*Revista Mexicana de Astronomía y Astrofísica Serie de Conferencias (RMxAC)*, **59**, 161–166 (2025)
© 2025: Instituto de Astronomía, Universidad Nacional Autónoma de México
https://doi.org/10.22201/ia.14052059p.2025.59.25

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Griffsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

# GERry : FINDING ELECTROMAGNETIC COUNTERPARTS IN GRAVITATIONAL WAVES FOLLOW-UP CAMPAIGNS

D. O'Neill[1]

RESUMEN

El campo de la astrofísica es cada vez más rico en datos, especialmente en la astronomía del dominio temporal, donde el número de telescopios robóticos para observaciones de gran campo sigue incrementándose. En parte, esto se ha debido al creciente interés en detectar las contrapartidas electromagnéticas (EM) de los eventos de ondas gravitacionales (GWs), en particular la contrapartidas ópticas de las fusiones de estrellas binarias de neutrones (BNS) que dan como consecuencia a las kilonovas (KNe). Se han diseñado nuevos telescopios como GOTO específicamente para este propósito. Sin embargo, el seguimiento robótico de GWs presenta dos desafíos: cómo optimizar su estrategia de observación para maximizar la probabilidad de recuperar una KN en sus observaciones e identificar qué campañas de seguimiento pueden haber producido un evento detectable. En este trabajo presentamos cómo el 'código de recuperación electromagnética de ondas gravitacionales' (GERry ) puede ayudar en estos aspectos al proporcionar un método para permitir rápidamente un análisis en profundidad de la campaña de seguimiento de GWs. Los datos que proporciona se pueden utilizar para determinar si es probable que haya una contrapartida óptica entre las fuentes detectadas, así como para probar la eficacia de la estrategia de observación para un evento determinado.

ABSTRACT

The field of astrophysics is increasingly becoming data rich, especially in time domain astronomy, where the number of wide field robotic survey telescopes is continuing to accelerate. In part, this has been due to the rapidly growing interest in detecting the electro-magntic (EM) counterparts to gravitational wave (GW) events, in particular the optical counterpart to binary neutron star (BNS) events: Kilonovae (KNe). New telescopes such as GOTO have been designed specifically for this purpose. However, robotic GW follow-up presents two challenges: How to optimise your observing strategy in order to maximise the probability of recovering a KN in your observations, and, identifying which follow-up campaigns may have produced a detectable event. In this work I demonstrate how the 'Gravitational wave Electromagnetic RecoveRY code' (GERry ) can help in these respects by providing a method to rapidly provide an in-depth analysis of GW follow-up campaign. The data it provides can be used to determine whether an optical counterpart is likely to be among the detected sources as well as to test the effectiveness of the observing strategy for a given event.

*Key Words:* gravitational waves — methods: observational — surveys

## 1. INTRODUCTION

The first detection of gravitational waves paved the way for a new era of multi-messenger astronomy through the combination of traditional electromagnetic (EM) datasets complimented by gravitational wave (GW) data. However, challenges remain in the detection of the optical counterparts. The interferometric design of gravitational wave detectors such as the Laser Interferometric Gravitational-wave Observatory (LIGO) and Virgo detectors can often result in poor localisations (∼ 100s - 10,000 sq deg). This poor localisation coupled with the rapid nature of KNe, typically rising and fading within ∼3–5 days makes it exceedingly difficult for traditional telescopes to find and identify these counterparts due to their limited field of view (∼sq arcminutes). Any counterpart will have likely faded before being imaged and identified. Therefore wide field survey instruments such as ATLAS, ZTF, PanSTARRs with fields of view in the 1-10s sq deg range are much more capable of identifying any EM counterparts and yet these events remain elusive. This could be set to change in the near future with the advent of new robotic wide field surveys such as BlackGEM, GOTO, LSST, etc

### 1.1. *GOTO*

Advances in CCD technologies has resulted in increasing supply of affordable, sensitive, large-scale CCDs allowing for the construction of increasingly sophisticated and cost-effective wide field telescope

[1]University of Warwick, Coventry, CV4 7AL.

