CrossMark

Data-driven Detection of Multimessenger Transients

Iftach Sadeh

DESY, Platanenallee 6, D-15738 Zeuthen, Germany

Abstract

The primary challenge in the study of explosive astrophysical transients is their detection and characterization using multiple messengers. For this purpose, we have developed a new data-driven discovery framework, based on deep learning. We demonstrate its use for searches involving neutrinos, optical supernovae, and gamma-rays. We show that we can match or substantially improve upon the performance of state-of-the-art techniques, while significantly minimizing the dependence on modeling and on instrument characterization. Particularly, our approach is intended for near- and real-time analyses, which are essential for effective follow-up of detections. Our algorithm is designed to combine a range of instruments and types of input data, representing different messengers, physical regimes, and temporal scales. The methodology is optimized for agnostic searches of unexpected phenomena, and has the potential to substantially enhance their discovery prospects.

Unified Astronomy Thesaurus concepts: Transient detection (1957); Time series analysis (1916); Gamma-ray transient sources (1853); Neutrino astronomy (1100); Neural networks (1933); Gamma-ray bursts (629)

1. Introduction

Multimessenger astronomy explores the universe by studying phenomena involving the electromagnetic, weak, strong, and gravitational forces, utilizing a variety of instruments (Huerta et al. 2019; Meszaros et al. 2019). Successful campaigns hinge on speedy and comprehensive coordination of observations. Examples include association of gravitational waves (GWs) from the binary neutron star merger, GW170817, with a short γ-ray burst (GRB; Abbott et al. 2017), as well as evidence for neutrino emission from the flaring blazar, TXS 0506+056 (Aartsen et al. 2018). The main observational strategies are:

1. real-time detection of signals in multiple channels;
2. near- and late-time follow-up for direct association of events;
3. archival stacking/population studies;
4. correlation of multiple low-significance observables, which combined may result in meaningful detections.

Such cross-domain analyses must reconcile differences in instrument sensitivities, spatial and temporal coverage, and measurement uncertainties. In the following we present a new data-driven deep learning (DL) framework, designed to tackle these challenges.

Machine learning is a computational technique, where learning from examples takes the place of explicit functional modeling. DL is a type of machine learning, based on artificial neural networks (ANNs) (LeCun et al. 2015; Goodfellow et al. 2016). ANNs are computational models composed of *neurons*, inspired by biological brains. Individual neurons perform simple transformations on vectors of inputs, using weight and bias parameters that are modified during training. Abstract representations of data sets can be encoded by arranging neurons in multiple interconnected layers, using nonlinear *activation functions*. DL utilizes deep and wide layouts of ANN layers, able to represent complex models, avoiding the need for explicit feature engineering by domain experts. Effective training of these large architectures has become achievable due to advances in optimization algorithms and in computational resources. A prominent type of DL is the recurrent neural network (RNN). Unlike *feedforward* ANNs, where information passes through a network in one direction, RNNs include cyclic connections; this allows them to effectively process sequential data, such as time series. A variant of RNNs called a long short-term memory network (LSTM) has the ability to simultaneously retain information spanning different time-scales. Such networks have proven very successful for, e.g., speech recognition and natural language translation (Graves & Jaitly 2014; Sutskever et al. 2014).

Deep learning has been used for a variety of astronomical analyses (see e.g., Carleo et al. 2019; Muthukrishna et al. 2019, and citations therein). The nominal approach has been to perform *supervised learning*, based on labeled data. Examples include GW waveforms (George & Huerta 2018a, 2018b), neutrino detector data (Choma et al. 2018), and optical images (Khan et al. 2019). When used for source detection, the inferred rate of *false positives* of these methods strongly depends on the *completeness of the training data*. For instance, one must verify that training includes all possible sources of systematics, including *glitches* (non-Gaussian noise; Zevin et al. 2017; George et al. 2018; Wei & Huerta 2020). One must also take care that such systematics are distributed in realistic proportion to each other and to true signal events. In principle, reliability of detection may be improved by imposing additional constraints (e.g., coincidence between detectors, as introduced by George & Huerta 2018a). However, it remains challenging to directly interpret classification outputs as detection probabilities (Gebhard et al. 2019).

An alternative strategy is to use data-driven *anomaly detection*. The latter is the task of identifying data that differ in some respect from a reference sample (Pimentel et al. 2014). Anomaly detection has been used in the past in different contexts. For instance, LSTMs have been utilized to detect hardware failure in medical and industrial data sets, assuming Gaussian anomaly distributions (Malhotra et al. 2015). For astronomy, traditional learning methods have mostly been employed, such as principal component analysis (Williamson et al. 2019) or *isolation forests* (Pruzhinskaya et al. 2019).

We present a novel use of anomaly detection for the discovery and characterization of astrophysical transients,

utilizing LSTM–RNNs. The networks are coupled to a statistical pipeline used to interpret the results. Our framework facilitates the combination of data sets of various types. It therefore enables derivation of realistic joint (multimessenger) probability distributions. These may be used for discovery, or for setting limits in cases of non-detection. Training samples may nominally be derived from an experiment in situ, or from historical data. Compared to existing approaches, this circumvents the potential pitfall of relying on unrepresentative reference data sets. Our method mitigates uncertainties on instrumental modeling and on physical backgrounds (e.g., galactic foregrounds). It also avoids the need for the explicit characterization of observing conditions, or of artifacts (e.g., stars in the field of view). Finally, this data-driven approach allows for agnostic searches, as minimal assumptions need to be taken on the properties of putative sources.

We illustrate our algorithm for three analyses, which represent the primary detection strategies of multimessenger astronomy: real-time source detection, population studies, and the correlation of (low-significance) observables. Specifically, we present examples for the study of high-energy transients, as observed with neutrinos, optical data, and γ-rays.

2. Network Architecture and Inference Pipeline

Our software pipeline and chosen RNN architecture are shown in Figure 1. The RNN may be decomposed into two elements, an *encoder* and a *decoder* (Cho et al. 2014). The RNN accepts $\tau_{RNN} = \tau_{enc} + \tau_{dec}$ time steps as input. The first τ_{enc} encoder steps represent the background interval, before a transient appears. A potential signal event is searched for within the following τ_{dec} decoder steps. For each of the example analyses described below, the value of τ_{RNN} is chosen according to the expected properties of signals (e.g., physical timescales), accounting for the structure of the available data. Each step receives a collection of (analysis-specific) η inputs.

In our nominal approach we employ an anomaly detection technique. We utilize sequences of input data, $S(\tau_{RNN}, \eta)$, which for training correspond to the response of an instrument in the absence of signal events. The RNN is used to predict the expected background of the experiment, $B(\tau_{dec}, \eta)$. Transients are then detected as significant divergences from these predictions. We define a unique test statistic (TS), which encapsulates the difference (*mean squared error*) between S and B for each analysis.

Additionally, we employ a complementary classification approach (illustrated for the γ-ray analysis example). In this case, the RNN is used to to directly classify transient events, rather than to predict the background. Correspondingly, the network is trained using labeled examples of both background and putative signal data. The discovery TS is based on the ratio between the background and signal classification probabilities (Cranmer et al. 2015; Goodfellow et al. 2016).

The TS output of the RNN is coupled to a pipeline that is used to derive the significance of detection. Nominally, we do not assume a specific statistical model for the background or for the signal. We therefore generate multiple realizations of the input background sample, from which we derive cumulative distributions of TS. The latter are used to calibrate the relationship between TS-values and p-values for a background-only test hypothesis.

We train the RNNs using *Adam* optimization with a *learning rate* of 0.01. For each of the three anomaly detection analyses, we have 10^4 background sequences of τ_{RNN} steps. For the

single classification example, an additional 10^4 signal events are used as well. The inputs are independently normalized, such that their nominal range of values for the background training sample is mapped to the interval, [0, 1]. Data are randomly split into batches of 64 sequences for training, where 20% of events are set aside for validation. In order to mitigate over-fitting, we apply 30% *dropout* training regularization (random masking of units). The process of training for the various analyses lasts several minutes on a laptop–CPU, with corresponding sub-sec inference (well below the requirements on the relevant real-time applications).

We compared several configurations of RNN hyper-parameters (internal configuration parameters of the network). For instance, this included doubling the number of LSTM layers, increasing the batch size, varying the learning rate, etc. In addition, we performed analysis-dependent systematic checks, as described below. We found no significant variation in the results. The robustness in performance is due to our design choice of a simple RNN, and to the fact that we rely on data-driven calibration of test statistics post-training.

In the following, we illustrate our framework using concrete examples. We restrict the discussion to the general features of the method in each case, such as relevant timescales, RNN inputs, and significance of detection. For a comprehensive description of the data sets, definition of test statistics, source modeling, and systematic tests, see the Appendix.

3. Results

3.1. Neutrino Point-source Search

A diffuse TeV–PeV flux of astrophysical neutrinos has been discovered by IceCube (Aartsen et al. 2013). While the exact nature of the emission remains elusive, its apparent isotropy suggests that it originates from relatively weak extragalactic sources. Possible sources of ultra high-energy cosmic rays (UHECRs) and neutrinos include long- and short-duration gamma-ray bursts (GRBs), as well as core-collapse supernovae (CC-SNe) with choked jets or shock breakouts (Kashiyama et al. 2013; Senno et al. 2018). The energy density of the astrophysical neutrinos is comparable to that of the isotropic γ-ray background and to that of UHECRs (Meszaros et al. 2019). This indicates that multimessenger interpretation may lead to breakthroughs in our understanding of cosmic ray (CR) accelerators. It may elucidate the connection between supernovae (SNe) and GRBs, and may shed light on the nature of their central engine. Searches commonly take the form of either spatio-temporal clustering of neutrinos, or their direct association with steady or transient sources (Aartsen et al. 2015).

