ESR 3: Real time analysis strategies for reconstruction, exotic physics, and market analysis

Leon Bozianu

Anna Sfyrla, Steven Schramm

27.11.2023

We acknowledge funding by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086 (SMARTHEP).

Presentation Outline

- **Qualification task**
  - **Object detection in ATLAS Trigger**
- **Secondment**
  - **Recurrent & Bayesian neural networks**
- **Other activities**

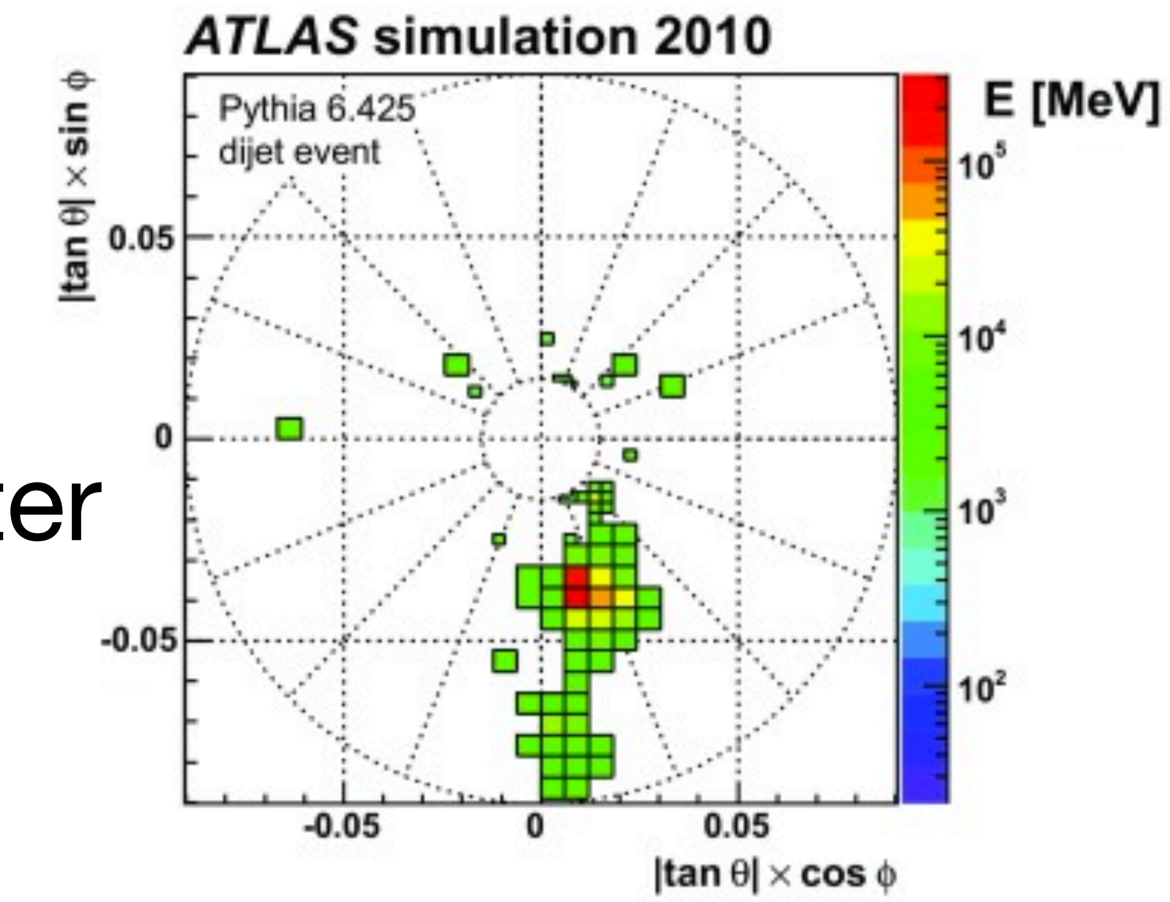
Qualification Task

https://arxiv.org/abs/1603.02934

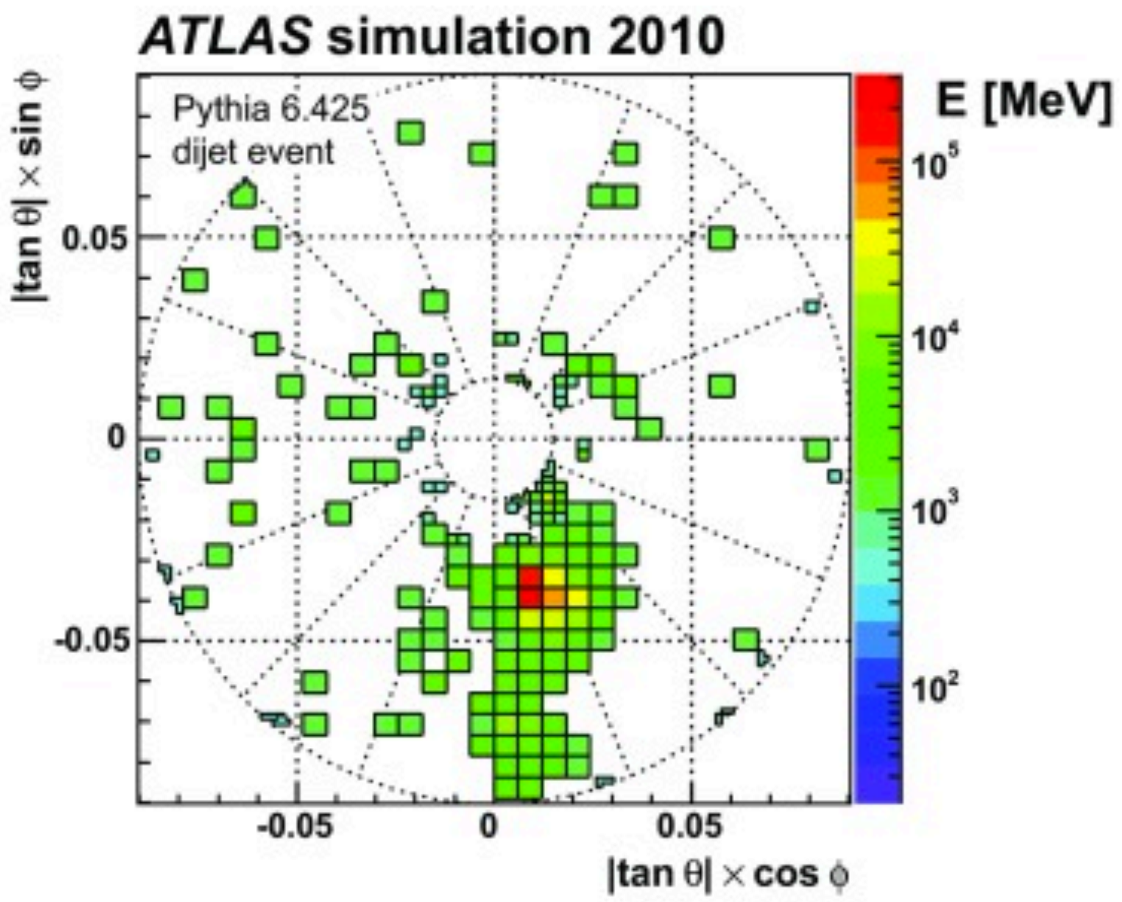
Topoclustering

Making of jet constituents

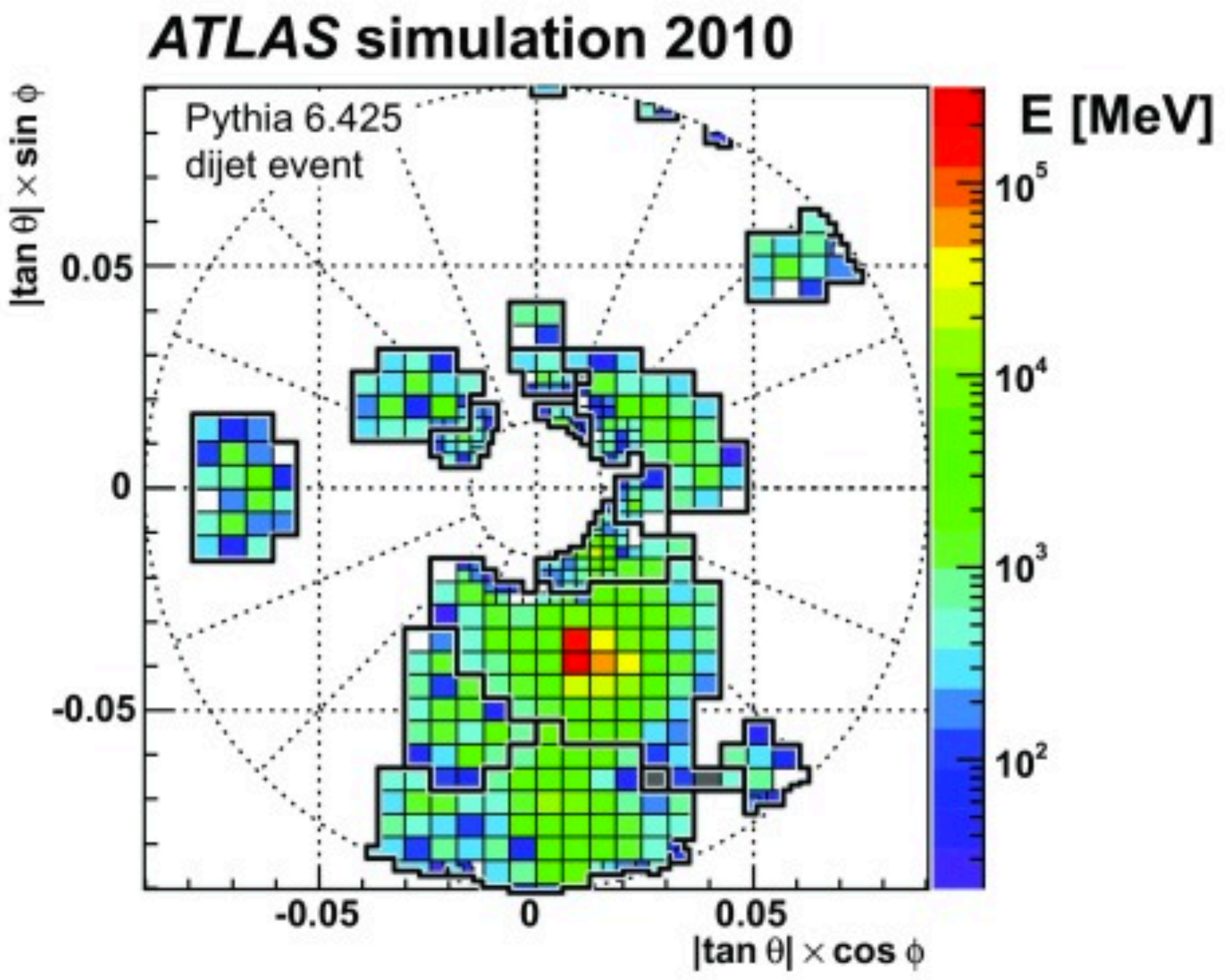
- ATLAS uses the **topoclustering** algorithm to cluster calorimeter cells together.
- The algorithm is **iterative**, it checks each cell in turn.
- It then checks all the neighbouring cells.
- This guarantees you “find” everything, but is very **slow**!
- The clusters go through several **post-processing** steps and are then used to make **jets**.



(a) Cells passing selection in Eq. (3)



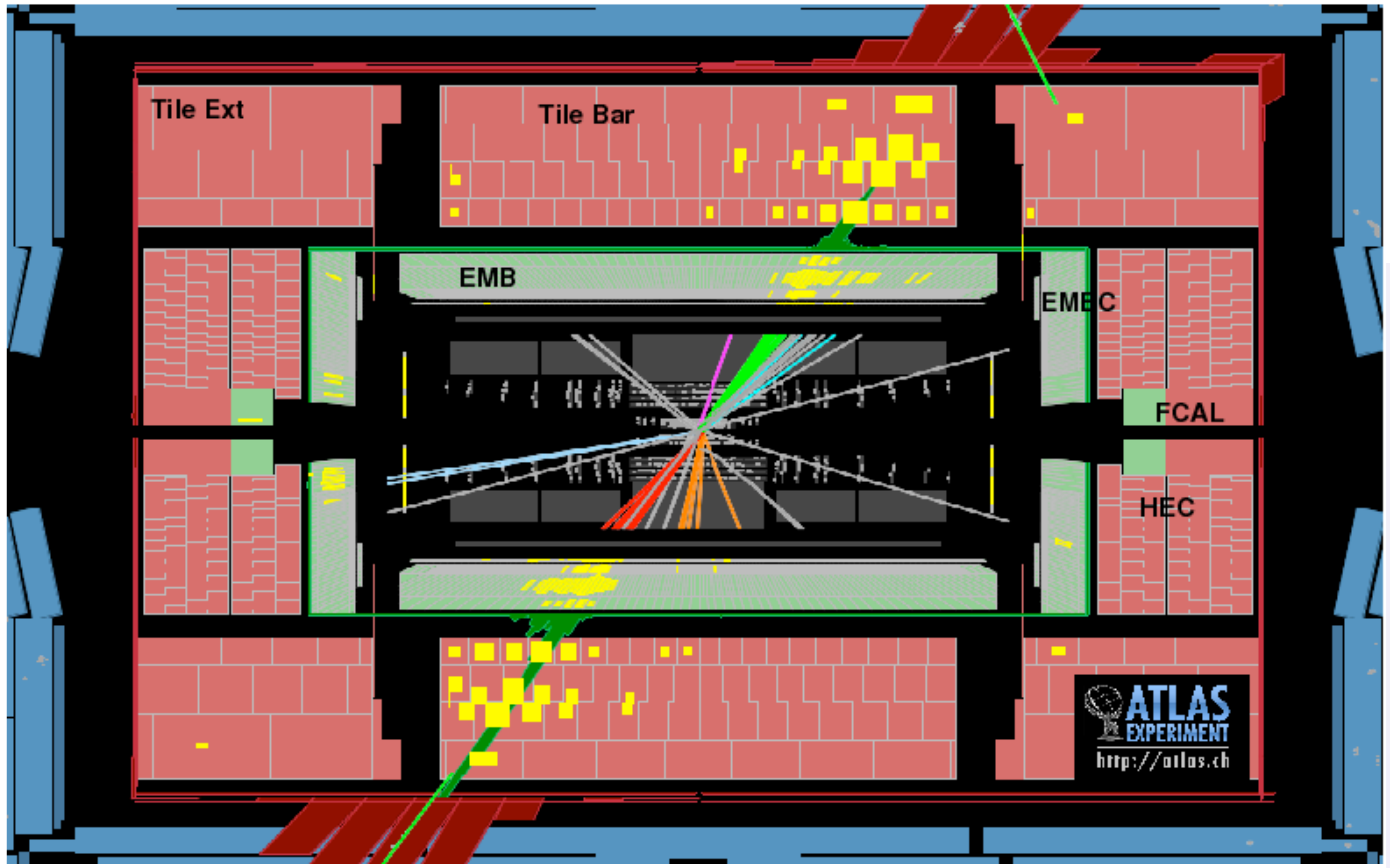
(b) Cells passing selection in Eq. (4)



