

## PHANGS-ALMA: Arcsecond CO(2-1) Imaging of Nearby Star-Forming Galaxies

ADAM K. LEROY,<sup>1,2</sup> EVA SCHINNERER,<sup>3</sup> ANNIE HUGHES,<sup>4,5</sup> ERIK ROSOLOWSKY,<sup>6</sup> JÉRÔME PETY,<sup>7,8</sup> ANDREAS SCHRUBA,<sup>9</sup> ANTONIO USERO,<sup>10</sup> GUILLERMO A. BLANC,<sup>11,12</sup> MÉLANIE CHEVANCE,<sup>13</sup> ERIC Emsellem,<sup>14,15</sup> CHRISTOPHER M. FAESI,<sup>16</sup> CINTHYA N. HERRERA,<sup>7</sup> DAIZHONG LIU,<sup>3</sup> SHARON E. MEIDT,<sup>17</sup> MIGUEL QUEREJETA,<sup>10</sup> TOSHIKI SAITO,<sup>3</sup> KARIN M. SANDSTROM,<sup>18</sup> JIAYI SUN (孙嘉懿),<sup>1</sup> THOMAS G. WILLIAMS,<sup>3</sup> GAGANDEEP S. ANAND,<sup>19</sup> ASHLEY T. BARNES,<sup>20</sup> ERICA A. BEHRENS,<sup>21,1</sup> FRANCESCO BELFIORE,<sup>22</sup> SAMANTHA M. BENINCASA,<sup>1,2</sup> IVANA BEŠLIĆ,<sup>20</sup> FRANK BIGIEL,<sup>20</sup> ALBERTO D. BOLATTO,<sup>23</sup> JAKOB S. DEN BROK,<sup>20</sup> YIXIAN CAO,<sup>24</sup> RUPALI CHANDAR,<sup>25</sup> JÉRÉMY CHASTENET,<sup>17,18</sup> I-DA CHIANG (江宜達),<sup>18</sup> ENRICO CONGIU,<sup>12</sup> DANIEL A. DALE,<sup>26</sup> SINAN DEGER,<sup>27</sup> COSIMA EIBENSTEINER,<sup>20</sup> OLEG V. EGOROV,<sup>13,28</sup> AXEL GARCÍA-RODRÍGUEZ,<sup>10</sup> SIMON C. O. GLOVER,<sup>29</sup> KATHRYN GRASHA,<sup>30</sup> JONATHAN D. HENSHAW,<sup>3</sup> I-TING HO,<sup>3</sup> AMANDA A. KEPLEY,<sup>31</sup> JAEYEON KIM,<sup>13</sup> RALF S. KLESSEN,<sup>29,32</sup> KATHRYN KRECKEL,<sup>13</sup> ERIC W. KOCH,<sup>33,6</sup> J. M. DIEDERIK KRUIJSSEN,<sup>13</sup> KIRSTEN L. LARSON,<sup>27</sup> JANICE C. LEE,<sup>34,27</sup> LAURA A. LOPEZ,<sup>1,2</sup> JOSH MACHADO,<sup>1</sup> NESS MAYKER,<sup>1,2</sup> REBECCA McELROY,<sup>35</sup> ERIC J. MURPHY,<sup>31</sup> EVE C. OSTRIKER,<sup>36</sup> HSI-AN PAN,<sup>3</sup> ISMAEL PESSA,<sup>3</sup> JOHANNES PUSCHNIG,<sup>20</sup> ALESSANDRO RAZZA,<sup>12</sup> PATRICIA SÁNCHEZ-BLÁZQUEZ,<sup>37</sup> FRANCESCO SANTORO,<sup>3</sup> AMY SARDONE,<sup>1,2</sup> FABIAN SCHEUERMANN,<sup>13</sup> KAZIMIERZ SLIWA,<sup>3</sup> MATTIA C. SORMANI,<sup>29</sup> SOPHIA K. STUBER,<sup>3</sup> DAVID A. THILKER,<sup>38</sup> JORDAN A. TURNER,<sup>26</sup> DYAS UTOMO,<sup>31</sup> ELIZABETH J. WATKINS,<sup>13</sup> AND BRADLEY WHITMORE<sup>39</sup>

<sup>1</sup> Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, Ohio 43210, USA

<sup>2</sup> Center for Cosmology and Astroparticle Physics, 191 West Woodruff Avenue, Columbus, OH 43210, USA

<sup>3</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

<sup>4</sup> CNRS, IRAP, 9 Av. du Colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France

<sup>5</sup> Université de Toulouse, UPS-OMP, IRAP, F-31028 Toulouse cedex 4, France

<sup>6</sup> Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada

<sup>7</sup> Institut de Radioastronomie Millimétrique (IRAM), 300 Rue de la Piscine, F-38406 Saint Martin d'Hères, France

<sup>8</sup> Sorbonne Université, Observatoire de Paris, Université PSL, CNRS, LERMA, F-75014, Paris, France

<sup>9</sup> Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße 1, D-85748 Garching, Germany

<sup>10</sup> Observatorio Astronómico Nacional (IGN), C/Alfonso XII, 3, E-28014 Madrid, Spain

<sup>11</sup> Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>12</sup> Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile

<sup>13</sup> Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany

<sup>14</sup> European Southern Observatory, Karl-Schwarzschild Straße 2, D-85748 Garching bei München, Germany

<sup>15</sup> Univ Lyon, Univ Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France

<sup>16</sup> University of Massachusetts—Amherst, 710 N. Pleasant Street, Amherst, MA 01003, USA

<sup>17</sup> Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium

<sup>18</sup> Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

<sup>19</sup> Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

<sup>20</sup> Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

<sup>21</sup> University of Virginia, 530 McCormick Rd, Charlottesville, VA 22904, USA

<sup>22</sup> INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50157, Firenze, Italy

<sup>23</sup> Department of Astronomy, University of Maryland, College Park, MD 20742, USA

<sup>24</sup> Aix Marseille Université, CNRS, CNES, LAM (Laboratoire d'Astrophysique de Marseille), F-13388 Marseille, France

<sup>25</sup> Ritter Astrophysical Center, University of Toledo, 2801 W. Bancroft St., Toledo, OH, 43606

<sup>26</sup> Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

<sup>27</sup> Caltech-IPAC, 1200 E. California Blvd. Pasadena, CA 91125, USA

<sup>28</sup> Sternberg Astronomical Institute, Lomonosov Moscow State University, Universitetsky pr. 13, Moscow, Russia

Corresponding author: Adam K. Leroy

leroy.42@osu.edu

<sup>3</sup> This paper represents a collective effort by the PHANGS-ALMA team. Please see a description of contributions of individual team members in Appendix A.

<sup>29</sup>Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany

<sup>30</sup>Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

<sup>31</sup>National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903-2475, USA

<sup>32</sup>Universität Heidelberg, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany

<sup>33</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

<sup>34</sup>Gemini Observatory/NSF's NOIRLab, 950 N. Cherry Avenue, Tucson, AZ, 85719, USA

<sup>35</sup>Sydney Institute for Astronomy, School of Physics A28, The University of Sydney, NSW 2006, Australia

<sup>36</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544 USA

<sup>37</sup>Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, E-28040, Spain

<sup>38</sup>Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA

<sup>39</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

## ABSTRACT

We present PHANGS–ALMA, the first survey to map CO  $J = 2 \rightarrow 1$  line emission at  $\sim 1'' \sim 100$  pc spatial resolution from a representative sample of 90 nearby ( $d \lesssim 20$  Mpc) galaxies that lie on or near the  $z = 0$  “main sequence” of star-forming galaxies. CO line emission traces the bulk distribution of molecular gas, which is the cold, star-forming phase of the interstellar medium. At the resolution achieved by PHANGS–ALMA, each beam reaches the size of a typical individual giant molecular cloud (GMC), so that these data can be used to measure the demographics, life-cycle, and physical state of molecular clouds across the population of galaxies where the majority of stars form at  $z = 0$ . This paper describes the scientific motivation and background for the survey, sample selection, global properties of the targets, ALMA observations, and characteristics of the delivered ALMA data and derived data products. As the ALMA sample serves as the parent sample for parallel surveys with VLT/MUSE, HST, AstroSat, VLA, and other facilities, we include a detailed discussion of the sample selection. We detail the estimation of galaxy mass, size, star formation rate, CO luminosity, and other properties, compare estimates using different systems and provide best-estimate integrated measurements for each target. We also report the design and execution of the ALMA observations, which combine a Cycle 5 Large Program, a series of smaller programs, and archival observations. Finally, we present the first 1'' resolution atlas of CO emission from nearby galaxies and describe the properties and contents of the first PHANGS–ALMA public data release.

### 1. INTRODUCTION

This paper presents PHANGS–ALMA, an Atacama Large Millimeter/submillimeter Array (ALMA) survey aimed at studying the physics of molecular gas across the nearby galaxy population. PHANGS–ALMA is a key component of the multiwavelength observational campaign conducted by the Physics at High Angular resolution in Nearby Galaxies (PHANGS) project<sup>1</sup>. Combining a Cycle 5 ALMA Large Program, a suite of smaller programs, and data from the ALMA archive, PHANGS–ALMA mapped the CO  $J = 2 \rightarrow 1$  emission, hereafter CO(2–1), from a cleanly-selected sample of 90 of the nearest, ALMA-accessible, massive, star-forming galaxies. The resulting CO(2–1) data have high spatial and spectral resolution, good surface brightness sensitivity, full flux recovery, and good coverage of the area of active star formation in each target.

These characteristics make PHANGS–ALMA the first “cloud scale,”  $\sim 100$  pc, survey of molecular gas across a local galaxy sample that is representative of where stars form in the  $z = 0$  universe. Though the data are

suitable for many scientific applications, the survey was designed with the broad goals of quantifying the physics of star formation and feedback at the scale of individual giant molecular clouds (GMCs) and connecting these measurements to galaxy-scale properties and processes.

With these goals in mind, the PHANGS team has followed up PHANGS–ALMA with a suite of multiwavelength programs that span the spectrum from far-UV to radio, aiming to sample all stages of the star formation and feedback cycle. “PHANGS–MUSE” is obtaining optical integral field spectroscopy using the VLT/MUSE instrument to measure the properties of ionized gas and stellar populations at resolution matched to ALMA (PI: E. Schinnerer; E. Emsellem et al. in preparation). “PHANGS–HST” is using HST/WFC3 five-filter broad-band imaging to find and characterize stellar clusters and associations (PI: J. Lee; Lee et al. 2021). Other programs include new, high resolution far-UV mapping by AstroSAT (PI: E. Rosolowsky), new ground-based narrow-band H $\alpha$  imaging using the MPG 2.2m/WFI and Du Pont/DirectCCD instruments (PIs: G. Blanc, I-T. Ho; A. Razza et al. in preparation), and new HI imaging using the VLA and MeerKAT (PI: D. Utomo; D. Utomo et al. in preparation).

<sup>1</sup> <http://phangs.org/>

This paper begins with an overview of the background, design, and goals of PHANGS-ALMA (§2). Because PHANGS-ALMA served as the parent sample for many of the multi-wavelength efforts described above, we discuss the sample selection in some detail in §3. We also present our best estimates for the integrated properties of our target galaxies in §4. In §5, we describe the ALMA observations. A full description of our data processing pipeline is presented in a companion paper (Leroy et al. 2021), and we summarize our approach in §6. In §7, we describe the properties of the science-ready data products. Then we present an atlas of the PHANGS-ALMA data in §8. We give a brief summary in §9.

## 2. SCIENTIFIC MOTIVATION

### 2.1. *Background*

#### 2.1.1. *Previous Surveys of Molecular Gas in Galaxies*

Much of our knowledge about the behavior of the molecular interstellar medium (ISM) in  $z = 0$  galaxies has been established by CO surveys that either integrate over whole galaxies (e.g., the FCRAO survey, Young et al. 1995; AMIGA, Lisenfeld et al. 2011; COLD GASS, Saintonge et al. 2011; ALLSMOG, Bothwell et al. 2014; xCOLD GASS Saintonge et al. 2017, and JINGLE, Saintonge et al. 2018) or resolve the large-scale structure of galaxy disks but do not distinguish individual molecular clouds (e.g., BIMA SONG, Helfer et al. 2003; the Nobeyama CO Atlas, Kuno et al. 2007; HERACLES, Leroy et al. 2009, the JCMT NGLS, Wilson et al. 2012; CARMA STING, Rahman et al. 2012; ATLAS-3D CO, Alatalo et al. 2013; Davis et al. 2014; CARMA EDGE Bolatto et al. 2017; NRO COMING, Sorai et al. 2019; and ALMAQUEST, Lin et al. 2019).

These surveys have demonstrated a close link between molecular gas and star formation, showing that the location and rate of recent star formation in a galaxy tracks the distribution of molecular gas (e.g., Wong & Blitz 2002; Kennicutt et al. 2007; Bigiel et al. 2008; Leroy et al. 2008; Schruba et al. 2011). Yet despite this good overall correspondence, observations reveal important variations in the amount of star formation per unit molecular gas both among different types of galaxies (e.g., Saintonge et al. 2011; Leroy et al. 2013b; Davis et al. 2014; Huang & Kauffmann 2015, and see Figure 1) and within different regions of the same galaxy (e.g., Longmore et al. 2013; Leroy et al. 2013b; Meidt et al. 2013; Momose et al. 2013; Leroy et al. 2017; Utomo et al. 2017; Brownson et al. 2020). The normalized CO emission of galaxies also varies, with CO emission appearing fainter relative to starlight or tracers of recent star formation in low mass and early-type galaxies (e.g., Young & Scoville 1991; Young et al. 1996; Schruba et al. 2012; Hunt et al. 2015; Saintonge et al. 2016, 2017). This change in brightness arises from both changes in molecular gas content and changes in CO emissivity per unit

of molecular gas mass, but the relative magnitude of these effects is uncertain (for a review see Bolatto et al. 2013b).

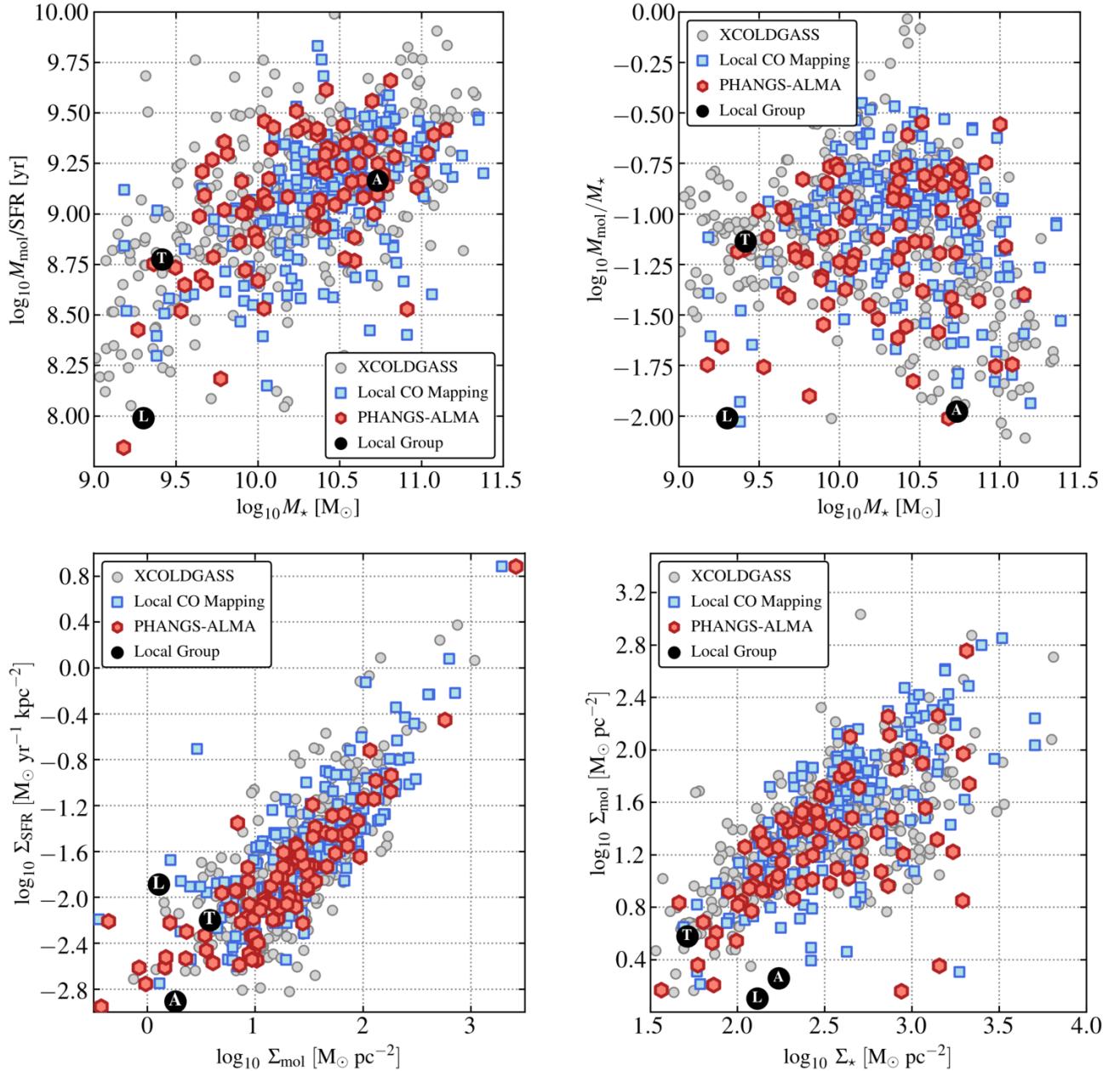
Figure 1 illustrates some of these global trends using data from PHANGS-ALMA (red, see §4 for details), xCOLD GASS (Saintonge et al. 2017) as the largest homogeneous unresolved survey, and a compilation of local CO mapping surveys (COMING, the Nobeyama CO Atlas, HERACLES, and a HERA follow up survey; Sorai et al. 2019; Kuno et al. 2007; Leroy et al. 2009, and A. Schruba et al. in preparation). The top panels show how the ratios of molecular gas mass to star formation rate and molecular gas mass to stellar mass change across the local galaxy population. The lower panels show the relationship between the average surface densities of molecular gas, stars, and star formation. Together, the four panels of Figure 1 demonstrate the good overall correspondence between star formation, stellar mass, and molecular gas in galaxies, but also illustrate important variations in the molecular content of galaxies normalized by size, star formation rate, or stellar mass. The abundance, structure, and ability of molecular gas to form stars varies across the  $z = 0$  galaxy population.

#### 2.1.2. *Key Physics at or Near Cloud Scales*

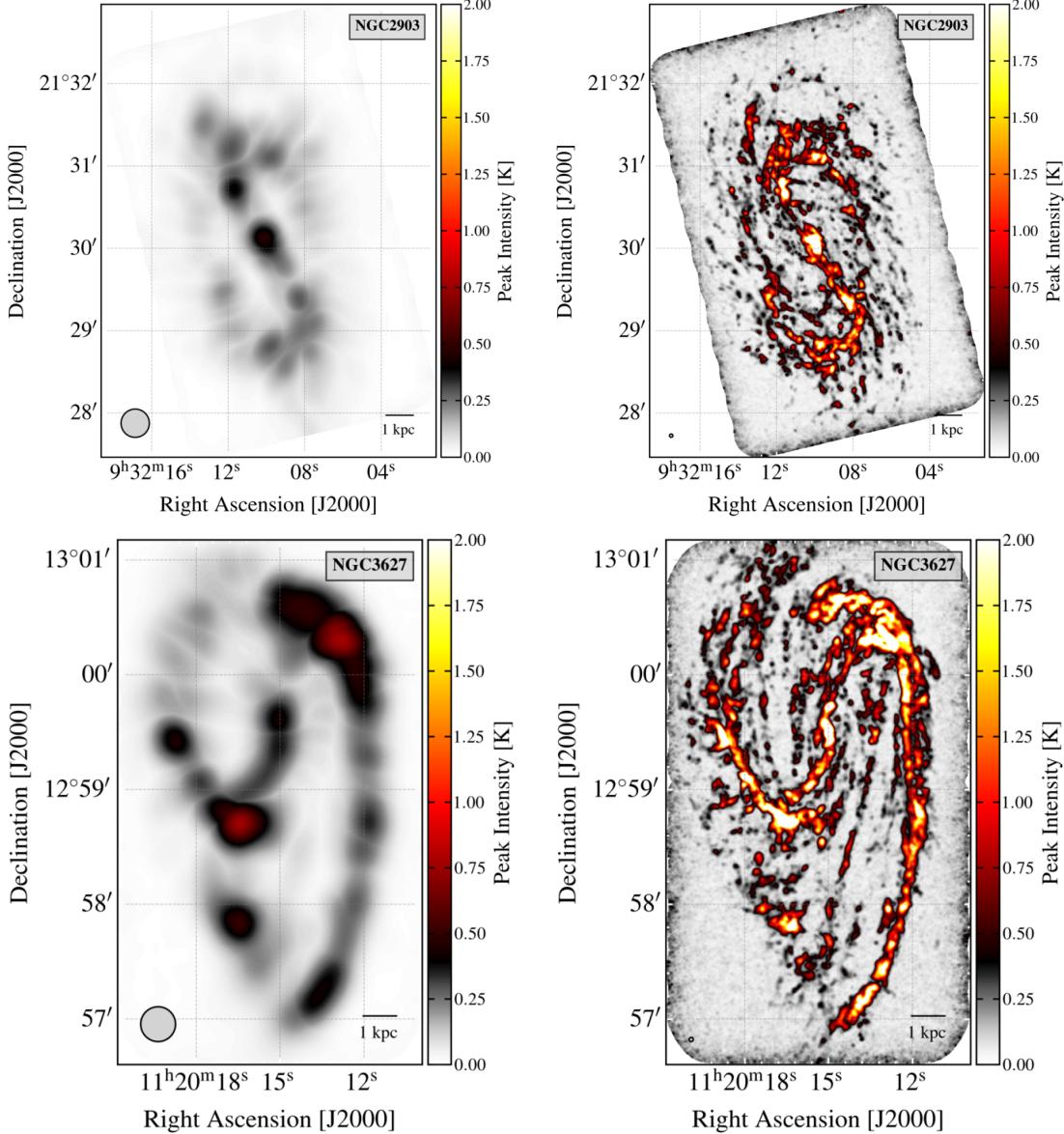
Unfortunately, low resolution observations offer limited insight into the physical state of molecular gas. In the Milky Way and its Local Group neighbors, most of the molecular gas resides in GMCs with masses  $\sim 10^4$ – $10^7$   $M_\odot$ . These clouds have sizes of tens of parsecs and appear dominated by supersonic turbulence (e.g., Solomon et al. 1987; Blitz et al. 2007; Fukui & Kawamura 2010; Roman-Duval et al. 2010; Gratier et al. 2012; Heyer & Dame 2015; Rice et al. 2016; Miville-Deschénes et al. 2017; Schruba et al. 2019).

GMCs do not fill the galaxy disk. In low resolution extragalactic observations like those mentioned in the previous section, the CO emission from GMCs is diluted with nearby non-CO emitting regions. That is, the intrinsic distribution of CO emission in galaxies is strongly clumped on scales much smaller than the  $\gtrsim$  kiloparsec resolution of the previous generation of large CO mapping surveys (e.g., Leroy et al. 2013a). Figure 2 illustrates this phenomenon by showing CO emission from two PHANGS-ALMA targets at two resolutions: 1 kpc and 120 pc resolution. The sharp, clumpy structure that is striking in maps at high resolution (right panels), blurs into faint, low-contrast structures when observed at kiloparsec resolution (left panels).

Current models of star formation predict a link between star formation, feedback, and gas properties on the scale of individual GMCs, which can be inferred using high resolution observations. For example, the mean density and density distribution within a cloud may set the characteristic timescale for star formation (e.g., Padoan & Nordlund 2002; Hennebelle & Chabrier 2011; Krumholz & McKee 2005; Federrath & Klessen 2012;



**Figure 1. Global trends in the molecular gas content of galaxies relative to their stellar mass and star formation rates. PHANGS-ALMA aims to link these global trends to local properties and local physics in the molecular gas.** Each point in each panel shows an individual galaxy from PHANGS-ALMA (red), xCOLD GASS (gray; Saintonge et al. 2017), or previous-generation CO mapping of local galaxies (blue; Kuno et al. 2007; Leroy et al. 2009; Sorai et al. 2019, A. Schruba et al. in preparation). The *top left* panel shows the ratio of molecular gas mass,  $M_{\text{mol}}$ , to SFR as a function of stellar mass,  $M_{\star}$ . The *top right* panel shows the ratio of  $M_{\text{mol}}$  to  $M_{\star}$  as a function of  $M_{\star}$ . The *bottom left* panel shows the average SFR surface density,  $\Sigma_{\text{SFR}}$ , within the half-light radius as a function of the molecular gas mass surface density,  $\Sigma_{\text{mol}}$ . The *bottom right* panel shows  $\Sigma_{\text{mol}}$  as a function of the average stellar surface density inside the half-light radius,  $\Sigma_{\star}$ . All four panels assume a fixed, Galactic CO-to-H<sub>2</sub> conversion factor of  $\alpha_{\text{CO}} = 4.35 \text{ M}_{\odot} \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$  (Bolatto et al. 2013b) and adopt a single CO (2-1)-to-CO (1-0) ratio,  $R_{21} = 0.65$  (den Brok et al. 2021; Leroy et al. 2013b) when needed. For local galaxies, we estimate SFR and  $M_{\star}$  using WISE mid-IR and GALEX UV following Leroy et al. (2019). In the bottom two panels, we assume that the effective radius for stellar mass also represents the half-light radius for CO emission and SFR. For more details see §4. For reference we show three Local Group galaxies with previous GMC-scale CO mapping: the LMC (L), M31 (A), and M33 (T) in each panel (values from Nieten et al. 2006; Fukui et al. 2008; Druard et al. 2014; Jameson et al. 2016; Leroy et al. 2019).



**Figure 2. CO emission at the kiloparsec resolution of previous surveys and the “cloud scale” resolution of PHANGS-ALMA.** The *left* panels shows CO(2-1) emission measured by PHANGS-ALMA for NGC 2903 and NGC 3627 convolved to 1 kpc resolution, roughly corresponding to the resolution of previous large CO mapping surveys. The *right* panels show the CO(2-1) emission from the same galaxies at the typical 120 pc resolution of PHANGS-ALMA. Circles show the beam in each panel, and each map shows the maximum intensity of emission along a line of sight at  $12.5 \text{ km s}^{-1}$  velocity resolution. The high resolution view shows clumpy structures corresponding to individual massive molecular clouds. The high resolution images also show a strong influence of dynamical features; both galaxies have strong stellar bars and spiral arms.

Krumholz & Dekel 2012). The strength of self-gravity in the cloud relative to turbulence and magnetic fields may affect the efficiency with which gas is converted into stars (e.g., Padoan et al. 2012, 2017; Burkhardt 2018; Kim et al. 2020b). The density, turbulence, and self-gravity may also determine how the new-born stars cluster (e.g., Kruijssen 2012; Hopkins 2013; Grudić et al.

2020; Krumholz & McKee 2020). The local gas (column) density distribution may also interact with sources of stellar feedback to determine whether a cloud is disrupted or not and how much gas and radiation leaves the system (e.g., Thompson et al. 2005; Walch et al. 2015; Thompson & Krumholz 2016; Geen et al. 2016; Raskutti et al. 2016, 2017; Reissl et al. 2018; Kim et al.

2018, 2019; Geen et al. 2021). Because the timescale, efficiency, and spatial clustering of star formation and feedback have a qualitative impact on the GMC-scale gas properties (e.g., Hopkins et al. 2013; Gentry et al. 2017; Keller & Kruijssen 2020), star formation, stellar feedback, and GMC properties form a complex, multi-scale system with many types of physics at play.

Dynamical processes acting on  $\sim$ 10–1000 pc scales also play a key role in setting the abundance, structure, and ability of molecular gas to form stars (for recent reviews see Dobbs et al. 2014; Krumholz 2014; Chevance et al. 2020c). As a concrete example, the high resolution images in Figure 2 show the unmistakable imprint of a stellar bar and spiral arms in both galaxies, which is far less obvious in the low resolution maps of these targets. Spiral arm passage may collect individual quiescent molecular clouds into large star-forming associations (e.g., Koda et al. 2009; Meidt et al. 2015), or trigger phase changes from the atomic to molecular medium (e.g., Dobbs et al. 2014). This can trigger star formation (e.g., Egusa et al. 2017) and/or organize the star-forming structures (e.g., Schinnerer et al. 2017; Elmegreen et al. 2018; Tress et al. 2020a; Kim et al. 2020c). Meanwhile, gas flows along stellar bars and arms may prevent collapse of the streaming gas, suppressing star formation (e.g., Meidt et al. 2013). These same streaming motions can also redistribute the gas, fuel star formation in the inner parts of the galaxy, or even trigger nuclear starbursts (e.g., Kenney et al. 1992; Sakamoto et al. 1999; Sheth et al. 2005; Schmidt et al. 2016). Collisions between gas clouds may also trigger star formation (e.g., Tan 2000; Inoue & Fukui 2013; Fukui et al. 2020).

A further advantage of high resolution observations is that they give access to the temporal domain of interstellar processes. Star formation is a dynamic process, with clouds evolving rapidly under the influence of both gravity (e.g., Elmegreen 2000) and “stellar feedback” (e.g., Lee et al. 2016; Kruijssen et al. 2019; Chevance et al. 2020b), a term used to describe the combined influence of ionizing photons, direct and indirect radiation pressure, gas heating, stellar winds, and supernova explosions (e.g., Lopez et al. 2011, 2014; Dale 2015; Rahner et al. 2017, 2019). Over the last decade it has been recognized that the details of the various stellar feedback mechanisms have a large impact on molecular gas properties, star formation rates, and galaxy evolution (e.g., Hopkins et al. 2012; Agertz et al. 2013; Walch et al. 2015; Klessen & Glover 2016; Kim & Ostriker 2017). But many such details remain poorly constrained by observations. The interplay between turbulence and gravity also remains imperfectly understood, with models variously positing short-lived clouds in near free-fall (e.g., Vazquez-Semadeni 1994; Elmegreen 2000; Hartmann et al. 2001a,a; Ibáñez-Mejía et al. 2016), star formation proceeding at a steady pace (e.g., Krumholz & McKee 2005), or steadily accelerating star formation within clouds (e.g., Murphy et al. 2011; Lee et al. 2016).

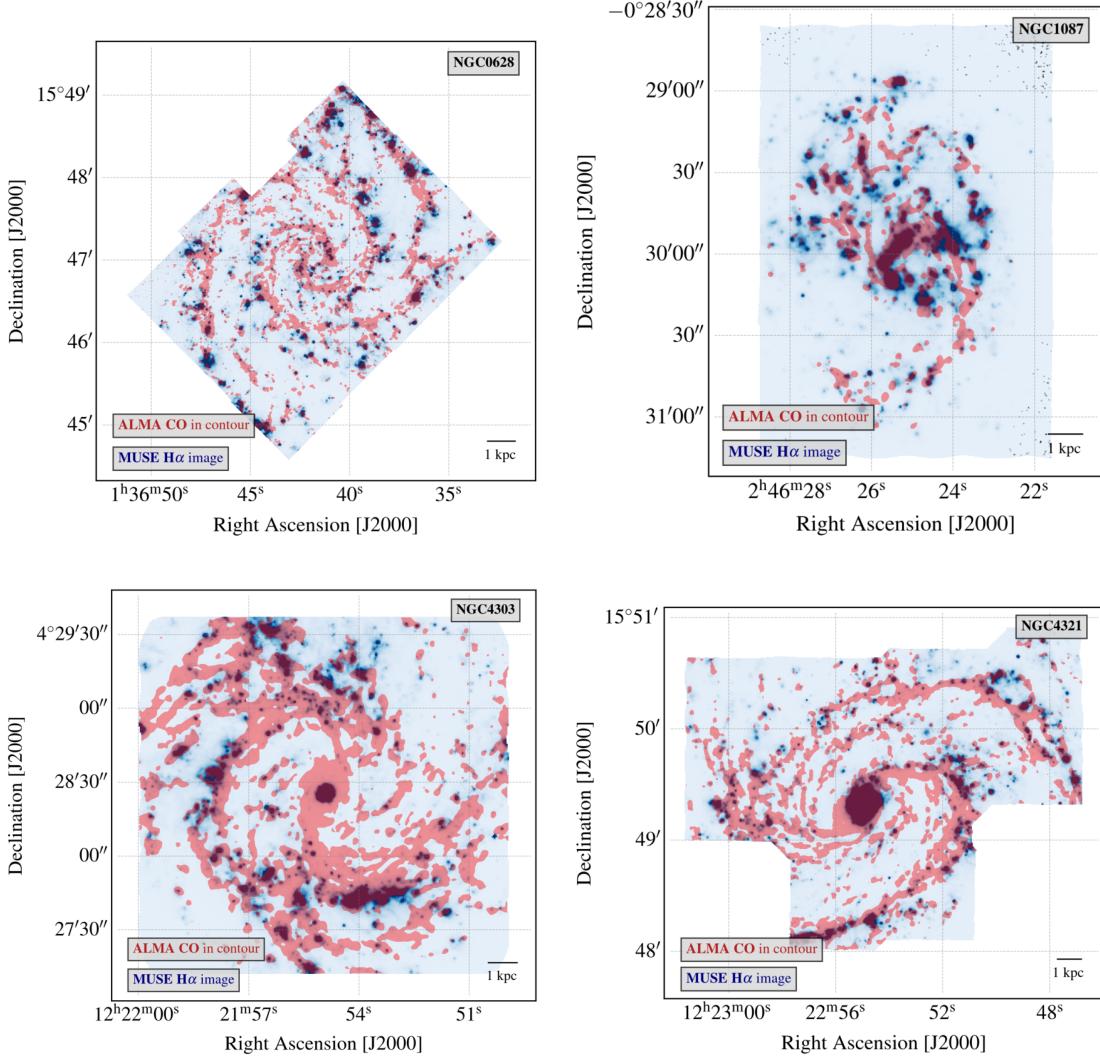
When observed at sufficient resolution, molecular gas, H II regions, stellar clusters, and other tracers of star formation and feedback visibly separate (e.g., Kawamura et al. 2009; Onodera et al. 2010; Schruba et al. 2010; Gratier et al. 2012; Schinnerer et al. 2013; Corbelli et al. 2017). To illustrate this effect, we overplot the distribution of  $\text{H}\alpha$  and CO emission for four PHANGS–ALMA targets in Figure 3. The distributions of CO, tracing GMCs, and of  $\text{H}\alpha$ , tracing recent star formation, mostly track each other at large scales, but are clearly distinct at spatial resolution of a few times 10 pc. Comparing GMCs to stellar tracers with well-understood ages or lifespans allows one to infer the timescales for star formation and feedback from high resolution imaging (e.g., Kawamura et al. 2009; Kruijssen & Longmore 2014; Corbelli et al. 2017; Kruijssen et al. 2018). This offers the prospect to build a picture of the evolutionary sequence of star formation (e.g., Murray 2011; Lee et al. 2016; Kruijssen et al. 2019), to constrain the feedback mechanisms responsible for cloud destruction (e.g., Chevance et al. 2020a), and to assess the fraction of non-star-forming molecular material (e.g., Schinnerer et al. 2019; Chevance et al. 2020b; Kim et al. 2020a).

### 2.1.3. Cloud Scale Surveys Before ALMA

High resolution CO observations measure the physical state of the gas, probe crucial dynamical processes, and constrain the timescales for star formation and feedback. So far, most cloud scale studies of normal, non-starburst galaxies have targeted members of the Local Group (see review by Fukui & Kawamura 2010). During the past two decades, there have been high spatial resolution, wide-field mapping surveys of the CO emission in the Magellanic Clouds (Fukui et al. 1999; Mizuno et al. 2001; Wong et al. 2011), M31 (Nieten et al. 2006; Rosolowsky 2007; Schruba et al. 2021), M33 (Engargiola et al. 2003; Rosolowsky et al. 2007; Onodera et al. 2012; Druard et al. 2014), and various Local Group dwarf galaxies (e.g., Leroy et al. 2006) as well as small-area mapping of Local Group targets with ALMA (e.g., Rubio et al. 2015; Schruba et al. 2017; Wong et al. 2019).

Unfortunately, the range of galaxy types and dynamical environments in the Local Group is limited. With one massive early-type spiral and a modest number of dwarf galaxies, the Local Group is not representative of the galactic environments where most star formation occurs at  $z = 0$ . Local Group galaxies also do not harbor the environmental extremes found in more distant galaxies. While their proximity offers significant advantages in terms of resolution and surface brightness sensitivity, observations in LMC, M33, and M31 cannot capture the full range of behavior seen in Figure 1.

The lack of diversity in the Local Group is problematic because we know from the low resolution surveys discussed above that the amount and behavior of molecular gas is closely linked to properties of the host galaxy. The balance between atomic gas and molecular gas depends



**Figure 3.** At high spatial resolution, tracers of massive young stars, here H $\alpha$ , and cold star-forming gas, here CO, visibly separate, providing statistical constraints on timescales for star formation and feedback. Each figure shows the CO(2-1) peak intensity maps from PHANGS-ALMA in a red contour plotted over H $\alpha$  emission measured from VLT-MUSE integral field unit observations in blue (PHANGS-MUSE, E. Emsellem et al. in preparation; see also Kreckel et al. 2016, 2018). The CO(2-1) traces cold, often star-forming gas. The H $\alpha$  traces ionizing photons produced by young stars. Though the two track one another on large scales, they show distinct distributions on small scales, reflecting that they trace different phases of the star formation process (e.g., Kawamura et al. 2009; Schruba et al. 2010; Kruijssen & Longmore 2014). The joint distributions of these two tracers, as well as stellar clusters identified by *Hubble*, constrain the timescales associated with different phases of the star formation process. In turn, these inferred timescales offer observational constraints on topics from stellar feedback to the interplay of gravity and turbulence (e.g., Chevance et al. 2020c).

sensitively on the interstellar pressure and local dust content (e.g., Wong & Blitz 2002; Blitz & Rosolowsky 2006; Leroy et al. 2008; Wong et al. 2013; Schruba et al. 2018; Sun et al. 2020a). The distribution of molecular gas strongly reflects the structure of the stellar disk (e.g., Young et al. 1995; Regan et al. 2001; Leroy et al. 2008; Schruba et al. 2011). These trends also hold across the whole galaxy population (e.g., Young & Scoville 1991; Young et al. 1995; Saintonge et al. 2011, 2017), since

the molecular gas content, or at least the CO emission, of a galaxy depends strongly on its mass and metallicity (e.g., Schruba et al. 2012; Bothwell et al. 2014; Hunt et al. 2015; Saintonge et al. 2017).

Mapping CO emission at the scale of individual GMCs across a diverse, representative sample of star-forming galaxies is thus the logical step forward in this field. Before ALMA, however, mapping GMC-scale CO emission from a single normal star-forming galaxy required a ma-

ajor time investment. As a result, mapping studies — especially with the PdBI and OVRO interferometers — typically focused on bright, compact starburst galaxies (e.g., Downes & Solomon 1998) and nuclear regions hosting starbursts and active galactic nuclei (e.g., NUGA and MAIN; García-Burillo et al. 2003; Jogee et al. 2005). After a number of studies targeting individual galaxies or dwarf galaxies (e.g., Rosolowsky & Blitz 2005; Bolatto et al. 2008; Rahman et al. 2011), the CANON survey took an important first step toward synthetic cloud-scale imaging of a sample of normal galaxies beyond the Local Group. CANON mapped CO(1–0) emission at  $\sim 2''$  resolution over the inner regions of a sample of spiral galaxies (Koda et al. 2009; Donovan Meyer et al. 2012, 2013; Momose et al. 2013). This survey provided important evidence for variations in molecular gas properties as a function of galactic environment, and for the role of spiral arms in GMC formation and evolution.

Subsequently, the PdBI Arcsecond Whirlpool Survey (PAWS; Schinnerer et al. 2013; Pety et al. 2013) mapped M51 at  $1'' \approx 40$  pc resolution (improving on a similar effort at  $3''$  resolution using CARMA by Koda et al. 2009). At this high resolution, the contrast between the molecular gas in M51 and that of Local Group galaxies proved striking (e.g., Hughes et al. 2013a,b). Figure 4 illustrates a similar contrast. It shows that at fixed 150 pc resolution, the surface density and line width of molecular gas vary significantly and systematically as a function of location in the galaxy and among host galaxies. Analysis of the PAWS data helped establish that the cloud-scale structure of the molecular ISM depends on dynamical environment and host galaxy properties (Hughes et al. 2013a,b; Colombo et al. 2014b; Leroy et al. 2016) and showed that the local star formation activity in M51 depends on cloud-scale ISM structure (Meidt et al. 2013; Leroy et al. 2017). PAWS studies also demonstrated how high resolution imaging yields insight into the evolution and timescales of individual star-forming regions (Schinnerer et al. 2013; Meidt et al. 2015; Schinnerer et al. 2017).

In spite of these important first efforts, the sample size, field of view, and sensitivity of high physical resolution CO observations targeting normal, star-forming galaxies remained limited before ALMA, preventing a synthetic view of molecular gas properties across the full local galaxy population.

## 2.2. PHANGS-ALMA

ALMA has transformed our ability to observe molecular line emission from nearby galaxies. In  $\sim 1$ – $2$  hours of on-source main array time, ALMA can map CO(2–1) emission at  $1''$  resolution across a  $\sim 2' \times 2'$  field with two times better sensitivity than achieved by PAWS. For comparison, PAWS required almost 130 hours on-source to map the CO(1–0) emission from the inner  $4.5' \times 3'$  of M51 with the IRAM PdBI (Pety et al. 2013). The dramatically faster survey speed of ALMA provides the first

opportunity of surveying a large, representative sample of local galaxies at  $\sim 1''$  resolution.

PHANGS-ALMA applies these capabilities to image CO(2–1) emission across almost all massive, nearby, southern, star-forming galaxies (see §3). The key elements of the survey are:

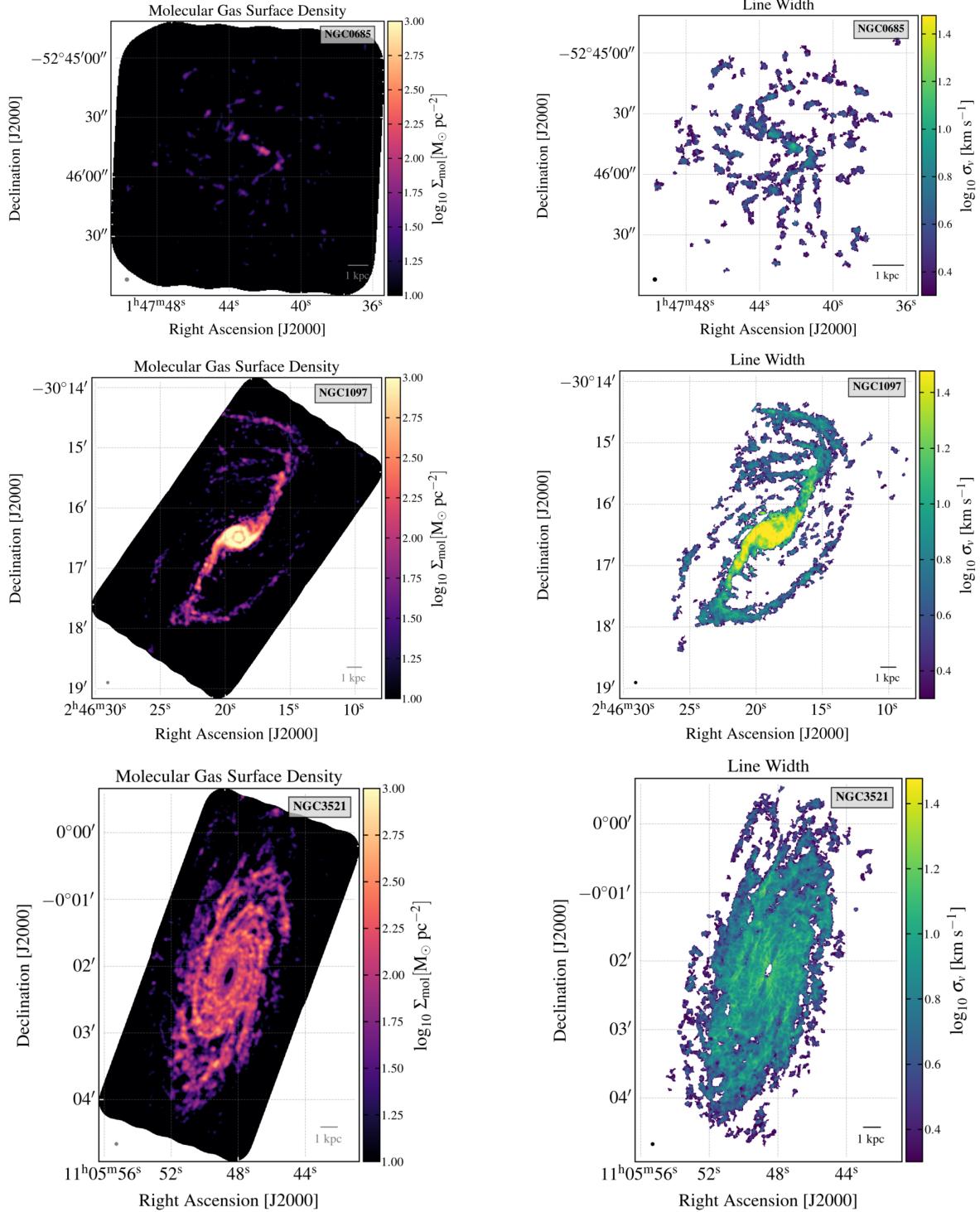
1. Observations of CO(2–1) emission across the region of active star formation of each galaxy with sensitivity to detect individual GMCs ( $M_{\text{mol}} > 10^5 M_{\odot}$ ) along each line of sight.
2. 90 targets with a simple selection function (§3) that selects most local massive galaxies on the star-forming main sequence (§4).
3.  $1'' \sim 100$  pc angular and physical resolution,  $\approx 2.5$  km s $^{-1}$  velocity resolution (§5), and inclusion of short-spacing data to ensure complete flux recovery.

The core of the survey is a Cycle 5 ALMA Large Program (PI: E. Schinnerer) that mapped CO(2–1) emission at  $\sim 1''$  resolution from 58 galaxies (§5). The Large Program built on several smaller pilot programs, and was supplemented by observations to complete the sample and extend the range of parameter space studied. Wherever feasible, archival ALMA CO(2–1) observations that match the PHANGS-ALMA observing strategy have been incorporated into the PHANGS sample for processing and analysis (see §3 and §5).

Because our selection strategy (§3) is simple, our targets include almost every massive, star-forming galaxy visible to ALMA within 12 Mpc. Our coverage of more distant targets, with  $12 \lesssim d \lesssim 17$  Mpc, is also good, but less complete due to the combination of distance uncertainties (where targets have a true distance  $< 17$  Mpc but a measurement of  $d > 17$  Mpc) and volume effects (the number of galaxies with distances  $< d$  grows  $\propto d^3$  precluding complete coverage in a reasonable amount of time).

This simple selection strategy yields a diverse sample of galaxy types (§4). Our targets span more than a decade in stellar mass, star formation rate, and specific star formation rate. They include strongly barred galaxies, grand design spirals, flocculent galaxies, and even some early-type galaxies. In the SFR– $M_{\star}$  space commonly used to discuss galaxy evolution, the PHANGS-ALMA targets provide good sampling of the local “main sequence” of star-forming galaxies (Noeske et al. 2007).

We refer to PHANGS-ALMA as a “cloud scale” spectroscopic imaging survey. This means that our resolution and sensitivity are well matched to the scale of an individual GMC. The  $\sim 100$  pc resolution of PHANGS-ALMA matches the thickness of the molecular disk in the Milky Way and other galaxies (Heyer & Dame 2015; Yim et al. 2020). Our beam also has roughly the same diameter as massive GMCs, which are often found to



**Figure 4. Cloud scale molecular gas properties depend on host galaxy and location in the galaxy.** PHANGS-ALMA CO(2-1)-based estimates of (left) molecular gas surface density,  $\Sigma_{\text{mol}}$ , and (right) effective line width,  $\sigma_v$ , for three galaxies. All three galaxies are shown at the same physical resolution of 150 pc and using the same logarithmic color stretch. The galaxies show striking differences in their overall surface density and line width, the morphology of their gas distribution at 150 pc scales, and in the distributions of gas and line width within each galaxy. Note that the line width is only measured within pixels detected at high confidence, while the integrated intensity is estimated over an area selected for completeness. See §7 for details on the calculation individual maps and §8 for similar images of the whole sample.

have radii of  $\sim 30$ – $60$  pc (e.g., Solomon et al. 1987; Bolatto et al. 2008; Colombo et al. 2014a; Freeman et al. 2017; Miville-Deschénes et al. 2017; Rosolowsky et al. 2021). The point source sensitivity of PHANGS–ALMA is also well matched to detecting individual GMCs, with a characteristic mass scale of  $M_{\text{mol}} \approx 10^5 M_{\odot}$  and power-law mass distribution (e.g., Fukui & Kawamura 2010). These characteristics make PHANGS–ALMA ideally suited to measure the demographics, motions, and organization of molecular gas (clouds) in galaxies.

The PHANGS–ALMA imaging is not designed to heavily resolve individual GMCs. Instead, we target cloud scale resolution across entire galaxies for a large sample. Although ALMA can achieve resolutions much better than  $1''$  at the  $\nu = 230$  GHz of CO(2–1), such observations have poor surface brightness sensitivity and require a prohibitive amount of time to detect CO emission from molecular clouds. Specifically, the integration time ( $t$ ) required to reach a fixed surface brightness sensitivity at a resolution ( $\theta$ ) scales as  $t \propto \theta^{-4}$ . Thus, even targeting an order of magnitude poorer sensitivity, ALMA could only survey one or two nearby galaxies at  $\sim 10$  pc resolution during the time required to map all 90 PHANGS–ALMA targets at  $\sim 100$  pc resolution.

### 2.3. Science goals

The science goals driving the PHANGS–ALMA survey design motivate this “cloud scale imaging” philosophy. We constructed the survey to address major open questions about the demographics of GMCs, the life cycle of star-forming regions, and the link between cloud scale physics, galactic scale processes and host galaxy properties.

The sample selection (§3) and observing strategy (§5) for PHANGS–ALMA were designed to address five core science goals:

1. *Measure the demographics of molecular clouds, and measure how GMC populations depend on host galaxy and location in a galaxy.*

Despite more than three decades studying GMCs in other galaxies, we lack a quantitative, observationally-grounded understanding of their demographics. Put another way, we still lack an answer to the question: “For a given set of local conditions inside a given host galaxy, what population of GMCs should be present?”

