ML for FEL background reduction

Using Robust Statistics and Machine Learning methods to help with data reduction in FEL data analysis pipeline

Alireza Sadri
Hamburg, 18.12.2019

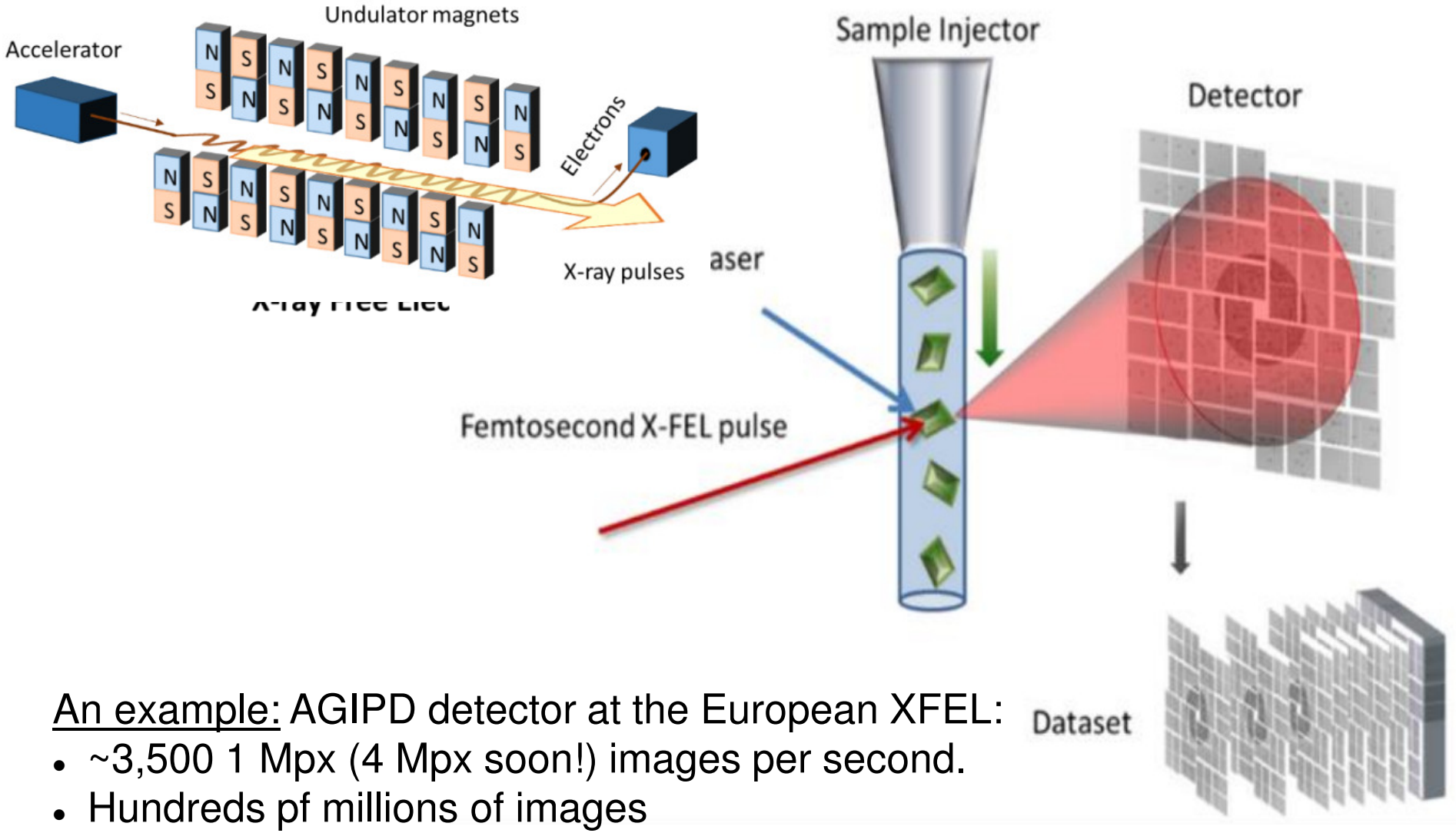
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY.

Agenda

- FEL experiments and data volume problem
- FEL hit and non-hit detection
- Detectors pixels performance
- Normal and Abnormal behavior
- Bad pixel masks
- Modelling the background
  - Statistical Models
  - Towards DL

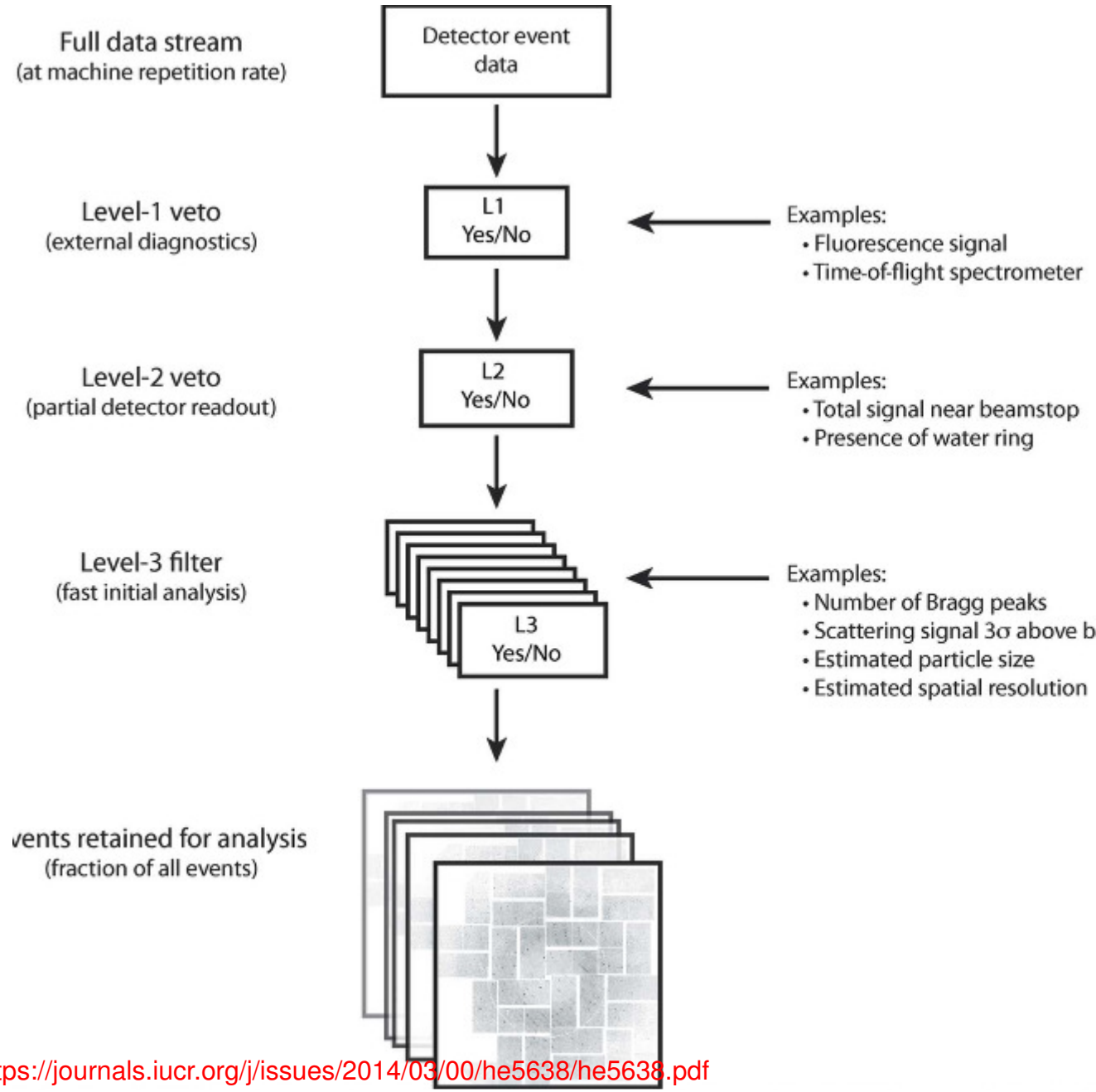
FEL Data volume challenge



An example: AGIPD detector at the European XFEL:

- ~3,500 1 Mpx (4 Mpx soon!) images per second.
- Hundreds pf millions of images
- We need maybe thousands of good images...

