

An Intertwined Short and Long GRB with 4-minute Separation

Corresponding Author: Professor Liang Li

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Astronomy.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Comments for the Author)

This manuscript has improved significantly from the previous version, and I thank the authors for their careful review of all referee comments and modifications to the paper. I have only several minor comments left for the the authors. I am happy to recommend the paper for publication once these are addressed.

“Previously reported quiescent GRBs also differ from GRB 160425A: precursor-main bursts, main-extended-emission bursts, and double-trigger long GRBs generally preserve a common episode-level classification, while short-GRB extended emission, late flares, and afterglow-onset cases are not ordinary long-like prompt emission episodes (see Table 3; and Methods).”

This sentence is a little confusing at present. I would separate it into 2-3 sentences, stating that GRB 160425A differs greatly from other bursts that challenge the canonical short versus long GRB division. And then stating the examples (perhaps with a little bit more description for what each of these sub-types are) you have would greatly improve readability.

Host Galaxy Section

This section has greatly improved from the previous version, and I appreciate the authors' attentiveness to my comments. I do think it gets quite repetitive, especially since the authors state the reasoning and description of the same statistical tests for every host property. I would just state this once, and say that this is how you will compare all host properties of GRB 160425A to short and long GRBs.

Another minor comment on this section - I would replace all the “It turns out that” to “We determine...” or “We find...”, since “it turns out...” is too colloquial.

(Remarks on code availability)

Reviewer #2

(Comments for the Author)

This manuscript presents the analysis of the prompt emission and afterglow of GRB 160425A and its host galaxy. This event presents two prompt emission episodes (G1 and G2) separated by a long quiescent period. By every tested metric G1 and G2 would be individually classified as a short GRB/merger origin and a long GRB/collapsar origin, demonstrating that GRB 160425A defies current classification schema. The manuscript further demonstrates that the afterglow and host galaxy properties of GRB 160425A also do not provide obvious clues to the progenitor of GRB 160425A. The authors have thoroughly addressed my previous comments and I recommend the manuscript for publication in Nature Astronomy. I have left a few minor comments below, but do not need to see the revised manuscript.

1) Please make sure to be consistent between using IGRB/LGRB and sGRB/SGRB, as the capitalization of the first letter

changes throughout the manuscript and in the tables and figures

2) I still think a dark border around the stars for GRB 160425A would increase the visibility on some of the figures, but this is up to the author's discretion.

3) Figure 1 panel c: The zorder for G2 should be higher, it is currently hidden under the grey IGRB points. Additionally, on the right side the histogram should probably match the color of the SGRB points on the left side (currently left side points for SGRBs are blue but the histogram for SGRBs is red, which is also the color as G1 on the left side).

4) Host Galaxy Observations and Properties: "We constructed a complete catalog of GRB normalized offset (R_{off}) using data from" - I believe the parentheses should say " (r_{off}) " or even better " $(r_{\text{off}} = R_{\text{off}}/R_{50})$ ". Similarly in the caption of Extended Data Figure 6 it should say " r_{off} " not just " r "

5) "GRB 160425A in Synchrotron Afterglow Emissions of Swift GRBs" - This section goes above and beyond in addressing my initial concerns and I am very impressed with the effort of uniformly modeling a large sample of GRBs. Extended Data Figure 3 is great. I have two minor comments on this section:

a) It would be good to specify if you modeled only the X-ray data of the comparison sample (which it sounds like you did), or whether you incorporated optical/radio data. You should also mention how that could be a limitation to the comparison to GRB 160425A. In the response it was mentioned that the fit for GRB 160425A was also done for only the X-ray data, so this could easily be addressed with no additional analysis.

b) No mention of the distribution of n_0 values is provided in the initial results paragraph (it is mentioned in the comparison paragraph), and whether there are differences between sGRBs or IGRBs.

Unrelated to this publication, but a more thorough investigation of the results of this comparison effort could easily warrant a separate paper that many would likely find useful.

(Remarks on code availability)

Reviewer #3

(Comments for the Author)

This manuscript makes a claim for a unique classification of a GRB. The main claim of the paper regards the conflicting classification of sub-bursts G1 and G2. The former appearing as a short-hard "merger-like" burst and the latter as a long-soft "collapsar-like" burst. There is a long ~ 250 s gap between the two sub-bursts. While this claim is explored in detail and may have useful implications for the classification of GRBs in general, I remain unconvinced that the paper merits publication in Nature Astronomy.

My main concern is that the two main features discussed: (i) a long quiescent gap mid-burst and (ii) the classification of the initial sub-burst as "merger-like", have both been previously reported on in the literature. Moreover, other bursts with somewhat ambiguous classification have been observed over the years. Finally, the evidence itself in favor of distinct classifications of G1 and G2 is not sufficiently robust.

I provide additional, more detailed comments below.

Detailed comments:

Several of the GRBs with long quiescent gaps shown for example in figure 1C have hard short isolated spikes (e.g. 060124A, Romano et al. 2006).

The main characteristics of the host galaxy lean more towards collapsar like than towards merger like (e.g. extended table 4).

E_p is one of the more differentiating properties between merger and collapsar like GRBs. However, E_p is only constrained from below by the BAT analysis. Moreover, the authors use their 'bending-point procedure' technique to make a very strong claim, that E_p of G1 is greater than 500keV, even though the spectral range of BAT only extends up to 150keV. A lot lies on this one claim (as shown in Fig. 5e, an E_p of 150keV would make G1 appear very much 'collapsar-like') which infers a lower limit on E_p substantially above the upper band of BAT.

The rejection of a precursor origin of G1 seems overstated. Bright precursors that outshine the main pulse are rare but have been seen before (e.g. Burlon et al. 2008).

The inferences from afterglow modeling are reported on in the main text as directly inferred parameters (e.g. values of density, Lorentz factor, ϵ_e) and uses their 'unusual values' as supporting evidence. In reality, there is usually much degeneracy in inferring those parameters from afterglow observations (as also shown in the methods section where errors are quoted), which should be addressed explicitly.

The chance coincidence probability of 10^{-12} is misleading. It doesn't account for the 'number of trials'. It calculates the probability of an event occurring at a specified time and position, instead of the probability of some such coincidence ever being detected. With thousands of BAT triggers over Swift's lifetime and many possible gap durations that would qualify as "suspicious," the trials factor is large.

The spectral lag of G1 is marginal and not strongly discriminating between merger and collapsar like.

An important classification of merger vs collapsar like GRBs is not discussed. This is the ratio of X-ray to gamma-ray fluence vs time. Merger like bursts tend to have a higher such ratio (e.g. Nysewander et al. 2009) implying a lower prompt efficiency. Considering this ratio for the G1 and G2 components separately can provide an independent test of classification (many of the currently adopted classifiers are not truly independent from each other, they essentially derive from a few main prompt observables such as duration and spectral peak – the fluence ratio mentioned above combines prompt with afterglow physics) and also directly test for the possibility of two afterglows underlying the bursts as opposed to one as in the fit performed by the authors.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #3

(Comments for the Author)

I appreciate the substantive work done by the authors to address my concerns from the previous report. In particular, re-analyzing GRBs 060124A, 051219A, 091024A, 110709B and 220627A and showing that the first episode in all cases is “long-soft” when treated individually, is a very useful test. There are still a few remaining points that I would like to see addressed before I can recommend this for publication in Nature Astronomy.

First, the estimation of physical parameters from the afterglow analysis. The authors mention testing different codes as a test of robustness. However, it is clear even from the quoted results, that there is much degeneracy. For instance, n_0 spans about two orders of magnitude. In addition, the statement that $\epsilon_e=0.62$ is unusually high, is undermined by the Gaussian jet model analysis of the authors, that finds a much more typical $\epsilon_e=0.28$. For comparison see typical values inferred for a wide GRB population, e.g. as in 2017MNRAS.472.3161B. Note that the larger ϵ_e value reported in the manuscript is not unique to this work and is commonly found by Bayesian population analysis (see 2022MNRAS.511.2848A). Considering these degeneracies, I suggest to significantly soften the wording and perhaps give some relevant context regarding the extreme nature of the underlying parameters.

The added joint-classification table is useful to addressing the robustness of the distinct G1 / G2 classifications. However, I think that the estimated probabilities are unrealistic. Consider for example the Norris row for G1. The 10^{-12} probability relies on a Gaussian tail extrapolation of log-normal fits to populations that overlap at the $\sim 10\%$ level. No individual burst can be classified at 10^{-27} confidence when the parent distributions overlap and considering the overall number of detected GRBs.

