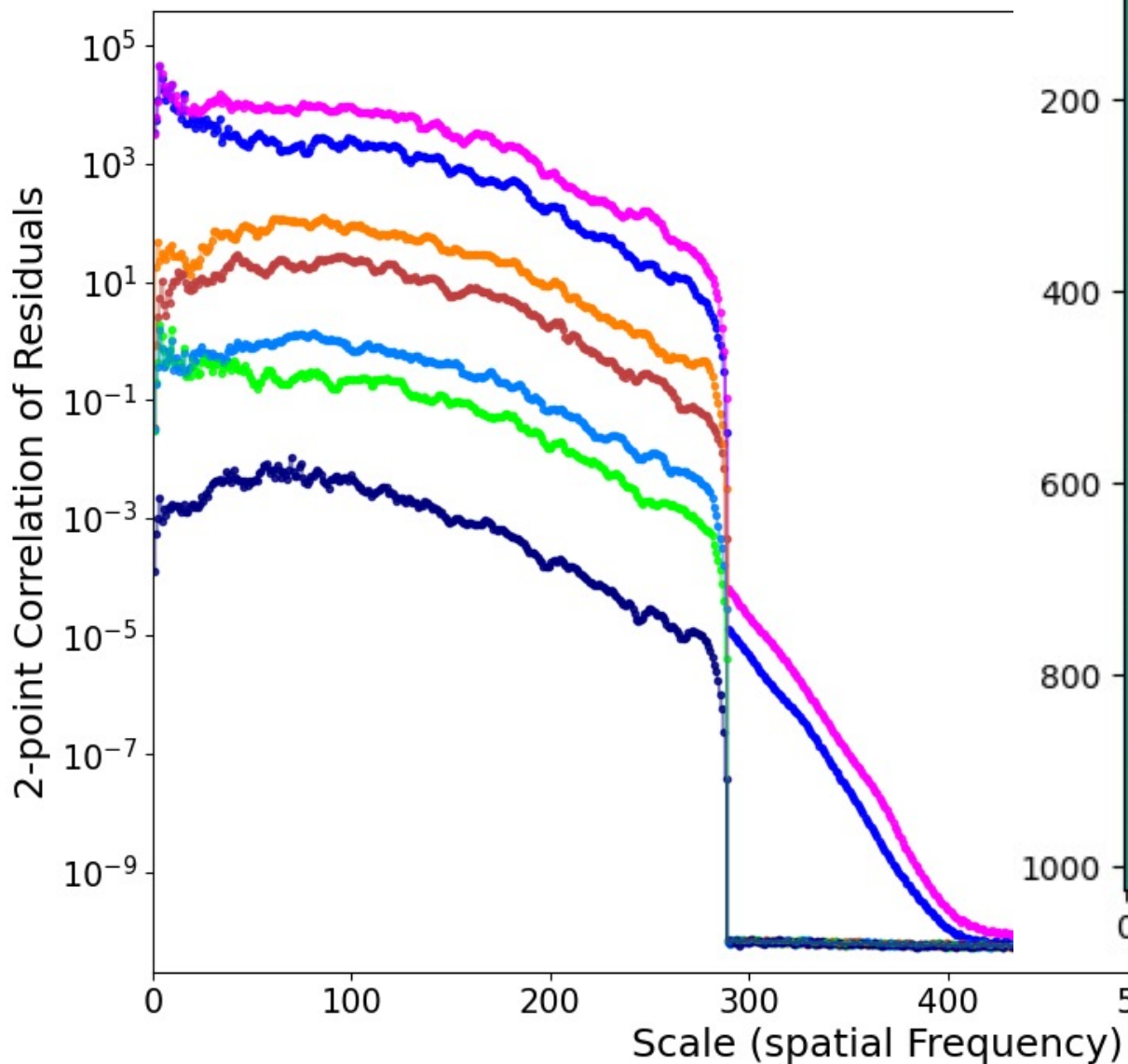
Errors above here are detectable  
(for general purposes)



