Astronomy & Astrophysics

HI4PI: a full-sky H I survey based on EBHIS and GASS[*]

HI4PI Collaboration: N. Ben Bekhti[1], L. Flöer[1], R. Keller[2], J. Kerp[1], D. Lenz[1], B. Winkel[1, 2,**]

and

J. Bailin[3, 4], M. R. Calabretta[5], L. Dedes[1], H. A. Ford[6], B. K. Gibson[7], U. Haud[8], S. Janowiecki[9], P. M. W. Kalberla[1,**],
F. J. Lockman[4], N. M. McClure-Griffiths[10, 5], T. Murphy[11], H. Nakanishi[12, 13, 14],
D. J. Pisano[15], and L. Staveley-Smith[9, 16]

[1] Argelander-Institut für Astronomie (AIfA), University of Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
[2] Max-Planck-Institut für Radioastronomie (MPIfR), Auf dem Hügel 69, 53121 Bonn, Germany
[3] Department of Physics and Astronomy, University of Alabama, PO Box 870324, Tuscaloosa, AL, 35487-0324, USA
[4] National Radio Astronomy Observatory, PO Box 2, Green Bank, WV, 24944, USA
[5] CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia
[6] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[7] E. A. Milne Centre for Astrophysics, University of Hull, Hull, HU6 7RX, UK
[8] Tartu Observatory, 61602 Tõravere, Tartumaa, Estonia
[9] International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
[10] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[11] Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia
[12] Graduate School of Science and Engineering, Kagoshima university, 1-21-35 Korimoto, 890-0065 Kagoshima, Japan
[13] Institute of Space and Astronautical Science, Japan Aerospace Exploring Agency, 3-1-1 Yoshinodai, Sagamihara, Kanagawa 252-5210, Japan
[14] SKA Organization, Jodrell Bank Observatory, Lower Withington, Macclesfield, Cheshire SK11 9DL, UK
[15] Department of Physics and Astronomy, West Virginia University, PO Box 6315, Morgantown, WV 26506, USA
[16] ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Australia

ABSTRACT

Context. Measurement of the Galactic neutral atomic hydrogen (H I) column density, N_{HI}, and brightness temperatures, T_B, is of high scientific value for a broad range of astrophysical disciplines. In the past two decades, one of the most-used legacy H I datasets has been the Leiden/Argentine/Bonn Survey (LAB).
Aims. We release the H I 4π survey (HI4PI), an all-sky database of Galactic H I, which supersedes the LAB survey.
Methods. The HI4PI survey is based on data from the recently completed first coverage of the Effelsberg-Bonn H I Survey (EBHIS) and from the third revision of the Galactic All-Sky Survey (GASS). EBHIS and GASS share similar angular resolution and match well in sensitivity. Combined, they are ideally suited to be a successor to LAB.
Results. The new HI4PI survey outperforms the LAB in angular resolution ($\vartheta_{FWHM} = 16\rlap{.}'2$) and sensitivity ($\sigma_{rms} = 43$ mK). Moreover, it has full spatial sampling and thus overcomes a major drawback of LAB, which severely undersamples the sky. We publish all-sky column density maps of the neutral atomic hydrogen in the Milky Way, along with full spectroscopic data, in several map projections including HEALPix.

Key words. surveys – ISM: atoms – techniques: spectroscopic

1. Introduction

Hydrogen is the most abundant element in space and as such it is a significant tracer for gas in the Universe (Burbidge et al. 1957). Because atomic neutral hydrogen (H I) is considered to trace Galactic interstellar matter quantitatively, its distribution is commonly used to supplement observations from the gamma-ray to the far-infrared regime, for example, to extract the cosmic infrared background (CIB) from Planck satellite data (Planck Collaboration XXX 2014, and references therein). The 21-cm hyperfine structure line of H I was predicted by van de Hulst (1945). With radio observations it can easily be observed in emission and absorption (e.g., Ewen & Purcell 1951; Radhakrishnan et al. 1972; Dickey et al. 1983, 2003). Soon after the discovery of the 21-cm line (Muller & Oort 1951; Ewen & Purcell 1951), surveys of the Milky Way (MW) H I distribution were undertaken with great success. For example, analyzing the H I data provided independent evidence of the MW being a spiral galaxy (van de Hulst et al. 1954; Oort et al. 1958). The early survey observations were made with relatively small dishes of diameter less than 10 m. A few years later, Weaver & Williams (1973) used the 85-foot Hat Creek telescope for a Galactic plane survey ($|b| \leq 10°$), the Berkeley Low-Latitude Survey of Neutral

---

Hydrogen, which was subsequently complemented by a survey for $|b| \geq 10°$ (north of declination $\geq -30°$) using the same instrument (Heiles & Habing 1974). The Bell Laboratories H I survey (declination $\geq -40°$; Stark et al. 1992) is another example of a sensitive large-area survey. It also offered lower side-lobe contamination compared to earlier datasets, but with only limited spectral resolution ($\delta v = 5.3$ km s^{-1}).

In the past two decades, the Leiden/Argentine/Bonn Survey (LAB; Bajaja et al. 1985; Kalberla et al. 2005) has been the prime source of information on H I. LAB was merged using the Leiden/Dwingeloo Survey (LDS; Burton & Hartmann 1994; Hartmann & Burton 1997) and the Instituto Argentino de Radioastronomía Survey (IAR; Arnal et al. 2000; Bajaja et al. 2005), both made with 25-m class telescopes. LAB was the first all-sky H I survey corrected for stray radiation (SR; Kalberla et al. 1980a,b). SR is the result of emission collected via the side-lobes of a radio antenna. Toward high galactic latitudes, where $N_{\rm H\,I}$ is low, the SR contribution can even exceed the true hydrogen column density. This is because a large fraction of the far side-lobe pattern can be directed at the luminous Milky Way disk in such a case.

Most recently, the GALFA-HI survey (Peek et al. 2011) was initiated exploiting the huge 300-m Arecibo dish and state-of-the-art spectrometers ($\delta v = 0.2$ km s^{-1}). While it excels in sensitivity, the Arecibo telescope is not fully steerable and the fraction of the sky it can access is therefore limited. Furthermore, there is currently no SR correction software available for H I observations made with the 300-m dish.

Galactic H I is so ubiquitous that in fact no single sight line exists through the Milky Way where it could not be detected after only a few seconds of integration with a modern radio telescope. Nonetheless it took a long time for the largest fully steerable radio telescopes in the world to begin the endeavor of imaging the full sky in the 21-cm line. Only with the advent of multi-beam receivers (Staveley-Smith et al. 1996; Klein et al. 2004; Keller et al. 2006) about two decades ago, was it finally feasible to conduct all-sky surveys, which still required thousands of hours of observing time. Another important aspect was the availability of cheap correlator technology to process the multi-beam signals.

Two such surveys, the Effelsberg-Bonn H I Survey (EBHIS; Kerp et al. 2011; Winkel et al. 2016a) and the Galactic All-Sky Survey (GASS, performed with the Parkes telescope; McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015), mapped the complete northern and southern hemisphere, respectively. Full spectroscopic data of the EBHIS were recently released, while GASS has been available to the scientific community since 2009 (McClure-Griffiths et al. 2009), with revisions published in 2010 (Kalberla et al. 2010) and 2015 (Kalberla & Haud 2015). There is a noteworthy difference between EBHIS and GASS: the Effelsberg spectroscopy back-ends allowed the mapping of not only the Galactic velocity regime but also the extra-galactic sky out to a redshift of $z \sim 0.07$. For the Southern sky there is a separate extra-galactic H I survey, the H I Parkes All-Sky Survey (HIPASS, Barnes et al. 2001).

The EBHIS and GASS Milky Way data provide an excellent database to approach a wealth of scientific questions. In the past it has been used to study the MW halo (Ford et al. 2008; Winkel et al. 2011; Ben Bekhti et al. 2012; Venzmer et al. 2012; Hernandez et al. 2013; Moss et al. 2013; For et al. 2014; Röhser et al. 2014; Hammer et al. 2015; Lenz et al. 2016; Kerp et al. 2016b) and the disk-halo interaction (Ford et al. 2010; McClure-Griffiths et al. 2010; Lenz et al. 2015; Röhser et al. 2016a). Likewise the MW disk material itself can be explored in much

greater detail than previously feasible (e.g., Haud 2013; Kalberla et al. 2016; Kalberla & Kerp 2016), also revealing spectacular Galactic super-shells (McClure-Griffiths et al. 2006; Moss et al. 2012). Most recently, Kerp et al. (2016a) reported the existence of a giant halo of H I around the Andromeda galaxy (M 31).

We now aim to substitute the all-sky LAB survey using the latest data from the EBHIS and GASS, both offering much higher angular resolution and sensitivity. Furthermore, LAB sampled the sky on a beam-by-beam grid only, which leads to difficulties in reconstructing small-scale features (see e.g., Kerp et al. 2011; Kalberla et al. 2016). The result of the merged EBHIS and GASS datasets is hereafter called the HI4PI Survey.

Since 2010, when the work on the third version of GASS began, EBHIS and GASS data reduction software was developed in close collaboration, with the ultimate goal of merging the two datasets. It turned out to be very beneficial that the software development for both surveys was done at the same site (Bonn), because intermediary and final data products were kept compatible early on. For example, a common HEALPix grid (Górski et al. 2005) was used for internal storage of spectra and we took care to use a common brightness temperature scale for both surveys.

In Sect. 2 we summarize the observing strategies and data processing for both surveys. The resulting full-sky H I column density map, $N_{\rm H\,I}$, is presented in Sect. 3, integrated over the full HI4PI velocity interval. Thus the H I also comprises extra-planar objects (e.g., intermediate- and high-velocity clouds) and nearby galaxies, which leads to a certain contamination of the Galactic $N_{\rm H\,I}$ map. This is discussed in Sect. 3.3. Section 4 describes how the new H I column density maps and data cubes can be retrieved. We conclude with a summary in Sect. 5. Furthermore, a comparison of EBHIS and GASS data quality and calibration consistency is made by inspecting the overlap area (declination range: $-5° \leq \delta \leq 0°$) of the two surveys. This is presented in Appendix A.

2. Observations and data processing

2.1. GASS observations

GASS observations were performed from January 2005 to November 2006, utilizing the 13-beam 21-cm feed array installed at the Parkes 64-m telescope (McClure-Griffiths et al. 2009). To avoid solar interference, which can cause ripples in the H I spectra, the measurements were carried out only during night time. Each area of the sky was covered twice, scanning in right ascension and declination, respectively. The feed array was rotated by 19°1 with respect to the scan direction to achieve a homogeneous distribution of the scan tracks. Data were recorded using a 2048-channel auto-correlator in in-band frequency-switching mode, storing one spectrum per feed and polarization every 5 s. The scan speed was 1° min^{-1}. The total bandwidth is 8 MHz which corresponds to a radial velocity coverage of $-470 \leq v_{\rm lsr} \leq 470$ km s^{-1}.