The host/afterglow association are most naturally associated with G2. Indeed, there are no afterglow detections in between G1 and G2. This means that the link between G1 and G2 relies only on the BAT position plus temporal coincidence, which in the revised paper, is estimated to have an associated probability of 10^{-5} . This is small but not negligible, and if true, this results in no challenge to progenitor models at all. I ask the authors to directly state what (if anything) links G1 to the host and to G2 other than the positional and timing consistency.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #3

(Comments for the Author)

The authors have clearly addressed my comments from the previous reports. I am happy to recommend the paper for publication.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Referee Report and Author Responses

GRB 160425A

Referee expertise

Referee #1: Multi-wavelength observations of SGRB-EEs

Referee #2: Radio follow-ups of GRB afterglow

Referee #3: Theory of GRB engines, progenitors, and post-explosion emission

Referees' comments

Referee #1 (Comments for the Author)

This manuscript has improved significantly from the previous version, and I thank the authors for their careful review of all referee comments and modifications to the paper. I have only several minor comments left for the the authors. I am happy to recommend the paper for publication once these are addressed.

Response

We appreciate the referee's careful review and positive assessment of the revised manuscript. We have considered the remaining comments and revised the manuscript accordingly. Our point-by-point responses are given below.

Comment 1.1

"Previously reported quiescent GRBs also differ from GRB 160425A: precursor-main bursts, main-extended-emission bursts, and double-trigger long GRBs generally preserve a common episode-level classification, while short-GRB extended emission, late flares, and afterglow-onset cases are not ordinary long-like prompt emission episodes (see Table 3; and Methods)."

This sentence is a little confusing at present. I would separate it into 2-3 sentences, stating that GRB 160425A differs greatly from other bursts that challenge the canonical short versus long GRB division. And then stating the examples (perhaps with a little bit more description for what each of these sub-types are) you have would greatly improve readability.

Response

The referee is right that the original sentence was too dense and that the different sub-types were not separated clearly enough. We have rewritten it as several shorter sentences and added brief descriptions of the relevant classes.

Comment 1.2

Host Galaxy Section

This section has greatly improved from the previous version, and I appreciate the authors' attentiveness to my comments. I do think it gets quite repetitive, especially since the authors state the reasoning and description of the same statistical tests for every host property. I would just state this once, and say that this is how you will compare all host properties of GRB 160425A to short and long GRBs.

Response

The Host Galaxy section did repeat the same statistical description for each host property. We have revised the section so that the method is introduced once at the beginning, followed by a statement that the same procedure is used for all host-galaxy comparisons between GRB 160425A and the short- and long-GRB samples. We now use the physical offset as the worked example; the other quantities are treated in the same way. This makes the section shorter and easier to follow.

Comment 1.3

Another minor comment on this section - I would replace all the "It turns out that" to "We determine..." or "We find...", since "it turns out..." is too colloquial.

Response

The referee is right that "it turns out that" is too colloquial for the manuscript. We have replaced this phrase throughout the text with more direct wording, such as "we find" or "we determine".

Referee #2 (Comments for the Author)

This manuscript presents the analysis of the prompt emission and afterglow of GRB 160425A and its host galaxy. This event presents two prompt emission episodes (G1 and G2) separated by a long quiescent period. By every tested metric G1 and G2 would be individually classified as a short GRB/merger origin and a long GRB/collapsar origin, demonstrating that GRB 160425A defies current classification schema. The manuscript further demonstrates that the afterglow and host galaxy properties of GRB 160425A also do not provide obvious clues to the progenitor of GRB 160425A. The authors have thoroughly addressed my previous comments and I recommend the manuscript for publication in Nature Astronomy. I have left a few minor comments below, but do not need to see the revised manuscript.

Response

We appreciate the referee's careful review and recommendation for publication in Nature Astronomy. The revised manuscript now makes the unusual classification of GRB 160425A clear: G_1 and G_2 resemble short/merger-origin and long/collapsar-origin GRBs, respectively, while the afterglow and host-galaxy properties do not point to a single obvious progenitor. We address the remaining minor comments below.

Comment 2.1

1) Please make sure to be consistent between using LGRB/LGRB and sGRB/SGRB, as the capitalization of the first letter changes throughout the manuscript and in the tables and figures

Response

We have checked the manuscript, tables, and figures, and now use the notation consistently. All instances of "lGRB" and "sGRB" have been changed to "LGRB" and "SGRB".

Comment 2.2

2) I still think a dark border around the stars for GRB 160425A would increase the visibility on some of the figures, but this is up to the author's discretion.

Response

We reconsidered the marker style and kept the current version so that the symbols remain consistent across the figures. If the editor or referee prefers the dark border, we can still make that change.

Comment 2.3

3) Figure 1 panel c: The zorder for G2 should be higher, it is currently hidden under the grey LGRB points. Additionally, on the right side the histogram should probably match the color of the SGRB points on the left side (currently left side points for SGRBs are blue but the histogram for SGRBs is red, which is also the color as G1 on the left side).

Response

We have replotted the figure to fix these issues. To make GRB 160425A easier to see, we removed GRB 041219A, whose point overlapped with GRB 160425A, and added GRB 050820A. We also changed the SGRB histogram on the right-hand side of panel c to blue, matching the SGRB points on the left and avoiding confusion with the red marker used for G_1 .

Comment 2.4

4) Host Galaxy Observations and Properties: "We constructed a complete catalog of GRB normalized offset (R_{off}) using data from" - I believe the parentheses should say " (r_{off}) " or even better " $(r_{\text{off}} = R_{\text{off}}/R_{50})$ ". Similarly in the caption of Extended Data Figure 6 it should say " r_{off} " not just " r ".

Response

The normalized offset should indeed be denoted by r_{off} , while R_{off} denotes the physical projected offset. We have corrected the notation throughout the main text and now define it at first use as $r_{\text{off}} = R_{\text{off}}/R_{50}$, where R_{50} is the host-galaxy half-light radius. We have also corrected the caption of Extended Data Figure 6 to use " r_{off} " rather than " r ".

Comment 2.5

5) "GRB 160425A in Synchrotron Afterglow Emissions of Swift GRBs" - This section goes above and beyond in addressing my initial concerns and I am very impressed with the effort of uniformly modeling a large sample of GRBs. Extended Data Figure 3 is great. I have two minor comments on this section:

a) It would be good to specify if you modeled only the X-ray data of the comparison sample (which it sounds like you did), or whether you incorporated optical/radio data. You should also mention how that could be a limitation to the comparison to GRB 160425A. In the response it was mentioned that the fit for GRB 160425A was also done for only the X-ray data, so this could easily be addressed with no additional analysis.

b) No mention of the distribution of n_0 values is provided in the initial results paragraph (it is mentioned in the comparison paragraph), and whether there are differences between sGRBs or lGRBs.

Unrelated to this publication, but a more thorough investigation of the results of this comparison effort could easily warrant a separate paper that many would likely find useful.

Response

a) We tested several public GRB afterglow codes, including Jetsimpy, VegasAfterglow, and afterglowpy. We based our fitting code on Jetsimpy because it was numerically stable for this sample and because the inferred parameters were less sensitive to the temporal coverage of the data.

For the reported comparison sample, we modeled only the X-ray afterglow data. We did not include optical or radio data because the optical observations come from many different telescopes and would require uniform reduction and calibration. This is a limitation of the comparison. As a

check, for bursts with good multiwavelength coverage, the parameters from X-ray-only fits usually differ by less than about 10% from those obtained with joint X-ray and optical fitting. We therefore regard the X-ray-only fits as adequate for the statistical comparison made here.

For GRB 160425A, we performed both X-ray-only and joint X-ray plus optical fits. The inferred physical parameters agree within the statistical uncertainties. For a uniform comparison with the full sample, the manuscript reports the X-ray-only fit for GRB 160425A.

b) We have added n_0 to the initial results paragraph and included the comparison between long and short GRBs.

We appreciate the encouragement. We will present the broader afterglow-modeling results in a separate paper, together with a comparison to prompt-emission properties. In that work, we will also test possible causal links between prompt-emission parameters and afterglow parameters.

Referee #3 (Comments for the Author)

This manuscript makes a claim for a unique classification of a GRB. The main claim of the paper regards the conflicting classification of sub-bursts G_1 and G_2 . The former appearing as a short-hard “merger-like” burst and the latter as a long-soft “collapsar-like” burst. There is a long 250 s gap between the two sub-bursts. While this claim is explored in detail and may have useful implications for the classification of GRBs in general, I remain unconvinced that the paper merits publication in Nature Astronomy.

My main concern is that the two main features discussed: (i) a long quiescent gap mid-burst and (ii) the classification of the initial sub-burst as “merger-like”, have both been previously reported on in the literature. Moreover, other bursts with somewhat ambiguous classification have been observed over the years. Finally, the evidence itself in favor of distinct classifications of G_1 and G_2 is not sufficiently robust.

I provide additional, more detailed comments below.

Detailed comments:

Response

We appreciate the referee’s careful assessment and clear statement of the main concern. The literature does contain GRBs with long quiescent intervals and an initial short-hard, merger-like episode. These events are usually discussed as short GRBs followed by late extended emission, i.e. SGRB (G_1) + extended emission (EE, G_2).

