CrossMark

An Unambiguous Separation of Gamma-Ray Bursts into Two Classes from Prompt Emission Alone

Christian K. Jespersen[1,2] , Johann B. Severin[1,2] , Charles L. Steinhardt[1,2] , Jonas Vinther[1,2] , Johan P. U. Fynbo[1,2] ,
Jonatan Selsing[1,2] , and Darach Watson[1,2]
[1] Cosmic Dawn Center (DAWN), Lyngbyvej 2, Copenhagen, Denmark; christian.jespersen@nbi.ku.dk
[2] Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark
Received 2019 December 15; revised 2020 May 22; accepted 2020 May 26; published 2020 June 15

Abstract

The duration of a gamma-ray burst (GRB) is a key indicator of its physical origin, with long bursts perhaps associated with the collapse of massive stars and short bursts with mergers of neutron stars. However, there is substantial overlap in the properties of both short and long GRBs and neither duration nor any other parameter so far considered completely separates the two groups. Here we unambiguously classify every GRB using a machine-learning dimensionality reduction algorithm, t-distributed stochastic neighborhood embedding, providing a catalog separating all Swift GRBs into two groups. Although the classification takes place only using prompt emission light curves, every burst with an associated supernova is found in the longer group and bursts with kilonovae in the short, suggesting along with the duration distributions that these two groups are truly long and short GRBs. Two bursts with a clear absence of a supernova belong to the longer class, indicating that these might have been direct-collapse black holes, a proposed phenomenon that may occur in the deaths of more massive stars.

Unified Astronomy Thesaurus concepts: Gamma-ray bursts (629); Astronomy data analysis (1858); Astronomy data visualization (1968); Light curve classification (1954); Supernovae (1668); High energy astrophysics (739)

Supporting material: machine-readable table

1. Introduction

The duration of a γ-ray burst (GRB) is a key indicator of its origin, with long-duration bursts typically associated with the core-collapse of a stripped massive star (Hjorth et al. 2003; Stanek et al. 2003) and the short-duration bursts with mergers of neutron stars (Berger et al. 2013; Tanvir et al. 2013; Ghirlanda et al. 2018). The dividing line is usually set at $T_{90} \approx 2$ s (Kouveliotou et al. 1993; Tavani et al. 1998; Paciesas et al. 1999). However, the long and short distributions are known to overlap substantially and neither duration nor any other parameter so far considered such as spectral hardness (Kouveliotou et al. 1993) or lag (Norris et al. 1986; Norris & Bonnell 2006) gives a clean separation of the progenitor type based on the prompt properties.

Considerable effort has been expended in trying to find a clean separation of burst types based on many other burst properties (Fruchter et al. 2006; Nakar 2007; Bromberg et al. 2011, 2013; Zhang et al. 2012). However, a complete separation has not yet been possible.

Ideally, classifying bursts should be done from the entirety of the light curve. However, the full light curve is a high dimensional data set, which makes comparing bursts hard because it is difficult to determine which information is most important. As a result, proposed classifications have previously relied on a small number of easily-described summary statistics such as duration, spectral hardness, and lag. This has led to significant progress, but not to a clean separation between short and long GRBs.

It has been hoped that with the right choice of summary statistics, there would be a clear separation into short and long GRBs. This work uses the dimensionality reduction algorithm t-distributed Stochastic Neighbor Embedding (t-SNE; Maaten & Hinton 2008; van der Maaten 2014) to instead reduce full GRB light curves to points in a two-dimensional space. The location that each light curve is mapped to in these two coordinates does not represent properties of that light curve, but rather can only be calculated from information about all of the light curves in the entire data set. Applying this technique to light curve data from the Swift satellite, we find a clear separation into two groups, with possible additional subgroups.

In Section 2, we describe t-SNE and the process of assembling a map from the Swift data set. The resulting map and implied classification into short and long GRBs is shown in Section 3, along with a comparison between each group with expected properties. In Section 4 we discuss how this classification works for GRBs of special interest. These results are discussed further in Section 5, including an application of this categorization to GRBs whose classification has been the subject of past debate.

2. Methods: Applying t-SNE to GRB Light Curves

The classification proposed here takes the entirety of the normalized Swift light curves from prompt emission and in an unsupervised way determines which GRBs should be considered similar based on the prompt data alone. This is done using t-SNE, a dimensionality reduction algorithm that can take complex, high-dimensional data and produce a faithful representation of that data in a low-dimensional space.