We search for clustering in two all-sky samples. The first comprises IceCube event lists of track-like muon neutrino candidates, taken between 2011 and 2012 (MJD 55694–56415; Aartsen et al. 2017). The second is of ANTARES muon neutrinos, observed within the same time period (Adrián-Martínez et al. 2014). The nominal inputs to our RNN are based on *neutrino event density* metrics (see Appendix A.2). The densities are defined with respect to a given location on the sky. For the IceCube sample, the data are split into four logarithmically spaced bins in the energy proxy, between 10 GeV and 8 PeV. For ANTARES the metrics are inclusive in energy. The data are integrated over 24 hr time periods, which effectively avoids dependence of event rates on R.A. (Aartsen et al. 2015). The response of the IceCube and ANTARES detectors depends
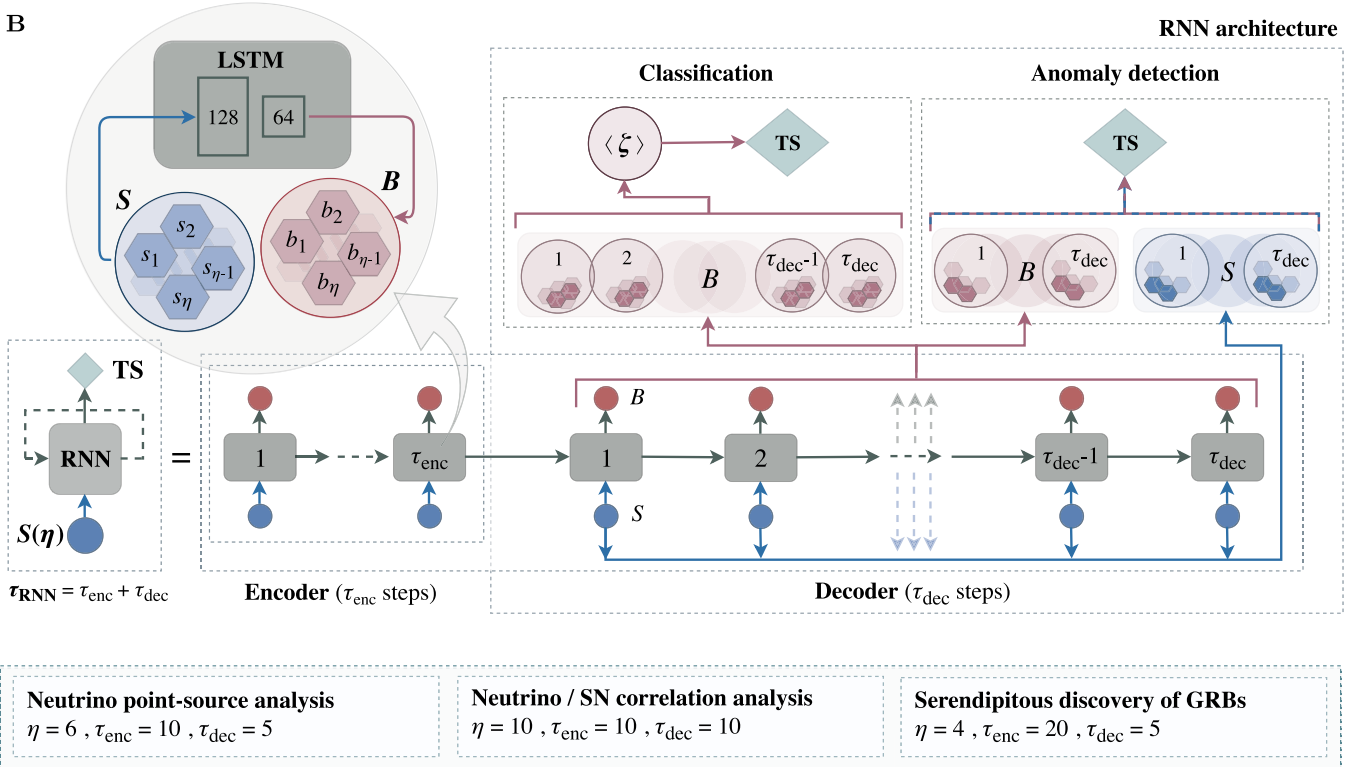
A
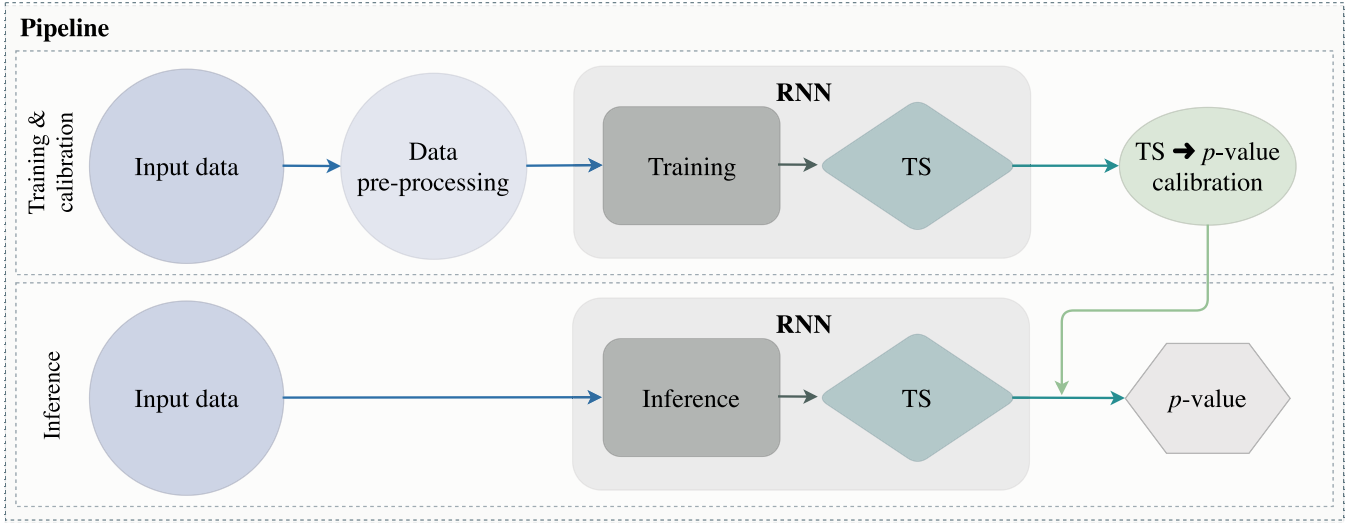
B

Figure 1. Schematic illustrations of the software pipeline, and of the architecture of the RNN. (A) The pipeline comprises two main phases, *training/calibration* and *inference*. Training includes a pre-processing stage for generation of background simulations. These data are used to train the RNN and calculate test statistics (TS). The latter are mapped to *p*-values as part of the *calibration* phase. Inference includes evaluation of TS- and *p*-values, using the trained RNN and the pre-calculated calibration. (B) The network may be decomposed into an *encoder* ($\tau_{\rm enc}$ time-steps) and a *decoder* ($\tau_{\rm dec}$ time-steps), where the decoder represents the search interval. The RNN is made up of LSTMs (green rectangles). Each LSTM comprises two layers of 128 and 64 hidden units. The input data, $S(\tau_{\rm dec}, \eta)$ (blue circles), make up η numbers for each time step (blue hexagons). Similarly, the outputs of the LSTMs are indicated as $B(\tau_{\rm dec}, \eta)$ (red circles). For *anomaly detection*, the decoder inputs and outputs, S and B, are directly used to calculate the TS for discovery. For *classification*, the decoder outputs are fed into logits, $\zeta(\tau_{\rm dec})$ (proxies for classification probabilities), which are used to derive the corresponding TS. The particular set of parameters used for each one of the example analyses (η, $\tau_{\rm enc}$, and $\tau_{\rm dec}$) are are shown in the bottom panel. The specific choices are described in detail in the text.

on the zenith of the observed event, which is therefore added as input to the RNN. In cases where an IceCube source is not within the field of view of ANTARES, we use background data instead. In total, we have $\eta = 6$ inputs per time step. The collection of inputs is derived for $\tau_{\rm RNN} = 15$ days periods. The first $\tau_{\rm enc} = 10$

days are assumed to only contain background events. Transient signals are searched for within the next $\tau_{\rm dec} = 5$ days interval. In total, the RNN receives $6 \times 15 = 90$ inputs.

The RNN is trained to predict the neutrino event densities in each of the five days being probed, for a particular sky position.
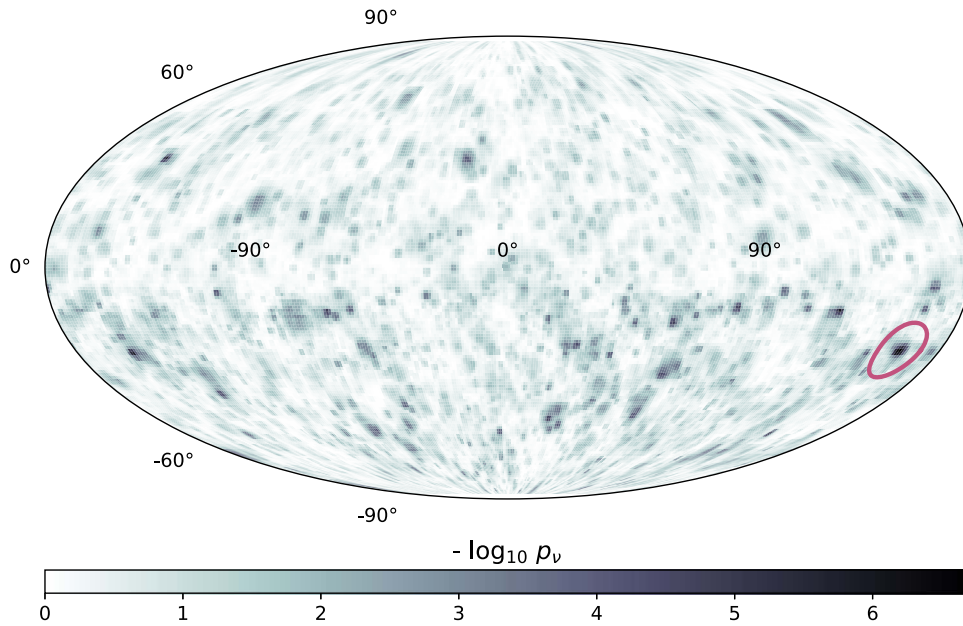
Figure 2. Results of the neutrino point-source search analysis. The sky-map in equatorial coordinates displays the pre-trials p-values, p_ν, for time-dependent neutrino point sources, utilizing IceCube and ANTARES data from 2011 to 2012. The most significant point ({R.A., Decl.} = {163°, −26°.5}) is indicated by the red oval. It has $p_\nu = 1.9 \times 10^{-7}$ and $p_\nu = 0.1$ pre- and post-trials, respectively, corresponding to 5.2σ and 1.6σ significance. The result does not correspond to any astronomical object of interest, and is consistent with a statistical fluctuation.

