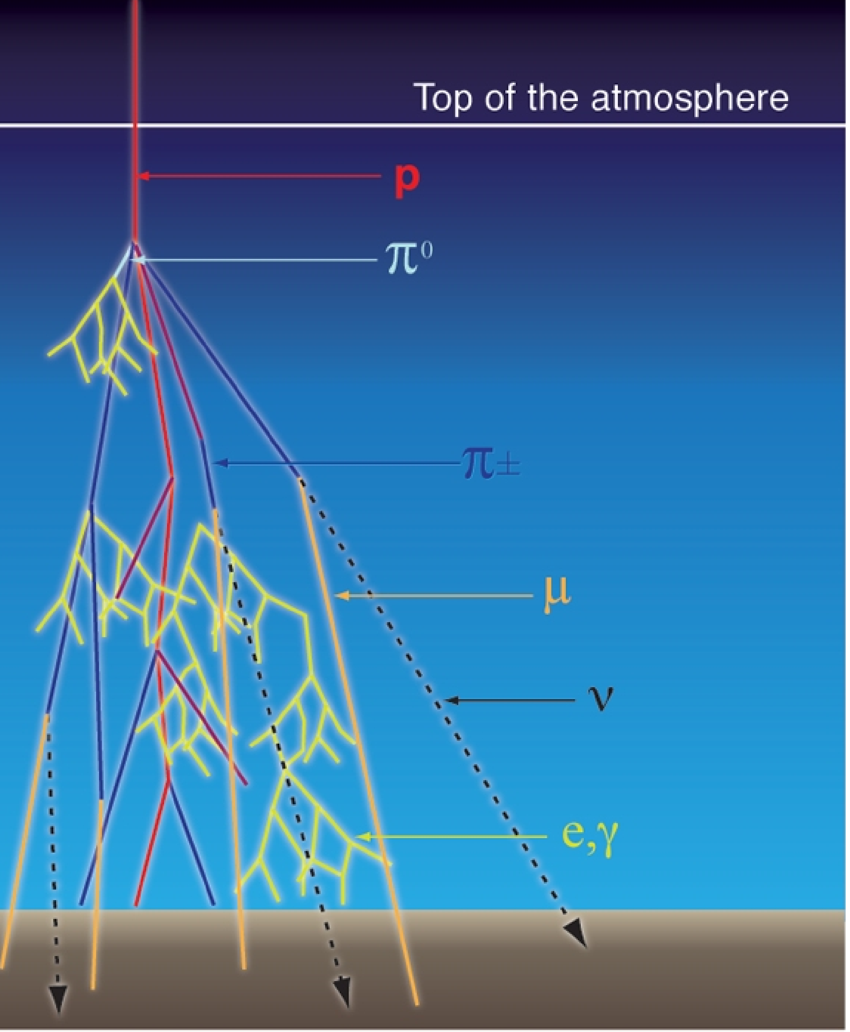
Radio measurements of cosmic rays: the road from LOFAR to SKA

Arthur Corstanje

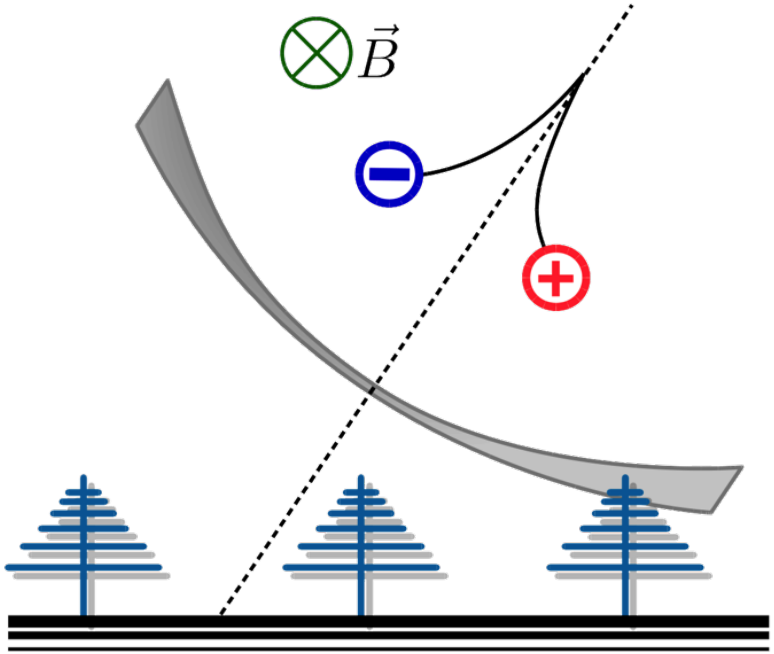
Vrije Universiteit Brussel,
Radboud University Nijmegen

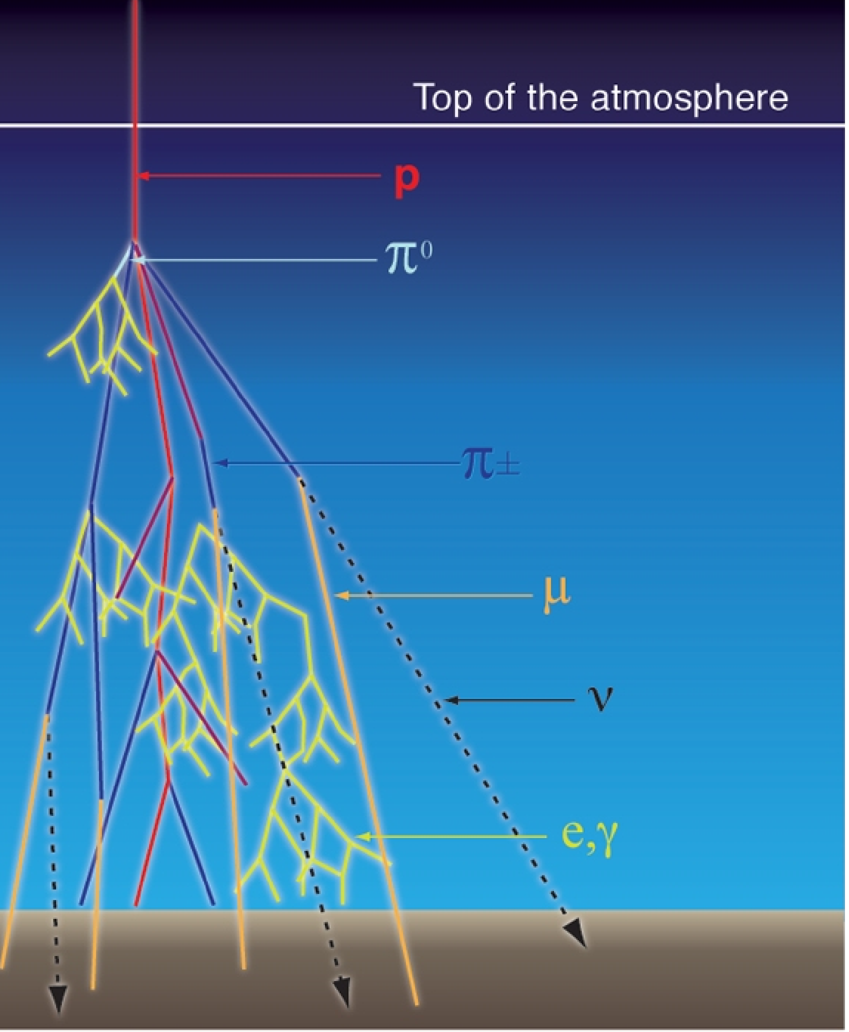
Seminar KIT,
June 4, 2025



Extensive air showers

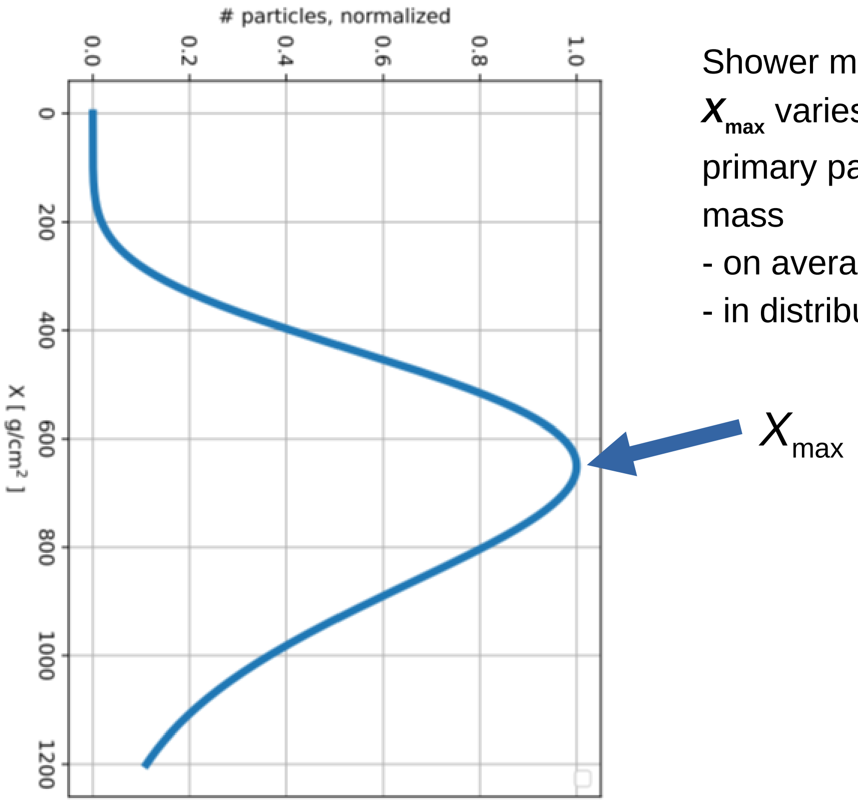
Radio signals from air showers





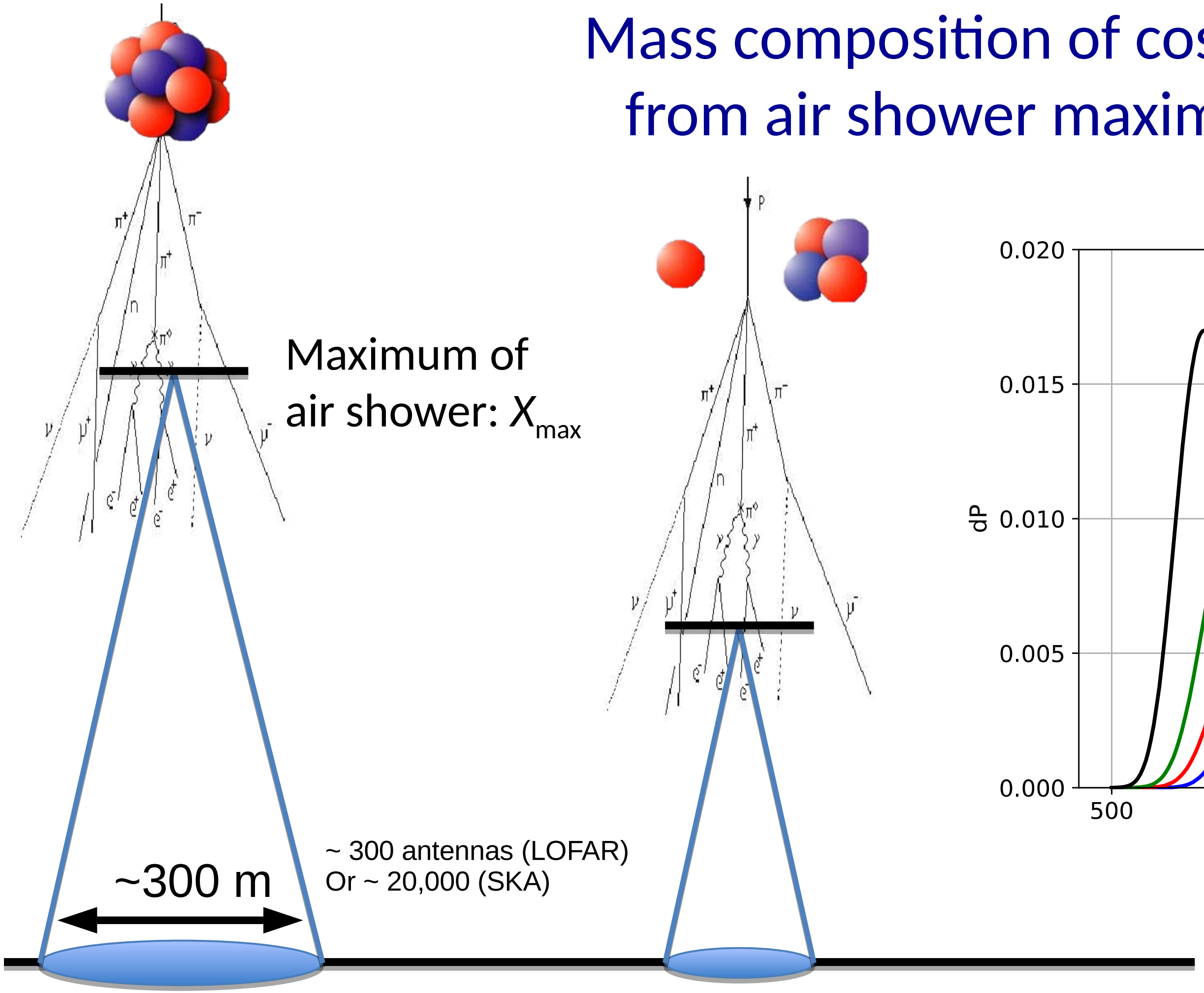
Extensive air showers

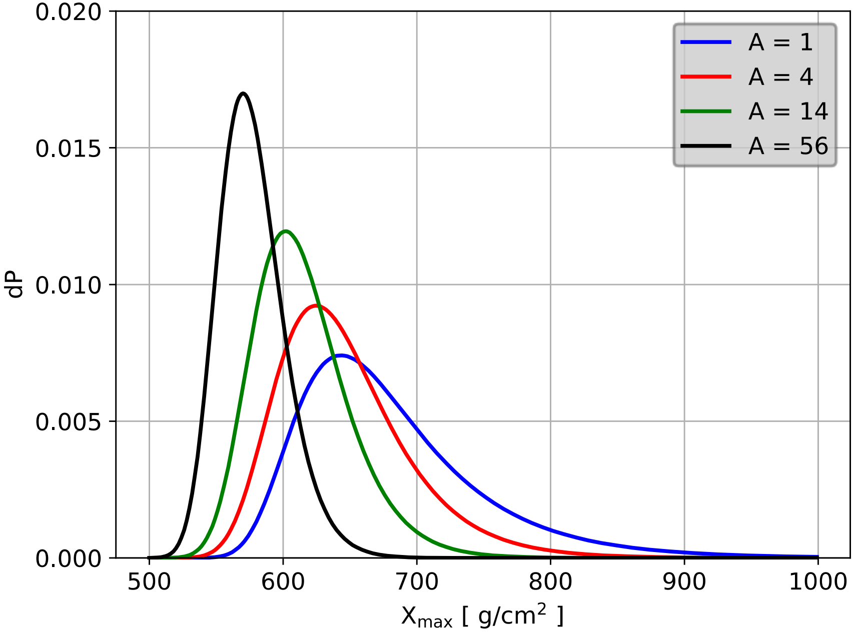
Longitudinal profile of number of particles



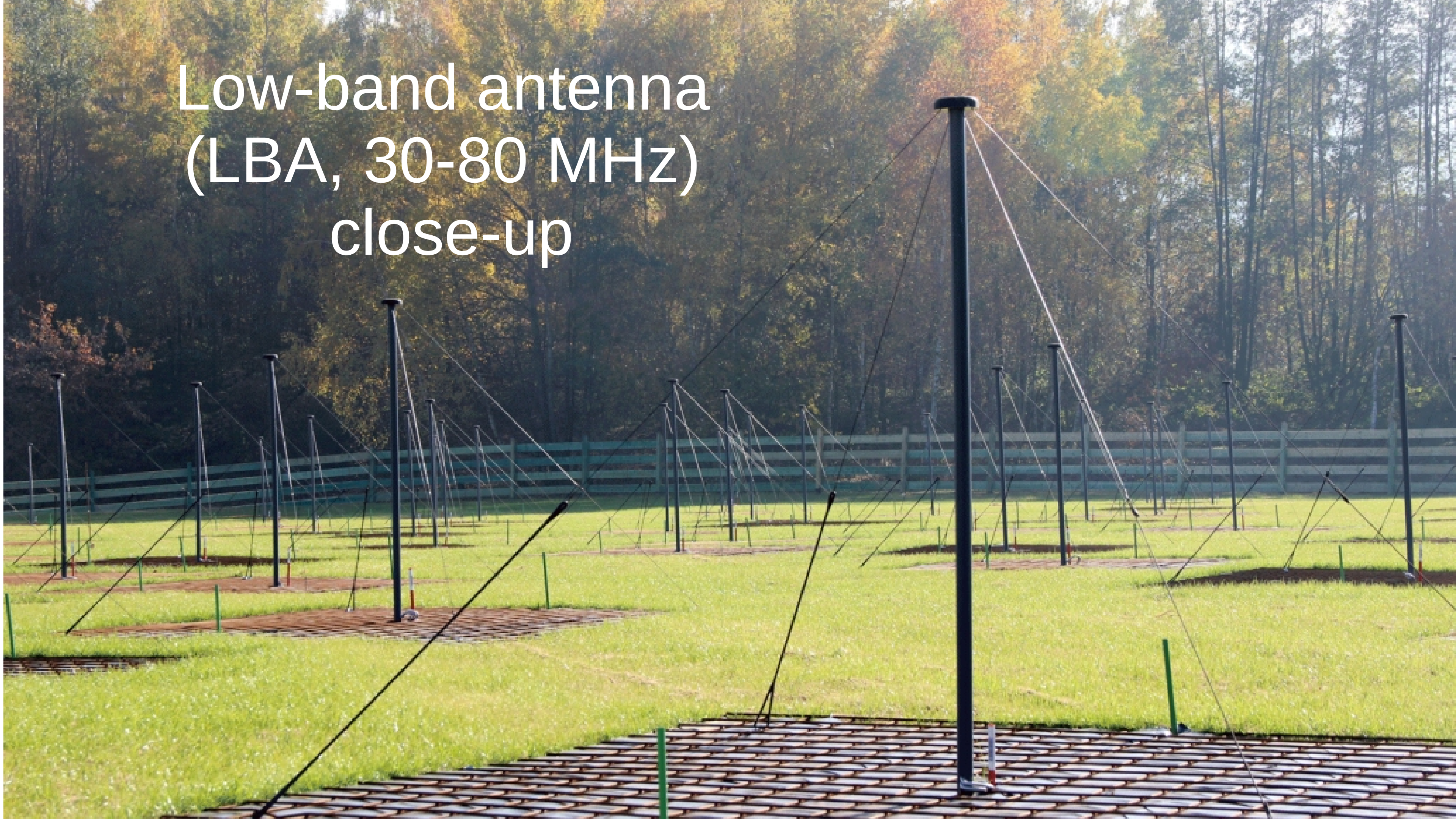
Shower maximum X_{max} varies with primary particle mass
- on average
- in distribution

