

# High Performance Computing softwares for astronomy and astrophysics

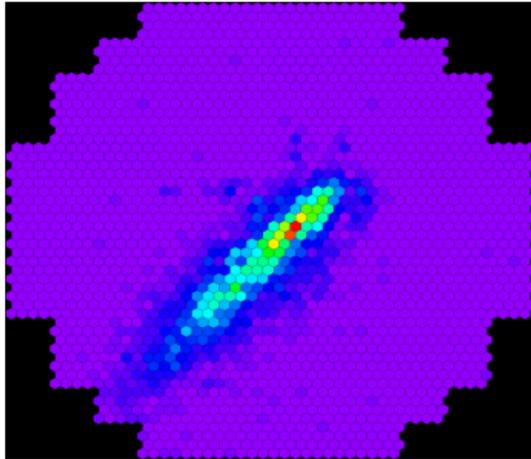
Pierre Aubert, Thomas Vuillaume, Jean  
Jacquemier, Florian Gaté, Gilles Maurin  
LAPP, CNRS

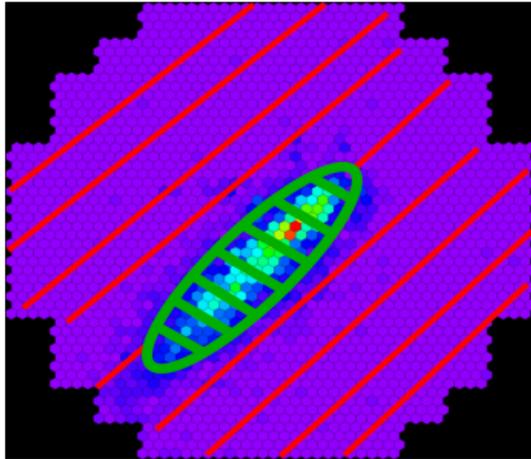
# CTA challenges

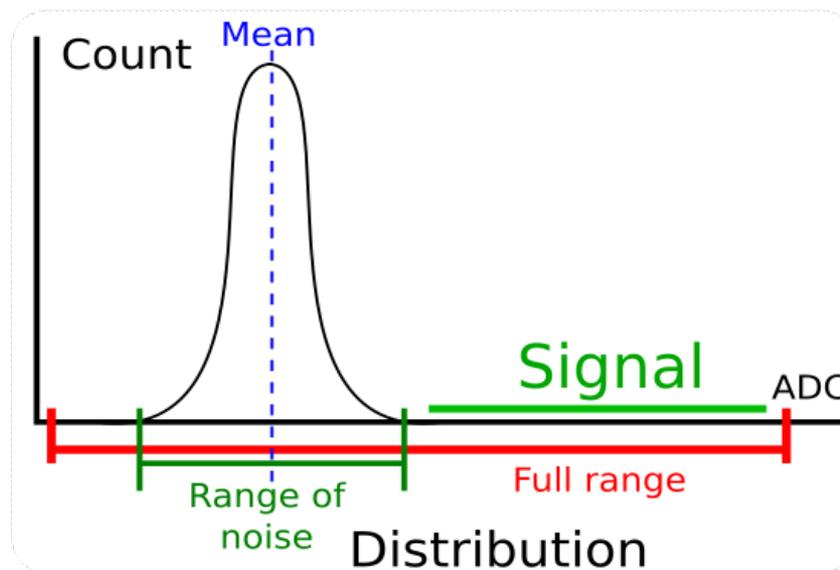
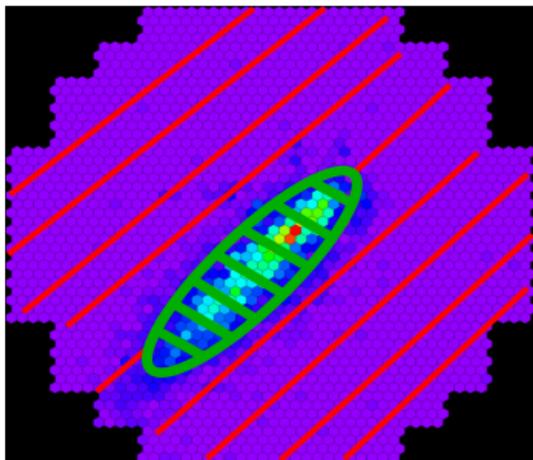
- Data volumes:
  - Storage
  - Transfer
  - Computing
- Complex reconstruction
  - Discrimination between particles
  - Reconstruction of incident photon direction and energy