2.2. EBHIS observations

EBHIS was conducted with the Effelsberg 100-m telescope, utilizing the seven-beam 21-cm receiver. Measurements were carried out between end of 2008 and early 2013. State-of-the-art

fast Fourier transform spectrometers (FFTS; Stanko et al. 2005; Klein et al. 2012), allowed us to simultaneously cover 100 MHz of bandwidth with 16 384 spectral channels. The resulting frequency and velocity channel width is slightly larger than for GASS. Owing to the large bandwidth, both the Galactic and extra-galactic sky could be observed simultaneously. For the first data release of EBHIS (Winkel et al. 2016a) the Galactic velocity regime ($-600 \le v_{lsr} \le 600$ km s^{-1}) was published. The FFTS can record spectra with integration (dump) times of ~100 ms. To reduce the raw data volume, H I spectra were stored every 500 ms, which is an order of magnitude faster than in the case of GASS. The short integration times are beneficial for RFI mitigation of burst-like events (see Winkel et al. 2016a, for details). EBHIS was scanned along right ascension (scan speed: 4° min^{-1}). Analogous to GASS, a 19°.1 feed rotation angle was used to ensure homogeneous angular sampling.

2.3. Data reduction

Pre-processing of GASS data was performed using the *Livedata* package (Barnes et al. 2001), but for radio-frequency interference (RFI) mitigation, flux calibration, baseline fitting, and SR correction, dedicated software was used. For details, we refer to the GASS data reduction papers (McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015).

In comparison with GASS, EBHIS was processed using entirely different data reduction software and techniques (Winkel et al. 2010, 2016a, and references therein). It was decided to not utilize the frequency-switching technique during processing. This was done to improve brightness temperature calibration by accounting for its frequency dependence, which is not easily possible for frequency switching (see Winkel et al. 2012, for details). Also, the very different RFI environment and spectrometers made it necessary to develop a well-adapted RFI mitigation tool (Flöer et al. 2010; Winkel et al. 2016a). Baselines were fitted using 2D polynomials instead of 1D polynomials as in the case of GASS.

The absolute flux calibration for EBHIS and GASS is performed in the same way as for the LAB survey. It is based on the ideas developed by Kalberla et al. (1982) and uses IAU standard line calibrators S 6, S 7, and S 8. This ensures a well-matching brightness temperature- and flux density calibration of all these data sets. Also the stray-radiation correction follows the same approach for all three surveys, originally developed by Kalberla et al. (1980a,b) for the 100-m telescope at Effelsberg and later improved upon and adapted to different telescopes (Kalberla et al. 2005, 2010; Winkel et al. 2016a). For both surveys, a convolution-based data gridding technique is applied. For EBHIS, the gridding software, cygrid[1], was recently published under open-source license and we refer to Winkel et al. (2016b) for more details on the underlying algorithm. Originally, for the first GASS data release, the *Gridzilla* software, which is part of *Livedata*, was used for gridding (Barnes et al. 2001). The second and third data releases also use a Gaussian-convolution based gridding software (Kalberla et al. 2010) but, unlike cygrid and Gridzilla, it lacks full support of the FITS (Hanisch et al. 2001) world coordinate system standard (WCS; Greisen & Calabretta 2002; Calabretta & Greisen 2002). WCS is widely used to define map projections in FITS images.

---

[1] https://github.com/bwinkel/cygrid

Table 1. Comparison of basic parameters of selected H I surveys of the Milky Way.

	LAB	GASS	EBHIS	HI4PI	Unit
δ	Full	$\le 1°$	$\ge -5°$	Full	
ϑ_{FWHM}	36′	16′.2	10′.8	16′.2	
$\|v_{lsr}\|$	$\le 460^{\dagger}$	≤ 470	≤ 600	$\le 600^{*}$	km s^{-1}
Δv	1.03	0.82	1.29	1.29	km s^{-1}
δv	1.25	1.00	1.49	1.49	km s^{-1}
σ_{rms}	80	55*	90	~43	mK
N_{HI}^{lim}	3.9	2.5	4.7	~2.3	10^{18} cm^{-2}
S_{HI}^{lim}	16.1	2.1	1.8	~2.0	Jy km s^{-1}

Notes. The table quotes the declination range, δ, angular resolution ϑ_{FWHM}, velocity interval, v_{lsr}, channel separation, Δv, spectral resolution, δv, and brightness temperature noise level, σ_{rms}. The bottom two rows quote theoretical 5σ detection limits (velocity-integrated intensity) integrating over a Gaussian profile of 20 km s^{-1} line width (FWHM) for surface brightness and point sources, respectively. Adapted from Winkel et al. (2016a). $^{(\dagger)}$ Northern (LDS) part: $v_{lsr} \le 400$ km s^{-1}. $^{(*)}$ Southern (GASS) part: $|v_{lsr}| \le 470$ km s^{-1}. $^{(\star)}$ Re-scaled intensity calibration (see Kalberla & Haud 2015).

2.4. Merging the two data sets: HI4PI

Important survey parameters of EBHIS and GASS were compiled in Winkel et al. (2016a, their Table 1). For the reader's convenience, we reproduce it here and add a column for HI4PI (Table 1). Nominally, EBHIS has $30-40\%$ higher brightness temperature noise than GASS. However, smoothing EBHIS to GASS angular resolution and decreasing the spectral resolution of GASS to match the one of EBHIS yields almost the same average noise level of ~43 mK in the two independent products used for HI4PI. Therefore, HI4PI offers a convenient and coherent data base of H I, which is homogeneous in all basic parameters across the two hemispheres.

We note that because the extra-galactic part of EBHIS is less sensitive than HIPASS, the EBHIS team is currently conducting observations for a second coverage of the sky, which will bring EBHIS and HIPASS to a very similar sensitivity level. Like GASS, EBHIS will then have two orthogonal scan directions.

In Winkel et al. (2016a) it was demonstrated that EBHIS and GASS yield very consistent column densities and brightness temperatures (see also Appendix A). With respect to the LAB survey no significant bias in the intensity scales could be identified, implying that the new HI4PI dataset can safely be used as a drop-in replacement for the former LAB data. To demonstrate the improvement in image fidelity of HI4PI compared to the LAB survey, we show in Fig. 1 N_{HI} maps of four regions containing very distinct objects (see also Sect. 3.3). Owing to the better angular resolution, HI4PI reveals significantly finer structure in the H I gas. It also contains fewer image artifacts. For example, in the top left panel LAB appears somewhat blocky, which is due to insufficient spatial sampling (causing aliasing) and in the bottom left panel there are clearly several outlying pixels.

Both surveys, EBHIS and GASS, have a considerable overlap ($-5° \lesssim \delta \lesssim 0°.5$), which allows us to assess the quality of the data by direct comparison. This is presented in detail in Appendix A, one result being that in some regions in the
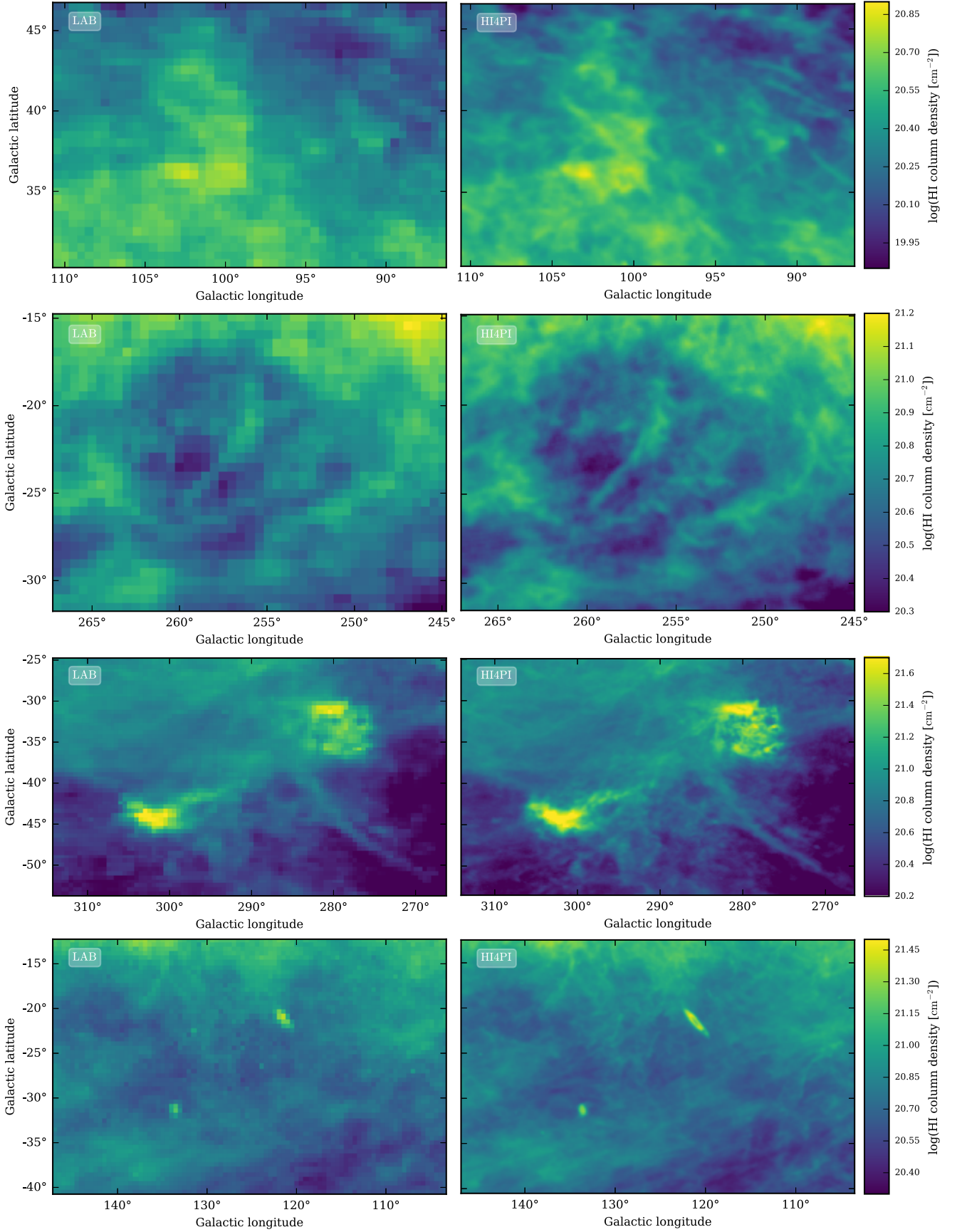
Fig. 1. Comparison between LAB and HI4PI column densities for selected regions, which are further discussed in Sect. 3.3. The *top row* shows the Draco cloud, the *second and the third row* contain a part of the Leading Arm of the Magellanic Cloud system, including the SMC and LMC, respectively. The *bottom row* displays the environment of M 31 and M 33.
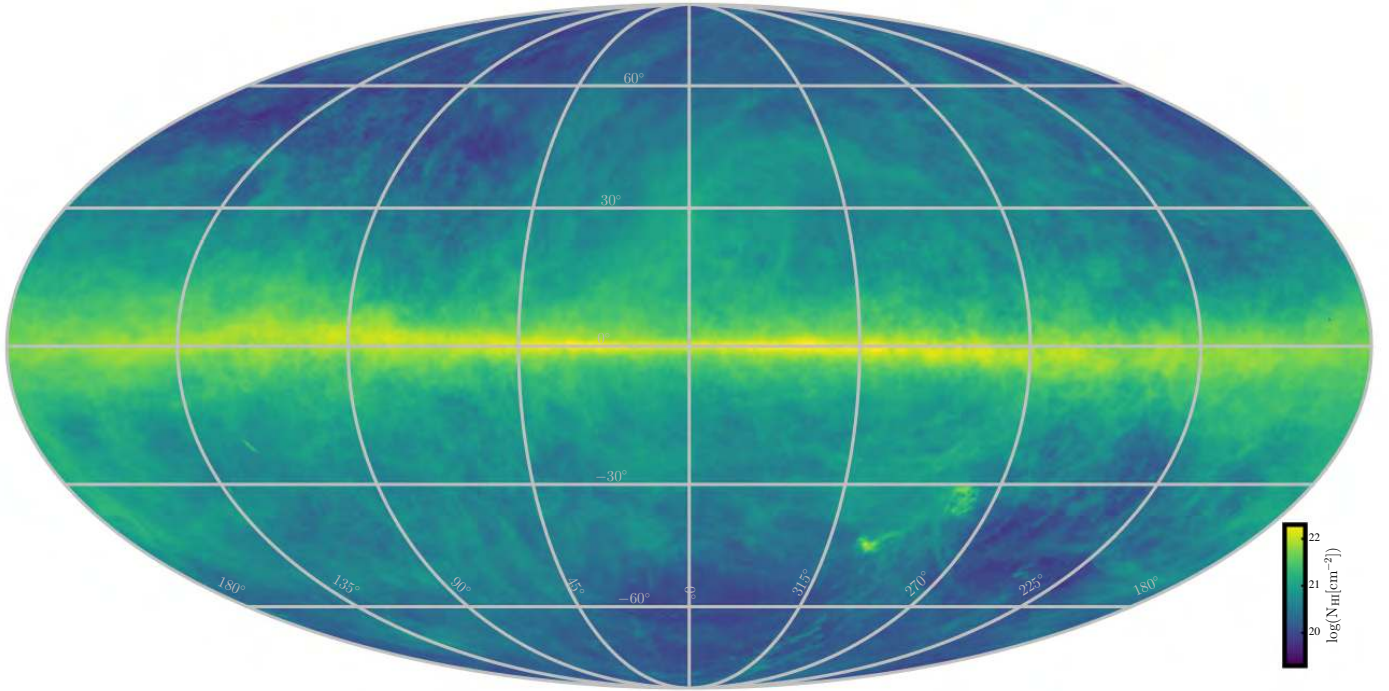
Fig. 2. HI4PI: all-sky column density map of H I gas from EBHIS and GASS data as integrated over the full velocity range $-600 \leq v_{\mathrm{lsr}} \leq 600 \mathrm{~km\,s^{-1}}$. The map is in Galactic coordinates using Mollweide projection.