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Gritsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

arrays. Motivated by this, the rapidly growing interest in GW optical counterparts as well as the aforementioned challenges they pose, the Gravitational wave Optical Transient Observer (GOTO) was born (Steeghs et al. 2022). GOTO has specifically designed to maximise the chances of recovering an EM counterpart signal. It consists of 4 mounts located across two sites, with two mounts in the northern hemisphere at Roque de los Muchachos Observatory in La Palma and the other pair in the southern hemisphere at Siding Spring Observatory in New South Wales, Australia. Each mount consists of 8 0.4m Telescopes each with a field of view of ∼5.8 sq deg. Allowing for a small amount of overlap between each of the telescopes, the field of view of a single mount is therefore approximately 44 sq deg. Images are currently taken in the L-band, a broad filter that covers Sloan r',g' and Johnson V (4000 − 7000Å). When a GW wave trigger is received, GOTO will automatically start imaging the 90% contour skymap, images are 4x60s mean-combined image stacks which typically an image depth of ∼20.5±0.5 mag depending on conditions and moon brightness. This setup should therefore be able to recover any 2017gfo-like KNe happening within $<$ 100 Mpc with possible detections out to ∼150 Mpc which is comparable LIGO's O4 estimated detection threshold for a BNS event.

## 2. MOTIVATION

At the onset of LIGO's O4a observing run, the northern node of GOTO was completed and operational while the southern node was still under construction. During this time, many low significance GW alerts had been released which were promptly followed up by GOTO where possible. It quickly became apparent that the number of sources detected for each gravitational wave event could span anywhere between 1 to 4 orders of magnitude, in particular this was a problem for single GW detector events which have localisations spanning half the sky. One example is that of S230529ay[2], an event at 197±62 Mpc with $P_{NSBH} = 62\%$, $P_{BNS} = 30\%$. While not statistically significant, its relatively low FAR of 1/160 years still garnered some interest. GOTO did not perform a dedicated follow-up campaign but instead continued in the default all-sky imaging mode. However, due to the very large localisation area, much of the 90% contour area was imaged regardless. Within 1 week post GW trigger, GOTO detected ∼50,000 sources requiring human vetting, which would be a significant time investment. Naturally the feasibility of such a task was questioned and whether it is likely that a KNe would ever have been detected from such a follow-up campaign. Another important consideration is the observing strategy one wishes to employ when observing GW skymaps. Typically the primary choice when considering a strategy is speed vs depth, shorter exposures will result in shallower images, but increases the rate at which you can cover the skymap. Intuitively one might favour longer exposures for further GW events as any counterpart will be fainter, and thus require deeper images to recover the signal. However, with increasing distance comes a lower GW signal-to-noise ratio (SNR) which in turn, results in poorer localisation requiring shorter image exposures to fully map. Finding which strategy to adopt can be difficult even before considering other factors such as airmass, observing conditions etc.

GERRY was developed to address these challenges. It provides a rapid method of quantitatively assessing GW follow-up performance and provide users with a means to quickly determine the probability of detecting an EM counterpart ($P_{tr}$ ) in a given GW follow-up campaign. Running the code on simulated observations of skymaps also allows the user to compare the effectiveness of different observing strategies allowing for improvement and refinement which is critical if we want to maximise the probability of detecting these EM counterparts.

## 3. METHOD

### 3.1. *Healpix Alchemy*

Upon the detection of a GW signal, LIGO releases a localisation skymap typically generated by the rapid localisation code BAYESTAR. These localisation skymaps utilise Hierarchical Equal Area isoLatitude Pixelization (HEALPix, Górski et al. (2005)), which models the sky as a sphere and subdivides the sky into equal sized pixels. The size of these pixels depend on the resolution. Also contained within each pixel is the probability and GW distance estimates at the position of the pixel. In the past these 'flat' skymaps were the primary format being used by the community. However, more recently 'Multi-Order Coverage' (MOCs) skymaps (Fernique et al. 2022) have become the default. Instead of a single resolution across the entire sky, GW MOC files instead reserve higher resolutions for areas of the sky with higher GW localisation probabilities resulting in significantly less pixels/memory being used for areas of the sky with negligible probabilities leading to a significant reduction in file sizes

[2] https://gracedb.ligo.org/superevents/S230529ay/view/

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Gritsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

and the computational resources required to process them.

