



# PROBING X-RAY ABSORPTION AND OPTICAL EXTINCTION IN THE INTERSTELLAR MEDIUM USING *CHANDRA* OBSERVATIONS OF SUPERNOVA REMNANTS

DILLON R. FOIGHT<sup>1</sup>, TOLGA GÜVER<sup>2</sup>, FERYAL ÖZEL<sup>3</sup>, AND PATRICK O. SLANE<sup>1</sup>

<sup>1</sup>Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138, USA

<sup>2</sup>Istanbul University, Science Faculty, Department of Astronomy and Space Sciences, Beyazıt, 34119, Istanbul, Turkey

<sup>3</sup>Department of Astronomy, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

Received 2015 April 27; revised 2016 April 27; accepted 2016 May 17; published 2016 July 22

## ABSTRACT

We present a comprehensive study of interstellar X-ray extinction using the extensive *Chandra* supernova remnant (SNR) archive and use our results to refine the empirical relation between the hydrogen column density and optical extinction. In our analysis, we make use of the large, uniform data sample to assess various systematic uncertainties in the measurement of the interstellar X-ray absorption. Specifically, we address systematic uncertainties that originate from (i) the emission models used to fit SNR spectra; (ii) the spatial variations within individual remnants; (iii) the physical conditions of the remnant such as composition, temperature, and non-equilibrium regions; and (iv) the model used for the absorption of X-rays in the interstellar medium. Using a Bayesian framework to quantify these systematic uncertainties, and combining the resulting hydrogen column density measurements with the measurements of optical extinction toward the same remnants, we find the empirical relation  $N_{\text{H}} = (2.87 \pm 0.12) \times 10^{21} A_{\text{V}} \text{ cm}^{-2}$ , which is significantly higher than the previous measurements.

*Key words:* dust, extinction – ISM: supernova remnants – X-rays: ISM

## 1. INTRODUCTION

The linear relationship between optical extinction ( $A_{\text{V}}$ ) and hydrogen column density ( $N_{\text{H}}$ ) has long been observed and utilized to estimate the X-ray or optical brightness for new sources or to fit the broadband spectrum of X-ray sources. It is also used to obtain distance estimates for X-ray sources using their measured column densities (see, e.g., Durant & van Kerkwijk 2006; Güver et al. 2010; Ratti et al. 2010; Nielsen et al. 2012; Soria et al. 2012).

The photoelectric absorption by interstellar material causes rapid attenuation of soft X-rays in the spectra of Galactic sources. Measuring the extent of this attenuation yields information about the total column density along the line of sight to the source. Although this is commonly expressed in terms of the equivalent hydrogen column density  $N_{\text{H}}$  in the soft X-ray band (0.1–10 keV), it is predominantly caused by abundant heavier elements such as O, Ne, Si, Mg, and Fe. Optical extinction is caused by grains of the same elements. Because dust tends to follow the metal distribution in the interstellar medium (ISM), when averaging over many different lines of sight and over distances that are larger than the clumping scale of the ISM, it is reasonable to assume an approximate linear relationship between  $A_{\text{V}}$  and  $N_{\text{H}}$ .

There have been a number of different studies that have sought to accurately determine the relation between optical extinction and the hydrogen column density empirically. The methods vary across these studies, and the results show discrepancies greater than the statistical errors for each. Reina & Tarengi (1973) used X-ray binaries and extended sources and found a relation of  $N_{\text{H}} = 1.85 \times 10^{21} \times A_{\text{V}} \text{ cm}^{-2}$  (hereafter  $N_{\text{H}}$  is in units of  $\text{cm}^{-2}$  and  $A_{\text{V}}$  is in magnitudes); Gorenstein (1975) used supernova remnants (SNRs) to find  $N_{\text{H}} = (2.22 \pm 0.14) \times 10^{21} \times A_{\text{V}}$ ; while Predehl & Schmitt (1995) used a combination of *ROSAT* point sources and SNRs and measured  $N_{\text{H}} = (1.79 \pm 0.03) \times 10^{21} \times A_{\text{V}}$ . Recently, Güver & Özel (2009) collected a sample of 22 SNRs, for which the hydrogen column density and optical extinction were previously

measured, and found  $N_{\text{H}} = (2.21 \pm 0.09) \times 10^{21} \times A_{\text{V}}$ . This had the advantage of using the high-quality data from modern X-ray telescopes such as *Chandra* and *XMM-Newton* and focusing on sources with little to no intrinsic absorption,<sup>4</sup> but had the disadvantage of being unable to account for systematic errors that may vary for each published value. For example, when using values from the literature, there is no way to account for the variety of choices that are made during the data processing pipeline. The selected regions or the particular emission models may be interesting for the objectives of a particular study, but less ideal for the determination of the hydrogen column density and the comparison to the optical extinction measurements to determine the slope of the relation.

In this study, we take advantage of the wealth of SNR data available in the *Chandra* archive to investigate and quantify the systematic errors present in the determination of the hydrogen column density from the analysis of the X-ray data. We analyze all of the observations from the *Chandra* archive using standardized procedures so that we can quantify the systematic errors on each measurement. Using only observations from the *Chandra* archive also ensures a completely uniform data set: all observations were performed using the ACIS detector, and each had spectra generated and analyzed using the same treatment for background subtraction and model fitting routines using the spectral analysis software *xspec* (version 12.8.1; Arnaud 1996, with NEIVERS 1.1). This consistent treatment of a uniform data set gives us the opportunity to quantify the existing systematic errors in a way that had not been possible before.

The uniformity of our data set also allows us to explore the uncertainties associated with fitting spectra with a variety of models. SNRs can be home to a wide range of plasma properties due to the wide range of ages, host environments,

<sup>4</sup> To illustrate this, consider a remnant with a 15 pc radius that has swept up  $100 M_{\text{Sun}}$  of ISM material (which is a large estimate for most SNRs). The column density through the remnant is approximately  $2.6 \times 10^{19} \text{ cm}^{-2}$ , which is at least a factor of 10 smaller than the  $N_{\text{H}}$  values presented in this paper.

and composition of the sampled gas. In general, however, SNRs exhibit continuum emission driven by thermal bremsstrahlung accompanied by emission lines produced by ejecta material or swept up ISM. In the 0.5–5.0 keV range, where we predominantly perform the spectral fits, these features manifest most visibly in magnesium, silicon, and sulfur lines (though features due to iron, oxygen, and neon are also possible within this range). Non-thermal features can also be present due to particle acceleration in SNR shocks. Due to the wide range in spectral features and plasma conditions, there also exist a range of models that can be used to fit X-ray spectra, which will be detailed in Section 2.

