Contribution Prospectives 2020
Listening to the transient sky in the LSST era

Working group: GT04 – *Physique des astroparticules*

Principal (corresponding) author:
Name: Julien Peloton
Institution: LAL
Email: peloton@lal.in2p3.fr
Phone: +33 1 64 46 84 90

Co-authors and endorsers:
Emille E. O. Ishida, Anais Möller (LPC - Clermont)

Alexandre Boucaud, Alexis Coleiro, Cyril Lachaud (APC)
Ada Nebot Gomez-Moran (CDS, ObAS)
Damien Dornic, Dominique Fouchez (CPPM)
Chris Arnault, Jean-Eric Campagne, Julius Hrivnac, Jérémy Neveu, Marc Moniez, Nicolas Leroy, Stéphane Plaszczynski (LAL)
Dominique Boutigny, Marina Ricci, Didier Verkindt (LAPP)
Emmanuel Gangler (LPC - Clermont)
Johan Bregeon (LPSC)
Johann Cohen-Tanugi, Frédéric Piron (LUPM)
Astrid Lamberts (Observatoire de la Côte d'Azur)

Abstract: In the rapidly growing multi-messenger astronomy (MMA) era, the key for success in joining information from different probes is to connect different communities with efficient frameworks. A collaboration led by LSST-France, and composed of scientists from several french scientific groups involved in the study of the high energy transient sky, is proposing a broker for the Large Synoptic Survey Telescope (LSST). The main goals for the broker, called Fink, are to collect and store all LSST alerts, add value (crossmatch with existing catalogs and preliminary classification) and redistribute them to the community through stable and standard channels (e.g. VOEvents). This interdisciplinary collaboration is a unique opportunity to create a synergy that will enable an efficient cross-correlation and characterization of alerts, maximizing the scientific return on LSST and related experiments over the next decade.

1 Introduction

The next decade will see the arrival of a new generation of instruments with a large increase of data flux. The KM3NET neutrino detectors will be deployed in the next few years and will soon start observations with sensitivity comparable to the IceCube detector. The gravitational wave observatory network is growing fast with the extension of LIGO/Virgo, and the deployment of KAGRA next year. The electromagnetic side will see the arrival of new major actors like the Square Kilometre Array telescopes (SKA, radio bands), the Chinese-French Space Variable Objects Monitor multiwavelength observatory (SVOM, gamma and X rays, and optical ground segment), the Cherenkov Telescope Array (CTA, gamma rays), the Large Synoptic Survey Telescope (LSST, visible bands), and Euclid (visible and near infra-red bands). Among these, LSST will provide unique information in the visible wavelength range. Thanks to its large field of view, sensitivity and cadence, it will produce a full picture of the Southern sky every few days with unprecedented depth and detail.

The scientific potential of all these experiments will only be fully exploited when combined, as demonstrated by the event GW170817 where a follow-up of unprecedented magnitude, involving more than 70 space and ground based telescopes and thousands of people, took place to draw a full multi-messenger astronomy (MMA) picture. Thanks to this observing campaign it was possible to link the gravitational wave signal to its kilonovae counterpart enabling a rich study of the nature of the merger. **Time domain MMA has become a scientific priority and the only means to learn about physical phenomena such as cosmic rays, gravitational waves and other high energy phenomena.**

To fully exploit the available data, the main challenge is the structuration of communities beyond individual experiments. For instance, inter-project coordination is necessary to constrain the source position in particular when the resolution for the source position is very large with sky error regions spanning on hundreds of square degrees for gravitational wave instruments. In addition, the coordination must be done in a timely manner due to the transient nature of various phenomena, with time frames spanning from seconds to weeks. Asides from the spatial and time constraints, all communities are facing challenges to correlate the very large volume of alerts from different telescopes and detectors (real-time and post-processing), extract the sources, and eventually trigger a repointing of one or several telescopes or detectors of the network. Hence the key for success for joining information of different probes with low latency is to connect them with efficient frameworks (communication and processing) developed and shared across different communities. **IN2P3 with its long tradition of collaborations and large projects offer an ideal place for such a cross-community effort, and in the context of LSST the french community is proposing a broker to serve the need of LSST-France as well as the different french MMA actors.**

2 Fink: a broker delivering LSST alerts to the MMA community

The advent of large surveys will enable the exploration of the Universe with an unprecedented amount of data from MMA probes. **This abundance of data comes with challenges:** new methods must be developed to deal with the large amounts of data, crossmatches of sources between different messenger signals must be efficiently obtained and resources must be timely coordinated for follow-up. **Fink[1], is an IN2P3 initiative being developed to face these challenges focusing on the data avalanche expected by the biggest survey of this upcoming decade, LSST.**