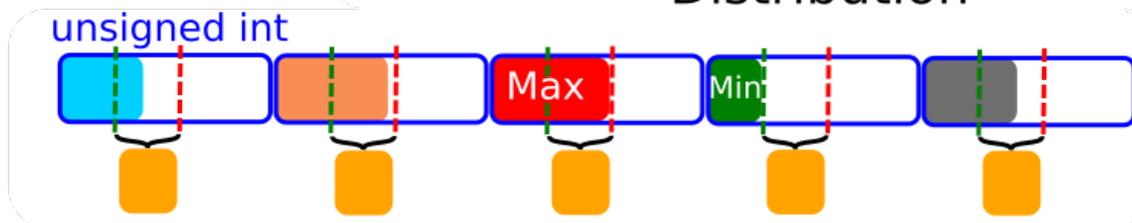
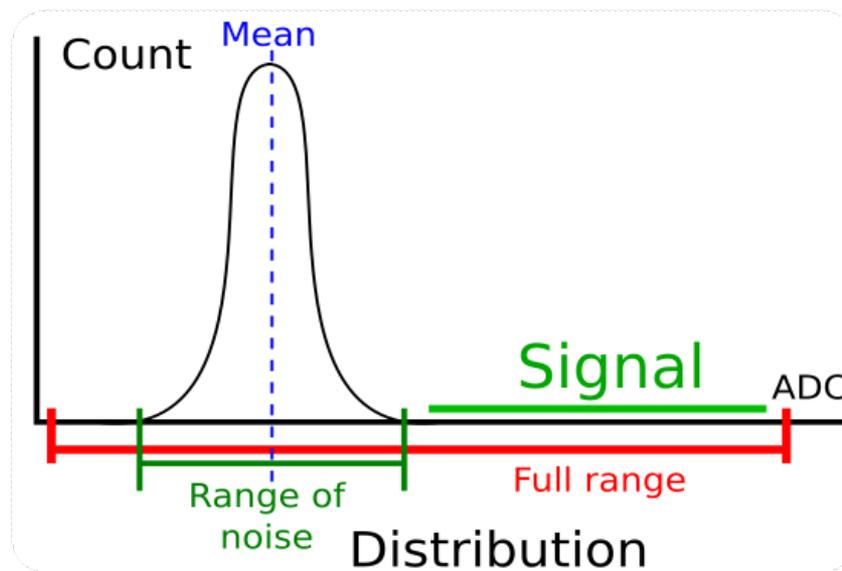
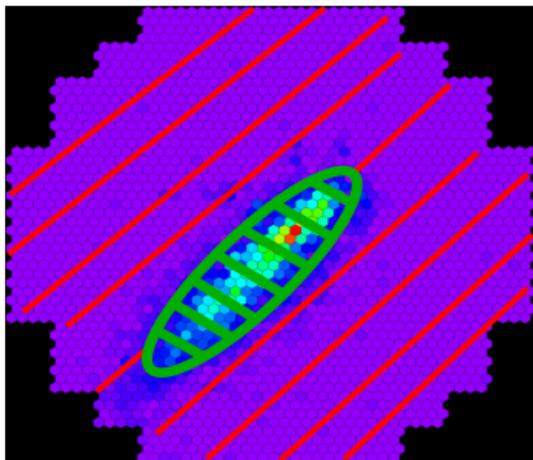
# Lossless data compression

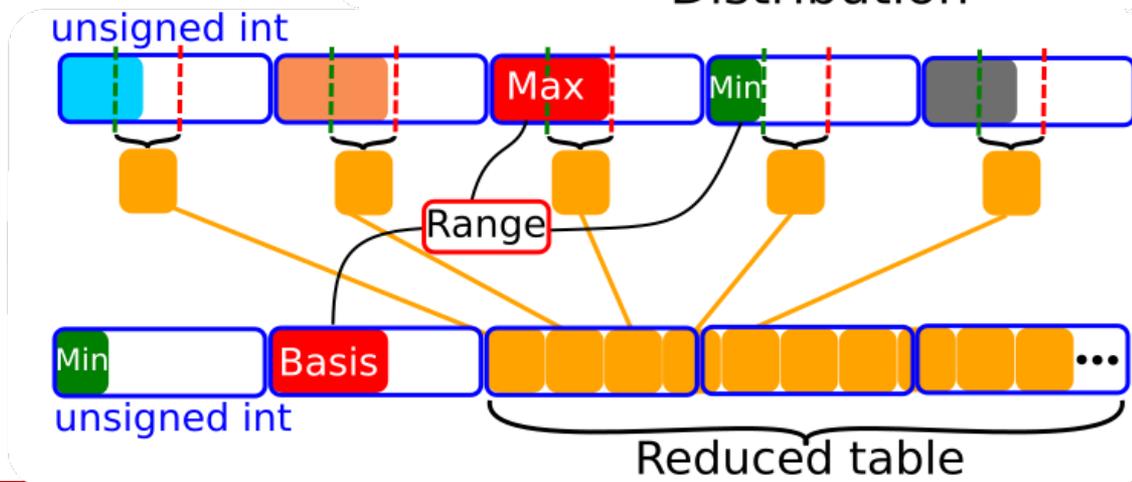
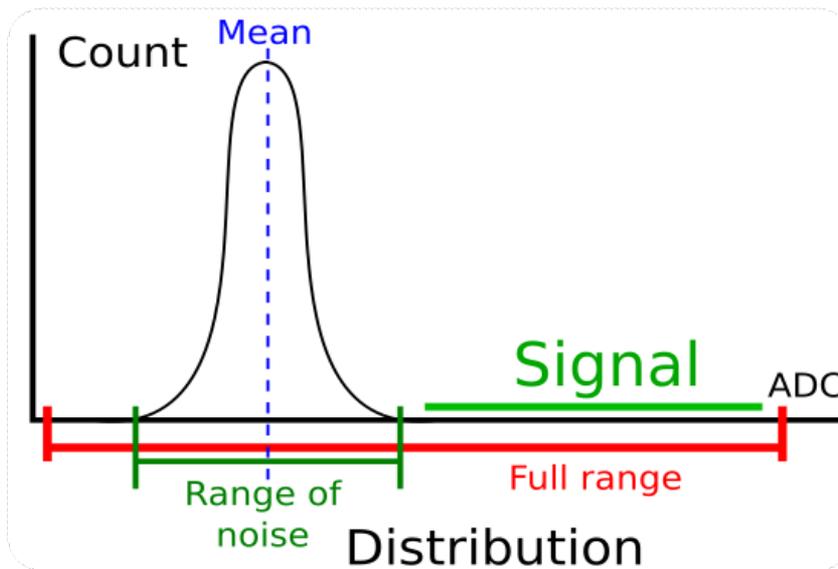
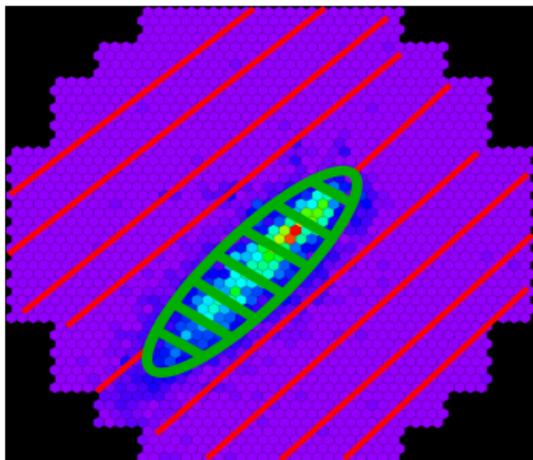
- Data transfer
- Data storage