GERRY is built on healpix-alchemy[3] (Singer et al. 2022), which allows the querying of GW MOCs using SQLalchemy in a PostgreSQL 14 database. healpix-alchemy utilises the mulitrange data type, and uses this to store HEALPix skymap tiles at a given resolution by a number range, also known as a range set. The range set corresponds to the minimum and maximum HEALPix pixel indices at a very high resolution(l=29). Given the sequential nature of these l=29 pixels, one HEALPix tile can therefore be represented by a single range set containing the minimum and maximum indices of the pixels which make up the larger tile which will take the form of e.g. [576460752303423488,577586652210266112). Therefore, the GW skymap can be broken down into a set of rangesets for each tile. These tiles are then stored in the PSQL14 database along with the tiles probability density values, distance estimate and distance uncertainty. The same process is performed for the images in the follow-up campaign, whereby the image is first turned into a MOC, the field of view of the image is broken down into healpix tiles with varying resolution below a maximum limit. They are organised such that the tiles fill the field of view as much as possible while still accurately representing the original field of view with larger resolutions (smaller tiles) at the edges and corners of the images and lower resolution tile (larger tiles) in the centre.

Once the skymap and image tiles have been read into the database, their rangesets can now be trivially queried against each other. If the pixel values in a skymap tile rangeset and an image tile rangeset overlap, then this means the the image tile overlaps the position of the skymap tile on sky. The size of the this overlap can also be found as the l=29 pixels in the rangeset have a fixed size ($3.6\times10^{-18}sr$), multiplying this value by the number of values common in both the skymap and image tile rangesets will give the area of the overlap. Healpix-Alchemy also adds functionality to union field images which can then be used to easily calculate the contained skymap tiles and using the encoded probability values contained within, calculate the total 2D sky probability imaged ($\mathrm{P}_{fov}$) in a follow-up campaign.

### 3.2. GERRY : *probing 3D probability space*

The issue with simply quoting $\mathrm{P}_{fov}$ values is that it is possible for many wide field survey instruments to image a significant fraction of the skymap and achieve high $\mathrm{P}_{fov}$ If they do not achieve sufficient depths to ever recover an EM counterpart, or image the sky long after any EM counterpart has faded, then this value can be somewhat meaningless as the chances of detecting with counterpart will be negligible. GERRY builds on Healpix-Alchemy by also utilising the encoded distance information to probe the 3D probability space while folding in the expected temporal evolution of the counterpart and returning meaningful statistics and probability values that directly tie to the probability of detecting an EM counterpart. The temporal evolution of the EM counterpart is estimated using model light curves in the desired filters. In all the following examples we will be using a MOSFiT model L-band light curve generated using the best fitting parameters to KN 2017gfo unless stated otherwise.

The basic flow of operations in GERRY is as follows:

- Find all instances where a skymap tile was imaged by an image tile, and identify what specific area of the tiles was covered and what time the skymap tile was imaged relative to the GW trigger time.
- Using the model light curve, estimate the brightness of the counterpart at the time the skymap tile was imaged by using the encoded GW distance and Milky Way reddening estimate at that position.
- Given the known depth of the image, calculate the fraction of probability that could be recovered at the position of the tile at that time by modelling the brightness probability distribution as a inverse cumulative distribution function and sampling the distribution at the image depth value.
- If a skymap tile was wholly or partially imaged multiple times, combine the probabilities from these overlapping observations and finally sum these probabilities over the entire imaged sky.

When multiple image tiles cover a skymap tile it is vital that it is determined if the image tiles cover the same or different areas of a skymap tile to facilitate precise area and probability calculations for a more meaningful and accurate result. GERRY identifies what specific areas of a skymap tile was covered by an image tile by breaking down the skymap tiles into polygons composed of unique image-skymap tile combinations.

$$P_{tr,poly_i} = P_{max,poly_i}(1-(P_{int_1,img_1}...\times P_{int_n,img_n})) \quad (1)$$

[3] https://github.com/skyportal/healpix-alchemy

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Gritsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

Where $P_{tr,poly}$ is the probability of detecting the transient at least once within this polygon, $P_{max,poly}$ is the maximum probability achievable from the polygon and $P_{int_n,img_n}$ is the recovered probability of this polygon from the $n^{th}$ image tile. $P_{tr}$ is then simply the sum of all $P_{tr,poly}$ values for all $N$ polygons:

$$P_{tr} = \sum^{N} P_{tr,poly_i} \quad (2)$$

The primary numerical output of GERRY are three probability values: $P_{fov}$ - the aforementioned 2D probability imaged, $P_{vol}$ - The sum of the maximum probabilities achieved across the entire imaged sky and finally, $P_{tr}$ - the probability of detecting the transient (e.g. KNe) at least once across the imaged sky.