2.1. t-Distributed Stochastic Neighbor Embedding

t-SNE takes a set of high dimensional vectors, $\{x_i\}$, and calculates the probability that each x_j should be considered a neighbor of x_i from the set of Euclidean distances $\{|x_j - x_i|\}$ and the perplexity, a hyperparameter that determines the sizes of the neighborhoods based on the density of the data in the respective regions and can be approximately interpreted as the typical number of neighbors that should be considered similar

when computing distances. Perplexity is formally defined (Maaten & Hinton 2008) as

$$\text{Perp}(P_i) = 2^{H(P_i)}, \qquad (1)$$

where P_i represents the conditional probability distribution over all other $\{x_j\}$ given $\{x_i\}$, and $H(P_i)$ is the Shannon entropy of P_i in bits. A higher perplexity increases the values for points further away and promotes global structure, whereas a lower perplexity is used to look for local structures and groupings of only a few vectors. The introduction of perplexity is necessary since dimensionality reduction cannot simultaneously preserve both small- and large-scale structure in lower dimensions, and perplexity controls their relative importance. It should be noted that generating probabilities from Euclidean distances requires that every x_i have identical dimensionality, with none missing.

These probabilities are then used to map the set of $\{x_i\} \rightarrow \{y_i\}$ in the lower-dimensional space, such that the probability of y_i and y_j being neighbors in the new mapping is as similar as possible to the probability that x_i and x_j were considered neighbors. This process depends upon random initialization, and running t-SNE on an identical data set can produce a variety of maps with similar topology (cf. Steinhardt et al. 2020, Figure 1). Thus, the axes of this low-dimensional space have no proper labels or meaning; unlike principal component analysis, an object further to the right on a t-SNE map is not in some sense more "x-like", but merely more similar to nearby objects and less similar to more distant ones. Essentially, the goal of t-SNE is to produce a simplified map that can be easily visualized while preserving the structure of the original data set, not to find a vector basis.

Although t-SNE is considered an unsupervised algorithm, in practice the final map depends upon not the data and hyperparameters alone, but also a series of human selections in preparing that data. Producing a data set of vectors with identical components requires a combination of removing individual objects for which some properties are poorly measured or unmeasured and instead removing some properties from the calculation altogether (or choosing an alternative metric for calculating neighbor probabilities). A choice of how to format and scale the raw data can change the relative weighting of various components in calculating neighbor probabilities. These human influences can often be minimized by running t-SNE on highly standardized data sets, such as the catalog produced by one survey or one observatory.

2.2. Data Preparation for the Swift Catalog

Here, t-SNE is run on the full Swift GRB catalog[3] (Lien et al. 2016) in an attempt to classify all detected bursts. Swift measures the light curve of each burst in four bands, 15–25, 25–50, 50–100, and 100–350 keV. The Swift data are released as four binned, background-subtracted light curves with temporal resolution 64 ms, and the start of the burst may vary slightly from the trigger time. Because bursts vary greatly in duration, there is a very wide range in the number of time bins with statistically significant flux, where the reported flux for Swift is measured as photon count $\text{cm}^{-2}\,\text{s}^{-1}$.

Therefore, it is essential to standardize the data set prior to running t-SNE, ideally in a manner that preserves meaningful differences but erases differences that should not have physical

origin. For example, a translation of all light curves by the same constant time offset should be ignored. A similar problem is tackled when analyzing the benchmark Modified National Institute of Standards and Technology handwritten digit data set with machine learning, since the data needs a large amount of preprocessing (e.g., normalization, deskewing, and noise removal) in order to be effectively analyzed (LeCun et al. 1998; Deng 2012). The goal is therefore to prepare data for t-SNE in a way that retains all of the useful information but removes irrelevant information that the analysis might otherwise use for classification. The end result of this process must be a set of vectors of identical dimension, for which it is still proper to expect that bursts with similar physical origin will be similar when compared component by component.

In our data, the main possible distractions would be (1) the total time-integrated flux, which would only return the known result that long GRBs typically but not always have higher fluence; (2) different lengths of light-curve measurements, which can depend upon noise in the tail of the Swift light curves and is therefore extremely important; and (3) the possibility of a trigger time offset common to all four Swift bands, although any relative time delay between different bands is still meaningful.