(c) All clustered cells

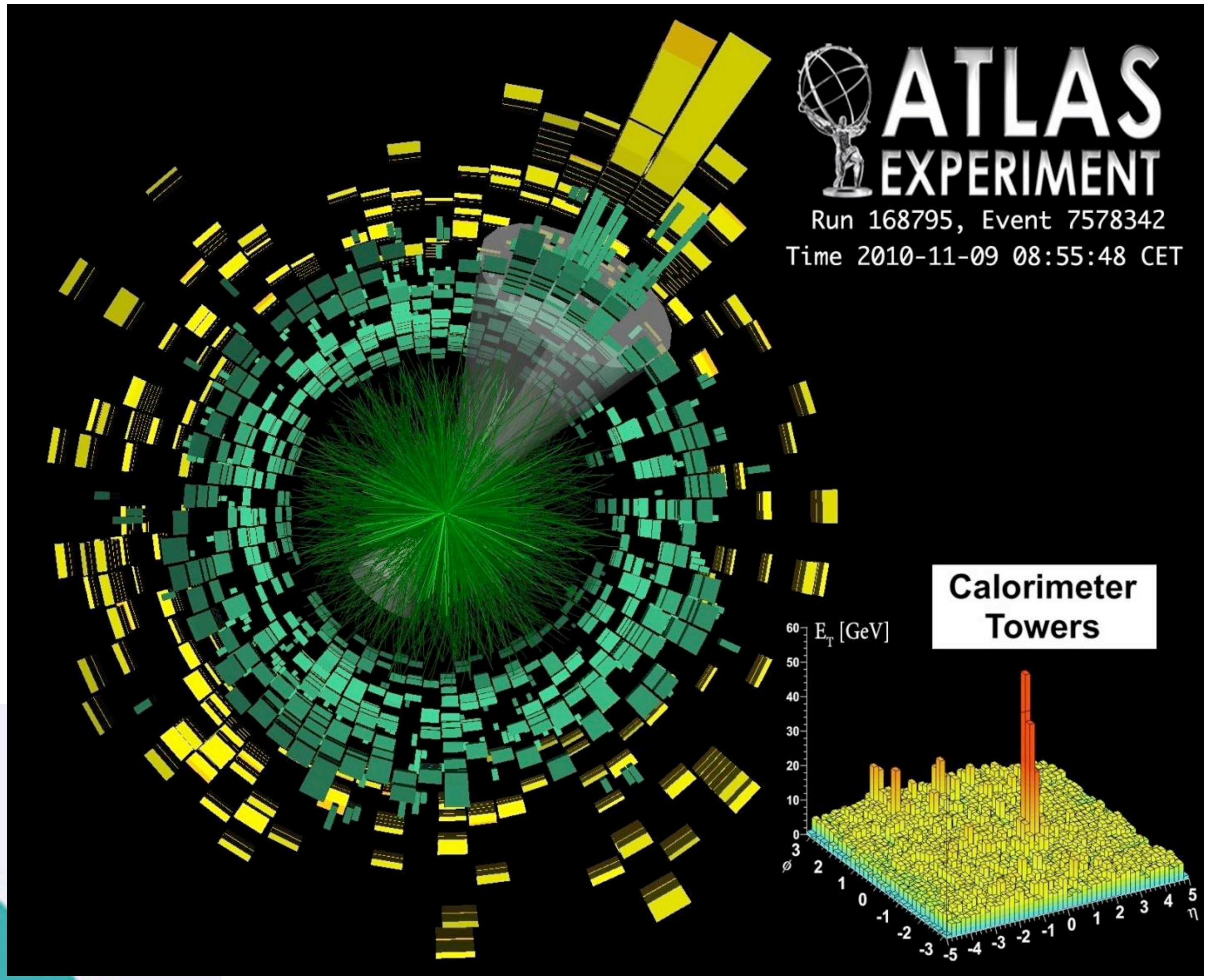
Topoclustering

You might have seen…



https://cds.cern.ch/record/1409965

https://cerncourier.com/a/atlas-observes-striking-imbalance-of-jet-energies-in-heavy-ion-collisions/

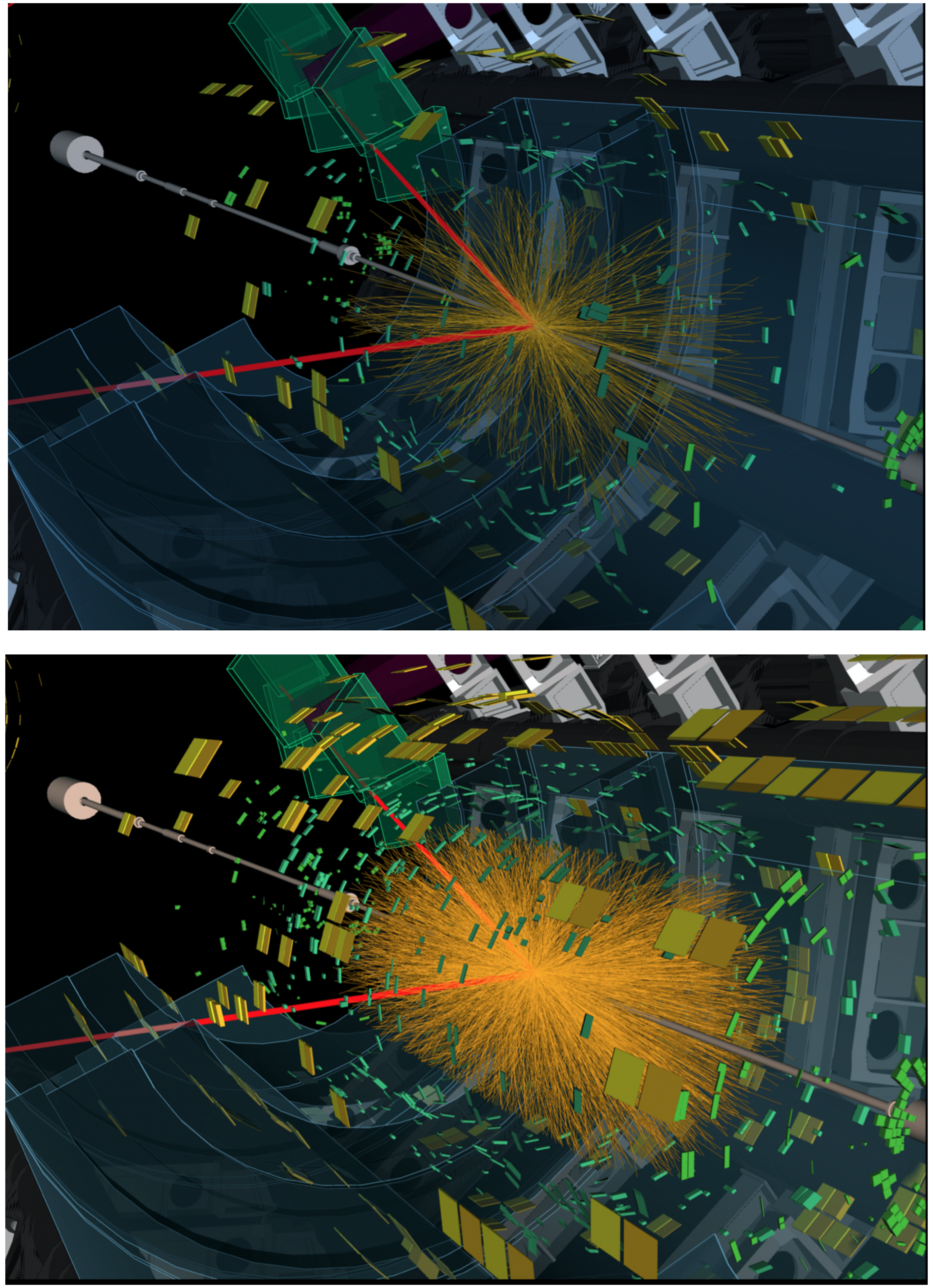


https://www.eurekalert.org/multimedia/662188

What's the motivation?

Why target this…

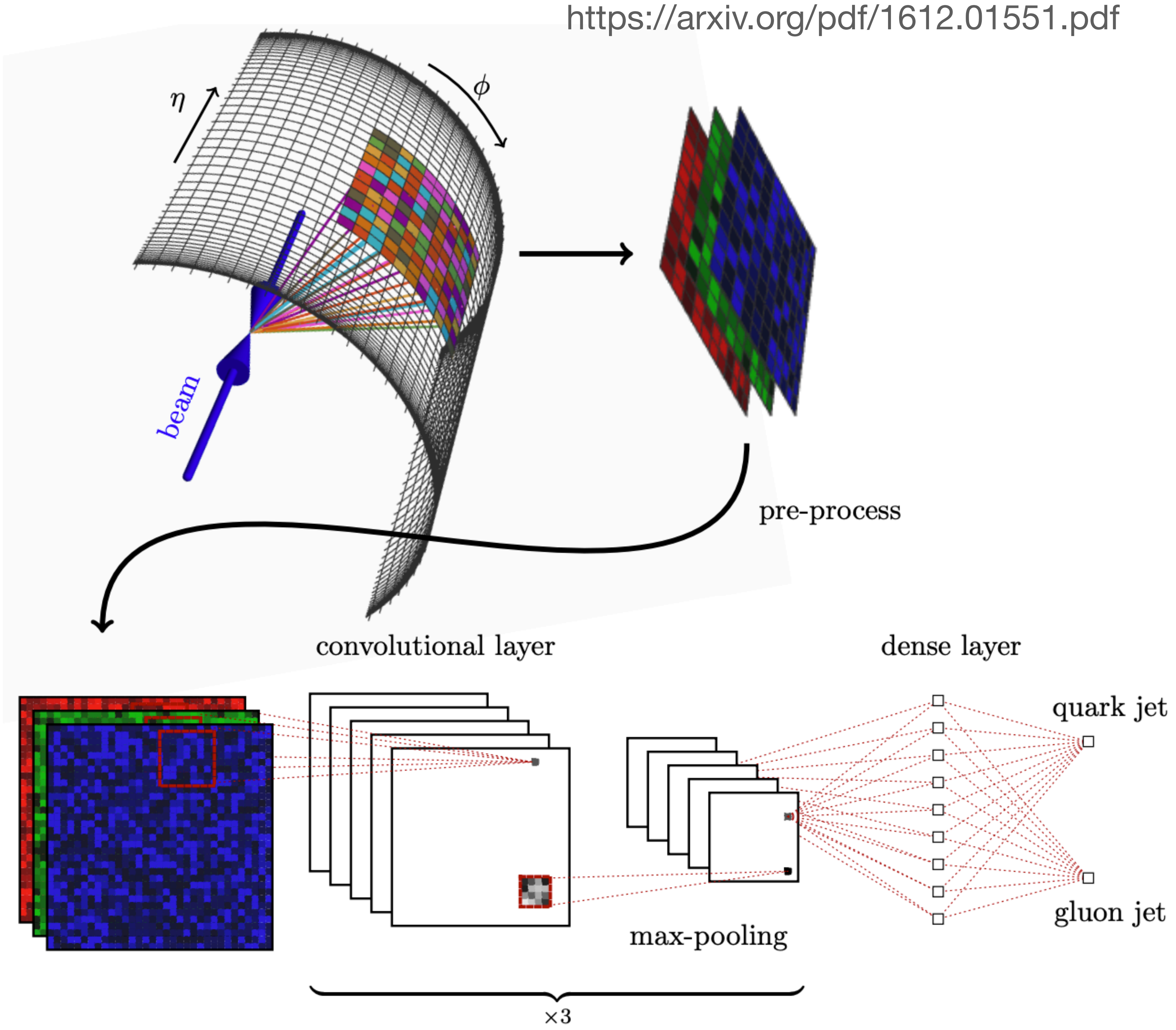
- **Topoclustering** is one of the most resource intensive algorithms in use in HLT.
- Crucial role in **jet** and **MET** reconstruction.
- Far worse **pile-up** conditions in HL-LHC.
- We pursue **faster** solutions with similar or improved performance.
- (And also less **energy consumption**).



