Data-driven detection of multi-messenger transients

- Multi-messenger science
- Deep learning transient detection
- Example analyses:
  - γ-ray transients with CTA
  - Neutrino point source search with IceCube & ANTARES
  - Neutrino emission correlation for core-collapse SNe with IceCube & LSST

I. Sadeh (DESY)

March 2021

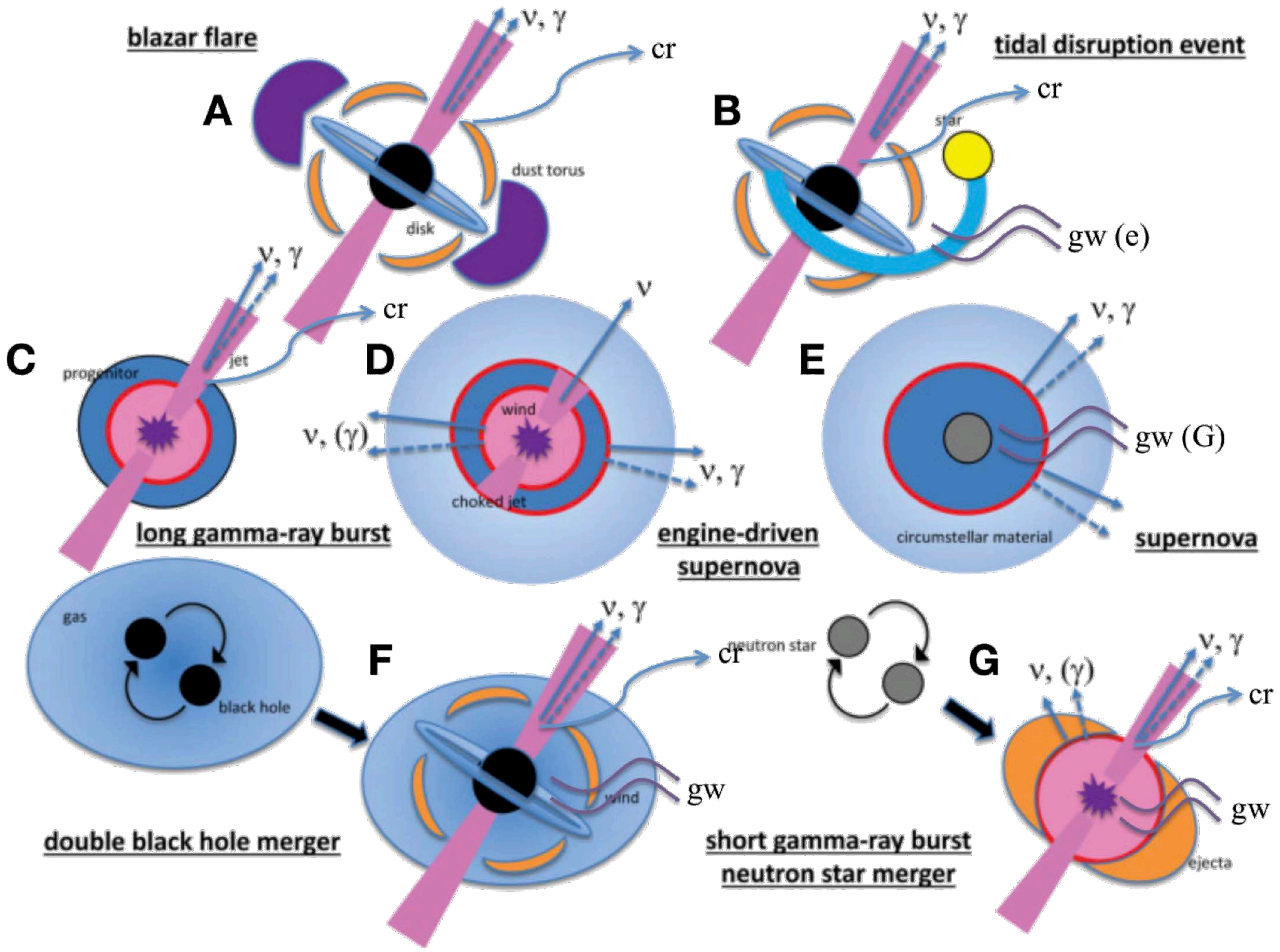
iftach.sadeh@desy.de

Sadeh (2020) arxiv:2005.06406

Multi-messenger & multi-wavelength science

- Some open questions
  - What are the sources of ultrahigh-energy cosmic rays (UHECRs)? • Does the spectrum suggest an acceleration cutoff energy?
  - What is the origin of the (TeV—PeV) cosmogenic neutrino background; of the non-blazar diffuse γ-ray background?
  - What are the origins of the heavy elements?
  - What are the explosion mechanisms of engine-driven SNe; what is the connection to GRBs?

- Wish list
  - More EM associations with GW sources
  - Detection of HE neutrinos (HENs) from GW/EM detected compact object mergers
  - Solid association of HENs with […]
  - Solid association of UHECR arrival directions with […]
  - Better UHECR composition measurements
  - …

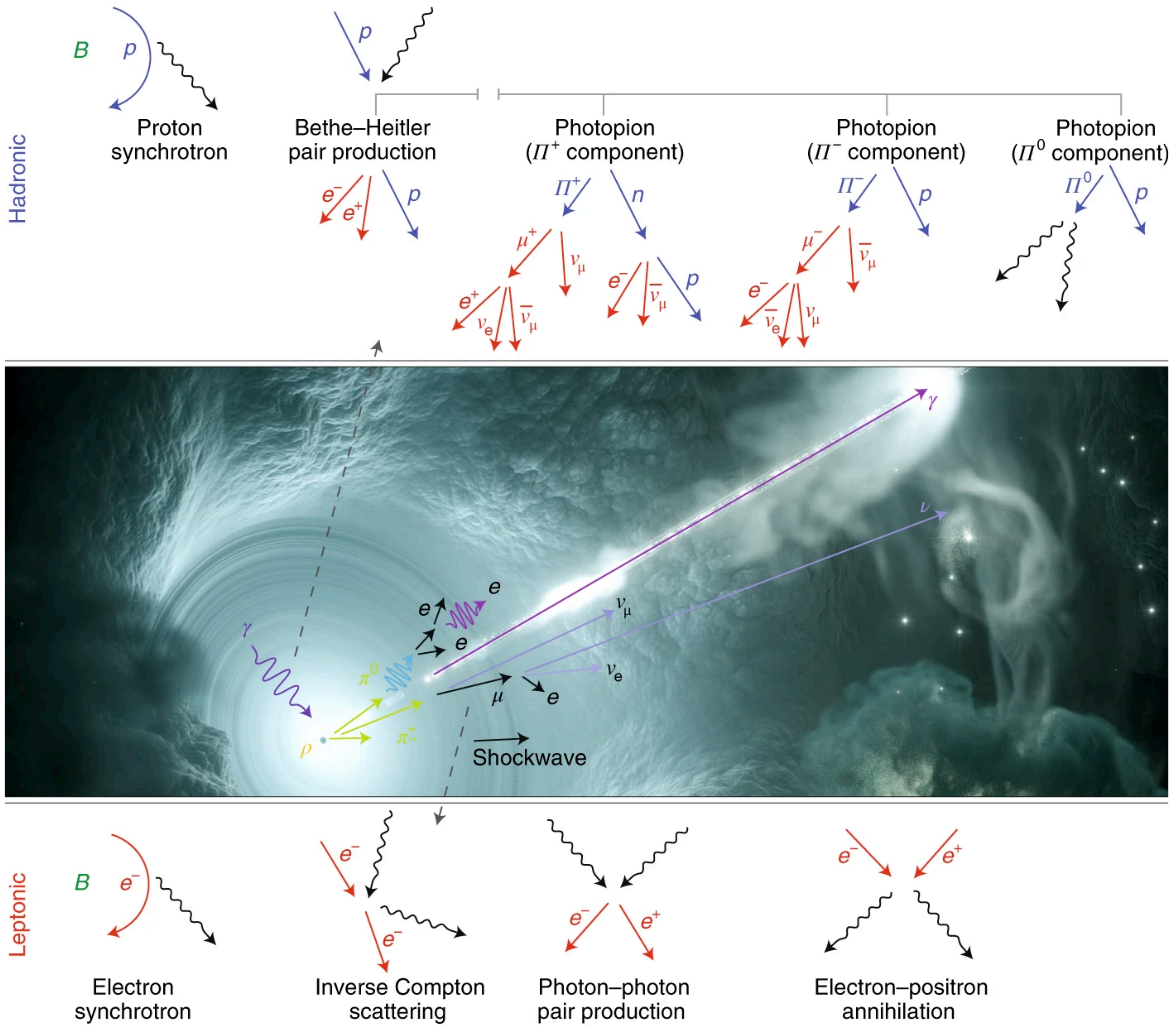


Meszaros et al (2019) arxiv:1906.10212

MMS transients

- Leptonic and(?) hadronic acceleration mechanisms
  - MMS ➜ explore astrophysical phenomena using distinct but complimentary information
  - CRs are deflected by magnetic fields, but their associated neutrinos point directly to their sources

