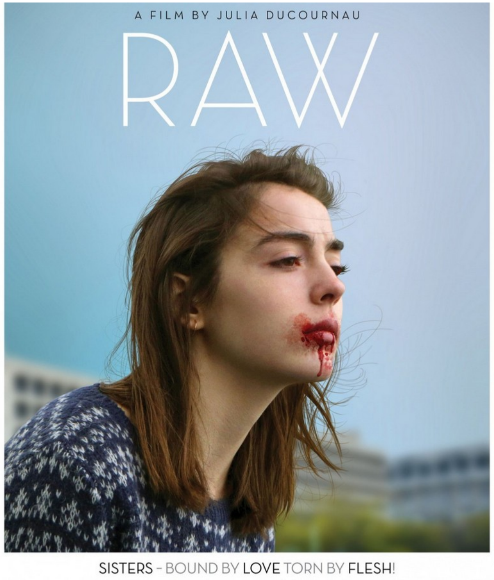
SII software school: Lesson I
Handling  data

T. Hassan on behalf of
the magic_spysii dev team

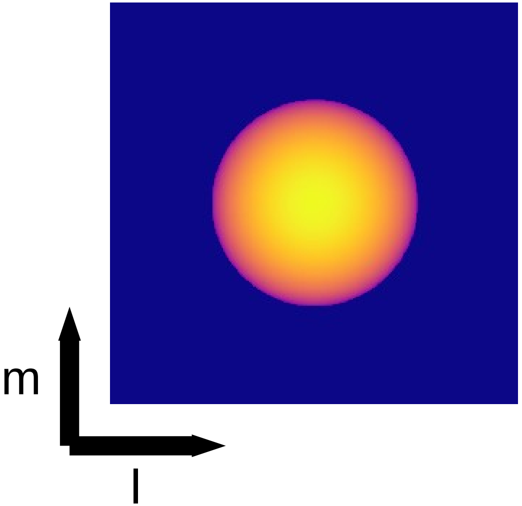
European Research Council
Established by the European Commission

Ciemat
Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas

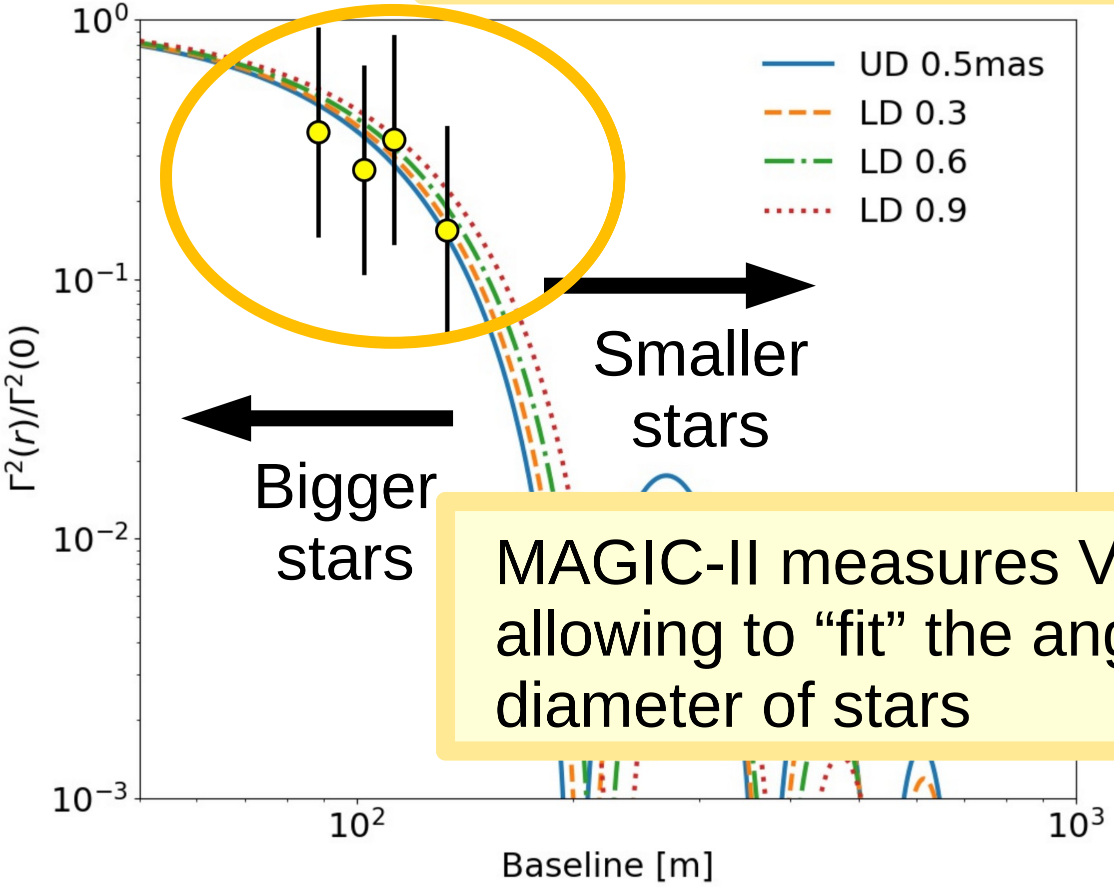
From last week: Analysis basics

• Van Cittert-Zernike Theorem:

Sky model



MAGIC-II observations



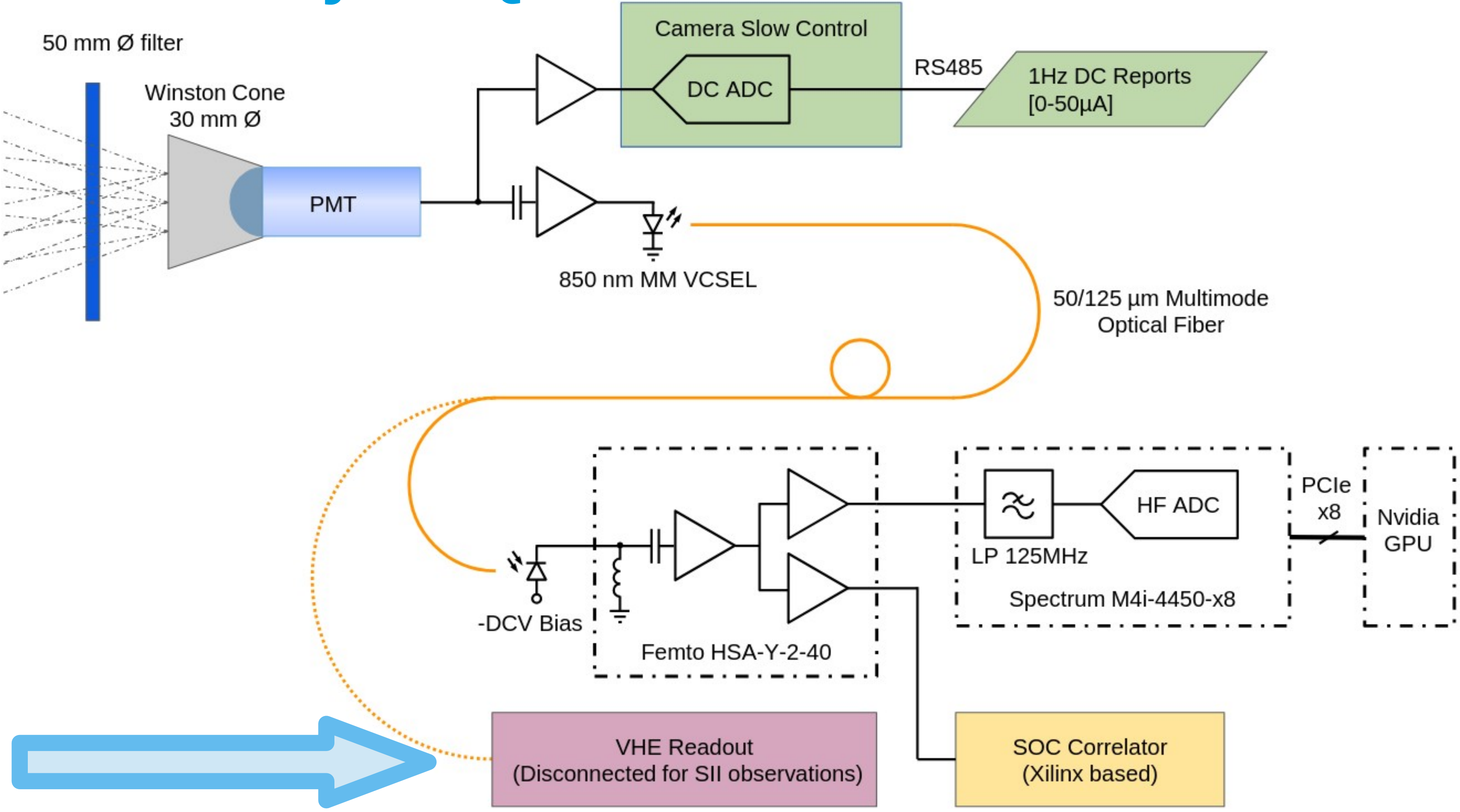
MAGIC-II measures V², allowing to “fit” the angular diameter of stars