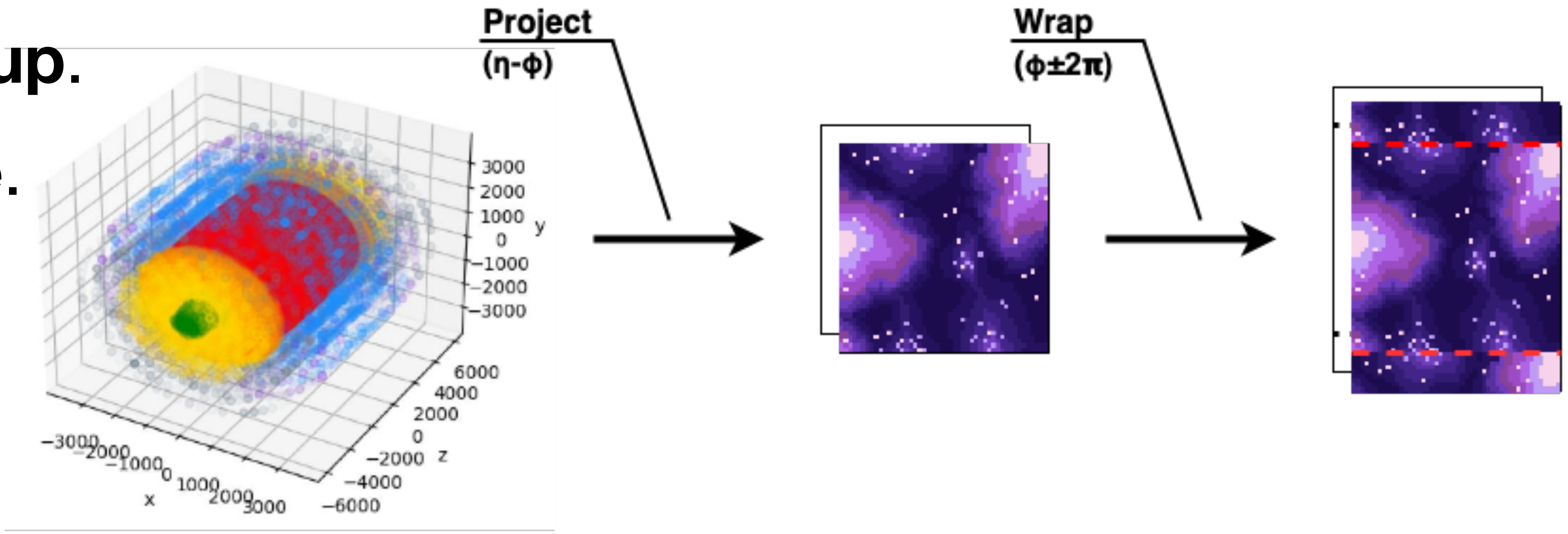
50 vs 200 p-p collisions per bunch crossing.

ML to the rescue?

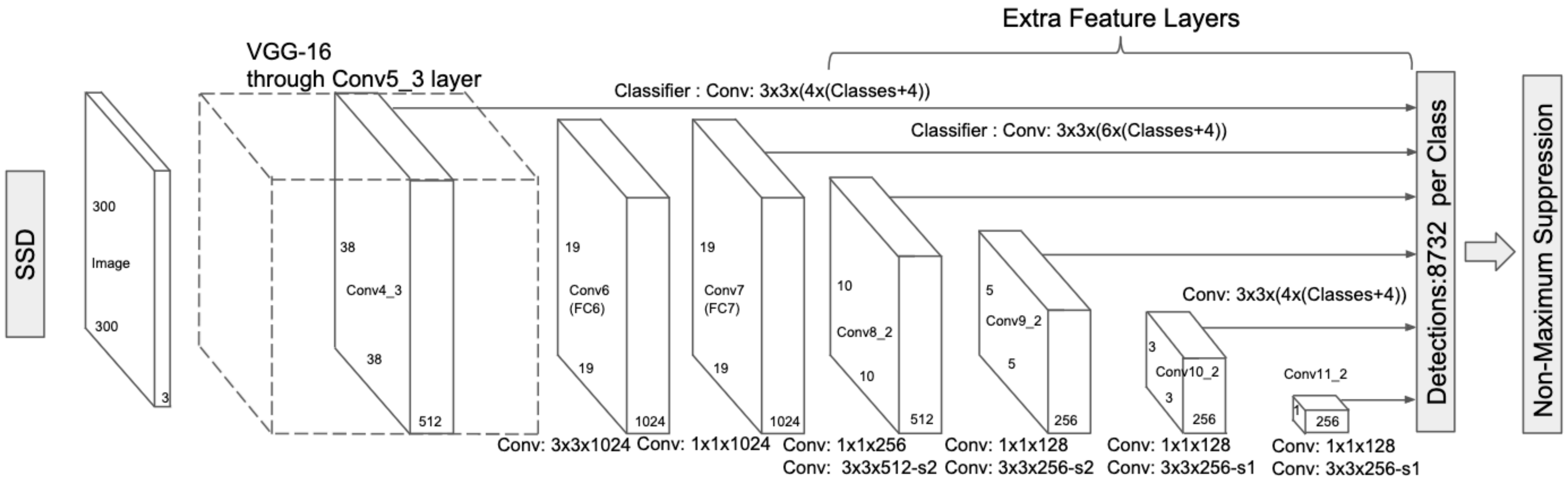
Can we use object detection to help



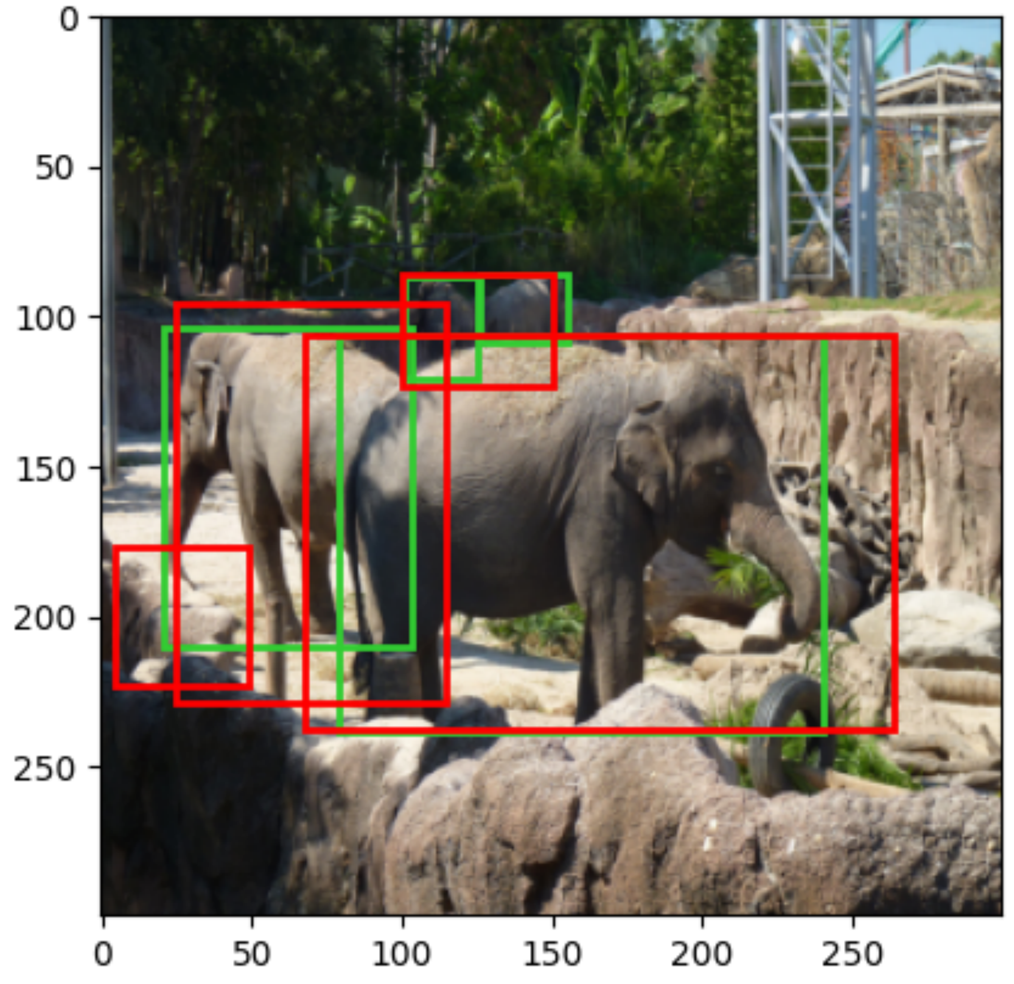
- Qualification task explores use of deep **convolutional** neural networks (CNNs) to locate important **regions** of the **calorimeter** in (η, ϕ)-space.
- Modern CNNs have **rapid inference** speed.
- Execution time **independent** of **pileup**.
- Could be ported to **faster hardware**.