As described above, this mostly reflects the technical obstacles to observing entire GMC populations before ALMA. These limitations induced GMC studies to focus on a handful of nearby galaxies, e.g., the LMC, M33, M31, and M51.

PHANGS–ALMA aims to change this situation by measuring the distributions of GMC mass, line width, surface density, internal pressure, and virial parame-

ter<sup>2</sup> in each region of each galaxy. Because we target a diverse sample of galaxies that represents where stars are forming at  $z = 0$ , we expect PHANGS–ALMA to provide a solid empirical foundation to understand the link between GMCs, host galaxy, and dynamical environment. These measurements will provide important constraints on GMC formation, destruction, and evolution (e.g., Jeffreson & Kruijssen 2018).

This will quantitatively connect GMC studies to models of galaxy evolution (e.g., Somerville & Davé 2015) and provide key benchmarks for numerical simulations aiming to “get the cold gas right” (e.g., Dobbs et al. 2019; Jeffreson et al. 2020; Tress et al. 2020b). First work on this topic using PHANGS–ALMA appears in Sun et al. (2018, 2020a,b), Herrera et al. (2020), and Rosolowsky et al. (2021).

This science goal drives us to observe a galaxy sample that spans the star-forming main sequence, to reach a resolution that approaches the scale of individual GMCs, and to achieve sensitivity to individual GMCs.

2. *Measure the star formation efficiency per free fall time,  $\epsilon_{\text{ff}}$ , at cloud scales. Measure how  $\epsilon_{\text{ff}}$  depends on the density, dynamical state, and turbulence in molecular clouds.*

Star formation is inefficient: only a small fraction of the mass of a cloud is converted to stars over the time it takes for the cloud to gravitationally collapse (e.g., Zuckerman & Evans 1974; McKee & Ostriker 2007). Over the last two decades, many analytic and numerical models have considered star formation in turbulent molecular clouds (e.g., following Padoan & Nordlund 2002 and Krumholz & McKee 2005). These models often treat the efficiency of star formation relative to direct collapse, i.e., the “star formation efficiency per free fall time,” as a crucial prediction (e.g., see a synthesis in Federrath & Klessen 2012, 2013).

Put more simply, much work over the last two decades views either the gravitational free fall time at the scale of an individual cloud ( $\tau_{\text{ff}} \propto \rho^{-0.5}$ , with  $\rho$  the gas volume density) or the turbulent crossing time ( $\tau_{\text{cross}} \propto l/\sigma$ , with  $\sigma$  the turbulent velocity dispersion) as the relevant timescale for star formation (e.g., Elmegreen 2000; Hartmann et al. 2001b; Mac Low & Klessen 2004; Krumholz & Dekel 2012; Padoan et al. 2016). In this view, the relevant efficiency for star formation is the fraction of gas converted to stars over the relevant timescale, e.g.,  $\epsilon_{\text{ff}} \equiv M_{\text{mol}}/\text{SFR}/\tau_{\text{ff}}$  or  $\epsilon_{\text{cross}} \equiv M_{\text{mol}}/\text{SFR}/\tau_{\text{cross}}$ .

These models are increasingly central to how we understand star formation in galaxies (see the review by Krumholz et al. 2019). Testing them requires estimating the key timescales,  $\tau_{\text{ff}}$  and  $\tau_{\text{cross}}$ , on the scales of interest.

<sup>2</sup> We adopt a simple virial parameter definition  $\alpha_{\text{vir}} = 2K/U_g$ , where  $K$  is the kinetic energy and  $U_g$  is the gravitational potential energy.

In turn, estimating these timescales requires measuring the density and velocity dispersion of cold gas at the scale of an individual GMC. This requires at least “cloud scale” resolution. Because such observations have been scarce, direct measurements of  $\epsilon_{\text{ff}}$  and tests of turbulent models have been mostly confined to studies of the Milky Way (Evans et al. 2014; Vutisalchavakul et al. 2016) and a handful of the nearest galaxies (e.g., Leroy et al. 2017; Ochsendorf et al. 2017; Schruba et al. 2019).

By making measurements of the mass surface density and line width of cold gas at cloud scales, PHANGS-ALMA yields estimates of  $\tau_{\text{ff}}$  and  $\tau_{\text{cross}}$ . Combining these with measurements of SFR and the total molecular gas reservoir,  $M_{\text{mol}}$ , allows us to make resolved estimates of  $\epsilon_{\text{ff}}$  across the whole local galaxy population.

This is the second core science goal of PHANGS-ALMA: to measure  $\epsilon_{\text{ff}}$  across the local galaxy population and quantify how  $\epsilon_{\text{ff}}$  depends on host galaxy properties and local conditions in the cold gas. Doing so, we aim to provide a benchmark and test for current and future models of star formation in molecular clouds. First work on this topic using the pilot PHANGS-ALMA data appears in Kreckel et al. (2018) and Utomo et al. (2018). Similar to the first goal, this science goal drives us to observe a diverse galaxy sample, to reach a resolution that approaches the scale of individual GMCs, and to achieve sensitivity to individual GMCs.

Combining these first two goals, we aim to link the observed global trends in molecular gas content and star formation within the molecular gas to local physics. We will measure how molecular cloud populations depend on local and global environment, and we will also measure how the properties of molecular clouds affect the star formation and feedback process. This will allow us to understand if global trends in the gas depletion time stem from underlying changes in the GMC population.

3. *Quantify the “violent cycling” between phases of the star formation process. Use this to constrain the life cycle of clouds and feedback.*

Several lines of evidence suggest that GMCs experience dramatic evolution and violent disruption on timescales of a few Myr to a few tens of Myr (e.g., Kawamura et al. 2009; Meidt et al. 2015; Lee et al. 2016; Corbelli et al. 2017; Kruijssen et al. 2019, among many others). The details of stellar feedback, its interaction with the ISM, the preconditions for star formation on cloud scales, and the dominant mechanism for cloud disruption all remain highly uncertain and areas of active theoretical research (e.g., Gatto et al. 2015; Jeffreson & Kruijssen 2018; Semenov et al. 2018).

A main way to constrain these physics is to measure the relative distributions of emission tracing different phases of the star formation process. At  $\sim 100$  pc resolution, GMCs, H II regions, and stellar clusters appear distinct from one another (e.g., Kawamura et al. 2009; Schruba et al. 2010, among many others, including a first

illustration of PHANGS-ALMA data in Kreckel et al. 2018). Figure 3 illustrates the dissimilar spatial distributions of H $\alpha$  and CO, which is thought to reflect the evolution of star-forming regions (e.g., Kawamura et al. 2009; Schruba et al. 2010; Kruijssen & Longmore 2014), e.g., from quiescent clouds to star-forming clouds to disrupted clouds. In the simplest terms, the fraction of clouds in different states maps to the timescales for a cloud to evolve through that state, though more sophisticated modeling techniques have been developed (e.g., Kruijssen et al. 2018), including treatment of gas flows along streamlines (e.g., Meidt et al. 2015; Egusa et al. 2017).

Despite many observations of H II regions and stellar clusters at  $\lesssim 100$  pc resolution, sensitive, wide area CO observations that isolate individual clouds have been scarcer and mostly focused on the Local Group. PHANGS-ALMA aims to change this situation, producing CO maps suitable to combine with H $\alpha$  maps, HST-based cluster catalogs, and integral field unit data to constrain the timescales and evolutionary sequence of GMCs across many environments. First work on this topic using the PHANGS-ALMA data appears in Kreckel et al. (2018), Schinnerer et al. (2019), and Chevance et al. (2020b,a).

This science goal requires PHANGS-ALMA to observe CO with high enough resolution to resolve the discrete distributions of molecular gas for comparison to high resolution maps of ionized gas and young stars. The required resolution varies, but is usually better than a few hundred parsecs (e.g., Chevance et al. 2020b). As for the first and second goal, the great diversity of galaxies observed in PHANGS-ALMA is instrumental for quantifying how the lifecycle of GMC evolution, star formation and feedback may vary with the galactic environment.

4. *Measure how the self-regulated, large scale structure of galaxy disks emerges from a medium made of individual clouds and star-forming regions.*

The disks of normal, star-forming galaxies at  $z = 0$  are often viewed as quasi-equilibrium systems. With that framework, vertical force balance is described by hydrostatic equilibrium with a dynamical pressure term balancing gravity (e.g., Elmegreen 1989; Blitz & Rosolowsky 2006; Ostriker et al. 2010, among many others). Radial equilibrium is often considered in terms of Toomre stability (e.g., Kennicutt 1989; Silk 1997; Thompson et al. 2005).

Past observations testing these models mostly had  $\sim$ kpc resolution, i.e., resolving galaxy disks but not breaking emission into individual star-forming regions (e.g., Martin & Kennicutt 2001; Wong & Blitz 2002; Boissier et al. 2003; Leroy et al. 2008; Colombo et al. 2018). Though simulations have explored how self-regulation emerges from a chaotic, high resolution view of the ISM spanning molecular clouds to galactic disks

(e.g., Kim et al. 2013; Orr et al. 2018), few observations had a comparable dynamic range. As a result, we lack a clear measurement of how the physical effects thought to be essential for self regulation — turbulent motions (e.g., Federrath et al. 2010; Padoan et al. 2016), supported by dynamical forces due to the galactic potential (e.g., Meidt et al. 2018, 2020), and gravitational collapse (e.g., Vazquez-Semadeni 1994; Ibáñez-Mejía et al. 2016) — relate to one another as a function of scale.

Concretely, PHANGS–ALMA aims to assess force (pressure) balance in the radial and vertical directions as a function of scale. This will place strong observational constraints on the dynamical state of the ISM and the scales on which the self-regulation of galactic disks sets in. First work on this topic using the PHANGS–ALMA data appears in Sun et al. (2020a).

This goal requires high physical resolution to break the molecular gas into individual clouds, high spectral resolution to assess the kinetic energy and other motions, and the inclusion of short-spacing data to allow studies that span a broad range of spatial scales.

### 5. Measure the motions, flows, and organization of cold gas in galaxies at $\sim 100$ –1,000 pc scales.

At  $\sim 100$  pc resolution, CO maps of massive disk galaxies reveal a highly structured medium (e.g., Schinnerer et al. 2013; Hirota et al. 2018; Sun et al. 2018; Schruba et al. 2021). Many galaxies show strikingly well-defined features associated with gas flows along bars, gas in spiral arms, and “feathers” associated with spiral arms (e.g., Lynds 1970; La Vigne et al. 2006; Corder et al. 2008; Schinnerer et al. 2017).

This structure is not captured either in low resolution studies of galaxy disks or GMC studies, which treat the gas as individual units. The last main goal of PHANGS–ALMA is to quantitatively characterize the structure and motions of the gas on scales bigger than a cloud but below the  $\sim$ kpc resolution at which disks appear relatively smooth.

Among metrics, we aim to quantify gas clumping (Leroy et al. 2013a), concentration into spiral arms (e.g., Foyle et al. 2010), and organization into filamentary structures (e.g., Jackson et al. 2010; Koch & Rosolowsky 2015; Zucker et al. 2018). Our goal is to approach these measurements in a quantitative, reproducible manner, similar to the techniques used to characterize density fields in studies of large scale structure (for a recent applications to CO maps see Grasha et al. 2018). These measurements will represent sophisticated benchmarks for simulations aiming to reproduce realistic cold gas structure or simulated CO emission.

The high-resolution kinematic information in PHANGS–ALMA also allows qualitatively new measurements related to these same phenomena. With high signal to noise, velocity resolution, and spatial resolution, we aim to measure streaming motions along spiral arms and bars, search for colliding gas flows,

signatures of gas inflow, and to identify when—and if—self-gravitating gas structures decouple from the global velocity field (e.g., Rosolowsky et al. 2003; Braine et al. 2018; Meidt et al. 2018; Herrera et al. 2020). The first application of the PHANGS data to measure detailed kinematic structure appears in Henshaw et al. (2020) and Lang et al. (2020).

Together, these five science goals inform the sample selection (§3) and observing strategy (§5) of PHANGS–ALMA. All can be met by a sensitive, wide area CO survey of a representative sample of galaxies that reaches cloud scale resolution. The rest of this paper describes our sample selection (§3), current best-estimate properties of the selected galaxies (§4), observation design and execution (§5), processing pipeline (§6), and the resulting data (§7).

## 3. SAMPLE SELECTION

PHANGS–ALMA aims to obtain cloud-scale CO maps for all ALMA-visible, massive, star-forming disk galaxies out to the near side of the Virgo Cluster. In this section, we discuss the motivation (§3.1), implementation (§3.2), and uncertainties associated with our selection strategy (§3.3). Further, we present several extensions to the main sample in §3.4.

### 3.1. Requirements

Our science goals (§2) require us to observe molecular gas at “cloud scales” and associate it with multi-wavelength signatures of star formation, feedback, and galactic structure. With that in mind, we selected our main sample according to the following criteria, summarized in Table 1.

1. **Close enough that  $1'' \leq 100$  pc.** We targeted galaxies with an estimated distance  $d \leq 17$  Mpc. Our core science goals require resolving molecular gas into individual cloud-sized resolution elements, a requirement associated with a fixed physical resolution of  $\sim 100$  pc.
2. **Not highly inclined.** We selected galaxies with inclination  $i < 75^\circ$ . This allows us to distinguish individual gas clouds and dynamical features. It also allows for clean association between CO and emission at other wavelengths, e.g., H $\alpha$  and the near-IR continuum.
3. **Visible to ALMA.** We considered targets with declination between  $\delta = -75^\circ$  and  $\delta = +25^\circ$ .
4. **Relatively massive.** We targeted galaxies with stellar mass  $\log_{10} M_\star [\mathrm{M}_\odot] \gtrsim 9.75$ . Our adopted mass cutoff translates to  $\sim 2$  times the mass of the LMC or M33 and lies about one order of magnitude below the knee in the local galaxy mass function,  $M^*$  where  $\log_{10} M^* [\mathrm{M}_\odot] \approx 10.65$  at  $z = 0$  (e.g., Baldry et al. 2008). In star-forming galaxies,

**Table 1.** PHANGS-ALMA Selection

Quantity	Value
Selection criteria for main sample <sup>a</sup>	
Declination	$-75^\circ < \delta < +25^\circ$
Inclination	$i < 75^\circ$
Distance	$d < 17 \text{ Mpc}$
$\log_{10} M_\star [\text{M}_\odot]$	$> 9.75$
$\log_{10} \text{SFR}/M_\star [\text{yr}^{-1}]$	$> -11$
Main sample selection <sup>b</sup>	75 galaxies
Extensions <sup>c</sup>	15 galaxies
Monte Carlo results (§3.3) <sup>d</sup>	
Without distance uncertainties	
... expected sample size	$82 \pm 4$
... false positive rate	$13\% \pm 3\%$
... false negative rate	$16\% \pm 3\%$
... correct selection rate	$87\% \pm 3\%$
With distance uncertainties	
... expected sample size	$76 \pm 6$
... false positive rate	$42\% \pm 4\%$
... false negative rate	$47\% \pm 5\%$
... correct selection rate	$58\% \pm 4\%$

<sup>a</sup>These are the selection criteria used to design the sample. As discussed in §3 and §4, some selected sample members no longer meet the selection criteria because we have improved our estimates of their properties. See the text for more details.

<sup>b</sup>The main sample is quoted as 74 galaxies in a number of our earlier works, because NGC 1068 met our selection criteria but was excluded due to previous archival mapping. As of the writing of this paper, we refer to the main sample as having 75 galaxies and we do include NGC 1068, which has new CO (2-1) 7-m+TP mapping (PI: M. Querejeta).

<sup>c</sup>These are members of PHANGS-ALMA known to not meet the selection criteria. Their scientific focus is contrasting with the properties and resolution of the main sample.

<sup>d</sup>Galaxies rejected “by eye” are entirely excluded from this calculation.

stellar mass correlates with star formation activity, gas fraction, molecular-to-atomic gas ratio, and metallicity (e.g., see review by Blanton & Moustakas 2009). By adopting this mass threshold, we aimed to capture a wide range of galaxy properties but to avoid focusing on low metallicity, low

mass galaxies where detecting CO can be a major challenge (e.g., Bolatto et al. 2013b; Hunt et al. 2015; Schruba et al. 2017).

At both low and high redshift, the shape of the galaxy mass function and the star-forming main sequence implies that most stars form in galaxies within one dex of  $M^*$  (e.g., Karim et al. 2011; Leslie et al. 2020), where  $M^*$  refers to the characteristic mass scale in the galaxy mass function ( $\log_{10} M^* \approx 10.6$  at  $z = 0$ , e.g., Weigel et al. 2016). We thus expect that our target mass range captures conditions representative of much of the secular build-up of galaxies.

**5. Actively star-forming.** We targeted galaxies with specific star formation rate  $\text{SFR}/M_\star > 10^{-11} \text{ yr}^{-1}$ . This selects galaxies close to the  $z = 0$  star-forming main sequence (e.g., Blanton & Moustakas 2009) and removes passive, non-star-forming galaxies that are less likely to have massive cold gas reservoirs. Our selection includes starburst galaxies with high  $\text{SFR}/M_\star$ . However such systems are rare in the  $z = 0$  universe (e.g., Sanders & Mirabel 1996) and they are mostly excluded by our distance cut.

### 3.2. Implementation

We worked on the selection of targets for the main PHANGS-ALMA sample from 2015 to 2017. The quantities that we used in the sample selection process were extracted from public databases or derived from public images. We caution that the property estimates that we used while making the sample selection may no longer represent our best estimates of certain galaxy properties for some targets. In particular, the distances to nearby galaxies can have large uncertainties and our best estimates for our targets’ distances have significantly evolved since the original selection process. We have also revised our approaches to estimate stellar mass and SFR since the original sample selection. We describe our current best estimates of galaxy properties in §4.

For selection, we implemented our sample criteria in the following way:

**1. Super-sample:** We considered objects classified as galaxies in LEDA, and required that they have either a deprojected rotation velocity  $> 120 \text{ km s}^{-1}$  or an absolute  $B$  magnitude  $M_B < -18 \text{ mag}$ . These represent less stringent cuts than those imposed on distance, mass, or specific star formation rate below. We do not expect that these criteria had a significant impact on the final sample selection.

**2. Distance:** Distance represents the dominant uncertainty in our selection. For the original selection, we used the median redshift-independent

distance from NED<sup>3</sup>. Our requirement of  $1'' \approx 100$  pc entails that our targets are too close for accurate Hubble flow distances, and very few of these nearby galaxies have high-quality, redshift-independent distances (see §4.2).

3. **Orientation:** We adopted positions and photometric inclination from HyperLEDA (Makarov et al. 2014). These inclinations can be uncertain, e.g., due to the uncertain handling of disk thickness in high-inclination cases or ambiguity in the geometry of the galaxy. They introduce a modest uncertainty into the selection.
4. **SFR and  $M_\star$  from WISE:** Our selection depended on stellar mass,  $M_\star$ , and specific star formation rate,  $\text{SFR}/M_\star$ . We estimated  $M_\star$  by carrying out photometry on WISE band 1 ( $3.4 \mu\text{m}$ ) images from the unWISE reprocessing (Lang 2014) of the WISE all-sky survey (Wright et al. 2010). We estimated the SFR using unWISE band 4 ( $22 \mu\text{m}$ ) images.

During our original sample selection, we translated the WISE band 1 luminosity to  $M_\star$  assuming a fixed mass-to-light ratio of  $\Upsilon_\star^{3.4} \approx 0.53 \text{ M}_\odot \text{ L}_\odot^{-1}$ . This is roughly consistent with Meidt et al. (2014), McGaugh & Schombert (2014), Querejeta et al. (2015), and other results from *Spitzer*'s S<sup>4</sup>G survey (Sheth et al. 2010). We converted from WISE band 4 to SFR using a factor  $C \approx 10^{-42.7}$  to convert from  $\nu L_\nu$  in  $\text{erg s}^{-1}$  to  $\text{M}_\odot \text{ yr}^{-1}$ . This agrees well with Kennicutt & Evans (2012) and Jarrett et al. (2013); for more details on both notation and appropriateness of this value see Leroy et al. (2019). We verified during sample selection that our WISE-based approach yielded SFRs consistent with estimates based on the IRAS Revised Bright Galaxy Survey (Sanders et al. 2003) and other common approaches to estimate the SFR (e.g., Kennicutt & Evans 2012). Since our adopted conversions between WISE luminosity and SFR or  $M_\star$  were linear, our original selection can be stated as a WISE band 1 luminosity cut combined with a WISE band 4 to WISE band 1 color cut.

The WISE band 1 photometry becomes uncertain at low Galactic latitude,  $b$ , due to the presence of foreground stars and the limited angular resolution of WISE. Our selection is therefore less accurate at low  $|b|$ . The adopted conversion factors also affected our sample selection. We did not include a UV or another “unobscured” term in the SFR estimate when selecting our target sample. This intro-

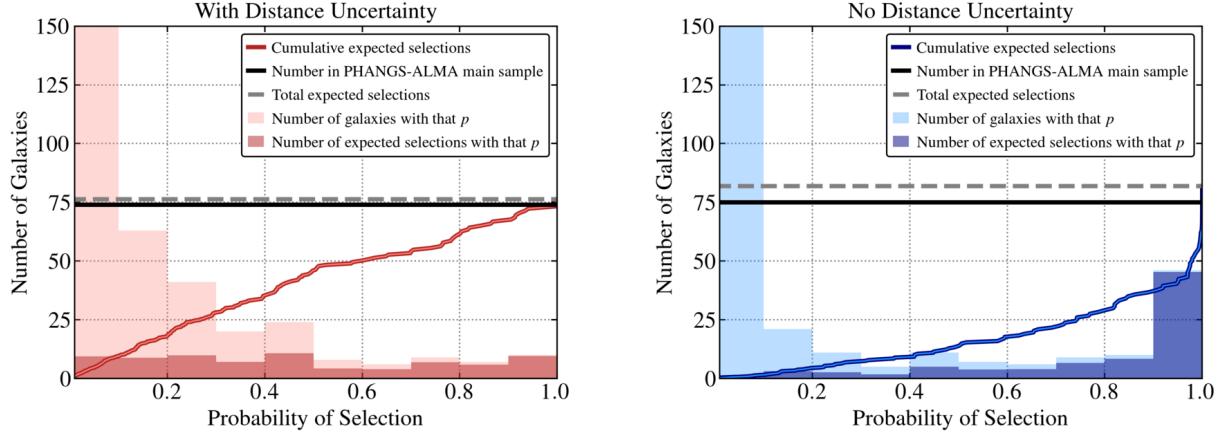
duced some bias against dwarf galaxies and other galaxies with little or no dust. We do not expect this to be a significant concern for relatively high mass “main sequence” galaxies. We also adopted a single  $\Upsilon_\star^{3.4}$ , which likely led us to overestimate the stellar mass of low mass, high  $\text{SFR}/M_\star$  galaxies (see §3.3). Most importantly, because we use stellar mass to select the sample, uncertainty in distance also affected this step of the selection.

5. **Rejection of incorrect selections:** After applying the above criteria, we identified a few cases where a target’s true inclination appeared to be nearly edge-on based on visual inspection of WISE and optical images. Visual inspection likewise revealed a few other targets with highly concentrated nuclear IR emission that is likely to be dominated by an AGN or a compact starburst. We rejected these targets as being unsuitable for wide-area CO mapping. Another handful of potential targets, usually at low Galactic latitude, are located directly behind bright Milky Way stars, making multi-wavelength analysis impractical. The overall list of manually excluded galaxies is ESO 138-010, ESO 494-026, IC 5201, NGC 1055, NGC 4802, NGC 6221, NGC 6875, PGC 18855, and PGC 54411.

This implementation yielded 75 primary PHANGS–ALMA targets. Of these, 18 were observed as part of several pilot programs, which we list in Table 2. Another 54 were observed as part of the ALMA Large Program “100,000 Molecular Clouds Across the Main Sequence: GMCs as the Drivers of Galaxy Evolution” (2017.1.00886.L, P.I. E. Schinnerer), which was carried out during Cycles 5 and 6. Two further galaxies, NGC 1365 and NGC 5236 (M83), meet our selection criteria and have been targeted for wide-area CO mapping by other programs (see Table 2). We include these galaxies in our sample for most science analysis, using a version of the archival data that we reprocessed using the PHANGS–ALMA pipeline. A final galaxy, NGC 1068, meets our criteria but was excluded from the Large Program due to the presence of archival CO (3-2) mapping (García-Burillo et al. 2014). New PHANGS CO (2-1) 7-m+TP mapping (PI: M. Querejeta) appear in this paper. We now formally include NGC 1068 in the main sample, bringing the main sample size to 75 galaxies.

We supplement this main sample with several extensions that relax one or more of the selection criteria, which we describe in §3.4, where we also note several public data sets with similar properties to PHANGS–ALMA. These extensions also leverage archival ALMA data, including the CO(2-1) observations of NGC 7793 presented in Grasha et al. (2018) and CO(2-1) observations of Centaurus A (NGC 5128) closely related to the CO(1-0) observations presented by Espada et al. (2019).

<sup>3</sup> The NASA/IPAC Extragalactic Database (NED) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.



**Figure 5. Sample selection expectations and uncertainties based on a Monte Carlo calculation.** Histograms show the number of galaxies as a function of the probability that they meet the main PHANGS-ALMA sample selection criteria. We take the physical property estimates and associated uncertainties for  $\sim 15,000$  local galaxies from Leroy et al. (2019) as “true”. Then we repeatedly realize random versions of galaxy properties to estimate the probability of selecting each galaxy. The *left* panel includes uncertainties on distance, the *right* panel considers only uncertainties on stellar mass, star formation rate, and inclination. Light histograms show the total number of galaxies with that probability of selection. Dark histograms show the number of galaxies that we would expect to select from that bin based on these probabilities. The (red/blue) solid and gray dashed lines show the resulting cumulative distribution function and total expected number of selected galaxies. The solid black line shows the actual number of galaxies (75) selected for the PHANGS-ALMA main sample. Our sample contains roughly the expected number of galaxies, but once distance uncertainties are accounted for there is substantial uncertainty associated with the exact selection (see also Figure 6).

Combining these extensions with the main sample, PHANGS-ALMA currently consists of 90 nearby galaxies observed at similar physical resolution in CO(2–1). We present the list of all targets, along with best estimates of their global properties in §4.

### 3.3. Accuracy and Uncertainty in Selection

Although the selection criteria for PHANGS-ALMA are quite simple, uncertainties in parameter estimation lead to uncertainty in the sample selection. Distance remains the primary driver of uncertainty for PHANGS-ALMA, since the Hubble flow does not yield high quality distances to galaxies closer than  $\sim 50$  Mpc (e.g., see figure 1 in Leroy et al. 2019). Any uncertainty in distance also affects the inferred stellar mass, which is another of our key selection criteria. Secondary uncertainties stem from how we estimate stellar mass and the star formation rate.

To assess the uncertainty associated with our sample selection, we carry out a Monte Carlo exercise. We begin with estimates of stellar mass, star formation rate, inclination, and distance to  $\sim 15,000$  nearby galaxies drawn from Leroy et al. (2019). These property estimates leverage GALEX and WISE photometry, with calibrations pinned to the properties of SDSS galaxies estimated by Salim et al. (2018). The distances are drawn from the Tully et al. (2009) Extragalactic Distance Database. We note that the stellar masses, SFRs, and distances in Leroy et al. (2019) should all be superior to those that we used for selection (§3.2).

The Monte Carlo exercise proceeds as follows:

1. We adopt the catalog values as the “true” values. We exclude the galaxies removed from the target list by hand from any calculations.
2. We randomly perturb the inclination, mass, and star formation rate of each galaxy according to their uncertainties. We adopted a  $\pm 5^\circ$  uncertainty for the inclination, consistent with Lang et al. (2020), and we cap the value at  $90^\circ$ . For the SFR and stellar mass, we draw the uncertainties from Leroy et al. (2019). These are typically 0.1 dex for  $M_*$  and 0.15 dex for the SFR; this primarily reflects uncertainty in the stellar mass-to-light ratio and conversion from IR and UV luminosity to SFR.
3. We randomly shift the distance, with the magnitude of the shift set by the uncertainty in the distance to each galaxy, which depends on the quality of the distance indicator<sup>4</sup>. We adjust the stellar mass and star formation rate to reflect the new distance.

<sup>4</sup> See section 2.2 in Leroy et al. (2019). Roughly, we adopt 0.03 dex uncertainty for TRGB distances, 0.06 dex for other quality distances, and 0.125 dex for other distances.

**Table 2.** ALMA Projects Processed as Part of PHANGS–ALMA

Cycle	Project Code	P.I.	Galaxies	Notes
PHANGS–ALMA Projects				
1	2013.1.00650.S	E. Schinnerer	1	NGC 0628, pilot project
3	2015.1.00925.S	G. Blanc	9 <sup>a</sup>	Pilot project
3	2015.1.00956.S	A. K. Leroy	8	Pilot project
5	2017.1.00886.L	E. Schinnerer	54	Large Program co-P.I.s A.K. Leroy, G. Blanc, A. Hughes, E. Rosolowsky, A. Schruba
5	2017.1.00392.S	G. Blanc	9 <sup>a</sup>	Pilot project completion
5	2017.1.00766.S	M. Chevance	7 <sup>b</sup>	Early-type extension
6	2018.1.00484.S	M. Chevance	7 <sup>b</sup>	Early-type extension completion
6	2018.1.01651.S	A. K. Leroy	9 <sup>a</sup>	Pilot project completion
6	2018.1.01321.S	C. Faesi	3	7-m and total power, very close galaxies
6	2018.A.00062.S	C. Faesi	5 <sup>c,d</sup>	7-m and total power, very close galaxies
7	2019.1.01235.S	C. Faesi	5 <sup>c,d</sup>	7-m and total power, very close galaxies completion
7	2019.2.00129.S	M. Querejeta	1	7-m and total power, NGC 1068
Archival CO(2–1) Data Processed with PHANGS–ALMA				
1	2013.1.01161.S	K. Sakamoto	2	NGC 1365 and NGC 5236 (M83) 12-m, 7-m, and total power
1	2013.1.00803.S	D. Espada	1	NGC 5128 <sup>c</sup> 12-m, 7-m, and total power
3	2015.1.00782.S	K. Johnson	1	NGC 7793 <sup>d</sup> 12-m; see <a href="#">Grasha et al. (2018)</a>
5	2015.1.00121.S	K. Sakamoto	1	NGC 5236 (M83) 12-m, 7-m, and total power
6	2016.1.00386.S	K. Sakamoto	1	NGC 5236 (M83) 12-m, 7-m, and total power

<sup>a</sup>These three programs targeted the same set of 9 total galaxies.

<sup>b</sup>These two programs targeted the same set of 7 total galaxies.

<sup>c</sup>NGC 5128 is Centaurus A. Projects 2018.A.0062.S and 2019.1.01235.S targeted NGC 5128 using the 7-m and total power antennas. Archival project 2013.1.00803.S targeted the galaxy with 12-m, 7-m, and total power observations, but the CO(2–1) total power observations were not usable. See the closely related CO(1–0) observations in [Espada et al. \(2019\)](#).

<sup>d</sup>Projects 2018.A.0062.S and 2019.1.01235.S targeted NGC 7793 using the 7-m and total power antennas. Archival project 2015.1.00782.S observed this galaxy using only the 12-m antennas.

4. For each galaxy and each realization, we check whether the galaxy’s new properties would qualify for our selection.
5. For each realization of the full sample, we check how many false positives and false negatives have been created by perturbing our best estimates of the galaxy properties.

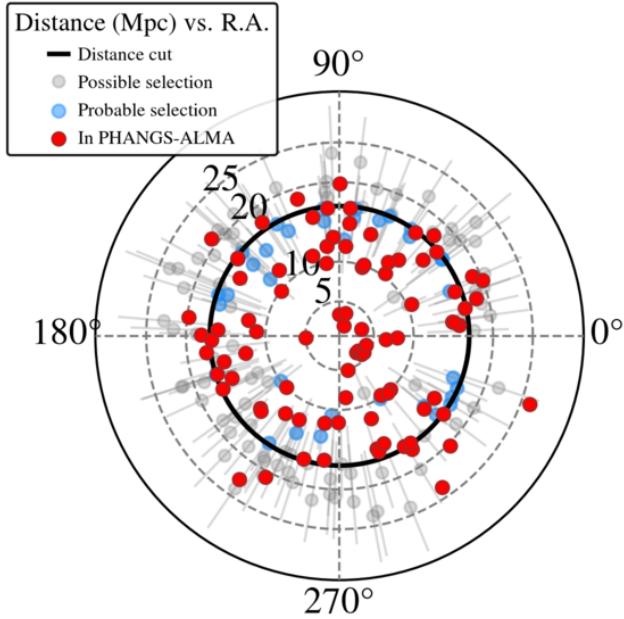
To do this, we first record if each set of true galaxy properties meets our selection criteria. Then we note whether each true selection would meet our criteria after perturbing the galaxy properties. The number of galaxies not selected but that have “true” properties that meet our selection criteria establishes the false negative rate.

To establish the false positive rate, we note how many selected sample members in the random realization would not have been selected if we used their “true” properties.

We repeat the exercise twice, each time using 10,000 realizations. In the first case, we impose distance uncertainties. In the second case we skip step #3, and only consider uncertainties unrelated to distance.

Based on the number of times a galaxy is selected over all 10,000 realizations, we assign each target a probability,  $p$ , of meeting our selection criteria. Figures 5, 6, and 7 visualize the results of this calculation, which we also summarize in the second part of Table 1.

Figure 5 shows histograms of the number of local galaxies that have probability  $p$  of matching our selec-



**Figure 6. Sample selection and location.** Right ascension and current best-estimate distance of PHANGS-ALMA targets (red points) and candidate targets that we did not select (blue and gray points). The overdensity of points towards  $180^\circ$  reflects the presence of the Virgo cluster. Blue points show targets close to meeting our selection criteria at their present distance; given the uncertainty in their stellar mass, SFR, and inclination we estimate that they are  $\gtrsim 50\%$  likely to meet our selection criteria. Gray points show galaxies where, given the uncertainty in distance, the targets are between 50% and 10% likely to have true values that meet our selection criteria. Red points show the actual PHANGS-ALMA sample. The dashed gray concentric circles indicate distances of 5, 10, 20 and 25 Mpc; our distance cutoff at 17 Mpc is indicated with a thick black line. Our selection did a good job of picking out all good candidates within  $\lesssim 10$  Mpc. The well-known uncertainties in distances to nearby galaxies leads to uncertainty in selection near our distance cutoff. Figure 5 gives another view of the uncertainty in selection.

tion. The total number of galaxies with  $p$  appears as a light shaded region. The expected number of selections in that bin, which is the sum of all  $p$  in that bin, appears as a dark shaded region. For example, 20 galaxies with a 5% chance of selection yield 1 expected selection. The cumulative distribution function (CDF) of selections ap-

pears as a solid line. This line shows the total number of galaxies with probability of selection  $< p$  expected to be selected. The total number of predicted selections, i.e., the last point in the CDF, appears as a dashed gray line. For comparison, the solid black line shows the total number of PHANGS-ALMA “main sample” selections, i.e., our total number of targets less the 15 extension galaxies that we know do not meet our original selection function.

There are several results in Figure 5 and Table 1 that are worth noting.

**PHANGS-ALMA selects about the expected number of galaxies:** First, the agreement between the dashed gray and black lines shows that our selected sample has about the expected size. Based on our Monte Carlo calculations, there should be  $\sim 75$  galaxies that meet our selection criteria. Specifically, as listed in Table 1, we select 76 and 82 galaxies on average, depending on whether we randomize the distance. We selected 75 targets for PHANGS-ALMA. In good agreement with this, without any Monte Carlo calculation, the [Leroy et al. \(2019\)](#) catalog yields 79 galaxies that meet our selection criteria.

**Distance represents the dominant source of uncertainty in sample selection:** Distance uncertainties are included in the left panel but not the right one of Figure 5. When including distance uncertainties, far fewer galaxies have high probabilities. We also expect many of our selections to be uncertain, which is reflected by their modest  $p$ . This demonstrates that distance uncertainties can easily shift galaxies into or out of our sample. By contrast, in the right panel, with no distance uncertainties, the sample selection appears clean, with most selected targets having high probabilities and relatively few ambiguous cases. Table 1 shows the same result. When distance uncertainties are included, the false positive and false negative rates are both much higher than in the case without distance uncertainties.

**Distance uncertainties lead to high false positive and false negative rates:** The uncertainty in distance leads to both a high false positive rate and a high false negative rate. Based on the Monte Carlo exercise, we estimate a false positive rate of  $\sim 40\%$  and a false negative rate of  $\sim 50\%$  (Table 1) when the distance uncertainty is fold into the model. This appears in the left panel of Figure 5 as a substantial contribution to the sample from bins with low  $p$ . A large fraction of our sample consists of galaxies with moderate  $p$  values, i.e., relatively uncertain selections. This is an unavoidable result of selecting local galaxies on mass and distance. It implies that as distance estimates improve, some of our targets will no longer meet our selection criteria, while other targets, not originally selected, will meet our criteria.

Figure 6 shows the location of our selected galaxies along with “possible” and “probable” selections. Here a

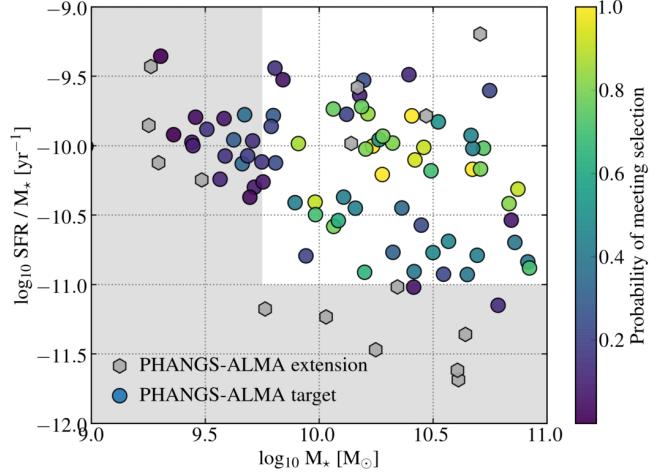
“possible” selection means  $0.5 \gtrsim p > 0.1$  (gray dots) with distance uncertainties. A “probable” selection refers to a galaxy with  $p > 0.5$  (blue dots) without distance uncertainties. We show 106 possible candidates and 31 probable candidates. As expected, the figure shows that almost all of these probable and possible selections hover near the distance cutoff of 17 Mpc. Put another way, there are a significant ( $> 100$ ) number of galaxies in the local Universe that have at least a moderate probability of having true properties that fulfil the PHANGS–ALMA selection criteria but were not included in our original main sample due to uncertainties in the estimation of nearby galaxy properties.

**Probability of re-selection and mass-to-light ratio for low mass galaxies:** In Figure 7, we examine the probability that the actual PHANGS–ALMA targets would be re-selected based on the Leroy et al. (2019) physical parameter estimates. We plot all PHANGS–ALMA targets in  $SFR/M_\star$  versus  $M_\star$  space and highlight our selection criteria using a shaded gray region. The color of each point indicates  $p$  for that galaxy from the Monte Carlo exercise above. Targets from the extension programs are also plotted in the figure as gray points.

Overall, the figure shows that our high mass, high SFR targets have a reasonably high chance of re-selection. However, the figure shows  $\sim 20$  low  $M_\star$  galaxies with low re-selection probability in the upper left part of the plot. The low  $M_\star$  and low re-selection probability for these targets reflect differences in how we estimated stellar mass during selection and the method used in Leroy et al. (2019), which is similar to what we use in §4. Our selection (§3.2) adopted a fixed WISE band 1 mass-to-light ratio. The Leroy et al. (2019) values used in this plot and the Monte Carlo analysis adopt a variable mass-to-light ratio. Their calibration yields moderately lower masses for the same WISE band 1 luminosity for low mass, high  $SFR/M_\star$  galaxies. Concretely, the median WISE1 mass-to-light ratio drops from  $\Upsilon_*^{3.4} \approx 0.5 M_\odot L_\odot^{-1}$  during selection to  $\Upsilon_*^{3.4} \approx 0.35 M_\odot L_\odot^{-1}$  on average in our current estimates. In practice, this highlights a modest systematic uncertainty in our selection by stellar mass, above and beyond the distance uncertainty. With our current best approach for mass estimation, PHANGS–ALMA actually extends down to  $\log_{10} M_\star [M_\odot] \approx 9.5$ .

**Notable Omissions:** This exercise revealed two clear omissions that appear almost certain to meet our criteria, NGC 3344 and NGC 3368. Another galaxy, NGC 4984, also appears likely to meet our criteria but the distance is quite uncertain. These three galaxies appear as blue dots near the 10 Mpc line in Figure 6.

**How to use this information:** For most PHANGS–ALMA users, the uncertainty in sample selection will have little impact. We adopt a simple selection function



**Figure 7. Probability of re-selection for PHANGS–ALMA targets.** Probability that actual PHANGS–ALMA targets would be re-selected, based on our Monte Carlo calculations and galaxy property estimates by Leroy et al. (2019). The gray region illustrates values excluded by our nominal selection criteria. The color of each point shows the fraction of realizations it is re-selected during the Monte Carlo calculations based on distance, inclination, mass, and specific star formation rate. Our high mass, high SFR targets usually meet our selection criteria. The variable WISE band 1 mass-to-light ratio used in Leroy et al. (2019) and adopted in §4 leads us to reassess some of our lower mass targets as likely having  $\log_{10} M_\star [M_\odot] \approx 9.5–9.75$ . Otherwise, most of deviation from  $p = 1$  reflects the persistent uncertainty in the distances to nearby galaxies.

and implement it in a reasonable way. Even if the selections are uncertain, we do not expect any important biases due to this uncertainty. For most applications, using all available data will represent a satisfactory approach. In these cases, the most important implication of this section is that one should always adopt the latest integrated galaxy parameter estimates. As our knowledge of distances to local galaxies improves, our estimates of their properties can change dramatically.

For those interested in a more rigorous sample selection, the uncertainty built into the sample selection implies that *one should re-construct the sample used for any science project using current best parameter estimates*. That is, best practice is to not consider the main sample as a fixed entity but instead to treat PHANGS–ALMA as a supersample from which rigorous subsamples using the most up-to-date parameter estimates can be drawn.

### 3.4. Extensions to the Main Sample

In addition to observations of our main sample, we have pursued several PHANGS–ALMA extension pro-

grams, which we also list in Table 2. These extensions adopt a similar observational setup as the core PHANGS-ALMA program but target galaxies that were missed or excluded by our initial selection.

The first major extension targets 7 early-type galaxies that have signs of active star formation (2017.1.00766.S, 2018.100484.S, P.I. M. Chevance) but specific star formation rate,  $\text{SFR}/\text{M}_\star$ , too low to qualify for our main sample. Molecular gas and star formation are present in a significant fraction of early-type galaxies (e.g., Young et al. 2011) and these systems represent a distinct environment for molecular cloud formation and evolution. This program pursues many of the same science goals as our main program with the goal of illuminating how the environments of early-type galaxies affect cold gas and star formation at cloud scales.

The other current major extension uses the ACA’s 7-m and total power facilities to target 8 galaxies with  $d \lesssim 4$  Mpc (2018.1.01321.S, 2018.A.00062.S, 2019.1.01235.S, P.I. C. Faesi). This sample serves two key goals. First, these local targets include lower mass galaxies. Galaxy mass and metallicity represent key drivers of ISM physics. However, the low surface brightness and small structure size of CO in low mass galaxies (e.g., see Hughes et al. 2013a; Druard et al. 2014; Faesi et al. 2018; Sun et al. 2018; Schruba et al. 2019) means that only observations of the nearest such systems are practical.

These very local targets also include several highly inclined, more massive galaxies, NGC 0253, NGC 4945, and the Circinus galaxy. These galaxies resemble other members of the main sample in their mass and star formation rate but were not selected due to inclination. The proximity of these targets means that telescopes with limited angular resolution still achieve high physical resolution, allowing cloud scale comparisons between molecular gas, far infrared emission, H I 21 cm emission, and radio continuum data. Furthermore, these tar-

gets are ideal for future ALMA follow-ups that achieve very high physical resolution over a wide area, building a bridge between PHANGS-ALMA and Milky Way and Local Group observations. Given the small set of such very nearby galaxies, we relaxed inclination cuts and added 3 galaxies. At present, we map these using only the 7-m and total power telescopes, producing data that effectively matches the properties of the other PHANGS-ALMA CO observations.

**Closely matched archival data:** The IRAM 30-m CO(2–1) map of M33 by Druard et al. (2014) (see also Gratier et al. 2010; Braine et al. 2018) closely resembles a PHANGS-ALMA map in terms of resolution and coverage of the whole area of active star formation. Though we do not treat it as a formal member of the PHANGS-ALMA sample, it serves as a valuable low-mass comparison galaxy for many PHANGS-ALMA analyses.

Several similar data sets exist targeting other CO transitions in nearby galaxies. For example, PAWS mapped CO(1–0) at 1'' resolution in M51 (Schinnerer et al. 2013), NOEMA has surveyed CO(1–0) emission from IC 342 at cloud scales (A. Schruba et al., in preparation), and CARMA (Schruba et al. 2021, see also Caldú-Primo & Schruba 2016) and the IRAM 30-m telescope (Nieten et al. 2006) have surveyed CO(1–0) in M31 at comparable resolution. The NANTEN CO(1–0) surveys of the Magellanic Clouds (Fukui et al. 1999; Mizuno et al. 2001) also achieve similar physical resolution as the PHANGS-ALMA maps. More recently, Kruijssen et al. (2019) presented an ALMA CO(1–0) map of NGC 300 at sufficient resolution to resolve GMCs (PI: A. Schruba). Similar to M33, we treat these as valuable complementary data sets but not members of the PHANGS-ALMA main sample.

