# Gamma-ray bursts as probes of high-redshift Lyman-α emitters and radiative transfer models

J.-B. Vielfaure, S. D. Vergani, M. Gronke, J. Japelj, J. T. Palmerio, J. P. U. Fynbo, D. B. Malesani, B. Milvang-Jensen, R. Salvaterra, N. R. Tanvir

# Gamma-ray bursts as probes of high-redshift Lyman-$\alpha$ emitters and radiative transfer models[⋆]

J.-B. Vielfaure[1,2], S. D. Vergani[2], M. Gronke[3,⋆⋆], J. Japelj[4], J. T. Palmerio[2], J. P. U. Fynbo[5,6], D. B. Malesani[7], B. Milvang-Jensen[5,6], R. Salvaterra[8], and N. R. Tanvir[9]

[1] Université de Paris, CNRS, AstroParticule et Cosmologie, 75013 Paris, France
[2] GEPI, Observatoire de Paris, PSL University, CNRS, 5 Place Jules Janssen, 92190 Meudon, France
e-mail: jean-baptiste.vielfaure@obspm.fr
[3] Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[4] Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[5] Cosmic Dawn Center (DAWN), Copenhagen Ø, Denmark
[6] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2100 Copenhagen Ø, Denmark
[7] DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 328, 2800 Kongens Lyngby, Denmark
[8] INAF – IASF/Milano, Via Corti 12, 20133 Milano, Italy
[9] Department of Physics & Astronomy and Leicester Institute of Space & Earth Observation, University of Leicester, University Road, Leicester LE1 7RH, UK

## ABSTRACT

*Aims.* We present the updated census and statistics of Lyman-$\alpha$ emitting long gamma-ray burst host galaxies (LAE-LGRBs). We investigate the properties of a subsample of LAE-LGRBs and test the shell model that is commonly used to fit Lyman-$\alpha$ (Ly$\alpha$) emission line spectra.
*Methods.* We considered all LGRB host galaxies with relevant publicly available information. We defined a golden sample of four LAE-LGRBs (GRBs 011211, 021004, 060926, and 070110) with afterglow and host galaxy observations allowing us to constrain the shell modeling of the Ly$\alpha$ line.
*Results.* The census results in 29 detected LAE-LGRBs. We present 5 new Ly$\alpha$ emission detections in host-galaxy spectra and the corresponding unpublished VLT/X-shooter data (GRBs 060926, 070110, 081121, 081222, and 100424A). From the comparison of the statistics and properties of LAE-LGRBs to those of LAE samples in the literature, we find evidence of Ly$\alpha$ suppression in dusty systems. The fraction of LAE-LGRBs in the overall LGRB hosts is lower than that found for Lyman-break galaxy (LBG) samples at similar redshift range. This result can arise because the selection criteria of the parent samples are different and the spectral observations of LGRB samples are shallower than those of LBG. However, we find that LAE-LGRBs are representative of Ly$\alpha$ emission from the bulk of UV-selected galaxies at $z \sim 2$. We find that the golden sample of LAE-LGRBs we studied consists of complex systems characterized by multiple emission blobs and by signs of possible galaxy interactions. The fitting procedure recovers the H I column densities ($N_{HI}$) measured from the afterglow spectra and the other properties described by the shell-model parameters in the two low-$N_{HI}$ cases, but it fails to do this in the other two cases with high $N_{HI}$. The afterglows of most LGRBs and LAE-LGRBs show $\log(N_{HI}/\mathrm{cm}^{-2}) > 20.3$, implying that statistically, the bulk of Ly$\alpha$ photons that is expected to be produced by massive stars in the star-forming region hosting the GRB will be surrounded by these opaque lines of sight. We therefore interpret our results in the context of more sophisticated models and of different dominant Ly$\alpha$-emitting regions. We also compare LAE-LGRBs to LAE Lyman continuum (LyC) leakers in the literature in terms of the properties that are identified as possible indirect indicators of LyC leakage. We find that only one LGRB (GRB 021004) would likely be a strong LyC leaker and discuss the validity of these indicators at high redshift. While our work shows that LGRBs are useful tools for probing LAEs and radiative transfer models, larger statistics are required to strengthen our findings.

*Key words.* gamma-ray burst: general – galaxies: star formation – galaxies: ISM – galaxies: evolution – galaxies: high-redshift – line: profiles

## 1. Introduction

Because of its brightness and rest-frame wavelength, the Lyman-$\alpha$ (Ly$\alpha$) emission line is one of the most frequently used features for detecting high-redshift galaxies (e.g., Ouchi et al. 2009; Sobral et al. 2015; Zitrin et al. 2015; Bagley et al. 2017). The natural connection of this line with the UV emission from star-forming regions makes it an interesting proxy for studying the escape of the Lyman continuum (LyC; <912 Å). Recent studies such as those of Verhamme et al. (2015, 2017) show that this line is one of the most reliable indirect indicators of ionizing photon leakage.

To escape a galaxy, the Ly$\alpha$ photons produced in star-forming regions have to pass through the gas in which they are embedded. As this radiation resonantly scatters in the presence of neutral hydrogen (H I) and is easily absorbed by dust, the journey of Ly$\alpha$ photons in the interstellar and circumgalactic

---

medium (ISM and CGM, respectively) can be complex. Nevertheless, different properties can favor their escape, such as low HI column densities, low dust content, or suitable ISM geometries and kinematics (e.g., Kunth et al. 1998; Shapley et al. 2003; Verhamme et al. 2008; Wofford et al. 2013; Henry et al. 2015; Rivera-Thorsen et al. 2015). As a consequence, the Lyα line reflects the signatures of the physical and dynamical properties of the gas and surrounding environment of the Lyα emitters. Interpreting the observed Lyα line is complex, and radiative transfer models taking the different sources of distortion of the intrinsic profile into account are necessary. A simple and successful model that is commonly used to reproduce the Lyα shape is the shell model (e.g., Ahn 2004; Verhamme et al. 2006; Schaerer et al. 2011; Gronke et al. 2015). It consists of a homogeneous expanding shell of neutral hydrogen and dust surrounding a central emitting source.

While successful in reproducing the line profile (see, e.g., Verhamme et al. 2008; Lidman et al. 2012; Yang et al. 2017; Gronke 2017), it is important to test whether the best-fit parameter values of the shell model correspond to the real characteristics of the Lyα-emitting galaxies. Orlitová et al. (2018) independently constrained five out of the seven shell-model parameters with ancillary data for 12 Green Pea (GP) galaxies at $z \sim 0.2$. Their study shows systematic discrepancies between parameters inferred from the modeling results and the observed data. In particular, the constrained model neither reproduces the observed blue peak of the line correctly nor, in half of the cases, the red peak. For the prediction of the parameters in the unconstrained case, the main discrepant values are the redshift, the intrinsic Lyα full width half maximum ($FWHM_i(Ly\alpha)$) and the velocity expansion of the shell. Similar discrepancies for the $FWHM_i(Ly\alpha)$ were also found by Hashimoto et al. (2015) for double-peak Lyα profiles of galaxies at $z \sim 2.2$. These studies emphasize that the shell model must be used with caution to interpret the Lyα line and retrieve physical properties, such as the HI column density ($N_{HI}$), to avoid misinterpretation. This also suggests that considering a homogeneous shell to describe star-forming regions and their surrounding gas might be too simplistic.

It is rare that the information needed to constrain the model parameters is available simultaneously, especially at high redshift. Two individual studies of lensed galaxies at redshift $z = 2.7$ allowed interpreting the Lyα line using a partially constrained shell model (Schaerer & Verhamme 2008; Dessauges-Zavadsky et al. 2010). The fitting of the Lyα line agrees well with the observation by Dessauges-Zavadsky et al. (2010), while it requires different expansion velocities for the front and back of the modeled shell by Schaerer & Verhamme (2008).

Gamma-ray bursts (GRBs) can be a useful additional tool for investigating Lyα emission and testing the shell model at high redshift. GRBs are the most extreme cosmic electromagnetic phenomena (see Gehrels & Razzaque 2013 for a review). Their brightness makes them powerful probes through the cosmic history because they can be detected up to the highest redshifts (the spectroscopic record holder is GRB 090423 at $z = 8.2$; Salvaterra et al. 2009; Tanvir et al. 2009). In the case of long GRBs (LGRBs), the energy powering the bursts is released during the core-collapse of massive stars (e.g., Hjorth et al. 2003). In addition, several studies indicate that LGRBs have the tendency to occur in dwarf galaxies with a high specific star-formation rate and prefer low-metallicity environments, typically subsolar (e.g., Perley et al. 2016a; Japelj et al. 2016; Graham & Fruchter 2017; Vergani et al. 2017; Palmerio et al.

2019). This makes LGRB hosts likely representative of the common galaxies at high redshift, including during the epoch of reionization (Salvaterra et al. 2011, 2013; Tanvir et al. 2019).

The bright afterglows associated with LGRBs provide ideal background lights to probe the ISM, CGM, and intergalactic medium (IGM) along the line of sight of this population of faint galaxies systematically and at any redshift. The absorption present in the afterglow spectra directly traces the environment of the star-forming regions and also traces outflows or inflows even for the faintest objects. When the afterglow has faded, the host galaxy can be directly observed through photometry and spectroscopy. This offers the interesting possibility of combining information on the cold and warm gas with the emission properties of the GRB host galaxy (e.g., Vergani et al. 2011; Chen 2012; Friis et al. 2015; Wiseman et al. 2017; Arabsalmani et al. 2018).

We update the statistics of Lyα-emitting (LAE) LGRB host galaxies and compare their properties with those of LGRB hosts in general and of LAEs and LyC leakers in the literature. We then select a golden sample of four LAE-LGRBs at $2 < z < 3.2$ with information on the emission properties of the host galaxies and on the ISM probed by the afterglow. We use this combined information to investigate the properties of these systems and test the Lyα radiative transfer modeling. In the literature Lyα-emitting galaxies are usually defined as Lyα emitters (LAEs) when their rest-frame Lyα emission equivalent width ($EW_0(Ly\alpha)$) is above a certain threshold (typically 20 Å) because historically they were selected from narrow band observations. In this study, we qualify Lyα-emitting galaxies as LAEs independently of their $EW_0(Ly\alpha)$.

The paper is organized as follows. In Sect. 2 we present a statistical study of LAEs in LGRB hosts. We also compare LAE-LGRBs to LGRB hosts in general and compare them to LAEs. We describe the physical properties of the host galaxies of our golden sample in Sect. 3 and the Lyα radiative transfer model results in Sect. 4. In Sect. 5 we discuss the differences between model predictions and observations, and we compare LAE-LGRBs to LyC-leaker galaxies in the literature. We draw our conclusions in Sect. 6.

All errors are reported at $1\sigma$ confidence unless stated otherwise. We consider a $\Lambda$CDM cosmology with the cosmological parameters provided in Planck Collaboration XIII (2016): $H_0 = 67.8 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$, $\Omega_m = 0.308$ and $\Omega_\Lambda = 0.692$.

## 2. LAE detections in LGRB systems

### 2.1. Previous studies and approach

The early studies of high-redshift GRB host galaxies, based on the first sample of five objects, (Kulkarni et al. 1998; Fynbo et al. 2002, 2003; Møller et al. 2002; Vreeswijk et al. 2004), seemed to indicate that all GRB hosts were LAEs. A subsequent systematic study, based on the larger TOUGH sample (Hjorth et al. 2012; 69 LGRB host galaxies), was carried out by Milvang-Jensen et al. (2012). They targeted a subsample of 20 LGRB hosts in the redshift range $z = 1.8 - 4.5$ for spectral observations with the FOcal Reducer and low dispersion Spectrograph (FORS1) of the Very Large Telescope (VLT) at the European Southern Observatory (ESO; Appenzeller et al. 1998). They found 7 LGRB hosts with significant Lyα emission ($3\sigma$ detection), corresponding to 35% of LAEs.

The first step of our work is to update the census of LAE-LGRBs. To this end, we considered two approaches: (i) a determination of the LAE statistics considering spectroscopic

samples of LGRB host galaxies or afterglows, and (ii) the search for LAE-LGRBs in the literature and from the host-galaxy spectra available in the ESO archive[1]. In both cases, we considered a minimum value of $z = 1.6$ for the LGRBs. This limit corresponds to the atmospheric UV cutoff at 310 nm and is the lower-redshift limit allowing the detection of the Ly$\alpha$ line in the VLT/X-shooter spectra. In our study, the VLT/X-shooter spectrograph (Vernet et al. 2011) is particularly interesting for its wide spectral coverage (from ∼300 nm to 2500 nm), allowing the simultaneous detection of absorption lines in the ISM of the host galaxy and associated nebular emission lines for a wide range of redshifts, with a medium spectral resolution of $R \sim 5000$. The samples considered for point (i) are the TOUGH, the X-shooter host-galaxy sample (Krühler et al. 2015, XHG in the following), and the X-shooter afterglow sample (Selsing et al. 2019; XAFT in the following).

We stress that the census presented in the following sections (especially in Sect. 2.5) may not reflect the general statistics of LAEs among LGRB host galaxies. In addition to issues concerning the completeness of the samples and of the observations, the spectra are generally not homogeneous in terms of exposure times, instruments, and observing conditions. Nonetheless, it is sometimes possible to define a common flux limit, as in the case of the Milvang-Jensen et al. (2012) sample. The case of afterglow spectra is even more complex because in addition to inhomogeneous follow-up and target brightness, the Ly$\alpha$ absorption along the GRB line of sight could affect the Ly$\alpha$ emission detection and profile.

## 2.2. Data reduction

To perform the census, we reduced archival data of several X-shooter spectra of GRB host galaxies. We describe here the method applied for the data reduction. All observations of the host galaxies, the telluric stars, and the spectrophotometric standards were reduced in the same way using version 2.8.5 of the X-shooter data reduction pipeline (Modigliani et al. 2010). Before processing the spectra through the pipeline, the cosmic-ray hits and bad pixels were removed following the method of van Dokkum (2001). Then, we subtracted the bias from all raw frames and divided them by the master flat field. We traced the echelle orders and calibrated the data in spatial and wavelength units using arc-line lamps. The flux calibration was performed using spectrophotometric standards (Vernet et al. 2009), and a correction for flexure was applied. Last, the sky-subtraction and the rectification and merging of the orders was performed to obtain the final 2D spectra. Additionally, the spectra were corrected for the Milky Way (MW) extinction using the extinction curve from Pei (1992). The $A_V$ values are obtained from the NASA Extragalactic Database (NED) and correspond to the extinction map of Schlafly & Finkbeiner (2011). The wavelengths of the extracted 1D spectra were converted into the vacuum reference and corrected for the Earth's rotation and revolution around the Sun (heliocentric correction). To optimally select the extraction regions, we chose the spatial extension of the brightest emission line and applied this 1D extraction throughout the whole spectrum. When the Ly$\alpha$ emission was detected, we selected the extraction region (for the corresponding arm) according to the spatial extension of this line. This line can be larger than the Balmer lines as a result of resonant scattering. Emission line fluxes were determined by fitting a Gaussian function to the data, setting the continuum flux density in

a region close to the emission line. We also numerically integrated the flux over the line width as a comparison to control the consistency of the values and uncertainties. For the asymmetric line profile of the Ly$\alpha$ line, we use a skewed Gaussian parameterized as described in Vielfaure et al. (2020). When lines of interest were not detected, we estimated a $3\sigma$ upper limit. For upper limits of nebular emissions, we used an FWHM in agreement with other nebular lines detected in the spectrum. For the Ly$\alpha$ line, similarly to Milvang-Jensen et al. (2012), we selected the same width for all upper limits, which is $900 \,\mathrm{km \, s^{-1}}$ centered at $300 \,\mathrm{km \, s^{-1}}$. The fluxes were corrected for slit loss by calculating the flux difference between the observation of a telluric star (close in time and space to the observation of the GRB host, and with the same instrumental setup) to the tabulated values[2] expected to be measured.

## 2.3. TOUGH sample

We first focus on the TOUGH sample of 69 GRBs. The TOUGH Ly$\alpha$ study (Milvang-Jensen et al. 2012) targeted the 20 GRBs with a redshift in the range $z = 1.8-4.5$, as known at the time of the Ly$\alpha$ observing campaign. Subsequently, as part of other TOUGH campaigns (Jakobsson et al. 2012; Krühler et al. 2012) and later work (Krühler et al. 2015), the redshift completeness of TOUGH has increased to 60 of the 69 sources (Krühler et al. 2015). With respect to Milvang-Jensen et al. (2012), this includes added redshifts for 11 TOUGH GRBs in the range $z = 1.8-4.5$ (GRBs 050714B, 050819, 050915A, 051001, 060805A, 060814, 070103, 070129, 070224, 070328, and 070419B); all these hosts have VLT/X-shooter spectra.

We reduced the X-shooter data as described in Sect. 2.2 to search for the detection of Ly$\alpha$ emission. We find no LAEs among these 11 additional host galaxies. The new statistic of LAEs among LGRB host galaxies of the TOUGH sample is therefore $23\% \pm 7\%$.

Taking into account the fact that the spectra have different flux limits (see Figs. 1 and 2), we determined the statistics of Ly$\alpha$ detection above a flux cut of $1.12 \times 10^{-17} \,\mathrm{erg \, s^{-1} \, cm^{-2}}$ (corresponding to the lowest Ly$\alpha$ detection of the TOUGH sample), a luminosity cut of $5.6 \times 10^{41} \,\mathrm{erg \, s^{-1}}$ (corresponding to the flux cut at the median redshift of the TOUGH sample, $z = 2.45$), and the fraction of LAEs among the sample with a rest-frame Ly$\alpha$ $EW > 20 \,\text{Å}$. To estimate the uncertainty on these statistics, based on the sample sizes, we performed a bootstrap method employing $10^6$ random resamples with replacement of the number of LAEs among the LGRBs considered for each cut. The results are reported in Table 1.

In principle, we can search for Ly$\alpha$ emission also at $z > 4.5$. This would add three further objects from the TOUGH sample. They lack host galaxy spectral observations, but these have afterglow spectra available in the literature (GRB 050904: Totani et al. 2006; GRB 060522: Tanvir et al. 2019; GRB 060927: Fynbo et al. 2009). They show no Ly$\alpha$ emission, but formal limits have not been determined.

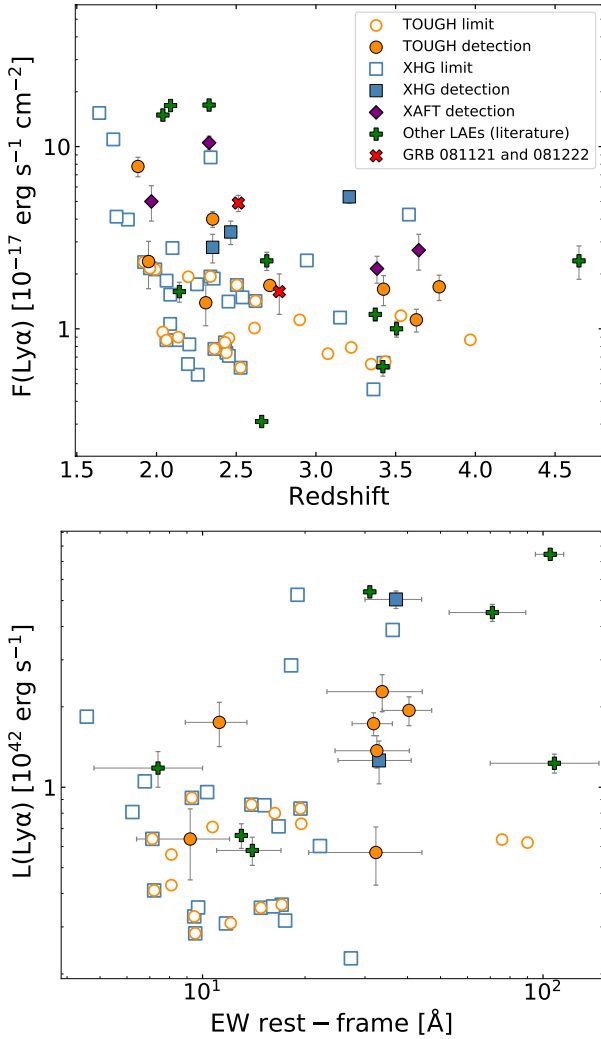## 2.4. XHG and XAFT samples

Krühler et al. (2015) presented the UVB-arm spectra for only three objects at $z > 1.6$ of their XHG sample. Therefore we

---

[2] The tabulated values of the magnitudes for the telluric stars have been taken from https://www.eso.org/sci/facilities/paranal/decommissioned/isaac/tools/spectra/Bstars.txt and http://simbad.u-strasbg.fr/simbad/
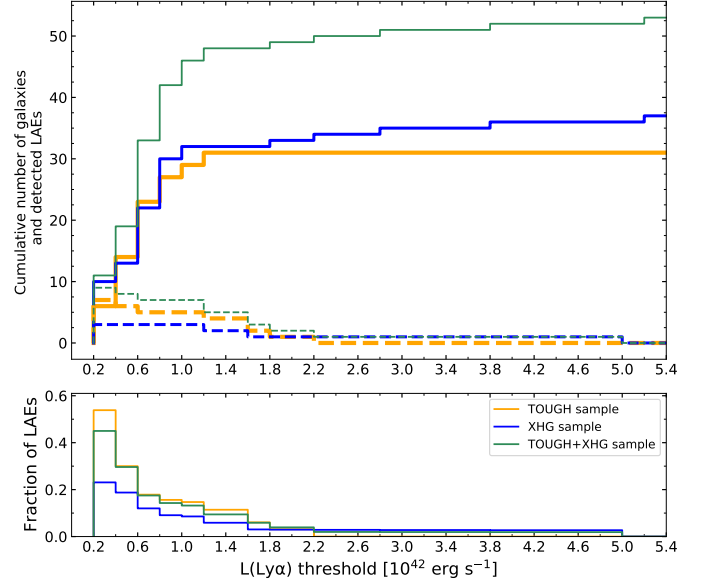
[1] http://archive.eso.org/cms.html

**Fig. 1.** Lyα emission in LGRB host galaxies. *Top panel*: Lyα fluxes of LAE-LGRBs as a function of redshift. We report the fluxes retrieved from the literature or determined in this paper (see the tables in Appendix A). For the TOUGH and XHG samples, we also plot the upper limits of the host galaxies without detected Lyα emission (empty symbols). *Bottom panel*: Lyα luminosity of the LAE-LGRBs as a function of rest-frame Lyα equivalent width ($EW_0(Ly\alpha)$). The sample and symbols are the same as in the top panel.



**Fig. 2.** Statistics of Lyα detection in LGRB host galaxy samples. *Top panel*: cumulative number of galaxies in the TOUGH (orange line), XHG (blue line), and merged samples (TOUGH+XHG, green) without overlapping GRBs as a function of Lyα luminosity detection threshold ($3\sigma$) of the spectra. The dashed lines represent the number of detected LAEs above the Lyα luminosity threshold in each sample. *Bottom panel*: fraction of detected LAEs for each sample as a function of Lyα luminosity threshold.

### 2.5. Overall detections

The results of the three samples presented above are detailed in the tables of Appendix A. By merging them and removing the overlapping cases, we find 14 LAEs out of 84 LGRB host and afterglow spectra. The statistics on the LAE-LGRBs (restricted to the host galaxy spectra) are summarized in Table 1.