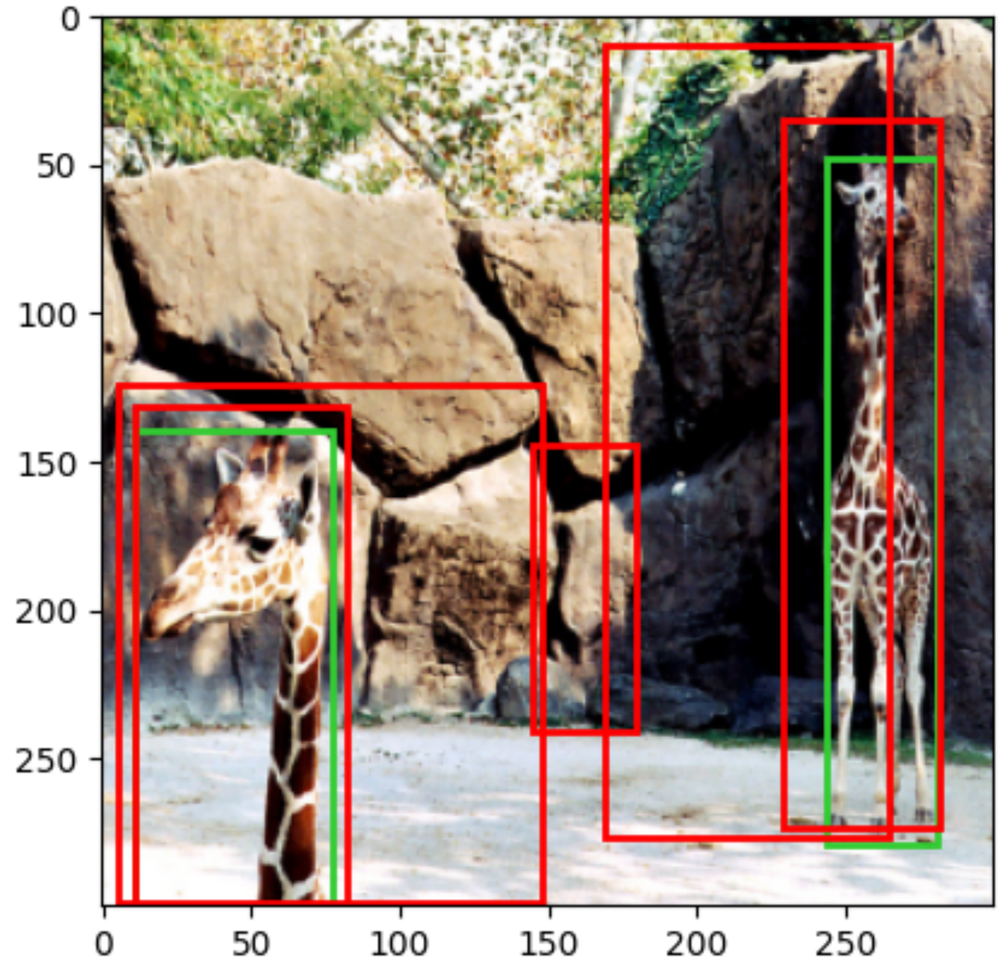


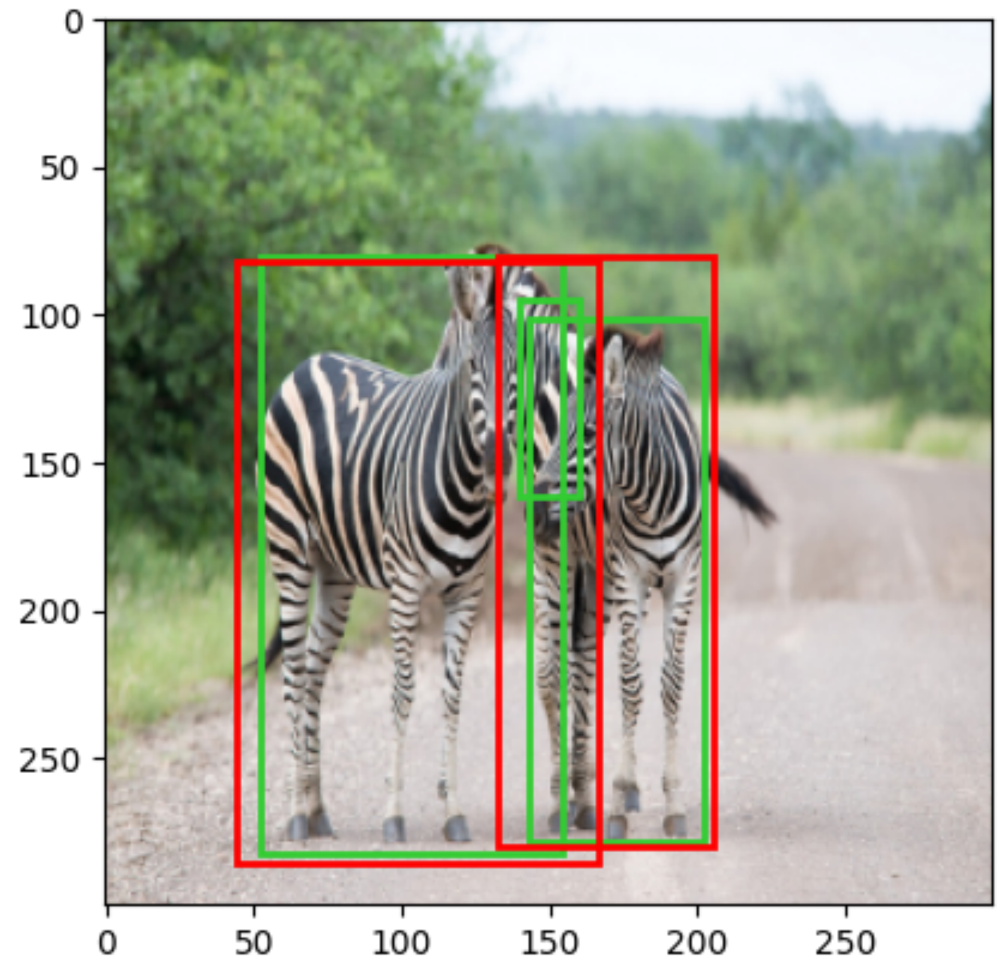
Aside: SSD Model on animals



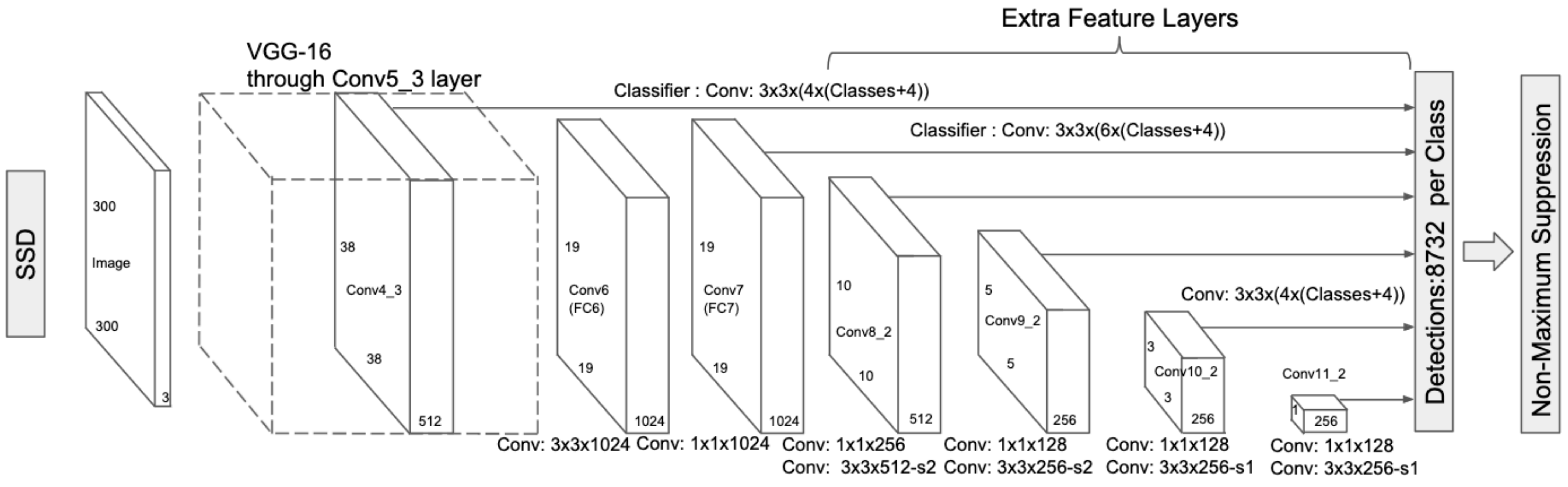
https://arxiv.org/abs/1512.02325



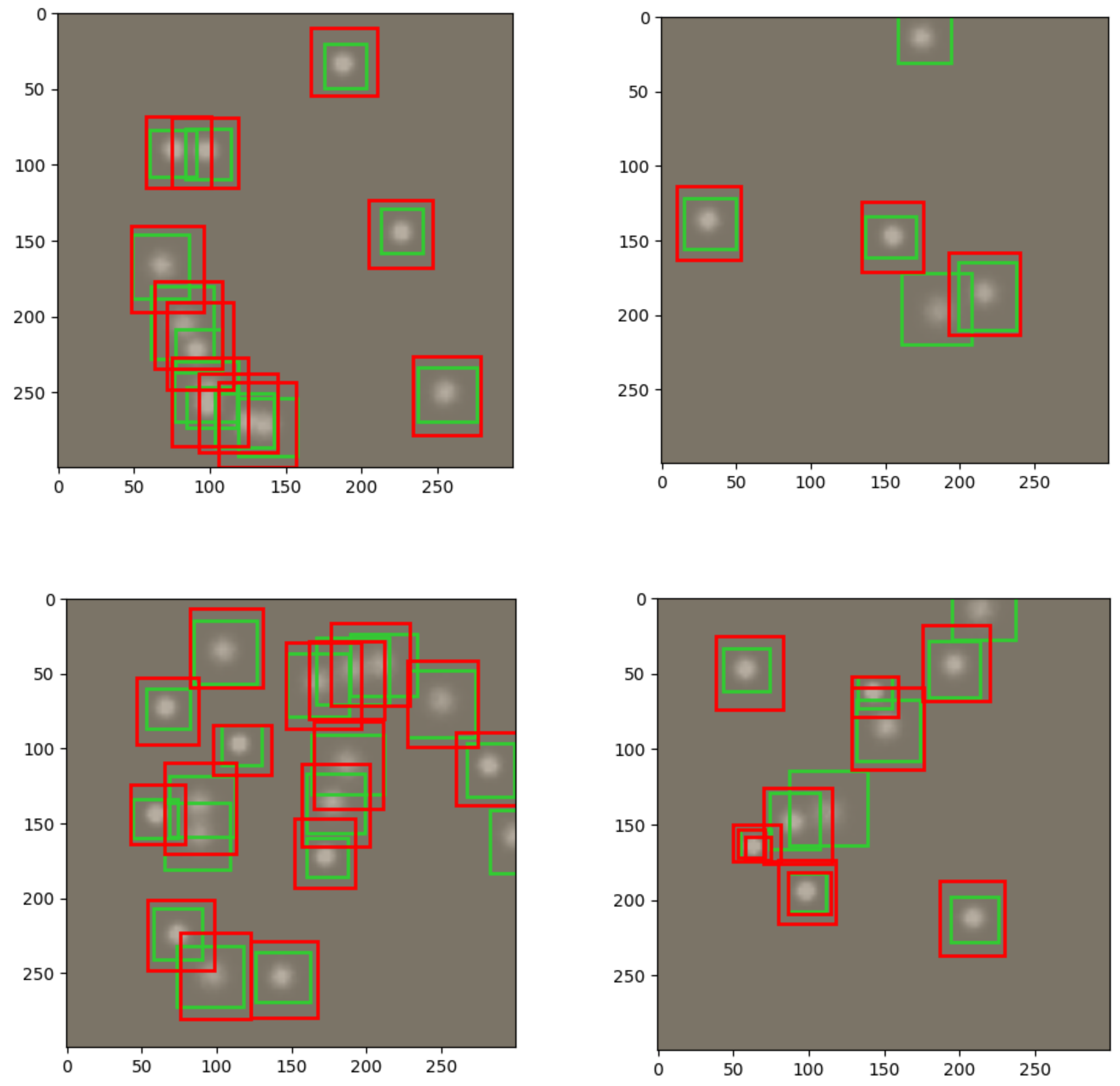




Aside: SSD Model on Gaussian blobs

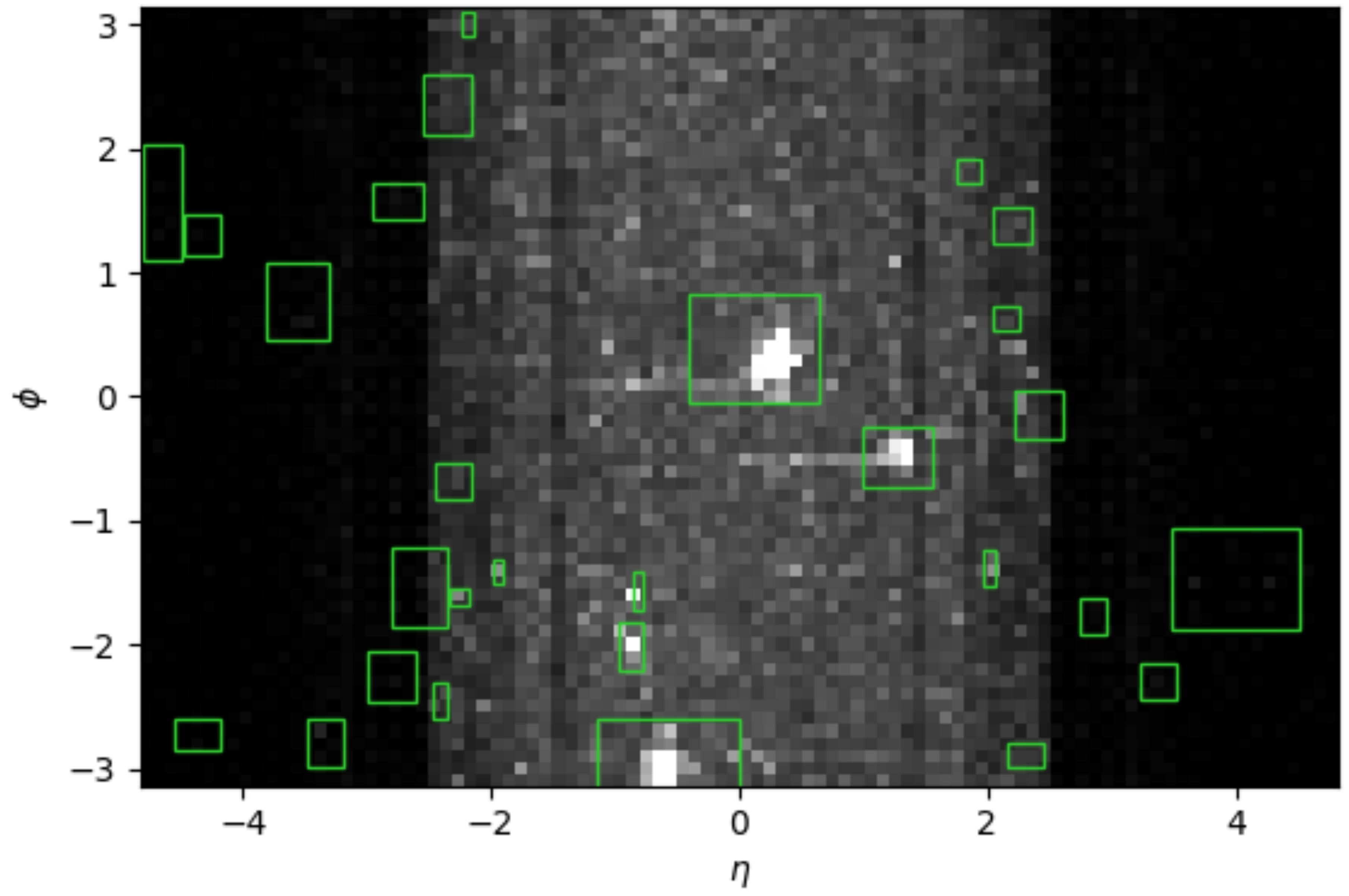


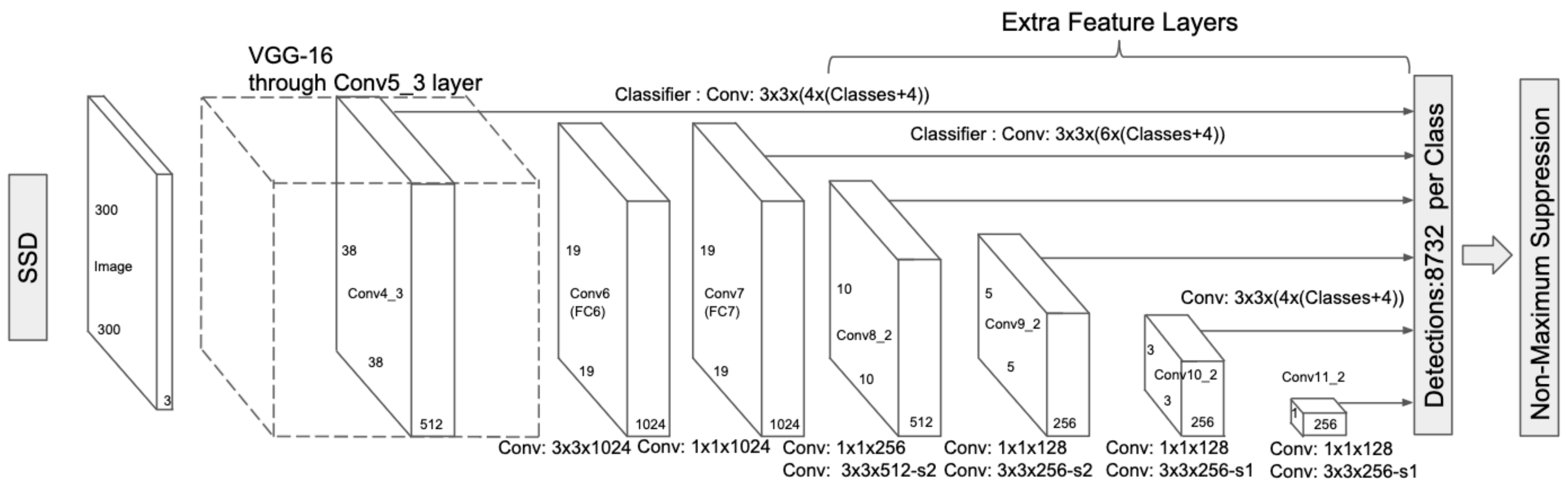