Interferometry DAQ

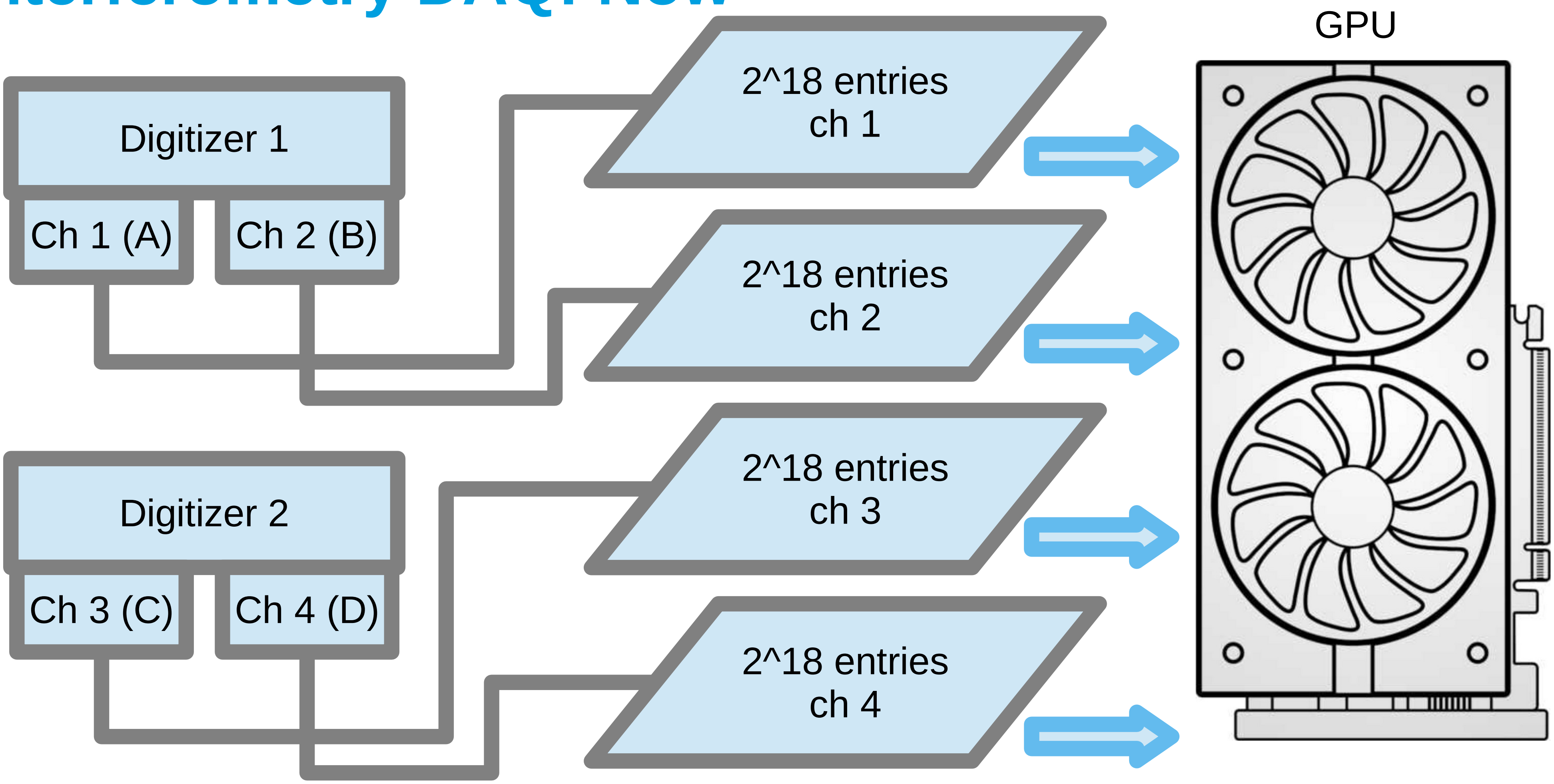
- The interferometry DAQ is a fully-independent machine within the MAGIC counting house, composed by:
  - Digitizers that sample the “interferometry pixels” at 500 Ms/s (2 ns)
  - GPUs that compute FFTs → correlation between the pixel signals
  - Disk space to allocate generated data

- Raw data files (*.bin) are stored in binary format (not the easiest format to work with), and copied to PIC (not on a nightly basis, so please be patient)

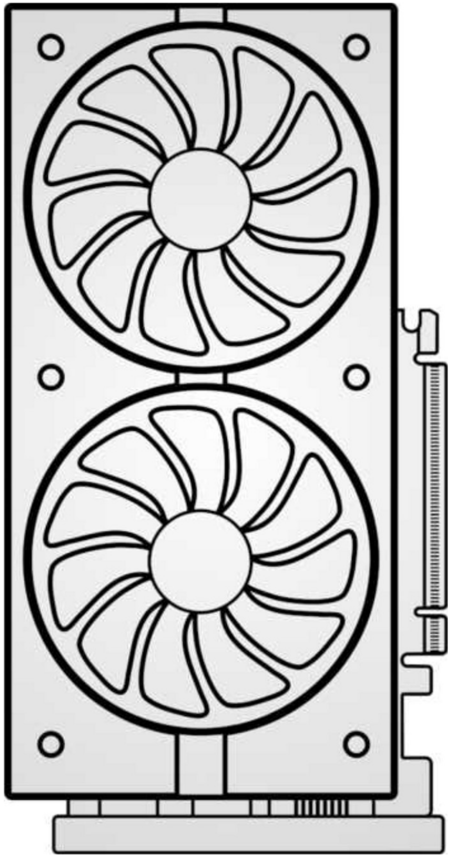
Interferometry DAQ



Interferometry DAQ: Now



Interferometry DAQ: Now



- Current DAQ computes:
  - Cross-correlation between each channel pair (via FFTs)
  - Auto-correlation of each channel
  - Certain useful parameters (mean and std dev of each channel)
- Current DAQ “accumulates” these, and writes the average correlation/mean/std dev every 500 cycles

Interferometry DAQ: In the past

- The DAQ has evolved **a lot**, and unfortunately for you, you need to have a bit of an understanding of what data you may expect
  - Originally there was 1 Digitizer
    - No accumulation
    - No several correlations (just 1)
    - Correlation arrays were stored every 2^18 entries (every 0.5 ms)

Interferometry DAQ: In the past

- The DAQ has evolved **a lot**, and unfortunately for you, you need to have a bit of an understanding of what data you may expect

  - Originally there was 1 Digitizer

  - Later on, a second digitizer was added (4-channel correlator). This was possible due to data “accumulation”. But:

    - **The 4 channels are not always the same pixels!**

      - Now there are 3 pixels from MAGIC and 1 from LST1

      - In the past… All kinds of combinations!

Interferometry DAQ: In the past

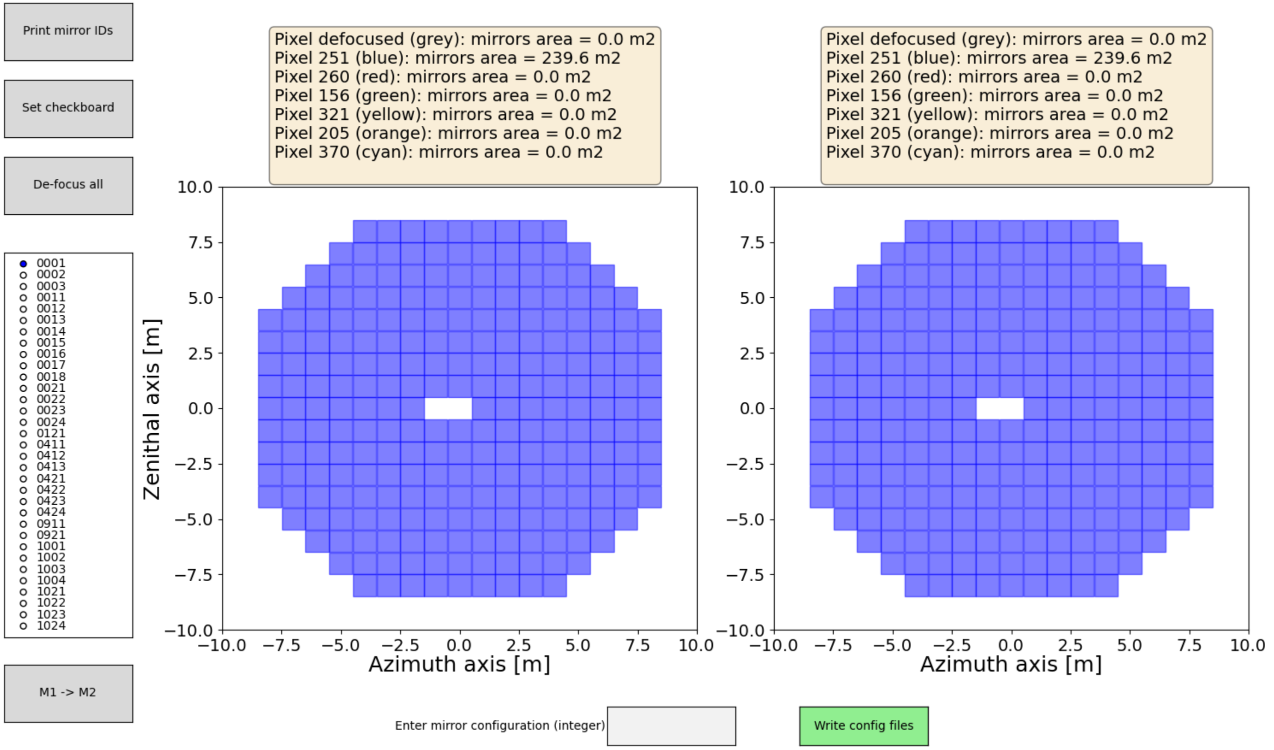
- The DAQ has evolved **a lot**, and unfortunately for you, you need to have a bit of an understanding of what data you may expect
  - Originally there was 1 Digitizer
  - Later on, a second digitizer was added (4-channel correlator). This was possible due to data “accumulation”.
  - The software “mostly” handles this by itself (its “hard” to make mistakes) but its important to understand that you may be adding data with a configuration different to what you have in mind.