To complete the census, the complementary LAE-LGRBs individually reported in the literature need to be added to this merged sample. For this aim, we selected from two GRB databases[3] that are maintained by D. Perley and J. Greiner all LGRBs detected until December 1, 2020, with a spectroscopic redshift $z > 1.6$. This resulted in 234 LGRBs (see Fig. 4), 193 of which have a spectrum covering the Lyα line based on which we can verify whether the Lyα emission is detected. For those that are not included in the sample presented in the previous sections, we searched the literature for articles or Gamma-ray Coordinates Network (GCN[4]) Circulars claiming a detection of the Lyα line. If unpublished spectra were available in the ESO archive, we retrieved and reduced them.

Overall, we found 13 additional LAE-LGRBs from the literature and 2 LAE-LGRBs from unpublished data: GRB 081121 at $z = 2.5134$ and 081222 $z = 2.770$. Both objects have X-shooter spectra (Prog. ID: 097.D-0672; PI: S.D. Vergani). We reduced the data following the procedure described in Sect. 2.2. Their Lyα emission is shown in Fig. 3. The measured Lyα fluxes of GRBs 081121 and 081222 corrected for Galactic extinction and slit loss are $F_{Ly\alpha} = (4.9 \pm 0.5) \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ and $F_{Ly\alpha} = (1.6 \pm 0.4) \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$, respectively. In addition to the Lyα line, GRB 081121 shows the [OII] doublet, [OIII]$\lambda5007$, and Hβ lines. Residual sky lines strongly contaminate the

reduced all the UVB spectra of the sample (following the procedure described in Sect. 2.2) and inspected them to search for Lyα emission. We find three LAEs among the 37 host galaxy spectra at $z > 1.6$, corresponding to $8\% \pm 6\%$ (see also Fig. 1). They are GRB 060926, GRB 070110, and GRB 100424A (also reported in Malesani et al. 2013). We focus in more detail on GRBs 060926 and 070110 in Sects. 3.3 and 3.4. The unpublished 1D and 2D spectra of the host of GRB 100424A, showing its Lyα detection, are reported in Fig. 3. The flux of the line corrected for Galactic extinction and slit loss is $F_{Ly\alpha} = (3.4 \pm 0.5) \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$. The flux limits are much less homogeneous than for the TOUGH sample. We applied the same cuts as for the TOUGH sample and summarize the statistics in Table 1.

Finally, focusing on the XAFT sample, Selsing et al. (2019) reported the detection of four LAEs (GRBs 121201A, 150915A, 151021A, and 170202A) among their X-shooter afterglow sample of 41 LGRBs. This corresponds to $10\% \pm 5\%$.

---

[3] http://www.astro.caltech.edu/grbox/grbox.php and http://www.mpe.mpg.de/~jcg/grbgen.html

[4] https://gcn.gsfc.nasa.gov/

**Table 1.** Fraction of LAEs among GRB host galaxies from the TOUGH and XHG samples.

| Sample | $f_{\text{tot}}$ | $f_{\text{f}}$ | $f_{\text{L}}$ | $f_{\text{EW}_0}$ |
|---|---|---|---|---|
| TOUGH | $23\% \pm 7\%$ (7/31) | $28\% \pm 8\%$ (7/25) | $39\% \pm 11\%$ (7/18) | $17\% \pm 7\%$ (5/30) |
| XHG | $8\% \pm 6\%$ (3/37) | $19\% \pm 13\%$ (3/16) | $19\% \pm 13\%$ (3/16) | $7\% \pm 4\%$ (2/28) |
| TOUGH + XHG | $17\% \pm 6\%$ (9/53) | $27\% \pm 9\%$ (9/33) | $36\% \pm 8\%$ (9/25) | $14\% \pm 5\%$ (6/42) |

**Notes.** The first column corresponds to the name of the sample: TOUGH, XHG, or merged samples without overlapping GRBs (TOUGH + XHG). $f_{\text{tot}}$: fraction of LAEs among the whole sample; $f_{\text{f}}$: fraction of LAEs among the sample with a flux cut of $1.12 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ $f_{\text{L}}$: fraction of LAEs among the sample with a luminosity cut of $5.6 \times 10^{41}$ erg s$^{-1}$ $f_{\text{EW}_0}$: fraction of LAEs among the sample with a rest-frame Ly$\alpha$ EW > 20 Å. Because the continuum is not detected, 7 Ly$\alpha$EW$_0$ measurements are missing for the TOUGH sample and 11 for the XHG sample. The fractions between brackets correspond to the number of LAEs over the size of the sample after applying the corresponding cuts. The uncertainty on the percentage of LAEs is calculated with a bootstrap method (see Sect. 2.3).
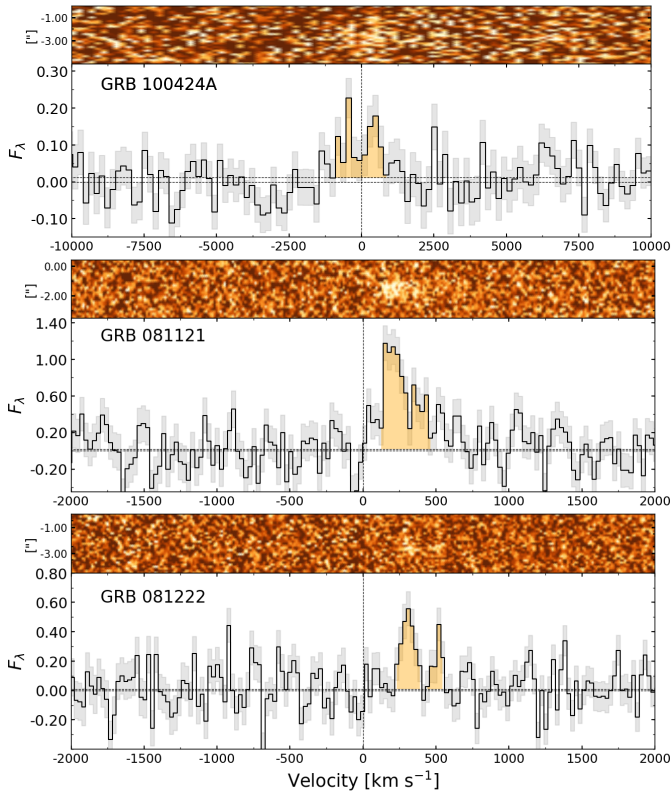


**Fig. 3.** Section of the 2D and 1D X-shooter-UVB spectra, not previously published, showing the Ly$\alpha$ emission line from the host galaxy of GRB 100424A (*top*), GRB 081121 (*middle*) and GRB 081222 (*bottom*). The lines are plotted in velocity frame centered at the systemic redshift of the galaxy (except for GRB 081222, centered on the redshift determined from the GRB afterglow absorption lines, as no other emission lines are detected). The flux density $F_\lambda$ is in units of $10^{-17}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$.

[OIII]$\lambda$4959 line and partially contaminate H$\beta$. We determine $F_{\text{[OII]}} = (1.4 \pm 0.4) \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$, $F_{\text{[OIII]}\lambda5007} = (9.0 \pm 0.4) \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$, and $F_{\text{H}\beta} \leq 1.2 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ (all corrected for Galactic extinction and slit loss).

In total, 29 LAE-LGRBs have been confirmed to date (see Figs. 1 and 4). Eleven of them have EW$_0$ > 20 Å, five are below this threshold, and the continuum flux measurement is lacking for the rest of them. Table B.1 summarizes the LAE-LGRBs and their Ly$\alpha$ fluxes. These fluxes should be considered as lower limits. It is well known (e.g., Fynbo et al. 1999; Steidel et al. 2011; Wisotzki et al. 2016; Leclercq et al. 2017) that a significant fraction of the Ly$\alpha$ emission of high-redshift galaxies is located in
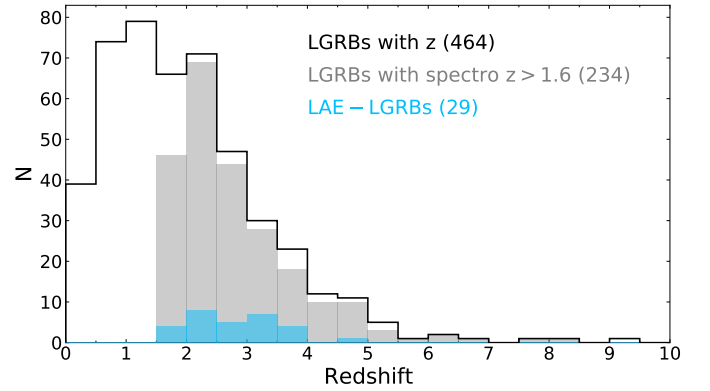


**Fig. 4.** Redshift distribution of the 29 LAE-LGRBs (blue) among the 464 LGRBs with a spectroscopic or photometric redshift measurement (black line; as of December 1 2020). The histograms are superimposed and not cumulative.

Ly$\alpha$ halos extending several kiloparsec away from the galaxy. In our measurements we considered only the flux falling in the slit, without any correction that would take the spatial distribution of the Ly$\alpha$ emission of our objects into account (it is unknown for almost all of them).

### 2.6. Comparison of LAE-LGRB properties to LGRBs that are not LAEs and LAEs that are not LGRB hosts

In principle, LGRB host galaxies are only selected based on the fact that they host an LGRB explosion. Therefore they are star-forming systems, with a young star formation, and preferentially with subsolar metallicity (see, e.g., Lyman et al. 2017; Palmerio et al. 2019). In the following, we consider LAE-LGRB statistics and compare it to LAE statistics in galaxy surveys at a similar redshift range to determine whether these systems have an enhanced or suppressed Ly$\alpha$ emission compared to the general population of star-forming galaxies. We also compare the properties of the LAE-LGRBs to those of LGRB host galaxies without Ly$\alpha$ emission (Fig. 5) to try to identify possible driving factors of Ly$\alpha$ escape. We stress that in some cases, only a small fraction of objects in our samples have available information that can be used for this comparison. We refer to the Tables in Appendix A for the values and references of the properties we discuss in the following and that we show in Fig. 5.

In the survey of galaxies at $z \sim 2$ presented by Steidel et al. (2004), 40% of galaxies with available spectra show Ly$\alpha$ emission (are LAEs) and 10% have EW$_0$(Ly$\alpha$) > 20 Å (Erb et al. 2014). Reddy et al. (2008) found 65% LAEs for Lyman-break
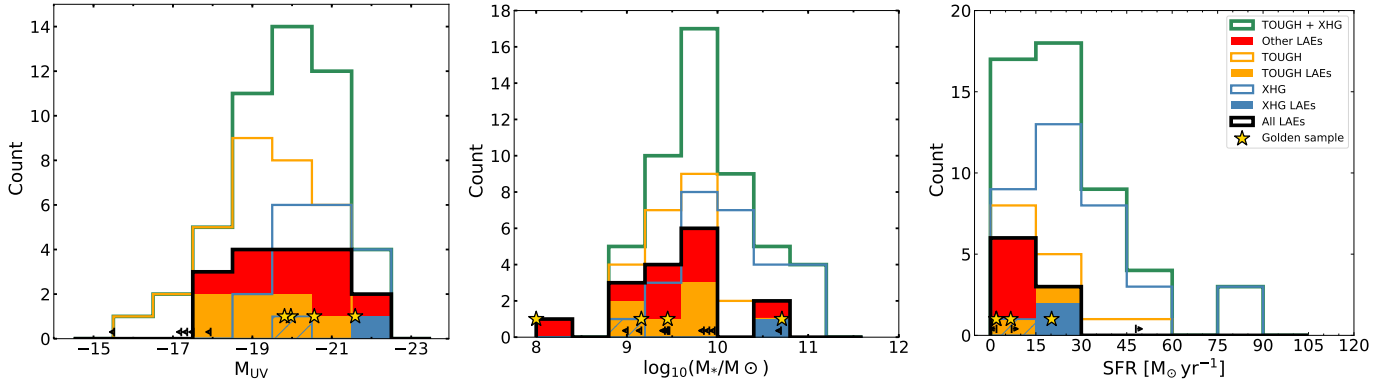
**Fig. 5.** Distribution of the $M_{UV}$ (*left panel*), stellar masses (*middle panel*), and SFR (*right panel*) of the overall LAE-LGRBs (solid black line). The colored areas inside the black histogram correspond to the LAEs detected in the TOUGH (yellow), XHG (blue), and other LAE-LGRBs outside these two samples (red). The dashed blue area shows GRB 070110, which is detected in the TOUGH and XHG samples both. The colored areas are cumulative and not superimposed. The overall TOUGH and XHG distributions are represented by the yellow and blue empty histograms, respectively. The green histogram is the sum of the TOUGH and XHG samples. The arrows represent limits, and the yellow stars are for the golden sample presented in Sect. 3.

galaxies (LBGs) at $z \sim 3$, and 23% with $EW_0(Ly\alpha) > 20$ Å. Pentericci et al. (2010) found a fraction of 50% LAEs in their selected GOOD-MUSIC LBG sample at $z \sim 3$, and 18% with $EW_0(Ly\alpha) > 20$ Å. The LAE statistics of the TOUGH and XHG samples presented in previous subsections of Sect. 2 is lower than that of those studies (see Table 1, column $f_{tot}$). However, the selection criteria of the parent sample are different and may contribute to this difference. Most of the studies listed above are based on galaxies with magnitudes brighter than $R = 25.5$ mag. This corresponds to the median value of the TOUGH sample, and about half of TOUGH LAEs are fainter galaxies. If we apply this selection to the TOUGH sample (hence reduced to 17 objects), we obtain a fraction of 22% ± 15% LAEs and 7% ± 7% with an $EW_0(Ly\alpha) > 20$ Å. These numbers should be considered as lower limits because the $Ly\alpha$ flux limit reached by the Steidel et al. (2003, 2004) spectroscopic survey is deeper than ours (typically $\sim 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ instead of $\sim 10^{-17}$ erg s$^{-1}$ cm$^{-2}$). The fraction of LAEs in the TOUGH sample could be as high as 47%, assuming that all the objects with flux limit values above that of the weakest detected $Ly\alpha$ are all LAEs. Therefore the lower statistics found for LGRB samples than for galaxy surveys could be explained by the different $Ly\alpha$ flux limits. The magnitude selection of Pentericci et al. (2010) reaches $R = 26$ mag. Applying the same cut to the TOUGH sample would not change the results compared to the $R = 25.5$ mag cut by much.

We do not find any particular correlation of the LGRB host galaxy properties (stellar mass, SFR, metallicity, UV magnitude, dust extinction, or H I column density along LGRB line of sight) with $Ly\alpha$ luminosity or $EW_0$, but our results are limited by the poor statistics. As shown in Fig. 5, whenever the information on the SFR of LAE-LGRBs is available, its value is below 30 $M_\odot$ yr$^{-1}$, except for GRB 080602, which has a lower limit of 48 $M_\odot$ yr$^{-1}$ (Palmerio et al. 2019). In stellar masses, LAE-LGRBs are associated preferentially with stellar masses $M_\star < 10^{10} M_\odot$, as also found by Pentericci et al. (2010) for LBG samples at $z \sim 3$. However, due to the poor statistics, we cannot conclude if this is an intrinsic properties of LAEs or a consequence of the characteristics of the parent samples.

The low fraction of LAEs in the XHG sample is likely due to the difference in the $M_{UV}$ and stellar mass range covered by the TOUGH and XHG samples. By construction, the latter is biased toward LGRBs with highly extinguished afterglows

(Krühler et al. 2015), which may originate in more massive and dusty galaxies on average. Of the 30 objects of the XHG sample for which information on $E(B - V)$ (even if with large errors) is available, 70% have $E(B - V)$ values higher than the most extinguished LAE-LGRB in the sample. This is consistent with the fact that dust is considered to have a significant impact on $Ly\alpha$ suppression (Neufeld 1990; Laursen et al. 2009) and that LAEs are preferentially found in systems with low extinction (e.g., Shapley et al. 2003; Pentericci et al. 2007, 2010; Matthee et al. 2016). The metallicities of the LAE-LGRBs also support this trend as their values are among the lowest values of the TOUGH and XHG sample (see Appendix A).

Matthee et al. (2021) predicted that the bulk of $Ly\alpha$ population of UV galaxies with $M_{1500} \sim -20 \pm 1$ at $z = 2$ should be below $\sim 2 \times 10^{42}$ erg s$^{-1}$, and the fraction of LAEs with $EW_0 > 25$ Å should be of 10−30%. Figure 6 shows that most of the TOUGH unbiased sample of LGRB hosts is below this limit. The reason is that LGRB samples have no preselection except for the LGRB explosions. Following the predictions from Matthee et al. (2021, see their Fig. 1), LAE-LGRBs are therefore representative of $Ly\alpha$ emission from the bulk of UV-selected galaxies at $z \sim 2$. Furthermore, the fraction of LAEs with $EW_0 > 25$ Å in the TOUGH sample is $\sim 25$%, in agreement with their expectations for UV-selected galaxies.

## 3. Golden sample

Gamma-ray bursts offer the rare possibility of combining the information about the absorbing and emitting gas of high-redshift galaxies. This is possible through afterglow and host galaxy spectroscopy, which are obtained after the afterglow has faded. This advantage can be useful to characterize LAE-LGRBs and to test $Ly\alpha$ models.

For this purpose, we selected a golden sample of LAE-LGRBs from the census presented in Sect. 2 fulfilling the following criteria: (i) The availability of afterglow and host galaxy spectra; (ii) The $Ly\alpha$ detection in the afterglow or host galaxy spectra with spectral resolution $R > 1000$ because lower resolution can lead to a misinterpretation of the line profile (e.g., Verhamme et al. 2015; Gronke et al. 2015); (iii) The detection of other host galaxy emission lines in order to determine the systemic redshift of the galaxy, the intrinsic $Ly\alpha$, and galaxy properties.

We stress the importance of X-shooter observations to fulfill point (iii) because its wide spectral coverage ranges from ~300 nm to 2500 nm.

We find only four objects fulfilling these criteria: GRBs 011211, 021004 (reported independently in the literature), 060926 (part of the XHG sample), and 070110 (part of both the XHG and TOUGH sample). The images of their fields, with the slits used to obtain the afterglow and host-galaxy spectra, are shown in Fig. 7. We reduced the X-shooter spectra of the four GRB hosts (see log of the observations in Table 2) and measured the emission line fluxes (reported in Tables 3 and 4) according to the procedure described in Sect. 2.2. These fluxes only represent the Ly$\alpha$ photons falling in the slit, without any correction to take the 2D spatial extension of the Ly$\alpha$ emission into account. If we were to consider a 2D Gaussian approximation, the resulting fluxes would be a factor ~2 higher on average. The Ly$\alpha$ fluxes of these host galaxies have previously been reported in Milvang-Jensen et al. (2012). Our measurements using X-shooter spectra provide lower values that are consistent within the uncertainties (at 2 or 3$\sigma$), however. This is a good agreement, considering that different observing techniques and slit position angles may cover different parts of the diffuse Ly$\alpha$ emission.

For each host galaxy we determined the star-formation rate (SFR; or we placed limits on it) and the Ly$\alpha$ properties by the following method. We used the H$\alpha$ flux (measured or converted from the H$\beta$ flux, assuming H$\alpha$ = 2.86 × H$\beta$; Osterbrock 1989) to derive the H$\alpha$ luminosity corrected for Milky Way extinction ($L(H\alpha_{obs}^{MWcor})$). We converted the H$\alpha$ luminosity into the SFR using the relation from Kennicutt (1998), scaled to the Chabrier (2003) initial mass function as

$$\text{SFR}(H\alpha) = 4.6 \times 10^{-42} \, L(H\alpha_{obs}^{MWcor}) \, M_\odot \, \text{yr}^{-1} \,. \qquad (1)$$

We cannot retrieve information on the dust extinction of the four host galaxies from the emission lines because only one of the Balmer lines (H$\alpha$ or H$\beta$) is detected in the spectra. Therefore the SFR values were determined without any correction for host-galaxy dust and should be considered as lower limits.

Under the assumption that the Ly$\alpha$ and the Balmer lines originate from the same regions and are produced by the same recombination process, it is possible to infer the intrinsic properties of the Ly$\alpha$ emission line from the H$\alpha$ or H$\beta$ lines. We determined the $FWHM_i(Ly\alpha)$ from a Gaussian fit of the H$\alpha$ (or H$\beta$) line profile that we corrected for instrumental dispersion. In case B recombination, the theoretical ratio between the Ly$\alpha$ and the H$\alpha$ lines is ~8.7 (Brocklehurst 1971). Using this value and assuming no extinction, we converted the Balmer flux into intrinsic Ly$\alpha$ flux ($F_i(Ly\alpha)$). We determined the $f_{esc}(Ly\alpha)$ as the ratio of the observed Ly$\alpha$ flux to the intrinsic flux. Following the consideration of Ly$\alpha$ fluxes reported above, the $f_{esc}(Ly\alpha)$ values should be considered as lower limits. The rest-frame Ly$\alpha$ equivalent width (EW$_0$(Ly$\alpha$)) is determined by dividing the rest-frame Ly$\alpha$ flux by the Ly$\alpha$ continuum level.

We detail the analysis of the data of each GRB in the following sections. The fluxes of the identified lines, the SFR, and the Ly$\alpha$ emission characteristics are reported in Tables 3 and 4.

In general, we remark that the Ly$\alpha$ emission line profile and velocity spread of our sample (see next sections) are typical of LAEs at $z \sim 2$. Similarly to those of the XLS-$z$2 sample of LAEs presented in Matthee et al. (2021), they show an asymmetric profile with a prominent peak redshifted from the systemic redshift of the host by $\gtrsim$200 km s$^{-1}$, and extending over $\gtrsim$500 km s$^{-1}$. However, the velocity shift of the ISM absorption lines detected in the afterglow spectra with respect to the systemic velocity of the galaxies ($\sim -100$ km s$^{-1}$) are lower than those of the XLS-$z$2
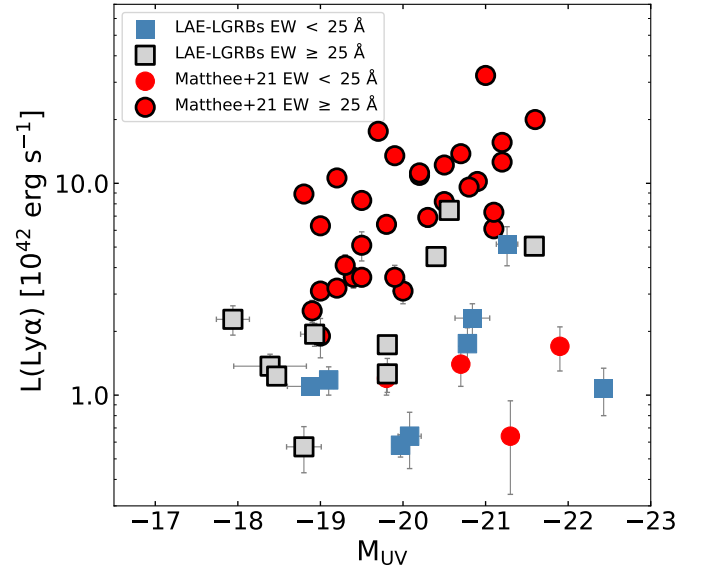


**Fig. 6.** Distribution of luminosities of LAE-LGRBs and LAEs of the XLS-$z$2 sample (Matthee et al. 2021) as a function of the galaxy UV absolute magnitude. Systems with EW$_0$ > 25 Å are shown with bold black contours. The remaining LAEs have a fainter or an undetermined EW$_0$.

sample ($\sim -260$ km s$^{-1}$), likely testifying to a more static environment.

