Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 2342–2353 (2020)
Advance Access publication 2020 May 15

doi:10.1093/mnras/staa1321

Using realistic host galaxy metallicities to improve the GRB X-ray equivalent total hydrogen column density and constrain the intergalactic medium density

Tony Dalton★ and Simon L. Morris

Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK

Accepted 2020 May 4. Received 2020 May 1; in original form 2020 March 3

ABSTRACT

It is known that the GRB equivalent hydrogen column density ($N_{\rm HX}$) changes with redshift and that, typically, $N_{\rm HX}$ is greater than the GRB host neutral hydrogen column density. We have compiled a large sample of data for GRB $N_{\rm HX}$ and metallicity [X/H]. The main aims of this paper are to generate improved $N_{\rm HX}$ for our sample by using actual metallicities, dust corrected where available for detections, and for the remaining GRB, a more realistic average intrinsic metallicity using a standard adjustment from solar. Then, by approximating the GRB host intrinsic hydrogen column density using the measured neutral column ($N_{\rm HI,\,IC}$) adjusted for the ionization fraction, we isolate a more accurate estimate for the intergalactic medium (IGM) contribution. The GRB sample mean metallicity is $= -1.17 \pm 0.09$ rms (or 0.07 ± 0.05 $\rm Z/Z_{sol}$) from a sample of 36 GRB with a redshift $1.76 \leq z \leq 5.91$, substantially lower than the assumption of solar metallicity used as standard for many fitted $N_{\rm HX}$. Lower GRB host mean metallicity results in increased estimated $N_{\rm HX}$ with the correction scaling with redshift as $\Delta\log(N_{\rm HX}\ {\rm cm}^{-2}) = (0.59 \pm 0.04)\log(1+z) + 0.18 \pm 0.02$. Of the 128 GRB with data for both $N_{\rm HX}$ and $N_{\rm HI,\,IC}$ in our sample, only six have $N_{\rm HI,\,IC} > N_{\rm HX}$ when revised for realistic metallicity, compared to 32 when solar metallicity is assumed. The lower envelope of the revised $N_{\rm HX} - N_{\rm HI,\,IC}$, plotted against redshift can be fit by $\log(N_{\rm HX} - N_{\rm HI,\,IC}\ {\rm cm}^{-2}) = 20.3 + 2.4\log(1+z)$. This is taken to be an estimate for the maximum IGM hydrogen column density as a function of redshift. Using this approach, we estimate an upper limit to the hydrogen density at redshift zero (n_0) to be consistent with $n_0 = 0.17 \times 10^{-7}{\rm cm}^{-3}$.

Key words: gamma-ray burst: general – galaxies: abundances – galaxies: high-redshift – intergalactic medium – cosmological parameters – X-rays: general.

1 INTRODUCTION

Gamma-ray bursts (GRBs) are among the most powerful explosions known in the Universe (see Schady 2017, hereafter S17) for a recent general review of GRB). Given the huge range of redshifts and distances for GRBs, and their high luminosities combined with the broad energy range of observed emissions, GRBs provide a valuable probe of all baryonic matter along the line of sight. X-ray absorption yields information on the total absorbing column density of the matter between the observer and the source because any element that is not fully ionized contributes to the absorption of X-rays (scattering by electrons becomes important at high energy above 10 keV (Wilms, Allen & McCray 2000, hereafter W00). Although the X-ray absorption cross-section is often dominated by metals, with hydrogen and helium contribution being minimal but not nil (Fig. 1, W00), it is typically reported as an equivalent hydrogen column density (in this paper $N_{\rm HX}$). $N_{\rm HX}$ consists of contributions from the local GRB environment, the intergalactic medium (IGM), and our own Galactic medium. However, X-ray absorption cannot reveal the redshift of the matter in the column due to a lack of spectral resolution and signal to noise. It is important to note that the common practice is to make the simplifying assumption that all X-ray absorption in excess of Galactic is at the redshift of the host, neglecting any IGM contribution (e.g. Watson et al. 2007; Starling et al. 2013). The GRB $N_{\rm HX}$ versus redshift relation has been investigated for many years. Early reports were based on small samples (e.g. Campana et al. 2010; Behar et al. 2011; Watson et al. 2013). A claimed strong correlation with redshift has recently been updated and confirmed with a much larger GRB sample by Rahin & Behar (2019). It has also been reported in many papers that the neutral intrinsic hydrogen column ($N_{\rm H\,I}$) in GRB has no significant correlation with redshift (e.g. Watson et al. 2007). Further, it was also noted in these papers, that $N_{\rm HX}$ exceeds $N_{\rm H\,I}$ in GRB, often by over an order of magnitude.

The cause of an $N_{\rm HX}$ excess over $N_{\rm H\,I}$, and the $N_{\rm HX}$ correlation with redshift is the source of much debate. One school of thought

★ E-mail: tonydalton@live.ie

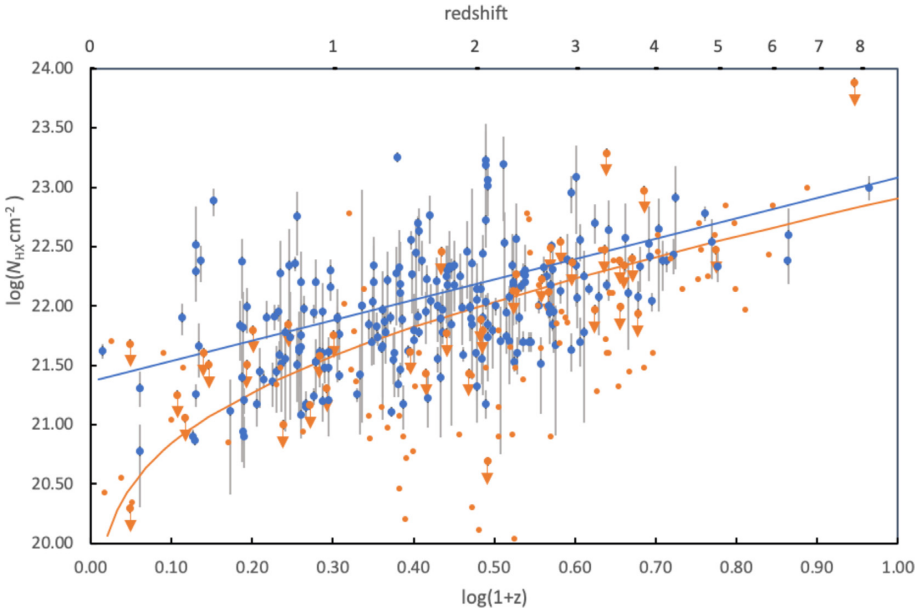

Figure 1. Distribution of intrinsic X-ray column densities (N_{HX} cm^{-2}) with redshift for the full 352 GRB *Swift* observed sample. The blue dots represent the GRB detections with error bars. The orange dots are best fits where the 90 per cent confidence interval includes zero. The orange dots with arrows are the upper limits where *Swift* repository has a best fit of zero. The blue line is the χ^2 best fit with for the GRB data with error bars. The orange line represents the integrated hydrogen density N_{HIGM} cm^{-2}) from a simple diffuse IGM model (see equation 2). The correlation statistics for the full 352 GRB sample are Pearson $r = 0.29$ and Spearman $\rho = 0.55$ (for the detection only sample (226 GRB) $r = 0.51$ and $\rho = 0.49$).

argues that the GRB host accounts for all the excess and evolution, e.g. a dense environment near the burst location (Campana et al. 2012, hereafter C12), ultra-ionized gas in the environment of the GRB (Schady et al. 2011, hereafter S11), dust extinction bias (Watson & Jakobsson 2012), dense Helium (He II) regions close to the GRB (Watson et al. 2013), and/or a host galaxy mass N_{HX} relation (Buchner, Schulze & Bauer 2017). Models for GRB N_{HX} being produced exclusively by gas intrinsic to the GRB host galaxy have required extreme conditions to be present within the absorbing material. The other school of thought argues that some of the excess N_{HX} and redshift correlation is due to the full integrated line of sight (LOS) including the diffuse IGM and intervening objects. Behar et al. (2011) modelled the effects of a cold, neutral, highly metal-enriched IGM model and showed that, at high redshift, this could produce the dominant excess X-ray absorption component. Starling et al. (2013, hereafter S13) modelled a more realistic warm IGM (WHIM) with temperature of between 10^5 and $10^{6.5}$ K and metallicity of $\sim$0.2 Z/Z$_{sol}$. Campana et al. (2015) used cosmological simulations to model the WHIM. Their results suggested that most of the excess N_{HX} absorption arises from discrete overdensities along the LOS to GRB, supporting the possibility of a significant contribution of the IGM to the N_{HX} – redshift relation.

All of the theories thus far have relied upon key assumptions (listed below) which, if unrealistic, will substantially affect the results.

1.1 Metallicity

It is known that GRB galaxy hosts have, on average, sub-solar metallicity, and that assuming solar metallicity in the X-ray fits introduces a systematic error, and generates an N_{HX} that is effectively, a minimum (S13; Krühler et al. 2015; Tanga et al. 2016). Further, models for the WHIM integrated gas density along the LOS heavily rely on the assumed gas metallicity of the WHIM. It is standard practice currently to assume solar metallicity when fitting models to GRB X-ray spectra. The main reasons for this historically were the small numbers of reliable metallicity measurements and poor constraints on any redshift metallicity evolution (S13). Even an assumption of solar metallicity, however, can lead to inconsistencies, as research improves the knowledge of solar abundances. The solar abundances of the key metals reported in the literature have undergone considerable changes in recent decades (for a useful review see Asplund et al. 2009). The X-ray fitting software XSPEC[1] (Arnaud 1996) is the most commonly used for GRB. Within XSPEC, the default solar abundances are those of Anders & Grevesse (1989). However, there are six other options for solar abundance in XSPEC. The more commonly used is that of W00. S13 noted that their results using W00 abundances were consistently higher than the N_{HX} reported in the UK Swift Science Data Centre[2] repository (hereafter *Swift*), which at the time were based on Anders & Grevesse (1989). These have since been updated using W00.