Mirror configurations

• Very related to this point (which pixels are connected to the correlator) it is important to understand mirror configurations

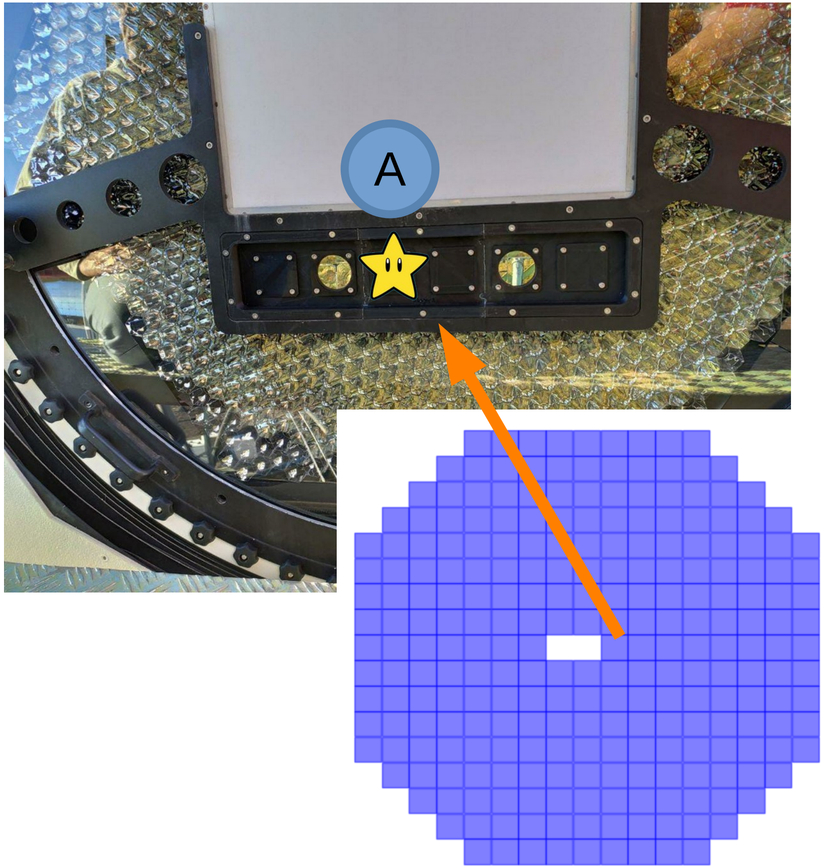
• Using AMC we can choose how to focus light into the interferometry pixels

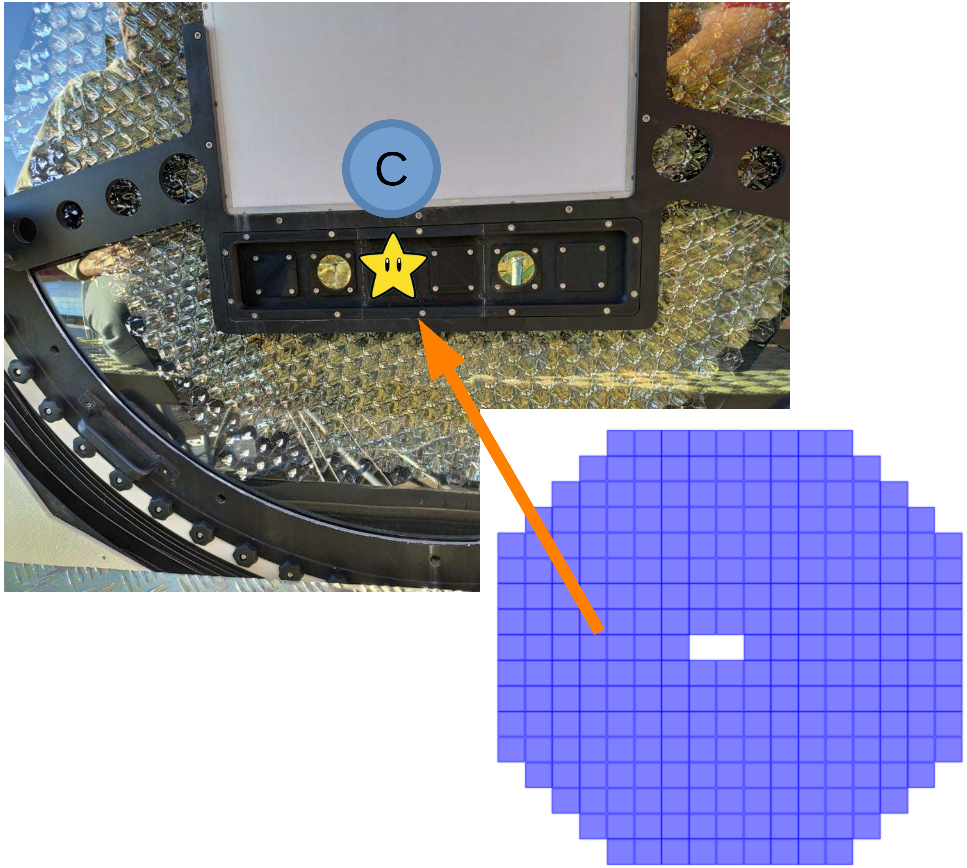
• Mirror configuration is added to the names of the raw data files. You may see a description of the different configurations in the MAGIC wiki



MAGIC-SII setup: The power of AMC

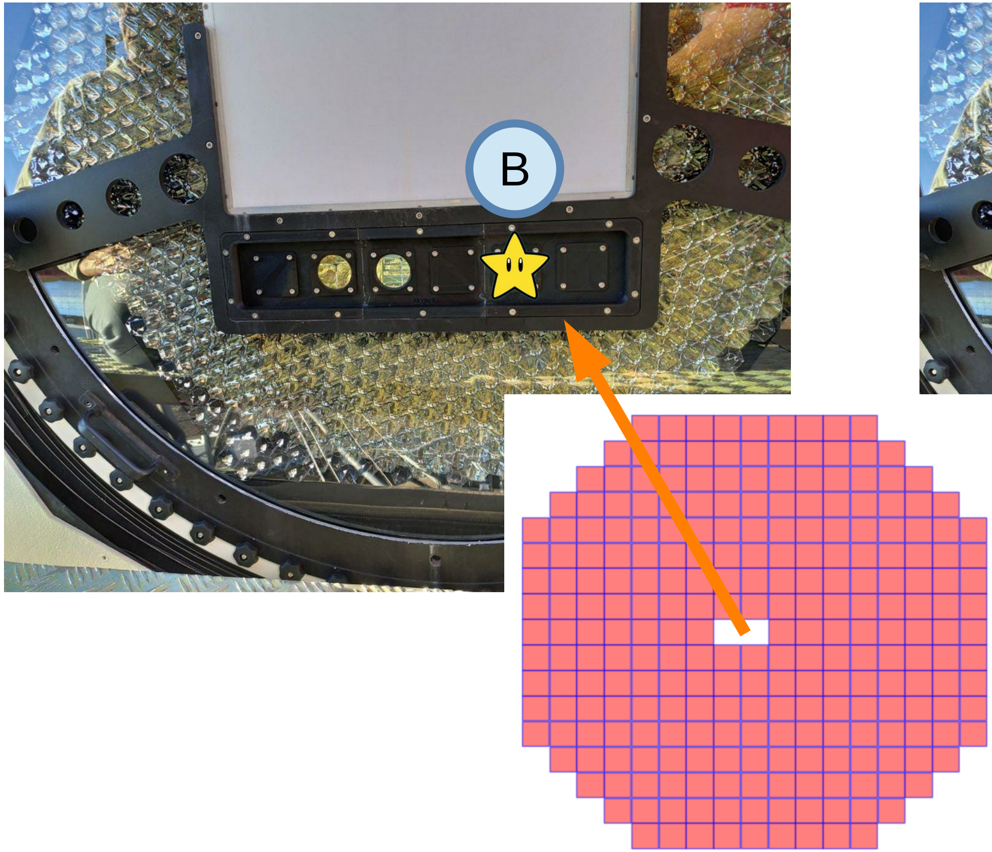
- This functionality adds enormous versatility to MAGIC: **full-mirror**