### 3.1. GRB 011211

The photometry of the afterglow and host galaxy of GRB 011211 has been published in Jakobsson et al. (2003). The detection of Ly$\alpha$ emission through narrow-band filters was reported by Fynbo et al. (2003). The host galaxy of GRB 011211 has a multicomponent morphology (Jakobsson et al. 2003; Fynbo et al. 2003). The GRB site is in the southeast part of the system, while the Ly$\alpha$ emission peaks in the central-north part of the system and extends over the entire system.

We present here the X-shooter observation of the host galaxy (Prog. ID: 084.A-0631; PI: S. Piranomonte, see Table 2), which has not previously been published. At the afterglow position, we clearly identify the [OIII]$\lambda\lambda$4959, 5007 doublet and Ly$\alpha$ line (see Fig. 8). From the [OIII]$\lambda\lambda$4959, 5007, we derive a redshift of 2.1434±0.0001. The Ly$\alpha$ peak is redshifted by 280±30 km s$^{-1}$ compared to the host galaxy redshift. The observed Ly$\alpha$ line properties are reported in Table 4. The spatial extension of the Ly$\alpha$ line in the 2D spectrum is 2″5, compared to 1″8 for the brightest nebular emission line [OIII]$\lambda$5007. From the 3$\sigma$ upper limit on the H$\alpha$ flux, we determine a SFR < 2.0 $M_\odot$ yr$^{-1}$. Perley et al. (2013) obtained an average dust attenuation of $A_V =$ 0.19$^{+0.70}_{-0.00}$ mag based on fitting the host galaxy spectral energy distribution (SED).

The afterglow spectrum shows the detection of the GRB Damped Ly$\alpha$ system (DLA) with $\log(N_{HI}/\text{cm}^{-2})$ = 20.4 ± 0.2 (Vreeswijk et al. 2006), with many associated absorption lines. From the low-ionization state lines (LIS) redshift reported by Vreeswijk et al. (2006), we determine the velocity of the ISM with respect to the systemic redshift of the host galaxy, $V_{LIS} = -160 \pm 180$ km s$^{-1}$, with high uncertainties due to the low resolution of VLT/FORS2 spectra. The Ly$\alpha$ emission peak would fall in the Ly$\alpha$ absorption trough but is not detected in the afterglow spectra, likely due to a combination of line faintness,
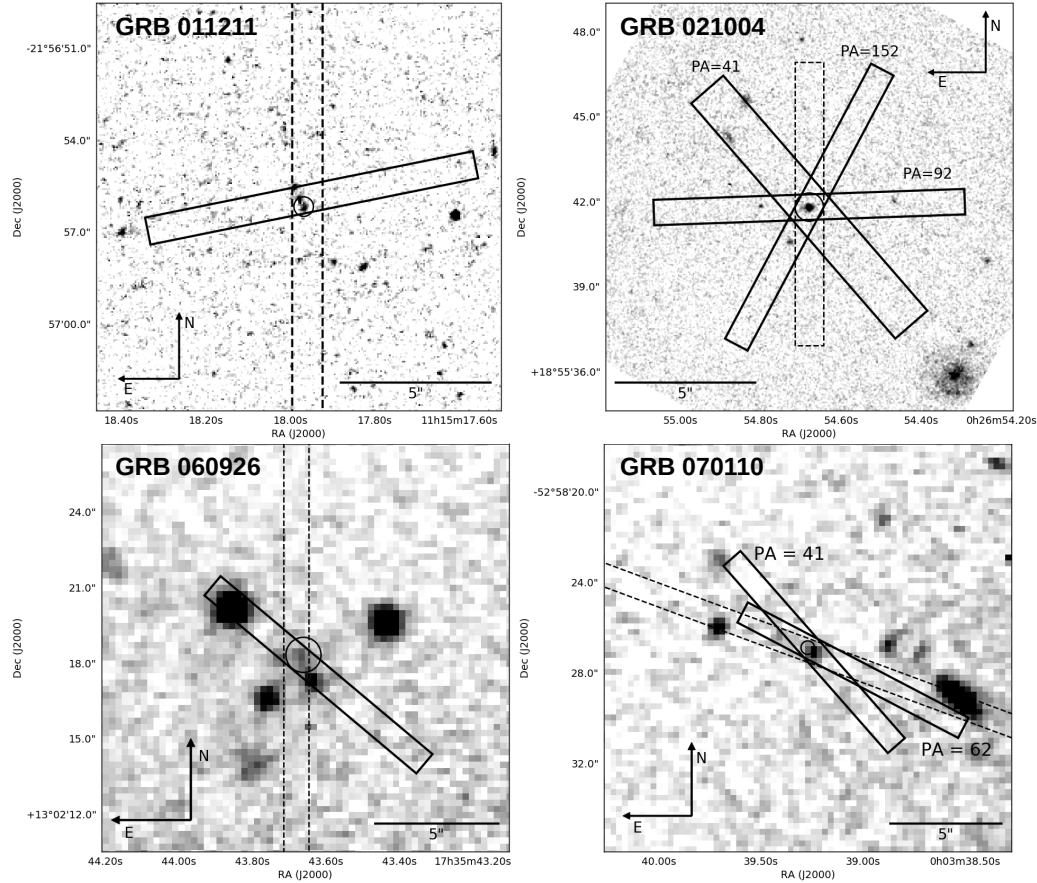
**Fig. 7.** Images of the fields of the four LAE-LGRBs of the golden sample. The slit positions of the afterglow (dashed lines) and host galaxy spectra (solid line) are superimposed. The afterglow position is represented by a circle (Jakobsson et al. 2003; Fox 2002; Holland et al. 2006; Malesani et al. 2007, for GRB 011211, 021004, 060926 and 070110, respectively). GRB 011211 field: HST/STIS image (Prog. ID: 8867; P.I.: S.R. Kulkarni); GRB 021004 field: HST/ACS WFC2 F814 image (Prog. ID: 9405; P.I.: A. Fruchter); GRB 060926A field: VLT/FORS1 image (Prog. ID: 079.A-0253(A); P.I.: P. Jakobsson); GRB 070110 field: VLT/FORS2 image (Prog. I.D.: 098.D-0416(A); PI: S. Schulze).

**Table 2.** Observing log of the observations of the four GRB host galaxy spectra of the golden sample.

| GRB host | Redshift | Exposure time (s) | | | Slit width | | | PA | Obs. date | Seeing | Airmass |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | UVB | VIS | NIR | UVB | VIS | NIR | (deg) | | (″) | |
| GRB 011211 | 2.1434 | $4 \times 1800$ | $4 \times 1800$ | $4 \times 900$ | 1.0 | 0.9 | 0.9 | 101 | 2010-03-18 | 0.5 | 1.4 |
| GRB 021004 | 2.3298 | $4 \times 1200$ | $4 \times 1200$ | $8 \times 650$ | 1.6 | 1.5 | 1.5 | 41 | 2009-11-21 | 1.3 | 1.4 |
| GRB 060926 | 3.2090 | $2 \times 1800$ | $2 \times 1800$ | $6 \times 600$ | 1.0 | 0.9 | 0.9 | 47 | 2010-04-18 | 0.8 | 1.3 |
| GRB 070110 | 2.3523 | $4 \times 1680$ | $4 \times 1680$ | $8 \times 840$ | 1.0 | 0.9 | 0.9 | 42 | 2010-10-29 | 0.8 | 1.4 |

**Notes.** The columns indicate the GRB name, redshift determined from emission lines (see Sect. 3), exposure time, slit width and position angle (PA) used in this study, as well as the observing date, average seeing, and airmass of the observations.

**Table 3.** Fluxes of Balmer and oxygen lines and derived properties.

| GRB host | $H\beta$ $[10^{-17}\ \mathrm{erg\ s^{-1}\ cm^{-2}}]$ | $H\alpha$ $[10^{-17}\ \mathrm{erg\ s^{-1}\ cm^{-2}}]$ | [OII]$\lambda 3726, 3729$ $[10^{-17}\ \mathrm{erg\ s^{-1}\ cm^{-2}}]$ | [OIII]$\lambda 4959, 5007$ $[10^{-17}\ \mathrm{erg\ s^{-1}\ cm^{-2}}]$ | [OIII]/[OII] | SFR $[M_\odot\ \mathrm{yr^{-1}}]$ |
|---|---|---|---|---|---|---|
| GRB 011211 | <0.4 [*] | <1.2 | <1.6 | $2.7 \pm 0.3$ | >1.7 | <2.0 |
| GRB 021004 | $1.2 \pm 0.2$ [*] | $3.3 \pm 0.4$ | <1.7 | $16.9 \pm 0.5$ | >10 | $6.7 \pm 0.9$ |
| GRB 060926 | $1.6 \pm 0.3$ | $4.6 \pm 0.9$ [*] | $2.3 \pm 0.4$ [#] | $3.5 \pm 0.4$ [†] | $1.5 \pm 0.4$ | $20.2 \pm 4.0$ |
| GRB 070110 | $0.3 \pm 0.2$ [*] | $0.9 \pm 0.4$ | $1.8 \pm 0.3$ [#] | $4.0 \pm 0.3$ | $2.2 \pm 0.4$ | $1.9 \pm 0.6$ |

**Notes.** Fluxes are corrected for Galactic extinction and slit loss. No correction for host-intrinsic extinction is applied. Upper limits are at $3\sigma$. [*]Flux estimated using $H\alpha = 2.86 \times H\beta$. [†][OIII]$\lambda 4959$ line not detected. We fixed its flux to one-third of the [OIII]$\lambda 5007$ flux. [#][OII]$\lambda 3729$ line falling on a sky line. We fixed its flux to two-thirds of the [OII]$\lambda 3726$ flux.

**Table 4.** Fluxes and properties of the observed Lyα emission line.

| GRB host | $F$(Lyα) $[10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}]$ | $\mathrm{EW_0}$(Lyα) [Å] | $FWHM_0$(Lyα) $[\mathrm{km\,s^{-1}}]$ | Lyα red-peak shift $[\mathrm{km\,s^{-1}}]$ | $f_{\mathrm{esc}}$(Lyα) | $F_i$(Lyα) $[10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}]$ |
|---|---|---|---|---|---|---|
| GRB 011211 | $1.6 \pm 0.2$ | $14 \pm 3$ | $249 \pm 60$ | $280 \pm 30$ | $>0.16$ | $<10.4$ |
| GRB 021004 | $16.9 \pm 0.3$ | $105 \pm 10$ | $228 \pm 20$ | $180 \pm 20$ | $0.59 \pm 0.08$ | $28.8 \pm 3.5$ |
| GRB 060926 | $5.3 \pm 0.4$ | $37 \pm 7$ | $417 \pm 40$ | $250 \pm 20$ | $0.13 \pm 0.03$ | $40.0 \pm 7.9$ |
| GRB 070110 | $2.8 \pm 0.5$ | $33 \pm 8$ | $412 \pm 80$ | $285 \pm 40$ | $0.36 \pm 0.09$ | $7.8 \pm 2.7$ |

**Notes.** The Lyα fluxes are corrected for Galactic extinction and slit loss. No correction for host-intrinsic extinction is applied. The $\mathrm{EW_0}$(Lyα) and $FWHM_0$(Lyα) are the rest-frame equivalent width and FWHM of the observed Lyα line. The Lyα red-peak shift is the velocity shift of the Lyα peak redward of the theoretical Lyα line center, as calculated from the systemic redshift of the galaxy (see Table 2). The escape fraction of Lyα photons ($f_{\mathrm{esc}}$(Lyα)) and the intrinsic flux of the Lyα line ($F_i$(Lyα)) are determined from Balmer lines as described in Sect. 3. We stress that these escape fractions are not corrected for dust extinction, as explained in Sect. 3.

spectral resolution, and noisy spectral region. From the power-law fitting of the afterglow, Jakobsson et al. (2003) determined $A_V = 0.08 \pm 0.08$ mag along the line of sight.

### 3.2. GRB 021004

The afterglow and host galaxy of GRB 021004 have been intensively observed (see Fynbo et al. 2005 and references therein). Møller et al. (2002), showed the presence of Lyα emission in the optical spectrum of the GRB afterglow, which was later confirmed by many other works (e.g., Mirabal et al. 2003; Starling et al. 2005). The host galaxy of GRB 021004 has a compact core with a faint second component, and the GRB site is located at the center of its host (Fynbo et al. 2005). The X-shooter observation of the host galaxy (Prog. ID: 084.A-0631; PI: S. Piranomonte, see Table 2) has previously been published in Vergani et al. (2011). In Fig. 7 we show the image of GRB 021004 field with the slits used to observe the host galaxy and the afterglow superimposed. Vergani et al. (2011) reported the detection of the [OIII]$\lambda5007$ emission line at the same redshift as the GRB from the galaxy eastward of the GRB host in the spectrum obtained with the slit position angle PA = 92°. Its proximity to the GRB host (projected distance of 14 kpc) may indicate a possible interaction between this galaxy and the GRB host.

For our analysis, we used the X-shooter spectrum obtained with slit PA of 41°. Because its slit width is larger, it minimizes the loss of Lyα flux. We reduced the spectrum following the procedure described in Sect. 2 and identify the [OIII]$\lambda\lambda4959, 5007$ doublet, Hα, and Lyα emission lines (see Fig. 8). Hβ falls exactly on a sky line, and the [OII]$\lambda3727$ doublet is not detected ($3\sigma$ upper limit of $1.7 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$). We derive a redshift of $z = 2.3298 \pm 0.0001$ for the host galaxy. We estimate the line fluxes and SFR following the procedure described in Sect. 3. We find an SFR $= 6.7 \pm 0.8\,M_\odot\,\mathrm{yr^{-1}}$ (without correction for host extinction). The Lyα emission from the host galaxy is an asymmetric line redshifted by $180 \pm 20\,\mathrm{km\,s^{-1}}$ with an $\mathrm{EW_0} = 105 \pm 10$ Å. The spatial extension of the Lyα line in the 2D spectrum is 4″7, compared to 2″0 for the brightest nebular emission line [OIII]$\lambda5007$. The SED fitting performed by de Ugarte Postigo et al. (2005) provides $A_V = 0.06^{+0.08}_{-0.06}$ mag, whereas Perley et al. (2013) reported $A_V = 0.42^{+0.09}_{-0.07}$ mag.

The Lyα emission is also detected in the afterglow spectrum at the same velocity and with a similar shape profile as that detected in the host galaxy spectrum. The afterglow spectrum shows also the Lyα absorption of the GRB sub-DLA system, with $\log(N_{\mathrm{HI}}/\mathrm{cm^{-2}}) = 19.5 \pm 0.5$ (Fynbo et al. 2005).

The VLT/UVES afterglow spectrum of GRB 021004 presents a plethora of absorption lines with complex velocity structures spanning thousands of $\mathrm{km\,s^{-1}}$ (Fiore et al. 2005; Chen et al. 2007; Castro-Tirado et al. 2010). Although a progenitor-star wind origin was claimed, this was firmly excluded (at least for some of these absorbing systems) by the detection of low-ionization transitions (see Chen et al. 2007). Here we focus on the lowest velocity component, associated with the GRB sub-DLA, therefore likely representing the cold and warm ISM gas of the GRB host galaxy. Using the shift of the AlII$\lambda1670$ and CII$\lambda1334$ LIS from the systemic redshift of the galaxy, we determine an ISM velocity $V_{\mathrm{LIS}} = -80 \pm 20\,\mathrm{km\,s^{-1}}$. From the spectral flux distribution of the afterglow at several epochs, de Ugarte Postigo et al. (2005) determined an average $A_V = 0.20 \pm 0.08$ mag along the GRB line of sight.

### 3.3. GRB 060926

In the VLT/FORS1 image (Prog. ID: 079.A-0253(A); P.I.: P. Jakobsson) shown in Fig. 7, an offset between the GRB position and the brightest part of the system is visible. Both are covered by the slit used for the host spectroscopic observations. In the X-shooter host galaxy observations obtained under program 085.A-0795(A) (presented here for the first time; PI: H. Flores; see Table 2), we identify the [OIII]$\lambda5007$, [OII]$\lambda3726$, Hβ, and Lyα lines (see Fig. 8). These lines, except for Lyα, are used to derive a redshift of $z = 3.2090 \pm 0.0001$ for the host galaxy, which is consistent with the redshift derived in Krühler et al. (2015) with a different data set. Even with seeing conditions of $\sim$0″8, it is not possible based on the emission lines to separate the different parts of the host galaxy systems. The spatial extension of the Lyα line in the 2D spectrum is 2″7, compared to 1″4 for the brightest nebular emission line [OIII]$\lambda5007$.

The high redshift of the host means that the wavelength coverage of X-shooter does not allow us to observe the Hα line. We determine an SFR $= 20.2 \pm 4.0\,M_\odot\,\mathrm{yr^{-1}}$, assuming Hα $= 2.86 \times$ Hβ (Osterbrock 1989). As we do not have any information on the host extinction, the value inferred above should be considered as a lower limit of the SFR. The Lyα emission from the host galaxy is an asymmetric line redshifted by $250 \pm 20\,\mathrm{km\,s^{-1}}$ from the systemic redshift. The observed Lyα equivalent width and intrinsic properties are reported in Table 4.

The afterglow spectrum of GRB 060926 has been published in Fynbo et al. (2009). The Lyα line is clearly detected in the trough of the DLA. From the SiII$\lambda1526$, AlII$\lambda1670$, and CII$\lambda1334$ LIS redshift, we determine an ISM velocity of $V_{\mathrm{LIS}} = -30 \pm 180\,\mathrm{km\,s^{-1}}$. The HI column density derived from the fit
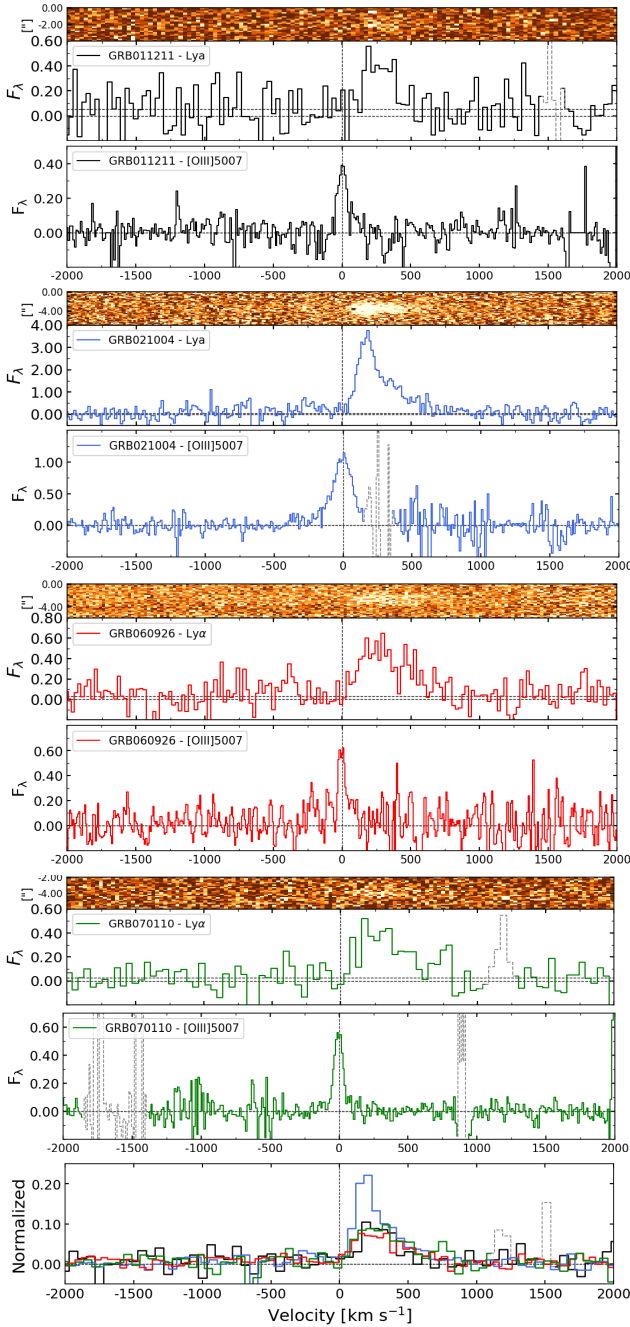
**Fig. 8.** Lyα and [OIII]λ5007 emission lines from the X-shooter spectra of the four GRB host galaxies GRB 011211, 021004, 060926, and 070110 *from top to bottom*. The lines are centered at the systemic redshift of the galaxy determined from the [OIII] emissions lines. The flux density $F_\lambda$ is in units of $10^{-17}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$. *Bottom panel*: superposition of the four Lyα lines rebinned to 0.8 Å and normalized to the Lyα line flux.

of the Lyα absorption (Fynbo et al. 2009) is one of the highest values probed by a GRB afterglow, $\log(N_{\rm HI}/{\rm cm}^{-2}) = 22.6 \pm 0.15$.

### 3.4. GRB 070110

The near-infrared X-shooter spectrum of GRB 070110 host galaxy has been studied in Krühler et al. (2015), but the UVB spectrum has not been previously published. The [OIII]λλ4959, 5007, [OII]λ3726, Hα, and Lyα lines are identi-

fied in the host galaxy spectrum (see Fig. 8), allowing a redshift determination of $z = 2.3523 \pm 0.0001$. The measured Hα flux corresponds to an SFR = $1.9 \pm 0.6\,M_\odot\,{\rm yr}^{-1}$. The Lyα emission from the host galaxy shows an asymmetric profile and is redshifted by $285 \pm 40\,{\rm km\,s}^{-1}$ from the systemic redshift (see also Table 4 for other Lyα properties). The spatial extension of the Lyα line in the 2D spectrum is 2″1, compared to 2″0 for the brightest nebular emission line [OIII]λ5007.

The afterglow spectrum of GRB 070110 has been published in Fynbo et al. (2009). They reported the detection of the Lyα emission line in the trough of the GRB-DLA. From the SiIIλ1526, AlIIλ1670, and CIIλ1334 LIS redshift, we determine an ISM velocity $V_{\rm LIS} = -20 \pm 50\,{\rm km\,s}^{-1}$. The HI column density derived from the fit of the Lyα absorption (Fynbo et al. 2009) is $\log(N_{\rm HI}/{\rm cm}^{-2}) = 21.7 \pm 0.10$. Troja et al. (2007) determined $A_V = 0.08 \pm 0.08$ mag along the GRB afterglow line of sight.

## 4. Modeling the Lyman-alpha line

### 4.1. Description of the model

Following Vielfaure et al. (2020), we modeled the observed Lyα line with the shell-model fitting pipeline described in Gronke et al. (2015). This simplistic model consists of an expanding, homogeneous, spherical shell composed of uniformly mixed neutral hydrogen and dust (Ahn et al. 2003; Verhamme et al. 2006). The expanding geometry was motivated by the HI outflows, which appear to be ubiquitous at both low- and high-redshifts (e.g., Shapley et al. 2003; Wofford et al. 2013; Rivera-Thorsen et al. 2015; Chisholm et al. 2015). The Lyα emitting source was placed at the center of the sphere filled with ionized gas and surrounded by a neutral and dusty shell. This source would correspond to the star-forming region hosting the LGRB. This assumption is reasonable considering that LGRBs are produced by massive stars and are found in the central star-forming region of their host (typically at ∼1 kpc from the center; e.g., Lyman et al. 2017). The photons are collected in all directions after their scattering through the neutral shell.