To remove these, light curves are normalized in every band by the total time-integrated flux across all bands of that specific burst (removing (1)), zero-padded in order to ensure a common axis and length (partly removing (2);[4] see Sacchi & Ulrych 1998; Shen & Wang 2006), then concatenated before taking the discrete-time Fourier transform (DTFT) of the light curve measured in each individual band (removing (3)). Because t-SNE relies on Euclidean distances, omitting these steps would group bursts almost solely by effects which are known not to produce a meaningful classification, e.g., omitting the normalization would effectively group bursts nearly solely by the magnitude of the flux, which would render only a known result; that long GRBs typically, but not always, have higher fluence than short GRBs (e.g., Ghirlanda et al. 2009), but instead here bursts are grouped using the entirety of the remaining information.

It should be noted that the DTFT is not necessarily an optimal preprocessing solution for separating GRBs. Other preprocessing techniques that preserve the information content in the light curve would yield equally valid representations of GRB classes, though possibly not as cleanly separated. As explained in the previous paragraph, the main motivation for the DTFT is that it suppresses some of the biases in the light curves, while building on previous successful descriptors (duration, hardness, and spectral lag) indicating that the general shape of the light curve could be a defining characteristic.

The DTFT also does not formally guarantee that similar light curves end up as close neighbors, but rather that light curves with similar DTFT end up close together, as t-SNE acts on Fourier components rather than flux measurements. Although in principle one could imagine constructing a function that would render a DTFT similar to a GRB light curve despite looking dissimilar, this is unlikely to occur in real data. Thus, having similar DTFTs is equivalent to having similar shapes, since the DTFT in essence is just a change in basis.

---

[3] Available at https://swift.gsfc.nasa.gov/results/batgrbcat/.

[4] The remainder of (2) is completely accounted for by the fact that an erroneous bias in duration would only, at worst, include noise-dominated measurements that do not affect the FT significantly, and do not add any significant fluence.
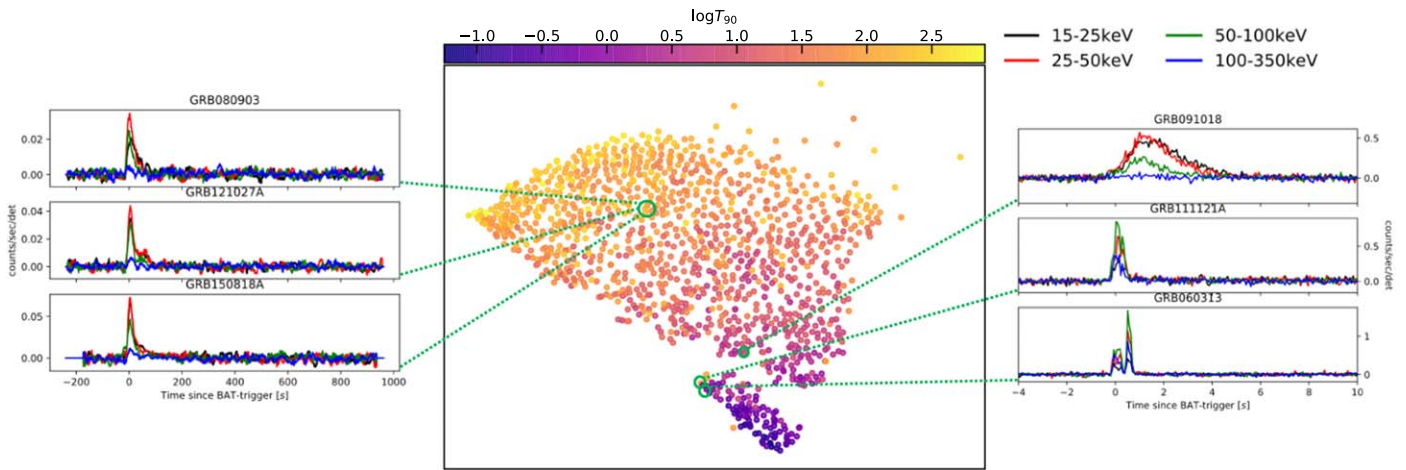
Figure 1. t-SNE mapping of Swift light curves, colored based on duration $\log(T_{90})$. Several sample light curves in the four observed bands are shown, with similar light curves placed as near neighbors and dissimilar light curves placed further apart. A clear separation into two groups is visible, with the smaller, bottom group (referred to as type-S) generally but not uniformly of shorter duration (see Figure 2). The axes resulting from t-SNE have no clear physical interpretation or units; only the structure is meaningful.