Early 2019, the LSST Data Management has made a call for Letters of Intent to propose community brokers that will be responsible to handle the large volume of alerts issued by the

[1] `https://fink-broker.readthedocs.io/en/latest/`

telescope each night. At the end of the selection process (mid-2020), only a few of the proposals will be retained to analyse the full set of LSST alerts, and the selected brokers will start operating with the survey around 2022, for the next 10 years. These brokers will have a strategic position for the analysis of the transient sky with LSST, in particular by defining the science cases to be covered. They will also participate in important decision making processes by inter-operating inside networks of communication and follow-up telescopes, enabling the discovery of new objects and exploring new physics. **By investing now, the French community at large would secure a prominent place in this global effort, guaranteeing the scientific return on both LSST and collaborating scientists and experiments for the next decade.**

In response to the call, the French community started an interdisciplinary collaboration to propose such a community broker. The collaboration is **led by LSST-France and includes CNRS/IN2P3 and European experts in supernovae, micro-lensing, and multi-messenger astronomy. It also houses a large number of R&D in new technologies, with a special focus on developing and applying novel big data and streaming techniques.** More than 30 researchers and engineers from 12 different IN2P3 laboratories and associates co-authored the successful initial Letter of Intent to LSST. The main project is Fink, a broker framework being developed to ingest the alerts from LSST and redistribute them with added values. Fink is designed to be able to filter, aggregate, enrich, and transform alerts for further consumption or coordination of resources. In particular, the broker is designed to perform efficient crossmatching with different catalogues and alerts from MMA surveys, relying on state-of-the-art big data solutions coupled with state-of-the-art machine learning tools for the specific science cases mentioned above, and beyond.

The daunting volume of alerts sent by LSST is one of the major differences with respect to previous experiments [1], making inoperative traditional streaming and processing tools. In Fink we based the broker on the cluster computing framework, Apache Spark[2] [2], which has proven to be successful in many contexts to efficiently analyze large amounts of data. To operate efficiently, the broker operates in two stages: the first part provides fast, scalable, fault-tolerant solution to decode the stream, and archive it regardless the underlying volume of data, while the second part provides fast downstream access to the alert database for Artificial Intelligence applications, astrophysics software programs, and other services. Furthermore, the infrastructure heavily relies on widely used services and tools, to benefit from standards in place and ensure interoperability with the different actors. On the one hand, Fink users will be able to receive alerts tailored to their specific needs thanks to a tight integration with services developed and maintained at the Centre de Données astronomique de Strasbourg, including for instance the latest information from crossmatched surveys. On the other hand, interoperability between alerts generated by LSST and others will be ensured by Fink by using standard protocols largely used by the community such as VOEvents and VOEvent transport protocol. This framework will be paramount to guarantee optimal exploitation of the data gathered by LSST in real time and efficient communication with the different MMA communities.

The current design integrates state of the art elements in data processing and analysis, while accommodating for further developments expected within the 10 years lifetime of the LSST survey and beyond to insure a stable, long lasting solution for coordination of alerts in MMA landscape. In order to ensure that it will continue to take advantage of new technologies and theoretical developments, it is imperative to maintain a high level of collaboration and coordination between the different scientific communities involved in the study of the transient sky.

[2] `http://spark.apache.org/`

3 Fink science themes

3.1 Gamma-ray Bursts and Gamma rays

The gamma ray sky is nowadays very well covered by a number of instruments both on ground and in space: The Fermi Space Telescope and the H.E.S.S. telescopes are two of these major instruments in which IN2P3 is deeply involved. The future is also well prepared with major French contributions to SVOM and CTA. The study of Gamma-ray Bursts (GRBs) has seen major breakthroughs in the past 2 years with events such as GW170817, H.E.S.S. detection of GRB180720B afterglow and MAGIC early detection of GRB190114C: these events show that an in-depth understanding of GRBs must rely on multi-wavelength and multi-messenger observations. A key point to achieve such observations is quick and precise localization, including a good estimate of the redshift that can only be achieved through spectroscopy, that remains one of the main bottlenecks. In this task the Fink broker shall be helpful to better exploit LSST data by crossmatching GRBs alerts from large field of view gamma observatories to the LSST optical transient alert stream: counterparts could be identified early, properly classified and the best candidates proposed for follow-up to spectroscopic facilities. As during the baseline survey LSST will monitor half of the sky every few days, the timescale for joining data can extend to a few days after the GRB trigger, allowing the characterization of possible optical afterglow. Besides, Fink could also be a major tool in trying to extract from the LSST alert stream the best candidates of orphan GRB optical afterglows: counterparts could then be searched for in existing gamma ray data, and could also constitute a target of opportunity by one of the gamma-ray observatories of the network.