- Evidence for high-energy astrophysical neutrino sources (hadronic signatures)
  - Association (Fermi ➜ MAGIC) of the flaring γ-ray AGN, TXS 0506+056, with a ~0.3 PeV neutrino at ~3σ
  - Association of the radio-emitting tidal disruption event, AT2019dsg, with a ~0.2 PeV neutrino at ~3σ

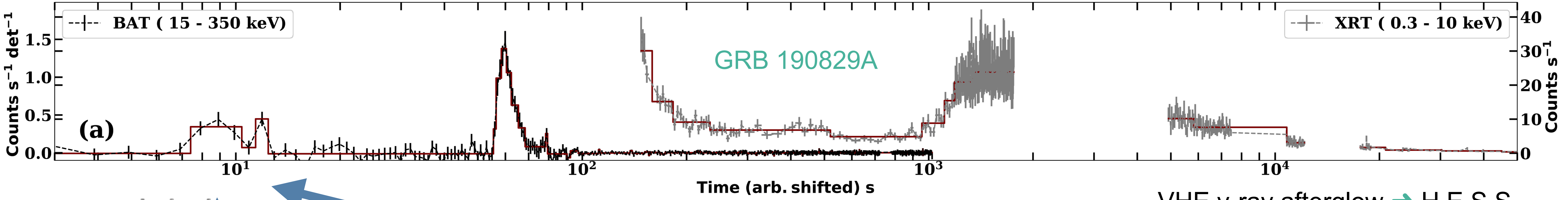


Elena Pian, [Nature Astronomy 3 (2019) 24]

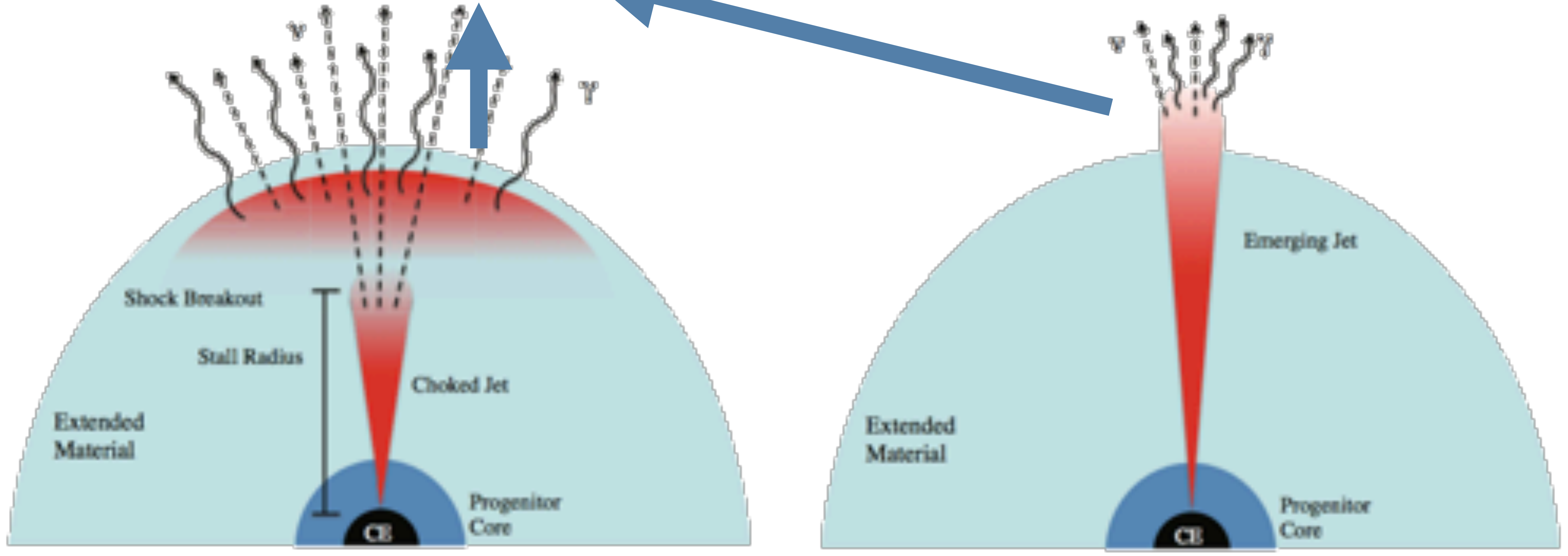
VHE emission from γ-ray bursts

- Recent first time detection of GRB afterglows the at very-high energies (Cherenkov telescopes)
  - GRB 180720B - H.E.S.S. (T_0 + 10 hr)
  - GRB 190114C - MAGIC (T_0 + 57s)
  - GRB 190829A - H.E.S.S. (T_0 + 4.3 hr)
    - ➜ Low-luminosity GRB with possible shock-breakout + jetted prompt emission

Chad et al (2020) [arxiv:2001.00648]



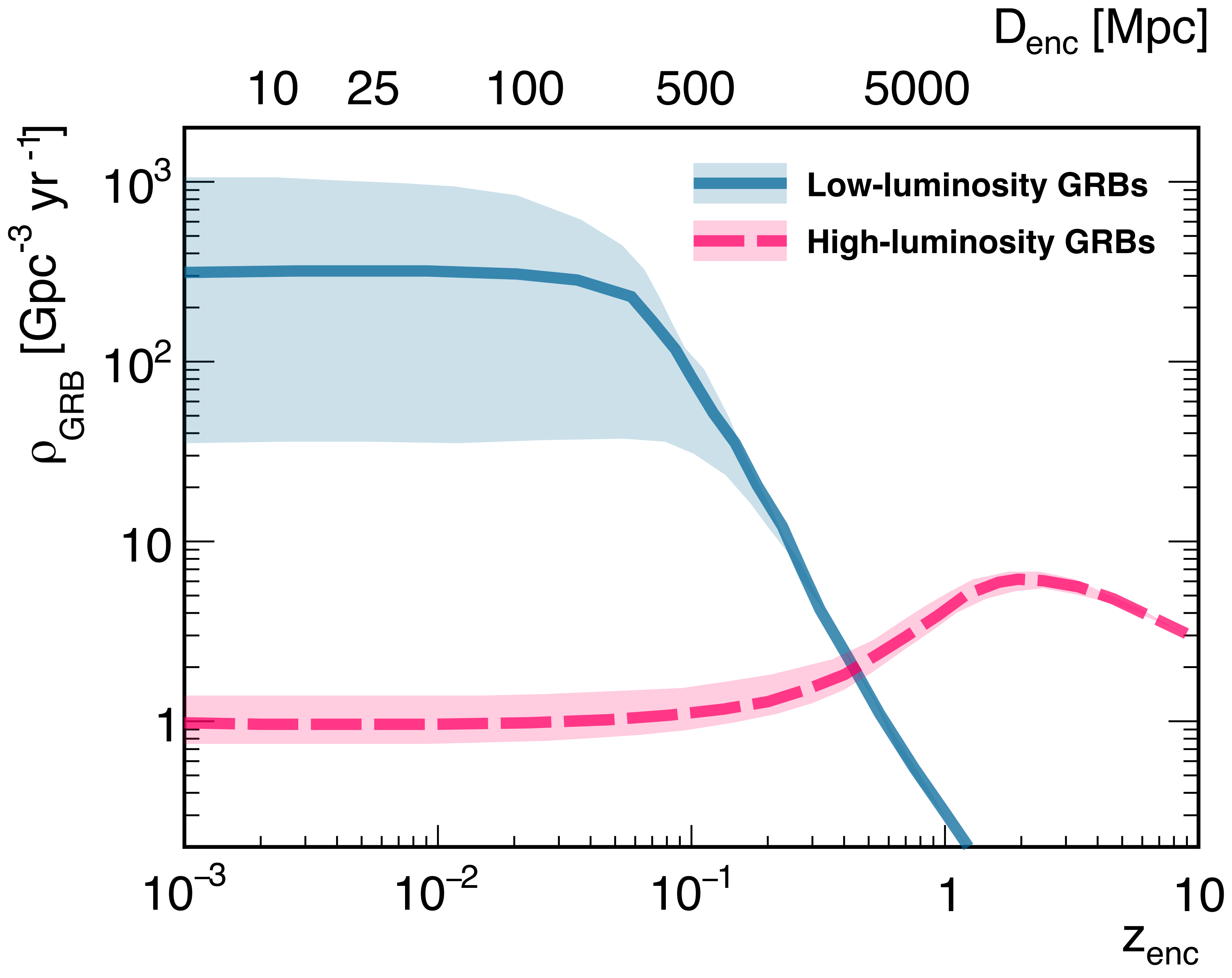
VHE γ-ray afterglow ➜ H.E.S.S.