Previous attempts at classification have typically focused on summary data, such as duration (typically T_{90}, the duration containing 90% of the statistically significant flux, or T_{100}, containing 100% of the flux), spectral lag, flux or fluence, and hardness. These are properties that can be calculated from the light curves in each band, and therefore part of what is available to t-SNE, but represent only a small subset of the full data set. Here, the full Swift light curves are encoded then used, and neighbor probabilities are ultimately produced from ≈ 30000-dimensional vectors, encoding far more information than summary data.

2.3. Summary of Classification Scheme

Using the background-subtracted light curves, Swift bursts can be classified with the following procedure:

1. The 64 ms binned light curve in each band is limited to the interval out to T_{100}.[5] Afterwards they are padded with zeros, placing light curves of different length on a standard basis.
2. Individual light curves are normalized by the total fluence, obtained as the numerical integral of the flux across all bands, preserving spectral hardness.
3. The resulting light curves are concatenated and the DTFT is applied, producing vectors suitable for input into t-SNE.
4. t-SNE is applied to the vectors x_i, consisting of the Fourier amplitudes, with perplexity set to 30 (a choice specific to the Swift catalog), terminating once no improvement is made on the cost function.
5. The resulting map is examined, in the hopes that bursts will be clearly separated into clusters that can be interpreted as arising from a similar physical origin.

The success of this procedure is evaluated in the following sections.

3. Classification

The procedure described in Section 2.3 indeed maps the GRBs in the Swift catalog to a two-dimensional space (Figure 1). As a result, light curves with similar Fourier amplitudes end up as close neighbors, which translates to light curves with similar shape being grouped close together, as per the discussion in Section 2.2.

Even more promisingly, there is discernible structure in the resulting map. The exact mapping can change depending upon random seeds or the order in which data are presented to t-SNE, but structure such as grouping and topology remains (see Figure 1 and the related discussion in Steinhardt et al. 2020). Bursts are divided into two groups, with a clear separation between the larger group at the top and a smaller group at the bottom as presented in Figure 2(a). Further substructure exists as well, discussed in Section 3.1.

A closer examination of the duration (T_{90}) of each burst indicates that the t-SNE map has grouped objects approximately (but not entirely) by duration (Figure 1). The larger group at the top is generally of longer duration, while the smaller group at the bottom is shorter. Both groups have possibly (skewed) Gaussian distributions in log duration (Figure 2(b)), with some overlap in the $T_{90} \sim 1$–10 s range.
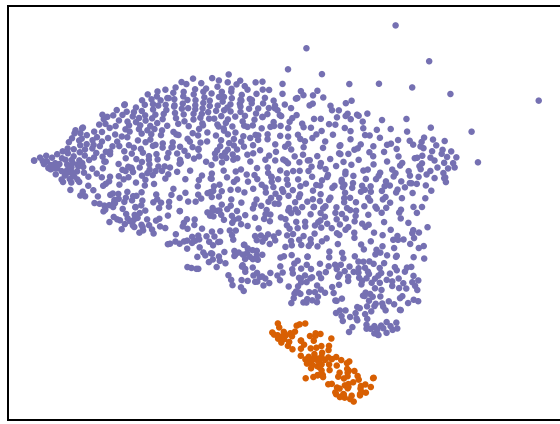
In short, the mapping cleanly divides Swift GRBs into two groups, with the duration distributions of the two groups similar to what might be expected for a classification into short and long bursts. To avoid confusion with preexisting classifications, the two t-SNE groups will be referred to as type-S and type-L.

A full list of our classification of Swift GRBs is given in Table 1.
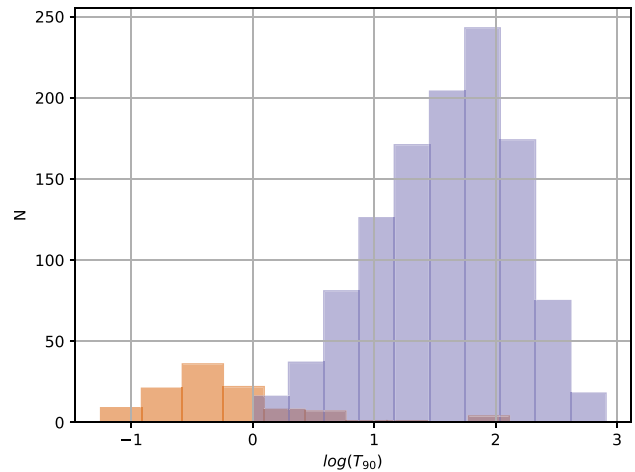
3.1. Comparison with Other Classifications

Although the classification of many individual GRBs as short or long is uncertain, particularly for bursts of intermediate duration, many bursts can be unambiguously classified as short or long on the basis of other observations. It is therefore necessary to confirm that the classification proposed here matches these previous results.