#### 4. PROPERTIES OF THE OBSERVED SAMPLE

**Table 3.** PHANGS-ALMA Positions, Orientations, and Distances

Galaxy	$\alpha_{J2000}$	$\delta_{J2000}$	$v_{\text{LSR}}$	P.A.	$i$	$d$
			(km s <sup>-1</sup> )	(°)	(°)	(Mpc)
ESO097-013 <sup>X</sup>	14 <sup>h</sup> 13 <sup>m</sup> 09.9 <sup>s</sup>	-65°20'21"	430.3 ± 5.2	36.7 ± 10.0	64.3 ± 5.0	4.20 ± 0.77 (1)
IC 1954	3 <sup>h</sup> 31 <sup>m</sup> 31.1 <sup>s</sup>	-51°54'18"	1039.1 ± 5.0	63.4 ± 0.2	57.1 ± 0.7	12.80 ± 2.17 (2,3,4)
IC 5273	22 <sup>h</sup> 59 <sup>m</sup> 26.7 <sup>s</sup>	-37°42'10"	1286.0 ± 5.0	234.1 ± 2.0	52.0 ± 2.1	14.18 ± 2.14 (3,4)
IC 5332	23 <sup>h</sup> 34 <sup>m</sup> 27.5 <sup>s</sup>	-36°06'04"	699.3 ± 2.3	74.4 ± 10.0	26.9 ± 5.0	9.01 ± 0.40 (5)
NGC 0247 <sup>X</sup>	0 <sup>h</sup> 47 <sup>m</sup> 08.6 <sup>s</sup>	-20°45'38"	148.8 ± 1.6	167.4 ± 10.0	76.4 ± 5.0	3.71 ± 0.13 (6)
NGC 0253 <sup>X</sup>	0 <sup>h</sup> 47 <sup>m</sup> 33.1 <sup>s</sup>	-25°17'18"	235.4 ± 2.4	52.5 ± 10.0	75.0 ± 5.0	3.70 ± 0.12 (6)
NGC 0300 <sup>X</sup>	0 <sup>h</sup> 54 <sup>m</sup> 53.5 <sup>s</sup>	-37°41'04"	155.5 ± 7.3	114.3 ± 10.0	39.8 ± 5.0	2.09 ± 0.09 (6)
NGC 0628	1 <sup>h</sup> 36 <sup>m</sup> 41.7 <sup>s</sup>	+15°47'01"	650.8 ± 5.0	20.7 ± 1.0	8.9 ± 12.2	9.84 ± 0.63 (6)
NGC 0685	1 <sup>h</sup> 47 <sup>m</sup> 42.8 <sup>s</sup>	-52°45'43"	1346.6 ± 5.0	100.9 ± 2.8	23.0 ± 43.4	19.94 ± 3.01 (3,4)
NGC 1068 <sup>X</sup>	2 <sup>h</sup> 42 <sup>m</sup> 40.7 <sup>s</sup>	-0°00'48"	1130.1 ± 6.4	72.7 ± 10.0	34.7 ± 5.0	13.97 ± 2.11 (3,4)

**Table 3** *continued*

**Table 3** (*continued*)

Galaxy	$\alpha_{J2000}$	$\delta_{J2000}$	$v_{\text{LSR}}$	P.A.	$i$	$d$
			(km s $^{-1}$ )	( $^{\circ}$ )	( $^{\circ}$ )	(Mpc)
NGC 1087	2 <sup>h</sup> 46 <sup>m</sup> 25.2 <sup>s</sup>	-0 $^{\circ}$ 29'55"	1501.5 $\pm$ 5.0	359.1 $\pm$ 1.2	42.9 $\pm$ 3.9	15.85 $\pm$ 2.22 (7)
NGC 1097	2 <sup>h</sup> 46 <sup>m</sup> 18.9 <sup>s</sup>	-30 $^{\circ}$ 16'29"	1257.5 $\pm$ 5.0	122.4 $\pm$ 3.6	48.6 $\pm$ 6.0	13.58 $\pm$ 2.05 (3,4)
NGC 1313 <sup>X</sup>	3 <sup>h</sup> 18 <sup>m</sup> 15.7 <sup>s</sup>	-66 $^{\circ}$ 29'52"	451.2 $\pm$ 7.2	23.4 $\pm$ 10.0	34.8 $\pm$ 5.0	4.32 $\pm$ 0.17 (6)
NGC 1300	3 <sup>h</sup> 19 <sup>m</sup> 41.0 <sup>s</sup>	-19 $^{\circ}$ 24'40"	1545.4 $\pm$ 5.0	278.0 $\pm$ 1.0	31.8 $\pm$ 6.0	18.99 $\pm$ 2.86 (3,4)
NGC 1317	3 <sup>h</sup> 22 <sup>m</sup> 44.3 <sup>s</sup>	-37 $^{\circ}$ 06'14"	1930.5 $\pm$ 5.0	221.5 $\pm$ 2.9	23.2 $\pm$ 7.5	19.11 $\pm$ 0.85 (6)
NGC 1365	3 <sup>h</sup> 33 <sup>m</sup> 36.4 <sup>s</sup>	-36 $^{\circ}$ 08'25"	1613.3 $\pm$ 5.0	201.1 $\pm$ 7.5	55.4 $\pm$ 6.0	19.57 $\pm$ 0.78 (6)
NGC 1385	3 <sup>h</sup> 37 <sup>m</sup> 28.6 <sup>s</sup>	-24 $^{\circ}$ 30'04"	1476.8 $\pm$ 5.0	181.3 $\pm$ 4.8	44.0 $\pm$ 7.6	17.22 $\pm$ 2.60 (3,4)
NGC 1433	3 <sup>h</sup> 42 <sup>m</sup> 01.5 <sup>s</sup>	-47 $^{\circ}$ 13'19"	1057.4 $\pm$ 5.0	199.7 $\pm$ 0.3	28.6 $\pm$ 6.0	18.63 $\pm$ 1.84 (8)
NGC 1511	3 <sup>h</sup> 59 <sup>m</sup> 36.6 <sup>s</sup>	-67 $^{\circ}$ 38'02"	1331.0 $\pm$ 5.0	297.0 $\pm$ 2.1	72.7 $\pm$ 1.2	15.28 $\pm$ 2.26 (2)
NGC 1512	4 <sup>h</sup> 03 <sup>m</sup> 54.1 <sup>s</sup>	-43 $^{\circ}$ 20'55"	871.4 $\pm$ 5.0	261.9 $\pm$ 4.2	42.5 $\pm$ 6.0	18.83 $\pm$ 1.86 (8)
NGC 1546	4 <sup>h</sup> 14 <sup>m</sup> 36.3 <sup>s</sup>	-56 $^{\circ}$ 03'39"	1243.8 $\pm$ 5.0	147.8 $\pm$ 0.4	70.3 $\pm$ 0.6	17.69 $\pm$ 2.02 (7)
NGC 1559	4 <sup>h</sup> 17 <sup>m</sup> 36.6 <sup>s</sup>	-62 $^{\circ}$ 47'00"	1275.2 $\pm$ 5.0	244.5 $\pm$ 3.0	65.4 $\pm$ 8.4	19.44 $\pm$ 0.45 (9)
NGC 1566	4 <sup>h</sup> 20 <sup>m</sup> 00.4 <sup>s</sup>	-54 $^{\circ}$ 56'17"	1483.3 $\pm$ 5.0	214.7 $\pm$ 4.1	29.5 $\pm$ 10.6	17.69 $\pm$ 2.02 (7)
NGC 1637	4 <sup>h</sup> 41 <sup>m</sup> 28.2 <sup>s</sup>	-2 $^{\circ}$ 51'29"	698.9 $\pm$ 1.6	20.6 $\pm$ 10.0	31.1 $\pm$ 5.0	11.70 $\pm$ 1.01 (10)
NGC 1672	4 <sup>h</sup> 45 <sup>m</sup> 42.5 <sup>s</sup>	-59 $^{\circ}$ 14'50"	1318.3 $\pm$ 5.0	134.3 $\pm$ 0.4	42.6 $\pm$ 12.9	19.40 $\pm$ 2.93 (3,4)
NGC 1809	5 <sup>h</sup> 02 <sup>m</sup> 05.0 <sup>s</sup>	-69 $^{\circ}$ 34'05"	1290.4 $\pm$ 5.0	138.2 $\pm$ 8.9	57.6 $\pm$ 23.6	19.95 $\pm$ 5.63 (2)
NGC 1792	5 <sup>h</sup> 05 <sup>m</sup> 14.3 <sup>s</sup>	-37 $^{\circ}$ 58'50"	1175.9 $\pm$ 5.0	318.9 $\pm$ 0.9	65.1 $\pm$ 1.1	16.20 $\pm$ 2.44 (3,4)
NGC 2090	5 <sup>h</sup> 47 <sup>m</sup> 01.9 <sup>s</sup>	-34 $^{\circ}$ 15'02"	898.2 $\pm$ 5.0	192.5 $\pm$ 0.6	64.5 $\pm$ 0.2	11.75 $\pm$ 0.84 (11)
NGC 2283	6 <sup>h</sup> 45 <sup>m</sup> 52.8 <sup>s</sup>	-18 $^{\circ}$ 12'39"	821.9 $\pm$ 5.0	-4.1 $\pm$ 1.0	43.7 $\pm$ 3.6	13.68 $\pm$ 2.06 (3,4)
NGC 2566	8 <sup>h</sup> 18 <sup>m</sup> 45.6 <sup>s</sup>	-25 $^{\circ}$ 29'58"	1609.6 $\pm$ 5.0	312.0 $\pm$ 2.0	48.5 $\pm$ 6.0	23.44 $\pm$ 3.53 (7)
NGC 2775	9 <sup>h</sup> 10 <sup>m</sup> 20.1 <sup>s</sup>	+7 $^{\circ}$ 02'17"	1339.2 $\pm$ 5.0	156.5 $\pm$ 0.1	41.2 $\pm$ 0.6	23.15 $\pm$ 3.49 (3,4)
NGC 2835	9 <sup>h</sup> 17 <sup>m</sup> 52.9 <sup>s</sup>	-22 $^{\circ}$ 21'17"	867.3 $\pm$ 5.0	1.0 $\pm$ 1.0	41.3 $\pm$ 5.3	12.22 $\pm$ 0.93 (5)
NGC 2903	9 <sup>h</sup> 32 <sup>m</sup> 10.1 <sup>s</sup>	+21 $^{\circ}$ 30'03"	547.0 $\pm$ 5.0	203.7 $\pm$ 2.0	66.8 $\pm$ 3.1	10.00 $\pm$ 1.99 (2,3,4)
NGC 2997	9 <sup>h</sup> 45 <sup>m</sup> 38.8 <sup>s</sup>	-31 $^{\circ}$ 11'28"	1076.9 $\pm$ 5.0	108.1 $\pm$ 0.7	33.0 $\pm$ 9.0	14.06 $\pm$ 2.80 (7)
NGC 3059	9 <sup>h</sup> 50 <sup>m</sup> 08.2 <sup>s</sup>	-73 $^{\circ}$ 55'20"	1236.5 $\pm$ 5.0	-14.8 $\pm$ 2.9	29.4 $\pm$ 11.0	20.23 $\pm$ 4.04 (7)
NGC 3137	10 <sup>h</sup> 09 <sup>m</sup> 07.5 <sup>s</sup>	-29 $^{\circ}$ 03'51"	1086.6 $\pm$ 5.0	-0.3 $\pm$ 0.5	70.3 $\pm$ 1.2	16.37 $\pm$ 2.34 (7)
NGC 3239	10 <sup>h</sup> 25 <sup>m</sup> 04.9 <sup>s</sup>	+17 $^{\circ}$ 09'49"	748.3 $\pm$ 3.2	72.9 $\pm$ 10.0	60.3 $\pm$ 5.0	10.86 $\pm$ 1.05 (12)
NGC 3351	10 <sup>h</sup> 43 <sup>m</sup> 57.8 <sup>s</sup>	+11 $^{\circ}$ 42'13"	774.7 $\pm$ 5.0	193.2 $\pm$ 2.0	45.1 $\pm$ 6.0	9.96 $\pm$ 0.33 (6)
NGC 3489 <sup>X</sup>	11 <sup>h</sup> 00 <sup>m</sup> 18.6 <sup>s</sup>	+13 $^{\circ}$ 54'04"	692.1 $\pm$ 3.1	70.0 $\pm$ 10.0	63.7 $\pm$ 5.0	11.86 $\pm$ 1.63 (2,13)
NGC 3511	11 <sup>h</sup> 03 <sup>m</sup> 23.8 <sup>s</sup>	-23 $^{\circ}$ 05'12"	1096.7 $\pm$ 5.0	256.8 $\pm$ 0.8	75.1 $\pm$ 2.2	13.94 $\pm$ 2.10 (3,4)
NGC 3507	11 <sup>h</sup> 03 <sup>m</sup> 25.4 <sup>s</sup>	+18 $^{\circ}$ 08'08"	969.4 $\pm$ 5.0	55.8 $\pm$ 1.3	21.7 $\pm$ 11.3	23.55 $\pm$ 3.99 (2)
NGC 3521	11 <sup>h</sup> 05 <sup>m</sup> 48.6 <sup>s</sup>	-0 $^{\circ}$ 02'09"	798.0 $\pm$ 5.0	343.0 $\pm$ 0.6	68.8 $\pm$ 0.3	13.24 $\pm$ 1.96 (2)
NGC 3596	11 <sup>h</sup> 15 <sup>m</sup> 06.2 <sup>s</sup>	+14 $^{\circ}$ 47'13"	1187.9 $\pm$ 5.0	78.4 $\pm$ 1.0	25.1 $\pm$ 11.0	11.30 $\pm$ 1.03 (6)
NGC 3599 <sup>X</sup>	11 <sup>h</sup> 15 <sup>m</sup> 26.9 <sup>s</sup>	+18 $^{\circ}$ 06'37"	836.8 $\pm$ 20.2	41.9 $\pm$ 10.0	23.0 $\pm$ 5.0	19.86 $\pm$ 2.73 (2,13)
NGC 3621	11 <sup>h</sup> 18 <sup>m</sup> 16.3 <sup>s</sup>	-32 $^{\circ}$ 48'45"	724.3 $\pm$ 5.0	343.8 $\pm$ 0.3	65.8 $\pm$ 1.8	7.06 $\pm$ 0.28 (5)
NGC 3626	11 <sup>h</sup> 20 <sup>m</sup> 03.8 <sup>s</sup>	+18 $^{\circ}$ 21'25"	1470.7 $\pm$ 5.0	165.2 $\pm$ 2.0	46.6 $\pm$ 6.0	20.05 $\pm$ 2.34 (2,13)
NGC 3627	11 <sup>h</sup> 20 <sup>m</sup> 15.0 <sup>s</sup>	+12 $^{\circ}$ 59'29"	715.4 $\pm$ 5.0	173.1 $\pm$ 3.6	57.3 $\pm$ 1.0	11.32 $\pm$ 0.48 (6)
NGC 4207	12 <sup>h</sup> 15 <sup>m</sup> 30.4 <sup>s</sup>	+9 $^{\circ}$ 35'06"	606.6 $\pm$ 5.0	121.9 $\pm$ 2.0	64.5 $\pm$ 6.0	15.78 $\pm$ 2.34 (2)
NGC 4254	12 <sup>h</sup> 18 <sup>m</sup> 49.6 <sup>s</sup>	+14 $^{\circ}$ 25'00"	2388.2 $\pm$ 5.0	68.1 $\pm$ 0.5	34.4 $\pm$ 1.0	13.10 $\pm$ 2.01 (14)
NGC 4293	12 <sup>h</sup> 21 <sup>m</sup> 12.8 <sup>s</sup>	+18 $^{\circ}$ 22'57"	926.2 $\pm$ 5.0	48.3 $\pm$ 2.0	65.0 $\pm$ 6.0	15.76 $\pm$ 2.38 (7)
NGC 4298	12 <sup>h</sup> 21 <sup>m</sup> 32.8 <sup>s</sup>	+14 $^{\circ}$ 36'22"	1138.1 $\pm$ 5.0	313.9 $\pm$ 0.7	59.2 $\pm$ 0.8	14.92 $\pm$ 1.36 (5)
NGC 4303	12 <sup>h</sup> 21 <sup>m</sup> 54.9 <sup>s</sup>	+4 $^{\circ}$ 28'25"	1559.8 $\pm$ 5.0	312.4 $\pm$ 2.5	23.5 $\pm$ 9.2	16.99 $\pm$ 3.02 (7)
NGC 4321	12 <sup>h</sup> 22 <sup>m</sup> 54.9 <sup>s</sup>	+15 $^{\circ}$ 49'20"	1572.3 $\pm$ 5.0	156.2 $\pm$ 1.7	38.5 $\pm$ 2.4	15.21 $\pm$ 0.50 (11)
NGC 4424	12 <sup>h</sup> 27 <sup>m</sup> 11.6 <sup>s</sup>	+9 $^{\circ}$ 25'14"	447.4 $\pm$ 5.0	88.3 $\pm$ 2.0	58.2 $\pm$ 6.0	16.20 $\pm$ 0.69 (6)
NGC 4457	12 <sup>h</sup> 28 <sup>m</sup> 59.0 <sup>s</sup>	+3 $^{\circ}$ 34'14"	886.0 $\pm$ 5.0	78.7 $\pm$ 2.0	17.4 $\pm$ 6.0	15.10 $\pm$ 2.00 (2,13)
NGC 4459 <sup>X</sup>	12 <sup>h</sup> 29 <sup>m</sup> 00.0 <sup>s</sup>	+13 $^{\circ}$ 58'43"	1190.1 $\pm$ 11.0	108.8 $\pm$ 10.0	47.0 $\pm$ 5.0	15.85 $\pm$ 2.18 (2,13)
NGC 4476 <sup>X</sup>	12 <sup>h</sup> 29 <sup>m</sup> 59.1 <sup>s</sup>	+12 $^{\circ}$ 20'55"	1962.7 $\pm$ 1.3	27.4 $\pm$ 10.0	60.1 $\pm$ 5.0	17.54 $\pm$ 2.41 (2,13)
NGC 4477 <sup>X</sup>	12 <sup>h</sup> 30 <sup>m</sup> 02.2 <sup>s</sup>	+13 $^{\circ}$ 38'11"	1362.2 $\pm$ 33.1	25.7 $\pm$ 10.0	33.5 $\pm$ 5.0	15.76 $\pm$ 2.38 (7)

**Table 3** *continued*

**Table 3** (*continued*)

Galaxy	$\alpha_{J2000}$	$\delta_{J2000}$	$v_{\text{LSR}}$	P.A.	$i$	$d$
			(km s $^{-1}$ )	( $^{\circ}$ )	( $^{\circ}$ )	(Mpc)
NGC 4496A	12 <sup>h</sup> 31 <sup>m</sup> 39.3 <sup>s</sup>	+3 <sup>°</sup> 56'23"	1721.8 $\pm$ 5.0	51.1 $\pm$ 4.1	53.8 $\pm$ 3.5	14.86 $\pm$ 1.06 (11)
NGC 4535	12 <sup>h</sup> 34 <sup>m</sup> 20.3 <sup>s</sup>	+8 <sup>°</sup> 11'53"	1953.6 $\pm$ 5.0	179.7 $\pm$ 1.6	44.7 $\pm$ 10.8	15.77 $\pm$ 0.37 (11)
NGC 4536	12 <sup>h</sup> 34 <sup>m</sup> 27.1 <sup>s</sup>	+2 <sup>°</sup> 11'18"	1794.6 $\pm$ 5.0	305.6 $\pm$ 2.3	66.0 $\pm$ 2.9	16.25 $\pm$ 1.12 (6)
NGC 4540	12 <sup>h</sup> 34 <sup>m</sup> 50.9 <sup>s</sup>	+15 <sup>°</sup> 33'06"	1286.5 $\pm$ 5.0	12.8 $\pm$ 4.3	28.7 $\pm$ 28.7	15.76 $\pm$ 2.38 (7)
NGC 4548	12 <sup>h</sup> 35 <sup>m</sup> 26.5 <sup>s</sup>	+14 <sup>°</sup> 29'47"	482.7 $\pm$ 5.0	138.0 $\pm$ 2.0	38.3 $\pm$ 6.0	16.22 $\pm$ 0.38 (11)
NGC 4569	12 <sup>h</sup> 36 <sup>m</sup> 49.8 <sup>s</sup>	+13 <sup>°</sup> 09'46"	-225.6 $\pm$ 5.0	18.0 $\pm$ 2.0	70.0 $\pm$ 6.0	15.76 $\pm$ 2.38 (7)
NGC 4571	12 <sup>h</sup> 36 <sup>m</sup> 56.4 <sup>s</sup>	+14 <sup>°</sup> 13'02"	343.0 $\pm$ 5.0	217.5 $\pm$ 0.6	32.7 $\pm$ 2.1	14.90 $\pm$ 1.07 (15)
NGC 4579	12 <sup>h</sup> 37 <sup>m</sup> 43.5 <sup>s</sup>	+11 <sup>°</sup> 49'06"	1516.7 $\pm$ 5.0	91.3 $\pm$ 1.6	40.2 $\pm$ 5.6	21.00 $\pm$ 2.03 (16)
NGC 4596 <sup>X</sup>	12 <sup>h</sup> 39 <sup>m</sup> 55.9 <sup>s</sup>	+10 <sup>°</sup> 10'34"	1883.3 $\pm$ 7.5	120.0 $\pm$ 10.0	36.6 $\pm$ 5.0	15.76 $\pm$ 2.38 (7)
NGC 4654	12 <sup>h</sup> 43 <sup>m</sup> 56.6 <sup>s</sup>	+13 <sup>°</sup> 07'36"	1051.5 $\pm$ 5.0	123.2 $\pm$ 1.0	55.6 $\pm$ 5.9	21.98 $\pm$ 1.14 (11)
NGC 4689	12 <sup>h</sup> 47 <sup>m</sup> 45.6 <sup>s</sup>	+13 <sup>°</sup> 45'46"	1614.2 $\pm$ 5.0	164.1 $\pm$ 0.3	38.7 $\pm$ 2.7	15.00 $\pm$ 2.26 (2,3,4)
NGC 4694	12 <sup>h</sup> 48 <sup>m</sup> 15.0 <sup>s</sup>	+10 <sup>°</sup> 59'01"	1168.4 $\pm$ 5.0	143.3 $\pm$ 2.0	60.7 $\pm$ 6.0	15.76 $\pm$ 2.38 (7)
NGC 4731	12 <sup>h</sup> 51 <sup>m</sup> 01.2 <sup>s</sup>	-6 <sup>°</sup> 23'34"	1483.6 $\pm$ 5.0	255.4 $\pm$ 2.0	64.0 $\pm$ 6.0	13.28 $\pm$ 2.11 (7)
NGC 4781	12 <sup>h</sup> 54 <sup>m</sup> 23.8 <sup>s</sup>	-10 <sup>°</sup> 32'14"	1248.3 $\pm$ 5.0	290.0 $\pm$ 1.3	59.0 $\pm$ 3.8	11.31 $\pm$ 1.18 (7)
NGC 4826	12 <sup>h</sup> 56 <sup>m</sup> 43.6 <sup>s</sup>	+21 <sup>°</sup> 41'00"	409.7 $\pm$ 5.0	293.6 $\pm$ 1.2	59.1 $\pm$ 0.9	4.41 $\pm$ 0.19 (5)
NGC 4941	13 <sup>h</sup> 04 <sup>m</sup> 13.1 <sup>s</sup>	-5 <sup>°</sup> 33'06"	1116.0 $\pm$ 5.0	202.2 $\pm$ 0.6	53.4 $\pm$ 1.1	15.00 $\pm$ 5.00 (7)
NGC 4951	13 <sup>h</sup> 05 <sup>m</sup> 07.7 <sup>s</sup>	-6 <sup>°</sup> 29'38"	1176.1 $\pm$ 5.0	91.2 $\pm$ 0.5	70.2 $\pm$ 2.2	15.00 $\pm$ 4.19 (2)
NGC 4945 <sup>X</sup>	13 <sup>h</sup> 05 <sup>m</sup> 27.3 <sup>s</sup>	-49 <sup>°</sup> 28'04"	559.3 $\pm$ 2.7	43.8 $\pm$ 10.0	90.0 $\pm$ 5.0	3.47 $\pm$ 0.12 (6)
NGC 5042	13 <sup>h</sup> 15 <sup>m</sup> 31.0 <sup>s</sup>	-23 <sup>°</sup> 59'02"	1385.6 $\pm$ 5.0	190.6 $\pm$ 0.8	49.4 $\pm$ 8.6	16.78 $\pm$ 2.53 (3,4)
NGC 5068	13 <sup>h</sup> 18 <sup>m</sup> 54.7 <sup>s</sup>	-21 <sup>°</sup> 02'19"	667.2 $\pm$ 5.0	342.4 $\pm$ 3.2	35.7 $\pm$ 10.9	5.20 $\pm$ 0.22 (5)
NGC 5128	13 <sup>h</sup> 25 <sup>m</sup> 27.6 <sup>s</sup>	-43 <sup>°</sup> 01'09"	549.5 $\pm$ 5.7	32.2 $\pm$ 10.0	45.3 $\pm$ 5.0	3.69 $\pm$ 0.13 (6)
NGC 5134	13 <sup>h</sup> 25 <sup>m</sup> 18.5 <sup>s</sup>	-21 <sup>°</sup> 08'03"	1749.1 $\pm$ 5.0	311.6 $\pm$ 2.0	22.7 $\pm$ 6.0	19.92 $\pm$ 2.69 (7)
NGC 5236	13 <sup>h</sup> 37 <sup>m</sup> 00.9 <sup>s</sup>	-29 <sup>°</sup> 51'56"	509.4 $\pm$ 2.1	225.0 $\pm$ 10.0	24.0 $\pm$ 5.0	4.89 $\pm$ 0.18 (6)
NGC 5248	13 <sup>h</sup> 37 <sup>m</sup> 32.0 <sup>s</sup>	+8 <sup>°</sup> 53'07"	1163.0 $\pm$ 5.0	109.2 $\pm$ 3.5	47.4 $\pm$ 16.3	14.87 $\pm$ 1.32 (7)
NGC 5530	14 <sup>h</sup> 18 <sup>m</sup> 27.3 <sup>s</sup>	-43 <sup>°</sup> 23'18"	1183.2 $\pm$ 5.0	305.4 $\pm$ 1.0	61.9 $\pm$ 2.6	12.27 $\pm$ 1.85 (3,4)
NGC 5643	14 <sup>h</sup> 32 <sup>m</sup> 40.8 <sup>s</sup>	-44 <sup>°</sup> 10'29"	1191.3 $\pm$ 5.0	318.7 $\pm$ 2.0	29.9 $\pm$ 6.0	12.68 $\pm$ 0.54 (6)
NGC 6300	17 <sup>h</sup> 16 <sup>m</sup> 59.5 <sup>s</sup>	-62 <sup>°</sup> 49'14"	1102.1 $\pm$ 5.0	105.4 $\pm$ 2.3	49.6 $\pm$ 5.8	11.58 $\pm$ 1.75 (3,4)
NGC 6744	19 <sup>h</sup> 09 <sup>m</sup> 46.1 <sup>s</sup>	-63 <sup>°</sup> 51'27"	832.3 $\pm$ 5.0	14.0 $\pm$ 0.2	52.7 $\pm$ 2.2	9.39 $\pm$ 0.42 (5)
NGC 7456	23 <sup>h</sup> 02 <sup>m</sup> 10.3 <sup>s</sup>	-39 <sup>°</sup> 34'10"	1192.3 $\pm$ 5.0	16.0 $\pm$ 2.9	67.3 $\pm$ 4.3	15.70 $\pm$ 2.33 (2)
NGC 7496	23 <sup>h</sup> 09 <sup>m</sup> 47.3 <sup>s</sup>	-43 <sup>°</sup> 25'40"	1639.2 $\pm$ 5.0	193.7 $\pm$ 4.2	35.9 $\pm$ 6.0	18.72 $\pm$ 2.82 (3,4)
NGC 7743 <sup>X</sup>	23 <sup>h</sup> 44 <sup>m</sup> 21.1 <sup>s</sup>	+9 <sup>°</sup> 56'02"	1687.3 $\pm$ 5.4	86.2 $\pm$ 10.0	37.1 $\pm$ 5.0	20.32 $\pm$ 2.80 (2,13)
NGC 7793 <sup>X</sup>	23 <sup>h</sup> 57 <sup>m</sup> 49.8 <sup>s</sup>	-32 <sup>°</sup> 35'28"	222.1 $\pm$ 2.3	290.0 $\pm$ 10.0	50.0 $\pm$ 5.0	3.62 $\pm$ 0.15 (6)

NOTE— X — extension member; centers from Salo et al. (2015), Jarrett et al. (2003), or LEDA (Paturel et al. 2003; Makarov et al. 2014); orientations and velocities from Lang et al. (2020) or Sheth et al. (2010) or LEDA; distance reference key: 1—Karachentsev et al. (2004) 2—Tully et al. (2016) 3—Shaya et al. (2017) 4—Kourkchi et al. (2020) 5—Anand et al. (2021) 6—Tully et al. (2009) 7—Kourkchi & Tully (2017) 8—Anand et al. (2021) 9—Huang et al. (2020) 10—Leonard et al. (2003) 11—Freedman et al. (2001) 12—Barbarino et al. (2015) 13—Tonry et al. (2001) 14—Nugent et al. (2006) 15—Pierce et al. (1994) 16—Ruiz-Lapuente (1996)

**Table 4.** PHANGS-ALMA Physical Properties

Galaxy	$\log_{10} M_{\star}$	Src	$R_e$	$l_{\star}$	$\log_{10} \text{SFR}$	Src	$\log_{10} L_{\text{CO}}$	Corr.	$\log_{10} M_{\text{HI}}$
	( $M_{\odot}$ )		(kpc)	(kpc)	( $M_{\odot}$ yr $^{-1}$ )		(K km s $^{-1}$ pc $^2$ )		( $M_{\odot}$ )
NGC 0247 <sup>X</sup>	9.53	I	5.0	3.3	-0.75	FUVW4	6.79	1.42	9.24
NGC 0253 <sup>X</sup>	10.64	I	4.7	2.8	0.70	FUVW4	8.96	1.00	9.33
NGC 0300 <sup>X</sup>	9.27	I	2.0	1.3	-0.82	FUVW4	6.61	1.50	9.32

**Table 4** *continued*

**Table 4** (*continued*)

Galaxy	$\log_{10} M_\star$	Src	$R_e$	$l_\star$	$\log_{10} \text{SFR}$	Src	$\log_{10} L_{\text{CO}}$	Corr.	$\log_{10} M_{\text{HI}}$
	( $M_\odot$ )		(kpc)	(kpc)	( $M_\odot \text{ yr}^{-1}$ )		(K km s $^{-1}$ pc $^2$ )		( $M_\odot$ )
NGC 0628	10.34	I	3.9	2.9	0.24	FUVW4	8.41	1.73	9.70
NGC 0685	10.07	I	5.0	3.1	-0.38	W4ONLY	7.87	1.25	9.57
NGC 1068 <sup>x</sup>	10.91	I	0.9	7.3	1.64	FUVW4	9.23	1.30	9.06
NGC 1097	10.76	I	2.6	4.3	0.68	FUVW4	8.93	1.31	9.61
NGC 1087	9.94	I	3.2	2.1	0.11	FUVW4	8.32	1.06	9.10
NGC 1313 <sup>x</sup>	9.26	I	2.5	2.1	-0.14	FUVW4	...	...	9.28
NGC 1300	10.62	I	6.5	3.7	0.07	FUVW4	8.50	1.28	9.38
NGC 1317	10.62	W	1.8	2.4	-0.32	FUVW4	8.10	1.28	...
IC 1954	9.67	I	2.4	1.5	-0.44	FUVW4	7.78	1.10	8.85
NGC 1365	11.00	I	2.8	13.1	1.24	FUVW4	9.49	1.36	9.94
NGC 1385	9.98	I	3.4	2.6	0.32	FUVW4	8.37	1.09	9.19
NGC 1433	10.87	I	4.3	6.9	0.05	FUVW4	8.47	1.38	9.40
NGC 1511	9.92	I	2.4	1.7	0.35	FUVW4	8.22	1.09	9.57
NGC 1512	10.72	I	4.8	6.2	0.11	FUVW4	8.26	1.45	9.88
NGC 1546	10.37	I	2.2	2.1	-0.08	FUVW4	8.44	1.13	8.68
NGC 1559	10.37	I	3.9	2.4	0.60	NUVW4	8.66	1.11	9.52
NGC 1566	10.79	I	3.2	3.9	0.66	FUVW4	8.89	1.22	9.80
NGC 1637	9.95	I	2.8	1.8	-0.20	W4ONLY	7.98	1.10	9.20
NGC 1672	10.73	I	3.4	5.8	0.88	FUVW4	9.05	1.25	10.21
NGC 1809	9.77	I	4.5	2.4	0.76	NUVW4	7.49	4.24	9.60
NGC 1792	10.62	I	4.1	2.4	0.57	FUVW4	8.95	1.11	9.25
NGC 2090	10.04	W	1.9	1.7	-0.39	FUVW4	7.67	1.47	9.37
NGC 2283	9.89	W	3.2	1.9	-0.28	W4ONLY	7.69	1.16	9.70
NGC 2566	10.71	W	5.1	4.0	0.93	W4ONLY	9.06	1.13	9.37
NGC 2775	11.07	I	4.6	4.1	-0.06	FUVW4	8.40	1.29	8.65
NGC 2835	10.00	W	3.3	2.2	0.10	FUVW4	7.71	1.72	9.48
NGC 2903	10.64	I	3.7	3.5	0.49	FUVW4	8.76	1.18	9.54
NGC 2997	10.73	W	6.1	4.0	0.64	FUVW4	8.97	1.25	9.86
NGC 3059	10.38	W	5.0	3.2	0.38	W4ONLY	8.59	1.07	9.75
NGC 3137	9.88	W	4.1	3.0	-0.30	FUVW4	7.60	1.35	9.68
NGC 3239	9.18	I	3.1	2.0	-0.41	FUVW4	< 6.62 <sup>ul</sup>	1.54	9.16
NGC 3351	10.37	I	3.0	2.1	0.12	FUVW4	8.13	1.55	8.93
NGC 3489 <sup>x</sup>	10.29	I	1.3	1.4	-1.59	FUVW4	6.89	1.37	7.40
NGC 3511	10.03	I	4.4	2.4	-0.09	FUVW4	8.15	1.07	9.37
NGC 3507	10.40	I	3.7	2.3	-0.00	FUVW4	8.34	1.17	9.32
NGC 3521	11.03	I	3.9	4.9	0.57	FUVW4	8.98	1.18	9.83
NGC 3596	9.66	I	1.6	2.0	-0.52	NUVW4	7.81	1.13	8.85
NGC 3599 <sup>x</sup>	10.04	I	1.7	2.0	-1.35	FUVW4	< 6.70 <sup>ul</sup>	1.35	...
NGC 3621	10.06	W	2.7	2.0	-0.00	FUVW4	8.13	1.27	9.66
NGC 3626	10.46	I	1.8	2.1	-0.68	NUVW4	7.75	1.14	8.89
NGC 3627	10.84	I	3.6	3.7	0.59	FUVW4	8.98	1.16	9.09
NGC 4207	9.72	I	1.4	0.7	-0.72	FUVW4	7.71	1.03	8.58
NGC 4254	10.42	I	2.4	1.8	0.49	FUVW4	8.93	1.15	9.48
NGC 4293	10.52	I	4.7	2.8	-0.30	FUVW4	8.12	1.57	7.67
NGC 4298	10.04	I	3.0	1.6	-0.34	FUVW4	8.26	1.09	8.87
NGC 4303	10.51	I	3.4	3.1	0.73	FUVW4	9.00	1.40	9.67
NGC 4321	10.75	I	5.5	3.6	0.55	FUVW4	9.02	1.25	9.43

**Table 4** *continued*

**Table 4** (*continued*)

Galaxy	$\log_{10} M_\star$	Src	$R_e$	$l_\star$	$\log_{10} \text{SFR}$	Src	$\log_{10} L_{\text{CO}}$	Corr.	$\log_{10} M_{\text{HI}}$
	( $M_\odot$ )		(kpc)	(kpc)	( $M_\odot \text{ yr}^{-1}$ )		(K km s $^{-1}$ pc $^2$ )		( $M_\odot$ )
NGC 4424	9.93	I	3.7	2.2	-0.53	FUVW4	7.59	1.16	8.30
NGC 4457	10.42	I	1.5	2.2	-0.52	FUVW4	8.21	1.15	8.36
NGC 4459 <sup>X</sup>	10.68	W	2.1	3.3	-0.65	FUVW4	7.46	2.41	...
NGC 4476 <sup>X</sup>	9.81	W	1.2	1.2	-1.39	FUVW4	7.05	1.09	...
NGC 4477 <sup>X</sup>	10.59	W	2.1	2.1	-1.10	FUVW4	6.76	1.58	...
NGC 4496A	9.55	I	3.0	1.9	-0.21	FUVW4	7.55	1.15	9.24
NGC 4535	10.54	I	6.3	3.8	0.34	FUVW4	8.61	1.78	9.56
NGC 4536	10.40	I	4.4	2.7	0.53	FUVW4	8.62	1.06	9.54
NGC 4540	9.79	I	2.0	1.4	-0.78	FUVW4	7.69	1.16	8.44
NGC 4548	10.70	I	5.4	3.0	-0.28	FUVW4	8.16	2.00	8.84
NGC 4569	10.81	I	5.9	4.3	0.12	FUVW4	8.81	1.40	8.84
NGC 4571	10.10	I	3.8	2.0	-0.54	FUVW4	7.88	1.55	8.70
NGC 4579	11.15	I	5.4	4.4	0.33	FUVW4	8.79	1.38	9.02
NGC 4596 <sup>X</sup>	10.59	I	2.7	3.8	-0.96	FUVW4	6.72	1.83	...
NGC 4654	10.57	I	5.6	4.0	0.58	FUVW4	8.84	1.18	9.75
NGC 4689	10.24	I	4.7	3.0	-0.39	W4ONLY	8.22	1.19	8.54
NGC 4694	9.90	I	1.9	1.6	-0.81	FUVW4	7.41	1.30	8.51
NGC 4731	9.50	I	7.3	3.0	-0.22	FUVW4	7.29	2.52	9.44
NGC 4781	9.64	I	2.0	1.1	-0.32	FUVW4	7.82	1.05	8.94
NGC 4826	10.24	I	1.5	1.1	-0.69	FUVW4	7.79	1.28	8.26
NGC 4941	10.18	I	3.4	2.2	-0.35	FUVW4	7.80	1.27	8.49
NGC 4951	9.79	I	1.9	1.9	-0.46	FUVW4	7.65	1.22	9.21
NGC 4945 <sup>X</sup>	10.36	W	4.5	1.6	0.19	W4ONLY	8.77	0.97	8.92
NGC 5042	9.90	I	3.3	2.4	-0.22	FUVW4	7.69	1.84	9.29
NGC 5068	9.41	I	2.0	1.3	-0.56	FUVW4	7.26	1.38	8.82
NGC 5134	10.41	I	2.9	2.1	-0.34	FUVW4	7.98	1.14	8.92
NGC 5128	10.97	W	4.7	4.1	0.09	FUVW4	8.40	0.98	8.43
NGC 5236	10.53	I	3.5	2.4	0.62	FUVW4	8.84	1.14	9.98
NGC 5248	10.41	I	3.2	2.0	0.36	FUVW4	8.77	1.14	9.50
ESO097-013 <sup>X</sup>	10.53	W	1.9	1.8	0.61	W4ONLY	8.42	1.40	9.81
NGC 5530	10.08	W	3.4	1.7	-0.48	W4ONLY	7.89	1.34	9.11
NGC 5643	10.34	W	3.5	1.6	0.41	W4ONLY	8.56	1.06	9.12
NGC 6300	10.47	W	3.6	2.1	0.29	W4ONLY	8.46	1.12	9.13
NGC 6744	10.72	W	7.0	4.8	0.38	FUVW4	8.27	2.75	10.31
IC 5273	9.73	I	2.5	1.3	-0.27	FUVW4	7.63	1.14	8.95
NGC 7456	9.65	I	4.4	2.9	-0.43	FUVW4	7.13	2.02	9.28
NGC 7496	10.00	I	3.8	1.5	0.35	FUVW4	8.33	1.15	9.07
IC 5332	9.68	I	3.6	2.8	-0.39	FUVW4	7.09	2.26	9.30
NGC 7743 <sup>X</sup>	10.36	I	2.9	1.9	-0.67	FUVW4	7.50	2.65	8.50
NGC 7793 <sup>X</sup>	9.36	I	1.9	1.1	-0.57	FUVW4	7.23	1.34	8.70

<sup>w</sup>We quote a  $5\sigma$  upper limit on  $L_{\text{CO}}$  constructed using a broad velocity window across the whole map.

NOTE—  $X$  — extension member;

A main goal of PHANGS-ALMA is to relate cloud scale gas properties, star formation timescales, and kinematics to the properties of the host galaxy and location within the galaxy (§2). To do this, we require estimates of galaxy properties. In this section, we report our current best estimate galaxy properties: orientation

(§4.1), distance (§4.2), stellar mass (§4.3), size (§4.4), star formation rate (§4.5), CO luminosity (§4.6), and H I masses (§4.7). Then we summarize the properties of the sample and show PHANGS-ALMA targets on two common scaling relations, the main sequence of star-forming galaxies and the size–mass relation (§4.8). Readers who

are only interested in the sample and not the provenance of the property estimates may wish to skip to §4.8.

We report the properties of the PHANGS–ALMA targets in Tables 3 and 4. These properties represent current best estimates. In many cases, these estimates have been derived or refined *after* sample selection, e.g., from rotation curve fitting using the PHANGS–ALMA data (Lang et al. 2020). As a result, they do not perfectly agree with those used for sample selection. In other cases, PHANGS–ALMA papers use several estimates that sometimes have different zero points and scales. We note the translation between these systems whenever possible. In each case, this section presents our preferred values for scientific analysis.

For stellar mass, size, star formation rate, CO luminosity calculation, we also make these estimates for a larger sample of 261 local galaxies that have CO maps. These include the PHANGS targets, the targets of the COMING survey (Sorai et al. 2019), the targets of the Nobeyama nearby galaxy atlas (Kuno et al. 2007), the targets of HERACLES (Leroy et al. 2009) and follow-up programs (A. Schruba et al. in preparation), and the targets of the JCMT Nearby Galaxy Legacy Survey (Wilson et al. 2012). In this section, we use this larger sample for methodology tests and a few comparisons. The measurements appear as blue dots in Figure 1. These data are treated exactly the same as the PHANGS–ALMA data when estimating stellar mass, SFR, and other properties.

#### 4.1. Orientation and Galaxy Center

We adopt position angles and inclinations from Lang et al. (2020). They use the CO kinematics derived from the PHANGS–ALMA data to constrain the position angle of each target. For 48 targets, they also obtain a kinematic fit for the inclination. In cases where the CO kinematics do not sufficiently constrain the inclination, Lang et al. (2020) identified preferred photometric estimates. These photometric orientation estimates come from the S<sup>4</sup>G analysis of Spitzer/IRAC 3.6 $\mu$ m imaging by Salo et al. (2015) when available, and from 2MASS NIR imaging work by Jarrett et al. (2003) when not available. Lang et al. (2020) also identify a preferred systemic velocity for each target. For targets not considered by Lang et al. (2020), we default to orientation parameters from S<sup>4</sup>G (Sheth et al. 2010; Muñoz-Mateos et al. 2015; Salo et al. 2015) when available and HyperLEDA (Paturel et al. 2003; Makarov et al. 2014) when not available. Whenever we become aware of kinematic-based estimates, we update our adopted orientations to reflect these. Based on comparison of the S<sup>4</sup>G and HyperLEDA orientation parameters, we adopt a  $\pm 5^\circ$  typical uncertainty for the inclination and a  $\pm 10^\circ$  typical uncertainty for the position angle (consistent with the uncertainty estimates by Lang et al. 2020, in the cases with well-measured orientations).

Following Lang et al. (2020), we adopt photometric centers from Salo et al. (2015) when available. These centers leverage sensitive near-infrared imaging and should accurately reflect the center of stellar mass in the galaxy. When these are not available, we adopt central positions from the 2MASS Large Galaxy Atlas (Jarrett et al. 2003). When neither are available, we use the optically defined central position from HyperLEDA or NED. The choice of photometric center generally matters at the level of a few arcseconds or less. We adopt a fiducial uncertainty of 1''.

#### 4.2. Distance

We adopt distance estimates from Anand et al. (2021). Their work compiles a mixture of literature estimates and new tip of the red giant branch (TRGB) distances based on *Hubble Space Telescope* (HST) observations. The new distances are derived from observations carried out as part of PHANGS–HST (Lee et al. 2021). We reproduce the Anand et al. (2021) distance estimates in Table 3, where we also note the original references for the literature distances. We recommend citing the original reference when adopting these distances.

In total, Anand et al. (2021) compile distance estimates and associated uncertainties for 117 galaxies, including all 90 PHANGS–ALMA targets. They evaluate the available distance estimates for each target, select the highest quality estimate, and assign an associated uncertainty. Approximately 40% of these distances come from high quality primary distance indicators, either TRGB- or Cepheid-based distances. Group-based distances, results from a numerical action method, and Tully–Fisher estimates account for most of the rest of the distances. A handful of distances come from other direct techniques, mostly surface brightness fluctuations. The numerical action method (e.g., Shaya et al. 2017) may be the least familiar of these. This method assigns distance based on a galaxy's position and velocity using a sophisticated three dimensional model of gravitationally induced flows in the local volume. It can be roughly thought of as a vastly-improved version of the Virgocentric flow-corrected Hubble flow distance.

The location of PHANGS–ALMA targets, combined with the distances from Anand et al. (2021), are shown in Figure 6 and we show the distribution of distances in Figure 16. Perhaps the main thing to see from these figures is that PHANGS–ALMA spans a large dynamic range in distances. This makes accounting for distance effects, e.g., by convolving data to a common physical scale before scientific analysis, crucial. The figures also show that, as expected, our targets cluster near our 17 Mpc distance cutoff with several targets beyond the nominal cutoff. This simply reflects the uncertainty in distances and larger volume present at large radius.

The distances for PHANGS galaxies provided by Anand et al. (2021) build on many time- and effort-intensive studies, and we also refer to the initial/original

studies when quoting the distances (Table 3). Also note PHANGS-ALMA has made heavy use of the distance compilations by the Extragalactic Distance Database (EDD; Tully et al. 2009), the closely related Cosmic-Flows projects (Courtois & Tully 2012; Tully et al. 2016), HyperLEDA (Paturel et al. 2003; Makarov et al. 2014), and NED. For non PHANGS-ALMA targets not considered by Anand et al. (2021), we utilize these databases following the method described by Leroy et al. (2019).

### 4.3. Stellar Mass

Stellar mass informs our selection and plays a wide-ranging role in the PHANGS-ALMA scientific analysis. The stellar plays a major role setting gravitational potential in the disk across most of the survey area. As a result, maps that trace stellar mass are crucial to define distinct dynamical environments in galaxies (e.g., Colombo et al. 2014b). For galaxies without high-quality optical spectroscopy (e.g., from PHANGS-MUSE), we also use stellar mass to predict the metallicity of a galaxy and to place the galaxy in the context of the larger galaxy population.

We use near-infrared maps, combined with a radially varying mass-to-light ratio,  $\Upsilon_*^{3.4}$ , to estimate stellar mass and stellar disk size in our targets<sup>5</sup>. Whenever possible, we use the high quality IRAC 3.6  $\mu\text{m}$  data from S<sup>4</sup>G (Sheth et al. 2010). When these are not available, we use WISE1 3.4  $\mu\text{m}$  maps from Leroy et al. (2019). The wavelength coverage of the two bands heavily overlap. We expect both to be dominated by light from old stars and minimally affected by dust extinction. Never the less dust emission can contaminate the band (e.g., Meidt et al. 2012; Simonian & Martini 2017) and the age of the stellar population still affects the mass-to-light ratio at these wavelengths (e.g., Bell & de Jong 2001). We prefer to use the IRAC imaging for practical reasons. Although the WISE maps cover every galaxy, the versions produced by Leroy et al. (2019) have 7.5'' spatial resolution, which presents a limitation when attempting to mask foreground stars. They also have poorer signal-to-noise ratio than the IRAC data, though this is less problematic since telescopes easily detect most of the emission in most targets. These mass calculations assume a Chabrier (2003) stellar initial mass function (via Bruzual & Charlot 2003; Salim et al. 2016, 2018).

In both cases, we apply the mass-to-light ratio prescription from Leroy et al. (2019), which we develop in more detail. This estimator combines UV, near-IR, and mid-IR emission to estimate the specific star formation rate, which is a strong driver of  $\Upsilon_*^{3.4}$ . The numerical estimate is calibrated to match the results of the GALEX-

SDSS-WISE Legacy Catalog (GSWLC; Salim et al. 2016, 2018) studying the full SDSS main galaxy sample. Adopting this mass-to-light ratio places our measurements on an approximately matched system with  $> 10,000$  other local galaxies and the full SDSS main galaxy sample studied by Salim et al. (2016, 2018). To carry out this estimate we also constructed matched 15'' resolution maps for WISE1, WISE3, WISE4, GALEX FUV, and GALEX NUV emission. In §4.5, we use these to help estimate the star formation rate of our targets. We discuss the exact procedure and compare the implied  $\Upsilon_*^{3.4}$  to other stellar mass estimates used in PHANGS in the next subsection.

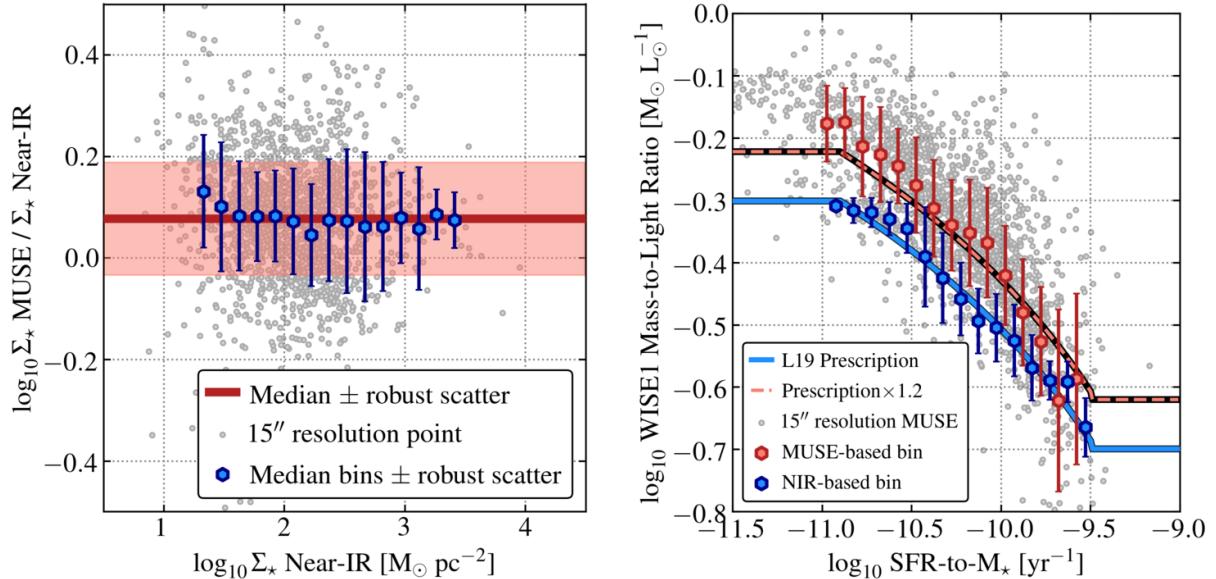
We process the S<sup>4</sup>G IRAC images before using them to measure stellar masses or sizes. First, we apply the S<sup>4</sup>G masks described by Muñoz-Mateos et al. (2015) outside the central 10'' of each target. These stars remove likely foreground stars and artifacts from the image. We do not mask the central part of each galaxy to avoid blanking bright nuclei; even when these regions are affected by saturation, we need them in our calculations. Then we mask some remaining visible artifacts and stars by hand. These are usually away from the galaxy but relevant to background subtraction. Then, we estimate and subtract a local background. This is usually calculated using the mode in a circular annulus 1.5–3  $r_{25}$  in radius, but we adjusted the range based on visual inspection. Finally, we interpolated to fill in masked pixels with the median value at that galactocentric radius (i.e., the value suggested by the galaxy's radial profile). After that, we convolved the masked, background subtracted, interpolated IRAC images to 7.5'' resolution for processing in a way that matched the WISE data.

Leroy et al. (2019) describe the background subtraction and masking for the WISE images. For stellar mass estimates, we begin with the WISE1 sky-subtracted images. We apply the Gaia- and 2MASS-based star masks described in Leroy et al. (2019), but only outside 0.5  $r_{25}$ . Then, we manually inspected each image and blanked any remaining bright stars or image artifacts by hand. This caught some low-level background artifacts missed by the automatic masking in Leroy et al. (2019) and also ensured that we only blanked objects that were foreground stars in the galaxy itself. Then, as with the IRAC data, we interpolate across masked regions using the median of pixels at the same galactocentric radius. A handful of cases lacked IRAC images and are so heavily covered by foreground stars that WISE struggles to recover the galaxy. In these cases, we used only the median profile.

#### 4.3.1. Mass-to-Light Ratio Estimates

We convert from IRAC 3.6  $\mu\text{m}$  or WISE 3.4  $\mu\text{m}$  intensity to stellar mass surface density,  $\Sigma_*$ , by multiplying the measured intensity with a local estimate of the mass-

<sup>5</sup> We adopt the same mass-to-light ratio,  $\Upsilon_*^{3.4}$  in units of  $\text{M}_\odot \text{ L}_\odot^{-1}$ , for WISE1 at 3.4  $\mu\text{m}$  and IRAC1 at 3.6  $\mu\text{m}$ . See Leroy et al. (2019) for details on the conversions.



**Figure 8. Stellar mass estimates and near-IR mass-to-light ratios compared to results from PHANGS–MUSE.** Comparison between our near infrared-based stellar mass estimates and stellar masses based on full spectral fitting of PHANGS–MUSE data in the wavelength range 4850–7000Å for 19 targets (E. Emsellem et al. in preparation). The *left* panel shows the ratio between MUSE-based  $\Sigma_*$  and near IR-based  $\Sigma_*$  at 15'' resolution. Gray points show individual lines of sight. Blue points show median and robustly-estimated scatter for data binned by  $\Sigma_*$ . The red line and shaded region show the overall median ratio of  $\sim 0.08$  dex, i.e., a factor of 1.2, with about  $\pm 0.1$  dex  $\approx \pm 30\%$  scatter. The *right* panel shows the near-IR mass-to-light ratio,  $\Upsilon_*^{3.4}$ , as a function of local specific star formation rate,  $\text{SFR}/M_*$ , estimated at 15'' resolution. Individual gray points show the  $\Upsilon_*^{3.4}$  implied by the MUSE fitting. Red and blue points with error bars show the binned  $\Upsilon_*^{3.4}$  in the region with MUSE coverage for our near-IR data (blue) and implied by MUSE (red, binning the gray points). Lines show the adopted  $\Upsilon_*^{3.4}$  prescription from Leroy et al. (2019) and a version scaled by 1.2×, which describes the binned MUSE data reasonably well. Our fiducial stellar masses follow the blue line showing the Leroy et al. (2019) prescription in order to retain the link to the SDSS via Salim et al. (2016, 2018).

to-light ratio,  $\Upsilon_*^{3.4}$ , so that

$$\frac{\Sigma_*}{1 \text{ M}_\odot \text{ pc}^{-2}} \approx 350 \left( \frac{\Upsilon_*^{3.4}}{0.5} \right) \left( \frac{I_{3.6\mu\text{m}}}{1 \text{ MJy sr}^{-1}} \right) \cos i \quad (1)$$

for IRAC1 data or

$$\frac{\Sigma_*}{1 \text{ M}_\odot \text{ pc}^{-2}} \approx 330 \left( \frac{\Upsilon_*^{3.4}}{0.5} \right) \left( \frac{I_{3.4\mu\text{m}}}{1 \text{ MJy sr}^{-1}} \right) \cos i \quad (2)$$

for WISE data. The  $\cos i$  factor accounts for the inclination of the galaxy and  $\Upsilon_*^{3.4}$  is the near-infrared mass-to-light ratio in units of  $\text{M}_\odot \text{ L}_\odot^{-1}$ . The slightly different prefactors between Equations 1 and 2 reflect differences in the bandpass of the two instruments. We assume that modulo the different pre-factors, the same mass-to-light ratio  $\Upsilon_*^{3.4}$  describes both 3.4  $\mu\text{m}$  to 3.6  $\mu\text{m}$ . We predict  $\Upsilon_*^{3.4}$  from an empirical fit relating  $\Upsilon_*^{3.4}$  to the local SFR-to-WISE1 color from Leroy et al. (2019),

$$\Upsilon_*^{3.4} [\text{M}_\odot \text{ L}_\odot^{-1}] = \begin{cases} 0.5 & \text{if } Q < a \\ 0.5 + b (Q - a) & \text{if } a < Q < c \\ 0.2 & \text{if } Q > c \end{cases} \quad (3)$$

where  $Q = \log_{10} \text{SFR}/(\nu_{\text{W1}} L_\nu, \text{W1})$  and  $\nu_{\text{W1}}$  refers to the frequency corresponding to the 3.4  $\mu\text{m}$  central wavelength of the WISE1 band. Here  $a = -11.0$ ,  $b = -0.375$ , and  $c = -10.2$ .