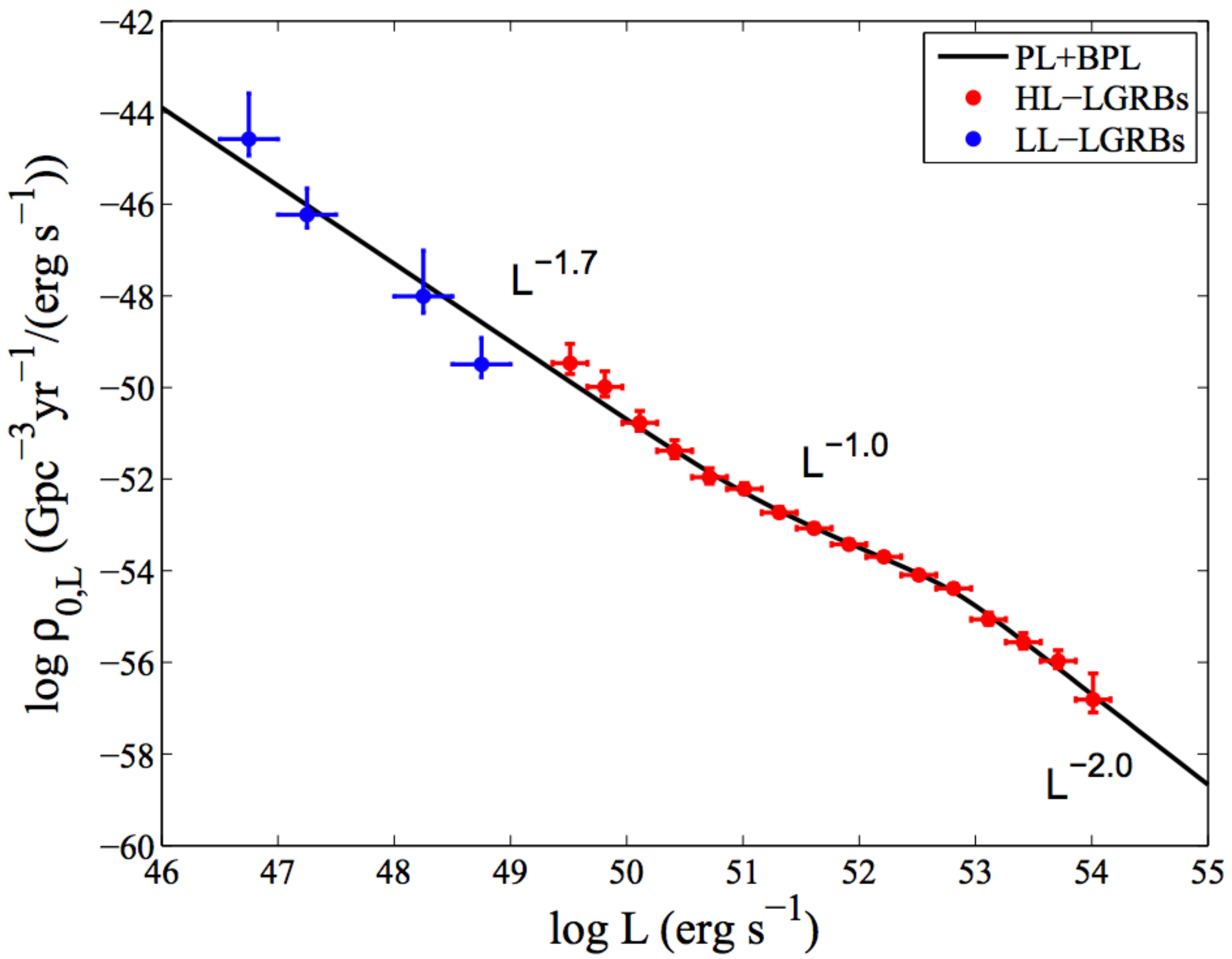
Prompt stage L_{iso} ~ 10^{48}-10^{49} [erg s^{-1}]

Low luminosity GRBs

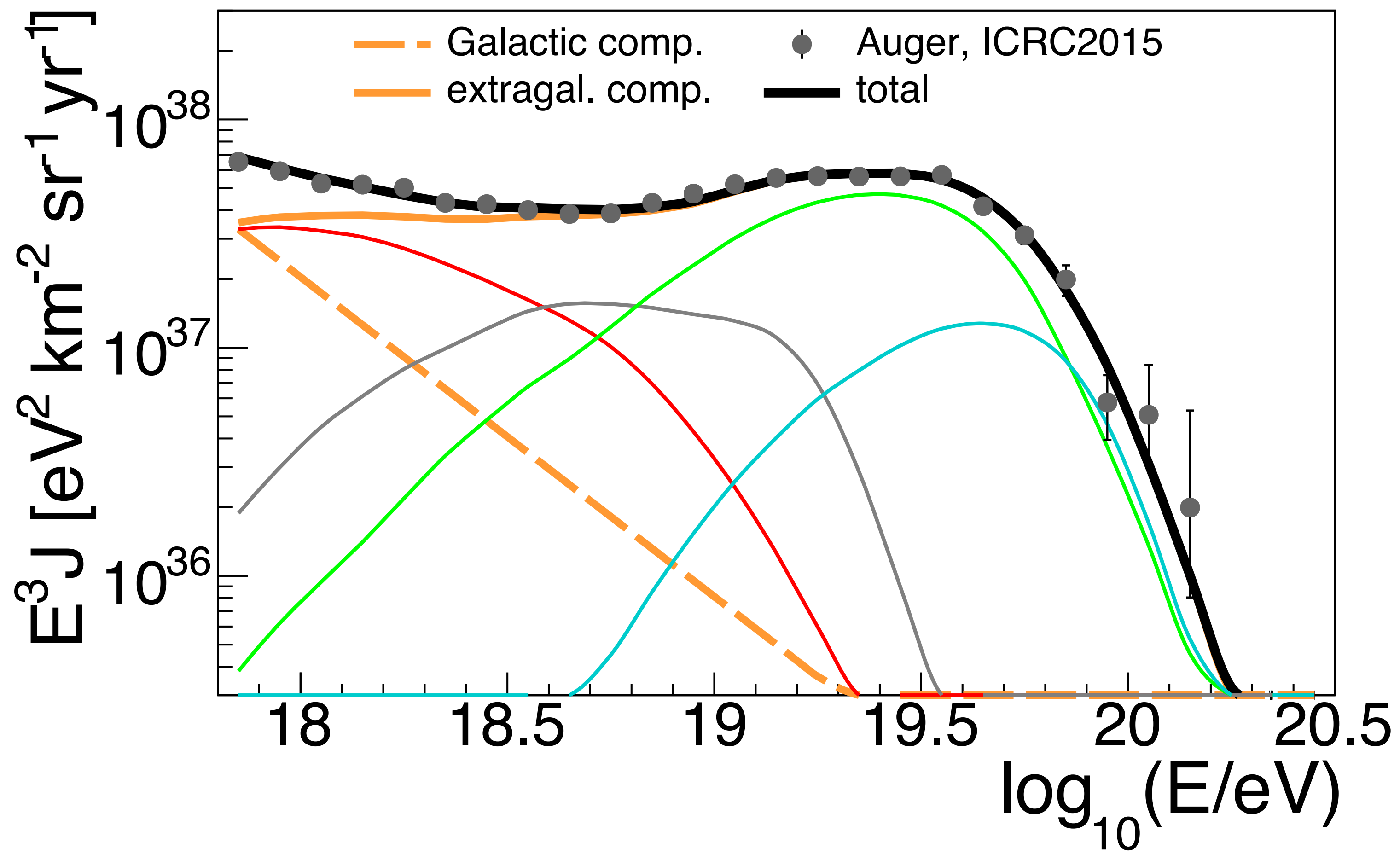
- Physical origins
  - ~1% of SN Ic, broad-line relativistic SNe
  - Possible connections to choked jets & shock-breakouts
  - Possible association with UHECRs & neutrinos
- Phenomenology
  - Typically having isotropic-equivalent luminosities, $10^{46} < L_{\gamma,iso} < 10^{48-50}$ [erg s^{-1}] • Low peak X-ray & Lorentz factor values • Only ~15 known events associated with SNe
  - Expected high event rates ➜ probe SNe & GRB physics • Targets for serendipitous γ-ray detection