Mass composition of cosmic rays, from air shower maximum X_{max}



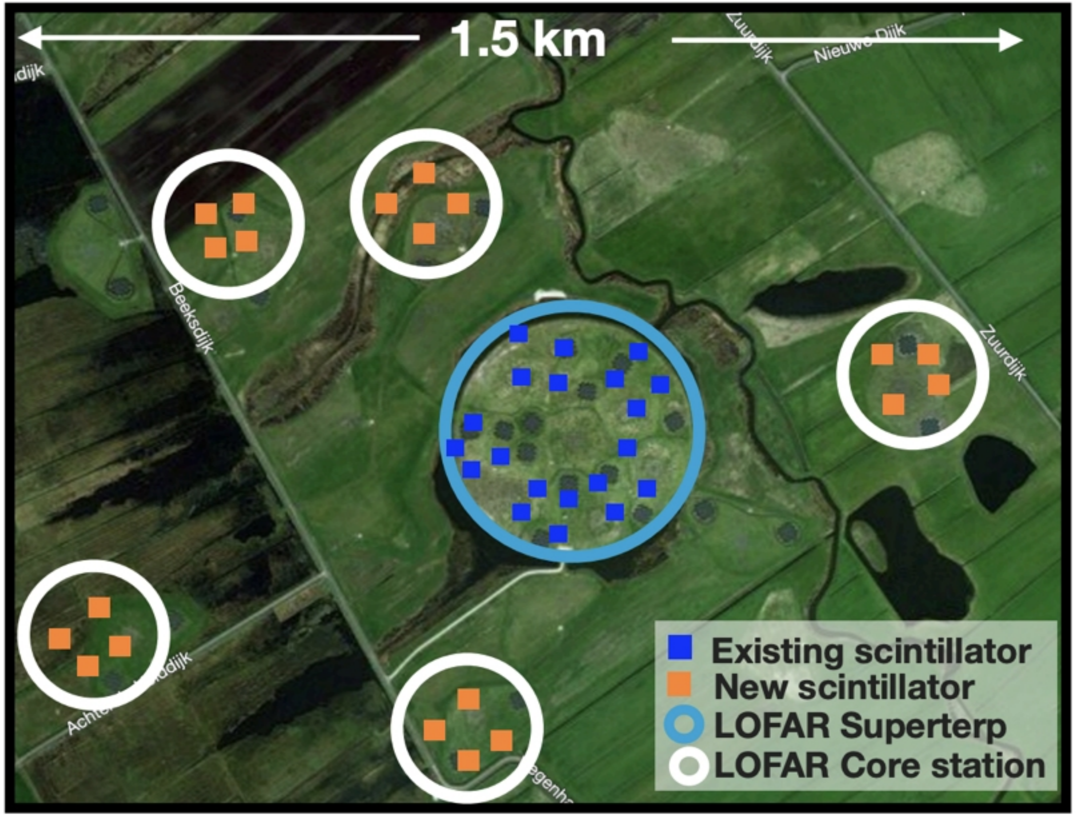



The LOFAR Superterp


Low-band antenna
(LBA, 30-80 MHz)
close-up

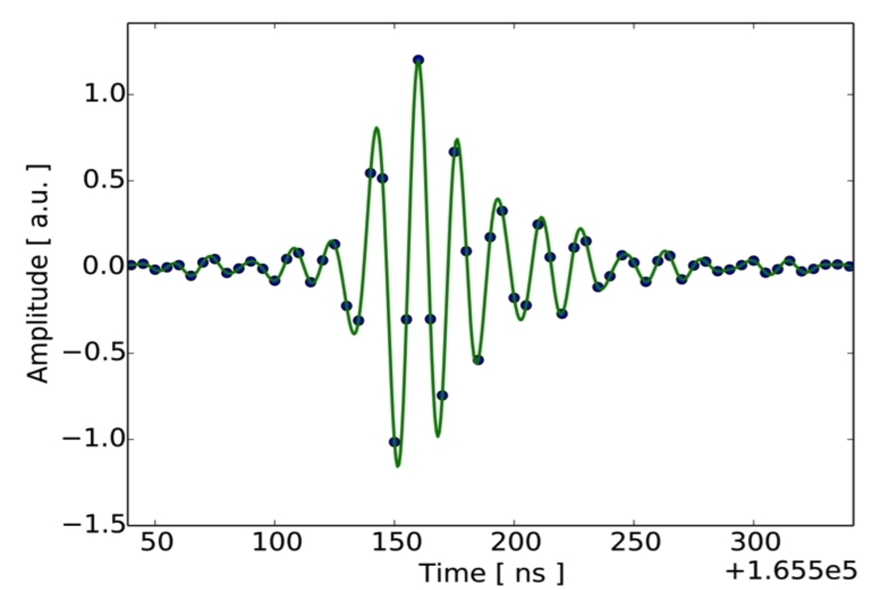
Trigger using particle detector array (LORA)

Figs by K. Mulrey

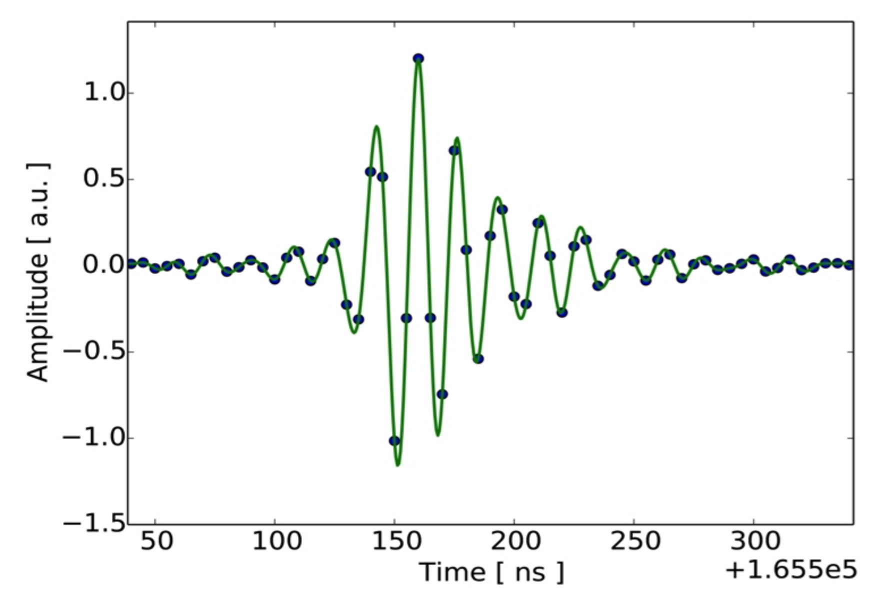


Used to trigger a readout of the transient buffers of each antenna

Pulses in LOFAR antennas



Pulses in LOFAR antennas



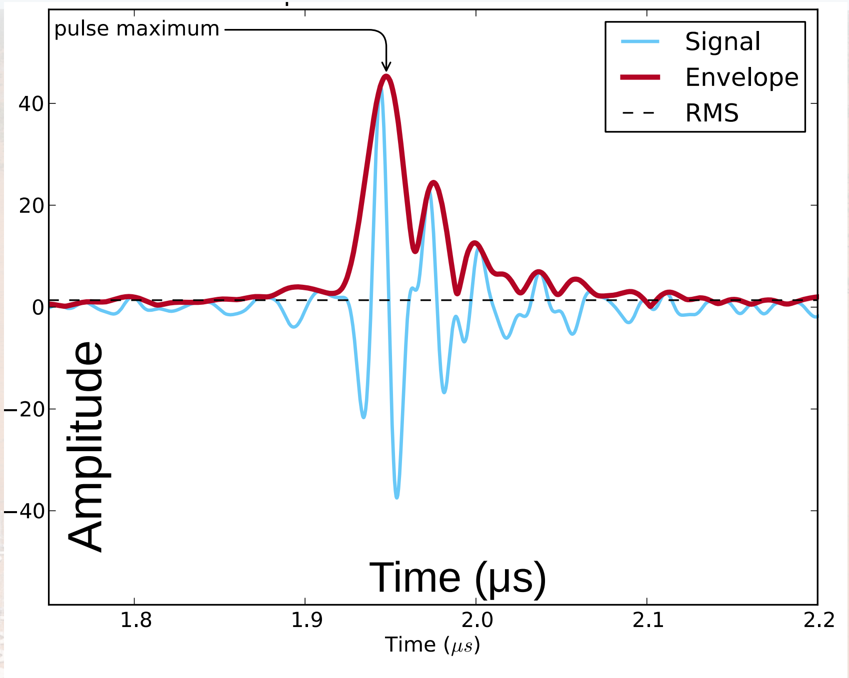
Our raw material:

One pulse in every dipole
(unprocessed time series)

- Amplitude & integrated power
- Arrival time
- Polarization
- Shape / spectrum

Pulse analysis steps

- Measure arrival time using Hilbert envelope
  - Sub-sample accuracy, < 1 ns LOFAR
  - Used for **arrival direction**
  - Contains info on shower development



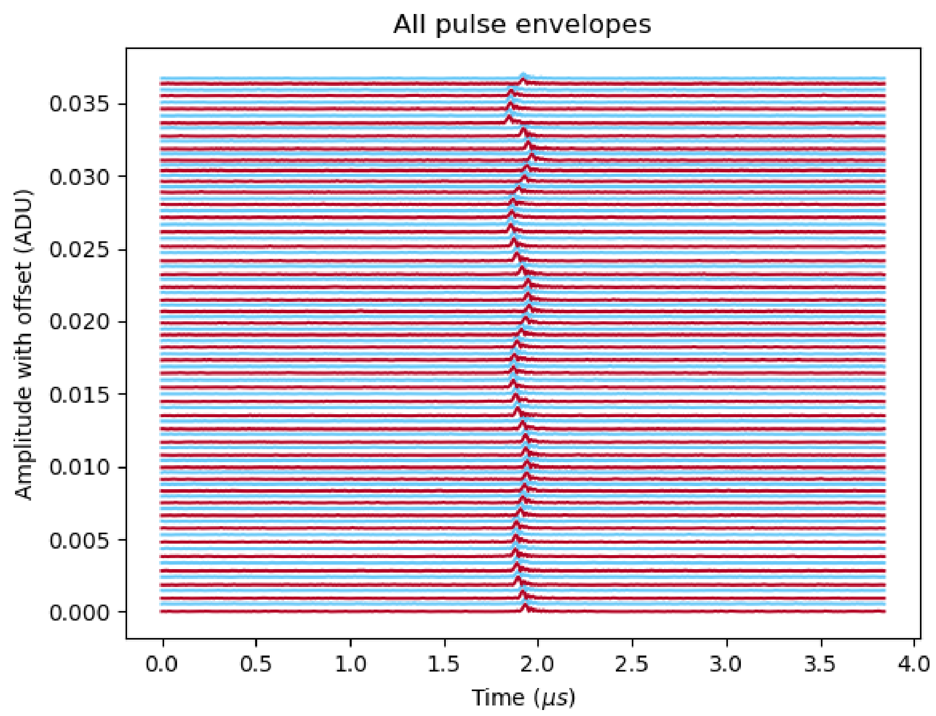
- Measure integrated power (energy fluence)
  - To construct **radio footprint** and compare to simulations

- Polarization signature
- Pulse shape

Pulse analysis steps: arrival time

- Measure arrival time using Hilbert envelope
  - Pulse detections in one LOFAR station

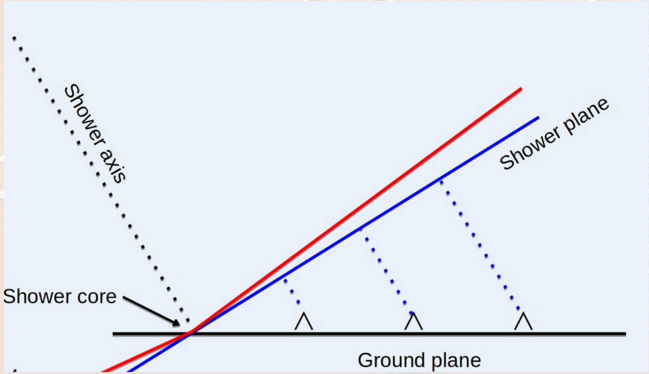
- **Requires**: up-to-date timing calibration per antenna
  - From observatory, ideally
  - From phases of narrow-band RFI (A&A 2016, arXiv:1603.08354)
  - From drone-based calibration signal

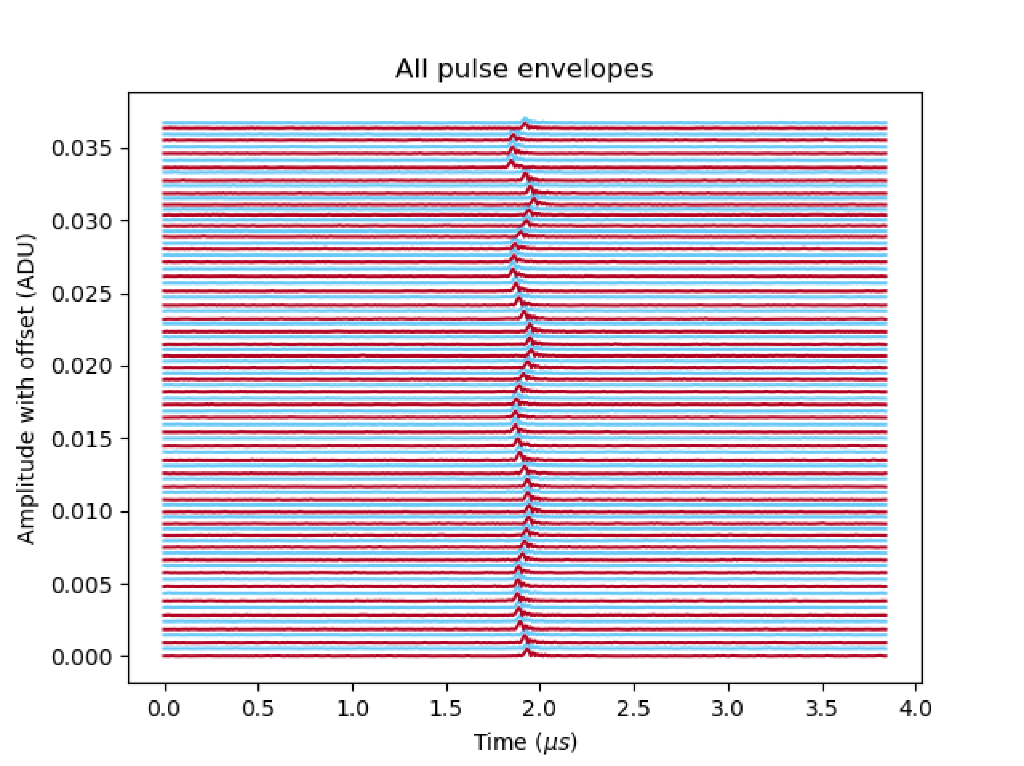


Pulse analysis steps: beamforming (far field)

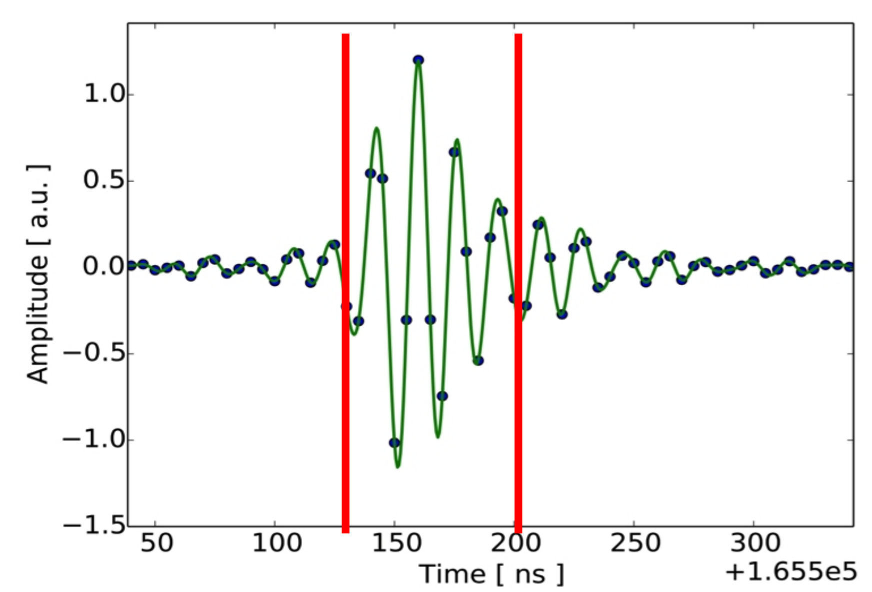
- Beamforming by time-shifting each trace according to geometric delay

- Far-field approximation works well for small stations (~ 30 m)
  - Used to identify pulse location in trace
  - Generally speaking, more complex wavefront shape: **hyperboloid**





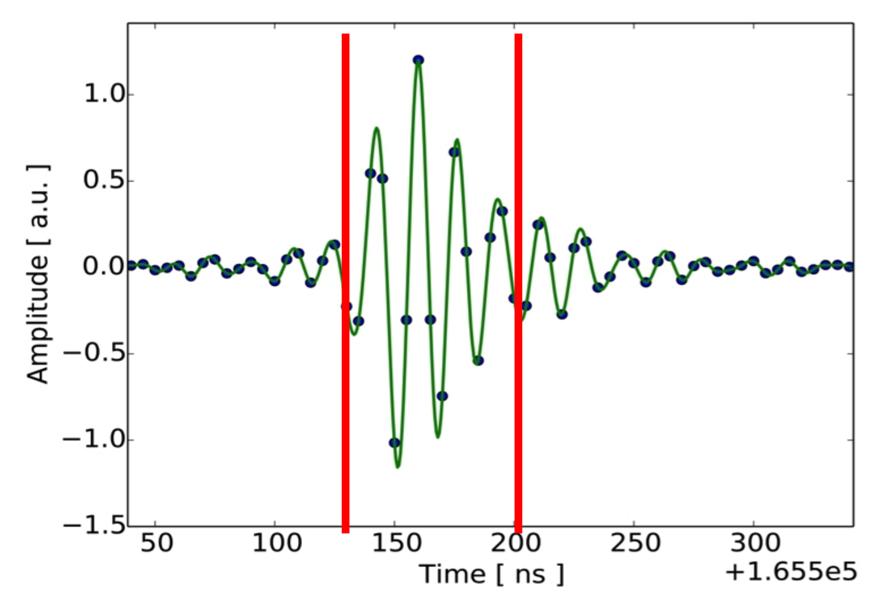
Measuring energy fluence



Integrate ‘power’ in a time window

- Width of time window not too large (noise)