The input required for GERRY is a csv file containing the image corners/vertices in ra/dec, the MJD or JD when the image was taken, the filter, image depth and image depth uncertainty. The primary driver of errors in the returned $P_{tr}$ and $P_{vol}$ values is the uncertainty in the image depths. Image depth will typically vary across an image. In order to simulate this, GERRY will need to be given an image depth uncertainty value for every image. A script is included that will extract the image corners and generate the required csv file from a directory of fits images, however the script will only work on square/rectangular images. The image depths and depths uncertainties will still have to be added manually to the csv file. The images will be converted into MOCs with the maximum resolution set by the user. We strongly recommend avoiding resolutions lower than 10, especially for instruments with smaller (<1 deg) field-of-views. At these lower resolutions, the ability of the larger moc tiles to approximate the image field of view will deteriorate towards the edge of the images resulting in larger errors in the calculated error of the field images. In most cases a value of 11 or 12 should be sufficient. In order to accurately determine the effect of reddening the transmission curve for all the filters used should also be provided. The model light curve of the transient is also required in all of the filters used in the imaging. Finally the GW skymap must be using the new MOC format.

## 4. RESULTS

In this section I will briefly show the output from GERRY using an example GW event from O3 as well as how it can be used with simulations to optimise observing strategy.

### 4.1. *O3 example - S190425z*

S190425z was the only statistically significant BNS merger to be detected during LIGO's O3 observing period. The estimated distance to the merger was 157±43 Mpc, it had the potential to produce a detectable, albeit faint, EM counterpart. It was only detected by a single LIGO inferometer as well as Virgo resulting in relatively poor localisation of ∼9000 deg$^2$.

A GOTO prototype (GOTO-4) was in operation during this time which consisted of a single mount with 4 telescopes at Roque de los Muchachos observatory in La Palma. Observations of the skymap started at 2019-04-25 20:40 to 2019-04-28 05:15 (+0.5 to 2.8 days post trigger) and resulted in 1948 L-band images covering 2155 deg$^2$. Figure 2 shows a detailed breakdown of the follow-up campaign along with relevant sky-position information such as reddening, image depth, and recovered probability. The total field-of-view probability covered was$P_{fov}$ = 32%.

GERRY was run to analyse the follow-up campaign and using the aforementioned MOSFiT 2017gfo light curve the resulting volume probed was $P_{vol}$ =$4.45^{+2.9}_{-2.01}$% with a final value for $P_{tr}$ =$7.72^{+4.07}_{-3.07}$%. Therefore, it is perhaps not surprising that no promising counterpart was detected in this follow-up campaign.

Figure 1 shows the expected light curve evolution along with the image depths achieved over the course of the campaign. Top left plot clearly shows that while images were taken at the expected KN peak, they still did not achieve the sufficient depth required to recover a 2017gfo-like KNe. This is primarily due to distance, as the top right plot shows that the majority of the imaged probability suffered little dust reddening.

### 4.2. *Optimising strategy*

The primary decision when considering strategy for surveys such as GW follow-up is speed vs image depth. More distant BNS mergers require both increasing image exposures due to the fainter counterparts as well as faster skymap coverage due to poorer localisation. Given that both of these requirements cannot be satisfied, knowing which is the optimal strategy as well as identifying possible parameter thresholds which may require different strategies is critical. By comparing the resulting $P_{tr}$ values for each strategy for our simulations, we can compare their performance.