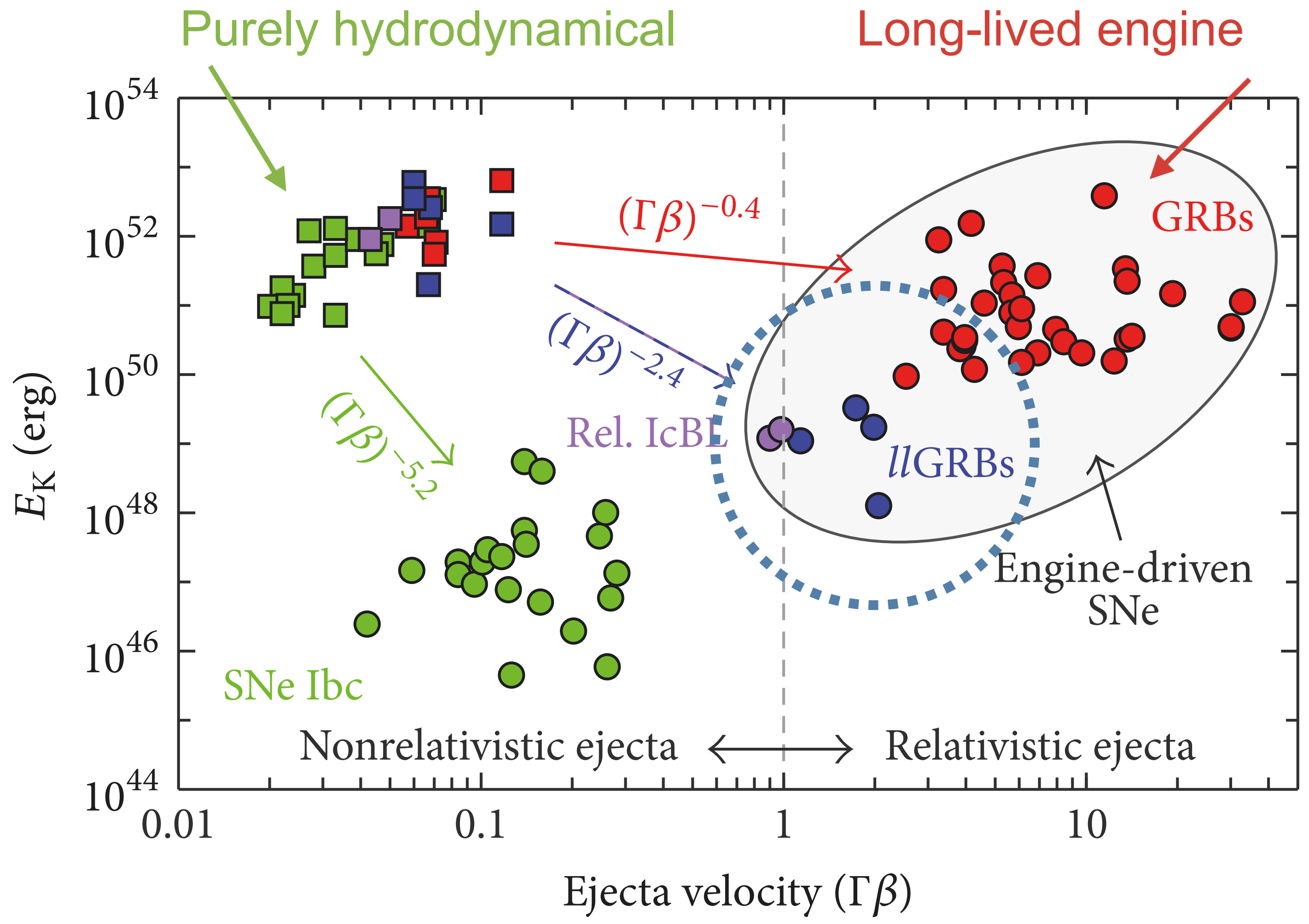
Liang et al (2007) arxiv:0605200



Sun et al (2015) arxiv:1509.01592



Boncioli et al (2018) arxiv:1808.07481



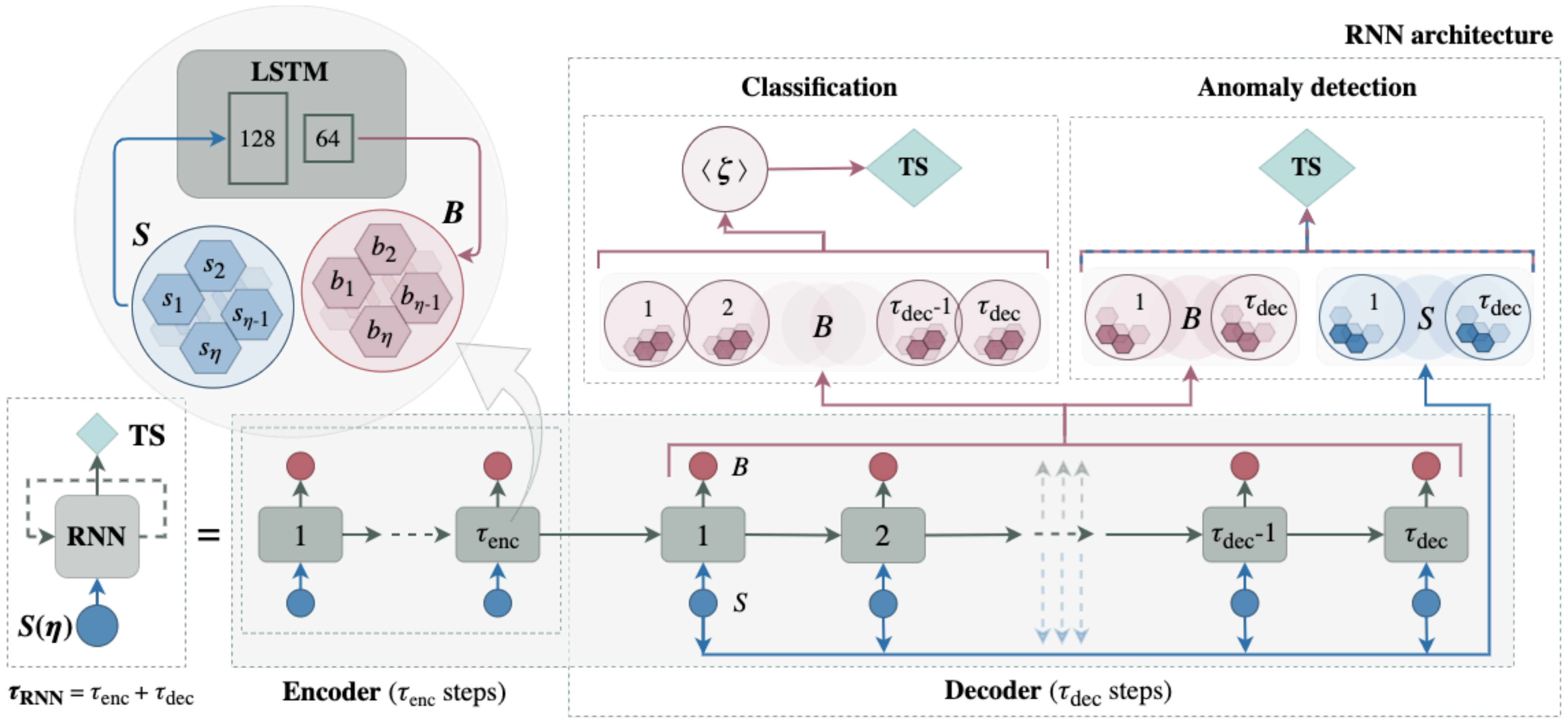
Cano et al (2016) arxiv:1604.03549

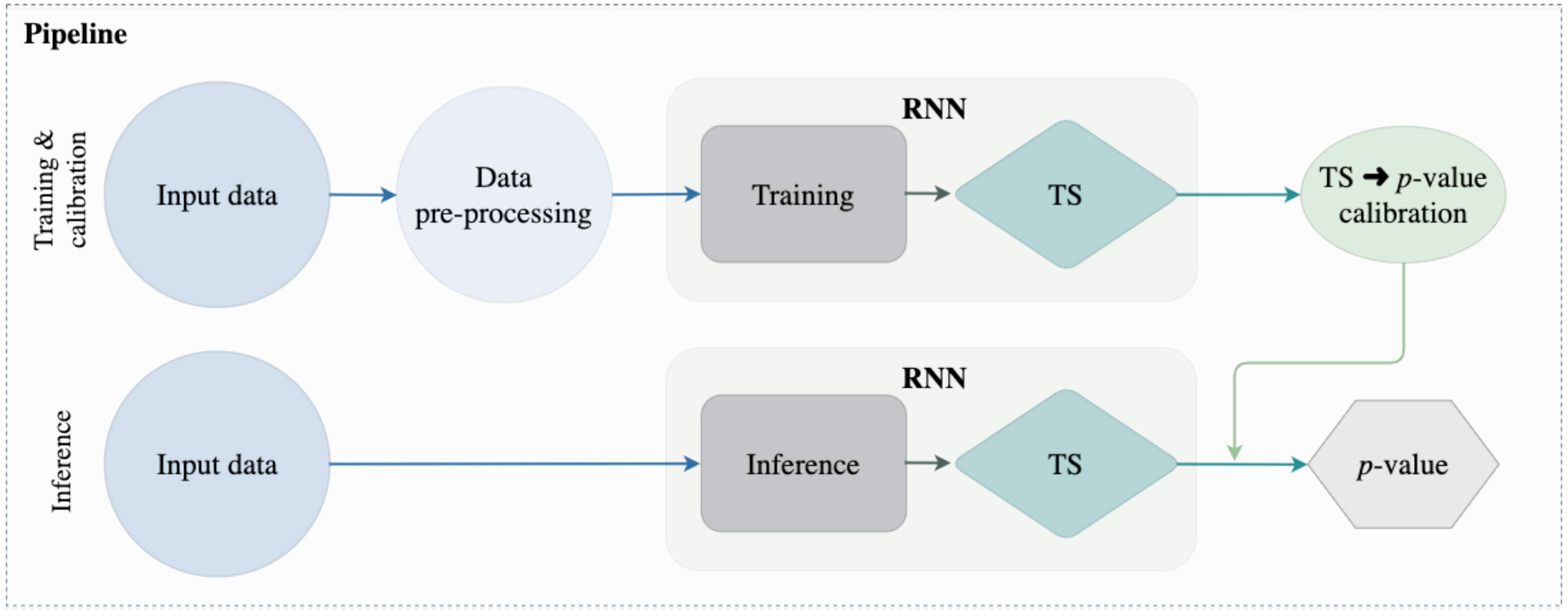
MMS transient detection

- **MMS observations**
  - Strategies
    - Real-time detection of signals in multiple channels
    - Near- and late-time follow-up for direct association of events
    - Archival stacking/population studies
    - Correlation of multiple low-significance observables, which combined may result in meaningful detections
  - Challenges
    - Uncertainties on instrument simulations (e.g., detector efficiency)
    - Uncertainties on physical backgrounds (e.g., galactic foregrounds)
    - Precise modelling of observing conditions (e.g., clouds, night-sky background)
    - Subtraction of artefacts (e.g., stars, satellites)
    - Extremely quick follow-up with multiple MMS/MWL facilities is necessary
- **Machine learning anomaly detection approach**
  - Training exclusively with real data ➜ mitigates systematics (no imperfect simulations used)
  - Does not require explicit physical modelling of perspective sources (generally not well constrained)
  - Facilitates data-fusion of inputs from different experiments
  - Extremely fast for evaluation, enabling quick response and coordination between facilities

Recurrent neural networks for transient detection

- **Two methodologies for source detection**
  - Anomaly detection
    - Train an RNN to predict a time-series of the expected background
    - Compare the predictions to the true time series ➜ identify a transient event as an anomalous flare
  - Classification
    - Train an RNN to classify a time series as background or signal, using labels
    - Training requires both background data and signal data (➜ introduces some model dependence)
- **Calibration pipeline**
  - The results are calibrated statistically ➜ significance / p-value estimates for discovery