The shell model is defined by seven parameters that we briefly detail here. Four parameters describe the shell properties: (i) the radial expansion velocity ($V_{\rm exp}$), (ii) the HI column density ($N_{\rm HI}$), (iii) the dust optical depth ($\tau_{\rm d}$) at wavelengths in the vicinity of Lyα, and (iv) the effective temperature of the gas ($T$). The radial expansion velocity is physically linked to the broadening and the peak shift of the Lyα line. For an outflowing shell, the velocity expansion regulates the blue peak intensity by scattering the photons in the red wing of the line. A higher velocity of the gas increases these effects, but suppresses the interaction of the photons with the neutral gas. See, for example, Verhamme et al. (2006) for more details about this effect. The $V_{\rm exp}$ of the neutral gas can be determined from the velocity shift of the LIS lines ($V_{\rm LIS}$) when available. The effect of increasing $N_{\rm HI}$ is to broaden the line by shifting the Lyα photons from the line center. The dust grains absorb and scatter the Lyα photons. An increasing dust content will favor this process and affect the intensity of the line, but it can also affect its shape. An increasing $N_{\rm HI}$ boosts the interaction of the photons with dust so that both parameters are linked. The effect of the effective temperature is complex. Its impact on the width of the line and the position of the peak depends on the other parameter values. We considered a Gaussian profile for the intrinsic Lyα emission and an adjacent flat UV continuum, therefore three additional parameters were used: (v) the redshift of the emitter ($z$) corresponding to the center of the intrinsic Lyα line, (vi) the intrinsic
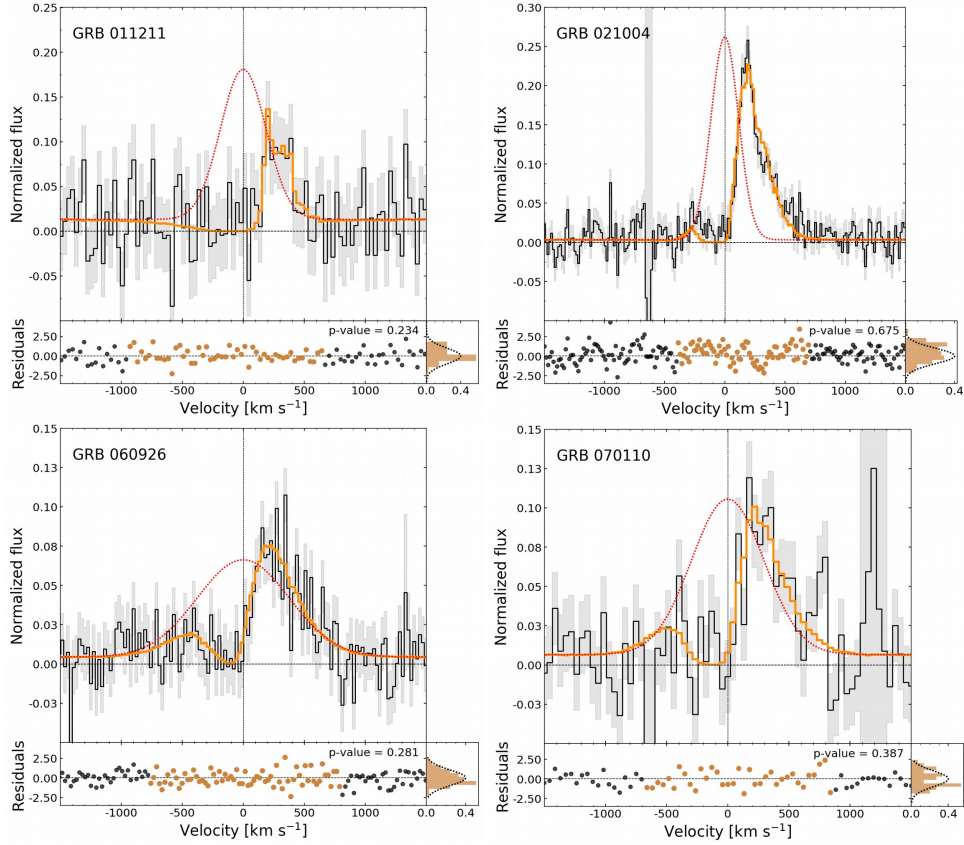
**Fig. 9.** Best fit of the Lyα line obtained by the unconstrained shell model. For each LGRB host, we show in the *top panels* the data (solid black line) with their error bars (grey), the best fit (dashed orange line), and the intrinsic Lyα emission predicted by the models (dotted red line). *Bottom panels*: the dots correspond to the normalized residuals between the observation and the model, and brown dots show the residuals covering the fitted Lyα line. The distribution of the residuals is projected in the *right bottom panels*. We also superimpose a Gaussian distribution with $\mu = 0$ and $\sigma = 1$ (dotted black line) as a reference for comparison. The p-values correspond to the Shapiro-Wilk test.

equivalent width of the Lyα line ($EW_i(Ly\alpha)$), and (vii) the intrinsic full width half maximum of the Lyα line ($FWHM_i(Ly\alpha)$). We refer to Verhamme et al. (2006) and Gronke et al. (2015) for a more detailed description of the model and parameters.

The observation of LGRB host galaxy and afterglow allows us to constrain the shell-model parameters even for faint star-forming galaxies. However, while host galaxy observations provide information for the overall galaxy, the afterglow constrains parameters along the LGRB line of sight. By using the shell model with afterglow observations, we therefore assume that the LGRB region is the main source of Lyα emission in the host, and that the LGRB line of sight probes gas properties characteristic of the environment surrounding the LGRB region. In the following, to differentiate between the parameters constrained by these two types of observation, we annotate with $X^{\mathrm{OA}}$ the parameters related to optical afterglow observations and $X^{\mathrm{HG}}$ the host galaxy parameters. As a first step, we fit the observed Lyα lines without constraints, as would typically be done for the majority of high-$z$ LAE observations, where most of the parameters can hardly be constrained observationally. Then, we fit the Lyα profiles using the observational constraints and also discuss the effects of constraining the parameters with only the host galaxy or afterglow observations.

### 4.2. Unconstrained Lyα profile fitting

To evaluate the predictions of the shell model, we fit the observed Lyα lines of the four LGRBs of the golden sample using

an improved model grid and process originally described in Gronke et al. (2015), without considering any prior. The automatic fitting succeeds in reproducing the line profile satisfactory in all four cases. In Fig. 9 we present for each spectrum the best fit corresponding to the lowest $\chi^2$. In all cases except for GRB 011211, the models fit a blue peak that is not clearly seen in the spectra, but is consistent with the uncertainties. The distributions of normalized residuals presented at the bottom of each spectrum are consistent with a unit Gaussian distribution, indicating a good agreement between the models and the observations.

While the unconstrained shell model succeeds in reproducing the observed profiles, the comparison between the best shell parameters returned by the fitting (Table 5) and the corresponding values determined from the observations (Table 6) reveals important discrepancies for GRBs 060926 and 070110. Their observed H I column density is higher than the fit results, and the $FWHM_i(Ly\alpha)$ fitting values are largely overestimated.

The Lyα escape fraction ($f_{\mathrm{esc}}(Ly\alpha)$; see Table 7) is also a byproduct of the shell model and is found to be consistent (within the uncertainties) with the value determined from the observations (see Table 4) only in the case of GRB 070110. Particularly, $f_{\mathrm{esc}}(Ly\alpha)$ is discrepant for GRBs 011211 and 021004, whereas the fit and best-fitting parameters agree with the observations. The disagreement on $f_{\mathrm{esc}}(Ly\alpha)$ also implies that the Hα flux and subsequent SFR obtained by the fit are inconsistent with the values determined from the observations (see Tables 3 and 7). For GRB 011211, the Hα flux is not well constrained

**Table 5.** Best-fit results of the unconstrained shell model.

| GRB host | $\Delta z$ | $\log(N_{HI}/cm^{-2})$ | $V_{exp}$ [km s$^{-1}$] | $\log(T/K)$ | $\tau_d$ | $FWHM_i(Ly\alpha)$ [km s$^{-1}$] | $EW_i(Ly\alpha)$ [Å] |
|---|---|---|---|---|---|---|---|
| | | | | Shell model results from unconstrained fitting | | | |
| GRB 011211 | $30^{+45}_{-45}$ | $19.9^{+0.3}_{-0.3}$ | $112^{+33}_{-27}$ | $4.52^{+0.66}_{-1.08}$ | $2.8^{+1.6}_{-1.7}$ | $440^{+130}_{-150}$ | $63^{+38}_{-31}$ |
| GRB 021004 | $5^{+35}_{-35}$ | $19.6^{+0.1}_{-0.1}$ | $122^{+6}_{-6}$ | $4.99^{+0.15}_{-0.15}$ | $0.365^{+0.104}_{-0.081}$ | $270^{+22}_{-24}$ | $157^{+18}_{-13}$ |
| GRB 060926 | $0^{+45}_{-45}$ | $19.5^{+0.2}_{-0.1}$ | $164^{+35}_{-29}$ | $3.59^{+1.04}_{-0.57}$ | $0.52^{+0.99}_{-0.40}$ | $890^{+115}_{-115}$ | $63^{+35}_{-20}$ |
| GRB 070110 | $0^{+45}_{-45}$ | $19.7^{+0.3}_{-0.2}$ | $182^{+58}_{-45}$ | $4.37^{+0.94}_{-1.07}$ | $1.14^{+2.00}_{-0.91}$ | $700^{+145}_{-160}$ | $53^{+32}_{-25}$ |

**Notes.** $\Delta z = (z_{fit} - z_{host}) \times c$, where $z_{fit}$ is the Ly$\alpha$ redshift of the best fit, $z_{host}$ is the redshift derived from the X-shooter emission lines (Table 6), and c is the speed of light; $\log(N_{HI}/cm^{-2})$ is the HI column density of the shell; $V_{exp}$ is the shell expansion velocity; $\log(T/K)$ is the temperature of the gas; $\tau_d$ is the dust optical depth; $FWHM_i(Ly\alpha)$ and $EW_i(Ly\alpha)$ are the intrinsic FWHM and equivalent width of the Ly$\alpha$ line, respectively.

**Table 6.** Observationally determined values of the parameters of the shell model.

| GRB host | Redshift$^{HG}$ | $\log(N_{HI}^{OA}/cm^{-2})$ | $V_{LIS}^{OA}$ [km s$^{-1}$] | $A_v^{OA}$ [mag] | $\tau_d^{OA}$ | $FWHM_i^{HG}(Ly\alpha)$ [km s$^{-1}$] | $EW_i^{HG}(Ly\alpha)$ [Å] | Refs |
|---|---|---|---|---|---|---|---|---|
| GRB 011211 | 2.1434 | $20.4 \pm 0.20$ | $-160 \pm 180$ | $0.08 \pm 0.08$ | $0.52 \pm 0.52$ | $50^{+10~(*)}_{-10}$ | <64 | (1) |
| GRB 021004 | 2.3298 | $19.5 \pm 0.50$ | $-80 \pm 20$ | $0.20 \pm 0.08$ | $1.29 \pm 0.52$ | $200^{+20}_{-20}$ | $160 \pm 40$ | (2) |
| GRB 060926 | 3.2090 | $22.6 \pm 0.15$ | $-30 \pm 180$ | – | – | $50^{+20}_{-50}$ | $275 \pm 85$ | |
| GRB 070110 | 2.3523 | $21.7 \pm 0.10$ | $-20 \pm 50$ | $0.08 \pm 0.08$ | $0.52 \pm 0.52$ | $30^{+10}_{-30}$ | $75 \pm 35$ | (3) |

**Notes.** We refer to Sect. 3 for the value determinations, except for the dust optical depth, $\tau_d$, which is converted from $A_V$ by considering a dust albedo of $A = 0.5$ (see Verhamme et al. 2006; Gronke et al. 2015). The asterisk stands for value determined from the FWHM of the [OIII]$\lambda$5007 line. "OA" is for the parameters related to optical afterglow observations and "HG" for the host galaxy parameters.
**References.** The $A_V$ values are reported in (1) Jakobsson et al. (2003), (2) de Ugarte Postigo et al. (2005), and (3) Troja et al. (2007); see Sect. 3 for more details.

**Table 7.** Properties derived from the best-fit results (unconstrained model).

| GRB host | $f_{esc}(Ly\alpha)$ (1) | $F_i(Ly\alpha)$ (2) | $F(H\alpha)$ (3) | $SFR(H\alpha)$ (4) |
|---|---|---|---|---|
| | Values obtained from the unconstrained fitting results | | | |
| GRB 011211 | $0.03^{+0.07}_{-0.02}$ | $53.1^{+124.5}_{-37.5}$ | $6.1^{+14.3}_{-4.3}$ | $10.1^{+24.0}_{-7.2}$ |
| GRB 021004 | $0.27^{+0.07}_{-0.05}$ | $62.6^{+16.3}_{-11.6}$ | $7.2^{+1.9}_{-1.4}$ | $14.7^{+3.9}_{-2.9}$ |
| GRB 060926 | $0.69^{+0.19}_{-0.22}$ | $7.8^{+7.5}_{-7.5}$ | $0.9^{+0.4}_{-0.4}$ | $3.9^{+1.8}_{-1.8}$ |
| GRB 070110 | $0.43^{+0.32}_{-0.24}$ | $7.0^{+5.3}_{-4.4}$ | $0.8^{+0.6}_{-0.5}$ | $1.6^{+1.3}_{-1.1}$ |

**Notes.** Flux units are $10^{-17}$ erg s$^{-1}$ cm$^{-2}$. (1): Escape fraction of Ly$\alpha$ photons calculated by the models according to the method described in Gronke et al. (2015). (2): Intrinsic Ly$\alpha$ flux determined from the flux of the observed Ly$\alpha$ line and the Ly$\alpha$ escape fraction calculated by the models. (3): H$\alpha$ flux inferred from the flux of the observed Ly$\alpha$ line and the $f_{esc}(Ly\alpha)$ provided by the best-fit results, calculated using $F(H\alpha) = F(Ly\alpha)/8.7$ (case B of the theory of the recombination (Brocklehurst 1971)). (4): SFR (in $M_\odot$ yr$^{-1}$) derived using the H$\alpha$ flux of column (3) and following the method described in Sect. 3.

(only $3\sigma$ upper limit), but a lower H$\alpha$ flux would imply an increased difference. We caution, however, that the Ly$\alpha$ escape fractions derived from the spectral fitting procedure is uncertain because it depends on the fitted value of $\tau_d$, which mostly has only minor effects on the spectral shape (see discussion in Gronke et al. 2015). The disagreement could also come from the lack of extinction correction of the observed H$\alpha$ flux in our calculation of $f_{esc}(Ly\alpha)$ and SFR (see Sect.3). An $A_V \approx 0.6$ mag for both GRBs 011211 and 021004 (considering an SMC extinction curve; Japelj et al. 2015) would result in consistent H$\alpha$ and $f_{esc}(Ly\alpha)$ values within $1\sigma$.

### 4.3. Constrained Ly$\alpha$ profile fitting

We then proceeded to fit the Ly$\alpha$ using the observational constraints described in Sect. 3 and listed in Table 6, as priors. Using the afterglow and host galaxy spectra as described in Sect. 3, we are able to constrain up to six out of the seven parameters taken as input by the model. The host galaxy observations provide the information on the observed Ly$\alpha$ line and the Balmer lines used to constrain the intrinsic Ly$\alpha$ emission ($FWHM_i^{HG}(Ly\alpha)$ and $EW_i^{HG}(Ly\alpha)$). They also provide the nebular emission lines used to constrain the redshift ($z^{HG}$). The afterglow observations constrain the gas properties: the hydrogen column density ($N_{HI}^{OA}$), the velocity of the gas ($V_{LIS}^{OA}$), and the information on the dust content ($\tau_d^{OA}$). It was not possible to constrain the HI temperature. The effect of this parameter on the final profile is complex and usually degenerated with the other parameters (Gronke et al. 2015). We let this parameter free for the four cases. For GRB 060926 we lack information about the dust extinction, and for GRB 011211 we have only upper limits for H$\alpha$. To be consistent with the constraints on the $N_{HI}^{OA}$ values derived from the afterglow, the dust extinctions were taken from the afterglow SED fitting reported in Sect. 3. The grid we used does not extend up to $\log(N_{HI}/cm^{-2}) = 22.6$, so that this value is not firmly constrained for GRB 060926, but constrained to the highest value of the grid, which is 21.8. We report the results of the fitting in Table 8. In Fig. 10, the intrinsic Ly$\alpha$ derived from the model and the best fit obtained with the lowest $\chi^2$ are superimposed to the data.

As expected, we find a good agreement between the fitting and the observations of the two hosts with lower $N_{HI}^{OA}$ (GRBs 011211 and 021004). As shown by the residuals at the bottom of Fig. 10, the fit is worse than in the unconstrained case

**Table 8.** Results of the shell model fit with constrained parameters.

| | Shell model results from constrained fitting | | | | | | |
|---|---|---|---|---|---|---|---|
| GRB host | $\Delta z$ | $\log(N_{HI}/cm^{-2})$ | $V_{exp}$ [km s$^{-1}$] | $\log(T/K)$ | $\tau_d$ | $FWHM_i(Ly\alpha)$ [km s$^{-1}$] | $EW_i(Ly\alpha)$ [Å] |
| GRB 011211 | $20^{+45}_{-45}$ | $20.1^{+0.2}_{-0.1}$ | $74^{+27}_{-24}$ | $3.91^{+0.8}_{-0.75}$ | $0.50^{+0.29}_{-0.24}$ | $30^{+50}_{-20}$ | $45^{+14}_{-18}$ |
| GRB 021004 | $5^{+45}_{-45}$ | $19.6^{+0.1}_{-0.1}$ | $118^{+5}_{-6}$ | $4.99^{+0.16}_{-0.15}$ | $0.402^{+0.103}_{-0.093}$ | $227^{+11}_{-12}$ | $161^{+13}_{-11}$ |
| GRB 060926 | $70^{+45}_{-45}$ | $21.5^{+0.1}_{-0.1}$ | $3^{+1}_{-1}$ | $2.94^{+0.16}_{-0.11}$ | $0.003^{+0.004}_{-0.002}$ | $70^{+20}_{-20}$ | $23^{+9}_{-6}$ |
| GRB 070110 | $35^{+45}_{-45}$ | $21.5^{+0.1}_{-0.1}$ | $3^{+1}_{-1}$ | $3.04^{+0.12}_{-0.16}$ | $0.008^{+0.010}_{-0.005}$ | $53^{+20}_{-20}$ | $33^{+25}_{-13}$ |

**Notes.** Same as Table 5 for models constrained with $z$, $N_{HI}$, $V_{exp}$, $\tau_d$, $FWHM_i(Ly\alpha)$, and $EW_i(Ly\alpha)$.
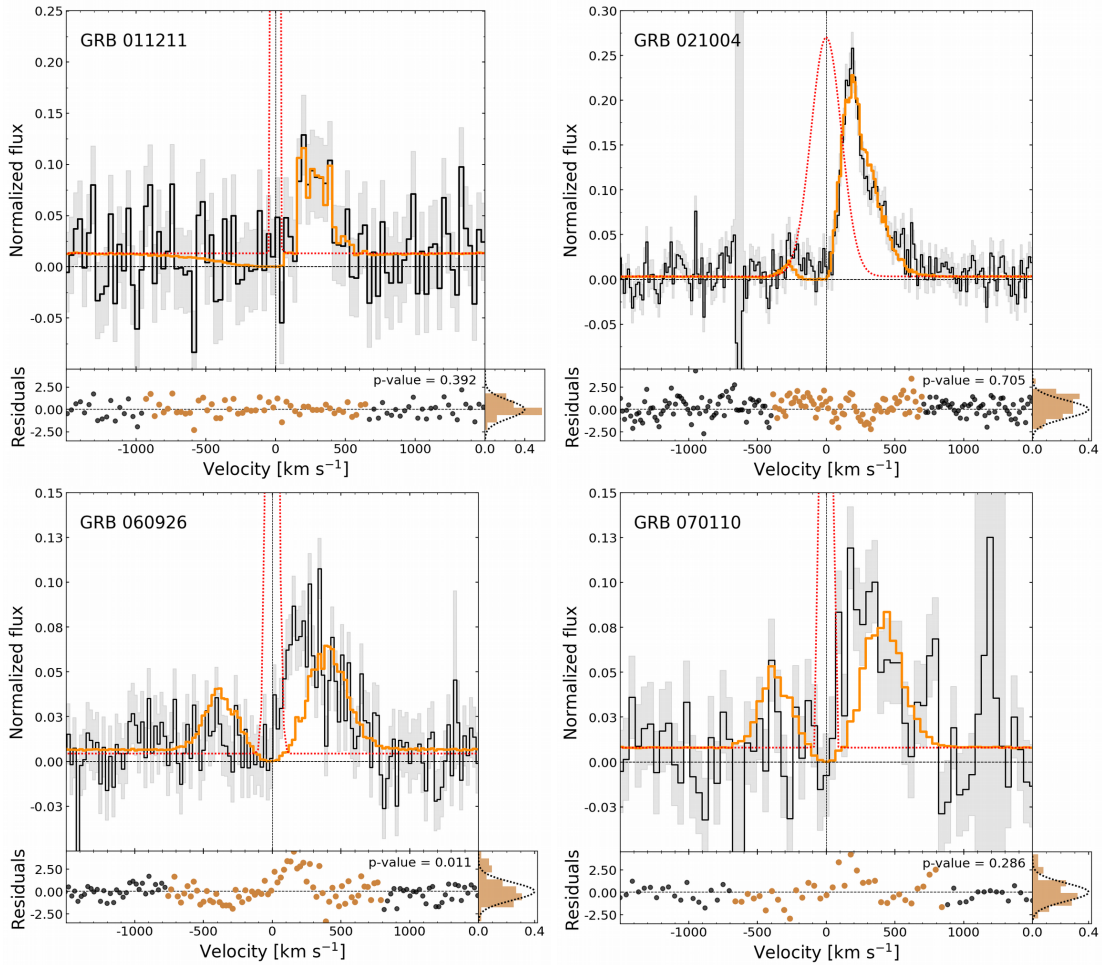


**Fig. 10.** Same as Fig. 9, but with the shell model parameters $z$, $N_{HI}$, $V_{exp}$, $\tau_d$, $FWHM_i(Ly\alpha)$, and $EW_i(Ly\alpha)$ constrained by the values determined from the observations.

for GRB 070110 and fails to reproduce the Ly$\alpha$ profile and parameters for GRB 060926. In these two cases, the models tend to fit a double-peak profile resulting from an higher red peak velocity shift ($\sim 400$ km s$^{-1}$) than observed ($\sim 280$ km s$^{-1}$) and quasi-static H I shell ($3 \pm 1$ km s$^{-1}$). By relaxing the constraints, without considering priors for $FWHM_i(Ly\alpha)^{HG}$, we can fit the profile of GRB 070110 in a better way, but the best-fitting $FWHM_i(Ly\alpha)$ becomes extremely narrow with an unrealistic value of $8^{+6}_{-3}$ km s$^{-1}$.