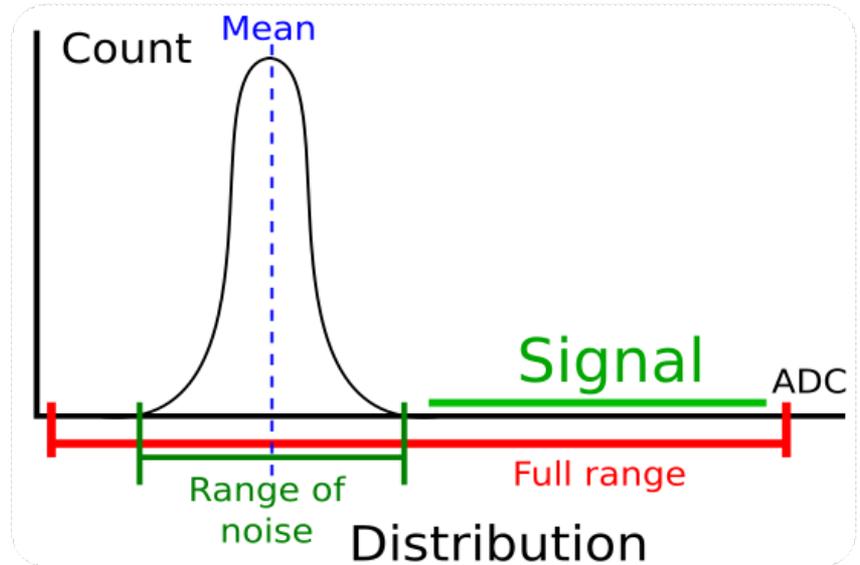
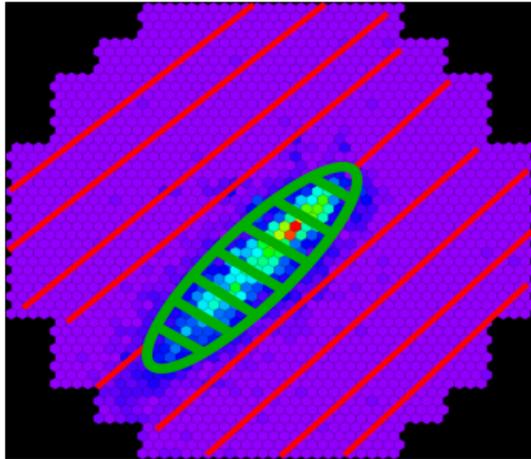




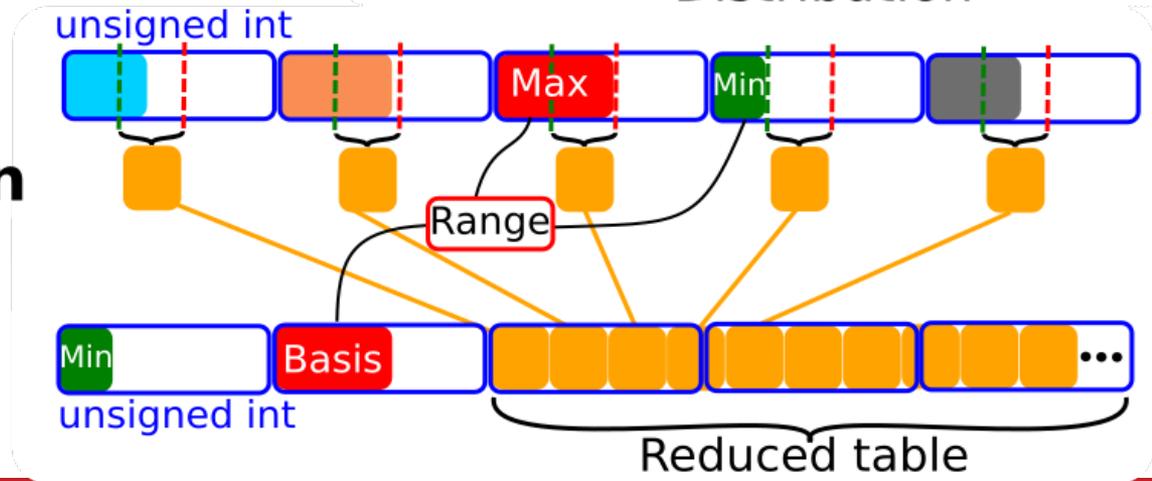








**Publication ongoing**



# Raw data compression and speed

## Test File

- 475 MB

- Up to 99% ADC values

	Compression ratio	Time	File size (MB)
LZMA (7Z)	4.84	7 min 48s	98
Advanced Polynomial Compression	3.74	3.7 s	127

# Raw data compression and speed

## Test File

- 475 MB

- Up to 99% ADC values

	Compression ratio	Time	File size (MB)
LZMA (7Z)	4.84	7 min 48s	98
Advanced Polynomial Compression	3.74	3.7 s	127
Advanced Polynomial Compression + LZMA	4.84	24.646 s	98

# Raw data compression and speed

## Test File

- 475 MB
- Up to 99% ADC values

	Compression ratio	Time	File size (MB)
LZMA (7Z)	4.84	7 min 48s	98
Advanced Polynomial Compression	3.74	3.7 s	127
Advanced Polynomial Compression + LZMA	4.84	24.646 s	98