Some comments have been made in literature as to how N_{HX} scales with metallicity e.g. S13 stated that N_{HX} scales approximately with metallicity, and Krongold & Prochaska (2013) stated that the X-ray estimated oxygen column density has a linear dependence on metallicity and density. Metallicity is the main focus of this paper and Section 3 examines the impact of the assumed metallicity on the derived N_{HX} in some detail.

1.2 Location of excess absorption

It is standard practice, when fitting models to GRB spectra, to assume all absorption in excess of the Galactic contribution is at the redshift of the GRB. X-ray optical depth is a function of frequency

[1] https://heasarc.gsfc.nasa.gov/xanadu/xspec/
[2] www.swift.ac.uk/xrt_spectra

or energy, due to the frequency dependence of the cross-section (Morrison & Mccammon 1983). The scaling relation between the observed amount of X-ray absorption for GRB and redshift was found by Campana et al. (2014) to be approximated by

$$N_H\,(z=0) = N_H\,(z)\,/(1+z)^a, a = 2.4 \quad (1)$$

The error in X-ray column density produced by assuming the total absorption is at the GRB redshift arises from the difference in redshift between the GRB and any intervening contributor. Hence, the potential error in $N_{\rm HX}$ increases with redshift of the GRB, dependant on the amount of IGM absorption, its location, and any error in the scaling law assumed. The IGM hydrogen column estimation is highly uncertain as the metal pollution is very poorly determined (e.g. Fumagalli 2014; Maiolino & Mannucci 2019).

1.3 Neutral fraction

The value found for the column density is almost always determined assuming a 100 per cent neutral absorbing gas (e.g. Behar et al. 2011; S11; S13). An ionized absorber would have a lower cross-section at X-ray energies, and a larger column density would be required to produce the same opacity. Therefore, the neutral assumption could cause $N_{\rm HX}$ to be underestimated (S11).

1.4 Galactic absorption

Column densities for GRB reported have normally had the Galaxy contribution ($N_{\rm HGal}$) removed. The most common references for the $N_{\rm HGal}$ are the Leiden Argentine Bonn (LAB) H I survey (Kalberla et al. 2005) and Willingale et al. (2013).

In conclusion, $N_{\rm HX}$, based on an assumed solar metallicity and 100 per cent neutral absorbing gas should be considered as a lower limit. Further, the inconsistent use or lack of reporting of the assumed metallicity, neutral fraction, and scaling factors can add uncertainty to any analysis using published data.

The hypothesis of this paper is that the IGM contributes to the total hydrogen column density, with the contribution increasing with redshift, as observed in GRB $N_{\rm HX}$, and that by correcting the GRB $N_{\rm HX}$ using a more realistic GRB intrinsic metallicity, and estimating the host $N_{\rm H}$ using the measured neutral intrinsic $N_{\rm H\,I}$ adjusted for ionization fraction (from optical spectra (rest-frame UV) of GRB afterglow), we can isolate a more accurate $N_{\rm HIGM}$ contribution.

The objectives of this paper are:

(i) A review of the literature on the metallicities of GRB host environments to obtain improved values to use when estimating NHX,

(ii) To present a revised GRB NHX, using these more realistic metallicities and hence to update the NHX – redshift relation.

(iii) To isolate the IGM contribution to the total NHX in GRB, by using GRB ionized corrected NH I (NH I, IC) as an estimate of the GRB intrinsic NHX and plotting NHX – NH I, IC against redshift, after the improved metal corrections have been used.

(iv) To compare an estimated NHIGM based on a simple model of the IGM with our lower envelope for NHX based on realistic metallicities and with the intrinsic $N_{\rm H\,I,\,IC}$ removed.

Section 2 sets out the methodology, data selection approach, and the data used. Section 3 presents the results with a discussion and an analysis. Section 4 sets out the main conclusions. The Appendix gives further details on the Section 3 analysis, including the metallicity and the resulting fractional increase in $N_{\rm HX}$ with redshift. Throughout this paper, the term 'metallicity' is used synonymously with metal abundance [X/H][3]. Where relevant, the ΛCDM cosmology variables used are $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm m} = 0.3$, and $\Omega_{\Lambda} = 0.7$ unless otherwise stated.

2 METHODOLOGY AND DATA SELECTION

The full sample used here consists of all *Swift* X-ray Telescope (XRT) (Burrows et al. 2005) observed GRB with spectroscopic redshift available up to 2019 July 31, plus GRB090429B which has a photometric redshift of 9.4 (Cucchiara et al. 2011). The vast majority of the $N_{\rm HX}$ data is taken from the *Swift* repository to ensure a homogeneous data set (S13). Alternative sources were used only where detections with measured errors were available, and where *Swift* has only reported column density lower limits consistent with zero, or where the errors reported in the alternative source were smaller (e.g. Arcodia, Campana & Salvaterra 2016 used a stricter methodology by selecting specific time intervals when hardness ratios were constant to minimize spectral variations). For all sources, we endeavoured to ensure the methodology and selection criteria were consistent in terms of confidence level and XSPEC models used. Data from the *Swift* repository for $N_{\rm HX}$ were taken from the Photon Counting Late Time mode, as they are most likely to be a more stable, final value since spectral slope evolution is more prevalent at early times, leading to poor quality fits using a single power law (Page, K., private communication). All $N_{\rm HX}$ error bars reported are the 90 per cent confidence range, unless otherwise stated.

We follow the *Swift* repository reporting conventions for $N_{\rm HX}$ i.e. we treat as an upper limit the cases where the best fit $N_{\rm HX}$ is zero. Further, where the lower limit of the 90 per cent confidence interval includes zero, we use the best-fitting $N_{\rm HX}$ but use a different symbol for these objects in our figures.

Where we refitted spectra for analysis, XSPEC v12.10.1 was used (Arnaud 1996). Spectra were fitted with a power law in the X-ray band from 0.3 to 10.0 keV, which is suitable for the vast majority of GRB and again is consistent with the *Swift* repository (S13). A fixed Galactic component is taken from *Swift* based on Willingale et al. (2013). The model used in XSPEC was *tbabs∗ztbabs∗po* where the initial assumption we want to use is that all absorption in excess of Galactic is at the host redshift. *tbabs* is the galactic ISM absorption model, *ztbabs* is the same model placing the absorption at a fixed redshift and *po* is the power-law intrinsic spectral model. Isotopic abundances from W00 were used with the assumption of solar metallicity initially. In Section 3, where we examine more realistic metallicities for GRBs, the XSPEC model *tbvarabs* was used instead of *ztbabs* which allows the individual metal abundances to be varied from solar values. Cash statistics (Cstat) were used in XSPEC as this is required for spectra with low count rates, and is consistent with the *Swift* repository (Cash 1979). In all refits, we took the best fits based on minimum reduced χ-squared.

The selection of a sample can introduce bias. Perley et al. (2016) found that GRB with measured redshifts tend to be found in brighter galaxies, which could produce such a bias. However, Rahin & Behar (2019) compared the $N_{\rm HX}$ versus redshift trend for that paper's (smaller) unbiased sample with the full *Swift* sample and found no notable difference. In, Section 3, where we require that GRB have both optical and X-ray spectra, this requirement can introduce a selection effect against dim or highly dust extinguished GRBs (S11). It is estimated that between 25 and 40 per cent of GRBs

[3] [X/H] = log(X/H)-log(X/H)$_{\rm solar}$, where X is the metal element, and H is Hydrogen

Table 1. The GRB full sample. Column (1) GRB identification, (2) spectroscopic redshift (photometric for GRB090429B), (3) $\log(N_{\rm HX}\ {\rm cm}^{-2})$, (4) Refs for $\log(N_{\rm HX}\ {\rm cm}^{-2})$(note all are from the *Swift* repository if no ref given), (5) log $(N_{\rm HI}\ {\rm cm}^{-2})$ (all from Tanvir et al. 2019). Those with 'IC' have been corrected for ionization fraction, (6) [X/H], (7) refs for [X/H].

GRB	z	$\log(N_{\rm HX}\ {\rm cm}^{-2})$	$N_{\rm HX}$ Ref.	$\log(N_{\rm H\,I,IC}\ {\rm cm}^{-2})$	[X/H]	[X/H] Ref.
090926A	2.11	$21.74^{+0.19}_{-0.17}$	1	21.55 ± 0.10	-1.97 ± 0.11	2
09 0809	2.74	$21.85^{+0.27}_{-0.85}$	–	21.70 ± 0.20	-0.86 ± 0.13	2
08 0210	2.64	$22.32^{+0.21}_{-0.32}$	–	21.90 ± 0.10	-1.21 ± 0.16	3
09 0313	3.38	$22.64^{+0.13}_{-0.18}$	–	21.30 ± 0.20	-1.40 ± 0.30	3
120909A	3.93	$22.41^{+0.16}_{-0.24}$	–	21.70 ± 010·	-1.06 ± 0.20	2

Table 1 Refs: (1) (Zafar et al. 2018), (2) (Bolmer et al. 2019), (3) (Arabsalmani et al. 2018)

are undetected in the optical wavelength range as a result of dust extinction (e.g. Greiner et al. 2011). Therefore, the conclusions in that section may not apply to dust extinguished or dim GRBs.