Data volume



Data correction by calibration constants

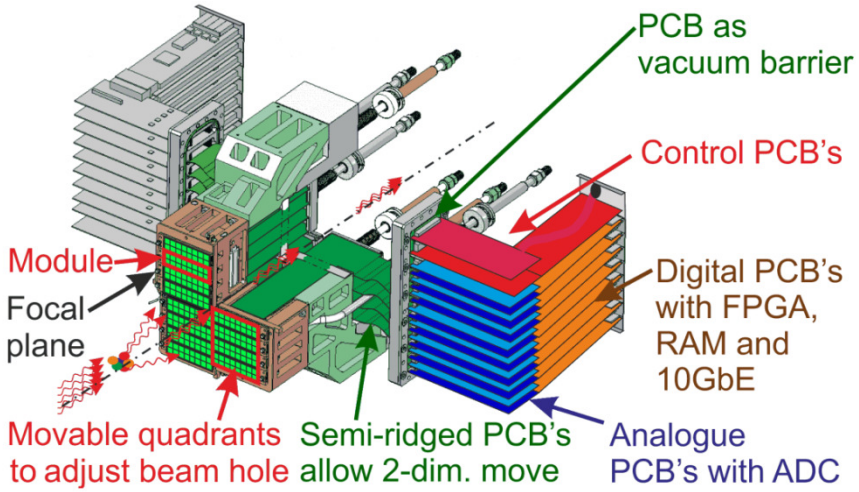
- ✓ Bad pixel mask
- ✓ Pedestal correction
- ✓ Gain correction

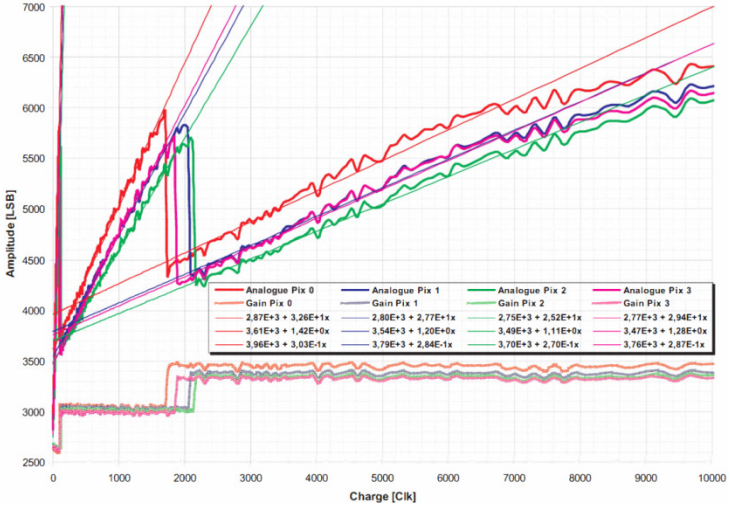
Runtime pre-processing

- ✓ Artefacts detection and segmentation
- ✓ Common mode removal

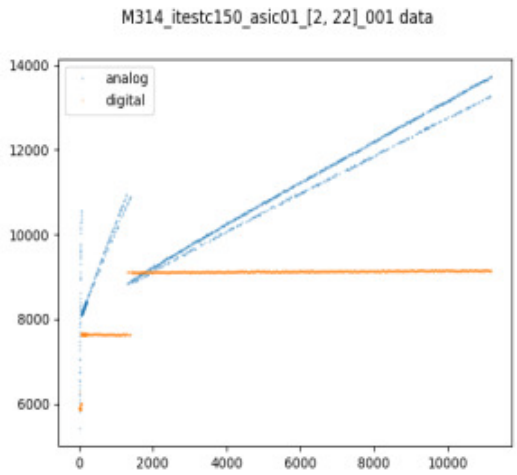
https://journals.iucr.org/j/issues/2014/03/00/he5638/he5638.pdf