**Placing PHANGS–ALMA in a larger context:** Leroy et al. (2019) derived Equation (3) from an empirical fit relating the measurements and fitting results in the GSWLC. Salim et al. (2016) fit multi-band NUV, optical, and IR photometry using the CIGALE population synthesis code (Boquien et al. 2019) and the Bruzual & Charlot (2003) models to estimate SFR and  $M_*$  for  $\gtrsim 650,000$  SDSS galaxies. They showed excellent agreement with the stellar masses in the previous standard JHU-MPA Value Added Catalog (see Kauffmann et al. 2003; Tremonti et al. 2004)<sup>6</sup>. Using a similar framework, Leroy et al. (2019) also estimated  $M_*$  and SFR for  $\sim 15,000$  galaxies likely to lie within 50 Mpc. Thus, adopting this approach to estimate stellar mass (and SFR below) maximizes our ability to place PHANGS–ALMA objects in a larger context. In addi-

<sup>6</sup> [https://www.sdss.org/dr14/spectro/galaxy\\_mpajhu/](https://www.sdss.org/dr14/spectro/galaxy_mpajhu/)

tion to matching the SDSS main galaxy sample via the GSWLC, this specifically places PHANGS-ALMA on a nearly-matched mass scale to the GASS and xCOLD GASS surveys (Saintonge et al. 2017; Catinella et al. 2018). These also built on SDSS stellar mass estimates from the JHU-MPA Value Added Catalog and leveraged combinations of GALEX and WISE for SFR estimates. Because we also construct these estimates for all local galaxies with a CO map, we are in a position to compare these galaxies, PHANGS-ALMA, and a larger set of literature galaxy property measurements.

**Using  $SFR/M_\star$  to predict the mass-to-light ratio:** Equation (3) specifically uses the ratio of SFR-to-WISE1 luminosity, i.e., a “specific star formation rate-like” quantity, to predict  $\Upsilon_\star^{3.4}$ . In the GSWLC as well as other theoretical and empirical studies (e.g., Bell et al. 2003; Kannappan et al. 2013; Telford et al. 2020) some cognate of the specific star formation rate is a strong predictor of the age of the stellar population and thus variations in the near infrared mass-to-light ratio. Because calculating the specific star formation rate requires knowing  $M_\star$ , this creates a degenerate situation and we predict  $\Upsilon_\star^{3.4}$  from the ratio of SFR-to-WISE1 luminosity. Based on comparison with the GSWLC, this predicts  $\Upsilon_\star^{3.4}$  with  $\sim 0.1$  dex scatter. See appendix A of Leroy et al. (2019) for more details.

**Radial Profile-based  $\Upsilon_\star^{3.4}$**  — We use Equation (3) to predict a radial profile of  $\Upsilon_\star^{3.4}$ . We first match the resolution of these all data sets at  $15''$ . Then, we calculate radial profiles of WISE1 emission and SFR, either from FUV+WISE4, NUV+WISE4, or only WISE4 emission as available. Then we apply Equation (3) to all rings. We smooth the resulting profile with a 2 kpc kernel to avoid noise in the estimate. In the outer parts of galaxies where we do not recover SFR at sufficient signal to noise, we adopt the median  $\Upsilon_\star^{3.4}$  calculated outside  $0.5 r_{25}$ .

This radial profile-based approach reflects the reality that stellar populations change as a function of galactocentric radius (e.g., Watkins et al. 2016; Dale et al. 2020, among many other examples). Because of the use of low resolution azimuthal averages, it will be less sensitive to mass-to-light ratio variations associated with young stars or dust, e.g., in spiral arms or bars (see Querejeta et al. 2015). Overall, we found that the azimuthal averages offered a good compromise between isolating distinct regions of a galaxy and achieving good signal to noise. This refines the approach in Leroy et al. (2019). They used global colors that trace  $SFR/M_\star$  to estimate  $\Upsilon_\star^{3.4}$ . The largest difference arises for targets with powerful nuclear sources, either starburst or AGN, which can affect the global colors. Any such bright, contaminating source can bias the global colors to suggest high  $SFR/M_\star$  and effectively obscure an underlying older stellar population.

**Validation against PHANGS-MUSE:** PHANGS-MUSE (E. Emsellem et al. in preparation;

P.I. E. Schinnerer) mapped 19 PHANGS-ALMA targets with the MUSE integral field unit on the Very Large Telescope. PHANGS-MUSE delivers estimates of the stellar mass based on fitting the  $4850-7000\text{\AA}$  spectrum of individual regions using stellar population synthesis models. Specifically, the fitting uses the E-MILES libraries (Vazdekis et al. 2016) and adopts the same Chabrier (2003) initial mass function (IMF) as the Salim et al. (2016, 2018) fits, so that the IMF is matched across both treatments. These fits offer stellar mass estimates independent of the near-IR based approaches above. The PHANGS-MUSE data cover only a limited number of targets and do not span the entire stellar disk of their targets. Therefore, we make use of these to explore the uncertainty in our  $\Upsilon_\star^{3.4}$  and  $\Sigma_\star$  values by comparing to an independent estimate with a distinct set of systematic uncertainties. After convolving the PHANGS-MUSE internal release v2 to  $15''$  resolution, we match the astrometric grid of our IRAC- and WISE-based measurements to compare  $\Sigma_\star$  from MUSE to our estimate. We also calculate the  $\Upsilon_\star^{3.4}$  implied by MUSE by dividing the MUSE-based stellar mass by the WISE1 intensity.

We compare the MUSE results to our fiducial measurements in Figure 8. The left panel shows overall excellent agreement between our near-IR based  $\Sigma_\star$  and the MUSE-based values. Over two orders of magnitude in  $\Sigma_\star$ , the ratio between the two methods remains roughly fixed, with only a modest offset and moderate scatter. Specifically, we find a median offset of  $\sim 0.08$  dex or a factor of  $\sim 1.2$ , and  $\sim 0.11$  dex or 30% point-to-point scatter.

The right panel of Figure 8 shows that the MUSE fitting results in a similar trend of  $\Upsilon_\star^{3.4}$  versus  $SFR/M_\star$  as what we adopt for the near-IR data. The PHANGS-MUSE fitting implies a steady decline in  $\Upsilon_\star^{3.4}$  as a function of increasing  $SFR/M_\star$ . The slope of the decline resembles that applied to the near-IR maps (blue). The higher  $\Sigma_\star$  found in the right panel means that the actual  $\Upsilon_\star^{3.4}$  values implied are shifted to  $\sim 20\%$  higher values than those we predict based on Leroy et al. (2019). Mostly, the offset appears to be a simple multiplicative shift. The “reddest,” lowest  $SFR/M_\star$  regions in the MUSE data also show higher  $\Upsilon_\star^{3.4}$ , approaching  $1 M_\odot L_\odot^{-1}$ , than implied by our adopted formula.

Overall, Figure 8 suggests excellent agreement between two independent ways to estimate the mass. The PHANGS-MUSE results provide a key validation of our adopted  $\Upsilon_\star^{3.4}$  treatment. To transfer from the “PHANGS near-IR” approach used to compute sample properties to the PHANGS-MUSE system one would need to scale the near-IR based stellar mass or  $\Sigma_\star$  up by a factor of  $\sim 1.2$ . Alternatively, one could scale the PHANGS-MUSE mass down by a similar factor.

#### 4.3.2. Checks on Stellar Mass Estimates

**Table 5.** Comparison Among Stellar Mass Estimates

Other Estimate	$\log_{10}$ Median <sup>a</sup> [dex]	$\log_{10}$ Scatter [dex]
Integrated photometry	-0.07	0.06
WISE1 vs. IRAC1	0.00	0.03
WISE1 + fixed $\Upsilon_*^{3.4}$	0.06	0.13
$S^4G$ ICA + fixed $\Upsilon_*^{3.4}$	0.00	0.11
PHANGS–MUSE <sup>b</sup>	0.08	0.11

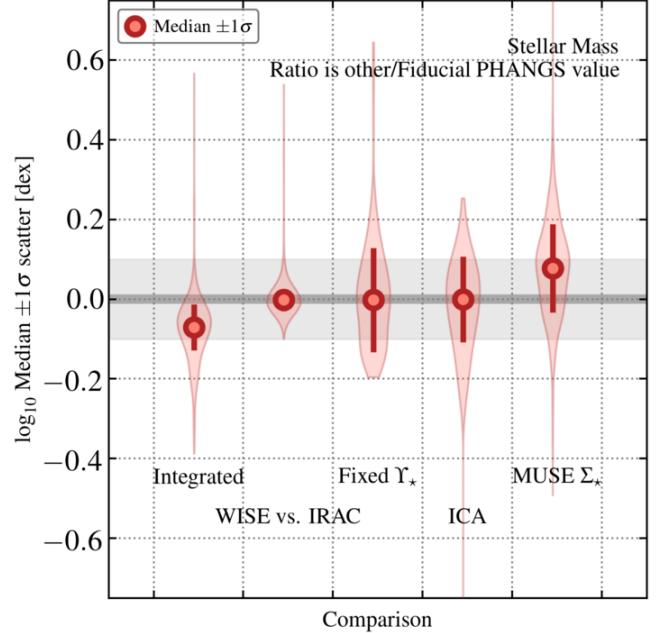
<sup>a</sup> Reported as  $\log_{10}$  of the median of the other estimate over our estimate for galaxies where both are available.

<sup>b</sup> For PHANGS–MUSE only the statistics refer to a line-of-sight by line-of-sight comparison at 15'' resolution. For all other cases we refer to comparison of integrated  $M_*$ .

NOTE—“Integrated photometry”: Our fiducial estimate divided by those from Leroy et al. (2019) based only on integrated photometry. “WISE1 vs. IRAC1”: comparing results using WISE1 vs. IRAC1 and otherwise treating the data identically. “WISE1 + fixed  $\Upsilon_*^{3.4}$ ”: Using WISE1 but with only a fixed mass-to-light ratio of  $0.35 M_\odot L_\odot^{-1}$ , not the radially varying or galaxy-dependent prescription used in our fiducial estimates. “ $S^4G$  ICA + fixed  $\Upsilon_*^{3.4}$ ”: ‘old stellar emission’ maps from Querejeta et al. (2015) using a fixed  $\Upsilon_*^{3.4}$  of  $0.5 M_\odot L_\odot^{-1}$ . The table reports comparisons for a sample of local galaxies that have CO mapping including all PHANGS–ALMA targets. The results restricting to just PHANGS–ALMA are almost identical. See also Figure 9.

Figure 9 and Table 5 provide further checks on our  $M_*$  estimates. Our fiducial masses use IRAC or WISE near-IR maps with the radially varying estimate of the mass-to-light ratio described above. In Figure 9 we adjust this approach and compare alternative estimates to our values. From left to right, we compare to (1) masses calculated only from integrated colors and luminosities by Leroy et al. (2019), (2) masses calculated using only IRAC or WISE data, i.e., checking whether adopting one data set or the other makes a difference; (3) masses calculated without any variable mass-to-light ratio, i.e., with a single fixed  $\Upsilon_*^{3.4}$  of  $0.35 M_\odot L_\odot^{-1}$  applied to all data; and (4) masses calculated applying a fixed mass-to-light ratio of  $0.5 M_\odot L_\odot^{-1}$  to the contaminant-corrected “old stellar”<sup>7</sup> maps of Querejeta et al. (2015). The final column, (5), compares the surface density,  $\Sigma_*$ , estimated in our adopted approach to that derived from PHANGS–MUSE spectral modeling. Unlike the other cases, which consider whole galaxies, this comparison treats each 15'' line of sight independently.

<sup>7</sup> We refer to the first component of their independent component analysis (ICA) as ‘old stellar emission’ following the interpretation in Meidt et al. (2012) and Querejeta et al. (2015). For brevity, we label this “ICA” in the plots and table.



**Figure 9. Comparisons among stellar mass estimates.** Median (point),  $\pm 1\sigma$  scatter (error bar), and distribution (shaded violin plot) for the ratios of stellar mass estimated in various ways. In each case, we consider the ratio of our adopted estimate, which uses IRAC- or WISE-based near-IR data and a radially varying mass-to-light ratio estimated following Equation 3, to some other estimate. From left to right we show “Integrated” as comparison between our estimates and those from Leroy et al. (2019) using only integrated photometry; “WISE vs. IRAC” comparing results using IRAC vs. WISE and treating the data identically. The results are highly consistent except for a few outliers. “Fixed  $\Upsilon_*^{3.4}$ ”: Showing results using *only* WISE1 intensity with no varying mass-to-light ratio. “ICA”: Using independent component analysis-based ‘old stellar emission’ maps from Querejeta et al. (2015) with a fixed mass-to-light ratio. “MUSE  $\Sigma_*$ ”: 15'' resolution comparison between our mass surface density estimates and PHANGS–MUSE spectral fitting results, available for a subset of targets. The gray line and shaded region show perfect agreement and  $\pm 0.1$  dex scatter. See also Table 5.

Table 5 summarizes the numerical results. Qualitatively, we see that shifting from the estimates based on integrated photometry from Leroy et al. (2019) to resolved estimates of  $\Upsilon_*^{3.4}$  here increases the overall stellar mass by only  $\sim 20\%$  on average but can matter much more in a few cases. These tend to be cases with strong nuclear concentrations, but also a few cases where our more careful treatment of the data here has improved the masking. Beside a few edge cases, it makes a negligible difference to the overall  $M_*$  whether we use IRAC or WISE as long as the treatment of the data remains the

same. The adoption of variable  $\Upsilon_*^{3.4}$  does matter. Without the variable  $\Upsilon_*^{3.4}$ , masses scatter by  $\pm 0.13$  dex or 35% relative to our derived stellar masses. Similarly, if we integrate the contaminant-corrected ‘old stellar emission’ component maps from Querejeta et al. (2015) and apply a fixed  $\Upsilon_*^{3.4}$ , we find  $\pm 0.11$  dex or 30% scatter relative to our adopted values.

In Figure 9 and Table 5 we have already chosen our typical mass-to-light ratios to remove systematic offsets between our adopted approach and the comparison data. Thus two key points from the comparison are:

1. We find  $\Upsilon_*^{3.4} = 0.35 \text{ M}_\odot \text{ L}_\odot^{-1}$  on average.
2. When using the ‘old stellar emission’ maps derived by Querejeta et al. (2015) from combining IRAC  $3.6\mu\text{m}$  and  $4.5\mu\text{m}$  emission, we suggest to apply  $\Upsilon_*^{3.4} = 0.5 \text{ M}_\odot \text{ L}_\odot^{-1}$  for consistency with our adopted system. The difference between this value and the average  $\Upsilon_*^{3.4} = 0.35 \text{ M}_\odot \text{ L}_\odot^{-1}$  reflects the removal of contaminants from the Querejeta et al. (2015) maps.

Although we set the median to agree, the scatter between our adopted system and either the Querejeta et al. (2015) or fixed  $\Upsilon_*^{3.4}$  is still significant. In general, our checks show that any substantial shift in methodology induces  $\sim 0.1$  dex, or  $\approx 25\%$ , scatter in the measurements. This uncertainty dramatically exceeds any statistical uncertainty associated with the photometry and we suggest adopting  $\pm 0.1$  dex as a realistic systematic uncertainty on stellar masses of individual galaxies (in good agreement with the level of uncertainty due to age variations found in the IRAC-based ICA maps by Meidt et al. 2014). It also seems reasonable to adopt comparable uncertainty to describe the overall calibration of our stellar mass scale.

Despite these uncertainties, we emphasize again that (1) our calibration scheme anchors our mass scale to the Salim et al. (2016, 2018) and Leroy et al. (2019) estimates, and through these to work on the SDSS main galaxy sample, and (2) that our adopted  $\Upsilon_*^{3.4}$  treatment seems to qualitatively agree with results from PHANGS–MUSE, while showing a mild 0.08 dex systematic offset.

#### 4.4. Stellar Disk Size

The size of the stellar disk plays a key role in many analyses. Using the same  $\Sigma_*$  values used to calculate  $M_*$ , we derive two size estimates for each target. First, we measure the effective radius,  $R_e$ , that contains half of the stellar mass of the galaxy. The effective radius is a standard measure of galaxy size but can be sensitive to stellar bulges, nuclear mass concentrations, and other details of inner galaxy structure. Therefore to complement  $R_e$ , we measure the scale length of the stellar disk,  $l_*$ , outside the galaxy center. This disk scale length is related to the metallicity gradient (e.g., Sánchez et al.

2014) and stellar disk scale height (e.g., Kregel et al. 2002; Salo et al. 2015; Sun et al. 2020a), among other applications. In Table 4 we also report the 25<sup>th</sup> magnitude isophotal  $B$ -band radius,  $r_{25}$ , from RC3 (de Vaucouleurs et al. 1991) via HyperLEDA.

##### 4.4.1. Measurement of $R_e$ and $l_*$

We measure  $R_e$  and  $l_*$  from the 7.5'' resolution estimates of  $\Sigma_*$  described in §4.3. Using these data, we construct a radial profile using the orientations in Table 3. In order to lower our sensitivity to the adopted inclination and resolution, we consider only points within  $\pm 75^\circ$  of the major axis, except for the central 30''. Within the central 30'', we include all position angles<sup>8</sup>. Then, we measure the azimuthally averaged intensity using the mean within 30'' and the median intensity in each ring at larger radii. The use of the median suppresses any unmasked foreground stars at the expense of also losing some sensitivity to galactic structure. We use the mean in the central aperture to retain full sensitivity to any bright nuclear emission when calculating  $R_e$ . Next, we subtract a background value from the radial profile, which is just the median of all radial bins outside 2  $r_{25}$ . Finally, we integrate these profiles assuming azimuthal symmetry and use this integral to identify the half-mass radius,  $R_e$ , the radius encompassing 90% of the mass,  $R_{90}$ , and the stellar scale length,  $l_*$ . The integral for calculating  $R_e$  extends out to 2  $r_{25}$ , which we verified by eye in each case to represent an apparent convergence of the radial profile. We report the estimated  $R_e$  and  $l_*$  in Table 3.

To fit  $l_*$ , we identify a fitting range over which the radial profile appears exponential. By default, this range spans from 0.4 to 1.0  $r_{25}$ , i.e., excluding roughly the central two scale lengths. We examined each profile by eye and adjusted this range to reflect where the disk appears exponential over a large range of radii outside the galaxy center. This approach does a good job of matching sophisticated disk/bulge decompositions (e.g., Salo et al. 2015). This also closely resembles the approach adopted by Sánchez et al. (2014), a study that we use to estimate metallicity gradients.

##### 4.4.2. Translation Between Different Size Estimates

We measure effective radii and stellar scale lengths. The 25<sup>th</sup> magnitude  $B$ -band isophotal radius also remains widely used, particularly for studies of local galaxies. In Table 6 and Figure 10 we report translations between different size measurements for our galaxies: half-mass radius,  $R_e$ , profile-based scale length,  $l_*$ ,

<sup>8</sup> We consider this necessary to ensure that we remain sensitive to any massive nuclear features, which can strongly affect the estimated  $R_e$ . Based on visual inspection of all profiles we do not see strong imprints of this change in procedure in the derived profiles, and as shown in §4.4.3 our estimates match previous profile-based size estimates very well.

and isophotal radius,  $r_{25}$  (de Vaucouleurs et al. 1991; Makarov et al. 2014, we adopt from RC3 via HyperLEDA). For reference, for a pure exponential disk,  $l_*$ , should relate to the effective radius via

$$R_e^{\text{disk}} = 1.68 l_* \quad \text{or} \quad l_* = 0.60 R_e^{\text{disk}}. \quad (4)$$

In practice, we find that our measured  $R_e$  and  $l_*$  deviate from this ideal case. We show this in Figure 11 and report the typical ratio and scatter for our measurements in Table 6. We find that in practice, the relationship in our sample can be better described via

$$R_e^{\text{disk}} = 1.41 l_* \quad \text{or} \quad l_* = 0.71 R_e^{\text{disk}}. \quad (5)$$

The difference between these two formulae simply reflects that our targets are not ideal exponential disks, and that many galaxies harbor compact inner structures. These drive  $R_e$  to smaller values but do not affect  $l_*$ . As Figure 11 shows, even Equation (5) still misses a significant minority of outliers that have high  $l_*$  compared to  $R_e$ , reflecting strong central mass concentrations relative to an exponential disk.

The 25<sup>th</sup> magnitude  $B$ -band isophotal radius,  $r_{25}$ , remains in common use when studying nearby galaxies. Table 6 and Figure 10 show that  $r_{25}$  exhibits a well-defined scaling with our two size measures:

$$R_e \approx r_{25}/3.08 \quad \text{and} \quad l_* \approx r_{25}/4.3 \quad (6)$$

on average, both with  $\sim 0.14$  dex scatter.

#### 4.4.3. Checks on Stellar Disk Sizes

We compare our estimated sizes to literature size estimates for the same galaxies and show the results in Table 6 and Figure 10. We compare our measured sizes to:

1. The 25<sup>th</sup> magnitude  $B$ -band isophotal radius from RC3 via HyperLEDA (de Vaucouleurs et al. 1991; Makarov et al. 2014). This remains the most widely used measure of galaxy size for studies of local galaxies, and the results of this comparison have already been used to derive Equation 6.
2. Effective radii estimated by the 2MASS Large Galaxy Atlas (LGA; Jarrett et al. 2003). They measured  $R_e$  at  $J$ ,  $H$ , and  $K$  band for many of our galaxies and a large sample of additional galaxies. These are profile-based sizes using near-IR data and we expect our measurements to match them well.
3. Effective radii estimated from the S<sup>4</sup>G near-IR data by Muñoz-Mateos et al. (2015). These are also profile-based measurements, in many cases using the same data with slightly distinct processing, assumptions, and orientations. We expect to match these values well.

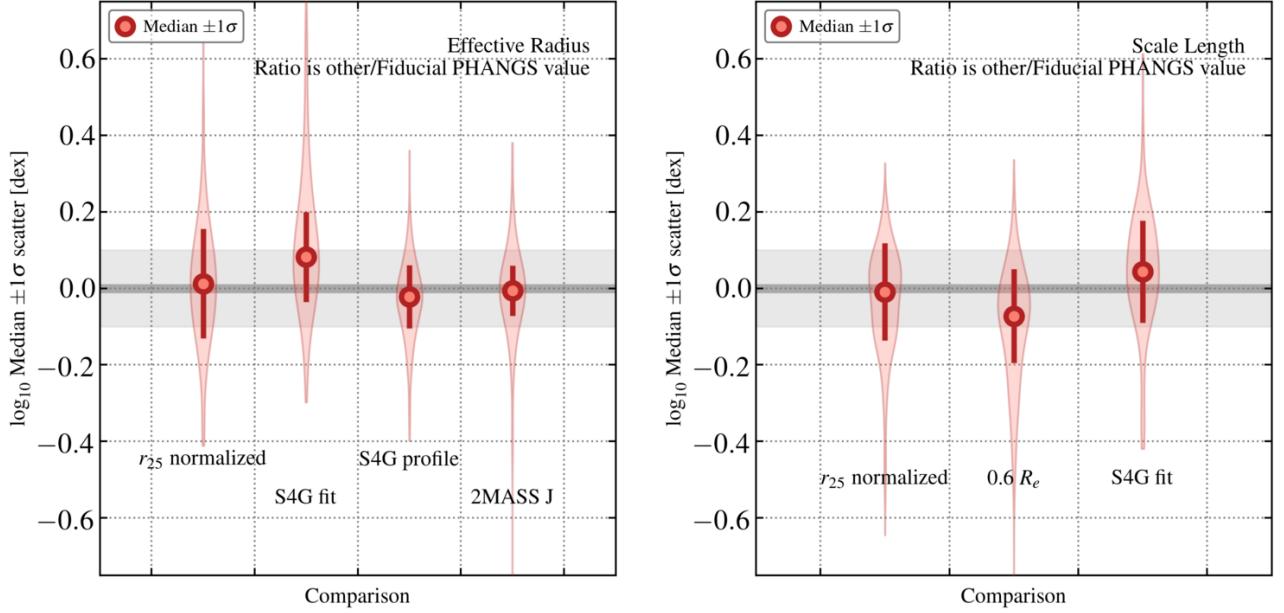
4. The results of single- and multi-component fits to the S<sup>4</sup>G data by Salo et al. (2015). They present one-component Sérsic profiles, which yield an effective radius. They also fit exponential disk components to most targets. When they fit an `expdisk` component, their  $h_r$  captures the exponential scale length and represent the most direct comparison for our  $l_*$ .

Overall, Figure 10 and Table 6 show that our measured sizes match measurements from other sources well. As expected, we match previous profile-based estimates of the size using near-IR data well. In 2MASS, we find best agreement with the  $J$ -band size, with a median ratio of 2MASS-to-PHANGS  $R_e$  of 0.99. We find slightly smaller sizes in 2MASS  $H$  and  $K$  band, likely reflecting the poorer sensitivity in those bands, but the agreement is also excellent for those bands. Our match to the Muñoz-Mateos et al. (2015) sizes is also good, we find an 0.95 ratio on average.

On average, the effective radii from the Sérsic model fits by Salo et al. (2015) are  $\sim 21\%$  larger than our  $R_e$ . This likely reflects the fact that fitting a single Sérsic profile to the entire galaxy de-emphasizes the contribution of any distinct, massive central component. By-eye checks of galaxies with strongly discrepant size estimates between us and Salo et al. (2015) reveal that such galaxies often have central mass concentrations. These same discrepancies arise comparing the Salo et al. (2015) values to Muñoz-Mateos et al. (2015) ones. Those two measurements use the same data, but the single-component fitting approach of Salo et al. (2015) tends to yield larger  $R_e$  than the profile fits of Muñoz-Mateos et al. (2015), a 26% higher  $R_e$  for matched galaxies, on average, with  $\sim 0.1$  dex scatter.

Despite the good match on average, we find just under  $\sim 0.1$  dex or about 25% scatter between our measurements and those of Muñoz-Mateos et al. (2015) or Jarrett et al. (2013), and we adopt  $\pm 0.1$  dex as a realistic  $1\sigma$  uncertainty on  $R_e$ . The scatter between estimates reflects differences in the adopted orientations and sensitivity to many aspects of the exact method used. The method of defining the total flux, choice to focus on the major axis or whole galaxy, choice to apply a radially varying mass-to-light ratio, and a host of other minor decisions all affect the derived  $R_e$ . To some extent, this reflects that galaxies are not axisymmetric, thin, tilted disks with a single orientation and a simple radial profile. It also highlights how  $R_e$ , which depends on an estimate of the total flux and focuses on the inner, often complex, part of the galaxy is a less stable size measure than an isophotal radius for local galaxies.

Figure 10 and Table 6 show that our stellar scale lengths agree reasonably well with those from Salo et al. (2015). Their sophisticated two dimensional decomposition tends to yield  $\sim 11\%$  larger scale lengths than our fits to the radial profiles. Our approach focuses on fitting



**Figure 10. Stellar disk sizes in different systems.** Comparisons between our estimated  $R_e$  and  $l_\star$  and literature measurements; see also Table 6. The left panel compares our measured  $R_e$  to literature size measurements: the isophotal radius  $r_{25}$  scaled according to Equation 6 to predict  $R_e$ , estimates from Salo et al. (2015, “S<sup>4</sup>G fit”) and Muñoz-Mateos et al. (2015, “S<sup>4</sup>G profile”), and estimates from 2MASS. The S<sup>4</sup>G profile-based sizes and 2MASS J-band sizes show good agreement with our estimates. The right panel shows results for our scale length  $l_\star$  fits. Here we have only the Salo et al. (2015) literature values to compare to. We do find overall agreement, but see more scatter than for  $R_e$ .

the part of the profile roughly in the range 0.4–1.0  $r_{25}$  that looks approximately exponential. This might lead us to emphasize more cleanly declining parts of the disk and avoid regions with flatter profiles. Regardless, the overall agreement with Salo et al. (2015) is also good, and  $l_\star$  is likely somewhat more uncertain than  $R_e$ .

#### 4.5. Star Formation Rate

We estimate the star formation rate (SFR) using GALEX FUV, GALEX NUV maps, WISE3 (12  $\mu\text{m}$ ), and WISE4 (22  $\mu\text{m}$ ) maps. We adopt the prescription suggested by Leroy et al. (2019) and calibrated to match results from the population synthesis modeling in the GSWLC (Salim et al. 2016, 2018). In brief:

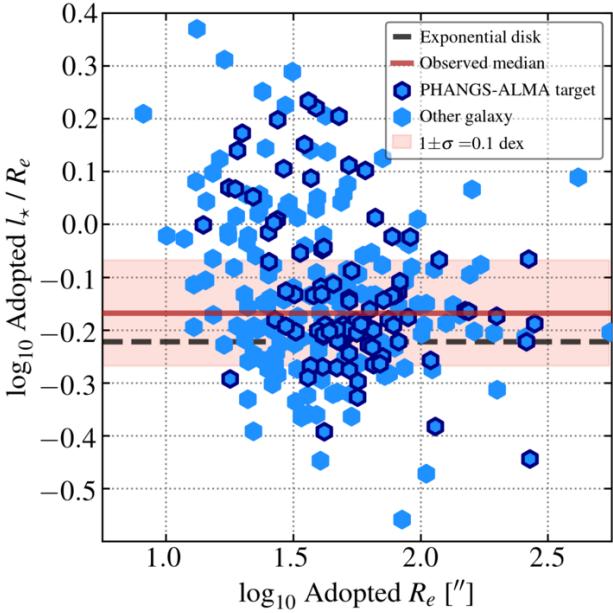
1. We adopt a linear combination of FUV and WISE4 (22  $\mu\text{m}$ ) light whenever both bands are available. If FUV is not available but NUV is, we use NUV emission. If no GALEX data are available, we use only WISE4. In this paper we use WISE3 emission only to construct alternative estimates used to check our data (see below). Table 4 list the combination of bands used to estimate SFR for each galaxy.
2. For each band, we convert from luminosity,  $\nu L_\nu$ , to SFR using a conversion factor,  $C$ .
3. The conversion factor for NUV and FUV are the same no matter which IR band is used (see Salim

et al. 2007 and Leroy et al. 2019). In practice this results from fits to the Salim et al. (2016) results. Physically, this reflects an assumption that the recent star formation history does not vary too much across our sample. The conversion factor for WISE3 or WISE4 depends on the band with which it is combined.

4. We add together the UV (if present) and mid-IR terms to obtain the total SFR. The choice to use a linear combination sacrifices some accuracy in exchange for a more stable estimator.

##### 4.5.1. Checks on Star Formation Rates

Tables 7 and 8 and Figure 12 illustrate the uncertainty in our estimated SFR. Table 7 reports the scatter and bias implied by adopting a single linear calibration and attempting to match the full Salim et al. (2018) data set. Here “bias” refers to the rms scatter in the implied calibration factor,  $C$ , as a function of  $M_\star$  and SFR. To calculate this, Leroy et al. (2019) binned the Salim et al. (2018) data into cells of SFR and  $M_\star$ , solved for  $C$ , in each cell, and then calculated the rms scatter across all populated cells in the SFR- $M_\star$  plane. Thus, this number captures the degree of systematic uncertainty implied by applying a single value of  $C$  to the whole galaxy population. This is discussed in detail by Leroy et al. (2019), but here we emphasize two key points. First, using WISE4 only shows notably higher bias than



**Figure 11. Scale length and effective radius.** The fit exponential scale length,  $l_*$ , divided by the effective radius,  $R_e$ , measured from the same profile. The black dashed line shows expectations for a pure exponential disk (Eq. 4). Points above this line show some central concentration compared to a pure exponential disk, and this central concentration drives  $R_e$  to smaller values. The  $l_*$  fits should better capture the exponential portion of the profile. The axes are correlated, and the detailed trend is less important than the qualitative point: that many galaxies with small  $R_e$  have bright nuclear concentrations, and that the fit  $l_*$  indicates a much larger disk in some of these cases. That larger  $l_*$  will be the relevant quantity for disk structure calculations or metallicity gradient estimation.

WISE4 combined with a UV band. This reflects the fact that low mass galaxies, in particular, often show less obscuration compared to high mass galaxies. The importance of a hybrid approach to star formation rate tracers has been reviewed by Kennicutt & Evans (2012) and Calzetti (2013).

The other important point from Table 7 relates to the use of WISE3. WISE3 data have better sensitivity and resolution than the WISE4 data, and below we will use them to aperture correct our CO data. However, as Table 7 shows, the appropriate calibration to translate WISE3 to SFR shows strong trends as a function of stellar mass and specific star formation rate. This is consistent with the significant contribution of PAHs to the WISE3 band and the known dependence of PAH emission on metallicity (e.g., Engelbracht et al. 2006) and radiation field (e.g., Chastenet et al. 2019). This is a strong effect, greater than a factor of 2. We do

**Table 6. Comparison Among Size Estimates**

Other Estimate	Median <sup>a</sup> [dex]	$\log_{10}$ Scatter [dex]
Effective radius $R_e$ <sup>b</sup>		
$r_{25}$ (255 <sup>c</sup> )	3.08	0.14
2MASS $J$ (158)	0.99	0.07
$S^4G$ profile fit (203)	0.95	0.08
$S^4G$ Sérsic fit (199)	1.21	0.12
Scale length $l_*$ <sup>d</sup>		
$r_{25}$ (253)	4.3	0.13
$S^4G$ $h_*$ fit (69)	1.11	0.14
Our $R_e$ (253)	1.41	0.12

<sup>a</sup>Reported as median of the other estimate over our estimate for galaxies where both are available.

<sup>b</sup>Effective radius calculated from stellar mass radial profile considering all data.

<sup>c</sup>Number of galaxies available for that comparison.

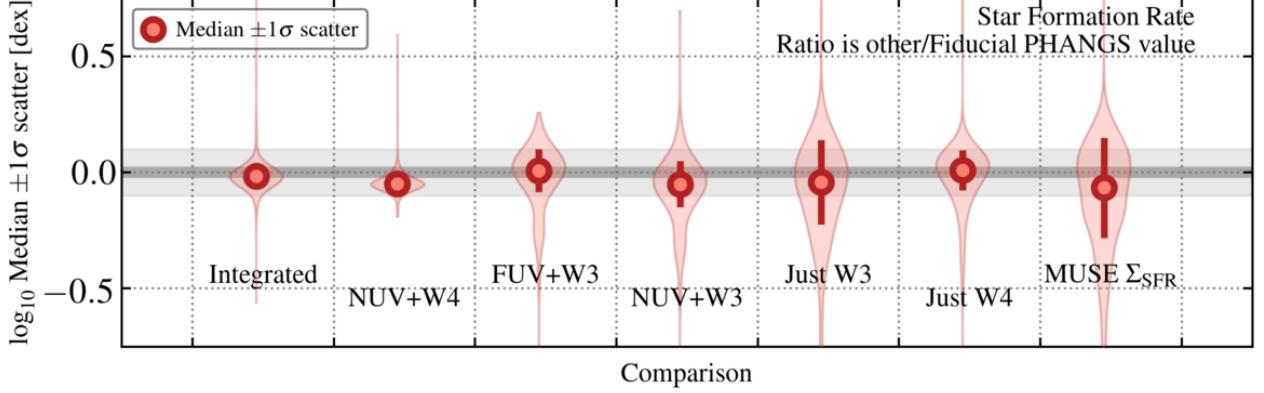
<sup>d</sup>Stellar disk scale length considering only the exponential part of the profile.

NOTE—The table reports comparisons for a sample of local galaxies that have CO mapping including all PHANGS-ALMA targets. The results restricting to just PHANGS-ALMA are almost identical (Figure 10).

note that this will mainly affect extremes in the galaxy population, i.e., low mass, low metallicity galaxies and very low and high SFR/ $M_*$  galaxies.

For some PHANGS-ALMA science applications, the higher resolution and improved signal to noise of the WISE3 data will be important. Indeed, we used WISE3 to define the field of view for exactly these reasons. When we need such a ratio to be quantitative, we will generally estimate a WISE3-to-WISE4 ratio for that galaxy or location and use this as an empirical correction to apply the better-calibrated WISE4-based SFR prescriptions.

Figure 12 and Table 8 mostly bear out these trends within our own data. In this figure we vary the adopted band combination and compare to our fiducial SFR estimate. We also compare the value from Leroy et al. (2019) based only on integrated photometry. Both the table and the figure show that all band combinations have consistent calibrations on average. However, if we do not hybridize WISE with some UV band, then we find a higher galaxy-to-galaxy scatter,  $\sim 0.1$  dex using only WISE4. If we substitute WISE3 for WISE4, we also see  $\sim 0.1$  dex scatter relative to our fiducial estimates. Finally, while using WISE3 alone yields a matched result on average, though, we find almost 0.2 dex galaxy-to-galaxy scatter relative to our fiducial estimates.



**Figure 12. Comparisons among star formation rate estimates.** Median (point),  $\pm 1\sigma$  scatter (error bar), and distribution (shaded violin plot) for the ratios among star formation rates estimated in various ways. In each case, we consider the ratio of some other estimate to our adopted estimate. From left to right, we compare our fiducial values (usually FUV+WISE4, see text) to “Integrated”: results using integrated photometry from Leroy et al. (2019), “NUV+W4”: results preferring the GALEX NUV band instead of FUV and hybridizing with WISE4; “FUV+W3”: results using WISE3 instead of WISE4; “NUV+W3” same, but now using NUV and WISE3; “Just W3”: estimates based on only WISE3 emission; “Just W4”: estimates based on only WISE4 emission; “MUSE  $\Sigma_{\text{SFR}}$ ”: estimates based on Balmer-decrement corrected H $\alpha$  emission from PHANGS–MUSE. Unlike the other estimates, which are all compared for whole galaxies, the MUSE  $\Sigma_{\text{SFR}}$  comparison considers each 15'' line of sight. See also Table 8.

#### 4.5.2. UV+IR and PHANGS–MUSE Balmer-Decrement

For 19 of our targets, PHANGS–MUSE (E. Emsellem et al. in preparation) offers an independent, high-quality estimate of  $\Sigma_{\text{SFR}}$  from H $\alpha$  recombination line emission corrected for extinction using the Balmer decrement. An in-depth exploration of the differences between these two methods is presented by F. Belfiore et al. (in preparation). Here we present a brief comparison of  $\Sigma_{\text{SFR}}$  estimated from the PHANGS–MUSE H $\alpha$  and H $\beta$  maps at 15'' resolution and the UV+IR  $\Sigma_{\text{SFR}}$  estimates available for the whole sample.

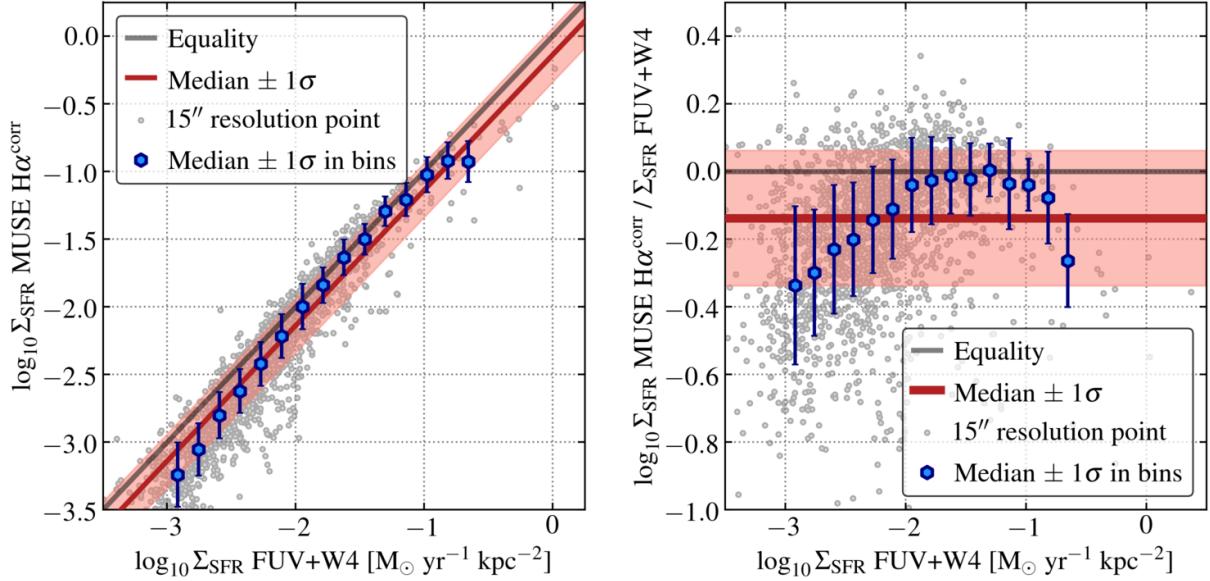
For this comparison, we compare our best  $\Sigma_{\text{SFR}}$  estimates, which are from FUV+WISE4 for all 19 MUSE targets, to  $\Sigma_{\text{SFR}}$  estimated from H $\alpha$  and H $\beta$  measured from PHANGS–MUSE at 15'' resolution. We use the two recombination lines to estimate  $A_{\text{H}\alpha}$ , the extinction affecting the H $\alpha$  line, by assuming a screen geometry, a Cardelli et al. (1989) extinction curve, and Case B recombination (see F. Belfiore et al. (in preparation) for more details). Then we place the two 15'' resolution  $\Sigma_{\text{SFR}}$  maps on the same astrometric grids and compare results.

Figure 13 shows the results of this comparison. As the left panel shows, the two estimates correlate extremely well, showing a Spearman rank correlation coefficient of  $\approx 0.95$  despite using completely different approaches to estimate  $\Sigma_{\text{SFR}}$ . However, the estimators do yield numerically different estimates. On average the FUV+WISE4 estimate is 20–30% higher than the H $\alpha$ -based estimate. Specifically, treating all points equally yields a median H $\alpha$ -to-FUV+WISE4 ratio of  $\approx 0.73$ , or  $-0.14$  dex, with

$\approx 0.2$  dex, or 60% rms scatter about the median. Considering the binned measurements, shown in blue, we find a median ratio of 0.84 or  $-0.07$  dex with  $\approx 0.1$  rm scatter in the median value from bin-to-bin. Given the independent approaches to estimate  $\Sigma_{\text{SFR}}$ , this offset still represents reasonable agreement, and any average offset could be accounted for by adjusting the empirically-derived coefficient on the WISE4 term in the FUV+WISE4 indicator.

The Figure shows that the offset and scatter between the two tracers does reflect an average offset and a systematic trend, such that the two tracers agree well at  $\Sigma_{\text{SFR}} \gtrsim 10^{-2} \text{ M}_\odot \text{ yr}^{-1} \text{ kpc}^2$  but show increasing divergence below this value. Specifically, at low  $\Sigma_{\text{SFR}}$ , the FUV+WISE4 SFR indicator predicts a higher  $\Sigma_{\text{SFR}}$  than the H $\alpha$ -based estimate. Figure 13 illustrates a systematic trend with  $\Sigma_{\text{SFR}}$ , and F. Belfiore et al. (in preparation) show that a similar trend holds as a function of the ratio  $\Sigma_{\text{SFR}}/\Sigma_*$ , a quantity that traces the local specific star formation rate.

The sense of the trend agrees with the expectation that the 22 $\mu\text{m}$  emission captured in the WISE4 band may contain significant emission not directly associated with star formation, and that this contamination will be stronger in regions with low star formation activity (e.g., see Groves et al. 2012; Leroy et al. 2012; Boquien et al. 2016; Simonian & Martini 2017). Given the multiwavelength coverage in PHANGS, it may be possible to correct these low-intensity sight-lines using physical or empirical prescriptions for the IR cirrus (e.g., Leroy et al. 2012; Davis et al. 2014). Alternatively, high resolution mid-IR observations from the *James Webb Space*



**Figure 13. Comparisons between  $\Sigma_{\text{SFR}}$  estimated from UV+IR and from PHANGS–MUSE Balmer decrement-corrected  $\text{H}\alpha$ .** For each  $15''$  sightline in the field of view of PHANGS–MUSE, we compare  $\Sigma_{\text{SFR}}$  estimated from  $\text{H}\alpha$  extinction-corrected using the Balmer decrement,  $\text{H}\alpha^{\text{corr}}$  ( $y$ -axis) to  $\Sigma_{\text{SFR}}$  estimated from combining FUV and mid-IR  $22\mu\text{m}$  emission ( $x$ -axis). The  $\text{H}\alpha+\text{H}\beta$  measurements are only available over a limited field-of-view in 19 targets, but represent a high quality, independent reference. The *left* panel shows the two estimates plotted against one another. The *right* panel shows the ratio of the two estimates. The two  $\Sigma_{\text{SFR}}$  estimates correlate well, but show a  $\approx 20\text{--}30\%$  systematic offset. We also observe a systematic trend in the ratio as a function of  $\Sigma_{\text{SFR}}$ , with the FUV+WISE4 estimate likely somewhat overestimating  $\Sigma_{\text{SFR}}$  at low values. For a more detailed exploration of these results see F. Belfiore et al. (in preparation).

*Telescope* offer a path forward to morphologically and structurally isolate emission associated with star forming regions (e.g., Calzetti et al. 2005; Boquien et al. 2016).

We refer the reader to F. Belfiore et al. (in preparation) for more details but emphasize three main points from the comparison. First, the overall scaling between the two tracers appears very good, and the median offset of 20–30% represents good agreement given the independent approaches. Second, the clear systematic trends suggest that the FUV+WISE4 maps be treated with caution at low  $\Sigma_{\text{SFR}} \sim 10^{-3} \text{ M}_\odot \text{ yr}^{-1} \text{ kpc}^2$ . A cirrus correction to the WISE4 data might improve the situation. However, third, the coefficients on the FUV+WISE4 estimator have been set to ensure average consistency with the GSWLC and the larger SDSS sample Salim et al. (2016, 2018). For this work, we use the coefficients “as is” and note the comparison to PHANGS–MUSE as an important indicator of the systematic uncertainty.

#### 4.6. $\text{CO}(2\text{--}1)$ Luminosity

We use the PHANGS–ALMA data to estimate the integrated CO luminosity of each target. To do this, we integrate a  $17''$  resolution version of the ALMA CO(2–1) cube. Working with a low resolution version of the data makes signal identification easier among. We apply a

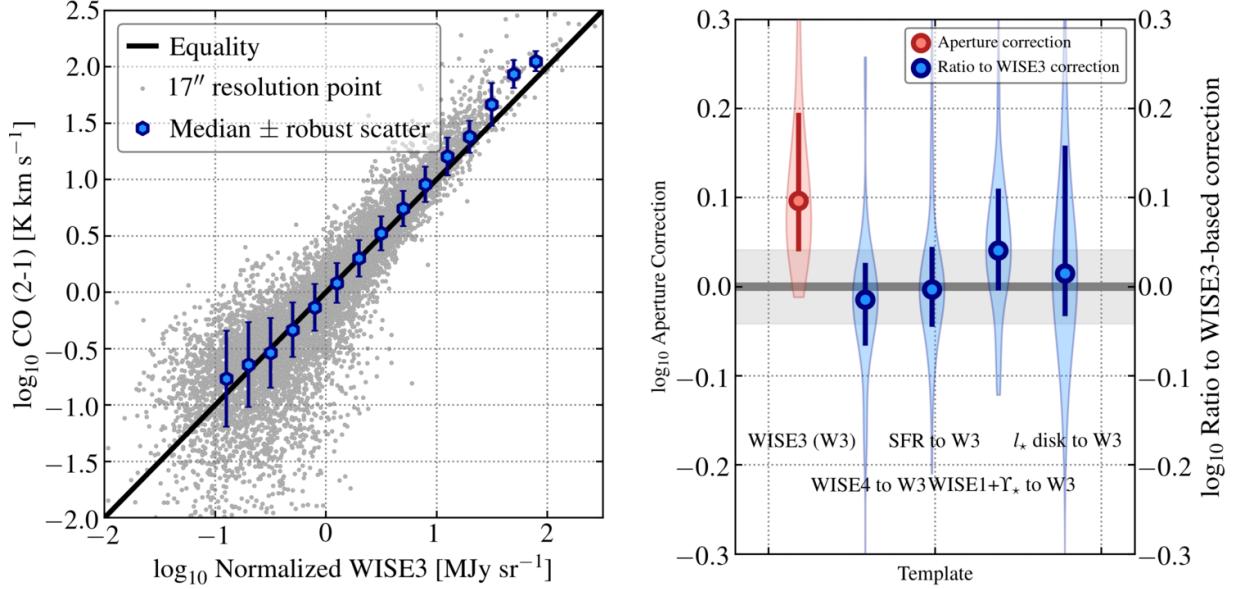
wide mask to the data before integrating<sup>9</sup>. We find that the results derived using this mask agree well with those applying no masking at all. Using no mask only adds more noise to the measurement. We report the integrated CO(2–1) luminosities for each galaxy with in Tables 4.

##### 4.6.1. Aperture Correction

Our maps only cover the area of active star formation, defined by an IR intensity contour (§5). Though we expect most CO emission to lie within our target area, faint CO emission can continue well into the outer disks of galaxies (e.g., Young et al. 1995; Braine et al. 2007; Schruba et al. 2011). To account for this, we derive an aperture correction that can be applied to our measured integrated CO(2–1) to yield the global  $L_{\text{CO}}$  for the galaxy.