Perhaps the strongest association is that between supernovae and long GRBs, with every GRB with detection of an

---

[5] The DTFT would be unaffected by the additional extra bins containing solely noise, and thus the limited information between T_{90} and T_{100} can be included.

(a)

(b)

Figure 2. (a) The clear separation into two groups (purple and orange) strongly suggests a classification of GRBs into two distinct types. (b) The distributions of duration in $\log(T_{90})$ of type-L and type-S bursts are approximately normal and similar to those expected for a long and short classification.

Table 1
Classification of Swift GRBs as Either Type-S or Type-L Based on the Separation in the DTFT-based t-SNE Map

GRB	T_{90} (s)	Type
GRB 190727B	39.2	L
GRB 190719C	185.8	L
GRB 190718A	704.0	L
GRB 190706B	43.6	L
GRB 190701A	38.4	L
…	…	…

(This table is available in its entirety in machine-readable form.)

associated supernova unambiguously being long (Hjorth & Bloom 2012; Cano et al. 2017). Eleven bursts in the Swift catalog, GRB 060218A, GRB 071112C, GRB 100316D, GRB 111209A, GRB 111228A, GRB 120714B, GRB 120729A, GRB 130215A, GRB 130831A, GRB 161219B, and GRB 171010A have clearly detected associated supernova that are well-characterized by spectroscopy (Bufano et al. 2012; Cano et al. 2014, 2017; Kann et al. 2019; Klose et al. 2019). Due to corrupted files for GRB 060218A, not having BAT light-curve data available for GRB071112C[6] and GRB171010A,[7] these are not included. The remaining eight are all classified as type-L by the t-SNE map (Figure 4).

It has also been suggested that there may be a link between the hardness of the GRB spectrum and classification (Kouveliotou et al. 1993; Ghirlanda et al. 2009). It is known that this does not produce a clean, unambiguous separation between short and long GRBs. However, there is still likely a strong correlation between hardness and type, particularly since the hardness of a burst should be closely related to its physical origin.

The t-SNE mapping indeed shows that bursts with similar light curves tend to have similar hardness (Figure 3). Bursts with harder spectra group together in tight clusters on the t-SNE map, and the hardest bursts lie in a tight cluster and are
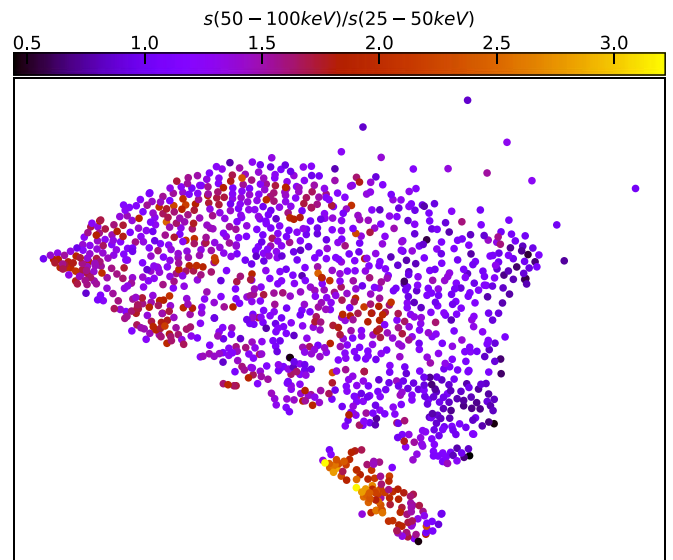

Figure 3. Coloring bursts on the t-SNE mapping by the flux ratio of the 50–100 and 25–50 keV bands, a proxy for hardness, indicates several clusters of hard bursts. The hardest bursts are all classified as type-S, but some type-S bursts do not have hard spectra, and several tight clusters of type-L bursts also exhibit harder spectra. A reasonable interpretation is that there are multiple classes both of type-S and -L bursts with different physical origins, with some of these origins producing harder spectra.