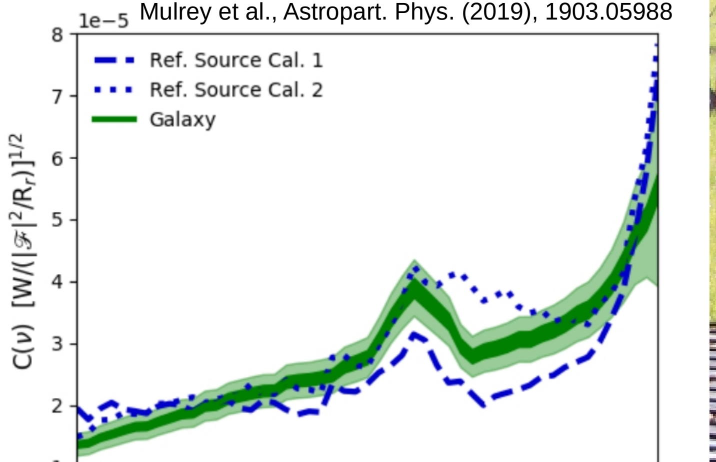
Measuring energy fluence

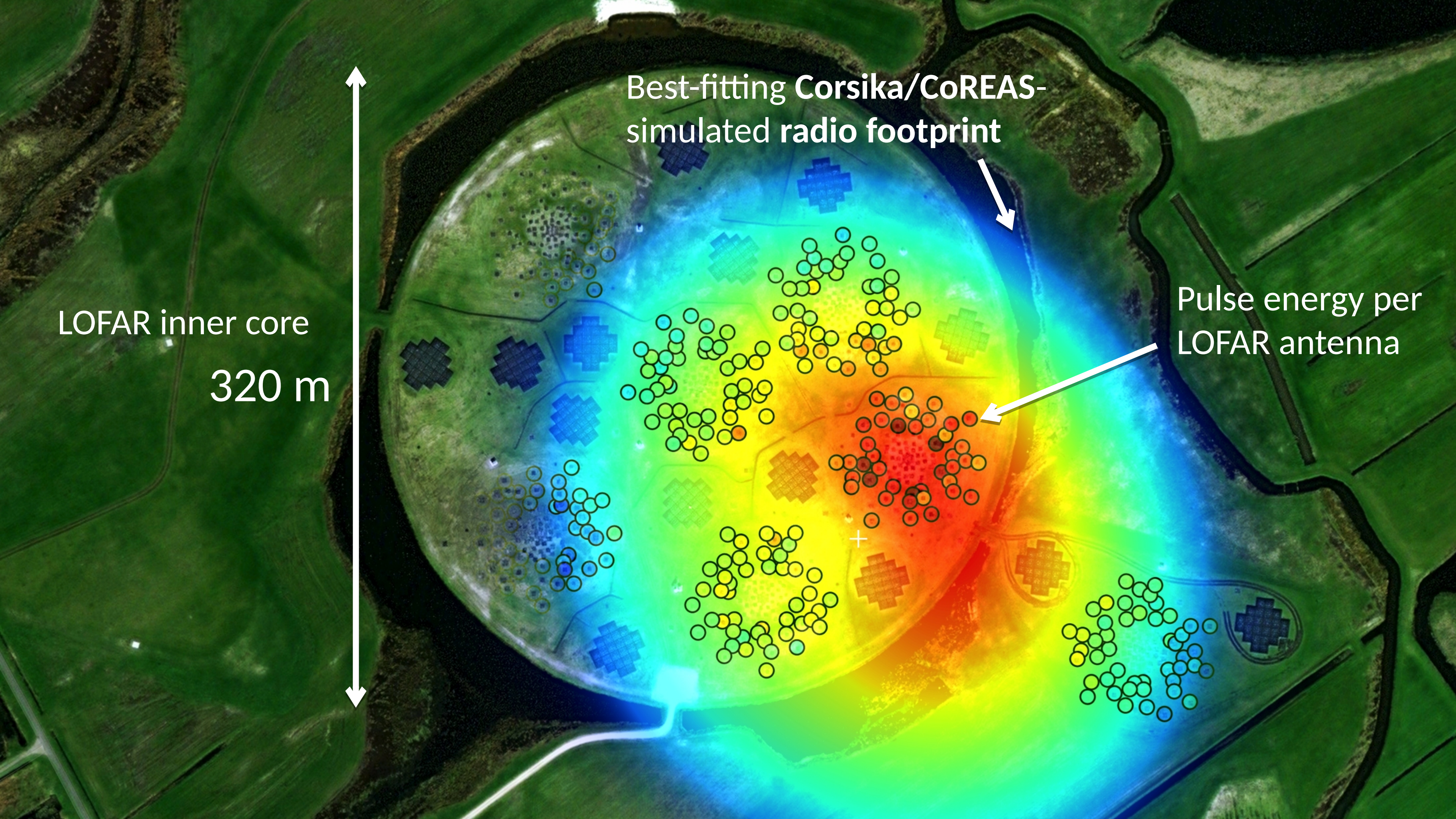


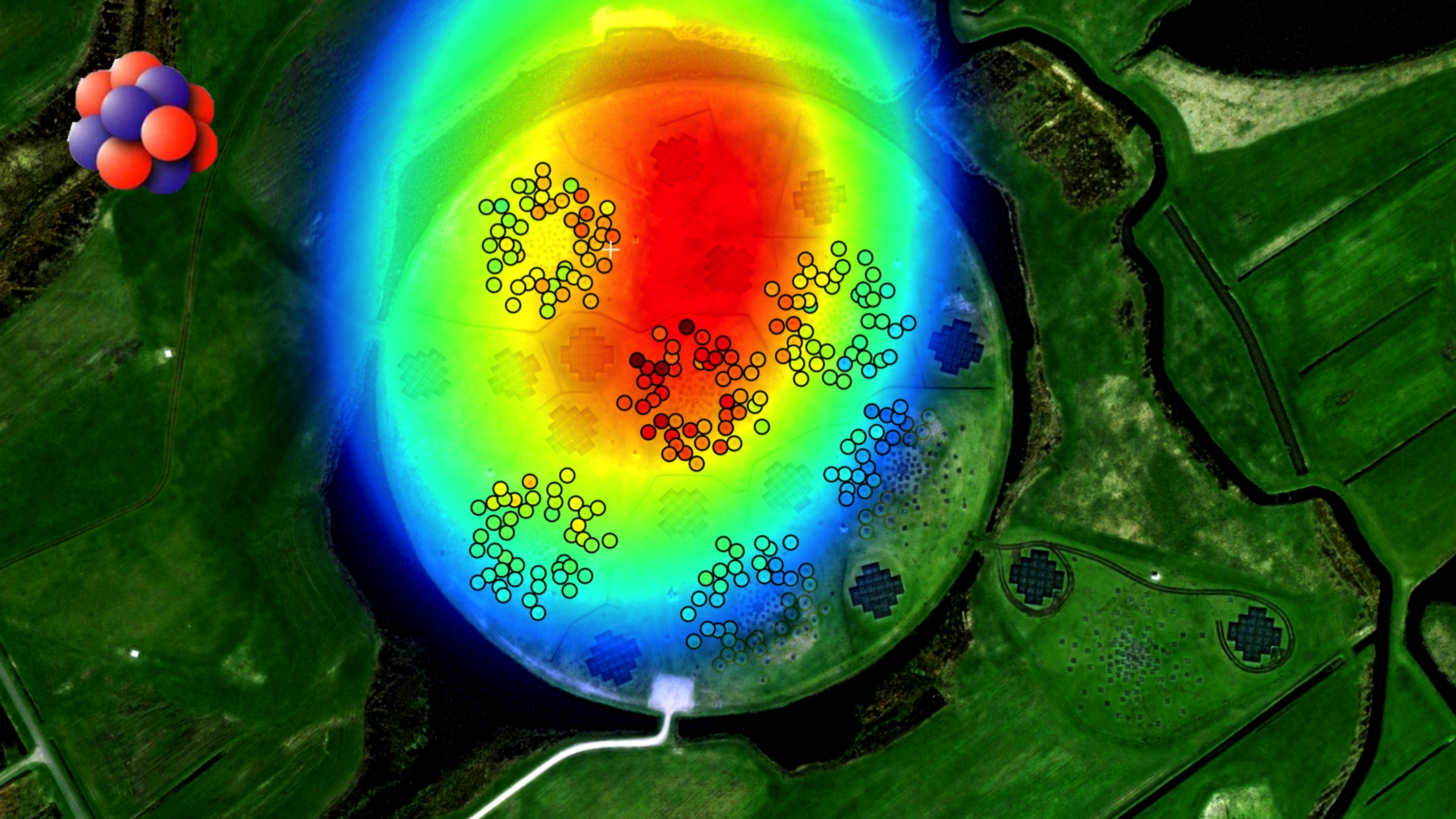
Integrate 'power' in a time window

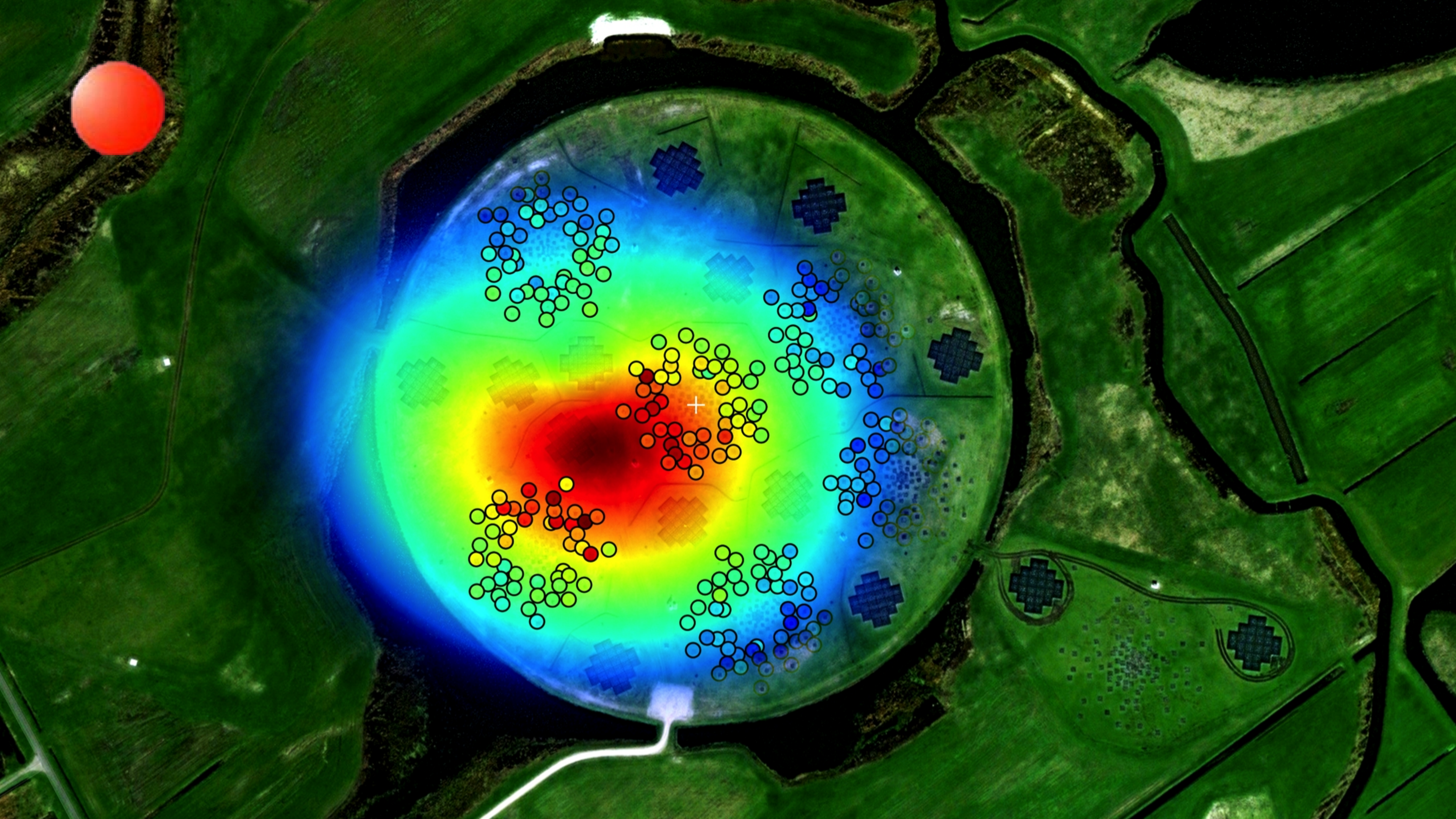
- **Requires**: absolute gain calibration per antenna

Mulrey et al., Astropart. Phys. (2019), 1903.05988




Best-fitting **Corsika/CoREAS**-
simulated **radio footprint**
LOFAR inner core
320 m
Pulse energy per
LOFAR antenna

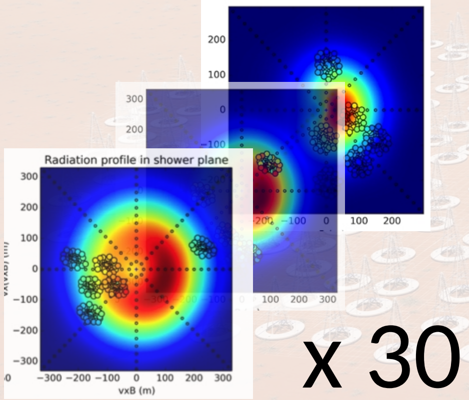


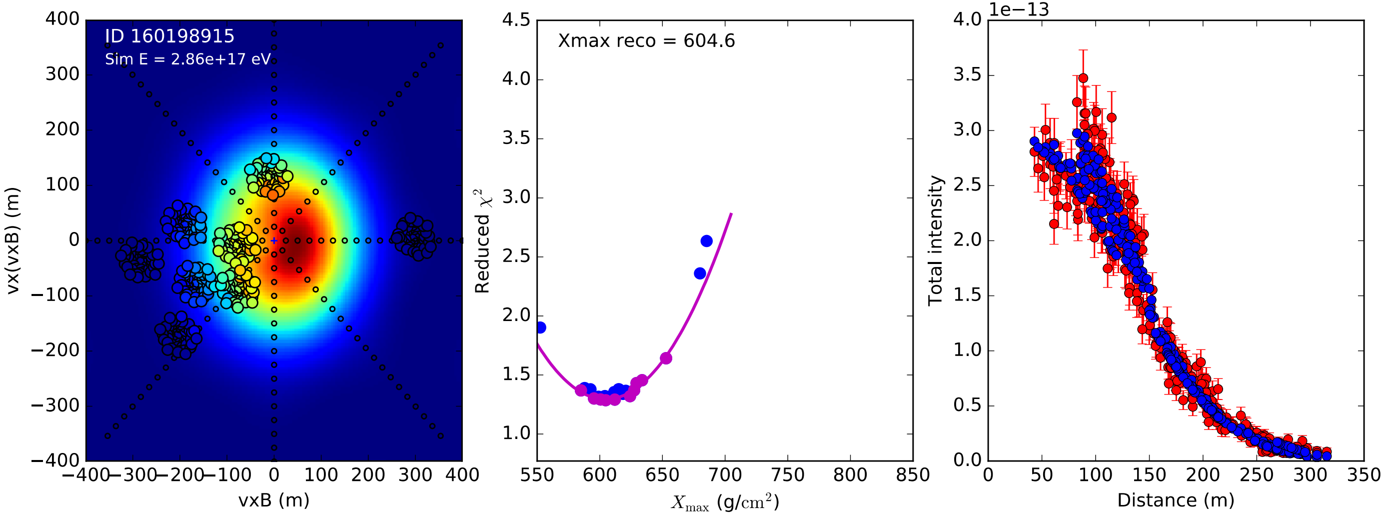


Matching simulated footprints to data

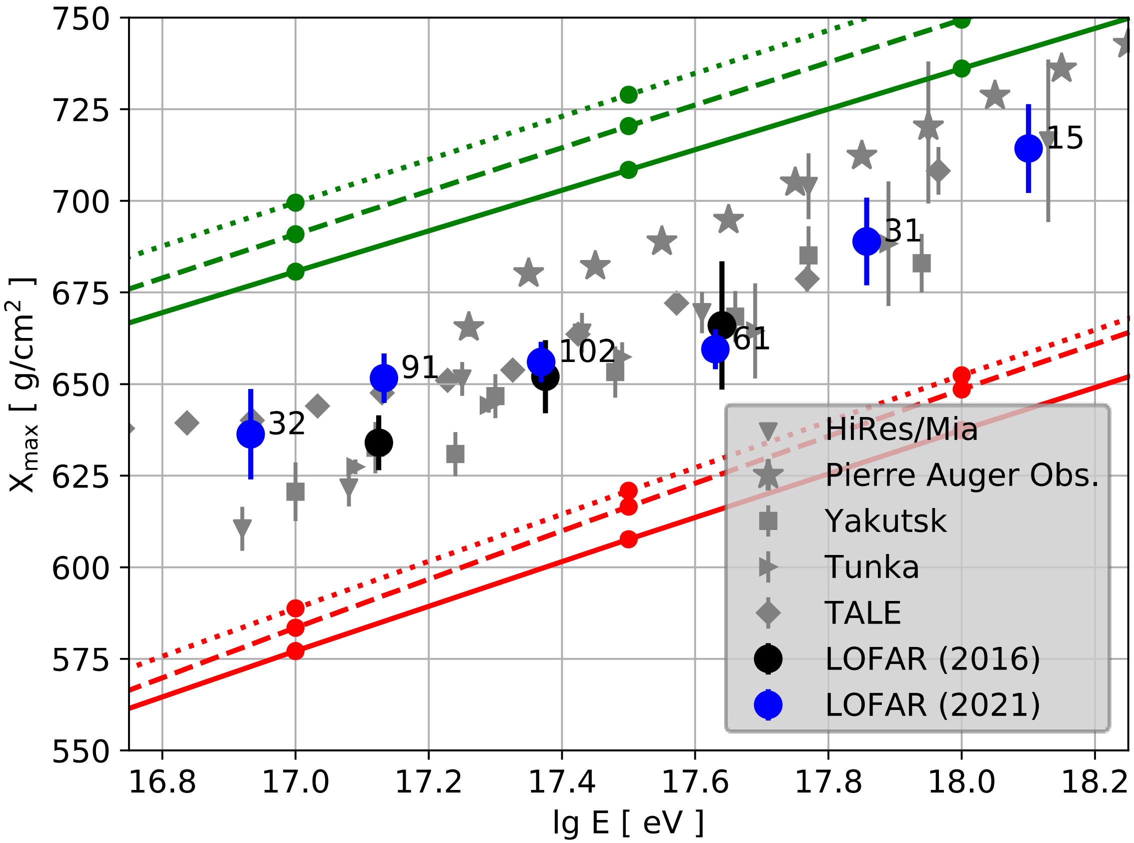
- Simulate about 30 showers per measured shower
- Fit them to data, observe X_{max} of best fit

- Resolution (@LOFAR) about 20 g/cm^2
- Systematic uncertainties < 9 g/cm^2
- In line with state of the art in the field





Result: Average X_{max} versus primary energy

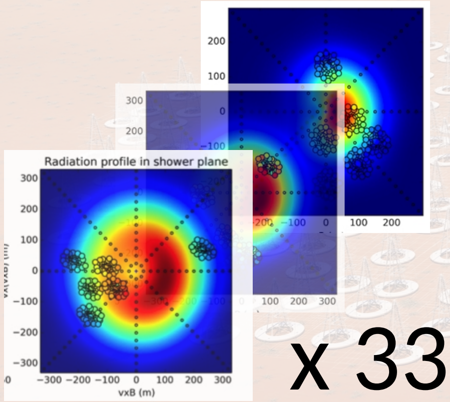


- Green lines: average X_{max} for pure proton composition
- Red lines: average X_{max} for pure iron composition

Corstanje et al., Phys Rev D 103, 102006 (2021)
arXiv: **2103.12549**

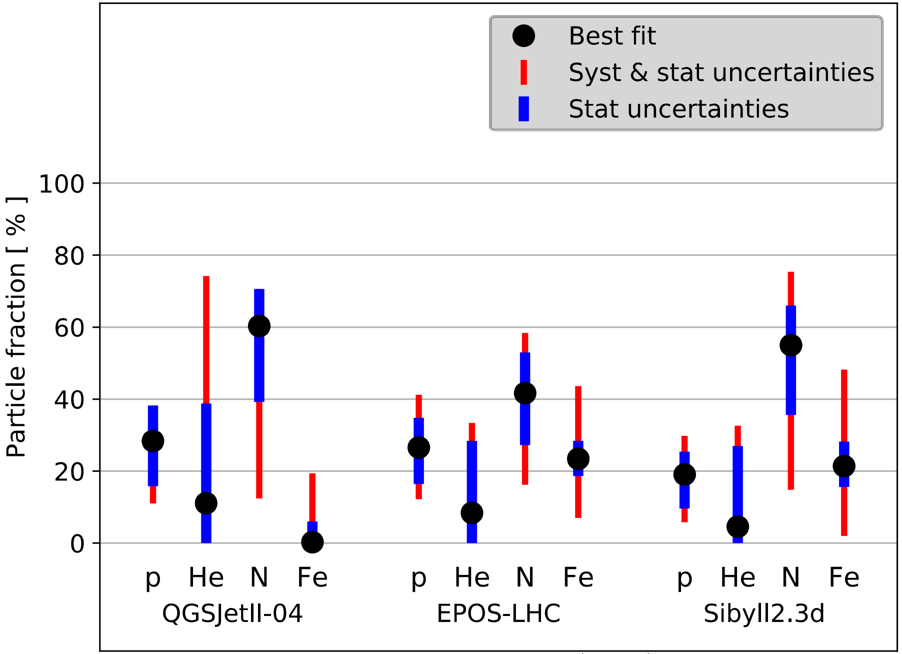
Results on mass composition

- Light-mass component (p+He) of 23 to 39% at best fit
- Still considerable (correlated) uncertainties, some inevitable
  - overlap of X_{max} distributions
  - Hadronic interaction models