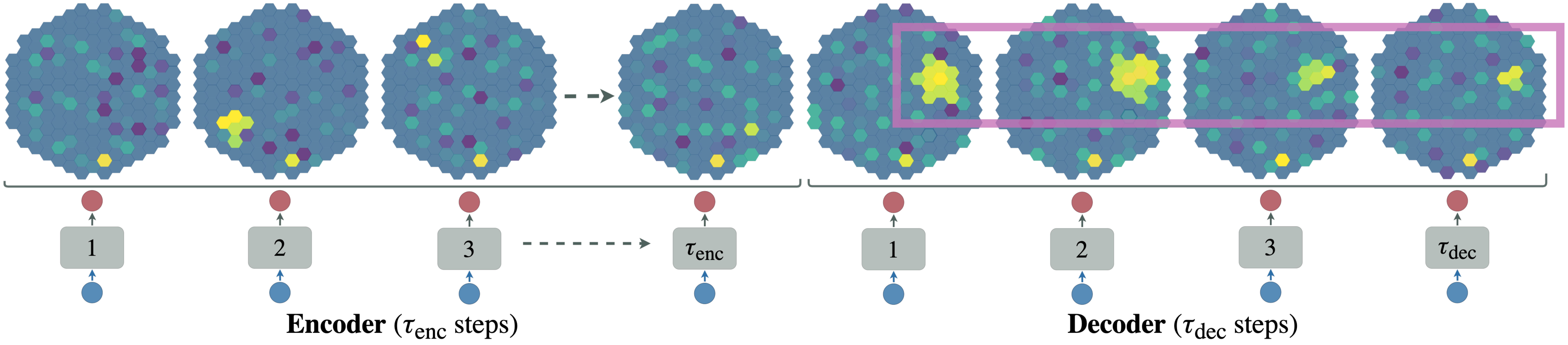




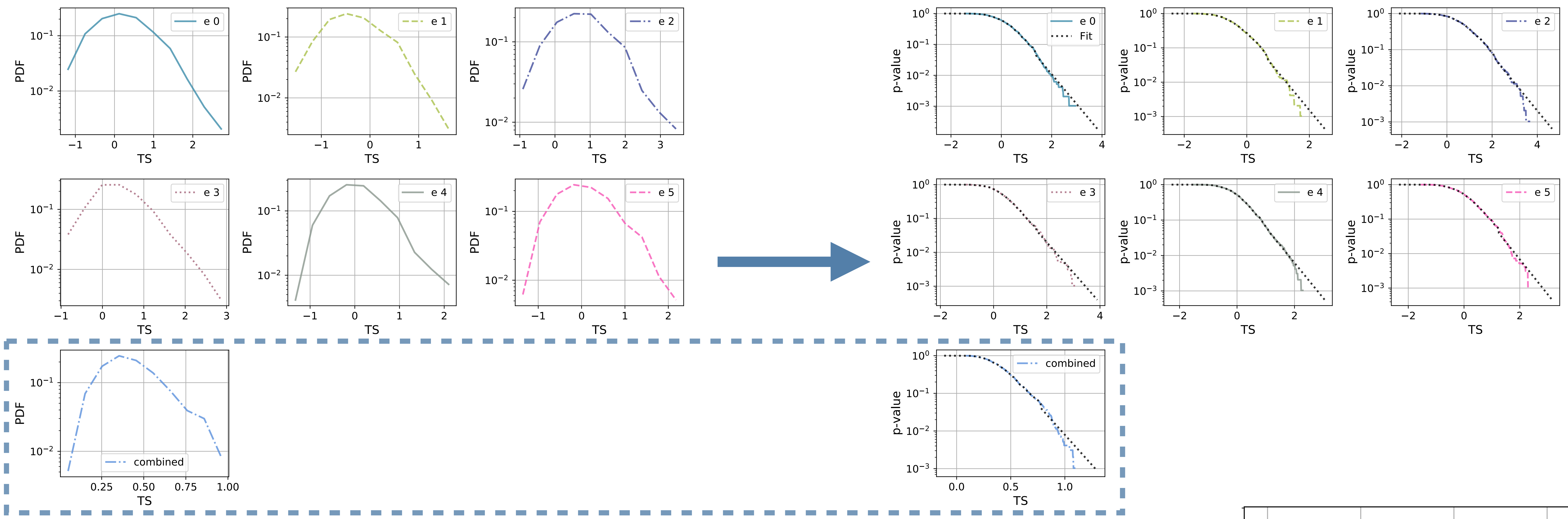
Sadeh (2020) arxiv:2005.06406

γ-ray transients

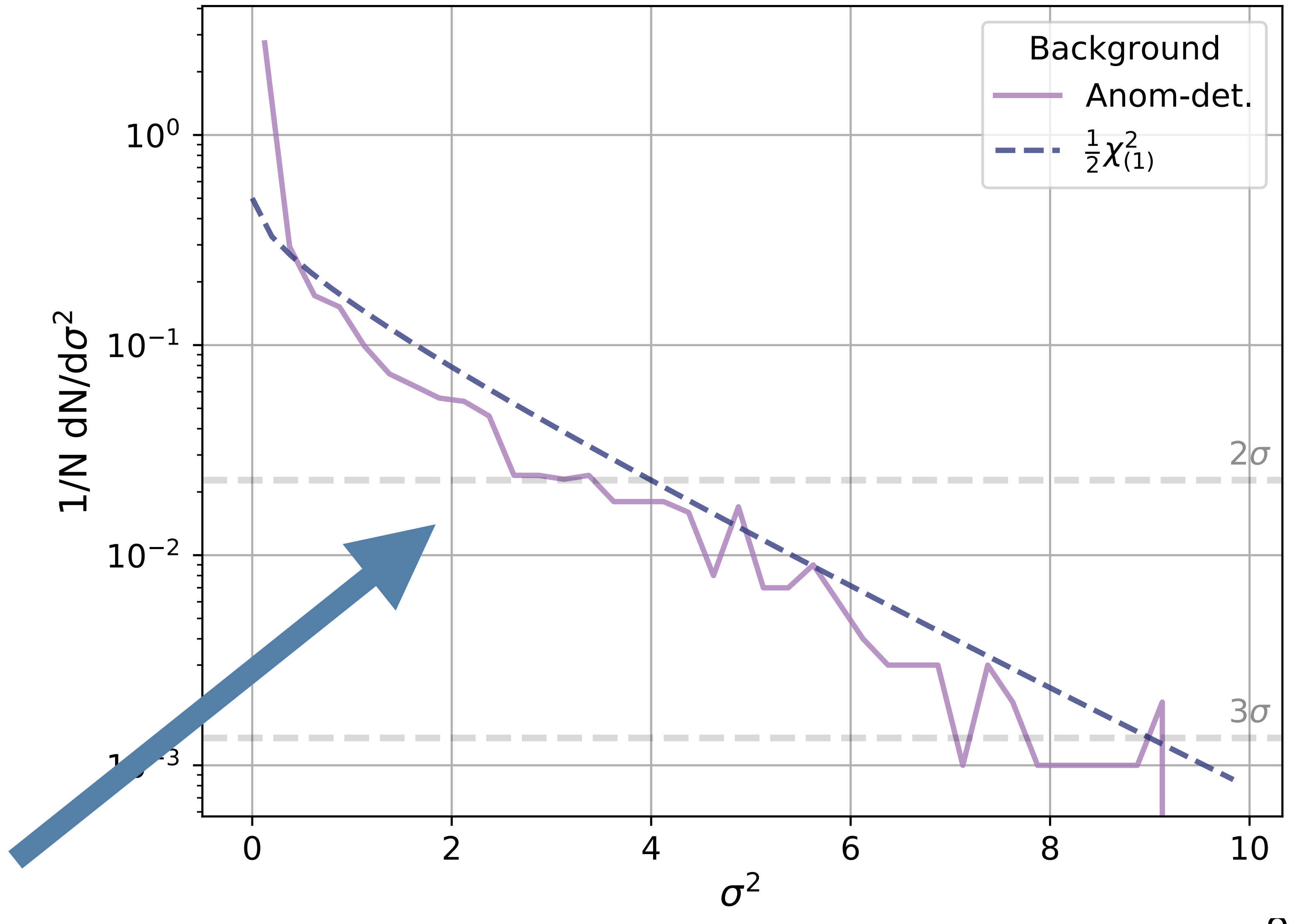
- **Example for the Cherenkov telescope array (CTA)**
  - Methodology
    - Train an RNN to predict a time-series of γ-ray event counts (binned in time & energy bins)
    - Add “auxiliary” input data, which affect the γ-ray rates (e.g., zenith of observation)
    - Compare the predictions to the true γ-ray rates, and identify a transient event as an anomalous flare
  - Training strategy
    - Anomaly detection: training exclusively on background data ➜ no-source in the region of interest; data potentially scrambled in time
    - Classification: also use simulations of GRBs ➜ simple spectral and temporal templates