# Compression

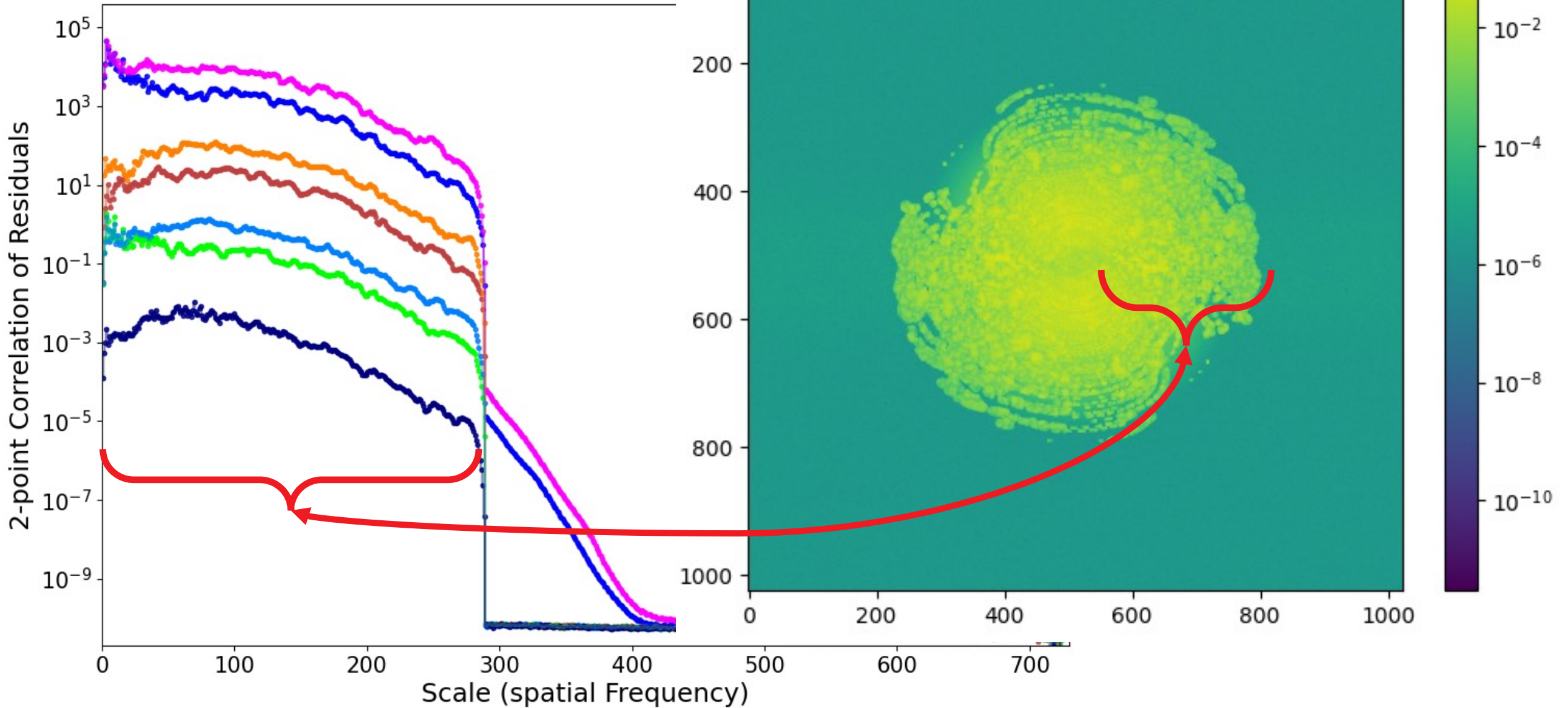


# Compression