Making use of the high spatial resolution and large number of counts in the *Chandra* archival data, we also investigate the systematic uncertainties associated with differing lines of sight toward larger remnants. We assume that all absorption is due to interstellar gas and dust along a line of sight, with no intrinsic absorption from the optically thin remnant. This assumption allows there to be differing values for  $N_{\text{H}}$  for different regions in a given SNR due to real differences in gas along a line of sight, which is supported by deep observations of large remnants such as W49B (Sasaki et al. 2013) and CTB 109 (also known as SNR G109.1–1.0; Keohane et al. 2007), for which many regions can be fit. When the archival observations have sufficient duration to have multiple high-count regions, we are able to fit multiple regions to understand how the value of  $N_{\text{H}}$  varies over the remnant. By fitting multiple regions with a range of models, we can determine the magnitude of the total systematic error on  $N_{\text{H}}$  measurements in the direction of the SNR. We note that any dust that is intrinsic to the SNR contributes negligible extinction. The integrated column density of swept up dust is small, and that of dust condensed from the stellar ejecta is smaller still. Moreover, SNR shocks are very efficient dust destroyers (Temim et al. 2015), making any contribution to the overall extinction exceedingly small.

Finally, and possibly most importantly, many of the previous analyses of *Chandra* and other SNR data used the solar abundances of Anders & Grevesse (1989) derived from observations of the Sun and meteorites. Wilms et al. (2000) showed that the abundances in the ISM can differ from solar values, which can affect the fit value of  $N_{\text{H}}$ . It is also important to consistently use an ISM absorption model that features improved calculations of the absorption cross sections when seeking to improve systematic errors in the hydrogen column density determination. We use the *wilms* abundance table and *tbabs* absorption model within *xspec* in order to accurately model the interstellar absorption to measure the hydrogen column density. The *tbabs* model in *xspec* adopts the cross sections from Verner & Yakovlev (1995). This standard that we adopt for the entire sample leads to larger ( $\sim 30\%$ ) column densities than those reported in the earlier studies.

In Section 2, we discuss the variety of models used to fit our data, as well as the results of simulated data sets. Sections 3 and 4 detail the processes we used to process the data, as well as our approach to identifying systematic errors and determining total uncertainties. Finally, in Section 5 we use our new data set to derive the slope of the  $N_{\text{H}} - A_{\text{V}}$  relation.

## 2. MODELS AND SIMULATED RESULTS

The models we used to fit each thermal region (*raymond*, *mekal*, *nei*, and *ps shock*) were chosen to span the range of complexity one might expect from a region of plasma in an

**Table 1**  
Recovered  $N_{\text{H}}$  Values from Simulated Data

Assumed $N_{\text{H}}$ ( $10^{22} \text{ cm}^{-2}$ )	$T_s$ (keV)	$T_{\text{es}}$ (keV)	$\tau_0$ ( $10^{11}$ )	Fit $N_{\text{H}}$ ( $10^{22} \text{ cm}^{-2}$ )	
0.5	0.3	0.15	10	$0.456 \pm 0.094$	
			5	$0.412 \pm 0.135$	
			1	$0.469 \pm 0.052$	
	0.6	0.3	10	$0.525 \pm 0.030$	
			5	$0.458 \pm 0.023$	
			1	$0.499 \pm 0.039$	
	0.9	0.45	10	$0.514 \pm 0.020$	
			5	$0.492 \pm 0.022$	
			1	$0.436 \pm 0.032$	
	1.0	0.3	0.15	10	$0.969 \pm 0.139$
				5	$0.922 \pm 0.132$
				1	$0.998 \pm 0.143$
0.6		0.3	10	$1.007 \pm 0.056$	
			5	$0.960 \pm 0.056$	
			1	$0.988 \pm 0.047$	
0.9		0.45	10	$0.952 \pm 0.051$	
			5	$0.973 \pm 0.048$	
			1	$0.942 \pm 0.061$	

SNR. The *raymond* model (Raymond & Smith 1977) is the simplest, modeling a hot, diffuse gas in collisional equilibrium. Similarly, the *mekal* model also calculates spectra of a plasma in collisional equilibrium, but with improved handling of the Fe-L complex (Kaastra et al. 1996). The *nei* and *ps shock* models (Borkowski et al. 2001) are more complicated, providing spectra for a non-equilibrium ionization (NEI) collisional plasma (*nei*), and constant temperature NEI plasma heated by a plane-parallel shock (*ps shock*). For non-thermal regions, the models *powerlaw* and/or *srcut* were used. Both models produce a spectrum we would expect from synchrotron emission from a power-law distribution of shocked electrons interacting with an SNR’s magnetic field, with an additional exponential cutoff of the distribution for the *srcut* model (Reynolds 1998; Reynolds & Keohane 1999).

As discussed previously, differing lines of sight can sample different paths through the ISM, so there is a chance for real variations in  $N_{\text{H}}$  measurements across an SNR. There can also be small systematic variations introduced by fitting the regions with the models discussed above. Each region may be composed of multiple plasma components, so fitting to a single plasma model may introduce a systematic error or a bias. To show this, we performed fits to simulated spectra produced with a range of  $N_{\text{H}}$ . We produced the simulated spectra using the *sedov* model in *xspec* (version 12.8.1 using NEIVERS 1.1), which models the total emission from an SNR undergoing adiabatic expansion (Borkowski et al. 2001). The complete set of simulated spectra cover a grid of inputs for shock temperature ( $T_s = 0.3\text{--}0.9$  keV), electron temperature ( $T_{\text{es}} = 0.15\text{--}0.45$  keV), ionization age ( $\tau_0 = 1 - 10 \times 10^{11} \text{ s cm}^{-3}$ ), and  $N_{\text{H}}$  ( $0.5\text{--}1.0 \times 10^{22} \text{ cm}^{-2}$ ). For each parameter set, we simulated 200 spectra and then binned these to have at least 50 counts per bin. We fit each spectrum with a *ps shock* model, using the *tbabs* absorption model with the *wilms* abundance table. We present the average of the best-fit values for each parameter set in Table 1, along with the standard deviation of all fit values.

The results in Table 1 show that while a single plasma model approximately recovers the assumed  $N_{\text{H}}$  value from a complex plasma, it can introduce significant statistical errors and

suggests a possible bias. In particular, the  $N_{\text{H}}$  values we measure from the simulated data are close to but often lower than the assumed values. This, in fact, is not surprising: fitting a single-temperature model to a simulation that is meant to represent the entire remnant (i.e., a Sedov spectrum which integrates over a range of temperatures) cannot adequately describe the spectrum or yield a very accurate value of the hydrogen column density.

To minimize these effects in the analysis of the actual data, we take a three-pronged approach. First, we extract spectra from as small regions of the remnant as possible to avoid creating complicated, multi-temperature regions with multiple plasma components. Second, we perform fits with thermal, non-thermal, and mixed thermal/non-thermal models to capture the spectral characteristics of the regions correctly. Third, we use Bayesian statistical tools (discussed in Section 4.2) to combine measurements from different spatial regions as well as from different spectral model results to assess any systematic uncertainties in the  $N_{\text{H}}$  measurement for a given remnant. It is important to note, however, that the simulated spectra results on their own are not the sole motivation for the Bayesian analysis; if we believed that the simulated results were immediately comparable to the *Chandra* data results, then we could use the simulated results to establish the magnitude of a bias term. In Section 5, we will use the total uncertainties we determine for the  $N_{\text{H}}$  toward each remnant to more accurately constrain the  $N_{\text{H}}/A_{\text{V}}$  relationship.