3.2 Neutrino

Neutrino astronomy allows us to study the most energetic non-thermal sources in the Universe. Despite observations of cosmic rays up to ultra-high energies and observations of γ-rays and astrophysical neutrinos, we do not yet know where or how these particles are accelerated. Neutrino astronomy is clearly a key to directly answer this question. Recently, coincident observations of neutrinos and γ-rays from the blazar TXS 0506+056 presented evidence of the first extragalactic neutrino source. However, this cannot be the entire story: multiple independent analyses indicate that only a small fraction of the diffuse neutrino flux can come from γ-ray blazars. The current lack of established neutrino point sources, despite a firm detection of a diffuse neutrino flux, indicates a dominant population of low-luminosity extragalactic sources. The deep sensitivity of LSST will be a strong asset to be able to identify and characterize the potential electromagnetic counterparts. KM3NeT will achieve a precision of $< 0.1°$ for the muon neutrino tracks at very high energies. One of the main expectations is the quite low angular resolution, $< 1.5°$, of the cascade events (ν_e-ν_τ + neutral current interaction of ν_μ). This is a factor of 10 in improvement compared to IceCube, allowing KM3NeT to perform a very efficient all-flavour neutrino astronomy. But despite the great effort in optimising the angular performance, the error boxes are still large so that generally few candidates remains as potential sources. LSST will be able to provide a complete optical survey of the sky visible by KM3NeT, and the Fink broker will be a very efficient tool to investigate all the possible candidates found.

3.3 Gravitational waves

Recently, the event GW170817 was the first gravitational wave observation which has been confirmed by non-gravitational means. The signal seen by the LIGO/Virgo gravitational wave detectors arose from the merger of two neutron stars, and it was accompanied by the detection of a gamma ray burst detected by the Fermi Gamma-ray Space Telescope shortly after, and finally scrutinized across the whole electromagnetic spectrum by dozens of other telescopes for months. **And yet**

despite these exciting first results, there are many holes in the puzzle that need to be filled, where LSST has a key role to play over the next decade. Thanks to its deep large field of view, LSST will be able to catch the very first glimpse of electromagnetic counterparts from gravitational wave alerts. In addition, by having the capability to quickly trigger additional follow-ups with its network, LSST will be able to give further constraints on time evolution models of the possible counterpart. The techniques planned in the Fink broker will be very effective to dig in all the possible candidates found within one field but also using efficiently the large volume of information from external observatories. Conversely this "traditional" sequence will be complemented in the next decade by detection of electromagnetic counterparts and then searches of low signal-to-noise gravitational waves. Fink will enable these searches in a timely manner using the LSST optical data stream.

3.4 Supernova

LSST is expected to discover 10^6 supernovae during its 10 year survey[3], orders of magnitude more than current surveys discovering thousands of supernovae. These stellar explosions may be visible, not only in multiple electromagnetic wavelengths, but also through other messengers. In particular core-collapse supernovae, which explode through mechanisms not completely understood, may produce diverse signals like neutrinos, gravitational waves and gamma rays. **In preparation for LSST, the supernova community has started to focus on disentangling possible interesting candidates to follow-up from others and enabling science with limited resources. In particular, photometric classification has become an importan research area.** This type of classification takes the evolution of brightness with time of supernovae and provides a type without the need of spectroscopy. **Fink is designed to have in its science module state-of-the-art photometric classifiers [4, 5]**. These classifiers will provide classification scores therefore ranking of potentially interesting supernovae which can enable appropriate follow-up for promising candidates in the case of early classification. Furthermore, these methods can be adapted to include MMA signatures and/or be extended to other astrophysical phenomena beyond supernovae.

3.5 Microlensing

IN2P3 has been a co-discoverer of microlensing as a technique to search for dark matter in the Milky-Way halo, and some endorsers of the present proposal were pioneers in this activity as founding members of the EROS collaboration, by designing and operating one of the earliest microlensing alert systems and enabling follow-up collaborations [6]. The systematic search for distorsions to the point-source point-lens rectilinear relative motion microlensing events allows to trigger follow-up from Earth and space that are essential (minutes to months timescale): (1) to detect parallax effects and source extension effects - providing extra-information about the mass and location of the lens, (2) to discover and characterize multiple lenses and multiple source systems, (3) to discover planetary systems, (4) to use the microlensing as a natural telescope, sometimes able to resolve a stellar disk, when a caustic line crosses the surface of a source. All of these cases dramatically benefit from time-critical added value with the capability to trigger fast cadence photometric observations through specific networks like MicroFUN [7], or astrometric and spectroscopic follow-up observations (through spatial telescopes and large spectrometers). However the search for short time-scale distortions in microlensing events is not practicable with LSST alone, and a systematic harvest – as required by the planet hunters – will need a collaboration with follow-up facilities triggered by the LSST brokers.