all categorized as type-S. However, clusters of hard bursts exist both in the type-S and type-L groups. In general, most type-S bursts are harder than most type-L bursts, but hardness alone is insufficient to determine whether a burst is type-S or type-L. Rather, the t-SNE map (Figure 4, bottom) at lower perplexity suggests that there might be substructure in both the type-S and type-L classes. The cause for this is currently undetermined, but could be associated with spectral hardness. Another measure of this feature could be the peak energy in $\nu f(\nu)$, $E_{\rm peak}$. However, Swift frequently is unable to reliably measure $E_{\rm peak}$ due to the relatively low energy bandpass of Swift-BAT. As a result, there is less structure in maps colored by $E_{\rm peak}$ than in maps using spectral hardness ratio compared across the Swift-BAT bandpass, although hardness also is an incomplete diagnostic.

---

6 Since its emission overlapped with 071112B.
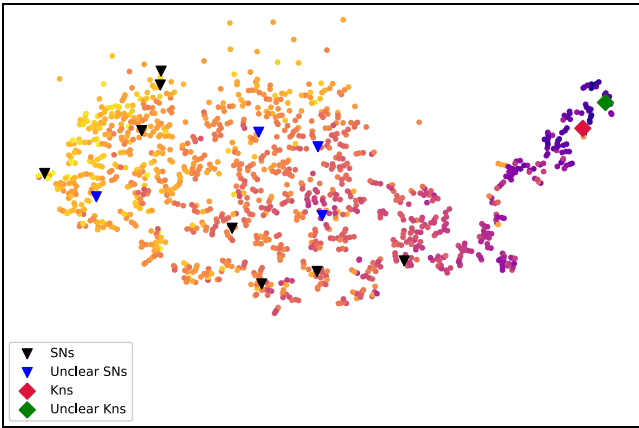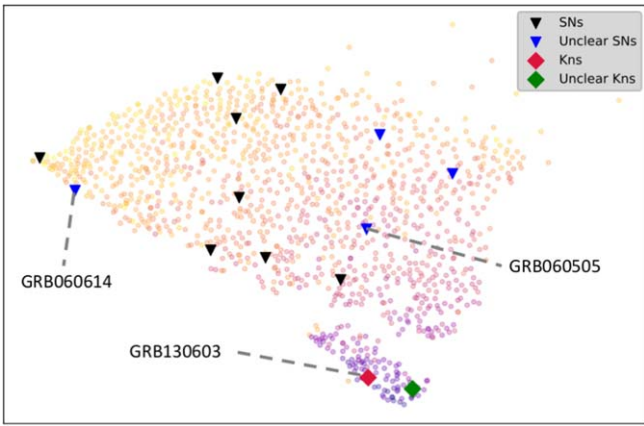7 GRB171010A is a Fermi burst that was followed up by Swift.

Figure 4. (Top) The locations of several GRBs that have been the focus of recent studies are indicated on the t-SNE mapping used throughout. Note that a complete list of the t-SNE classification into long and short GRBs is available in Table 1. (Bottom) Classification of GRBs using t-SNE run at a perplexity of 5, lower than the map shown in the other figures. The lower perplexity is chosen to emphasize substructure.

4. GRBs of Special Interest

The clear separation on the t-SNE map into two groups, namely type-S and type-L, provides labels for all GRBs, including ones that were previously difficult to classify.

In addition to providing this broad classification, there is considerable additional information provided by the t-SNE map. Because each burst is placed close to other, similar bursts, it is possible to determine both whether an object is typical (many neighbors) or atypical (fewer neighbors, and often on the edges of a cluster). If some bursts have additional information available such as an observed afterglow or host galaxy, it may be natural to ascribe similar properties to neighboring bursts for which those observations are unavailable.

A full discussion of the various groups indicated is beyond the scope of this paper. However, here we briefly describe the t-SNE classification of several GRBs that have previously been the subject of debate, as different analysis techniques have disagreed on their proper categorization.

4.1. Supernova-less Long GRBs

GRB 060505 and GRB 060614 are of special interest, since they are long-duration GRBs with no observed optical

counterpart (Della Valle et al. 2006; Fynbo et al. 2006). They have therefore been suggested to have their own progenitor mechanism, constituting their own class of Extended Emission Short Gamma Ray Bursts (EE sGRB; Norris & Bonnell 2006; Ofek et al. 2007). The t-SNE map classifies GRB 060614 as a type-L burst. It lies close to GRB 111209A, which has an associated luminous SN (Kann et al. 2019), and in the same cluster as GRB 111209A. This suggests that GRB 060614 would not have a separate progenitor mechanism, but is instead a more standard long GRB.