https://arxiv.org/abs/1512.02325

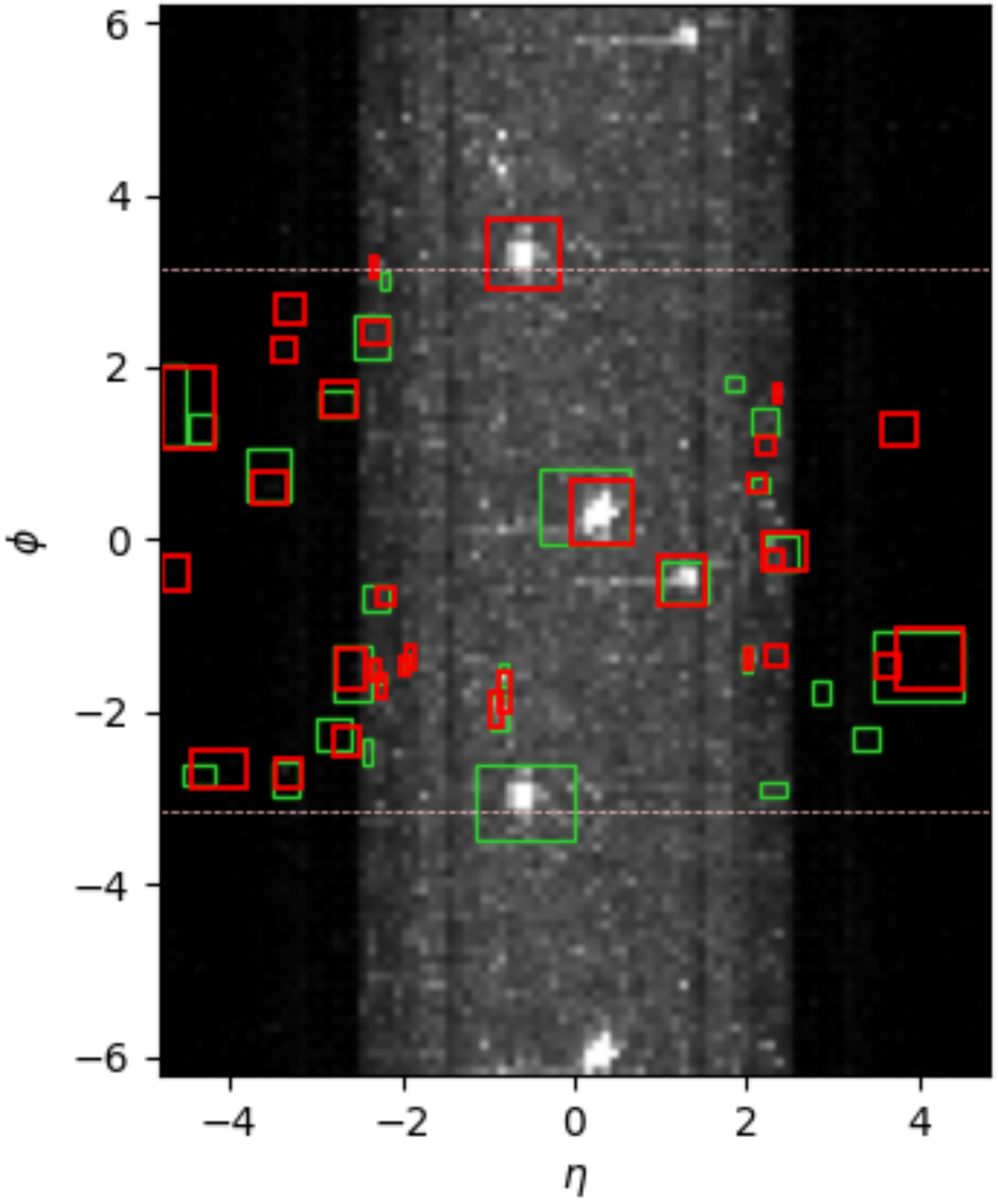


SSD on physics

Does it work in the calorimeter?



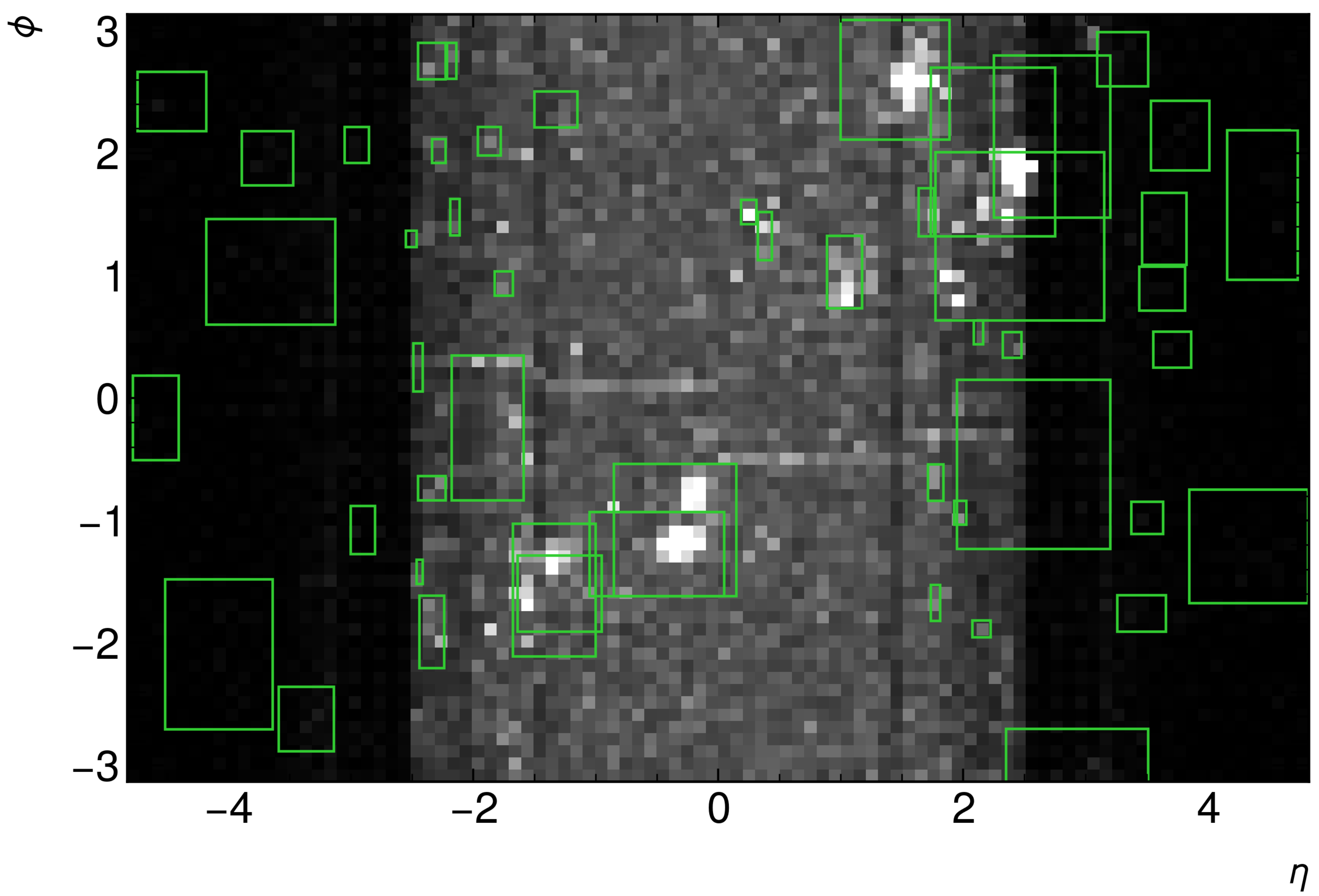


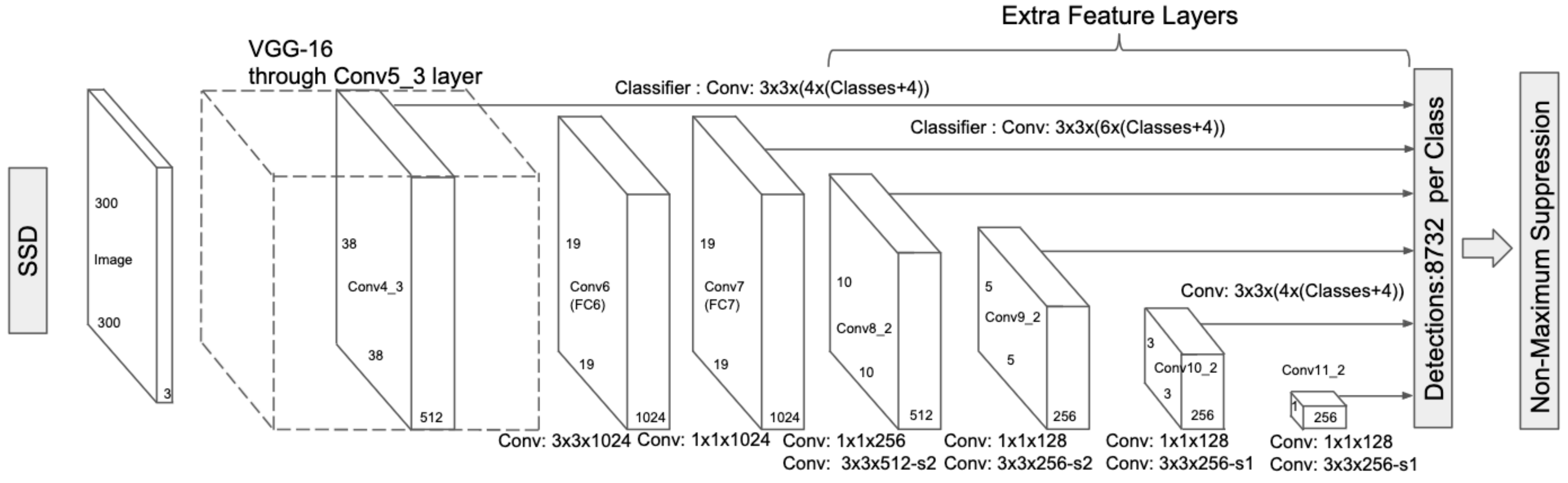


https://arxiv.org/abs/1512.02325

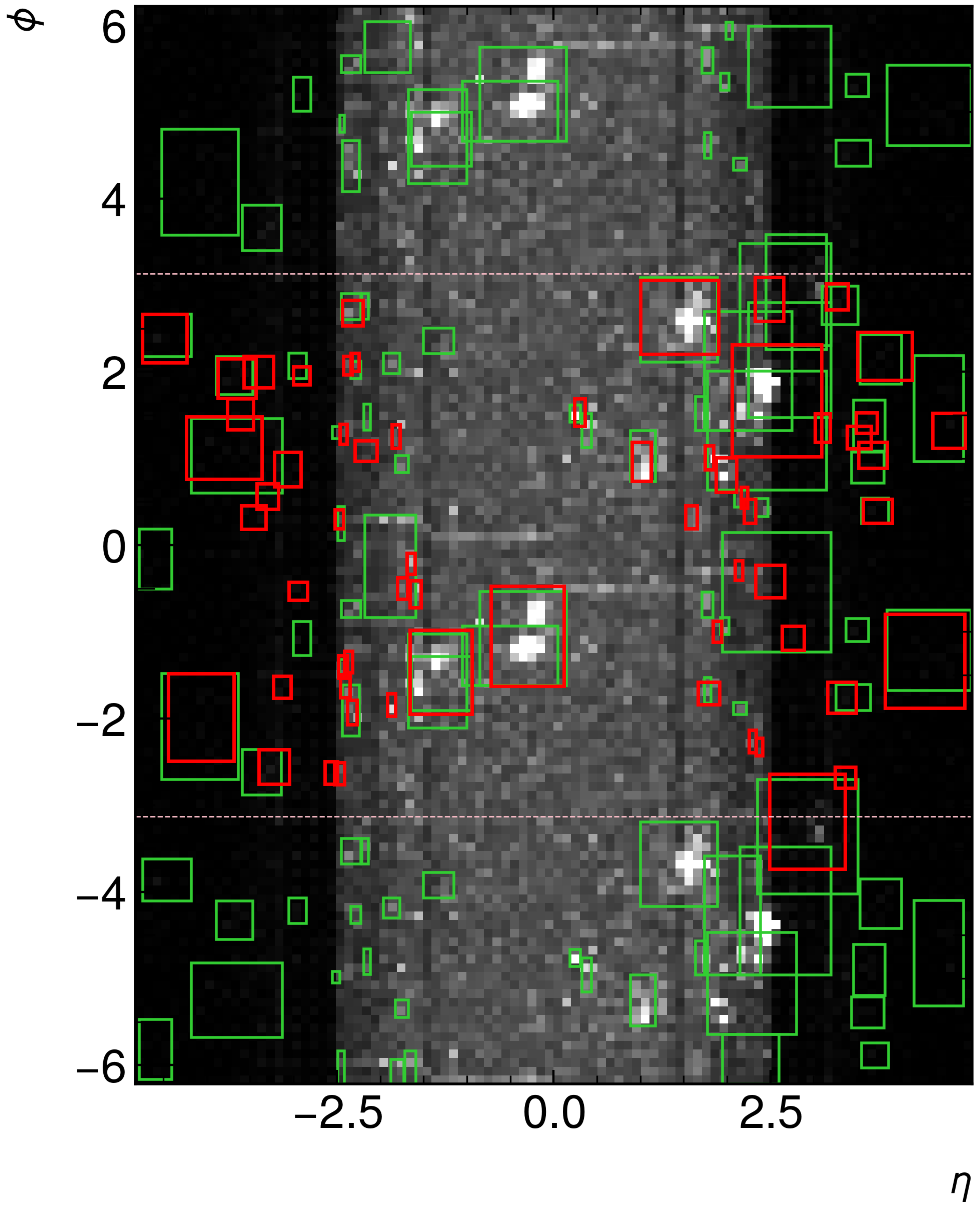
SSD on physics

Does it work in the calorimeter?