### 3. CHANDRA DATA PIPELINE AND PROCESSING

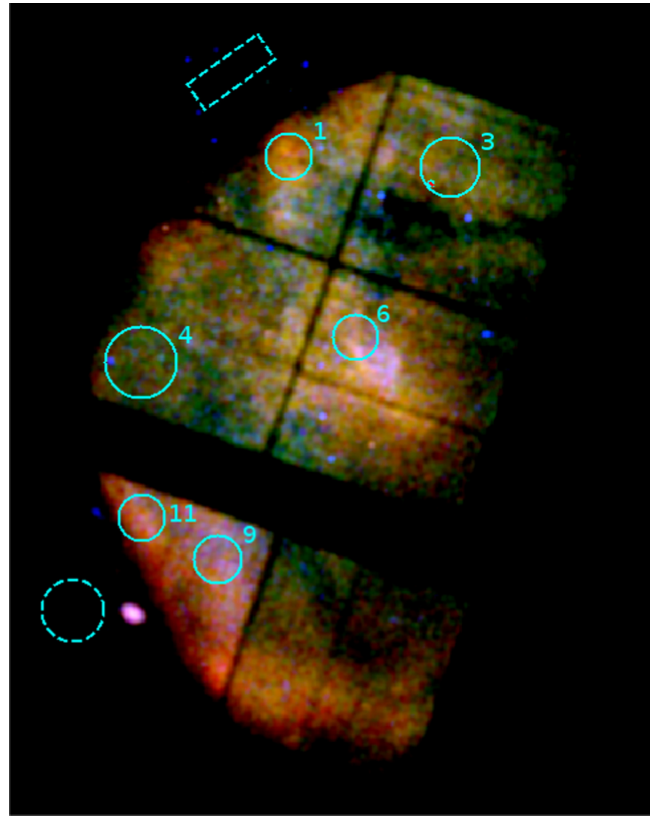
#### 3.1. Pre-processing

For each of the selected SNRs, we downloaded archival data from the *Chandra* archive. For remnants with multiple pointings and exposures, we chose one observation to eliminate any need to co-add multiple exposures, and placed preference on newer observations that were long enough to provide multiple regions with sufficient counts for spectral analysis. We ran the data through two pipeline scripts that were designed to automate the processing procedures. All scripts that utilized *CIAO* tools used version 4.5. Our first script reapplied the latest calibration and produced a new level = 1 event file. Because the majority of remnants were taken with exposure mode FAINT or VFAINT, we cleaned the ACIS background using procedures that were appropriate for these modes. Finally, we filtered the level = 1 file for bad grades and applied the good time intervals to generate the new level = 2 event file.

#### 3.2. Region Selection and Data Processing

We used the new level = 2 event file to select regions for spectral extraction. We selected regions using ds9 and chose areas that had  $\geq 10,000$  counts, as well as sampling a variety of lines of sight across the remnant. In some remnants, it was preferable or necessary to select regions with non-thermal emission. We processed these using procedures identical to the thermal regions, but fit them with non-thermal models. In general, regions were selected using the morphology of the remnant as a guide, such that the regions would contain one variety of plasma (the importance of which is highlighted in Section 2). We extracted a spectrum from each region and then grouped it to a minimum of 25 counts per bin.

As a typical example of the regions selected for a remnant, we show SNR G109.1–1.0 in Figure 1, with a number of the

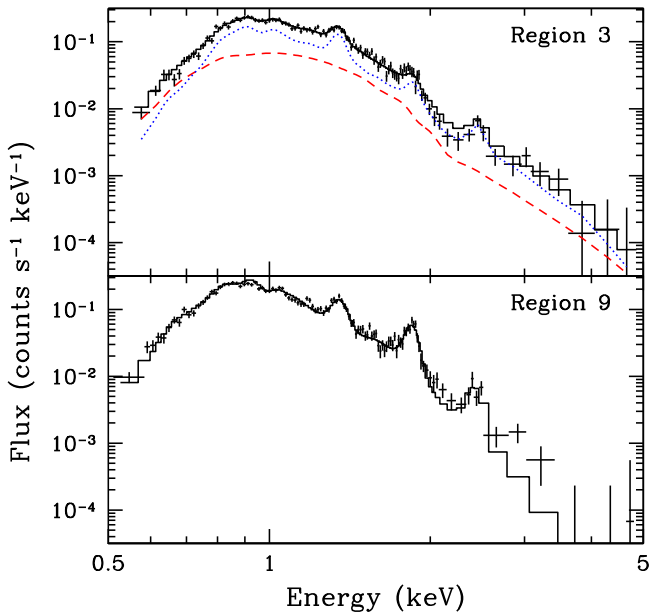


**Figure 1.** *Chandra* image of SNR G109.1–1.0 taken with the ACIS-I detector (data from OBSID 1901; red: 0.5–1.1 keV; green: 1.1–1.6 keV; blue: 1.6–5.0 keV). The numbered circles show 6 of the 11 regions that were selected for spectrum extraction from this remnant. The dashed rectangle and circle shows the regions selected for background subtraction.

selected regions marked with circles and the regions selected for local background shown with a dashed circle and rectangle. In the full analysis of G109.1–1.0, we used a total of 11 regions. However, for clarity and simplicity, we chose to highlight only six in this figure. We show the resultant binned spectra from two of these regions in Figure 2, again as two representative examples of the types of spectra we encountered in the analyses. Note that small contributions from model components (such as the power-law component in Region 3 of Figure 2) may appear insignificant, but are important for producing acceptable fits.

We fit a variety of models (discussed in Section 2) to each grouped spectrum. We fit each model with default ISM abundances, but it was often necessary to allow some of the plasma abundances to vary in order to properly fit regions with ejecta-enriched plasma. The number and variety of the free abundances were allowed to change from model to model for a given region, and only the plasma models' abundances were allowed to vary. For non-thermal regions, this process was the same, but simplified by the fact that there were no free abundance parameters possible for those models.

Despite our best efforts to select regions that would be composed of only one variety of plasma (and thus could be fit using one model), sometimes this was infeasible (for example, if the regions had to be large in order to contain an adequate number of counts). In such cases, we needed to add other components to the models. Most commonly, this required adding a non-thermal model to the thermal model to



**Figure 2.** Thermal spectra extracted from two different regions of SNR G109.1–1.0 (corresponding to Regions 3 and 9 shown in Figure 1) in the 0.5–5.0 keV range with best-fit model components overlaid. The Region 3 spectrum exhibits weak emission lines (from the thermal component shown in the dotted blue line) with some power-law component contribution (shown as the red dashed line), while the Region 9 spectrum has stronger lines and no power-law component (thermal model shown as solid black line).

compensate for regions near non-thermal filaments or a central pulsar wind nebula. However, in some cases we added Gaussian features to fit spectral lines that were either not properly fit or non-existent in the thermal model, or to account for ejecta mixing into a non-thermal region, causing emission lines to appear on top of a power-law spectrum. For the most part, the data were good enough that we could find regions that allowed for a single model to produce an acceptable fit, but when included, these additional components did not make statistically significant changes in the best-fit value of  $N_{\text{H}}$ .