As part of anomaly detection, these predictions are meant to correspond to the background. Potential transient signals must therefore be removed from the training data set. This is done by scrambling neutrino events in R.A. and time of detection.

The test statistics for detection are based on differences between the predictions of the network and the true data. For our particular definitions of observables, one can not assume a priori that a particular statistical model (e.g., Poissonian counts) would be valid. We therefore calculate detection probabilities in a more general approach, using simulations. We consider the zenith to be an *auxiliary RNN parameter*, as it is not directly used to derive detection probabilities. Rather, we bin the data in zenith into 90 intervals of 2° width, in which the detector response is approximately uniform. We then generate ∼10^6 scrambled background sequences for each bin and evaluate them with the RNN. The derived distributions of test statistics are used to parameterize the correspondence between TS- and p-values, accounting for all spatial and temporal trials.

We proceed to use the trained network to evaluate the original (non-scrambled) data, carrying out an all-sky grid search with 0.3° spacing. The corresponding spatial distribution of p-values is shown in Figure 2. The most significant position is {R.A., Decl.} = {163°, −26.5°}. It has $p_\nu = 1.9 \times 10^{-7}$ (pre-trials) and $p_\nu = 0.1$ (post-trials), corresponding to 5.2σ and 1.6σ significance. The result does not correspond to any astronomical object of interest, and is consistent with a statistical fluctuation. We also use the outputs of the RNN to perform a correlation analysis between the IceCube and ANTARES events (see Appendix A.2), finding no significant result.

Our conclusions on the existence of neutrino sources are consistent with previous studies of these data, which did not detect any source ($p_\nu = 0.6$, by Aartsen et al. 2015). In this case, however, the analysis is done without the need to explicitly define likelihood functions for the background or sources. The study is performed on a combined IceCube and ANTARES data set. Another advantage of our approach, is that there is no need to model the relative response between the two neutrino observatories. This example also illustrates the flexibility of the methodology regarding observables. Our choice of inputs and test statistics is motivated by the properties of the data sets, but is by no means unique. However, our framework is designed to provide self-consistent detection probabilities in general, given a set of primary (e.g., neutrino densities) and auxiliary (e.g., zenith) parameters. Finally, we note that the limitations of the public data sets constrain us to ⩾ 1 day time bins. Provided that the full data-streams of the experiments become available, the same approach would be applicable for real-time searches on shorter timescales.

3.2. Correlation Analysis between Neutrinos and CC-SNe

An alternative to auto-correlation analyses is to search for cross-correlation between different messengers. An important physical example is that of CC-SNe with relativistic outflows (Murase & Ioka 2013; Cano et al. 2017). CC-SNe may accelerate UHECRs, which produce γ-rays and neutrinos, for instance, via pp or $p\gamma$ interactions. A direct connection between neutrinos and their SN counterparts can be made for events that also exhibit high-energy emission, e.g., GRBs. However, γ-rays are not always observed in association with CC-SNe. For example, depending on the environment within and around a source, γ-rays may become attenuated due to $\gamma\gamma$ interactions (Boncioli et al. 2019). A few percent of CC-SNe are estimated to harbor undetectable jets, which are not powerful enough to punch through their progenitors and winds. Such events are associated with choked jets, and possibly shock breakouts (Kashiyama et al. 2013). Conversely, those jets that are successful are mostly launched off-axis with respect to the observer, and thus are also undetected (Denton & Tamborra 2018). Finally, even when the high-energy emission is beamed toward Earth, individual SNe are unlikely to be identified as sources of neutrinos, as illustrated in Figures 3(A)
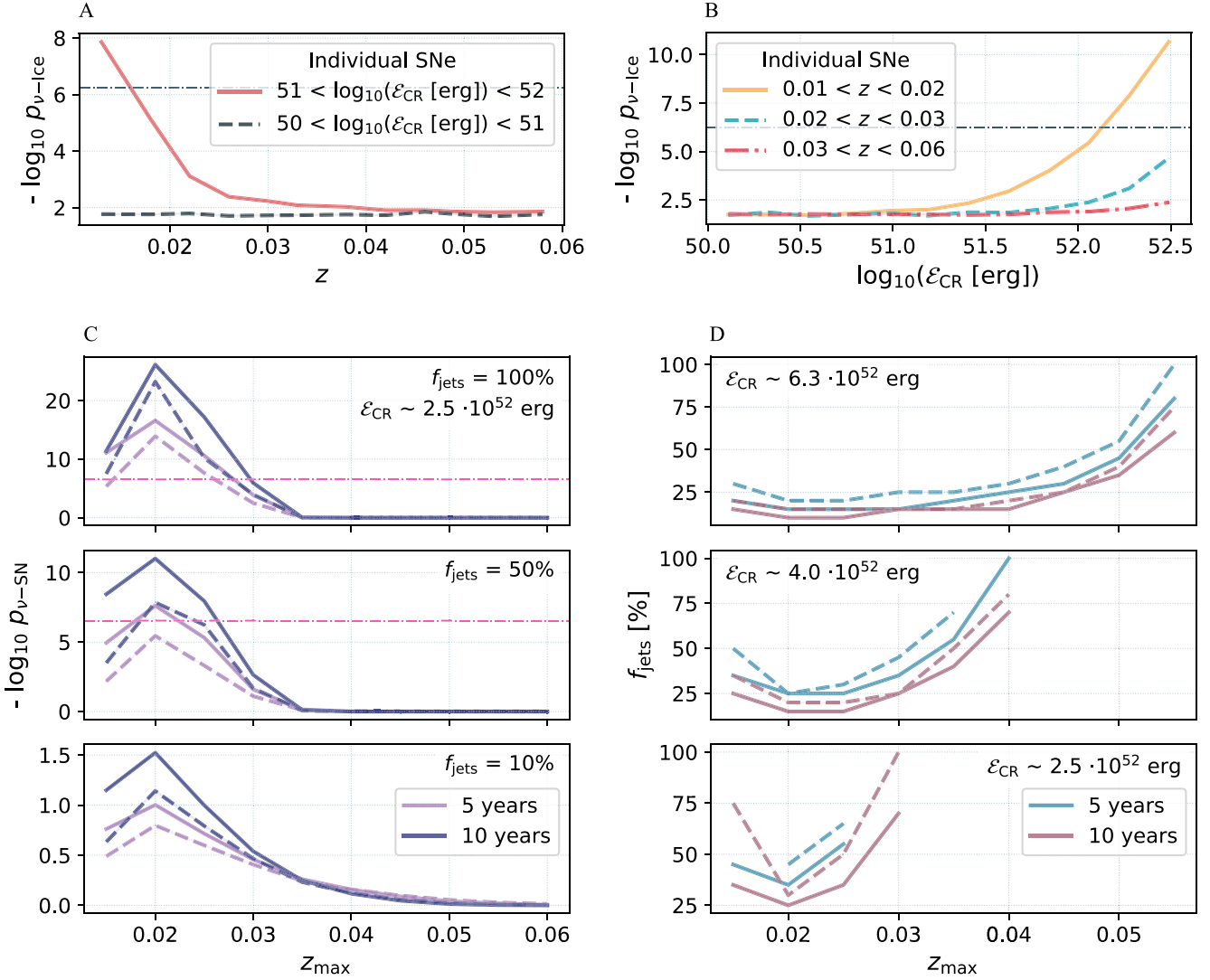
Figure 3. Results of the correlation analysis between neutrinos and CC-SNe. (A)–(B): The top 1% (most significant) of the distribution of p-values, $p_{\nu-\mathrm{Ice}}$, for neutrinos associated with individual SNe (not corrected for trials), (A) as a function of the redshift, z; and (B) as a function of the energy of a source deposited into cosmic rays, $\mathcal{E}_{\mathrm{CR}}$. The dashed–dotted horizontal lines highlight the value corresponding to the pre-trials 5σ detection significance of a single event. The combined values of $p_{\nu-\mathrm{Ice}}$ from different samples of SNe serve as the basis for the stacking analysis. (C): Post-trials p-values, $p_{\nu-\mathrm{SN}}$, for the stacked sample of neutrinos, as a function of the maximal redshift of observed SNe included in the analysis, z_{max}. Dashed and full lines, respectively, correspond to $\sigma_{\mathrm{SN}}^{\mathrm{min}} = 5\sigma$ and 3σ for the optical detection threshold for individual SNe. The dashed–dotted horizontal line in the top panel highlights the value corresponding to 5σ significance for the stacked search. As indicated, we consider 5 and 10 yr LSST surveys, where it is assumed that all SNe have $\mathcal{E}_{\mathrm{CR}} = 2.5 \times 10^{52}$ erg. The three panels illustrate the results for different values of f_{jets}, the fraction of SNe for which neutrino emission is observable. In general, $f_{\mathrm{jets}} \ll 1$ is expected for beamed emission, and $f_{\mathrm{jets}} \sim 1$ for shock breakouts. (D): The sensitivity of the stacking analysis, expressed as the minimal value of f_{jets}, for which $\geqslant 5\sigma$ detection is achievable. Different values of $\mathcal{E}_{\mathrm{CR}}$ are compared in the three panels, where dashed and full lines, respectively, correspond to $\sigma_{\mathrm{SN}}^{\mathrm{min}} = 5\sigma$ and 3σ.

–(B). This motivates performing a stacking analysis, combining observations from many events.

We simulate a neutrino stacking analysis, intended to identify an accumulated neutrino over-density in spatio-temporal coincidence with CC-SNe (Senno et al. 2018). (This should be distinguished from the more direct cross-correlation approach, which involves optical follow-up of specific neutrino events; Morgan et al. 2019.) The purpose of the RNN is to provide a background model for the joint neutrino and optical observation. Optical transients are simulated according to projected observations with the Wide-Fast-Deep survey of the upcoming Vera C. Rubin Observatory (previously referred to as the Large Synoptic Survey Telescope, LSST; LSST Science Collaboration et al. 2009). These SNe are only used to define time-intervals, during which neutrino flares may occur.