GRB 060505 lies far away from the cluster of type-S GRBs and clusters with type-L GRBs. This indicates that GRB 060505 was a long GRB. However, it does not group cleanly with the confirmed SNe, suggesting there may have been a different physical origin.

4.2. Short GRB Association with Kilonovae

GRB 130603B is a short GRB with a kilonova (KN) counterpart (Tanvir et al. 2013). GRB 130603B lies in the type-S GRB group close to GRB 160821B, which has also been theorized to have had an associated KN (Lamb et al. 2019; Troja et al. 2019). This suggests that GRB 160821B may indeed be associated with a KN. Another proposed KN candidate is GRB 050509B, which is unfortunately removed from our catalog due to having a duration of 1 time bin or less, making the DTFT meaningless.

Since only one burst with an optical KN counterpart is included in the catalog, extrapolation is not very useful. As more KN are discovered, this method will become an increasingly important feature of the t-SNE map. GRB170817A, the burst associated with GW170817 is not a Swift burst and thus is not included in our analysis (Abbott et al. 2017; Goldstein et al. 2017).

5. Discussion

The application of the dimensionality reduction algorithm t-SNE to GRB light curves observed by Swift is used to group bursts based upon similarities with their neighbors. The resulting map (Figure 2(a)) suggests that bursts should be classified into two broad groups. An analysis of the duration distribution, optical counterparts, and previous proposed classifications strongly suggests that these two groups correspond to short and long GRBs. If so, this technique would provide a complete, unambiguous separation of the GRB population into short and long bursts. A natural next step is to determine which properties best correlate with classification and use this to develop astrophysical models, although such a study is beyond the scope of this Letter.

The t-SNE algorithm can be tuned to instead focus on substructure, and it is possible that further subgroups exist within these two populations (Figure 4). Substructure might be associated with distinctions between bursts belonging to these groups, perhaps indicating that a long GRB could be produced by multiple physical causes. Because relatively few bursts have known counterparts or other additional information, it is difficult to connect these groups with physical origins. However, the few bursts known to have a common origin, such as supernovae, are indeed mapped to nearby locations. Similarly, the confirmed kilonova GRB 130603B is a near neighbor of GRB 160821B (Lamb et al. 2019), proposed as a kilonova candidate.

If these substructures indeed correspond with physical origin, then the t-SNE map can be used to select bursts for targeted follow-up. For example, it would be straightforward to determine whether a new burst is a kilonova candidate based upon whether it is placed close to GRB 130603B.

With small modifications to avoid recalculating the entire map for a new object, t-SNE could also easily be used in an automatic classification pipeline along with a suitable clustering algorithm. However, GRBs are sufficiently rare to warrant individual human attention, with good choice of parameters, a t-SNE map will have sufficient separation that human visual classification is both possible and likely preferable.

5.1. Implications for Anomalous GRBs

This study demonstrates that GRB 060614, long-debated as a possible extended emission short GRB, is indeed a long-duration type GRB. Its lack of a supernova down to 5 mag below SN1998bw presents a real challenge to standard jet-driven supernova models and suggests a direct-collapse black hole, as originally postulated by Fynbo et al. (2006). The same is true for GRB 060505, which was at least 6 mag fainter than SN1998bw and is found unequivocally to belong to the long-duration category. The most compelling interpretation would be that GRB 060505 and GRB 060614 were direct-collapse black holes (Adams et al. 2017; Liu et al. 2019), a possibility that has been theorized but not yet observed.

The methods developed in this work can be used to cleanly divide all Swift light curves into two distinct classes. Based upon the duration distribution and other properties it is tempting to label type-S and type-L GRBs as short and long. However, this classification is entirely empirical, based upon the distribution of light-curve properties rather than on astrophysics. One of the drawbacks of t-SNE is that it is not immediately clear which properties of the full light curves have been most influential in this classification or how those properties might relate to underlying GRB astrophysics. Whether these indeed have distinct astrophysical origins, perhaps one arising from collapse and the other from collision, will therefore require additional follow-up studies of each group. It is hoped that this process will be far more straightforward now that every GRB, even at duration $\sim$2 s, can be unambiguously classified.

