AGATA Data Processing
Adapted from D. Bazzacco 2012 EGAN school lecture

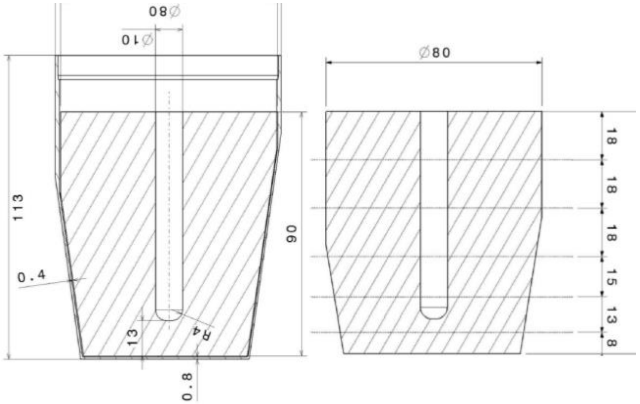
2018 –NEDA campaign

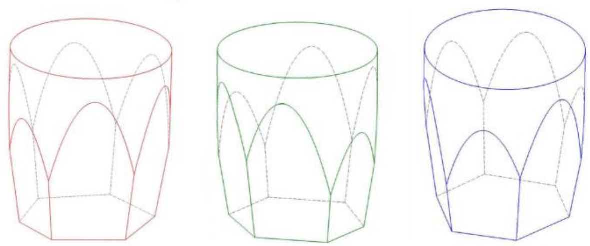
E.Clément

Topics

- General structure of data acquisition
- Data processing model == data analysis model
- Narval (ADA/C/C++) and actors ➔ emulators
- Data flow ➔ adf
- Survey of actors

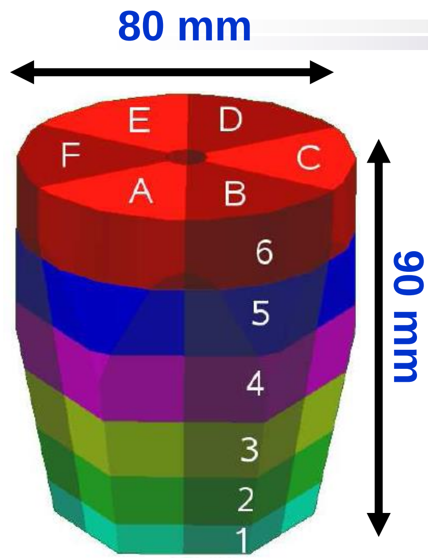
AGATA detectors





(d) A - red (e) B - green (f) C - blue

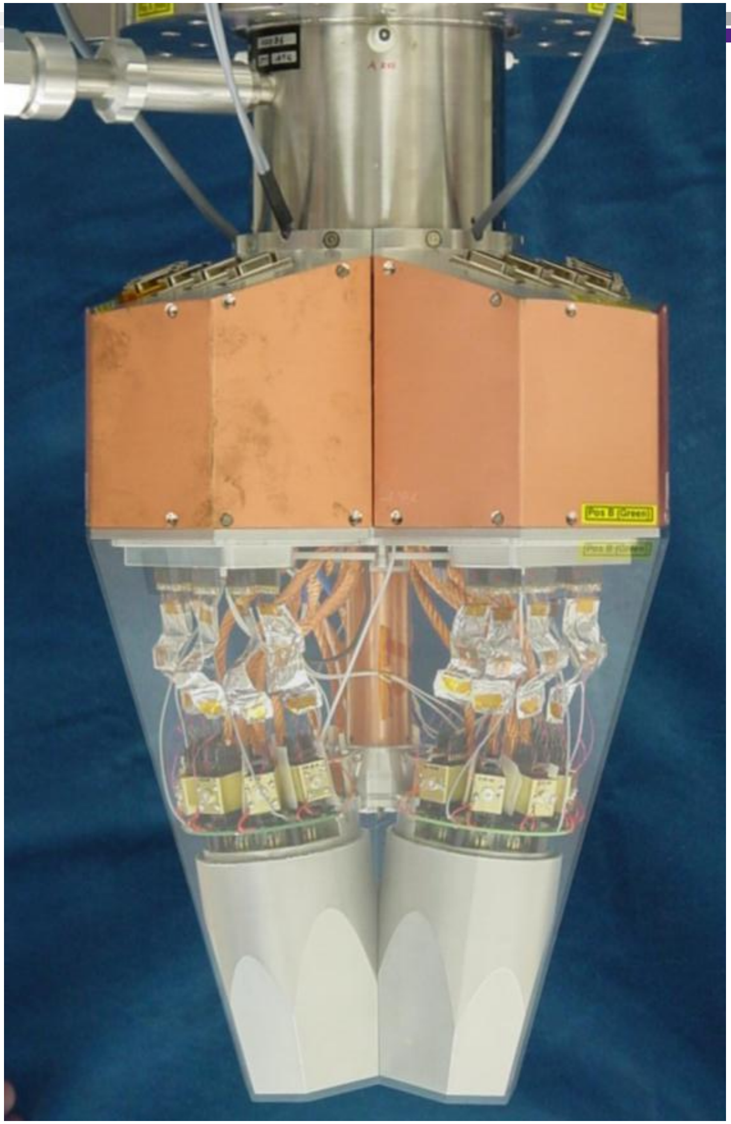
Volume ~370 cc   Weight ~2 kg
(shapes are volume-equalized to 1%)



6x6 segmented cathode
Cold FET for all signals

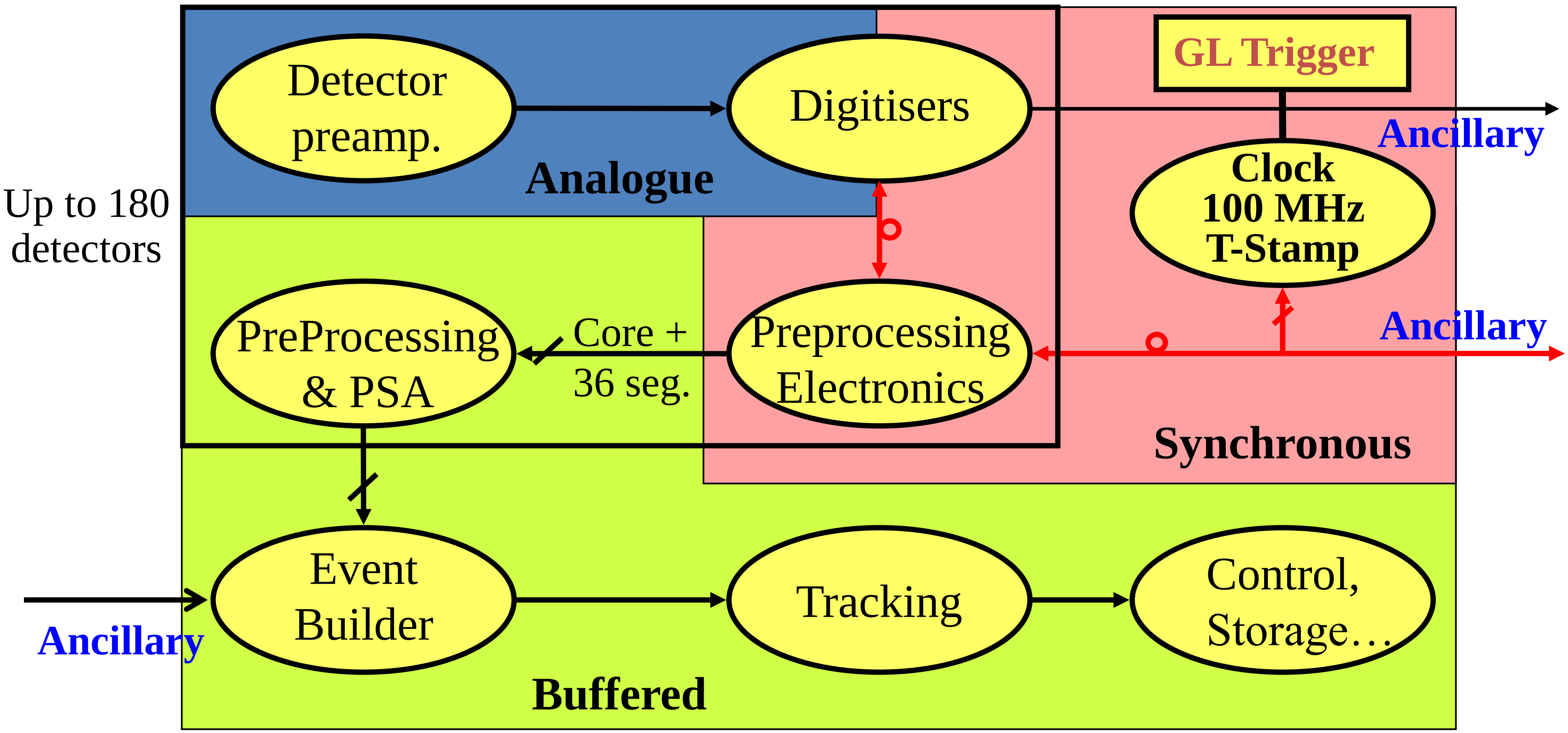
Energy resolution
Core: 2.35 keV
Segments: 2.10 keV
(FWHM @ 1332 keV)
A. Wiens et al. NIM A 618 (2010) 223
D. Lersch et al. NIM A 640(2011) 133

38 high-resolution signals / detector



AGATA Triple/Double Cryostat
Manufactured by CTT

Structure of Electronics and DAQ



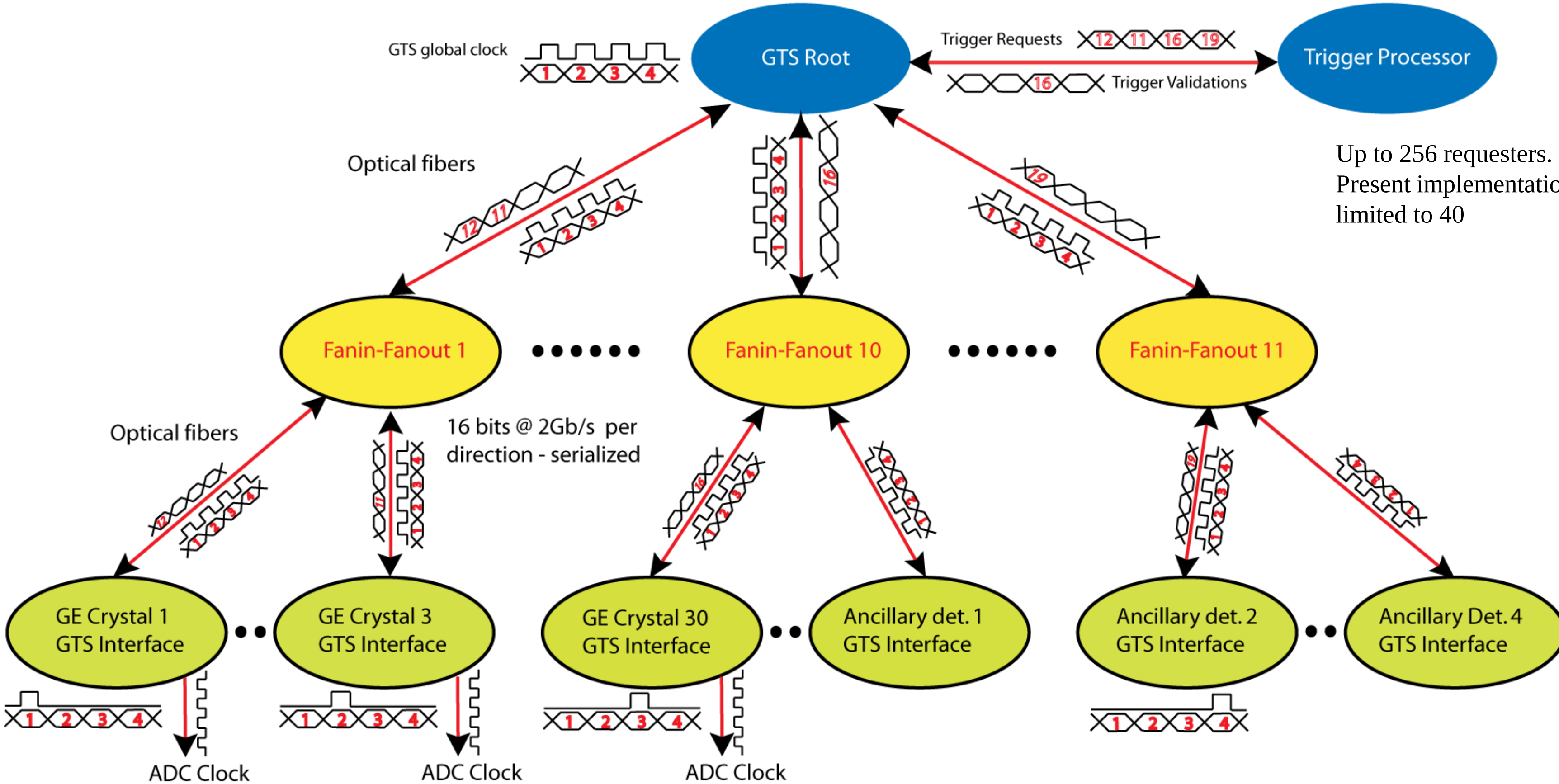
Other detectors

1. interface to GTS via mezzanine
2. merge time-stamped data into event builder (merger)

GTS : the system coordinator

All detectors operated on the same 100 MHz clock

Downwards	100 MHz clock + 48 bit Timestamp (updated every 16 clock cycles)
Upwards	trigger requests, consisting of address (8 bit) and timestamp (16 bit)
	max request rate 10 MHz total, 1 MHz/detector
Downwards	validations/rejections, consisting of request + event number (24 bit)



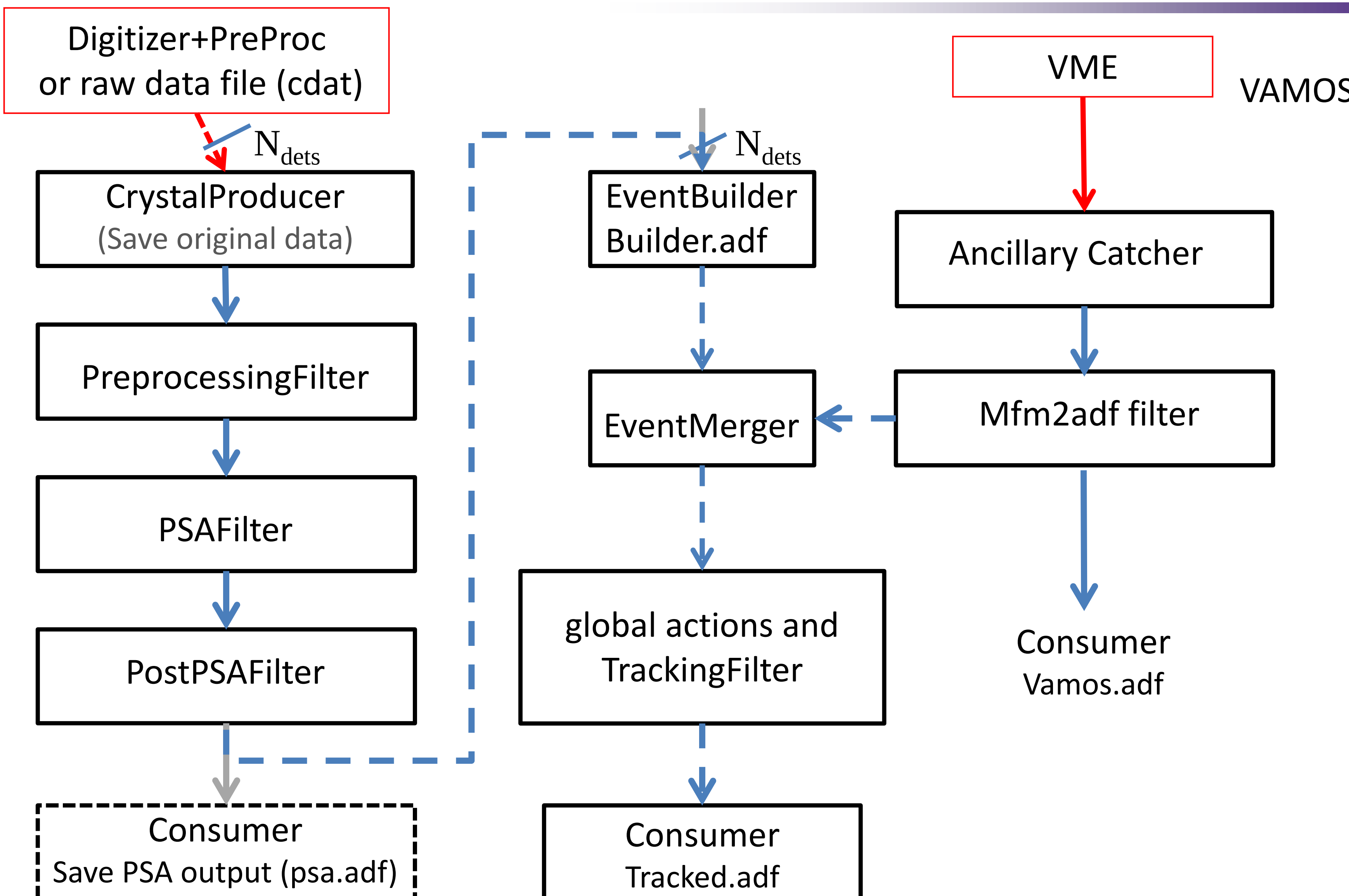
Up to 256 requesters.
Present implementation limited to 40