All data for $N_{\rm H\,I}$ were taken from Tanvir et al. (2019). Of the 140 objects in their sample, we have used 128 which have $N_{\rm HX}$ data in our analysis. In Section 3, we adjust the $N_{\rm H\,I}$ for the ionization fraction for $\log(N_{\rm H\,I}\ {\rm cm}^{-2}) < 20$. The redshift range in our full GRB $N_{\rm HX}$ sample is from 0.03 to 9.4, while the $N_{\rm H\,I}$ sample range is from 1.6 to 6.73. The lower $N_{\rm H\,I}$ redshift cut-off is due to the requirement that the observed wavelength of the Ly α absorption line be in the visible/UV band.

In Section 3.5, we analyse the impacts on $N_{\rm HX}$ of using metallicities that more realistically reflect the LOS absorption through the entire host galaxy. X-ray absorption is dominated by the metals and H and He are relatively unimportant. Below 1 keV, C, N, O, and Ne are the main absorbers, while above 1 keV, Si, S, and Fe dominate (W00). W00 also note that interpreting X-ray observations is subject to the uncertainties remaining in the atomic data. Data for metallicities in Table 1 are all UV/optical absorption line based. Absorption metallicities measure the metal enrichment of gas along the LOS from the GRB through the galaxy.

Table 1 contains the data for the full GRB sample for redshift, $N_{\rm HX}$, $N_{\rm H\,I,\,IC}$ and metallicity where data are available. We list in the table extract, a sample of five GRB with data for all columns (see the online version of this paper for the complete table with all values listed). (See Appendix A1 an investigation about whether a flux limited sample would introduce any substantial bias).

3 RESULTS AND ANALYSIS

In this section, we examine the distributions of GRB $N_{\rm HX}$ and $N_{\rm H\,I}$ with redshift, and the use of adjusted $N_{\rm H\,I,\,IC}$ as an approximation for the GRB host intrinsic contribution to the total integrated hydrogen column density to isolate the IGM column density. We then examine GRB host metallicity to derive a more accurate metallicity to use in XSPEC fitting to get an improved $N_{\rm HX}$. Finally, we replot the distributions of revised $N_{\rm HX}$ and $N_{\rm HX} - N_{\rm H\,I,\,IC}$ with redshift to get the lower envelope of GRB $N_{\rm HX}$ as a step towards constraining $N_{\rm HIGM}$.

3.1 $N_{\rm HX}$ and redshift

Fig. 1 shows the distribution of $N_{\rm HX}$ with redshift (based on the standard assumptions of solar metallicity and that all the absorption, $N_{\rm HX}$ is at the GRB redshift) for the full *Swift* observed sample with known spectroscopic redshift (with the exception of GRB090429B where the redshift is a photometric estimate). Where an estimate of the actual $N_{\rm HX}$ was available from the *Swift* repository but the 90 per cent confidence interval included zero, these are plotted with a yellow dot. Yellow dots with arrows are the upper limits where the *Swift* repository has a best fit of zero $N_{\rm HX}$. A relationship or dependence between $N_{\rm HX}$ and redshift has been reported in several papers over the last decade (e.g. Behar et al. 2011; S13). The correlation statistics for the full 352 GRB sample are Pearson $r = 0.29$ and Spearman $\rho = 0.55$. For the detection only sample (226 GRB), the correlation results are Pearson $r = 0.51$ and Spearman $\rho = 0.49$. Both samples pass the null hypothesis test, indicating that the correlation seen is significant. However, when we used an error weighted least-squares fit to a linear model (χ^2), the reduced χ^2 was large, indicating either that a simple linear redshift relation is not a realistic model or that there is an additional substantial source of scatter. We checked for the impact of outliers with high $N_{\rm HX}$ and very small error bars. To bring the reduced χ^2 close to 1 would have required the removal of over 10 per cent of the sample.

We include in Fig. 1 a simple model of the diffuse IGM following S13 (equation 5 in that paper) based on

$$N_{\rm HIGM} = (n0c/H0) \int_0^z (1+z')^2/((1+z')^3\Omega_{\rm M} + \Omega_\Lambda)^{1/2} {\rm d}z', \quad (2)$$

where n_0 is the hydrogen density at redshift zero, taken as $1.7 \times 10^{-7}{\rm cm}^{-3}$ from Behar et al. (2011).

The solution to the integral from Shull & Danforth (2018) (equation 4 in that paper) is

$$(2/(3\Omega_M))\{(\Omega_M(1+z)^3 + \Omega_\Lambda)^{1/2} - 1\}. \quad (3)$$

In Fig. 1, we can see the $N_{\rm HIGM}$ model runs through the GRB data points. If it is to represent the diffuse IGM only, we would expect all the GRB to be above this curve (if there were no measurement errors). Given the large error bars for many GRB, the IGM hypothesis could still be plausible where a small fraction, 10 per cent approximately given the 90 per cent confidence, are below the curve. A much higher fraction than 10 per cent are below the IGM curve in Fig. 1. However, the IGM model is admittedly very simple and therefore could poorly represent the real Universe. Also, not all LOS will be at the mean density.

We note that our model is based on the mean hydrogen density as a simple model, so the metallicity uncertainty in the IGM does not affect it directly. In our next paper, the metallicity in different phases of the IGM will be reviewed in detail. We examine this further in Section 3.3 onwards.

3.2 $N_{\rm H\,I}$ review with redshift

In this section, we review the most recent substantial GRB $N_{\rm H\,I}$ sample from Tanvir et al. (2019) which consists of new measurements combined with those from literature. We examine this latest sample for any relations between $N_{\rm H\,I}$ and redshift, or with $N_{\rm HX}$.

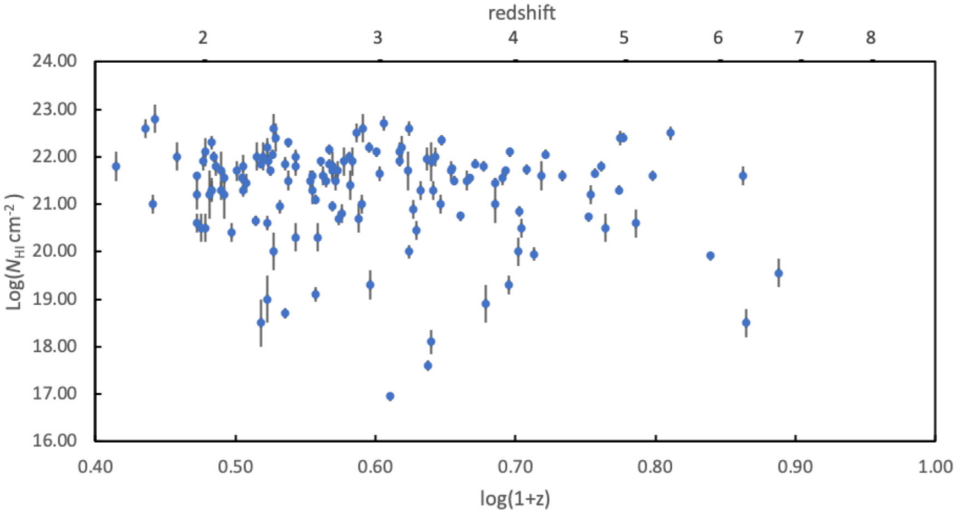

Figure 2. Distribution of $N_{\rm H\,I}$ for GRB with redshift. No strong trends with redshift are visible.

Optical spectroscopy enables the approximate location of any neutral hydrogen absorber to be identified. GRB hosts are typically found to have high column densities of cold neutral gas, with a large fraction of GRB hosts containing a DLA system ($\log(N_{\rm HI}$ cm$^{-2}) >$ 20.3) or sub-DLA ($19.0 < \log(N_{\rm H\,I}$ cm$^{-2}) <$ 20.3) (S11). Much of the neutral gas component is found at a few hundred parsecs from the GRB (Ledoux et al. 2009).

Fig. 2 shows the distribution of the GRB $N_{\rm H\,I}$ sample with redshift. Where not specified in literature, we have assumed the errors are Gaussian and correspond to one standard deviation. The Pearson r is -0.15 and Spearman ρ -0.10. Both fail the null hypothesis test i.e. there is no statistically significant correlation. The lack of a detectable redshift correlation for $N_{\rm H\,I}$ is in contrast with the clear redshift correlation for $N_{\rm HX}$. Clearly, this does not provide support for the argument that redshift evolution in the GRB host properties is responsible for the redshift correlation for $N_{\rm HX}$. (See Appendix A2 for further review of any $N_{\rm H\,I}$ correlation with $N_{\rm HX}$.

3.3 Using $N_{\rm H\,I,IC}$ as proxy for the GRB intrinsic contribution to $N_{\rm HX}$

Part of the aim of this paper is to attempt to isolate an IGM contribution to $N_{\rm HX}$. Here, we investigate the plausibility of assuming that the host intrinsic hydrogen column density is equal to the measured ionized corrected intrinsic neutral column ($N_{\rm H\,I,\,IC}$) and examine the resulting residual column's dependence on redshift. To do this, we first make an ionization correction (IC) to the hydrogen column density as measured by $N_{\rm H\,I}$ using the approach described in Fumagalli, O'Meara & Prochaska (2016) who observed that neutral fraction is a function of $N_{\rm H\,I}$. The neutral fraction drops rapidly from $\sim$0.7 at $\log(N_{\rm H\,I}$ cm$^{-2}) \sim 20$ to $\sim$0.02 at $\sim\log(N_{\rm H\,I}$ cm$^{-2}) \sim 18$ with a 0.3 dex characteristic error. As the vast majority of GRB are in hosts with high column densities DLAs, only 11 out of 128 GRB sample required an IC.