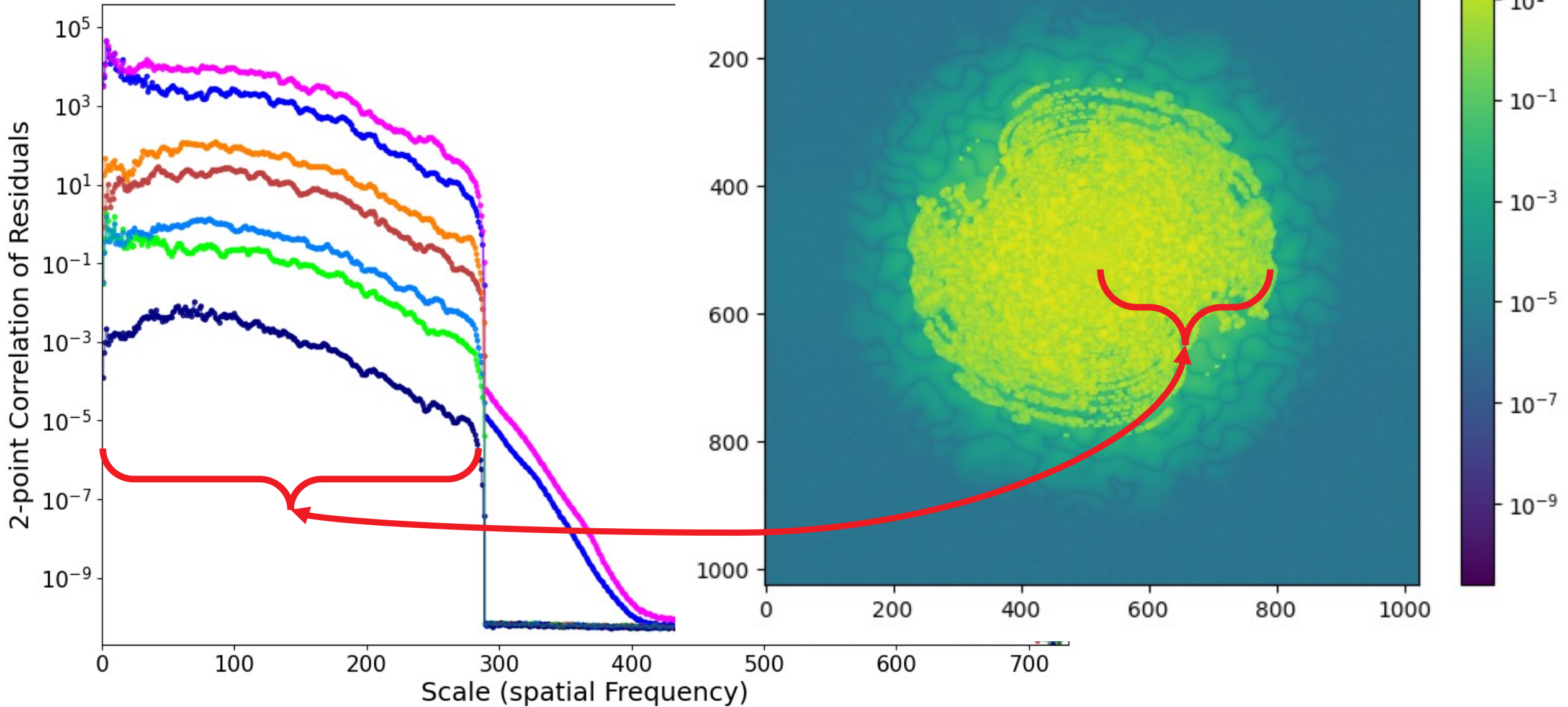
# Compression