Architecture of the system

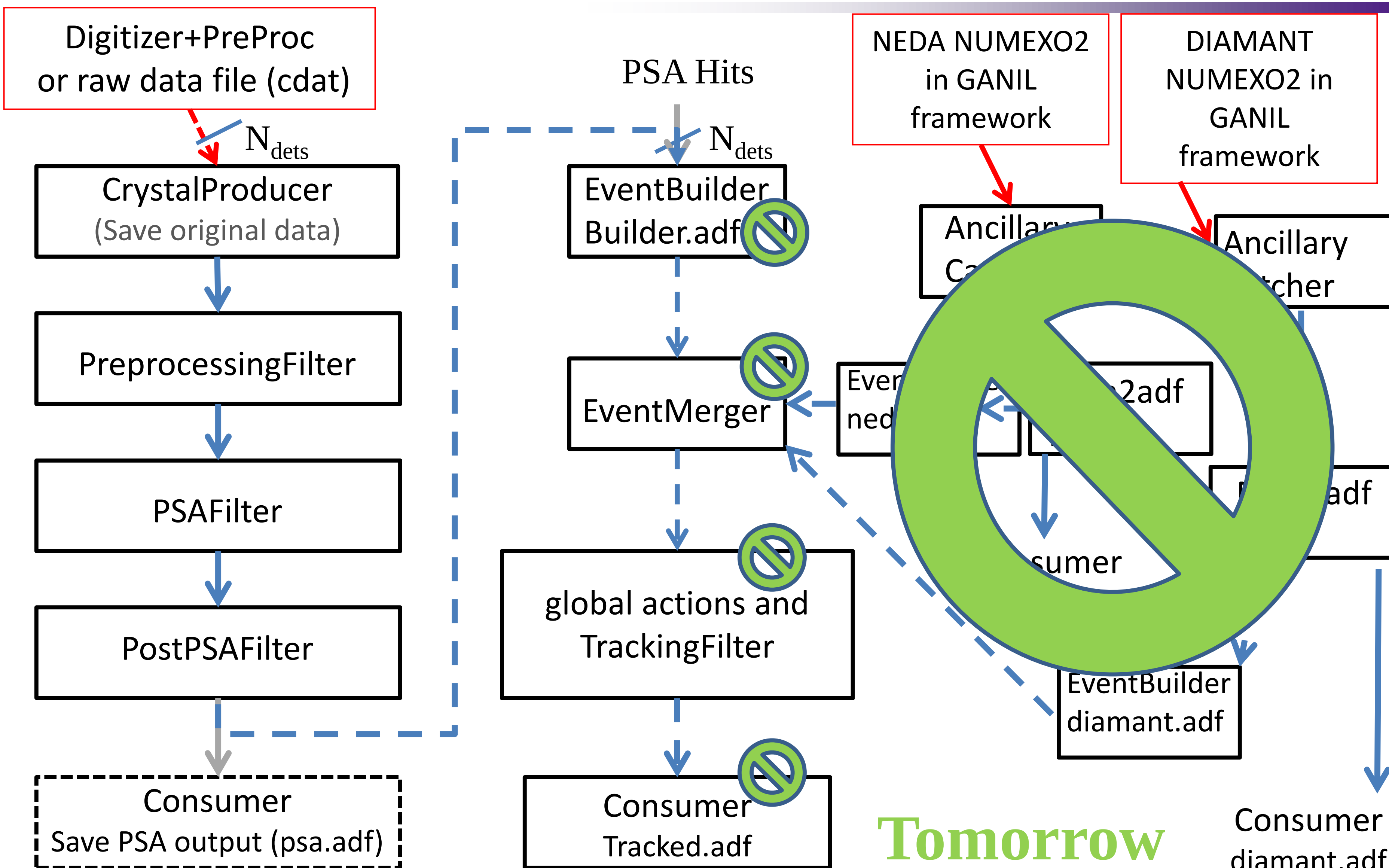
- **Local Level :** **where the individual detectors don't know of each other.**
  - Electronics and computing follow a model with minimum coupling among the individual detectors, which are operated independently as long as possible
  - Electronics is almost completely digital, operated on the same 100 MHz clock
  - Data processing (in the electronics and in the front-end computers) is the same for all detectors and proceeds in parallel
  - **Every chunk of produced data is tagged with a 48 bit number (time stamp) giving the absolute time (with a precision of 10 ns) since the last hard-reset of the system ➔ roll around takes place every 32.5 days.**

00A 00B 00C 01A 01B 01C….

PSA Hits



00A 00B 00C 01A 01B 01C….



Data processing and replay

- Offline = Online
- A series of programs organized in the style of Narval actors
- Director of the actors
  - **Online** : **Narval (now DCOD)**
  - **Offline** : narval emulator(s)
    - Depending on the available computing resources
    - Single computer or Farm with distributed processing
- The system is complex and “difficult” to manage
  - Very large number of channels NumberOfCrystals*35 1330 Channels
  - No chance to take care of them individually
  - Rely on automatic procedures
  - Hope that the system is stable
  - Result depends on average performance

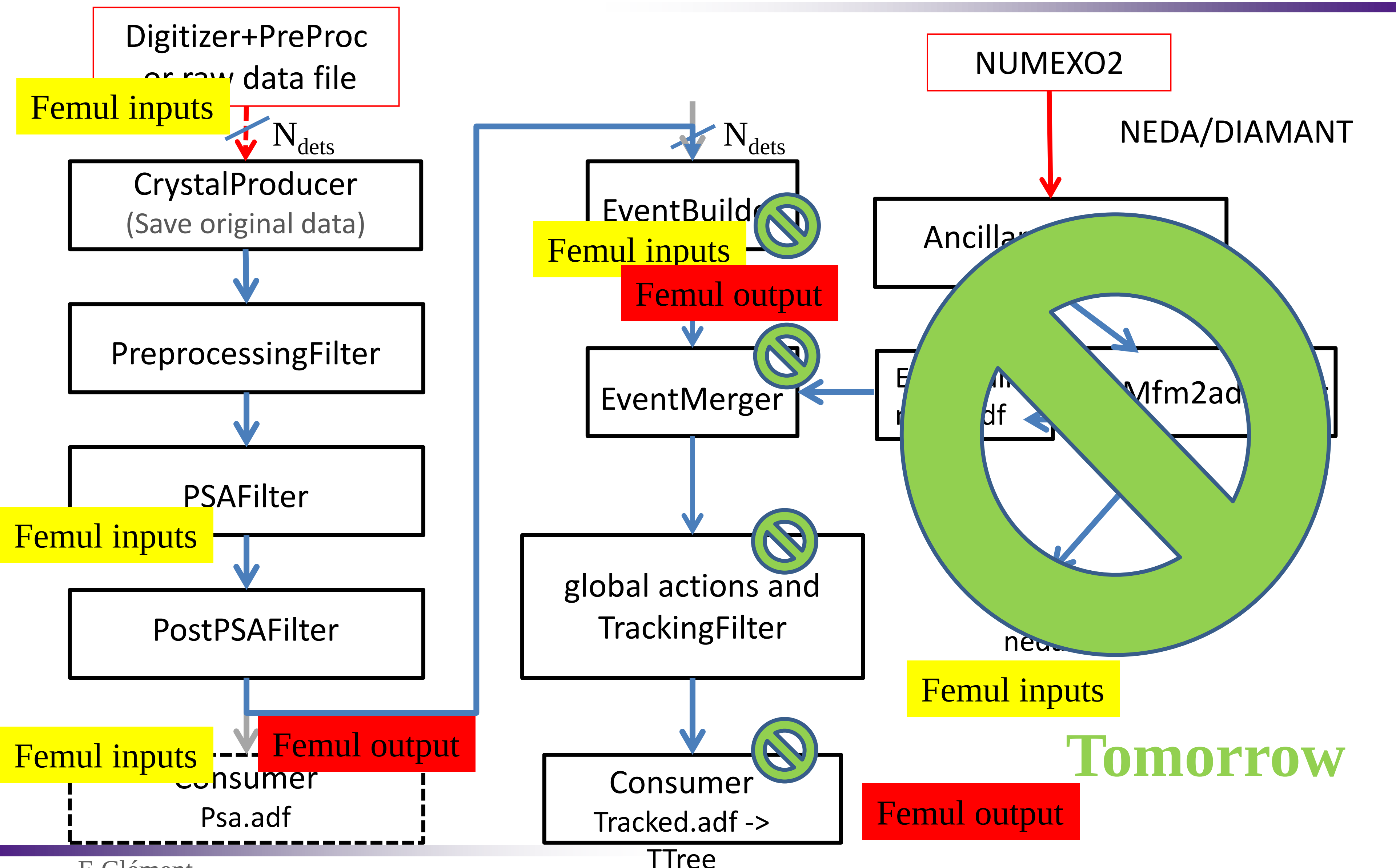
Use of actors, managed by the framework (DCOD or femul)

- Creation
  - process_config and object creation …
- **process_initialise**
  - Read parameters from confPath/Actor.conf (e.g. Conf/00A/PSAFilter.conf)
  - Generate configuration-dependent internal objects, open data files
- process_start (Important for Producers/Consumers)
- **process_block (or adf::ProcessBlock) 🡸 the event loop**
  - **where the actual work is done**
- process_stop (Important for Producers/Consumers)
- Destruction
  - Save spectra and close files
- A fair amount of repeated code (boilerplate) saved by inheriting from proper base classes (NarvalInterface, NarvalProducer, NarvalFilter, NarvalConsumer provided by the adf library.

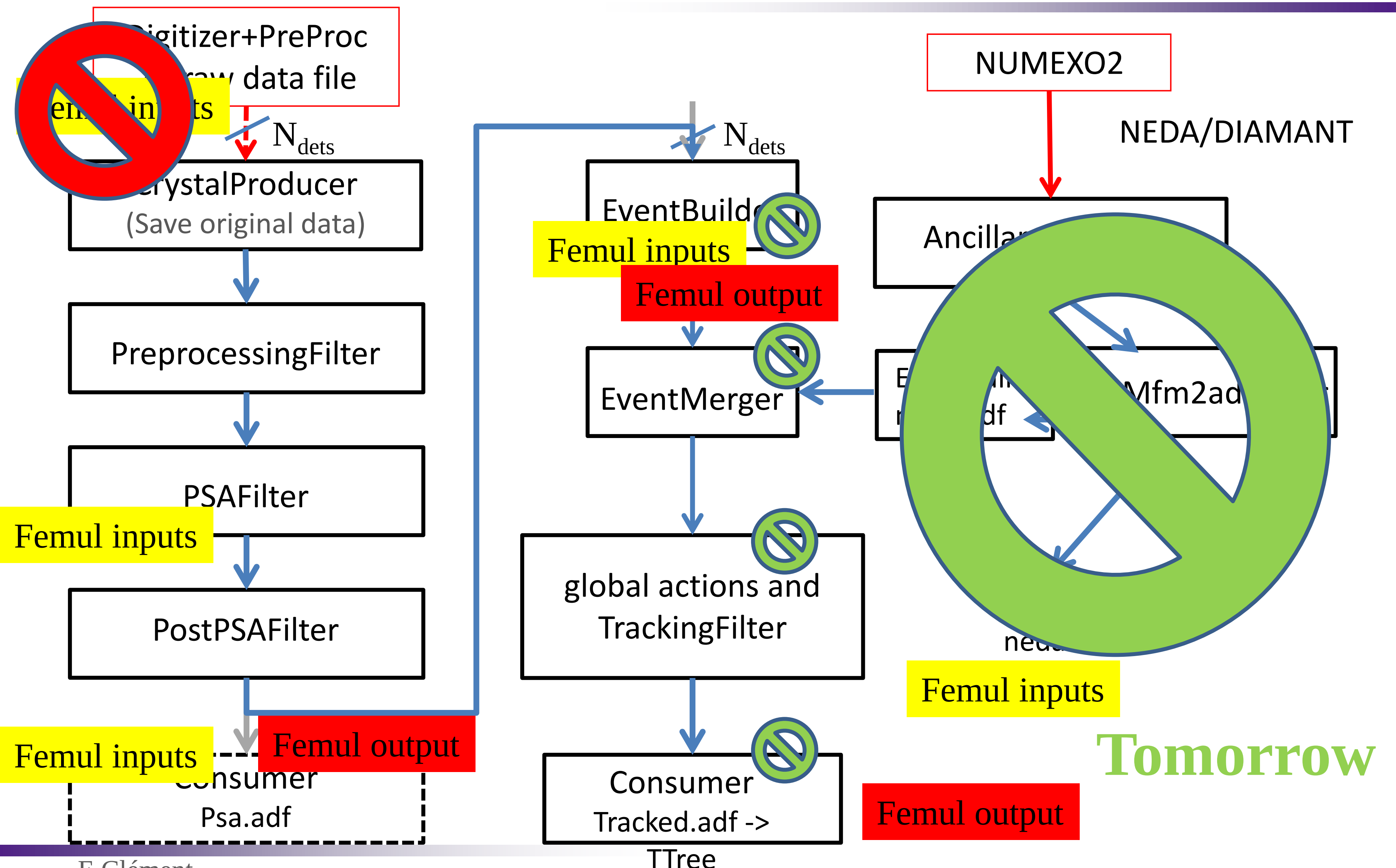
Emulator(s)

- femul
  - Originally intended to help developing and debugging the user libraries which is very hard to do in a distributed processing environment like Narval.
  - Gradually developed as a full emulation of the Narval framework with the limitation of being a single process running a specific machine (threads used to distribute the work on the available cores).
  - Configuration (detectors, actors, ..) specified by a “topology” file
  - Generation of configuration parameters via gen_conf.py

00A 00B 00C 01A 01B 01C....



00A 00B 00C 01A 01B 01C….



Structure of analysis directories

- The directory where you produce your data (e.g. /agatadisks/exptname (EXXX) /(Config EXXX)/run_XXX_date) contain some standard sub-directories
  - **Conf** → configuration of actors, calibrations, ... for each detector

    00A, 00B, 00C ... 14B, 14C Ancillary(ies) Global

    with minimal differences between online and offline
  - **Data** → data and spectra produced during the experiment

    00A, 00B, 00C ... 14B, 14C Ancillary(ies) Global

    Online writes data here
    Offline replay takes data from here
  - **Out** → data and spectra produced during data replay

    00A, 00B, 00C ... 14B, 14C Ancillary(ies) Global

    Offline writes data here

Conf, Data and Out are often symbolic links to actual directories

Typical configuration directories

Conf/12A

- CrystalProducer.conf
- CrystalProducerATCA.conf
- PreprocessingFilter.conf
- PreprocessingFilterPSA.conf
- PSAFilter.conf
- PostPSAFilter.conf
- xdir_1325-1340.cal
- xinv_1325-1340.cal
- BasicAFC.conf
- BasicAFP.conf

Conf/Global

- EventBuilder.conf
- EventMerger.conf
- TrackingFilter.conf
- CrystalPositionLookUpTable
- BasicAFC.conf

Binary spectra

- Simple C-style multidimensional (max 6) arrays written mostly in binary format
- For historical reasons the format is not recorded in the file.
- Often written as part of the file name:
  - Prod__4-38-32768-UI__Ampli.spec
    Is a file dump of an array defined as
    unsigned integer Ampli[4][38][32768]
    containing 4*38 = 152 spectra of 32768 channels
  - No difference between spectra and matrices; the type is only an hint to how to interpret them
- The viewers TkT and Mat can decode and interpret the format. The user can always override the program.
- Other programs (e.g. RecalEnergy) can interpret the spectrum length and type but the user has to specify the number of spectra to act upon.