Neutrino signals, based on our IceCube sample, are integrated over these intervals.

The inputs to the RNN are of two types. For a given time step, the first consists of the IceCube neutrino densities (as for the previous example). The second type consists of LSST signal-to-noise metrics (Zackay & Ofek 2017) in five optical bands (see Appendix A.3). The zenith of observation is also added as an auxiliary parameter. We thus have $\eta = 10$ inputs per time step in total. SNe light curves evolve over days–weeks. We therefore construct $\tau_{\mathrm{RNN}} = 20$ days data sequences, with the first $\tau_{\mathrm{enc}} = 10$ days representing the background, and the next $\tau_{\mathrm{dec}} = 10$ the search period. In total the RNN receives $10 \times 20 = 200$ inputs.

We begin by creating background samples. These represent periods of joint neutrino and optical observations, in which no

transient exists in either messenger. The neutrino data are derived by the scrambling procedure described above. The optical data are chosen based on the true peak time of simulated SNe. The background samples are used to train the RNN. For a particular sky position, we then derive individual test statistics for source detection with each messenger, which are calibrated to p-values as a function of zenith.

In the next step of the analysis, we simulate stacked neutrino signal samples. For the source model, we assume that optically detected SNe are in fact GRBs, and that neutrinos are emitted during the short prompt phase of explosions. The fluence of muon neutrinos is modeled as $\mathcal{F}_{\nu_\mu} \propto \mathcal{E}_{CR}/D_L^2$, where $\mathcal{E}_{CR}$ is the energy deposited in CRs, and D_L is the luminosity distance to the source (Waxman & Bahcall 1999; Senno et al. 2018). We consider different values for f_{jets}, the fraction of SNe for which the neutrino emission is directed toward the Earth. Depending on the physical model, the emission may e.g., be collimated (relativistic jets; $f_{jets} \ll 1$), or quasi-spherical (shock break-outs; $f_{jets} \sim 1$). We also account for SN misclassifications, in which cases neutrino signals are not injected. For this simplified example, we assume that CC-SNe may be identified from their photometric light curves with a 10% fake-rate (Muthukrishna et al. 2019).

The TS of the stacked signal is computed by averaging the individual neutrino detection significance of all events passing a selection threshold, σ_{SN}^{min}. This threshold is defined by the detection significance of the corresponding optical SN. In order to derive the final stacked detection significance, we simulate the background hypothesis; we impose the nominal optical detection procedure, but do not inject any signals into the neutrino data. We generate $\sim 10^7$ such realizations of a full LSST survey, accounting for random coincidence between observables.

The results of the analysis are shown in Figure 3. We compare different values of f_{jets} and $\mathcal{E}_{CR}$, different optical detection thresholds ($\sigma_{SN}^{min} = 3\sigma, 5\sigma$), and different durations of LSST surveys (5, 10 yr). The search would be sensitive in the case of high f_{jets} and $\mathcal{E}_{CR}$ values, in accordance with Senno et al. (2018). Using our method, it would be possible to e.g., significantly constrain shock breakout scenarios. If no signal is detected, the sensitivity curves could be used to derive physical limits on neutrino emission.

For the neutrino point-source search, the RNN was used to derive joint detection probabilities for two experiments. Similarly in this case, we avoid having to impose a specific relationship between two messengers. In addition, our approach provides a statistical framework to optimize searches, and to derive limits on non-detections. In this illustrative example, we use the RNN to tune the redshift range and to relax the detection threshold of individual events. This circumvents the need to a priori refine the selection criteria of SNe. While we do not explicitly incorporate SN classification or redshift estimation as part of our pipeline, such extensions are feasible (Muthukrishna et al. 2019). They are planned for future work, paving the way for real-time applications.

3.3. Serendipitous Discovery of GRBs

For SNe engines of sufficient power, relativistic jets manage to break out of the progenitor. Depending on their inclination, these may be observed as long GRBs. Observations of GRBs at high energies are very interesting, e.g., for the study of acceleration mechanisms in relativistic shocks. In many cases, γ-ray emission extending up to GeV energies has been detected (Ackermann et al. 2013). Recently, emission of up to hundreds of TeV has been detected for the first time with ground-based Cherenkov Telescopes, which were following up alerts from other instruments (Abdalla et al. 2019; Acciari et al. 2019).

For the current example, we simulate an uninformed search scenario for the upcoming Cherenkov Telescope Array (CTA; Acharya et al. 2017). We focus on low-luminosity bursts (LL-GRBs), a sub-class of long GRBs (Virgili et al. 2009), which have been connected to SNe having mildly relativistic outflows (Cano et al. 2017), and are potential sources of UHECRs and neutrinos (Murase & Ioka 2013; Boncioli et al. 2019). The true properties of LL-GRBs are not well constrained by observations. Our sample therefore encompasses a wide range of spectral and temporal parameters. As such, it may be used to illustrate the detection prospects of such bursts, rather than to make precise predictions on observable rates. We correspondingly choose simple models, simulating the spectra of LL-GRBs as either simple power laws (PLs) in energy and time, or as PLs having exponential cutoffs.

We simulate observations for the Northern array of CTA using the *ctools* analysis framework (Knödlseder et al. 2016). The simulations consists of γ-ray-like events. These correspond to true γ-rays, as well as to cosmic rays and electrons passing the nominal CTA selection and classification cuts. The inputs to the RNN are event counts, n_γ, within $\eta = 4$ logarithmically spaced energy bins between 30 and 200 GeV. They are integrated within circular regions of interest (RoIs), over 1 s periods (in order to probe the prompt emission phase of long GRBs). We construct the RNN to represent $\tau_{RNN} = 25$ s of data. The first $\tau_{enc} = 20$ steps correspond to the background, and the final $\tau_{dec} = 5$ steps to the putative signal. In total, the RNN receives $4 \times 25 = 100$ inputs.

For anomaly detection, the RNN is trained to predict γ-ray-like counts in the absence of signals. In addition, we employ a classification approach. Here, the network is trained with examples of both background and signal events, where signals are injected over the $\tau_{dec} = 5$ step interval.

Figure 4(A) shows the real-time discovery potential of the RNN. As expected, bursts with lower redshifts and higher luminosities are more likely to be detected. In general, a large fraction of the parameter space is accessible. In Figure 4(B) we compare our algorithm with the likelihood-based method for source detection of *ctools*. The RNN performs similarly or better. For anomaly detection, this is achieved without relying on instrument response functions.

Contrary to anomaly detection, classification requires some assumption on sources as part of training. However, the performance of the method is shown to be robust to this constraint. In the current example, we simulate the intrinsic spectra of LL-GRBs as exponentially cutoff PLs. The classifier is trained exclusively with simple PL examples. However, nonetheless it is able to generalize, and outperforms the likelihood approach by achieving higher detection rates. We emphasize that for both RNN configurations, training will primarily utilize real CTA data, once it becomes available, rather than simulations. (For instance, for classification a hybrid approach is possible, where simulated signals are injected into background sequences from real data.) Correspondingly, the dependence on instrumental modeling is minimized using our framework.
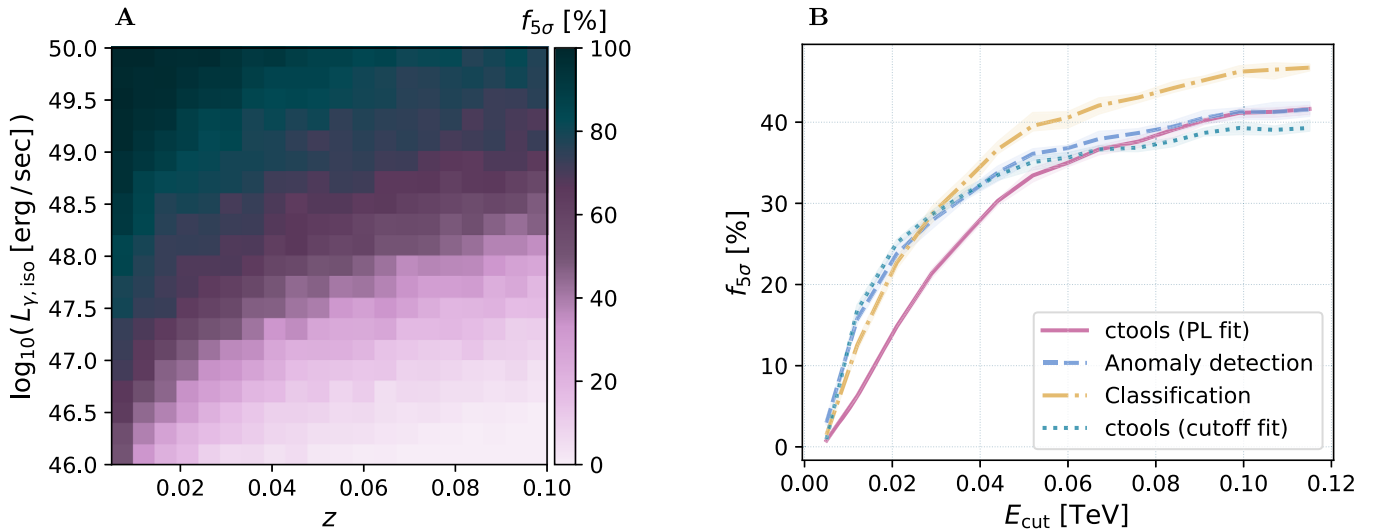
Figure 4. Results of the serendipitous GRB search analysis. (A): The probability to detect a GRB with at least 5σ significance, $f_{5\sigma}$, derived with anomaly detection, after accounting for trials. The simulated sample includes different combinations of redshift, z, and isotropic equivalent luminosity, $L_{\gamma,\mathrm{iso}}$, spanning the expected properties of LL-GRBs, assuming PL spectral models (see Equation (A12)). (B): Dependence of $f_{5\sigma}$ on the cutoff energy, E_{cut}, for bursts simulated as exponentially cutoff PLs (see Equation (A14)). The shaded regions correspond to 1σ statistical uncertainties on the values of $f_{5\sigma}$. Two alternative models are assumed for detection with the likelihood method of *ctools*, an exponentially cutoff PL, and a simple PL, as indicated. While the classification method was exclusively trained using simple PL examples, it effectively generalizes and identifies sources having cutoff spectra.