GRB 160425A does not fit this picture well. The G_2 episode appears after a quiescent interval of about four minutes, has a larger fluence than G_1 , has a comparable peak flux, and shows a strong single-pulse profile. These are not the usual properties of extended-emission tails in short GRBs, so a standard SGRB+EE interpretation is difficult.

The duration of the quiescent interval is also unusual. Short GRBs with extended emission may show a gap between the hard spike and the soft tail, but the gap is normally shorter than in GRB 160425A. In Figure 1c, we compiled more than 170 multi-episode Fermi and Swift GRBs with quiescent intervals, including both short and long GRBs. Only a few events have gaps comparable to or longer than that of GRB 160425A, for example GRB 041219A, GRB 060124, GRB 091024A, GRB 110709B, and GRB 220627A. These are collapsar-like long-duration GRBs, and their separated sub-bursts remain long-GRB-like when analysed individually.

Previous statistical studies also suggest that different emission components within one GRB, including precursors, prompt emission, extended soft emission, and late X-ray flares, often reflect repeated activity of the same central engine and share related observational properties (e.g., Hu et al. 2014, ApJ, 789, 145). In short GRBs with extended emission, the hard spike and the extended component also tend to remain short-GRB-like in diagnostics such as spectral lag; see, for example, Norris & Bonnell (2005, ApJ, 643, 266; their Table 4) and Kaneko et al. (2015, MNRAS, 452, 824; their Table 2). GRB 160425A is different: G_1 has a lag consistent with zero, whereas G_2 has a significant positive lag.

The host-galaxy offset is also closer to the long-GRB host distribution. This is not decisive and does not rule out an extreme short-GRB origin, so we treat the offset only as contextual information. Still, it does not make a standard SGRB+EE interpretation easier.

There is a related physical difficulty. A compact binary merger can produce a short GRB and, in some cases, extended emission. It is harder to explain a second prompt-like episode that is delayed by about four minutes, is more energetic than the first, and has a strong single-pulse profile. These properties are not easy to accommodate as ordinary extended emission from a short burst.

Ambiguous GRBs have been known for many years. Most of them, however, mainly expose the limitation of the traditional $T_{90} = 2$ s boundary. GRB 211211A and GRB 200826A are examples: they challenge the duration criterion, but they do not contain two well-separated prompt-like episodes with opposite classifications in the same event. We now compare GRB 160425A with the main classes of ambiguous or peculiar GRBs in Table 3 and in the Methods section “Uniqueness of GRB 160425A in the Peculiar GRB Population”.

From this comparison, we find no previously reported event with the same combination of properties: (1) a short-like prompt episode and a long-like prompt episode in one event; (2) a quiescent interval of about four minutes; and (3) standard prompt-emission diagnostics that consistently classify G_1 as short-like and G_2 as long-like. This is the specific sense in which GRB 160425A differs from earlier ambiguous events.

The referee also asks whether the evidence for different classifications of G_1 and G_2 is robust. Our conclusion does not rely on a single diagnostic. We use standard prompt-emission classifiers that have been used in GRB studies for decades: T_{90} , hardness ratio, spectral lag, minimum variability timescale, and two-parameter empirical relations such as the Amati and Norris relations. Similar diagnostics have been widely used for other peculiar events, including GRB 211211A in Troja et al. (2022, *Nature*, 612, 228; see their Extended Data Fig. 2) and GRB 200826A in Zhang et al. (2021, *Nature Astronomy*, 5, 911).

For GRB 160425A, these diagnostics point in the same direction. G_1 is short, hard, rapidly variable, and has a lag consistent with zero. G_2 is longer, softer, more slowly variable, and has a significant positive lag. Each diagnostic has limitations, but the same short-like versus long-like contrast appears in several prompt-emission observables. We therefore regard the episode-level empirical classification as robust, while still avoiding a definitive progenitor claim.

We added a quantitative joint classification test based on the comparison samples used in the classification figures. For each diagnostic, we define the reliability as the posterior support for the favored class under equal short/long priors. For an observed value x_i , we estimate the likelihoods $p(x_i|S)$ and $p(x_i|L)$ from the short- and long-GRB samples and compute

$$R_i(S) = \frac{p(x_i|S)}{p(x_i|S) + p(x_i|L)}, \quad R_i(L) = \frac{p(x_i|L)}{p(x_i|S) + p(x_i|L)}.$$

For two-dimensional diagrams, including hardness–duration, MVT–duration, Amati, and EHD, we evaluate the likelihoods in the plotted parameter space using the short- and long-GRB distributions. For diagnostics with an explicit boundary, such as $T_{90} = 2$ s, $f = 1.5$, and $\varepsilon = 0.03$, we propagate the measurement uncertainties across the boundary and calculate the probability that the value falls on the relevant side. We do not multiply all probabilities directly, because several diagnostics are correlated. For example, Amati, EHD, and ε all depend on E_p , E_{iso} , and/or T_{90} . We therefore first combine correlated diagnostics within each evidence group by averaging their log-odds, and then multiply the odds of the independent evidence groups.

Diagnostic group	$P(G_1 = \text{short})$	$P(G_2 = \text{long})$	Main information used
T_{90} duration	0.7475	$1 - 2.7 \times 10^{-8}$	$T_{90} = 2$ s boundary
Amplitude parameter	0.9998	> 0.999999999999	f and f_{eff}
Hardness–duration	0.9214	$1 - 5.9 \times 10^{-7}$	T_{90} –hardness plane
MVT–duration	0.9904	0.999994	rest-frame T_{90} –MVT plane
Spectral lag	0.8882	$1 - 1.9 \times 10^{-9}$	lag measurements in BAT bands
Energy-related diagnostics	$1 - 2.5 \times 10^{-7}$	$1 - 1.2 \times 10^{-9}$	Amati, EHD, and ε
Norris relation	$1 - 1.0 \times 10^{-12}$	0.8214	luminosity–lag relation
Combined prompt-emission result	> 0.999999999999	> 0.999999999999	group-compressed odds product

The calculation gives $\log_{10}$ odds = 26.69 for classifying G_1 as short-like and $\log_{10}$ odds = 49.31 for classifying G_2 as long-like. In one-sided Gaussian-equivalent terms, these are 10.8σ for G_1 and 14.8σ for G_2 . The joint probability for $G_1 = \text{short}$ and $G_2 = \text{long}$ is therefore $P(G_1 = \text{short}, G_2 = \text{long}) > 0.999999999999$, equivalent to about 10.8σ . We treat the host-galaxy and afterglow information separately, as contextual evidence, rather than multiplying them into this prompt-classification posterior.

The point-by-point replies follow.

Comment 3.1

Several of the GRBs with long quiescent gaps shown for example in figure 1C have hard short isolated spikes (e.g. 060124A, Romano et al. 2006).

Response

The referee is right that several GRBs in Figure 1c have long quiescent gaps and visually isolated early episodes, including GRB 041219A, GRB 110709B, GRB 060124A, GRB 091024A, and GRB 220627A. An isolated early episode, however, is not necessarily a short-hard, merger-like component.

Except for GRB 060124A, whose Konus-Wind observations of the G_2 emission reported in the literature are not publicly available, we have already discussed these comparison bursts in the previous revision, in our earlier response to the referee, and in the supplemental material (Appendix material 2.pdf).

We summarize the case-by-case comparison below.

For GRB 060124A, Romano et al. (2006) reported a precursor G_1 about 570 s before the main burst G_2 . This precursor is isolated, but it is not a short-hard spike in the usual classification sense. In their light curve, the precursor extends approximately from $t = -3$ s to $t = 13$ s, corresponding to a duration of about 16 s (see Figure 1 of Romano et al. 2006 and Figure 1 below). Our *Swift*/BAT re-analysis gives a consistent result, $T_{90} = 16.38 \pm 3.07$ s. Its hardness ratio is 0.53 ± 0.19 , which is soft under the definition used in this work. We also obtain a spectral lag of $\tau = 205 \pm 74$ ms and MVT = 1519 ± 270 ms (see Table 1 in this reply), both typical of long-GRB populations. The first episode of GRB 060124A is therefore better described as a long-soft isolated precursor, not as a short-hard spike like G_1 of GRB 160425A.

GRB 041219A is similar. Its first episode is also better interpreted as a long-GRB precursor than as a short-hard event. The duration is about 7 s, above the traditional 2 s boundary (see

Table 2 of this reply). Its fluence and E_p are also much smaller than those of the following main emission episodes, and its spectral lag is typical of long GRBs. These properties differ from G_1 of GRB 160425A, which is short, hard, sharply peaked, and nearly lag-free.

The other long-gap events highlighted in Figure 1c, GRB 091024A, GRB 110709B, and GRB 220627A, are double-trigger or double-main long GRBs. Their first emission episodes are not short-hard components. The first episode has $T_{90} \simeq 72.6$ s in GRB 091024A, $T_{90} \simeq 55.6$ s in GRB 110709B, and $T_{90} \simeq 136.7$ s in GRB 220627A. Other prompt-emission diagnostics also point to long-soft prompt emission rather than short-hard emission (see Tables 3, 4, and 5 of this reply).