Some Useful programs

- TkT spectrum viewer (& al) tailored to composite detectors
  - A nightmare concerning programming-style, but contains ~all I need(ed) to analyze AGATA data. Virtually no documentation.
- RecalEnergy
  - Analysis of spectra looking for peaks
- SortPsaHits
  - Sort of PSA hits (special format) to determine neutron damage correction parameters
- gen_conf.py
  - Unified procedure to produce configuration files for all actors
- solveTT.py
  - Optimize time alignment of "equal" detectors

Local Actors

- Producers
  - CrystalProducer
    - Readout of electronics, or get raw data from file
    - Local event builder
    - Save original data to be able to replay experiment
    - Raw projections
- Filters
  - PreprocessingFilter ➔ PreprocessingFilterPSA
    - Energy calibrations
    - Retrigger and Time calibration
    - Cross talk correction
    - Amplitude calibration and time alignment of traces
    - Improved pile-up rejection
  - PSAFilter ➔ PSAFilterGridSearch
    - Decomposition of calibrated experimental traces by comparison with a calculated signal basis
    - In principle more than one algorithm available but only one used in practice.
  - PostPSAFilter
    - Neutron deficit corrections
    - Recalibrations of energy and time
    - *Smearing of positions (not recommended)*
- Consumer
  - Save PSA hits for "Global Level"-only processing

Done

Global Actors

- Event Builder
  - Essentially a producer with multiple inputs
  - Assemble event fragments from Ge detector der) a
    Auxiliary detectors (Merger)
- Filters
  - TrackingFilter
    - Global histograms (time, energy …)
    - Grouping and further filtering
    - Format data as needed by the specific Tracking
    - Histograms of tracked gammas
- Consumer
  - Write tracked data

Tomorrow

CrystalProducer

- Reads the data from:
  - The PCI express driver connected to the ATCA/GGP electronics
  - Raw data files (event_mezzdata.cdat) for the offline (

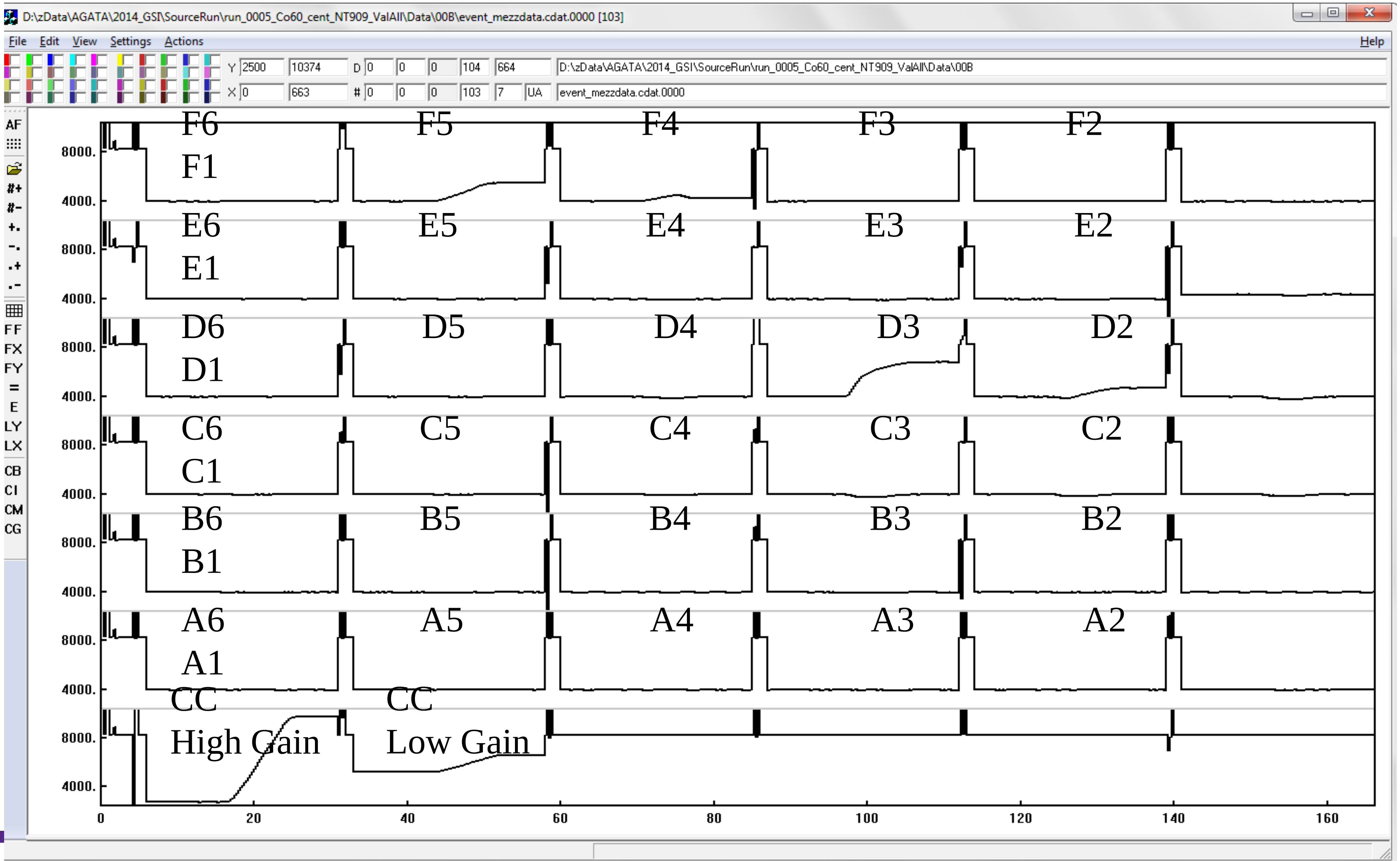
- Acts as a local event builder to check and merge the sub-events coming from the ATCA/GGP readout (or from the raw data file)
  - For the online has to work in nonblocking-mode to avoid blocking the ACQ
  - This feature was implemented with boost::threads
  - Damiano has now implemented the “select” and threads have been removed
  - The local event builder and the management of the mezzanine data should be completely rewritten …
- Can write the original data with the format of the ATCA/GGP electronics, and various other formats (e.g. only the energies for calibrations)
- Prepares data:crystal frames, without using adf (version using adf exists, maybe slower)

- Not much to do for the users both for Data Acquisition and Data Replay

What does it read and write

- Raw event: data from 7 mezzanines hosted in two ATCA Carrier Boards or data from the single GGP per anode
- length (for 100 samples traces): 16+(8+100)*6 = 664 short (2 bytes) words
- **Event length: 7*664 = 9296 bytes**
- **Local event builder** to assemble data from 2 readout ports, according to mapping specified in CrystalProducerATCA.conf contains the mapping
- Write original data (event built, compressed→ 3.6 kB)
  - data files typically split after **1 M events ~ 3.6 GB**
  - event_mezzdata.cdat.0000 event_mezzdata.cdat.0001 …
- Generate raw spectra for amplitudes and baselines
- Format data into an **data:crystal** adf frame and send it to the data flow.

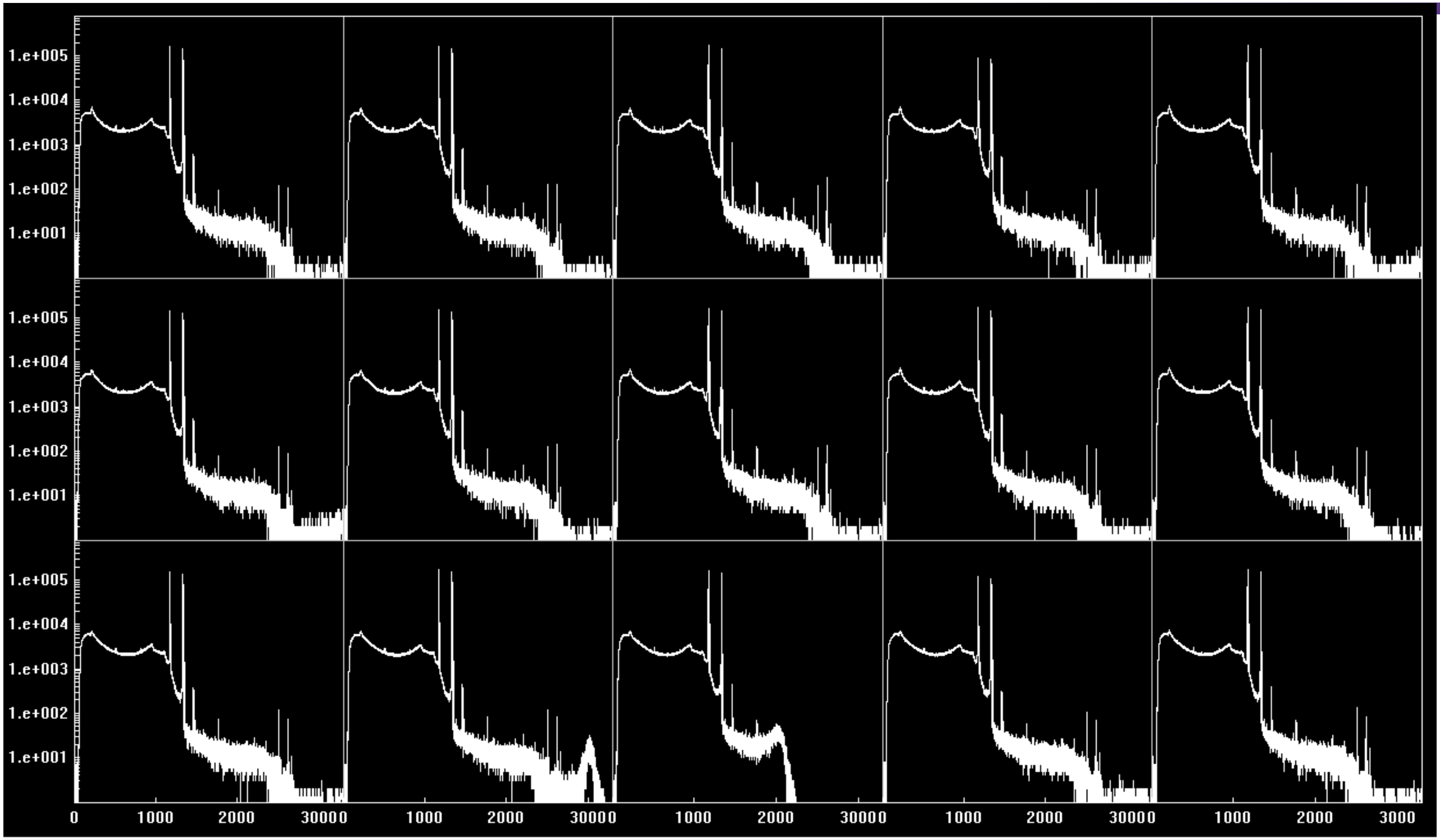
Raw data from EDAQ



Digitizers

- 100 MS/s
  - max frequency correctly handled is 50 MHz (Nyquist)
- 14 bits
  - Effective number of bits is ~12.5 (SNR ~75 db)
- 2 cores (one board) and 36 segments (in 6/3 boards)
- Core 2 ranges 5, 20 MeV nominal
- Segments taken either at high gain (5 MeV) or at low gain (20 MeV)

Range of digitizers 1

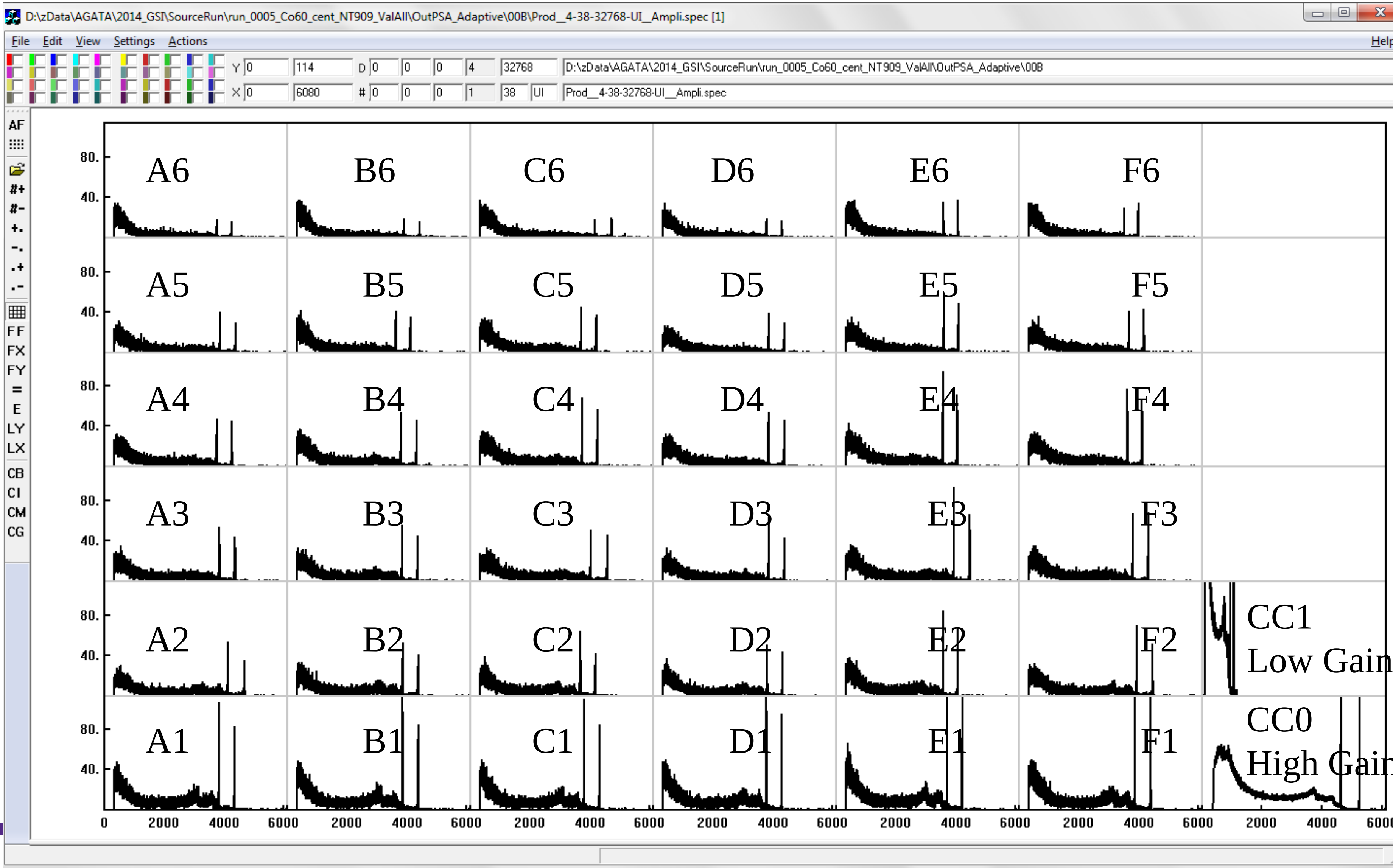


Some cores and segments saturate below 3 MeV on the high-gain range

Gain in the digitizers should be reduced by ~30% → Low gain for segment to be set in the digitizers

Prod__4-38-32768-UI__Ampli.spec

[1]-38-32768 used for energy calibrations



PreprocessingFilter

- Performs
  - Energy calibrations and cross talk corrections
  - Analysis of traces
    - Calculation of T0 from core (Digital CFD or linear fit of the first samples)
    - Time calibrations and shifts
    - Vertical normalization of traces
    - Define the net-charge segments
  - Reformats the data as data:ccrystal frames
  - After Preprocessing:
    - energies are stored in units of keV
    - times are in units of samples (10 ns) (but time calibration parameters are in ns)
    - positions are given in mm, when they show up after the PSA

Calibrations

- Energy
  - Gain-only, no offset coefficient needed because of the way the amplitude is generated in the preprocessing electronics.
  - Challenged by the Differential Non Linearity (DNL) of the ADC
- Cross-talk
  - 36*36=1296 coefficients to correct capacitive coupling correlations between segments and core
  - Used also to recover up to one broken or missing segment per crystal
- Time
  - Local → 36 coefficients to alignment of segment to the core
    great influence on the performance of PSA
  - Global → alignment of crystals and other detectors
    important to reduce random coincidences
- Response function of the system
  - To match the PSA calculated signals to the real data
- Neutron damage
  - Recover energy resolution of the segments using info from PSA.