overlap there are mild differences between EBHIS and GASS. Thus, a mere concatenation of the two surveys (e.g., at $\delta = 0°$) could lead to discontinuities in the H I distribution. A simple linear interpolation between EBHIS and GASS, however, is sufficient to mitigate these differences. In the overlap area (limited by δ_{lo} and δ_{hi}) we calculate the merged brightness temperature as

$$T_{\mathrm{B}}^{\mathrm{HI4PI}}(\alpha, \delta) = \frac{\delta - \delta_{\mathrm{lo}}}{\delta_{\mathrm{hi}} - \delta_{\mathrm{lo}}} T_{\mathrm{B}}^{\mathrm{EBHIS}}(\alpha, \delta) + \frac{\delta_{\mathrm{hi}} - \delta}{\delta_{\mathrm{hi}} - \delta_{\mathrm{lo}}} T_{\mathrm{B}}^{\mathrm{GASS}}(\alpha, \delta), \quad (1)$$

and we choose $\delta_{\mathrm{lo}} = -5°$ and $\delta_{\mathrm{hi}} = 0°$ as the declination range in which the interpolation is to be applied.

3. Moment maps

3.1. N_{HI} map

By integrating spectroscopic data in velocity, one can infer the $N_{\mathrm{H\,I}}$ column densities,

$$N_{\mathrm{H\,I}} \left[\mathrm{cm^{-2}} \right] = 1.823 \times 10^{18} \int \mathrm{d}v \, T_{\mathrm{B}}(v) \left[\mathrm{K\,km\,s^{-1}} \right], \quad (2)$$

where $T_{\mathrm{B}}(v)$ is the brightness temperature profile of the H I gas (e.g., Wilson et al. 2013). The resulting $N_{\mathrm{H\,I}}$ map covers the full sky (4π) and is presented in Fig. 2 on a logarithmic intensity scale. In the figure one can see the warp and flaring of the MW disk (Kalberla & Kerp 2009, and references therein). Furthermore, because of integrating across HI4PI's full velocity range ($-600 \leq v_{\mathrm{lsr}} \leq 600 \mathrm{~km\,s^{-1}}$), the $N_{\mathrm{H\,I}}$ map does not only contain MW disk material but also features residing in the MW halo: the intermediate- and high velocity clouds (IVC,

HVC) and cloud complexes, as well as extra-galactic objects such as the Magellanic Clouds (LMC and SMC) and M 31. This contamination usually needs to be considered when working with the $N_{\mathrm{H\,I}}$ data and is further discussed in Sect. 3.3. For completeness, Appendix B contains a version of the $N_{\mathrm{H\,I}}$ map using a linear intensity scale (Fig. B.1).

We note that Eq. (2) is only correct in the optically thin limit. For regions of high H I volume density (usually cold gas, having low spin temperatures), mainly at low Galactic latitudes, self-absorption occurs, such that Eq. (2) provides only a lower limit on $N_{\mathrm{H\,I}}$ (e.g., Radhakrishnan 1960; Gibson et al. 2005; Braun et al. 2009; Martin et al. 2015). H I absorption spectroscopy can in principle be used to overcome this shortcoming. For example, in the framework of the on-going THOR survey (The H I, OH, Recombination line survey of the Milky Way), Bihr et al. (2015) calculate the optical depth toward the giant molecular cloud W 43 to estimate the true H I column density. Using Galactic continuum emitters for the H I absorption technique, however, has the intrinsic issue that only the fraction of H I gas in front of the continuum emitter can be accounted for. Another approach is to study H I absorption features toward extra-galactic sources. In both cases, however, one can only derive opacities for a relatively small set of sight lines. Several studies were made to infer a correction factor, f, from such samples (see e.g. Dickey et al. 2000; Lee et al. 2015, and references therein), where f is a function of the thin gas-approximated $N_{\mathrm{H\,I}}$ value.

Although the aforementioned absorption studies provide very interesting results, it is not yet feasible to perform a proper opacity correction with the full HI4PI data set. Empirically inferred correction factors show too much spread from region to region, preventing application to the full MW disk. The situation may change, once high-resolution absorption-line surveys of the full MW disk are published.

3.2. Composite maps: combining N_{HI} and radial velocities

In Fig. 3 we display a composite map of the first two image moments: the N_{HI} column density distribution (Moment-0) and the intensity-weighted radial velocity (Moment-1). The former was calculated using Eq. (2) from the full spectral dataset (data cube); for the latter a mask was applied to avoid numerical artifacts caused by individual noise peaks in low-column density regions. This mask was constructed by applying a Gaussian filter along the spectral axis followed by a simple thresholding. The combination of both Moment maps was done in the hue-saturation-value (HSV) color space, where Moment-1 is directly used for hue and the (log-scale) N_{HI} determines the "brightness" (a combination of saturation and value). To visually enhance intermediate- and high velocity cloud features, the brightness was furthermore boosted for larger radial velocities, $|v_{\mathrm{lsr}}|$. The two-dimensional color wedge in Fig. 3 depicts the transfer function.

A slightly modified version of the composite map is shown in Appendix B (Fig. B.2) where the chosen radial velocity interval is narrower to enhance the kinematics of the MW disk material. We point out that features with radial velocities beyond the visualized velocity range were not removed from the figure but are assigned one of the two extreme colors (red and orange, respectively). Therefore one has to be careful with the interpretation of the map. For example the Magellanic System has much higher radial velocity than suggested by the color coding. This clipping also occurs in Fig. 3 but to a much lower degree.

Both, Moment-0 and Moment-1, marginalize over velocity and as such a lot of information present in the HI4PI data cubes is lost. This is a severe issue in particular within the Galactic plane where many sight lines usually reveal multiple components with very different radial velocities. As a consequence, all presented Moment maps offer only limited scientific information and one should be careful with the interpretation, especially of the composite maps displayed in Figs. 3 and B.2.

3.3. Contamination by extragalactic objects

By providing the HI4PI total column density map we also want to facilitate cross-correlation studies from the gamma-ray to the far-infrared regime. One aspect of this is that such observations need to be corrected for foreground absorption caused by neutral hydrogen, or interstellar matter quantitatively traced by H I. Toward the high Galactic latitude sky, significant absorption by H I is not commonly expected. Thus, one could be inclined to consider the total column density map, integrated in radial velocity over $|v_{\mathrm{lsr}}| \leq 600 \, \mathrm{km \, s^{-1}}$ ($470 \, \mathrm{km \, s^{-1}}$ for GASS; compare Sect. 4), as a quantitative tracer of the Milky Way's interstellar matter. But this expectation is not entirely correct: the HyperLEDA database[2] (Makarov et al. 2014) contains about 5800 entries (status: June 2016) with H I flux densities above $1 \, \mathrm{Jy \, km \, s^{-1}}$ populating this radial velocity regime, ~500 of which are most likely associated with galaxies. Only a fraction is connected to very bright H I emission: the number of objects drops to about 600 (87 galaxies) if the flux density threshold is increased to $50 \, \mathrm{Jy \, km \, s^{-1}}$. Nevertheless, one should consider the contamination of the HI4PI column density map with non-Milky Way objects.