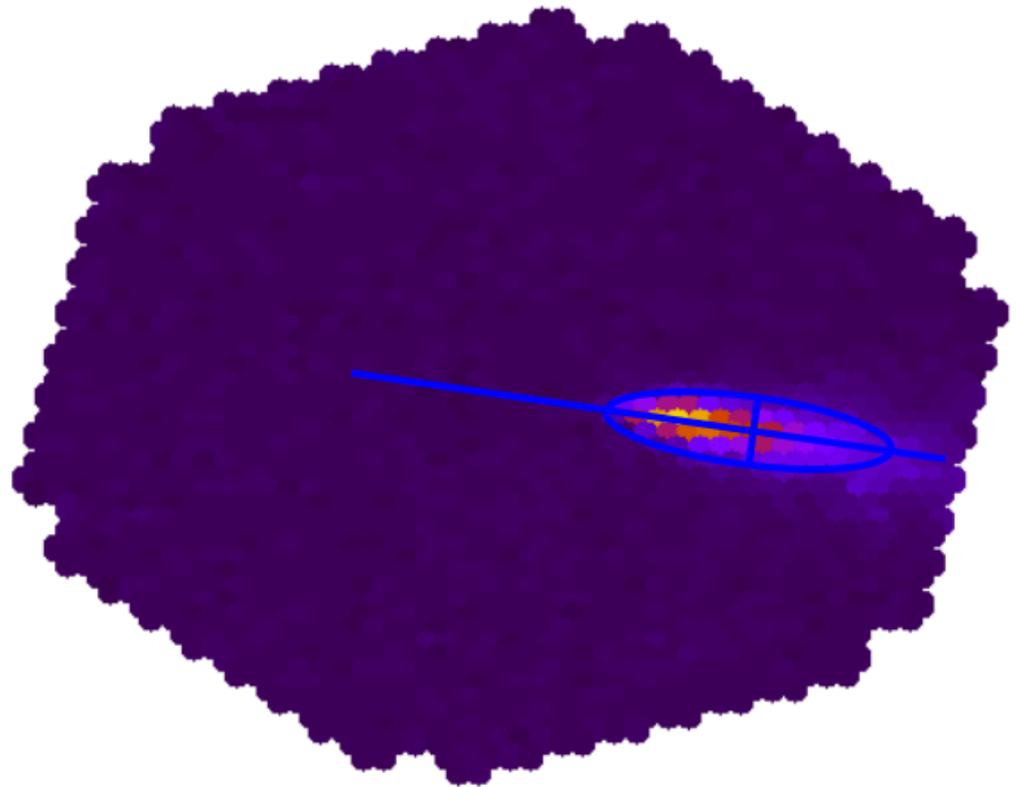
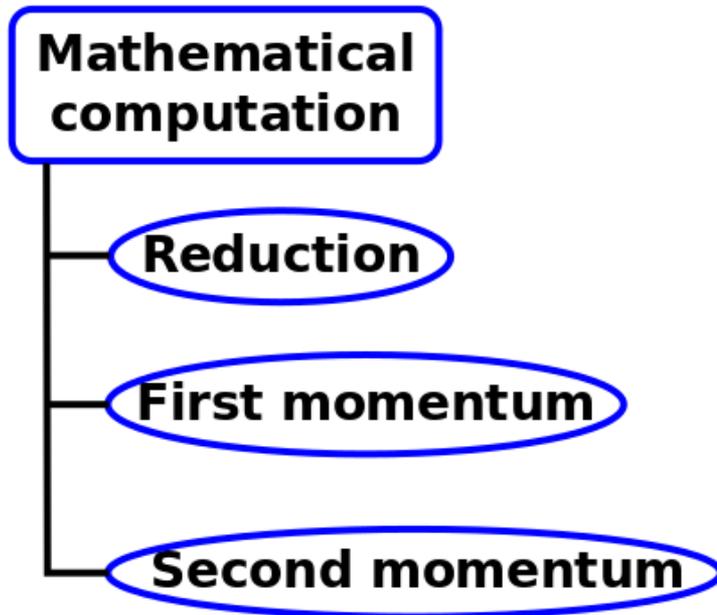
Same Compression ratio

19 × faster

# High Performance Computing

- Improve existing algorithms
- Develop new ones

# Hillas method



# Hillas with optimized data format

**Optimized data format**

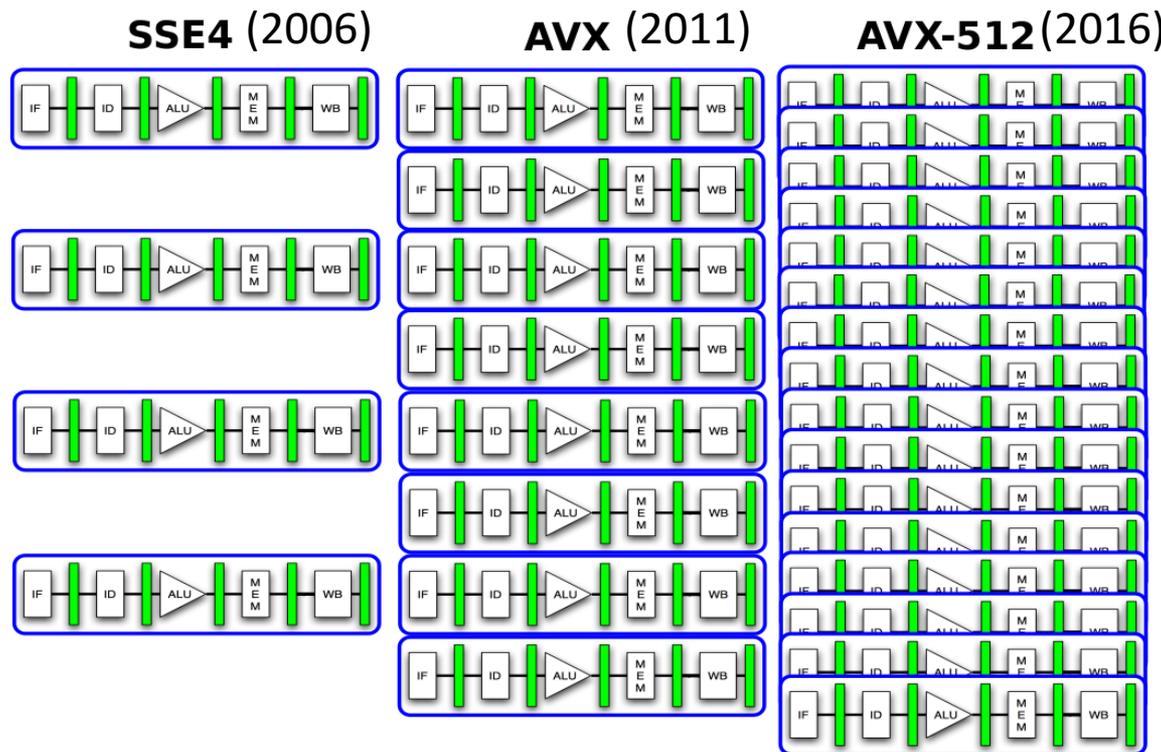
**Allows CPU data prefetching**

**Contiguous data**

**Cache friendly**

	Speed (cy/el)	Speed up
Initial data format (H.E.S.S. ROOT)	2125.5	1
<b>Optimized data format</b>	53.1365	<b>40.0</b>