We also performed two sets of fitting for which we separately constrained the parameter values determined from the host galaxy and the afterglow observations. We find that the profiles of the fitting constrained with the host galaxy observations ($z^{HG}$,

$FWHM_i^{HG}(Ly\alpha)$ and $EW_i^{HG}(Ly\alpha)$) are very similar to the unconstrained cases and reproduce the data well (see Appendix C). Differently from the unconstrained case, the blue peak is absent, which is due to the constraint on $FWHM_i^{HG}(Ly\alpha)$. The predictions for the properties of the gas are also similar to the unconstrained cases. Especially the $N_{HI}$ values are lower than observed through the LGRB line of sight. The fitting constrained with the afterglow observations ($z^{HG}$, $N_{HI}^{OA}$, $V_{LIS}^{OA}$, and $\tau_d^{OA}$) only gives very similar results to the fully constrained case (see Appendix D). The models successfully reproduce the low $N_{HI}$ cases, but fail to reproduce the two cases with high $N_{HI}^{OA}$ values. These results agree with previous finding that the most constraining parameters are $N_{HI}$ and $V_{exp}$ (e.g., Gronke et al. 2015).
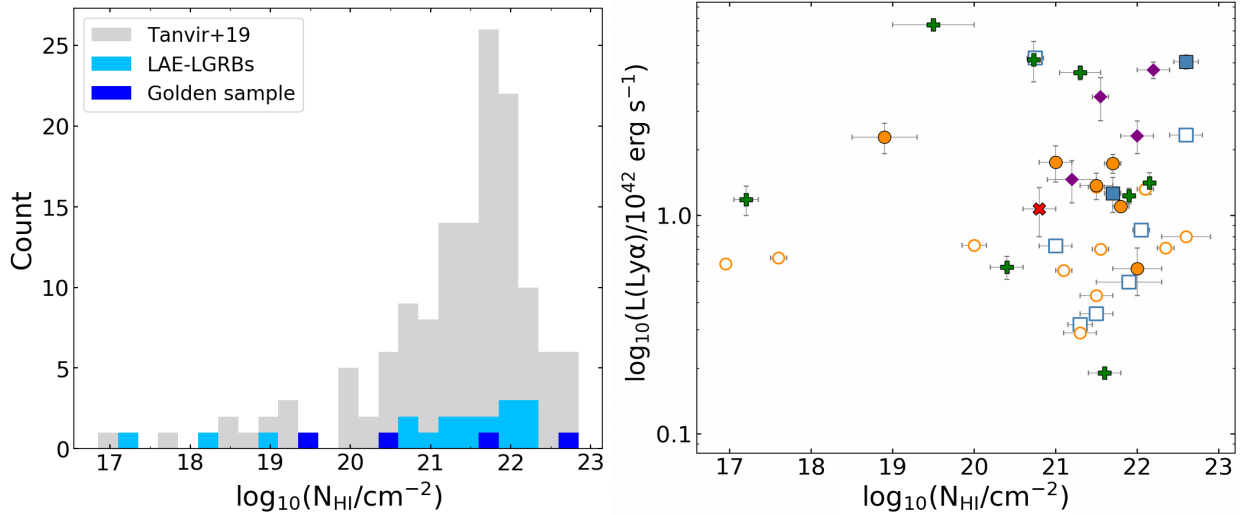
**Fig. 11.** Lyα emission and $N_{HI}$ values in LGRB host galaxies. *Left panel*: $N_{HI}$ distribution of the 140 LGRBs reported in Tanvir et al. (2019) (gray). The 22 of the 29 LAE-LGRBs that contain this information are shown in light blue, except for the GRB in the golden sample, which are represented in dark blue. The different histograms are superimposed and not cumulative. *Right panel*: Lyα luminosity distribution as function of $N_{HI}$ column density. The symbols are the same as in Fig. 1.

## 5. Discussion

### 5.1. Implications of the shell-model fits

The unconstrained fitting succeeds in reproducing all Lyα profiles of the golden sample (and GRB 191004B; see Vielfaure et al. 2020). However, if we compare the parameter values resulting from the fitting with those observed, we find important discrepancies for GRBs 060926 and 070110. A fit performed by constraining the parameter values with the observed values does not succeed in reproducing the observed line profile of these two cases.

The $FWHMs_i(Lyα)$ provided by the unconstrained models for GRBs 060926 and 070110 are very large and disagree with the $FWHM_i(Lyα)^{HG}$ we determined assuming that the Balmer photons and Lyα photons are produced by recombination in the same regions. These high values are usually returned by the shell-model fitting to reproduce the line profile of the emitters whose Lyα peaks and wings are substantially shifted from the systemic redshift. A similar discrepancy has also been found by Hashimoto et al. (2015) in the study of $z \sim 2$ LAEs and in the Yang et al. (2016) and Orlitová et al. (2018) studies on low-redshift GP galaxies. Orlitová et al. also reported a discrepancy between the systemic redshift derived from the models and the observed redshifts, whereas they agree in our case. Hashimoto et al. (2015) proposed that other sources of Lyα emission than star formation, such as gravitational cooling, could account for the large intrinsic width of the Lyα emission line. Orlitová et al. (2018) tried to explain the FWHM discrepancy considering complex kinematic structures (as testified by the broad Hα and Hβ wings observed in their GP spectra), but with negative results. They concluded that the discrepancy is a model failure and not a physical effect. From our data we cannot recover possible kinematic structures. A likely cause for the mismatch is a widening of the intrinsic Lyα spectrum due to radiative transfer effects (as discussed, e.g., in Yang et al. 2016). Gronke et al. (2018) supported this scenario (also see the recent study by Li et al. 2021, where ≫100 km s$^{-1}$ observed spectra were modeled with narrow intrinsic spectra). The authors post-processed a hydrodynamic simulation of a

galactic disk using Lyα radiative transfer and showed that the spectrum is indeed widened compared to Hα due to ISM turbulence (cf. their Fig. 1, where the intrinsic spectrum of width ~100 km s$^{-1}$ is widened to ~400 km s$^{-1}$ due to radiative transfer effects within the gaseous disk). For Gronke et al. (2018), the subsequent effect of the "shell" can be taken quite literally as they consider cosmic-ray driven outflows that are likely smooth and cold (e.g., Girichidis et al. 2018), like a shell. However, in general, this can also be due to CGM effects or a lack of numerical resolution leading to a "shell" by smoothing out otherwise multiphase structure (Gronke et al. 2018). This might in principle explain the effect seen here where the intrinsic width obtained from the shell-model fitting is much wider than the width inferred from the observations.

Another important discrepant value is $N_{HI}$. The fit returns lower values than are determined from the afterglow spectroscopy in half of our sample. The two GRBs of the discrepant cases have much higher $N_{HI}^{OA}$ than the other GRBs in the golden sample, with $\log(N_{HI}^{OA}/cm^{-2}) > 21.5$ as opposed to values of $\log(N_{HI}^{OA}/cm^{-2}) \sim 20$. The escape of Lyα emission is unfavorable when the photons pass through a static high-HI density medium. In the shell model, $N_{HI}$ impacts the Lyα peak shift, with low $N_{HI}$ values corresponding to small shifts and vice versa. When we constrain its value, the model is unable to reproduce the line profile for GRBs 060926 and 070110, and results in a peak shifted to higher velocities (~400 km s$^{-1}$).

When all the LGRBs with a measured $N_{HI}^{OA}$ (reported in Tanvir et al. 2019) are considered, these high $N_{HI}^{OA}$ values are common among LAE-LGRBs (Fig. 11). Even if the fraction of LAE-LGRBs from sub-DLA/Lyman-limit systems ($\log(N_{HI}^{OA}/cm^{-2}) < 20.3$; 22% ± 11%) appears to be higher (considering also the poor statistics) than that from DLA ($\log(N_{HI}^{OA}/cm^{-2}) > 20.3$; 15% ± 4%), there is no strong evidence for Lyα suppression associated with high $N_{HI}^{OA}$. Moreover, the Lyα luminosity does not appear to depend on $N_{HI}^{OA}$ (see Fig. 11, right panel).

The discrepancy among the observed $N_{HI}^{OA}$ values and those found by the fitting, as well as the high $N_{HI}^{OA}$ values found among LAE-LGRBs, could be explained by the difference between

the medium probed by Ly$\alpha$ photons and the afterglow emission. LGRB lines of sight may go through regions with high $N_{HI}^{OA}$, while Ly$\alpha$ photons may have escaped through lower $N_{HI}$ lines of sight from the GRB young star-forming region. This would imply an anisotropic environment. The theoretical studies of resonant line transfer through simplified anisotropic geometries (e.g., Dijkstra & Kramer 2012; Behrens et al. 2014; Eide et al. 2018) as well as turbulent medium with ionized channels (e.g., Kimm et al. 2019; Kakiichi & Gronke 2021) show that Ly$\alpha$ spectra are shaped by the lowest-density pathways. This predicts that in general, the column density probed by Ly$\alpha$ is lower than or equal to the density observed along the line of sight, in agreement with what we find for the LAE-LGRB modeling presented here. Placing constraints on the HI anisotropy is important also for the consequences this can have on the ionizing escape fraction of galaxies (e.g., Vielfaure et al. 2020). Already from this work where we find $\log(N_{HI}^{OA}(\text{LGRB})/N_{HI}(\text{Ly}\alpha)) \sim 2$–3 in two of the cases, we can speculate that it is difficult to explain this high anisotropy by purely turbulent driving. For reasonable Mach numbers $\mathcal{M} \lesssim 5$, studies show only ~1 dex difference (for ~10% of the sight lines; Safarzadeh & Scannapieco 2016), which suggests the need of radiative or mechanical feedback causing higher anisotropies (Kimm et al. 2019; Kakiichi & Gronke 2021; Cen 2020).

The study of LGRB lines of sight shows that from a statistical point of view, low-density channels surrounding star-forming regions should be rare (Tanvir et al. 2019). This is consistent with LGRB progenitors being massive stars typically formed in dense molecular clouds and residing in gas-rich star-forming galaxies (e.g., Jakobsson et al. 2006). The observation of the time-variability of fine-structure transitions in LGRB afterglow spectra has shown that the distance of the dominant absorbing clouds ranges from ~50 pc to 1 kpc (e.g., Vreeswijk et al. 2013). The neutral gas probed by LGRB afterglows should therefore be connected to the star-forming regions in which LGRBs explode. Because these regions are also expected to be the main birthplaces of Ly$\alpha$ photons, the properties of LGRB lines of sight could represent the average conditions of the environment surrounding Ly$\alpha$ emission regions. If we consider the proportion of sub-DLA LGRBs on average as representative of the low-density channels through which the Ly$\alpha$ radiation can favorably escape, they would correspond to 12% of all the possible lines of sight. Interestingly, previous studies of anisotropic Ly$\alpha$ escape show that these low opening angles can in principle be sufficient to set the properties of the isotropically emergent spectrum (Dijkstra et al. 2016; Eide et al. 2018; Kakiichi et al. 2018).

The LGRB sites indicate regions with high star-formation activity, and therefore regions that produce a significant number of Ly$\alpha$ photons. However, another possibility to explain the $N_{HI}$ discrepancy is that most of the observed Ly$\alpha$ photons may not originate from the young star-forming regions probed by GRBs (as also proposed by Vreeswijk et al. 2004 for GRB 030323). One way to investigate these two scenarios would be through galaxy simulations, studying the global, the line of sight, and the Ly$\alpha$ emission properties and comparing them to those determined by LGRB afterglow and host observations. This is beyond the scope of this paper and will be the object of a further paper.

The situation is even more complex if we consider the morphological properties of the host galaxies in more detail. The images of the host galaxies reveal a patchy and irregular morphology (see Fig. 7). The HST images obtained for GRB 021004 (Fynbo et al. 2005) show that the GRB position is superimposed on the strongest UV emitting region. This is not the case for GRB 011211 (Fynbo et al. 2003), where

the Ly$\alpha$ emission envelops the entire galaxy, but the peak of emission is clearly separated from the GRB position. Interestingly, this does not prevent the match between the shell model and the observations. A complex morphology is also found for other LAE-LGRB hosts (GRB 000926 Fynbo et al. 2002, GRB 050315 and GRB 061222A Blanchard et al. 2016). These irregular and clumpy systems may represent mergers (e.g., Conselice et al. 2003) or just clumps of star formation. The presence of a companion galaxy ~14 kpc away from the GRB 021004 host and the high velocity absorptions found in its afterglow spectra may suggest interaction of the two galaxies and the presence of outflowing gas. In this light, the success of a very simple configuration, as the one of the shell model, in reproducing the Ly$\alpha$ profiles and the observed parameters is maybe surprising.

### 5.2. Comparison to LyC leakers

In Fig. 12 we compare properties of LAE-LGRBs to those of known LyC leakers in the literature. For this comparison, we consider the LyC leakers at $z \sim 3.1$ reported by Fletcher et al. (2019) and Ion2 ($z \sim 3.2$, de Barros et al. 2016; Vanzella et al. 2016). We also consider those detected in the local Universe analyzed in Verhamme et al. (2017), as well as the green peas studied in Izotov et al. (2018a,b). The compared properties are the Ly$\alpha$ escape fraction, the rest-frame Ly$\alpha$ equivalent width, the Ly$\alpha$ red peak velocity ($v_{peak}$), and the [OIII]/[OII] ratio ($O_{32}$). The last three properties have been proposed in the literature to correlate with ionizing and Ly$\alpha$ photon leakage. Verhamme et al. (2017) have pointed out rest-frame EW(Ly$\alpha$) > 70 Å, $f_{esc}(\text{Ly}\alpha) > 0.3$, $v_{peak} < 150$ km s$^{-1}$ and $O_{32} > 4$ as good indirect indicators of high LyC leakage ($f_{esc}(\text{LyC}) > 5\%$). The $O_{32}$ ratio is a proxy of the ionization state of the gas in the star-forming regions and has been proposed as a marker of density-bounded HII regions (Jaskot & Oey 2014; Nakajima et al. 2013; Nakajima & Ouchi 2014; Stasinska et al. 2015). Its positive correlation with the escape fraction of ionizing and Ly$\alpha$ photons have been found in several studies for local GPs or high-redshift LAEs (e.g., Nakajima & Ouchi 2014; Izotov et al. 2016, 2018a,b; de Barros et al. 2016; Verhamme et al. 2017; Fletcher et al. 2019).

The comparisons show that except for GRBs 060707 and 060605, which have a very high $v_{peak}$, the LAE-LGRBs fall in the same parameter space as LyC leakers and follow the correlations between the indirect indicators found by Verhamme et al. (2017). Following their study, we can cut these plots into two regions corresponding to strong LyC leakers ($f_{esc}(\text{LyC}) > 5\%$, red rectangle) and weak LyC leakers ($f_{esc}(\text{LyC}) < 5\%$, blue rectangle). All LAE-LGRBs fall in the category of the weak LyC leakers except for GRB 021004, which appears systematically to agree well with the region of the strong leakers. In panel b of Fig. 12, we also superimpose the distribution of [OIII]/[OII] ratio for the GRB sample of Krühler et al. (2015) and GRB 081121 reported here. For the majority of the GRBs, this ratio is about two, with seven cases at $O_{32} > 4$. GRB 021004 has a high value of $O_{32} > 10$ and is the strongest LAE of our golden sample (with $f_{esc}(\text{Ly}\alpha)$ of 60%) in agreement with potential high $f_{esc}(\text{LyC})$. Nevertheless, the Ly$\alpha$ profile of GRB 021004 is a single peak with no residual flux at the Ly$\alpha$ line center. This is not the typical line shape observed for confirmed LyC emitters, which have the tendency to show double- or triple-peak profiles (Verhamme et al. 2017; Rivera-Thorsen et al. 2017; Vanzella et al. 2020). Our shell-model fitting also suggests that the column densities are too high
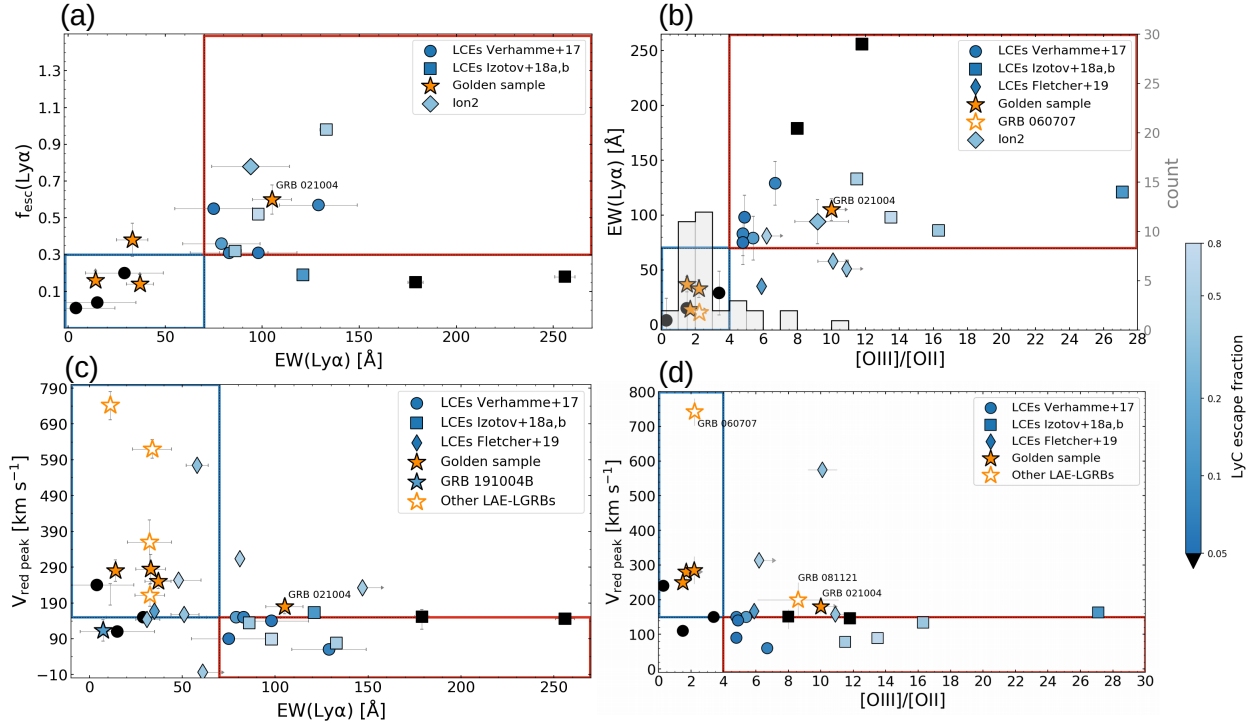
**Fig. 12.** Comparison of the properties of the four LAE-LGRBs from the golden sample (orange stars) to other LAE-LGRBs and Lyman continuum emitters (LCEs) reported in Verhamme et al. (2017), Izotov et al. (2018a,b), and Fletcher et al. (2019). The white stars with orange contours represent LAE-LGRBs for which information on indirect indicators is available. The color bar codes the $f_{esc}$(LyC) of the LCEs. In the four panels the blue rectangle represents the space parameter of low LyC leakage ($f_{esc}$(LyC) < 5%), while the red rectangle marks the space parameter of high LyC leakage ($f_{esc}$(LyC) > 5%), as defined by Verhamme et al. (2017). *Panel a*: escape fraction of Ly$\alpha$ photons as a function of the rest-frame Ly$\alpha$ EW. *Panel b*: correlation between the rest-frame Ly$\alpha$ EW and the [OIII]/[OII] ratio. The distribution of [OIII]/[OII] ratio for the LGRB host galaxies from Krühler et al. (2015) is also superimposed. *Panel c*: relation between the Ly$\alpha$ red peak velocity shift and the rest-frame Ly$\alpha$ EW. *Panel d*: correlation between the Ly$\alpha$ red peak velocity shift and the [OIII]/[OII] ratio.

to allow LyC photons to escape. However, this model already predicted a similar HI column density ($\log(N_{HI}/cm^{-2}) \approx 19-20$) for four green pea galaxies out of the five with detected LyC emission reported in Yang et al. (2017). This suggests that LyC emission can escape through holes in the ISM even if the Ly$\alpha$ photons probe denser neutral gas. The only LAE-LGRB for which LyC leakage has been detected along the LGRB line of sight ($f_{esc}$(LyC) = $0.35^{+0.10}_{-0.11}$) is GRB 191004B (Vielfaure et al. 2020). However, Fig. 12 (panel c) shows that it does not fall in the high escape fraction region, but in the lower left area. The reasons could be that (i) $f_{esc}$(LyC) is lower at the scale of the galaxy than along the LGRB line of sight, and (ii) the indicators of strong LyC leakage evolves with redshift. As a comparison, the LyC emitters from Fletcher et al. (2019) are found out of the high escape region (red rectangle). They show lower rest-frame EW(Ly$\alpha$) and higher $v_{peak}$ than the local LyC emitters, whereas their escape fraction of ionizing photons is significantly higher ($f_{esc}$(LyC) = 15−60%). This could also suggest that strong LyC leakers span a wider parameter space than predicted by the study of local LyC emitters. Overall, it is clear that this type of studies is still limited by the poor statistics, and the current results show the difficulty of characterizing LyC leakage based on these properties alone.

## 6. Conclusions

We have studied the Ly$\alpha$ emission of LGRB host galaxies. First, we provided a new census of LAEs among LGRB host galaxies. To date, there are 29 LAE-LGRBs. The fraction of LAEs

among LGRB hosts varies from ∼10% to 40% depending on the sample and threshold considered. These statistics are lower than those found for LBG samples at similar redshift range, but they become comparable when only LAEs with rest-frame EW(Ly$\alpha$) > 20 Å are taken into account. These results can be explained by the different selection criteria of the parent samples and by the shallower spectral observations of LGRB samples compared to LBG samples. We compared the properties of LAE-LGRBs to those of LGRB hosts in general and find evidence of Ly$\alpha$ emission suppression in dusty hosts. We showed that because LGRB hosts are not selected following usual criteria and techniques used for Ly$\alpha$ emission searches in star-forming galaxies, they probe the regime of Ly$\alpha$ emission of the bulk of the UV-selected galaxy population at intermediate (and possibly high) redshifts.

We then selected a subsample of four LGRBs that allowed the combination of the emission properties of the host galaxies with the information on the ISM probed by the afterglow. We fit the Ly$\alpha$ emission of these galaxies using the shell model that we constrained and tested taking advantage of ancillary observational properties. We find that without priors on the parameters, the shell model succeeds in reproducing the four profiles. However, the properties predicted by the model differ from the observed properties for the two cases (GRBs 060926 and 070110) that have a high $N_{HI}^{OA}$. Similarly, constraining the model parameters using the values determined from the observations, the shell model succeeds in reproducing the two lower $N_{HI}^{OA}$ cases, but fails for the two highest $N_{HI}^{OA}$ cases. These results may arise because the $N_{HI}$ of the gas probed by the LGRB

afterglow ($N_{\mathrm{HI}}^{\mathrm{OA}}$(LGRB)) is not representative of that encountered by the Ly$\alpha$ photons ($N_{\mathrm{HI}}$(Ly$\alpha$)); specifically, we find $N_{\mathrm{HI}}$(Ly$\alpha$) $\leq$ $N_{\mathrm{HI}}^{\mathrm{OA}}$(LGRB). This is coherent with the picture that Ly$\alpha$ photons escape preferentially through low-column density channels and that the shell model relies on a too simplistic gas configuration. A compatible scenario would be provided by a turbulent medium widening the intrinsic Ly$\alpha$ spectrum, and an anisotropic HI distribution allowing the escape of Ly$\alpha$ photons through lower density channels than are probed by the LGRB sightline. This description is consistent with mechanical and radiative feedback mechanisms produced by starburst and supernova that clear out channels from the active regions within the galaxies. Nevertheless, we stress that the afterglows of most LGRBs and LAE-LGRBs show log($N_{\mathrm{HI}}^{\mathrm{OA}}$/cm$^{-2}$) > 20.3. This implies that the Ly$\alpha$ photons produced by massive stars will likely predominantly be surrounded by high-density gas, and low-density channels should be rare ($\sim$10%). Another possibility is that even if dense star-forming regions produce a high fraction of Ly$\alpha$ photons, most of the Ly$\alpha$ escaping photons come from less dense regions throughout the galaxy. We will investigate both scenarios further in the future, taking advantage of galaxy simulations.