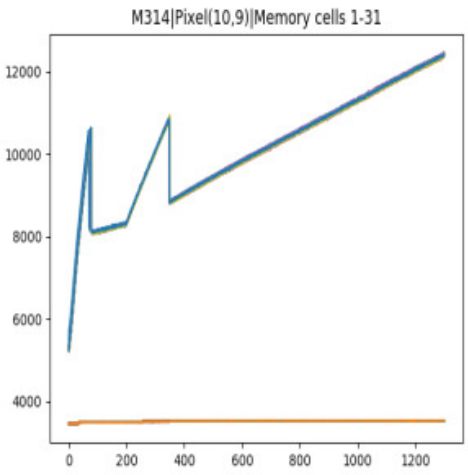
A detector's pixel: AGIPD

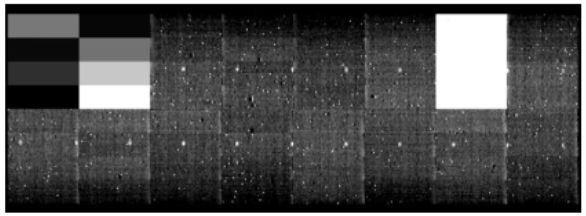




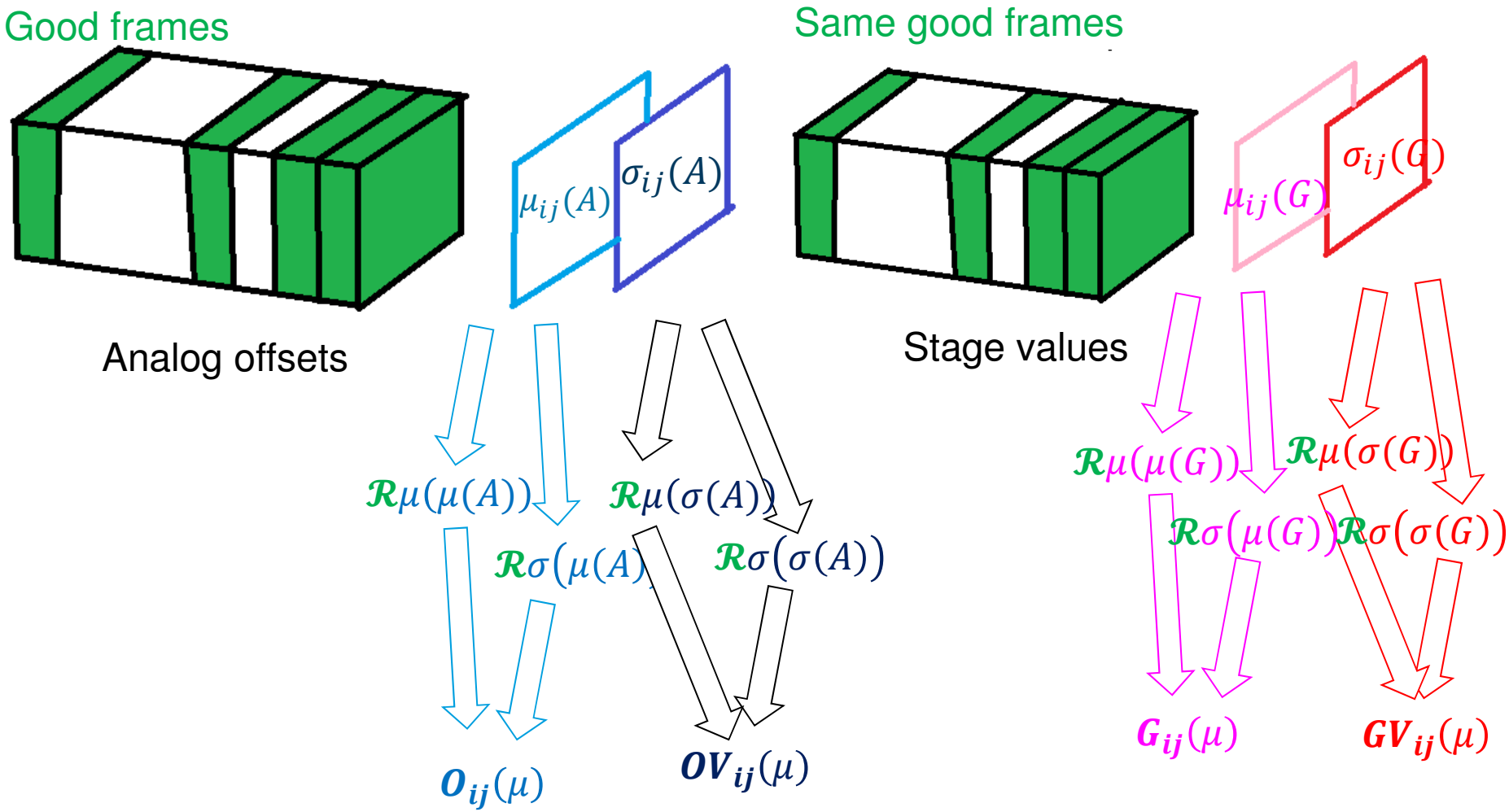
Abnormalities in pixels







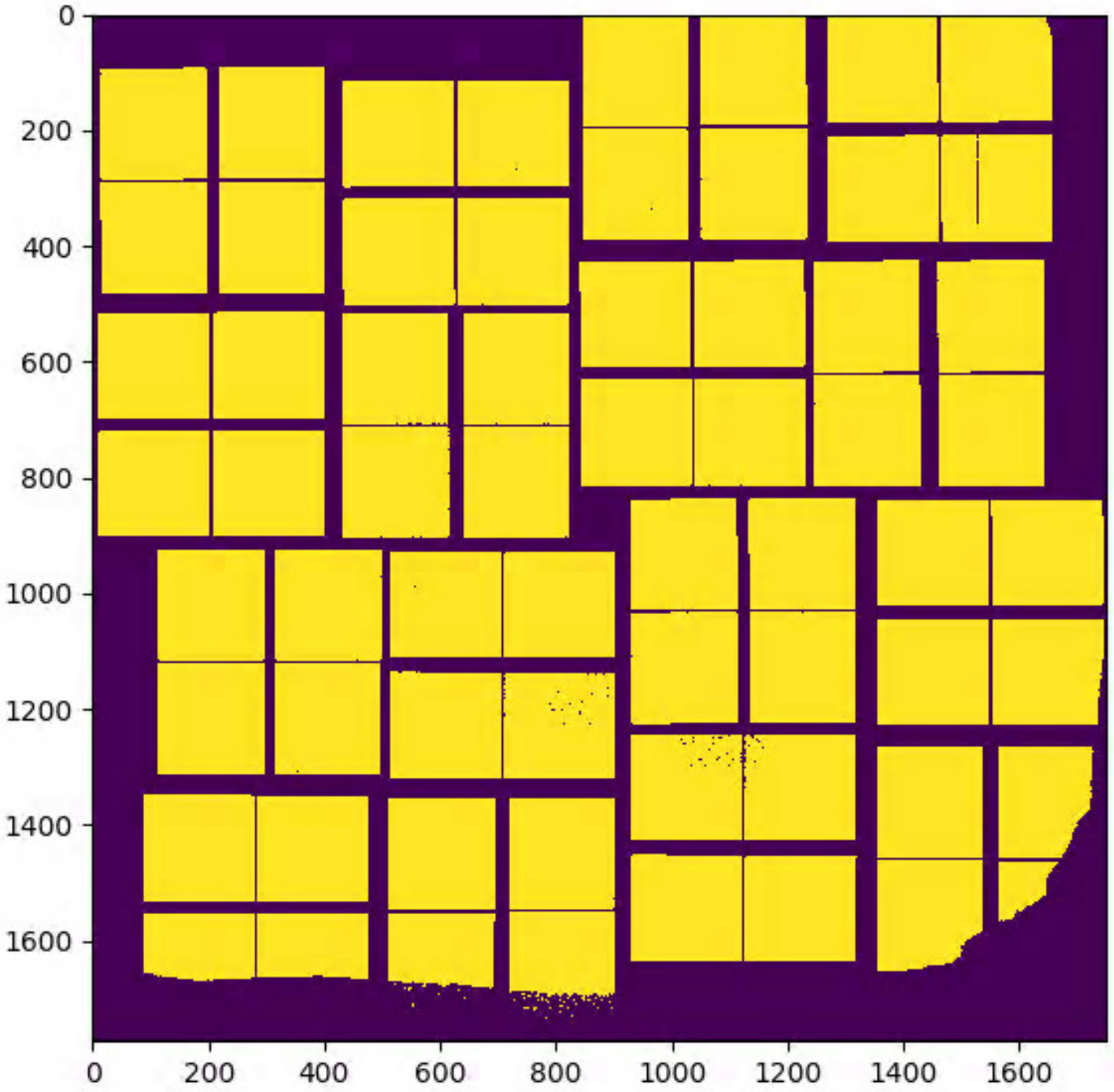
Abnormality maps



$\mathcal{R}$ means a robust function that trusts only normal pixels

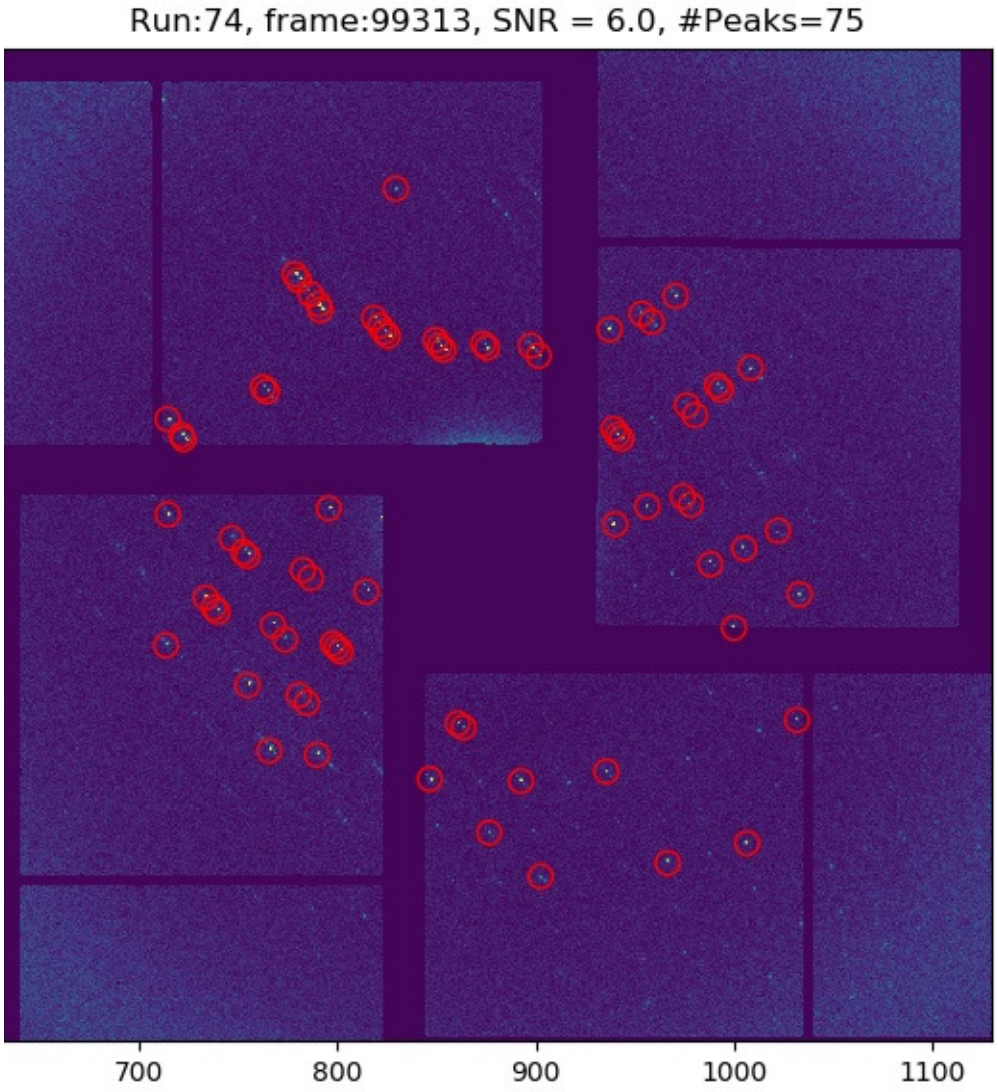
Bad pixel mask

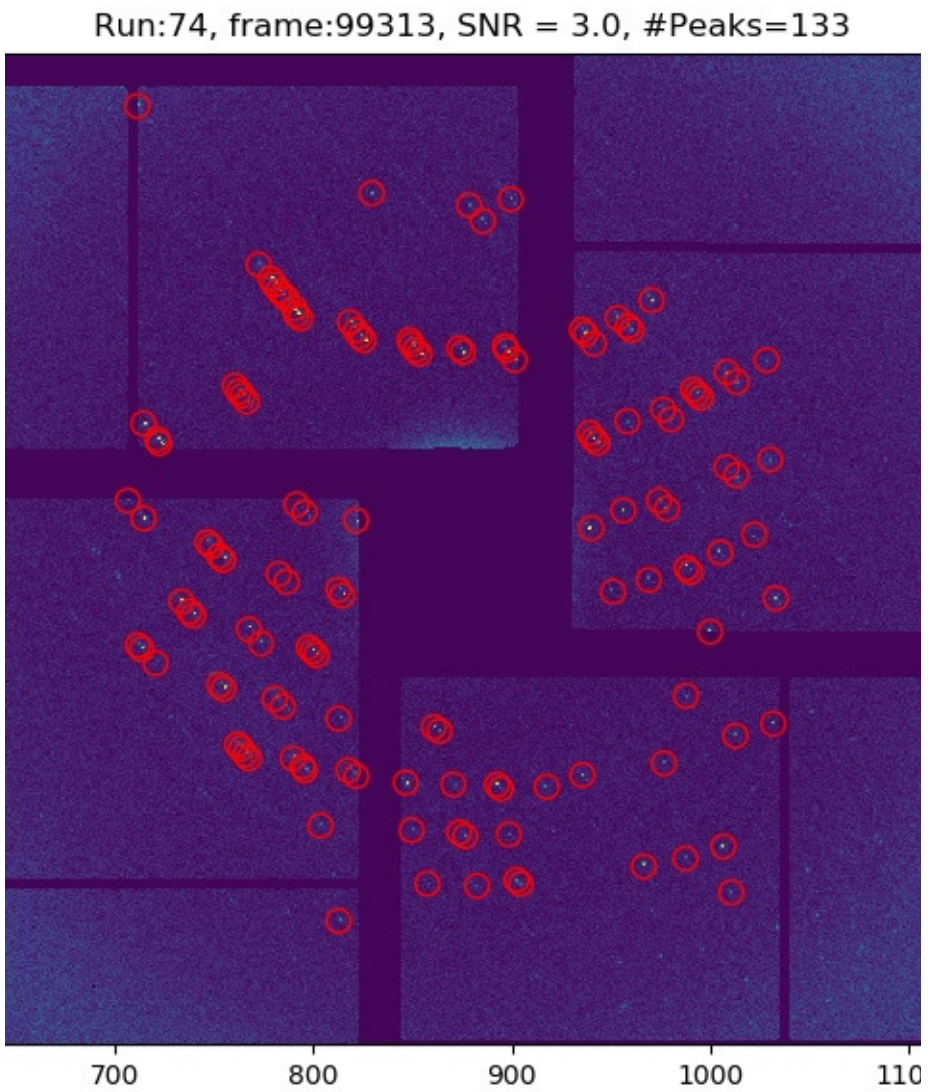
The tool is complete and tested for AGIPD (EuXEFL) and CSPAD (Old LCLS)