The comparison events in Figure 1c show that long quiescent gaps can occur in GRBs, but they do not provide examples of a hard, short, isolated first episode comparable to G_1 of GRB 160425A.

Table 1. Individual burst properties of GRB 060124A.

Diagnostic	GRB 060124A- G_1		
	Value	Classification	Ref.
Time interval (15-350 keV) [$t_1 \sim t_2$]	[-6.94~15.59] (s)		This paper
Duration (15-350 keV) [T_{90}]	$16.38^{+3.07}_{-3.07}$ (s)	Long	This paper
Hardness ratio [$\frac{S(25-50 \text{ keV})}{S(15-25 \text{ keV})}$]	0.53 ± 0.19	Long	This paper
Minimum variability timescale (MVT)	1519 ± 270 (ms)	Long	This paper
Time lag [25-50 keV $\sim$ 15-25 keV]	(205 ± 74) (ms)	Long	This paper

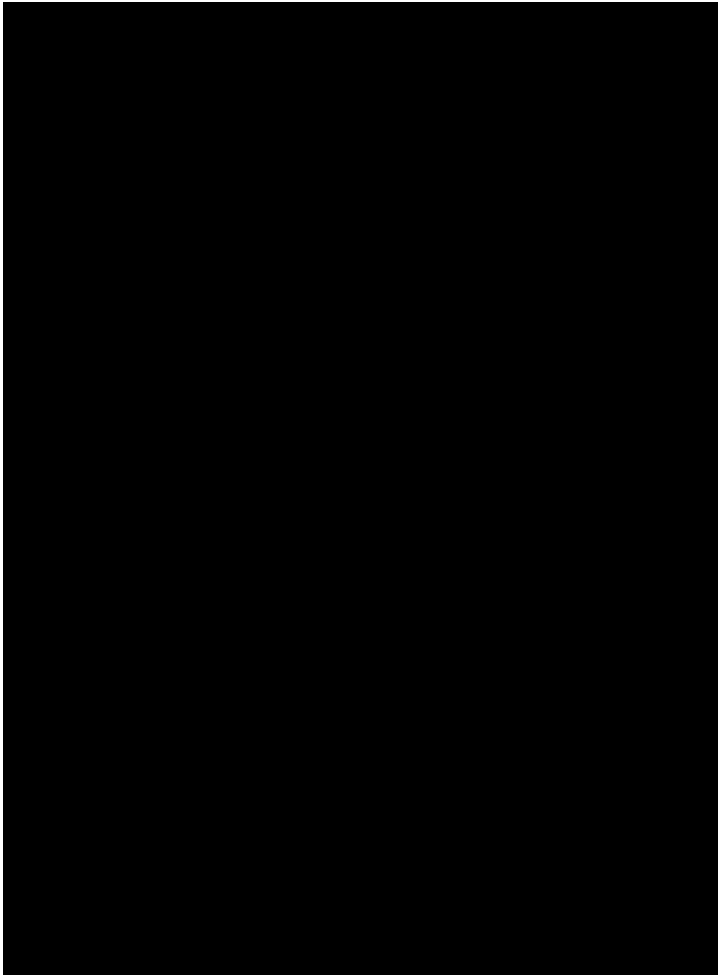


Figure 1. The light curve of GRB 060124A. Figure is taken from Romano et al. (2006).

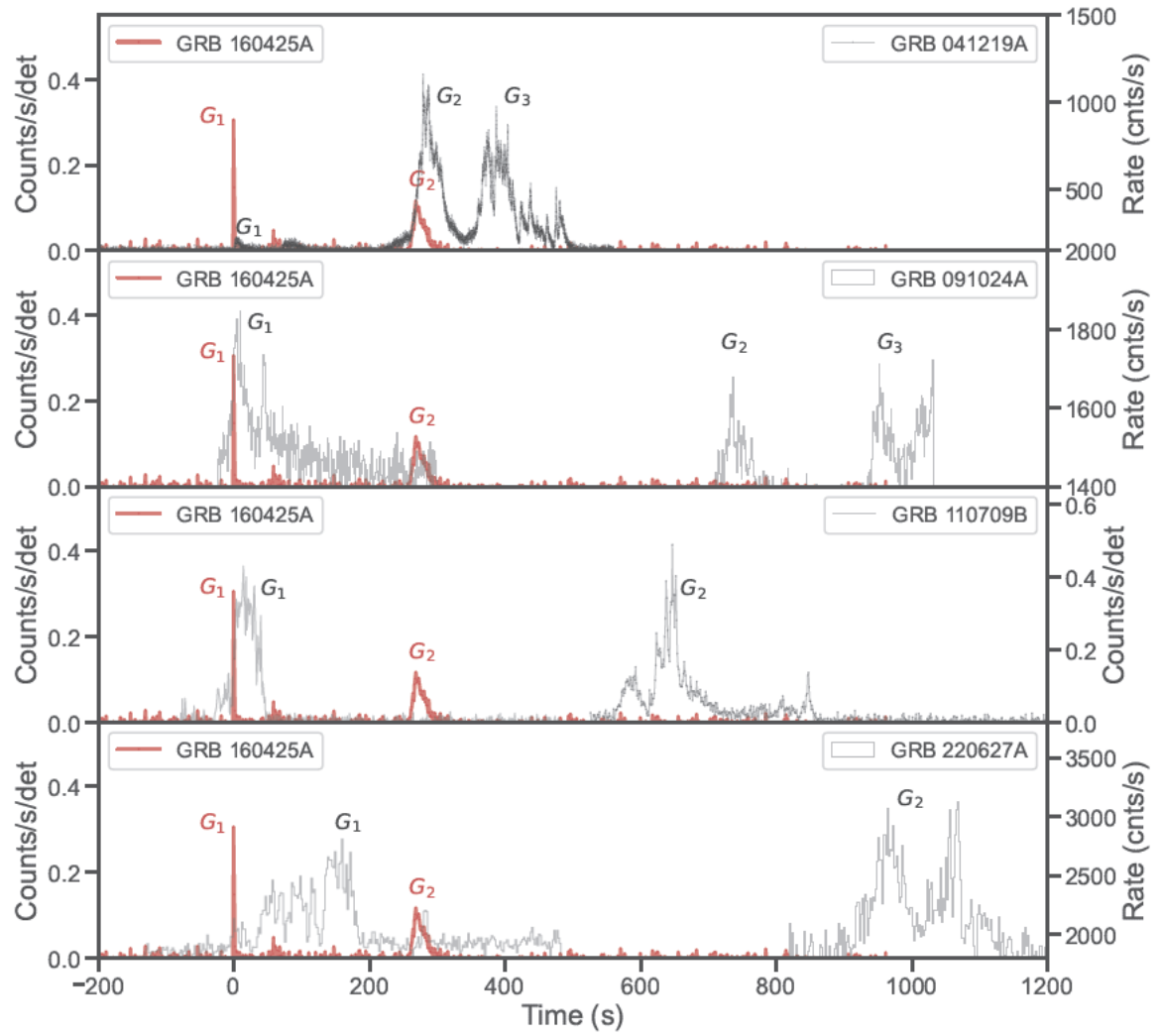


Figure 2. The light curve of GRB 160425A (red) compared with other GRBs (grey) with long quiescent gap.

Table 2. Individual burst properties of GRB 041219A.

Diagnostic	GRB 041219A- G_1		GRB 041219A- G_2		GRB 041219A- G_3	
	Value	Classification	Value	Classification	Value	Classification
Time interval (20-200 keV) [$t_1 \sim t_2$]	[0~7] s		[261~327] s		[356~414] s	
T_{90} (20-200 keV)	~ 7 s	Long	~ 66 s	Long	~ 58 s	Long
Time lag [50-100 keV $\sim$ 15-25 keV]	(1200 $\pm$ 80) ms	Long	(110 $\pm$ 10) ms	Long	(150 $\pm$ 10) ms	Long
Spectral photon index [α]	$-0.45^{+0.37}_{-0.30}$		$-1.50^{+0.08}_{-0.06}$		$-1.76^{+0.09}_{-0.08}$	
Spectral peak energy [E_p]	145.4 $^{+79.7}_{-48.1}$ (keV)		568.3 $^{+310.0}_{-205.2}$ (keV)		363.6 $^{+193.1}_{-99.5}$	
Fluence [S_γ](20 keV-8 MeV)	7.1×10^{-6} (erg cm $^{-2}$)		2.6×10^{-4} (erg cm $^{-2}$)		1.0×10^{-4} (erg cm $^{-2}$)	

Note. Data collected from McBreen et al. 2006, A&A, 455,433.