ORCID iDs

Christian K. Jespersen https://orcid.org/0000-0002-8896-6496
Johann B. Severin https://orcid.org/0000-0002-4942-562X
Charles L. Steinhardt https://orcid.org/0000-0003-3780-6801
Jonas Vinther https://orcid.org/0000-0002-4418-752X
Johan P. U. Fynbo https://orcid.org/0000-0002-8149-8298
Jonatan Selsing https://orcid.org/0000-0001-9058-3892
Darach Watson https://orcid.org/0000-0002-4465-8264

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, ApJL, 848, L12
Adams, S. M., Kochanek, C. S., Gerke, J. R., Stanek, K. Z., & Dai, X. 2017, MNRAS, 468, 4968
Berger, E., Fong, W., & Chornock, R. 2013, ApJL, 774, L23
Bromberg, O., Nakar, E., & Piran, T. 2011, ApJL, 739, L55
Bromberg, O., Nakar, E., Piran, T., & Sari, R. 2013, ApJ, 764, 179
Bufano, F., Pian, E., Sollerman, J., et al. 2012, ApJ, 753, 67
Cano, Z., de Ugarte Postigo, A., Pozanenko, A., et al. 2014, A&A, 568, A19
Cano, Z., Wang, S.-Q., Dai, Z.-G., & Wu, X.-F. 2017, AdAst, 2017, 8929054
Della Valle, M., Chincarini, G., Panagia, N., et al. 2006, Natur, 444, 1050
Deng, L. 2012, ISPM, 29, 141
Fruchter, A. S., Levan, A. J., Strolger, L., et al. 2006, Natur, 441, 463
Fynbo, J. P. U., Watson, D., Thöne, C. C., et al. 2006, Natur, 444, 1047
Ghirlanda, G., Nappo, F., Ghisellini, G., et al. 2018, A&A, 609, A112
Ghirlanda, G., Nava, L., Ghisellini, G., Celotti, A., & Firmani, C. 2009, A&A, 496, 585
Goldstein, A., Veres, P., Burns, E., et al. 2017, ApJL, 848, L14
Hjorth, J., & Bloom, J. S. 2012, The Gamma-Ray Burst—Supernova Connection (Cambridge: Cambridge Univ. Press), 169
Hjorth, J., Sollerman, J., Møller, P., et al. 2003, Natur, 423, 847
Kann, D. A., Schady, P., Olivares E., F., et al. 2019, A&A, 624, A143
Klose, S., Schmidl, S., Kann, D. A., et al. 2019, A&A, 622, A138
Kouveliotou, C., Meegan, C. A., Fishman, G. J., et al. 1993, ApJL, 413, L101
Lamb, G. P., Tanvir, N. R., Levan, A. J., et al. 2019, ApJ, 883, 48
LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, IEEE, 86, 2278
Lien, A., Sakamoto, T., Barthelmy, S. D., et al. 2016, ApJ, 829, 7
Liu, J., Zhang, H., Howard, A. W., et al. 2019, Natur, 575, 618
Maaten, L. v. d., & Hinton, G. 2008, Journal of Machine Learning Research, 9, 2579
Nakar, E. 2007, PhR, 442, 166
Norris, J. P., & Bonnell, J. T. 2006, ApJ, 643, 266
Norris, J. P., Share, G. H., Messina, D. C., et al. 1986, ApJ, 301, 213
Ofek, E. O., Cenko, S. B., Gal-Yam, A., et al. 2007, ApJ, 662, 1129
Paciesas, W. S., Meegan, C. A., Pendleton, G. N., et al. 1999, ApJS, 122, 465
Sacchi, W., & Ulrych, T. J. 1998, ITSP, 46, 31
Shen, Y., & Wang, A. 2006, ApOpt, 45, 1102
Stanek, K. Z., Matheson, T., Garnavich, P. M., et al. 2003, ApJL, 591, L17
Steinhardt, C. L., Weaver, J. R., Maxfield, J., et al. 2020, ApJ, 891, 136
Tanvir, N. R., Levan, A. J., Fruchter, A. S., et al. 2013, Natur, 500, 547
Tavani, M., Kniffen, D., Mattox, J., Paredes, J., & Foster, R. 1998, ApJL, 497, L89
Troja, E., Castro-Tirado, A. J., Becerra González, J., et al. 2019, MNRAS, 489, 2104
van der Maaten, L. 2014, Journal of Machine Learning Research, 15, 3221
Zhang, F.-W., Shao, L., Yan, J.-Z., & Wei, D.-M. 2012, ApJ, 750, 88