PreprocessingFilterPSA.conf

laboratoire commun CEA/DSM CNRS/IN2P3

#segm/core	%d(id)	%f(tfall)	%f(trise)	%f(egain)	%f(emink)	%f(tmove)
segm	0	4600	1000	0.162979	10.0	9.891
segm	1	4600	1000	0.170079	10.0	10.463
segm	2	4600	1000	0.163820	10.0	10.361
segm	3	4600	1000	0.169401	10.0	9.485
segm	4	4600	1000	0.158867	10.0	7.971
segm	5	4600	1000	0.155504	10.0	10.077
segm	6	4600	1000	0.170291	10.0	9.050
segm	7	4600	1000	0.165092	10.0	9.263
segm	8	4600	1000	0.145804	10.0	7.420
segm	9	4600	1000	0.168806	10.0	8.448
segm	10	4600	1000	0.143493	10.0	6.188
segm	11	4600	1000	0.159609	10.0	10.510
segm	12	4600	1000	0.153815	10.0	10.251
segm	13	4600	1000	0.155996	10.0	9.448
segm	14	4600	1000	0.168760	10.0	9.537
segm	15	4600	1000	0.175860	10.0	8.866
segm	16	4600	1000	0.185031	10.0	13.873
segm	17	4600	1000	0.157300	10.0	10.564
segm	18	4600	1000	0.169683	10.0	9.836
segm	19	4600	1000	0.168100	10.0	9.683
segm	20	4600	1000	0.170233	10.0	9.677
segm	21	4600	1000	0.174663	10.0	9.472
segm	22	4600	1000	0.174109	10.0	8.942
segm	23	4600	1000	0.165021	10.0	11.498
segm	24	4600	1000	0.152862	10.0	10.267
segm	25	4600	1000	0.169911	10.0	11.067
segm	26	4600	1000	0.165142	10.0	9.910
segm	27	4600	1000	0.159595	10.0	9.393
segm	28	4600	1000	0.168353	10.0	8.347
segm	29	4600	1000	0.167807	10.0	11.519
segm	30	4600	1000	0.163006	10.0	9.945
segm	31	4600	1000	0.159887	10.0	10.383
segm	32	4600	1000	0.155449	10.0	9.600
segm	33	4600	1000	0.143345	10.0	8.487
segm	34	4600	1000	0.150043	10.0	7.216
segm	35	4600	1000	0.176351	10.0	12.197
core	0	4200	1000	0.347055	20.0	25.000
core	1	4200	1000	0.069358	20.0	-0.285

tntf 2097152

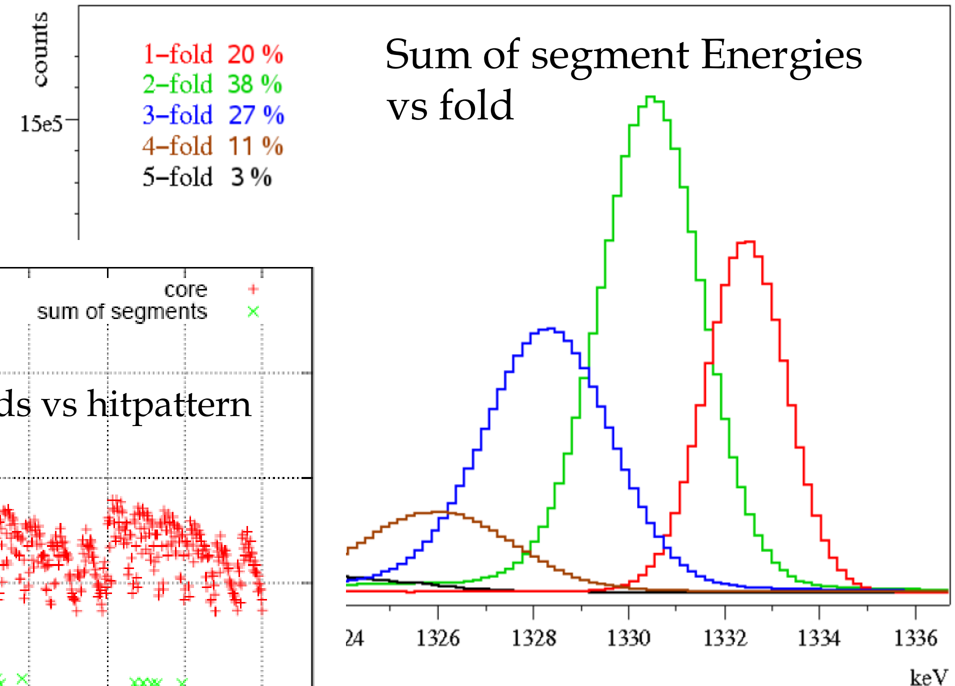
Cross-Talk

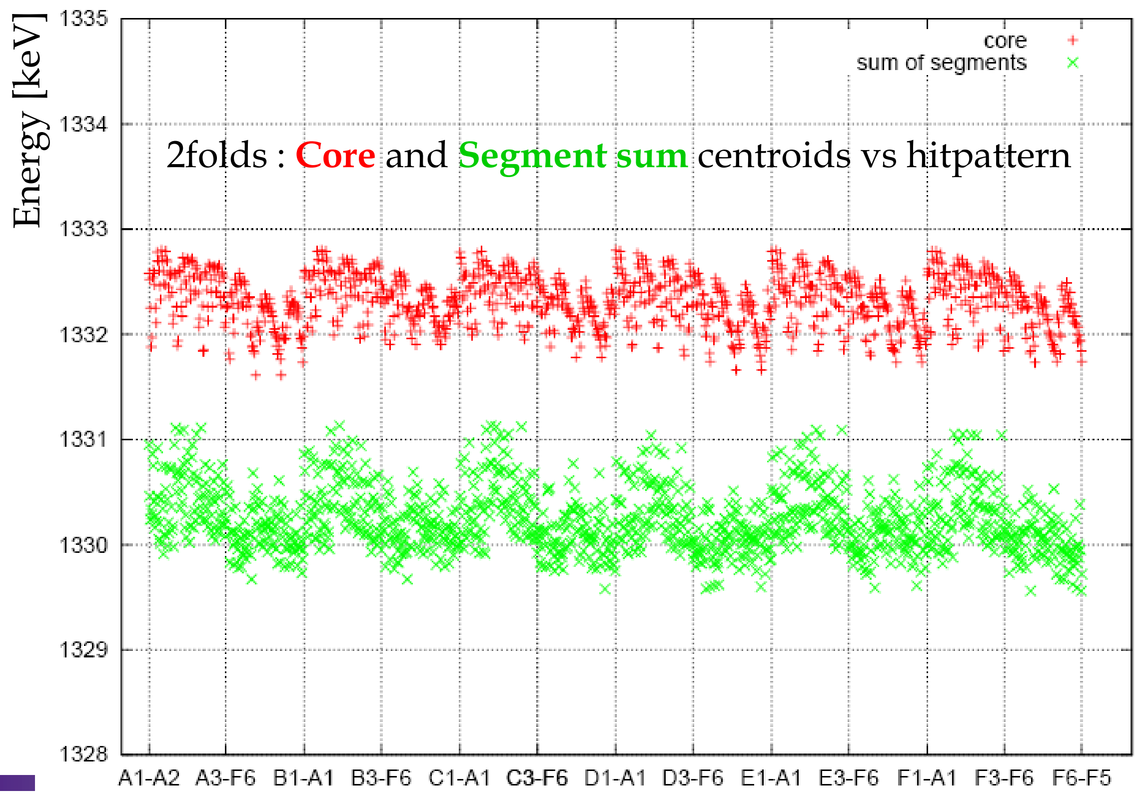
- Proportional
  - Affects energy spectra of multi-segment events
  - xTalkSort …
- Differential
  - Affects rise-time of signals and therefore PSA
  - Not yet fully characterized

Crosstalk correction: Motivation

laboratoire commun CEA/DSM CNRS/IN2P3

- Crosstalk is present in <u>any</u> segmented detector
- Creates strong energy shifts proportional to fold
- Tracking needs <u>segment energies</u> !



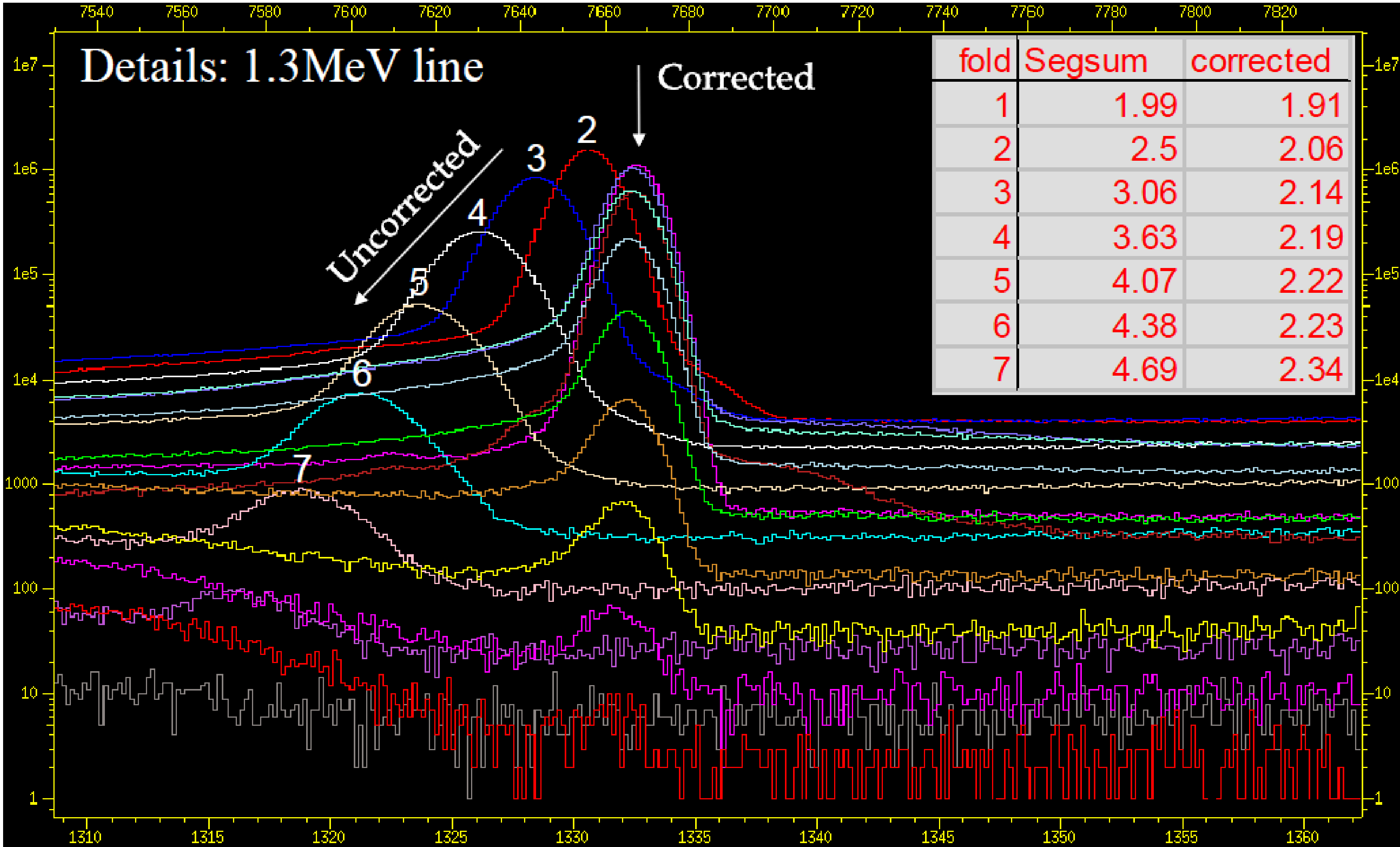


.All possible 2fold combinations

Cross talk correction: Results

laboratoire commun CEA/DSM CNRS/IN2P3

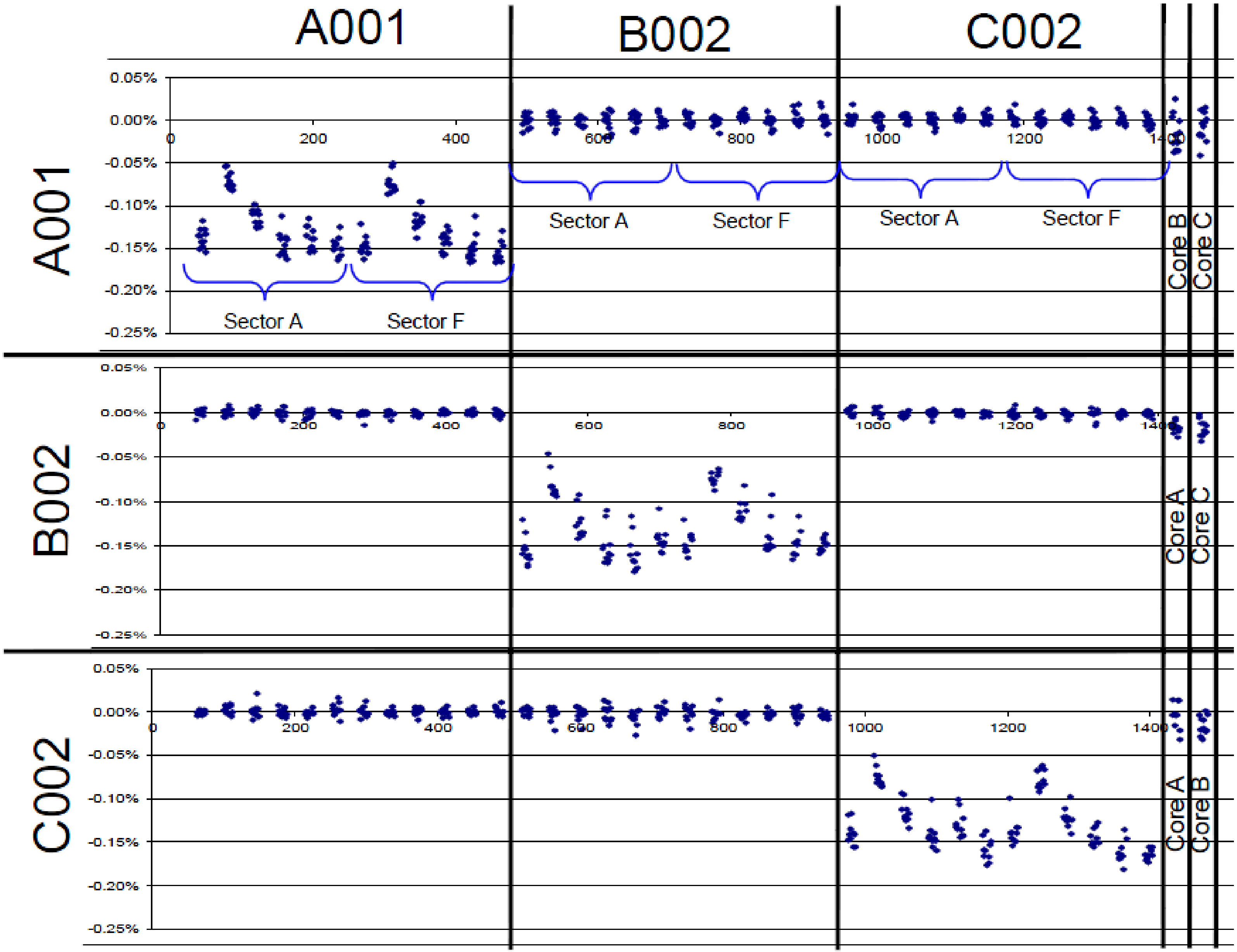
FWHM 60keV: 1.20 → 1.02 !