https://arxiv.org/abs/1512.02325

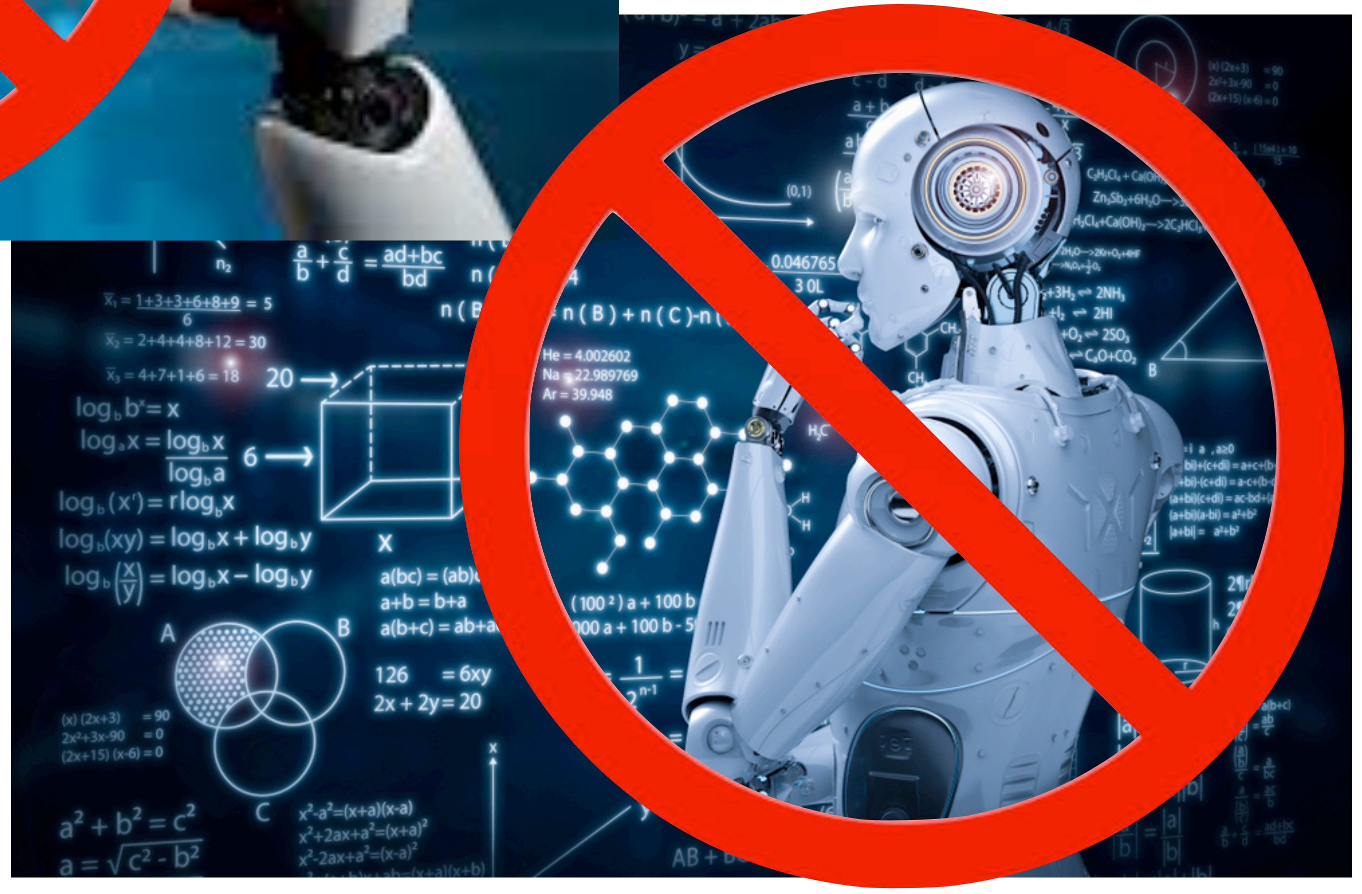


Secondment

Lightbox, Geneva

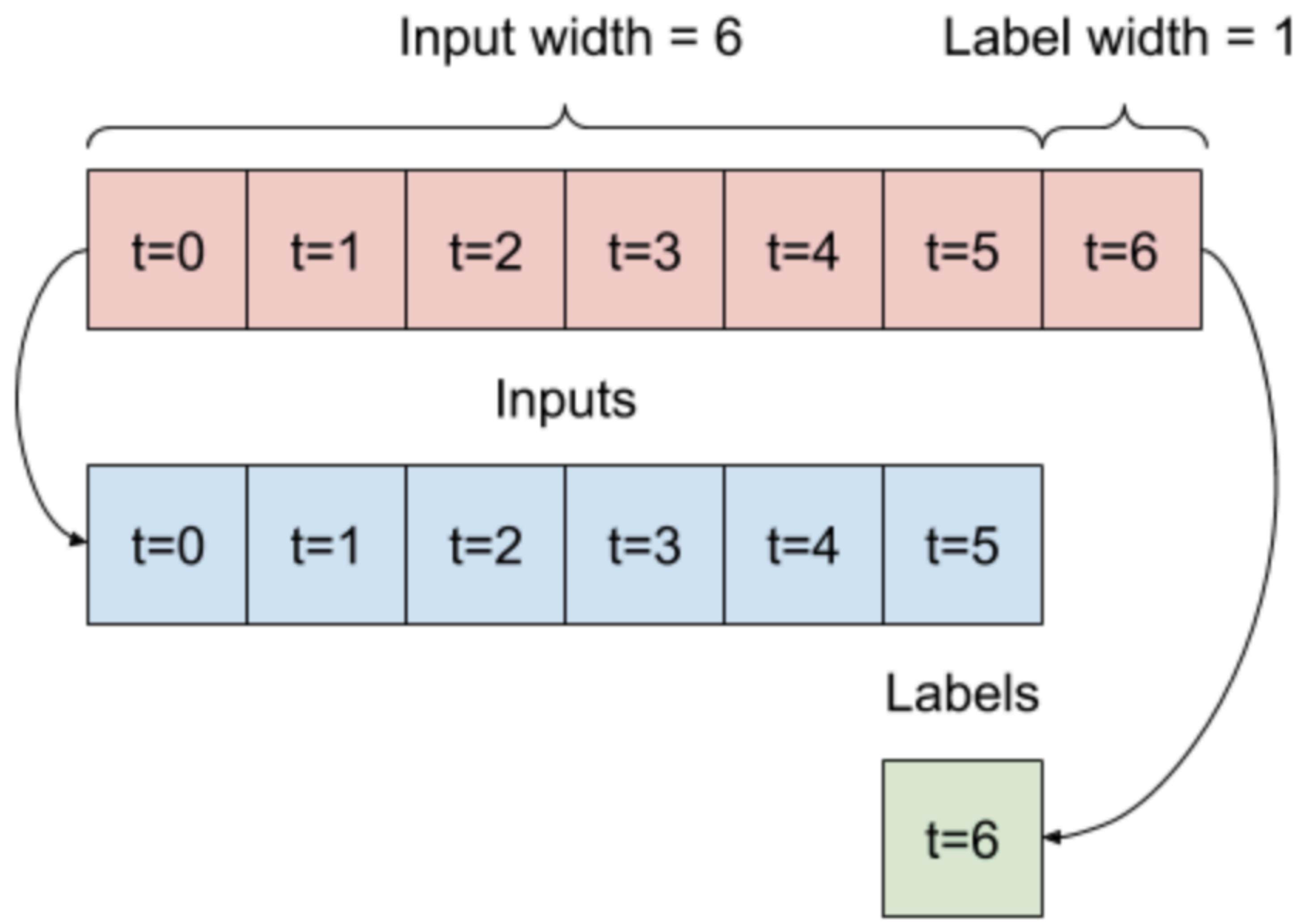
Market Analysis for Financial Firm

- **NOT prediction of the future!**
- Using ML and other data science tools to give signals/inform a financial strategy.
- Already widely adopted.
- A lot of different advice/practices compared to physics and academia.
- Time series data - handle with care!



https://www.tapinto.net/towns/elizabeth/articles/3-applications-of-machine-learning-and-ai-in-finance-2d5c4825-e0ad-4dda-bec6-22e91842eaa7 https://www.tensorflow.org/tutorials/structured_data/time_series

Time Series Data



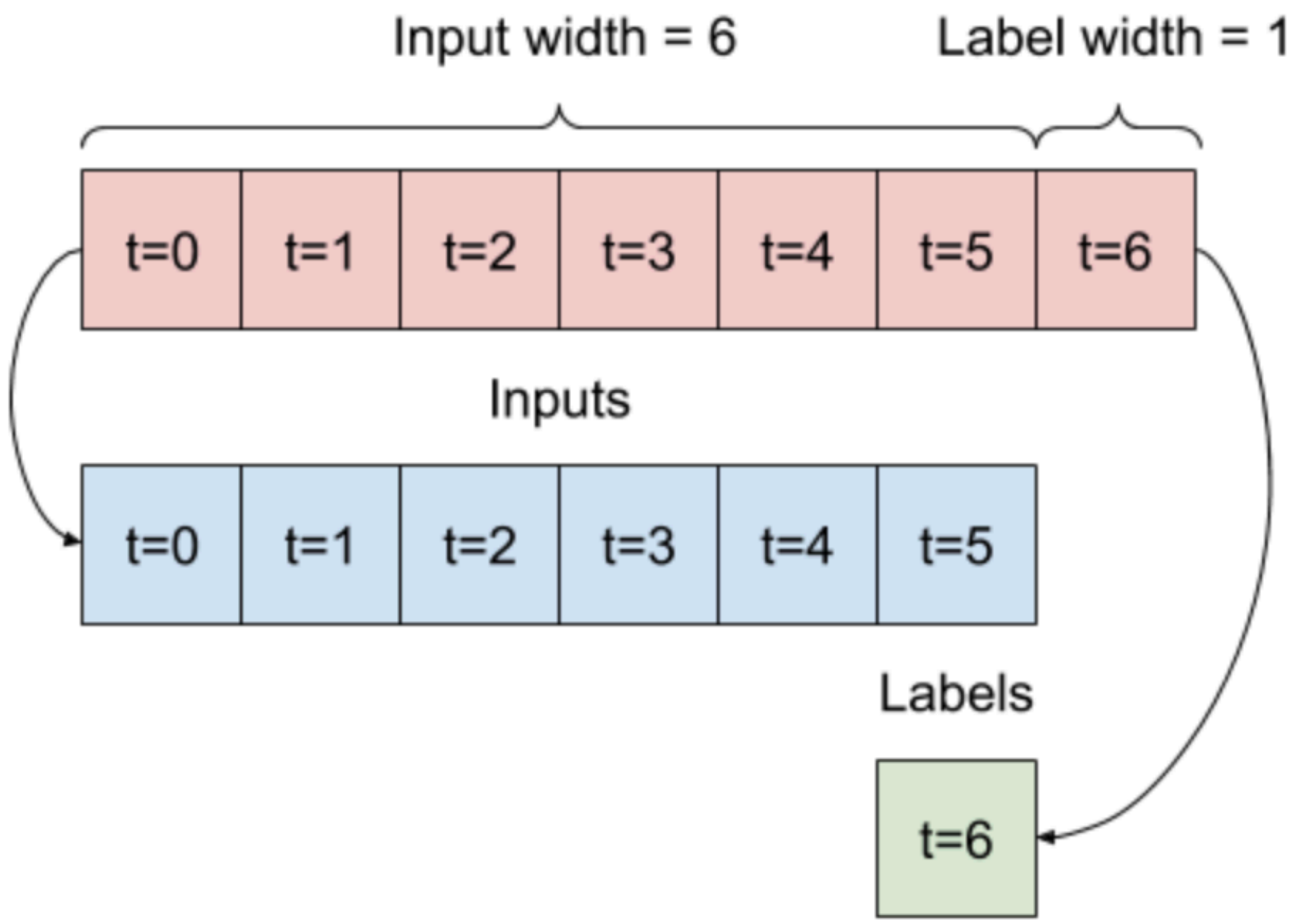
- Focus on **time series data**, different underlying (stocks, commodities, FX).
- Split into **training** (past) and **validation/test** (future) data.
- In a **recurrent neural network** (RNN) we use a 14 step **window** as input.
- Forecast **horizon** can change, most common choice is one step.

https://groww.in/blog/how-to-read-candlestick-charts

Time Series Data

- Focus on **time series data**, different underlying (stocks, commodities, FX).
- Split into **training** (past) and **validation/test** (future) data.
- In a **recurrent neural network** (RNN) we use a 14 step **window** as input.
- Forecast **horizon** can change, most common choice is one step.

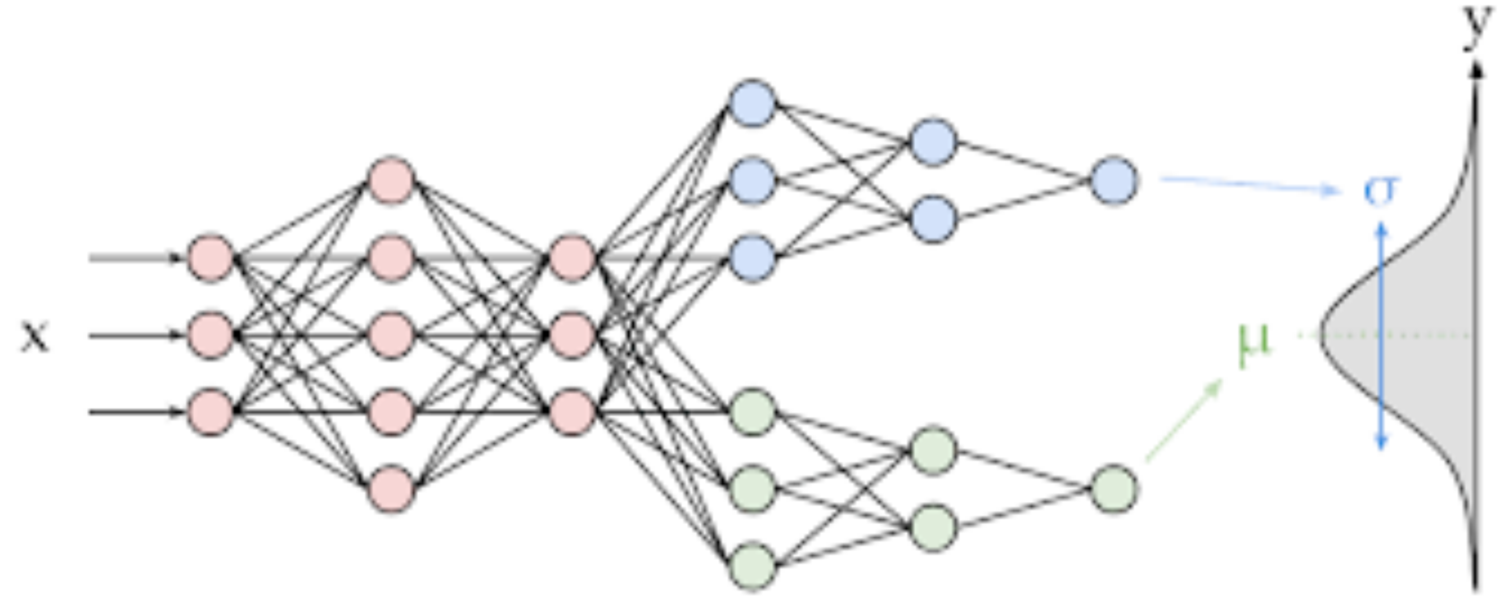
https://www.tensorflow.org/tutorials/structured_data/time_series