The advantage of the RNN is particularly evident when only PL models are used as part of the *ctools* likelihood fit. This represents a realistic strategy for blind searches, for which simple assumptions are generally made. Both our algorithms are comparatively agnostic to the properties of sources. They therefore enable real-time detection of a wider range of transients compared to standard techniques. This illustrates the merits of the methodology for unbiased searches, and the potential for unexpected discoveries.

4. Discussion

Our method is optimized to minimize the bias on (real-time) detection of transients. It is distinguished from previous works, which have either incorporated explicit examples of signal events (mostly for classification), or have modeled the characteristics of backgrounds. Conversely, we use a data-driven strategy, which is less susceptible to the pitfalls of unrepresentative training data sets. The predictions of the network are relatively robust against theoretical and systematic uncertainties on sources and instruments. Considering the simple architecture, searches may be conducted on different timescales. This may be done by modifying the value of τ_{RNN}; by relating one or several RNN steps to other temporal intervals; or by constructing (multiple) tests statistics that span different scales. In all cases the calibration phase facilitates correct derivation of the final p-values for detection, accounting for trials. As illustrated for the SNe correlation study, the framework may also be used to derive limits from non-detections, though this requires instrument modeling.

Our approach is relatively generalizable, enabling model-independent combination of observables. In principle, the framework facilitates sophisticated schemes of data fusion, where different data formats may be integrated consistently. However, in our nominal approach, this is not necessary. Subsequently, simple RNN architectures may be used. This leads to robust predictions that do not depend on extensive optimization of hyper-parameters.

In a realistic scenario of real-time searches, various systematic effects may initially be detected as transients. Experience with a particular instrument is necessary in order to suppress these spurious signals. Different strategies may be employed to this effect. For example:

1. incorporating more sophisticated architectures, such as Bayesian networks (Shen et al. 2019), which may improve the calibration phase;
2. using data that include these systematics for training (e.g., the scrambling procedure for our neutrino point-source analysis);
3. injecting known glitches as background events during training;
4. correlating multiple observables that are susceptible to different systematics (e.g., real-time combination of γ-rays and optical data), or cross-correlating RoIs across the field of view;
5. adding informative auxiliary observables to the RNN;
6. performing selection cuts pre-processing (e.g., image quality cuts);
7. performing filtering post-processing (e.g., incorporating an additional network that would be trained either to classify systematic-induced events, or to perform anomaly detection with regards to known glitches; George et al. 2018);
8. performing selection cuts post-processing, based on source modeling (e.g., general constraints on timescales and energetics).

In any case, even after accounting for systematics, true detections could correspond to different physical scenarios. In order to correctly characterize transients, multiwavelength/multimessenger follow-up will be essential. One of the primary applications of our method will be the effective identification of candidates for follow-up.

We would like to thank the following people for numerous useful discussions: D. Biehl, D. Boncioli, Z. Bosnjak, A.

Franckowiak, O. Gueta, T. Hassan, D. Horan, M. Krause, F. Longo, G. Maier, M. Nievas Rosillo, I. Oya, A. Palladino, E. Pueschel, R. R. Prado, L. Rauch, and W. Winter. This Letter has also gone through internal review by the CTA Consortium.

We use the publicly available, all-sky point-source IceCube data for years 2011–2012 (Aartsen et al. 2017; see https://icecube.wisc.edu/science/data/PS-3years). We also use publicly available data from the ANTARES experiment for the same time period (Adrián-Martínez et al. 2014; see http://antares.in2p3.fr/publicdata2012.html.) We utilize simulations from the Photometric LSST Astronomical Time Series Classification Challenge (PLAsTiCC), an open data challenge to classify simulated astronomical time-series data (Kessler et al. 2019; see https://plasticc.org; PLASTICC Team & PLASTICC Modelers 2019). We also use *SNCosmo*, an open source library for supernova cosmology analysis (see Barbary et al. 2016). This research makes use of *ctools* (Knödlseder et al. 2016), a community-developed analysis package for Imaging Air Cherenkov Telescope data. *ctools* is based on *GammaLib*, a community-developed toolbox for the high-level analysis of astronomical gamma-ray data (see http://cta.irap.omp.eu/ctools/; http://cta.irap.omp.eu/gammalib/). We also use CTA-IRFs, provided by the CTA Consortium and Observatory (version *prod3b-v1*; see http://www.cta-observatory.org/science/cta-performance/). This research makes use of the open source software, *tensorflow* (see Abadi et al. 2015).

Appendix
Statistical Analysis

A.1. Network Architecture and Inference Pipeline

Our software pipeline and chosen RNN architecture are shown in Figure 1. The RNN is implemented using *tensorflow* (Abadi et al. 2015). It may be decomposed into two elements, an *encoder* and a *decoder* (Cho et al. 2014). The RNN accepts $\tau_{RNN} = \tau_{enc} + \tau_{dec}$ time steps as input. The different steps are implemented as LSTM cells. A cell is composed of a pair of LSTM layers, respectively, comprising 128 and 64 hidden units (the set of parameters tuned during training). Each step receives a collection of (analysis-specific) η inputs. The inputs are independently normalized, such that their nominal range of values for the background training sample is mapped to the interval, [0, 1].

In our nominal approach we employ an anomaly detection technique. We utilize sequences of input data, $S(\tau_{RNN}, \eta)$, which for training correspond to the response of an instrument in the absence of signal events. The RNN is used to predict the expected background of the experiment, $B(\tau_{dec}, \eta)$, having the same data structure as $S(\tau_{RNN}, \eta)$. Transients are then detected as significant divergences from these predictions. Training involves minimizing a *loss function*, which is defined as the *mean squared error* between B and S for each set of $\tau_{RNN}-\eta$.

For classification, an external layer is appended to the output decoder. The latter maps $B(\tau_{dec}, \eta)$ into *logits*, $\zeta(\tau_{dec})$. Each of these is a proxy for the RNN probability density function (PDF) for data of a given step to belong to the signal class (Goodfellow et al. 2016). Training proceeds by minimizing a *cross-entropy* loss function for the different steps, where the final probability is taken as the average, $\zeta_{dec} = \langle \zeta \rangle_{\tau_{dec}}$. The corresponding TS is based on the ratio between the background and signal PDFs for a given value of ζ_{dec} (Cranmer et al. 2015). As for anomaly detection, calibration of test statistics into

p-values is performed once following the training stage, using simulations.

In the following we detail the data reduction, source modeling, and definition of test statistics used for the example analyses presented in this study. A short summary is given in Figure A1.

A.2. Neutrino Point-source Search

We search for clustering in two publicly available all-sky (Decl. $\in \{-85°, 85°\}$) neutrino samples. The first comprises IceCube event lists of track-like muon neutrino candidates, taken between 2011 and 2012 (MJD 55694–56415), with the IC-86I detector configuration (Aartsen et al. 2017). The second is of ANTARES muon neutrinos, observed within the same time period (Adrián-Martínez et al. 2014). Both samples include time-stamps, reconstructed event directions, and angular uncertainties. In the case of IceCube, energy-proxy metrics and tabulated instrument response functions (IRFs) estimates are available as well. These public data sets are limited in scope. It is therefore not feasible to, e.g., reliably correct for instrumental dead-time, or to estimate the efficiency of event reconstruction and selection cuts. An advantage of our methodology, is that such explicit corrections are not always necessary, so long as unbiased self-consistent observables can be derived.

The nominal inputs to our RNN are based on *neutrino event density* metrics with respect to a given location on the sky, $\boldsymbol{x}$. We define the density for a given neutrino as

$$s_\nu(\boldsymbol{x}) = \min\left(\frac{\delta_\nu}{\Delta_{\nu,\boldsymbol{x}}}, \ 1\right). \quad (A1)$$

Here δ_ν is the angular uncertainty of the event, and $\Delta_{\nu,\boldsymbol{x}}$ is the angular distance between the position of the neutrino and $\boldsymbol{x}$. We note that this definition of observables, while well motivated, is not unique. However, our framework is designed to provide self-consistent detection probabilities, given such informative input parameters.

We integrate the event densities over 24 hr time periods, which effectively avoids dependence of the event rates on R.A. (Aartsen et al. 2015). For the IceCube sample, the data are split into four logarithmically spaced bins in the energy proxy, E_ν, between 10 GeV and 8 PeV. For ANTARES the summation is inclusive in energy. We therefore have five density metrics as RNN inputs,

$$S_{\nu-Ice}^{\epsilon}(\boldsymbol{x}) = \sum_{1 \, day, \, \Delta E_\nu^\epsilon} s_\nu(\boldsymbol{x}),$$
$$\text{and } S_{\nu-ANT}(\boldsymbol{x}) = \sum_{1 \, day} s_\nu(\boldsymbol{x}), \quad (A2)$$