In addition, there are intermediate-velocity clouds (Wakker 2001) and high-velocity clouds (Wakker & van Woerden 1997)

---

[2] http://leda.univ-lyon1.fr/

---

in the halo of the Milky Way (which account for the non-galaxy objects in the HyperLEDA catalog). In Fig. 3, IVC and HVC complexes, as well as several galaxies can be easily identified by eye. IVCs are thought to populate the inner halo. In the FIR, it has been found that the dust opacity and the dust temperature of IVCs differ significantly from local gas properties (Planck Collaboration XXIV 2011, their Fig. 18). However, they have similar molecular content, densities, and dust-to-gas as the MW disk gas (see e.g. Wakker 2004; Röhser et al. 2016b). Therefore, it is usually, but not always, appropriate to include the IVC sky into the H I column density when analyzing absorption processes (of external radiation), while other galaxies and HVCs often have to be treated distinctly. In the following, we will discuss these two classes of objects and their impact in more detail.

High-velocity clouds cover a significant fraction of the H I sky (see Putman et al. 2012, their Fig. 1). HVCs differ in chemical composition from IVC and Milky Way gas (Gibson et al. 2001) and appear to have insignificant amounts of dust (Wakker & Boulanger 1986; Bates et al. 1988; Boulanger et al. 1996; Planck Collaboration XXIV 2011; Saul et al. 2014; Lenz et al. 2016) and molecular gas (Richter et al. 2001). For soft X-rays $E \leq 0.5$ keV the photoelectric attenuation is mainly caused by hydrogen and helium (e.g., Snowden et al. 1990; Herbstmeier et al. 1994, 1995; Wilms et al. 2000). Thus the HI4PI column density map accounts quantitatively very well for the Galactic photoelectric absorption of redshifted X-ray sources. For higher X-ray photon energies, however, the metallicity of the interstellar gas needs to be taken into account. Here, using HI4PI N_{HI} as a unique measure for the X-ray attenuation would be an oversimplification. For a proper treatment of the photoelectric absorption cross section one needs to weight the different gas columns and their metallicity properly. Therefore, we recommend downloading HI4PI data cubes from the archive in order to evaluate the gas columns for the Milky Way and HVC velocity regimes separately. As an example, in Fig. 1 (row 1) the Draco cloud (Goerigk et al. 1983; Mebold et al. 1985; Rohlfs et al. 1989; Miville-Deschênes et al. 2016) is shown, which adds considerably to the amount of N_{HI} observed.

Extragalactic objects in the HI4PI database are all local galaxies. Prominent examples are the Magellanic Clouds, M 31, M 33 as well as the M 81/M 82 and Sculptor galaxy groups (e.g., Yun et al. 1994; Kerp et al. 2016a). All these extragalactic objects contribute significantly to, or even dominate, the HI4PI total column density map at certain positions. While more compact galaxies are easily identified, the contamination by extended objects such as the Magellanic Clouds often demands a more detailed investigation. Furthermore, the H I profile of many galaxies overlaps in velocity with MW disk material. As an example, Fig. 1 displays a part of the Leading Arm region (row 2; Brüns et al. 2005; Venzmer et al. 2012; For et al. 2013) of the Magellanic Clouds (row 3; Kim et al. 2003; Brüns et al. 2005) and the M 31/M 33 region (row 4; Braun et al. 2003; Thilker et al. 2004; Wolfe et al. 2016; Kerp et al. 2016a).

In conclusion, we emphasize that the published HI4PI N_{HI} map is not free of contamination with extra-galactic objects and HVCs. It may be possible to remove a majority of them from the data with some extra work. However, for objects that overlap in velocity with MW disk material this task gets challenging and more error-prone. Therefore, we defer this topic to a future paper. In the meantime, we recommend using the full spectral data toward problematic areas of interest (to distinguish between the
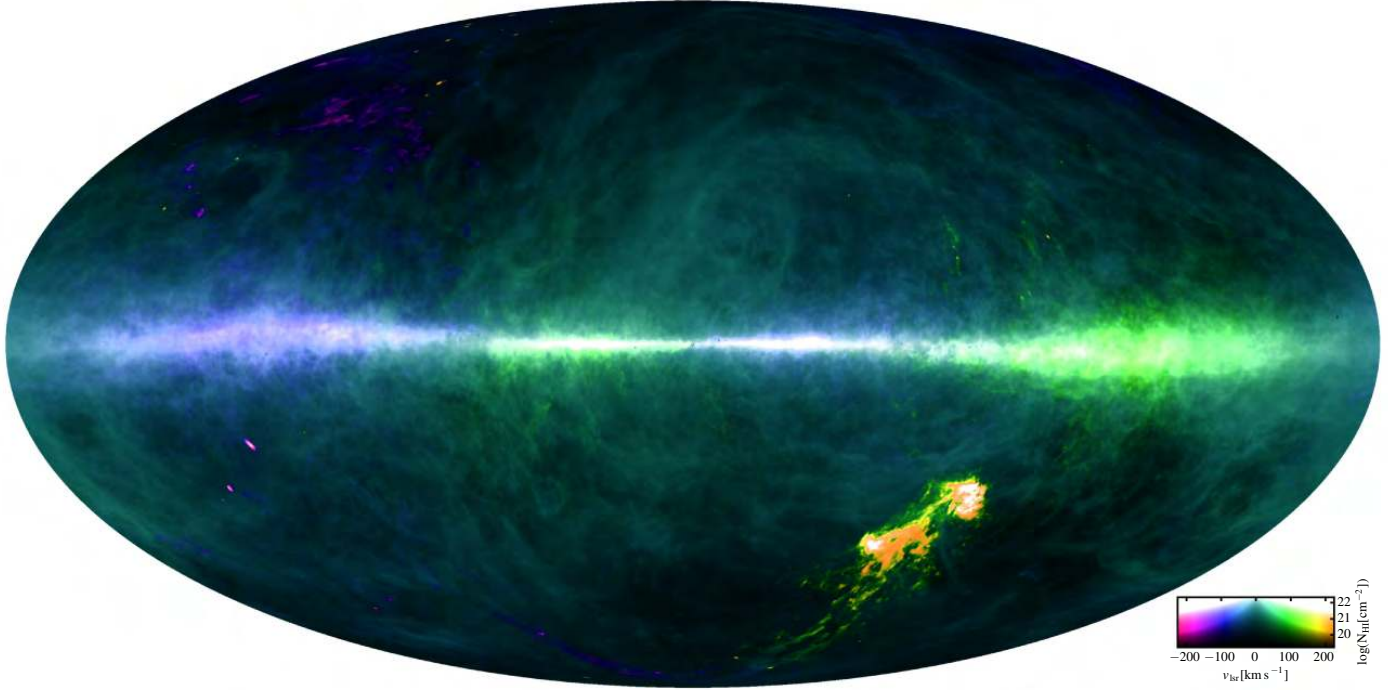
Fig. 3. All-sky composite map of the first two moments of the H I distribution using the same coordinate system and map projection as Fig. 2. The Moment-0 (the column density) determines brightness (log-scale), while the Moment-1 (intensity-weighted radial velocities) is encoded as different hues. The brightness is furthermore modified according to the velocity scale, to render the weaker intermediate- and high-velocity gas more visible. We note that radial velocities in excess of the displayed color scale are subject to clipping: the true radial velocities of red and orange features may be larger than what the colors denote.

different components) and making use of supplementary information where available.

4. Data products

The individual GASS and EBHIS datasets are already available to the scientific community. GASS spectral profiles and data cubes can be accessed via the Bonn H I-survey server[3]. The data cubes, however, are limited in size to $10° × 10°$ to minimize the computing effort and reduce the traffic on the server. EBHIS profiles are also available on the H I-survey server. Furthermore pre-computed data cubes of size $20° × 20°$ and all-sky column density maps of EBHIS are available on CDS[4].

The new data products are made available on CDS, in a similar manner as previously for EBHIS, but for the Galactic instead of the Equatorial coordinate system. We note that for certain science cases the EBHIS and GASS databases may still be better suited: the original EBHIS data offer higher angular resolution and GASS has better spectral resolution than what is now published in HI4PI.

We provide H I column density maps as FITS images in various map projections (according to WCS; Greisen & Calabretta 2002; Calabretta & Greisen 2002) such as Mollweide (MOL) and Hammer-Aithoff (AIT). The H I column density map in Fig. 2 is based on the Mollweide projection, for example. Furthermore, a version of the $N_{\rm HI}$ distribution on the HEALPix grid (Górski et al. 2005) is available. HEALPix is widely used

for all-sky datasets. For this we chose an nside parameter of 1024, which leads to a pixel size of 3′44.

For the column density maps, we chose to integrate over the full velocity range of each of the two surveys (EBHIS: $|v_{\rm lsr}| ≤ 600\,{\rm km\,s^{-1}}$, GASS: $470\,{\rm km\,s^{-1}}$). One could have used the smaller interval (GASS) also for the northern hemisphere, but then several features (e.g., M 31 and M 33, see Fig. 1 bottom panel) would have been affected.

Spectral data are presented in two flavors: (1) FITS images (data cubes), again in various projections and (2) as FITS binary table containing spectra on the HEALPix grid. Because a full-sky HI4PI data cube has a size of about 25 GB, we also offer smaller $20° × 20°$ subcubes, which in many cases should be more convenient to work with. For the same reason, the HEALPix binary table is split into smaller files, each containing the spectra for one of the 192 HEALPix pixels of the nside=4 representation.

5. Summary

We have presented the new HI4PI all-sky H I survey, that we constructed from the recent EBHIS and GASS data sets. It is freely available to the scientific community via CDS. HI4PI supersedes the LAB survey for any astronomical study that needs accurate H I profiles or column density values of the Milky Way. The various formats in which we release the data sets will make it easy to use HI4PI as a drop-in replacement for the LAB survey: (1) all-sky column density maps as FITS images and HEALPix-grid binary FITS tables, (2) all-sky FITS data cubes, and subcubes of size $20° × 20°$, and (3) spectral profiles in HEALPix-grid binary FITS tables. All of the FITS images (data cubes) are provided in various WCS projections for the Galactic coordinate system.

[3] https://www.astro.uni-bonn.de/hisurvey/
[4] http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/585/A41

The high data quality of the HI4PI constituents, EBHIS and GASS, has already been demonstrated (Kalberla & Haud 2015; Winkel et al. 2016a). Both are corrected for stray radiation. The joint database excels in angular resolution and sensitivity compared to any previous all-sky H I survey of the Milky Way.