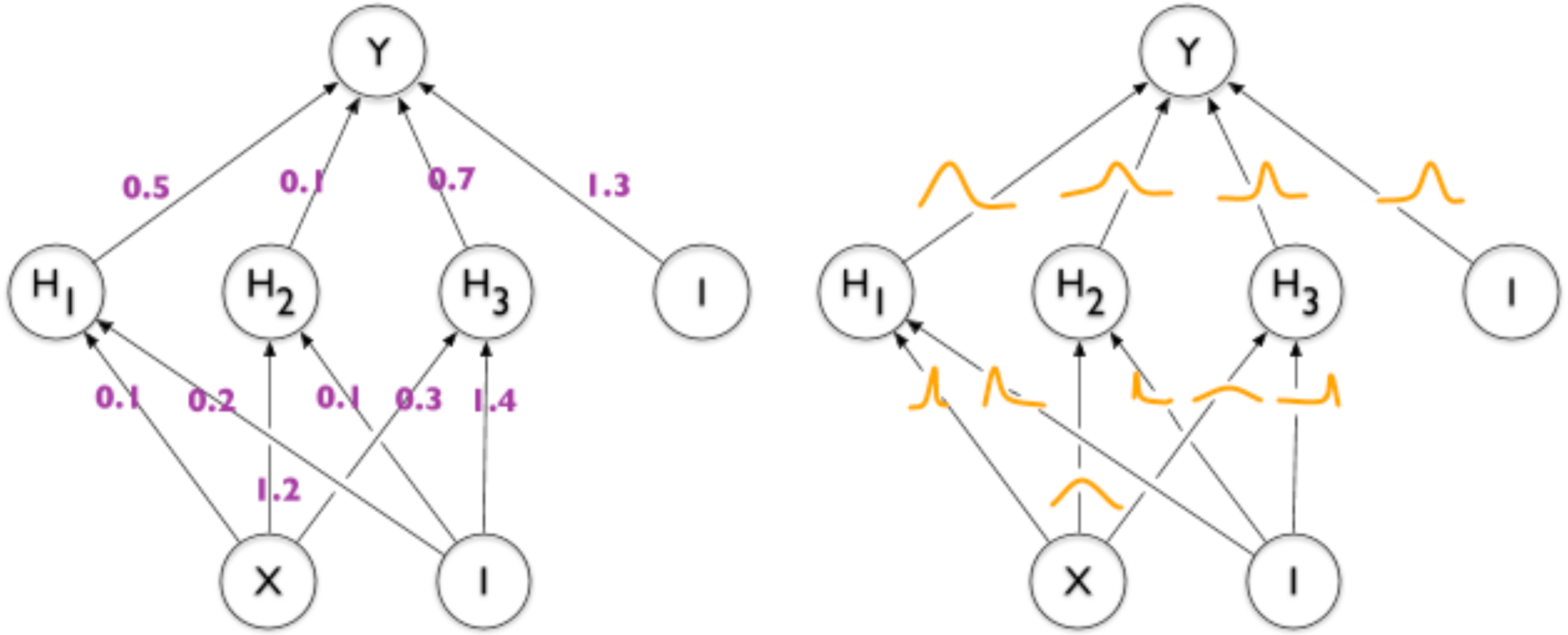
https://www.axi.com/int/blog/education/trading-indicators

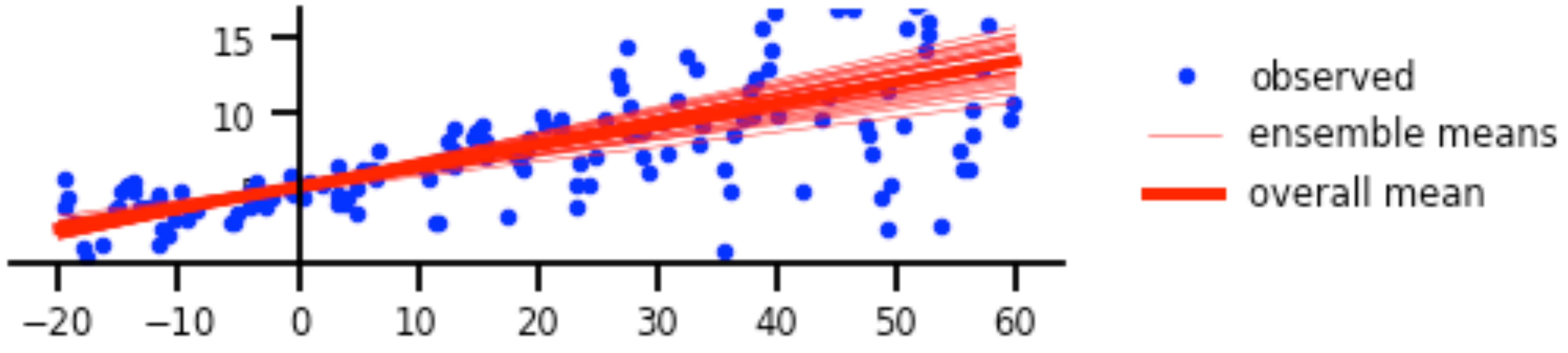
Bayesian Neural Networks

A quick detour



- **Bayesian** neural networks (BNNs) model the **uncertainty** associated to a prediction.
- Rather than a weights leading to a single **deterministic** output their weights define PDFs.
- The model output can also be a '**distribution**'.
- Very robust to **noisy** data, hard to **overtrain**.





https://www.tensorflow.org/probability/examples/Probabilistic_Layers_Regression
https://brendanhasz.github.io/2019/07/23/bayesian-density-net.html
https://neptune.ai/blog/bayesian-neural-networks-with-jax

Our work

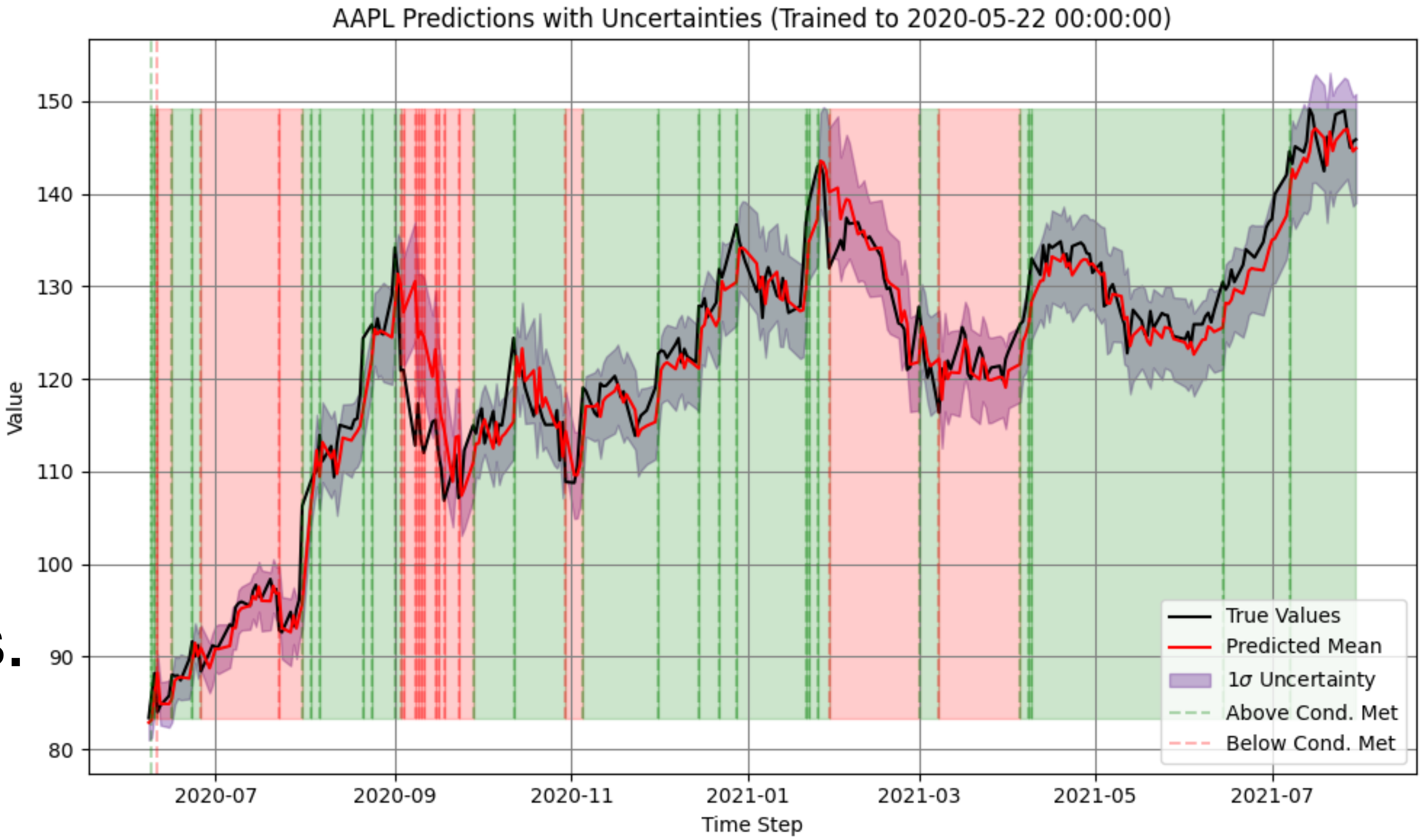
Combine LSTM + Bayesian Output

- We use the uncertainty estimation of BNNs to guide a small LSTM backed model (~20k parameters).
- At each time step **model outputs** a Gaussian (or student's-T) **distribution.**
- Use σ as a **quality/confidence** score. Use this to filter buy and sell decisions.
- Predictions highly **dependent** on **training size** (training sample regime) and **forecast window** .
- (And a small excursion into transfer learning…)

Our work

Combine LSTM + Bayesian Output

- We use the uncertainty estimation of BNNs to guide a small LSTM backed model (~20k parameters).
- At each time step **model outputs** a Gaussian (or student's-T) **distribution.**
- Use σ as a **quality/confidence** score. Use this to filter buy and sell decisions.
- Predictions highly **dependent** on **training size** (training sample regime) and **forecast window** .
- (And a small excursion into transfer learning…)



Conclusion

- Object detection in ATLAS trigger, promising though requires **more validation** and rigorous **training** + **optimising**.
- Secondment **project prototype** finished. Now work on improvements in **robustness** before any **potential deployment**.
- More tests with **different** time series **data** (frequency, duration, horizon, etc.)
- To join ATLAS physics analysis with ML/TLA focus.

Conclusion

- Object detection in ATLAS trigger, promising though requires **more validation** and rigorous **training** + **optimising**.
- Secondment **project prototype** finished. Now work on improvements in **robustness** before any **potential deployment**.
- More tests with **different** time series **data** (frequency, duration, horizon, etc.)
- To join ATLAS physics analysis with ML/TLA focus.

Thanks for listening

Backup

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 64)	17408
dense (Dense)	(None, 64)	4160
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 16)	528
dense_3 (Dense)	(None, 8)	136
dense_variational (DenseVaiational)	(None, 2)	54
distribution_lambda (DistributionLambda)	((None, 1), (None, 1)	0

Total params: 24366 (95.18 KB)
Trainable params: 24366 (95.18 KB)
Non-trainable params: 0 (0.00 Byte)

Parabolic SAR

The PSAR indicator is used to determine trend direction and potential reversals in price. Establish trend first the, If the trend is up, buy when the indicator moves below the price. If the trend is down, sell when the indicator moves above the price. A buy signal occurs when the PSAR moves from above to below the price, while a sell signal occurs when the dots move from below to above the price.