A sample of 60 simulated O4 BNS skymaps with distances <600 Mpc were selected and a 3 day ob-

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Gritsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

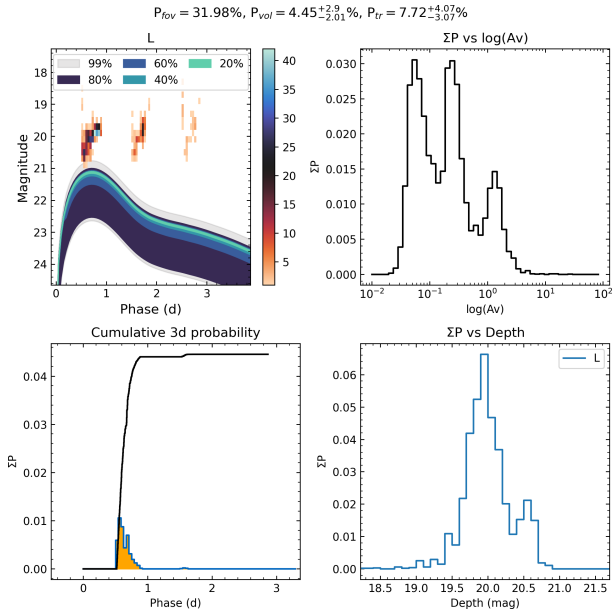


Fig. 1. Top left: The light curve distribution of the model light curve over all parts of the image skymap after accounting for variations in distance and reddening. The squares are a 2D histogram showing the density of images with respect to time and depth. Top right: The reddening distribution over the GW localisation probability. Bottom right: The cumulative 3D probability probed with respect to time. The overwhelming majority of the probability was recovered in the first night when the counterpart would have been brightest. Bottom right: The probability distribution with respect to Depth.

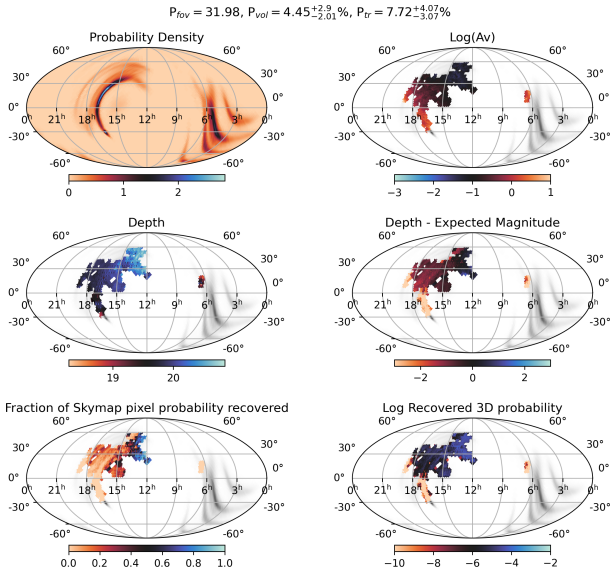


Fig. 2. Results from GERRY showing pointing specific information such as reddening, depths achieved, depths relative to estimated counterpart magnitude and probability recovered.

serving campaign was simulated using the GOTOTILE python package. Users can input the field of view of their instrument, location, number of mounts, exposure times and how long the skymap should be observed. The simulations were run for 2 exposure times, the first was a 4x60s exposure as well as an additional 60s for overheads per pointing resulting in a total pointing time of 300s. The second was 6x90s exposure with the same additional 60s for overheads. These simulations however can not account for sub-optimal observing conditions therefore it is assumed all observations occurred during bright time resulting in estimated depths of 20.3 mag for the 300s scenario and 21.3 mag for the 600s scenario. In both cases we adopt an image depth uncertainty of 0.3 mag. The results are shown in Figure 3. While both exposure scenarios offer similar levels of performance for distances within 100 Mpc, the longer t600s strategy starts to outperform t300s at distances >150 Mpc. Indeed at a distance of 200 Mpc, 600s offers almost a 50% increase in $P_{tr}$ and at 300 Mpc, a factor of 5 increase in performance. While it would be tempting to simply adopt the t600s in all scenarios, given the similarity in performance at distances <100 Mpc, the t300s option would be preferable in this nearby distance range as an increase in cadence can help constrain source rise times, helping to separate any fast rising EM counterpart from more common SNe.

Since the above simulations now give us the average recovery curve over distance, we can use this curve along with the estimated BNS merger rates to estimate the number of KNe that we could expect to recover (assuming the KNe will have a 2017gfo-like light curve) using the 2 different strategies. This is shown in Figure 4.

### 4.3. *Performance*

Using the S190425z as an example, for a 1948 image follow-up campaign using an image MOC resolution of 10, the runtime of GERRY using a typical laptop with 4 cores, 16GB RAM is ∼7 minutes with an additional 1 minute if the results are plotted. The execution time scales linearly with the number of input images. For each increase in the MOC resolution of the input images typically results in a doubling of image tiles and execution times.