We determined if a spectral fit was good both by visually checking the final fit, as well as by using the  $\chi^2/\nu$  fit statistic. We show in Figure 3 an example of an acceptable fit compared to a poor fit. We did not place a hard upper limit on the fit statistic because some large  $\chi^2/\nu$  values were dominated by a single feature that was not captured by the model but did not affect the inferred  $N_{\text{H}}$ . Nevertheless, nearly all of the fits used in our final analysis had  $\chi^2/\nu \leq 2$ .

## 4. RESULTS

### 4.1. Results of The Spectral Analysis

For each SNR region, we included all models that produced acceptable fits. As a typical example, we show in Table 2 the complete fit results for six of the regions of G109.1–1.0. It is evident from this table that *vnei* and *vpshock* models with thawed abundances were commonly the best models for remnants with thermal regions, with a *powerlaw* component added in when necessary. The *raymond* and *mekal* models ranged from fairly successful to very poor, which presumably reflects significant NEI contributions in the data. For SNRs with completely non-thermal regions, the *powerlaw* model was the default model with consistent success. (Two of the remnants, G120.1+1.4 and G04.5+6.8, were better fit by an

*srcut* model.) All errors presented on the fit  $N_{\text{H}}$  values in Table 2 are 90% confidence ranges. The complete fit results for all remnants used in our analysis have been released as a Zenodo data set (<http://zenodo.org/record/17183#.V22YjpMrKqA>).

### 4.2. Bayesian Analysis to Determine $N_{\text{H}}$ and Its Uncertainty

Each of the fits within a region have formal (and often asymmetric) uncertainties, as we showed in Table 2 for SNR G109.1–1.0. From these fits, the dispersion in the measurements arising from spatial sampling or from different continuum models can be calculated. For a number of remnants, this contribution is comparable to the formal uncertainties. In others, however, the difference between the various measurements of  $N_{\text{H}}$  is significantly larger than the formal uncertainties, pointing to systematic uncertainties originating from spatial sampling, from the choice of model, or both. For the purposes of our analysis, we do not distinguish between the variance of the ISM along different lines of sight and the systematic error introduced by the incomplete plasma models and we combine these different types of systematic error into a single term (see, e.g., Sinervo 2003). Our aim in this section is to determine the most likely  $N_{\text{H}}$  value for each remnant, as well as a measure of the combined formal and systematic uncertainties in this quantity. We accomplish this by finding the parameters of the underlying  $N_{\text{H}}$  distribution for each remnant that is consistent with our sample of measurements. We use this distribution to find the most likely value of the hydrogen column density and its uncertainty.

We start with a parametric form of the  $N_{\text{H}}$  distribution and use the set of measurements from the individually fitted regions in order to estimate its parameters. We take the assumed underlying distribution to be a Gaussian

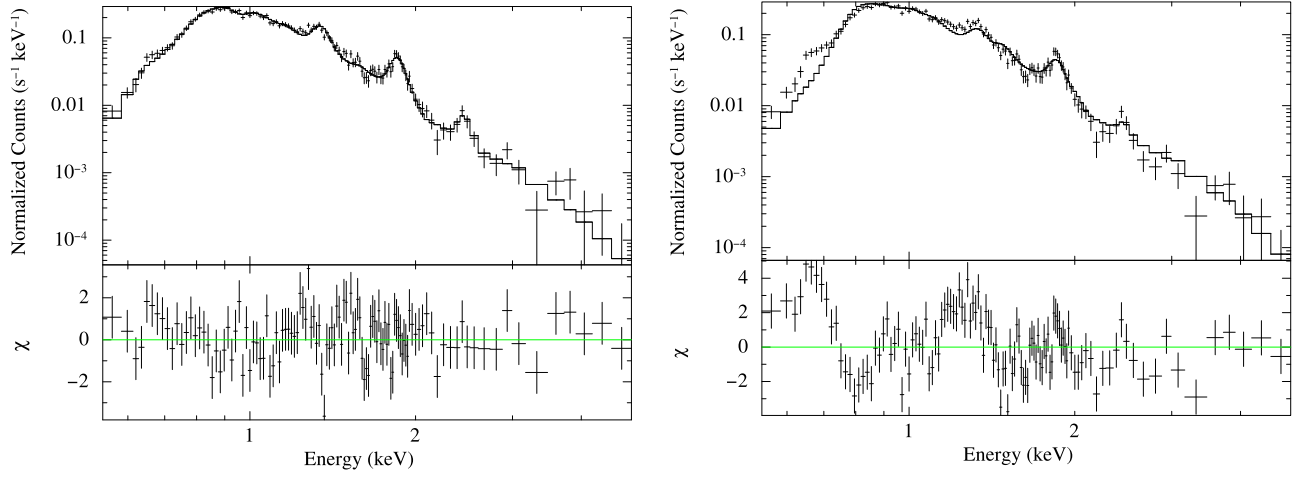
$$P(N_{\text{H}}; \sigma, N_{\text{Hc}}) = C \exp\left[-\frac{(N_{\text{H}} - N_{\text{Hc}})^2}{2\sigma^2}\right], \quad (1)$$

with a mean  $N_{\text{Hc}}$  and a standard deviation  $\sigma$  that can be different for each remnant. In this and the following expressions,  $C$  is a proper normalization constant such that

$$\int_0^{\infty} P(N_{\text{H}}; \sigma, N_{\text{Hc}}) dN_{\text{H}} = 1. \quad (2)$$

We also need to model the individual measurement uncertainties  $P_i(N_{\text{H}})$ , where  $i$  represents a particular SNR region/model combination that yields a single measurement of  $N_{\text{H}}$  for that region. In general, the  $\chi^2$  surface for each measurement is not simple or symmetric around the minimum, leading to asymmetric formal errors in these individual measurements. However, around the minimum, it is accurate to represent the likelihood using two half Gaussians with different dispersions that smoothly connect at the most likely column density  $N_{\text{H}0,i}$  for each measurement; i.e.,

$$P_i(\text{data}|N_{\text{H}}) = \begin{cases} C_i \exp\left[-\frac{(N_{\text{H}} - N_{\text{H}0,i})^2}{2\sigma_{-,NH,i}^2}\right], & N_{\text{H}} < N_{\text{H}0,i} \\ C_i \exp\left[-\frac{(N_{\text{H}} - N_{\text{H}0,i})^2}{2\sigma_{+,NH,i}^2}\right], & N_{\text{H}} > N_{\text{H}0,i} \end{cases}, \quad (3)$$



**Figure 3.** Two example model fits to a spectrum of SNR G109.1–1.0 extracted from Region 6. The acceptable fit (left panel) is a *vnei* model, whereas the unacceptable fit (right panel) is a *vraymond* model in this particular example. Each model has Mg, Si, and S abundances as free fit parameters.