with energy bins, $\epsilon \in \{1, 2, 3, 4\}$. The response of the IceCube and ANTARES detectors depends on the zenith of the observed event, which therefore is added as an auxiliary input to the RNN. We nominally use the IceCube events to determine the zenith for a particular sequence. Considering the time of arrival and the location of the observatories, we calculate the corresponding zenith for the same putative source with ANTARES. In cases where an IceCube source is not within the field of view of ANTARES, we use background data
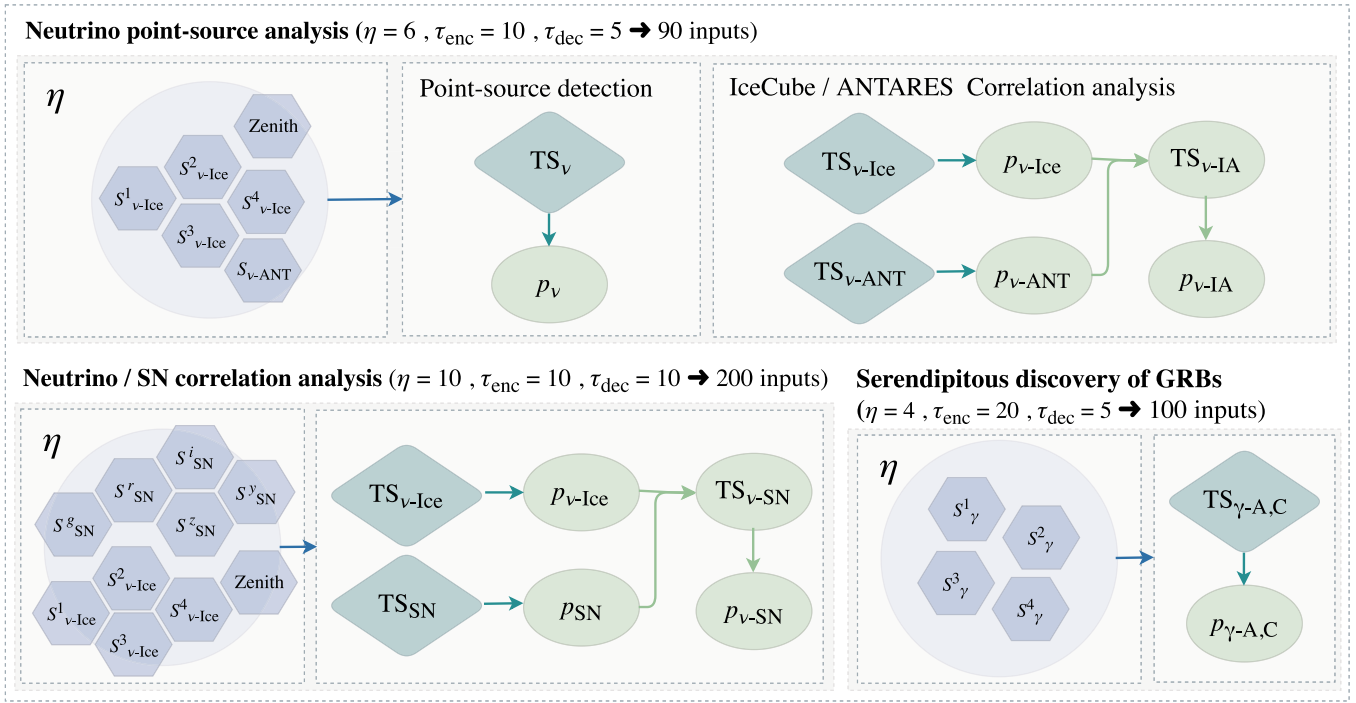
Figure A1. Summary of the specific RNN and pipeline configurations used for the three analysis examples. The RNN architectures are defined by the number of encoder and decoder steps, τ_{enc} and τ_{dec}, and by the type and number of inputs per RNN cell, η (blue hexagons), as indicated. The calibration phases of the pipeline are defined by the particular choices of TS and their corresponding p-values for each analysis. The specific notation for each of the analyses is defined in the text.

instead. In total, we have $\eta = 6$ inputs per time step. The collection of inputs is derived for $\tau_{RNN} = 15$ days periods. The first $\tau_{enc} = 10$ days are assumed to only contain background events. Transient signals are searched for within the next $\tau_{dec} = 5$ days interval. In total, the RNN receives $6 \times 15 = 90$ inputs.

We construct a background sample of 10^4 τ_{RNN} sequences, where potential transient signals are removed by scrambling the events in R.A. and time of detection (Aartsen et al. 2015). The RNN is trained to predict the neutrino event densities in each of the five days being probed, for a particular $\boldsymbol{x}$. The outputs of the network of the five event densities for a given day are denoted by $B_{\nu-\mathrm{Ice}}^{\epsilon(1-4)}$ and $B_{\nu-\mathrm{ANT}}$ for the two experiments. We use these to calculate a combined test statistic,

$$\mathrm{TS}_\nu(\boldsymbol{x}) = \mathrm{TS}_{\nu-\mathrm{Ice}}(\boldsymbol{x}) + \mathrm{TS}_{\nu-\mathrm{ANT}}(\boldsymbol{x}), \qquad (\mathrm{A3})$$

where

$$\mathrm{TS}_{\nu-\mathrm{Ice}}(\boldsymbol{x}) = -\sum_{5\,\mathrm{days},\,\epsilon} \log_{10}\left(\frac{B_{\nu-\mathrm{Ice}}^\epsilon}{S_{\nu-\mathrm{Ice}}^\epsilon}\right),$$

$$\mathrm{and}\quad \mathrm{TS}_{\nu-\mathrm{ANT}}(\boldsymbol{x}) = -\sum_{5\,\mathrm{days}} \log_{10}\left(\frac{B_{\nu-\mathrm{ANT}}}{S_{\nu-\mathrm{ANT}}}\right). \qquad (\mathrm{A4})$$

High values of TS_ν correspond to large discrepancies between (background) predictions and the corresponding true data. TS-values are calibrated to p-values, p_ν, as a function of the auxiliary parameter (the zenith, binned into 90 intervals of $2°$ width).

Note that we take a generalizable approach, where the two data sets are combined on equal footing. More complicated schemes are also possible, e.g., by introducing relative weights in the definition of TS_ν. Similarly, different timescales may be

probed, either by modifying the definition of τ_{RNN}, or that of TS_ν.

We also use the outputs of the RNN to perform a correlation test between the IceCube and ANTARES events over the entire period of the data set for each spatial position. For this purpose, we consider $\mathrm{TS}_{\nu-\mathrm{Ice}}$ and $\mathrm{TS}_{\nu-\mathrm{ANT}}$ individually. As part of the calibration stage, we derive the relation between TS-values and p-values for each statistic independently. This is done using simulations (as for TS_ν above).

Our metrics for the correlation analysis are Pearson coefficients, based on the p-values of the two samples, $p_{\nu-\mathrm{Ice}}$ and $p_{\nu-\mathrm{ANT}}$. Explicitly, we define,

$$\mathrm{TS}_{\nu-\mathrm{IA}}(\boldsymbol{x}) = \frac{\sum_t(\rho_{\mathrm{Ice}}^t - \langle\rho_{\mathrm{Ice}}\rangle)(\rho_{\mathrm{ANT}}^t - \langle\rho_{\mathrm{ANT}}\rangle)}{\sqrt{\sum_t(\rho_{\mathrm{Ice}}^t - \langle\rho_{\mathrm{Ice}}\rangle)^2}\sqrt{\sum_t(\rho_{\mathrm{ANT}}^t - \langle\rho_{\mathrm{ANT}}\rangle)^2}}. \tag{A5}$$

Here $\rho_{\mathrm{Ice}}^t(\boldsymbol{x}) \equiv -\log_{10}p_{\nu-\mathrm{Ice}}$, and $\rho_{\mathrm{ANT}}^t(\boldsymbol{x}) \equiv -\log_{10}p_{\nu-\mathrm{ANT}}$. The summation is over the entire period of the data set, with t representing a particular 5 days interval for a given sample. We proceed to derive the p-value for detection of a correlation signal, $p_{\nu-\mathrm{IA}}$. For this purpose, we independently scramble ρ_{Ice}^t and ρ_{ANT}^t in time and R.A. This procedure is used to create multiple realizations of $\mathrm{TS}_{\nu-\mathrm{IA}}(\boldsymbol{x})$, for which the two samples are uncorrelated. Using these background distributions, we account for spurious correlations, as well as for the number of spatial and temporal grid points.

A.3. Correlation Analysis between Neutrinos and CC-SNe

We simulate a neutrino stacking analysis. The general strategy is to identify an accumulated neutrino over-density (based on $\mathrm{TS}_{\nu-\mathrm{Ice}}$) in spatio-temporal coincidence with CC-

SNe (Senno et al. 2018). This should be distinguished from the more direct cross-correlation approach, which involves optical follow-up of specific neutrino events (Morgan et al. 2019). A direct analysis is in principle preferred, as it enables the detailed study of a particular source (Aartsen et al. 2018). However, it has two main disadvantages for the case of CC-SNe. First, one must have high confidence that a given neutrino is astrophysical, which generally constrains events to very-high energies. Additionally, the sensitivity is limited by the irreducible contamination of unassociated SNe within the uncertainty region of the neutrino. Given the weak nature of neutrino sources and the large number of potential counterparts, indirect population studies become competitive.

SNe for our stacking analysis are simulated according to projected observations with the upcoming LSST Wide-Fast-Deep survey (LSST Science Collaboration et al. 2009). We utilize the public PLAsTiCC data set, which was created as part of an open data challenge in preparation for LSST (Kessler et al. 2019). The data set represents the projected cadence and observing constraints for LSST. For instance, it includes a prototype scheduler for science program optimization; realistic environmental conditions, such as weather and seeing; maintenance downtime; and instrumental artifacts.

The inputs to the RNN are of two types. For a given time step, the first consists of the IceCube neutrino densities, $S_{\nu-\text{Ice}}^{\epsilon}(\boldsymbol{x})$. The second type consists of LSST signal-to-noise metrics in several optical bands, $b \in \{g, r, i, z, y\}$. These are defined as

$$S_{\text{SN}}^{b}(\boldsymbol{x}) = \frac{m_b}{\delta m_b}, \qquad (A6)$$

where m_b and δm_b, respectively, stand for an observed magnitude in band, b, and the corresponding uncertainty. As above, the zenith of observation is added as an auxiliary parameter. We thus have $\eta = 10$ inputs per time step in total. We construct $\tau_{\text{RNN}} = 20$ days data sequences, with the first $\tau_{\text{enc}} = 10$ days representing the background, and the next $\tau_{\text{dec}} = 10$ days the search period. In total the RNN receives $10 \times 20 = 200$ inputs.

We begin by creating background samples. These represent periods of joint neutrino and optical observations, in which no transient exists in either messenger. The scrambling procedure described for the neutrino point-source analysis is used to derive $S_{\nu-\text{Ice}}$. We extract S_{SN} from simulation periods that lack transient signals, based on the true peak time of SNe light curves. The simulated LSST observations include gaps in observations in some/all bands. We account for missing inputs by randomly interleaving background optical data in such gaps. The background samples are used to train the RNN, using 10^4 τ_{RNN} sequences.