3.6 Anomaly detection

Perhaps one of the most anticipated outcomes of the MMA era is the identification of unforeseen astrophysical events. As an example, potentially unforeseen cataclysmic events leading to new mechanisms of particle acceleration or electromagnetic counterparts of GW can provide important insights into their progenitor systems. However, given the volume of data which will accompany them, it is very unlikely that such events will be serendipitously identified. Additionally, some of these events may be on very short time scales, making the identification task harder. Nevertheless, we cannot underestimate the impact of accurate and responsive anomaly detection in MMA science. It would enable crossmatches with alert brokers dealing with different regions of the electromagnetic spectrum and subsequent coordination of follow-up observations for confirmation. In order to fully exploit the potential of LSST for new discoveries, Fink is designed to have a specific anomaly detection module, based on contemporary adaptive machine learning techniques which will be specifically designed to optimize the use of domain knowledge [8]. This will allow Fink to deliver increasingly more accurate anomaly scores during the first years of the survey.

4 Preparing the astronomy communities for the next 10 years of MMA

In order to allow this unprecedented interconnection between different communities, it is important to foster interdisciplinary activities in all stages of scientific training. This includes: summer/winter schools detailing a range of astronomical probes and their theoretical backgrounds; joint gatherings tackling scientific aspects of current and future experiments and activities focusing on the construction of a common language across technical subjects (data management, reduction, storage and distribution; machine learning basics; new paradigm of programming techniques) and scientific applications[3]. The same strategy should also guide the development of long term activities of the permanent research staff. This includes specifically designed venues where social and professional connections can be established between researchers working on Science, Engineering, Technology and Math (STEM) related subjects - in academia as well as in the industry sector. Once these connections are established, there will be a natural avenue for the adaptation of state of the art techniques to the many different fields within the IN2P3 community involved in MMA.

At the highest level, inter-institute efforts are already in place such as the series of workshops Transient Sky 2020[4] organised by the *Programme National Hautes Energies*. For a few years, these events successfully gathered actors from different communities, and concrete collaborations and partnerships often emerged from the discussions. But the existence of these efforts in the long term must be protected by official inter-institute agreements ensuring long term logistic support and funding. **Despite the undeniable effort necessary to put these structures in place, or sustain the continuation of the activities already under development, such initiatives hold the key to ensure the long term scientific impact of interdisciplinary activities developed within IN2P3 in general and Fink in particular (during the duration of LSST).**

[3] An example of this type of event, in the international level, which has received strong support from CNRS and the French community in the last few years is the Cosmostatistics Initiative (COIN).

[4] `https://indico.in2p3.fr/event/19471/`

References

[1] E. Bellm, R. Blum, M. Graham, *et al.* Plans and Policies for LSST Alert Distribution. *Large Synoptic Survey Telescope (LSST) Data Management*, 2019.

[2] Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, *et al.* Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing. In *Presented as part of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12)*, pages 15–28, San Jose, CA, 2012. USENIX.

[3] LSST Science Collaboration, Paul A. Abell, Julius Allison, *et al.* LSST Science Book, Version 2.0. *arXiv:0912.0201*, Dec 2009.

[4] E. E. O. Ishida, R. Beck, S. González-Gaitán, *et al.* Optimizing spectroscopic follow-up strategies for supernova photometric classification with active learning. *MNRAS*, 483(1):2–18, Feb 2019.

[5] Anais Möller and Thibault de Boissière. SuperNNova: an open-source framework for Bayesian, Neural Network based supernova classification. *arXiv:1901.06384*, Jan 2019.

[6] Bruno Mansoux. *Première analyse des données relatives aux bras spiraux de la voie lactée dans l'expérience EROS 2*. PhD thesis, Paris 7, 1997.

[7] Microlensing Follow-Up Network. `http://www.astronomy.ohio-state.edu/~microfun/`.

[8] Emille E. O. Ishida, Matwey V. Kornilov, Konstantin L. Malanchev, *et al.* Active Anomaly Detection for time-domain discoveries. *arXiv e-prints*, page arXiv:1909.13260, Sep 2019.