$\text{RPSAR} = \text{Prior PSAR} +$
$[\text{Prior AF}\,(\text{Prior EP-Prior PSAR})]$
$\text{FPSAR} = \text{Prior PSAR} -$
$[\text{Prior AF}\,(\text{Prior PSAR-Prior EP})]$

where:

$\text{RPSAR} = \text{Rising PSAR}$

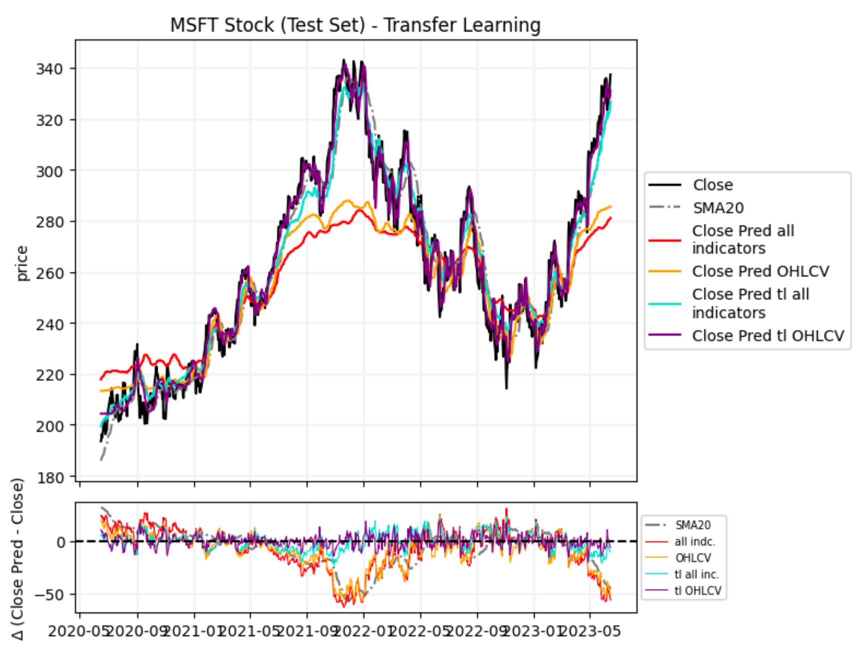
$\text{AF} =$ Acceleration Factor, it starts at 0.02 and increases by 0.02, up to a maximum of 0.2, each time the extreme point makes a new low (falling SAR) or high (rising SAR)

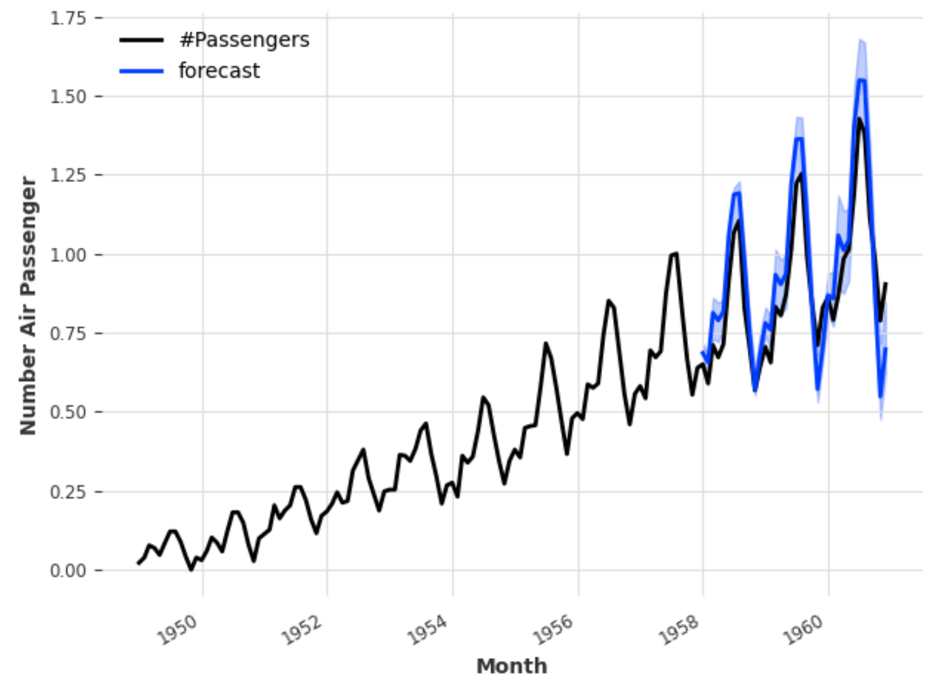
$\text{FPSAR} = \text{Falling PSAR}$

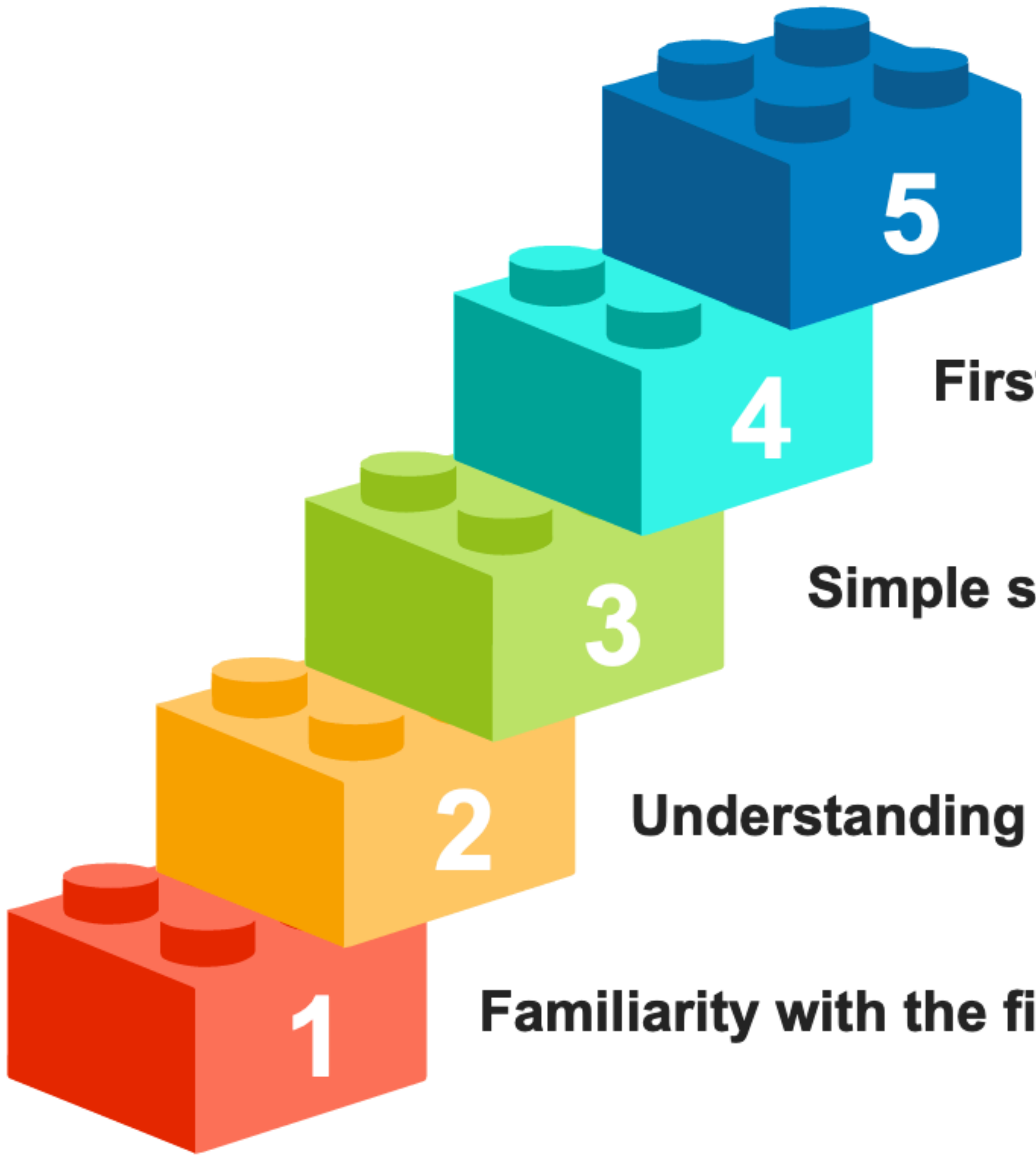
$\text{EP} =$ Extreme Point, the lowest low in the current downtrend (falling SAR) or the highest high in the current uptrend (rising SAR)

Our work

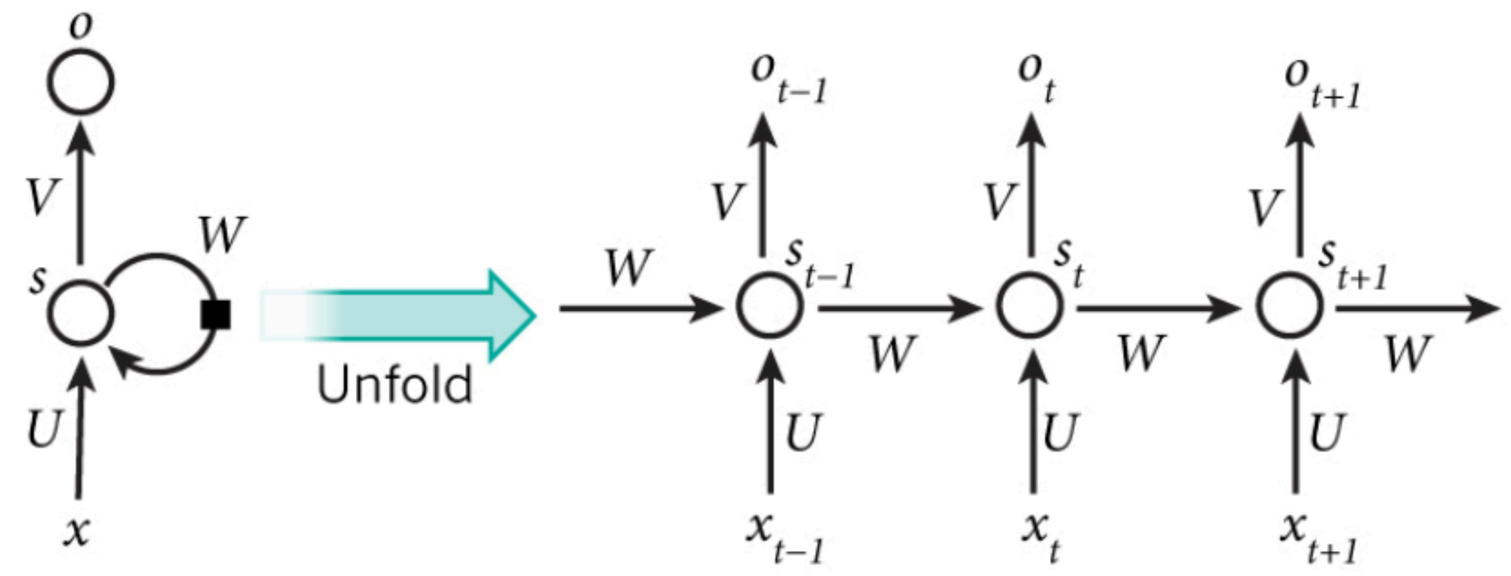
A bit of trial and error

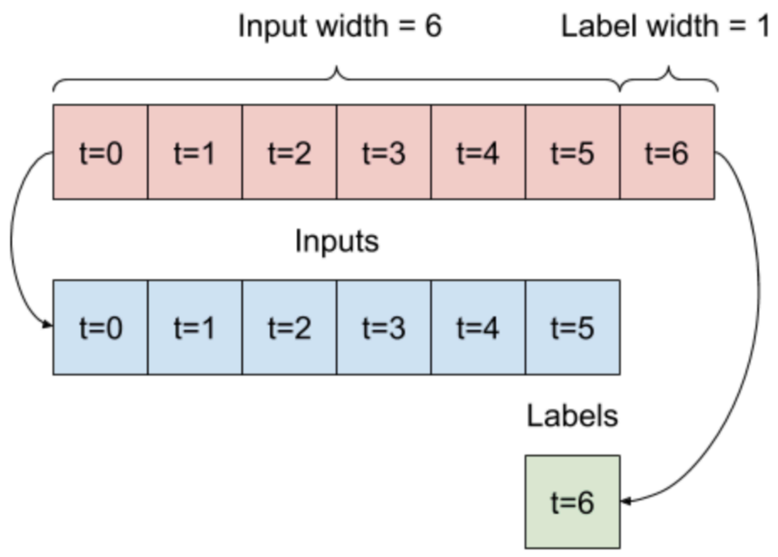


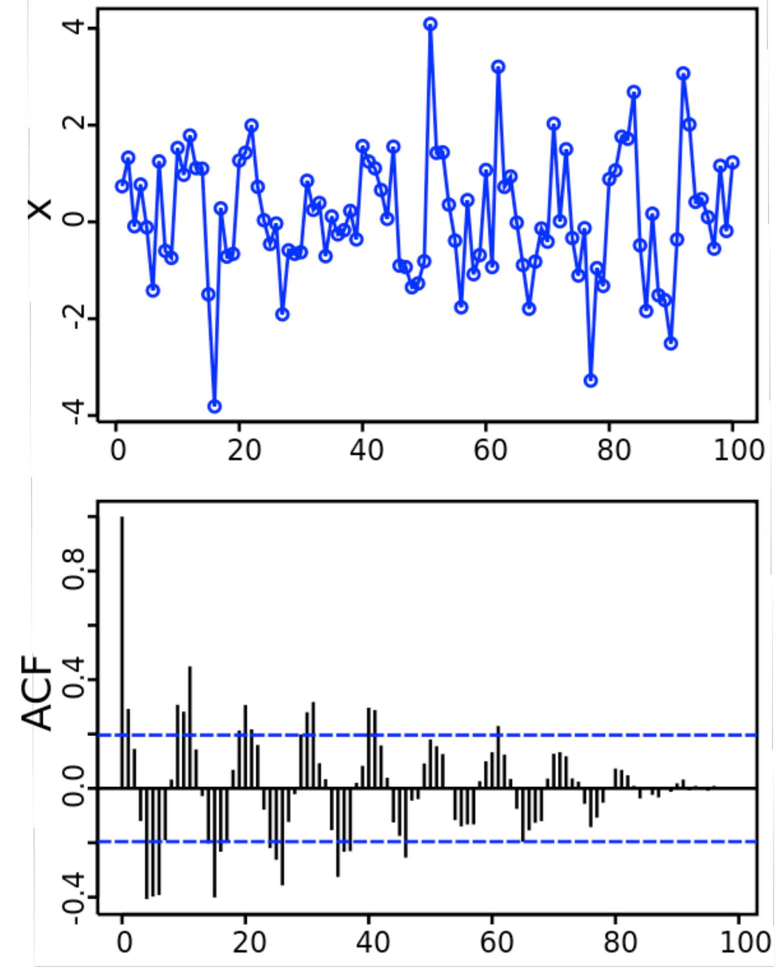




5. Progress to RNNs (LSTM, GRU,α-RNN)
4. First, straightforward BDTs and FC-NNs
3. Simple statistical tests (PCA, Autocorr.)
2. Understanding the technical indicators
1. Familiarity with the financial data







https://otexts.com/fpp3/acf.html
https://arxiv.org/pdf/1412.3555.pdf
https://www.barchart.com/education/technical-indicators

Miscellaneous activities

Additional engagements

In Geneva:

- University course in statistical methods
- University course in scientific computing in physics
- 3rd Symposium on AI for Industry, Science and Society
- High performance computing cluster training

At CERN:

- ATLAS Control Room Shifter Training for trigger + DQ
- PyRoot + Scikit-HEP masterclasses/tutorials
- ECSB Lecture Series
- Radioactivity safety course

Elsewhere:

- CHIPP PhD Winter School, Leukerbad Jan. 2023
- 3rd COMCHA School, Oviedo Oct. 2023