**Table 2**  
Example Fit Values for SNR G109.1–0.0

Region	Model	Thawed Elements	Fit $N_{\text{H}}$ ( $10^{22} \text{ cm}^2$ )	$\chi^2/\nu$	
1 (NE)	<i>vnei</i> + powerlaw	Mg, Si	$1.07^{+0.07}_{-0.07}$	146.94/97	
	<i>vpshock</i> + powerlaw	Mg, Si, S	$1.05^{+0.08}_{-0.07}$	143.81/96	
	<i>nei</i> + powerlaw	None	$0.95^{+0.09}_{-0.09}$	132.71/111	
3 (N)	<i>vnei</i> + powerlaw	Mg, Si	$0.93^{+0.10}_{-0.16}$	130.55/109	
	<i>pshock</i> + powerlaw	None	$0.95^{+0.09}_{-0.04}$	131.83/111	
	<i>vpshock</i> + powerlaw	Mg, Si	$0.89^{+0.09}_{-0.06}$	126.88/109	
	4 (E)	<i>vnei</i> + powerlaw	Mg, Si	$0.90^{+0.07}_{-0.07}$	160.49/124
		<i>vpshock</i> + powerlaw	Mg, Si	$0.93^{+0.08}_{-0.07}$	153.14/124
6 (Center)		<i>vnei</i>	Mg, Si, S	$0.86^{+0.08}_{-0.07}$	138.73/110
	<i>vpshock</i>	Mg, Si, S	$0.71^{+0.08}_{-0.05}$	130.71/110	
9 (SE Inner)	<i>vnei</i>	Mg, Si, S	$1.24^{+0.06}_{-0.07}$	139.21/117	
	<i>vpshock</i>	Mg, Si, S	$1.23^{+0.05}_{-0.05}$	137.94/117	
11 (SE Outer)	<i>nei</i>	None	$0.87^{+0.06}_{-0.06}$	142.52/111	
	<i>vnei</i>	Mg, Si	$0.88^{+0.08}_{-0.07}$	142.15/109	
	<i>pshock</i>	None	$0.88^{+0.08}_{-0.06}$	130.27/111	
	<i>vpshock</i>	Mg, Si	$0.89^{+0.08}_{-0.08}$	129.88/109	

where “data” stands for the most likely column density  $N_{\text{H}0,i}$  and the two associated uncertainties  $\sigma_{-,N_{\text{H},i}}$  and  $\sigma_{+,N_{\text{H},i}}$  for the  $i$ th SNR region/model combination.

We want to calculate the quantity  $P(\sigma, N_{\text{Hc}}|\text{data})$ , which measures the posterior likelihood of the parameters of the  $N_{\text{H}}$  distribution, given the observations. Using Bayes’ theorem, we can write this as

$$P(\sigma, N_{\text{Hc}}|\text{data}) = C_2 P(\text{data}|\sigma, N_{\text{Hc}}) P(\sigma) P(N_{\text{Hc}}), \quad (4)$$

where  $C_2$  is a normalization constant and  $P(\sigma)$  and  $P(N_{\text{Hc}})$  are the priors over the values of the Gaussian dispersion,  $\sigma$ , and the peak of the  $N_{\text{H}}$  distribution,  $N_{\text{Hc}}$ . Here, “data” stands for the

ensemble of the  $N_{\text{H}0,i}$ ,  $\sigma_{-,N_{\text{H},i}}$  and  $\sigma_{+,N_{\text{H},i}}$  values for a particular SNR. We take a flat prior over the Gaussian dispersion  $\sigma$  between  $\sigma_{\text{min}}$  that is equal to 0.1 times the smallest formal uncertainty obtained from a spectral fit for each remnant and  $\sigma_{\text{max}}$  that is equal to 10 times the largest difference between two  $N_{\text{H}}$  measurements for each remnant:

$$P(\sigma) = \begin{cases} 0, & \sigma \leq \sigma_{\text{min}} \\ \frac{1}{\sigma_{\text{max}} - \sigma_{\text{min}}}, & \sigma_{\text{min}} < \sigma < \sigma_{\text{max}} \\ 0, & \sigma > \sigma_{\text{max}}. \end{cases} \quad (5)$$

Similarly, we take a flat prior over the centroid of the  $N_{\text{H}}$  distribution  $N_{\text{Hc}}$  that spans the range from 0.1 times the smallest  $N_{\text{H}}$  measurement to 10 times the largest  $N_{\text{H}}$  measurement per source. These limits ensure that the particular minimum and maximum values of the prior distribution do not affect the results.

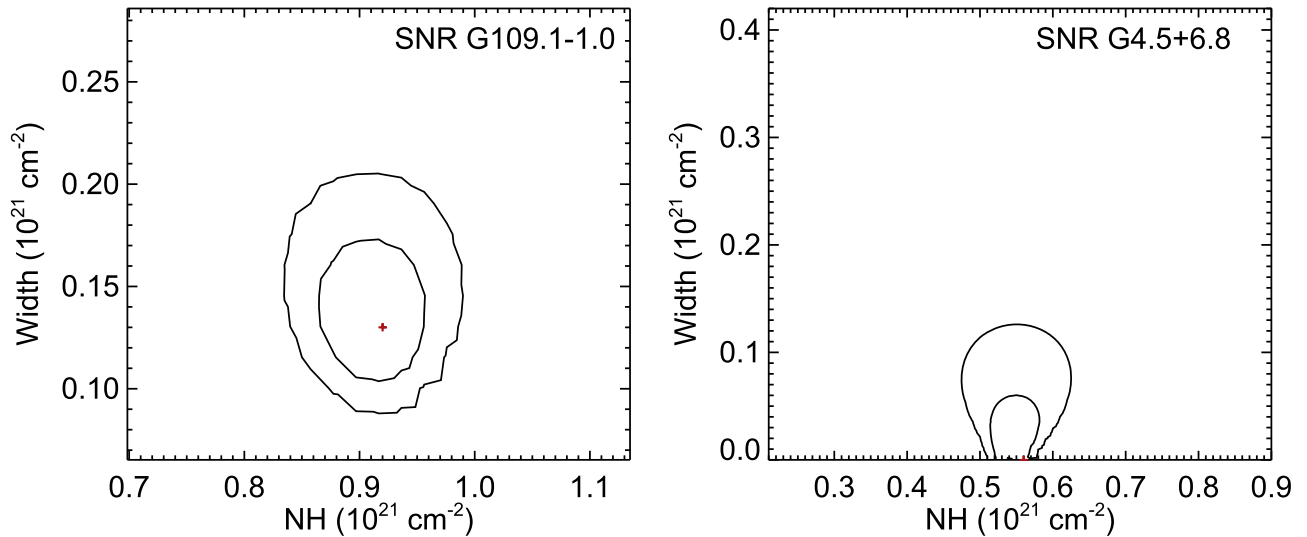
In Equation (4), the quantity  $P(\text{data}|\sigma, N_{\text{Hc}})$  measures the likelihood that we will make a particular set of measurements for the column density given the values of the parameters of the column density distribution. We need to estimate this quantity, given the measurement likelihoods given in Equation (3). We will assume that each measurement is independent, so that

$$P(\text{data}|\sigma, N_{\text{Hc}}) = \prod_i \int dN_{\text{H}} P_i(\text{data}|N_{\text{H}}) P(N_{\text{H}}; \sigma, N_{\text{Hc}}). \quad (6)$$