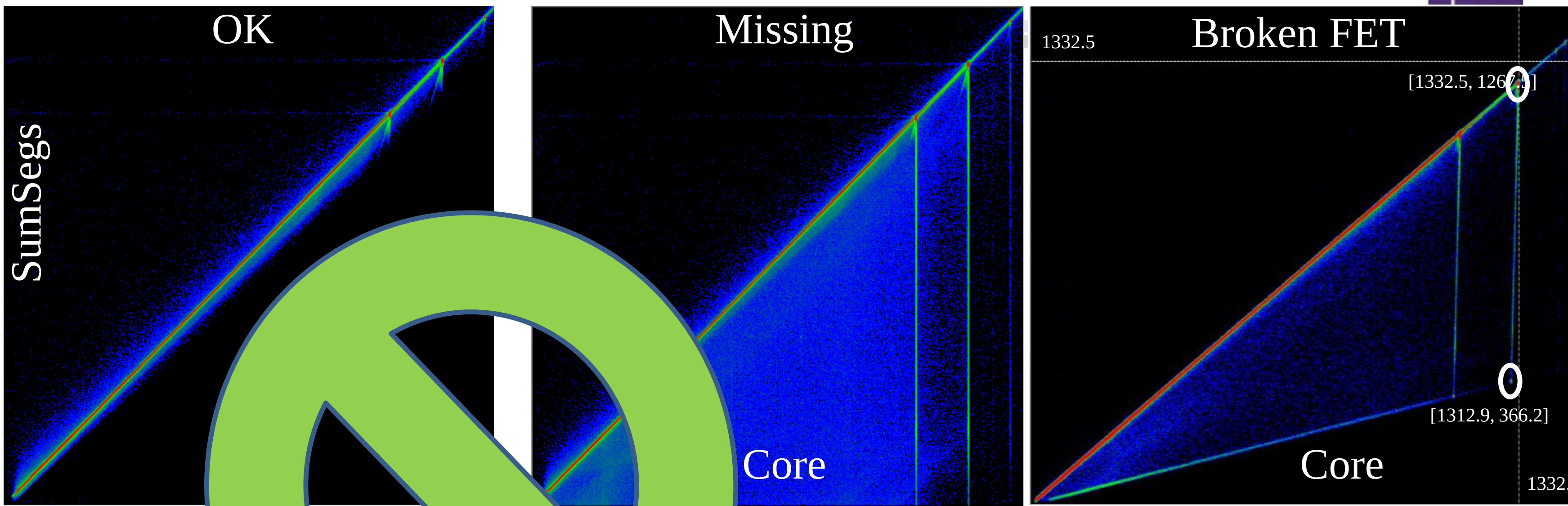
fold	Segsum	corrected
1	1.99	1.91
2	2.5	2.06
3	3.06	2.14
4	3.63	2.19
5	4.07	2.22
6	4.38	2.23
7	4.69	2.34

B. Bruyneel et al, Nucl. Instr. and Meth. A 608, (2009) 99

Cross talk in AGATA Triple Cluster

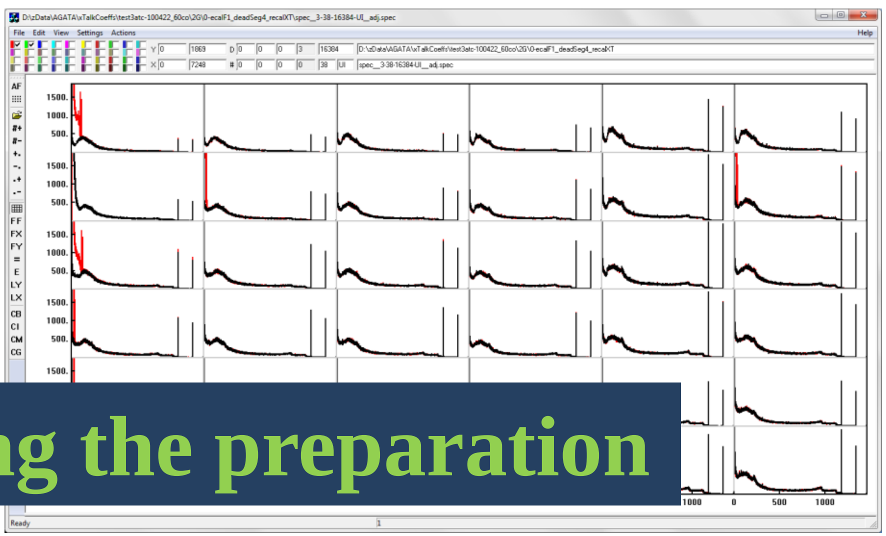


Missing/Broken segments



Read the Cook... by Rosa

- Exploiting the ...ncy
  ∑Se...
  - Compensate loss ...
  - Assign energy to missing segment
    $E_{missing} = E_{core} - Sum_{Other}$
  - Remove ghost peaks in neighbours and determine energy of missing segme... talk
  - Final energy of missing segment using again SumSegs==Core



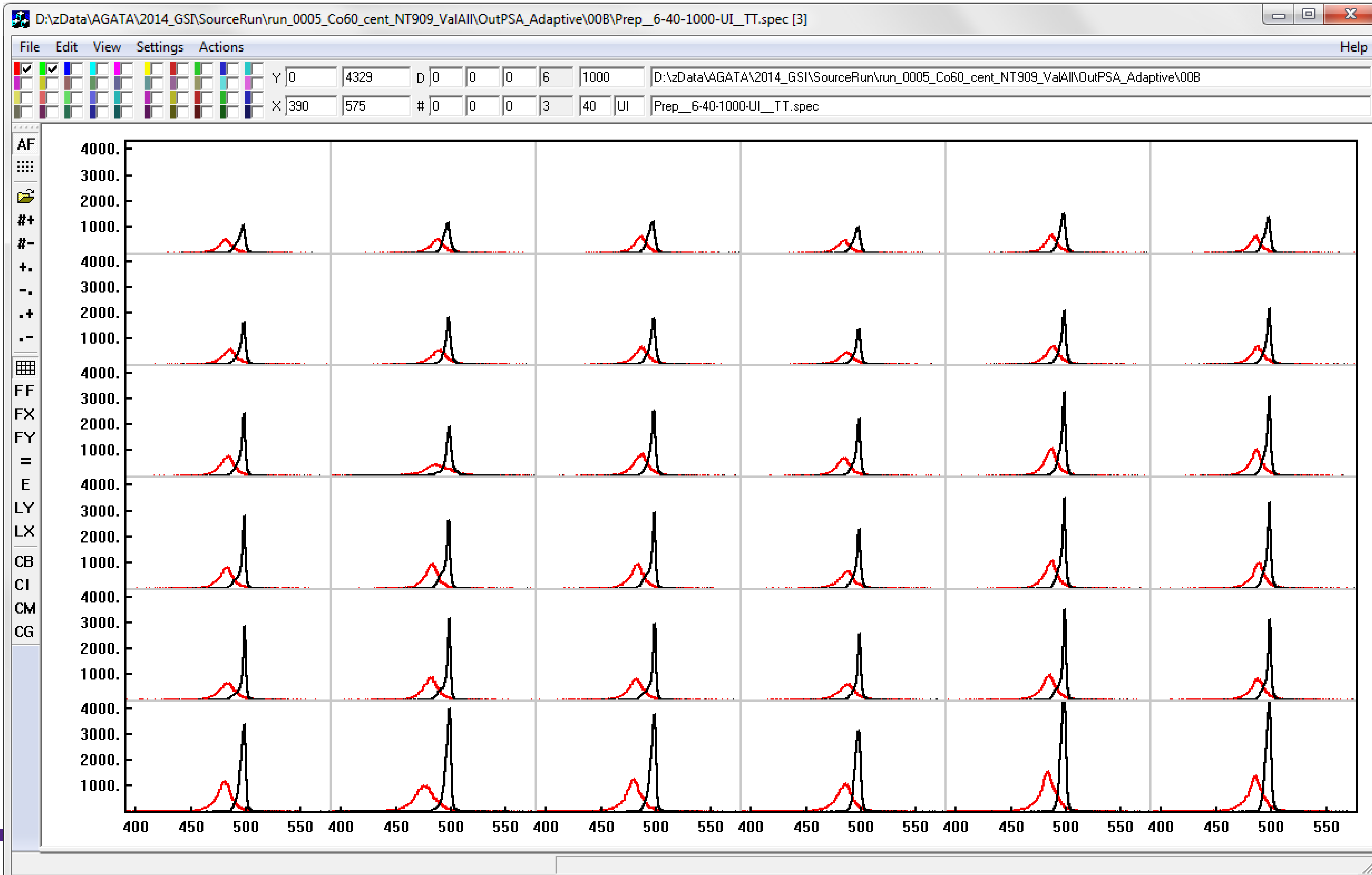
Recovery of broken segment

Done for you during the preparation

Time calibrations at Local Level

- Need to align the traces because the hardware does not compensate all time lags
- Need to re-trigger the event of a whole inside the capture time. Important is the absolute time of the first sample (timestamp)
  - Straight line fit or Digital CFD of the signal rise-time using info from core and segments.

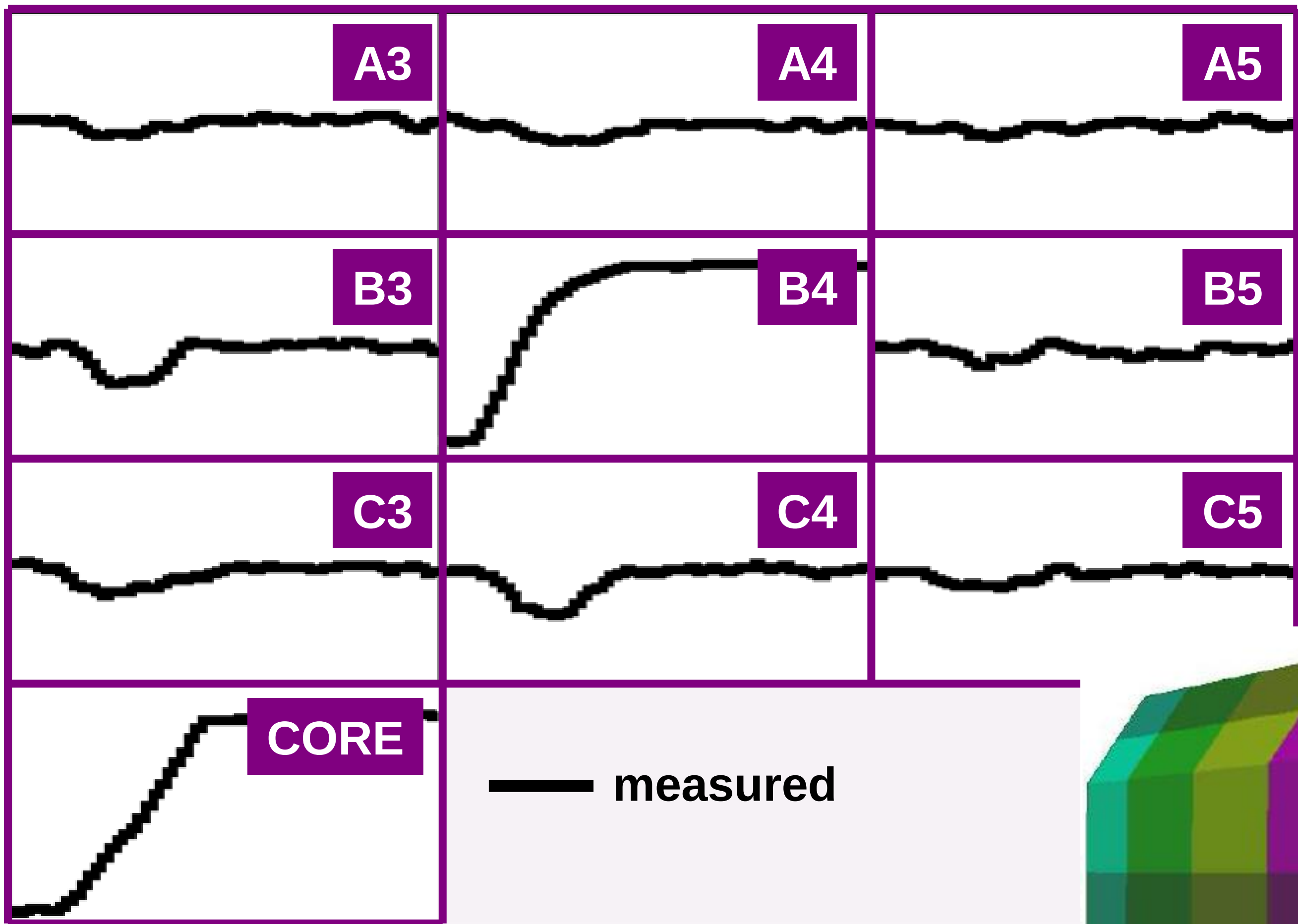
Time alignment of segments re Core

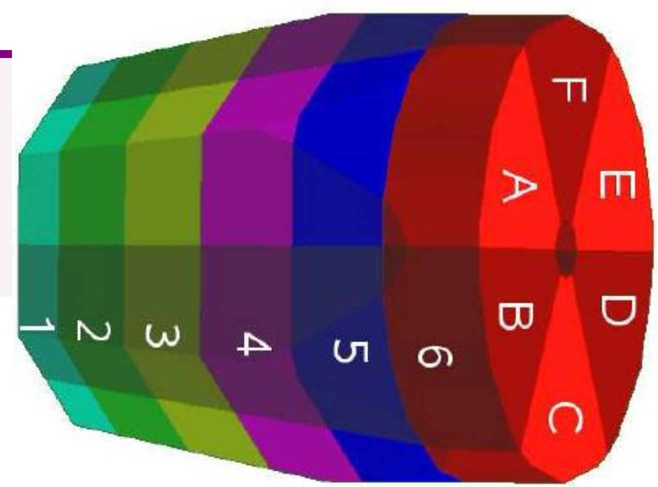


PSAFilter

- Signal decomposition
- Implemented algorithm is the Grid Search
  - As a full grid search
  - As a coarse/fine search (AGS)
- Reduces size of data by factor 20
- Provides the parameters for the correction of neutron damage (can also perform it)
- Must be expanded to improve timing
- **Takes ~95 % of total CPU time**
- Is the critical point for the processing speed of online and offline analyses