Table 3. Individual burst properties of GRB 091024A

Diagnostic	GRB 091024A- G_1		GRB 091024A- G_2		GRB 091024A- G_3	
	Value	Classification	Value	Classification	Value	Classification
Time interval (10-1000 keV) [$t_1 \sim t_2$]	[-3.8~67.8] (s)		[622.7~664.7] (s)		[838.8~1070.2] (s)	
Duration (10-1000 keV) [T_{90}]	72.6 ± 1.8 (s)	Long	44.5 ± 5.4 (s)	Long	150.0 ± 10 (s)	Long
Hardness ratio [$\frac{S(50-300\text{keV})}{S(10-50\text{keV})}$]	1.01 ± 0.20	Long	1.07 ± 0.36	Long	0.63 ± 0.25	Long
Minimum variability timescale (MVT)	1577 ± 477 (ms)	Long	6340 ± 2661 (ms)	Long	3987 ± 1009 (ms)	Long
f - parameters [F_p/F_b]	1.18 ± 0.04	Long	1.27 ± 0.04	Long	1.33 ± 0.04	Long
Time lag [100-300 keV $\sim$ 25-50 keV]	(540 ± 13) (ms)	Long	(380^{+380}_{-580}) (ms)	Long	(100^{+30}_{-130}) (ms)	Long
Spectral photon index [α]	-0.92 ± 0.07		-1.17 ± 0.07		-1.38 ± 0.02	
Spectral peak energy [E_p]	412^{+69}_{-53} (keV)		371^{+111}_{-71} (keV)		278^{+22}_{-18} (keV)	
Bolometric (1-10 ⁴ keV) fluence [S_γ] (erg cm ⁻²)	$(1.81 \pm 0.07) \times 10^{-5}$		$(0.79 \pm 0.04) \times 10^{-5}$		$(6.73 \pm 0.09) \times 10^{-5}$	
Peak luminosity (1-10 ⁴ keV) [$L_{\gamma,p}$] (erg s ⁻¹)	$(5.3 \pm 0.7) \times 10^{51}$		$(4.6 \pm 2.6) \times 10^{51}$		$(1.05 \pm 0.22) \times 10^{52}$	
Isotropic energy (1-10 ⁴ keV) [$E_{\gamma,iso}$]	$(9.0^{+0.4}_{-0.3}) \times 10^{52}$ (erg)	Long	$(5.0^{+0.4}_{-0.5}) \times 10^{52}$ (erg)	Long	$(3.1^{+0.07}_{-0.06}) \times 10^{53}$	Long
ϵ -parameter [$E_{\gamma,iso,52}/E_{p,z,2}^{5/3}$]	$0.25^{+0.07}_{-0.05}$	Long	$0.16^{+0.08}_{-0.05}$	Long	$1.65^{+0.22}_{-0.18}$	Long
EHD-parameter [$E_{p,i,2} E_{iso,51}^{-0.4} T_{90,i}^{-0.5}$]	$0.24^{+0.04}_{-0.03}$	Long	$0.35^{+0.11}_{-0.07}$	Long	$0.07^{+0.006}_{-0.005}$	Long
Amati relation [$E_{p,z} \propto E_{\gamma,iso}^{0.5}$]	On the relation	Long	On the relation	Long	On the relation	Long
Norris relation [$L_{\gamma,p,iso} \propto \tau^{-1.14}$]	On the relation	Long	On the relation	Long	On the relation	Long

Note. For more detailed information, please see Li et al. (2026), ApJS, 283, 47.

Table 4. Individual burst properties of GRB 110709B

Diagnostic	GRB 110709B- G_1		GRB 110709B- G_2	
	Value	Classification	Value	Classification
Time interval (15-350 keV) [$t_1 \sim t_2$]	[-28~55] (s)		[550.0~865.0] (s)	
Duration (15-350 keV) [T_{90}]	55.6 $\pm$ 3.2 (s)	Long	259.2 $\pm$ 8.8 (s)	Long
Hardness ratio [$\frac{S(50-100\text{keV})}{S(25-50\text{keV})}$]	0.85 $\pm$ 0.07	Long	0.77 $\pm$ 0.06	Long
Minimum variability timescale (MVT)	225 $\pm$ 65 (ms)	Long	144 $\pm$ 77 (ms)	Long
f - parameters [F_p/F_b]	1.34 $\pm$ 0.14	Long	1.39 $\pm$ 0.08	Long
Time lag [50-100 keV $\sim$ 25-50 keV]	(201 $\pm$ 52) (ms)	Long	(107 $\pm$ 77) (ms)	Long
Spectral photon index [α]	-1.09 $\pm$ 0.09		-1.12 $\pm$ 0.01	
Spectral peak energy [E_c]	182 $\pm$ 53 (keV)		116 $\pm$ 9 (keV)	
BAT (15-150 keV) fluence [S_γ]	(8.95 $^{+0.29}_{-0.62}$) $\times 10^{-6}$ (erg cm $^{-2}$)		(1.34 $^{+0.05}_{-0.07}$) $\times 10^{-6}$ (erg cm $^{-2}$)	

Note. For more detailed information, please see Li et al. (2026), ApJS, 283, 47.

Table 5. Individual burst properties of GRB 220627A

Diagnostic	GRB 220627A- G_1		GRB 220627A- G_2	
	Value	Classification	Value	Classification
Time interval (10-1000 keV) [$t_1 \sim t_2$]	[-1.28~138.99] (s)		[-8.84~135.82] (s)	
Duration (10-1000 keV) [T_{90}]	136.71 $\pm$ 1.28 (s)	Long	126.98 $\pm$ 8.84 (s)	Long
Hardness ratio [$\frac{S(50-300\text{keV})}{S(10-50\text{keV})}$]	1.67 $\pm$ 0.40	Long	0.76 $\pm$ 0.20	Long
Minimum variability timescale (MVT)	3715 $\pm$ 1841 (ms)	Long	15691 $\pm$ 14138 (ms)	Long
f - parameters [F_p/F_b]	1.47 $\pm$ 0.03	Long	1.23 $\pm$ 0.02	Long
Time lag[100-300 keV $\sim$ 25-50 keV]	(162.5 $\pm$ 82.6) (ms)	Long	(184.1 $\pm$ 91.9) (ms)	Long
Spectral photon index [α]	-0.89 $\pm$ 0.04		-0.88 $\pm$ 0.07	
Spectral peak energy [E_p]	334.08 $\pm$ 33.32 (keV)		327.936 $\pm$ 58.84 (keV)	
Fluence [S_γ](10-1000 keV)	(5.54 $\pm$ 0.01) $\times 10^{-5}$ (erg cm $^{-2}$)		(1.08 $\pm$ 0.02) $\times 10^{-5}$ (erg cm $^{-2}$)	
Peak luminosity (10-1000 keV) [$L_{\gamma,p,iso}$]	(8.98 $^{+1.15}_{-1.25}$) $\times 10^{52}$ (erg s $^{-1}$)[1-s]		(3.68 $^{+0.74}_{-1.33}$) $\times 10^{52}$ (erg s $^{-1}$)[1-s]	
Isotropic energy (1-10 ^{4} keV) [$E_{\gamma,iso}$]	(4.92 $^{+0.46}_{-0.42}$) $\times 10^{53}$ (erg)	Long	(1.78 $^{+0.13}_{-0.09}$) $\times 10^{54}$ (erg)	Long
ϵ -parameter [$E_{\gamma,iso,52}/E_{p,z,2}^{5/3}$]	2.68 $^{+0.45}_{-0.42}$	Long	1.80 $^{+0.49}_{-0.41}$	Long
EHD-parameter [$E_{p,i,2}E_{iso,51}^{-0.4}T_{90,i}^{-0.5}$]	(1.01 $^{+0.16}_{-0.14}$) $\times 10^{-1}$	Long	(4.58 $^{+0.23}_{-0.23}$) $\times 10^{-1}$	Long
Amati relation [$E_{p,z} \propto E_{\gamma,iso}^{0.5}$]	On the relation	Long	On the relation	Long
Norris relation [$L_{\gamma,p,iso} \propto \tau^{-1.14}$]	On the relation	Long	On the relation	Long

Note. For more detailed information, please see Li et al. (2026), ApJS, 283, 47.

Comment 3.2

The main characteristics of the host galaxy lean more towards collapsar like than towards merger like (e.g. extended table 4).

Response

Several host-galaxy properties of GRB 160425A, especially the projected offset, normalized offset, and fractional light, are indeed closer to the long-GRB/collapsar-like host distributions. For this reason, we do not use the host galaxy to argue for a purely merger-like origin.

Extended Data Table 4, however, does not give a clean host-galaxy classification. The redshift and SFR lie slightly closer to the short-GRB host distributions, whereas R_{off} , r_{off} , and F_{light} are closer to the long-GRB host distributions. The half-light radius R_{50} is essentially non-discriminating, since the short- and long-GRB probabilities are nearly the same. None of these host-galaxy comparisons is statistically decisive, so we do not treat the host properties alone as a firm collapsar-like or merger-like classifier.