References

Arnal, E. M., Bajaja, E., Larrarte, J. J., Morras, R., & Pöppel, W. G. L. 2000, A&AS, 142, 35

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Bajaja, E., Cappa de Nicolau, C. E., Cersosimo, J. C., et al. 1985, ApJS, 58, 143

Bajaja, E., Arnal, E. M., Larrarte, J. J., et al. 2005, A&A, 440, 767

Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, MNRAS, 322, 486

Bates, B., Catney, M. G., & Keenan, F. P. 1988, Ap&SS, 146, 195

Begum, A., Chengalur, J. N., Karachentsev, I. D., & Sharina, M. E. 2005, MNRAS, 359, L53

Behnel, S., Bradshaw, R., Citro, C., et al. 2011, Computing in Science Engineering, 13, 31

Ben Bekhti, N., Winkel, B., Richter, P., et al. 2012, A&A, 542, A110

Bihr, S., Beuther, H., Ott, J., et al. 2015, A&A, 580, A112

Boulanger, F., Abergel, A., Bernard, J.-P., et al. 1996, A&A, 312, 256

Braun, R., Thilker, D., & Walterbos, R. A. M. 2003, A&A, 406, 829

Braun, R., Thilker, D. A., Walterbos, R. A. M., & Corbelli, E. 2009, ApJ, 695, 937

Brüns, C., Kerp, J., Staveley-Smith, L., et al. 2005, A&A, 432, 45

Burbidge, E. M., Burbidge, G. R., Fowler, W. A., & Hoyle, F. 1957, Rev. Mod. Phys., 29, 547

Burton, W. B., & Hartmann, D. 1994, Ap&SS, 217, 189

Calabretta, M. R., & Greisen, E. W. 2002, A&A, 395, 1077

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Dickey, J. M., Kulkarni, S. R., van Gorkom, J. H., & Heiles, C. E. 1983, ApJS, 53, 591

Dickey, J. M., Mebold, U., Stanimirovic, S., & Staveley-Smith, L. 2000, ApJ, 536, 756

Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 585, 801

Ewen, H. I., & Purcell, E. M. 1951, Nature, 168, 356

Flöer, L., Winkel, B., & Kerp, J. 2010, in RFI Mitigation Workshop, 42

For, B.-Q., Staveley-Smith, L., & McClure-Griffiths, N. M. 2013, ApJ, 764, 74

For, B.-Q., Staveley-Smith, L., Matthews, D., & McClure-Griffiths, N. M. 2014, ApJ, 792, 43

Ford, H. A., McClure-Griffiths, N. M., Lockman, F. J., et al. 2008, ApJ, 688, 290

Ford, H. A., Lockman, F. J., & McClure-Griffiths, N. M. 2010, ApJ, 722, 367

Gibson, B. K., Giroux, M. L., Penton, S. V., et al. 2001, AJ, 122, 3280

Gibson, S. J., Taylor, A. R., Higgs, L. A., Brunt, C. M., & Dewdney, P. E. 2005, ApJ, 626, 195

Goerigk, W., Mebold, U., Reif, K., Kalberla, P. M. W., & Velden, L. 1983, A&A, 120, 63

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Greisen, E. W., & Calabretta, M. R. 2002, A&A, 395, 1061

Hammer, F., Yang, Y. B., Flores, H., Puech, M., & Fouquet, S. 2015, ApJ, 813, 110

Hanisch, R. J., Farris, A., Greisen, E. W., et al. 2001, A&A, 376, 359

Hartmann, D., & Burton, W. B. 1997, Atlas of Galactic Neutral Hydrogen

Haud, U. 2013, A&A, 552, A108

Heiles, C., & Habing, H. J. 1974, A&AS, 14, 1

Herbstmeier, U., Kerp, J., & Moritz, P. 1994, in Rev. Mod. Astron. 7, ed. G. Klare, 151

Herbstmeier, U., Mebold, U., Snowden, S. L., et al. 1995, A&A, 298, 606

Hernandez, A. K., Wakker, B. P., Benjamin, R. A., et al. 2013, ApJ, 777, 19

Hunter, J. 2007, Computing in Science Engineering, 9, 90

Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python, [Online; accessed 2015-06-27]

Kalberla, P. M. W., & Haud, U. 2015, A&A, 578, A78

Kalberla, P. M. W., & Kerp, J. 2009, ARA&A, 47, 27

Kalberla, P. M. W., & Kerp, J. 2016, A&A, in press DOI: 10.1051/0004-6361/201629113

Kalberla, P. M. W., Mebold, U., & Reich, W. 1980a, A&A, 82, 275

Kalberla, P. M. W., Mebold, U., & Velden, L. 1980b, A&AS, 39, 337

Kalberla, P. M. W., Mebold, U., & Reif, K. 1982, A&A, 106, 190

Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775

Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, A&A, 521, A17

Kalberla, P. M. W., Kerp, J., Haud, U., et al. 2016, ApJ, 821, 117

Keller, R., Nalbach, M., Müller, K., et al. 2006, Multi-Beam Receiver for Beam-Park Experiments and Data Collection Unit for Beam Park Experiments with Multi-Beam Receivers, Tech. Rep., Max-Planck-Institut für Radioastronomie

Kerp, J., Winkel, B., Ben Bekhti, N., Flöer, L., & Kalberla, P. M. W. 2011, Astron. Nachr., 332, 637

Kerp, J., Kalberla, P. M. W., Ben Bekhti, N., et al. 2016a, A&A, 589, A120

Kerp, J., Lenz, D., & Röhser, T. 2016b, A&A, 589, A123

Kim, S., Staveley-Smith, L., Dopita, M. A., et al. 2003, ApJS, 148, 473

Klein, B., Bartsch, G., Keller, R., & Leushacke, L. 2004, in 35th COSPAR Scientific Assembly, ed. J.-P. Paillé, COSPAR Meeting, 35, 3453

Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, A&A, 542, L3

Lee, M.-Y., Stanimirović, S., Murray, C. E., Heiles, C., & Miller, J. 2015, ApJ, 809, 56

Lenz, D., Kerp, J., Flöer, L., et al. 2015, A&A, 573, A83

Lenz, D., Flöer, L., & Kerp, J. 2016, A&A, 586, A121

Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13

Martin, P. G., Blagrave, K. P. M., Lockman, F. J., et al. 2015, ApJ, 809, 153

McClure-Griffiths, N. M., Ford, A., Pisano, D. J., et al. 2006, ApJ, 638, 196

McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al. 2009, ApJS, 181, 398

McClure-Griffiths, N. M., Madsen, G. J., Gaensler, B. M., McConnell, D., & Schnitzeler, D. H. F. M. 2010, ApJ, 725, 275

Mebold, U., Cernicharo, J., Velden, L., et al. 1985, A&A, 151, 427

Miville-Deschênes, M.-A., Salomé, Q., Martin, P. G., et al. 2016, A&A, submitted [arXiv:1602.03340]

Moss, V. A., McClure-Griffiths, N. M., Braun, R., Hill, A. S., & Madsen, G. J. 2012, MNRAS, 421, 3159

Moss, V. A., McClure-Griffiths, N. M., Murphy, T., et al. 2013, ApJS, 209, 12

Muller, C. A., & Oort, J. H. 1951, Nature, 168, 357

Oort, J. H., Kerr, F. J., & Westerhout, G. 1958, MNRAS, 118, 379

Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, ApJS, 194, 20

Planck Collaboration XXIV. 2011, A&A, 536, A24

Planck Collaboration XXX. 2014, A&A, 571, A30

Putman, M. E., Peek, J. E. G., & Joung, M. R. 2012, ARA&A, 50, 491

Radhakrishnan, V. 1960, PASP, 72, 296

Radhakrishnan, V., Brooks, J. W., Goss, W. M., Murray, J. D., & Schwarz, U. J. 1972, ApJS, 24, 1

Richter, P., Sembach, K. R., Wakker, B. P., & Savage, B. D. 2001, ApJ, 562, L181

Rohlfs, R., Herbstmeier, U., Mebold, U., & Winnberg, A. 1989, A&A, 211, 402

Röhser, T., Kerp, J., Winkel, B., Boulanger, F., & Lagache, G. 2014, A&A, 564, A71

Röhser, T., Kerp, J., Ben Bekhti, N., & Winkel, B. 2016a, A&A, 592, A142

Röhser, T., Kerp, J., Lenz, D., & Winkel, B. 2016b, ArXiv e-prints [arXiv:1609.06540]
Saul, D. R., Peek, J. E. G., & Putman, M. E. 2014, MNRAS, 441, 2266
Snowden, S. L., Cox, D. P., McCammon, D., & Sanders, W. T. 1990, ApJ, 354, 211
Stanko, S., Klein, B., & Kerp, J. 2005, A&A, 436, 391
Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, ApJS, 79, 77
Staveley-Smith, L., Wilson, W. E., Bird, T. S., et al. 1996, PASA, 13, 243
Terlouw, J. P., & Vogelaar, M. G. R. 2015, Kapteyn Package, version 2.3, Kapteyn Astronomical Institute, Groningen, available from http://www.astro.rug.nl/software/kapteyn/
Thilker, D. A., Braun, R., Walterbos, R. A. M., et al. 2004, ApJ, 601, L39
van de Hulst, H. C. 1945, Ned. tijd. natuurkunde, 11, 210
van de Hulst, H. C., Muller, C. A., & Oort, J. H. 1954, Bull. Astron. Inst. Netherlands,, 12, 117
van der Walt, S., Colbert, S., & Varoquaux, G. 2011, Computing in Science Engineering, 13, 22

Venzmer, M. S., Kerp, J., & Kalberla, P. M. W. 2012, A&A, 547, A12
Wakker, B. P. 2001, ApJS, 136, 463
Wakker, B. P. 2004, in High Velocity Clouds, eds. H. van Woerden, B. P. Wakker, U. J. Schwarz, & K. S. de Boer, Astrophys. Space Sci. Libr., 312, 25
Wakker, B. P., & Boulanger, F. 1986, A&A, 170, 84
Wakker, B. P., & van Woerden, H. 1997, ARA&A, 35, 217
Weaver, H., & Williams, D. R. W. 1973, A&AS, 8, 1
Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914
Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2013, Tools of Radio Astronomy (Berlin, Heidelberg: Springer-Verlag)
Winkel, B., Kalberla, P. M. W., Kerp, J., & Flöer, L. 2010, ApJS, 188, 488
Winkel, B., Ben Bekhti, N., Darmstädter, V., et al. 2011, A&A, 533, A105
Winkel, B., Kraus, A., & Bach, U. 2012, A&A, 540, A140
Winkel, B., Kerp, J., Flöer, L., et al. 2016a, A&A, 585, A41
Winkel, B., Lenz, D., & Flöer, L. 2016b, A&A, 591, A12
Wolfe, S. A., Lockman, F. J., & Pisano, D. J. 2016, ApJ, 816, 81
Yun, M. S., Ho, P. T. P., & Lo, K. Y. 1994, Nature, 372, 530

Appendix A: Comparing EBHIS and GASS

The two surveys, EBHIS and GASS share a relatively large area on the sky ($\sim$1800 deg^2 or 4.4%). This gives us the opportunity to directly compare both datasets, with the aim of assessing the data quality in the HI4PI survey. For single-dish H I data sets there are typically four major effects to be considered: (1) flux-density (or brightness-temperature) calibration uncertainties, (2) accuracy of the radial velocity scale, (3) baseline problems caused by uncertainties in the determination of the instrumental bandpass and system temperature curves but also from residual errors in the stray-radiation correction, and (4) residual radio frequency interference not identified by the RFI flagger or improperly removed after detection. These aspects have been discussed in detail previously (McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015; Winkel et al. 2016a). For convenience, we will briefly repeat the most important findings here and complement these with additional analyses with a focus on comparing the overlap region of EBHIS and GASS.