Pulse Shape Analysis concept

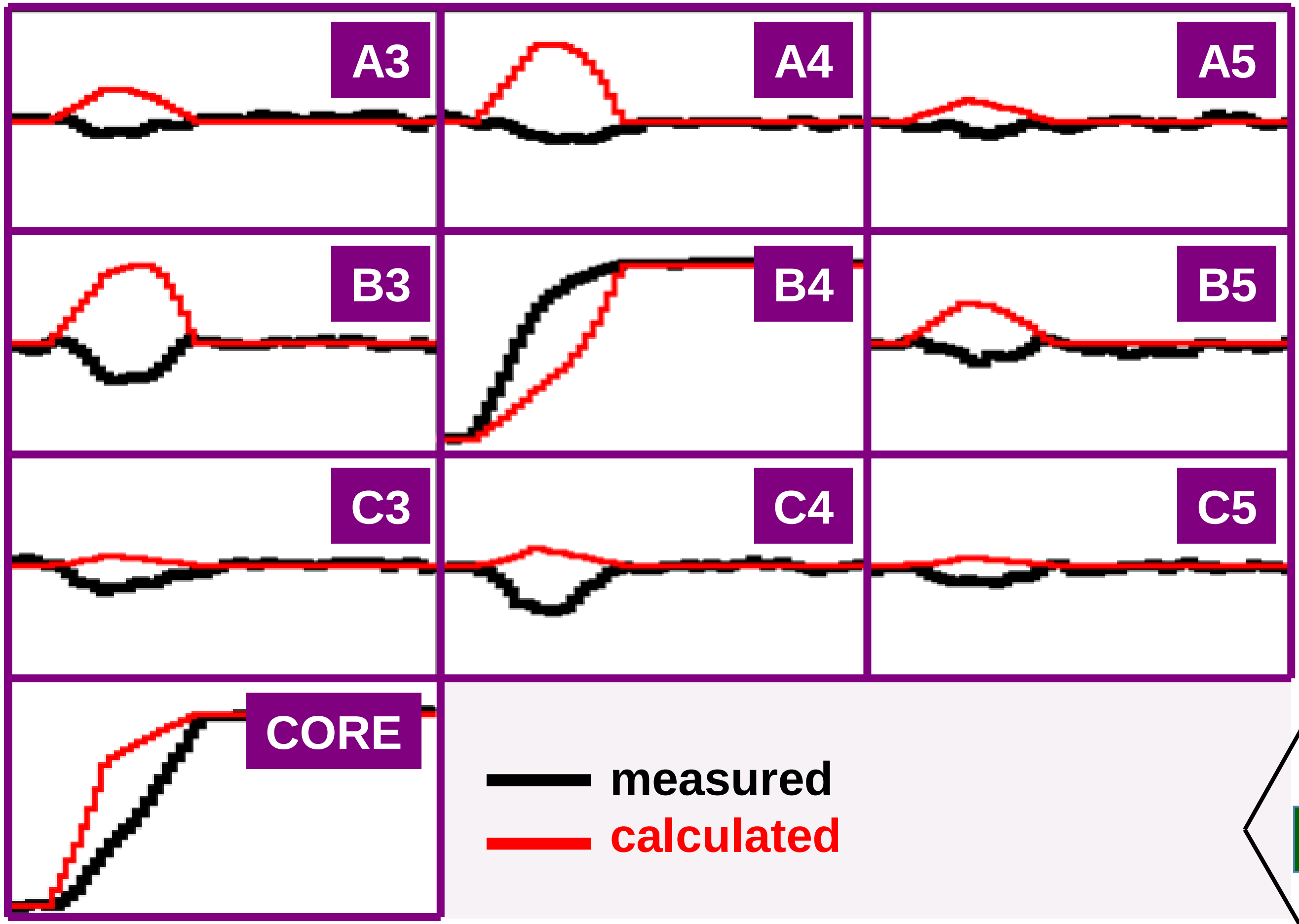




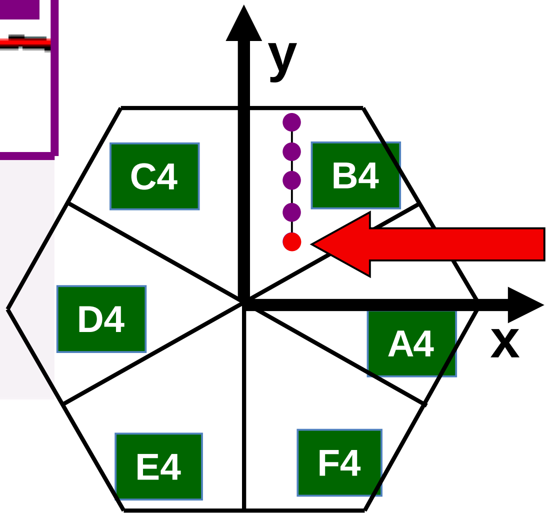
791 keV deposited in segment B4

Courtesy Francesco Recchia

Pulse Shape Analysis concept

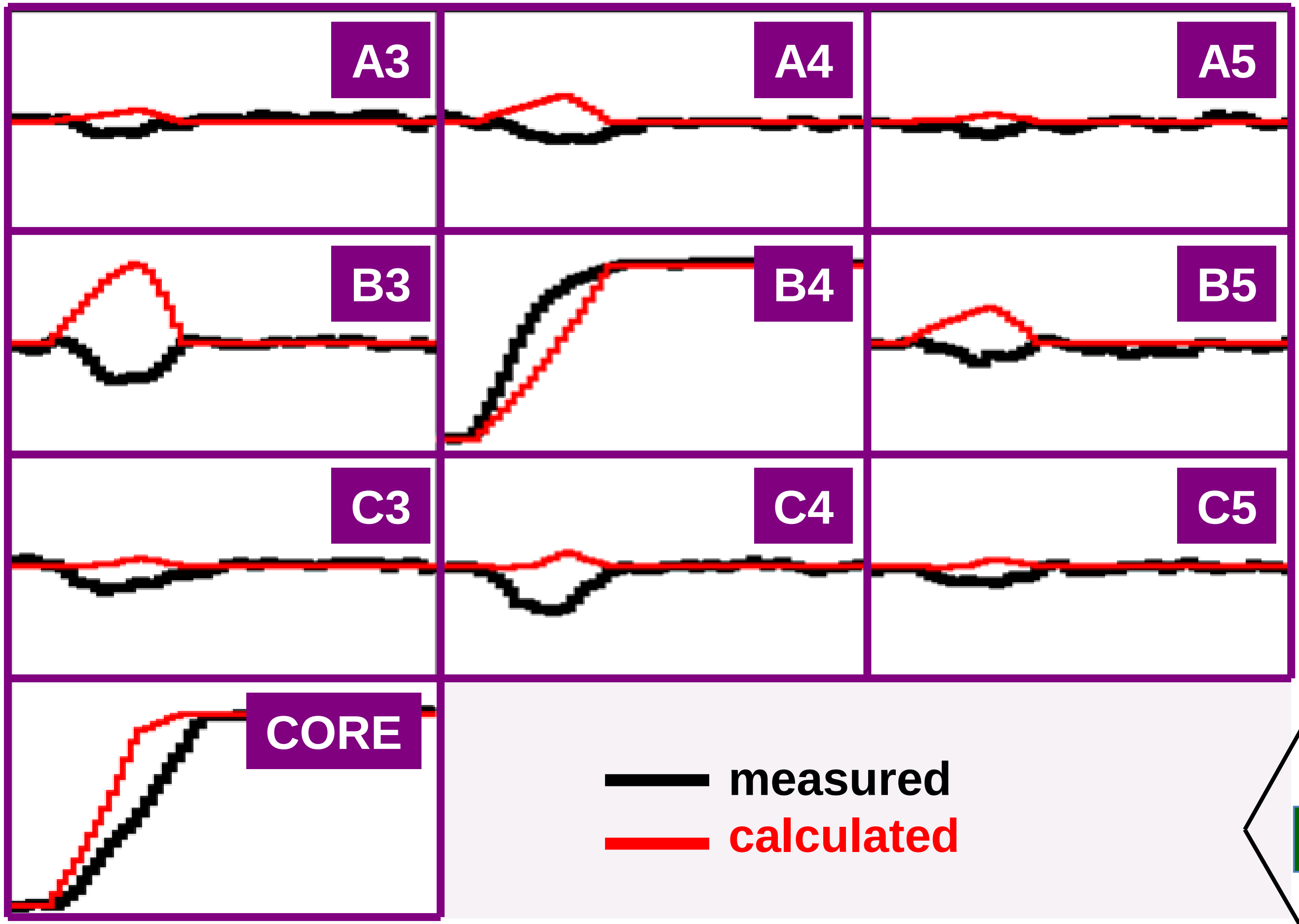


(10,10,46)

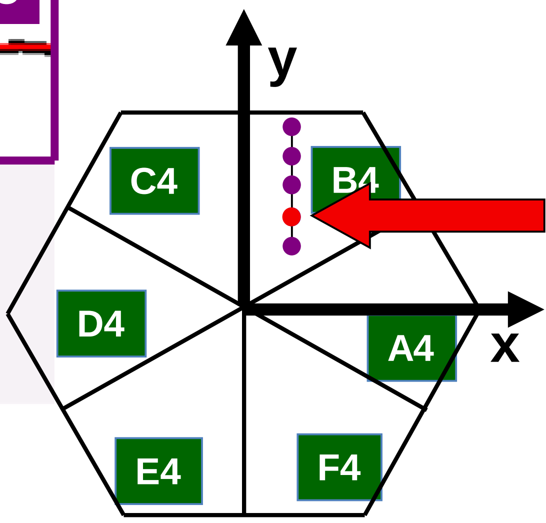


791 keV deposited in segment B4

Pulse Shape Analysis concept

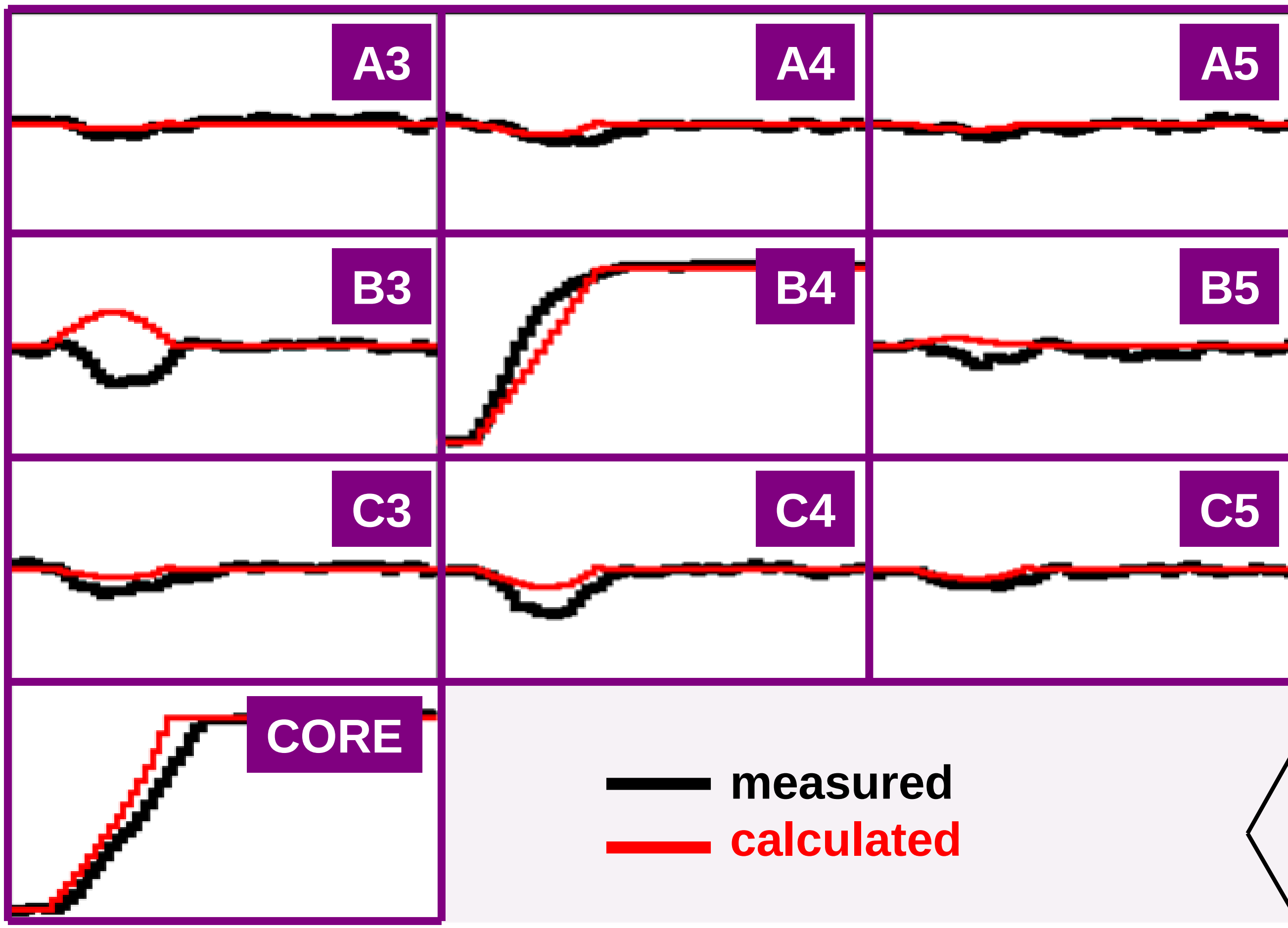


(10,15,46)



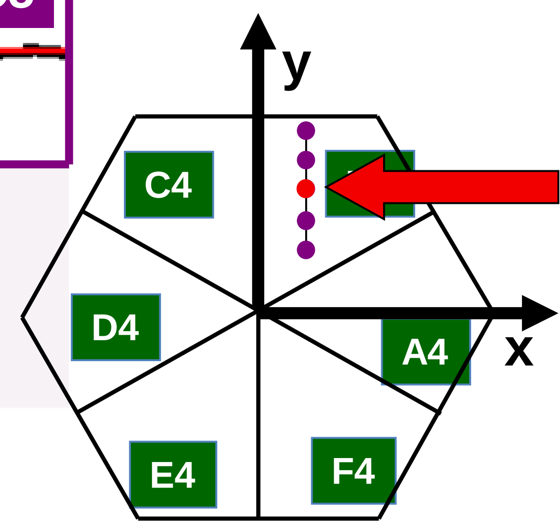
791 keV deposited in segment B4

Pulse Shape Analysis concept



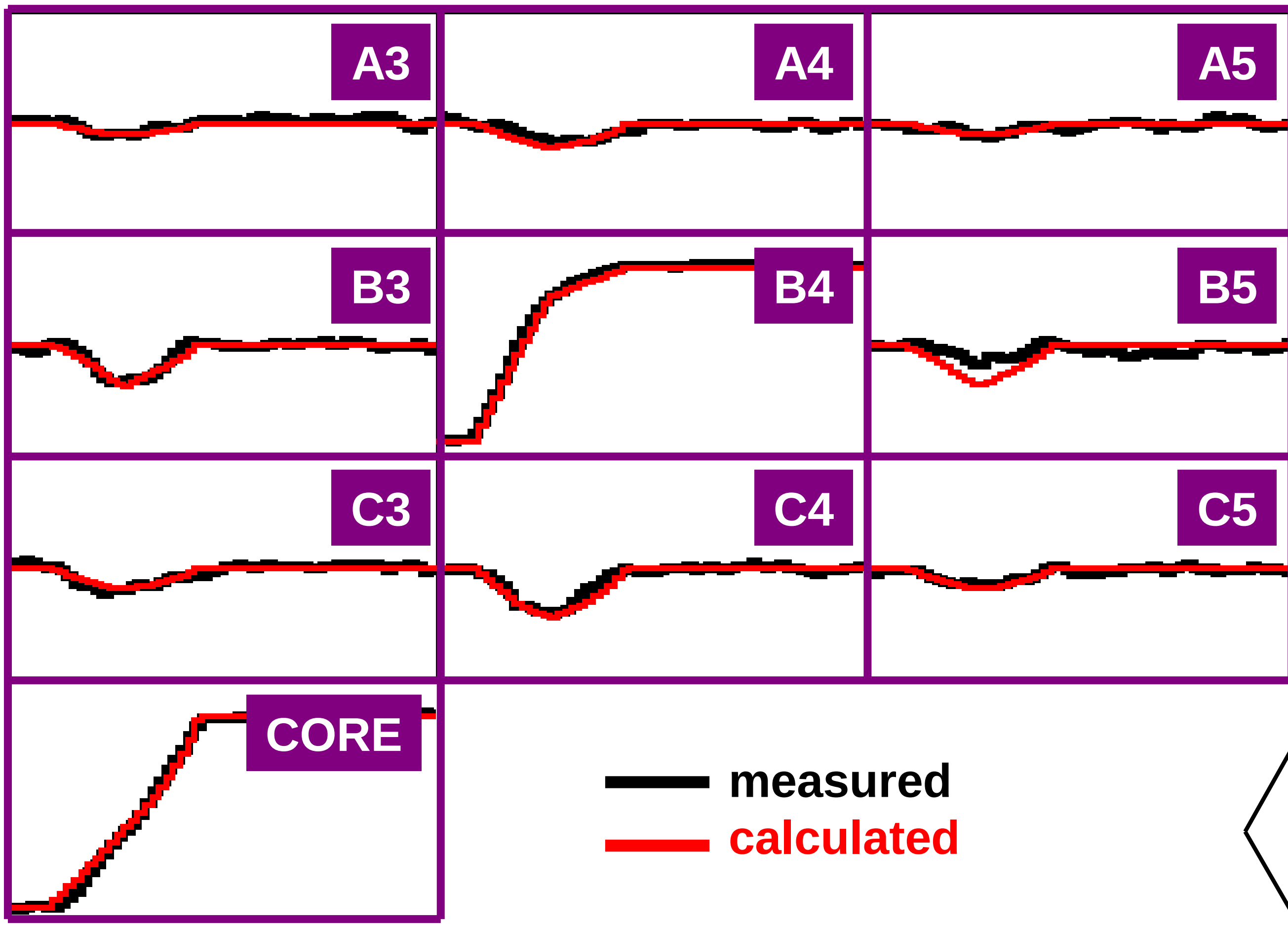
791 keV deposited in segment B4

(10,20,46)