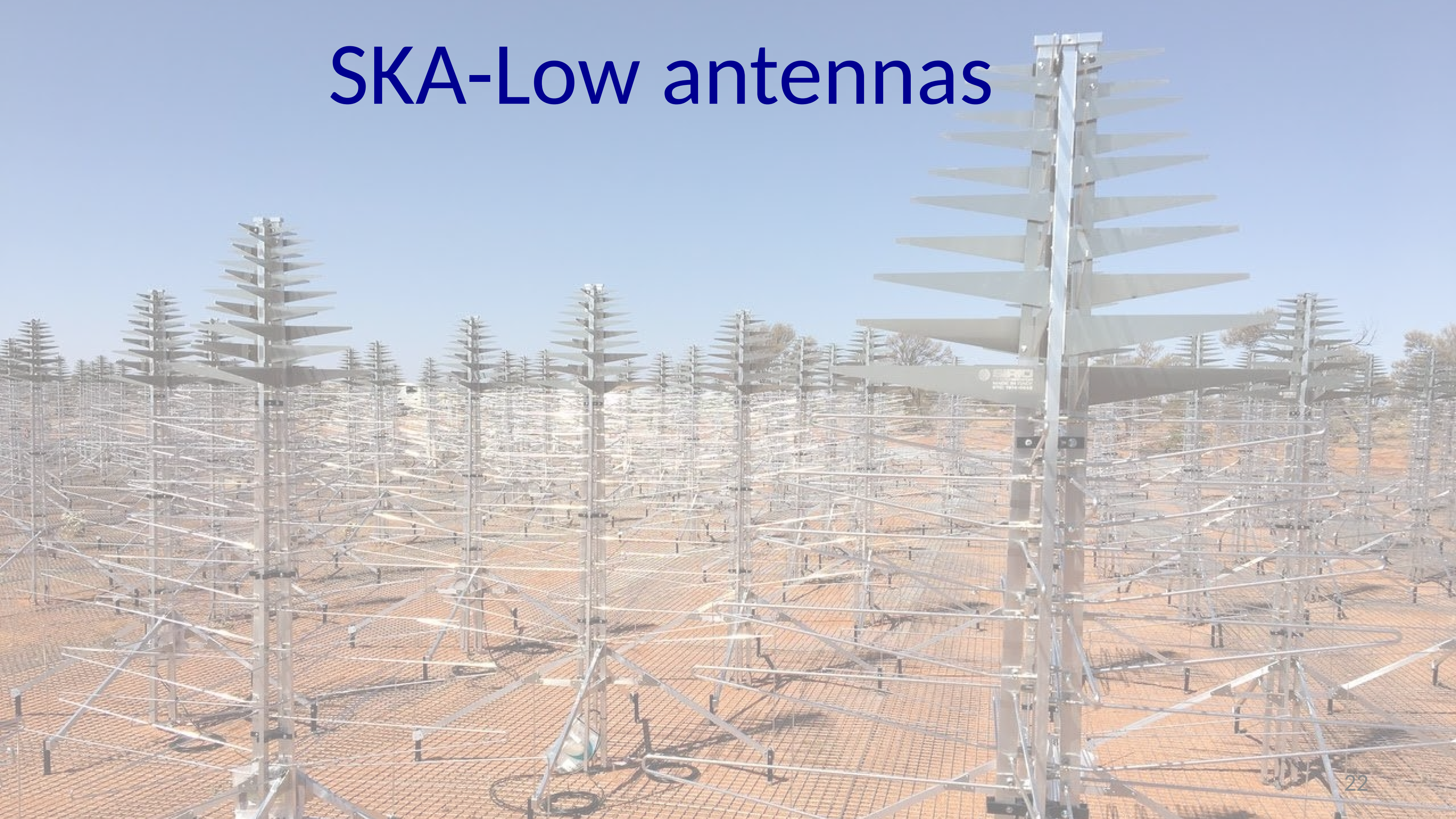
x 334

Main coverage in lg E: **17.39 +/- 0.32**



Corstanje et al., Phys Rev D 103, 102006 (2021), arXiv: **2103.12549**

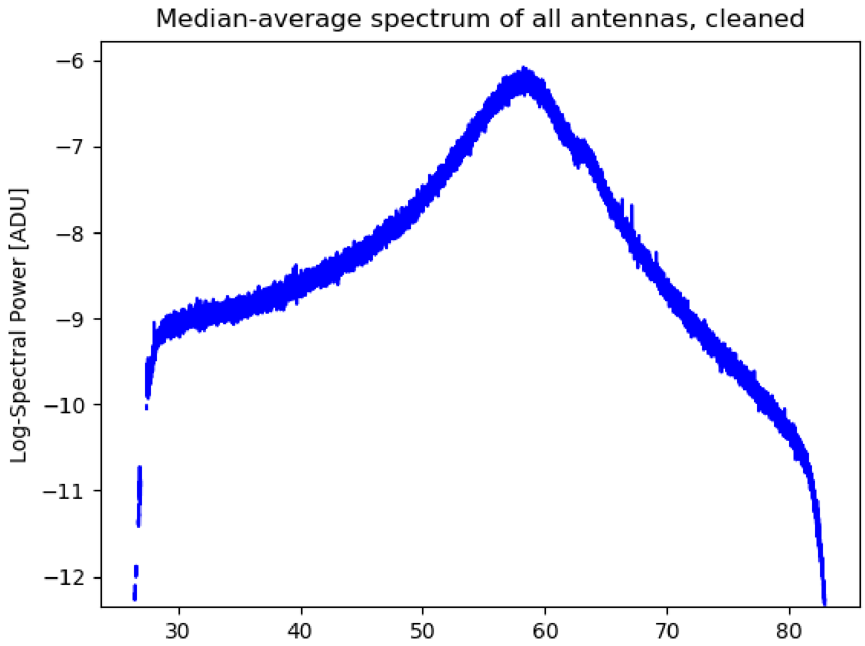
SKA-Low antennas

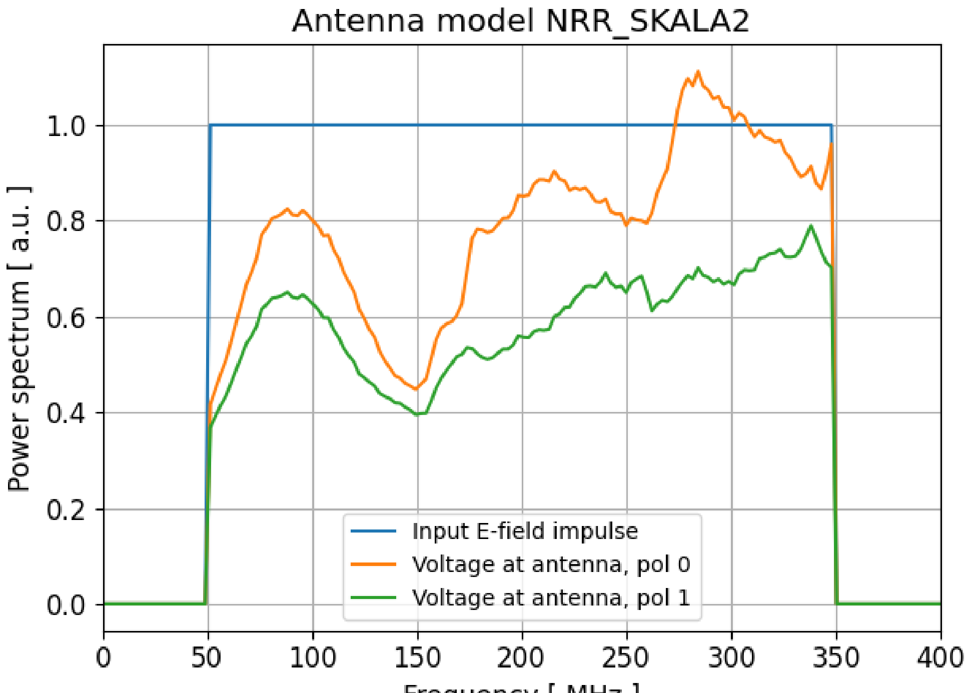


SKA-Low antennas

LOFAR: resonance ~ **58 MHz**
(note log scale!)

SKA: much flatter response
over **300 MHz bandwidth** (lin scale)

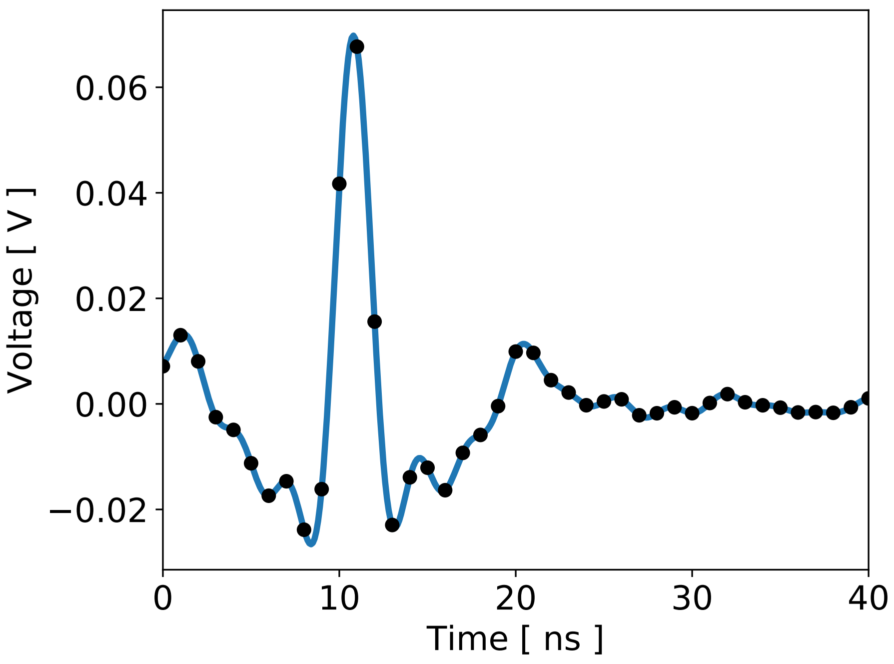




Pulses in SKA-Low antennas

Use raw voltage signals per antenna

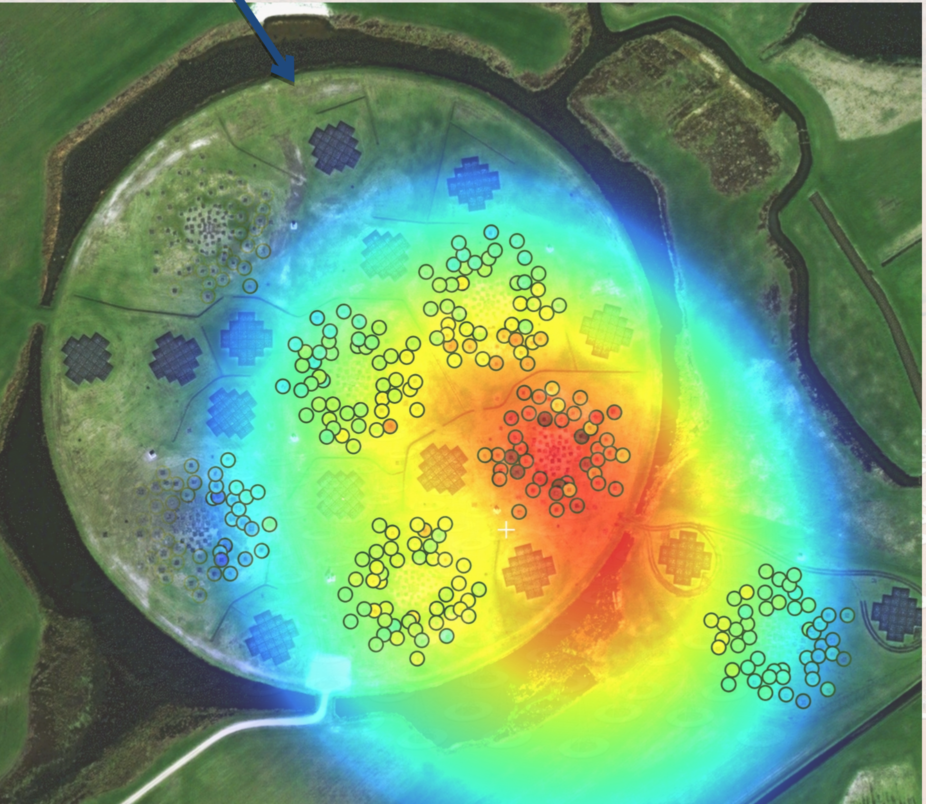
After offline filtering stage (e.g. de-dispersion)

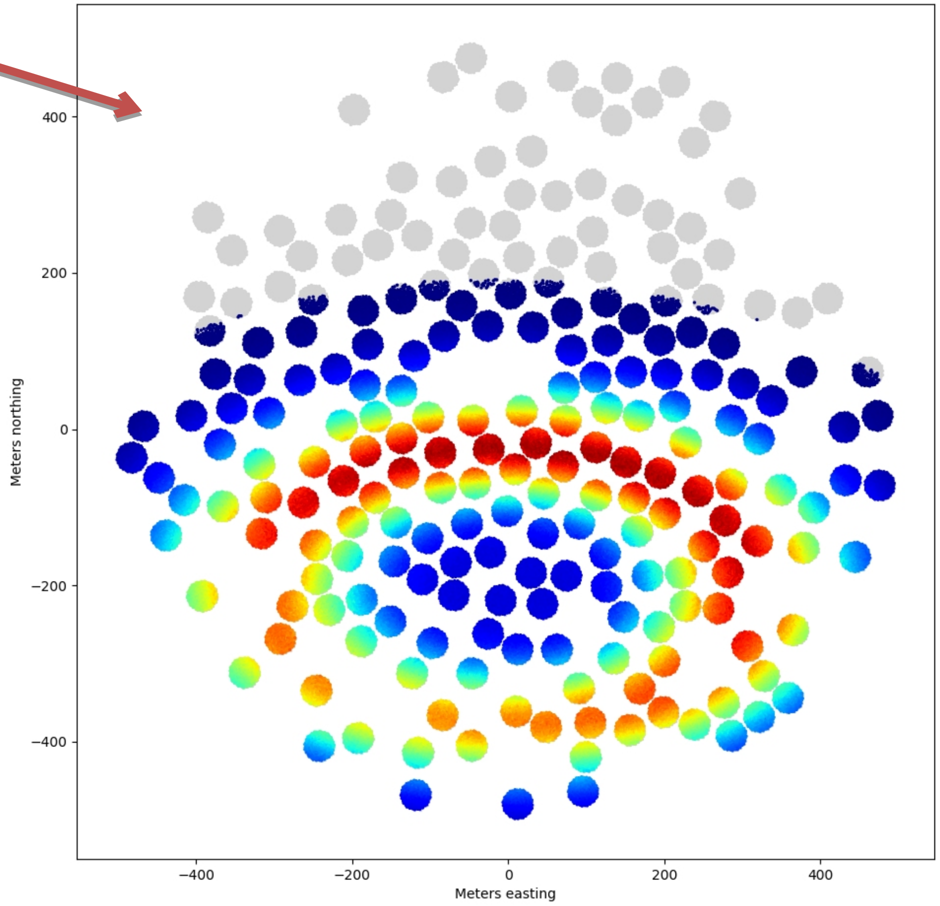


SKA-Low, a really dense array

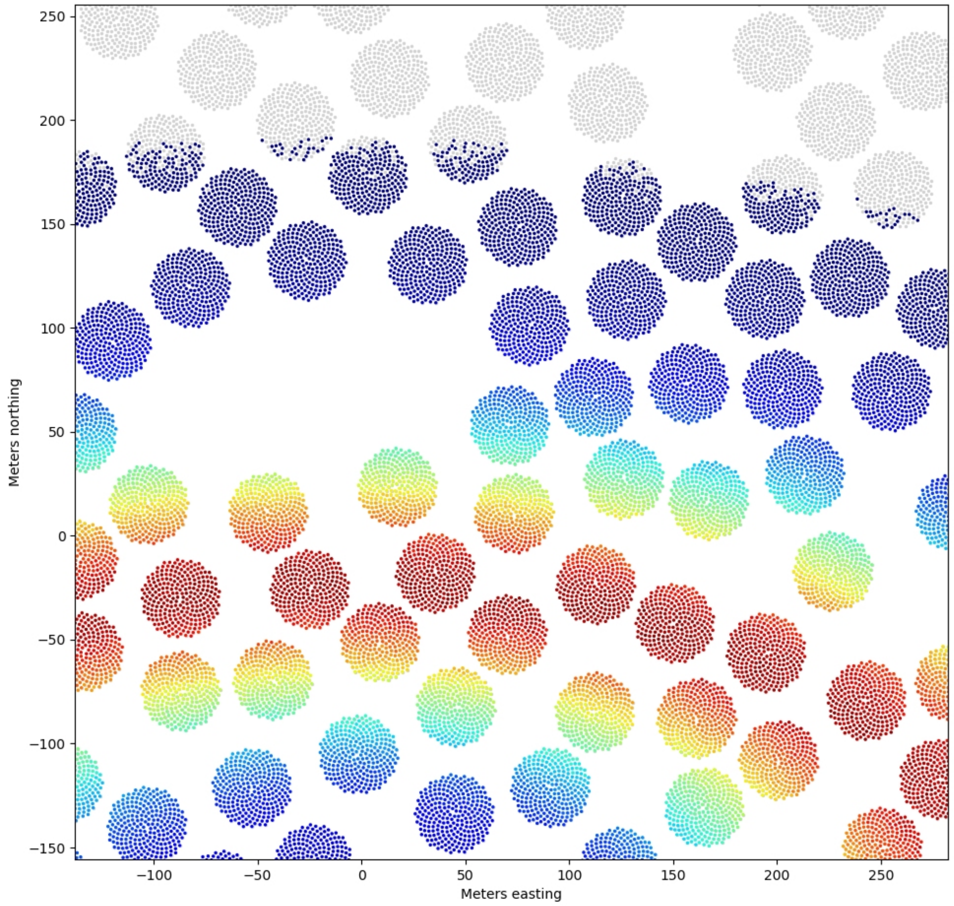
LOFAR

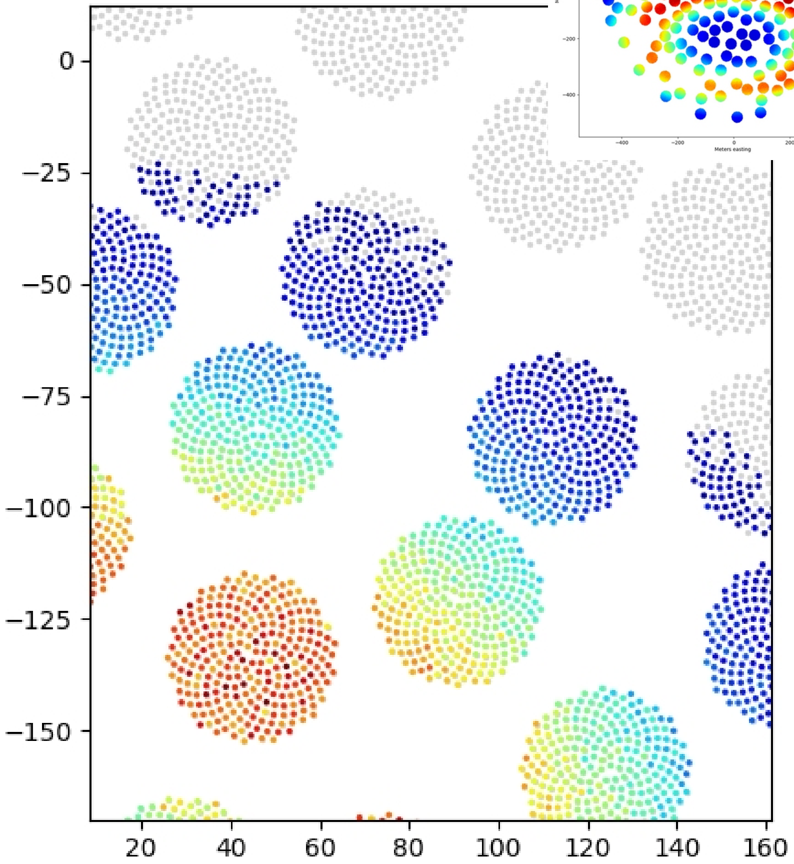
SKA antenna layout

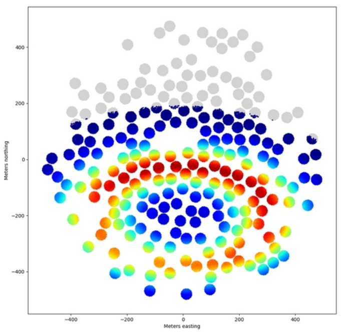




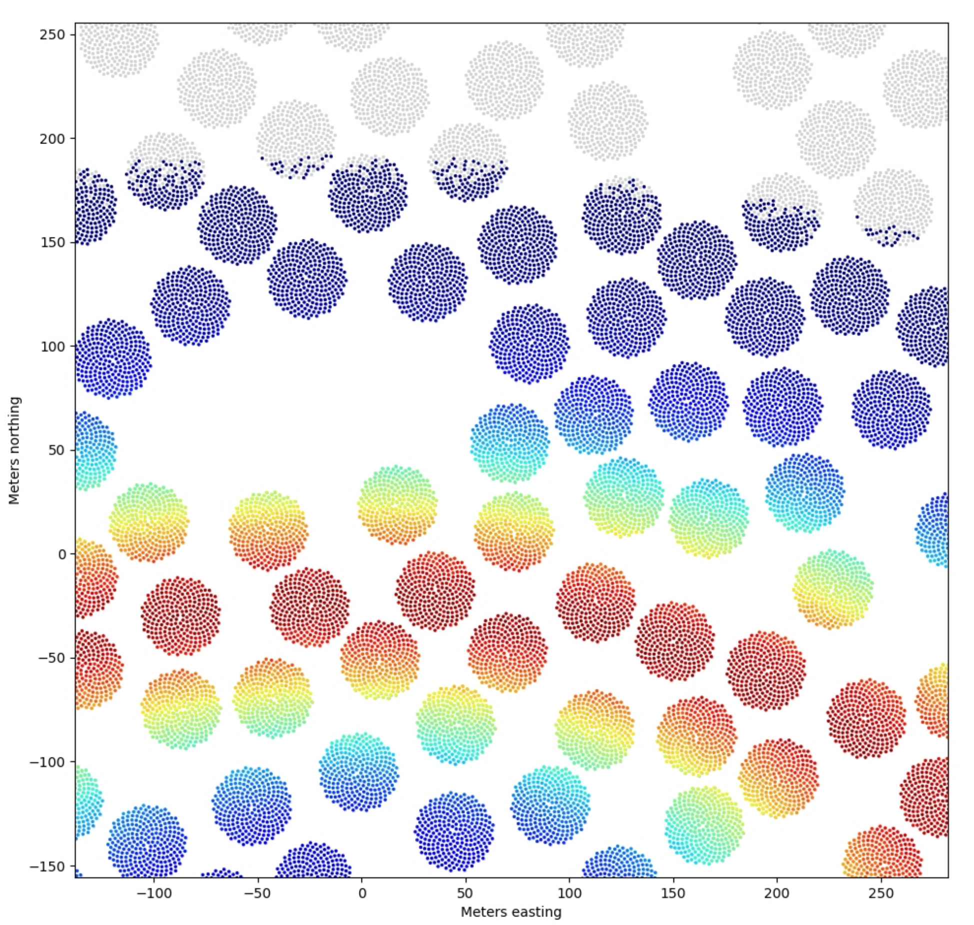
SKA-Low, a really dense array!