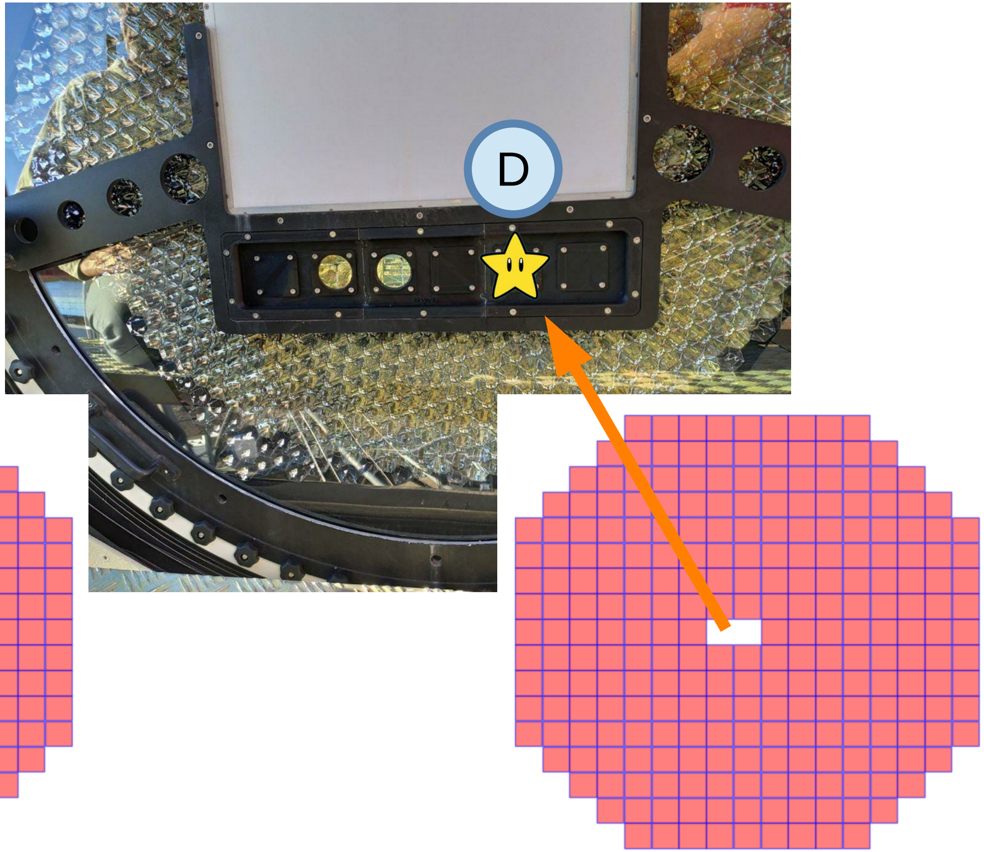




MAGIC-SII setup: The power of AMC

- This functionality adds enormous versatility to MAGIC: **full-mirror**

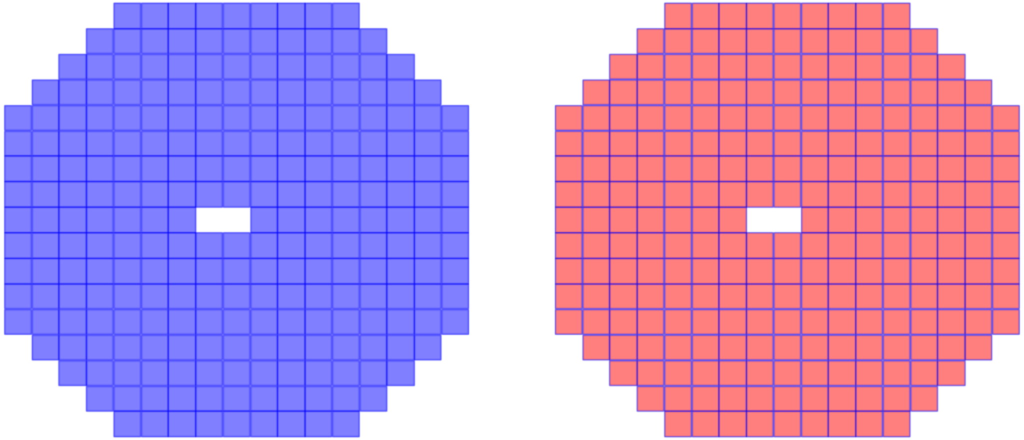


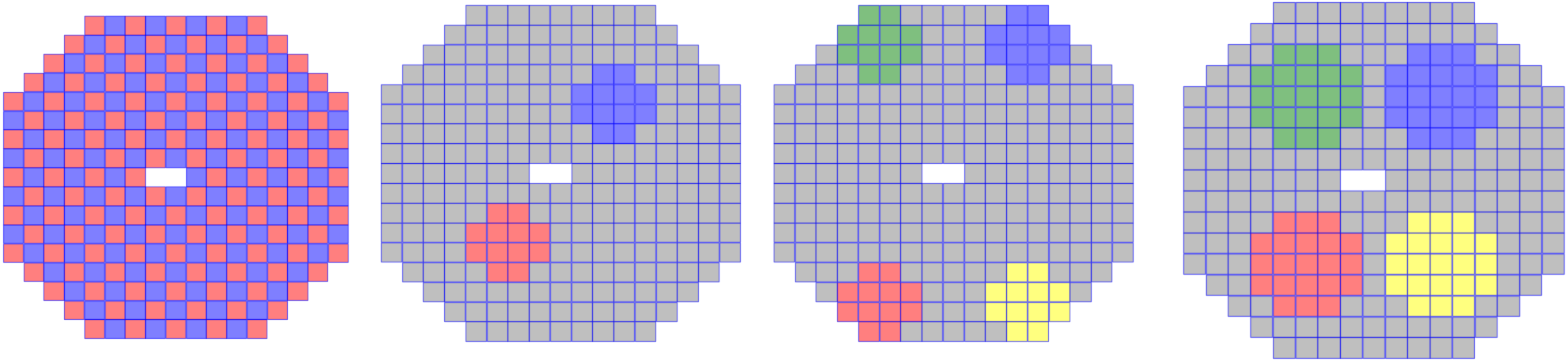


Mirror configurations

- Diversity of mirror configurations:

(which means, diversity of data!)





RAW data: naming convention

- Naming convention is described within the documentation:
  - If you use “recent” data, most of the data will be almost identical (we now tinker less)
  - In any case, always make sure you don’t use weird mirror configurations that may just introduce noise (or problems) to your analysis
  - **Tarek’s hands on**: check evolution of data in the archive

```
Spectrum_XCorr_acc-6C4A-LST1_500MSa_Buff_50Ohm_200mV_5min_Benetnasch_10002_20002_202
```

In this example:

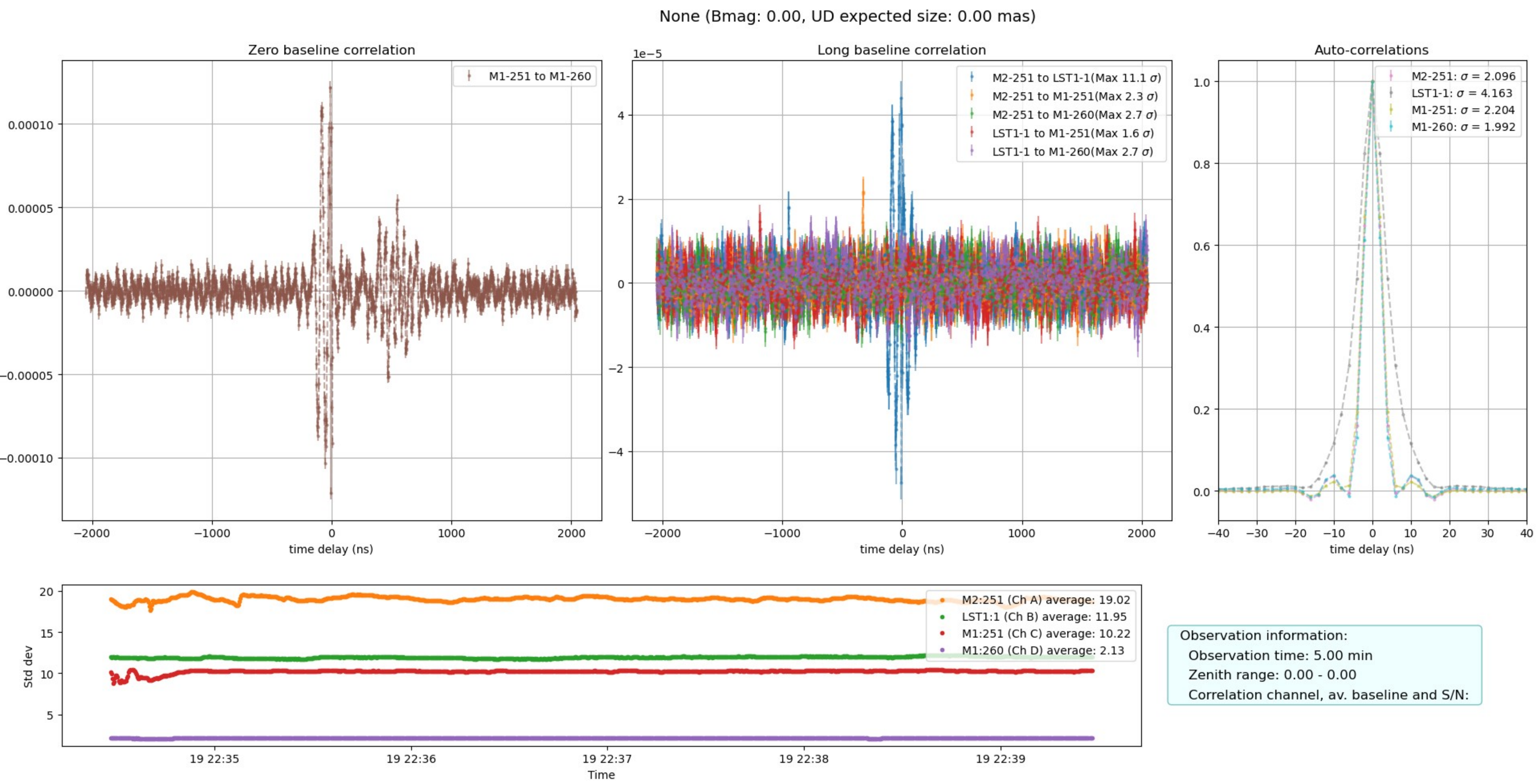
- ‘Spectrum_XCorr’ is the name of the digitizer used and type of data stored in the file. This is the most common type of data, and contains the cross-correlations of the input signals.
- ‘acc’ highlights the fact that correlation arrays are computed and accumulated in time. Those correlation arrays stored are the result of an averaging in time of 1/4th of a second. This dramatically reduces the size of the binary files, as well as the GPU load during acquisition. This approach allowed to use the same server, before used to correlate 2 input signals, to correlate 4 input signals (**Carlos Delgado’s idea!**).
- ‘6C4A’ shows different correlations stored in the file. This tag means these files contain 6 cross-correlations and 4 auto-correlations. In the past there is data with 6 correlations (and no auto-correlations) and even data with just 1 correlation.
- ‘LST1’ indicates that one pixel from LST1 is connected to the system. Be aware this does not mean LST1 is taking data simultaneously, only that one pixel from LST1 is connected via optical fiber to the correlator.
- ‘500MSa’ is the sampling rate of the digitizer, 500 MSa/s.
- ‘Buff’ describes the acquisition mode of the digitizer, which is buffered mode (unnecessary technical detail).
- ‘50Ohm’ is the impedance of the digitizer, 50 Ohm. There exists **1MOhm data that should not be used**.
- ‘200mV’ is the dynamic range of the digitizer (can also be 500mV or 1V). Most of our observations are 200mV, except technical runs that you will very likely never use or care about (lucky you).
- ‘5min’ is the duration of the run, 5 minutes.
- ‘Benetnasch’ is the name of the target that Super Arehucas sends the DAQ. It is used by the software to identify the target and, together with the time stamps, know where the telescopes are pointing at (needed for knowing where in the time-delay array we expect the correlation signals).
- ‘10002_20002’ are the mirror configurations used for M1 and M2 (this naming convention needs to evolve, as now we have more telescopes!). You may find the different mirror configurations used in the MAGIC wiki.
- ‘20240319T223928’ is the date and time of the start of the run, in the format YYYYMMDDTHHMMSS.