Diffraction pattern of a hit

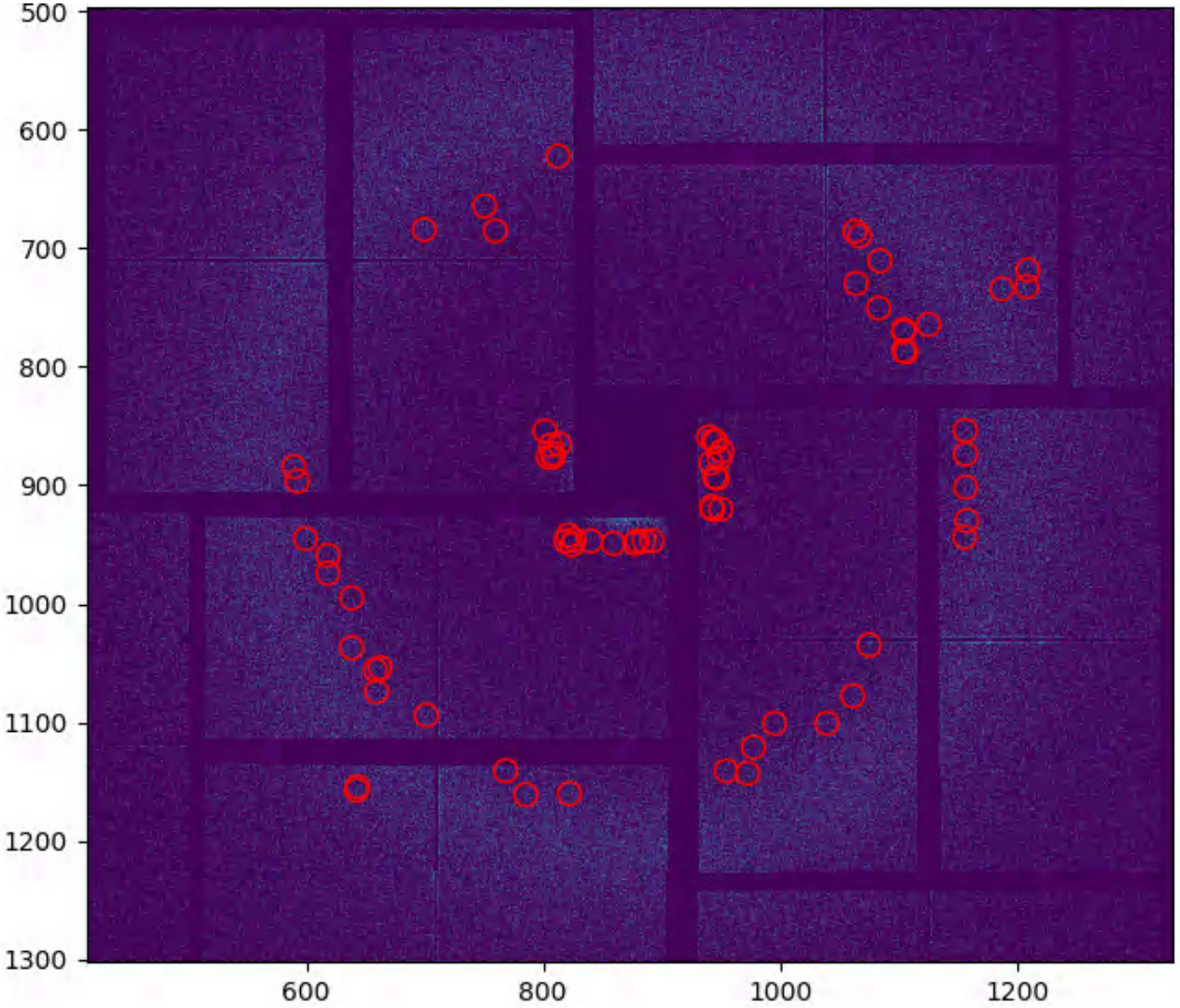
Data from CXI-DB32 – LCLS 2015 – CSPAD detector