After considering several options, we chose to calculate this aperture correction by using the WISE3 map as a template for CO emission. We calculate the sum of WISE3 emission over the whole galaxy and the sum of

<sup>9</sup> This is not exactly the same as the “broad mask” described in Sections 6 and 7. The mask that we use here is designed to also apply to the HERA data used for checking the fluxes and to compare to other literature CO measurements. We do not expect this detail to have any important effect on the derived CO luminosities.



**Figure 14. Calculations of aperture corrections to infer total CO luminosity.** PHANGS-ALMA covers most of the area of active star formation but not the entire disk of each target. We derive an aperture correction to correct from our measured  $L_{\text{CO}}$  to the global  $L_{\text{CO}}$  for each target. To do this, we use WISE3 intensity as a template for CO intensity. The *left* panel shows the motivation for this choice, plotting CO(2-1) intensity as a function of WISE3 intensity at 17'' resolution within the PHANGS-ALMA coverage. A single normalization factor has been derived and applied to correct each galaxy for differences in the median WISE3-to-CO(2-1) ratio. After this normalization, the two intensities show a linear correlation coefficient of 0.88, exhibiting a stunningly tight correlation across three decades in intensity. The *right* panel shows the derived aperture corrections for WISE3 (red) and the ratio between aperture corrections derived using other templates and that from WISE3. WISE3 represents the clear best option, but if we adopted another reasonable choice like WISE4, stellar mass, or star formation rate then this would imply  $\sim 10\%$  scatter (gray band) on average and a bias of  $\lesssim 10\%$ .

WISE3 emission over just the region with CO coverage. In Table 4 we report this WISE3-based aperture correction, but note that we leave it to the user to actually apply the correction. The  $L_{\text{CO}}$  that we report reflects only the direct integration of our maps.

Table 9 and Figure 14 show the motivation for choosing WISE3. Inside the PHANGS-ALMA footprint, we construct several possible “templates” that could be used for the aperture correction. Here a template refers to a quantity that might linearly trace CO intensity. We consider WISE3 intensity, WISE4 intensity,  $\Sigma_{\text{SFR}}$  from FUV+WISE4, and  $\Sigma_{\star}$  estimated from WISE1+ $\Upsilon_{\star}^{3.4}$ . We also consider three plausible versions of an exponential disk model: one adopting a scale length of  $0.2 r_{25}$  (Young et al. 1995; Leroy et al. 2008; Bolatto et al. 2017), one adopting our fit exponential scale length, and one adopting  $0.6 R_e$  (see Equation 4). These size-based corrections resemble those used, e.g., by Lisenfeld et al. (2011) and Saintonge et al. (2011).

Table 9 shows that WISE3 exhibits the strongest linear correlation coefficient with CO intensity of any considered quantity. After removing a single median WISE3-to-CO ratio from each galaxy, the WISE3 emission exhibits a linear correlation coefficient of 0.82 with CO(2-1) intensity across the PHANGS sample. This

stunningly strong correlation has also been seen and investigated in detail by Chown et al. (2021) in the EDGE-CALIFA CO(1-0) data. We confirm their finding that  $I_{\text{CO}}$  shows a strong, linear correlation with WISE3 within galaxies. After removing a galaxy-to-galaxy offset, the correlation between WISE3 emission and CO(2-1) intensity is even stronger than that relating CO to  $\Sigma_{\text{SFR}}$  estimated from FUV and WISE4 emission. Figure 14 shows this correlation. After galaxy-by-galaxy normalization, WISE3 tracks CO(2-1) emission almost linearly with  $< 0.1$  dex scatter across three decades in intensity. We do caution that as discussed above, the ratio of WISE3 emission to SFR shows strong galaxy-to-galaxy variations. These variations correlate with metallicity and specific star formation rate (Leroy et al. 2019). Our application avoids much of this issue by removing an overall galaxy-to-galaxy scaling. The PHANGS-ALMA data offer the possibility to explore the empirical and physical nature of the WISE3-CO correlation in much more detail in future works.

WISE3 offers an outstanding option to derive aperture corrections and we derive a median correction of  $\sim 1.25$  for PHANGS-ALMA, with most targets yielding corrections between 1.1 and 1.5. In other words, the WISE3 emission implies that our maps usually cover  $\sim 70\text{--}90\%$

**Table 7.** Adopted SFR Calibrations and Uncertainties

Band	$\log_{10} C^a$ $\left( \frac{\text{M}_\odot \text{ yr}^{-1}}{\text{erg s}^{-1}} \right)$	RMS Scatter <sup>b</sup> (dex)	RMS Bias <sup>c</sup> (dex)
FUV <sup>d</sup>	-43.42	0.1	0.09 <sup>d</sup>
NUV <sup>d</sup>	-43.24	0.1	... <sup>d</sup>
WISE4 with FUV	-42.73	0.17	0.06
WISE4 with NUV	-42.79	0.18	0.08
WISE4 alone	-42.63	0.17	0.15
WISE3 with FUV	-42.79	0.21	0.44
WISE3 with NUV	-42.86	0.22	0.47
WISE3 alone	-42.70	0.2	0.37

<sup>a</sup>  $C$  is the factor to convert  $\nu L_\nu$  to SFR (notation follows Kennicutt & Evans 2012).

<sup>b</sup> Galaxy-to-galaxy scatter of  $C$ , in dex, from comparison to Salim et al. (2018). See Leroy et al. (2019) for more details.

<sup>c</sup> “Bias” here refers to the systematic scatter, in dex, among the empirical calibration coefficients to be applied to the WISE data across the SFR- $M_*$  plane. This is calculated by binning the Salim et al. (2018) data into cells of SFR and  $M_*$ , calculating  $C$  for each cell, and then measuring the rms scatter across the binned data. This indicates the amount of systematic uncertainty in the empirical calibration of the tracer. See Leroy et al. (2019) for more details.

<sup>d</sup> The coefficients applied to FUV and NUV do not depend on which IR band is used.

NOTE—From Leroy et al. (2019) based on the GSWLC (Salim et al. 2016, 2018). These coefficients represent a best effort to anchor our SFR estimates to both a large population of local galaxies (Leroy et al. 2019) and the very large SDSS-based catalog of Salim et al. (2016, 2018).

<sup>d</sup> The bias in the calibration coefficient to convert extinction-corrected FUV emission from the Salim et al. (2018) fits to SFR for cells in SFR- $M_*$  space with at least 50 galaxies. The NUV calibration factor is extrapolated from FUV, so no independent bias is calculated.

of the total CO emission. Table 9 and the right panel of Figure 14 show that we would have arrived at similar aperture corrections using only WISE4 emission or our FUV+WISE4-based  $\Sigma_{\text{SFR}}$  estimates. Our stellar mass maps imply slightly higher corrections, on average, but  $\Sigma_*$  also correlates notably less well with  $I_{\text{CO}}$  than the SFR and mid-IR tracers. The right panel in Figure 14 shows that had we adopted one of these other plausible templates, we might expect the aperture correction to scatter by  $\pm 10\%$  (the gray band) relative to our preferred WISE3 values. We consider that the true uncertainty on the aperture correction is slightly less than this because of the strength of the WISE3-CO correlation.

Aperture corrections that assume an exponential disk structure have the large advantage that they can be applied with almost no multi-wavelength data. However, simple exponential disk models show a notably worse correlation with  $I_{\text{CO}}$  than any of the other templates. Among these, the template using our fitted  $l_*$  and the one using a scale length of 0.2  $r_{25}$  offer approximately equally good corrections. On average, these disk mod-

**Table 8.** Comparison Among Star Formation Rate Estimates

Other Estimate	$\log_{10}$ Median <sup>a</sup> [dex]	$\log_{10}$ Scatter [dex]
Integrated photometry	-0.02	0.02
NUV+WISE4	-0.05	0.01
FUV+WISE3	0.01	0.09
NUV+WISE3	-0.05	0.10
WISE3 only	-0.04	0.18
WISE4 only	0.0	0.09

<sup>a</sup> Reported as median of the other estimate over our estimate for galaxies where both are available.

NOTE—All columns report  $\log_{10}$  median and scatter among galaxy-integrated SFR estimates. Each row compares a different indicator to our adopted SFR. “Integrated photometry”: our fiducial estimated divided by those from Leroy et al. (2019) based only on integrated photometry. The other entries refer to changes in the adopted band combination. The table reports comparisons for a sample of local galaxies that have CO mapping including all PHANGS-ALMA targets. The results restricting to just PHANGS-ALMA are almost identical. See also Figure 12.

**Table 9.** PHANGS-ALMA CO(2–1) Aperture Corrections

Template	Correction <sup>a</sup>	Correlation <sup>b</sup>
WISE3 (adopted)	1.25 (1.10–1.57)	0.82
WISE4	1.16 (1.04–1.59)	0.77
SFR (FUV+W4) <sup>c</sup>	1.19 (1.07–1.84)	0.79
WISE1+ $\Upsilon_*^{3.4}$	1.41 (1.19–1.86)	0.53
Exponential disk (0.2 $r_{25}$ )	1.16 (1.07–1.55)	0.43
Exponential disk ( $l_*$ )	1.36 (1.10–1.90)	0.41
Exponential disk (0.6 $R_e$ )	1.18 (1.05–1.47)	0.18

<sup>a</sup> Median and 16<sup>th</sup> to 84<sup>th</sup> percentile range for the aperture correction to be applied to the PHANGS-ALMA data.

<sup>b</sup> Linear correlation coefficient relating the template band to CO intensity in the region with coverage. Each galaxy is first normalized by the median template-to-CO ratio.

<sup>c</sup> Best SFR estimate. This uses FUV+WISE4 when available. Otherwise we adopt NUV+WISE4 if available and WISE4 otherwise.

NOTE—This table reports results only for PHANGS-ALMA. The correction gives the implied aperture corrections. Correlation indicates how well each band linearly predicts CO(2–1) emission for each individual galaxy inside the coverage.

els yield about the same magnitude of correction factor as our other templates. However, Table 9 implies that they offer less precision, and Figure 14 shows that the derived correction using an  $l_*$  scale length exponential

disk scatters by 10–20% relative to the WISE3 correction.

#### 4.6.2. Checks on CO(2–1) Luminosity

Figure 15 compares our derived CO fluxes and intensities to previous CO(2–1) mapping using the IRAM 30-m telescope<sup>10</sup>. In total, PHANGS-ALMA covers 18 galaxies that have also been mapped in the CO(2–1) line using the HERA array receiver on the IRAM 30-m telescope (Schuster et al. 2007). These maps are part of the HERACLES survey (Leroy et al. 2009, 2013b) and a follow-up survey (A. Schruba et al. in preparation). They cover wider area than the ALMA maps, but with higher noise levels. The calibration of the IRAM 30-m maps depends on “chopper wheel” observations and a main beam efficiency correction, while both the ALMA interferometric and total power data are pinned to monitoring of ALMA calibrators with monitored fluxes.

The IRAM 30-m maps have native resolution of  $13.3''$ , so we must smooth the PHANGS-ALMA data to make a direct comparison. For this exercise, we used a  $17''$  version of both cubes. We aligned both cubes for each galaxy onto the same astrometric and spectroscopic grid. We then apply an identical mask to each galaxy in each survey to create matched resolution, matched velocity-interval integrated intensity maps, along with associated uncertainty maps.

The top left panel in Figure 15 compares the integrated fluxes between the IRAM and ALMA maps. For the calculation in the top left panel, we apply the WISE3-based aperture corrections discussed above. We find a median PHANGS-to-HERA ratio of 0.97, or  $-0.01$  dex, with a robustly-estimated galaxy-to-galaxy scatter of 0.06 dex, or about 15%. The right panel shows galaxy integrated ratios. We calculate these only integrating over the area where both maps have coverage. These also show good agreement. The statistical uncertainty in these ratios is low,  $< 1\%$  in almost all cases, so the observed scatter reflects mostly calibration or other systematic uncertainties. The 16<sup>th</sup> to 84<sup>th</sup> percentile range, illustrated in gray, spans from 0.87 ( $-0.06$  dex) to 1.06 (0.025 dex) with the median 0.97 ( $-0.01$  dex). These suggest an overall  $\lesssim 5\%$  difference in the calibration scale with about  $\pm 10$ –15% calibration uncertainty in any given galaxy. We mark four galaxies with strongly discrepant fluxes in Figure 15: NGC 3351 (PHANGS-to-HERA ratio of 0.74), NGC 3627 (1.24), NGC 4579 (0.80), and NGC 4689 (1.32). The discrepancy in NGC 3627, which was observed by the IRAM 30-m under non-ideal conditions shortly after HERA commissioning, has been noted and explored by den Brok et al. (2021). They also noted a similar discrepancy

between new and old IRAM 30-m mapping projects for NGC 5194 (M51) which is not in our sample. den Brok et al. suggest that issues with the HERA data most likely account for these strong outliers.

The bottom rows in Figure 15 explore the pixel-to-pixel correspondence between the PHANGS and HERA data at  $17''$  resolution. Here we only consider where both maps have coverage and  $I_{\text{CO}} > 1 \text{ K km s}^{-1}$ <sup>11</sup> in both maps, a total of 1,161 independent  $17''$  lines of sight. Again, we find good overall agreement, here spanning two decades in CO intensity. The bottom right panel shows the distribution of ratios for individual pixels. Here we find a median ratio of 0.93 ( $-0.03$  dex) and a robustly estimated scatter of  $\pm 0.11$  dex or about  $\pm 28\%$ . Thus, the pixel data indicate a slight offset in overall flux scale, now with PHANGS  $\sim 7\%$  fainter than HERA.

Based only on the statistical errors associated with individual data, we expect a scatter of about  $\pm 0.05$  dex or about  $\pm 12\%$ , a range that we indicate by a blue bar in Figure 15. This is significantly smaller than the measured scatter, again indicating the presence of systematic uncertainties. The blue histogram shows a model that treats these systematic uncertainties as a  $\pm 20\%$  point-to-point multiplicative uncertainty and also incorporates the measured statistical errors. Overall this matches the measured scatter well.

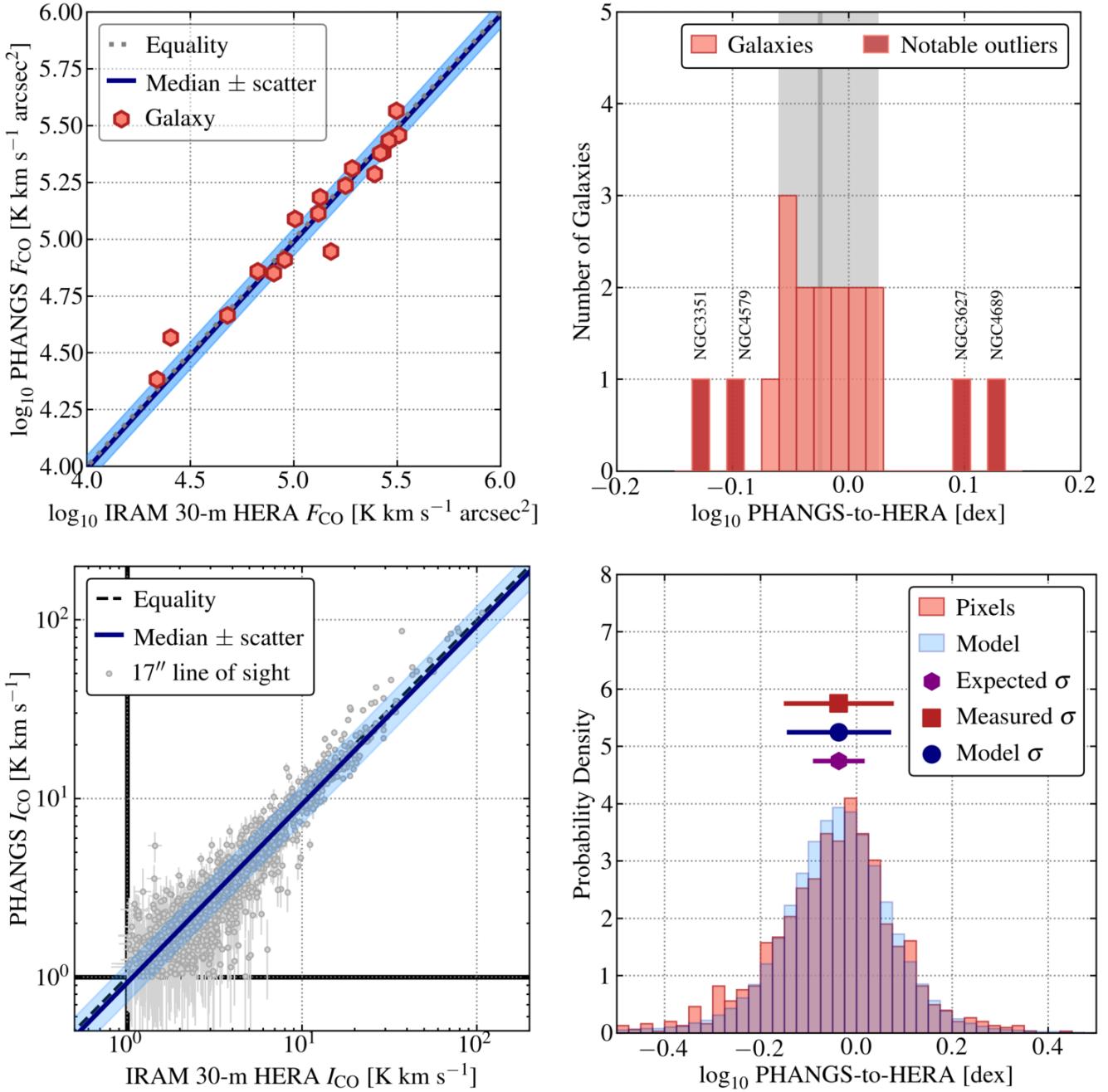
This  $\pm 20\%$ , or  $\pm 0.11$  dex, point-to-point scatter is somewhat higher than the galaxy-to-galaxy scatter seen in the top row. Indeed, if we subtract the galaxy-to-galaxy scatter and the statistical scatter from the measured scatter in quadrature we are left with about  $\pm 0.06$ –0.07 dex (about  $\pm 15\%$ ) point-to-point unaccounted scatter. Our best estimate is that this reflects uncertainties in the image reconstruction in both data sets. The HERA array receiver data suffer from variable pixel gains (see Leroy et al. 2009; den Brok et al. 2021) while the ALMA data have been partially reconstructed from interferometric imaging. Based on a detailed analysis and visual inspection of the 5 galaxies in den Brok et al. (2021), the HERA pixel gain uncertainties appear to be responsible for the most serious issues, and we primarily attribute this unaccounted scatter to the HERA data. This will appear, e.g., as mapping artifacts like striping or variations in the gain across the map. Such issues are common in single dish maps, especially those made with array receivers, and the identified magnitude of uncertainty,  $\pm 15\%$ , agrees with that inferred from jackknife tests carried out on the HERACLES data (Leroy et al. 2009).

Overall, this comparison paints the following picture:

1. The two data sets agree very well to first order.

<sup>10</sup> This part of the work is based on observations carried out with the IRAM 30m telescope. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain).

<sup>11</sup> This corresponds to  $S/N \gtrsim 5$  in the HERACLES maps, which represent the limiting signal to noise.



**Figure 15. Comparisons between CO(2-1) measurements from PHANGS-ALMA and previous IRAM 30-m CO(2-1) mapping.** All four panels compare measurements from our CO(2-1) data to CO(2-1) observations targeting the same galaxies using the HERA array receiver on the IRAM 30-m telescope (from HERACLES and follow-up work; Leroy et al. 2009, 2013b, and A. Schruba et al. in preparation). The *top left* panel compares integrated fluxes; in this figure we have applied our WISE3-based aperture correction to our PHANGS measurements. The *bottom left* panel compares line-integrated intensity between the two data sets at 17'' resolution. The right hand panels show ratios between the two data sets, the *top right* one for galaxy-integrated ratios and the *bottom right* one for pixel-by-pixel ratios. Overall the agreement is good. The median galaxy-integrated ratio is 0.95 ( $-0.02$  dex) with  $0.057$  dex, or about 14%, galaxy-to-galaxy scatter. We mark four strong outliers, listed in the text. The median pixel-wise ratio is 0.92 ( $0.037$  dex) with  $0.11$  dex, or about 29%, scatter. The model shown in the bottom panel includes the measured statistical uncertainty and adopts a 20% gain uncertainty. We attribute most of the non-statistical scatter to gain uncertainties associated with the IRAM 30-m observations, though formally these represent only relative calibration and image construction uncertainties.

2. There appears to be a  $\sim 3\text{--}7\%$  systematic offset in the flux scale between PHANGS and the HERA data, with the PHANGS data showing slightly lower flux. Our best estimate is that the ALMA data are better calibrated.
3. We observe  $\sim 15\%$  galaxy-to-galaxy scatter in the ratio of ALMA-to-HERA fluxes, which reflects the combined calibration uncertainty in the ALMA and IRAM 30-m data. This is largely consistent with the 10–15% calibration uncertainty expected for the HERA data (Leroy et al. 2009; den Brok et al. 2021) and the 5–10% uncertainty associated with ALMA band 6 observations.
4. At  $17''$  resolution, the point-to-point scatter between the two data sets is  $\sim 0.11$  dex. This is higher than expected from only the statistical noise or the galaxy-to-galaxy flux calibration uncertainties. We consider that this 0.11 dex scatter reflects roughly equal contributions from statistical uncertainty, flux calibration uncertainty, and point-to-point gain variations in the single dish map. The latter arises from pixel-to-pixel gain variations in the array receiver (see den Brok et al. 2021) and uncertainties in interferometric image reconstruction.

#### 4.7. Literature H I Masses

We draw integrated H I masses from the literature, specifically the homogenized compilation by HyperLEDA. We derive the H I mass from HyperLEDA’s 21 cm flux, applying no correction for opacity effects. Over the next year, we expect new and archival 21 cm line imaging using the VLA (P.I. D. Utomo) to improve our estimates of the H I mass for our targets.

#### 4.8. PHANGS-ALMA Sample Properties

Figures 16 and 17 show distributions of the derived sample properties reported in Tables 3 and 4. Figure 18 shows the PHANGS-ALMA sample in two commonly-considered parameter spaces for galaxy evolution: SFR– $M_*$  and  $R_e$ – $M_*$  space. In both plots, we distinguish the survey extension targets from the original sample. In addition to the quantities discussed already in this section, we show the ratios of  $L_{\text{CO}}$ -to-SFR and SFR-to- $M_*$ . We apply the aperture correction to  $L_{\text{CO}}$  before constructing these ratios. We also show the distribution of morphological T-type code (from HyperLEDA Makarov et al. 2014). Table 10 summarizes the distributions of properties for the PHANGS-ALMA sample.

**PHANGS-ALMA on Scaling Relations:** In Figure 18 we plot PHANGS-ALMA in red hexagons along with galaxies from the xCOLD GASS survey (blue points, from Saintonge et al. 2017) and a large sample of galaxies with  $d < 50$  Mpc in SFR– $M_*$  (top) and  $R_e$ – $M_*$

**Table 10.** PHANGS-ALMA Observed Property Summary

Property	Percentile				
	5	16	50	84	95
$i$ [ $^{\circ}$ ]	23	30	47	65	72
$d$ [Mpc]	3.7	9.9	15.1	19.3	20.7
$\log_{10} M_*$ [ $M_{\odot}$ ]	9.45	9.74	10.36	10.72	10.94
$\log_{10} \text{SFR}$ [ $M_{\odot} \text{ yr}^{-1}$ ]	-0.89	-0.57	-0.09	0.58	0.74
$\log_{10} R_e$ [kpc]	0.17	0.29	0.53	0.68	0.78
$\log_{10} L_{\text{CO}}^a$	6.97	7.66	8.29	8.95	9.11
Morphological T-type	-1.7	1.2	4.0	6.0	7.0
$\log_{10} \text{SFR}/M_*$ [ $\text{yr}^{-1}$ ]	-11.27	-10.82	-10.18	-9.86	-9.62
$\log_{10} L_{\text{CO}}/\text{SFR}^a$	7.71	7.96	8.31	8.56	8.63
$\Delta_{\text{MS}}$ [dex]	-1.02	-0.48	0.01	0.41	0.57
$\Delta_R$ [dex]	-0.40	-0.26	-0.07	0.09	0.16

<sup>a</sup> $L_{\text{CO}}/\text{SFR}$  has units of  $\text{K km s}^{-1} \text{ pc}^2 (\text{M}_{\odot} \text{ yr}^{-1})^{-1}$  and  $L_{\text{CO}}$  has units of  $\text{K km s}^{-1} \text{ pc}^2$ . Add  $\log_{10} \alpha_{\text{CO}}^{2-1}$  to the quoted value to convert to  $M_{\text{mol}}$  or molecular gas depletion time,  $\tau_{\text{dep}}^{\text{mol}}$ . A fiducial Milky Way  $\alpha_{\text{CO}}^{2-1} = 6.7 \text{ M}_{\odot} \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$  (Bolatto et al. 2013a, with  $R_{21} = 0.65$ ) implies  $\log_{10} \alpha_{\text{CO}}^{2-1} = 0.83$ .

NOTE—See Figures 16 and 17. Values report percentiles in the full set of targets, including extensions and archival data.

space (bottom). In both plots we indicate fiducial scaling relations for late-type galaxies. In the top panel, this is the result of a linear fit to  $\sim 4,000$  local galaxies with  $M_*$  and SFR estimated from WISE and GALEX data that closely resemble our calculations (Leroy et al. 2019). We use this to define the star-forming main sequence as

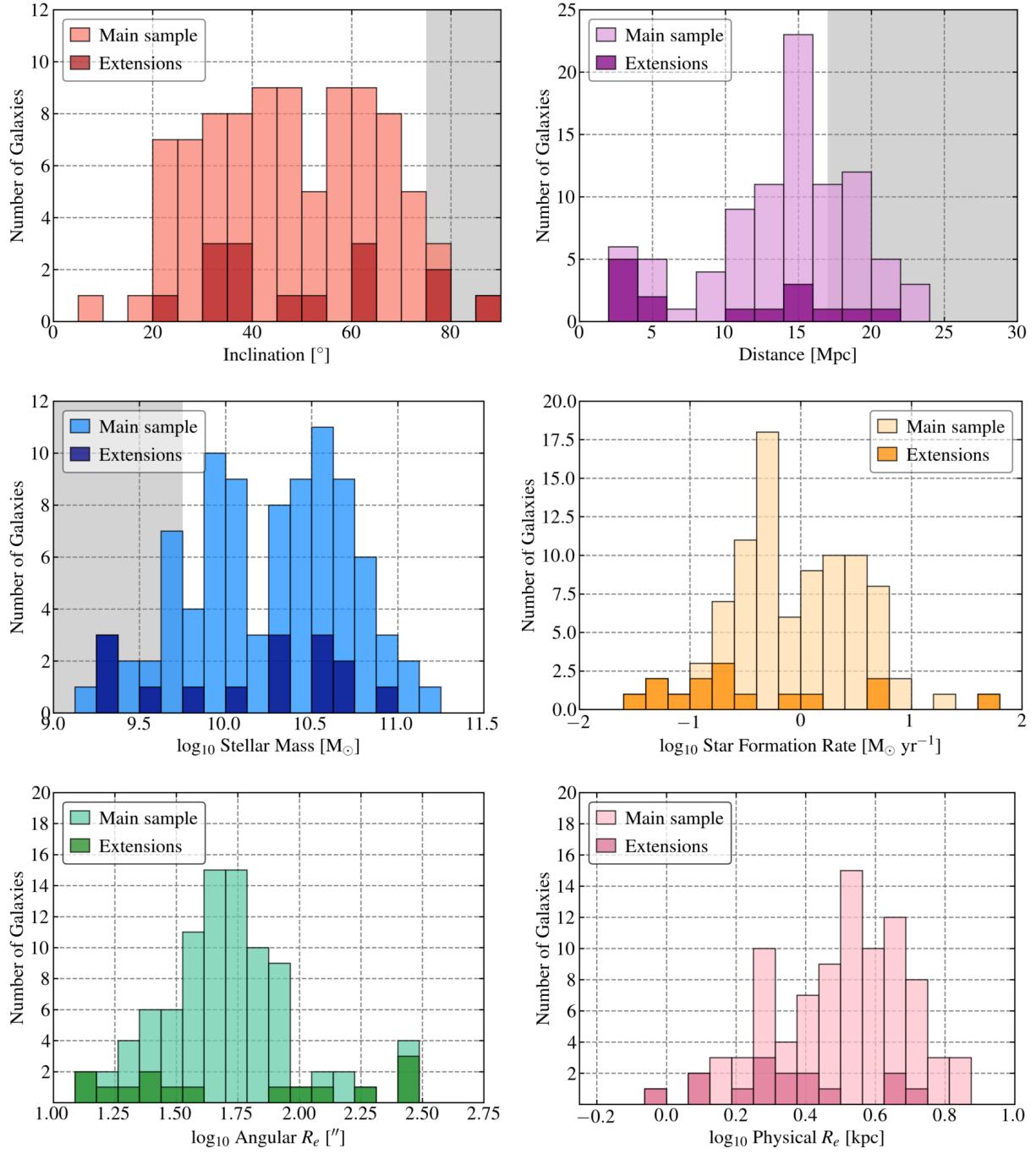
$$\log_{10} \frac{\text{SFR}_{\text{MS}}}{1 \text{ M}_{\odot} \text{ yr}^{-1}} = -0.32 \left( \log_{10} \frac{M_*}{1 \text{ M}_{\odot}} - 10.0 \right) - 10.17. \quad (7)$$

This is almost identical to the main sequence fit by Catinella et al. (2018) for the GASS sample, which has slope  $-0.344$  and a value of  $-10.17$  at  $\log_{10} M_* [\text{M}_{\odot}] = 10$ .

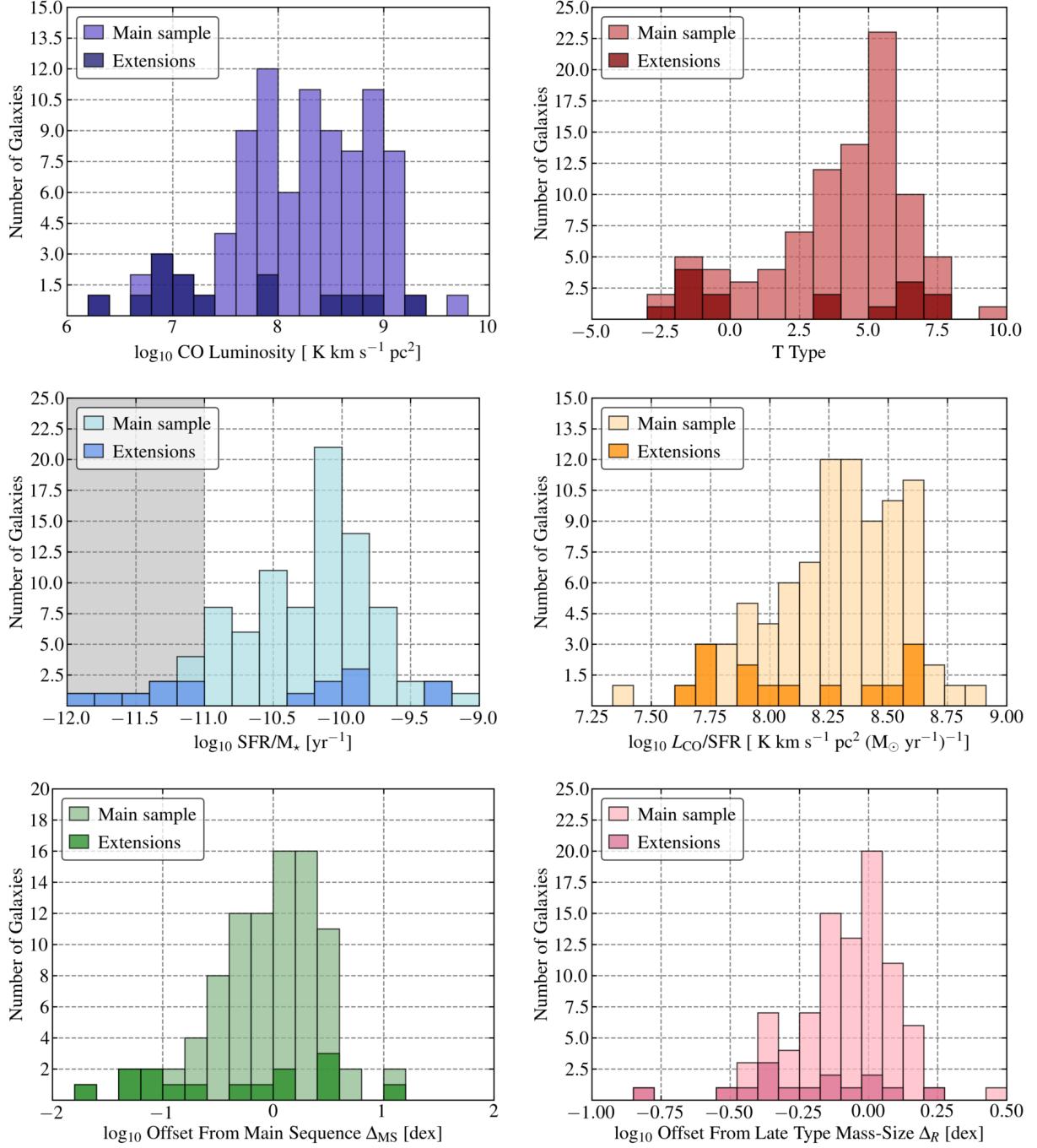
In the bottom panel, we show the mass-size relation found by Lange et al. (2015) for late-type galaxies as a black line. We plot their fiducial  $r$ -band relation, which uses the functional form defined by Shen et al. (2003). Note that the relation applies specifically to late-types but we do not restrict the galaxies plotted by morphology or Sérsic index. The large population of xCOLD GASS outliers at high  $M_*$  simply reflects the inclusion of early-type galaxies in that survey.

We calculate offsets from both scaling relations for each PHANGS-ALMA galaxy and show the distributions of  $\Delta_{\text{MS}}$  and  $\Delta_R$  in Figure 17. These are calculated as

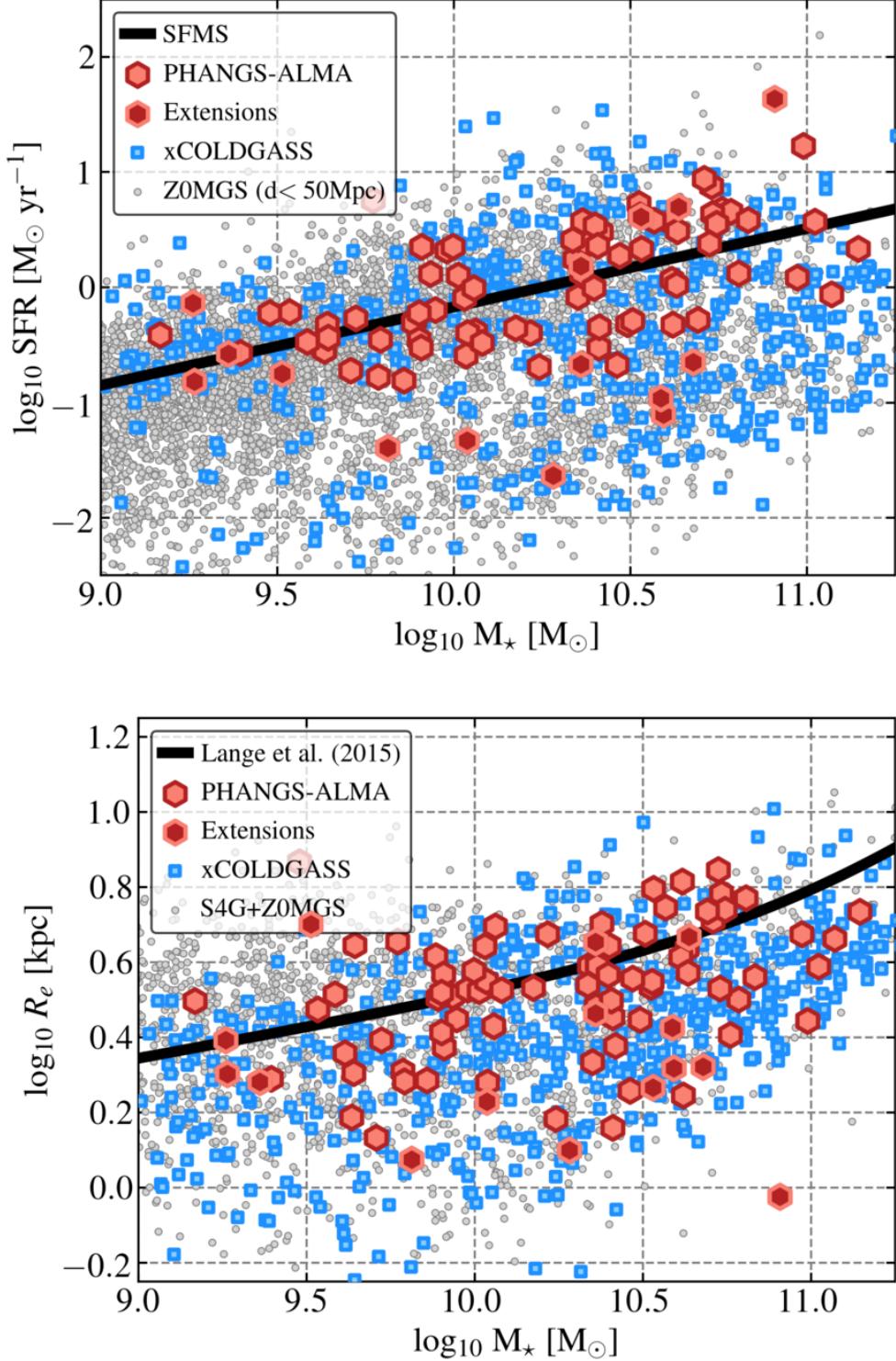
$$\Delta_{\text{MS}} \text{ [dex]} = \log_{10} \text{SFR} - \log_{10} \text{SFR}_{\text{MS}} (M_*) , \quad (8)$$



**Figure 16. Histograms showing PHANGS-ALMA sample properties.** The *top row* shows inclination and distance. Inclinations come mostly from Lang et al. (2020) and we adopt distances from Anand et al. (2021). The *second row* shows stellar mass (§4.3) and star formation rate (§4.5). The *third row* shows the stellar half mass radii (§4.4) in both angular and physical units. In panels relevant to sample selection, gray shading marks the region excluded by the main sample selection (§3). Note that our best estimates of each quantity have improved since selection, so that we selected some galaxies that we now believe miss the selection criteria. The darker histogram shows properties of galaxies in survey extensions that focus on early-type and very nearby galaxies.



**Figure 17. Histograms showing PHANGS-ALMA sample properties.** The *top row* shows CO luminosity (§4.6) and morphological T-type code from HyperLEDA (Paturel et al. 2003; Makarov et al. 2014). The *middle row* shows the ratio of SFR-to- $M_*$ , i.e., the specific star formation rate, and the ratio of  $L_{CO}$ -to-SFR, which can be scaled by  $\alpha_{CO}$  to calculate the molecular gas depletion time. The *bottom row* shows offsets from two scaling relations: the SFR- $M_*$  “star-forming main sequence” and the mass-size relation. Aperture corrections have been applied to the  $L_{CO}$  estimates used in the plots (§4.6.1). In panels relevant to sample selection, gray shading marks the region excluded by the main sample selection (§3). Note that our best estimates of each quantity have improved since selection, so that we selected some galaxies that we now believe miss the selection criteria. The darker histogram shows properties of galaxies in survey extensions that focus on early-type and very nearby galaxies.



**Figure 18. PHANGS-ALMA on the star-forming main sequence and size-mass relation.** PHANGS-ALMA targets in SFR- $M_{\star}$  and  $R_e$ - $M_{\star}$  space. PHANGS-ALMA targets appear as red hexagons. For comparison, galaxies observed as part of the xCOLD GASS survey are represented by blue squares (Saintonge et al. 2017). We show a large sample of local galaxies as gray dots. In the SFR- $M_{\star}$  space these are all galaxies with  $d < 50 \text{ Mpc}$  and measurement from Leroy et al. (2019). In the lower panel, we combined masses from Leroy et al. (2019) with sizes from Muñoz-Mateos et al. (2015), with the latter being the limiting quantity. Black lines show scaling relations relevant to late-type galaxies. In the top panel we show the star-forming main sequence as a solid black line, with the specific formula from Leroy et al. (2019). In the bottom panel we show the  $r$ -band result for late-type galaxies from Lange et al. (2015). We show the distribution of offsets from both relations in Figure 17.

and

$$\Delta_R [\text{dex}] = \log_{10} R_e - \log_{10} R_e^{\text{L15}} (M_\star) , \quad (9)$$

where  $R_e^{\text{L15}} (M_\star)$  refers to the effective radius predicted from the Lange et al. (2015) relation.

We also refer the reader back to Figure 1 where we place PHANGS-ALMA on several other scaling relations: molecular gas depletion time versus stellar mass, molecular gas mass fraction versus stellar mass, and molecular gas surface density versus star formation rate surface density and stellar mass surface density. Those figures also show xCOLD GASS and the full sample of local CO mapping targets considered in this section.

### PHANGS-ALMA Sample Properties:

PHANGS-ALMA aimed to select all relatively massive, nearby galaxies close to the “main sequence” of star-forming galaxies (§3). Despite the small volume available within  $d < 17$  Mpc and  $-75^\circ\delta < +25^\circ$ , Figures 16 through 18 and Table 10 show that the PHANGS-ALMA sample does achieve that goal. We achieve good sampling of more than a decade in stellar mass  $9.5 \lesssim \log_{10} M^* \lesssim 10.75$ . We also achieve good sampling of SFR. Our sample property estimates show good agreement with previous results, and our sample scatters around well-established scaling relations for disk galaxies. Our galaxies are mostly spirals, including both early- and late-type spirals, but not many irregular galaxies. We also cover a decade in specific star formation rate and offset from the star-forming main sequence, but a somewhat narrower range, a factor of roughly three, in  $L_{\text{CO}}/\text{SFR}$ . Despite uncertainties in distance and evolving distance estimates, most of our targets do lie within 19 Mpc, with  $\sim 16\%$  scattered out to  $\sim 21$  Mpc. And by construction our targets are overwhelmingly at moderate inclination, allowing resolved multiwavelength studies.

The figures, including the contrast with xCOLD GASS, also highlight the shortcomings of the sample for studying galaxy evolution. We do not include many extreme starbursts due to the simple fact that these tend to be rare. The classic local starburst galaxies, the members of the IRAS bright galaxy sample (e.g., Sanders et al. 2003), or GOALS sample (Armus et al. 2009) mostly lie at much greater distances. We do include many galaxies that host starburst nuclei, including some classic examples like NGC 0253 and NGC 4945. But those nuclear bursts are not strong enough to displace their parent galaxy far from the star-forming main sequence. We also include only a limited number of early-type galaxies. Massive early-type galaxies are relatively rare in the local volume and in any case these were excluded by our sample selection, though our extensions have begun to build in this direction. Finally, we do not push far into the regime of dwarf galaxies, with good coverage in mass down to only  $\log_{10} M_\star [\text{M}_\odot] \approx 9.5$ . At lower masses, conversion factor effects become large and

the required integration times rise, which would require adapting our observing strategy.

All of these directions are interesting and important. Starburst galaxies host high density, high pressure gas (e.g., Downes & Solomon 1998). Dwarf galaxies host sparse, faint, and isolated molecular clouds (e.g., Fukui & Kawamura 2010), while early-type galaxies show evidence for stabilized molecular gas disks with little or no star formation (Davis et al. 2014; Davis & McDermid 2017). However, we emphasize that most stars form in “main sequence” galaxies within a decade of  $\log_{10} M^* [\text{M}_\odot] \approx 10.6$  (e.g., Karim et al. 2011; Leslie et al. 2020). PHANGS-ALMA does a very good job of providing the first cloud-scale survey of a representative sample of such star-forming main sequence galaxies.

## 5. OBSERVATIONS

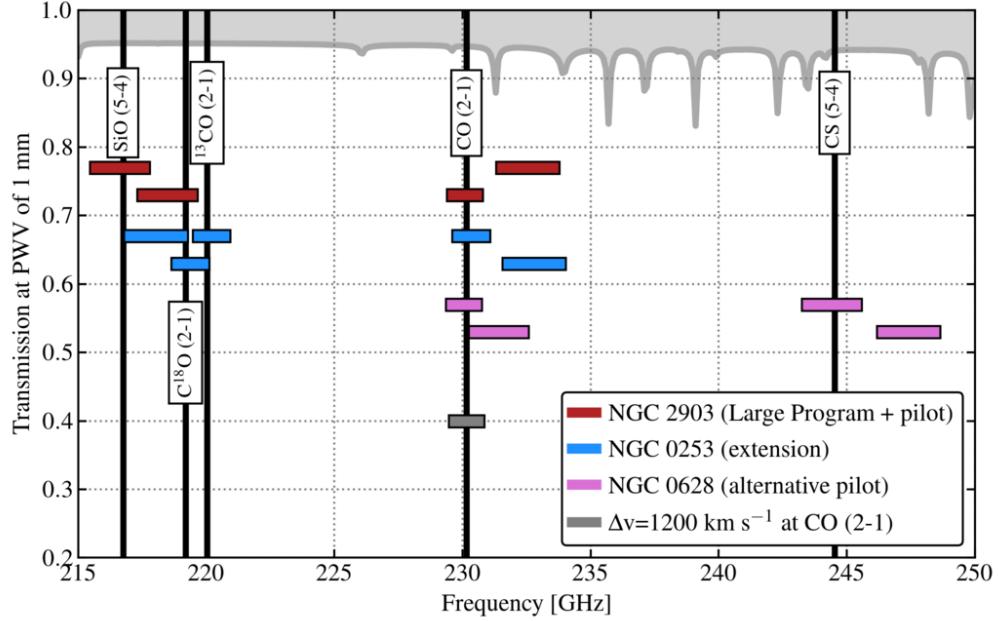
To meet the Science Goals described in §2, PHANGS-ALMA mapped the distribution and kinematics of molecular gas with high sensitivity, high physical resolution, and good fidelity across 90 nearby galaxies. To do this, we observed the  $J = 2 \rightarrow 1$  transition of CO arising from the galactic disk in each target. We used the ALMA main array of 12-m dishes in a compact configuration, which yields  $\sim 1''$  resolution at Band 6 ( $\nu \approx 230$  GHz), corresponding to  $\sim 100$  pc linear resolution, or about the size of a massive GMC at the typical distance to our targets. We covered the area of active star formation in each target using large mosaics that consisted of  $\sim 100$ – $450$  12-m pointings per galaxy. Despite short integration times for each pointing, we still achieve good  $u$ – $v$  coverage and sensitivities with the main array. Because nearby galaxies are extended and complex structured, we supplement the main array observations using both parts of the Morita Atacama Compact Array (ACA), the 7-m array and the single dish, or “total power” (TP), antennas.

This section describes and motivates our adopted strategy (§5.1). We also report details related to the actual execution of the observations (§5.2).

### 5.1. Observing Strategy

#### 5.1.1. Spectral Setup and Choice to Target CO(2–1)

We target the  $J = 2 \rightarrow 1$  rotational transition of CO, CO(2–1). This line is readily excited at temperatures and densities that characterize molecular clouds, and its spatial distribution in nearby galaxies correlates well with the fundamental CO(1–0) transition (e.g., see Braine & Combes 1992; Leroy et al. 2009, 2013b; Koda et al. 2020; den Brok et al. 2021). This choice represents a key practical element of our observing strategy. Given typical observed CO(2–1)-to-CO(1–0) line ratios of  $R_{21} \approx 0.5$ – $1$  (den Brok et al. 2021; Leroy et al. 2013b) and considering the array configurations and system temperatures at ALMA, mapping CO(2–1) is a factor of  $\sim 2$ – $4$  times faster than mapping CO(1–0) to the



**Figure 19. PHANGS-ALMA spectral setup.** Colored horizontal bars show the spectral windows observed using ALMA’s Band 6 receiver. The top, red bars show the setup for Large Program, the follow-up completion programs, and the two main pilot surveys (see Table 2 for details). This covers CO(2–1), C<sup>18</sup>O(2–1), and allocates the remaining bandwidth to continuum; though we mark SiO(5–4), it is generally too faint to be detected by our observations. Added flexibility for programs started from Cycle 6 onwards allowed us to also observe <sup>13</sup>CO(2–1) during some extensions to the project, as shown in the middle, blue bars. Pink bars show an alternative setup used for early pilot observations of NGC 0628. In this case we observe the high critical density line CS(5–4) instead of the rare isotopologue C<sup>18</sup>O(2–1). The gray shaded region at the top of the figure indicates the transmission of the atmosphere at a typical precipitable water vapor (PWV) level of 1 mm. Solid black lines indicate the Doppler-shifted frequency of observed spectral lines at a typical recessional velocity of  $v_{\text{LSR}} = 500 \text{ km s}^{-1}$ .

same mass surface density sensitivity and angular resolution. Moreover, the  $\nu = 230.538 \text{ GHz}$  rest frequency of <sup>12</sup>CO(2–1) lies in a favorable part of the atmosphere for ALMA, and can be observed effectively almost any time that ALMA is on sky.

We visualize the spectral setup of our observations in Figure 19. We observed CO(2–1) with a bandwidth of 937.5 MHz and a channel width of  $\Delta\nu = 244 \text{ kHz}$ . After taking into account the online Hanning smoothing, this corresponds to a native channel width of  $\sim 0.6 \text{ km s}^{-1}$  and a total velocity coverage of  $\sim 1200 \text{ km s}^{-1}$ . This channel width easily resolves emission lines from individual molecular clouds, which have typical full line width of  $\sim 10 \text{ km s}^{-1}$ . The velocity coverage easily encompasses the full emission from each of our targets. In practice, we image the cubes at a velocity resolution of  $\sim 2.5 \text{ km s}^{-1}$  to reduce data volume and improve the signal to noise and quality of the deconvolution.

We allocated the other correlator resources in several different ways over the course of the project. During the Large Program and the two main pilot projects (see Table 2), we configured the spectral setup to observe the mm continuum with the other three windows. We configured each to have the maximum possible 1.875 GHz

bandwidth. We placed one of these windows to cover the C<sup>18</sup>O(2–1) line at a spectral resolution of 2 MHz, corresponding to  $\sim 5\text{--}6 \text{ km s}^{-1}$  after taking the Hanning smoothing into account. The C<sup>18</sup>O isotopologue is several hundred times less abundant than <sup>12</sup>C<sup>16</sup>O (e.g., Wilson & Rood 1994a) and the C<sup>18</sup>O(1–0) line is often 30–100 times fainter than CO(1–0) in local galaxies (e.g., Jiménez-Donaire et al. 2019). Therefore the C<sup>18</sup>O line represents a target of opportunity that may be recovered in a few bright regions or via stacking. Though we imaged these data, this line is not included in the first PHANGS-ALMA data release.