For a particular sky position, we derive individual test statistics for source detection with each messenger. For neutrinos, we use $\text{TS}_{\nu-\text{Ice}}(\boldsymbol{x})$ (here defined over $\tau_{\text{dec}} = 10$ days intervals). For optical SNe, we have

$$\text{TS}_{\text{SN}}(\boldsymbol{x}) = - \sum_{10 \text{ days}, b} \log_{10}\left(\frac{B_{\text{SN}}^{b}}{S_{\text{SN}}^{b}}\right), \qquad (A7)$$

where B_{SN}^{b} stands for the output of the network in a given band, b. In addition to high values of TS_{SN}, we impose a constraint on optical detections, that a SN is observed over at least three nights in different bands. The correspondence between $\text{TS}_{\nu-\text{Ice}}$

and TS_{SN} and their respective p-values, $p_{\nu-\text{Ice}}$ and p_{SN}, is derived from simulations as a function of zenith.

In the next step of the analysis, we simulate signal samples. The total number of simulated SNe up to a redshift of 0.06 is $\sim 2000 \text{ yr}^{-1}$, following the star formation rate (Bernstein et al. 2012). Of these, about 57% are detected with our RNN with at least 3σ significance, and 42% with 5σ. We assume that the optically detected SNe are in fact unobserved GRBs. The SNe are generally characterized by the peak time of their light curves, which occurs on average 13 days after the putative GRB (Cano et al. 2017). As the exact time of the emission is uncertain, we randomly generate it as a Poisson process. We consider different values for f_{jets}, the fraction of SNe for which the neutrino emission is directed toward the Earth. Depending on the physical model, the emission may e.g., be collimated (relativistic jets; $f_{\text{jets}} \ll 1$), or quasi-spherical (shock break-outs; $f_{\text{jets}} \sim 1$). We also account for misclassifications of CC-SNe, by introducing a 10% fake-rate, for which neutrino signals are not injected (Muthukrishna et al. 2019).

A neutrino flare is assumed to occur during the short prompt phase of the explosion. We relate the simulated flux of muon neutrinos to the energy deposited by SNe in CRs, $\mathcal{E}_{\text{CR}}$. We use a flat Λ CDM cosmology, with $\Omega_{\text{m}} = 0.3$ and the Hubble constant, $H_0 = 70 \text{ km s}^{-1} \text{Mpc}^{-1}$. The luminosity distance of optical SNe, D_{L}, is derived from their redshift. The fluence of simulated muon neutrinos is then given by $\mathcal{F}_{\nu_{\mu}} \propto \mathcal{E}_{\text{CR}}/D_{\text{L}}^2$ (Waxman & Bahcall 1999; Senno et al. 2018), assuming a PL spectrum for the parent CRs with a spectral index of -2, and a flare duration of ~ 10–100 s. We derive the expected observed neutrino signal from the spectrum using the IRFs of IceCube (Senno et al. 2018).

The test statistic of the stacked signal is defined as

$$\text{TS}_{\nu-\text{SN}}(\tau_{\text{LSST}}, \sigma_{\text{SN}}^{\min}) = \langle -\log_{10} p_{\nu-\text{Ice}} \rangle, \qquad (A8)$$

where the average is over all SNe having been detected during a survey of duration, τ_{LSST} (either 5 or 10 yr). In this context, detections are defined as events having an optical detection significance, σ_{SN}, higher than a given threshold, $\sigma_{\text{SN}}^{\min}$ (either 3σ or 5σ). In order to derive the stacked detection p-value, $p_{\nu-\text{SN}}$, we simulate the background hypothesis; we impose the nominal optical detection procedure, but do not inject any signals into the neutrino data. We generate $\sim 10^7$ such realizations of a full LSST survey of duration τ_{LSST}, accounting e.g., for misidentified SNe, and for random coincidence between observables (Senno et al. 2018).

A.4. Serendipitous Discovery of GRBs

Observations are simulated for the Northern array of CTA using *ctools* (Knödlseder et al. 2016), one of the proposed analysis frameworks for the observatory. The simulations produce *γ-ray-like* events. These correspond to true γ-rays, as well as to CRs and electrons, which pass all selection and classification cuts. *ctools* allows generation of background (CR, electron) sky-maps, as well as of background+signal observations, by inclusion of source spectral models. We use IRFs optimized for 30 minutes observations at zenith angles of $20°$. The RoI for the simulation is chosen as a circular area with a radius of $0.25°$. It is centered at the position of the putative source, and is displaced by $0.5°$ from the center of the camera.

The inputs to the RNN are γ-ray-like event counts, n_γ, within $\eta = 4$ logarithmically spaced energy bins, E_γ,
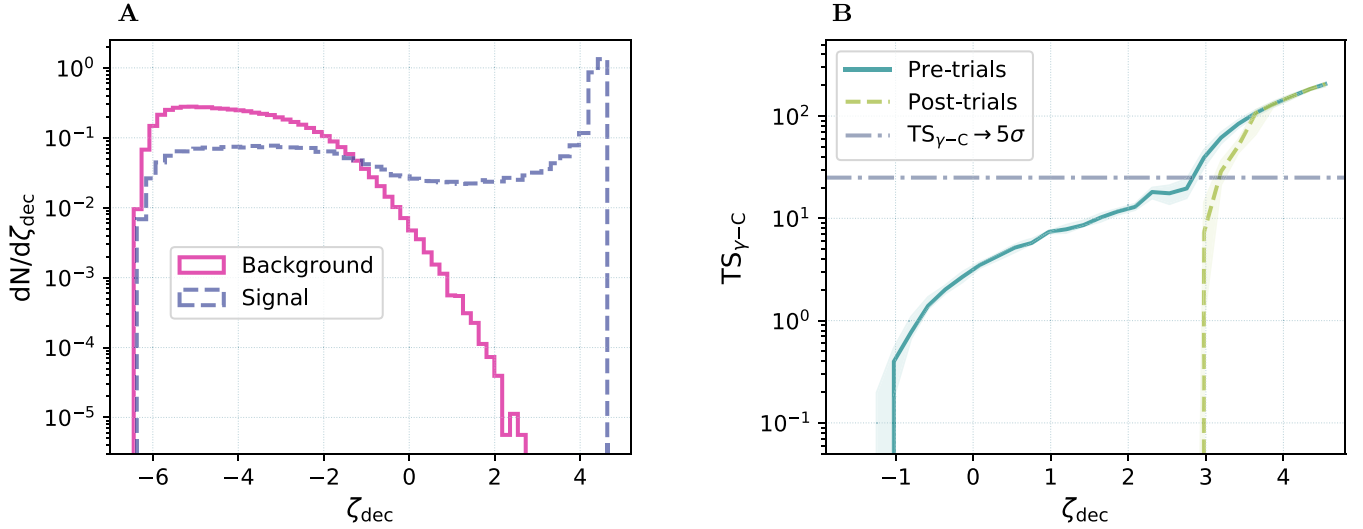
Figure A2. Calibration of the classification output of the RNN, ζ_{dec}, for the search for LL-GRBs. (A): The distributions of ζ_{dec} for the background and signal samples, from which the probability density functions, $P^{\mathrm{B}}_{\mathrm{CLAS}}$ and $P^{\mathrm{S}}_{\mathrm{CLAS}}$, are estimated. (B): Distributions of the test statistic for classification, $\mathrm{TS}_{\gamma-\mathrm{C}}$, as a function of ζ_{dec}, before and after the correction for trials, as indicated. The dashed–dotted horizontal line highlights the value of $\mathrm{TS}_{\gamma-\mathrm{C}}$ corresponding to a 5σ detection threshold.

between 30 and 200 GeV ($\epsilon \in \{1, 2, 3, 4\}$), integrated within 1 s intervals within the RoI,

$$S^{\epsilon}_{\gamma} = \sum_{1\,\mathrm{s},\,\Delta E^{\epsilon}_{\gamma}} n_{\gamma}. \qquad (A9)$$

For anomaly detection, we train the network with background-only *ctools* simulations. The RNN therefore predicts $B^{\epsilon\,(1-4)}_{\gamma}$, the background event counts in each of the energy bins for each of the steps. We construct the RNN to represent $\tau_{\mathrm{RNN}} = 25$ s of data. The first $\tau_{\mathrm{enc}} = 20$ steps correspond to the background, and the final $\tau_{\mathrm{dec}} = 5$ steps to the putative signal. In total, the RNN receives $4 \times 25 = 100$ inputs. We train the network using 10^4 background sequences.

Our framework allows for explicit statistical models to be used, avoiding the need for calibration with background simulations. We illustrate this for the LL-GRB analysis with Poissonian models of the signal and background+signal hypotheses. These are defined as $P_{\mathrm{pois}}(k|\lambda) = e^{-\lambda}\lambda^{k}/k!$, given an observed number of events, k, and a rate, λ. We approximate λ from the event count corresponding to a given hypothesis, integrating over the final $\tau_{\mathrm{dec}} = 5$ steps of the RNN for a given energy bin,

$$\lambda^{\epsilon}_{S} = \sum_{5\,\mathrm{s}} S^{\epsilon}_{\gamma}, \quad \text{and} \quad \lambda^{\epsilon}_{B} = \sum_{5\,\mathrm{s}} B^{\epsilon}_{\gamma}. \qquad (A10)$$

We then define the test statistic,

$$\mathrm{TS}_{\gamma-\mathrm{A}} = -2\sum_{\epsilon}\ln\left(\frac{P_{\mathrm{pois}}(S^{\epsilon}_{\gamma}|\lambda^{\epsilon}_{B})}{P_{\mathrm{pois}}(S^{\epsilon}_{\gamma}|\lambda^{\epsilon}_{S})}\right). \qquad (A11)$$

We directly derive the pre-trials p-values for anomaly detection, $p_{\gamma-\mathrm{A}}$, using Wilks's theorem (Wilks 1938), which was validated by background simulations).