Combining this last equation with Equation (4), we obtain the posterior likelihood

$$P(\sigma, N_{\text{Hc}}|\text{data}) = C P(\sigma) P(N_{\text{Hc}}) \times \prod_i \int dN_{\text{H}} P_i(\text{data}|N_{\text{H}}) P(N_{\text{H}}; \sigma, N_{\text{Hc}}), \quad (7)$$

where  $C$  is the overall normalization constant. We can use Equation (7) to determine the parameters  $\sigma$  and  $N_{\text{Hc}}$  given the individual region/model fits for each remnant. The dispersion ( $\sigma$ ) is then a measure of the systematic uncertainty associated with the remnant. If a remnant has a significant systematic uncertainty, that contribution is considered when we define the most likely value of the  $N_{\text{H}}$  for the SNR and its uncertainty.



**Figure 4.**  $p = 0.32$  and  $p = 0.05$  Bayesian credible regions for the central value and the dispersion of an underlying Gaussian distribution of  $N_{\text{H}}$  measurements for SNRs G109.1–1.0 (left) and G04.5+6.8 (right). The dispersion of the distribution reflects the systematic uncertainty in the measurements. The dispersion contours for G04.5+6.8 are consistent with 0.0, which indicates that systematic uncertainties are not significant. For G109.1–1.0, the most likely dispersion is not consistent with zero, indicating a level of systematic error comparable to the formal errors. These are included in the uncertainties in Table 3 following the discussion in Section 4.

In Figure 4, we show as examples the Bayesian credible regions over the parameters of the  $N_{\text{H}}$  distribution that correspond to the measurements for G109.1–1.0 and G4.5+6.8. We show with a dot the most likely values of the centroid and dispersion; i.e., the peak of the posterior likelihood given by Equation (7). In the case of G4.5+6.8, the best-fit value of the dispersion is very close to zero and this parameter is consistent with being zero within the  $p = 0.05$  credible region, indicating a negligible level of systematic uncertainty. On the other hand, for remnants such as G109.1–1.0, the best-fit value of the dispersion is not consistent with being zero, indicating the presence of systematic uncertainties arising from spatial sampling across the remnant or the differences among spectral models.

The previous Bayesian analysis to quantify the systematic error term yields two categories of remnants. In the first category the systematic dispersion ( $\sigma$ ) is consistent with zero, so we assign a zero systematic uncertainty, i.e., set  $P(N_{\text{H}}; \sigma, N_{\text{Hc}}) = \delta(N_{\text{H}} - N_{\text{Hc}})$  in Equation (6). This allows us to find a properly weighted average of the  $N_{\text{H}}$  measurements using their formal uncertainties and report the best-fit value and its formal uncertainty in Table 3. In the second category of remnants, we compute the posterior likelihood over their  $N_{\text{H}}$  by weighing each Gaussian distribution with a given centroid  $N_{\text{Hc}}$  and dispersion  $\sigma$  with the likelihood calculated in Equation (7) that those pair of parameters represent the observed  $N_{\text{H}}$  values:

$$P(N_{\text{H}}) = \int \int P(N_{\text{H}}; \sigma, N_{\text{Hc}}) P(\sigma, N_{\text{Hc}} | \text{data}) dN_{\text{Hc}} d\sigma. \quad (8)$$

We use this distribution to infer the best-fit values of  $N_{\text{H}}$  and its uncertainty for each source and report this in Table 3.

## 5. APPLICATION TO THE $N_{\text{H}}/A_{\text{V}}$ RELATION

We presented in the previous section the hydrogen column densities toward a large sample of SNRs that we measured using the *Chandra* data archive. We now combine these  $N_{\text{H}}$  measurements with the optical extinction measurements presented in Güver & Özel (2009) for our sample of SNRs to

**Table 3**  
Measured Column Densities for All SNRs

SNR	Number of Regions/Models	$N_{\text{H}}$ ( $10^{22} \text{ cm}^{-2}$ )	Error
G0.0+0.0	3	9.86	0.34
G04.5+6.8	4	0.54	0.011
G06.4–0.1	22	0.63	0.19
G53.6–2.2	2	0.67	0.03
G54.1+0.3	1	2.55	0.04
G69.0+2.7	1	0.45	0.02
G109.1–1.0	30	0.90	0.14
G111.7–2.1	8	1.17	0.16
G116.9+0.2	2	0.92	0.07
G119.5+10.2	1	0.38	0.11
G120.1+1.4	14	0.80	0.10
G130.7+3.1	6	0.54	0.01
G184.6–5.8	1	0.30	0.02
G260.4–3.4	3	0.35	0.15
G263.2–3.3	1	0.03	0.01
G327.6+14.6	2	0.16	0.01
G332.4–0.4	9	0.87	0.35

determine a relation between these two quantities that accounts for systematic errors. The optical extinction toward the remnants in our sample was determined through several different methods, which we list in Table 4. Most of these methods involve measuring the intensity ratio of two emission lines, for which the intrinsic ratio is known. Comparing the observed intensity ratio to the intrinsic ratio yields a reddening, which is then converted into a measurement of the optical extinction. These methods are challenging, in general, because a high signal-to-noise ratio spectrum is necessary to obtain the intensity ratios.

Several emission line pairs are frequently used for this purpose. The  $\text{H}_{\alpha}$  (6563 Å) to  $\text{H}_{\beta}$  (4861 Å) line ratio, referred to as the Balmer decrement method, is one of the most well known and reliable ones among these pairs. Other emission line ratios from the SII multiplet (Miller 1968) and FE[II] IR transitions (Oliva et al. 1989) are also used in a similar way.

**Table 4**  
 $A_V$  Values from Güver & Özel (2009)

SNR	$A_V$ (mag)	Error <sup>a</sup> (mag)	Method	References
G0.0+0.0	29	2	Nearby Stars	(1)
G04.5+6.8	2.5	0.9	FeII Ratio	(2)
G06.4−0.1	3.57	0.47	$H_\alpha/H_\beta$	(3)
G53.6−2.2	3.57	0.47	$H_\alpha/H_\beta$	(3)
G54.1+0.3	8.0	0.70	Nearby Stars	(4)
G69.0+2.7	2.48	-	$H_\alpha/H_\beta$	(5)
G109.1−1.0	3.15	0.65	$H_\alpha/H_\beta$	(6)
G111.7−2.1	5.0	0.40	SII ratio	(7)
G116.9+0.2	2.70	0.5	$H_\alpha/H_\beta$	(8)
G119.5+10.2	1.27	0.41	Extinction Map	(9)
G120.1+1.4	1.86	0.12	Nearby Stars	(10)
G130.7+3.1	2.11	-	Nearby Stars	(11)
G184.6−5.8	1.55	0.186	$Ly\alpha$ Absorption	(12)
G260.4−3.4	2.60	-	Nearby Stars	(13)
G263.2−3.3	0.38	-	$H_\alpha/H_\beta$	(14)
G327.6+14.6	0.34	-	HI/GC	(15)
G332.4−0.4	4.70	0.90	FeII Ratio	(2)

**Note.**

<sup>a</sup> Errors designated as “-” are taken as 15%.