In one early set of observations targeting NGC 0628, also illustrated in Figure 19, we targeted CS(5–4) instead of C<sup>18</sup>O(2–1). The main practical difference is that this places the main CO(2–1) line in the lower sideband instead of the upper sideband, which affects the spectral variation of the noise (see Leroy et al. 2021).

ALMA began allowing more flexible Band 6 spectral setups in Cycle 6 after the execution of most PHANGS-ALMA observations. Therefore, in the cases of some of the PHANGS-ALMA extensions using the ACA 7-m array, we also targeted the <sup>13</sup>CO(2–1) line. The <sup>13</sup>CO isotopologue is only  $\sim 30\text{--}120$  times less abundant than

$^{12}\text{CO}$  (e.g., Wilson & Rood 1994b; Milam et al. 2005), and the  $^{13}\text{CO}$  line is only  $\sim 10$  times fainter than  $^{12}\text{CO}$  (e.g., Roman-Duval et al. 2016; Cormier et al. 2018). This makes the  $^{12}\text{CO}/^{13}\text{CO}$  ratio a useful probe of optical depth and excitation (e.g., Cormier et al. 2018). As shown in Figure 19, observing  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  together requires using two somewhat narrower 937.5 MHz windows.

Finally, we note that ALMA employs “Doppler setting” rather than online Doppler tracking and records data in the topocentric frame. Our observed bandwidth is large enough that there is never any problem keeping the line emission from the galaxy in the full spectral window. However, this does imply that we must employ spectral regridding to construct the final velocity cubes.

### 5.1.2. Target Angular Resolution and Inclusion of Short Spacing Observations

We target a FWHM resolution of  $1''$ , which corresponds to  $\sim 100$  pc resolution at 20 Mpc or  $\sim 80$  pc at the outer edge of our selection function (see §3). We picked this resolution so that an individual beam matches the size of an individual massive GMC or association of massive molecular clouds. During Cycle 5, when the PHANGS-ALMA Large Program was approved, this resolution at Band 6 and  $\nu \approx 230$  GHz corresponded to the C43-2 array configuration. In practice, because ALMA executes observations within some tolerance, many of our observations also occurred in configurations similar to C43-1 or C43-3. Several factors related to data processing mean that our final data products have typical angular and physical resolutions slightly coarser than  $1''$  (see §5.2).

Our galaxies typically have full optical sizes of several arcminutes (§4). CO emission often shows extended, complex structure across the full area of a galaxy. The C43-2 main configuration recovers emission only from scales of  $\lesssim 10''$ . Thus short- and zero-spacing information is important to accurately reconstruct the intensity distribution (e.g., see the case of M51; Koda et al. 2009; Pety et al. 2013, and see details for PHANGS-ALMA data processing in Leroy et al. 2021). Therefore, we observed all targets with both components of the Morita Atacama Compact Array (ACA), the compact array of 7-m dishes and the 12-m single-dish telescopes used for “total power.”

The ratio of observing time among the main array, 7-m array, and total power antennas followed the standard observatory recommendations. For our fiducial configuration C43-2, the ALMA Cycle 5 technical handbook recommended an observing time ratio of 1:5:8.5 among the 12-m main array, the 7-m array, and the total power antennas.

In extension projects, we observed a subset of very nearby targets using only the ACA 7-m and total power

arrays<sup>12</sup>. These projects targeted galaxies within 5 Mpc, and typically at  $\sim 3\text{--}4$  Mpc, so that the  $\sim 7''$  resolution of the 7-m array at  $\nu = 230$  GHz already corresponds to 100–140 pc, reasonably matched to the main array resolutions for more distant targets.

### 5.1.3. Target Sensitivity

PHANGS-ALMA integrates for 20–30 s per pointing, quickly covering each field in each large mosaic with a modest integration time. To arrive at this strategy, we targeted a fiducial rms noise level of  $7.5 \text{ mJy beam}^{-1}$  or better at the frequency of the redshifted CO(2–1) line and a  $5 \text{ km s}^{-1}$  channel width. At our nominal  $1''$  target resolution, this target sensitivity corresponds to an rms brightness temperature sensitivity of  $\sim 0.17 \text{ K}$ .

We can also express our target sensitivity as a mass surface density sensitivity via:

$$\Sigma_{\text{mol}} = \alpha_{\text{CO}}^{1-0} R_{21}^{-1} I_{\text{CO}(2-1)} \cos i. \quad (10)$$

Here  $\alpha_{\text{CO}}^{1-0}$  is the CO(1–0) conversion factor,  $R_{21}$  is the CO(2–1)-to-CO(1–0) line ratio in Kelvin units,  $i$  is the inclination of the galaxy, and  $I_{\text{CO}(2-1)}$  is the line-integrated CO(2–1) intensity in  $\text{K km s}^{-1}$ . For a standard Milky Way  $\alpha_{\text{CO}}^{1-0} = 4.35 \text{ M}_\odot \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$  (Bolatto et al. 2013b), a CO(2–1)-to-CO(1–0) ratio of  $R_{21} = 0.65$  (den Brok et al. 2021; Leroy et al. 2013b), and a galaxy with inclination  $i$ ,

$$\Sigma_{\text{mol}} [\text{M}_\odot \text{ pc}^{-2}] = 6.7 I_{\text{CO}(2-1)} [\text{K km s}^{-1}] \cos i. \quad (11)$$

Then, for our target per-channel sensitivity of  $0.17 \text{ K}$  and if we assume a total line width of  $\Delta v = 5 \text{ km s}^{-1}$  for the spectral line then the  $1\sigma$  noise in  $\Sigma_{\text{mol}}$  is  $\Sigma_{\text{mol}} \approx 5.7 \text{ M}_\odot \text{ pc}^{-2}$  before any inclination correction.

For comparison, typical GMC surface densities range from a few times  $10 \text{ M}_\odot \text{ pc}^{-2}$  to a few times  $100 \text{ M}_\odot \text{ pc}^{-2}$ . Therefore, this sensitivity level allows us to detect every location where a GMC fills a reasonable fraction of the synthesized beam at good significance.

Under the same assumptions for  $\alpha_{\text{CO}}^{1-0} = 4.35 \text{ M}_\odot \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$ ,  $R_{21} = 0.65$ , and fiducial line width, our target sensitivity corresponds to  $1\sigma$  point mass sensitivity of about  $4 \times 10^3 \text{ M}_\odot$ ,  $1.5 \times 10^4 \text{ M}_\odot$ , and  $4.4 \times 10^4 \text{ M}_\odot$  at distances of 5, 10, and 17 Mpc, respectively. Our chosen target sensitivity is motivated by the canonical value for the characteristic mass of GMCs in the inner Milky Way,  $M_{\text{mol}} \gtrsim 10^5 \text{ M}_\odot$  (e.g., Blitz 1993), and thus, even in our most distant targets, we expect to detect individual moderately massive GMCs ( $M_{\text{mol}} \gtrsim 10^5 \text{ M}_\odot$ ) at  $2\text{--}3\sigma$  significance. Note that this  $\Delta v = 5 \text{ km s}^{-1}$  is the fiducial value used for sensitivity calculations. As described above, we observed at

<sup>12</sup> Projects 2018.1.01321.S, 2018.A.00062.S, and 2019.1.01235.S in Table 2.

finer spectral resolution than this and our nominal data products are constructed using  $\sim 2.5$  km s $^{-1}$  channels.

The achieved surface brightness sensitivity depends sensitively on the final beam size of the image, and our final resolutions tend to be slightly coarser than 1''. Therefore, in practice our images tend to have somewhat better sensitivity than this 0.17 K. We discuss the achieved properties of the PHANGS–ALMA data more below.

This target sensitivity drove our main array integration times to low values,  $\approx 12$ –24 s per pointing. Integrating on a single pointing under typical Band 6 conditions, achieving 7.5 mJy beam $^{-1}$  per 5 km s $^{-1}$  channel takes ALMA  $\sim 30$  sec using 43 antennas. After accounting for the effect of overlapping mosaic pointings, the actual nominal integration time per field dropped to  $\sim 12$  sec. Quantization of ALMA integration times and scheduling block creation tended to increase this per field integration time to  $\approx 18$ –24 s, which set our final sensitivity (see §7.1.2).

#### 5.1.4. Target Area

Our targets all have large extent compared to the 30'' primary beam of the 12-m ALMA antennas at  $\nu = 230$  GHz. We cover them using mosaics consisting of many individual pointings, typically 100–300 per galaxy. To select the exact target area, we use mid-IR emission, which we take to indicate the presence of recent star formation and thus the likely presence of molecular gas (see §4.6.1 for more discussion on this topic).

Specifically, when designing the field of view for the Large Program, we aimed to cover the full area of each galaxy that shows WISE3 intensity above 0.5 MJy sr $^{-1}$  at 7.5'' resolution. The pilot programs targeted the area with WISE4 intensity above 1 MJy sr $^{-1}$  at 15'' resolution, which yielded qualitatively similar results. Follow the tight scaling between mid-IR intensity and CO intensity discussed in §4.6.1, this design implies that we will cover all parts of the galaxy where  $I_{\text{CO}}^{2-1} \gtrsim 1$  K km s $^{-1}$  at low resolution. Roughly, this means we cover all region where we expect molecular gas to dominate the cold gas budget in the galaxy (e.g., see Leroy et al. 2008) and also all regions where the average intensity over a large area exceeds our  $1\sigma$  sensitivity at high resolution.

As illustrated in Figure 20, we defined the actual coverage region for each galaxy as a rotated rectangle, with the position angle and extent matched to the WISE contour by eye. In a few cases, a rectangle was not a good match to the exact WISE contour, e.g., because a small amount of mid-IR emission would force us to dramatically expand the mosaic and thus the integration time. In these cases, we made a judgment call to adjust the exact field of view to a more practical value.

In practice, as §4.6.1, Table 9, and Figure 20 show, our adopted field of view typically encloses 80% of the WISE3 emission and almost 90% of the WISE4 emission from each target. The aperture corrections in Tables 4

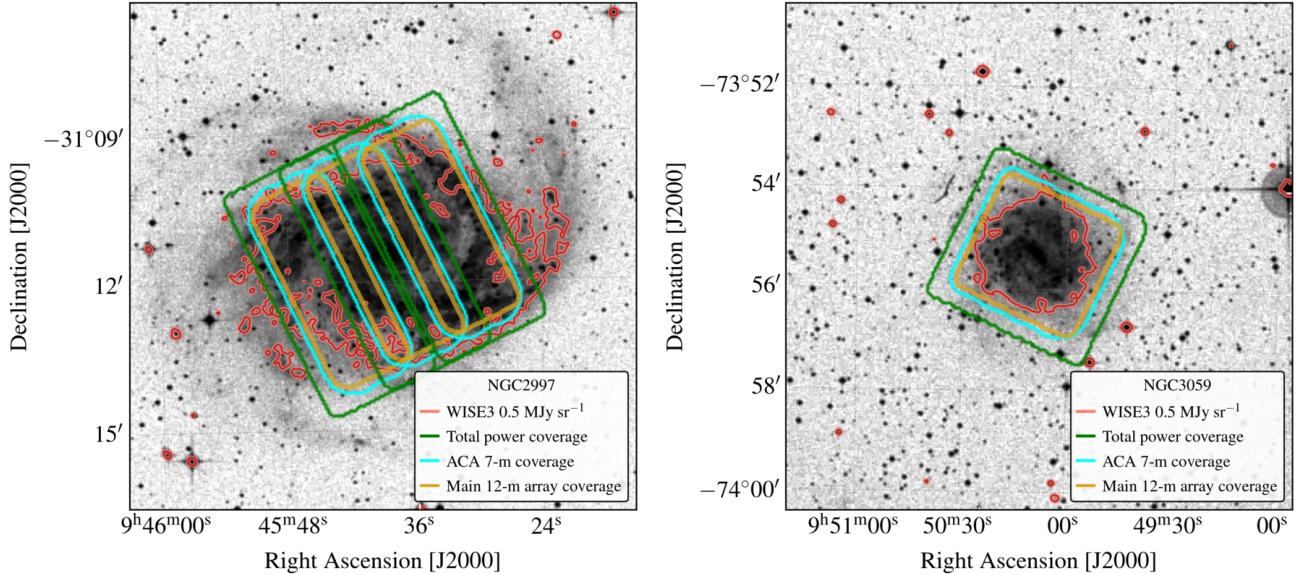
give a more detailed estimate of how well our field of view captures the area of interest for each target. These values should indicate the fraction of total CO emission captured by our field of view. Qualitatively, our maps are more compact than those made by array receivers on single-dish telescopes (e.g., HERACLES; Leroy et al. 2009), but they do cover almost all of the active star formation and molecular gas in each target.

Within each rectangle, the ALMA observing tool places hexagonally-packed fields with pointings spaced by  $\lambda/\sqrt{3}D$  or  $\sim 13''$ . During each of our observing cycles, ALMA allowed no more than 150 pointings per “Science Goal.” This restriction placed a maximum size on individual mosaics, and also defined the field of view for the 7-m and TP observations. Many of our targets required two or even three of these maximum-sized 150 pointing mosaics to cover the mid-IR defined regions. For these targets, we observed the galaxy in several “parts.” All parts shared the same spectral setup, but we defined each part as a separate Science Goal with a distinct field of view. The fields of view of different Science Goals targeting the same galaxy overlapped one another by  $\sim 10''$  along the shared edge to ensure uniform sensitivity. The separate Science Goals were scheduled, observed, and mostly reduced separately. The left panel in Figure 20 illustrates a three-part case, NGC 2997. In this galaxy, ALMA observed each of the three rectangular fields separately. The figure demonstrates the overlap between fields that we use to achieve a good final image of the whole galaxy. The right panel shows NGC 3059, where a single  $< 150$  pointing mosaic covers the whole galaxy. Both panels show the rotated rectangular fields that we use and the final coverage of the different arrays, with the total power data slightly extended relative to the ACA 7-m data and both covering the footprint of the 12-m data.

As we discuss below, observing the galaxies in separate parts led to uneven resolution in some cases. Our processing matches the angular resolution of the parts at the coarsest common beam, which led to a modest overall loss in resolution in our data products. In practice, a modest difference in surface brightness sensitivity between the individual parts of a mosaic is likely to be the most noticeable consequence of these multi-part observations.

Following standard ALMA practice, the total power observations used an “on-the-fly mapping” approach to cover the area of interest. Individual observations used multiple antennas to cover the same area. The total power observations were descended from the same Science Goals as the interferometric observations. As a result, when a galaxy was observed in multiple parts spread across several Science Goals, the total power observations were similarly divided to cover individual parts of the galaxy.

#### 5.1.5. Notes on Archival Observations



**Figure 20. PHANGS-ALMA coverage illustrated.** Digitized Sky Survey images of two of our targets: (left) NGC 2997, which we observe in three “parts,” each covered by separate ALMA Science Goals, and (right) NGC 3059, which we covered with a single Science Goal. Red contours show the WISE Band 3 ( $\lambda = 12 \mu\text{m}$ )  $0.5 \text{ MJy sr}^{-1}$  contour, which we use to define our target area. Contours indicate the area covered by the total power, 7-m, and 12-m array mosaics.

We periodically reviewed the ALMA archive to find all observations that match our PHANGS-ALMA setup. We searched for observations of nearby,  $d \lesssim 20$  Mpc, galaxies targeting CO(2–1) and covering all or most of the galaxy’s star-forming area at  $\sim 100$  pc resolution. As of our latest review, in autumn 2020, we identified four galaxies that closely matched the PHANGS-ALMA setup and included these in our sample. The targets are NGC 1365, NGC 5128 (Centaurus A), NGC 5236 (M83), and NGC 7793; and we list the project codes and PIs in Table 2. We include NGC 1365 “as is” from the archive. For Centaurus A, the CO(2–1) part of the total power data in the archive was not usable due to a problem in the observations. We obtained new ACA mapping in ALMA Cycle 7 (P.I. C. Faesi, see Table 2). NGC 5236 was observed over two cycles using 9 separate Science Goals. The data have large extent, but otherwise match the PHANGS-ALMA setup exactly. The archival NGC 7793 observations obtained only 12-m main array data. We obtained 7-m and total power data in ALMA Cycle 7 (P.I. C. Faesi, see Table 2).

For this version of a PHANGS-ALMA data release, we considered only the CO(2–1) line and only processed wide-area maps of star-forming galaxies. By including the CO(1–0) and CO(3–2) lines and relaxing the areal coverage requirements, many more archival observations of molecular line emission in nearby galaxies could be included. The PHANGS-ALMA data processing pipeline (Leroy et al. 2021) can straightforwardly handle these data, and we intend to return to this in future work.

## 5.2. Execution of Observations

Including all pilots and extensions, observations for PHANGS-ALMA spanned six years, from early 2013 until the end of 2019. Tables 11, 12, and 13 report details of the individual interferometric and total power observations that passed observatory quality assurance and were included in our data processing. In the tables, we report the target name, the relevant project code, the date of the observation, the duration of the observation, the number of antennas used, the elevation range, and the precipitable water vapor (PWV) associated with the observation. For the interferometric observations, we also report the minimum, maximum, and median baseline length associated with the observation. For the total power observations, we report the Jy/K gain factor supplied by the observatory.

We visualize the dates and the properties of individual observations in Figures 21, 22, and 23. Table 14 summarizes the observations, giving the typical duration, observing conditions, and  $u-v$  coverage with each array. We also report the total number of observations, 184 for the 12-m array, 479 for the 7-m array, and 823 for the total power antennas. Taking into the account the duration of each observation, the total array time for the survey, including pilots, archival observations<sup>13</sup>, extensions, and the Large Program, is 177 hours in the main array, 652 hours in the 7-m array, and 855 hours of

<sup>13</sup> Currently the archival total power observations are not included in these plots and logs.

**Table 11.** PHANGS–ALMA Total Power Observation Log

Target	Project	Start	End	Min. El.	Max. El.	<i>N</i> <sub>ant</sub>	PWV	Gain
		(MJD)	(MJD)	(°)	(°)		(mm)	(Jy/K)
ESO097-013	2018.1.01321.S	58551.230	58551.281	43.0	47.0	3	1.80	41.2
ESO097-013	2018.1.01321.S	58551.282	58551.333	47.0	48.0	3	1.80	41.1
IC1954	2017.1.00886.L	58124.013	58124.062	60.0	61.0	3	1.10	44.4
IC1954	2017.1.00886.L	58126.046	58126.095	57.0	61.0	3	2.00	44.3
IC1954	2017.1.00886.L	58132.986	58133.036	60.0	61.0	3	2.30	44.5
IC1954	2017.1.00886.L	58133.036	58133.085	56.0	61.0	3	2.20	44.6
IC1954	2017.1.00886.L	58133.951	58134.000	56.0	61.0	3	3.10	43.9
IC1954	2017.1.00886.L	58134.001	58134.051	59.0	61.0	3	3.00	44.2
IC1954	2017.1.00886.L	58136.028	58136.077	56.0	61.0	3	1.80	44.1
IC5273	2017.1.00886.L	58115.863	58115.920	70.0	75.0	3	2.80	43.5

NOTE—This table is a stub. The full version of the table appears as a machine-readable table in the online version of the paper. Columns give: Target—the target of the observations; Project—the ALMA project code; Start and End—the beginning and end of the observations reported as Modified Julian Day (00:00 on Jan 1, 2018 is MJD 58119); Min. and Max. El.—the minimum and maximum elevation during the observation; *N*<sub>ant</sub>—the number of antennas participating in the observations; PWV—the precipitable water vapor during the observation; Gain—the Jansky per Kelvin calibration for the observation.

**Table 12.** PHANGS–ALMA 7-m Observation Log

Target	Project	Start	End	Min. El.	Max. El.	<i>N</i> <sub>ant</sub>	PWV	Min. <i>u</i> – <i>v</i>	Med. <i>u</i> – <i>v</i>	Max. <i>u</i> – <i>v</i>
		(MJD)	(MJD)	(°)	(°)		(mm)	(m)	(m)	(m)
CIRCINUS.1	2018.1.01321	58501.398	58501.428	43.5	60.7	12	2.67	7.9	20.3	37.4
CIRCINUS.2	2018.1.01321	58501.429	58501.459	44.6	70.5	12	2.45	8.2	20.9	36.8
IC1954	2017.1.00886	58055.221	58055.281	52.3	63.3	11	1.14	8.2	21.1	43.8
IC1954	2017.1.00886	58056.211	58056.270	53.5	61.2	11	1.15	8.3	21.0	43.7
IC1954	2017.1.00886	58055.281	58055.338	49.0	75.8	11	1.08	7.3	20.9	43.9
IC5273	2017.1.00886	58032.085	58032.136	63.1	80.4	11	0.51	8.6	24.0	47.7
IC5273	2017.1.00886	58037.970	58038.020	55.5	77.1	10	1.40	7.5	20.8	46.0
IC5273	2017.1.00886	58038.045	58038.095	51.8	85.0	10	1.38	8.8	21.9	47.5
IC5332	2015.1.00925	57559.388	57559.445	65.5	77.0	8	0.84	8.7	21.2	42.7
IC5332	2015.1.00925	57559.445	57559.501	54.7	73.2	8	0.75	8.0	21.2	43.1

NOTE—This table is a stub. The full version of the table appears as a machine readable table in the online version of the paper. Columns give: Target—the target of the observations; Project—the ALMA project code; Start and End—the beginning and end of the observations reported as Modified Julian Day (00:00 on Jan 1, 2018 is MJD 58119); Min. and Max. El.—the minimum and maximum elevation during the observation; *N*<sub>ant</sub>—the number of antennas participating in the observations; PWV—the precipitable water vapor during the observation; Min., Med., and Max. *u*–*v*—minimum, median, and maximum baseline.

total power observations (typically using 2–3 antennas at a time). As Figure 21 illustrates, the bulk of these observations occurred in 2016–2019, coinciding with the execution of the two main pilot programs and the Large Program.

### 5.2.1. Calibration

For the interferometric data, calibration followed the standard ALMA procedures with no special requirements. The observatory selected the primary and secondary calibrators that were used to calibrate the bandpass, flux scale, and phase response of each antenna.

The cycle time also followed standard observatory procedures, and the observations employed the standard ALMA calibrations to measure and correct for atmospheric water vapor and to measure the system temperature.

For the total power data, we provided the observatory with reference “OFF” positions. We chose these to be well-separated from the galaxy, typically 5'–10' away and always well outside the optical radius of the galaxy. We verified from mid-IR WISE and optical imaging that the OFF position did not coincide with any emission from the galaxy or a neighboring galaxy. For CO(2–1)

**Table 13.** PHANGS-ALMA 12-m Observation Log

Target	Project	Start	End	Min. El.	Max. El.	$N_{\text{ant}}$	PWV	Min. $u-v$	Med. $u-v$	Max. $u-v$
		(MJD)	(MJD)	( $^{\circ}$ )	( $^{\circ}$ )		(mm)	(m)	(m)	(m)
IC1954	2017.1.00886	58299.543	58299.575	55.2	64.1	44	0.67	13.1	87.9	303.4
IC5273	2017.1.00886	58256.379	58256.409	54.6	61.9	45	1.24	12.4	82.5	284.4
IC5273	2017.1.00886	58264.398	58264.428	59.2	70.5	44	1.24	13.8	86.6	297.8
IC5332	2015.1.00925	57558.316	57558.360	51.1	65.6	43	0.71	13.1	201.0	641.0
NGC0628	2012.1.00650	56316.907	56316.949	36.3	83.1	29	3.74	12.4	118.2	367.9
NGC0628	2012.1.00650	56781.615	56781.672	46.1	59.8	30	2.34	13.5	171.1	509.5
NGC0628	2012.1.00650	56781.687	56781.743	36.7	60.1	30	2.56	12.0	154.7	532.4
NGC0628	2012.1.00650	56782.487	56782.546	28.7	75.6	32	1.75	16.9	139.6	455.7
NGC0628	2012.1.00650	56794.721	56794.769	19.8	68.1	32	0.44	12.7	161.8	638.9
NGC0628	2012.1.00650	56994.130	56994.189	29.4	55.9	33	0.62	11.6	75.3	279.4

NOTE—This table is a stub. The full version of the table appears as a machine readable table in the online version of the paper. Columns give: Target—the target of the observations; Project—the ALMA project code; Start and End—the beginning and end of the observations reported as Modified Julian Day (00:00 on Jan 1, 2018 is MJD 58119); Min. and Max. El.—the minimum and maximum elevation during the observation;  $N_{\text{ant}}$ —the number of antennas participating in the observations; PWV—the precipitable water vapor during the observation; Min., Med., and Max.  $u-v$ —minimum, median, and maximum baseline.

the recessional velocity of our galaxies means that they are well-separated from any foreground Milky Way emission.

In a subset of targets the CO(2–1) total power observations showed significant contamination by an atmospheric ozone feature at 229.575 GHz. In cases where the ozone feature overlapped the emission from the galaxy, this sometimes contaminated the total power data to the point where they could not be used. We present a discussion of this effect in the appendix of Leroy et al. (2021) and in Herrera et al. (2020), and a more detailed summary is presented in two memos by A. Usero<sup>14</sup>. This telluric ozone contamination reflects imperfect sky subtraction. To improve the situation, we worked with the observatory to re-observe the most strongly affected cases using a reference fixed in elevation, rather than in the equatorial frame. These observations were also conducted with more stringent requirements on the weather, ensuring less atmospheric contamination. We mark the affected targets in Table 16. See Leroy et al. (2021) for more details on the effect and processing.

### 5.2.2. Observing Time and Conditions

Figure 22 visualizes the conditions of our individual observations, which are also summarized in Table 14 and reported in detail in Tables 11, 12, and 13. In brief, individual observations typically lasted 1–1.4 h across all arrays, and most observations took place at mean elevation of 50–70° with precipitable water vapor (PWV)  $\sim 0.5$ –2 mm. The tables and figures report more exact

**Table 14.** PHANGS-ALMA CO(2–1) Observing Conditions

Quantity	Array		
	12-m	7-m	TP
Number of observations	184	479	823
Total time [h]	177	652	855
Median conditions during observations <sup>a</sup> ...			
... duration [h]	$1.0^{+0.2}_{-0.2}$	$1.4^{+0.1}_{-0.2}$	$1.2^{+0.1}_{-0.4}$
... PWV <sup>b</sup> [mm]	$1.4^{+0.8}_{-0.7}$	$1.0^{+0.8}_{-0.5}$	$0.9^{+1.0}_{-0.4}$
... mean elevation [°]	$60^{+9}_{-7}$	$58^{+10}_{-7}$	$54^{+16}_{-9}$
... min. baseline [m]	$13^{+2}_{-1}$	$8^{+0.7}_{-0.8}$	...
... median baseline [m]	$97^{+34}_{-14}$	$21^{+2}_{-1}$	...
... max. baseline [m]	$388^{+99}_{-91}$	$44^{+4}_{-5}$	...

<sup>a</sup>Values quote median and error bars give 16%–84% range.

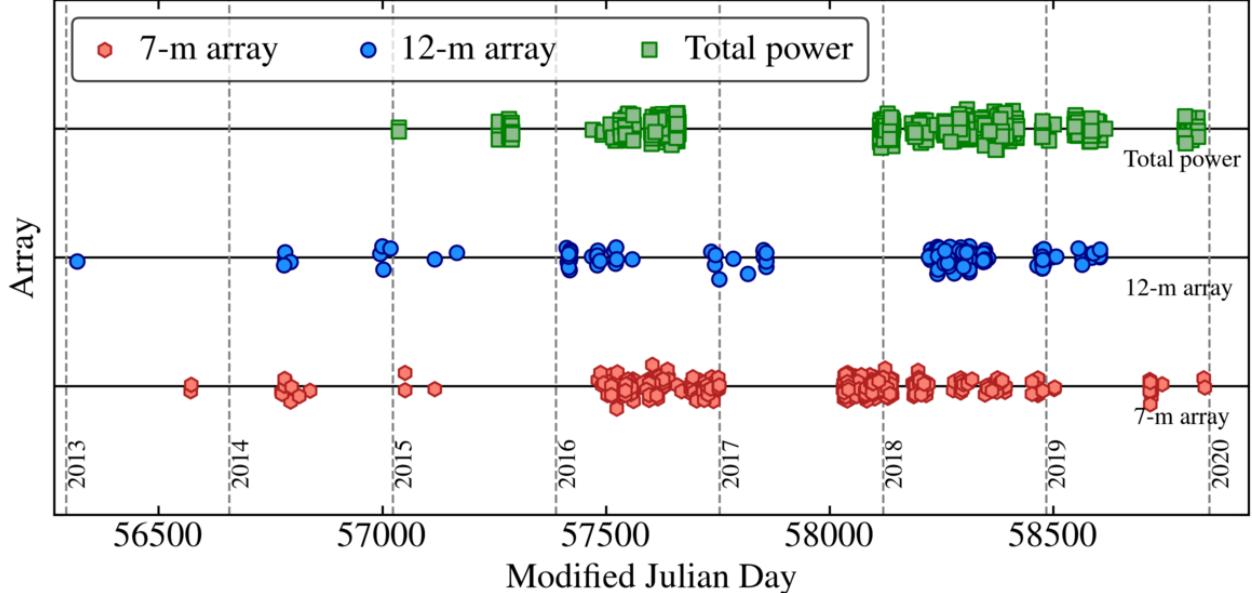
<sup>b</sup>Precipitable water vapor.

numbers, but overall this reflects almost ideal scheduling on behalf of the array. The high elevation imply limited shadowing, and minimal atmospheric contamination, and the PWV of the observations is almost always less than the  $\sim 1.8$  mm associated with ALMA’s 5<sup>th</sup> octile of weather conditions, the nominal threshold for 230 GHz (Band 6) observations.

### 5.2.3. Achieved $u-v$ coverage and synthesized beam

Tables 13 through 14 report the minimum, maximum, and median  $u-v$  baseline length of each observation. We visualize this in Figure 23, which shows these values for each observation. On average, the 12-m baselines span

<sup>14</sup> Available at <https://sites.google.com/view/phangs/publications>.



**Figure 21. PHANGS-ALMA observing dates (including archival observations for the 12-m and 7-m arrays but not the total power data).** Date of observations of PHANGS-ALMA observations with the 12-m, 7-m, and total power arrays. We include all pilot and archival observations processed as part of the survey described in this paper.

13–388 m with median baseline length 97 m. In terms of wavelength, this translates to a range of  $\sim 10$ –3000  $\text{k}\lambda$ , corresponding to  $\sim 20''$ –0.7'', with a mean of  $\sim 75$   $\text{k}\lambda$ , corresponding to  $\sim 2.8''$ . The 7-m data typically span baselines 8–44 m with a mean of 21 m. This maps to 6–34  $\text{k}\lambda$ , corresponding to 33''–6'', with a mean of 16  $\text{k}\lambda$ , corresponding to  $\sim 13''$ . Figure 23 shows that there is some scatter in the median 12-m baseline length across our data, and significant, about  $\pm 25\%$ , scatter in the maximum baseline for the 12-m data. However, the 7-m coverage remains stable, consistent with the fixed positions of the ACA, and in all cases there is good overlap between the 12-m baseline range, the 7-m baseline range, and the diameter of the total power antennas.

The baseline range only gives the approximate response of the array. Figures 24, 25, and 26 show the  $u$ – $v$  coverage, distribution of collecting area vs. baseline distance, and achieved synthesized beam in more detail. All of these figures focus on one example galaxy, NGC 3059, which also appears in Figure 20.

Figures 24 and 25 show the  $u$ – $v$  plane coverage for one 7-m pointing (red) and one 12-m pointing (blue). In Figure 24, we see that the 7-m data, which combine several distinct observations, achieve good rotation synthesis. The 12-m data do not, but the figure shows the excellent instantaneous  $u$ – $v$  coverage of ALMA. The 12-m baselines do an outstanding job of filling out the  $u$ – $v$  plane even without long enough integrations to achieve rotation synthesis. In Figure 25 we plot the same data as in Figure 24, but now showing the distribution of collecting area as a function of  $u$ – $v$  baseline length for individ-

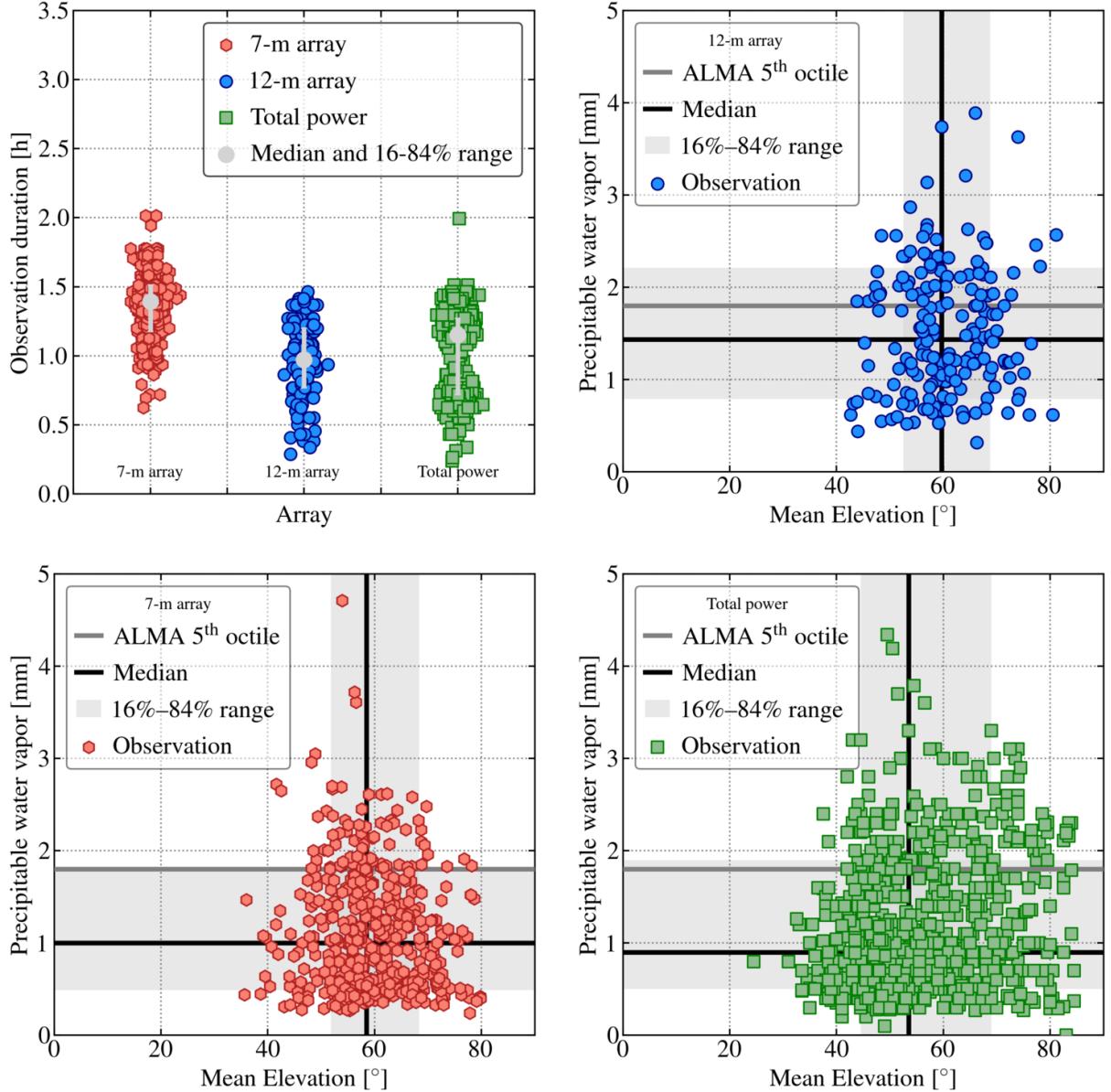
ual integrations. We use collecting area as a proxy for sensitivity, assuming fixed atmospheric conditions and integration length. The figure shows a relatively even distribution of sensitivity as a function of  $\log_{10} u$ – $v$  distance, and so spatial scale, out to  $\sim 300$  m. Similar to Figure 23, but now in sharper detail, we see good sensitivity to a range of spatial scales and overlap between the three types of telescopes (note that we only indicate the  $u$ – $v$  range of the total power data, not the collecting area).

Figure 26 shows the sampling of the  $u$ – $v$  coverage. Here we show the synthesized beam for a single channel in NGC 3059. This is the same target used for Figures 24 and 25, but the beams here represent the average result across the whole mosaic because CASA does not track spatial variations of the beam. We show the beam for both the 7-m array alone and the combined 12-m+7-m array.

Both beams show a reasonable Gaussian core, but the 7-m-only beam shows significantly worse positive and negative sidelobes. This is consistent with the significantly worse performance of 7-m-only imaging compared to 12-m+7-m imaging discussed by Leroy et al. (2021). Overall, the 12-m+7-m synthesized beam, which corresponds to our key data product, looks symmetric without clearly visible pathologies in the sidelobes.

### 5.3. Mosaic sensitivity

Figure 27 illustrates the combined primary beam response, i.e., the mosaic sensitivity, of the 7-m data and 12-m+7-m data for one channel in one ALMA data cube.



**Figure 22. PHANGS-ALMA observing conditions for individual observations.** *Top left:* Duration, from start to finish, of individual observing blocks observed with the 12-m array, 7-m array, and total power telescopes. Gray points and error bars show the median and 16%–84% range of the duration of individual observations, which are typically 1–1.5 h. The remaining three panels show the mean elevation and precipitable water vapor (PWV) measured in the atmosphere for each individual block. Gray regions show the 16%–84% range of each quantity and black lines show the median. A thick gray line shows ALMA’s 5<sup>th</sup> octile of PWV,  $\sim 1.8$  mm, the nominal threshold for Band 6 (230 GHz) observations.

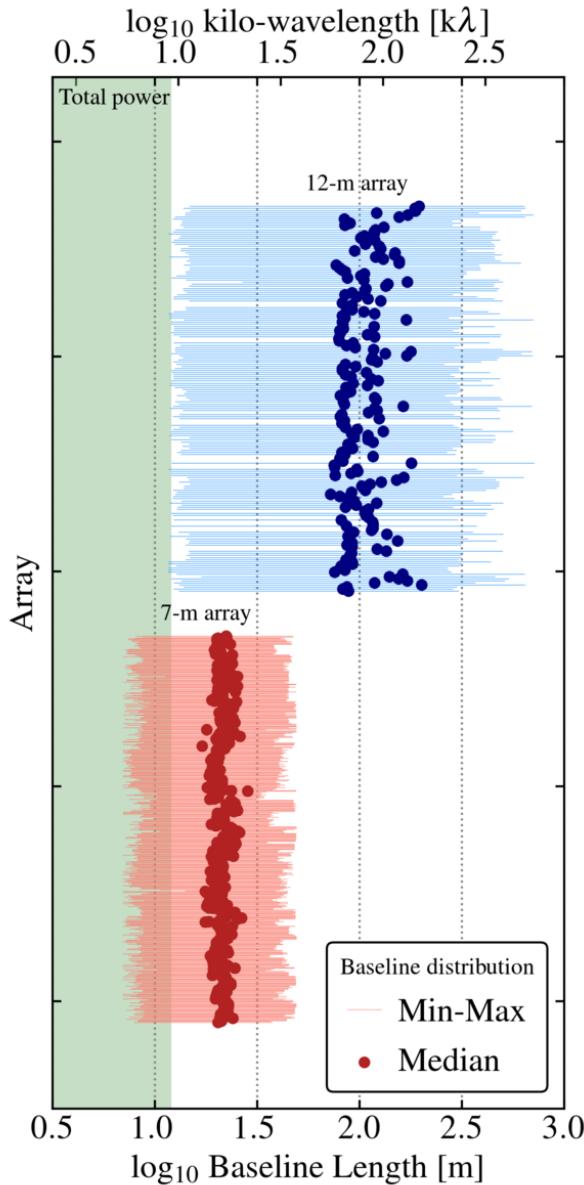
The figure shows that ALMA achieves an even sensitivity across the area of the galaxy and demonstrates the Nyquist-sampled pointing spacing used by ALMA. The outer 0.5 response contour indicates the typical extent of our imaging, though in a few cases we use a more stringent primary beam cutoff. This even coverage, at the few percent level, is typical of our targets, with only a

few cases showing more uneven coverage due to flagging or issues with the execution of observations.

## 6. DATA PROCESSING

### 6.1. Summary

Leroy et al. (2021) describe the reduction and data processing for PHANGS-ALMA. That paper explains our approaches to calibration,  $u-v$  data processing,



**Figure 23. PHANGS–ALMA baseline lengths.** Illustration of  $u-v$  baseline lengths in the 12-m (blue, upper) and 7-m (red, lower) arrays. Each line shows one observation, with the circle showing the median value and the lines showing the full minimum to maximum baseline length in the data. A green shaded region indicates the 12-m diameter of the total power antennas. For a more detailed look at the distribution of collecting area vs. baseline length in one case, see Figure 25.

imaging, post-processing of the imaged data, short-spacing correction, data product creation, and quality assurance. It includes an overview, illustrations of each step, and detailed discussions of our motivation for each step. The accompanying `CASA+python` data pipeline

is also publicly available<sup>15</sup>. Herrera et al. (2020) describe our calibration and imaging procedure for the total power data. Given these detailed presentations elsewhere, we only summarize the procedure here.

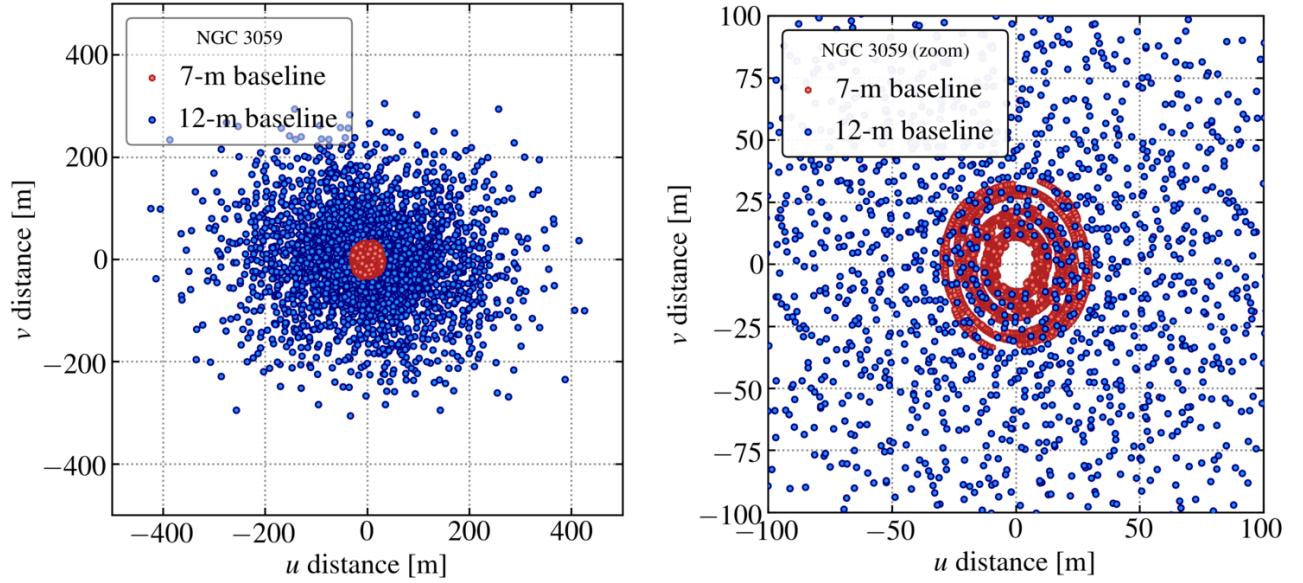
Figure 28, reproduced exactly from Leroy et al. (2021), summarizes the overall workflow of the PHANGS–ALMA processing pipeline. For the interferometric data, we download each observation and apply the ALMA observatory pipeline calibration. Based on inspection of the  $u-v$  data, we found that, in general, additional by-hand flagging does not appreciably improve the quality of the final data products. We also performed several checks related to the overall calibration. We found that the internal stability of the total power appears excellent, with rms scatter of  $\sim 3\%$  from observation to observation, and that the fluxes of the secondary calibrators derived from the 7-m and 12-m array observations agree well with expectations from the ALMA calibrator database. As discussed above, we did identify a minor, 2–5% net magnitude, issue regarding the overall calibration of a subset of the total power data. We expect this to be fixed in future releases of PHANGS–ALMA but at the moment this represents a “known issue.”

Next, we staged the calibrated  $u-v$  data for imaging. This step involves continuum subtraction, spectral regridding and rebinning, and an empirical reweighting of the data. For the current release of PHANGS–ALMA, we carried out these steps in `CASA` version 5.6.1, though our tests showed little influence of the adopted `CASA` version on our final results.

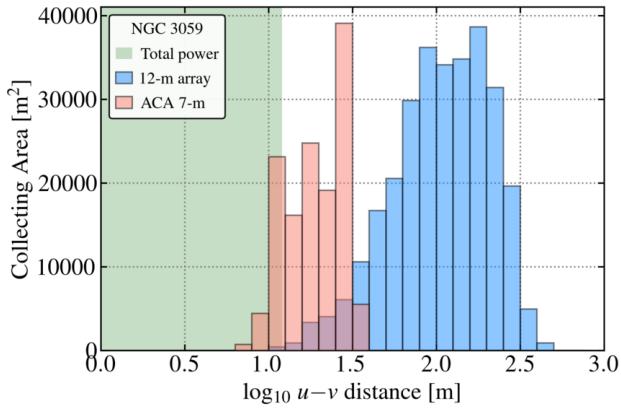
We imaged the staged  $u-v$  data using the `CASA` task `tclean`. We use a two stage deconvolution process. In the first stage, we used the “multiscale clean” (Cornwell 2008) deconvolution algorithm with no clean mask or a very extended clean mask, and we imposed a residual threshold of 4 times the rms noise in the cube. In other words, we carry out a multiscale clean and clean the image until the maximum residual, i.e., not deconvolved emission, has signal to noise of 4. Then, we construct a more restrictive clean mask based on the location of emission in the current version of the deconvolved image. Within this more restrictive clean mask, we continue deconvolving using a classic “Hogbom clean” (Högblom 1974). We run this second stage of deconvolution, stopping when the residual, i.e., not deconvolved emission, within the clean mask has a peak  $\leq 1\sigma$  or the amount of deconvolved emission changes by a negligible amount over successive calls to `clean`. In practice, the signal-to-noise criterion is essentially always reached.

After imaging, we correct the image cube for the response of the primary beam. The image products often have elliptical synthesized beams at arbitrary position

<sup>15</sup> URL provided as soon as that paper is published and sent to the arxiv.

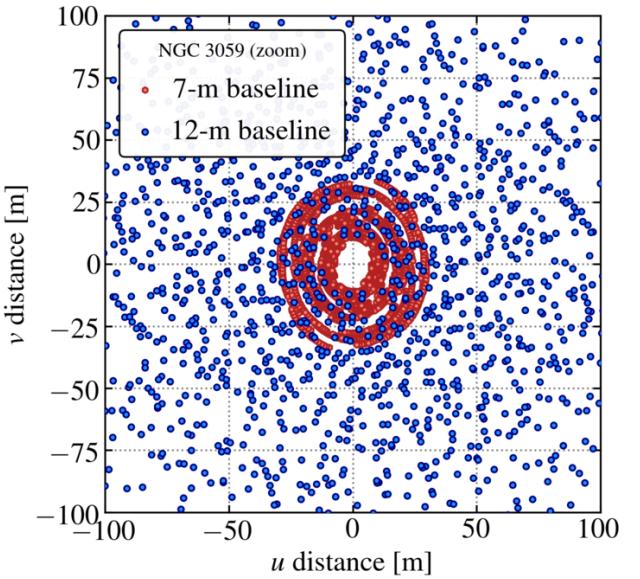


**Figure 24. PHANGS-ALMA  $u-v$  coverage example.** Example plots of  $u-v$  plane coverage for two pointings in NGC 3059, one covered by the ACA 7-m array (red points) and one covered by the main 12-m array (blue points). Both panels show the same data, with the *right* hand panel showing a zoom-in to the inner part of the  $u-v$  plane. Figure 25 shows the collecting area vs. baseline length for these data and Figure 26 show the synthesized beams for NGC 3059.



**Figure 25. PHANGS-ALMA distribution of collecting area vs.  $u-v$  distance for individual integrations.** For the same pointings in Figure 24, we show the distributions of collecting area vs.  $\log_{10} u-v$  baseline length for individual integrations from the 12-m array (blue) and 7-m array observations. The figure illustrates good overlap between the three components of ALMA and good sensitivity to a range of spatial scales. Note that the shaded total power region only indicates the  $u-v$  range of the total power antennas, not the amount of collecting area.

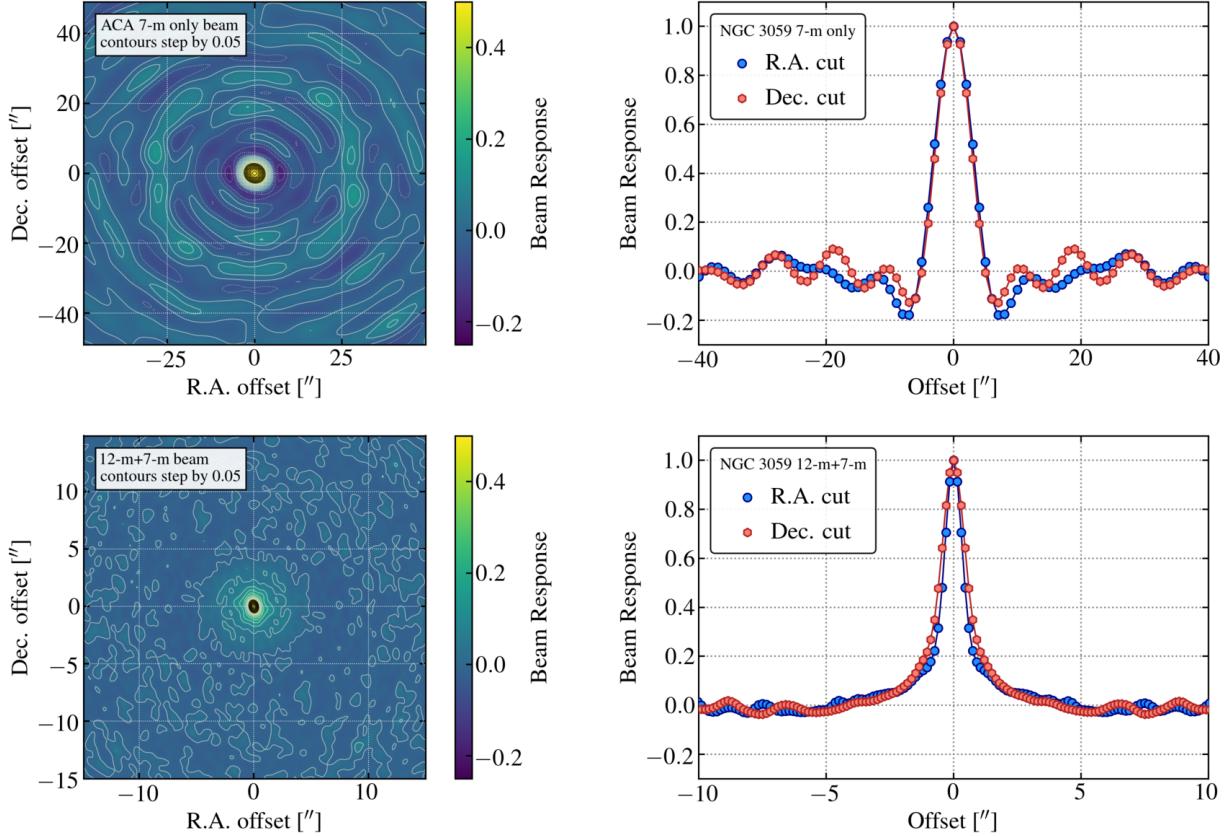
angles. We convolve these so that the final images have a round, Gaussian synthesized beam shape. For galaxies



observed in multiple separate parts by several different Science Goals, we image the data separately and then combine the different parts via a noise-weighted linear mosaicking operation at this stage. Before doing this, we smooth all of the parts of a galaxy to share a common synthesized beam. At this stage, each galaxy has a single deconvolved, primary beam-corrected interferometric image cube with a round synthesized beam.