Significance calibration for anomaly detection

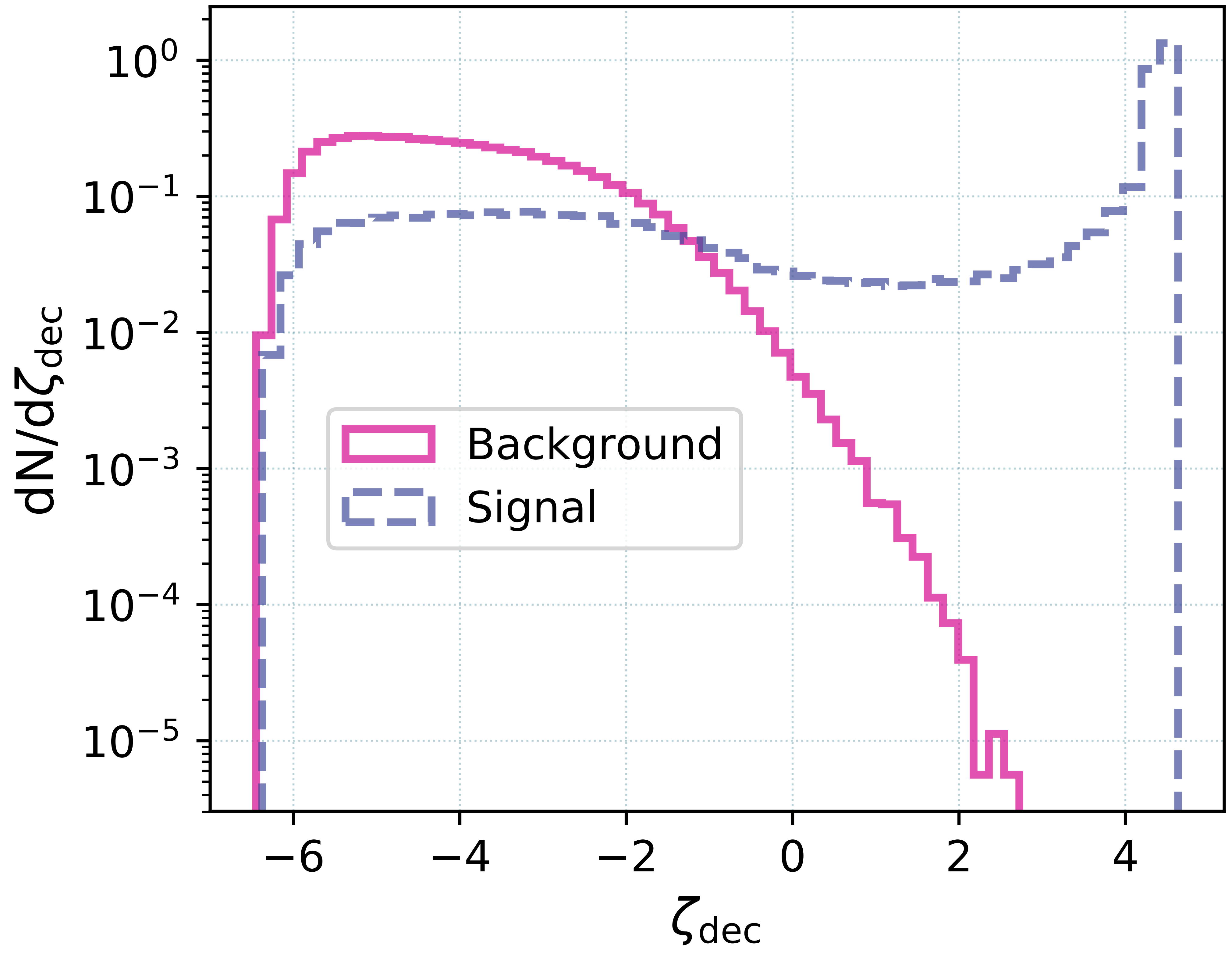


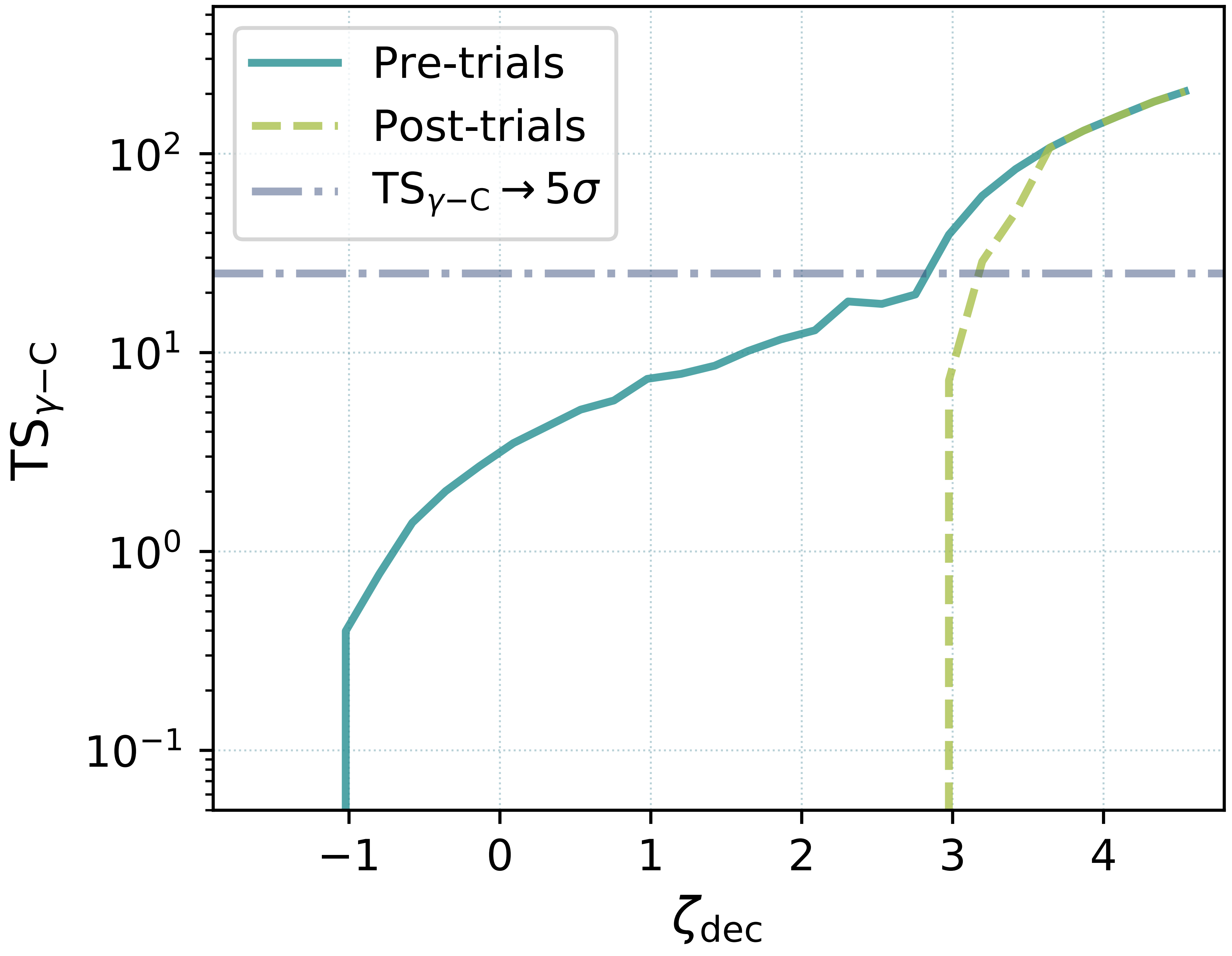
- In this example, the outputs of the RNN are γ-ray event counts in 6 energy bins
- Calibration procedure
  - Calculate a test statistic (TS) for each metric (based on the normalised difference between the RNN predictions and the ground truth)
  - Map TS ➜ p-values from TS distribution
  - Derive combined TS from the logarithms of individual p-values
  - Map combined TS ➜ combined p-value from distribution
- The combined TS distribution is compared to the expected background hypothesis



Significance calibration for classification

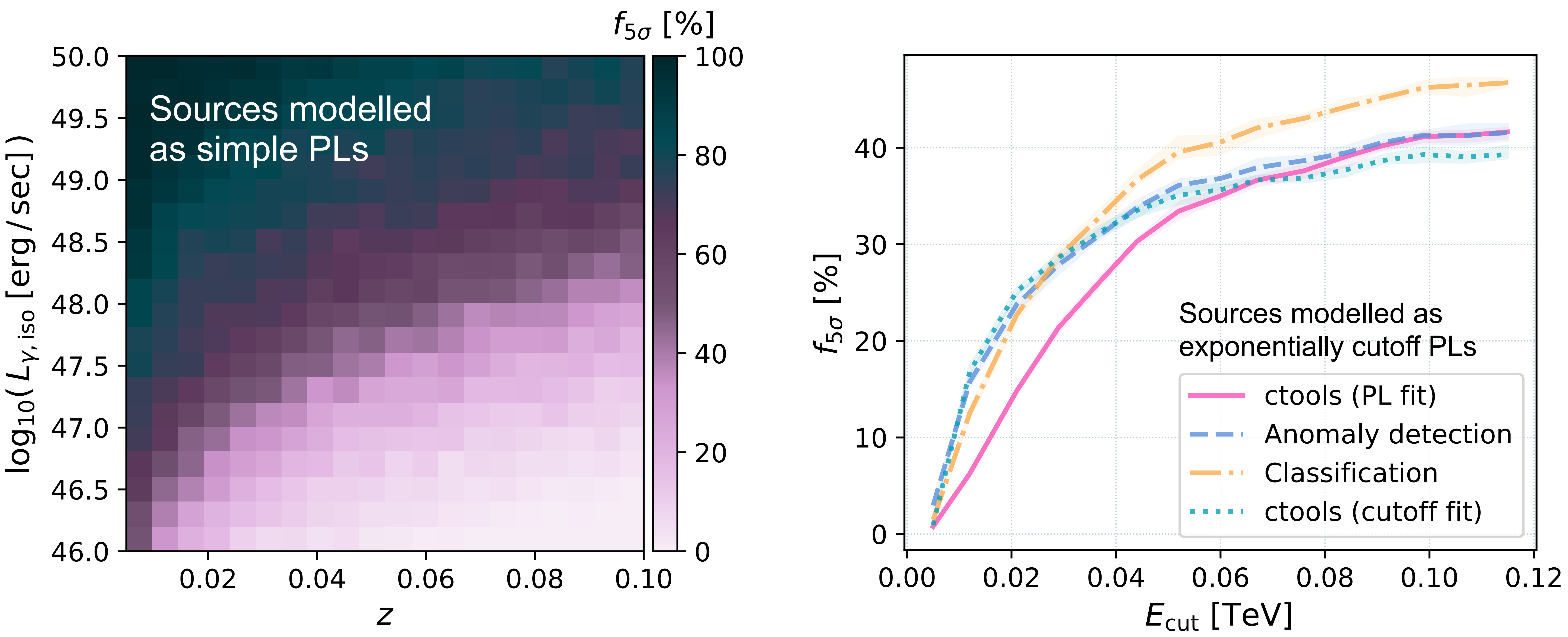
- In this example, the output of the RNN is a classification estimator, ζ_{dec}
- ζ_{dec} is evaluate for the background and signal samples individually
- The TS is derived from the ratio of the distributions as a function of ζ_{dec}
- TS ➜ p-value mapping is based on Wilks' theorem