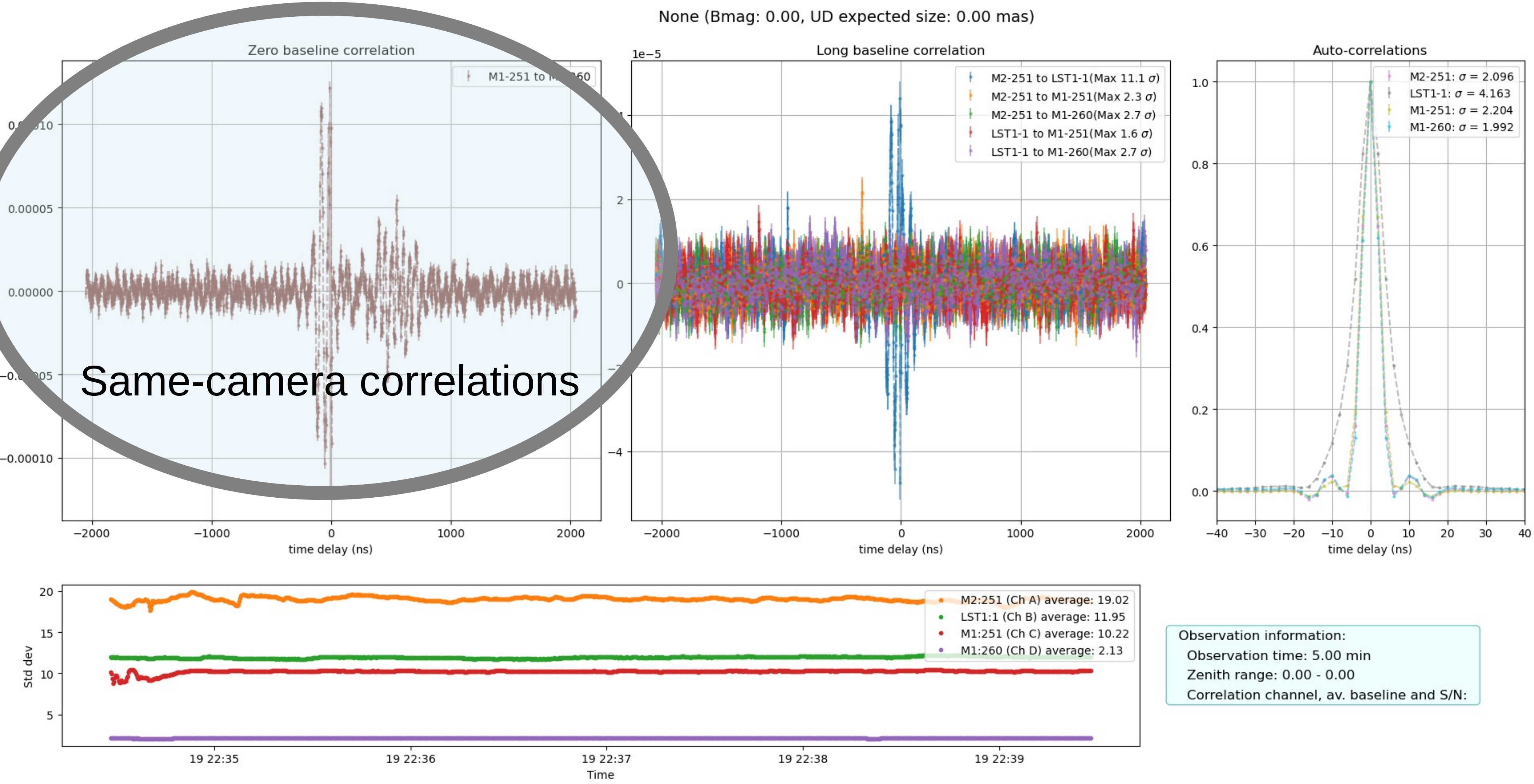
RAW data: check data content

- To access raw data, you may either directly check data within the mics, or copy raw data to your machine

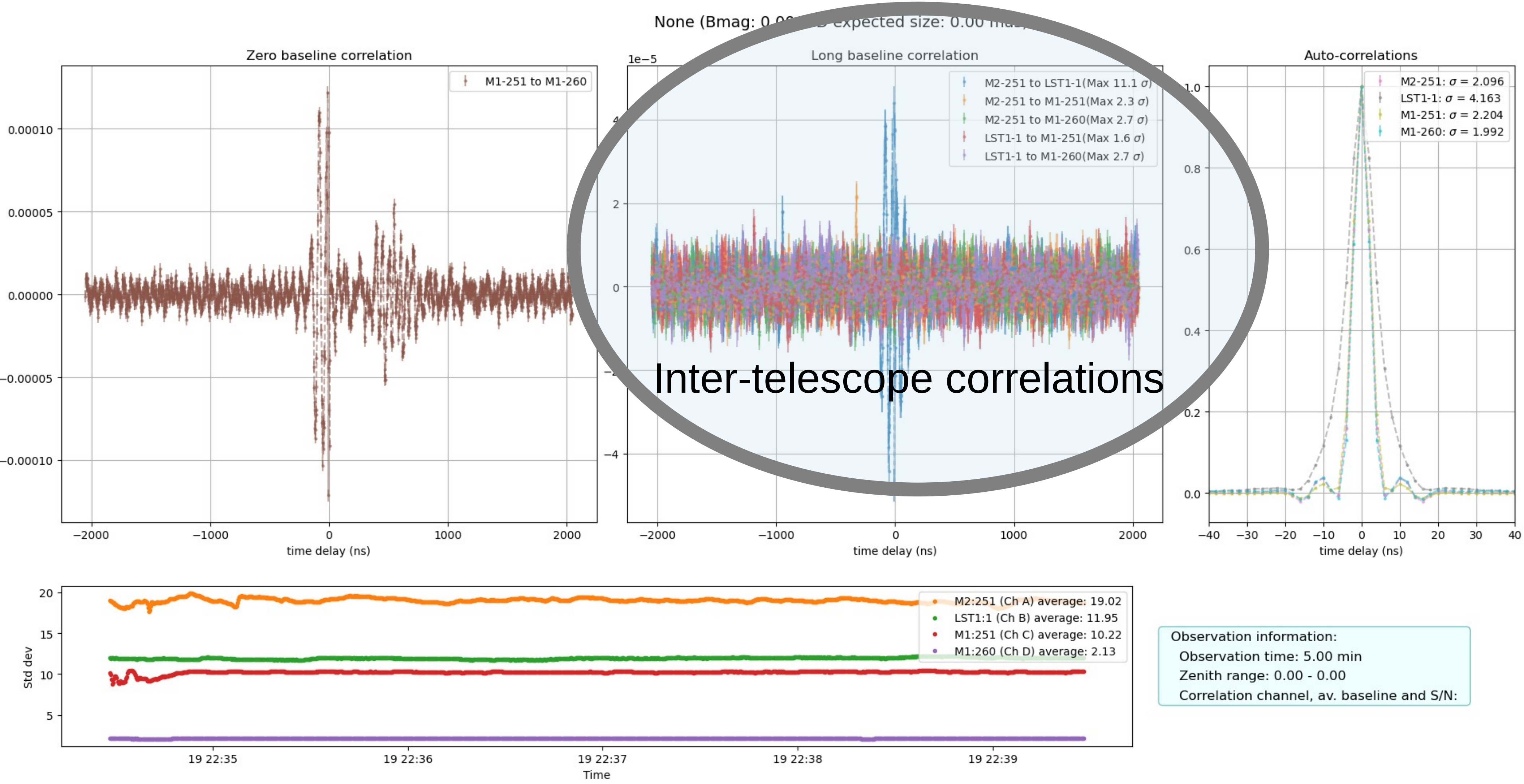
  - Producing interactive plots remotely is generally a pain (slow) but certainly useful in certain situations

  - **Tarek's hands on**: show how one may use "reduce_binaries_6C" on recent data (you may slowly follow documentation later!)