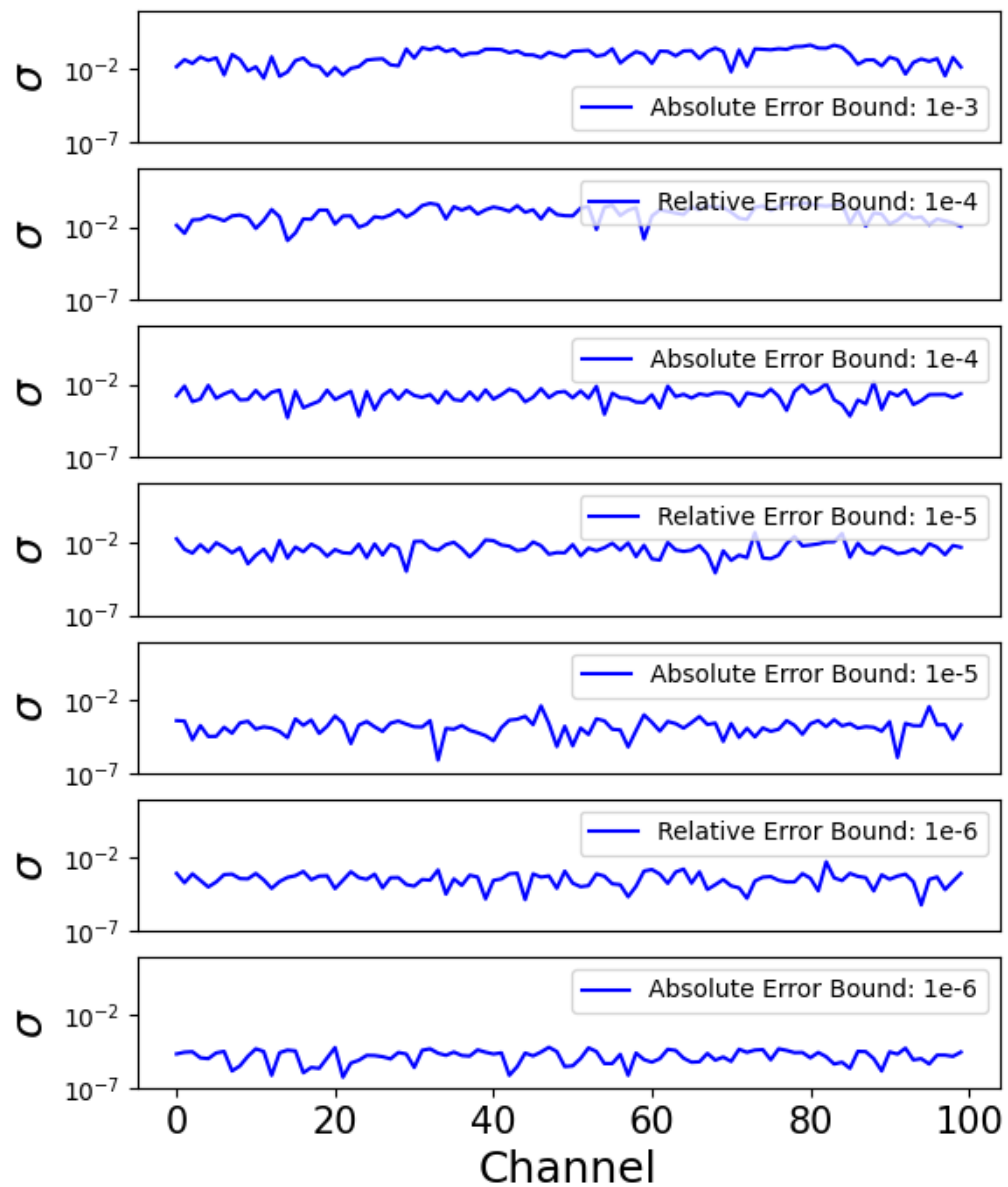
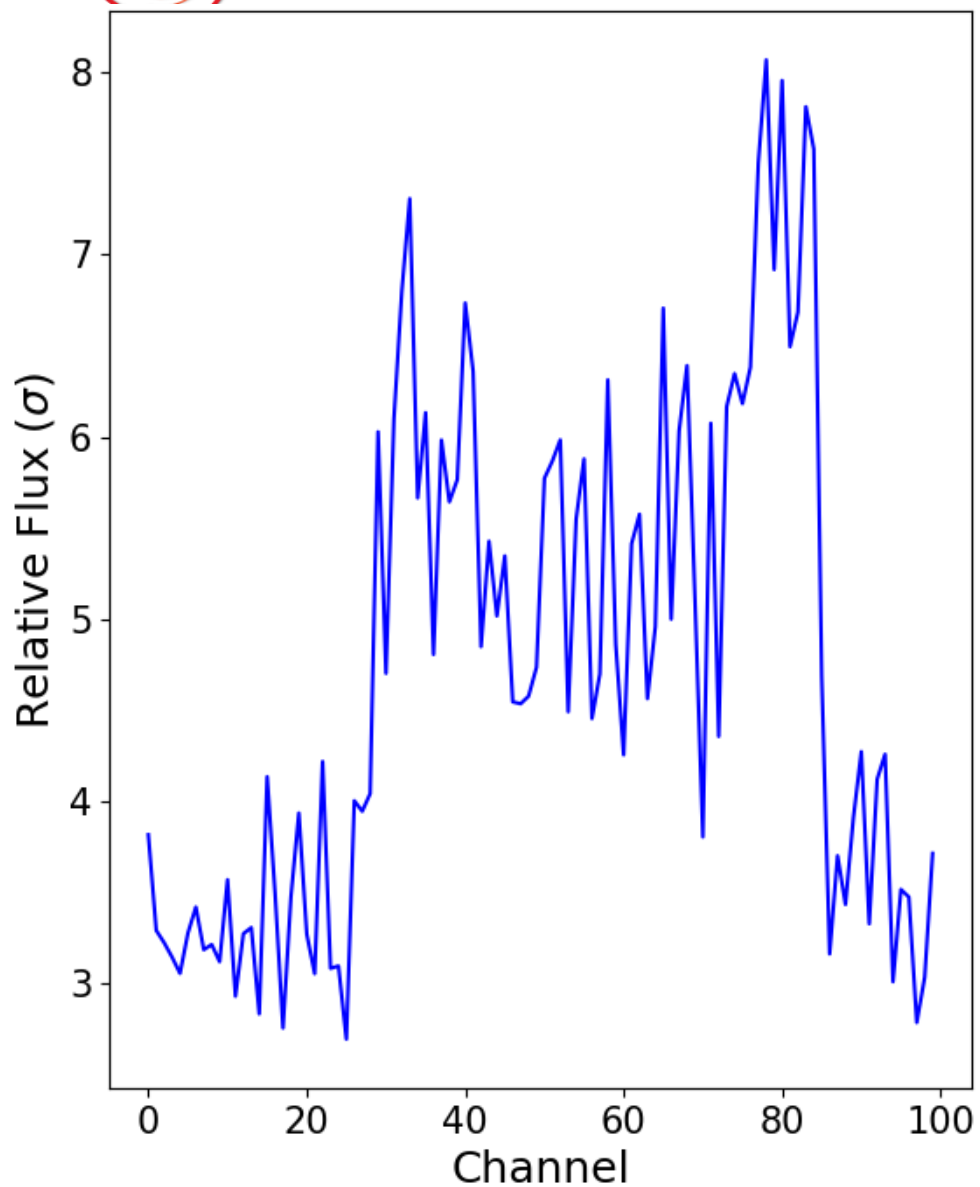