# Vectorisation - CPU Architecture



# Reduction Optimization

**Reduction : Sum of all vector elements**

	Speed (cy/el)	Speed up
Classical	<b>2.69842</b>	1
Vectorized (GCC, SSE4)	0.702845	3.8

# Reduction Optimization

**Reduction : Sum of all vector elements**

	Speed (cy/el)	Speed up
Classical	<b>2.69842</b>	1
Vectorized (GCC, SSE4)	0.702845	3.8
Intrinsic Vectorized SSE4	0.226675	11.9
<b>Intrinsic Vectorized AVX</b>	<b>0.11379</b>	<b>23.7</b>

## Hillas with optimized data format

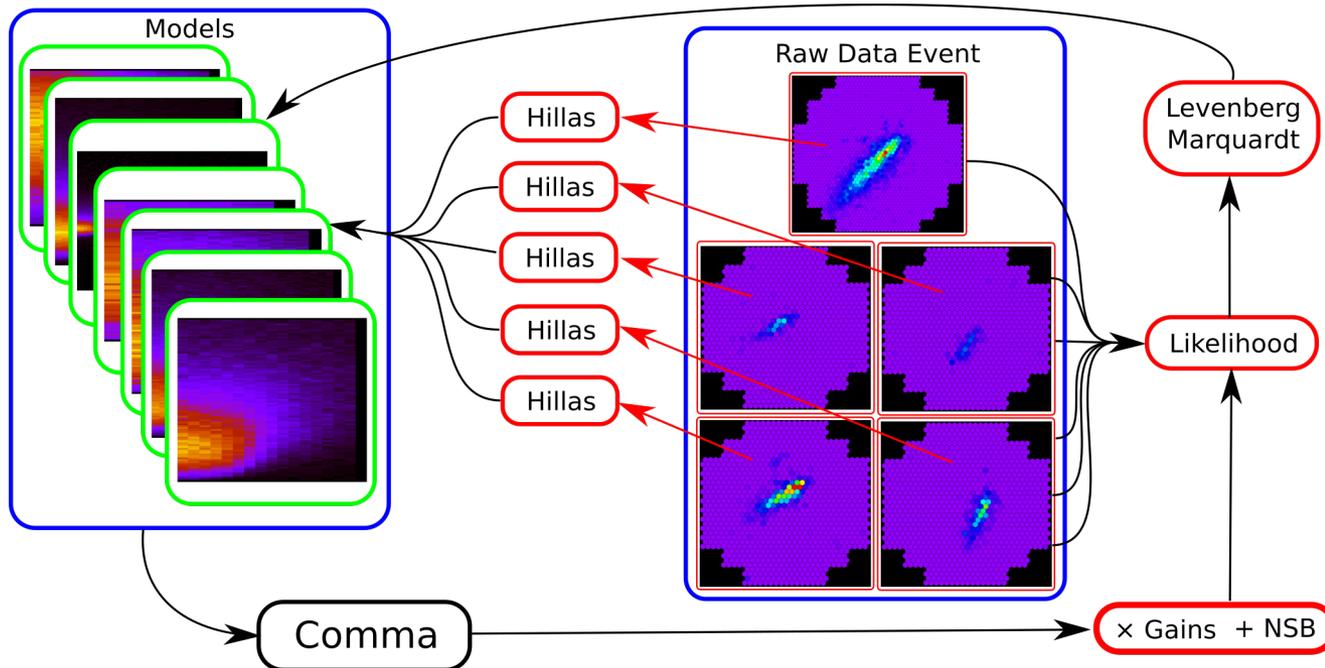
Data format  
**×40**

Vectorization of  
the reduction

	Speed (cy/el)	Speed up
Initial data format	2125.5	1
<b>Optimized data format</b>	53.1365	<b>40.0</b>
+ Intrinsic Vectorized SSE4	6.39931	332
<b>+ Intrinsic Vectorized AVX</b>	<b>2.98499</b>	<b>712</b>

# Template analysis

- Used in H.E.S.S. (Model++)



- Very good physics results, tested and approved

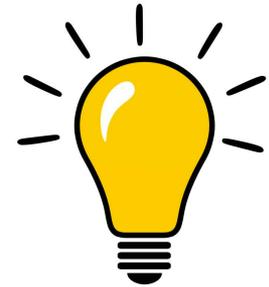
# Template analysis for CTA

## Main issues

- Data organized per telescopes and not per events
- Pixel-to-pixel comparison is expensive