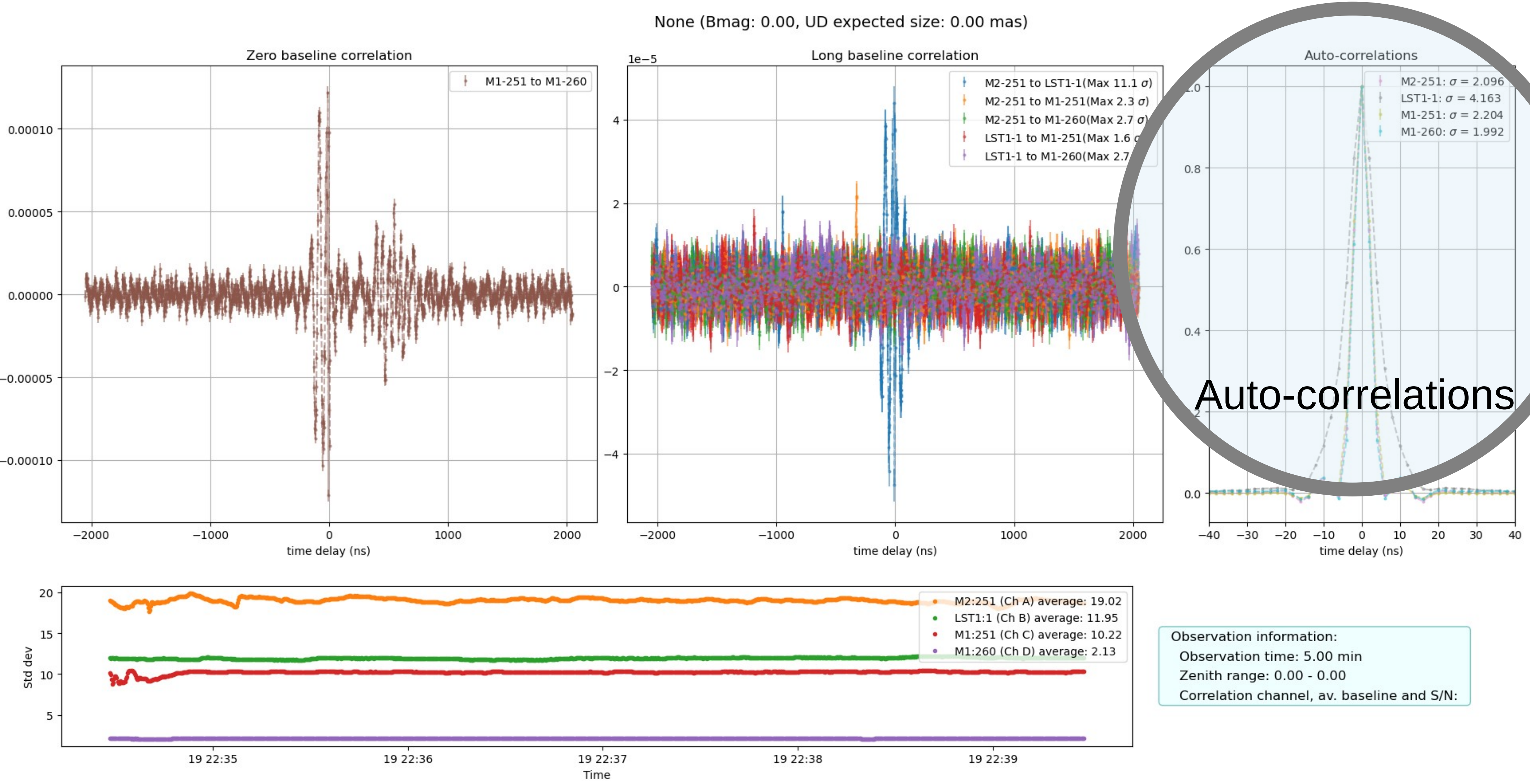

None (Bmag: 0.00, UD expected size: 0.00 mas)
Zero baseline correlation
M1-251 to M1-260
time delay (ns)
Long baseline correlation
1e−5
M2-251 to LST1-1(Max 11.1 σ)
M2-251 to M1-251(Max 2.3 σ)
M2-251 to M1-260(Max 2.7 σ)
LST1-1 to M1-251(Max 1.6 σ)
LST1-1 to M1-260(Max 2.7 σ)
time delay (ns)
Auto-correlations
M2-251: σ = 2.096
LST1-1: σ = 4.163
M1-251: σ = 2.204
M1-260: σ = 1.992
time delay (ns)
Std dev
M2:251 (Ch A) average: 19.02
LST1:1 (Ch B) average: 11.95
M1:251 (Ch C) average: 10.22
M1:260 (Ch D) average: 2.13
19 22:35
19 22:36
19 22:37
19 22:38
19 22:39
Time
Observation information:
Observation time: 5.00 min
Zenith range: 0.00 - 0.00
Correlation channel, av. baseline and S/N:

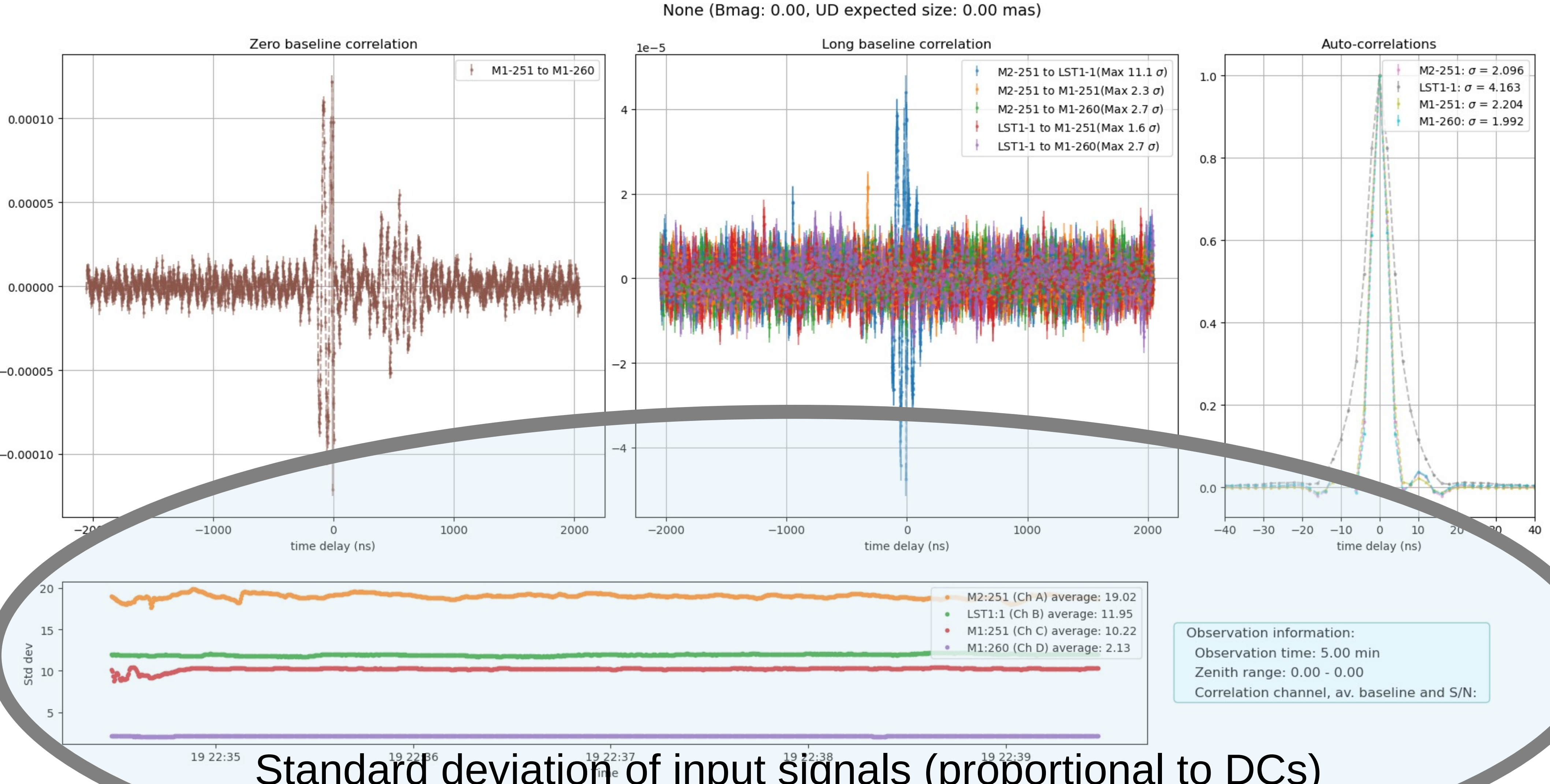


Same-camera correlations

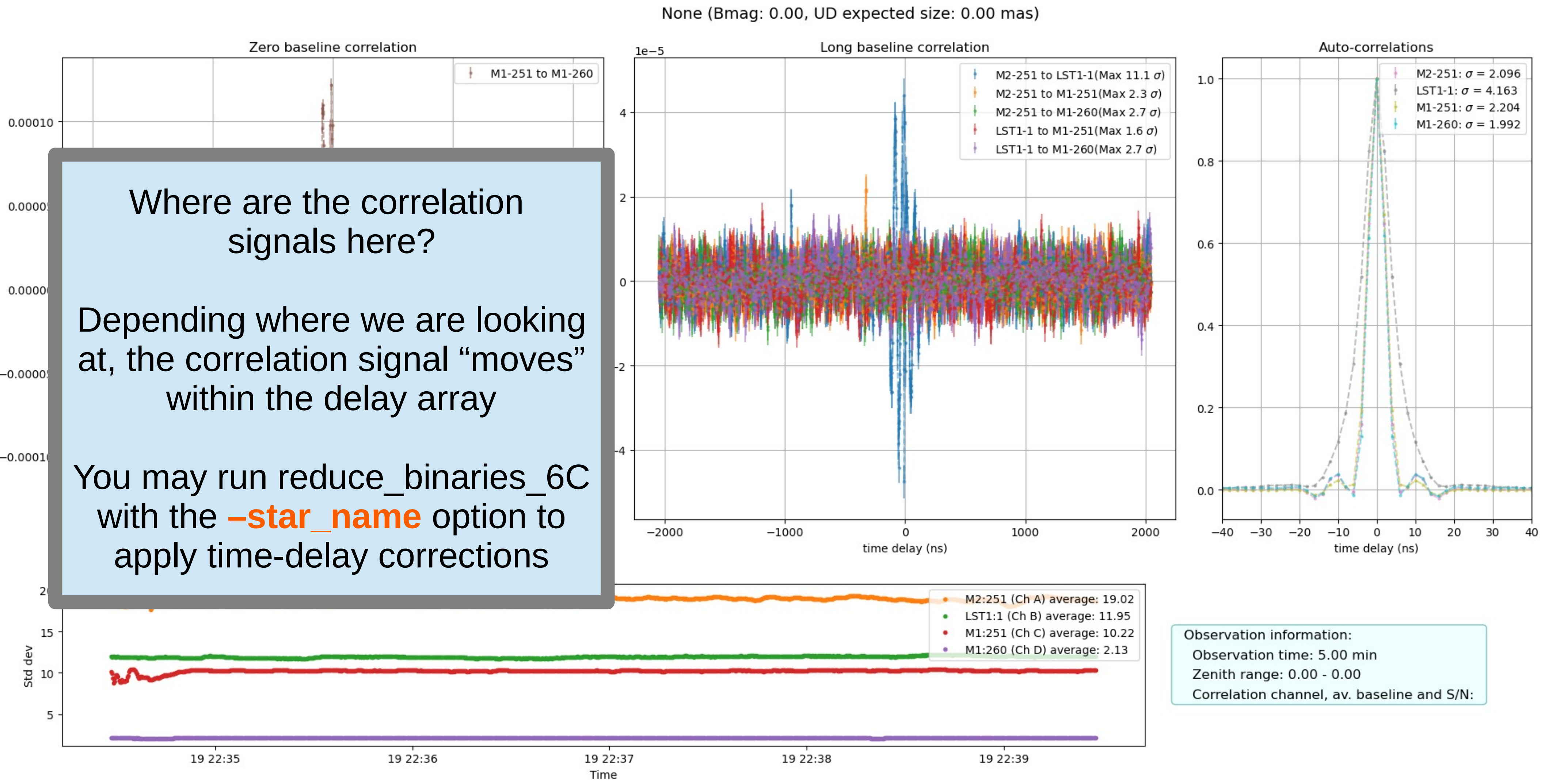


Inter-telescope correlations


None (Bmag: 0.00, UD expected size: 0.00 mas)
Zero baseline correlation
M1-251 to M1-260
time delay (ns)
Long baseline correlation
1e−5
M2-251 to LST1-1(Max 11.1 σ)
M2-251 to M1-251(Max 2.3 σ)
M2-251 to M1-260(Max 2.7 σ)
LST1-1 to M1-251(Max 1.6 σ
LST1-1 to M1-260(Max 2.7
time delay (ns)
Auto-correlations
M2-251: σ = 2.096
LST1-1: σ = 4.163
M1-251: σ = 2.204
M1-260: σ = 1.992
time delay (ns)
Auto-correlations
M2:251 (Ch A) average: 19.02
LST1:1 (Ch B) average: 11.95
M1:251 (Ch C) average: 10.22
M1:260 (Ch D) average: 2.13
Std dev
Time
Observation information:
Observation time: 5.00 min
Zenith range: 0.00 - 0.00
Correlation channel, av. baseline and S/N:



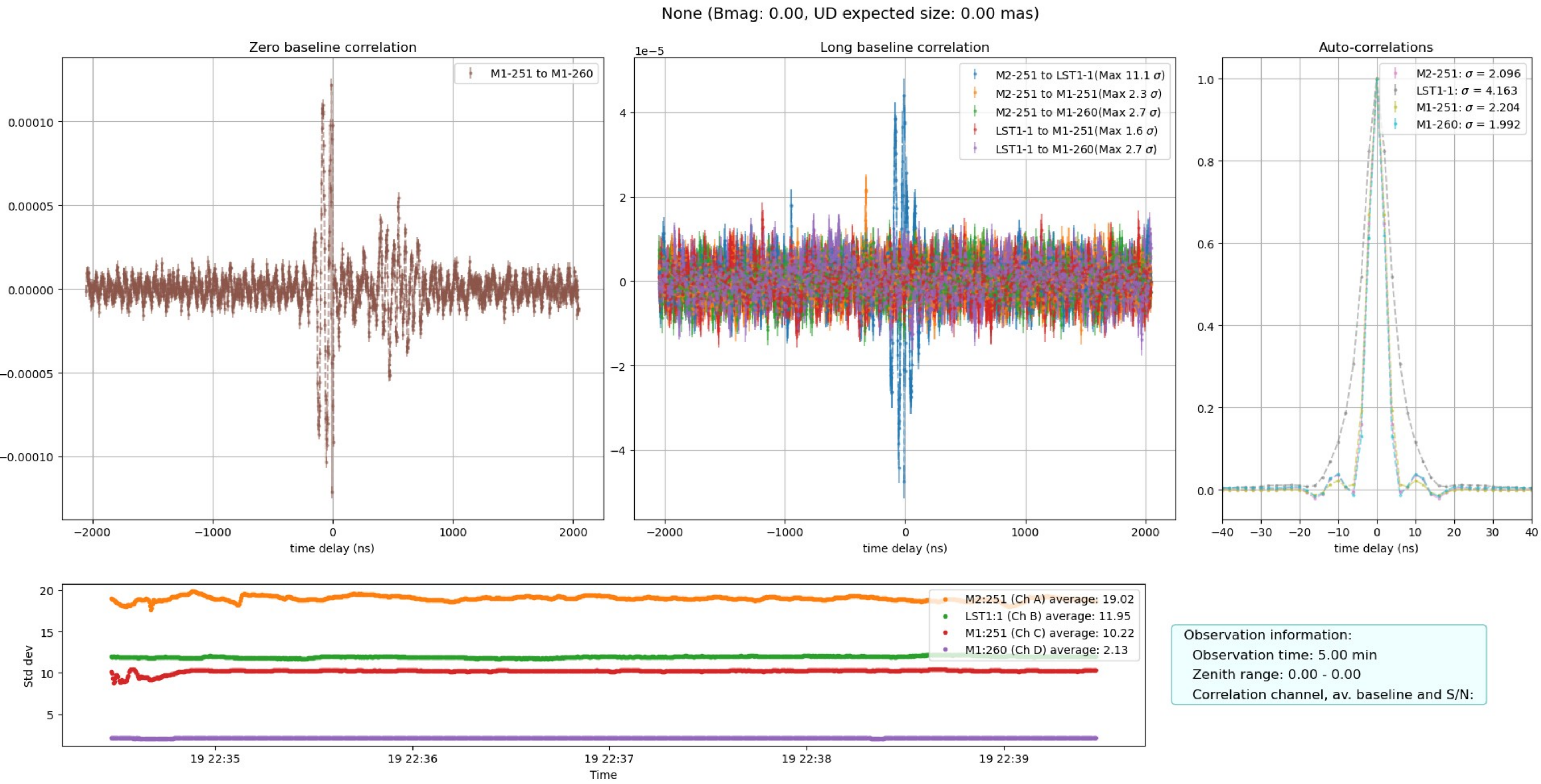
Standard deviation of input signals (proportional to DCs)