SKA-Low, a really dense array!

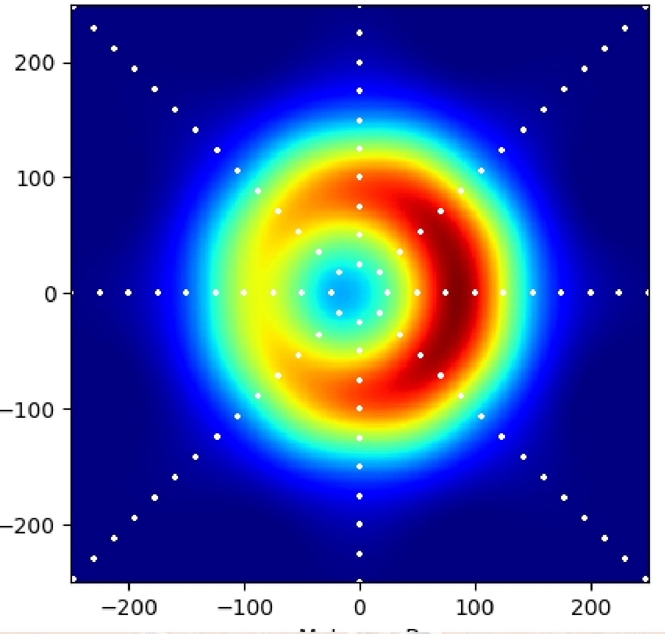


So, simulate 60,000 antennas??
(200 antennas takes 1.5 CPU-day)

Now what?

Interpolation from star-shape polar grid

(**50-350 MHz**, no antenna model)



Simulating 60,000 antennas is intractable

But: signals will not vary much across ~ 2 meters (redundancy)

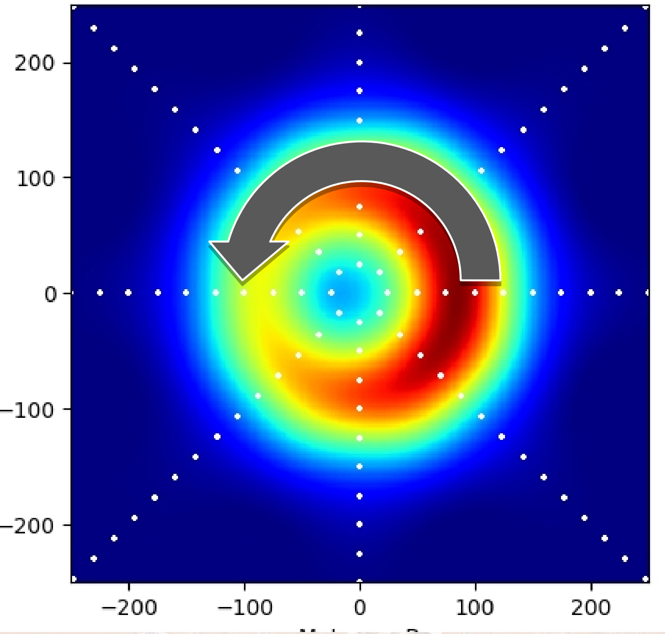
So, an **interpolation** routine would be helpful

Requires **high accuracy**
(in line with measuring with 60,000 antennas)

and **full signal traces**

Interpolation based on Fourier series

(**50-350 MHz**, no antenna model)



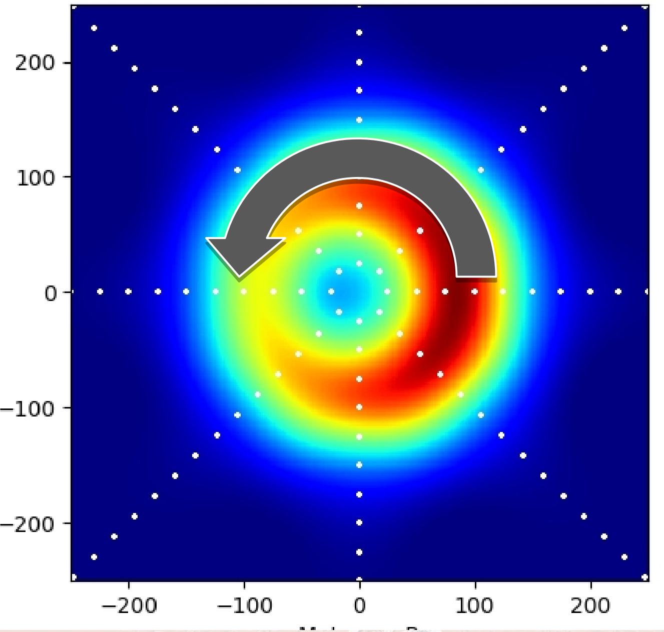
Idea:

- At each radial position, use Fourier series (FFT) to describe the angular variations
  - To a large extent: ~ cos(phi)
  - If purely sum of geomagnetic and charge excess, ampli f(x, y) ~ 1 + C cos(phi), C = charge excess fraction
  - Energy: $f^2(x, y) \sim 1 + \frac{1}{2} C^2 + 2C \cos(\phi) + \frac{1}{2} C^2 \cos(2\phi)$

- Express as sum of cos & sin for clarity

Interpolation based on Fourier series

(**50-350 MHz**, no antenna model)



Idea:

- At each radial position, use Fourier series (FFT) to describe the angular variations
  - To a large extent: ~ cos(phi)

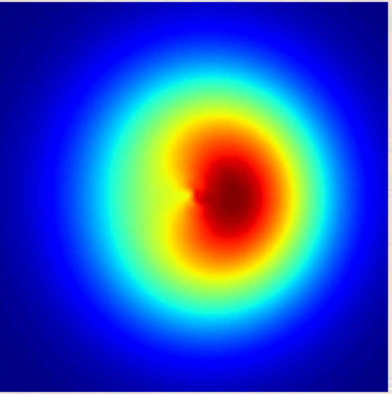
- Interpolate radially ALL the **8** components of the Fourier series
- Polar coordinates on grid, **r** gives Fourier components, evaluate at **phi**

- At each point, you have a sum
  $I(r, \mathrm{phi}) = \mathbf{I_0\,(r)} + \Sigma\, \mathbf{c_k(r)} \cos(k*\mathrm{phi}) + \Sigma\, \mathbf{s_k(r)} \sin(k*\mathrm{phi})$

- Use **cubic splines** in radial direction on the components, i.e. $\mathbf{I_0(r)}$, $\mathbf{c_k(r)}$, and $\mathbf{s_k(r)}$.

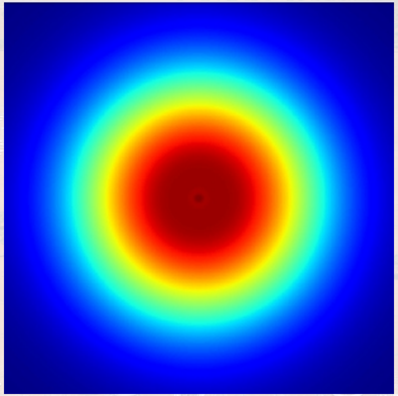
Decomposition in angular Fourier modes

CR Footprint



=

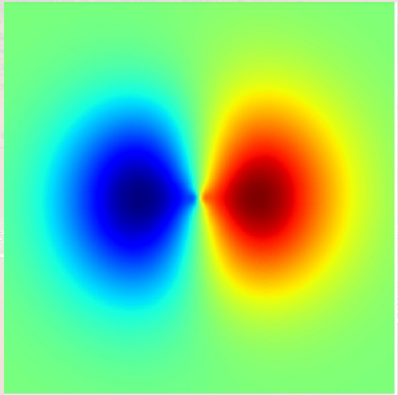
(100 %)



“monopole”

\+

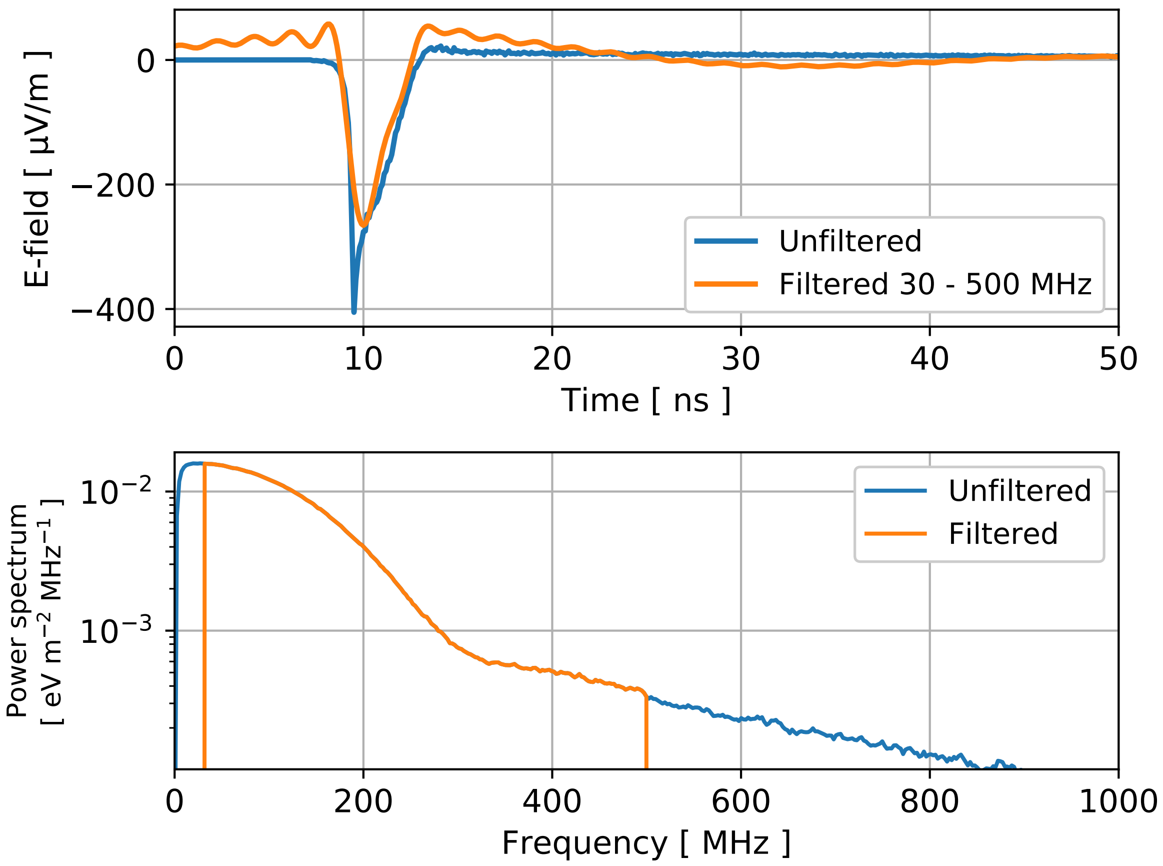
~ 20 %



“dipole”

\+ Higher order terms ~ 2 %

Interpolating full traces: Amplitude spectra

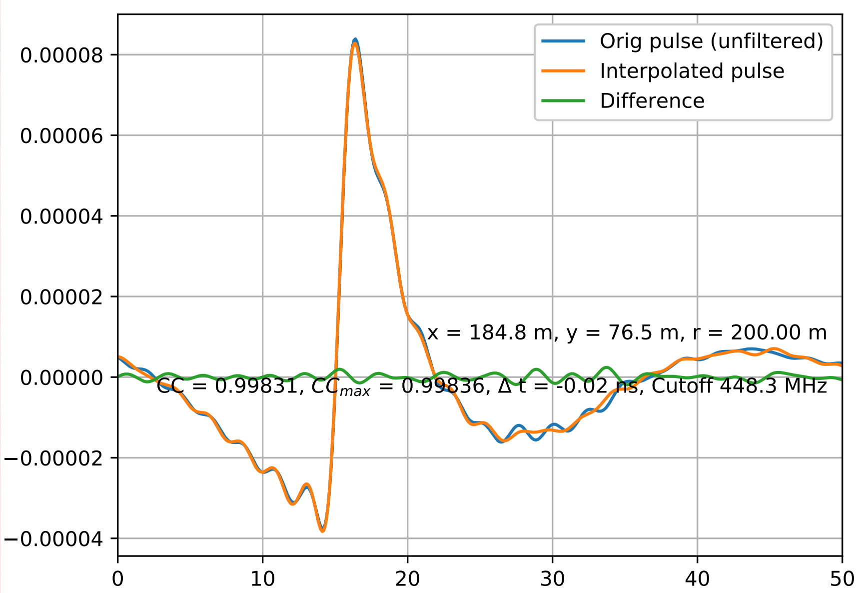


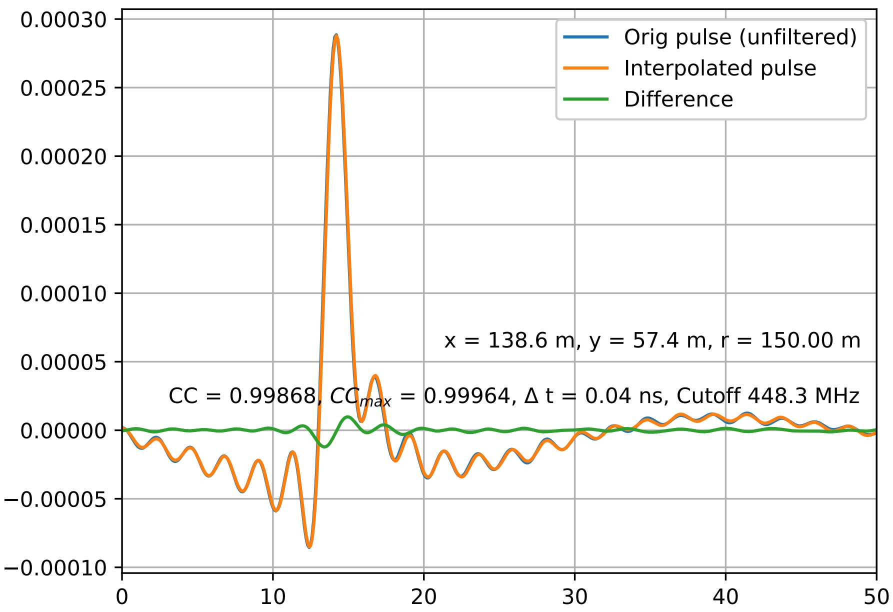
Each single-valued function at an antenna location can be passed into the interpolator (!)
For example:

- Arrival time
- Amplitude at frequency **f**

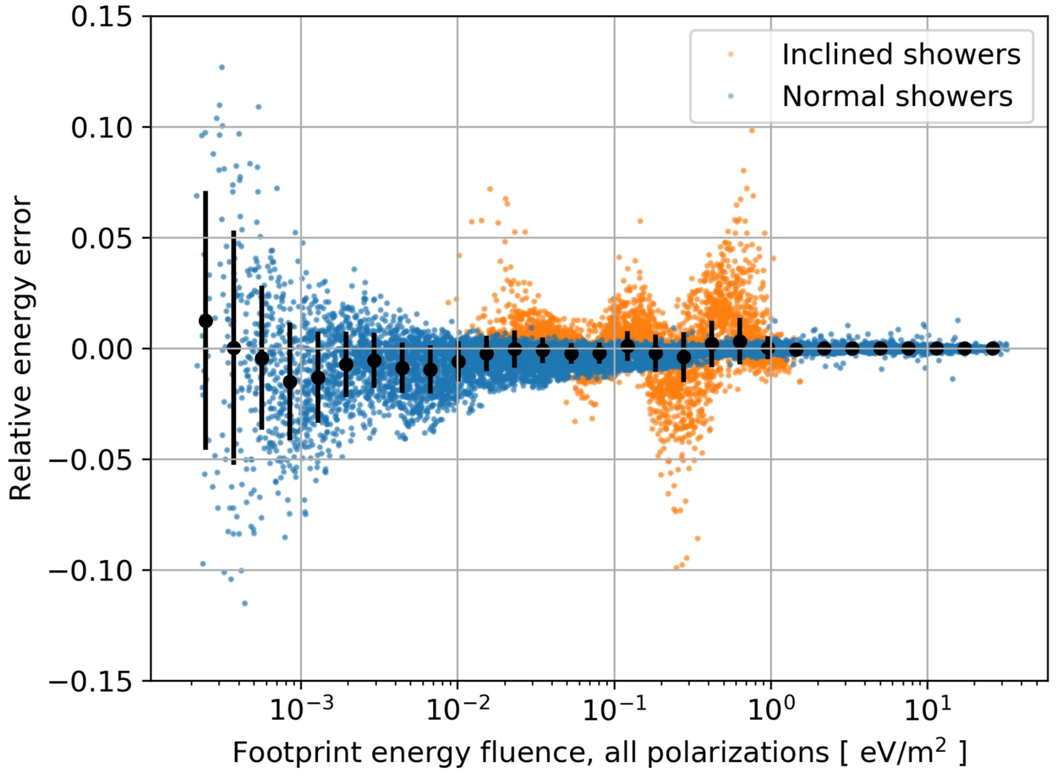
Performance results: pulse traces

30 to 500 MHz;
differences hard to see





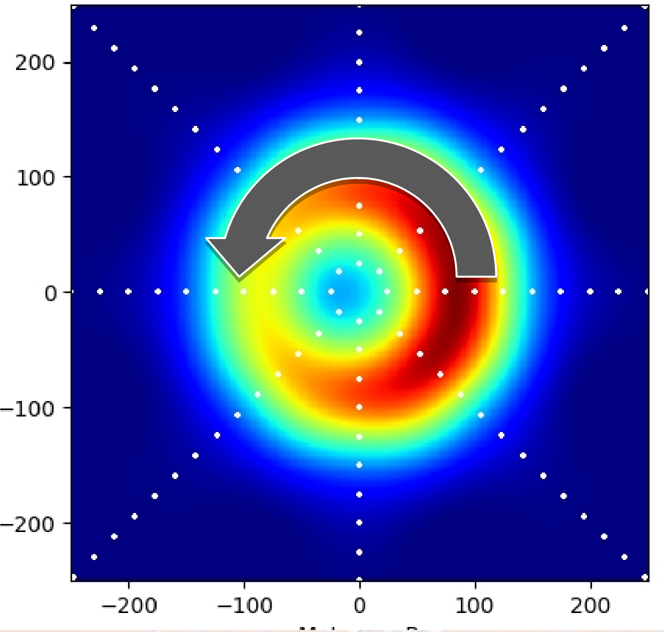
Performance results: fluence accuracy



- 30 to 500 MHz
- Accuracy better than 1 %, < 0.1 % for strong pulses
- Inclined showers had sub-optimal simulated antenna coverage
- When accuracy degrades, thinning artifacts etc become important
- “SKA-worthy” accuracy

Interpolation method

(**50-350 MHz**, no antenna model)



- Paper:
  Corstanje et al., JINST 18 P09005, 2023
  arXiv: 2306.13514

- Code:
  github.com/nu-radio/cr-pulse-interpolator

- Simulate 208 antennas, interpolate to ~ 60,000 (takes ~ 3 hours)

What makes SKA-Low unique for CR measurements?