# Template analysis for CTA

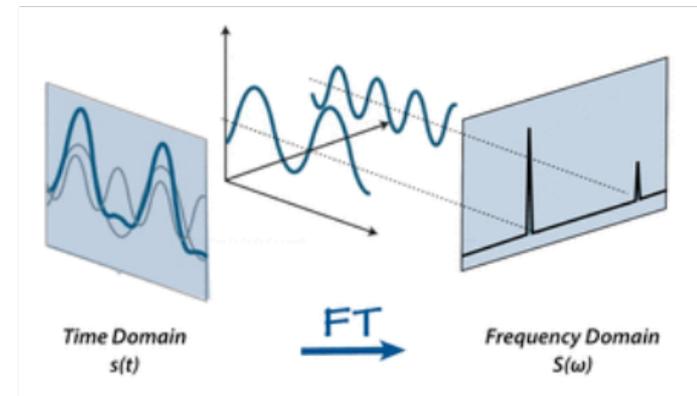
## A solution: data reduction



- Easier to re-organise data per events
- Comparing keys is much faster

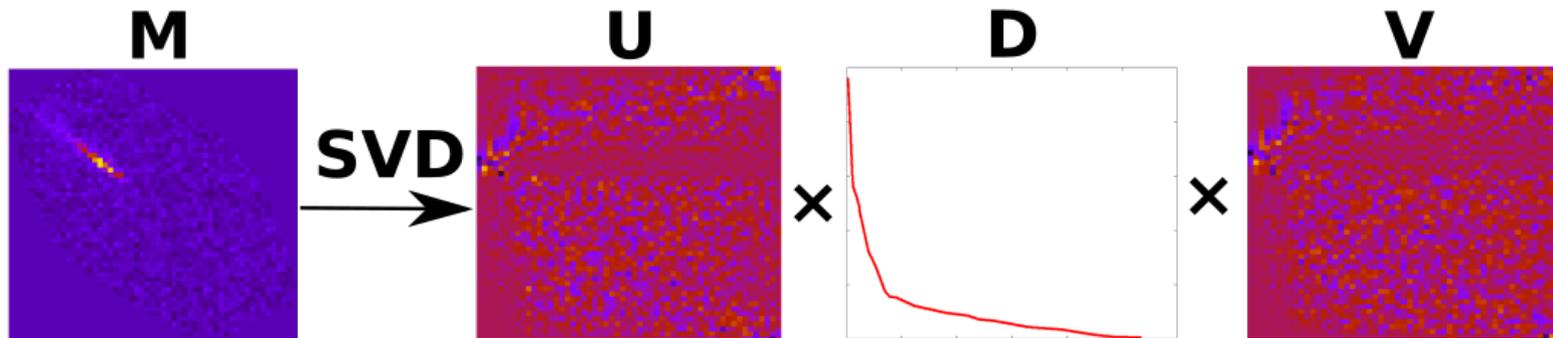
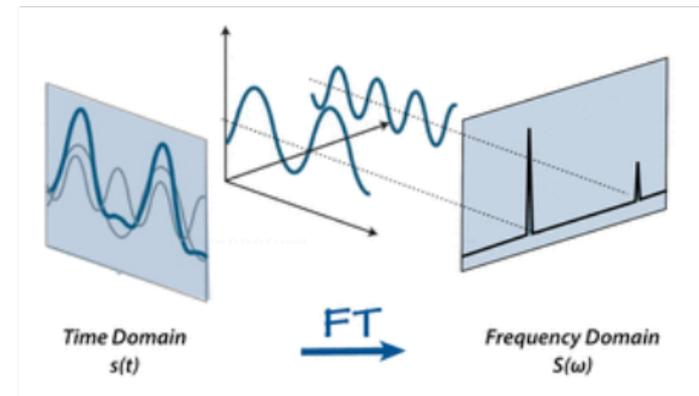
# Singular Value Decomposition

## Analogy with Fourier Transform



# Singular Value Decomposition

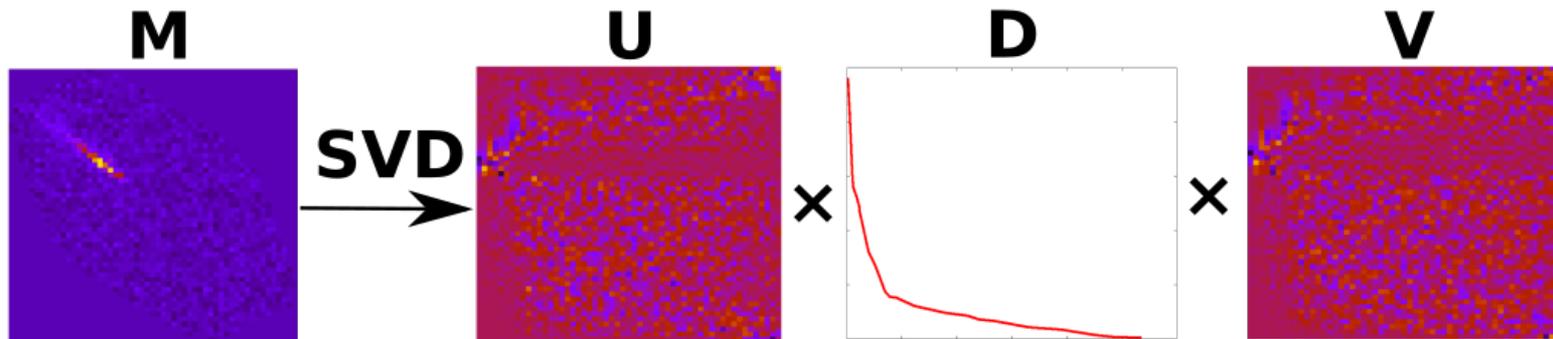
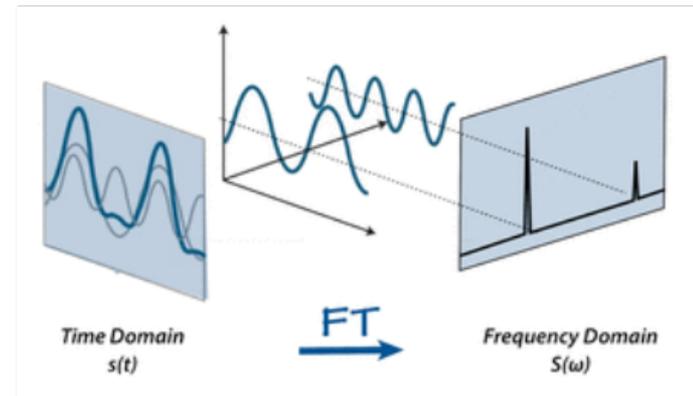
Analogy with Fourier Transform