This is why we treat the host as an indirect, population-level diagnostic. It describes the environment and the location of the afterglow within the host, but it does not probe the prompt-emission region or central-engine activity as directly as the prompt observables do. For GRB 160425A, the classification is based mainly on the episode-resolved prompt emission: duration, hardness ratio, pulse morphology, MVT, spectral lag, and empirical prompt-emission relations. These diagnostics separate G_1 and G_2 into short-like and long-like regimes.

GRB 160425A may also fall outside the standard Type I/Type II channels. If the event comes from a non-standard or dynamically linked progenitor configuration, its host properties need not match either the short-GRB or the long-GRB host population cleanly. It may resemble one class in some environmental properties and the other class in others, as seen in Extended Data Table 4.

Comment 3.3

E_p is one of the more differentiating properties between merger and collapsar like GRBs. However, E_p is only constrained from below by the BAT analysis. Moreover, the authors use their ‘bending-point procedure’ technique to make a very strong claim, that E_p of G_1 is greater than 500 keV, even though the spectral range of BAT only extends up to 150 keV. A lot lies on this one claim (as shown in Fig. 5e, an E_p of 150 keV would make G_1 appear very much ‘collapsar-like’) which infers a lower limit on E_p substantially above the upper band of BAT.

Response

E_p is a useful diagnostic for distinguishing merger-like and collapsar-like GRBs. *Swift*/BAT alone cannot directly measure a precise E_p if the spectral peak lies well above the BAT band.

The robust BAT result is that the G_1 spectrum shows no break or turnover within the BAT band. A single power law fits the BAT data, and there is no evidence for a spectral peak below the upper BAT boundary. Thus, independent of any extrapolation method, the BAT data require E_p for G_1 to be above the BAT band, namely $E_p > 150$ keV.

The value $E_p \sim 500$ keV comes from our bending-point procedure (Li & Wang 2026, ApJS, 283, 18). In brief, GRB spectra are usually described by a Band function, with two power-law segments connected around E_p . Even when the BAT-band fit is close to a single power law, weak curvature information may remain in the data. The bending-point procedure fixes E_p at a series of values above the BAT band, refits the other Band parameters, and checks where the improvement in AIC, BIC, DIC, and reduced χ^2 reaches a plateau. For G_1 , these curves bend around $E_p \sim 500$ keV. We use this as an estimated lower limit, not as a direct BAT measurement.

To address the concern about extrapolating beyond the BAT band, Extended Data Fig. 5e shows the location of G_1 in the Amati plane for three representative values: $E_p = 150$ keV, 500 keV, and 3648 keV. The most conservative case, $E_p = 150$ keV, uses only the upper edge of the BAT band and does not rely on the bending-point estimate. Even in this limiting case, G_1 lies near the boundary of the 2σ region of the short-GRB Amati relation and remains well separated from the 2σ region of the long-GRB relation. Extended Data Figure 5e, and Figure 3 in this reply, show this separation directly. The Amati comparison therefore still favors a merger-like rather than a collapsar-like classification for G_1 under the most conservative assumption.

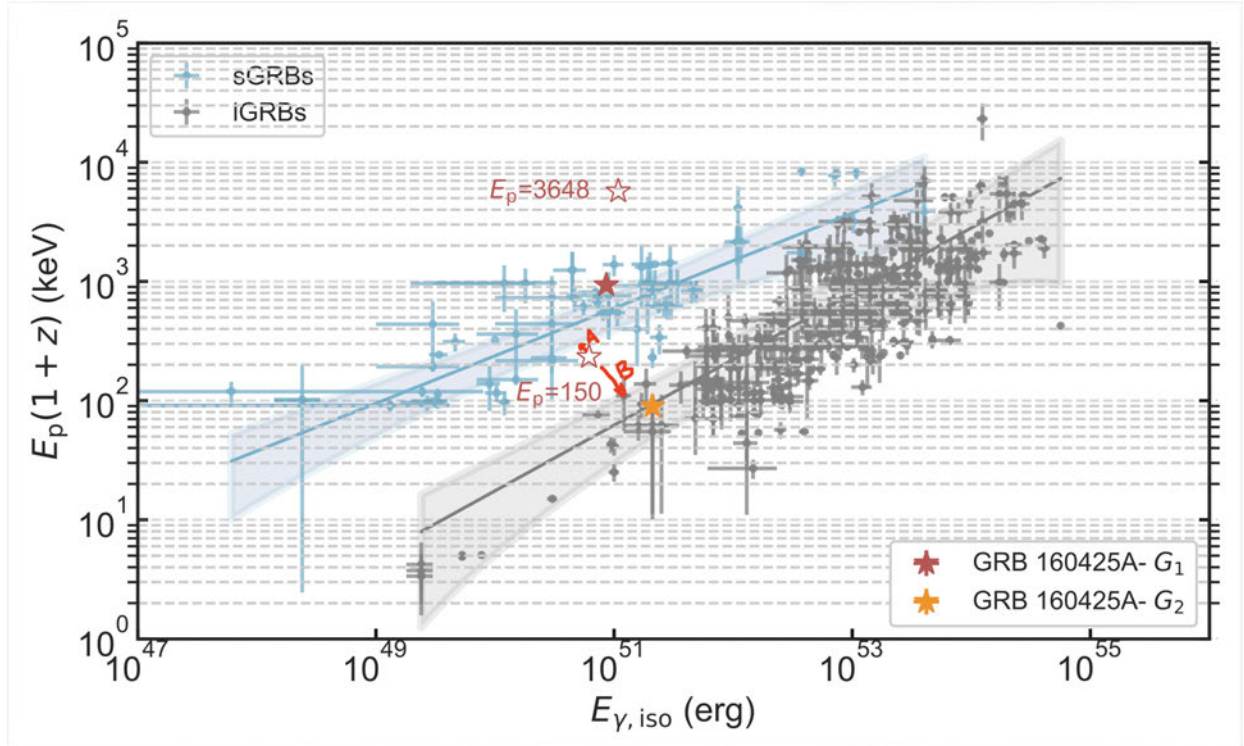


Figure 3. The location of G_1 ($=150$ keV) in the Amati correlation plane. Arrow A indicates the distance to the $2\text{-}\sigma$ boundary of the short-duration (merger-like) GRBs, while arrow B indicates the distance to the $2\text{-}\sigma$ boundary of the long-duration bursts (collapsar-like) GRBs. As shown in the figure, A is much smaller than B, which implies that even if we assume that the E_p value of G_1 is the conservative lower limit of 150 keV (i.e., the upper energy boundary of the *Swift*/BAT band), G_1 's position on the Amati correlation is more inclined towards the short burst and is consistent with the event.

Comment 3.4

The rejection of a precursor origin of G1 seems overstated. Bright precursors that outshine the main pulse are rare but have been seen before (e.g. Burlon et al. 2008).

Response

Bright or energetically significant precursors have been reported. Burlon et al. (2008), for example, found that a small fraction of precursors can have an energy comparable to, or even larger than, that of the main burst.

There are two differences between those events and GRB 160425A. In the Burlon et al. (2008) sample, some precursors are energetically significant, but their peak luminosity is still not higher than that of the main burst. In GRB 160425A, G_1 has a much higher peak flux than G_2 . This is unusual for a standard precursor-main-burst structure, although we now treat it only as supporting evidence.

For classification, the known bright precursors remain long-GRB-like in their prompt-emission properties. Several events discussed by Burlon et al. (2008), including GRB 050820A, GRB 061121, GRB 070411, GRB 071010B, GRB 070306, and GRB 061007, are included in our recent pulse-resolved study of long GRBs with precursors (Li et al. 2026, ApJS, 283, 47). In these cases, the precursor episode remains consistent with the long-GRB population when tested with episode-level diagnostics such as T_{90} , spectral lag, and MVT. Bright precursors therefore exist, but they do not show the short-like classification found for G_1 in GRB 160425A.

For GRB 160425A, the difficulty with a standard precursor interpretation comes from the combination of G_1 properties: $T_{90} = 1.7^{+0.5}_{-0.4}$ s, a hard spectrum, a sharply peaked pulse, $\text{MVT} = 18.5 \pm 3.1$ ms, and a spectral lag consistent with zero. These are the usual prompt-emission properties of short-like GRBs, not of long-GRB precursors. By contrast, G_2 is long-like in duration, MVT, and lag, and it is consistent with long-GRB empirical prompt-emission relations.

Comment 3.5

The inferences from afterglow modeling are reported on in the main text as directly inferred parameters (e.g. values of density, Lorentz factor, ϵ_e) and uses their ‘unusual values’ as supporting evidence. In reality, there is usually much degeneracy in inferring those parameters from afterglow observations (as also shown in the methods section where errors are quoted), which should be addressed explicitly.

Response

Parameter degeneracy is an intrinsic limitation of GRB afterglow modeling. The inferred physical parameters are not unique in a strict sense. We have revised the manuscript to state this more clearly and removed wording that could imply that the best-fitting values are uniquely determined.