Of the 128 GRB with data for both $N_{\rm HX}$ and $N_{\rm H\,I}$ in our sample, 96 have $N_{\rm HX} > N_{\rm H\,I,\,IC}$. In Fig. 3, these are plotted, with the remaining 32 placed at the bottom of the figure for completeness. 32 of the GRB from this sub-sample (blue dots) are detections for both column densities. Where an estimate of $N_{\rm HX}$ was available from the *Swift* repository but the confidence interval included zero, the object was plotted with a orange dot. The orange dots with arrows are the upper limits where *Swift* repository has a best fit of zero $N_{\rm HX}$. The correlation statistics for the sub-sample of 32 GRB detections with $N_{\rm HX} > N_{\rm H\,I}$ are Pearson $r = 0.75$ and Spearman $\rho = 0.69$ (for the full 128 sub-sample taking limits as detections one gets $r = 0.55$, $\rho = 0.59$.) The $N_{\rm HX} - N_{\rm H\,I,\,IC}$ relation with redshift is much more significant than for a $N_{\rm HX}$ alone. The reduced $\chi^2 = 1.02$ for a linear fit with the form $N_{\rm HX} - N_{\rm H\,I,\,IC}$ α $(1 + {\rm z})^{3.5\pm0.1}$. It can be argued that this result supports the case for using $N_{\rm H\,I,\,IC}$ as a proxy for the GRB intrinsic hydrogen column density, leaving the major remaining column density contribution being from the IGM. A caveat is that a large part of the sub-sample has been excluded as the X-ray or UV column density was not measured. The fraction excluded because $N_{\rm H\,I,\,IC} > N_{\rm HX}$ (32/128) is a cause for concern, as $N_{\rm HX}$ is supposed to be the total column density. The large error bars on $N_{\rm HX}$ may account for some of these. We can also see that the original $N_{\rm HIGM}$ model is higher than the majority of the estimated intervening absorption. This may indicate that the IGM model is too simple e.g. it ignores LOS variation, or that the parameters used in the simple model need to be adjusted. However, the result could well be due to unrealistic assumptions of metallicity and ionization for the GRB host galaxy. The next section examines the effects of assumptions about GRB host metallicity.

3.4 GRB host metallicity

GRB typically occur in sub-solar metallicity galaxy host environments (S13; Cucchiara et al. 2015; Krühler et al. 2015). Our sub-sample of 36 GRB with a range in redshift of $1.76 \leq {\rm z} \leq 5.91$, and [X/H] from -2.18 to 0.25 is plotted in Fig. 4. Where data from multiple sources were available, we took those with the smallest reported errors. Further, we only used data with detections and error bars, and excluded those with lower limits only. This resulted in omitting most of the 55 GRB from Cucchiara et al. (2015, hereafter C15). The Pearson and Spearman correlation coefficients are $r = -0.24$ and $\rho = -0.27$. However, both correlations fail the null hypothesis test, indicating there is no statistically significant correlation between GRB metallicity and redshift. The blue line is the best linear fit to the data is

$$[{\rm X/H}] = (-1.01 \pm 0.04) - (0.09 \pm 0.01){\rm z} \quad (4)$$

This possible mild metallicity evolution of GRB intrinsic gas with

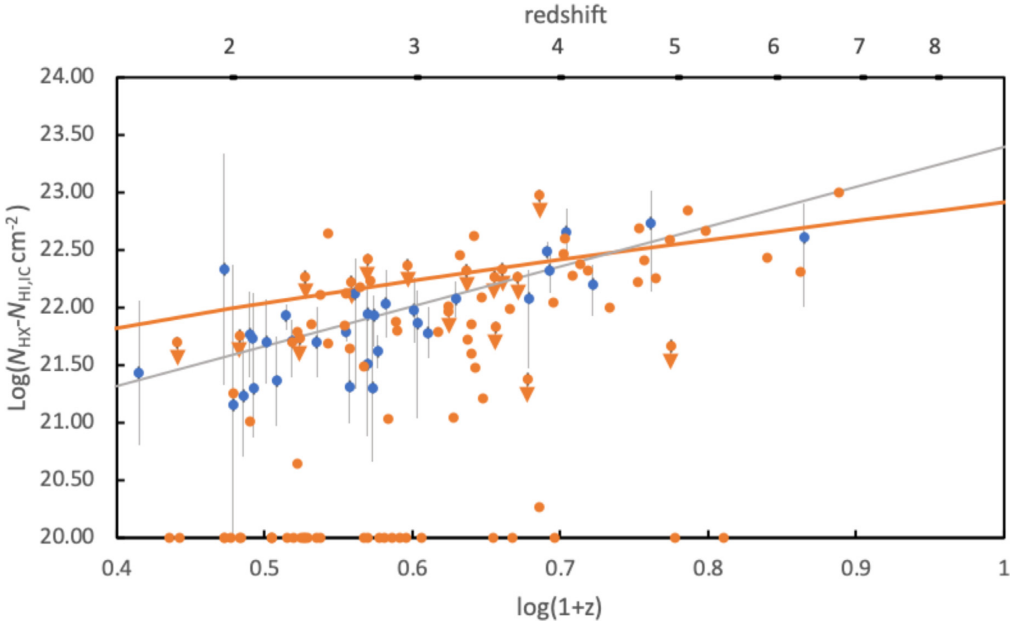

Figure 3. Distribution of total $N_{\rm HX}$ minus the localized $N_{\rm H\,I,\,IC}$ (which is being used as a proxy of the intrinsic GRB hydrogen column density) in the sub-sample of 128 GRB with data on both $N_{\rm HX}$ and $N_{\rm H\,I,\,IC}$. The blue dots are GRB detections for both $N_{\rm HX}$ and $N_{\rm H\,I,\,IC}$. The orange dots are objects for which the 90 per cent confidence interval includes zero. The orange dots with arrows are the $N_{\rm HX}$ upper limits where *Swift* repository has a best fit of zero. Where $N_{\rm HX} - N_{\rm H\,I,\,IC} < 0$ they are placed at 20.0 on the y-axis. The orange line is the integrated diffuse IGM $N_{\rm HIGM}$, model described by equation (2). A power law fit to the $N_{\rm HX} - N_{\rm H\,I,\,IC}$ versus redshift trend scales as $(1+z)^{3.5+/-0.1}$ (grey line with reduced $\chi^2 = 1.02$). Pearson and Spearman correlation coefficients are 0.75 and 0.69, respectively, for the GRB detections, and 0.55 and 0.59 for the full sample where $N_{\rm HX} > N_{\rm H\,I,\,IC}$ *taking limits as detections.*

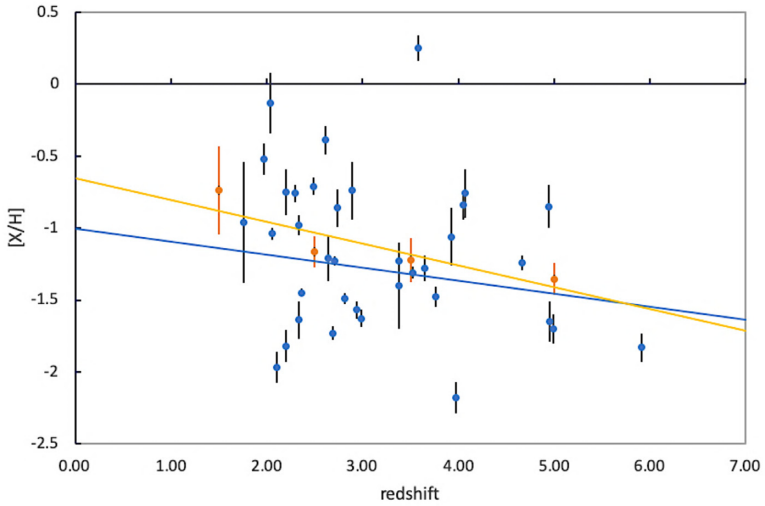

Figure 4. Distribution of the combined GRB absorption-based metallicities (blue dots) with redshift. The GRB absorption sample mean metallicity is -1.17 ± 0.09 (or 0.07 ± 0.05 Z/Z$_{\rm sol}$). The orange dots are the weighted average metallicity over specific redshift bins of bins $\Delta z = 1$ (except for z = 4 to 6 as there is only one GRB with z > 5) weighted by the total $N_{\rm H\,I}$. The Pearson and Spearman correlation coefficients are $r = -0.24$ and $\rho = -0.27$, both correlations failing the null hypothesis tests, indicating there is no statistically significant correlation between GRB metallicity and redshift. The blue line is the χ^2 linear fit to the blue dot data. The orange line is the best linear fit to the orange dots for the weighted average [X/H] and shows possible evolution (see Appendix A4 for more discussion).

redshift is noted in some of the literature. For non-GRB absorption systems, stronger evolution is seen. For example, De Cia et al. (2018) reported evolution with a slope of ∼0.32z. However, GRB metallicity does not appear to evolve as much, if at all, based on our sample. This is consistent with C15, for example.