To study the common survey area, a data cube was produced for each of EBHIS and GASS covering $-4°.5 \leq \delta \leq 0°.5$ using exactly the same angular and spectral grid for both surveys, smoothed to the HI4PI resolution. This gives us the possibility to compare $N_{\mathrm{H\,I}}$ (by integrating the data cube along the spectral axis) and position-velocity slices that allow a better visualization of certain problems with the data compared to position-position diagrams.

In Fig. A.1 the resulting $N_{\mathrm{H\,I}}$ values are compared for a large portion of the overlap region. Each of the four stripes contains a 4-h interval in right ascension. The upper panels of each stripe display the GASS column density, $N_{\mathrm{H\,I}}^{\mathrm{GASS}}$, the lower panels the relative difference, $(N_{\mathrm{H\,I}}^{\mathrm{EBHIS}} - N_{\mathrm{H\,I}}^{\mathrm{GASS}})/N_{\mathrm{H\,I}}^{\mathrm{EBHIS}}$ (in percent). In the difference plots significant deviations of up to 10$-$15% are revealed toward some regions. Furthermore, a rectangular pattern is prominently visible, which we can attribute to the GASS. In the GASS column density maps, it is hardly visible, though. The pattern is related to the GASS scanning strategy, which is made up from the two orthogonal scan directions along right ascension and declination. The stripes are likely caused by correlator or receiver failures (see Kalberla & Haud 2015, their Sect. 2.7.). Only the most severe degradations could be eliminated, low level scanning problems remained.

Some of the other $N_{\mathrm{H\,I}}$ differences are related to the $5° \times 5°$ survey fields of EBHIS (marked with small black arrows in Fig. A.1). Far-side-lobe SR contamination is well known to produce patches that are related to the individual observing sessions (compare Kalberla et al. 2005, 2010; Winkel et al. 2016a). These SR contributions are typically at a level below 40 mK, occasionally up to 100 mK. This is close to the average noise level of $\sim$43 mK in HI4PI. So far, we were unable to identify the origin of the remaining SR features, which would be necessary to improve the SR correction software. We also note that SR contamination would be most severe in low-column density regions, but it appears that at higher Galactic latitudes the residual contamination is rather insignificant (compare Martin et al. 2015).

More worrying, however, is that also other medium-scale structures (several degrees but smaller than the $5° \times 5°$ fields) appear in the difference maps, which are neither clearly correlated with the observed column density nor with the scan strategy. At the moment, it is not even clear which of the two surveys is responsible for these structures. Again, the second sky coverage of EBHIS will potentially shed some light onto this issue.

To visualize the baseline quality and other effects, position-velocity (p-v) diagrams are well-suited. Figures. A.2 and A.3

display the difference between EBHIS and GASS (top left panel), the GASS data with full intensity scale (top right panel), and EBHIS and GASS data with a zoomed-in intensity scale (bottom row), which allows us to trace faint brightness temperature features.

As mentioned above, EBHIS data were smoothed to the GASS angular resolution, such that both surveys have comparable H I feature sizes and brightness temperature noise. However, the different gridding kernel size leads to a different small-scale correlation in the noise, an effect which is explained in detail in Winkel et al. (2016b, their Appendix 1). It can be seen in Figs. A.2 and A.3: noise grains in GASS appear smaller. However, this should not have any negative effect on analyses using the data.

In the following, we will discuss the HI4PI data quality in more detail based on the three Figs. A.1 to A.3 and previous work.

A.1. Brightness temperature calibration

Winkel et al. (2016a) demonstrated the overall very consistent brightness temperature calibration of EBHIS with respect to the established LAB survey data, by comparing $N_{\mathrm{H\,I}}$ and T_{B} pixelwise. They report on a nearly perfect one-to-one correlation for both quantities

$$N_{\mathrm{H\,I}}^{\mathrm{EBHIS}} = 1.0023(4) \cdot N_{\mathrm{H\,I}}^{\mathrm{LAB}} - 0.3(3) \times 10^{18}\ \mathrm{cm}^{-2}, \qquad (A.1)$$

$$T_{\mathrm{B}}^{\mathrm{EBHIS}} = 1.0000(6) \cdot T_{\mathrm{B}}^{\mathrm{LAB}} + 12.2(5)\ \mathrm{mK}. \qquad (A.2)$$

Similar results had previously been obtained by Kalberla & Haud (2015) for GASS and LAB[5]. Martin et al. (2015) compared selected small-area regions at high Galactic latitudes observed by EBHIS, GASS, and the Green Bank Telescope and also find a consistent calibration scale between the three data sets. Furthermore, they discuss the SR correction quality for the low-column density regime in detail.

Another noteworthy analysis presented by Winkel et al. (2016a) is the study of typical systematic uncertainties of the $N_{\mathrm{H\,I}}$ and T_{B} ensemble distributions. For this, the authors compared the relative difference between EBHIS and LAB column densities and brightness temperatures as a function of $N_{\mathrm{H\,I}}$ and T_{B}, respectively. They could show that the ensemble scatter can be explained by thermal noise plus a 2.5% contribution, which they attribute to intensity calibration errors; see their Sect. 5.2. We note that the thermal noise is a function of radial velocity and usually higher in regions of high brightness temperature. The resulting uncertainty in $N_{\mathrm{H\,I}}$ is a strong function of $N_{\mathrm{H\,I}}$ itself (see Winkel et al. 2016a, their Fig. 15). For low column densities ($\log N_{\mathrm{H\,I}} \lesssim 20.25$) the 1σ (68%) confidence level is about 6%, while for high column densities ($\log N_{\mathrm{H\,I}} \gtrsim 21.5$) the uncertainty is below 2.5$-$3%.

Therefore, the seemingly large deviations in the difference maps in Fig. A.1 are still consistent with the findings presented in Winkel et al. (2016a). Most of the high-column-density regions in the overlap area show less than 3% deviation. Even in the lower column density regions the deviations rarely exceed 6%. To guide the eye, we have marked the ±3% and ±6% levels with contours in Fig. A.1. To reduce clutter due to the very small-scale rectangular pattern, we spatially smoothed the differences prior to contour computation.