In our analysis, the quoted uncertainties come from the posterior distributions in the Bayesian fits, so they include the parameter degeneracies within the adopted afterglow model. We also repeated the fits with different priors. The posterior distributions show the expected degeneracies, but the inferred parameter ranges remain mutually consistent within the quoted uncertainties. We further tested several public afterglow codes, including Jetsimpy, VegasAfterglow, and afterglowpy, and obtained the same overall physical interpretation.

The conclusion does not rely on the exact numerical value of any single parameter. It comes from the allowed parameter ranges and from their comparison with the uniformly modeled sample. Under the same modeling framework, GRB 160425A occupies parameter ranges that differ from

the typical comparison sample, and this statement is not removed by the degeneracies. We have clarified in the manuscript that the results should be interpreted in terms of allowed ranges rather than exact best-fitting values.

Comment 3.6

The chance coincidence probability of 10^{-12} is misleading. It doesn't account for the 'number of trials'. It calculates the probability of an event occurring at a specified time and position, instead of the probability of some such coincidence ever being detected. With thousands of BAT triggers over Swift's lifetime and many possible gap durations that would qualify as "suspicious," the trials factor is large.

Response

The referee is correct that our previous wording was too strong. The quoted value of $\sim 10^{-12}$ is not a false-alarm probability for the full *Swift*/BAT archive. It is a local probability for two independent GRB detections to fall within the chosen 250-s window and within a BAT-scale localization area. It does not include the possible choices of trigger pairs, time separations, and search windows across the mission. We have revised the manuscript to state this limitation explicitly.

A simple trials correction gives the scale of the effect. If we estimate the expected number of unrelated GRB pairs with $\Delta t \lesssim 250$ s and an angular separation comparable to the BAT localization scale over the *Swift*/BAT observing period, the probability is no longer 10^{-12} . Depending on whether we use the *Swift*/BAT trigger rate or the all-sky GRB rate, the survey-level value is of order 10^{-7} – 10^{-5} . We therefore no longer quote 10^{-12} as the probability of ever finding such a pair in the BAT data set.

This change does not affect the main conclusion, because the chance-coincidence estimate is not the basis of the classification. The host association comes from the fading X-ray/optical afterglow and its position on the $z = 0.555$ host galaxy. The short-like and long-like classifications of G_1 and G_2 come from the episode-resolved prompt-emission diagnostics. We now use the coincidence calculation only as a supporting check: a random superposition is still unlikely, but the probability argument is not treated as decisive evidence.

Comment 3.7

The spectral lag of G1 is marginal and not strongly discriminating between merger and collapsar like.

Response

The measured lag for G_1 of GRB 160425A is 38 ± 39 ms. The central value is small, and the uncertainty is comparable to the value itself. Statistically, the lag of G_1 is therefore consistent with zero.

This is the behaviour expected for short-duration, merger-like GRBs. Large short-GRB samples usually show spectral lags consistent with zero rather than significant positive lags. The lag of G_1 is therefore not a strong stand-alone proof of a merger-like origin, but it is consistent with the short-like classification. By contrast, G_2 has a much larger and statistically significant positive lag

of 1152 ± 264 ms, typical of long-duration GRBs.

We also use the Norris lag-luminosity relation, which combines spectral lag with peak luminosity. In the $(L_{\gamma, \text{p, iso}} - \tau)$ plane, G_2 follows the long-GRB lag-luminosity relation, whereas G_1 deviates from it and lies in the short-GRB region (see Extended Data Figure 1b). This two-parameter comparison gives the same interpretation as the direct lag measurement: G_1 is short-like and G_2 is long-like.

Comment 3.8

An important classification of merger vs collapsar like GRBs is not discussed. This is the ratio of X-ray to gamma-ray fluence vs time. Merger like bursts tend to have a higher such ratio (e.g. Nysewander et al. 2009) implying a lower prompt efficiency. Considering this ratio for the G_1 and G_2 components separately can provide an independent test of classification (many of the currently adopted classifiers are not truly independent from each other, they essentially derive from a few main prompt observables such as duration and spectral peak – the fluence ratio mentioned above combines prompt with afterglow physics) and also directly test for the possibility of two afterglows underlying the bursts as opposed to one as in the fit performed by the authors.

Response

The ratio between prompt gamma-ray fluence and X-ray afterglow brightness can provide an additional check. It combines prompt-emission and afterglow information and is not identical to the prompt-only diagnostics used in the main classification analysis.

We re-examined Nysewander et al. (2009). That paper compares the optical and X-ray afterglow brightness at 11 hr with the prompt gamma-ray fluence or $E_{\gamma, \text{iso}}$, and also discusses the optical-to-X-ray ratio F_R/F_X (see their Figure 8; see also Figure 4 in this reply). We did not find a direct diagnostic in the form of an X-ray-to-gamma-ray fluence ratio as a function of time, nor a clear conclusion that merger-like GRBs generally have a higher F_X/F_γ ratio than collapsar-like GRBs. Nysewander et al. instead find that afterglow brightness scales with prompt fluence, with broadly similar proportionality for short and long GRBs given the large scatter. We therefore use this comparison cautiously.

We also considered whether an F_X/S_γ -type diagnostic could be applied to GRB 160425A. A component-resolved application is not possible without model assumptions. *Swift*/XRT began observing the field about 208 s after the BAT trigger, after G_1 and before or during G_2 . There is no independent X-ray afterglow measurement immediately after G_1 . The early X-ray emission also overlaps in time with the G_2 interval and the following flare/transition phase. Assigning one part of the X-ray light curve to a G_1 afterglow and another part to a G_2 afterglow would require assuming a two-afterglow model from the start.

For this reason, we do not use separate $F_X/S_{\gamma, G_1}$ and $F_X/S_{\gamma, G_2}$ ratios as classification diagnostics. These ratios would not be independent measurements of two afterglows; they would mainly reflect how the same observed X-ray emission is assigned to the two prompt episodes. We instead use the X-ray data as event-level context. The post-flare X-ray decay is well described by a single external-shock afterglow component, and the optical and X-ray data do not require a second afterglow. A weak second component cannot be excluded, but the present data do not justify fitting

two independent afterglows.

Prompt-afterglow fluence ratios are useful at the population level, but for GRB 160425A they cannot be robustly computed separately for G_1 and G_2 because the X-ray observations begin after G_1 and overlap G_2 and the early-flare phase. This limitation does not affect the main classification result, which is based on episode-resolved prompt-emission properties.