Finally, we compared the properties of the LAE-LGRBs with those of LyC leakers in the literature. Specifically, we considered the commonly used Ly$\alpha$ indirect indicators of LyC leakage at low redshift and find that only one LAE-LGRB (GRB 021004) has the values of a strong LyC leaker. However, LyC reported in the literature at $z \sim 3$ span a wider range of parameter values than those commonly considered to indicate strong LyC leakage, which makes the use of these indicators at least at high redshift doubtful.

While our work demonstrated that LGRBs can be used to probe LAEs and radiative transfer models, our results are limited by the small statistics. Further LAE-LGRB host galaxy observations will help solve this issue. To this purpose, we have been awarded X-shooter time to observe 11 more LAE-LGRBs reported in the census. In the future, it will be possible to extend these studies to higher redshift. New space missions optimized for high-redshift GRB detections, such as the Gamow Explorer (White 2020) and THESEUS[5] (Amati et al. 2018) projects, in synergy with ELT and JWST observations, will enable us to perform similar studies at $z > 3.5$ that will have better statistics.

# References

Ahn, S.-H. 2000, ApJ, 530, L9

Ahn, S.-H. 2004, ApJ, 601, L25
Ahn, S.-H., Lee, H.-W., & Lee, H. M. 2003, MNRAS, 340, 863
Amati, L., O'Brien, P., Götz, D., et al. 2018, Adv. Space Res., 62, 191
Appenzeller, I., Fricke, K., Fürtig, W., et al. 1998, The Messenger, 94, 1
Arabsalmani, M., Le Floc'h, E., Dannerbauer, H., et al. 2018, MNRAS, 476, 2332
Bagley, M. B., Scarlata, C., Henry, A., et al. 2017, ApJ, 837, 11
Behrens, C., Dijkstra, M., & Niemeyer, J. C. 2014, A&A, 563, A77
Berger, E., & Rauch, M. 2008, GCN, 8542, 1
Blanchard, P. K., Berger, E., & Fai Fong, W. 2016, ApJ, 817, 144
Bolmer, J., Ledoux, C., Wiseman, P., et al. 2019, A&A, 623, A43
Brocklehurst, M. 1971, MNRAS, 153, 471
Castro-Tirado, A. J., Møller, P., García-Segura, G., et al. 2010, A&A, 517, A61
Cen, R. 2020, ApJ, 889, L22
Chabrier, G. 2003, PASP, 115, 763
Chen, H.-W. 2012, MNRAS, 419, 3039
Chen, H., Prochaska, J. X., Ramirez-Ruiz, E., et al. 2007, ApJ, 663, 420
Chisholm, J., Tremonti, C. A., Leitherer, C., et al. 2015, ApJ, 811, 149
Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183
Cucchiara, A., Fumagalli, M., Rafelski, M., et al. 2015, ApJ, 804, 51
D'Avanzo, P., Perri, M., Fugazza, D., et al. 2010, A&A, 522, A20
de Barros, S., Vanzella, E., Amorín, R., et al. 2016, A&A, 585, A51
D'Elia, V., Fynbo, J. P. U., Izzo, L., et al. 2019, GCN, 25956, 1
Dessauges-Zavadsky, M., D'Odorico, S., Schaerer, D., et al. 2010, A&A, 510, A26
de Ugarte Postigo, A., Castro-Tirado, A. J., Gorosabel, J., et al. 2005, A&A, 443, 841
Dijkstra, M., & Kramer, R. 2012, MNRAS, 424, 1672
Dijkstra, M., Gronke, M., & Sobral, D. 2016, ApJ, 823, 74
Eide, M. B., Gronke, M., Dijkstra, M., & Hayes, M. 2018, ApJ, 856, 156
Elíasdóttir, Á., Fynbo, J. P. U., Hjorth, J., et al. 2009, ApJ, 697, 1725
Erb, D. K., Steidel, C. C., Trainor, R. F., et al. 2014, ApJ, 795, 33
Fiore, F., D'Elia, V., Lazzati, D., et al. 2005, ApJ, 624, 853
Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, ApJ, 878, 87
Fox, D. W. 2002, GCN, 1564, 1
Friis, M., De Cia, A., Krühler, T., et al. 2015, MNRAS, 451, 167
Fynbo, J. U., Møller, P., & Warren, S. J. 1999, MNRAS, 305, 849
Fynbo, J. P. U., Møller, P., Thomsen, B., et al. 2002, A&A, 388, 425
Fynbo, J. P. U., Jakobsson, P., Møller, P., et al. 2003, A&A, 406, L63
Fynbo, J. P. U., Gorosabel, J., Smette, A., et al. 2005, ApJ, 633, 317
Fynbo, J. P. U., Jakobsson, P., Prochaska, J. X., et al. 2009, ApJ, 185, 526
Gehrels, N., & Razzaque, S. 2013, Front. Phys., 8, 661
Girichidis, P., Naab, T., Hanasz, M., & Walch, S. 2018, MNRAS, 479, 3042
Graham, J. F., & Fruchter, A. S. 2017, ApJ, 834, 170
Gronke, M. 2017, A&A, 608, A139
Gronke, M., Bull, P., & Dijkstra, M. 2015, ApJ, 812, 123
Gronke, M., Girichidis, P., Naab, T., & Walch, S. 2018, ApJ, 862, L7
Hartoog, O. E., Malesani, D., Fynbo, J. P. U., et al. 2015, A&A, 580, A139
Hashimoto, T., Verhamme, A., Ouchi, M., et al. 2015, ApJ, 812, 157
Heintz, K. E., Ledoux, C., Fynbo, J. P. U., et al. 2019, A&A, 621, A20
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, ApJ, 809, 19
Hjorth, J., Sollerman, J., Møller, P., et al. 2003, Nature, 423, 847
Hjorth, J., Malesani, D., Jakobsson, P., et al. 2012, ApJ, 756, 187
Holland, S. T., Barthelmy, S. D., Burrows, D. N., et al. 2006, GCN, 5612, 1
Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016, MNRAS, 461, 3683
Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, MNRAS, 474, 4514
Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, MNRAS, 478, 4851
Jakobsson, P., Hjorth, J., Fynbo, J. P. U., et al. 2003, A&A, 408, 941
Jakobsson, P., Hjorth, J., Fynbo, J. P. U., et al. 2004, A&A, 427, 785
Jakobsson, P., Björnsson, G., Fynbo, J. P. U., et al. 2005, MNRAS, 362, 245
Jakobsson, P., Fynbo, J. P. U., Ledoux, C., et al. 2006, A&A, 460, L13
Jakobsson, P., Hjorth, J., Malesani, D., et al. 2012, ApJ, 752, 62
Japelj, J., Covino, S., Gomboc, A., et al. 2015, A&A, 579, A74
Japelj, J., Vergani, S. D., Salvaterra, R., et al. 2016, A&A, 590, A129
Jaskot, A. E., & Oey, M. S. 2014, ApJ, 791, L19
Kakiichi, K., & Gronke, M. 2021, ApJ, 908, 30
Kakiichi, K., Ellis, R. S., Laporte, N., et al. 2018, MNRAS, 479, 43
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kimm, T., Blaizot, J., Garel, T., et al. 2019, MNRAS, 486, 2215
Kruehler, T., & Schady, P. 2017, Figshare, https://doi.org/10.6084/m9.figshare.4776886.v3
Krühler, T., Malesani, D., Milvang-Jensen, B., et al. 2012, ApJ, 758, 46
Krühler, T., Malesani, D., Fynbo, J. P. U., et al. 2015, A&A, 581, A125
Kulkarni, S. R., Djorgovski, S. G., Ramaprakash, A. N., et al. 1998, Nature, 393, 35
Kunth, D., Mas-Hesse, J. M., Terlevich, E., et al. 1998, A&A, 334, 11
Laursen, P., Sommer-Larsen, J., & Andersen, A. C. 2009, ApJ, 704, 1640

Leclercq, F., Bacon, R., Wisotzki, L., et al. 2017, A&A, 608, A8
Li, Z., Steidel, C. C., Gronke, M., & Chen, Y. 2021, MNRAS, 502, 2389
Lidman, C., Hayes, M., Jones, D. H., et al. 2012, MNRAS, 420, 1946
Lyman, J. D., Levan, A. J., Tanvir, N. R., et al. 2017, MNRAS, 467, 1795
Malesani, D., Jaunsen, A. O., & Vreeswijk, P. M. 2007, GCN, 6015, 1
Malesani, D., Xu, D., Fynbo, J. P. U., et al. 2013, GCN, 14291, 1
Matthee, J., Sobral, D., Oteo, I., et al. 2016, MNRAS, 458, 449
Matthee, J., Sobral, D., Hayes, M., et al. 2021, MNRAS, 505, 1382
Milvang-Jensen, B., Fynbo, J. P. U., Malesani, D., et al. 2012, ApJ, 756, 25
Mirabal, N., Halpern, J. P., Chornock, R., et al. 2003, ApJ, 595, 15
Modigliani, A., Goldoni, P., Royer, F., et al. 2010, in The X-shooter pipeline,
    SPIE Conf. Ser., 7737, 773728
Møller, P., Fynbo, J. P. U., Hjorth, J., et al. 2002, A&A, 396, L21
Nakajima, K., & Ouchi, M. 2014, MNRAS, 442, 900
Nakajima, K., Ouchi, M., Shimasaku, K., et al. 2013, ApJ, 769, 3
Neufeld, D. A. 1990, ApJ, 350, 216
Orlitová, I., Verhamme, A., Henry, A., et al. 2018, A&A, 616, A60
Osterbrock, D. E. 1989, Astrophysics of Gaseous Nebulae and Active Galactic
    Nuclei
Ouchi, M., Ono, Y., Egami, E., et al. 2009, ApJ, 696, 1164
Palmerio, J. T., Vergani, S. D., Salvaterra, R., et al. 2019, A&A, 623, A26
Pei, Y. C. 1992, ApJ, 395, 130
Pentericci, L., Grazian, A., Fontana, A., et al. 2007, A&A, 471, 433
Pentericci, L., Grazian, A., Scarlata, C., et al. 2010, A&A, 514, A64
Perley, D. A., Cenko, S. B., Bloom, J. S., et al. 2009, AJ, 138, 1690
Perley, D. A., Levan, A. J., Tanvir, N. R., et al. 2013, ApJ, 778, 128
Perley, D. A., Tanvir, N. R., Hjorth, J., et al. 2016a, ApJ, 817, 8
Perley, D. A., Krühler, T., Schulze, S., et al. 2016b, ApJ, 817, 7
Planck Collaboration XIII. 2016, A&A, 594, A13
Prochaska, J. X., Chen, H.-W., Dessauges-Zavadsky, M., & Bloom, J. S. 2007,
    ApJ, 666, 267
Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS, 175, 48
Rivera-Thorsen, T. E., Hayes, M., Östlin, G., et al. 2015, ApJ, 805, 14
Rivera-Thorsen, T. E., Dahle, H., Gronke, M., et al. 2017, A&A, 608, L4
Safarzadeh, M., & Scannapieco, E. 2016, ApJ, 832, L9
Salvaterra, R., Valle, M. D., Campana, S., et al. 2009, Nature, 461, 1258
Salvaterra, R., Ferrara, A., & Dayal, P. 2011, MNRAS, 414, 847
Salvaterra, R., Maio, U., Ciardi, B., & Campisi, M. A. 2013, MNRAS, 429, 2718
Savaglio, S. 2006, New J. Phys., 8, 195
Schaerer, D., & Verhamme, A. 2008, A&A, 480, 369
Schaerer, D., Hayes, M., Verhamme, A., & Teyssier, R. 2011, A&A, 531, A12
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schulze, S., Chapman, R., Hjorth, J., et al. 2015, ApJ, 808, 73

Selsing, J., Malesani, D., Goldoni, P., et al. 2019, A&A, 623, A92
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588,
    65
Sobral, D., Matthee, J., Darvish, B., et al. 2015, ApJ, 808, 139
Starling, R. L. C., Wijers, R. A. M. J., Hughes, M. A., et al. 2005, MNRAS, 360,
    305
Stasinska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, A&A, 576, A83
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728
Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, ApJ, 604, 534
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2011, ApJ, 736, 160
Tanvir, N. R., Fox, D. B., Levan, A. J., et al. 2009, Nature, 461, 1254
Tanvir, N. R., Fynbo, J. P. U., de Ugarte Postigo, A., et al. 2019, MNRAS, 483,
    5380
Thöne, C. C., Campana, S., Lazzati, D., et al. 2011, MNRAS, 414, 479
Totani, T., Kawai, N., Kosugi, G., et al. 2006, PASJ, 58, 485
Troja, E., Cusumano, G., O'Brien, P. T., et al. 2007, ApJ, 665, 599
van Dokkum, P. G. 2001, PASP, 113, 1420
Vanzella, E., de Barros, S., Vasei, K., et al. 2016, ApJ, 825, 41
Vanzella, E., Caminha, G. B., Calura, F., et al. 2020, MNRAS, 491, 1093
Vergani, S., Piranomonte, S., Petitjean, P., et al. 2011, Astron. Nachr., 332,
    292
Vergani, S. D., Palmerio, J., Salvaterra, R., et al. 2017, A&A, 599, A120
Verhamme, A., Schaerer, D., & Maselli, A. 2006, A&A, 460, 397
Verhamme, A., Schaerer, D., Atek, H., & Tapken, C. 2008, A&A, 491, 89
Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, A&A, 578, A7
Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, A&A, 597, A13
Vernet, J., Kerber, F., Mainieri, V., et al. 2009, Proc. IAU, 5, 535
Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, A&A, 536, A105
Vielfaure, J. B., Vergani, S. D., Japelj, J., et al. 2020, A&A, 641, A30
Vreeswijk, P. M., Ellison, S. L., Ledoux, C., et al. 2004, A&A, 419, 927
Vreeswijk, P. M., Smette, A., Fruchter, A. S., et al. 2006, A&A, 447, 145
Vreeswijk, P. M., Ledoux, C., Raassen, A. J. J., et al. 2013, A&A, 549, A22
White, N. E. 2020, Yamada Conference LXXI: Gamma-ray Bursts in
    the Gravitational Wave Era 2019, Proc. of the conference held 28
    October - 1 November 2019 in Yokohama, Kanagawa, Japan, eds. T.
    Sakamoto, M. Serino, & S.Sugta, http://yokohamagrb2019.wikidot.
    com/proceedings, 51
Wiseman, P., Perley, D. A., Schady, P., et al. 2017, A&A, 607, A107
Wisotzki, L., Bacon, R., Blaizot, J., et al. 2016, A&A, 587, A98
Wofford, A., Leitherer, C., & Salzer, J. 2013, ApJ, 765, 118
Yang, H., Malhotra, S., Gronke, M., et al. 2016, ApJ, 820, 130
Yang, H., Malhotra, S., Gronke, M., et al. 2017, ApJ, 844, 171
Zitrin, A., Labbé, I., Belli, S., et al. 2015, ApJ, 810, L12

## Appendix A: Tables of the properties of the objects in the samples

The following tables present the LGRB host galaxies and afterglows from the TOUGH and XHG samples, the literature, and this work, for which the Lyα emission line has been detected or an upper limit has been estimated. *GRB* and *Redshift* are for the name of the LGRB and its redshift. When observations of the host galaxy or the optical afterglow are available, we provide the name of the spectrographs used in the columns *Hosts* and *OA*, respectively. "no" indicates that no observation is available. "F1"/"F2" is for VLT/FORS1/2 and "xsh" for VLT/X-shooter spectrograph. *Lyα Host* and *Lyα OA* inform about the detection of the Lyα line in the host or afterglow spectra, respectively. No information is provided when no spectra are available to verify the presence of the line. *F(Lyα)* and L(Lyα) correspond to the Lyα line flux and luminosity (respectively) retrieved from the literature (see references in column *Refs*) or derived in this work. *EW(Lyα)* corresponds to the rest-frame Lyα equivalent width. $N_{HI}$ is for the neutral hydrogen column density determined from

the GRB afterglow. $M_{UV}$ is the UV magnitude of the LGRB host galaxy. $M_*$ is its stellar mass. *R* is for the apparent magnitude of the host in the *R* band. *SFR* is the star-formation rate. $\log(sSFR/yr^{-1})$ is the specific star-formation rate. *E(B-V)* is the host extinction. $12 + \log(O/H)$ is the oxygen abundance. *Z* is the metallicity determined from the absorption lines detected in the GRB afterglow spectrum.

**References:** (1): Milvang-Jensen et al. (2012); (2): This work; (3): Schulze et al. (2015); (4): Perley et al. (2016a); (5): Tanvir et al. (2019); (6):Krühler et al. (2015); (7): Palmerio et al. (2019); (8): Selsing et al. (2019); (9): Vielfaure et al. (2020); (10): Jakobsson et al. (2003); (11): Fynbo et al. (2005); (12): Vreeswijk et al. (2004); (13): Jakobsson et al. (2004); (14): D'Avanzo et al. (2010); (15): Perley et al. (2013); (16): de Ugarte Postigo et al. (2005); (17): Kruehler & Schady (2017); (18): Jakobsson et al. (2006); (19): Bolmer et al. (2019); (20): Heintz et al. (2019); (21): Cucchiara et al. (2015); (22): Savaglio (2006); (23): Vreeswijk et al. (2006); (24): Prochaska et al. (2007); (25): Thöne et al. (2011); (26): Elíasdóttir et al. (2009); (27): Chen et al. (2007).

**Table A.1.** LGRBs in the TOUGH sample with a Lyα-emission detection (in bold) or an upper limit.

| GRB | Redshift | Host | OA | Lyα Host | Lyα OA | EW(Lyα) [Å] | F(Lyα) [$10^{-17}$ erg s$^{-1}$ cm$^{-2}$] | L(Lyα) [$10^{42}$ erg s$^{-1}$] | Refs |
|---|---|---|---|---|---|---|---|---|---|
| **050315** | 1.9500 | F1 | no | yes | ... | 9.2 ± 2.8 | 2.34 ± 0.68 | 0.64 ± 0.19 | (1) |
| 050401 | 2.8983 | F1 | F2 | no | no | < 16.3 | < 1.12 | < 0.80 | (1) |
| 050714B | 2.4383 | xsh | F1 | no | no | < 17.08 | < 0.74 | < 0.36 | (2) |
| 050730 | 3.9686 | F1 | UVES | no | no | ... | < 0.87 | < 1.32 | (1) |
| 050819 | 2.5042 | xsh | no | no | ... | < 9.30 | < 1.74 | < 0.91 | (2) |
| 050820A | 2.6147 | F1 | UVES | no | no | < 8.1 | < 1.01 | < 0.56 | (1) |
| 050908 | 3.3467 | F1 | F1 | no | no | ... | < 0.64 | < 0.64 | (1) |
| 050915A | 2.5273 | xsh | F1 | no | no | < 9.45 | < 0.61 | < 0.33 | (2) |
| 050922C | 2.1992 | F1 | AlFOSC | no | no | ... | < 1.93 | < 0.70 | (1) |
| 051001 | 2.4296 | xsh | F1 | no | no | < 7.21 | < 0.84 | < 0.41 | (2) |
| 060115 | 3.5328 | F1 | F1 | no | no | ... | < 1.18 | < 1.35 | (1) |
| 060526 | 3.2213 | F1 | F1 | no | no | < 19.5 | < 0.79 | < 0.73 | (1) |
| 060604 | 2.1357 | F1, xsh | AlFOSC | no | no | < 12.1 | < 0.90 | < 0.31 | (1) |
| **060605** | 3.7730 | F1 | PMAS | yes | no | 33.7 ± 10.5 | 1.70 ± 0.27 | 2.28 ± 0.36 | (1) |
| 060607A | 3.0749 | F1 | UVES | no | no | ... | < 0.73 | < 0.60 | (1) |
| **060707** | 3.4240 | F1, xsh | F1 | yes | no | 11.2 ± 2.3 | 1.65 ± 0.31 | 1.75 ± 0.33 | (1) |
| **060714** | 2.7108 | F1 | F1 | no | yes | ... (< 26.3) | 1.73 (< 8.1) | 1.10 (< 0.49) | (1), (18) |
| 060805A | 2.3633 | xsh | no | no | ... | < 14.84 | < 0.77 | < 0.35 | (2) |
| 060814 | 1.9223 | xsh | F1, F2 | no | no | < 7.12 | < 2.33 | < 0.64 | (2) |
| **060908** | 1.8836 | F1 | F1, F2 | yes | no | 40.4 ± 6.7 | 7.78 ± 0.95 | 1.94 ± 0.24 | (1) |
| 061110B | 3.4344 | F1 | F1 | no | no | < 10.7 | < 0.66 | < 0.71 | (1) |
| 070103 | 2.6208 | xsh | F2 | no | no | < 19.40 | < 1.42 | < 0.83 | (2) |
| **070110** | 2.3523 | F1, xsh | F1 | yes | yes | 31.8 ± 4.3 | 4.0 ± 0.4 | 1.73 ± 0.17 | (1) |
| 070129 | 2.3384 | xsh | F1, F2 | no | no | < 13.92 | < 1.94 | < 0.86 | (2) |
| 070224 | 1.9922 | xsh | no | no | ... | < 75.8 | < 2.12 | < 0.64 | (2) |
| 070328 | 2.0627 | xsh | no | no | ... | < 9.52 | < 0.87 | < 0.28 | (2) |
| 070419B | 1.9586 | xsh | F1 | no | no | < 90.06 | < 2.15 | < 0.62 | (2) |
| **070506** | 2.3090 | F1 | F1 | yes | no | 32.3 ± 11.8 | 1.39 ± 0.35 | 0.57 ± 0.14 | (1) |
| 070611 | 2.0394 | F1 | F2 | no | no | ... | < 0.96 | < 0.29 | (1) |
| **070721B** | 3.6298 | F1 | F2 | yes | yes | 32.5 ± 8.0 | 1.12 ± 0.16 | 1.37 ± 0.19 | (1) |
| 070802 | 2.4541 | F1, xsh | F2 | no | no | < 8.1 | < 0.89 | < 0.43 | (1) |

**Notes.** For GRB 060714, the values reported correspond to the Lyα detection from the afterglow spectrum (Jakobsson et al. 2006), but the upper limits derived from the host galaxy (Milvang-Jensen et al. 2012) are given between brackets.