---

5   After re-calibrating GASS to correct for an inconsistent intensity scale between LDS and IAR survey; see Kalberla & Haud (2015) and Winkel et al. (2016a) for details.
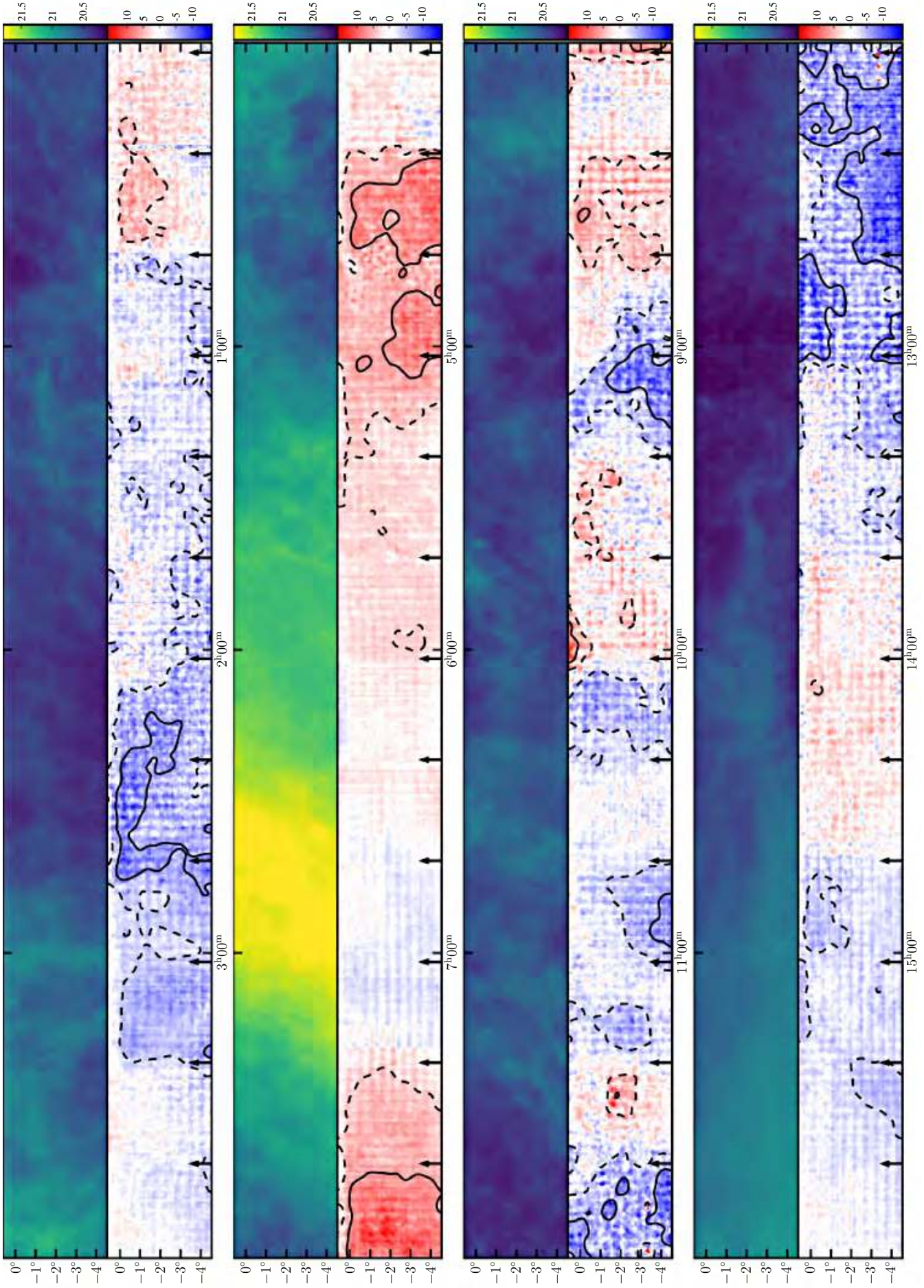
Fig. A.1. EBHIS *vs.* GASS column densities in the right ascension range 0^h to 16^h. The *upper panel* in each stripe shows $\log(N_\mathrm{H\,I}^\mathrm{GASS}[\mathrm{cm}^{-2}])$, the *lower panels* show the relative difference $(N_\mathrm{H\,I}^\mathrm{EBHIS} - N_\mathrm{H\,I}^\mathrm{GASS})/N_\mathrm{H\,I}^\mathrm{EBHIS}$ in percent. For convenience, each of the lower panels also contains contour levels for $\pm 3\%$ (dashed lines) and $\pm 6\%$ (solid lines), as well as black arrows to mark the EBHIS field limits.
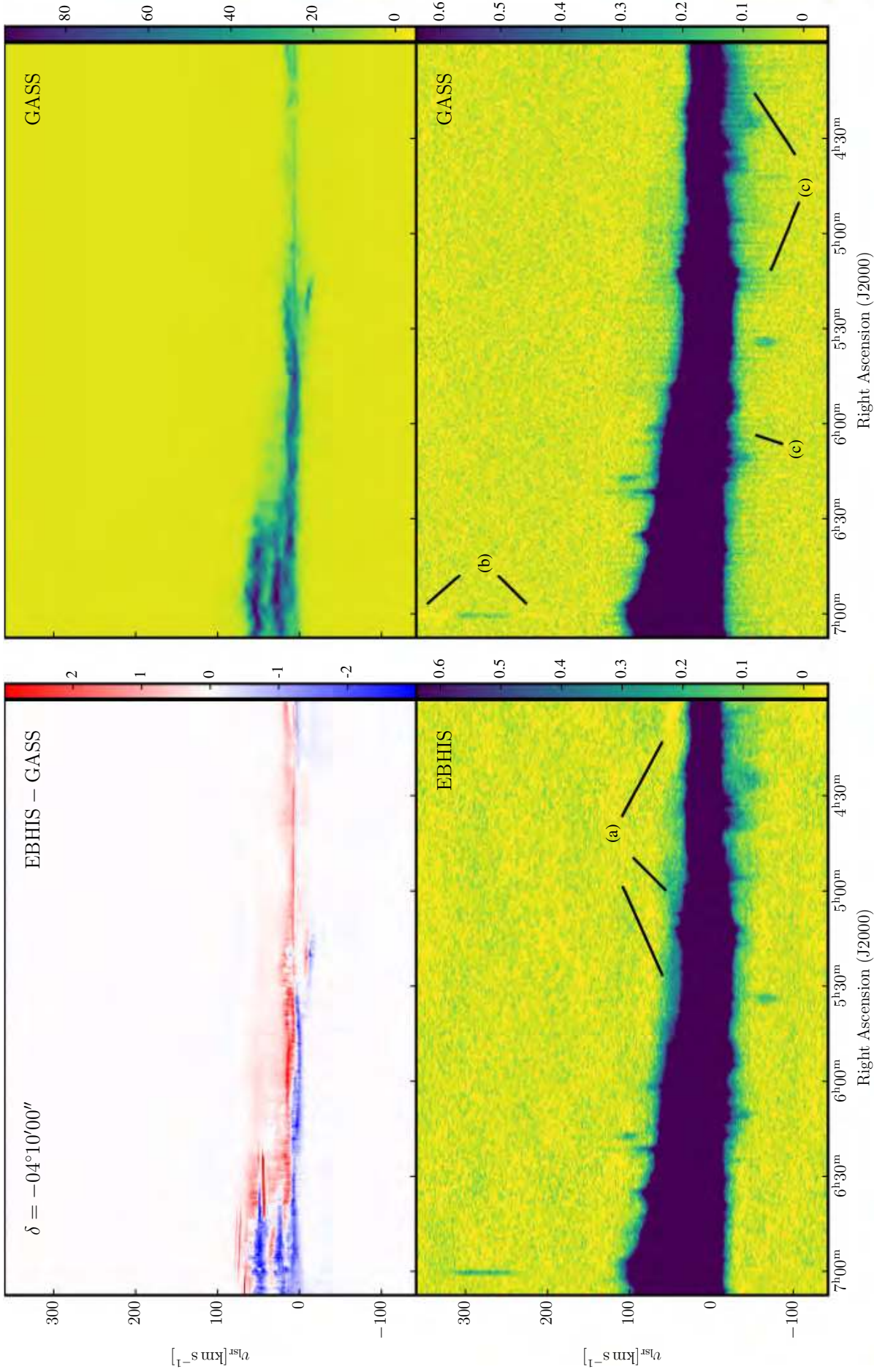
Fig. A.2. Typical EBHIS and GASS baseline problems. The *four panels* show a position-velocity slice of the overlap-region datacube. *Top left panel*: difference between EBHIS and GASS. *top right panel*: GASS data, and *bottom panels*: EBHIS and GASS data with a zoomed-in intensity scale. The labeled features are: **a)** larger-angular scale baseline fluctuations in direct vicinity of the Milky Way disk H I profile in EBHIS data; **b)** baseline deficiencies around non-disk objects in GASS, caused by improper flagging; and **c)** increase in baseline level associated to the rectangular pattern visible in Fig. A.1. All colorbars show brightness temperature, T_B, in units of Kelvin. See text for further discussion.
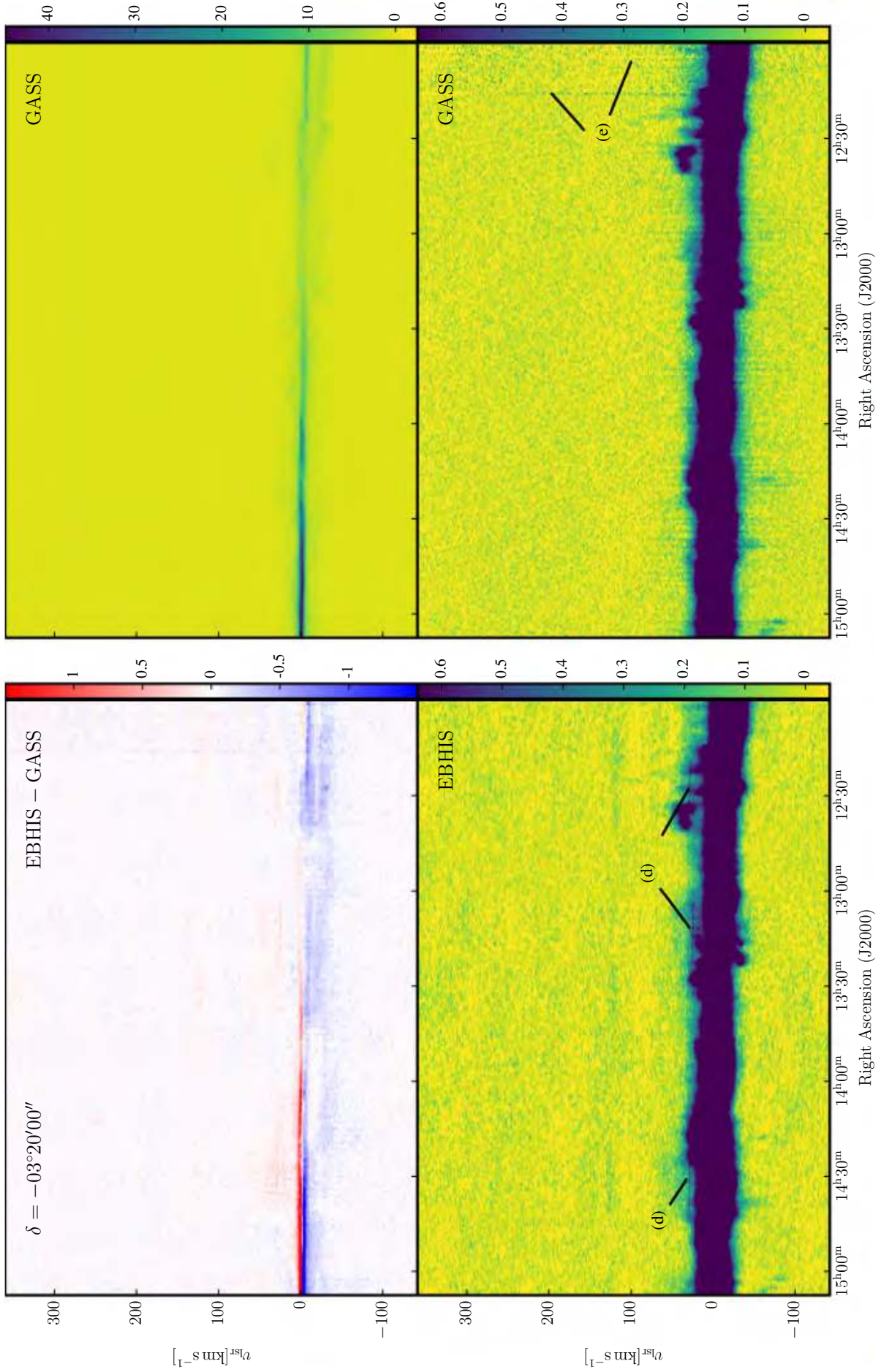
Fig. A.3. As Fig. A.2, showing typical EBHIS and GASS residual RFI features. The labeled features are: **d)** narrowband RFI within or very close to the Milky Way disk emission in EBHIS data; **e)** rare broadband events, not detected by the GASS RFI flagger. See text for further discussion.

A.2. Radial velocities

A striking feature in Figs. A.2 and A.3 are S-shaped artifacts in the difference spectra (top left panels). Today, we attribute this to a marginal radial velocity mismatch between EBHIS and GASS of about $\Delta v \approx -0.35 \, \mathrm{km \, s^{-1}}$, previously reported in Winkel et al. (2016a, see their Fig. A.3). When calculating the difference of the two data sets, a slight velocity shift will effectively act like a numerical derivative operator, which enhances the steepest gradients of each H I profile. Often these are located in the wings of the bright H I peaks, for example close to $v_{lsr} \approx 0 \, \mathrm{km \, s^{-1}}$ in Fig. A.3. In Fig. A.2 the H I profile is more complex, leading to a more complicated residual pattern in the difference panel. We could not identify the cause for the slight velocity shift so far. Winkel et al. (2016a) find an insignificant shift between LAB and EBHIS, while the shift between LAB and GASS is similar to the shift between EBHIS and GASS. This makes it somewhat more likely that the GASS data set is introducing the problem, and not EBHIS.