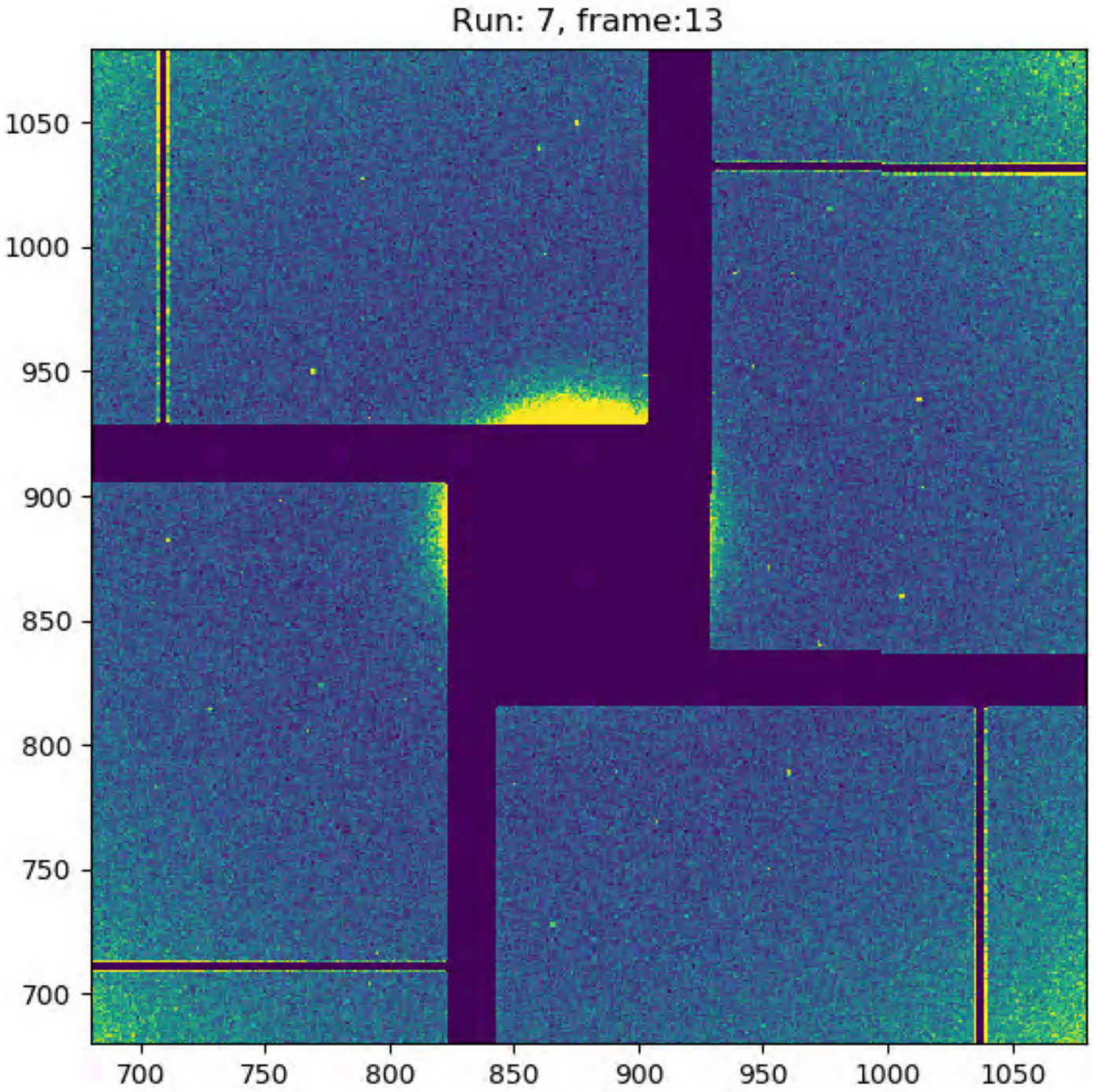


A non-hit with a too sensitive peakfinder

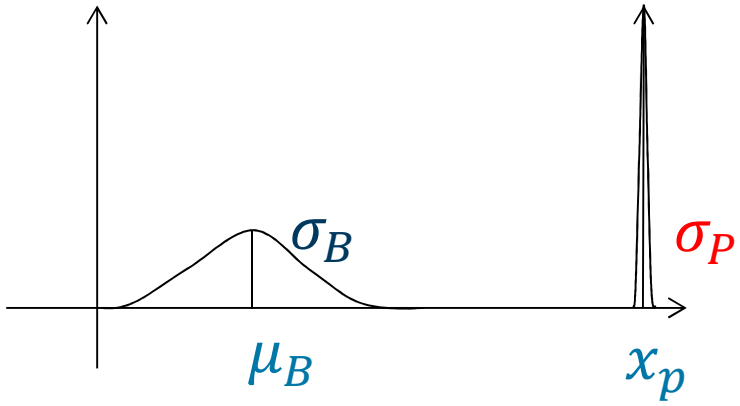
No data reduction this way



A weak pattern



A weak pattern

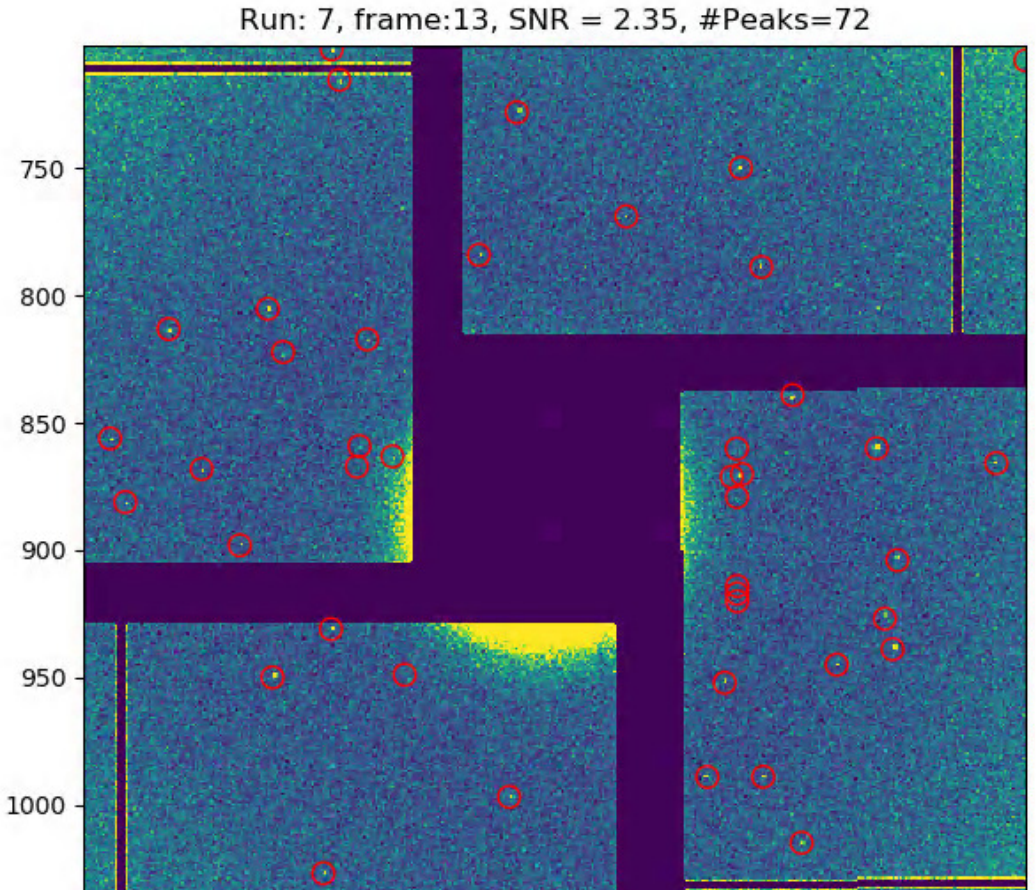


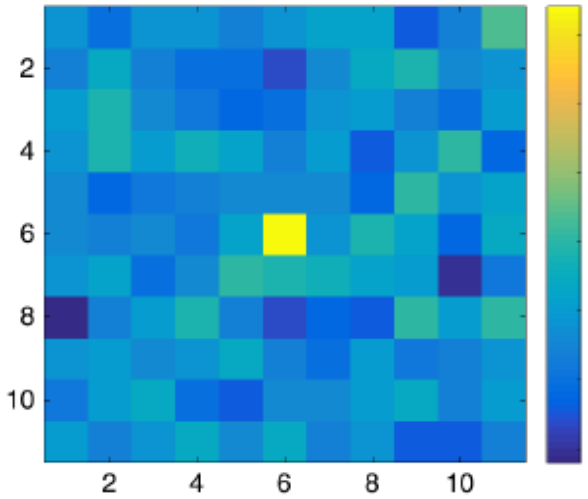
$$SNR = \frac{\mu_P - \mu_B}{\sigma_P + \sigma_B}$$