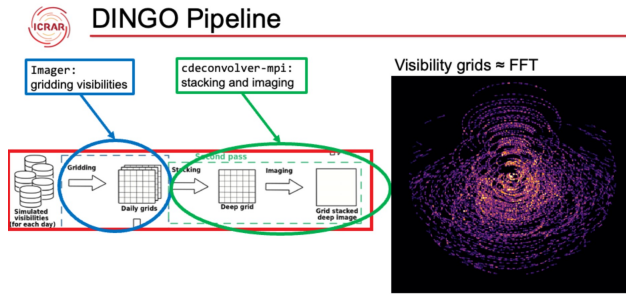
# Compression



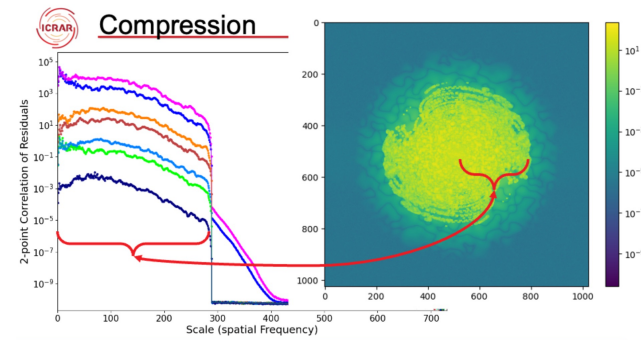
# Compression



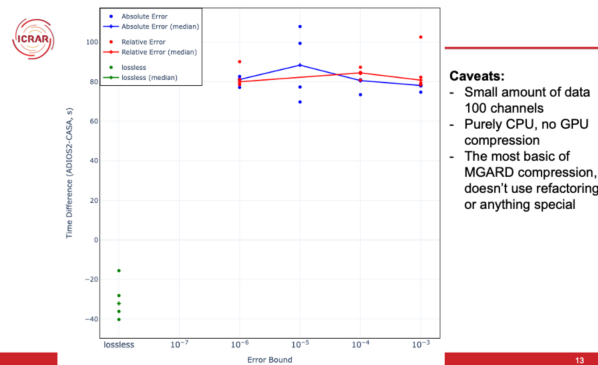
## Compressing raw visibilities



## Creating quantitative image quality metrics



## GPU-based/optimised compression





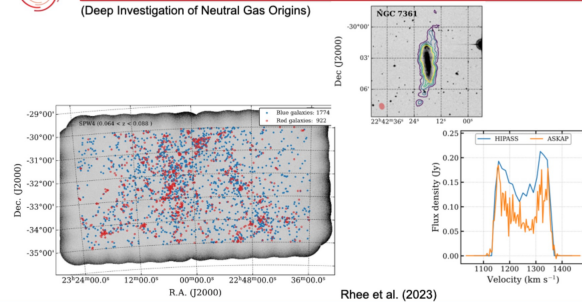
# Summary

## Radio Astronomy



### DINGO

(Deep Investigation of Neutral Gas Origins)



Rhee et al. (2023)