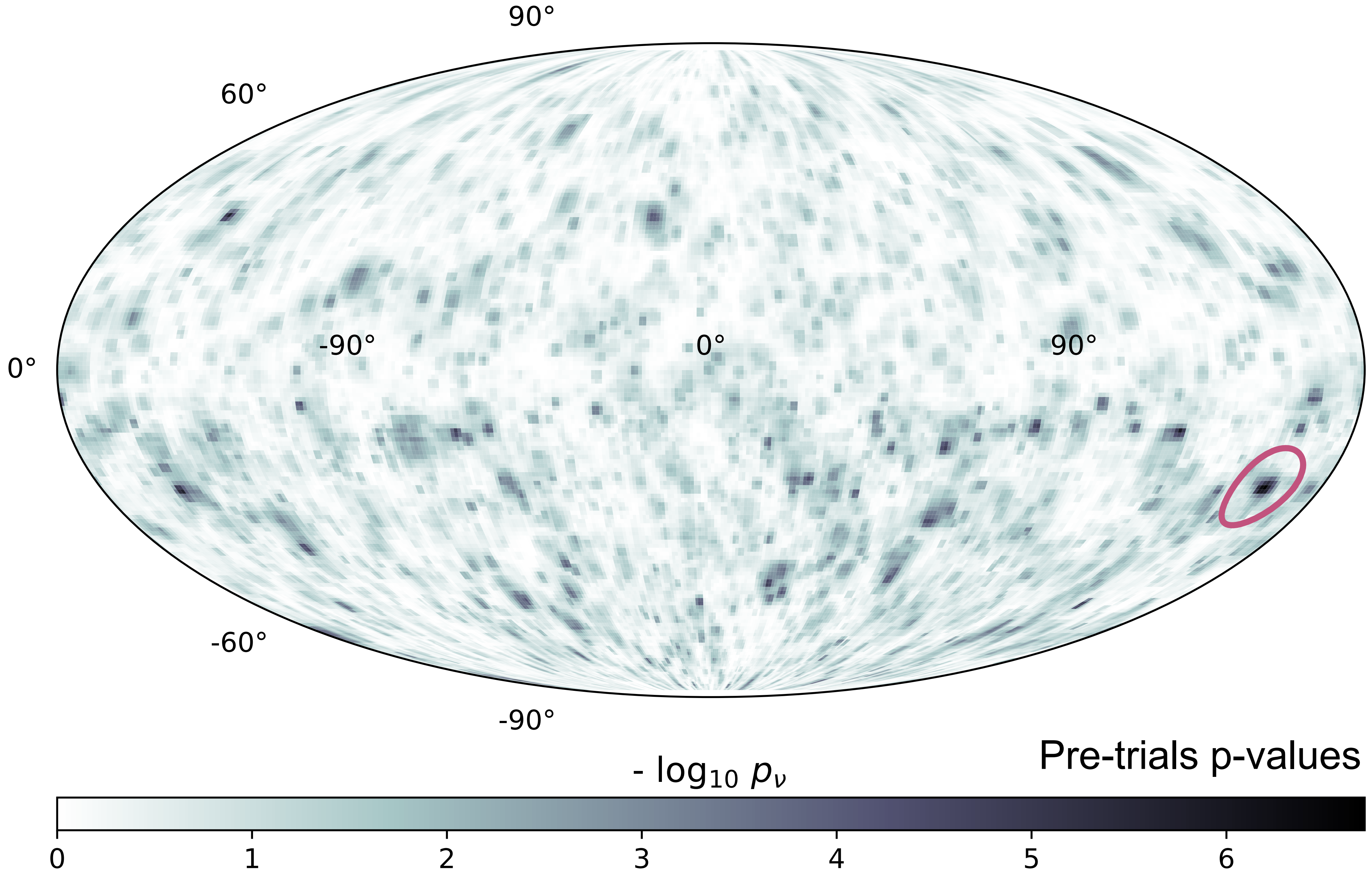
Serendipitous γ-ray transient detection

- Methodology
  - Shown here for a sample with expected properties for LL-GRBs, assuming either simple power-law (PL) or exponentially cutoff spectral PL models.
  - The reference detection rate (ctools) indicates a likelihood-based method, implemented as part of the ctools software package for CTA simulations
- Main takeaways
  - When simple PL models are fit the the data, both RNN methods perform better than the likelihood approach



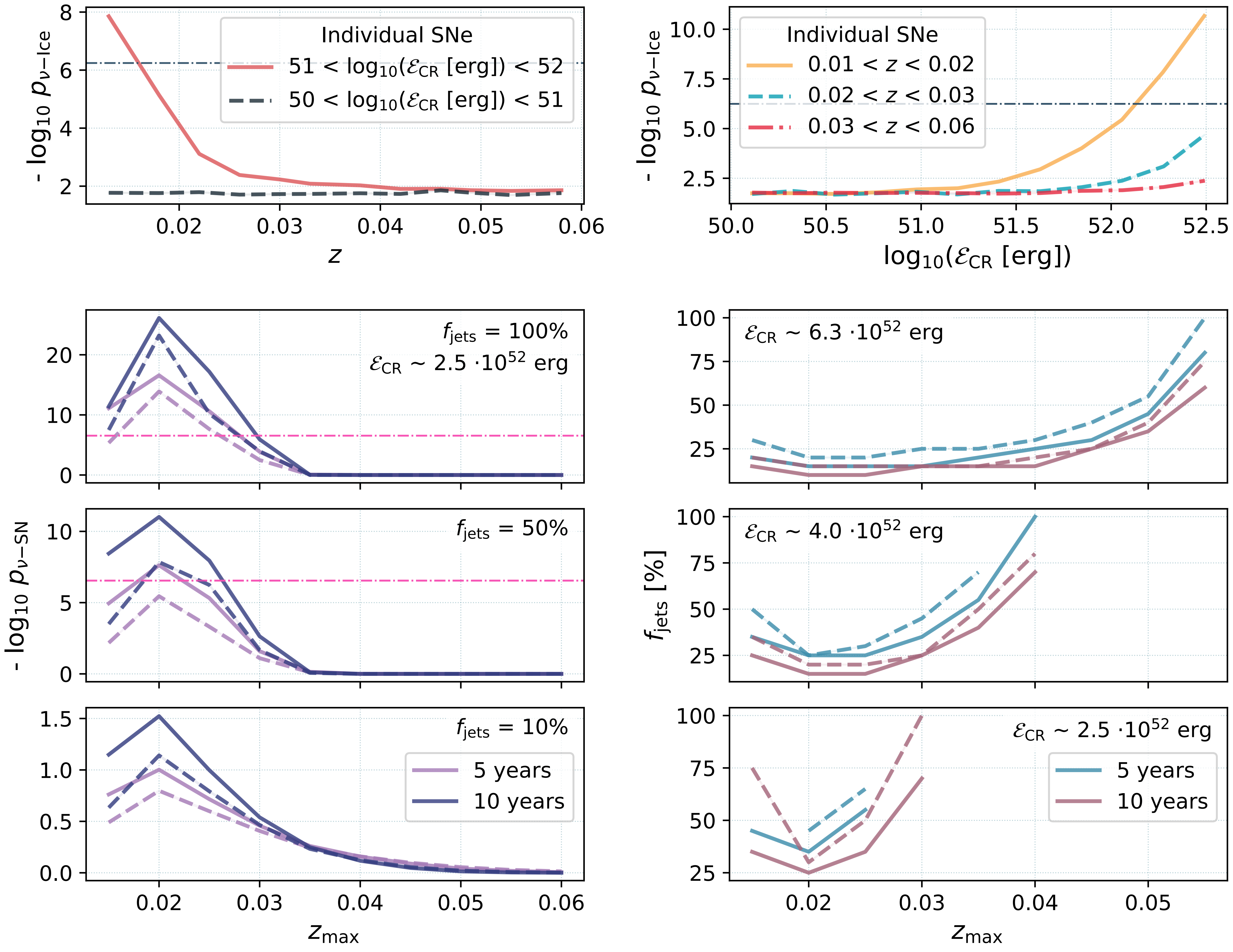
Neutrino point source search

- **Anomaly detection search for point sources in IceCube & ANTARES public datasets**
  - Methodology
    - Data from the two observatories are combined into a single RNN
    - Data are binned in 1-day intervals
    - Reference background dataset derived from same dataset, scrambled in arrival time and R.A.
    - RNN inputs are defined as neutrino event densities
    - TS ➜ p-value mapping includes trials correction (spacial and temporal)
  - Results
    - No source found (best post-trials significance ➜ 1.6σ)
    - No correlation found between IceCube & ANTARES
  - Main takeaways
    - No need to explicitly define the atmospheric neutrino background rates
    - No need to explicitly model the relative response between the two experiments
    - Trial factors are automatically taken into account as part of TS calibration



Neutrino / SNe correlation study

- **Anomaly detection search for neutrino emission (IceCube) correlated with core-collapse SNe (LSST)**
  - Methodology
    - Simulate observations for different LSST survey profiles
    - Detect SNe from the optical sims
    - Correlate with neutrino densities in spatio-temporal coincidence with the expected explosion time of the SNe
  - Main takeaways
    - Alternative to direct association (no need for individual VHE neutrino trigger)
    - Trial factors are automatically taken into account as part of TS calibration
    - No need for explicit combined likelihood formulation of the optical ⊕ neutrino signals
    - RNN is also used to derive limits in case of non-detection



Closing remarks

- Anomaly detection enables minimally-biased detection of transients ➜ most of the usual simulations & modelling for such analyses are not explicitly needed
- Simple neural network architectures are sufficient in many cases ➜ no need to parameter tuning
- Searches may be conducted on different time scales, and are robust to missing data
- The outputs of the network are consistently mapped to p-values for source detection ➜ no need for explicit likelihood formulations of individual / combined experiments
- It is relatively simple to combine multiple signals of different types into a single network, which internally models the join probability of the background-only hypothesis • In case of complicated signals (e.g., GW waveforms), standalone networks may be combined • Similarity for SN classifiers, etc. …
- The same network may also be used to derive limits on none-detection
- The quick response of the network facilitates efficient follow-up alert generation