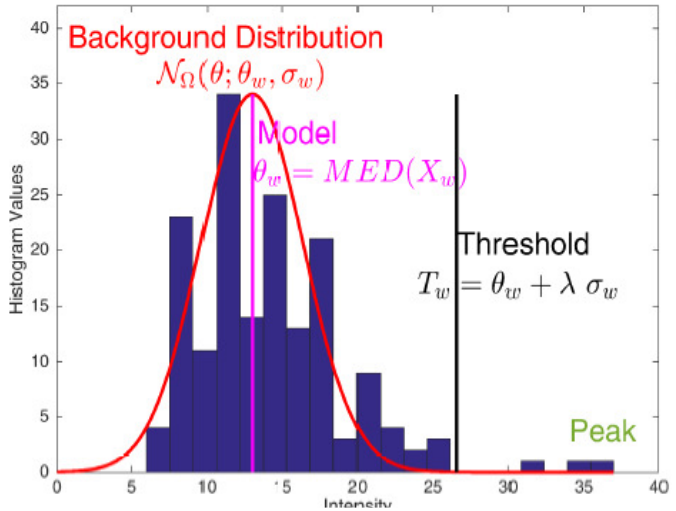
σ_P: pixel‘s uncertainty

⇑

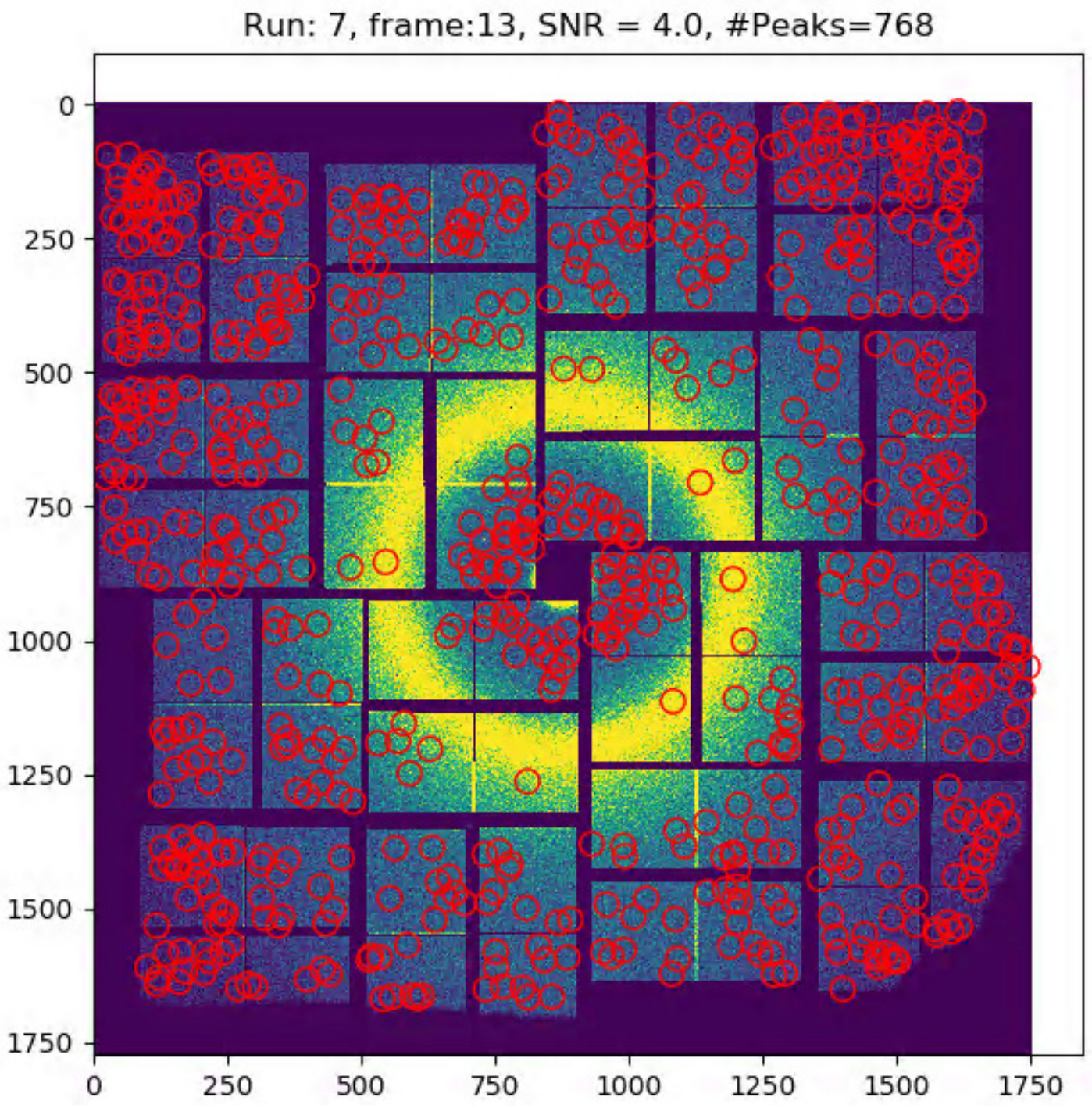
Can be learned



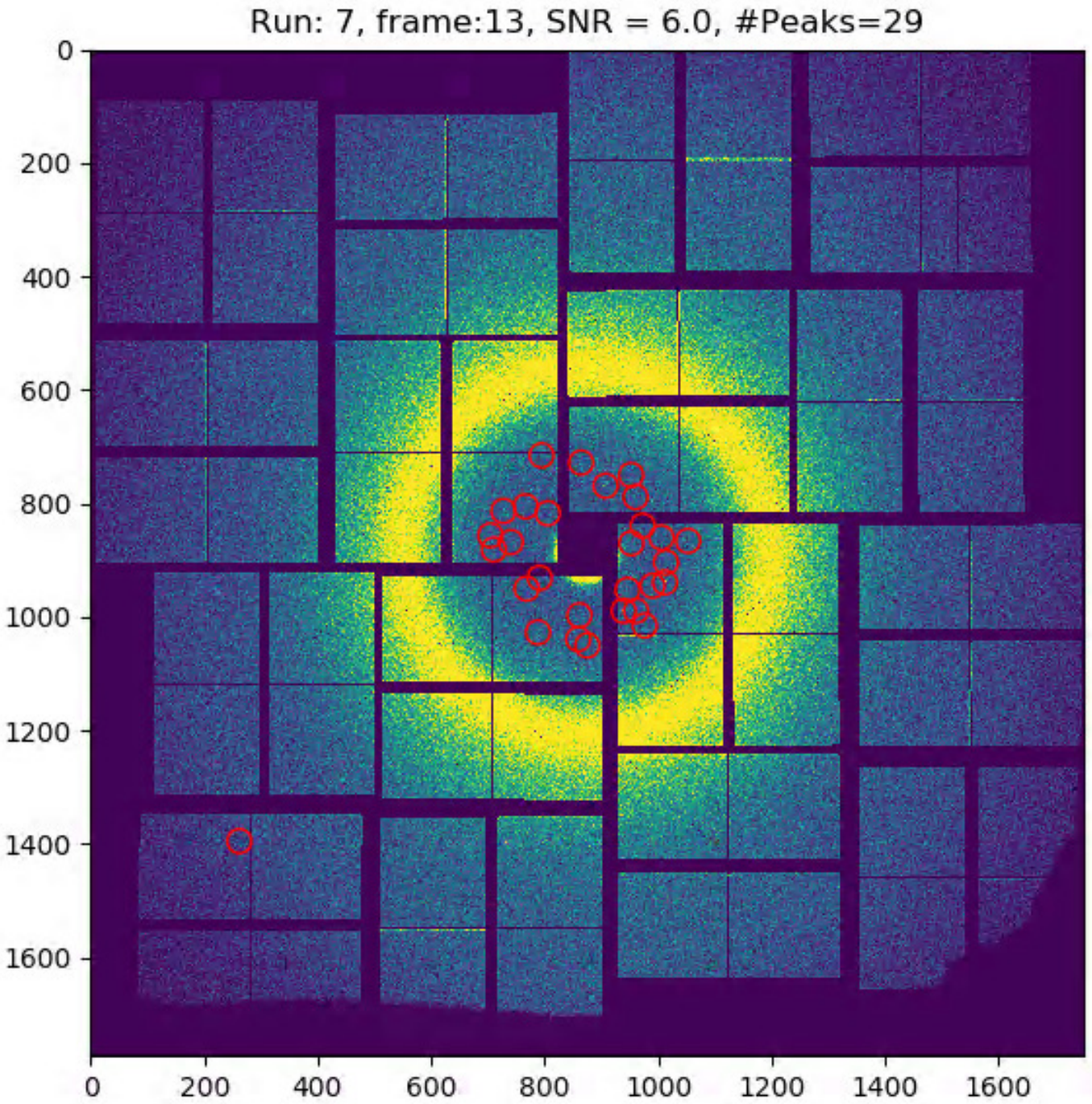




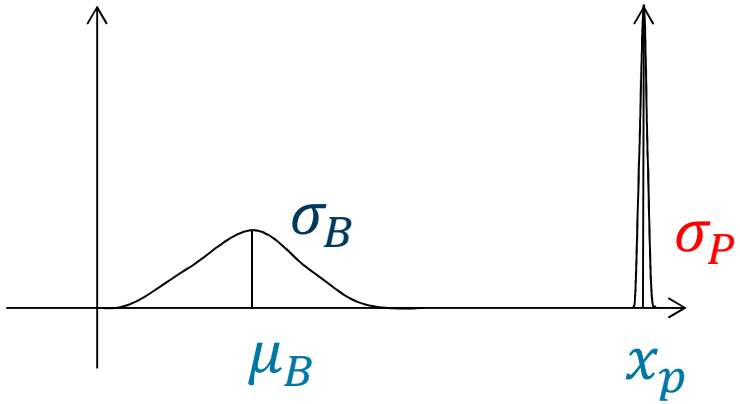
Low SNR and no σ_P uncertainty modelling



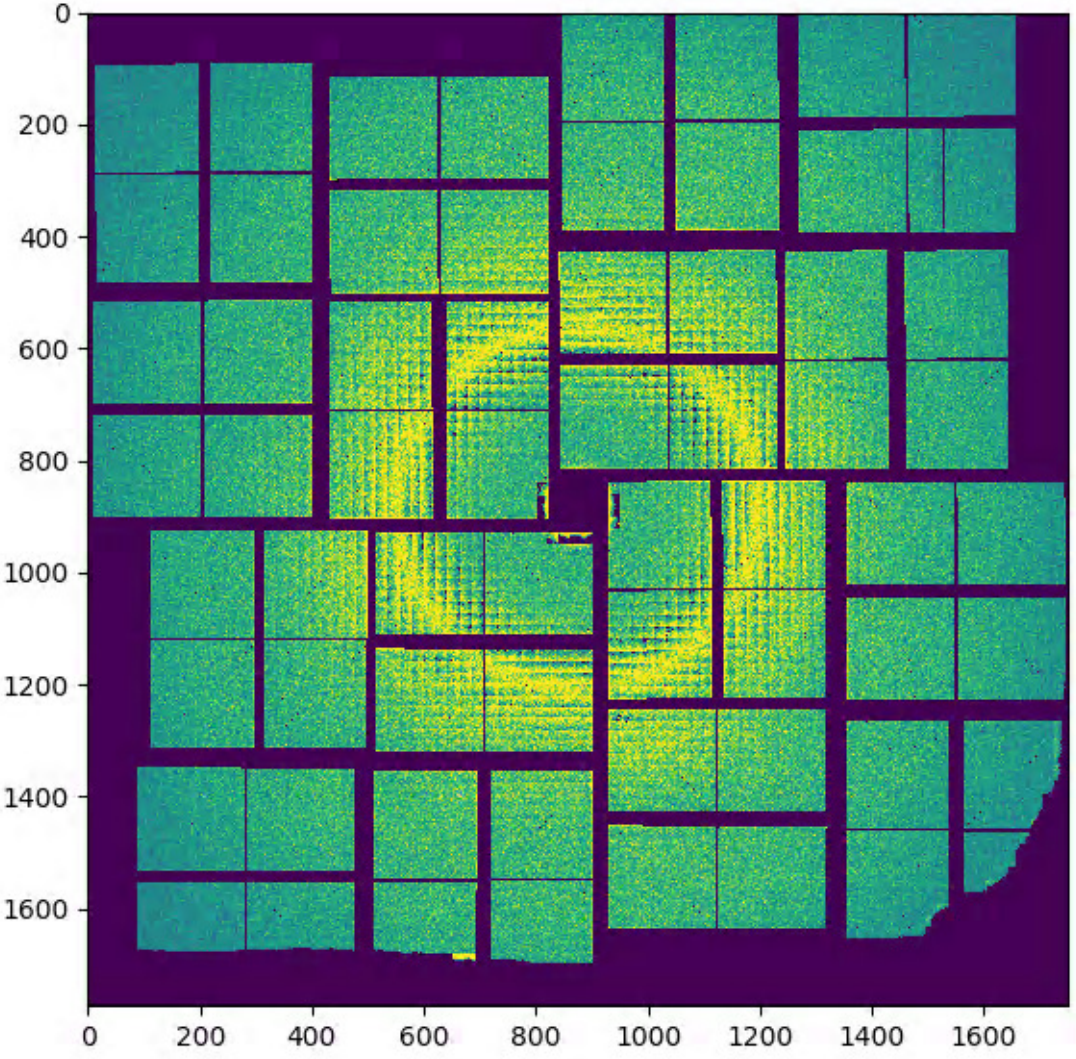
High SNR and no σ_P uncertainty modelling