**Table A.2.** Properties of the LGRBs in the TOUGH sample with a Lyα-emission detection (in bold) or an upper limit.

| GRB | $N_{\rm HI}$ [log( $N_{\rm HI}$ /$cm^{-2}$)] | $M_{\rm UV}$ [mag] | $M_*$ [log($M_*/M_\odot$)] | $R$ [mag] | Refs | SFR [$M_\odot\,yr^{-1}$] | log(sSFR/$yr^{-1}$) | E(B-V) | $12 + \log(O/H)$ | Z | Refs |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **050315** | ... | $-20.08 \pm 0.14$ | 9.7 | $24.51 \pm 0.15$ | (3),(4) | ... | ... | ... | ... | ... | ... |
| 050401 | $22.60 \pm 0.30$ | $-19.28 \pm 0.31$ | 9.61 | $26.19 \pm 0.31$ | (5),(3) | ... | ... | ... | ... | $-1.05 \pm 0.18$ | (24) |
| 050714B | ... | $-19.61 \pm 0.19$ | ... | $25.51 \pm 0.20$ | (3) | $12.9^{+14.0}_{-5.3}$ | ... | $0.21^{+0.28}_{-0.21}$ | ... | ... | (6) |
| 050730 | $22.10 \pm 0.10$ | $> -17.21$ | $< 9.46$ | $> 27.2$ | (5),(1) | ... | ... | ... | ... | $-1.96 \pm 0.11$ | (21) |
| 050819 | ... | $-21.13 \pm 0.09$ | ... | $23.99 \pm 0.09$ | (3) | $22^{+426}_{-15}$ | ... | $0.34^{+0.84}_{-0.34}$ | ... | ... | (6) |
| 050820A | $21.10 \pm 0.10$ | $-18.93 \pm 0.06$ | 8.95 | 24.8 | (5),(1),(17) | 0.9 | $-9.00$ | ... | ... | $-0.76 \pm 0.13$ | (6),(21) |
| 050908 | $17.60 \pm 0.10$ | $-18.18 \pm 0.27$ | $< 9.91$ | 27.7 | (5),(1) | ... | ... | ... | ... | ... | ... |
| 050915A | ... | $-20.47 \pm 0.15$ | 10.66 | $24.70 \pm 0.16$ | (3),(17) | 1.9 | $-10.39$ | $0.84^{+0.61}_{-0.60}$ | ... | ... | (6) |
| 050922C | $21.55 \pm 0.10$ | $> -17.95$ | $< 9.01$ | $> 26.29$ | (5),(1) | ... | ... | ... | ... | $-1.88 \pm 0.14$ | (21) |
| 051001 | ... | $-20.55 \pm 0.12$ | 10.08 | $24.53 \pm 0.13$ | (3),(4) | $110^{+124}_{-59}$ | $-8.04^{+0.33}_{-0.33}$ | $0.58^{+0.28}_{-0.28}$ | ... | ... | (6) |
| 060115 | $21.50 \pm 0.10$ | $-18.61 \pm 0.27$ | 9.43 | 27.1 | (5),(1) | ... | ... | ... | ... | $> -1.53$ | (21) |
| 060526 | $20.00 \pm 0.15$ | $> -17.36$ | 9.30 | $> 27.0$ | (5),(1) | ... | ... | ... | ... | $-1.09 \pm 0.24$ | (25) |
| 060604 | ... | $-19.23 \pm 0.17$ | ... | $25.62 \pm 0.18$ | (3) | $7.2^{+9.4}_{-3.6}$ | ... | $0.38^{+0.32}_{-0.27}$ | $8.10^{+0.28}_{-0.35}$ | ... | (6) |
| **060605** | $18.90 \pm 0.40$ | $-17.94 \pm 0.20$ | $< 9.97$ | $> 26.5$ | (5),(1) | ... | ... | ... | ... | ... | ... |
| 060607A | $16.95 \pm 0.03$ | $> -15.52$ | $< 9.45$ | $> 28.05$ | (5),(3) | ... | ... | ... | ... | ... | ... |
| **060707** | $21.00 \pm 0.20$ | $-20.78 \pm 0.06$ | 9.99 | $25.01 \pm 0.06$ | (5),(3) | $19.9^{+48.0}_{-14.3}$ | $-8.69^{+0.53}_{-0.56}$ | ... | ... | $\geq -1.69$ | (6),(21) |
| **060714** | $21.80 \pm 0.10$ | $-18.88 \pm 0.28$ | 9.25 | $26.46 \pm 0.28$ | (5),(3) | ... | ... | ... | ... | $\geq -0.97$ | (21) |
| 060805A | ... | $-19.79 \pm 0.13$ | ... | $25.26 \pm 0.14$ | (3) | $9.0^{+3.9}_{-2.5}$ | ... | $0.00^{+0.16}_{-0.00}$ | ... | ... | (6) |
| 060814 | ... | $-21.46 \pm 0.10$ | 10.00 | $22.96 \pm 0.11$ | (3),(17) | $47.5^{+72.5}_{-15.6}$ | $-8.32^{+0.40}_{-0.17}$ | $0.17^{+0.39}_{-0.17}$ | $8.46^{+0.10}_{-0.16}$ | ... | (7),(6) |
| **060908** | ... | $-18.93 \pm 0.17$ | 9.15 | $25.66 \pm 0.18$ | (3),(4) | ... | ... | ... | ... | ... | ... |
| 061110B | $22.35 \pm 0.10$ | $-19.82 \pm 0.29$ | $< 9.47$ | 26.0 | (5),(1) | ... | ... | ... | ... | $\geq -1.84$ | (21) |
| 070103 | ... | $-21.03 \pm 0.14$ | ... | $24.21 \pm 0.14$ | (3) | $43^{+162}_{-17}$ | ... | $0.00^{+0.50}_{-0.00}$ | ... | ... | (6) |
| **070110** | $21.70 \pm 0.10$ | $-19.81 \pm 0.11$ | $< 9.16$ | $25.19 \pm 0.11$ | (5),(3) | $8.9^{+10.9}_{-2.8}$ | $< -7.86$ | $0.00^{+0.38}_{-0.00}$ | ... | $\geq -1.32$ | (6),(21) |
| 070129 | ... | $-20.75 \pm 0.11$ | 10.15 | $24.23 \pm 0.12$ | (3),(4) | $20^{+28}_{-7}$ | $-8.85^{+0.38}_{-0.19}$ | $0.17^{+0.35}_{-0.17}$ | ... | ... | (6) |
| 070224 | ... | $-18.71 \pm 0.29$ | ... | $26.02 \pm 0.31$ | (3) | $3.2^{+6.5}_{-2.3}$ | ... | ... | ... | ... | ... |
| 070328 | ... | $-20.17 \pm 0.12$ | 9.65 | $24.55 \pm 0.13$ | (3),(17) | $8.4^{+130.0}_{-4.2}$ | $-8.76^{+1.25}_{-0.33}$ | $0.16^{+0.39}_{-0.16}$ | ... | ... | (6) |
| 070419B | ... | $-19.44 \pm 0.19$ | 9.84 | $25.20 \pm 0.20$ | (3),(4) | $21^{+35}_{-11}$ | $-8.52^{+0.42}_{-0.32}$ | $0.56^{+0.39}_{-0.30}$ | ... | ... | (6) |
| **070506** | $22.00 \pm 0.30$ | $-18.80 \pm 0.21$ | ... | $26.21 \pm 0.22$ | (5),(3) | ... | ... | ... | ... | $\geq -0.65$ | (21) |
| 070611 | $21.30 \pm 0.20$ | $> -17.52$ | ... | $> 27.27$ | (5),(3) | ... | ... | ... | ... | ... | ... |
| **070721B** | $21.50 \pm 0.20$ | $-18.39 \pm 0.44$ | $< 9.42$ | 27.5 | (5),(1) | ... | ... | ... | ... | $\geq -2.14$ | (21) |
| 070802 | $21.50 \pm 0.20$ | $-19.85 \pm 0.20$ | 9.69 | $25.25 \pm 0.21$ | (5),(3) | $24^{+11}_{-8}$ | $-8.31^{+0.16}_{-0.17}$ | $0.31^{+0.12}_{-0.12}$ | ... | $-0.46 \pm 0.38$ | (6),(26) |

**Table A.3.** LGRBs in the XHG sample with a Lyα-emission detection (in bold) or an upper limit determined in this work.

| GRB | Redshift | Host | OA | Lyα Host | Lyα OA | EW(Lyα) [Å] | F(Lyα) [$10^{-17}$ erg s$^{-1}$ cm$^{-2}$] | L(Lyα) [$10^{42}$ erg s$^{-1}$] |
|---|---|---|---|---|---|---|---|---|
| 050714B | 2.4383 | xsh, F1 | no | no | ... | < 17.1 | < 0.74 | < 0.36 |
| 050819 | 2.5042 | xsh | no | no | ... | < 9.3 | < 1.74 | < 0.91 |
| 050915A | 2.5275 | xsh | F1 | no | no | < 9.5 | < 0.61 | < 0.33 |
| 051001 | 2.4295 | xsh, F1 | no | no | ... | < 7.2 | < 0.84 | < 0.41 |
| 060204B | 2.3393 | xsh | no | no | ... | < 36.2 | < 8.73 | < 3.88 |
| 060604 | 2.1355 | xsh, F1 | AlFOSC | no | no | < 11.7 | < 0.87 | < 0.31 |
| 060707 | 3.4240 | xsh, F1 | F1 | no | no | ... | < 0.65 | < 0.73 |
| 060805A | 2.3633 | xsh | no | no | ... | < 14.8 | < 0.78 | < 0.35 |
| 060814 | 1.9223 | xsh | F1, F2 | no | no | < 7.1 | < 2.33 | < 0.64 |
| **060926** | 3.2090 | xsh | F1 | yes | yes | 37.0 ± 7.0 | 5.30 ± 0.40 | 5.05 ± 0.38 |
| 061202 | 2.2543 | xsh | no | no | ... | < 16.7 | < 1.76 | < 0.71 |
| 070103 | 2.6208 | xsh, F2 | no | no | ... | < 19.4 | < 1.42 | < 0.83 |
| **070110** | 2.3523 | xsh | F2 | yes | yes | 33.0 ± 8.0 | 2.80 ± 0.50 | 1.26 ± 0.23 |
| 070129 | 2.3384 | xsh | F1, F2 | no | no | < 13.9 | < 1.94 | < 0.86 |
| 070224 | 1.9922 | xsh | no | no | ... | ... | < 2.12 | < 0.64 |
| 070328 | 2.0627 | xsh | no | no | ... | < 9.5 | < 0.87 | < 0.28 |
| 070419B | 1.9586 | xsh | F1 | no | no | ... | < 2.15 | < 0.62 |
| 070521 | 2.0865 | xsh | no | no | ... | ... | < 1.54 | < 0.52 |
| 070802 | 2.4538 | xsh, F1 | F2 | no | no | < 9.7 | < 0.71 | < 0.36 |
| 071021 | 2.4515 | xsh | no | no | ... | ... | < 1.41 | < 0.70 |
| 080207 | 2.0856 | xsh | no | no | ... | < 16.1 | < 1.06 | < 0.36 |
| 080602 | 1.8204 | xsh | no | no | ... | < 10.3 | < 3.98 | < 0.96 |
| 080605 | 1.6410 | xsh | F2 | no | no | < 18.2 | < 15.29 | < 2.86 |
| 080804 | 2.2059 | xsh | UVES | no | no | < 17.5 | < 0.82 | < 0.32 |
| 081210 | 2.0631 | xsh | no | no | ... | < 22.1 | < 1.84 | < 0.60 |
| 081221 | 2.2590 | xsh | no | no | ... | < 27.3 | < 0.56 | < 0.23 |
| 090113 | 1.7494 | xsh | no | no | ... | ... | < 4.13 | < 0.90 |
| 090201 | 2.1000 | xsh | no | no | ... | ... | < 2.78 | < 0.95 |
| 090323 | 3.5832 | xsh | F2 | no | no | < 19.0 | < 4.24 | < 5.26 |
| **100424A** | 2.4656 | xsh | no | yes | ... | ... | 3.40 ± 0.50 | 1.72 ± 0.22 |
| 110818A | 3.3609 | xsh | xsh | no | no | ... | < 0.47 | < 0.50 |
| 111123A | 3.1513 | xsh | xsh | no | no | < 6.8 | < 1.15 | < 1.05 |
| 120118B | 2.9428 | xsh | no | no | ... | < 4.6 | < 2.37 | < 1.84 |
| 120119A | 1.7291 | xsh | xsh | no | no | ... | < 10.95 | < 2.33 |
| 120624B | 2.1974 | xsh | no | no | ... | ... | < 0.64 | < 0.25 |
| 120815A | 2.3587 | xsh | xsh | no | no | < 15.2 | < 1.89 | < 0.86 |
| 130131B | 2.5393 | xsh | no | no | ... | < 6.2 | < 1.49 | < 0.81 |

**Table A.4.** Properties of the LGRBs in the XHG sample with a Ly$\alpha$-emission detection (in bold) or an upper limit.

| GRB | $N_{HI}$ [log($N_{HI}$/cm$^{-2}$)] | $M_{UV}$ [mag] | $M_*$ [log($M_*$/$M_\odot$)] | $R$ [mag] | Refs | SFR [$M_\odot$ yr$^{-1}$] | log(sSFR/yr$^{-1}$) | E(B-V) | 12 + log(O/H) | Z | Refs |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 050714B | ... | $-19.61 \pm 0.19$ | ... | $25.51 \pm 0.20$ | (3) | $12.9^{+14.0}_{-5.3}$ | ... | $0.21^{+0.28}_{-0.21}$ | ... | ... | (6) |
| 050819 | ... | $-21.13 \pm 0.09$ | ... | $23.99 \pm 0.09$ | (3) | $22^{+426}_{-15}$ | ... | $0.34^{+0.84}_{-0.34}$ | ... | ... | (6) |
| 050915A | ... | $-20.47 \pm 0.15$ | 10.66 | $24.70 \pm 0.16$ | (3),(17) | 1.9 | -10.39 | $0.84^{+0.60}_{-0.61}$ | ... | ... | (6) |
| 051001 | ... | $-20.55 \pm 0.12$ | 10.08 | $24.53 \pm 0.13$ | (3),(4) | $110^{+124}_{-59}$ | $-8.04^{+0.33}_{-0.33}$ | $0.58^{+0.28}_{-0.28}$ | ... | ... | (6) |
| 060204B | ... | ... | 9.81 | ... | (17) | $78^{+85}_{-34}$ | $-7.92^{+0.32}_{-0.25}$ | $0.34^{+0.29}_{-0.23}$ | ... | ... | (6) |
| 060604 | ... | ... | ... | $25.62 \pm 0.18$ | (3) | $7.2^{+9.4}_{-3.6}$ | ... | $0.38^{+0.32}_{-0.27}$ | $8.10^{+0.28}_{-0.35}$ | ... | (6) |
| 060707 | $21.00 \pm 0.20$ | $-20.78 \pm 0.06$ | 9.99 | $25.01 \pm 0.06$ | (5),(3) | $19.9^{+48.0}_{-14.3}$ | $-8.69^{+0.53}_{-0.56}$ | ... | ... | $\geq -1.69$ | (6),(21) |
| 060805A | ... | $-19.79 \pm 0.13$ | ... | $25.26 \pm 0.14$ | (3) | $9.0^{+3.9}_{-2.5}$ | ... | $0.00^{+0.16}_{-0.00}$ | ... | ... | (6) |
| 060814 | $22.60 \pm 0.15$ | $-21.46 \pm 0.10$ | $10.03 \pm 0.10$ | $22.96 \pm 0.11$ | (3),(4) | $47.5^{+72.5}_{-15.6}$ | $-8.35^{+0.40}_{-0.17}$ | $0.17^{+0.00}_{-0.17}$ | $8.46^{+0.10}_{-0.16}$ | ... | (6),(7) |
| **060926** | ... | $-21.59 \pm 0.05$ | 10.71 | ... | (5) | $26^{+47}_{-17}$ | $-9.29^{+0.44}_{-0.45}$ | $0.58^{+0.34}_{-0.27}$ | ... | $\geq -1.32$ | (6),(21) |
| 061202 | ... | ... | 9.64 | ... | (17) | $43^{+60}_{-22}$ | $-8.01^{+0.38}_{-0.31}$ | $0.00^{+0.50}_{-0.00}$ | ... | ... | (6) |
| 070103 | ... | $-21.03 \pm 0.14$ | ... | $24.21 \pm 0.14$ | (3) | $43^{+62}_{-17}$ | ... | $0.00^{+0.38}_{-0.00}$ | ... | ... | (6) |
| **070110** | $21.70 \pm 0.10$ | $-19.81 \pm 0.11$ | <9.16 | $25.19 \pm 0.11$ | (5) | $8.9^{+10.9}_{-2.8}$ | > -8.37 | $0.17^{+0.35}_{-0.17}$ | ... | $\geq -1.32$ | (6),(21) |
| 070129 | ... | $-20.75 \pm 0.11$ | 10.15 | $24.23 \pm 0.12$ | (3),(4) | $20^{+28}_{-7}$ | $-8.85^{+0.38}_{-0.19}$ | ... | ... | ... | (6) |
| 070224 | ... | $-18.71 \pm 0.29$ | ... | $26.02 \pm 0.31$ | (3) | $3.2^{+6.5}_{-2.3}$ | ... | ... | ... | ... | (6) |
| 070328 | ... | $-20.17 \pm 0.12$ | 9.65 | $24.55 \pm 0.13$ | (3),(17) | $8.4^{+130.7}_{-4.2}$ | $-8.73^{+1.23}_{-0.30}$ | $0.16^{+0.39}_{-0.16}$ | ... | ... | (6) |
| 070419B | ... | $-19.44 \pm 0.19$ | 9.84 | $25.20 \pm 0.20$ | (3),(4) | $21^{+35}_{-11}$ | $-8.52^{+0.42}_{-0.39}$ | $0.56^{+0.39}_{-0.30}$ | ... | ... | (6) |
| 070521 | ... | ... | 10.5 | ... | (17) | $26^{+34}_{-17}$ | $-9.09^{+0.37}_{-0.47}$ | $0.71^{+0.05}_{-0.04}$ | ... | ... | (6) |
| 070802 | $21.50 \pm 0.20$ | $-19.85 \pm 0.20$ | 9.69 | $25.25 \pm 0.21$ | (5),(3) | $24^{+11}_{-8}$ | $-8.31^{+0.16}_{-0.17}$ | $0.31^{+0.12}_{-0.12}$ | ... | $-0.46 \pm 0.63$ | (6),(20) |
| 071021 | ... | ... | 10.10 | ... | (17) | $32^{+20}_{-12}$ | $-8.59^{+0.21}_{-0.20}$ | $0.19^{+0.16}_{-0.17}$ | ... | ... | (6) |
| 080207 | ... | ... | 11.09 | $25.84 \pm 0.18$ | (17) | $77^{+86}_{-38}$ | $-9.20^{+0.32}_{-0.29}$ | $0.66^{+0.28}_{-0.25}$ | $8.74^{+0.15}_{-0.15}$ | ... | (6) |
| 080602 | ... | ... | $9.40 \pm 0.10$ | ... | (7) | > 48 | > -7.72 | $0.58^{+0.29}_{-0.26}$ | $8.69^{+0.12}_{-0.21}$ | ... | (6),(7) |
| 080605 | ... | ... | $9.60 \pm 0.10$ | ... | (7) | $42.5^{+30.5}_{-18.2}$ | $-7.97^{+0.23}_{-0.24}$ | $0.26^{+0.10}_{-0.11}$ | $8.47^{+0.04}_{-0.04}$ | ... | (6),(7) |
| 080804 | $21.30 \pm 0.15$ | ... | 9.28 | ... | (5) | $15.2^{+41.2}_{-8.7}$ | $-8.10^{+0.57}_{-0.37}$ | $0.38^{+0.35}_{-0.35}$ | ... | $-0.75 \pm 0.16$ | (6),(21) |
| 081210 | ... | ... | 10.07 | ... | (17) | $15.3^{+111.7}_{-7.0}$ | $-8.89^{+0.92}_{-0.27}$ | $0.13^{+0.61}_{-0.13}$ | ... | ... | (6) |
| 081221 | ... | ... | 10.08 | ... | (17) | $35^{+106}_{-22}$ | $-8.54^{+0.61}_{-0.43}$ | $0.31^{+0.55}_{-0.31}$ | ... | ... | (6) |
| 090113 | ... | ... | 9.90 | ... | (17) | $17.9^{+10.1}_{-4.8}$ | $-8.65^{+0.19}_{-0.14}$ | $0.01^{+0.22}_{-0.01}$ | ... | ... | (6) |
| 090201 | ... | ... | 10.90 | ... | (4) | $48^{+30}_{-14}$ | $-9.22^{+0.21}_{-0.21}$ | $0.11^{+0.19}_{-0.11}$ | ... | ... | (6) |
| 090323 | $20.75 \pm 0.10$ | $-21.60 \pm 0.18$ | 10.30 | ... | (5),(17) | $24^{+53}_{-17}$ | $-8.92^{+0.50}_{-0.53}$ | ... | ... | ... | (6) |
| **100424A** | ... | ... | ... | > 24 | (8) | $21^{+20}_{-8}$ | ... | $0.13^{+0.18}_{-0.13}$ | $7.93^{+0.25}_{-0.18}$ | ... | (6) |
| 110818A | $21.90 \pm 0.40$ | $-21.68 \pm 0.05$ | 10.50 | ... | (5),(17) | $44^{+62}_{-26}$ | $-8.86^{+0.38}_{-0.39}$ | ... | $8.25^{+0.17}_{-0.25}$ | ... | (6) |
| 111123A | ... | $-22.05 \pm 0.03$ | 11.17 | ... | (3),(17) | $77^{+163}_{-52}$ | $-9.28^{+0.49}_{-0.49}$ | ... | $8.01^{+0.28}_{-0.28}$ | ... | (6) |
| 120118B | ... | ... | ... | ... | ... | $28^{+21}_{-11}$ | ... | $0.00^{+0.16}_{-0.00}$ | $7.89^{+0.23}_{-0.17}$ | ... | (6) |
| 120119A | $22.60 \pm 0.20$ | ... | 9.93 | ... | (5),(17) | $43^{+24}_{-14}$ | $-8.30^{+0.19}_{-0.17}$ | $0.35^{+0.16}_{-0.14}$ | $8.60^{+0.14}_{-0.14}$ | $-0.96 \pm 0.28$ | (6),(20) |
| 120624B | ... | ... | 10.84 | ... | (17) | $30^{+73}_{-13}$ | $-9.36^{+0.53}_{-0.24}$ | $0.21^{+0.50}_{-0.21}$ | $8.43^{+0.20}_{-0.27}$ | ... | (6) |
| 120815A | $22.05 \pm 0.10$ | ... | ... | ... | (5) | $2.3^{+2.7}_{-1.0}$ | ... | $0.06^{+0.34}_{-0.06}$ | ... | $-1.15 \pm 0.12$ | (6),(20) |
| 130131B | ... | ... | ... | ... | ... | $8.0^{+13.4}_{-5.0}$ | ... | ... | ... | ... | (6) |

**Table A.5.** LGRBs from the literature (other than TOUGH and XHG samples) and this work with a Lyα-emission detection. GRBs 070223 and 080810 are not reported in this table although they are reported in Appendix B.1 because no information about Lyα flux is available for these two GRB host galaxies.