It is well known that dust depletion affects the determination of metallicity in GRBs (e.g. Savaglio 2006; De Cia et al. 2013). In Bolmer et al. (2019), 22 GRB are studied at z > 2 for features including dust depletion measurements and any relation to redshift. Based on this sample, they found that, on average, the dust corrected metallicity [M/H] $= -1.09 \pm 0.50$ compared with -1.27 ± 0.37 for the uncorrected metallicity (0.08 Z/Z$_{\rm sol}$ versus 0.05 Z/Z$_{\rm sol}$, respectively). This is an average correction of 0.2 dex which is considerably lower than found by De Cia et al. (2018) (0.4–0.5 dex)

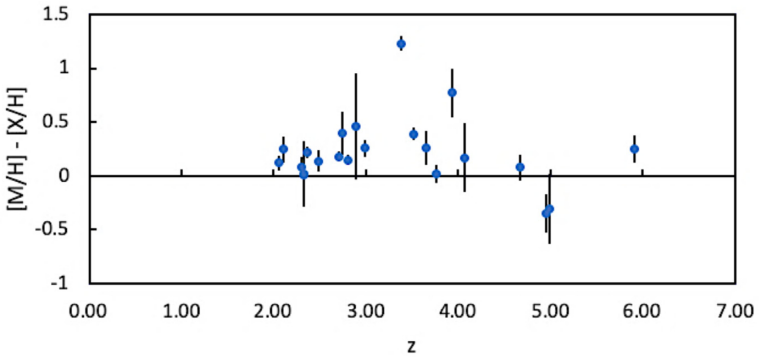

Figure 5. Plot of the dust correction to metallicities [M/H]-[X/H] by redshift for the Bolmer et al. (2019) GRB sample. The Pearson and Spearman coefficients are −0.14 and 0.03, respectively, and both fail the null hypothesis tests for a significant correlation. There is no detectable evolution

for non-GRB objects. In Fig. 5, we plot the dust correction [M/H]-[X/H] versus redshift for the Bolmer et al. (2019) sample, to see if there is any obvious evolution with redshift. No detectable evolution is seen. The Pearson and Spearman coefficients for Fig. 5 are −0.14 and 0.03, respectively and both fail the null hypothesis tests for a significant correlation.

Where actual dust corrections are not available, an argument can be made for using a standard dust correction to metallicity for XSPEC fitting for example, based on the Bolmer et al. (2019) mean value of 0.2 dex. This mean correction increases the average metallicity in our sample from 0.07 to 0.11 Z/Z_{sol}.

While this is an important correction to [X/H], the impact on revised N_{HX} is very small as the corrected metallicity is still <<solar [X/H]. Testing a sample of GRBs at redshift from 1 to 7, the change in log(N_{HX}) after making a dust correction was 0.03 to 0.06 dex. Further, how any dust correction is estimated and used in the literature is not always clear. In conclusion, given that the impact of an average dust correction to N_{HX} is very small, we do not consider that a standard dust correction to the metallicity adjustment is appropriate.

In conclusion, from our GRB metallicity review, the GRB absorption sample mean metallicity is equal to -1.17 ± 0.09 or (0.07 ± 0.05 Z/Z_{sol}). In further analysis therefore, we use the actual metallicity, dust corrected, for detections where available. While noting that some of the literature claims that there is possible mild redshift evolution in GRB host absorption, for the reasons outlined, we chose to use the average metallicity, without evolution or dust correction, of 0.07 Z/Z_{sol} for the remaining GRB. This is certainly a more realistic value than simply assuming solar metallicity in revisiting the N_{HX} for the full GRB sample in the next section.

3.5 Impact of metallicity assumptions on N_{HX}

We wish to examine the impact on GRB N_{HX} fits in XSPEC of using actual dust corrected metallicities for GRB detections where available. For the remaining GRB, we examine a more realistic average host metallicity than solar, and importantly, look at the variation with redshift. To do this, we used an XSPEC model *tbabs∗tbvarabs∗po* for the X-ray data from GRB151027A (a very high S/N GRB), varying the modelled host redshift between 0 and 10 and testing for metallicities $Z/Z_{sol} = 0.07$ (the mean from our sample in Section 3.4), and solar. A lower metallicity results in an increased fitted N_{HX}, with the increase varying with redshift (see Appendix A3 for more details). In order to see whether this correction is consistent for different GRB X-ray spectra, we plotted the fractional increase in fitted N_{HX} with redshift for a test sample of three high S/N GRB spectra with differing reported redshifts and N_{HX}. Again, we varied the redshifts between 0 and 10 and used metallicity = 0.07 Z/Z_{sol} compared with the value assuming solar metallicity. The fractional increase in N_{HX} with redshift is very similar for the three GRB). A power law-fit to the increase for GRB151027A is (orange line in Appendix Fig. A4)

$$\Delta\log(N_{HX}\mathrm{cm}^{-2}) = (0.59 \pm 0.04)\log(1+z) + 0.18 \pm 0.02. \quad (5)$$

A more accurate power law could be obtained from a combined fit for the three GRB. However, this fit is deemed sufficient for the purposes of analysing the impact of a more realistic general metallicity assumption when calculating N_{HX}.

3.6 GRB N_{HX} revised for realistic host metallicity

Using actual metallicities, dust corrected where available, and the above power law relation for the remaining GRB, we use the new N_{HX} for our full GRB sample and replot the relation with redshift in Fig. 6.

The Pearson and Spearman correlation coefficients are 0.59 and 0.61, respectively, for GRB detections in Fig. 6, which are stronger than prior to the correction for a low metallicity. The blue dots are GRB detections. The orange dots with arrows are upper limits where the fitted N_{HX} are 0, and orange dots are where the 90 per cent confidence interval includes zero. The orange line in Fig. 6 is the simple model IGM line from equation (2). The blue line is an estimate of the GRB lower envelope based on a requirement of having 90 per cent of detections including their error bars to be above the envelope.

$$\log(N_{HX}\mathrm{cm}^{-2}) = 20.3 + 2.4\log(1+z). \quad (6)$$

The green line is an estimate of the lower envelope based on the requirement that 99 per cent of all GRB measurements, ignoring error bars and treating any upper limits as detections, are above the envelope, using the rule of thumb in Campana et al. (2015) (note that they put N_{HX} at the top of the 90 per cent confidence interval where the 90 per cent confidence interval of a fit includes zero, whereas we use the Swift best estimate N_{HX} which is lower, except for those with best estimates equal to zero).

$$\log(N_{HX}\mathrm{cm}^{-2}) = 19.5 + 2.4\log(1+z). \quad (7)$$

The envelope fits may give an indication of the maximum N_{HIGM} potential contribution to N_{HX}. Both the envelopes have been assumed to scale with redshift as $(1+z)^{2.4}$ (Campana et al. 2014),

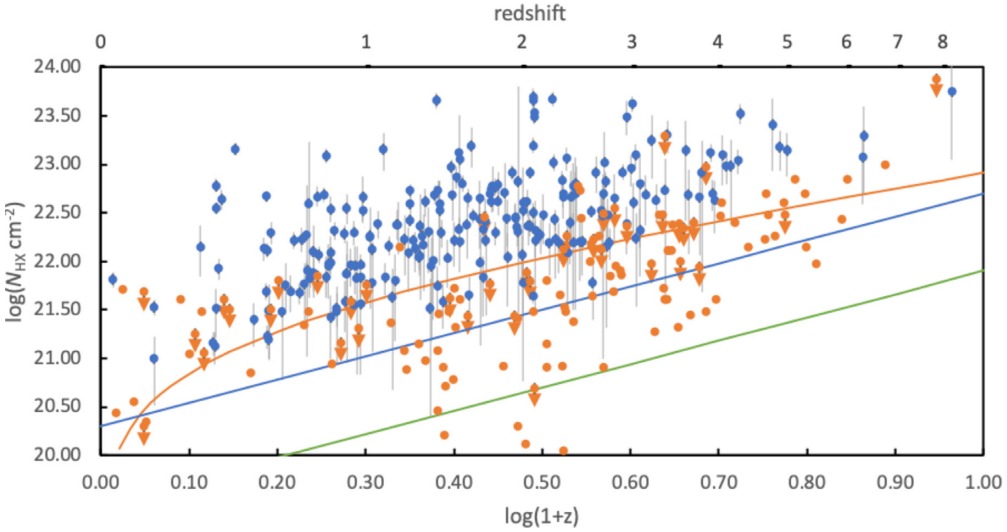

Figure 6. GRB $N_{\rm HX}$ revised using actual metallicities, dust corrected where available and a mean metallicity of 0.07 $Z/Z_{\rm sol}$ for the remaining GRB. The blue dots are GRB detections. The orange dots are objects for which the 90 per cent confidence interval includes zero. The orange dots with arrows are the upper limits where the *Swift* repository has a best fit of zero. The orange line is the simple model IGM line from equation (2), $N_{\rm HIGM}$. The blue line is the GRB lower envelope based on the requirement for 90 per cent of detections, including error bars to be above the envelope, with log ($N_{\rm HX}$ (z = 0) cm^{-2}) = 20.3. The green line is the envelope with the requirement that 99 per cent of all GRB, ignoring error bars, are above the envelope, which has log($N_{\rm HX}$ (z = 0) cm^{-2}) = 19.5 following the rule of thumb in Campana et al. (2015). Both these envelopes are plotted with assumed slopes of $(1 + z)^{2.4}$ (Campana et al. 2014). The Pearson and Spearman correlation coefficients are 0.59 and 0.61, respectively, for GRB detections and 0.62 and 0.62 for the full sample treating the limits as detections.

and we note that this may only be realistic for a cold absorber and not for a warm absorber (S13). The GRB LOS goes through a wide range of environments with different temperatures and densities. This will change the effective absorption cross-section at different frequencies. However, we will retain the cold absorber approximation for the current analysis. Using a χ^2 fit, the revised $N_{\rm HX}$ for the detections scale as $N_{\rm HX}$ α $(1 + z)^{1.94\pm0.04}$. However, a large reduced χ^2 indicates that the relationship is not a simple power law or that the data has a large additional source of scatter. We explore this further in Section 3.7. Of 226 GRB detections, only 11 are now below the $N_{\rm HIGM}$ curve, not taking error bars into account.

3.7 Revised GRB $N_{\rm HX} - N_{\rm HI,IC}$

As before, to isolate the IGM contribution to $N_{\rm HX}$, we subtract from the revised $N_{\rm HX}$ the GRB $N_{\rm HI}$ adjusted for an ionization correction, as a proxy for intrinsic hydrogen column density, as we did in Section 3.3, and plot the result against redshift in Fig. 7.