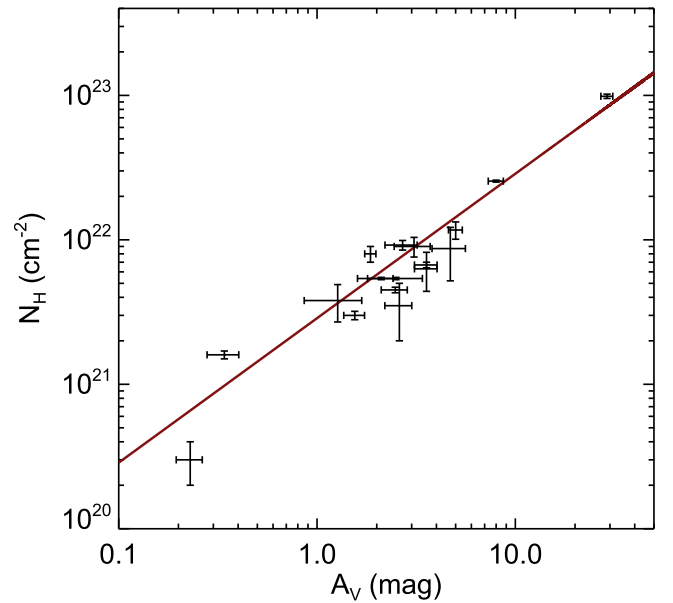
**References.** (1) Predehl & Truemper (1994), (2) Oliva et al. (1989), (3) Long et al. (1991), (4) Koo et al. (2008), (5) Hester & Kulkarni (1989), (6) Fesen & Hurford (1995), (7) Hurford & Fesen (1996), (8) Fesen et al. (1997), (9) Mavromatakis et al. (2000), (10) Ruiz-Lapuente (2004), (11) Fesen et al. (1988), (12) Sollerman et al. (2000), (13) Gorenstein (1975), (14) Manchester et al. (1978), (15) Raymond et al. (1995).

For all of these pairs of lines, the ratio depends only very weakly on the temperature and density of the emitting plasma, leading to minimal uncertainties in the calculation of the ratio (Osterbrock 1989; Lequeux 2005).

Another common method is to use stars near an SNR to estimate the extinction toward that remnant. This method relies on having an existing estimate of the distance to the remnant, and then identifying stars at similar distances. If suitable stars can be observed, the extinction measurements of those stars can be applied to the remnant (Ruiz-Lapuente 2004; Koo et al. 2008).

In Figure 5, we plot the hydrogen column density measurements and their uncertainties presented in Table 3 against the measurements of the optical extinction summarized in Table 4. Following Güver & Özel (2009), we assign a 15% error to the optical extinction measurements for which uncertainties have not been reported (denoted by “-” in Table 4). In addition, upon inspecting the uncertainties of individual measurements in Table 3, we note that a best-fit line between these two quantities will be heavily influenced by a small number of remnants where  $N_H$  measurements have very small (<5%) formal uncertainties. Even though the sample of regions and the range of models we considered in the spectral fits did not allow us to measure a systematic uncertainty for these remnants, considering the many possible sources of systematic uncertainty that are usually present in the determination of the hydrogen column densities, we assign a 5% error to the  $N_H$  measurement of SNRG130.7+3.1, SNRG69.0+2.7, SNRG004.5+06.8, SNRG53.6−2.2, SNRG54.1+0.3, and SNRG00.0+0.0 to avoid biasing the results.

Using these data, we obtain a best-fit linear relation between the hydrogen column density and the optical extinction that is



**Figure 5.**  $N_H$  and  $A_V$  measurements for the SNRs in the *Chandra* archive along with the best-fit linear model, which gives  $N_H = 2.87 \pm 0.12 \times 10^{21} A_V$ . The  $N_H$  measurements are presented in Table 3 and have a minimum error of 5%, while the  $A_V$  measurements (listed in Table 4) are taken from Güver & Özel (2009).

described by

$$N_H = 2.87 \pm 0.12 \times 10^{21} A_V \text{ cm}^{-2}, \quad (9)$$

where  $A_V$  is in magnitudes and the error represents the  $1\sigma$  uncertainty. We show the best-fit line in Figure 5.

We can test the effect of imposing a floor to the  $N_H$  uncertainties by also fitting to the raw values. This increases the best-fit value of the  $N_H/A_V$  to  $(2.92 \pm 0.11) \times 10^{21} \text{ cm}^{-2}$ , which is within  $1\sigma$  of the best-fit value presented above, but dominated by remnants with small errors. Finally, we note that the remnants with either very low or high  $N_H$  will have the greatest leverage on the final fit. To test the magnitude of this effect, we remove the points with the highest and lowest values of the hydrogen column density. This results in a best-fit value of  $(2.76 \pm 0.13) \times 10^{21} \text{ cm}^{-2}$ , again within  $1\sigma$  of the best-fit value using all of the data and a minimum error imposed on the  $N_H$  measurements.

These results indicate that a consistently analyzed sample of SNRs using ISM abundances for spectral fitting produces an  $N_H/A_V$  relation that is significantly higher than previous estimates. This has the effect of increasing the  $N_H$  derived from existing  $E(B-V)$  measurements. Conversely, using the new relation to estimate the brightness of an optical counterpart of an X-ray source will result in less extinction compared to estimates calculated with previous values of the  $N_H/A_V$  relation. As anticipated by Güver & Özel (2009) and further discussed by Watson (2011), the primary reason for this difference is the change in the abundances that are used for the ISM. The default metallicity library of the frequently used spectral analysis package *xspec* uses the solar abundances given by Anders & Grevesse (1989) and this is indeed the main library that is utilized in the studies of the spectral properties of X-ray sources. However, these values are known to be on average 45% higher than the updated values by Asplund et al. (2009) using the solar spectrum and by Wilms et al. (2000) for the ISM, and therefore result in a smaller  $N_H$  value for a given

amount of total absorption due to the matter in the ISM. We showed here in our systematic analysis that a change in the abundance table results in  $\approx 30\%$  change in the coefficient of the linear relation between the hydrogen column density and optical extinction.

As a final caveat, we note that when the hydrogen column density is measured using the ISM abundances (e.g., from Wilms et al. 2000), the relation we presented here should be used to predict or compare with the extinction in the optical band. On the other hand, if an  $N_{\text{H}}$  is found using solar abundances (e.g., from Anders & Grevesse 1989), then the relation reported in Güver & Özel (2009) should be employed for self consistency.