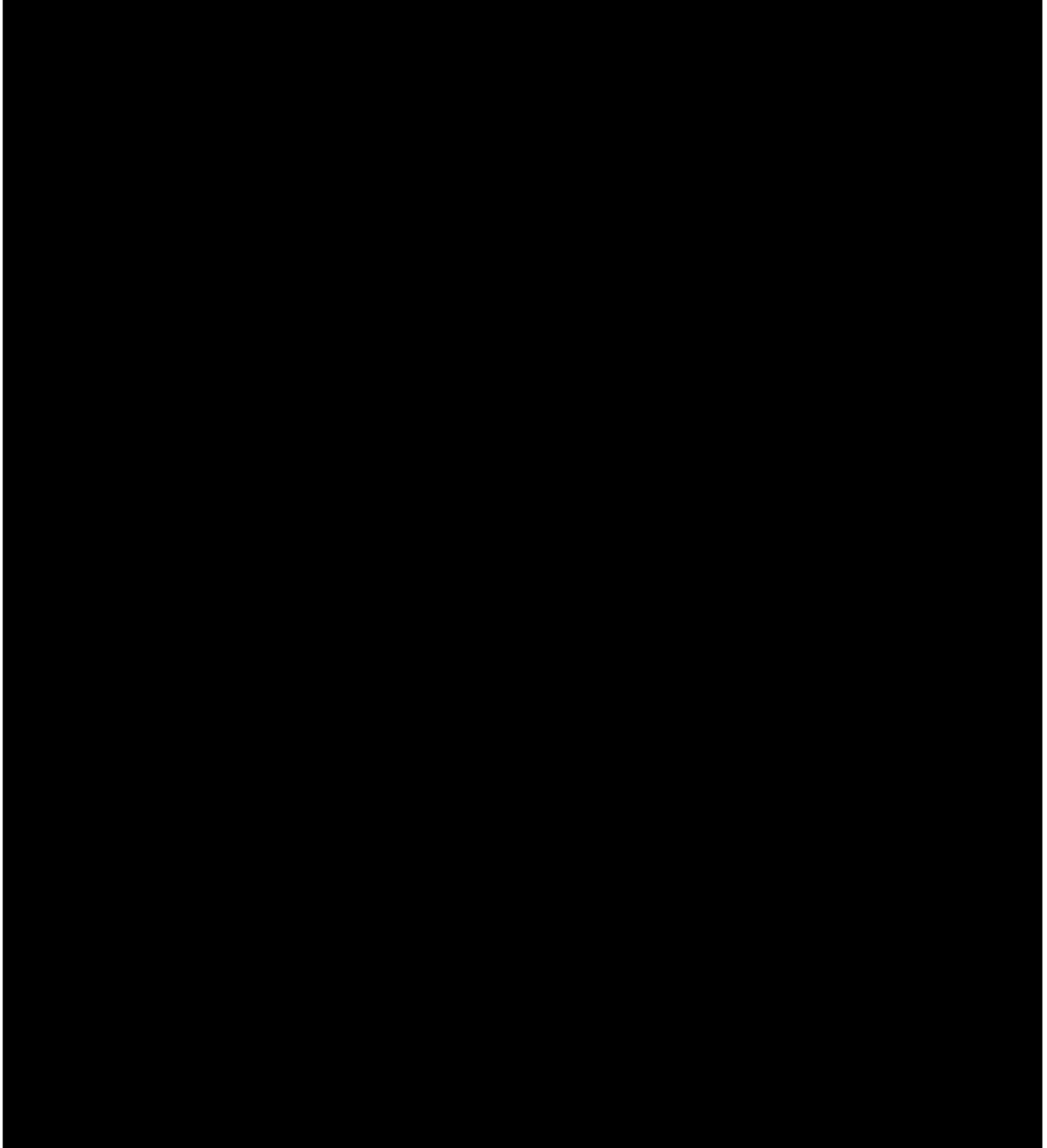


Figure 4. Histogram of the ratio of the observed fluxes at the optical and X-band after 11 hr (observed time). Figure is taken from Nysewander et al. 2009.

Referee Report and Author Responses

GRB 160425A

Referee expertise

Referee #1: Multi-wavelength observations of SGRB-EEs

Referee #2: Radio follow-ups of GRB afterglow

Referee #3: Theory of GRB engines, progenitors, and post-explosion emission

Referees' comments

Referee #3 (Comments for the Author)

Reviewer #3 (Comments for the Author):

I appreciate the substantive work done by the authors to address my concerns from the previous report. In particular, re-analyzing GRBs 060124A, 051219A, 091024A, 110709B and 220627A and showing that the first episode in all cases is “long-soft” when treated individually, is a very useful test. There are still a few remaining points that I would like to see addressed before I can recommend this for publication in Nature Astronomy.

Response

We thank the Reviewer for the careful reassessment, and we are glad the episode-resolved analysis of the comparison GRBs was useful. Each of the three remaining points has led to a concrete change in the manuscript, which we describe below.

Comment 3.1

First, the estimation of physical parameters from the afterglow analysis. The authors mention testing different codes as a test of robustness. However, it is clear even from the quoted results, that there is much degeneracy. For instance, n_0 spans about two orders of magnitude. In addition, the statement that $\epsilon_e=0.62$ is unusually high, is undermined by the Gaussian jet model analysis of the authors, that finds a much more typical $\epsilon_e=0.28$. For comparison see typical values inferred for a wide GRB population, e.g. as in 2017MNRAS.472.3161B. Note that the larger ϵ_e value reported in the manuscript is not unique to this work and is commonly found by Bayesian population analysis (see 2022MNRAS.511.2848A). Considering these degeneracies, I suggest to significantly soften the wording and perhaps give some relevant context regarding the extreme nature of the underlying parameters.

Response

We agree that the afterglow parameters are degenerate, and the two studies the Reviewer points to frame the issue well; we now cite both in the Methods. Beniamini & van der Horst (2017, MNRAS, 472, 3161) found a narrow distribution centred at $\epsilon_e \simeq 0.13\text{--}0.15$ from radio afterglow peaks. The Bayesian broadband modelling of Aksulu et al. (2022, MNRAS, 511, 2848) gave a much wider range, $\epsilon_e \sim 0.1\text{--}1$, in which values above 0.3 are common.

We have toned the manuscript down accordingly. The main text no longer calls ϵ_e , Γ_0 or n_0 unusually high, extreme or exceptionally low, and we now present these numbers for what they are: model-dependent estimates limited by model degeneracy.

Comment 3.2

The added joint-classification table is useful to addressing the robustness of the distinct G_1/G_2 classifications. However, I think that the estimated probabilities are unrealistic. Consider for example the Norris row for G_1 . The 10^{-12} probability relies on a Gaussian tail extrapolation of log-normal fits to populations that overlap at the $\sim 10\%$ level. No individual burst can be classified at 10^{-27} confidence when the parent distributions overlap and considering the overall number of detected GRBs.

Response

Following the Reviewer’s comment, we have rebuilt the classification probabilities directly from the comparison samples. No step of the calculation now extrapolates a fitted distribution beyond the range of the data.

In each classification diagram we now estimate the short- and long-GRB likelihoods with a kernel density estimate (KDE; Rosenblatt 1956, Ann. Math. Statist., 27, 832; Parzen 1962, Ann. Math. Statist., 33, 1065), using Gaussian kernels and Scott’s bandwidth rule (Scott 1992, *Multivariate Density Estimation*, Wiley) as implemented in SCIPY (Virtanen et al. 2020, Nat. Methods, 17, 261). The estimator places a kernel on each comparison burst, so the density we read at any position comes from where the bursts actually lie, not from a functional form we have imposed. Under equal priors, the support for the short-like class at a measured position x is then

$$R(S | x) = \frac{\hat{f}_S(x)}{\hat{f}_S(x) + \hat{f}_L(x)}, \quad (1)$$

where $\hat{f}_S$ and $\hat{f}_L$ are the short- and long-GRB densities. Because the kernels sit on the individual bursts, the two populations overlap in the estimate exactly as much as they overlap in the data, and that overlap goes straight into the likelihood ratio. We build the densities in the logarithmic coordinates in which each diagram is defined, from the same comparison samples shown in the figures. Spectral lag is the one exception: it is judged against a boundary rather than within a plane, so there we use the observed fractions of short and long bursts on either side of $|\tau| = 0.1$ s.

We impose two limits. A sample of N bursts cannot constrain a probability below the level of one object in $N + 1$, so we do not allow any diagnostic to exceed $N/(N + 1)$. We also wanted to know whether these classifiers work at all, so we cross-validated them (Stone 1974, J. R. Stat. Soc. B, 36, 111): each comparison burst was classified in turn from the remaining ones, and 94–100 per cent came back in their published class, diagnostic by diagnostic. That rate is the population overlap

the Reviewer refers to, measured on the samples themselves rather than assumed.

We are also careful not to count the same information twice. Amati, EHD and ε all derive from E_p and $E_{\gamma,iso}$, so we treat them as one energy group (0.99061 for G_1 , 0.97429 for G_2), and the three lag band pairs as one lag group (0.53690, 0.68389). Duration gets no group of its own, since it is already one axis of the hardness–duration plane. Within a group we average the log-odds instead of multiplying them, which is the conservative choice. The group-level probabilities are:

Diagnostic group	$P(G_1 = \text{short})$	$P(G_2 = \text{long})$
Hardness–duration	0.91248	0.99507
MVT–duration	0.99210	0.98925
Energy-related (Amati, EHD, ε)	0.99061	0.97429
Spectral lag (empirical side fractions)	0.53690	0.68389
Lag–luminosity relation	0.94118	0.87500
Combined	$1 - 3.9 \times 10^{-7}$	$1 - 9.3 \times 10^{-8}$

Combining the five groups gives $P(G_1 = \text{short}) = 1 - 3.9 \times 10^{-7}$ (4.9σ) and $P(G_2 = \text{long}) = 1 - 9.3 \times 10^{-8}$ (5.2σ). We would stress that no single diagnostic is stronger than 0.996. What produces the combined numbers is five independent lines of evidence agreeing with one another, each one held within the reach of its own comparison sample, and not the tail of any one distribution. The Methods now include a section, “Robustness of Classification”.

Comment 3.3

The host/afterglow association are most naturally associated with G_2 . Indeed, there are no afterglow detections in between G_1 and G_2 . This means that the link between G_1 and G_2 relies only on the BAT position plus temporal coincidence, which in the revised paper, is estimated to have an associated probability of 10^{-5} . This is small but not negligible, and if true, this results in no challenge to progenitor models at all. I ask the authors to directly state what (if anything) links G_1 to the host and to G_2 other than the positional and timing consistency.

Response

The direct answer: beyond the mutually consistent BAT localizations and the timing, two observations link G_1 , G_2 , and the host. First, MASTER-SAAO detected a marginal optical transient of $\sim 18.2 \pm 0.5$ mag (unfiltered) *during* the quiescent interval, in its first 30-s exposure at $T_0 + 168$ – 198 s, at a position within $0.4''$ of the enhanced XRT position of the afterglow and host (Lipunov et al., GCN 19344; Balanutsa et al., GCN 19351; Beardmore et al., GCN 19347). Second, *Swift*-XRT detected significant emission at the same position from $T_0 + 208.8$ s onward, ~ 50 s before the BAT onset of G_2 at $T_0 + 257.0$ s (Krimm et al., GCN 19343); this emission rises into G_2 and is most naturally its early development. Neither is an independent, component-resolved afterglow of G_1 . Both observations, with their caveats, are now recorded in the Methods.