- Very high antenna density, found nowhere else

  - e.g. Auger, GRAND etc. focus on highest energies, need very large area: antenna spacing ~ 50 – 200 m

- Wider frequency band than LOFAR, Auger/AERA

- Measuring air showers at the highest level of detail
  - Energy range 10^{16} to 10^{18} eV
  - More info on longitudinal distribution than just Xmax
  - Better constraints on proton fraction, mass composition
  - Reveals differences between hadronic interaction models at energies beyond LHC

The SKA layout

- The SKA antenna layout with particle detectors
- Placed quasi-randomly alongside antenna stations, respecting a 10 m distance to the nearest antenna

arXiv:**2504.16873**
submitted to Phys Rev D

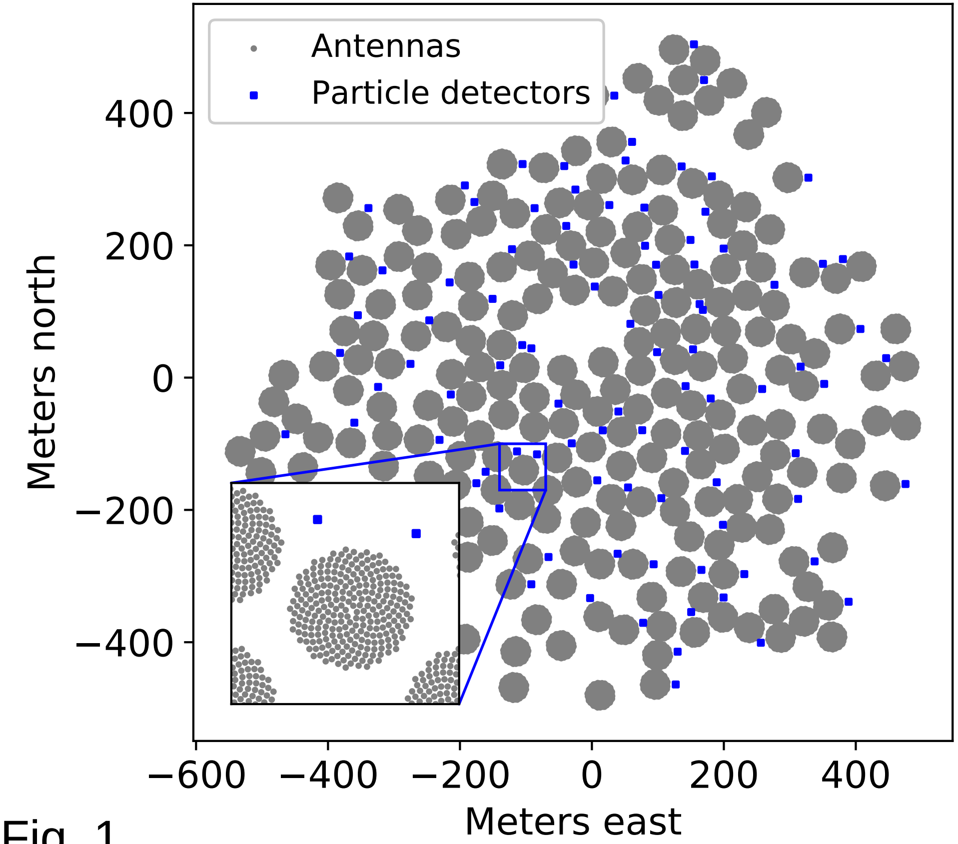


Fig. 1

Monte Carlo setup for SKA reconstruction

- Start from CoREAS traces
- Use pulse interpolation to get ~ 30,000 antenna traces
- Decimate number of antennas if desired (speed)

- Apply SKA Antenna model (NRR-SKALA4)
- De-disperse pulses (unit-gain filter)
- Add Galactic noise (NRR)
- Apply noise-whitening filter
- Measure fluence in **24** ns time window, **5 sigma** threshold

- Get fluence uncertainty from formula found using noise-Monte Carlo
  - No noise trials needed in simulated measurement
  - Still **memory intensive** with up to 30,000 antennas

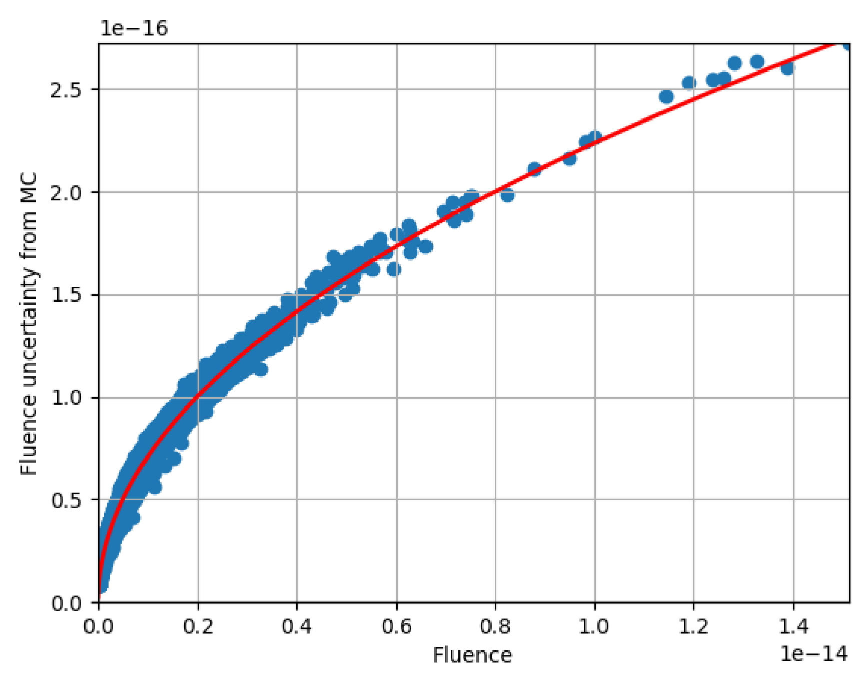
Fluence uncertainty versus (real) fluence

- Get fluence uncertainty from formula found using noise Monte Carlo
  - Determine fluence on simulated positions 1000 times with different noise realizations
  - For 30 showers
  - For 5 different energy levels to see different pulse shapes at various fluence levels etc

  - Formula: uncertainty = **sqrt(A * fluence + B)**

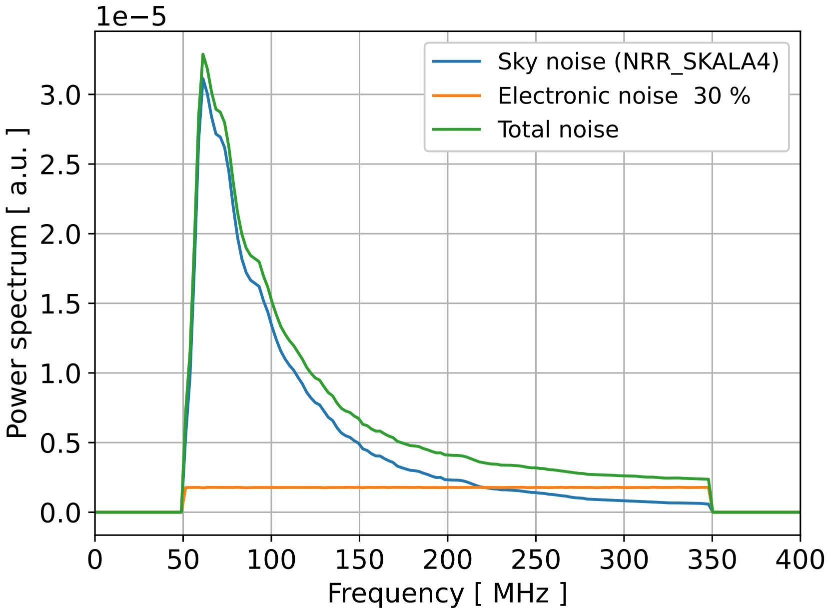
  - Gaussian uncorrelated noise:
    A = 4 sigma^2
    B = 2 N_s sigma^4

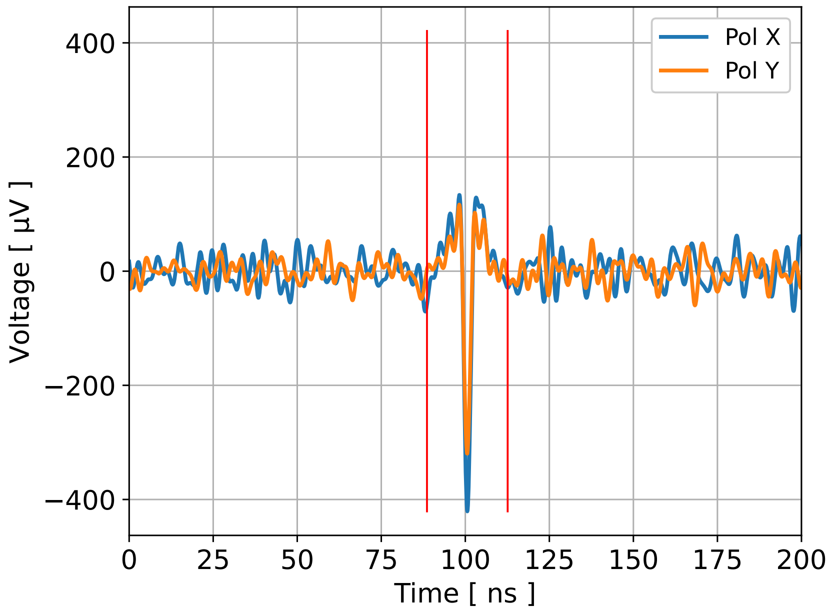
    N_s = number of samples in window
    sigma = stddev of noise



Noise spectrum and example time trace

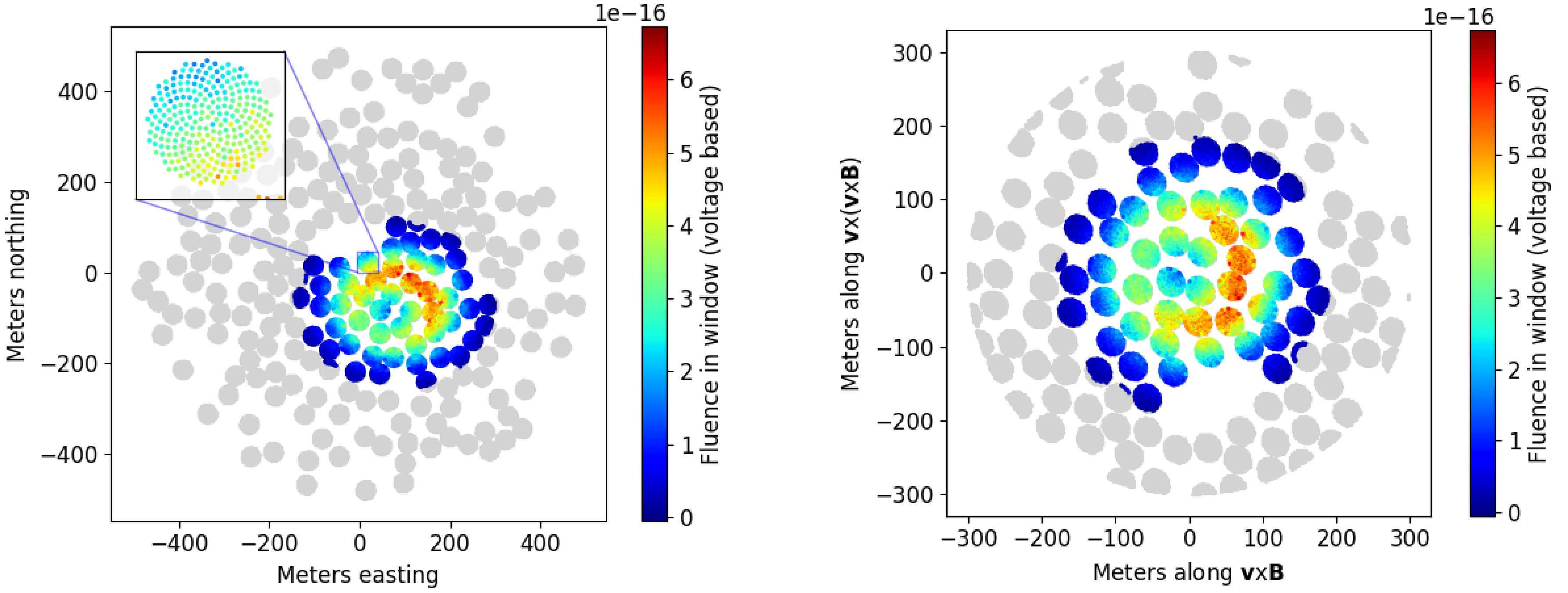
- Noise spectrum from the Galaxy, plus flat-spectrum thermal noise at 30% of Galactic noise energy
- Time trace of a pulse, after phase-compensation filter (hence symmetric)





Example footprint on ground and in shower plane

- Footprint of a zen 30 degree air shower

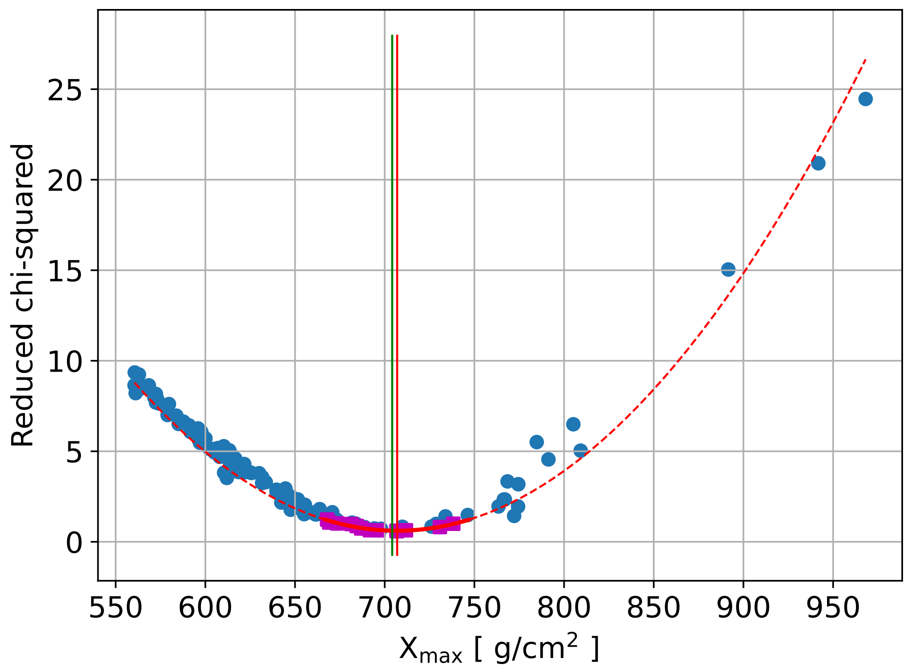


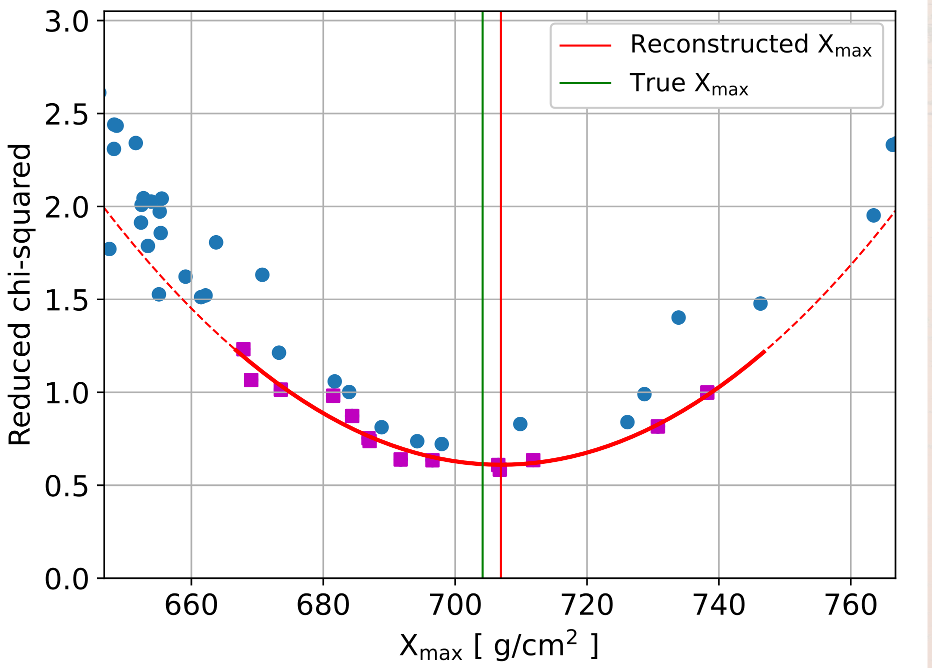
Simulation ensemble

- **140 showers** each: 50 p, 20 He, 20 C, 20 Si, 30 Fe
- From east; zenith angle 15, 30, 40 degrees
- Use one shower as mock data, use other 139 to reconstruct Xmax
- Do all showers in turn as ‘data’
- Determine reco-Xmax minus true Xmax
    - Average error: **bias**
    - Stddev error: **precision**

Fit quality, chi-squared vs Xmax

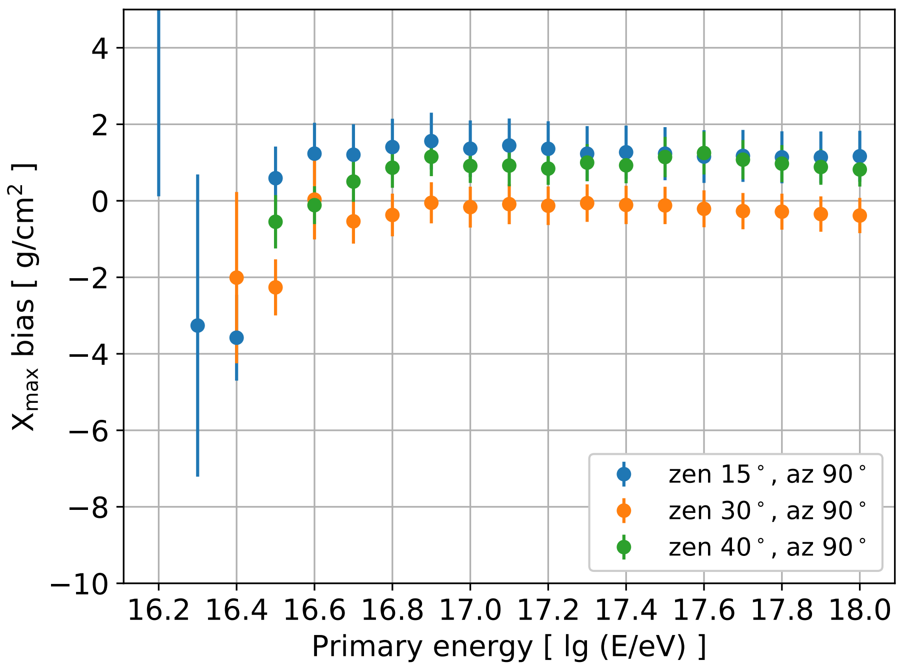
- Example fit quality plot, with Xmax reconstruction (parabola fit)