5

## SKA



### SKA



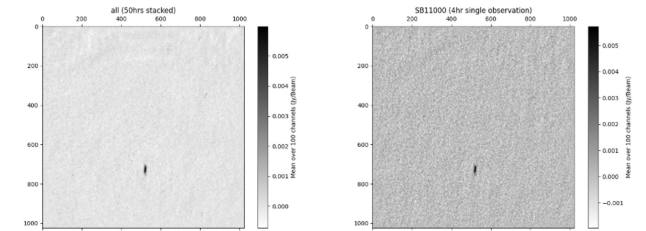
710 PB/year

2

## DINGO



### DINGO Science

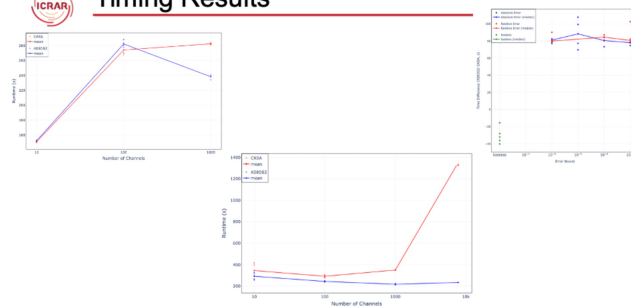


6

## Timing Results



### Timing Results

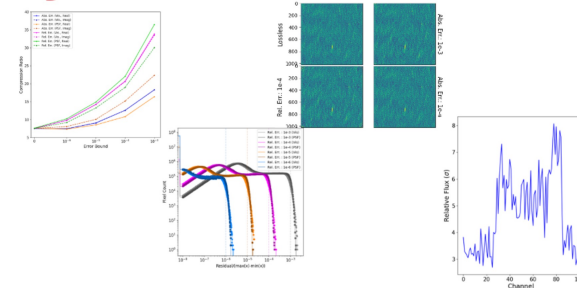