Of the 128 GRB with data for both $N_{\rm HX}$ and $N_{\rm HI}$ in our sample, 122 now have $N_{\rm HX} > N_{\rm HI,IC}$. The $N_{\rm HX} - N_{\rm HI,IC}$ for GRB detections now has a best-fitting power-law slope of $(1 + z)^{3.1+/-0.3}$ (grey line – with reduced $\chi^2 = 2.6$). The Pearson and Spearman correlation coefficients are 0.65 and 0.67, respectively, for the GRB detections, and 0.53 and 0.62 for the full sample with best fits being treated as detections, where upper limits are treated as detections where the best fit equals zero, and where $N_{\rm HX} > N_{\rm HI,IC}$.

This final figure is our best representation of the use of GRB X-ray spectral fits to potentially constrain the IGM hydrogen column density. The blue line is the GRB lower envelope based on the requirement for 90 per cent of detections (equation 6), including error bars, to be above the envelope. Using equation (2), the orange line is $N_{\rm HIGM}$ for a mean hydrogen number density $n_0 = 1.7 \times 10^{-7}$cm^{-3}.

In Fig. 3, 32 GRB had $N_{\rm HX} < N_{\rm HI,IC}$ compared to only six based on the revised $N_{\rm HX}$ for our updated GRB host metallicity in Fig. 7. Therefore, the more realistic GRB metallicity generates a more plausible $N_{\rm HX}$, if it is assumed to represent the total hydrogen column density, which hence must be greater than the intrinsic column density. Given this requirement, we examined the six GRB where the fitted $N_{\rm HX}$ was less than $N_{\rm HI,IC}$. For these objects, Table 2 lists: (1) GRB name; (2) log($N_{\rm HI}$ cm^{-2}); (3) Measured metallicity from literature if available; (4) revised log($N_{\rm HX}$ cm^{-2}); (5) Whether the revised $N_{\rm HX} > N_{\rm HI,IC}$); (6) log($N_{\rm HIGM}$ cm^{-2}) at the GRB redshift; and (7) whether the revised $N_{\rm HX} - N_{\rm HI,IC}$ is greater than $N_{\rm HIGM}$. We also include GRB180624A in the table, which had a revised $N_{\rm HX} - N_{\rm HI,IC}$ substantially below $N_{\rm HIGM}$.

We refitted each GRB in XSPEC using *tbabs∗tbvarabs∗po* using the actual reported metallicity, or 0.07 $Z/Z_{\rm sol}$ otherwise. As can be seen from Table 2, all GRB now have $N_{\rm HX} > N_{\rm HI,IC}$. Further, all show $N_{\rm HX}$ less $N_{\rm HI,IC}$ (GRB160227A within error bars) as proxy for the host intrinsic column density, being greater than $N_{\rm HIGM}$. The refitting using the actual redshift and metallicity (or 0.07 otherwise) gives a higher corrected $N_{\rm HX}$ as the power-law correction approximation marginally understates the actual relation between metallicity correction and redshift for redshift between $0.3 < \log(1 + z) < 0.8$ (see Appendix A3). Of the 67 GRB detections, five lie below the $N_{\rm HIGM}$ curve in Fig. 7. None lie below the $N_{\rm HIGM}$ curve after refitting for more realistic or actual metallicity.

In conclusion, by using actual metallicities, dust corrected where available, and a more realistic average GRB metallicity than the standard solar assumption for the remainder, we have shown that the revised larger $N_{\rm HX}$ is greater than an ionization corrected $N_{\rm HI}$ for our entire sample of 128 GRB, where measurements of both are available, together with a spectroscopic redshift. Further, the lower envelope of $N_{\rm HX} - N_{\rm HI,IC}$ is potentially a useful constraint on the IGM contribution to $N_{\rm HX}$. Finally, the metallicity revised $N_{\rm HX} - N_{\rm HI,IC}$ for detections are mostly above the simple model $N_{\rm HIGM}$ curve further suggesting that this is a useful constraint on the IGM hydrogen column density.

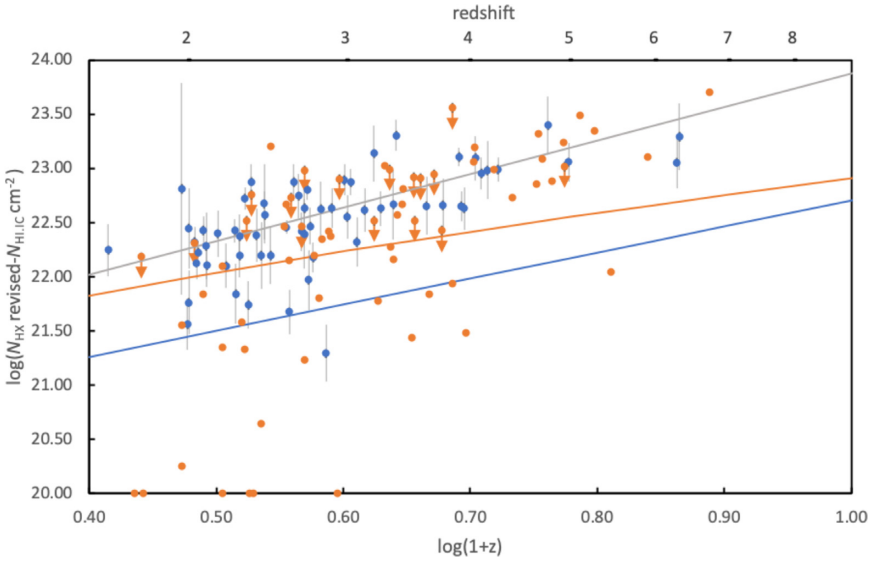

Figure 7. Distribution of GRB revised $N_{HX} - N_{H\,I,\,IC}$ with redshift. The blue dots are GRB detections and the orange with arrows are upper limits where the N_{HX} best fit is zero, or orange dots where the 90 per cent confidence interval of the x-ray fit includes zero. Where $N_{HX} < N_{H\,I,\,IC}$ the GRB are placed at the bottom of the figure for completeness. The blue line is the GRB lower envelope based on the requirement for 90 per cent of detections, including error bars to be above the envelope, and has $\log(N_{HX}\ (z = 0)\ cm^{-2}) = 20.3$. The orange line is the original simple model N_{HIGM} for $n_0 = 1.7 \times 10^{-7} cm^{-3}$. The $N_{HX} - N_{H\,I,\,IC}$ for GRB detections best fit has a power-law slope of $(1 + z)^{3.1\pm0.3}$ (grey line, reduced $\chi^2 = 2.6$). The Pearson and Spearman correlation coefficients are 0.65 and 0.67, respectively, for the GRB detections, and 0.53 and 0.62 for the full sample with limits treated as detections and where $N_{HX} < N_{H\,I,\,IC}$.

Table 2. Summary analysis for refitting of six GRB where the revised $N_{HX} <, N_{H\,I,\,IC}$ and one where the revised $N_{HX} - N_{H\,I,\,IC}$ is substantially below N_{HIGM}. (1) GRB name; (2) $\log(N_{H\,I}\ cm^{-2})$; (3) Measured metallicity from literature if available; (4) revised $\log(N_{HX}\ cm^{-2})$; (5) Whether the revised $N_{HX} > N_{H\,I,\,IC}$); (6) $\log(N_{HIGM}\ cm^{-2})$ at the GRB redshift; and (7) whether the revised $N_{HX} - N_{H\,I,\,IC}$ is greater than N_{HIGM}.

GRB	$\log(N_{H\,I}\ cm^{-2})$	Measured Z/Z_{sol}	$Log(N_{HX}\ cm^{-2})$	New $N_{HX} > N_{H\,I,IC}$?	$\log(N_{HIGM}\ cm^{-2})$ at GRB z	$N_{HX} - N_{HI,IC} >$ IGM?
050922C	21.55+/−0.10	0.15[a]	22.04*	Y	22.04	Y
120119A	22.60+/−0.20	0.11[b]	22.45+/−1.74	Y	21.91	Y
120815A	22.05+/−0.10	0.04[c]	22.16*	Y	22.09	Y
121027A	22.80+/−0.30	**	22.86+/−1.20	Y	21.92	Y
160227A	22.40+/−0.20	**	22.08+/−1.15	Y	22.10	Within error bars
181020A	22.20+/−0.10	0.27[c]	22.38+/−1.25	Y	22.22	Y
180624A	22.5+/−0.20	**	22.70+/−3.40	Y	22.21	Y

Notes: Z/Z_{sol} references: [a](Arabsalmani et al. 2018), [b](Heintz et al. 2019), [c](Bolmer et al. 2019), *$\log(N_{HX})$ lower error bar not with 90 per cent confidence,**metallicity unknown so 0.07 Z/Z_{sol} used in XSPEC fitting.

4 CONCLUSION

We compiled a large sample of all *Swift* X-ray Telescope observed GRB with spectroscopic redshifts up to 2019 July 31 (with a photometric redshift only for GRB090429B). Of this sample of 352 GRB with fitted X-ray equivalent hydrogen column densities, 128 have also have intrinsic neutral hydrogen column density measurements. We have also compiled a sample of absorption-based metallicity data. The main aims of this paper are to generate improved N_{HX} for our sample by using more realistic host metallicity and, by approximating the host intrinsic hydrogen column density as equal to the measured $N_{H\,I}$, with an ionization correction applied, to isolate the more accurate IGM column density contribution.

We analysed the impacts on N_{HX} of using metallicities that more realistically reflect the LOS absorption through the host galaxy than the standard use of the solar abundance. We discussed the possibility of using an average dust correction where actual measurements were not available but it had an insignificant effect on N_{HX}.