Next, we combine the total power and interferometric data using the CASA implementation of the “feather” algorithm (e.g., see Cotton 2017). Feathering combines the total power and interferometer data in the Fourier domain, ensuring that in the final image the low-spatial-frequency information, that is crucial to determine the overall flux of the image, is set by the total power data. Before this stage, we calibrate, baseline subtract, and image the total power data following the procedures described in Herrera et al. (2020) and Leroy et al. (2021). For a subset of galaxies, this processing included additional steps to deal with the terrestrial ozone contamination described above.

At this stage, we have our final data cubes, which we convert from units of  $\text{Jy beam}^{-1}$  to units of Kelvin. We use these to construct a series of data products for use in scientific analysis. First, we convolve the cubes to a series of fixed physical and angular resolutions, which allows the whole sample to be subject to rigorous comparative analysis. The exact suite of resolutions has varied over time. For the current release we create cubes with



**Figure 26. PHANGS–ALMA synthesized beam examples.** Synthesized beams for one channel in our NGC 3059 CO(2–1) cube produced by our CASA imaging for the (*top row*) ACA 7-m-only and (*bottom row*) 12-m+7-m combined arrays. The *left* column shows images of the synthesized beam response for each case, with contours spaced by 0.05. The *right* column shows response cuts along right ascension and declination through the synthesized beam.

FWHM synthesize beams of 2'', 7.5'', 11'', and 15'', as well as 60 pc, 90 pc, 120 pc, 150 pc, 500 pc, 750 pc, and 1 kpc. When convolving to fixed physical resolution, we take into account the distance to the galaxy as compiled by Anand et al. (2021), discussed in Section 4.2, and reported in Table 3.

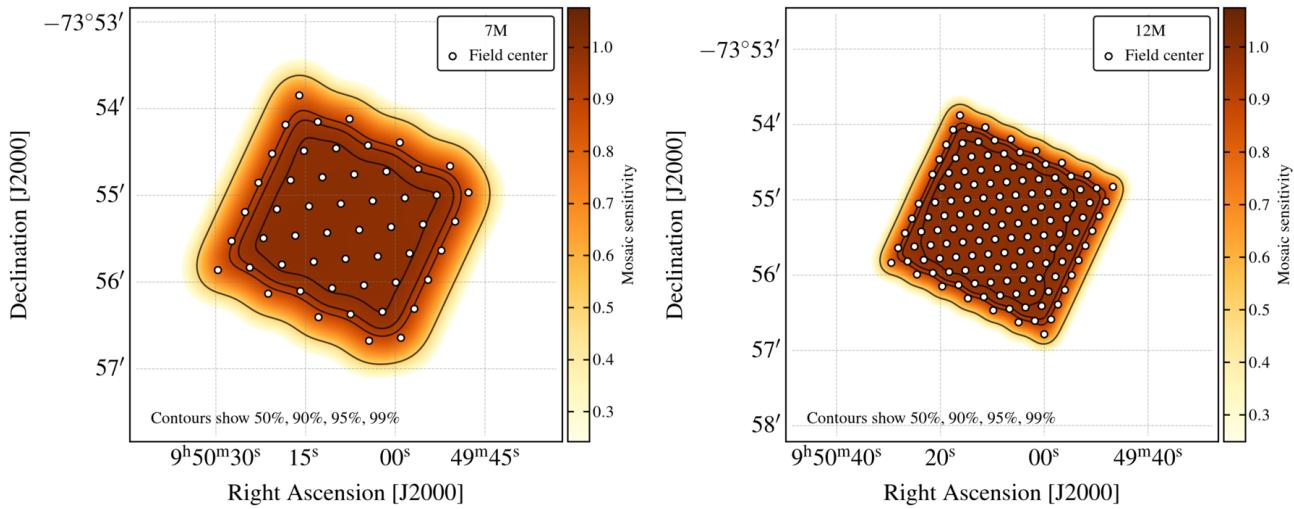
For each cube at each resolution, we construct a three dimensional noise model, bootstrapped from regions of the cube that contain little or no signal. Then, we combine the noise models and the data cubes to construct a series of masks that identify the likely location of real CO emission inside each cube. We construct two sets of masks, high confidence “strict” masks and high completeness “broad” masks. The strict masks are constructed for each cube at each resolution. Following Rosolowsky & Leroy (2006) and Leroy et al. (2021), these strict masks initially consider only locations in the cube where emission exceeds signal-to-noise of 4 over two consecutive velocity channels. Then, we extend these masks to include any adjacent regions where the emission exceeds a signal-to-noise of 2 over two consecutive velocity channels. We refer to these strict masks as

“high confidence” because this construction procedure renders them very likely to include mostly real emission. The broad masks are created from the union of strict masks across all resolutions. These masks aim for “high completeness” in the sense that they tend to include almost all emission in the cube, as determined by comparison to the integrated flux from the total power data.

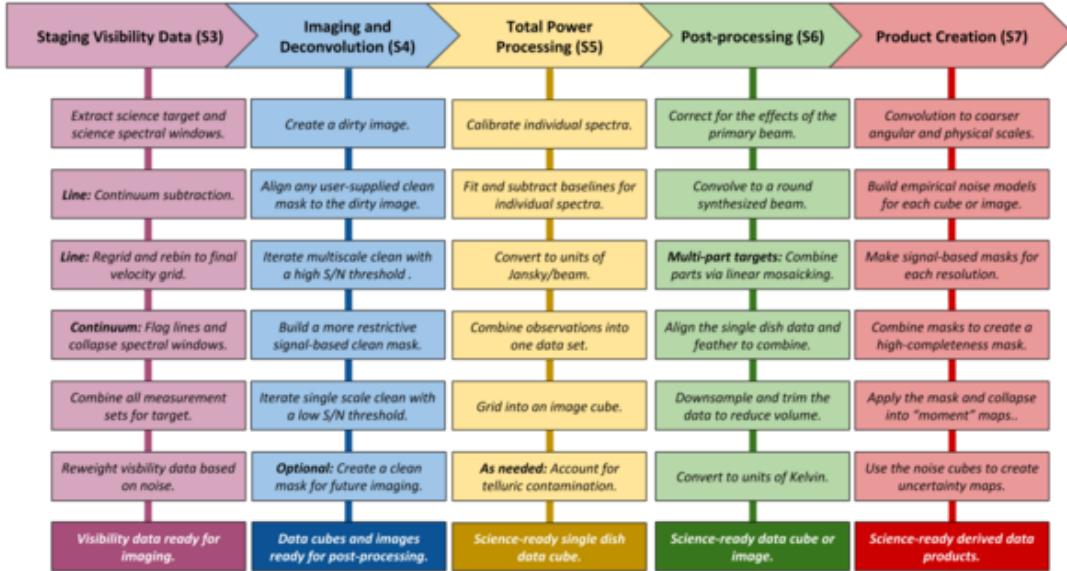
Finally, we apply the masks to the cubes and noise models and “collapse” the cubes to produce a series of two-dimensional maps and associated uncertainty maps. We discuss the specific products more below, but briefly these “moment maps” include images of integrated intensity, peak intensity along the line of sight, intensity-weighted mean velocity, and the spectral width of the CO line along the line of sight. We produce an associated uncertainty map for each moment map.

For more details on each step of this process see Leroy et al. (2021) and the associated, public PHANGS–ALMA post-processing pipeline.

## 6.2. Validation, known issues, and limitations



**Figure 27. Example of PHANGS-ALMA mosaic sensitivity.** Examples of mosaic sensitivity, meaning combined primary beam response of all observations, for the 7-m (left) and 12-m+7-m (right) observations of one channel in one PHANGS-ALMA target, NGC 3059. The image shows the primary beam response, which is proportional to the inverse of the noise in the final image, i.e., a response of 0.5 translates to 2 $\times$  higher noise. The black-and-white markers show the locations of individual pointings and contours show the 50%, 90%, and 95%, and 99% response. The figure illustrates the mosaic pointings and that the combined ALMA response is even across the field.



**Figure 28. PHANGS-ALMA Processing Pipeline.** Reproduced *exactly* from Leroy et al. (2021). The figure illustrates the processing steps used to process the  $u-v$  and total power data delivered by ALMA into final science-ready data products. Leroy et al. (2021) present the processing pipeline and discuss the motivation, implementation, and limitations of each step in detail.

We engaged in a mixture of automated and manual quality assurance of the data and validation of the PHANGS-ALMA pipeline. A broad cross-section of the team participated in several rounds of manual inspection, looking at the cubes, associated imaging products, and derived data products. For each version of the data

after early 2019, we ran automated regression tests to quantify changes from version to version and to ensure that we understood the reason for all changes. Finally, to validate our pipeline, we carried out end-to-end tests using the ALMA simulator in CASA. These tests allowed

us to verify that the pipeline output represents a close match to the known input to the simulator.

These tests, and other checks, showed that our data processing has yielded stable results over time, that the results pass manual inspections by experts, and that the output from the pipeline replicates known input to good precision. However, there are still several known limitations of our approach and issues related to data processing that we note here.

- 1. Both the spectral and angular resolution of the data could be improved at the cost of signal-to-noise.** The convolution to a round beam and matching beams between parts of multi-part mosaics both somewhat inflate the synthesized beam size. We also use a  $u-v$  weighting scheme with intermediate “robustness parameter.” All of these choices improve the surface brightness sensitivity, but they degrade the resolution by a modest amount. Our cubes also use a  $\sim 2.5$  km s $^{-1}$  channel width, which is coarser than the native spectral resolution. By altering these choices, future processing of the data could improve the spatial and spectral resolution at the cost of signal-to-noise. This may be of particular interest for bright regions like the centers of galaxies.
- 2. Deconvolution of 7-m-only imaging depends sensitively on the signal-to-noise and structure of the data.** Both our real data and end-to-end tests based on simulations show that the image reconstruction performs poorly using low or moderate signal-to-noise data from the 7-m array alone. The effect is explored in detail in Leroy et al. (2021), and it is largely alleviated by short-spacing correction. However, the 7-m-only data

for many galaxies appear unreliable. We also view the combined 7-m and total power data as less reliable than the full combined 12-m, 7-m, and total power data.

- 3. There is a small known bias in the total power flux calibration.** As mentioned above and discussed in Leroy et al. (2021), there was a small lag incorporating upgrades to the surface accuracy of the total power antennas into the observatory-provided calibration. As a result, a subset of our total power data sets apply an observatory-provided calibration known to be high by  $\sim 7\%$ . The net effect is that a subset of cubes has overall fluxes biased high by 2–5%. This will be fixed in future releases.
- 4. The data processing introduces some mild dependence of noise on frequency.** As discussed in Leroy et al. (2021), the regridding and rebinning parts of the  $u-v$  staging procedure introduce a mild spectral variation into the noise of the final data cubes. The overall magnitude is  $\sim 10\%$  across the full bandwidth of the cubes. The effect seems unavoidable in the current version of CASA, but we expect this to be addressed in the future.

Several other issues arose during processing that we have largely addressed, including the telluric ozone contamination of the total power data. To our knowledge the issues above represent the main outstanding issues and limitations of the data.

## 7. DATA PROPERTIES AND KEY PRODUCTS

### 7.1. Properties of the final cubes

**Table 16.** PHANGS-ALMA Cube Properties

Galaxy	Arrays	Resolution		Area Mapped		Noise			Completeness		Notes
		Angular	Physical	Angular	Physical	Native	150 pc	Native	Native	150 pc	
		( $''$ )	(pc)	(arcmin $^2$ )	(kpc $^2$ )	(mK)	(mK)	( $\frac{\text{mJy}}{\text{beam}^{-1}}$ )	(%)	(%)	
NGC 0247	7m+TP	8.51	153.1	80.7	93.9	23	23	73.9	7.2	7.2	
NGC 0253	7m+TP	8.37	150.2	98.7	114.4	36	36	108.2	85.8	85.8	
NGC 0300	7m+TP	8.18	82.8	57.0	21.1	35	13	100.4	36.7	51.3	
NGC 0628	12m+7m+TP	1.12	53.5	14.7	120.7	115	41	6.3	44.7	64.5	
NGC 0685	12m+7m+TP	1.69	163.0	4.8	162.4	40	40	4.8	36.5	36.5	
NGC 1068	7m+TP	8.69	588.6	6.5	106.6	18	...	58.9	96.6	...	
NGC 1097	12m+7m+TP	1.70	111.7	13.4	208.7	52	37	6.4	79.9	82.7	
NGC 1087	12m+7m+TP	1.60	123.1	6.8	145.4	66	52	7.3	64.7	67.5	
NGC 1313	7m	7.93	166.0	55.7	88.0	28	28	75.0	11.6	11.6	

**Table 16** *continued*

**Table 15.** PHANGS-ALMA CO(2–1) Cube Summary

Quantity	Value
Targets	
... 12-m+7-m+TP	81
... 12-m+7-m+TP, $\theta < 165$ pc <sup>a</sup>	78
... 7-m+TP or 7-m	9
... 7-m+TP or 7-m, $\theta < 165$ pc <sup>a</sup>	7
Properties of 12-m+7-m+TP data...	median <sub>-1<math>\sigma</math></sub> <sup>+1<math>\sigma</math></sup> (min–max)
Resolution...	
... angular ["]	1.3 <sub>-0.2</sub> <sup>+0.4</sup> (0.6–2.1)
... physical [pc]	98 <sub>-35</sub> <sup>+31</sup> (26–183)
1 $\sigma$ noise in the cube ...	
... mJy beam <sup>-1</sup> native res.	6.2 <sub>-1.8</sub> <sup>+2.4</sup> (4.4–8.6)
... mK native res.	85 <sub>-41</sub> <sup>+36</sup> (4.4–8.6)
... mK at 150 pc res.	50 <sub>-20</sub> <sup>+26</sup> (16–114)
... mass <sup>b,c</sup> native res. [ $10^3 M_{\odot}$ ]	21 <sub>-10</sub> <sup>+13</sup> (1.5–76)
... mass <sup>b,c</sup> at 150 pc res. [ $10^3 M_{\odot}$ ]	31 <sub>-12</sub> <sup>+16</sup> (10–69)
Implied 1 $\sigma$ surface brightness noise <sup>c</sup> ...	
... K km s <sup>-1</sup> native res.	0.30 <sub>-0.14</sub> <sup>+0.13</sup> (0.12–1.3)
... K km s <sup>-1</sup> at 150 pc res.	0.18 <sub>-0.07</sub> <sup>+0.09</sup> (0.06–0.41)
... $\Sigma_{\text{mol}}$ <sup>b,c</sup> native res. [ $M_{\odot}$ pc <sup>-2</sup> ]	2.0 <sub>-1.0</sub> <sup>+0.8</sup> (0.8–8.7)
... mass <sup>b,c</sup> at 150 pc res. [ $M_{\odot}$ pc <sup>-2</sup> ]	1.2 <sub>-0.5</sub> <sup>+0.6</sup> (0.4–2.7)
Completeness <sup>d</sup> ...	
... at native resolution	61 <sub>-24</sub> <sup>+20</sup> % (0–94%)
... at 150 pc resolution	67 <sub>-26</sub> <sup>+18</sup> % (0–100%)
Area mapped	
Physical area	
... total for survey [kpc <sup>2</sup> ]	10,650
... mean 12-m+7-m+TP [kpc <sup>2</sup> ]	124
... mean 7-m+TP [kpc <sup>2</sup> ]	71
Angular area	
... total for survey [arcmin <sup>2</sup> ]	1,050
... mean 12-m+7-m+TP [arcmin <sup>2</sup> ]	7.0
... mean 7-m+TP [arcmin <sup>2</sup> ]	54

<sup>a</sup>We adopt a 10% tolerance when convolving to a fixed physical resolution, so that all maps with  $\theta < 165$  pc have “150 pc” data products.

<sup>b</sup>When calculating mass, we adopt a Milky Way  $\alpha_{\text{CO}}^{1-0}$  and  $R_{21} = 0.65$ . See text and Equations 10 and 13.

<sup>c</sup>When calculating line-integrated quantities, including mass and surface brightness, we adopt a full line width of  $\Delta v = 5$  km s<sup>-1</sup>. This implies an improvement of  $\sqrt{2}$  in the surface brightness sensitivity.

<sup>d</sup>“Completeness” is defined here as the ratio of the sum of flux in the strictly masked moment 0 map to the direct sum of the whole cube or the broad map. Completeness is not random but correlates with the overall brightness of the galaxy. See §7.1.2.

**Table 16** *continued*

Table 16 (continued)

Galaxy	Arrays	Resolution		Area Mapped		Noise			Completeness		Notes
		Angular	Physical	Angular	Physical	Native	150 pc	Native	Native	150 pc	
		( $''$ )	(pc)	(arcmin $^2$ )	(kpc $^2$ )	(mK)	(mK)	( $\frac{\text{mJy}}{\text{beam}^{-1}}$ )	(%)	(%)	

Table 16 (continued)

Galaxy	Arrays	Resolution		Area Mapped		Noise			Completeness		Notes
		Angular	Physical	Angular	Physical	Native	150 pc	Native	Native	150 pc	
		( $''$ )	(pc)	(arcmin $^2$ )	(kpc $^2$ )	(mK)	(mK)	( $\frac{\text{mJy}}{\text{beam}^{-1}}$ )	(%)	(%)	
NGC 1300	12m+7m+TP	1.23	113.1	13.0	396.8	102	76	6.6	48.5	52.4	
NGC 1317	12m+7m+TP	1.59	147.1	1.7	52.1	46	46	4.9	89.2	90.0	
IC 1954	12m+7m+TP	1.56	97.1	4.0	55.6	45	27	4.7	64.1	70.7	
NGC 1365	12m+7m+TP	1.38	130.8	7.3	237.2	112	95	9.1	75.3	76.4	
NGC 1385	12m+7m+TP	1.27	105.9	6.8	171.3	82	55	5.7	60.9	65.5	
NGC 1433	12m+7m+TP	1.10	99.1	11.8	346.0	119	76	6.2	51.9	57.3	
NGC 1511	12m+7m+TP	1.45	107.1	3.6	71.2	67	45	6.0	75.6	81.0	
NGC 1512	12m+7m+TP	1.03	94.5	7.9	235.5	120	70	5.5	46.6	53.1	
NGC 1546	12m+7m+TP	1.28	109.6	2.8	73.9	62	41	4.4	88.4	91.6	
NGC 1559	12m+7m+TP	1.25	117.5	6.8	217.6	88	66	5.9	60.8	64.5	
NGC 1566	12m+7m+TP	1.25	107.6	11.9	314.3	90	64	6.1	81.1	92.3	
NGC 1637	12m+7m+TP	1.39	78.9	6.2	71.7	36	17	3.0	77.0	83.5	
NGC 1672	12m+7m+TP	1.93	181.7	8.1	258.2	85	...	13.7	72.8	...	
NGC 1809	12m+7m+TP	1.41	136.0	2.4	79.4	96	86	8.2	32.3 <sup>F</sup>	33.2 <sup>F</sup>	
NGC 1792	12m+7m+TP	1.92	150.9	9.1	203.0	46	46	7.3	88.5	88.5	
NGC 2090	12m+7m+TP	1.30	73.8	3.0	35.6	99	52	7.2	60.4	66.7	
NGC 2283	12m+7m+TP	1.31	87.0	5.4	85.4	84	51	6.2	34.5	36.5	
NGC 2566	12m+7m+TP	1.28	145.3	6.8	317.1	88	88	6.2	74.7	74.7	
NGC 2775	12m+7m+TP	1.09	122.6	5.5	248.0	134	108	6.9	27.7	33.0	
NGC 2835	12m+7m+TP	0.84	50.0	7.9	99.4	239	74	7.4	19.2	25.8	
NGC 2903	12m+7m+TP	1.45	70.5	15.5	131.4	71	37	6.6	78.1	83.7	
NGC 2997	12m+7m+TP	1.77	120.5	19.0	318.1	42	32	5.7	78.3	80.9	
NGC 3059	12m+7m+TP	1.22	119.9	7.5	259.3	95	75	6.1	58.6	61.9	
NGC 3137	12m+7m+TP	1.51	120.0	3.4	76.2	62	45	6.1	54.0	59.4	
NGC 3239	12m+7m+TP	1.28	67.5	2.9	28.5	132	73	9.4	1.7 <sup>F</sup>	0.9 <sup>F</sup>	
NGC 3351	12m+7m+TP	1.46	70.7	7.8	65.2	108	46	10.0	60.2	71.7	
NGC 3489	12m+7m+TP	0.75	42.9	0.7	8.1	100	27	2.4	50.1 <sup>F</sup>	58.4 <sup>F</sup>	
NGC 3511	12m+7m+TP	1.80	121.5	6.4	105.0	44	34	6.1	75.4	77.5	
NGC 3507	12m+7m+TP	1.36	155.2	5.1	239.9	79	79	6.3	40.0	40.0	
NGC 3521	12m+7m+TP	1.33	85.5	12.6	186.4	61	35	4.7	85.2	89.2	
NGC 3596	12m+7m+TP	1.22	66.9	4.6	49.3	128	67	8.3	47.4	59.9	
NGC 3599	12m+7m+TP	0.66	63.2	0.7	23.0	120	50	2.2	59.4 <sup>F</sup>	54.8 <sup>F</sup>	
NGC 3621	12m+7m+TP	1.82	62.4	12.2	51.4	39	16	5.6	83.4	87.8	
NGC 3626	12m+7m+TP	1.17	114.1	1.9	66.2	156	114	9.3	33.9 <sup>F</sup>	65.4 <sup>F</sup>	
NGC 3627	12m+7m+TP	1.63	89.2	11.8	128.4	80	45	9.1	80.8	85.3	
NGC 4207	12m+7m+TP	1.22	93.2	1.6	34.7	132	85	8.5	68.9	73.7	
NGC 4254	12m+7m+TP	1.78	113.1	12.7	184.1	64	45	8.7	79.4	82.4	
NGC 4293	12m+7m+TP	1.16	89.0	3.5	73.5	130	80	7.6	66.7	70.5	
NGC 4298	12m+7m+TP	1.59	114.7	4.2	79.8	38	25	4.1	82.2	85.7	
NGC 4303	12m+7m+TP	1.81	149.3	7.8	190.5	82	82	11.5	73.1	73.1	

Table 16 continued

Table 16 (continued)

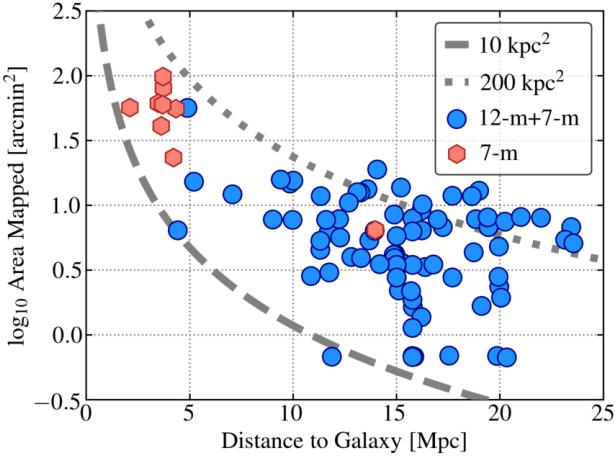
Galaxy	Arrays	Resolution		Area Mapped		Noise			Completeness		Notes
		Angular	Physical	Angular	Physical	Native	150 pc	Native	Native	150 pc	
		( $''$ )	(pc)	(arcmin $^2$ )	(kpc $^2$ )	(mK)	(mK)	( $\frac{\text{mJy}}{\text{beam}^{-1}}$ )	(%)	(%)	
NGC 4321	12m+7m+TP	1.67	122.9	13.7	269.0	89	69	10.7	69.2	72.2	
NGC 4424	12m+7m+TP	1.14	89.4	1.4	30.6	151	89	8.5	61.4	69.4	
NGC 4457	12m+7m+TP	1.11	81.3	2.2	42.7	101	52	5.4	73.6	80.7	
NGC 4459	12m+7m+TP	0.64	48.9	0.7	14.5	113	34	2.0	84.7 <sup>F</sup>	79.8 <sup>F</sup>	
NGC 4476	12m+7m+TP	0.87	73.9	0.7	18.2	70	32	2.3	85.1 <sup>F</sup>	97.9 <sup>F</sup>	
NGC 4477	12m+7m+TP	0.62	47.2	0.7	14.6	125	41	2.0	52.2 <sup>F</sup>	46.3 <sup>F</sup>	
NGC 4496A	12m+7m+TP	1.25	90.3	4.2	77.8	104	68	7.0	29.6	29.2	
NGC 4535	12m+7m+TP	1.56	119.1	7.9	167.2	83	60	8.6	60.6	66.9	
NGC 4536	12m+7m+TP	1.48	116.3	10.2	228.5	44	33	4.2	80.8	84.7	
NGC 4540	12m+7m+TP	1.37	104.8	1.9	39.4	111	76	9.0	45.5	52.6	
NGC 4548	12m+7m+TP	1.69	132.7	6.4	143.1	49	41	6.1	40.6	42.1	
NGC 4569	12m+7m+TP	1.69	128.9	6.3	132.8	47	38	5.9	81.6	83.2	
NGC 4571	12m+7m+TP	1.18	85.0	3.9	73.7	118	70	7.1	18.8	31.7	
NGC 4579	12m+7m+TP	1.79	182.7	8.2	304.5	46	...	6.4	60.3	...	
NGC 4596	12m+7m+TP	0.65	49.6	0.7	14.3	99	30	1.8	93.8 <sup>F</sup>	179.9 <sup>F</sup>	
NGC 4654	12m+7m+TP	1.72	182.8	8.1	330.1	52	...	6.6	82.7	...	
NGC 4689	12m+7m+TP	1.18	86.0	5.8	111.0	111	63	6.7	51.5	62.7	
NGC 4694	12m+7m+TP	1.17	89.3	1.1	23.9	143	82	8.4	24.8 <sup>F</sup>	28.9 <sup>F</sup>	
NGC 4731	12m+7m+TP	1.53	98.4	3.9	58.7	39	23	3.9	42.6 <sup>F</sup>	47.2 <sup>F</sup>	
NGC 4781	12m+7m+TP	1.31	71.7	5.4	58.0	46	20	3.4	65.4	75.6	
NGC 4826	12m+7m+TP	1.26	26.9	6.4	10.6	77	18	5.3	87.6	93.6	
NGC 4941	12m+7m+TP	1.59	115.3	3.5	66.2	40	28	4.4	57.0	63.3	
NGC 4951	12m+7m+TP	1.25	91.2	2.8	53.2	91	54	6.2	47.7	52.4	
NGC 4945	7m+TP	7.90	132.9	61.5	62.7	40	32	107.3	85.1	87.1	
NGC 5042	12m+7m+TP	1.33	107.8	3.5	83.4	79	54	6.0	28.1	31.2	
NGC 5068	12m+7m+TP	1.04	26.2	15.3	35.0	185	38	8.7	33.4	46.5	
NGC 5134	12m+7m+TP	1.23	119.0	2.8	94.7	87	66	5.7	41.3	45.6	
NGC 5128	7m+TP	8.13	145.5	60.0	69.2	35	35	100.9	76.8	76.8	
NGC 5236	12m+7m+TP	2.14	50.7	56.7	114.7	43	16	8.6	87.8	92.1	
NGC 5248	12m+7m+TP	1.29	93.1	8.5	159.4	92	62	6.6	71.8	76.2	
CIRCINUS	7m+TP	7.32	149.0	23.4	34.9	55	55	128.1	77.1	77.7	
NGC 5530	12m+7m+TP	1.13	66.9	5.7	72.0	114	56	6.2	47.5	57.3	
NGC 5643	12m+7m+TP	1.30	79.9	10.6	143.7	77	43	5.6	69.7	76.0	
NGC 6300	12m+7m+TP	1.08	60.4	7.8	88.2	116	54	5.8	64.7	73.0	
NGC 6744	12m+7m+TP	1.13	51.6	15.8	118.2	187	69	10.4	35.8	60.3	
IC 5273	12m+7m+TP	1.76	120.7	3.5	60.2	33	25	4.4	53.0	56.1	
NGC 7456	12m+7m+TP	1.68	127.8	2.2	45.7	40	32	4.9	31.6 <sup>F</sup>	33.9 <sup>F</sup>	
NGC 7496	12m+7m+TP	1.68	152.0	4.4	129.7	36	36	4.3	73.3	73.3	
IC 5332	12m+7m+TP	0.74	32.1	7.8	53.4	366	102	8.6	2.5 <sup>F</sup>	7.1 <sup>F</sup>	
NGC 7743	12m+7m+TP	0.64	63.2	0.7	23.6	76	26	1.3	58.1 <sup>F</sup>	60.8 <sup>F</sup>	
NGC 7793	7m+TP	7.99	140.2	41.2	45.7	34	34	94.9	32.3	32.3	

<sup>F</sup>Faint galaxy, completeness is unreliable. See text.

NOTE— Properties of the PHANGS-ALMA data cubes. Columns — name of the galaxy; best available array combination; physical and angular resolution of the native resolution cube; noise in surface brightness (mK) units at the native resolution and 150 pc resolution, if available; noise in mJy beam $^{-1}$  units in the native-resolution cube; completeness, defined as the ratio of flux inside the strict mask to the sum of the cube or the broad mask. See text. at native resolution and 150 pc resolution.

Table 15 summarizes the properties of our final cubes, which we report in detail in Table 16. For each target,

we note the best available array combination, angular



**Figure 29. Area Mapped.** Angular area mapped for each PHANGS–ALMA target as a function of distance to the target. Galaxies mapped using the 12-m+7-m array appear in blue, and those mapped with the 7-m-only appear in red. The dotted and dashed lines show physical areas of 200 kpc<sup>2</sup> and 10 kpc<sup>2</sup>.

and physical resolution, noise, and completeness. When translating angular to physical beam size and coverage area, we use the distances reported in Table 3 and compiled by Anand et al. (2021). The reported noise represents the median value in the three dimensional noise cube, and should be characteristic for the cube. Here we define completeness,  $f_{\text{flux}}$ , as

$$f_{\text{flux}} = \frac{\sum \text{strict mom0}}{\max(\sum \text{broad mom0}, \sum \text{cube})}, \quad (12)$$

that is, the ratio of flux inside the “strict,” high confidence, high signal-to-noise based mask to the total flux in the cube. The “max” in the denominator uses the larger of the sum inside the “broad” mask or the cube; we adopt this to help account for some instability in the direct sum of cubes with faint CO emission (see below). Thus, the completeness,  $f_{\text{flux}}$ , represents the fraction of the total flux recovered at good signal-to-noise.

Figure 29 and Table 15 summarize the area mapped. In total, PHANGS–ALMA surveyed  $\sim 1050$  arcmin<sup>2</sup>, slightly larger than the angular size of the Moon. At the distances to our targets, this translates to  $\sim 10,650$  kpc<sup>2</sup>. On average, the individual maps cover 124 kpc<sup>2</sup> for the 12-m+7-m+TP and 71 kpc<sup>2</sup> for 7-m+TP data. This corresponds to linear map sizes of  $\sim 11$  kpc or typical coverage out to  $r_{\text{gal}} \approx 5\text{--}6$  kpc. Because the 7-m-only targets tend to lie much closer, they cover on average 54 arcmin<sup>2</sup>, compared to 7 arcmin<sup>2</sup>, on average for the 12-m+7-m data.

### 7.1.1. Resolution

Figure 30 and Table 15 summarize the resolutions reported in Table 16. For our 81 12-m+7-m+TP data sets, the median angular resolution is 1.3'' with 70% of the data between 1.1–1.7''. At the distance to our targets, this translates to a median physical resolution of 98 pc with 70% of the data between 60 pc and 120 pc. The top left panel in Figure 30 shows that most of the 7-m-only data sets target very nearby galaxies,  $d < 5$  Mpc. As a result, they achieve similar physical resolutions to the 12-m+7-m+TP observations of our more distant sample members, with a typical resolution of  $\sim 150$  pc for these very nearby galaxies. The lone large-distance 7-m-only data set is NGC 1068, which has not yet been mapped in CO(2–1) using the 12-m array (but has been mapped in CO(3–2) using ALMA by García-Burillo et al. 2014).

We construct data products at a series of fixed physical resolutions, in addition to the native angular resolution and several coarser angular resolutions. For the 12-m+7-m data we adopt FWHM beam size  $\theta = 60, 90, 120$ , and 150 pc at the distance to the galaxy. Reflecting uncertainties in the distance determination, we allow a  $\pm 10\%$  tolerance when labeling a galaxy as having that resolution; that is, we label a map with  $\theta = 91$  pc or 110 pc as a “100 pc” resolution map. The bottom right panel of Figure 30 and Table 15 show the motivation for these choices: 78 of our 81 12-m+7-m galaxies and 7 of 9 7-m-only galaxies have  $\theta < 165$  pc, allowing us to construct a “150 pc resolution” map. Meanwhile, roughly half of our targets have physical resolution better than  $\theta = 120$  pc and 25% have physical resolution better than  $\theta = 90$  pc.

### 7.1.2. Sensitivity

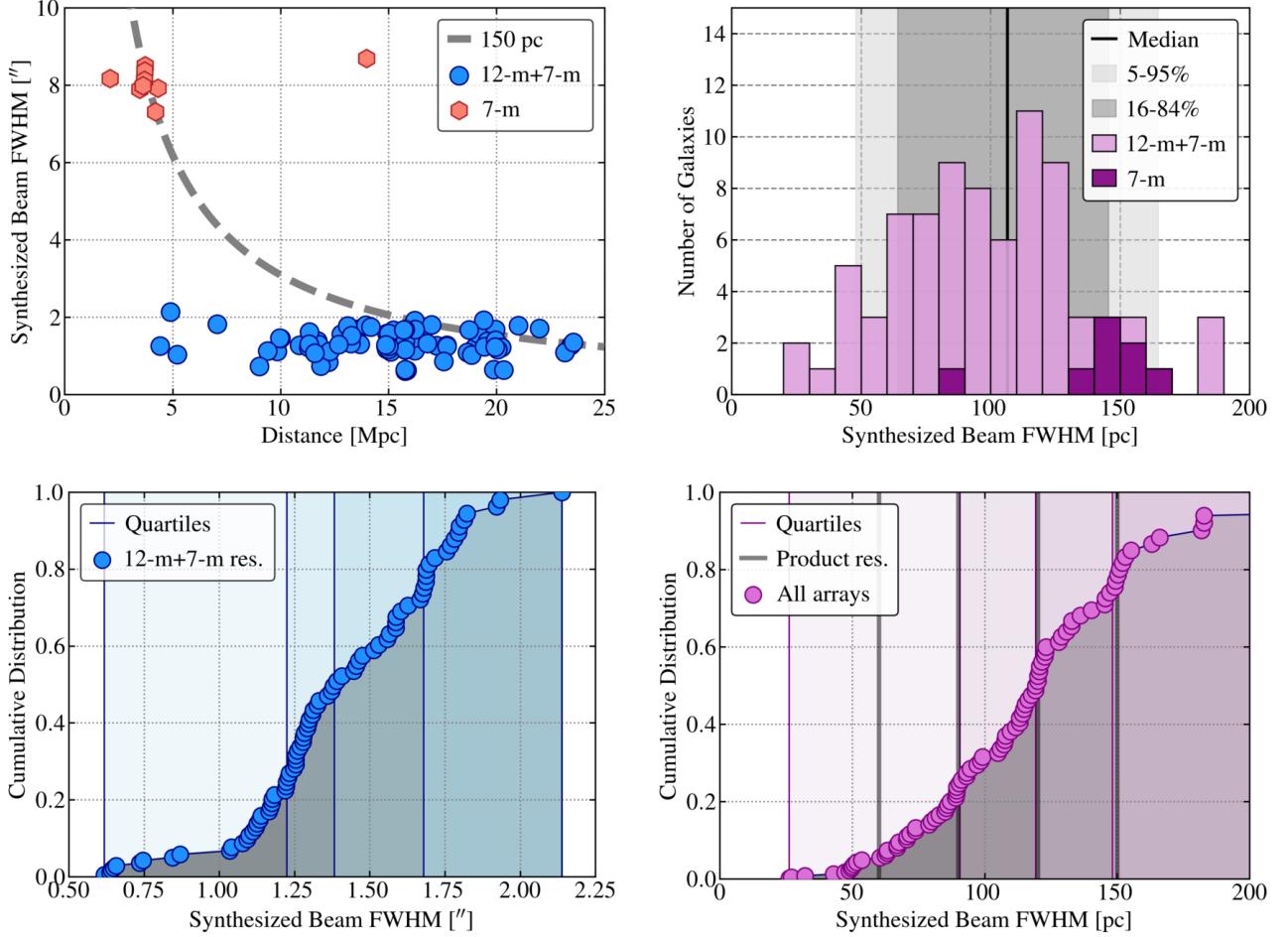
Figure 31 and Table 15 summarize the sensitivity of our data. The top left panel in Figure 31 shows the point source sensitivity in Jy beam<sup>-1</sup> units as a function of the distance to the source. In that panel, we plot two lines showing the cases where the noise and distance equate to  $1\sigma$  mass sensitivity of  $10^4$  and  $10^5$  M<sub>⊙</sub> for a line width of  $\Delta v = 5$  km s<sup>-1</sup>. As above, we translate from line flux to molecular mass via:

$$M_{\text{mol}} [\text{M}_\odot] = 2.63 \times 10^3 \frac{\alpha_{\text{CO}}^{1-0}}{\alpha_{\text{CO}}^{\text{MW}}} \frac{1}{R_{21}} D_{\text{Mpc}}^2 F_{\text{CO}}^{2-1} \quad (13)$$

where  $F_{\text{CO}}^{2-1}$  is the line-integrated CO(2–1) flux in Jy km s<sup>-1</sup> of a source at  $D_{\text{Mpc}}$  distance. Again  $\alpha_{\text{CO}}^{1-0}$  is the CO-to-H<sub>2</sub> conversion factor and  $R_{21}$  is the CO(2–1)-to-CO(1–0) line ratio in Kelvins, so that  $R_{21} = 1$  refers to the thermal case. For reference, if the flux is instead in units of K km s<sup>-1</sup> arcsec<sup>2</sup>, the mass can be calculated via

$$M_{\text{mol}} [\text{M}_\odot] = 1.02 \times 10^2 \frac{\alpha_{\text{CO}}^{1-0}}{\alpha_{\text{CO}}^{\text{MW}}} \frac{1}{R_{21}} D_{\text{Mpc}}^2 F_{\text{CO}}^{\prime\,2-1} \quad (14)$$

where now  $F_{\text{CO}}^{\prime\,2-1}$  is the line-integrated CO(2–1) flux in units of K km s<sup>-1</sup> arcsec<sup>2</sup>.



**Figure 30. Angular and Physical Resolution of PHANGS-ALMA.** The *top left* panel shows the FWHM angular resolution of each final data cube as a function of the distance to that galaxy. We mark 7-m data and 12-m+7-m data with separate symbols, and a dashed line shows 150 pc. Most of the 7-m+TP data target the nearest  $d < 5$  Mpc galaxies, and so achieve high,  $\sim 150$  pc resolution. The single large distance 7-m-only data set is NGC 1068. The *top right* panel shows the distribution of FWHM physical synthesized beam size. The *bottom row* shows the cumulative distributions of angular (*bottom left*) and physical (*bottom right*) beam sizes. Lines and shaded regions show the quartiles in each distribution. In the *bottom right* panel, solid vertical lines show the physical resolutions used to construct data products.

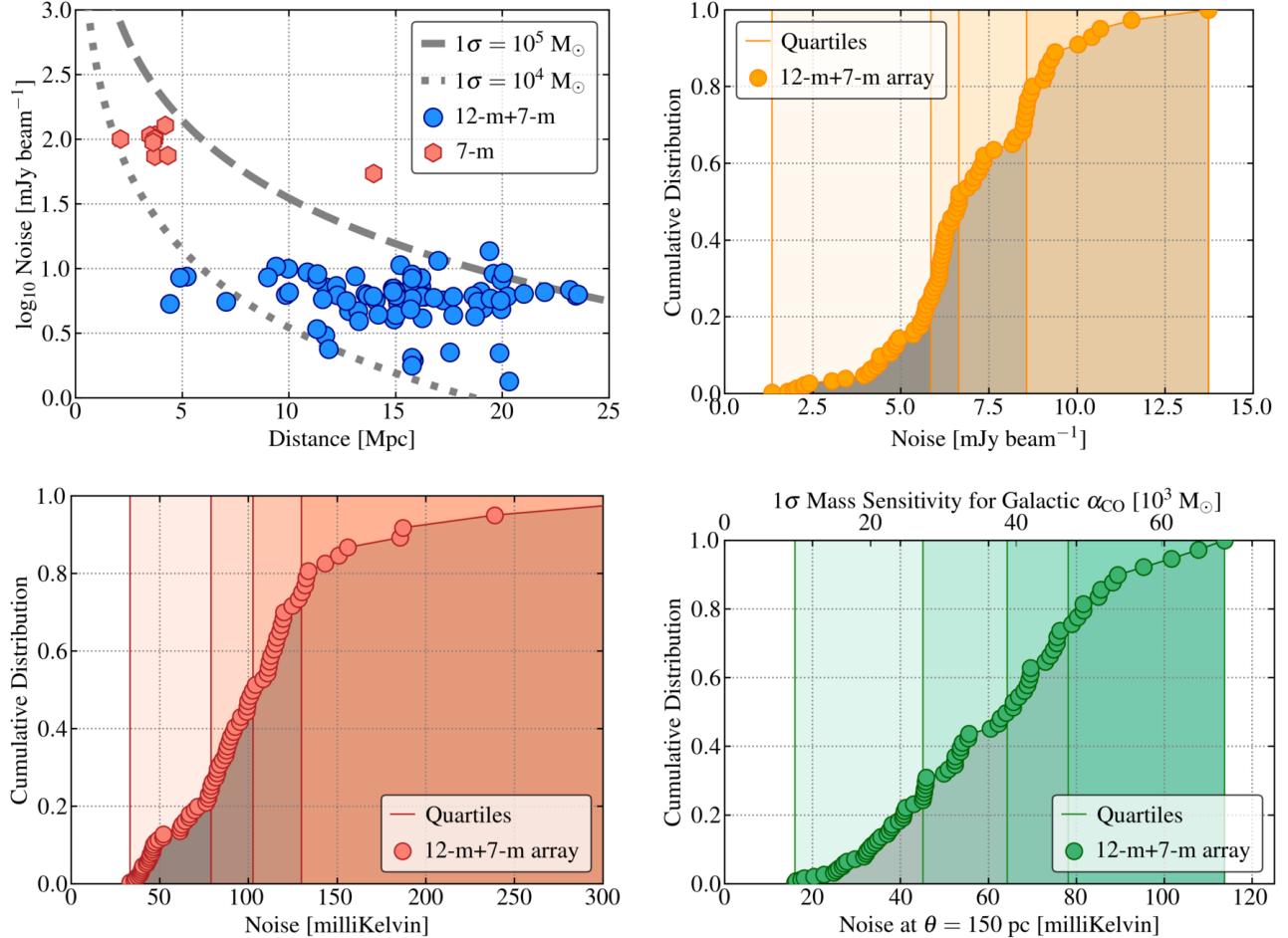
Note that when converting our sensitivity into mass units, e.g., in Figure 31 and Table 15, we assume a fiducial line width of  $5 \text{ km s}^{-1}$  for the line but we measure the initial noise in  $2.54 \text{ km s}^{-1}$  channels. Noise in  $\text{Jy beam}^{-1}$  improves by  $\sqrt{2}$  due to averaging before calculating  $F_{\text{CO}} = \sigma \times \Delta v$  with  $\sigma$  the noise in  $\text{Jy beam}^{-1}$  averaged across the line.

The table and figure show that our typical rms point source sensitivity is  $6.2 \text{ mJy beam}^{-1}$  per  $2.54 \text{ km s}^{-1}$  channel, about  $\sqrt{2}$  times better than our minimum sensitivity of  $7.5 \text{ mJy beam}^{-1}$  per  $5 \text{ km s}^{-1}$  channel (see §5.1). This is partially due to the convolutions during our post-processing, but mostly reflecting that our observations typically achieve  $\sim 24$  s per pointing, which translates to  $\sim 1$  min of effective integration time after accounting for overlap of mosaic fields. There is about

$\pm 30\%$  rms scatter in the point source sensitivity across our sample.

Our achieved sensitivity translates to a median point mass sensitivity of  $2 \times 10^4 \text{ M}_\odot$  for a Galactic  $\alpha_{\text{CO}}$  and  $R_{21} = 0.65$ . As discussed above, this means that, on average, we will detect individual GMCs, which are often defined to have mass  $> 10^5 \text{ M}_\odot$ , at signal-to-noise of 5 or higher.

Our median surface brightness sensitivity is  $85 \text{ mK}$  per  $2.54 \text{ km s}^{-1}$  channel, again better than the nominal target of  $170 \text{ mK}$  per  $5 \text{ km s}^{-1}$  channel. This surface brightness sensitivity depends sharply on the achieved angular resolution, with rms noise  $\propto \theta^{-2}$  even for fixed point source sensitivity. This number thus partially also reflects that our achieved median resolution is  $\sim 1.3''$  while the target was  $1''$ . Because the surface bright-



**Figure 31. Sensitivity of PHANGS-ALMA.** The *top left* panel shows the median rms noise in each  $\delta v \approx 2.5 \text{ km s}^{-1}$  channel of each final data cube (at their native resolution) as a function of the distance to that galaxy. We mark 7-m data and 12-m+7-m data with separate symbols. Dashed lines show the cases where the noise equates to a  $1\sigma$  mass sensitivity of  $10^4 \text{ M}_\odot$  or  $10^5 \text{ M}_\odot$ , assuming  $\alpha_{\text{CO}} = 4.35 \text{ M}_\odot \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$ ,  $R_{21} = 0.65$ , and  $\Delta v = 5 \text{ km s}^{-1}$ . Most of our cubes have  $1\sigma$  sensitivities between  $10^4 \text{ M}_\odot$  and  $10^5 \text{ M}_\odot$  per beam. The other panels show the cumulative distributions of noise in the 12-m+7-m data in (*top right*)  $\text{Jy beam}^{-1}$  units, (*bottom left*) surface brightness units, i.e., milliKelvin, at the native resolution, and (*bottom right*) surface brightness units at a common 150 pc resolution. The upper axis in the last panel also shows the corresponding  $1\sigma$  mass sensitivity assuming the same  $\alpha_{\text{CO}} = 4.35 \text{ M}_\odot \text{ pc}^{-2} (\text{K km s}^{-1})^{-1}$ ,  $R_{21} = 0.65$ , and  $\Delta v = 5 \text{ km s}^{-1}$ .

ness sensitivity depends on both point source sensitivity and achieved angular resolution, our data show a  $\pm 50\%$  scatter about the median value.

Again assuming a full line width  $\Delta v = 5 \text{ km s}^{-1}$ , this 85 mK per  $2.54 \text{ km s}^{-1}$  channel translates to  $1\sigma$  surface brightness sensitivity of  $\sim 0.30 \text{ K km s}^{-1}$ . Following Equation 11, this translates to median  $1\sigma$  mass surface density sensitivity of  $\Sigma_{\text{mol}} = 2.0 \text{ M}_\odot \text{ pc}^{-2}$ . As discussed above, typical resolved GMCs have surface densities of a few times  $10\text{--}100 \text{ M}_\odot \text{ pc}^{-2}$ . Thus our maps should recover CO emission everywhere where GMCs fill an appreciable fraction of the beam.

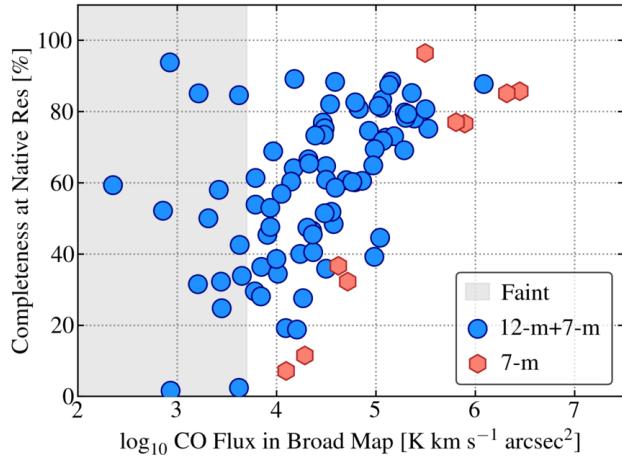
As discussed above, 150 pc is a common physical scale for most of the data. When we convolve to this scale, the typical surface brightness sensitivity is  $1\sigma = 50 \text{ mK}$ ,

again with  $\pm 50\%$  scatter. The corresponding  $1\sigma$  point mass sensitivity is  $3.1 \times 10^4 \text{ M}_\odot$ , and the  $1\sigma$  mass surface density sensitivity is  $\Sigma_{\text{mol}} = 1.2 \text{ M}_\odot \text{ pc}^{-2}$ .

Together these sensitivity numbers can all be roughly summarized as: the PHANGS-ALMA CO maps are sensitive to individual massive giant molecular clouds at the  $3\text{--}10\sigma$  level. This has been key to the already published studies of Sun et al. (2018, 2020b) and Rosolowsky et al. (2021), which have used PHANGS-ALMA to study populations of individual clouds across our sample.

### 7.1.3. Completeness

Table 15 and Figure 32 illustrate the completeness,  $c$ , of the data. Following Equation 12, we define “completeness” as the ratio of the flux detected within the



**Figure 32. Completeness of the PHANGS-ALMA high signal-to-noise maps at the native resolution.** We defined completeness as the ratio of flux within the high confidence “strict” mask to the integrated flux in the cube or the high completeness “broad” mask. Thus the number reflects the fraction of the total flux captured at moderate-to high signal-to-noise in each cube. The completeness correlates with total flux, to first order tracking the overall brightness of CO in the target.