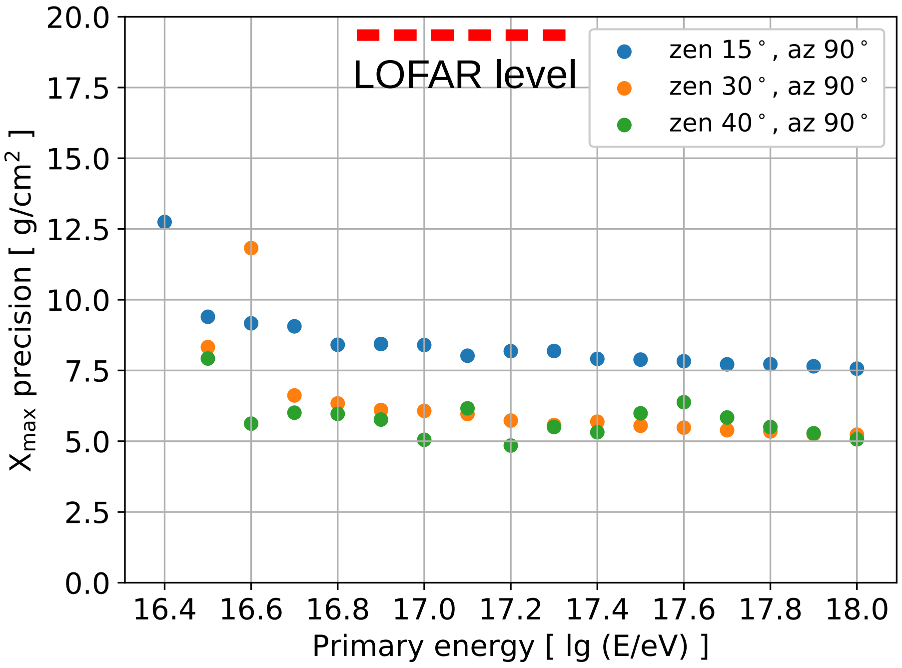




Bias and precision of Xmax reconstruction

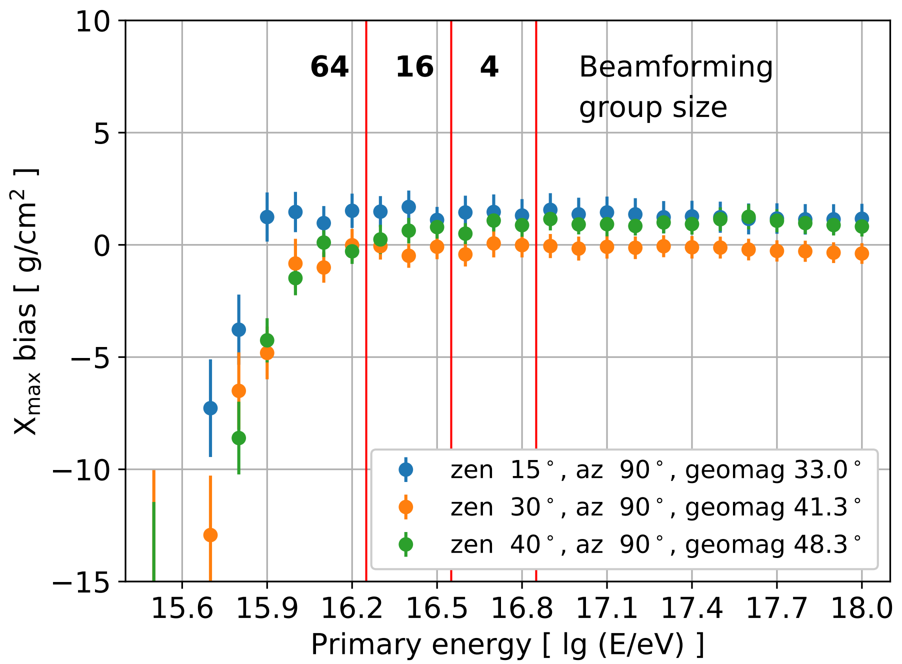
- Versus primary energy, with 1 in 4 antennas as “full array”
- Compare to 20 g/cm^2 precision on average with LOFAR

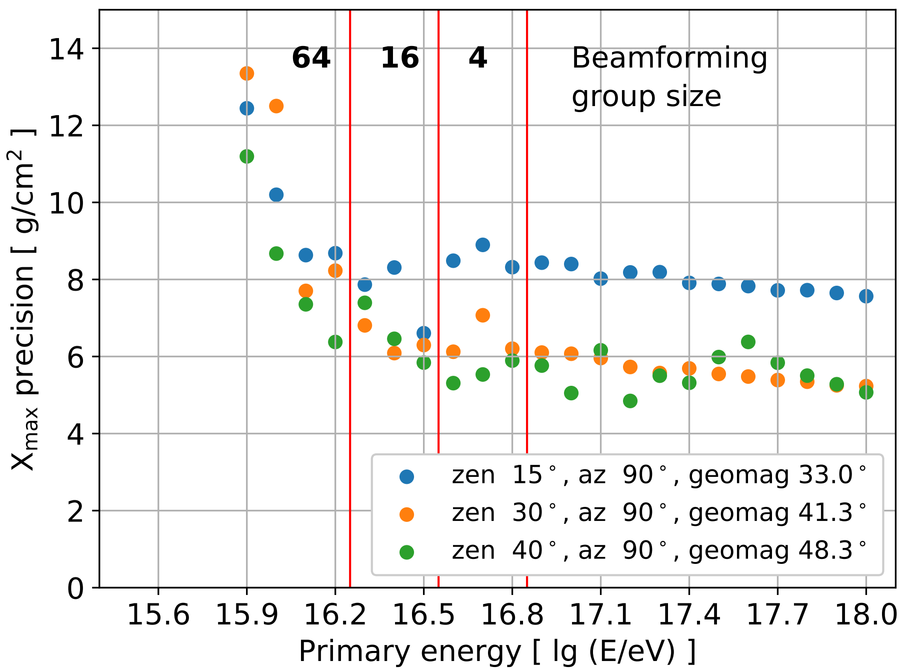




Bias and precision of Xmax reconstruction with (pseudo-)beamforming

- Results stay good down to 10^**16.0** eV
- Again taking 1 in 4 antennas as “full array”

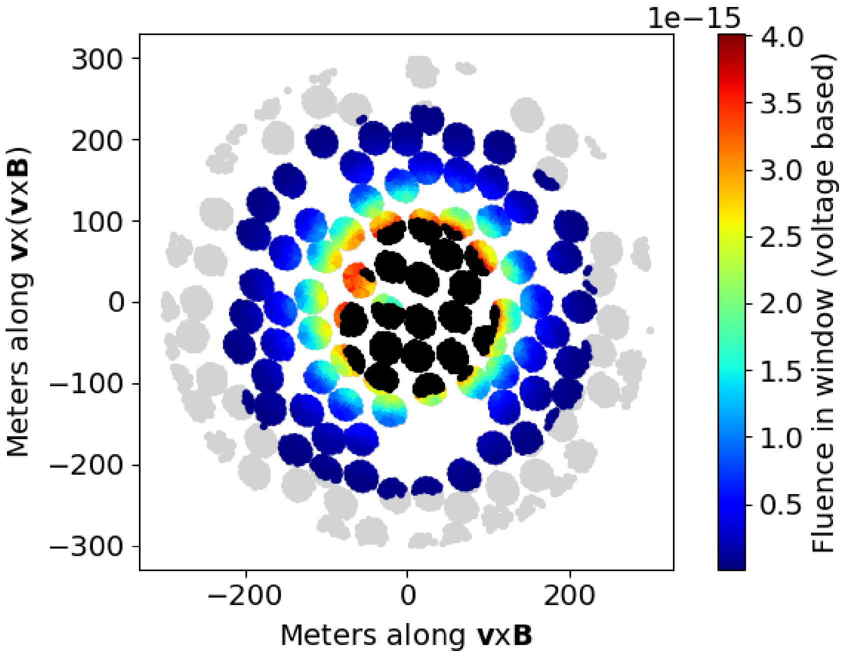




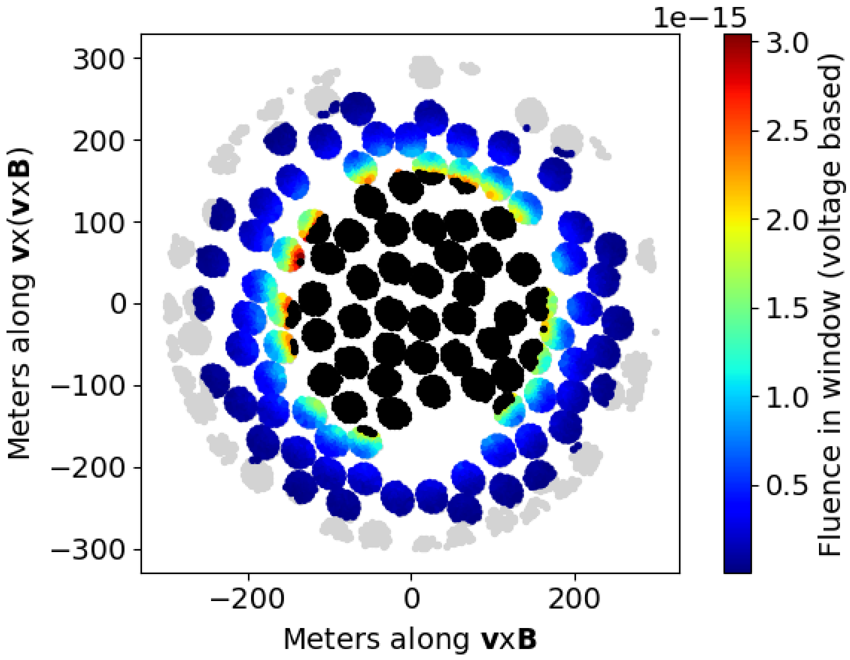
Example footprint with max dynamic range 32 sigma

- Cutting out all antennas where >= 1 polarization saturates the range

- 10^17.5 eV

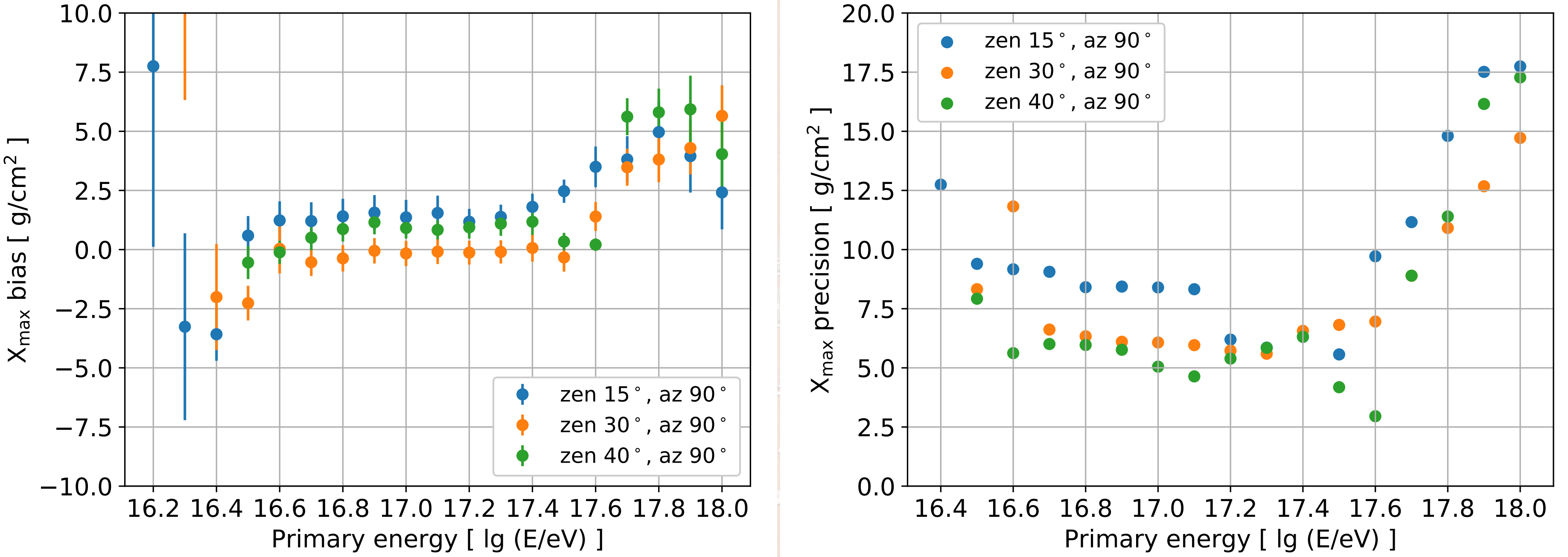


10^17.8 eV



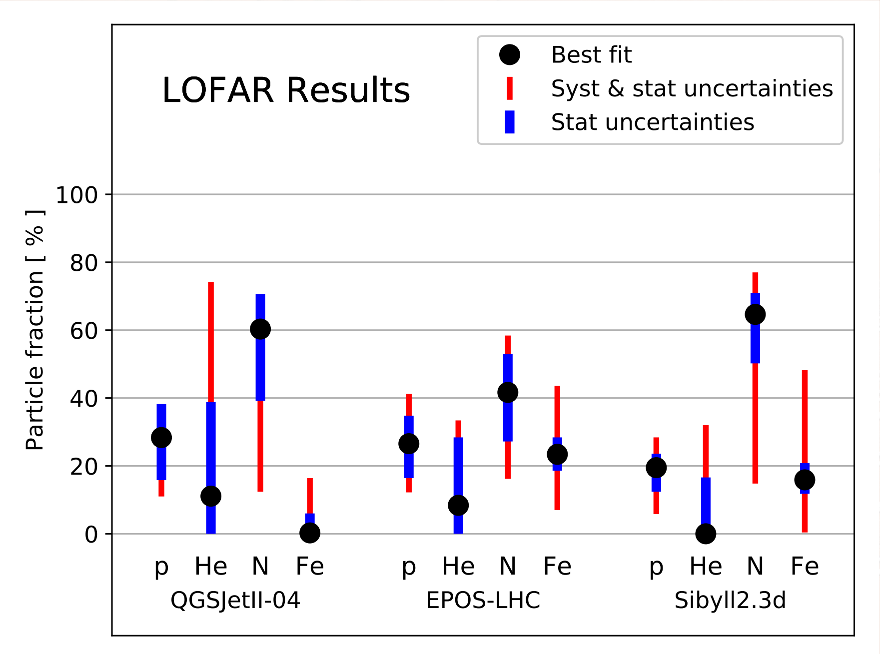
Bias & precision results with dynamic range 32 sigma

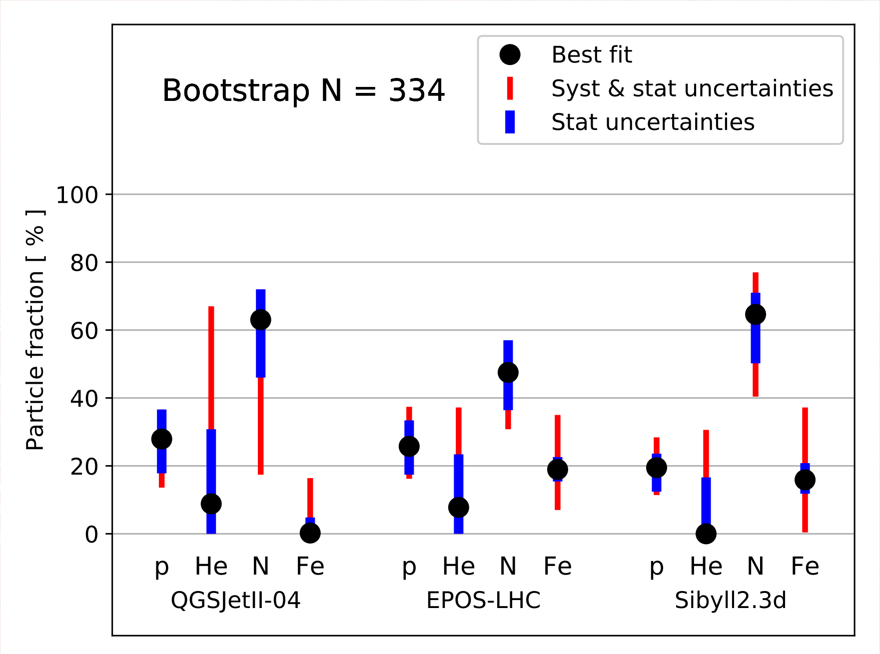
- Degradation at lg E >= 17.6
- Bias increases, has to be investigated (later, not here)
- Precision manageable, though newer analyses may be impacted more



Mass composition results

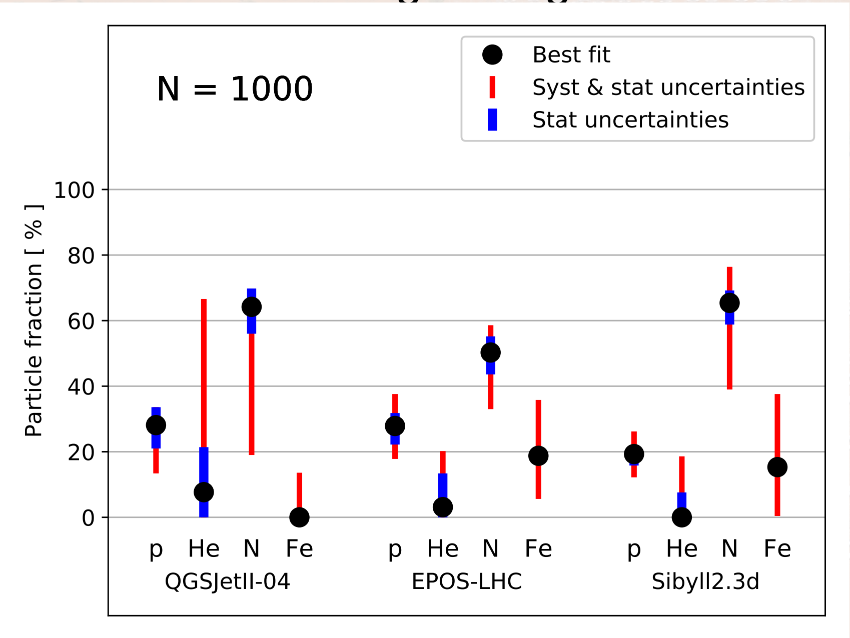
- How many showers needed to have:
  - Systematic uncertainties >> statistical uncertainties
  - A mass composition in narrow energy bins, improving over LOFAR (2021)

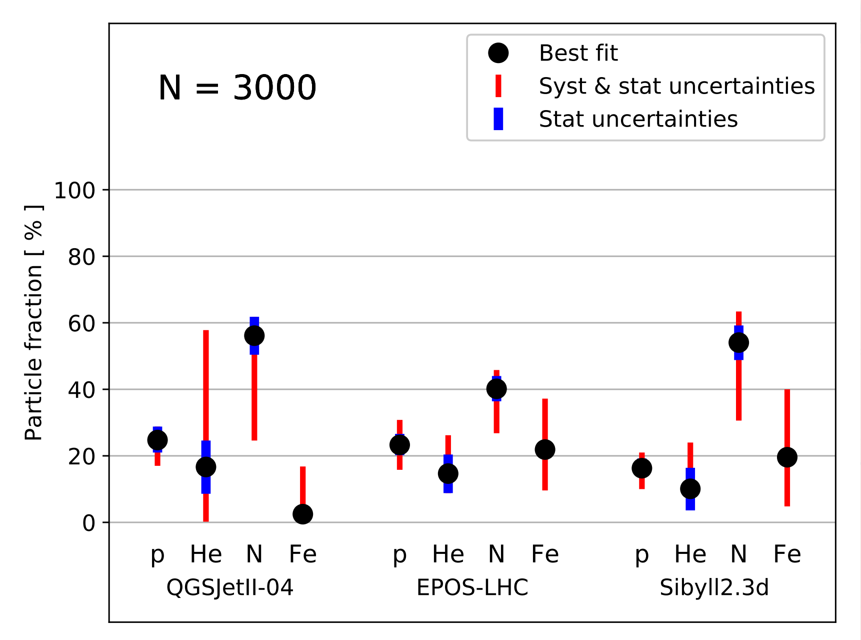




Mass composition results

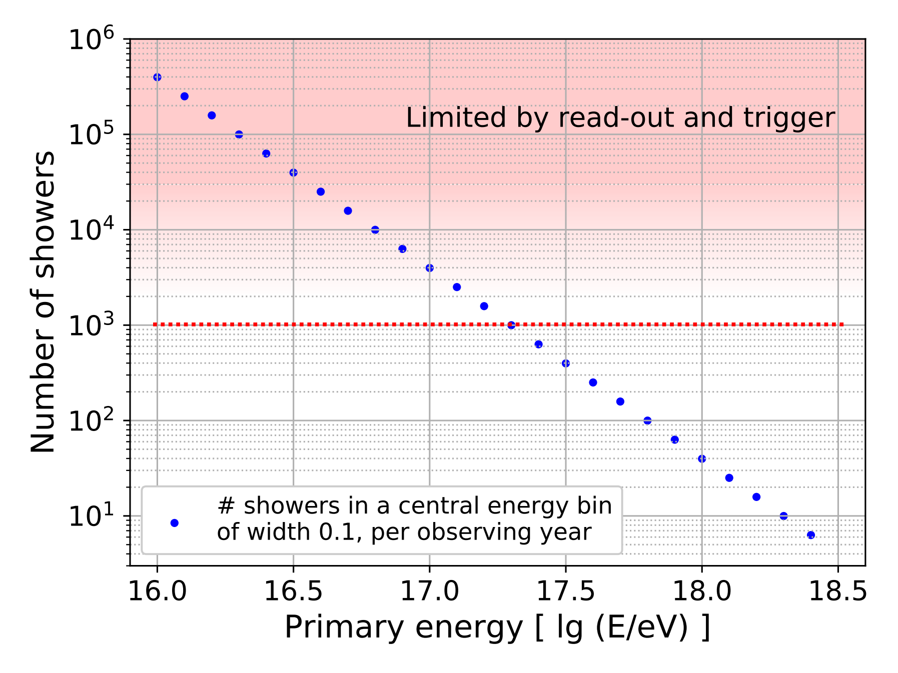
- Bootstrap: take best-fitting mass composition (LOFAR)
  - Do random drawing from Xmax distribution for that composition, add random Xmax errors +/- 7 g/cm2, account for syst errors +/- 9 g/cm2
  - Run that drawing through same mass composition analysis





Cosmic ray spectrum & expected event counts

- Event counts in an energy bin of width 0.1 in lg E/eV
- Over 1 observing year = 1 year of CR-mode uptime
- Nature-limited above **lg E ~ 17.3**, tech/obs limited below



Next steps

- Results presented here are just the beginning!