Our main findings and conclusions are:

1. While some of the literature notes that GRB metallicity shows a mild evolution with redshift, the Pearson, and Spearman correlation coefficients for our sample are −0.24 and −0.27, respectively, and both correlations fail the null hypothesis test, indicating there is no detected trend. Further, the large reduced $\chi 2$ of the fit means either that a linear model is not a good description of any potential relation or that there is a large additional source of scatter. Hence we do not find a statistically significant relation between GRB metallicity and redshift.

2. The GRB absorption sample mean metallicity is [X/H] = −1.17 ± 0.09 (or 0.07 ± 0.05 Z/Zsol). This is substantially lower than the assumption of solar metallicity used as standard for many fitted NHX.

3. We find that using a lower GRB host metallicity results in increasing the fitted NHX with the correction scaling with redshift. In order to determine this relation at mean metallicity 0.07 Z/Zsol, we plotted the fractional increase in NHX with redshift for some trial fits. We find that the fractional increase in NHX with redshift is very similar for a range of GRB fits. The power-law relation for

GRB151027A, used as a standard GRB for metallicity 0.07 Z/Zsol is $\Delta\log(\mathrm{NHX\ cm} - 2) = (0.59 \pm 0.04)\log(1 + z) + (0.18 \pm 0.02)$.

A more accurate power law could be obtained from a combined fit of a large sample of GRB. However, this is sufficient for our purposes of analysing the impacts of a more realistic general metallicity assumption for calculating NHX.

4. Using actual metallicities, dust corrected where available, and, for the remaining GRB, our power-law relationship for the mean GRB host metallicity of 0.07 Z/Zsol, we revised the NHX for our full GRB sample and replotted the relation with redshift. To more accurately isolate the IGM contribution to the total hydrogen column density, we subtracted from the revised NHX the GRB NH I after ionization correction, as a proxy for the intrinsic hydrogen X-ray column density, and plotted the result against redshift. Of the 128 GRB with data for both NHX and NH I in our sample, only six have NH I, IC greater than the revised NHX, compared to 32 when solar metallicity is assumed. Therefore, using more realistic GRB metallicities generates an improved NHX, if it is interpreted as representing the total hydrogen column density, which must be greater than the local neutral column density. The estimated NHX – NH I, IC for GRB detections now has a redshift dependence of $(1 + z)3.1+/-0.3$ for the GRB detections, compared with power laws of $3.5+/-0.1$ for NHX fitted assuming solar abundance. The Pearson and Spearman correlation coefficients are 0.65 and 0.67, respectively, for the GRB detections, and 0.53 and 0.62, respectively, for the full sample where NHX > NH I, IC.

5. The lower envelope of the revised NHX > NH I, IC plotted against redshift has $\mathrm{NHX}\,(z = 0) = 20.3\ \mathrm{cm} - 2$ for our GRB sample of revised NHX – NH I, IC. This is taken to be representative of the maximum IGM hydrogen column density, based on the requirement for 90 per cent of detections, including error bars, to be above the envelope. Using this approach, we estimate the IGM n0 to be equal to $1.7 \times 10^{-7}\ \mathrm{cm}^{-3}$ for the NHIGM curve which is consistent with that used by Behar et al. (2011) and S13.

X-ray spectroscopy at higher resolution and at higher signal-to-noise than is currently available would be required to detect absorption edges from individual ions in GRB. Such observations in the future, will provide valuable data on the distribution of the material along the line of sight, including its temperature, composition, density, and ionization state. The value found for the column density is almost always determined assuming a 100 per cent neutral absorbing gas. This neutral assumption would cause the N_{HX} to be underestimated if incorrect. Therefore, we can further improve the N_{HX} and GRBs as probes of the IGM when higher resolution X-ray spectroscopy becomes available. We plan to examine the properties of the IGM such as metallicity, temperature, and density in a subsequent paper to develop a better IGM model and compare it with the results of this paper.

ACKNOWLEDGEMENTS

We thank M. Fumagalli for useful comments, K. Page for information on the *Swift* repository, and R. Mahmaud and K. Arnaud for advice on XSPEC. We also thank the referee for providing valuable and constructive feedback. SLM also acknowledges support from the Science and Technology Facilities Council (ST/P000541/1).

REFERENCES

Anders E., Grevesse N., 1989, Geochim. Cosmochim. Acta, 53, 197
Arabsalmani M. et al., 2018, MNRAS, 473, 3312
Arcodia R., Campana S., Salvaterra R., 2016, A&A, 590, 1
Arnaud K.A., 1996, Astron. Data Anal. Softw. Syst., 101, 17
Asplund M., Grevesse N., Sauval A.J., Scott P., 2009, Annu. Rev. Astron. Astrophys., 47, 481
Behar E., Dado S., Dar A., Laor A., 2011, Astrophys. J., 734, 26
Bolmer J. et al., 2019, A&A, 623, A43
Buchner J., Schulze S., Bauer F.E., 2017, MNRAS, 464, 4545
Burrows D.N. et al., 2005, Space Sci. Rev., 120, 165
Campana S., Salvaterra R., Ferrara A., Pallottini A., 2015, A&A, 575, A43
Campana S., Thöne C.C., de Ugarte Postigo A., Tagliaferri G., Moretti A., Covino S., 2010, MNRAS, 402, 2429
Campana S. et al., 2012, MNRAS, 421, 1697 (C12)
Campana S. et al., 2014, MNRAS, 441, 3634
Cash W., 1979, Astrophys. J., 228, 939
Cucchiara A., Fumagalli M., Rafelski M., Kocevski D., Prochaska J.X., Cooke R.J., Becker G.D., 2015, AJ, 804, 51 (C15)
Cucchiara A. et al., 2011, AJ, 736, 7
De Cia A., Ledoux C., Petitjean P., Savaglio S., 2018, A&A, 611, A76
De Cia A., Ledoux C., Savaglio S., Schady P., Vreeswijk P.M., 2013, A&A, 560, A88.
Fumagalli M., 2014, Memor. Soc. Astron. Ital., 85, 355
Fumagalli M., O'Meara J.M., Xavier Prochaska J., 2016, MNRAS, 455, 4100
Fynbo J.P.U., Prochaska J.X., Sommer-Larsen J., Dessauges-Zavadsky M., Møller P., 2008, Proc. Int. Astron. Un., 4, 41
Greiner J. et al., 2011, A&A, 526, A30
Heintz K.E. et al., 2019, A&A, 621, 1
Kalberla P.M.W., Burton W.B., Hartmann D., Arnal E.M., Bajaja E., Morras R., Pöppel W.G.L., 2005, A&A, 440, 775
Krongold Y., Prochaska J.X., 2013, Astrophys. J., 774, 115
Krühler T. et al., 2015, A&A, 581, A125
Ledoux C., Vreeswijk P.M., Smette A., Fox A.J., Petitjean P., Ellison S., Fynbo J., Savaglio S., 2009, A&A, 506, 661
Maiolino R., Mannucci F., 2019, A&A, 27, 3
Morrison R., Mccammon D., 1983, Astrophys. J., 270, 119
Perley D.A. et al., 2016, Astrophys. J., 817, 8
Rafelski M., Neeleman M., Fumagalli M., Wolfe A.M., Prochaska J.X., 2014, Astrophys. J., 782, L29
Rahin R., Behar E., 2019, Astrophys. J., 885, 47
Savaglio S., 2006, New J. Phys., 8, 195
Schady P., 2017, R. Soc. Open Sci., 4, 170304 (S17)
Schady P., Savaglio S., Krühler T., Greiner J., Rau A., 2011, A&A, 525, A113 (S11)
Shull J.M., Danforth C.W., 2018, Astrophys. J., 852, L11
Starling R.L.C. et al., 2013, MNRAS, 431, 3159 (S13)
Tanga M., Schady P., Gatto A., Greiner J., Krause M.G.H., Diehl R., Savaglio S., Walch S., 2016, A&A, 595, A24
Tanvir N.R. et al., 2019, MNRAS, 483, 5380
Watson D., Hjorth J., Fynbo J.P.U., Jakobsson P., Foley S., Sollerman J., Wijers R.A.M.J., 2007, Astrophys. J., 660, L101
Watson D., Jakobsson P., 2012, Astrophys. J., 754
Watson D. et al., 2013, Astrophys. J., 768, 23
Willingale R., Starling R.L.C., Beardmore A.P., Tanvir N.R., O'brien P.T., 2013, MNRAS, 431, 394
Wilms J., Allen A., McCray R., 2000, Astrophys. J., 542, 914 (W00)
Zafar T. et al., 2018, MNRAS, 479, 1542

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table S1 Full On-line version

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A

A1 S/N bias review

As we are mainly using *Swift* detected GRB where a redshift is available, we wished to examine if an S/N limited sample would cause bias. As a proxy for S/N, we plotted the log(N_{HX}) versus both log of total error in N_{HX} and total error/N_{HX} for all detections in Fig. A1. The left-hand panel plots the X-ray equivalent hydrogen column density against the ratio of the total error/N_{HX} where the total error is the 90 per cent confidence range of the N_{HX} fit. The scatter appears random so any cut-off by total error/N_{HX} should not result in a bias in N_{HX}. The right-hand panel plots distribution of N_{HX} against the total error for the detection. There a clear strong correlation so any cut-off for an S/N limited sample based on total error could produce a bias towards low N_{HX} GRB values. As a test of the impact of using a flux limited sample, we restricted a sub-sample to total error/N_{HX} < 1 resulting in 163 GRB. This sub-sample had essentially the same properties as the GRB detection sample using the Pearson and Spearman correlation coefficients and the N_{HX} versus redshift trend. Based on these results, we chose to use our full samples and not limit by a minimum flux.