In addition to anomaly detection, we employ a classification approach. In this case, we require examples of signal events for training. To simulate the possible γ-ray signatures of LL-GRBs for CTA, we nominally model their spectra as simple PLs in energy, E, and time, t,

$$M_{\mathrm{PL}}(E, t) \propto E^{-\Gamma}t^{-T}. \qquad (A12)$$

The spectral and temporal indices, Γ and T, are parameters of the models. The true properties of these events are not well constrained, due to the scarcity of observations. We therefore consider a wide range of parameters. Motivated by bright bursts, detected at high energies with Fermi-LAT, we randomize over likely values, $1.9 < \Gamma < 2.7$ and $0.8 < T < 2$ (Ackermann et al. 2013). The flux normalization is randomly shifted with regard to these reference events in redshift and luminosity to the expected ranges for LL-GRBs (Acharya et al. 2017). The observed spectra are corrected for interactions with the extragalactic background light (EBL), which attenuates high-energy γ-rays (Franceschini et al. 2008).

The network is trained for classification using 10^4 background sequences and 10^4 signal sequences. Signals are injected over the $\tau_{\mathrm{dec}} = 5$ step interval, as part of the *ctools* simulation. The network is trained with a wide range of such models having different flux normalizations, corresponding to different signal-to-noise ratios with respect to the background. The signal models incorporate PL temporal decay. Correspondingly, late-time models are equivalent to early-time models with relatively lower flux normalization. The inclusive composition of the training sample therefore enhances generalization (time-invariance) of the RNN.

While the inputs to the RNN for classification are the same as for anomaly detection, the output in this case is a single number, ζ_{dec}. The latter takes low values for input sequences that do not include transient signals, and high values in the presence of high-significance signals. We use the training sample to derive $P^{\mathrm{B}}_{\mathrm{CLAS}}$ and $P^{\mathrm{S}}_{\mathrm{CLAS}}$, the approximated PDFs for the background and for the background+signal hypotheses (see Figure A2 (A)). Under the assumption that ζ_{dec} is monotonic with the ratio of PDFs (Cranmer et al. 2015), the classification test statistic can be defined as

$$\mathrm{TS}_{\gamma-\mathrm{C}}(\zeta_{\mathrm{dec}}) = -2\ln\left(\frac{P^{\mathrm{B}}_{\mathrm{CLAS}}(\zeta_{\mathrm{dec}})}{P^{\mathrm{B}}_{\mathrm{CLAS}}(\zeta_{\mathrm{dec}}) + P^{\mathrm{S}}_{\mathrm{CLAS}}(\zeta_{\mathrm{dec}})}\right). \qquad (A13)$$

We calibrate the relation between $\mathrm{TS}_{\gamma-\mathrm{C}}$ and the corresponding pre-trials $p_{\gamma-\mathrm{C}}$ with background and signal simulations, using

Wilks's theorem. The calibration is defined as a function of ζ_{dec}, as illustrated in Figure A2(B).

In a realistic scenario, an uninformed search would be performed for RoIs at different positions across the field of view of CTA. The search would be repeated multiple times, depending on the definition of τ_{dec} and on the amount of observing time. Correspondingly, the detection significance must be corrected for trials. We take, $p_{post} = 1 - (1 - p_{pre})^{n_{trial}}$, as the relation between pre- and post-trials probabilities. We use the number of trials, $n_{trial} \sim 3.6 \times 10^7$, corresponding to searches over 100 hr of observations, over the entire CTA field of view.

It is currently difficult to estimate the rate of false detections for CTA, as the observatory will explore a new regime of sensitivity (Acharya et al. 2017). Experience with real data will improve the fidelity of detections. As a first step, we performed systematic checks on the stability of the γ-ray-like event counts we use as input to the RNN. The counts are susceptible to fluctuation due to imperfect γ-ray reconstruction, and to uncertainties on the IRFs. In particular, energy dispersion below $\sim$50 GeV may result in migration between bins, and can change the energy threshold of the analysis. We studied these effects by comparing simulations where we vary the IRFs by their expected uncertainty (of up to 10%). We found that the propagated uncertainties on counts do not significantly affect our results. We also tested theoretical uncertainties on the effect of the EBL. Comparing different models of the EBL (Franceschini et al. 2008; Dominguez et al. 2011; Gilmore et al. 2012), we found negligible impact on the observed spectra of LL-GRBs. This is primarily due to the low redshift and energy regimes we consider in the current study.

We also simulate a separate signal sample, where LL-GRBs are modeled as PLs having exponential cutoffs,

$$M_{EC}(E, t) = M_{PL} \cdot \exp\left(-\frac{E}{E_{cut}}\right). \qquad (A14)$$

In this case, we scan a range of cutoff energies, $1 < E_{cut} < 120$ GeV. The cutoff models are not used for training. Rather, we utilize them to illustrate the robustness of our methods (see Figure 4). As discussed above, the networks perform well for different source types. For instance, they enable identification of GRBs having cutoff spectral models, despite only being trained with simple PL examples.

ORCID iDs

Iftach Sadeh ⬤ https://orcid.org/0000-0003-1387-8915

References

Aartsen, M. G., Abbasi, R., Abdou, Y., et al. 2013, Sci, 342, 1242856
Aartsen, M. G., Abraham, K., Ackermann, M., et al. 2017, ApJ, 835, 151
Aartsen, M. G., Ackermann, M., Adams, J., et al. 2015, ApJ, 807, 46
Aartsen, M. G., Ackermann, M., Adams, J., et al. 2018, Sci, 361, 147
Abadi, M., Agarwal, A., Barham, P., et al. 2015, tensorflow, https://www.tensorflow.org/
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, ApJL, 848, L13
Abdalla, H., Adam, R., Aharonian, F., et al. 2019, Natur, 575, 464
Acciari, V., Ansoldi, S., Antonelli, L. A., et al. 2019, Natur, 575, 455
Acharya, B. S., Agudo, I., Al Samarai, I., et al. 2017, arXiv:1709.07997
Ackermann, M., Ajello, M., Asano, K., et al. 2013, ApJS, 209, 11
Adrián-Martínez, S., Albert, A., André, M., et al. 2014, ApJL, 786, L5
Barbary, K., Biswas, R., Goldstein, D., et al. 2016, sncosmo/sncosmo v1.4.0, Zenodo, doi:10.5281/zenodo.168220
Bernstein, J. P., Kessler, R., Kuhlmann, S., et al. 2012, ApJ, 753, 152
Boncioli, D., Biehl, D., & Winter, W. 2019, ApJ, 872, 110
Cano, Z., Wang, S.-Q., Dai, Z.-G., & Wu, X.-F. 2017, AdAst, 2017, 8929054
Carleo, G., Cirac, I., Cranmer, K., et al. 2019, RvMP, 91, 045002
Cho, K., van Merrienboer, B., Gulcehre, C., et al. 2014, arXiv:1406.1078
Choma, N., Monti, F., Gerhardt, L., et al. 2018, arXiv:1809.06166
Cranmer, K., Pavez, J., & Louppe, G. 2015, arXiv:1506.02169
Denton, P. B., & Tamborra, I. 2018, ApJ, 855, 37
Dominguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556
Franceschini, A., Rodighiero, G., & Vaccari, M. 2008, A&A, 487, 837
Gebhard, T. D., Kilbertus, N., Harry, I., & Scholkopf, B. 2019, PhRvD, 100, 063015
George, D., & Huerta, E. A. 2018a, PhRvD, 97, 044039
George, D., & Huerta, E. A. 2018b, PhLB, 778, 64
George, D., Shen, H., & Huerta, E. A. 2018, PhRvD, 97, 101501
Gilmore, R. C., Somerville, R. S., Primack, J. R., & Dominguez, A. 2012, MNRAS, 422, 3189
Goodfellow, I., Bengio, Y., & Courville, A. 2016, Deep Learning (Cambridge, MA: MIT Press), http://www.deeplearningbook.org
Graves, A., & Jaitly, N. 2014, in Proc. 31st Int. Conf. on Machine Learning (ICML'14), Vol. 32, II–1764, http://dl.acm.org/citation.cfm?id=3044805.3045089
Huerta, E. A., Allen, G., Andreoni, I., et al. 2019, NatRP, 1, 600
Kashiyama, K., Murase, K., Horiuchi, S., Gao, S., & Mészáros, P. 2013, ApJL, 769, L6
Kessler, R., Narayan, G., Avelino, A., et al. 2019, PASP, 131, 094501
Khan, A., Huerta, E. A., Wang, S., et al. 2019, PhLB, 795, 248
Knödlseder, J., Mayer, M., Deil, C., et al. 2016, A&A, 593, A1
LeCun, Y., Bengio, Y., & Hinton, G. 2015, Natur, 521, 436 EP
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. 2015, in 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2015), 89, https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2015-56.pdf
Meszaros, P., Fox, D. B., Hanna, C., & Murase, K. 2019, NatRP, 1, 585
Morgan, R., Bechtol, K., & Kessler, R. 2019, ApJ, 883, 125
Murase, K., & Ioka, K. 2013, PhRvL, 111, 121102
Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019, PASP, 131, 118002
Pimentel, M. A., Clifton, D. A., Clifton, L., & Tarassenko, L. 2014, SigPr, 99, 215
PLASTICC Team & PLASTICC Modelers 2019, Unblinded Data for PLAsTiCC Classification Challenge v1, Zenodo, doi:10.5281/zenodo.2535746
Pruzhinskaya, M. V., Malanchev, K. L., Kornilov, M. V., et al. 2019, MNRAS, 489, 3591
Senno, N., Murase, K., & Meszaros, P. 2018, JCAP, 1801, 025
Shen, H., Huerta, E. A., Zhao, Z., Jennings, E., & Sharma, H. 2019, arXiv:1903.01998
Sutskever, I., Vinyals, O., & Le, Q. V. 2014, in Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14), Vol. 2 (Cambridge, MA: MIT Press), 3104, http://dl.acm.org/citation.cfm?id=2969033.2969173
Virgili, F., Liang, E., & Zhang, B. 2009, MNRAS, 392, 91
Waxman, E., & Bahcall, J. N. 1999, PhRvD, 59, 023002
Wei, W., & Huerta, E. A. 2020, PhLB, 800, 135081
Wilks, S. S. 1938, Ann. Math. Stat., 9, 60
Williamson, M., Modjaz, M., & Bianco, F. B. 2019, ApJL, 880, L22
Zackay, B., & Ofek, E. O. 2017, ApJ, 836, 187
Zevin, M., Coughlin, S., Bahaadini, S., et al. 2017, CQGra, 34, 064003