# Singular Value Decomposition

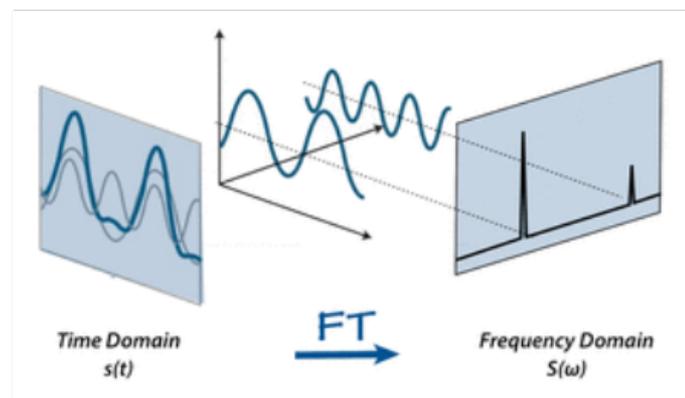
Analogy with Fourier Transform

**Scalable**  
**Already vectorized**



# Singular Value Decomposition

Analogy with Fourier Transform



**Scalable**  
**Already vectorized**

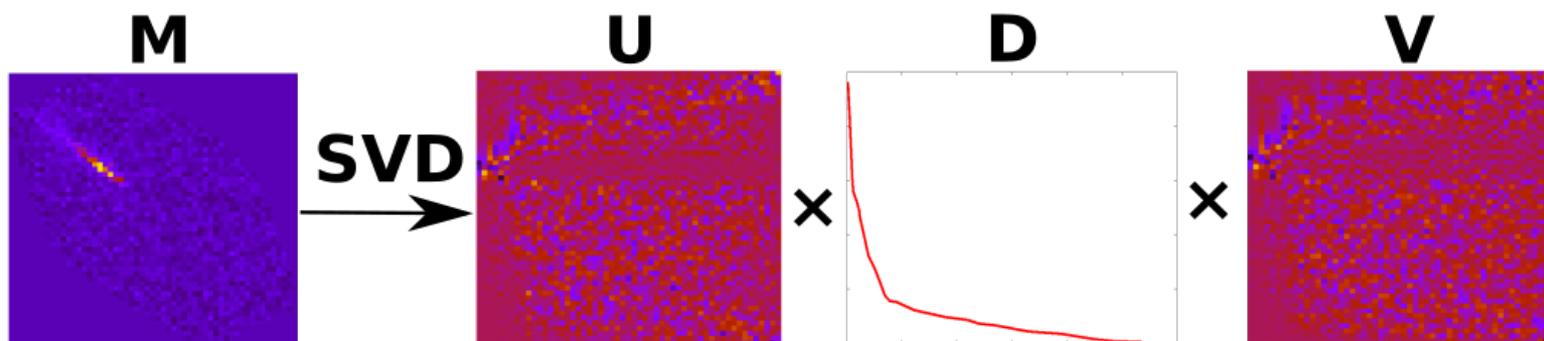
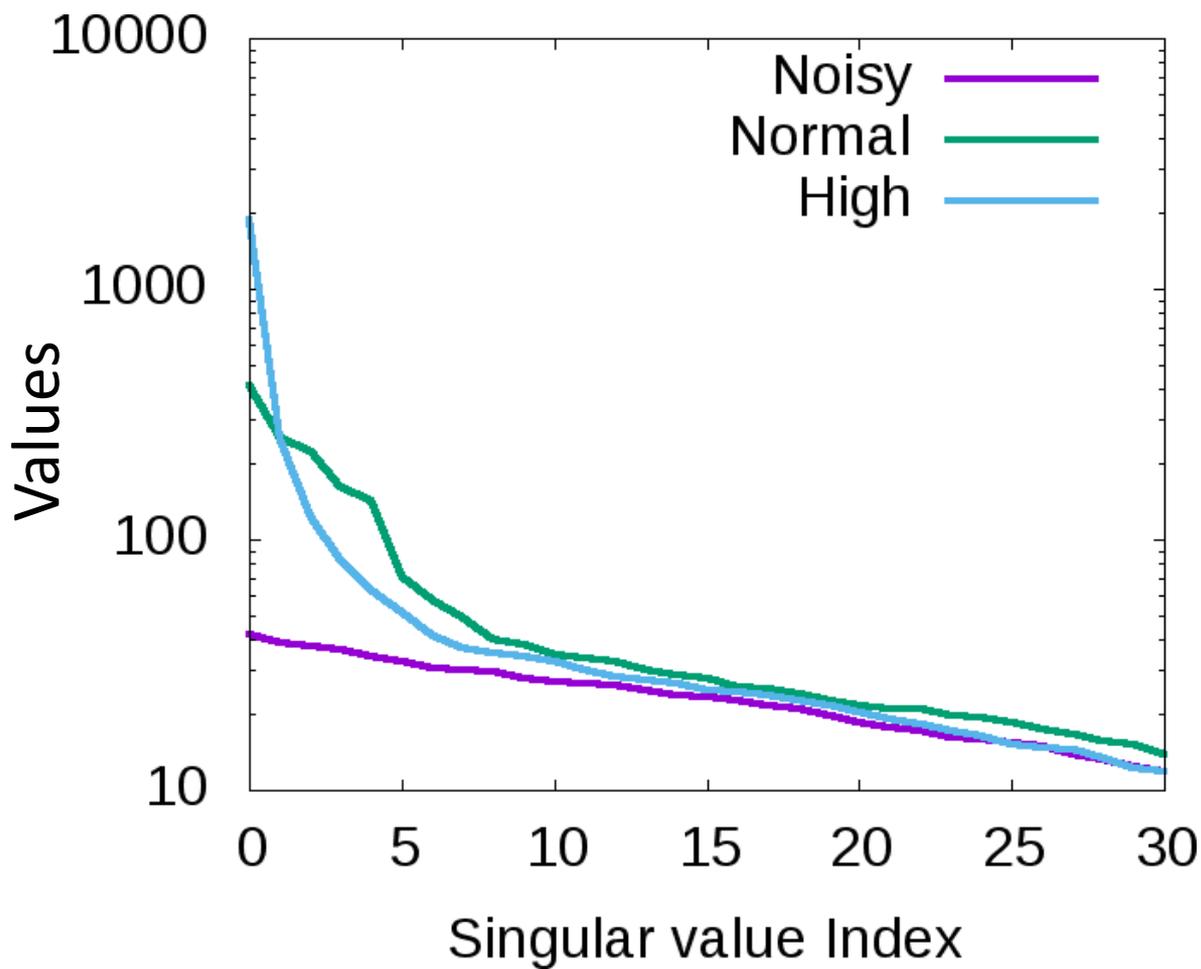
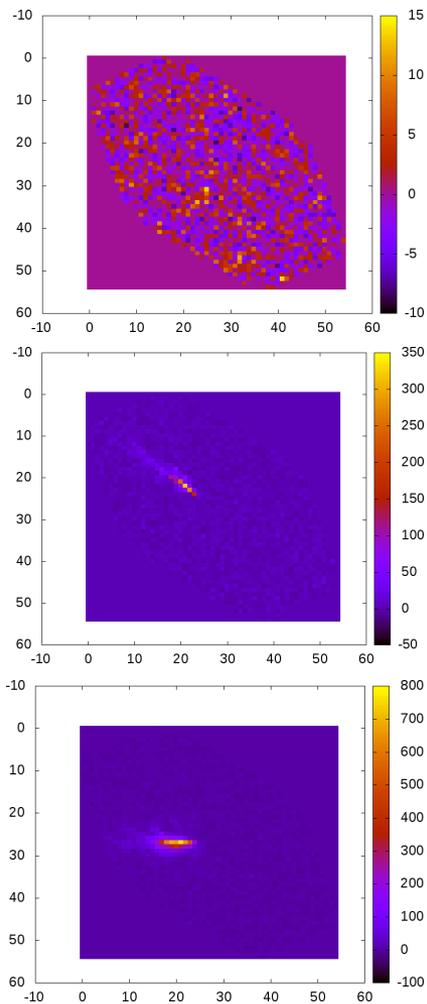
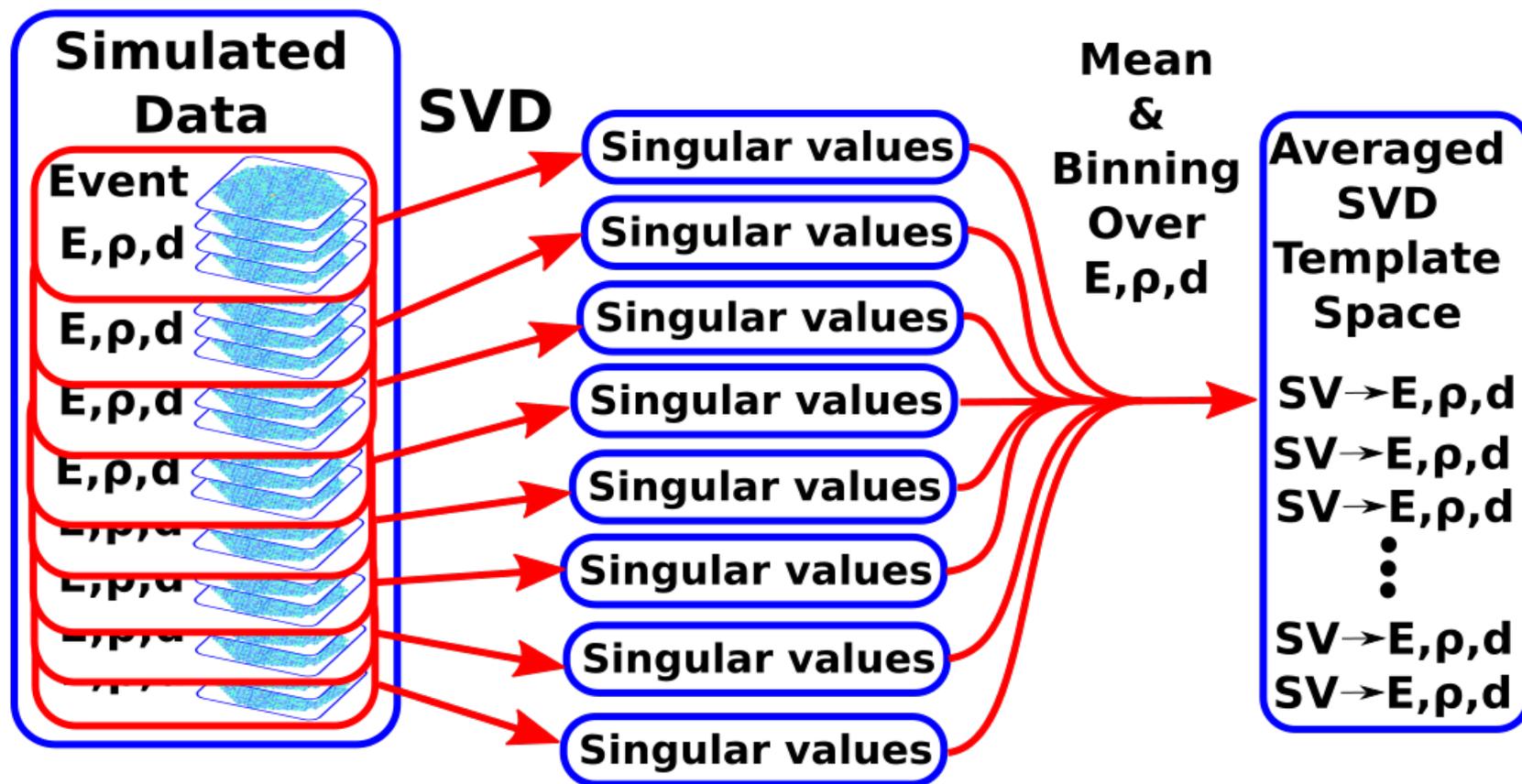


Image Search : Google Images, Wolfram



# SVD Analysis template creation



- Astronomy is entering the big data era and facing new challenges (data volumes, data rates, real-time...)
- High Performance Computing to the help
  - Compression
  - Vectorization for Hillas based reconstruction
  - New reconstruction methods (e.g. SVD)
    - Scalable and vectorised
    - Hopefully better physics