A2 N_{HX} and $N_{H\,I}$ correlation review

In Fig. A2, we plot N_{HX} and $N_{H\,I,\,IC}$ for the full sub-sample of 128 GRB with both N_{HX} and $N_{H\,I}$ data.

No strong correlation is detected using a null hypothesis test, with both Pearson and Spearman coefficients being 0.10. It can be argued that this result strengthens the case that it is the IGM that is causing the N_{HX} redshift relation and is not intrinsic to the GRB host.

A3 Power law for metallicity redshift scaled adjustment to N_{HX}

To examine the impact on GRB N_{HX} fits in XSPEC using a more realistic mean GRB host metallicity than solar, and the variation with redshift, we used an XSPEC model *tbabs∗tbvarabs∗po* for GRB151027A (a high S/N GRB). We varied the redshifts between 0 and 10 and used metallicity = 0.07 Z/Z_{sol} in fitting N_{HX}. Fig. A3 shows a clear increasing metallicity adjusted log(N_{HX}) with redshift (blue line for Z = 0.07). In order to determine this relationship, and see whether it is consistent for different GRB X-ray spectra, we plotted in Fig. A4 the fractional increase in N_{HX} with redshift for a test sample of three high S/N GRB spectra with differing reported redshifts and N_{HX}. The fractional increase in N_{HX} with redshift is very similar for the three GRB. The power-law relation from a best-fitting least-squares for GRB151027A is $\log(N_{HX}) = (0.59 \pm 0.04)\log(1+z) + (0.18 \pm 0.02)$.

We note that in Fig. A4, the power-law curve is higher at $\log(1+z) < 0.2$ and $\log(1+z) > 0.8$. As a result, using this relation to adjust N_{HX} will result in marginally higher values than actual at low and high redshift but is a reasonable approximation for our purposes.

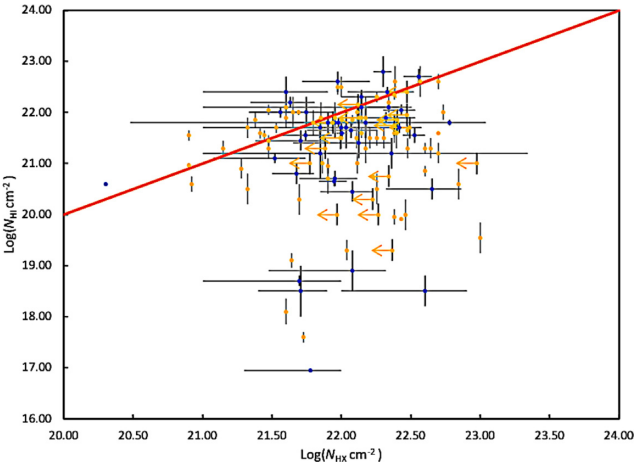

Figure A2. Plot of the log of the column densities for 128 GRB with both N_{HX} and $N_{H\,I}$ data from Tanvir et al. (2019) with ionization corrections. The blue dots are GRB for which the N_{HX} are detections. The orange dots are GRB best fits per Swift but where the 90 per cent confidence includes zero, and orange dots with arrows where the best-fitting N_{HX} = zero. The line shows where $N_{H\,I,IC}$ is equal to N_{HX}. There is no correlation.

A.4 $N_{H\,I}$ weighted average metallicity of GRB hosts (see Fig. 4)

The sample of metallicity measurements can be used to investigate the cosmic metallicity at different redshifts. A method to do this is to weight the average metallicity over a specific redshift bin with the total neutral hydrogen column density in the same redshift interval (C15). We used redshifts bins of $\Delta z = 1$ except for redshift 4 to 6 as there was only one GRB data point for metallicity greater than redshift 5. This weighted sample gives a marginally stronger evolution for metallicity than for the individual GRB.

$$[X/H]_{NHIweighted} = -0.65 \pm 0.07 - (0.15 \pm 0.06)_z \quad \text{(A1)}$$

Despite this possible mild evolution, in the high redshift range z > 4, all the GRB are metal enriched from 0.02 to 0.17 Z/Z_{sol},

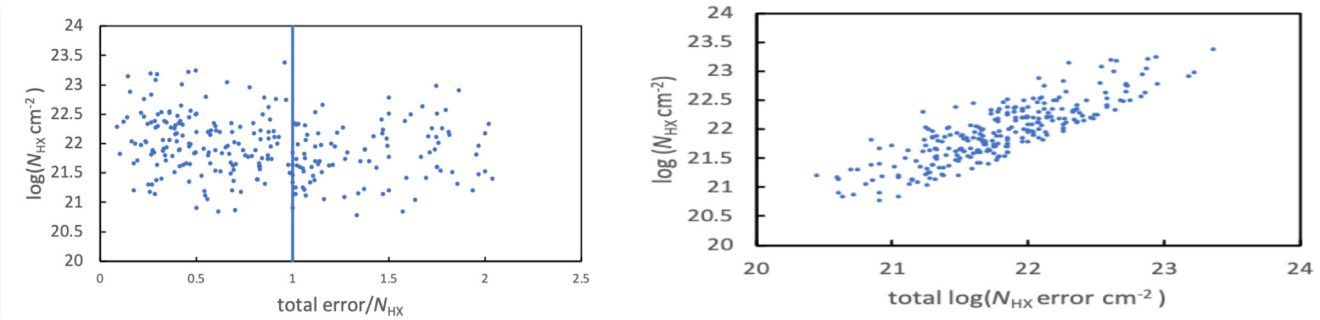

Figure A1. The left-hand panel shows the X-ray equivalent hydrogen column density against the ratio of the total error/N_{HX}. The blue vertical line is where total error/N_{HX} is 1. The scatter appears random. The right-hand panel shows the distribution of N_{HX} against the total error for the detection and shows a strong correlation. An S/N limited sample based on total error would introduce a bias to low N_{HX} GRB values.

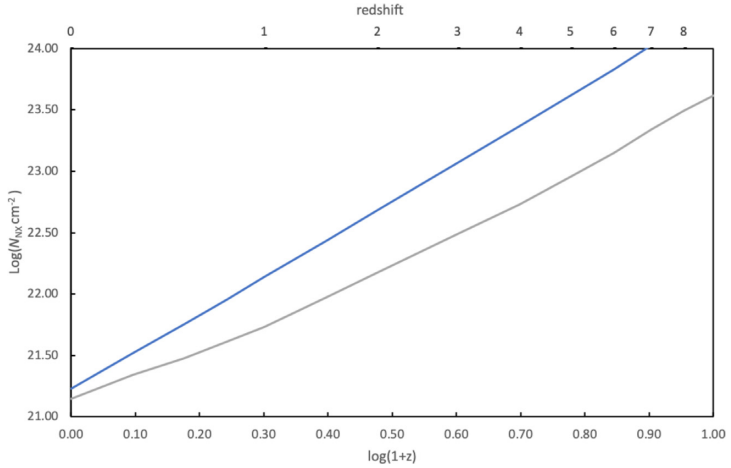

Figure A3. Impact on fitted N_{HX} for GRB151027A varying the host redshift between 0 and 10 and using metallicities of 0.07 Z/Z_{sol} (mean from our sample in Section 3.4), and solar. The blue line is with $Z = 0.07$ and the grey $Z = 1$. A lower metallicity results in increasing the fitted N_{HX} with the increased varying with redshift.

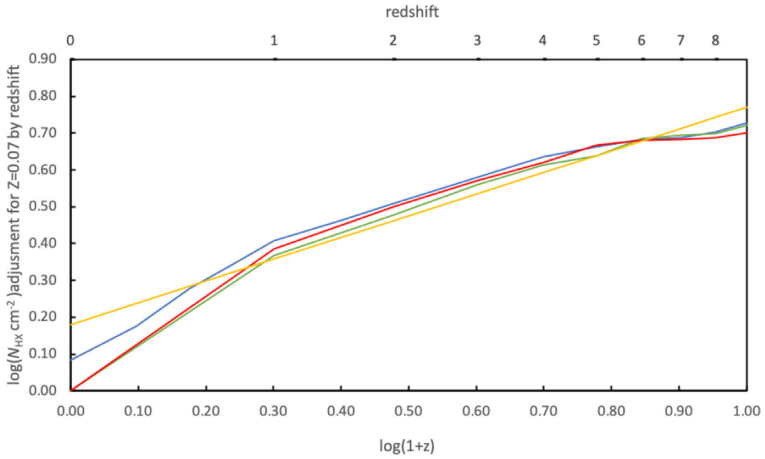

Figure A4. Comparison of the fractional increase in N_{HX} with redshift for three high S/N GRB, blue is GRB151027A, red is GRB150403A, and green is GRB120909A. The power-law relation yellow line is from a least-squares best fit for GRB151027A, $\log(N_{HX}) = (0.59 \pm 0.04)\log(1 + z) + (0.18 \pm 0.02)$.

suggesting that substantial amounts of metals were already present in galaxies the early Universe. The lack of clear redshift evolution is in contrast with quasars (QSO) which show strong evolution (Rafelski et al. 2014). A possible explanation for the lack of evolution is that GRBs may be located in different environments to quasars (Fynbo et al. 2008; C15). However, while this would affect the emission line metallicity, it should impact less on the absorption metallicity which is tracing the average galaxy LOS (Arabsalmani et al. 2018). We need far more GRB at high redshift to increase the statistical significance of metallicity evolution. For this paper, we will not use the $N_{H\,I}$ weighted average values as we wish to establish a metallicity to be applied to each GRB for the N_{HX} fitting.

This paper has been typeset from a Microsoft Word file prepared by the author.