| GRB | Redshift | Host | OA | Lyα Host | Lyα OA | EW(Lyα) [Å] | F(Lyα) [$10^{-17}$ erg s$^{-1}$ cm$^{-2}$] | L(Lyα) [$10^{42}$ erg s$^{-1}$] | Refs |
|---|---|---|---|---|---|---|---|---|---|
| 971214 | 3.4200 | Keck | no | yes | ... | 13.0 | 0.62 ± 0.07 | 0.66 ± 0.07 | (1) |
| 000926 | 2.0400 | AlFOSC | NOT | yes | no | 71.0 ± 18.0 | 14.90 ± 1.10 | 4.51 ± 0.33 | (1) |
| 011211 | 2.1434 | xsh | F1 | yes | no | 14.0 ± 3.0 | 1.60 ± 0.20 | 0.58 ± 0.07 | (2) |
| 021004 | 2.3298 | xsh | UVES | yes | yes | 105.0 ± 10.0 | 16.90 ± 0.30 | 7.45 ± 0.18 | (2) |
| 030323 | 3.3720 | no | F2 | ... | yes | 108.0 ± 38.0 | 1.20 ± 0.10 | 1.23 ± 0.10 | (1) |
| 030429 | 2.6600 | no | F1 | ... | yes | ... | 0.31 | 0.19 ± 0.00 | (1) |
| 061222A | 2.0880 | LRIS | no | yes | ... | 31.0 | 16.80 | 5.39 ± 0.00 | (1) |
| 071031 | 2.6918 | no | F2 | ... | yes | ... | 2.36 ± 0.27 | 1.41 ± 0.16 | (1) |
| 081121 | 2.5134 | xsh | LDSS3 | yes | no | ... | 4.90 ± 0.50 | 2.60 ± 0.26 | (2) |
| 081222 | 2.7700 | xsh | GMOS-S | yes | no | ... | 1.60 ± 0.40 | 1.07 ± 0.27 | (2) |
| 090205 | 4.6500 | no | F1 | ... | yes | ... | 2.36 ± 0.49 | 5.16 ± 1.08 | (1) |
| 121201A | 3.3830 | no | xsh | ... | yes | ... | 2.14 ± 0.36 | 2.31 ± 0.39 | (2) |
| 150915A | 1.9680 | no | xsh | ... | yes | ... | 5.00 ± 1.10 | 1.46 ± 0.32 | (2) |
| 151021A | 2.3300 | no | xsh | ... | yes | ... | 10.50 ± 0.90 | 4.63 ± 0.40 | (2) |
| 170202A | 3.6450 | no | xsh | ... | yes | ... | 2.70 ± 0.60 | 3.49 ± 0.78 | (2) |
| 191004B | 3.5055 | xsh | xsh | yes | yes | 7.4 ± 2.6 | 1.00 ± 0.10 | 1.18 ± 0.18 | (9) |

**Table A.6.** Properties of the LGRBs from the literature (other than TOUGH and XHG samples) and this work with a Lyα-emission detection.

| GRB | $N_{\rm HI}$ [log($N_{\rm HI}$/cm$^{-2}$)] | $M_{\rm UV}$ [mag] | $M_*$ [log($M_*/M_\odot$)] | $R$ [mag] | Refs | SFR [$M_\odot\,{\rm yr}^{-1}$] | log(sSFR/yr$^{-1}$) | E(B-V) | 12 + log(O/H) | $Z$ | Refs |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 971214 | ... | ... | ... | ... | | ... | ... | $0.44^{+0.06}_{-0.03}$ | ... | ... | (15) |
| 000926 | $21.30 \pm 0.25$ | $-20.40^{+0.07}_{-0.07}$ | 9.64 | ... | (5) | ... | ... | $0.19^{+0.13}_{-0.09}$ | ... | $-0.13 \pm 0.21$ | (15),(22) |
| 011211 | $20.40 \pm 0.20$ | $-19.97^{+0.10}_{-0.10}$ | 8.0 | $24.95 \pm 0.11$ | (5),(10) | $< 2.0$ | $< -7.70$ | $0.03^{+0.03}_{-0.03}$ | ... | $-0.90 \pm 0.50$ | (2),(10),(23) |
| 021004 | $19.50 \pm 0.50$ | $-20.56^{+0.10}_{-0.10}$ | 9.45 | $24.39 \pm 0.04$ | (5),(11) | $6.7 \pm 0.8$ | $> -8.62$ | $0.02^{+0.03}_{-0.03}$ | ... | $-1.30 \pm 0.50$ | (2),(16),(27) |
| 030323 | $21.90 \pm 0.07$ | $-18.47^{+0.10}_{-0.10}$ | $< 9.85$ | $28.0 \pm 0.3$ | (5),(12) | ... | ... | ... | ... | $-1.26 \pm 0.20$ | (12) |
| 030429 | $21.60 \pm 0.20$ | ... | ... | $> 26.3$ | (5),(13) | ... | ... | ... | ... | $\geq -1.13$ | (21) |
| 061222A | ... | ... | 9.55 | ... | (4) | ... | ... | $0.00^{+0.00}_{-0.00}$ | ... | ... | (15) |
| 071031 | $22.15 \pm 0.05$ | ... | ... | ... | (5) | ... | ... | ... | ... | $-1.85 \pm 0.12$ | (21) |
| 081121 | ... | ... | 9.24 | ... | (4) | $> 7.0$ | $> -8.35$ | ... | ... | ... | (2) |
| 081222 | $20.80 \pm 0.20$ | $-22.43^{+0.0}_{+0.0}$ | 9.61 | ... | (5) | ... | ... | ... | ... | ... | |
| 090205 | $20.73 \pm 0.05$ | $-21.26^{+0.13}_{+0.13}$ | $< 10.7$ | $26.4 \pm 0.3$ | (5),(14) | ... | ... | ... | ... | $\geq -0.57$ | (21) |
| 121201A | $22.00 \pm 0.20$ | $-20.84^{+0.21}_{+0.21}$ | ... | ... | (5) | ... | ... | ... | ... | ... | |
| 150915A | $21.20 \pm 0.30$ | ... | ... | ... | (5) | $13.0 \pm 0.6$ | ... | ... | ... | ... | (2) |
| 151021A | $22.20 \pm 0.20$ | ... | ... | ... | (5) | ... | ... | ... | ... | $-0.98 \pm 0.07$ | (19) |
| 170202A | $21.55 \pm 0.10$ | ... | ... | ... | (5) | ... | ... | ... | ... | $-1.28 \pm 0.09$ | (19) |
| 191004B | $17.20 \pm 0.15$ | $-19.10^{+0.0}_{+0.0}$ | ... | $> 26.6$ | (5),(9) | $2.85 \pm 1.85$ | ... | $0.01^{+0.01}_{-0.01}$ | ... | ... | (9) |

## Appendix B: Table summarizing the LAE-LGRB census.

**Table B.1.** 29 LGRBs with a detected Lyα emission line in the host galaxy or optical afterglow spectra.

| GRB | Redshift | Host | OA | Lyα Host | Lyα OA | Flux(Lyα) | Sample | Refs |
|---|---|---|---|---|---|---|---|---|
| 971214 | 3.4200 | Keck | no | yes | - | 0.6 ± 0.1 | - | (1), (2), (3), (8) |
| 000926 | 2.0400 | AlFOSC | NOT | yes | no | 14.9 ± 1.1 | - | (4), (8) |
| 011211 | 2.1434 | xsh | F1 | yes | no | 1.6 ± 0.2 | - | (2) |
| 021004 | 2.3298 | xsh | UVES | yes | yes | 16.9 ± 0.3 | - | (2) |
| 030323 | 3.3720 | no | F2 | - | yes | 1.2 ± 0.1 | - | (2), (6),(8) |
| 030429 | 2.6600 | no | F1 | - | yes | 0.3 (2σ) | - | (7) |
| 050315 | 1.9500 | F1 | no | yes | - | 2.3 ± 0.7 | TOUGH | (8) |
| 060605 | 3.7730 | F1 | PMAS | yes | no | 1.7 ± 0.3 | TOUGH | (8) |
| 060707 | 3.4240 | F1, xsh | F1 | yes | no | 1.7 ± 0.3 | TOUGH, XHG[a] | (8), (10) |
| 060714 | 2.7108 | F1 | F1 | no | yes | 1.7 | TOUGH[b] | (8), (10), (11) |
| 060908 | 1.8836 | F1 | F1, F2 | yes | no | 7.8 ± 1.0 | TOUGH | (8) |
| 060926 | 3.2090 | xsh | F1 | yes | yes | 5.3 ± 0.4 | XHG | (8), (10) |
| 061222A | 2.0880 | LRIS | no | yes | - | 16.8 | - | (8), (12) |
| 070110 | 2.3523 | F1, xsh | F1 | yes | yes | 4.0 ± 0.4 | TOUGH, XHG | (8), (10) |
| 070223 | 1.6295 | LRIS | no | yes | - | - | - | (9) |
| 070506 | 2.3090 | F1 | F1 | yes | no | 1.4 ± 0.4 | TOUGH | (8), (10) |
| 070721B | 3.6298 | F1 | F2 | yes | yes | 1.1 ± 0.2 | TOUGH | (8), (10) |
| 071031 | 2.6918 | no | F2 | - | yes | 2.4 ± 0.3 | - | (8), (10) |
| 080810 | 3.3604 | LRIS | AlFOSC | yes | yes | - | - | (10), (23) |
| 081121 | 2.5134 | xsh | LDSS3 | yes | no | 4.9 ± 0.5 | - | (20), (21) |
| 081222 | 2.7700 | xsh | GMOS-S | yes | no | 1.6 ± 0.4 | - | (17), (20) |
| 090205 | 4.6500 | no | F1 | - | yes | 2.4 ± 0.5 | - | (8), (13) |
| 100316A | 3.1600 | no | OSIRIS | no | yes | - | - | (17) |
| 100424A | 2.4656 | xsh | no | yes | - | 3.4 ± 0.5 | XHG | (15), (18) |
| 121201A | 3.3830 | no | xsh | - | yes | 2.1 ± 0.4 | XAFT | (16), (19) |
| 150915A | 1.9680 | no | xsh | - | yes | 5.0 ± 1.1 | XAFT | (16) |
| 151021A | 2.3300 | no | xsh | - | yes | 10.5 ± 0.9 | XAFT | (16) |
| 170202A | 3.6450 | no | xsh | - | yes | 2.7 ± 0.6 | XAFT | (16) |
| 191004B | 3.5055 | xsh | xsh | yes | yes | 1.0 ± 0.1 | - | (20), (22) |

**Notes.** *GRB* and *Redshift* are for the name of the LGRB and its redshift. When observations of the host galaxy or the optical afterglow are available, we provide the name of the spectrographs used in the columns *Hosts* and *OA*, respectively. "no" indicates that no observation is available. "F1"/"F2" is for VLT/FORS1/2 and "xsh" for VLT/X-shooter spectrograph. *Lyα Host* and *Lyα OA* inform about the detection of the Lyα line in the host or afterglow spectra, respectively. No information is provided when no spectra are available to verify the presence of the line. *Flux(Lyα)* corresponds to the Lyα line flux, in units of $10^{-17}$ erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$, retrieved from the literature (see references in column *Refs*) or derived in this work. *Samples* indicates to which sample the LGRB belongs. *Refs* is for the references where the Lyα detection is reported. [a] GRB 060707 is part of the XHG sample but the Lyα is not detected in the X-shooter spectrum. [b] For GRB 060714 the Lyα line is only detected at 2.5σ in the host galaxy observation of Milvang-Jensen et al. (2012), but is convincingly detected in the afterglow spectrum presented by Jakobsson et al. (2006), therefore we consider it as a detection and report the flux measured in the afterglow spectrum corrected for Galactic extinction.

**References:** (1): Kulkarni et al. (1998); (2): Fynbo et al. (2003); (3): Ahn (2000); (4): Fynbo et al. (2002); (5): Jakobsson et al. (2005); (6): Vreeswijk et al. (2004); (7): Jakobsson et al. (2004); (8): Milvang-Jensen et al. (2012); (9): Perley et al. (2016b); (10): Fynbo et al. (2009); (11): Jakobsson et al. (2006); (12): Perley et al. (2009); (13): D'Avanzo et al. (2010); (14): Hartoog et al. (2015); (15): Krühler et al. (2015); (16): Selsing et al. (2019); (17): Tanvir et al. (2019); (18): Malesani et al. (2013); (19): Cucchiara et al. (2015); (20): this work, (21): Berger & Rauch (2008), (22): D'Elia et al. (2019), (23): Wiseman et al. (2017)

# Appendix C: Results of the shell model with constraints from the GRB host galaxy.

**Table C.1.** Shell-model predictions with constraints on $z^{HG}$, $FWHM_i^{HG}(Ly\alpha)$, and $EW_i^{HG}(Ly\alpha)$

| GRB host | $\Delta z$ | $\log(N_{HI}cm^{-2})$ | $V_{exp}$ [km s$^{-1}$] | $\log(T/K)$ | $\tau_d$ | $FWHM_i(Ly\alpha)$ [km s$^{-1}$] | $EW_i(Ly\alpha)$ [Å] |
|---|---|---|---|---|---|---|---|
| GRB 011211 | $30^{+45}_{-45}$ | $19.9^{+0.2}_{-0.2}$ | $90^{+21}_{-26}$ | $4.17^{+0.9}_{-0.9}$ | $0.47^{+0.39}_{-0.27}$ | $50^{+10}_{-10}$ | $33^{+20}_{-16}$ |
| GRB 021004 | $5^{+45}_{-45}$ | $19.6^{+0.1}_{-0.1}$ | $120^{+5}_{-6}$ | $4.96^{+0.17}_{-0.14}$ | $0.399^{+0.114}_{-0.084}$ | $223^{+13}_{-12}$ | $165^{+14}_{-13}$ |
| GRB 060926 | $7^{+45}_{-45}$ | $19.8^{+0.1}_{-0.1}$ | $158^{+5}_{-5}$ | $4.97^{+0.17}_{-0.17}$ | $1.27^{+0.28}_{-0.23}$ | $148^{+7}_{-7}$ | $107^{+3}_{-2}$ |
| GRB 070110 | $3^{+45}_{-45}$ | $19.8^{+0.2}_{-0.1}$ | $92^{+13}_{-11}$ | $3.69^{+0.90}_{-0.55}$ | $0.083^{+0.138}_{-0.062}$ | $20^{+26}_{-12}$ | $20^{+10}_{-6}$ |

**Notes.** Same as Table 5.
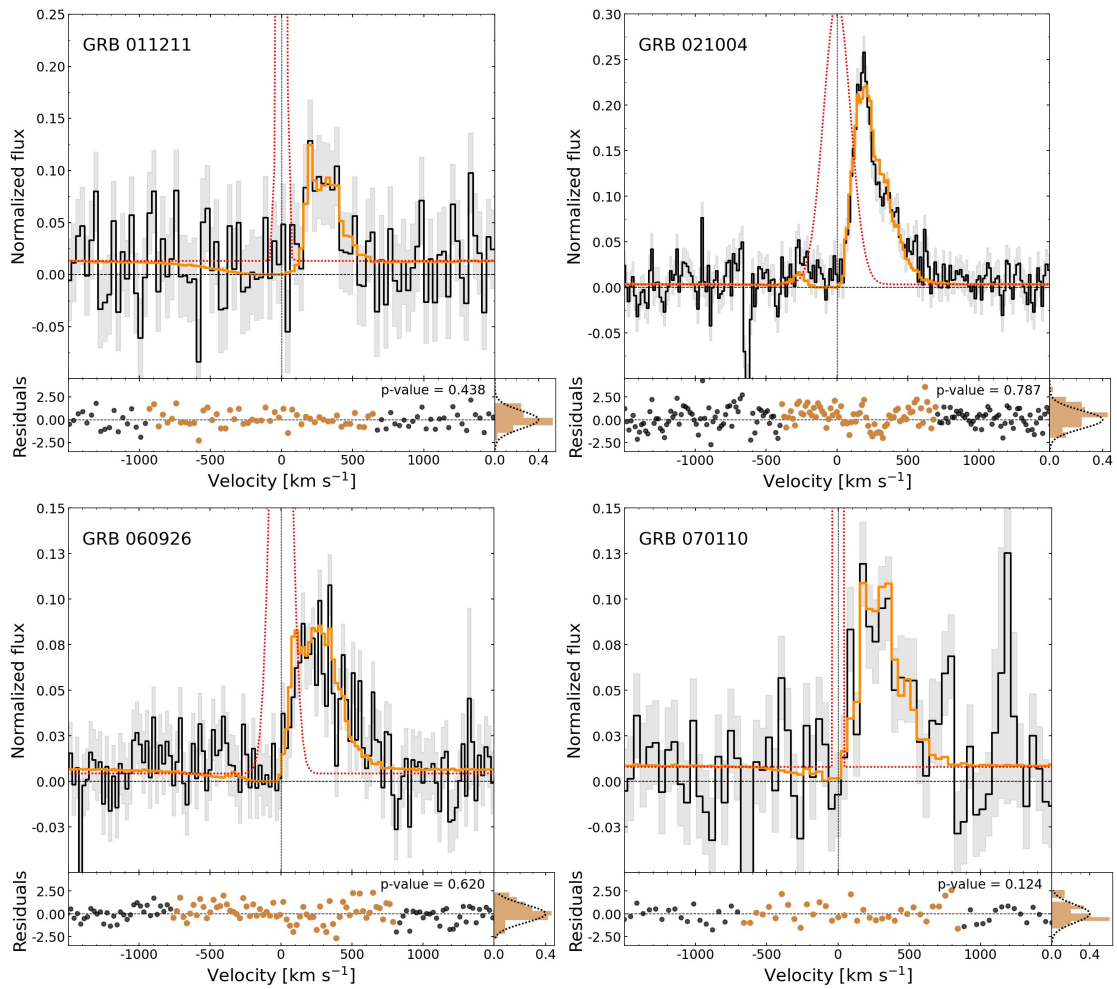


**Fig. C.1.** Same as Fig. 9 but with the shell model parameters $z^{HG}$, $FWHM_i^{HG}(Ly\alpha)$, and $EW_i^{HG}(Ly\alpha)$ constrained by the values determined from the observations of the GRB host galaxy.

## Appendix D: Results of the shell model with constraints from the GRB afterglow.

**Table D.1.** Shell-model predictions with constraints on $z^{HG}$, $N_{HI}^{OA}$, $V_{LIS}^{OA}$, and $\tau_d^{OA}$

| GRB host | $\Delta z$ | $\log(N_{HI}cm^{-2})$ | $V_{exp}$ [km s$^{-1}$] | $\log(T/K)$ | $\tau_d$ | FWHM$_i$(Ly$\alpha$) [km s$^{-1}$] | EW$_i$(Ly$\alpha$) [Å] |
|---|---|---|---|---|---|---|---|
| GRB 011211 | $17^{+45}_{-45}$ | $20.2^{+0.1}_{-0.1}$ | $56^{+47}_{-28}$ | $3.52^{+0.87}_{-0.52}$ | $0.65^{+0.51}_{-0.40}$ | $280^{+140}_{-165}$ | $66^{+60}_{-34}$ |
| GRB 021004 | $6^{+45}_{-45}$ | $19.6^{+0.1}_{-0.1}$ | $117^{+5}_{-4}$ | $4.97^{+0.17}_{-0.15}$ | $0.362^{+0.103}_{-0.075}$ | $269^{+15}_{-20}$ | $162^{+16}_{-15}$ |
| GRB 060926 | $60^{+45}_{-45}$ | $21.5^{+0.1}_{-0.1}$ | $3^{+1}_{-1}$ | $3.04^{+0.12}_{-0.15}$ | $0.003^{+0.004}_{-0.002}$ | $114^{+67}_{-44}$ | $22^{+8}_{-5}$ |
| GRB 070110 | $43^{+45}_{-45}$ | $21.3^{+0.2}_{-0.1}$ | $3^{+1}_{-1}$ | $3.07^{+0.12}_{-0.17}$ | $0.018^{+0.010}_{-0.007}$ | $10^{+9}_{-4}$ | $62^{+17}_{-17}$ |

**Notes.** Same as Table 5.
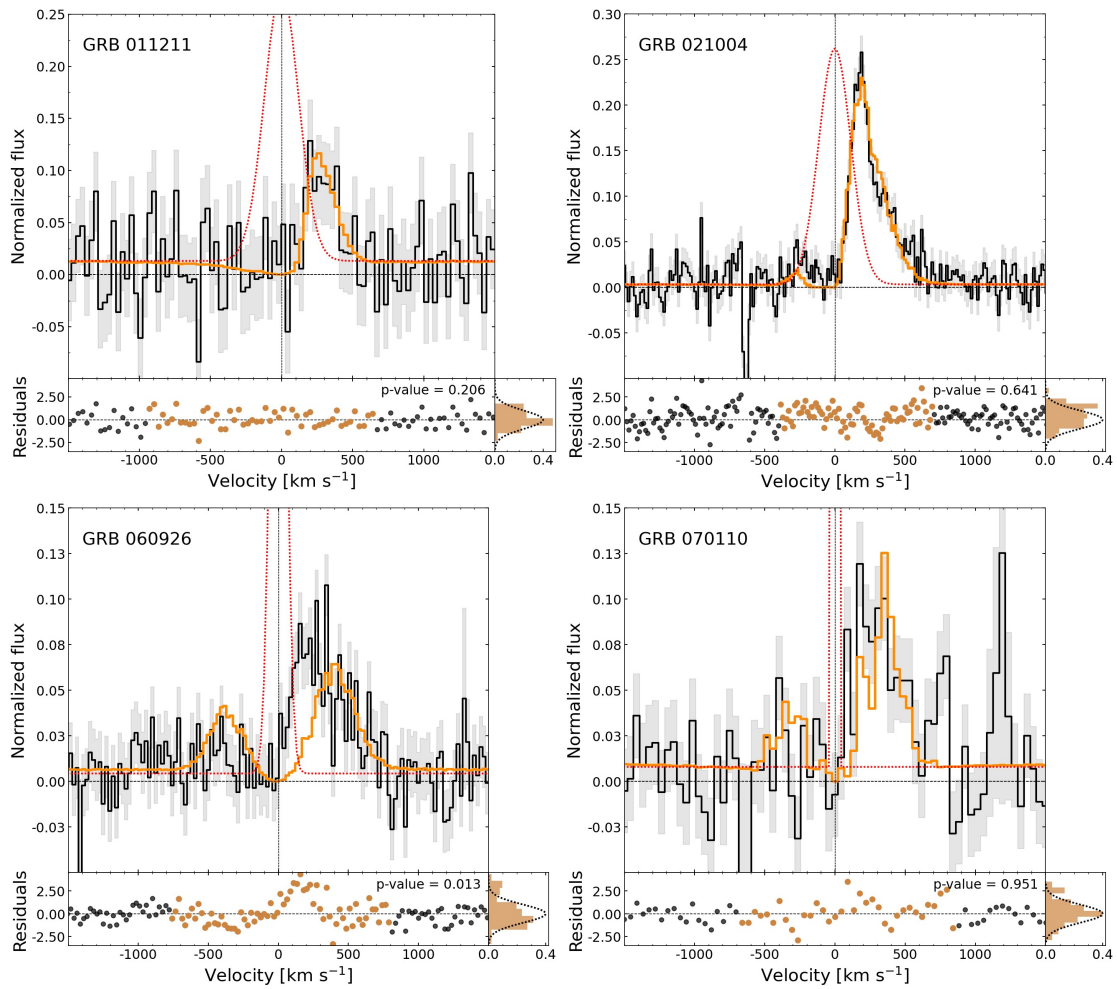


**Fig. D.1.** Same as Fig. 9, but with the shell model parameters $z^{HG}$, $N_{HI}^{OA}$, $V_{LIS}^{OA}$, and $\tau_d^{OA}$ constrained by the values determined from the observations of the GRB afterglow (and systemic redshift from the observations of the host galaxy).