z = 46 mm

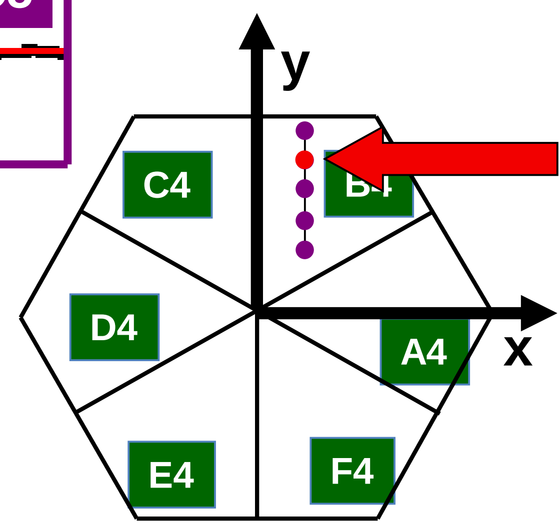
Pulse Shape Analysis concept



measured

calculated

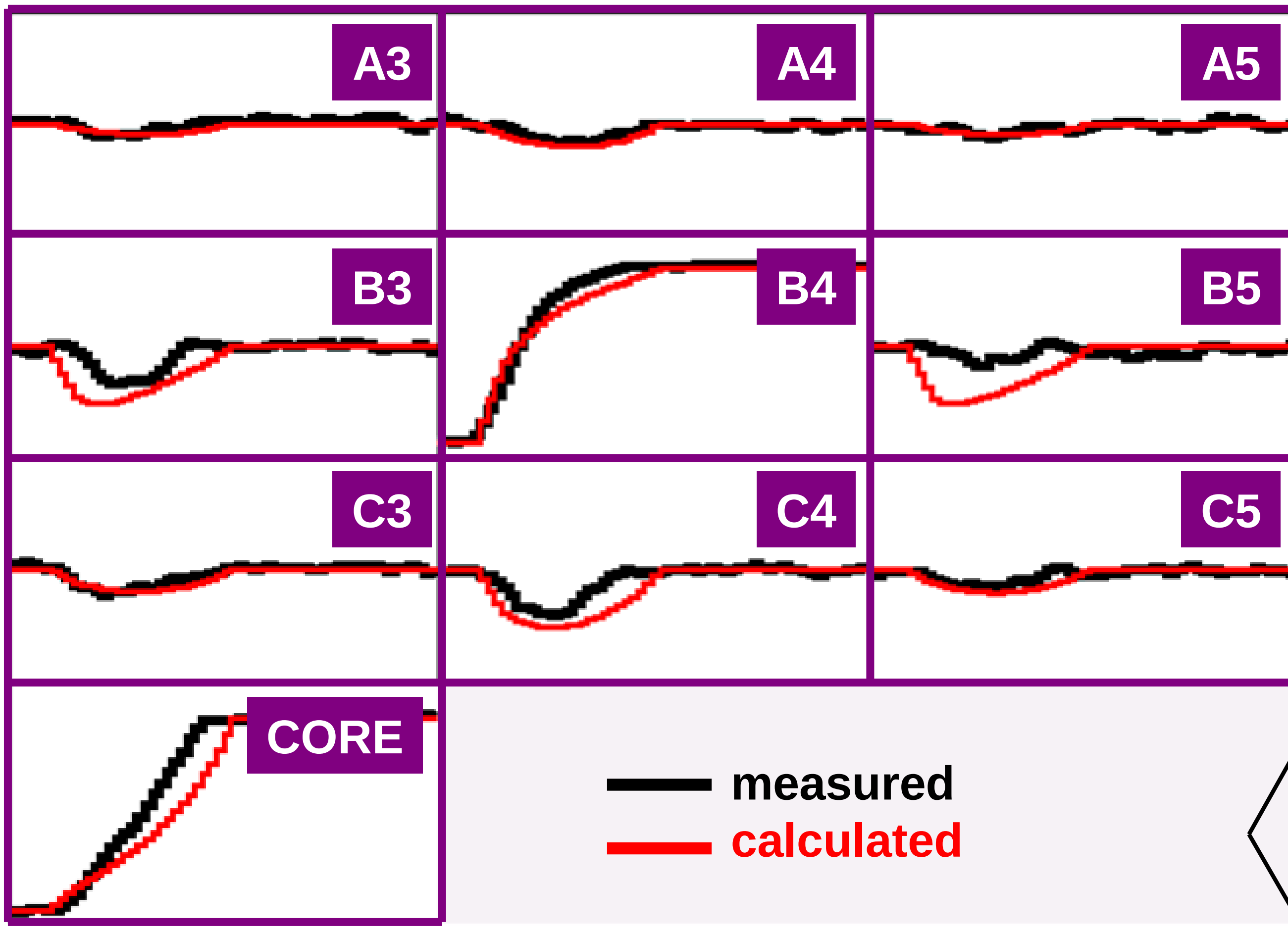
(10,25,46)



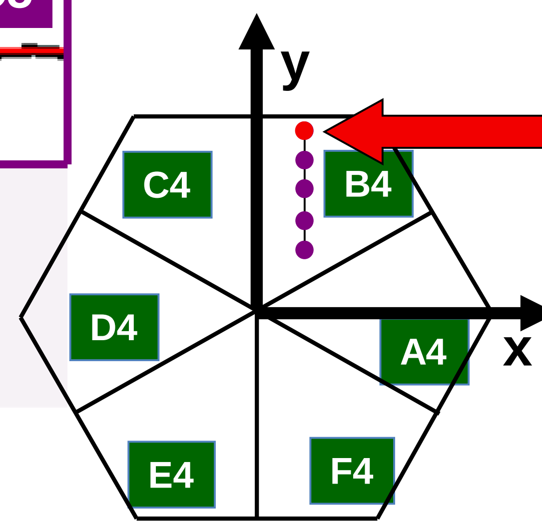
z = 46 mm

791 keV deposited in segment B4

Pulse Shape Analysis concept



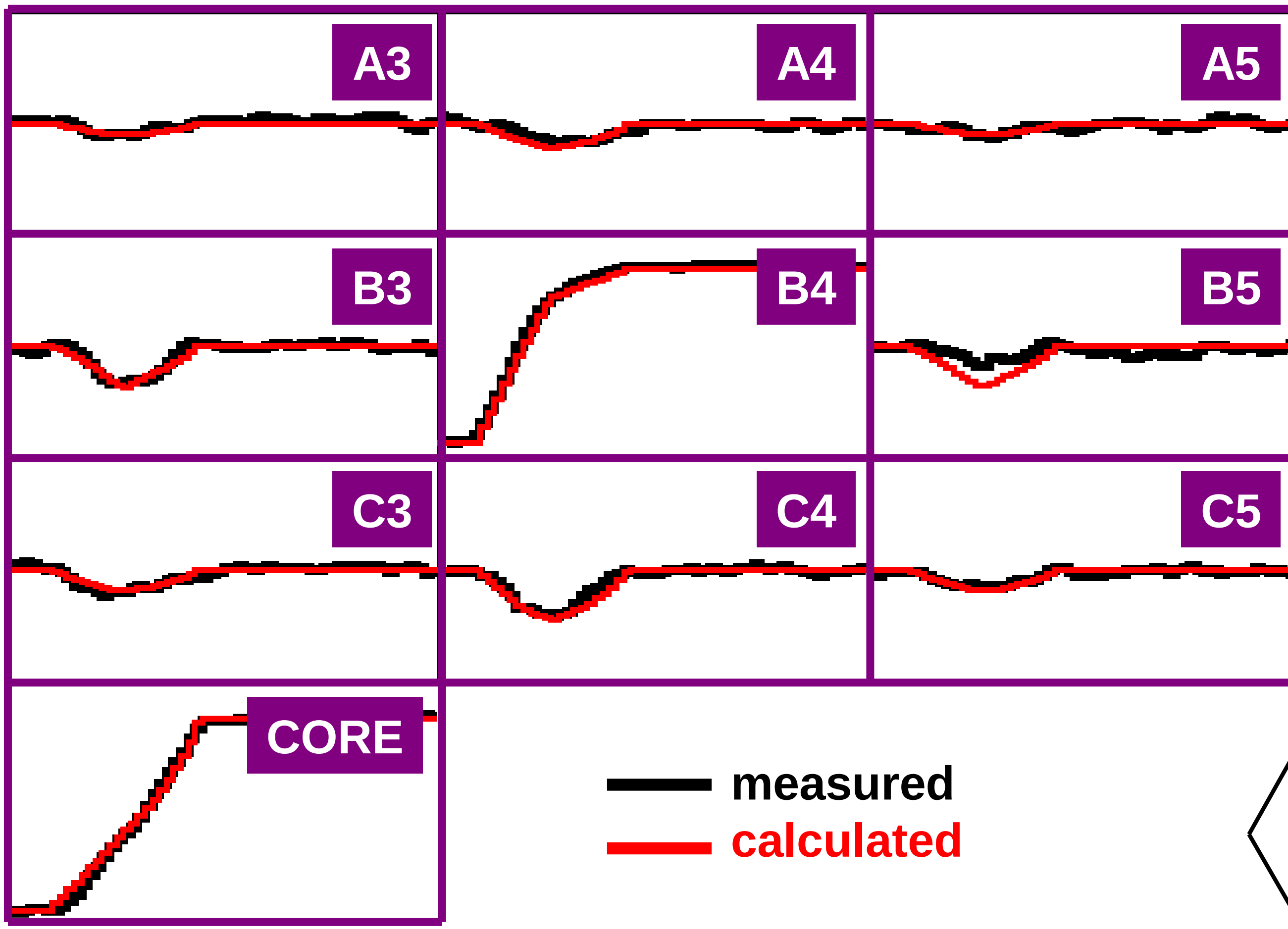
(10,30,46)



z = 46 mm

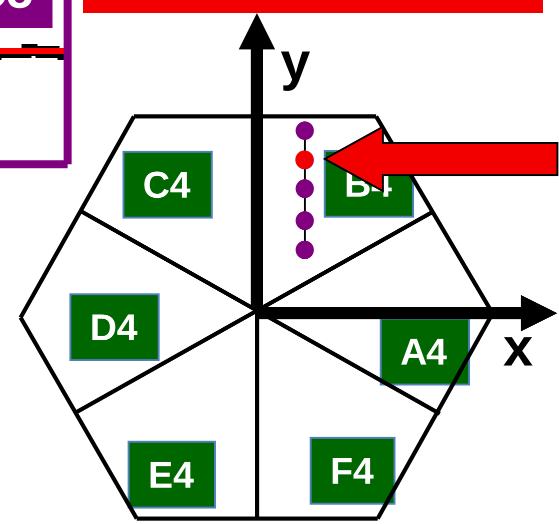
791 keV deposited in segment B4

Pulse Shape Analysis concept



Result of *Grid Search* Algorithm

(10,25,46)



z = 46 mm

791 keV deposited in segment B4

The Grid Search algorithm

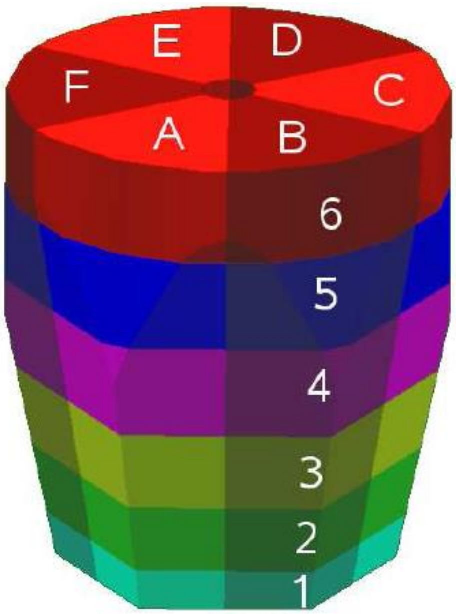
implemented in Narval as **PSAFilterGridSearch**

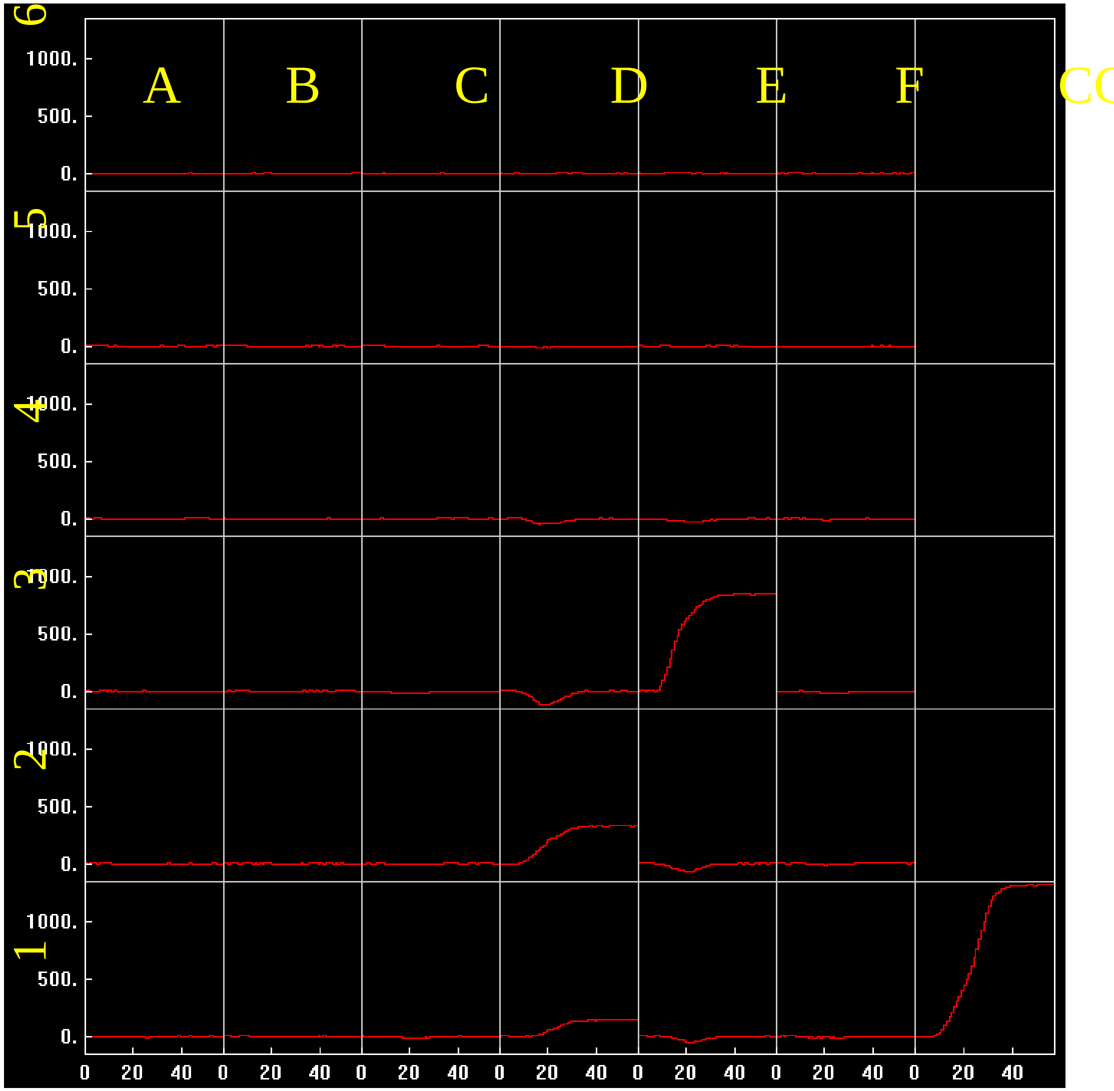
- Signal decomposition **assumes one interaction per segment**
- The decomposition uses the transients and a differentiated version of the net charge pulse
- Proportional and differential cross-talk are included using the xTalk coefficients of the preprocessing.
- The minimum energy of the “hit” segments is a parameter in the PreprocessingFilter → 10 keV
- No limit to the number of fired segments (i.e. up to 36)
- The number of used neighbours is a compile time parameter (usually 2 as Manhattan distance)
- The algorithm cycles through the segments in order of decreasing energy;
  the result of the decomposition is removed from the remaining signal
  → subtraction method at detector level

- Using ADL bases (Bart Bruyneel) with the neutron-damage correction model
  (the n-damage correction is actually performed in the PostPSAFilter)