- There is more to be measured than only Xmax and energy
  - Longitudinal distribution
  - Special showers (‘double bumps’)

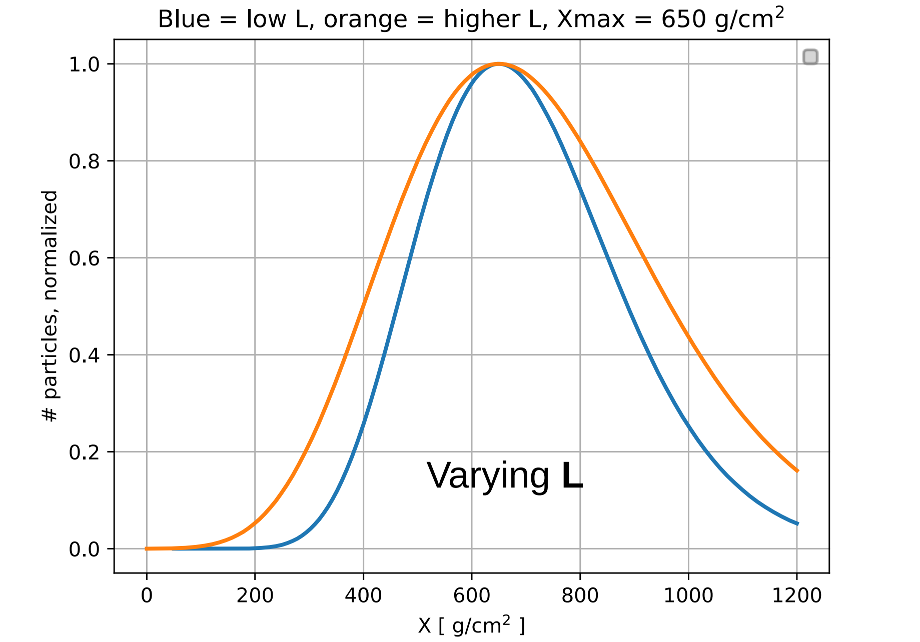
- New analysis techniques
  - Information field theory
  - Interferometry / beamforming
  - Other approaches that use timing, polarization, frequency spectra

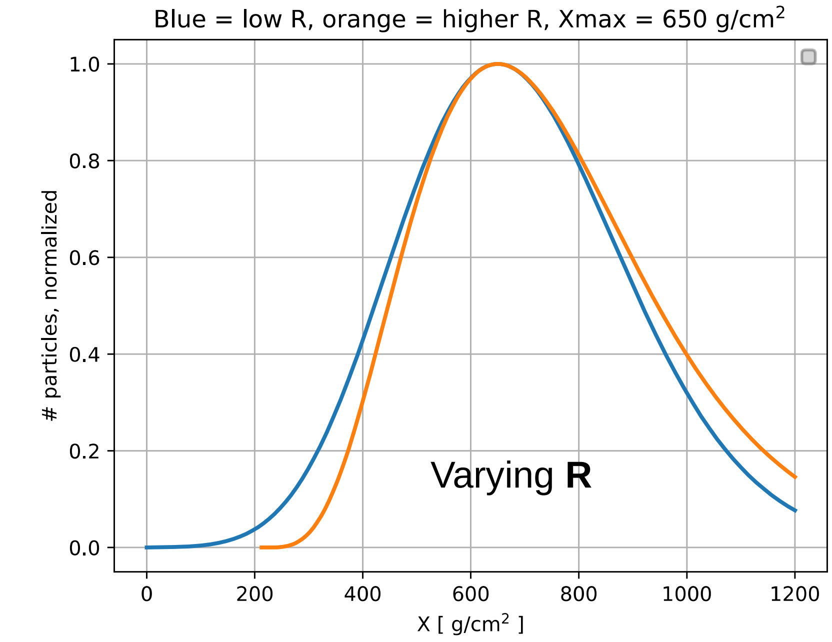
Beyond Xmax: Longitudinal distribution of particles

$$N(X) = \exp\left(-\frac{X - X_{\max}}{RL}\right)\left(1 - \frac{R}{L}(X - X_{\max})\right)^{\frac{1}{R^2}}$$

Parameter **L**: width (variance)

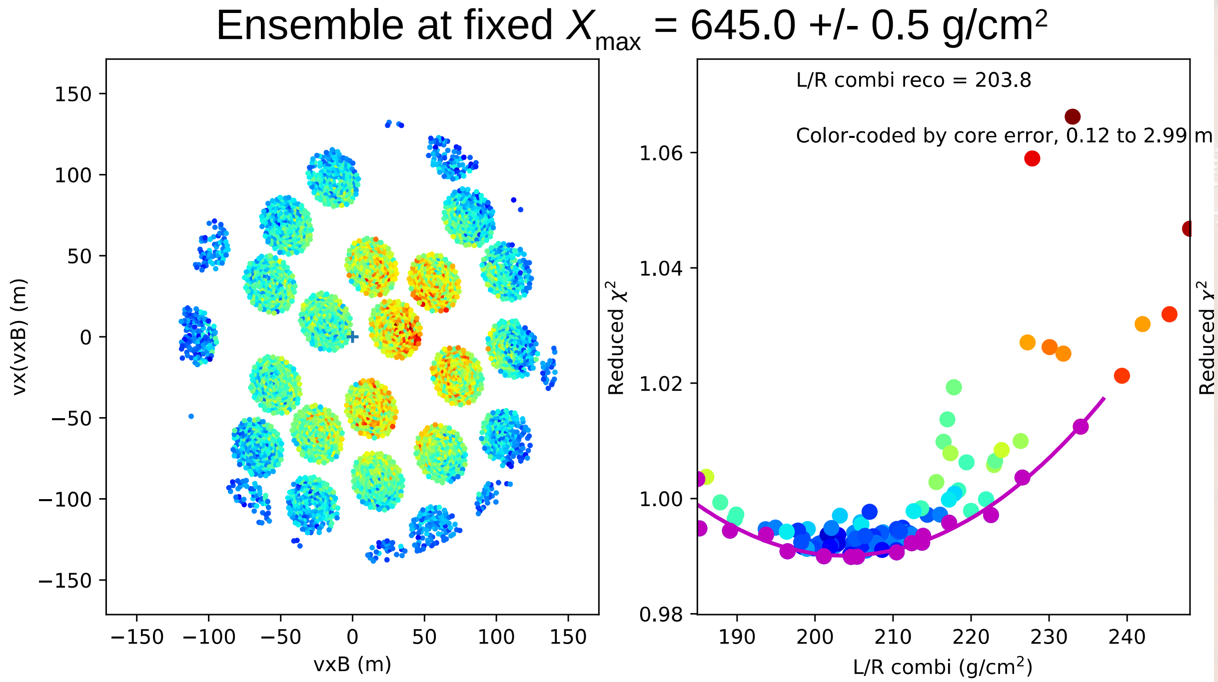
Parameter **R**: asymmetry (skewness)





Reconstruction result for L+R combination

50 – 100 MHz



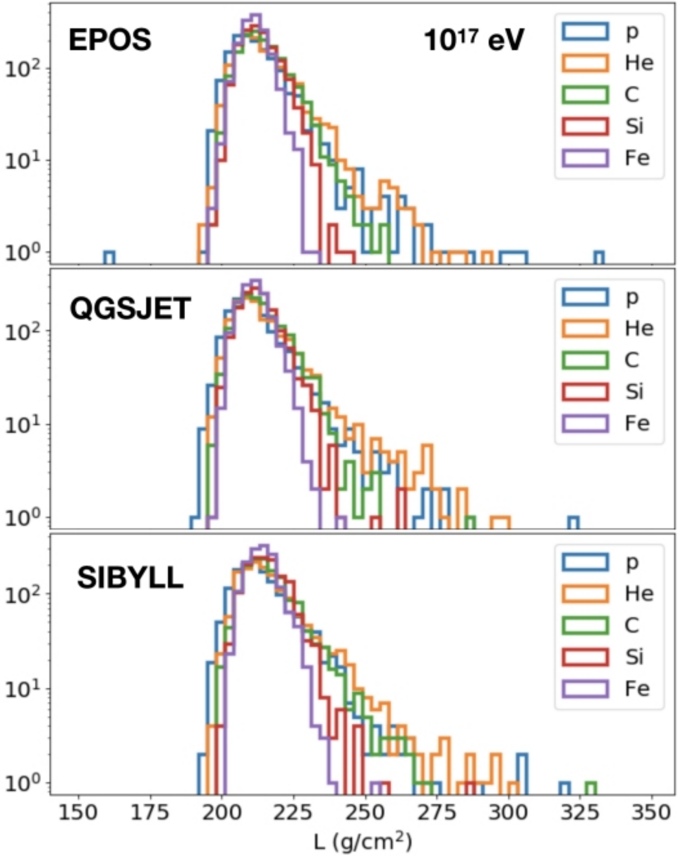
~ 5000 antennas triggered

Reconstructed "L+R" = 203.8 g/cm²,

True "L+R" = 204.2 g/cm²

Reconstruction Xmax, and L+R together still to do

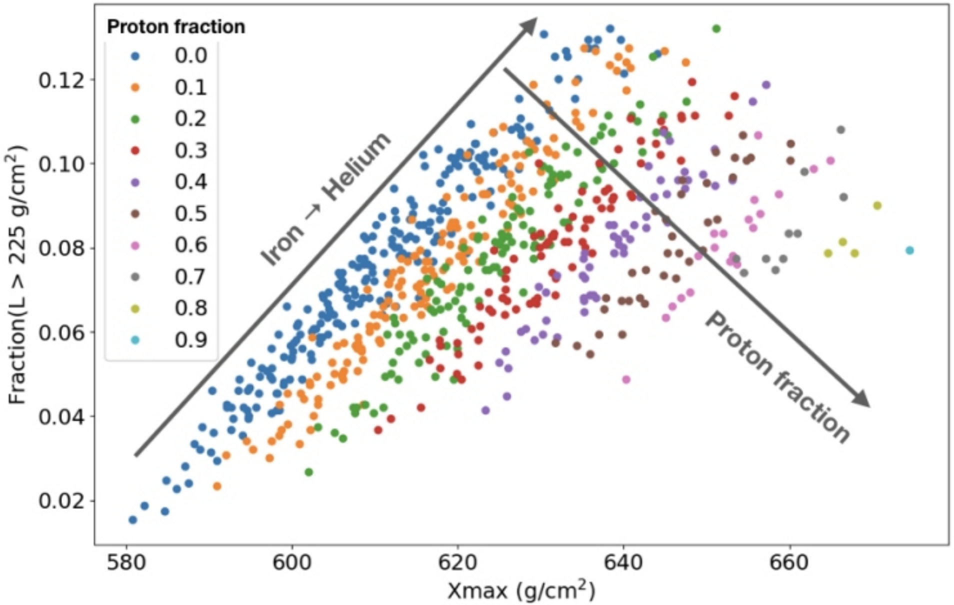
Distribution (tail) of L parameter: **proton fraction**



Tails are highest for **helium**, not protons

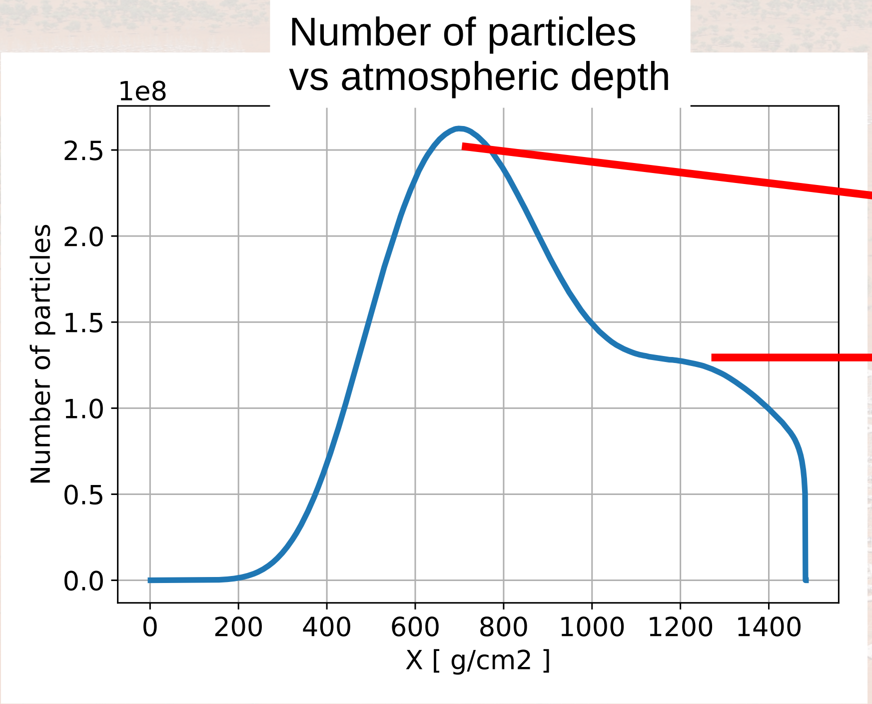
- Independent handle on proton fraction!

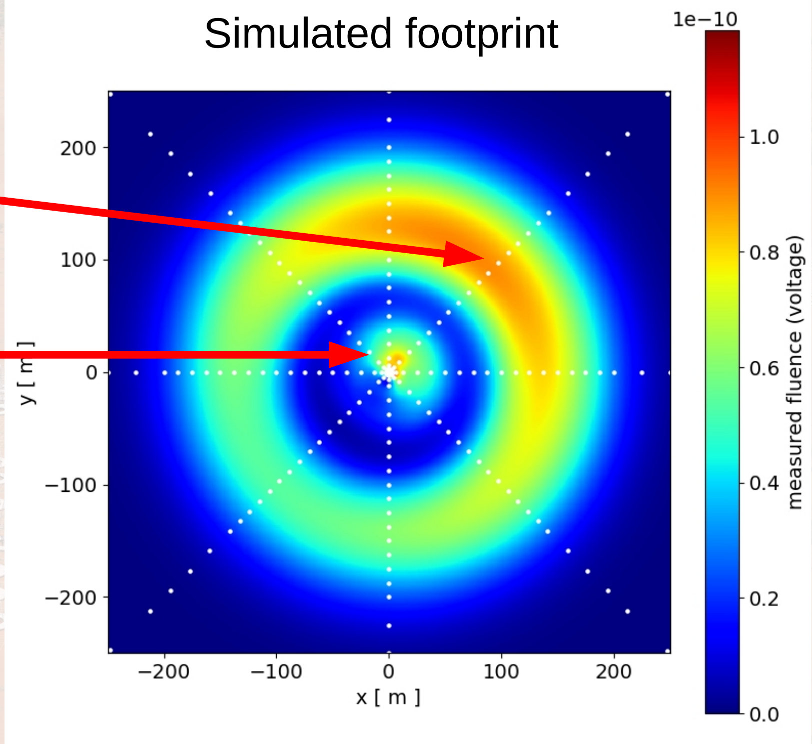
Fig. by S. Buitink



Outlier showers, 'double bumps'

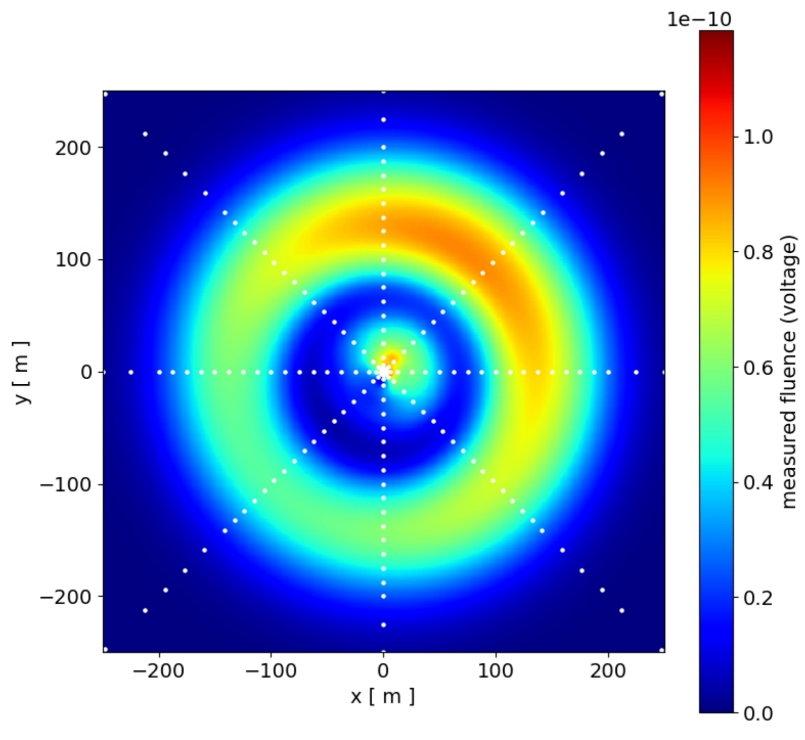
- Filter to 150 – 350 MHz band for sharper features
- Secondary shower visible separately (though uncommon)



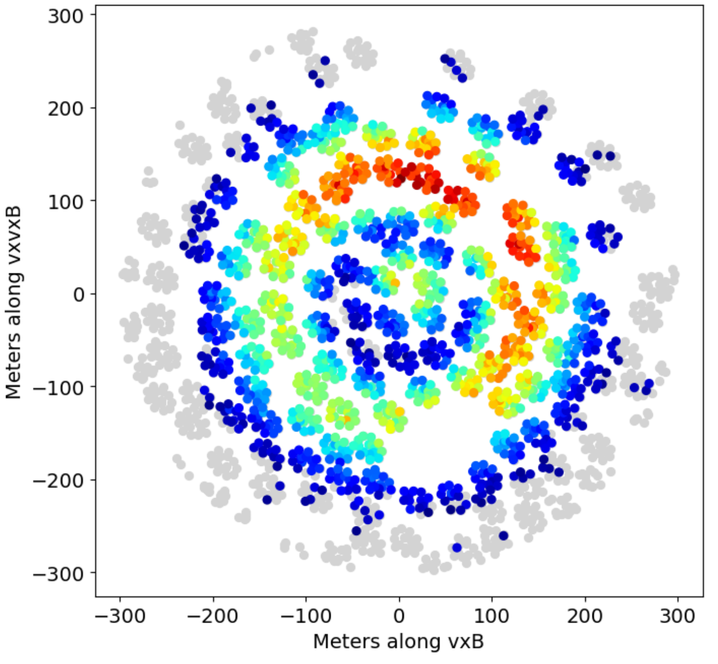


Outlier showers

- SNR boost by a factor **4** from beamforming patches of **16** antennas
- Further enhancement; clearly detectable double structure



E = 7 x 10^{16} eV



Summary

- SKA is the ‘ultimate’ dense radio array for cosmic-ray detection

- Using established techniques developed for LOFAR,
  precision on Xmax is 3x better than LOFAR, using known techniques
  (may improve later, measuring full shower evolution)

- Energy range extended to lower energies, **10^{16}** to **10^{18} eV** now available
- Mass composition analysis in narrow energy bins will be possible,
  for low energies quite soon

- New techniques promise to add new, independent information
  - Mass composition (p/He separation)
  - Hadronic physics

Backup