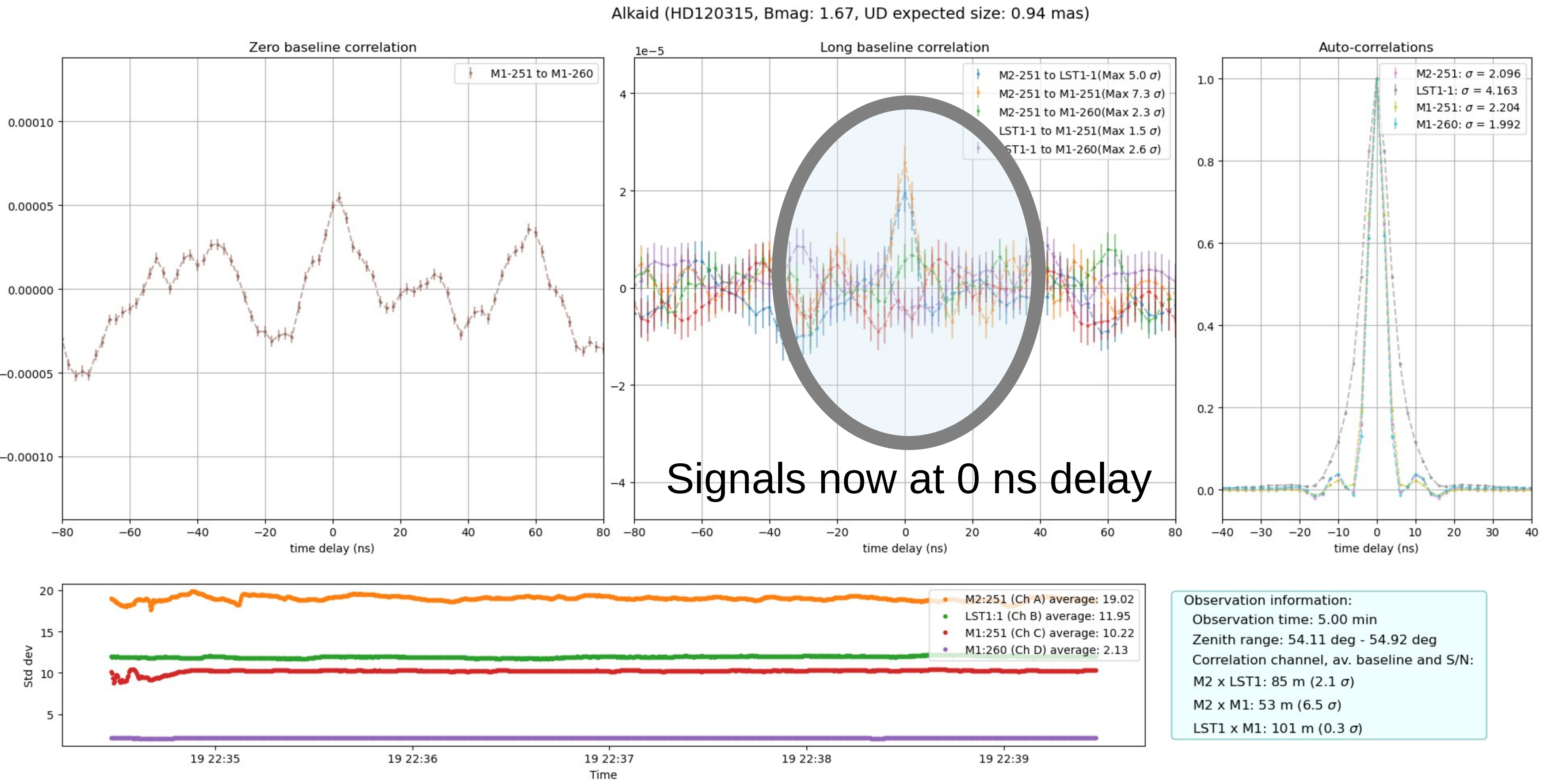
Where are the correlation signals here?

Depending where we are looking at, the correlation signal "moves" within the delay array

You may run reduce_binaries_6C with the **–star_name** option to apply time-delay corrections



None (Bmag: 0.00, UD expected size: 0.00 mas)
Zero baseline correlation
M1-251 to M1-260
time delay (ns)
Long baseline correlation
1e−5
M2-251 to LST1-1(Max 11.1 σ)
M2-251 to M1-251(Max 2.3 σ)
M2-251 to M1-260(Max 2.7 σ)
LST1-1 to M1-251(Max 1.6 σ)
LST1-1 to M1-260(Max 2.7 σ)
time delay (ns)
Auto-correlations
M2-251: σ = 2.096
LST1-1: σ = 4.163
M1-251: σ = 2.204
M1-260: σ = 1.992
time delay (ns)
Std dev
M2:251 (Ch A) average: 19.02
LST1:1 (Ch B) average: 11.95
M1:251 (Ch C) average: 10.22
M1:260 (Ch D) average: 2.13
Time
Observation information:
Observation time: 5.00 min
Zenith range: 0.00 - 0.00
Correlation channel, av. baseline and S/N:



Alkaid (HD120315, Bmag: 1.67, UD expected size: 0.94 mas)
Zero baseline correlation
M1-251 to M1-260
time delay (ns)
Long baseline correlation
1e−5
M2-251 to LST1-1(Max 5.0 σ)
M2-251 to M1-251(Max 7.3 σ)
M2-251 to M1-260(Max 2.3 σ)
LST1-1 to M1-251(Max 1.5 σ)
LST1-1 to M1-260(Max 2.6 σ)
Signals now at 0 ns delay
time delay (ns)
Auto-correlations
M2-251: σ = 2.096
LST1-1: σ = 4.163
M1-251: σ = 2.204
M1-260: σ = 1.992
time delay (ns)
Std dev
M2:251 (Ch A) average: 19.02
LST1:1 (Ch B) average: 11.95
M1:251 (Ch C) average: 10.22
M1:260 (Ch D) average: 2.13
Time
Observation information:
Observation time: 5.00 min
Zenith range: 54.11 deg - 54.92 deg
Correlation channel, av. baseline and S/N:
M2 x LST1: 85 m (2.1 σ)
M2 x M1: 53 m (6.5 σ)
LST1 x M1: 101 m (0.3 σ)

RAW data: check data content

- To access raw data, you may either directly check data within the mics, or copy raw data to your machine

    - Producing interactive plots remotely is generally a pain (slow) but certainly useful in certain situations.

    - **Tarek's hands on**: show how one may use "reduce_binaries_6C" on recent data (you may slowly follow documentation later!)

    - For those that want to go deeper into raw data, you may check: **notebooks/check_raw_data.ipynb**

Data synchronization: First step of every analysis

- The way interferometry analysis goes is:
  - Download/sync data
    - Copy to your $MAGICSIIDATA folder all data that you want to analize (binary files + reports)
    - Reduce binary files (*.bin) into pickle files (*.pkl)
  - Perform the analysis:
    - Pick a star, gather all data runs from that star (within $MAGICSIIDATA), calibrate them (using camera reports), combine them and reconstruct the diameter

Data synchronization: First step of every analysis

- The way interferometry analysis goes is:

sync_magic_sii_data

- Download/sync data
    - Copy to your $MAGICSIIDATA folder all data that you want to analize (binary files + reports)
    - Reduce binary files (*.bin) into pickle files (*.pkl)
- Perform the analysis:
    - Pick a star, gather all data runs from that star (within $MAGICSIIDATA), calibrate them (using camera reports), combine them and reconstruct the diameter

Data synchronization: First step of every analysis

- What does ***sync_magic_sii_data*** allow you to do?
  - Check data in the archive **not already present** in $MAGICSIIDATA
    - Bad nights (within magic_spysii/data/night_reject_list.txt) are automatically ignored
  - Sync (sym links) or download (copy) data (*.bin & *.dc) to $MAGICSIIDATA
    - Copying is needed if working from remote, sym links are possible only at PIC. This step may be performed via condor only at PIC.
  - Reduce data

Data synchronization: First step of every analysis

- **Hands on by Tarek** (remember, all this is explained in the documentation):
  - Show the structure of $MAGICSIIDATA and its content
  - Show how to check available data in the archive
  - Show how to synchronize a night
  - Show how to synchronize and reduce a night
  - Quick check: Once you have data from one night, one can take a look to the DC evolution of the night with ***check_night_reports***

IACT-SII: Lesson I wrap up

- Binary files contain signal statistics plus correlation arrays (over a broad delay range)

- Following *magic_spysii* documentation you should now be able to:

  - Download/synchronize raw data and camera reports into your $MAGICSIIDATA folder
  - Plot the content of individual or multiple “averaged” runs, and see clear correlation signals
  - Draw the DCs and Hvs of any synced night
  - Tinker with binary data, understand how it is stored and what kinds of binary data exist

IACT-SII: Lesson I – Proposed exercises

- Download/synchronize a couple of 2024 nights in which **Adhara** was observed
  - Draw (using reduce_binaries_6C) some runs and see if you understand the figures
  - Draw an average of the correlation signal combining several runs (--merge option). What happens when you add a decent amount of runs?
- Tinker with a single raw file:
  - Are you able to plot **a single** correlation array? (1/4th of a second integration)


BACKUP

ZBC: The constant of our system

- From HB&T, we know that the expected correlation:

$$\overline{c(d)} = \langle \Delta i_1(t) \Delta i_2(t) \rangle = e^2 A^2 \alpha^2 n^2 |\gamma_d(0)|^2 \Delta\nu \Delta f \qquad (4.28)$$

- After dividing by the flux, the remaining correlation should be:

$$\frac{\overline{c(d)}}{flux} \propto \frac{|\gamma_d(0)|^2}{\Delta \nu \Delta f}$$

- But as with our setup, our "flux" is the DCs (different gain than the correlation):

$$\frac{\overline{c(d)}}{\sqrt{DC^1 DC^2}} \propto \frac{|\gamma_d(0)|^2}{\sqrt{G^1{}_{DC} G^2{}_{DC}} \Delta \nu \Delta f}$$

Any time evolution in the DC gain will be a **systematic**