## 6. CONCLUSIONS

In this paper, we presented a comprehensive study of interstellar X-ray extinction using the *Chandra* SNR archive. We used standardized procedures and made use of the high energy and spatial resolution of the data set to assess the uncertainties in the measurement of the hydrogen column density from X-ray spectra. In contrast with earlier work, we modeled the interstellar extinction using the latest ISM model as well as interstellar abundances of Wilms et al. (2000). We also modeled SNR spectra with a variety of thermal, non-thermal, and mixed models to evaluate the effects of the continuum models on the measured hydrogen column density. In addition to assessing the uncertainties in this measurement arising from the range of models, we explored uncertainties due to the spatial variations within individual remnants and the different physical conditions of the remnants such as their compositions, temperatures, and non-equilibrium regions.

We used a Bayesian statistical analysis tools to determine the systematic uncertainties in the hydrogen column density for each remnant. We used the hydrogen column density measurements and their uncertainties in combination with the measurements of the optical extinction toward these same remnants to determine the relation between these quantities. We found a best-fit linear relation described by  $N_{\text{H}} = (2.87 \pm 0.12) \times 10^{21} A_{\text{V}} \text{ cm}^{-2}$ .

A couple of different avenues could lead to further progress in the determination of the empirical relation between the optical extinction and the hydrogen column density, as also discussed in Güver & Özel (2009) and Watson (2011). First, a larger number of high-quality optical or near-infrared spectra of SNRs need be obtained to increase the sample of precise optical extinction measurements. This is because while the extensive archives of X-ray satellites allow us to determine the hydrogen column density and its uncertainty accurately, optical/NIR SNR data, especially for strongly reddened regions in the Galaxy, are scarce. The second strategy targets different populations of sources that are bright in both optical and X-ray bands, such as blazars. The measurement of the hydrogen column density and the optical extinction in such a population would provide an independent measurement of the relation between these two quantities and help cross check the results.

This work was supported in part by *Chandra* award No. ARO-11007X. T.G. has been supported by Istanbul University: project Nos. 49429 and 57321. F.Ö. gratefully acknowledges partial support from NSF grant AST 1108753. P.S. acknowledges partial support from NASA contract NAS8-03060. This research has made use of data obtained from the *Chandra* Data Archive and the *Chandra* Source Catalog, and software provided by the *Chandra* X-ray Center (CXC) in the application packages CIAO, ChIPS, and Sherpa. The authors thank K. Arnaud for helpful discussions regarding the *tbabs* model in *xspecc*.

## REFERENCES

- Anders, E., & Grevesse, N. 1989, *GeCoA*, **53**, 197
- Arnaud, K. A. 1996, in ASP Conf. Ser. 101, *Astronomical Data Analysis Software and Systems V*, ed. G. H. Jacoby & J. Barnes (San Francisco, CA: ASP), 17
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Borkowski, K. J., Lyerly, W. J., & Reynolds, S. P. 2001, *ApJ*, **548**, 820
- Durant, M., & van Kerkwijk, M. H. 2006, *ApJ*, **650**, 1070
- Fesen, R. A., & Hurford, A. P. 1995, *AJ*, **110**, 747
- Fesen, R. A., Kirshner, R. P., & Becker, R. H. 1988, in IAU Coll. 101, *Supernova Remnants and the Interstellar Medium*, ed. R. S. Roger & T. L. Landecker (Cambridge: Cambridge Univ. Press), 55
- Fesen, R. A., Winkler, F., Rathore, Y., et al. 1997, *AJ*, **113**, 767
- Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, *Proc. SPIE*, 6270, 1
- Gorenstein, P. 1975, *ApJ*, **198**, 95
- Güver, T., & Özel, F. 2009, *MNRAS*, **400**, 2050
- Güver, T., Özel, F., Cabrera-Lavers, A., & Wroblewski, P. 2010, *ApJ*, **712**, 964
- Hester, J. J., & Kulkarni, S. R. 1989, *ApJ*, **340**, 362
- Hurford, A. P., & Fesen, R. A. 1996, *ApJ*, **469**, 246
- Kaastra, J. S., Mewe, R., Liedahl, D. A., et al. 1996, *A&A*, **314**, 547
- Keohane, J. W., Reach, W. T., Rho, J., & Jarrett, T. H. 2007, *ApJ*, **654**, 938
- Koo, B.-C., McKee, C. F., Lee, J.-J., et al. 2008, *ApJL*, **673**, L147
- Lequeux, J. 2005, *The Interstellar Medium* (Berlin: Springer-Verlag)
- Long, K. S., Blair, W. P., Matsui, Y., & White, R. L. 1991, *ApJ*, **373**, 567
- Lu, F. J., Wang, Q. D., Aschenbach, B., Durouchoux, P., & Song, L. M. 2002, *ApJL*, **568**, L49
- Manchester, R. N., Lyne, A. G., Goss, W. M., et al. 1978, *MNRAS*, **184**, 159
- Mavromatakis, F., Papamastorakis, J., Paleologou, E. V., & Ventura, J. 2000, *A&A*, **353**, 371
- Miller, J. S. 1968, *ApJL*, **154**, L57
- Nielsen, M. T. B., Voss, R., & Nelemans, G. 2012, *MNRAS*, **426**, 2668
- Oliva, E., Moorwood, A. F. M., & Danziger, I. J. 1989, *A&A*, **214**, 307
- Osterbrock, D. E. 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (Sausalito, CA: University Science Books)
- Predehl, P., & Schmitt, J. H. M. M. 1995, *A&A*, **293**, 889
- Predehl, P., & Truemper, J. 1994, *A&A*, **290**, L29
- Ratti, E. M., Bassa, C. G., Torres, M. A. P., et al. 2010, *MNRAS*, **408**, 1866
- Raymond, J. C., Blair, W. P., & Long, K. S. 1995, *ApJL*, **454**, L31
- Raymond, J. C., & Smith, B. W. 1977, *ApJS*, **35**, 419
- Reina, C., & Tarengi, M. 1973, *A&A*, **26**, 257
- Reynolds, S. P. 1998, *ApJ*, **493**, 375
- Reynolds, S. P., & Keohane, J. W. 1999, *ApJ*, **525**, 368
- Ruiz-Lapuente, P. 2004, *ApJ*, **612**, 357
- Sasaki, M., Plucinsky, P. P., Gaetz, T. J., & Bocchino, F. 2013, *A&A*, **552**, A45
- Sinervo, P. 2003, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, ed. L. Lyons, R. Mount, & R. Reitmeyer (London: Imperial College Press) 122
- Sollerman, J., Lundqvist, P., Lindler, D., et al. 2000, *ApJ*, **537**, 861
- Soria, R., Hakala, P. J., Hau, G. K. T., Gladstone, J. C., & Kong, A. K. H. 2012, *MNRAS*, **420**, 3599
- Temim, T., Dwek, E., Tchernyshyov, K., et al. 2015, *ApJ*, **799**, 158
- Verner, D. A., & Yakovlev, D. G. 1995, *A&AS*, **109**, 125
- Watson, D. 2011, *A&A*, **533**, A16
- Wilms, J., Allen, A., & McCray, R. 2000, *ApJ*, **542**, 914