A.3. Baseline quality

As discussed in Sect. 2, EBHIS and GASS use very different strategies to model the baselines. EBHIS uses 2D polynomial fitting of the time-frequency plane, while GASS does 1D polynomial fitting in each individual spectral dump. In both cases, H I emission (and absorption) needs to be identified and masked prior to the fit, otherwise the resulting baseline model would be strongly biased. The parametrization of these masks is implemented as an iterative process in both data reduction pipelines. A polynomial fit is performed, after which outliers in the residual spectrum (data minus fit) are searched for and flagged. This process is repeated until the solution has converged (for details we refer to McClure-Griffiths et al. 2009; Kalberla et al. 2010; Winkel et al. 2016a). Prior information, if available, can also help to improve the baseline solution, especially during the very first iteration. Both surveys used LAB data to define priors. In addition, EBHIS utilized the HyperLEDA database (Makarov et al. 2014) to mask extragalactic H I objects and NVSS (Condon et al. 1998) to flag strong continuum sources ($\geq$1.5 Jy). Because several H I galaxies have radial velocities falling into the MW velocity regime, using the HyperLEDA database is beneficial even for the MW velocity regime. This can be seen in Fig. A.2 (label b), where a baseline deficiency is visible around the galaxy pair HIZSS003A/B (Begum et al. 2005) in the GASS data.

When interpolating across any masked feature, the polynomial fitting becomes uncertain, in particular across the very broad velocity interval in which H I emission from the MW disk is present. This gives rise to the features (label a) in Fig. A.2. In EBHIS, with the 2D approach, such baseline defects are usually

stretched over several degrees, while in GASS baseline fitting artifacts are expected to be more localized features, as each input profile is individually processed.

The regular rectangular pattern discussed in Appendix A.1 is also visible in Fig. A.2, annotated with label (c). Technically, it is not caused by baseline fitting, but it produces features in the data mimicking typical baseline-fitting artifacts.

As each receiver feed's signal has to be processed individually before gridding, the baseline fitting must be applied to spectra noisier than in the final data cube. This is a problem for EBHIS, because the structures in the raw-data baselines are a function of the incident continuum irradiation (and therefore of time, see Winkel et al. 2016a). In Fig. A.2 one can see that the underlying baseline (noise floor) in EBHIS appears somewhat less flat than in GASS.

Unrelated to the baseline fitting itself, but causing similar features, is an improper SR correction. Baseline fitting errors and SR contributions have different origins and different characteristic shapes. A good fraction of the errors is found in distinct radial velocity intervals. Accordingly the corrections are independent from each other. However remaining problems on a low level and in velocity close to the MW emission line may be hardly distinguishable. At this point objective criteria for further improvements of the corrections are missing and it is not possible to increase the quality of the data without additional information such as from a second independent sky coverage. Feature (a) in Fig. A.2, as an example, could be due to baseline uncertainties but may also be caused by improper SR correction. For a thorough discussion of SR correction in both surveys we refer to Kalberla et al. (2010) and Winkel et al. (2016a).

A.4. Residual radio frequency interference features

Figure A.3 shows another type of artifact in both data sets: residual RFI. For EBHIS, it was reported by Winkel et al. (2016a) that residual narrow-band RFI exists in the data cubes. As this type of RFI is mostly constant in frequency for a complete observation, usually $5° \times 5°$-areas are affected (the EBHIS observing field size). Close to the brightest parts of the Milky-Way emission, the EBHIS RFI mitigation software had to be used with very conservative threshold levels. Otherwise, narrow Milky Way line emission peaks would sometimes be falsely identified as RFI. As a consequence, some of the narrow-band RFI events within or close to the MW disk emission were not properly removed from the data; see Fig. A.2 (top left panel, between $6^h25 \leq \alpha \leq 7^h$, $50 \leq v_{lsr} \leq 80 \, \mathrm{km \, s^{-1}}$) and Fig. A.3 (label d). In other cases, when the RFI was correctly flagged, the removal algorithm failed to subtract the RFI intensity with the appropriate amplitude, such that sometimes spurious features remain in the data. GASS has on average fewer remaining artifacts caused by RFI; an example is visible in Fig. A.3 (label e). However, the situation for EBHIS will improve once the second sky coverage is available.
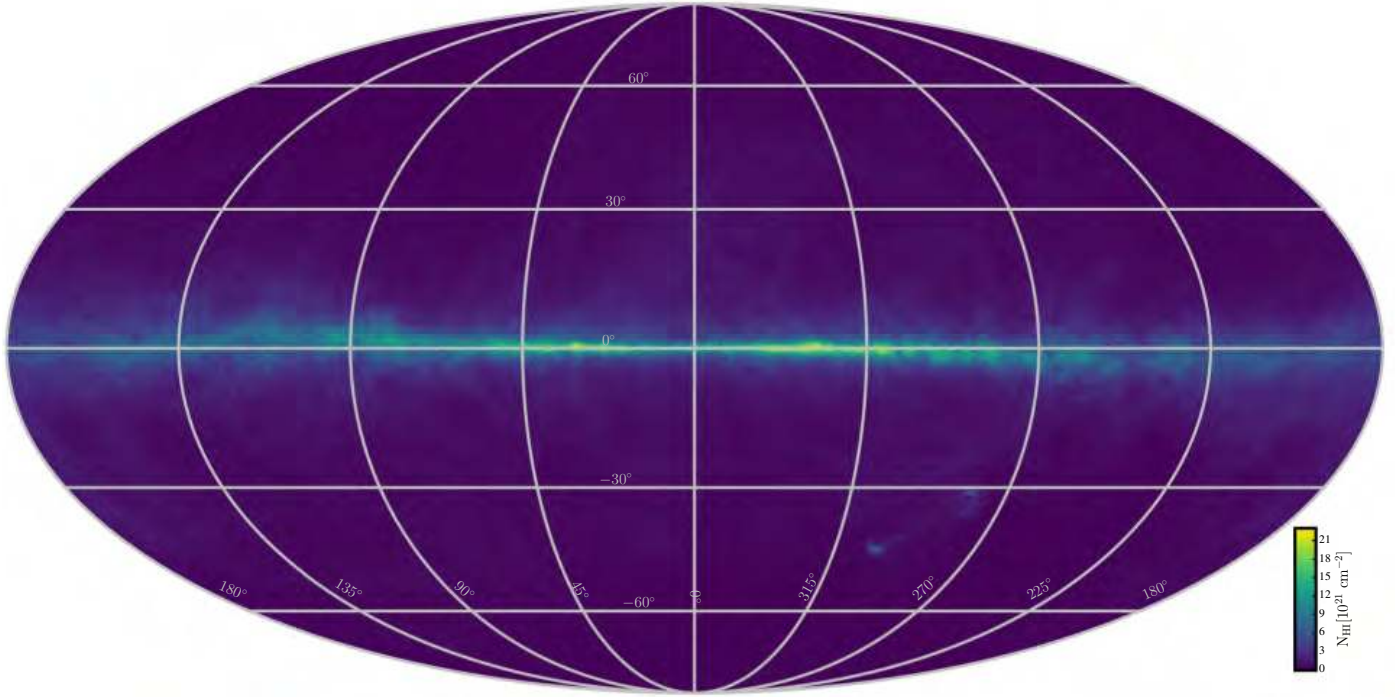
Appendix B: Supplementary figures

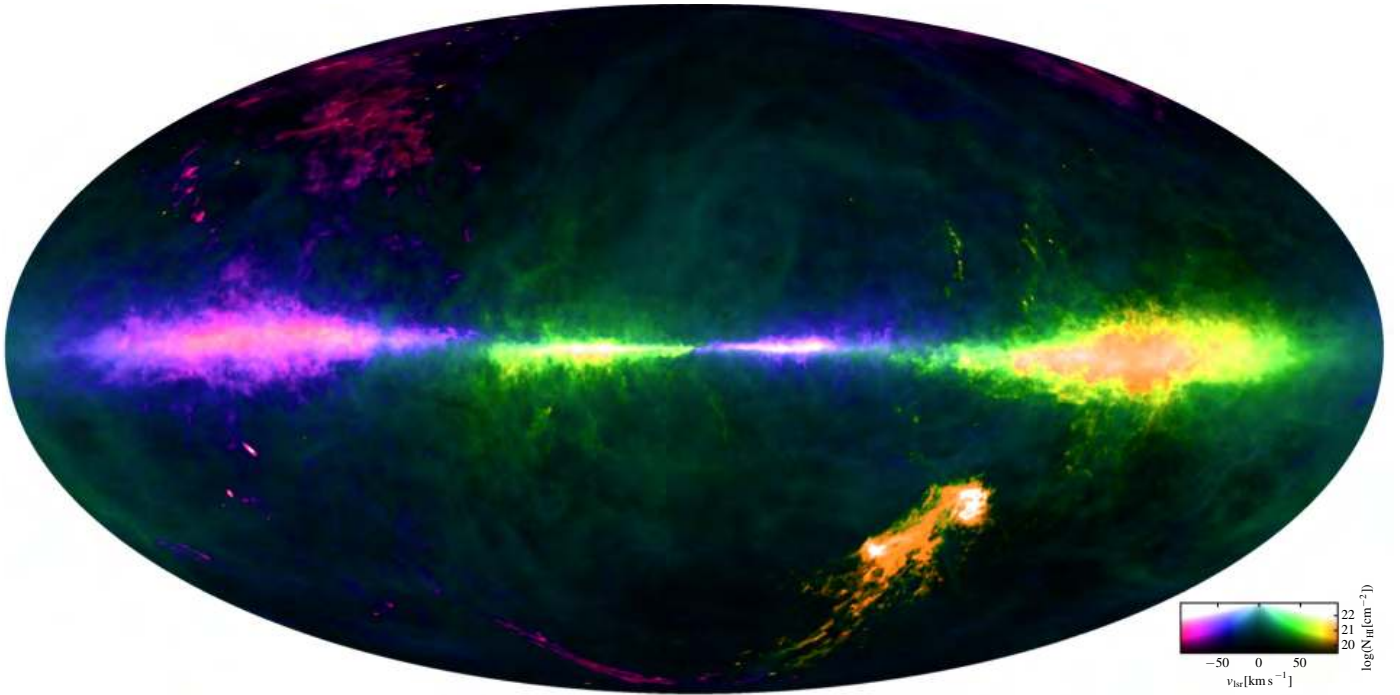**Fig. B.1.** As Fig. 2 but with linear intensity scale.

Fig. B.2. As Fig. 3 displaying a narrower velocity interval to enhance kinematic features in the Milky Way disk. Owing to the narrower velocity range, clipping effects are more severe than in Fig. 3.