28

## Compression Results



### Compression Results

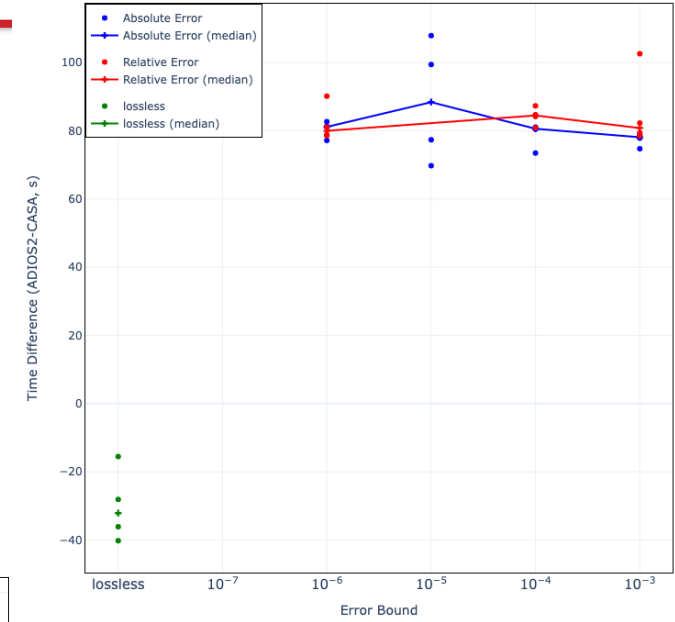
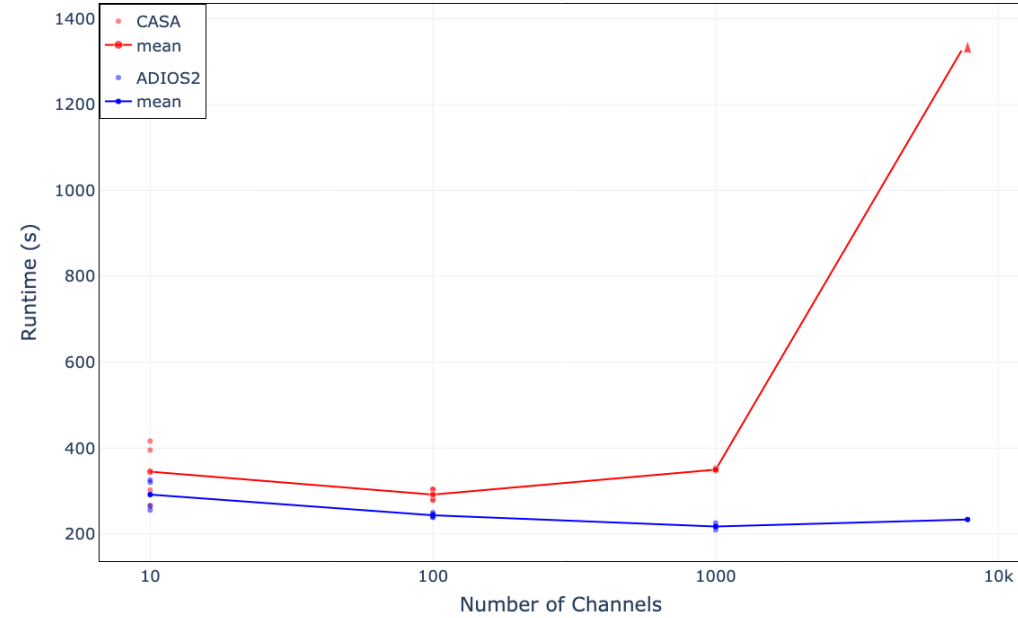
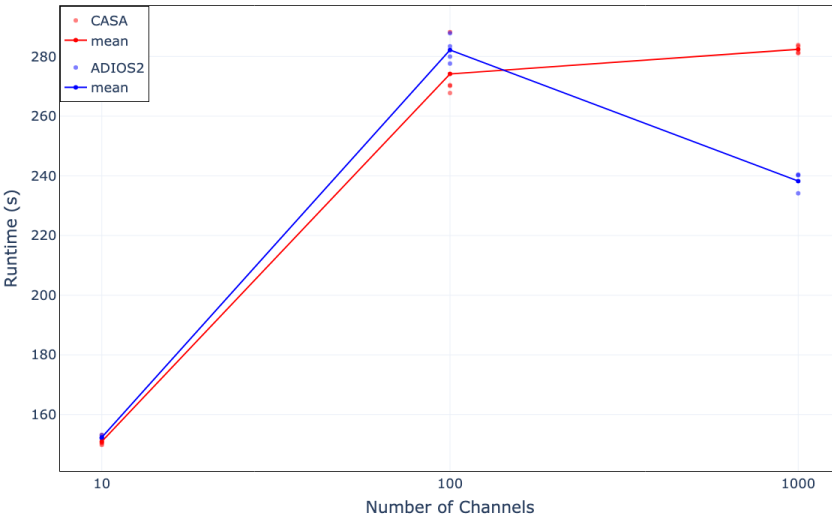


29





# Timing Results



# Compression Results