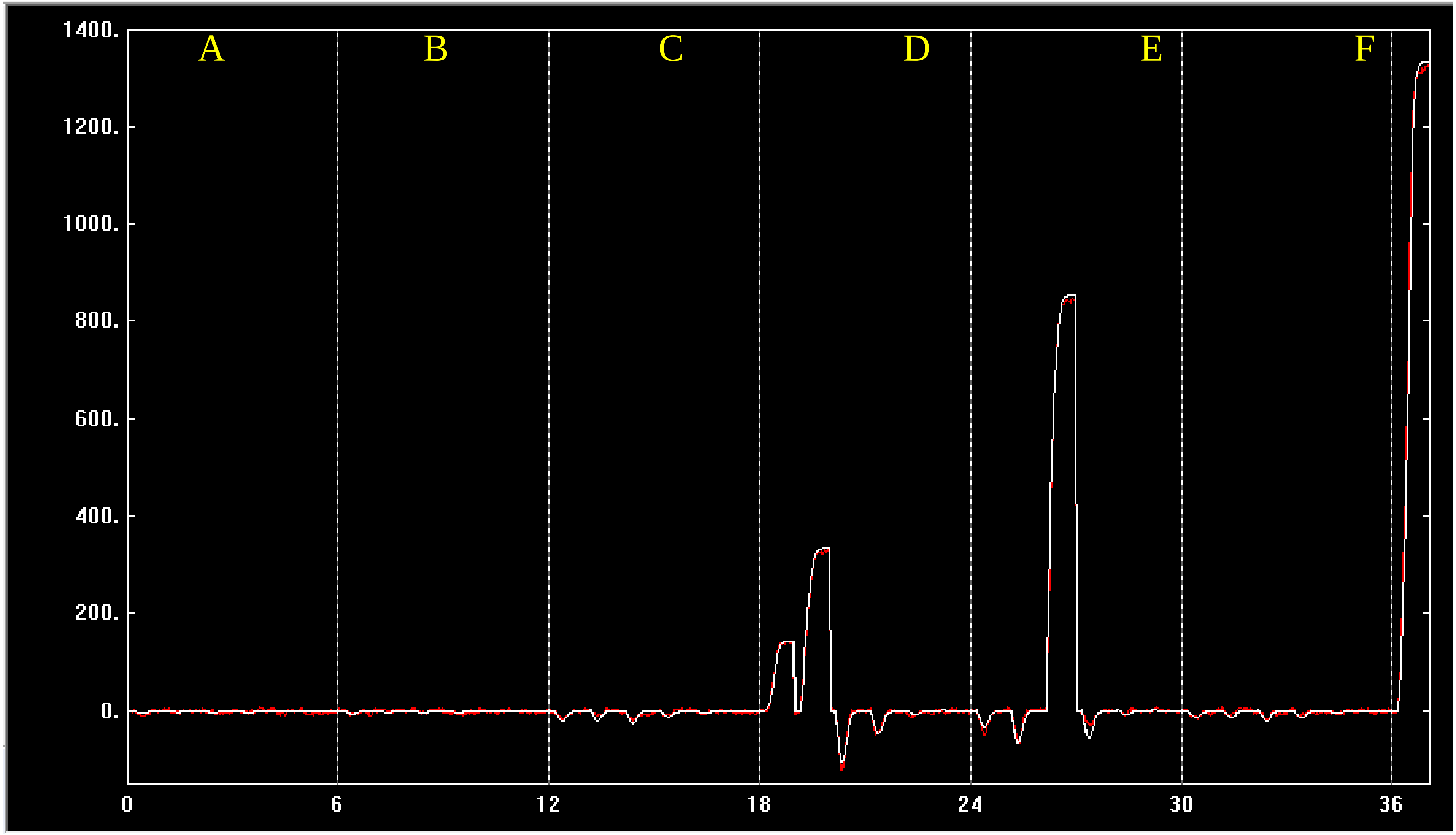
- Using 2 mm grids → ~48000 grid points in a crystal; 700-2000 points/segments
- Speed is **~ 150 events/s/core for the Full Grid Search**
  ~ 1000 events/s/core for the Adaptive Grid Search
- To speedup execution, parallelism has been implemented (using boost:threads) with blocks of ~100 events passed to N parallel threads of execution

Adaptive Grid Search in action



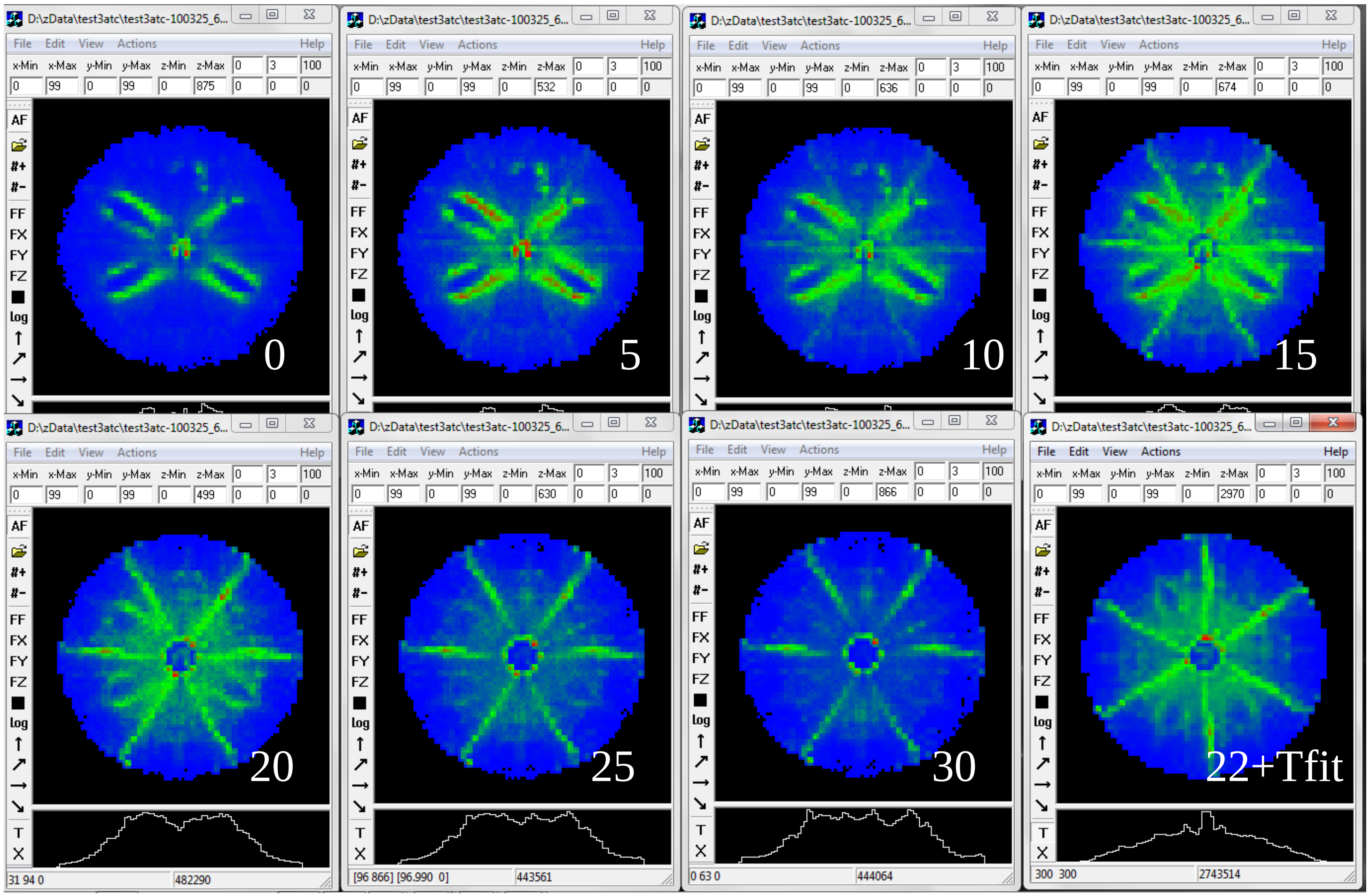


Adaptive Grid Search in action



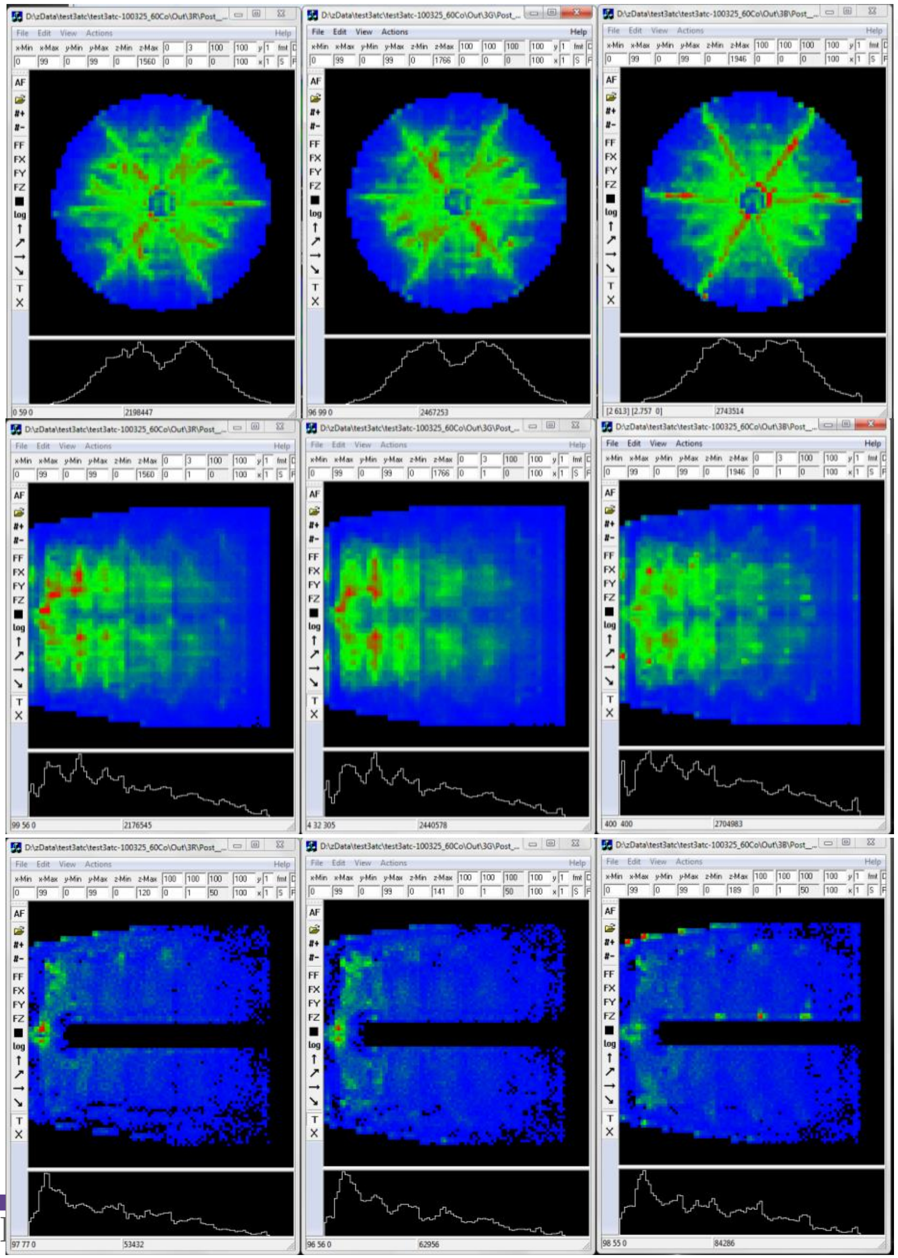
Final Result

Effect of time alignment



Effect of time position of the experimental trace
Similar behaviour when scanning over the risetime of the preamplifier, because a too-fast/too-slow response sees the experimental traces as slow/fast or late/early

Best time alignment and fit of T_0 of signal

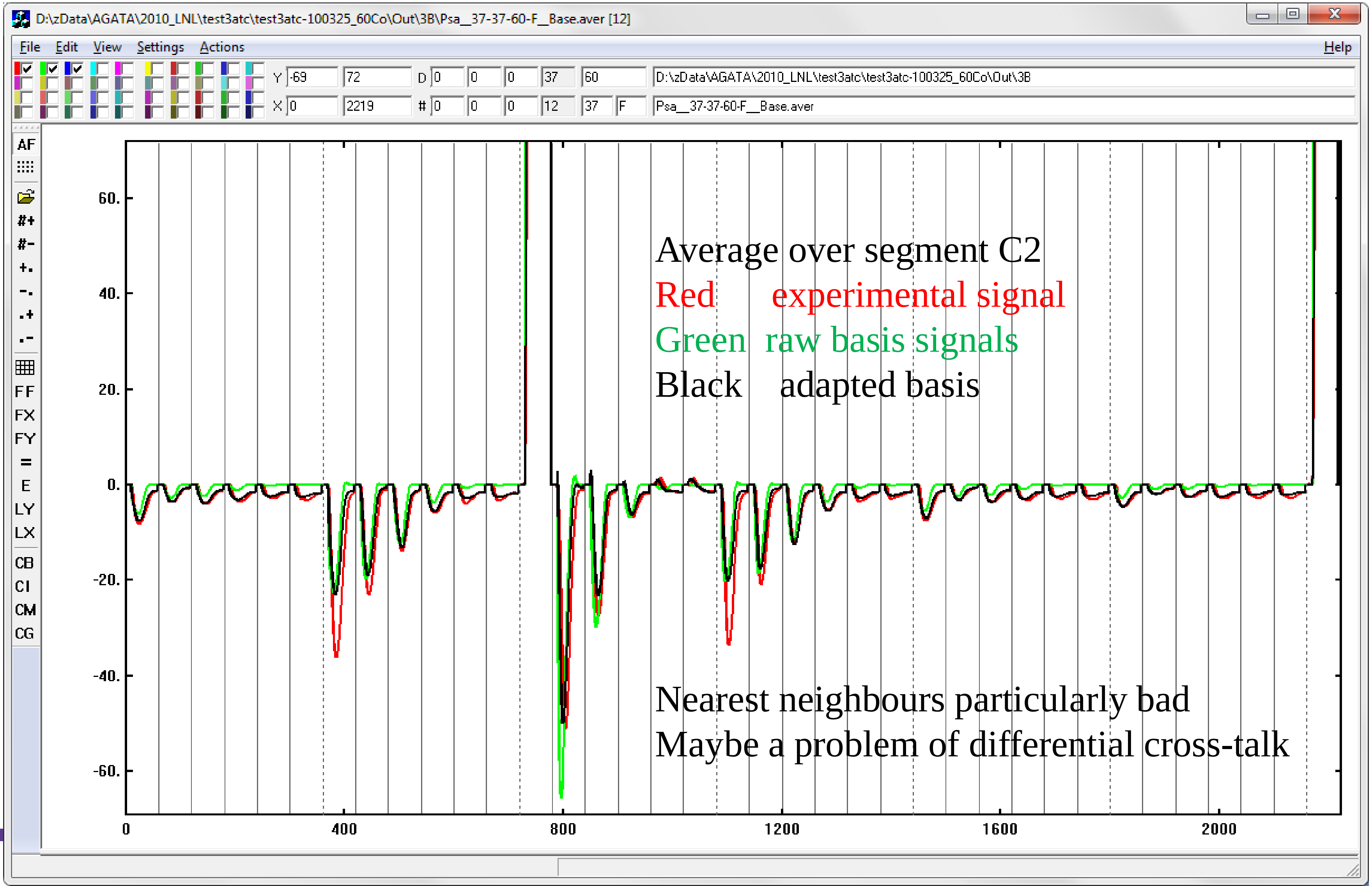


- Still a lot of clustering
- Difficult to get better results
- Is it Problem of:
  - Algorithm ?
  - Calculated basis ?
  - Preparation of data ?
  - with a different algorithm and a completely different way of preparing the data GRETINA has similar effects
  - Could be an intrinsic limit of the method

Cross talk and PSA

- Logically, cross talk is part of response function of the system
  - Proportional cross-talk applied to signal basis
  - Differential cross-talk applied in the same way but clearly not done well. Question of proportionality between the two.
  - Example

Response function and cross-talk applied to the signal basis



PostPSAFilter

- Final energy and time calibrations
- Recovery (partial) of neutron damage using info from PSA
  - Impact of neutron damage on:
    - energy resolution
    - Signal shape and PSA

See Next talk