## 5. CONCLUSION

Identifying GW EM counterparts is difficult due to the typical poor localisations as well as their inherent elusive nature. Here we have shown how GERRY

VII Workshop on Robotic Autonomous Observatories (October 16-20, 2023)
Editors: Maria Gritsevich, Alberto J. Castro-Tirado, Petr Kubánek, Shashi B. Pandey, and David Hiriart - DOI: https://doi.org/10.22201/ia.14052059p.2025.59.25

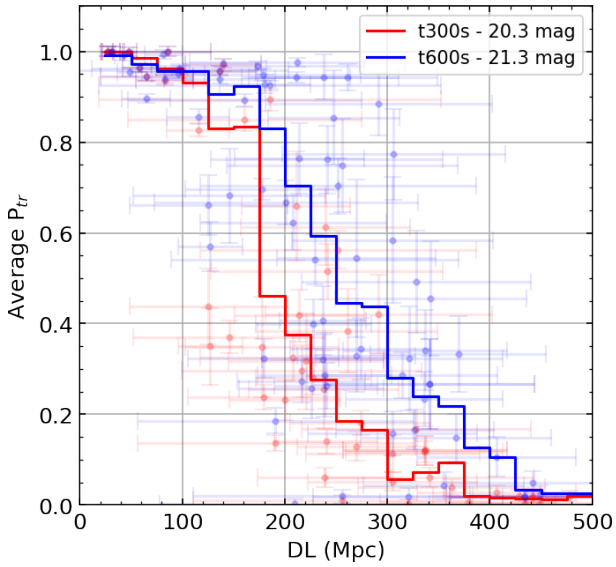


Fig. 3. Results from the follow-up simulations showing the effect of different exposure times on $P_{tr}$ . The blue and red points show individual results from the simulations, the lines show a weighted mean curve binned in 25Mpc intervals.

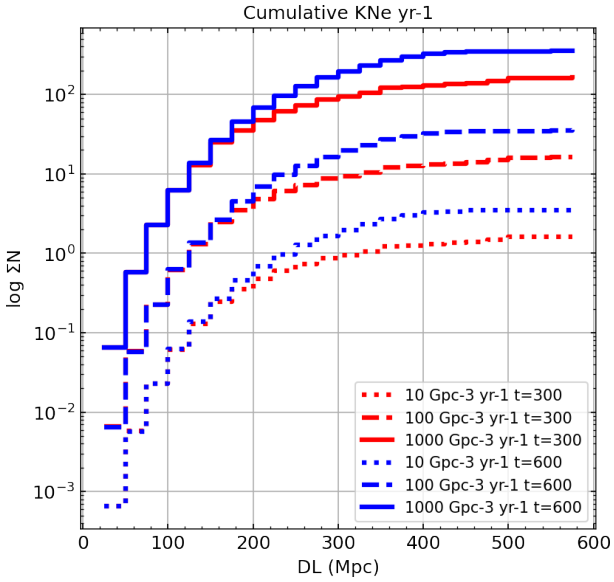


Fig. 4. The expected number of recovered 2017gfo-like KNe using both exposure strategies for different BNS merger rates.

can be a valuable tool in analysing GW follow-up campaigns by providing meaningful probability and follow-up statistics so that the users can determine whether a follow-up campaign should have produced any viable candidates. We have also shown how by leveraging these probability values, one can use simulated GW skymaps and observations to determine the most effective strategy for their instrumentation in order to maximise the potential of their future follow-up campaigns. Much of the work here will be presented in much more detail in an upcoming paper (O'Neill,D & Lyman, J in prep.) as well as showing some of the more advanced capabilities of GERRY and a study of the current and future suite of wide field survey instruments with regards to their expected performance in future GW observing periods. The Code is not yet public but can be found on GitHub when released.

## REFERENCES

Fernique, P., Nebot, A., Durand, D., et al. 2022, MOC: Multi-Order Coverage map Version 2.0, IVOA Recommendation 27 July 2022

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Singer, L. P., Parazin, B., Coughlin, M. W., et al. 2022, AJ, 163, 209

Steeghs, D., Galloway, D. K., Ackley, K., et al. 2022, MNRAS, 511, 2405