Pixels uncertainties



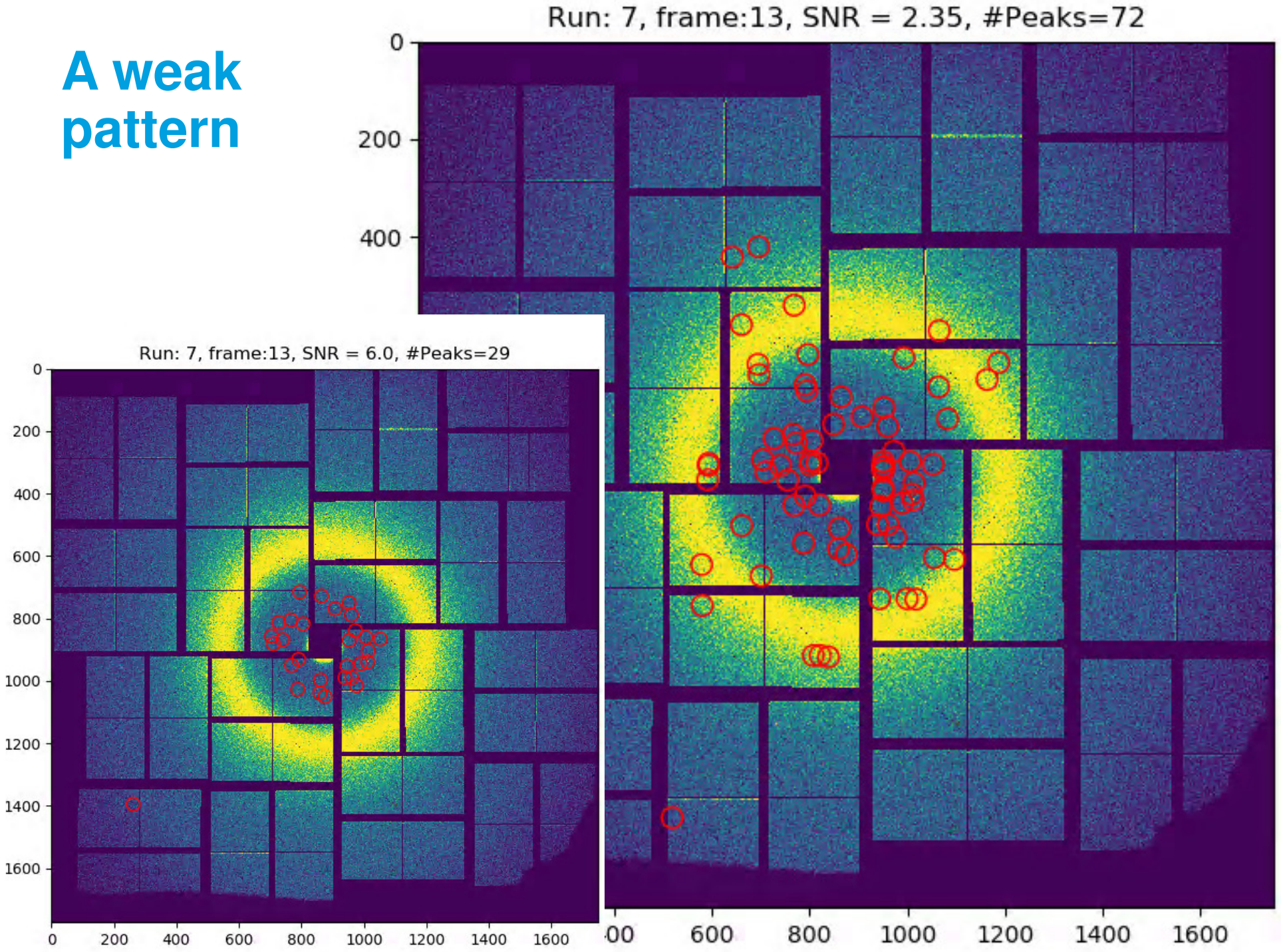
$$SNR = \frac{\mu_P - \mu_B}{\sigma_P + \sigma_B}$$



Prior σ_P:

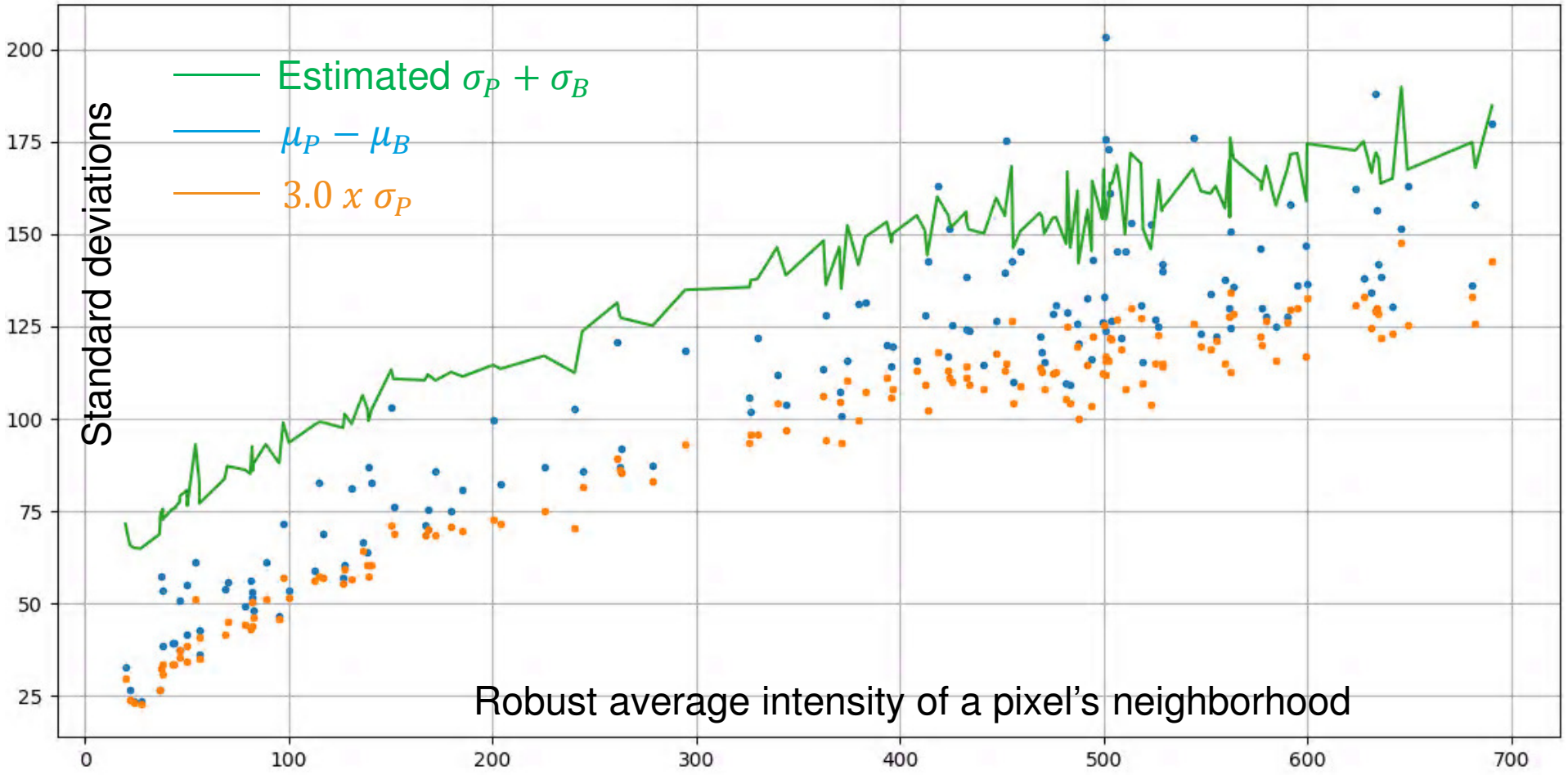
- Statistical
- How to use deep learning?

A weak pattern



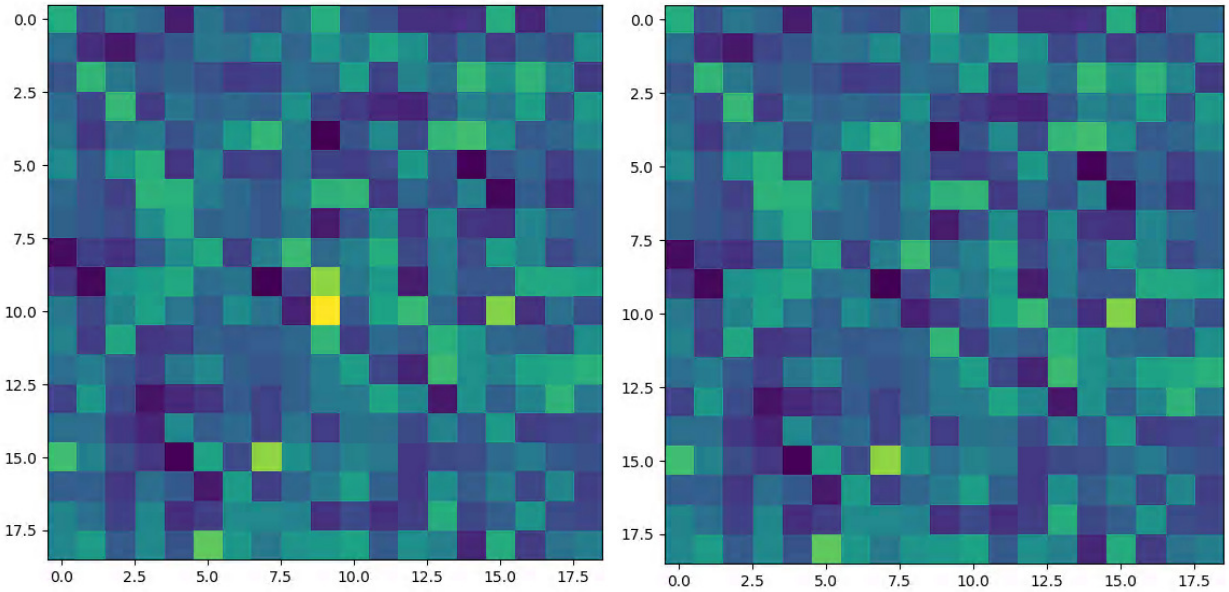
Pixels uncertainties statistical modeling: σ_P

- Image is for one pixel. Each data point is for an image. (25000 of them)
- Data: CXIDB32-LCLS-Nov-2014-CSPAD..
- Only frames where SNR of the pixel is above 3 are shown.

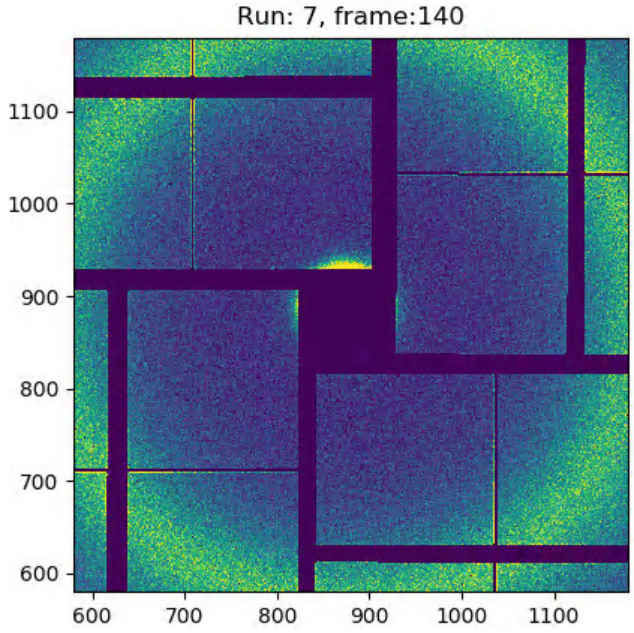


How to use deep learning?

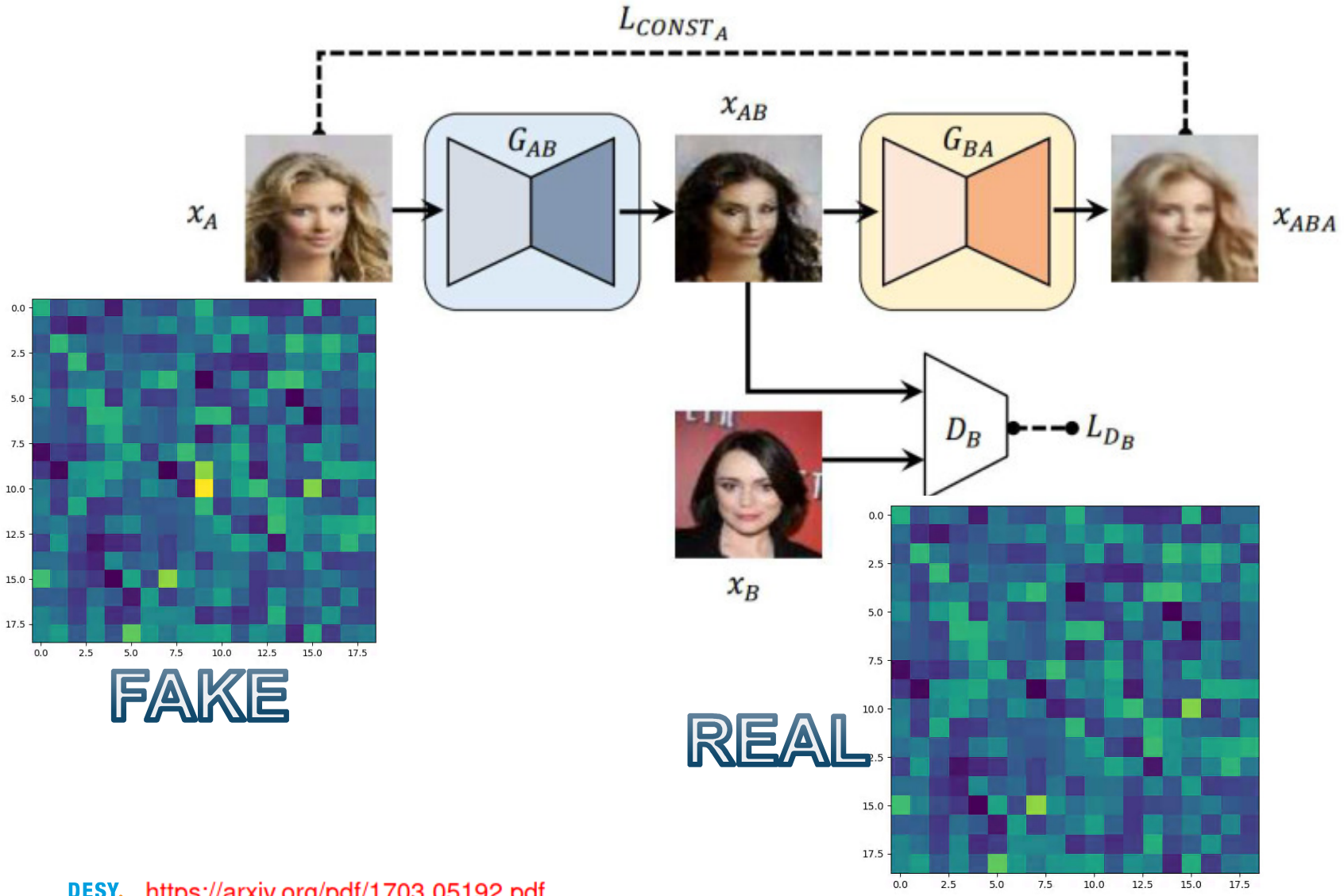
- Adversarial training



- I am trying pix2pix currently
- Questions:
  - How to divide image into parts ?
  - How many neural networks?
    - Make transfer learning possible?
  - How many Bragg peaks?
    - How large the training dataset?
- Anomaly detection using non-hits



Adversarial training by pix2pix



Thank you

Contact

DESY. Deutsches
Elektronen-Synchrotron

www.desy.de

Alireza Sadri
CFEL-DESY
E-mail Alireza.Sadr@desy.de