“strict” mask to the total flux in the cube, estimated from either a direct sum or using the “broad” mask. This completeness thus measures how much of the flux in the cube is detected at moderate signal to noise at a given sensitivity and resolution. Figure 32 shows completeness for the native resolution and Tables 15 and 16 also report values at 150 pc resolution.

At our native resolution, the 12-m+7-m+TP data have median completeness of  $\sim 61\%$ , but with a wide range. This improves to  $\sim 67\%$  at 150 pc resolution, but still shows a wide range. Figure 32 shows that the completeness variations track the overall flux of the target, with brighter galaxies yielding higher recovery. A few of the galaxies are so faint that a simple sum of the cube no longer yields a flux consistent with integrating over the broad, high completeness map. This typically occurs when  $F_{\text{CO}} \lesssim 10^{3.5} \text{ K km s}^{-1} \text{ arcsec}^2$ . We mark this regime in Figure 32 and note galaxies that have flux lower than this in Table 16. To avoid instability when calculating  $c$ , we use the maximum flux of either the direct sum of the cube or the broad map in Equation 12.

Excluding these “faint” galaxies increases the median completeness in the sample by  $\sim 5\%$ , but Figure 32 shows that there is still a trend with brightness at higher flux: we detect  $\sim 60\text{--}90\%$  of the flux at native resolution in high-flux galaxies. Meanwhile for galaxies with intermediate integrated fluxes, the completeness varies across almost the whole range of possible values, but is lower on average.

Overall these levels of flux recovery are very good, and in line with previous measurements on Local Group and related data by Leroy et al. (2016) and previous measurements in PHANGS-ALMA by Sun et al. (2018, 2020b). The completeness increases dramatically as we degrade the resolution of the data, with almost all area in PHANGS-ALMA detected at  $\sim 15''$  or  $\sim 1$  kpc resolution (recall that this is expected because PHANGS-ALMA targets regions with detectable mid-IR dust emission).

Note that Sun et al. (2020b) give equations for the selection function associated with our strict masking technique. These can be useful for modeling the properties of the emission not included in the strict mask.

## 7.2. Description of high level data products

Leroy et al. (2021) describe the creation of the PHANGS-ALMA data products, and we refer the reader to that paper for details. Here we provide a high level summary of the products that make up the PHANGS-ALMA data release:

1. **Cubes at multiple resolutions.** We convolve the native resolution, post-processed data cubes to a series of common physical and angular resolutions. In addition to the native resolution, we convolve all data to angular resolutions of  $\theta = 2'', 7.5'', 11'',$  and  $15''$  (FWHM). Almost all of the 12-m+7-m data can reach each of these angular resolutions, while the 7-m-only data sets can typically only reach  $11''$  and  $15''$ .

We also convolve each cube to a FWHM resolution corresponding to 60, 90, 120, 150, 500, 750, and 1000 pc using the distances compiled by Anand et al. (2021) and reported in Table 3. As discussed above the first four values roughly correspond to the quartiles of the distribution of physical resolutions for our native resolution 12-m+7-m data set.

During these convolutions we impose a 10% tolerance. That is, if the FWHM beam of a cube is already within  $\pm 10\%$  of the target resolution, we do not convolve but label the cube as belonging to that fixed-resolution subset. We do not alter the cube metadata in these cases, so that, e.g., the FITS header BMAJ and BMIN keywords still reflect the true values.

Because most scientific applications involve measuring surface brightness or integrating over an aperture larger than the beam, we present the PHANGS-ALMA cubes in units of Kelvin. Conversions to Jy beam $^{-1}$  are provided as part of the header metadata.

2. **Noise models for each cube.** For each data cube at each resolution, we construct a three dimensional noise model. This model gives our

best estimate of the rms noise at each position–position–velocity pixel. We construct the noise model by treating the spectral and spatial variations of the noise as separable and then using iterative outlier rejection to isolate the parts of the cube most likely to be signal free. When discussing the noise statistics above, we quote the median value across the whole noise model, but see Leroy et al. (2021) for illustrations of the spatial and spectral variations of the noise.

3. **“Strict” and “broad” masks for each cube.** Combining the cubes and the noise models, we construct two sets of three dimensional masks that identify the location of likely signal. Our “strict” masks are *high confidence*, meaning that they include only pixels with at least moderate signal to noise at the relevant resolution. The strict masks should be used in calculations that require secure detections. These are the basis of our completeness calculations above.

We also construct “broad” masks that include all regions of the cube likely to contain signal. These are *high completeness* masks, in the sense that we expect them to include essentially all emission from the galaxy. To build the broad masks, we use the union of the strict masks at all resolutions, from the native resolution up to 15'' and 1 kpc. The coarse resolution masks play a crucial role here because we detect emission along almost every line of sight in those cubes. Including the high resolution masks in the union tends to matter most for high line width, compact features, which mostly occur in galaxy centers. Within the cubes, the broad masks typically look like an extended, “puffed up,” version of the galaxy rotation curve in position–position velocity space.

Because the broad mask represents a union of all strict masks, there is only a single broad mask, which we apply across all resolutions. We reproject it onto different three dimensional astrometric grids as necessary.

4. **Integrated intensity maps.** We apply the masks to the cubes and integrate along the spectral axis to produce two-dimensional maps of line-integrated intensity (“moment 0”), in units of  $\text{K km s}^{-1}$ . We create versions of these maps for each resolution and for both the strict and the broad masks. The “strict” integrated intensity maps should be used as our best map of securely detected CO emission from each galaxy. The “broad” integrated intensity maps represent our most complete map of CO emission from each galaxy. The broad maps will include some negative pixels, reflecting that they include regions dominated by noise but likely to contain faint

emission. Integrating across these broad maps should yield the total flux in the cube to good approximation.

The integrated intensity,  $I_{\text{CO}}^{2-1}$  in units of  $\text{K km s}^{-1}$ , can be translated into an estimate of molecular gas mass surface density,  $\Sigma_{\text{mol}}$  in units of  $\text{M}_{\odot} \text{ pc}^{-2}$ , following Equation 10 or Equation 11 for standard assumptions. At the resolution of PHANGS–ALMA it remains ambiguous whether or not to apply an inclination correction,  $\cos i$ , when estimating  $\Sigma_{\text{mol}}$ . There is a long history of assuming individual molecular clouds to have an isotropic, spherically symmetric geometry (e.g., Solomon et al. 1987; Rosolowsky & Leroy 2006). However, our resolution is of the same order as the  $\sim 100$  pc disk scale height (Heyer & Dame 2015; Sun et al. 2020a). For now, the user will need to make their own best judgment on this topic.

For each integrated intensity map, we also propagate the noise model into maps of the associated statistical uncertainty in  $I_{\text{CO}}^{2-1}$  (see details in Leroy et al. 2021).

5. **Peak intensity maps.** We also calculate maps of the peak intensity,  $I_{\text{CO}}^{2-1}$  along each line of sight. These maps of peak intensity, also called “peak temperature” or “moment 8,” are useful to highlight the detailed structure of emission in the cube, though they have a less straightforward physical interpretation than the line-integrated intensity maps.

To calculate the peak intensity, we first calculate the full velocity range covered by the relevant mask for each sightline. Then for each line of sight, we find the maximum intensity at any velocity in this range and record this as the peak intensity.

We also create a version of the peak intensity map after convolving the spectral axis of the cube with a  $12.5 \text{ km s}^{-1}$  boxcar kernel. Because  $12.5 \text{ km s}^{-1}$  corresponds to a typical full line width for a bright molecular cloud, this spectral smoothing essentially represents running a matched filter across the spectral axis.

We present the peak intensity maps in units of Kelvin.

6. **Velocity fields.** We calculate the intensity-weighted mean velocity of emission, frequently referred to as “moment 1,” along each line of sight. We present two versions of the intensity-weighted mean velocity field. The first uses only emission within the strict mask. These strictly masked velocity maps present the mean velocity associated with all securely detected emission.

The second set of velocity maps consider more emission and so cover a larger area. Because

noise spikes can contribute spurious outliers to the velocity field, we apply several prior expectations when constructing this velocity field. These “moment 1 with prior” maps begin with the broad mask. We calculate the intensity-weighted mean velocity using the broad masks. This typically includes many spurious velocities associated with low signal-to-noise or even signal-free regions. Therefore, we blank all values from the map that either (1) have signal to noise less than two in the integrated intensity map or (2) deviate from a “prior expectation” velocity field by more than some tolerance,  $\pm 30 \text{ km s}^{-1}$  by default. For this release, we use the velocity field calculated using the strict mask at lower resolution.

We present all velocity fields in units of  $\text{km s}^{-1}$  and also calculate maps of associated statistical uncertainty by propagating errors from the noise model cubes.

**7. Line width maps.** We calculate two sets of line width maps, both using only the strict masks. First, we calculate the rms velocity dispersion about the intensity-weighted mean velocity, often referred to as “moment 2.” This calculation diverges in the presence of noise, so we only present the calculation for the strictly masked case.

The second moment can be sensitive to clipping and unstable in the presence of noise. We also calculate and report the “effective width” following the definition of Heyer et al. (2001) and discussed in Leroy et al. (2016) and Sun et al. (2018, 2020b). In this measurement, the line width  $\sigma_{\text{ew}} \equiv \int I_{\text{CO}} \text{d}v / (\sqrt{2\pi} I_{\text{CO}}^{\text{peak}})$ . In other words, the line width is defined as the ratio of the line integrated intensity to the peak intensity. In the limit of a well-resolved, high signal-to-noise Gaussian line profile the two line width measurements match. The effective width is more robust to the presence of noise but also sensitive to the requirement to resolve the line. The ratio between the two line width measurements can also indicate deviations from a single Gaussian spectral shape.

For both line width estimates, we present the result in units of  $\text{km s}^{-1}$  and also calculate and deliver associated maps of statistical uncertainty. For this delivery, we do not correct the measurements for the line spread function. This can be an important effect for narrow lines and should be addressed by the user (e.g., following Leroy et al. 2016; Koch et al. 2018; Sun et al. 2020a; Rosolowsky et al. 2021).

This already represents a rich set of data products that characterize the CO spectral line position-by-position across each target. We also provide several supporting

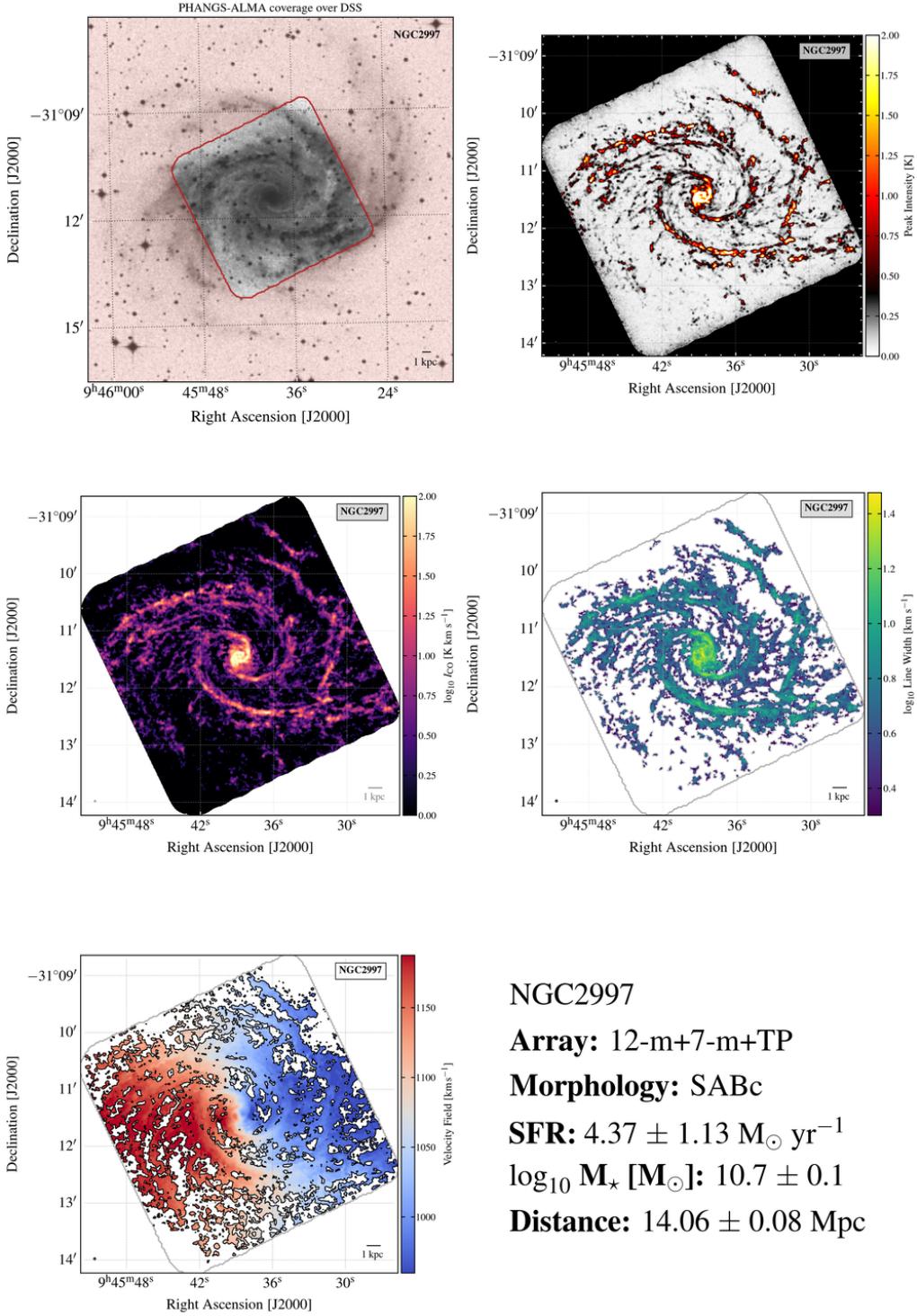
data products, including maps indicating the coverage of the ALMA observations. There are also several clear next steps. In the inner regions of galaxies and in regions of colliding flows (e.g., bar ends, see Bešlić et al. 2021), multi-component decomposition like that presented in Henshaw et al. (2020) represents a better treatment of the spectrum than single component methods. In the future, we also intend to use the rotation curves fit in Lang et al. (2020) to refine the “priors” use to construct the velocity fields (e.g., see Colombo et al. 2014b). Our next round of products will also likely include “shuffled” cubes appropriate for spectral stacking (e.g., see Schruba et al. 2011).

## 8. ATLAS

### Fig. Set A. Atlas of PHANGS-ALMA maps.

A figure set in the online version of this paper presents an atlas of data products for the PHANGS-ALMA targets, which we illustrate with an example in Figure 33. For each galaxy, we illustrate the area mapped by ALMA over the DSS optical image of the galaxy (*top left*). Then we show four ALMA data products:

1. The peak intensity map calculated over a  $12.5 \text{ km s}^{-1}$  window (*top right*) highlights detailed structure of each galaxy. Arms, bars, filamentary structure, and individual cloud complexes are all highly visible in these maps. Because the maps show the maximum over the full velocity extent of the mask, these offer a direct view of the contents of the data cubes.
2. The integrated intensity map (*middle left*) on a logarithmic stretch shows the distribution of CO (2-1) emission in the galaxy. Here we show the integrated intensity map constructed using the “broad”, i.e., high completeness, mask. Therefore these maps show almost all emission from the galaxy, though note that we do set the lower end of the color stretch to  $1 \text{ K km s}^{-1}$ , cloud to our nominal  $3\sigma$  limit (see Table 15).
3. A map of line width (*middle right*) on a logarithmic stretch. In the atlas, we show  $1\sigma$  line widths calculated using the effective width. Because the line width is challenging to calculate in the presence of noise, these maps use the “strict”, i.e., high confidence, masks.
4. The velocity field appears in the *bottom left* panel. We show the “moment 1 with prior”, i.e., the intensity-weighted mean velocity calculated within the broad, high completeness mask and then pruned using a low resolution prior estimate of the velocity. This represents our velocity field with the highest covering fraction and illustrates both large and small-scale systematic motions in the galaxy.



**Figure 33. PHANGS-ALMA Image Atlas.** PHANGS-ALMA data products for one galaxy, NGC 2997. *Top left:* ALMA mapping coverage, indicated by the unshaded area, over the DSS optical image. *Top right:* Peak intensity image. *Middle left:* Integrated intensity (“moment 0”) image on a logarithmic stretch, calculated summing over the “broad,” high completeness, mask (§7). *Middle right:* Line width, here  $1\sigma$  estimated from the effective width. *Bottom left:* intensity weighted mean velocity field constructed using a low resolution prior. The bottom right panel provides key information for the galaxy. Similar figures for all 90 PHANGS-ALMA targets appear as part of an online figure set.

Along with these images, each figure notes a few key properties of each galaxy along with the best available array combination.

## 9. SUMMARY

This paper presents PHANGS-ALMA, the first large  $\theta \approx 1'' \approx 100$  pc resolution CO(2–1) survey of a representative nearby galaxy population. The paper describes:

1. **The scientific motivation for PHANGS-ALMA (§2).** We summarize how PHANGS-ALMA builds on previous CO mapping efforts in the local universe, highlight that many key physics related to both star formation and galaxy evolution occur at or near the scale of individual giant molecular clouds (GMCs), and outline the high-level science goals of PHANGS-ALMA:
  - Measure the demographics of molecular clouds, and measure how GMC populations depend on host galaxy and location in a galaxy.
  - Measure the star formation efficiency per free fall time,  $\epsilon_{\text{ff}}$ , at cloud scales. Measure how  $\epsilon_{\text{ff}}$  depends on the density, dynamical state, and turbulence in molecular clouds.
  - Quantify the “violent cycling” between phases of the star formation process. Use this to constrain the life cycle of clouds and feedback.
  - Measure how the self-regulated, large scale structure of galaxy disks emerges from a medium made of individual clouds and star-forming regions.
  - Measure the motions, flows, and organization of cold gas in galaxies at 100–1,000 pc scales.
2. **The sample selection (§3).** PHANGS-ALMA attempts to target all relatively massive, actively star-forming galaxies within 17 Mpc that are not heavily inclined. This means that the sample selection is unbiased, easy to understand, and simple to attempt to reproduce. It also means that a secondary goal of PHANGS-ALMA is that ALMA provide a high quality CO map of all local galaxies. Our estimates of key galaxy properties, including distance and stellar mass, have evolved some since the original sample selection. We show that while uncertainties in galaxy properties do affect our exact sample, we include approximately the expected number of targets. We also show that the targets do a good job of spanning the  $z = 0$  main sequence of star-forming galaxies and the mass radius relation for late-type galaxies.
3. **Calculation and presentation of the properties of the sample galaxies (§4).** We present estimates the physical properties of the PHANGS-ALMA targets: size, mass, CO luminosity, and

star formation rate (SFR). We draw distances from the recent work by [Anand et al. \(2021\)](#). We compare estimates of these quantities using different methods and provide quantitative translations between different methods of estimating size, mass, and SFR. This includes new work on the size and stellar mass estimation.

We attempt to present numbers on a system that is self-consistent with work on the SDSS main galaxy sample by [Salim et al. \(2016\)](#) and [Salim et al. \(2018\)](#), which is itself broadly consistent with earlier work on the SDSS main galaxy sample. To verify this overall consistency, we show a good agreement between the properties of the PHANGS-ALMA targets and previous measurements of the main sequence of star-forming galaxies and the mass-radius relation.

PHANGS-ALMA typically covers only  $\sim 70\%$  of the star formation activity, and so presumably missed  $\sim 30\%$  of the molecular gas. To account for this, we provide aperture corrections that can be used to correct the measured CO luminosity to the full CO luminosity for comparison to other global properties or unresolved CO surveys. We consider several aperture correction templates, including an exponential disk model and several bands. We find that WISE3 12  $\mu\text{m}$  emission provides the best template for such an aperture correction, in good agreement with recent work by [Chown et al. \(2021\)](#).

4. **A summary of the PHANGS-ALMA observations (§5).** We observe the area of active star formation, as gauged from mid-infrared (WISE 12  $\mu\text{m}$ ) emission. To cover this area, we use large mosaics. In many cases where the galaxy exceeds the area covered by a single 150 pointing mosaic we break a galaxy into parts and observe each part separately, combining the data into a single cube in post-processing. The observations spanned 6 years and the typical elevation and atmospheric conditions were excellent,  $> 50^\circ$  and 1–1.4 mm of precipitable water vapor. Even though the total observing time on any given field using the 12-m array was short, the achieve  $u-v$  coverage for the combined 12-m+7-m data appears excellent and yields a synthesized beam without dramatic sidelobe features.
5. **A short summary of the data processing (§6).** We provide a short overview of the PHANGS-ALMA post-processing and product creation pipeline. The full pipeline is presented in [Leroy et al. \(2021\)](#) and will be publicly available along with the PHANGS-ALMA data products.
6. **A description of the properties of the PHANGS-ALMA cubes and data products**

(§7). We cover 81 targets with the combined 12-m, 7-m, and total power arrays and 9 targets with the 7-m array or combined 7-m plus total power arrays. A typical 12-m+7-m+TP map covers  $\sim 7$  arcmin $^2$  or 124 kpc $^2$ . In total, the survey maps  $\sim 1050$  arcmin $^2$  or 10,650 kpc $^2$ .

For the 12-m+7-m+TP data, the median angular resolution is 1.3'', the channel width is 2.54 km s $^{-1}$ , and the typical 1 $\sigma$  sensitivity is  $\sim 6.2$  mJy beam $^{-1}$  or 85 mK. These values translate to a median physical resolution of 100 pc and 1 $\sigma$  sensitivity to mass of  $\sim 2 \times 10^4$  M $_{\odot}$  for typical Milky Way conditions. Across almost the whole survey,  $\theta = 150$  pc represents a common physical resolution achievable for most data. At the native resolution, our cubes recover a median  $\sim 60\%$  of the total flux in the noise at good signal to noise.

## 7. An atlas showing the data for each target (§8).

A figure set shows the areal coverage, peak and integrated intensity, velocity field, and line width maps for each galaxy.

## ACKNOWLEDGMENTS

We thank the anonymous referee and the editors for a rapid, constructive review during a difficult time.

This work was carried out as part of the PHANGS collaboration.

The work of A.K.L., J.S., and D.U. was partially supported by the National Science Foundation (NSF) under Grants No.1615105, 1615109, and 1653300, as well as by the National Aeronautics and Space Administration (NASA) under ADAP grants NNX16AF48G and NNX17AF39G. C.M.F. acknowledges support from the NSF under Award No. 1903946. A.S. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1903834.

ER acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-03987, and computational support from Compute Canada.

D.L., T.S., E.S., C.M.F., K.S., and T.G.W. acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 694343).

CH, AH, and JP acknowledge support by the Programme National “Physique et Chimie du Milieu Interstellaire” (PCMI) of CNRS/INSU with INC/INP co-funded by CEA and CNES. AH acknowledges support by the Programme National Cosmology et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES.

AU and AG-R acknowledge support from the Spanish funding grants AYA2016-79006-P (MINECO/FEDER)

and PID2019-108765GB-I00 (MICINN). AU acknowledges support from the Spanish funding grant PGC2018-094671-B-I00 (MCIU/AEI/FEDER). MQ acknowledges support from the research project PID2019-106027GA-C44 from the Spanish Ministerio de Ciencia e Innovación.

M.C. and J.M.D.K. gratefully acknowledge funding from the German Research Foundation (DFG) through an Emmy Noether Research Group (grant number KR4801/1-1). M.C., J.M.D.K., and J.J.K. gratefully acknowledge funding from the DFG Sachbeihilfe (grant number KR4801/2-1). J.M.D.K. gratefully acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme via the ERC Starting Grant MUSTANG (grant agreement number 714907).

FB, ATB, IB, JdB, JP acknowledge funding from the European Union’s Horizon 2020 research and innovation programme (grant agreement No 726384/EMPIRE).

CE acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) Sachbeihilfe, grant number BI1546/3-1.

R.S.K., S.C.O.G. and M.C.S. acknowledge financial support from the DFG via the collaborative research center (SFB 881, Project-ID 138713538) “The Milky Way System” (subprojects A1, B1, B2, and B8). They also acknowledge subsidies from the Heidelberg Cluster of Excellence *STRUCTURES* in the framework of Germany’s Excellence Strategy (grant EXC-2181/1 - 390900948) and funding from the ERC via the ERC Synergy Grant *ECOGAL* (grant 855130).

KK and FS gratefully acknowledge funding from the DFG in the form of an Emmy Noether Research Group (grant number KR4598/2-1).

EW acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 (“The Milky Way System”, subproject P2).

PSB acknowledges support from the research project PID2019-107427GB-C31 from the Spanish Ministerio de Ciencia Innovación.

The work of E.C.O. was supported by the NSF under grant No. AST-1713949 and NASA under grant No. NNX17AG26G.

This paper makes use of the following ALMA data, which have been processed as part of the PHANGS-ALMA CO(2-1) survey:

ADS/JAO.ALMA#2012.1.00650.S,  
 ADS/JAO.ALMA#2013.1.00803.S,  
 ADS/JAO.ALMA#2013.1.01161.S,  
 ADS/JAO.ALMA#2015.1.00121.S,  
 ADS/JAO.ALMA#2015.1.00782.S,  
 ADS/JAO.ALMA#2015.1.00925.S,  
 ADS/JAO.ALMA#2015.1.00956.S,  
 ADS/JAO.ALMA#2016.1.00386.S,  
 ADS/JAO.ALMA#2017.1.00392.S,  
 ADS/JAO.ALMA#2017.1.00766.S,

ADS/JAO.ALMA#2017.1.00886.L,  
 ADS/JAO.ALMA#2018.1.00484.S,  
 ADS/JAO.ALMA#2018.1.01321.S,  
 ADS/JAO.ALMA#2018.1.01651.S,  
 ADS/JAO.ALMA#2018.A.00062.S,  
 ADS/JAO.ALMA#2019.1.01235.S,  
 ADS/JAO.ALMA#2019.2.00129.S,  
 ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ. The National Radio Astronomy Observatory is a facility of the National Sci-

ence Foundation operated under cooperative agreement by Associated Universities, Inc.

*Software:* `aply`, ALMA Calibration Pipeline (L. Davis et al. in preparation), `CASA` (McMullin et al. 2007), `numpy` (Oliphant 2006), `scipy` (Virtanen et al. 2020), `astropy` (Astropy Collaboration et al. 2013, 2018), IDL Astronomy User’s Library (Landsman 1993), `cprops` (Rosolowsky & Leroy 2006), GILDAS Pety (2005), PHANGS-ALMA Pipeline (Leroy et al. 2021), PHANGS-ALMA Total Power Pipeline (Herrera et al. 2020), `R` (R Core Team 2015), `spectral-cube` (Ginsburg et al. 2019), `radio-beam` (<https://github.com/radio-astro-tools/radio-beam>)

## REFERENCES

Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, *ApJ*, 770, 25

Alatalo, K., Davis, T. A., Bureau, M., et al. 2013, *MNRAS*, 432, 1796

Anand, G. S., Lee, J. C., Van Dyk, S. D., et al. 2021, *MNRAS*, 501, 3621

Armus, L., Mazzarella, J. M., Evans, A. S., et al. 2009, *PASP*, 121, 559

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

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123

Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, *MNRAS*, 388, 945

Barbarino, C., Dall’Ora, M., Botticella, M. T., et al. 2015, *MNRAS*, 448, 2312

Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, *ApJS*, 149, 289

Bešlić, I., Barnes, A. T., Bigiel, F., et al. 2021, *MNRAS* submitted

Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, 136, 2846

Blanton, M. R., & Moustakas, J. 2009, *ARA&A*, 47, 159

Blitz, L. 1993, in *Protostars and Planets III*, ed. E. H. Levy & J. I. Lunine, 125–161

Blitz, L., Fukui, Y., Kawamura, A., et al. 2007, in *Protostars and Planets V*, ed. B. Reipurth, D. Jewitt, & K. Keil, 81–96

Blitz, L., & Rosolowsky, E. 2006, *ApJ*, 650, 933

Boissier, S., Prantzos, N., Boselli, A., & Gavazzi, G. 2003, *MNRAS*, 346, 1215

Bolatto, A. D., Leroy, A. K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, *ApJ*, 686, 948

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013a, *ARA&A*, 51, 207

Bolatto, A. D., Warren, S. R., Leroy, A. K., et al. 2013b, *Nature*, 499, 450

Bolatto, A. D., Wong, T., Utomo, D., et al. 2017, *ApJ*, 846, 159

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, *A&A*, 622, A103

Boquien, M., Kennicutt, R., Calzetti, D., et al. 2016, *A&A*, 591, A6

Bothwell, M. S., Wagg, J., Cicone, C., et al. 2014, *MNRAS*, 445, 2599

Braine, J., & Combes, F. 1992, *A&A*, 264, 433

Braine, J., Ferguson, A. M. N., Bertoldi, F., & Wilson, C. D. 2007, *ApJL*, 669, L73

Braine, J., Rosolowsky, E., Gratier, P., Corbelli, E., & Schuster, K. F. 2018, *A&A*, 612, A51

Brownson, S., Belfiore, F., Maiolino, R., Lin, L., & Carniani, S. 2020, *MNRAS*, 498, L66

Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000

Burkhart, B. 2018, *ApJ*, 863, 118

Caldú-Primo, A., & Schruba, A. 2016, *AJ*, 151, 34

Calzetti, D. 2013, in *Secular Evolution of Galaxies*, ed. J. Falcón-Barroso & J. H. Knapen, 419

Calzetti, D., Kennicutt, Jr., R. C., Bianchi, L., et al. 2005, *ApJ*, 633, 871

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245

Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, *MNRAS*, 476, 875

Chabrier, G. 2003, *PASP*, 115, 763

Chastenet, J., Sandstrom, K., Chiang, I. D., et al. 2019, *ApJ*, 876, 62

Chevance, M., Kruijssen, J. M. D., Krumholz, M. R., et al. 2020a, arXiv e-prints, arXiv:2010.13788

Chevance, M., Kruijssen, J. M. D., Hygate, A. P. S., et al. 2020b, *MNRAS*, 493, 2872

Chevance, M., Kruijssen, J. M. D., Vazquez-Semadeni, E., et al. 2020c, *SSRv*, 216, 50

Chown, R., Li, C., Parker, L., et al. 2021, *MNRAS*, 500, 1261

Colombo, D., Hughes, A., Schinnerer, E., et al. 2014a, *ApJ*, 784, 3

Colombo, D., Meidt, S. E., Schinnerer, E., et al. 2014b, *ApJ*, 784, 4

Colombo, D., Kalinova, V., Utomo, D., et al. 2018, *MNRAS*, 475, 1791

Corbelli, E., Braine, J., Bandiera, R., et al. 2017, *A&A*, 601, A146

Corder, S., Sheth, K., Scoville, N. Z., et al. 2008, *ApJ*, 689, 148

Cormier, D., Bigiel, F., Jiménez-Donaire, M. J., et al. 2018, *MNRAS*, 475, 3909

Cornwell, T. J. 2008, *IEEE Journal of Selected Topics in Signal Processing*, 2, 793

Cotton, W. D. 2017, *PASP*, 129, 094501

Courtois, H. M., & Tully, R. B. 2012, *ApJ*, 749, 174

Dale, D. A., Anderson, K. R., Bran, L. M., et al. 2020, *AJ*, 159, 195

Dale, J. E. 2015, *NewAR*, 68, 1

Davis, S. W., Jiang, Y.-F., Stone, J. M., & Murray, N. 2014, *ApJ*, 796, 107

Davis, T. A., & McDermid, R. M. 2017, *MNRAS*, 464, 453

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, Herold G., J., et al. 1991, *Third Reference Catalogue of Bright Galaxies* (New York: Springer)

den Brok, J. S., Chatzigiannakis, D., Bigiel, F., et al. 2021, *MNRAS* submitted

Dobbs, C. L., Rosolowsky, E., Pettitt, A. R., et al. 2019, *MNRAS*, 485, 4997

Dobbs, C. L., Krumholz, M. R., Ballesteros-Paredes, J., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 3

Donovan Meyer, J., Koda, J., Momose, R., et al. 2012, *ApJ*, 744, 42

—. 2013, *ApJ*, 772, 107

Downes, D., & Solomon, P. M. 1998, *ApJ*, 507, 615

Druard, C., Braine, J., Schuster, K. F., et al. 2014, *A&A*, 567, A118

Egusa, F., Mentuch Cooper, E., Koda, J., & Baba, J. 2017, *MNRAS*, 465, 460

Elmegreen, B. G. 1989, *ApJ*, 338, 178

—. 2000, *ApJ*, 530, 277

Elmegreen, B. G., Elmegreen, D. M., & Efremov, Y. N. 2018, *ApJ*, 863, 59

Engargiola, G., Plambeck, R. L., Rosolowsky, E., & Blitz, L. 2003, *ApJS*, 149, 343

Engelbracht, C. W., Kundurthy, P., Gordon, K. D., et al. 2006, *ApJL*, 642, L127

Espada, D., Verley, S., Miura, R. E., et al. 2019, *ApJ*, 887, 88

Evans, II, N. J., Heiderman, A., & Vutisalchavakul, N. 2014, *ApJ*, 782, 114

Faesi, C. M., Lada, C. J., & Forbrich, J. 2018, *ApJ*, 857, 19

Federrath, C., & Klessen, R. S. 2012, *ApJ*, 761, 156

—. 2013, *ApJ*, 763, 51

Federrath, C., Roman-Duval, J., Klessen, R. S., Schmidt, W., & Mac Low, M. M. 2010, *A&A*, 512, A81

Foyle, K., Rix, H.-W., Walter, F., & Leroy, A. K. 2010, *ApJ*, 725, 534

Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, *ApJ*, 553, 47

Freeman, P., Rosolowsky, E., Kruijssen, J. M. D., Bastian, N., & Adamo, A. 2017, *MNRAS*, 468, 1769

Fukui, Y., Habe, A., Inoue, T., Enokiya, R., & Tachihara, K. 2020, *PASJ*, arXiv:2009.05077

Fukui, Y., & Kawamura, A. 2010, *ARA&A*, 48, 547

Fukui, Y., Mizuno, N., Yamaguchi, R., et al. 1999, *PASJ*, 51, 745

Fukui, Y., Kawamura, A., Minamidani, T., et al. 2008, *ApJS*, 178, 56

García-Burillo, S., Combes, F., Hunt, L. K., et al. 2003, *A&A*, 407, 485

García-Burillo, S., Combes, F., Usero, A., et al. 2014, *A&A*, 567, A125

Gatto, A., Walch, S., Low, M. M. M., et al. 2015, *MNRAS*, 449, 1057

Geen, S., Bieri, R., Rosdahl, J., & de Koter, A. 2021, *MNRAS*, 501, 1352

Geen, S., Hennebelle, P., Tremblin, P., & Rosdahl, J. 2016, *MNRAS*, 463, 3129

Gentry, E. S., Krumholz, M. R., Dekel, A., & Madau, P. 2017, *MNRAS*, 465, 2471

Ginsburg, A., Koch, E., Robitaille, T., et al. 2019, *radio-astro-tools/spectral-cube: Release v0.4.5*, doi:10.5281/zenodo.3558614

Grasha, K., Calzetti, D., Bittle, L., et al. 2018, *MNRAS*, 481, 1016

Gratier, P., Braine, J., Rodriguez-Fernandez, N. J., et al. 2010, *A&A*, 512, A68+

—. 2012, *A&A*, 542, A108

Groves, B., Krause, O., Sandstrom, K., et al. 2012, *MNRAS*, 426, 892

Grudić, M. Y., Kruijssen, J. M. D., Faucher-Giguère, C.-A., et al. 2020, *arXiv e-prints*, arXiv:2008.04453

Hartmann, L., Ballesteros-Paredes, J., & Bergin, E. A. 2001a, *ApJ*, 562, 852

—. 2001b, *ApJ*, 562, 852

Helfer, T. T., Thornley, M. D., Regan, M. W., et al. 2003, *ApJS*, 145, 259

Hennebelle, P., & Chabrier, G. 2011, *ApJL*, 743, L29

Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, *Nature Astronomy*, 4, 1064

Herrera, C. N., Pety, J., Hughes, A., et al. 2020, *A&A*, 634, A121

Heyer, M., & Dame, T. M. 2015, *ARA&A*, 53, 583

Heyer, M. H., Carpenter, J. M., & Snell, R. L. 2001, *ApJ*, 551, 852

Hirota, A., Egusa, F., Baba, J., et al. 2018, *PASJ*, 70, 73

Högbom, J. A. 1974, *A&AS*, 15, 417

Hopkins, P. F. 2013, *MNRAS*, 428, 1950

Hopkins, P. F., Kereš, D., Murray, N., Quataert, E., & Hernquist, L. 2012, *MNRAS*, 427, 968

Hopkins, P. F., Narayanan, D., & Murray, N. 2013, *MNRAS*, 432, 2647

Huang, C. D., Riess, A. G., Yuan, W., et al. 2020, *ApJ*, 889, 5

Huang, M.-L., & Kauffmann, G. 2015, *MNRAS*, 450, 1375

Hughes, A., Meidt, S. E., Colombo, D., et al. 2013a, *ApJ*, 779, 46

Hughes, A., Meidt, S. E., Schinnerer, E., et al. 2013b, *ApJ*, 779, 44

Hunt, L. K., García-Burillo, S., Casasola, V., et al. 2015, *A&A*, 583, A114

Ibáñez-Mejía, J. C., Mac Low, M.-M., Klessen, R. S., & Baczyński, C. 2016, *ApJ*, 824, 41

Inoue, T., & Fukui, Y. 2013, *ApJL*, 774, L31

Jackson, J. M., Finn, S. C., Chambers, E. T., Rathborne, J. M., & Simon, R. 2010, *ApJL*, 719, L185

Jameson, K. E., Bolatto, A. D., Leroy, A. K., et al. 2016, *ApJ*, 825, 12

Jarrett, T. H., Chester, T., Cutri, R., Schneider, S. E., & Huchra, J. P. 2003, *AJ*, 125, 525

Jarrett, T. H., Masci, F., Tsai, C. W., et al. 2013, *AJ*, 145, 6

Jeffreson, S. M. R., & Kruijssen, J. M. D. 2018, *MNRAS*, 476, 3688

Jeffreson, S. M. R., Kruijssen, J. M. D., Keller, B. W., Chevance, M., & Glover, S. C. O. 2020, *MNRAS*, 498, 385

Jiménez-Donaire, M. J., Bigiel, F., Leroy, A. K., et al. 2019, *ApJ*, 880, 127

Jogee, S., Scoville, N., & Kenney, J. D. P. 2005, *ApJ*, 630, 837

Kannappan, S. J., Stark, D. V., Eckert, K. D., et al. 2013, *ApJ*, 777, 42

Karachentsev, I. D., Karachentseva, V. E., Huchtmeier, W. K., & Makarov, D. I. 2004, *AJ*, 127, 2031

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, *ApJ*, 730, 61

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33

Kawamura, A., Mizuno, Y., Minamidani, T., et al. 2009, *ApJS*, 184, 1

Keller, B. W., & Kruijssen, J. M. D. 2020, *arXiv e-prints*, arXiv:2004.03608

Kenney, J. D. P., Wilson, C. D., Scoville, N. Z., Devereux, N. A., & Young, J. S. 1992, *ApJL*, 395, L79

Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531

Kennicutt, Jr., R. C. 1989, *ApJ*, 344, 685

Kennicutt, Jr., R. C., Calzetti, D., Walter, F., et al. 2007, *ApJ*, 671, 333

Kim, C.-G., & Ostriker, E. C. 2017, *ApJ*, 846, 133

Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2013, *ApJ*, 776, 1

Kim, J., Chevance, M., Kruijssen, J. M. D., et al. 2020a, *MNRAS* submitted, arXiv:2012.00019

Kim, J.-G., Kim, W.-T., & Ostriker, E. C. 2018, *ApJ*, 859, 68

—. 2019, *ApJ*, 883, 102

Kim, J.-G., Ostriker, E. C., & Filippova, N. 2020b, *arXiv e-prints*, arXiv:2011.07772

Kim, W.-T., Kim, C.-G., & Ostriker, E. C. 2020c, *ApJ*, 898, 35

Klessen, R. S., & Glover, S. C. O. 2016, *Saas-Fee Advanced Course*, 43, 85

Koch, E., Rosolowsky, E., & Leroy, A. K. 2018, *Research Notes of the American Astronomical Society*, 2, 220

Koch, E. W., & Rosolowsky, E. W. 2015, *MNRAS*, 452, 3435

Koda, J., Scoville, N., Sawada, T., et al. 2009, *ApJL*, 700, L132

Koda, J., Sawada, T., Sakamoto, K., et al. 2020, *ApJL*, 890, L10

Kourkchi, E., Courtois, H. M., Graziani, R., et al. 2020, *AJ*, 159, 67

Kourkchi, E., & Tully, R. B. 2017, *ApJ*, 843, 16

Kreckel, K., Blanc, G. A., Schinnerer, E., et al. 2016, *ApJ*, 827, 103

Kreckel, K., Faesi, C., Kruijssen, J. M. D., et al. 2018, *ApJL*, 863, L21

Kregel, M., van der Kruit, P. C., & de Grijs, R. 2002, *MNRAS*, 334, 646

Kruijssen, J. M. D. 2012, *MNRAS*, 426, 3008

Kruijssen, J. M. D., & Longmore, S. N. 2014, *MNRAS*, 439, 3239

Kruijssen, J. M. D., Schruba, A., Hygate, A. e. P. S., et al. 2018, *MNRAS*, 479, 1866

Krijssen, J. M. D., Schruba, A., Chevance, M., et al. 2019, *Nature*, 569, 519

Krumholz, M. R. 2014, *PhR*, 539, 49

Krumholz, M. R., & Dekel, A. 2012, *ApJ*, 753, 16

Krumholz, M. R., & McKee, C. F. 2005, *ApJ*, 630, 250

—. 2020, *MNRAS*, 494, 624

Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, *ARA&A*, 57, 227

Kuno, N., Sato, N., Nakanishi, H., et al. 2007, *PASJ*, 59, 117

La Vigne, M. A., Vogel, S. N., & Ostriker, E. C. 2006, *ApJ*, 650, 818

Landsman, W. B. 1993, in *Astronomical Society of the Pacific Conference Series*, Vol. 52, *Astronomical Data Analysis Software and Systems II*, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes, 246

Lang, D. 2014, *AJ*, 147, 108

Lang, P., Meidt, S. E., Rosolowsky, E., et al. 2020, *ApJ*, 897, 122

Lange, R., Driver, S. P., Robotham, A. S. G., et al. 2015, *MNRAS*, 447, 2603

Lee, E. J., Miville-Deschénes, M.-A., & Murray, N. W. 2016, *ApJ*, 833, 229

Lee, J. C., Whitmore, B. C., Thilker, D. A., et al. 2021, *ApJS* submitted, arXiv:2101.02855

Leonard, D. C., Kanbur, S. M., Ngeow, C. C., & Tanvir, N. R. 2003, *ApJ*, 594, 247

Leroy, A., Bolatto, A., Walter, F., & Blitz, L. 2006, *ApJ*, 643, 825

Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, 136, 2782

Leroy, A. K., Bolatto, A., Bot, C., et al. 2009, *ApJ*, 702, 352

Leroy, A. K., Bigiel, F., de Blok, W. J. G., et al. 2012, *ArXiv e-prints*, arXiv:1202.2873

Leroy, A. K., Lee, C., Schruba, A., et al. 2013a, *ApJL*, 769, L12

Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013b, *AJ*, 146, 19

Leroy, A. K., Hughes, A., Schruba, A., et al. 2016, *ApJ*, 831, 16

Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2017, *ApJ*, 846, 71

Leroy, A. K., Sandstrom, K. M., Lang, D., et al. 2019, *ApJS*, 244, 24

Leroy, A. K., Hughes, A., Liu, D., et al. 2021, *ApJS* submitted

Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, *ApJ*, 899, 58

Lin, L., Pan, H.-A., Ellison, S. L., et al. 2019, *ApJL*, 884, L33

Lisenfeld, U., Espada, D., Verdes-Montenegro, L., et al. 2011, *A&A*, 534, A102

Longmore, S. N., Bally, J., Testi, L., et al. 2013, *MNRAS*, 429, 987

Lopez, L. A., Krumholz, M. R., Bolatto, A. D., Prochaska, J. X., & Ramirez-Ruiz, E. 2011, *ApJ*, 731, 91

Lopez, L. A., Krumholz, M. R., Bolatto, A. D., et al. 2014, *ApJ*, 795, 121

Lynds, B. T. 1970, in *IAU Symposium*, Vol. 38, *The Spiral Structure of our Galaxy*, ed. W. Becker & G. I. Kontopoulos, 26

Mac Low, M.-M., & Klessen, R. S. 2004, *Reviews of Modern Physics*, 76, 125

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

Martin, C. L., & Kennicutt, Robert C., J. 2001, *ApJ*, 555, 301

McGaugh, S. S., & Schombert, J. M. 2014, *AJ*, 148, 77

McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, 45, 565

McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in *Astronomical Society of the Pacific Conference Series*, Vol. 376, *Astronomical Data Analysis Software and Systems XVI*, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127

Meidt, S. E., Schinnerer, E., Knapen, J. H., et al. 2012, *ApJ*, 744, 17

Meidt, S. E., Schinnerer, E., García-Burillo, S., et al. 2013, *ApJ*, 779, 45

Meidt, S. E., Schinnerer, E., van de Ven, G., et al. 2014, *ApJ*, 788, 144

Meidt, S. E., Hughes, A., Dobbs, C. L., et al. 2015, *ApJ*, 806, 72

Meidt, S. E., Leroy, A. K., Rosolowsky, E., et al. 2018, *ApJ*, 854, 100

Meidt, S. E., Glover, S. C. O., Krijssen, J. M. D., et al. 2020, *ApJ*, 892, 73

Milam, S. N., Savage, C., Brewster, M. A., Ziurys, L. M., & Wyckoff, S. 2005, *ApJ*, 634, 1126

Miville-Deschénes, M.-A., Murray, N., & Lee, E. J. 2017, *ApJ*, 834, 57

Mizuno, N., Rubio, M., Mizuno, A., et al. 2001, *PASJ*, 53, L45

Momose, R., Koda, J., Kennicutt, Jr., R. C., et al. 2013, *ApJL*, 772, L13

Muñoz-Mateos, J. C., Sheth, K., Regan, M., et al. 2015, *ApJS*, 219, 3

Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, *ApJ*, 737, 67

Murray, N. 2011, *ApJ*, 729, 133

Nieten, C., Neininger, N., Guélin, M., et al. 2006, *A&A*, 453, 459

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, 660, L43

Nugent, P., Sullivan, M., Ellis, R., et al. 2006, *ApJ*, 645, 841

Ochsendorf, B. B., Meixner, M., Roman-Duval, J., Rahman, M., & Evans, Neal J., I. 2017, *ApJ*, 841, 109

Oliphant, T. E. 2006, *A guide to NumPy*, Vol. 1 (Trelgol Publishing USA)

Onodera, S., Kuno, N., Tosaki, T., et al. 2010, *ApJL*, 722, L127

—. 2012, *PASJ*, 64, 133

Orr, M. E., Hayward, C. C., Hopkins, P. F., et al. 2018, *MNRAS*, 478, 3653

Ostriker, E. C., McKee, C. F., & Leroy, A. K. 2010, *ApJ*, 721, 975

Padoan, P., Haugbølle, T., & Nordlund, Å. 2012, *ApJL*, 759, L27

Padoan, P., Haugbølle, T., Nordlund, Å., & Frimann, S. 2017, *ApJ*, 840, 48

Padoan, P., & Nordlund, Å. 2002, *ApJ*, 576, 870

Padoan, P., Pan, L., Haugbølle, T., & Nordlund, Å. 2016, *ApJ*, 822, 11

Paturel, G., Petit, C., Prugniel, P., et al. 2003, *A&A*, 412, 45

Pety, J. 2005, in *SF2A-2005: Semaine de l’Astrophysique Francaise*, ed. F. Casoli, T. Contini, J. M. Hameury, & L. Pagani, 721–722

Pety, J., Schinnerer, E., Leroy, A. K., et al. 2013, *ApJ*, 779, 43

Pierce, M. J., Welch, D. L., McClure, R. D., et al. 1994, *Nature*, 371, 385

Querejeta, M., Meidt, S. E., Schinnerer, E., et al. 2015, *ApJS*, 219, 5

R Core Team. 2015, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria

Rahman, N., Bolatto, A. D., Wong, T., et al. 2011, *ApJ*, 730, 72

Rahman, N., Bolatto, A. D., Xue, R., et al. 2012, *ApJ*, 745, 183

Rahner, D., Pellegrini, E. W., Glover, S. C. O., & Klessen, R. S. 2017, *MNRAS*, 470, 4453

—. 2019, *MNRAS*, 483, 2547

Raskutti, S., Ostriker, E. C., & Skinner, M. A. 2016, *ApJ*, 829, 130

—. 2017, *ApJ*, 850, 112

Regan, M. W., Thornley, M. D., Helfer, T. T., et al. 2001, *ApJ*, 561, 218

Reissl, S., Klessen, R. S., Mac Low, M.-M., & Pellegrini, E. W. 2018, *A&A*, 611, A70

Rice, T. S., Goodman, A. A., Bergin, E. A., Beaumont, C., & Dame, T. M. 2016, *ApJ*, 822, 52

Roman-Duval, J., Heyer, M., Brunt, C. M., et al. 2016, *ApJ*, 818, 144

Roman-Duval, J., Israel, F. P., Bolatto, A., et al. 2010, *A&A*, 518, L74+

Rosolowsky, E. 2007, *ApJ*, 654, 240

Rosolowsky, E., & Blitz, L. 2005, *ApJ*, 623, 826

Rosolowsky, E., Engargiola, G., Plambeck, R., & Blitz, L. 2003, *ApJ*, 599, 258

Rosolowsky, E., Keto, E., Matsushita, S., & Willner, S. P. 2007, *ApJ*, 661, 830

Rosolowsky, E., & Leroy, A. 2006, *PASP*, 118, 590

Rosolowsky, E., Hughes, A., Leroy, A. K., et al. 2021, *MNRAS accepted*, arXiv:2101.04697

Rubio, M., Elmegreen, B. G., Hunter, D. A., et al. 2015, *Nature*, 525, 218

Ruiz-Lapuente, P. 1996, *ApJL*, 465, L83

Saintonge, A., Kauffmann, G., Wang, J., et al. 2011, *MNRAS*, 415, 61

Saintonge, A., Catinella, B., Cortese, L., et al. 2016, *MNRAS*, 462, 1749

Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, *ApJS*, 233, 22

Saintonge, A., Wilson, C. D., Xiao, T., et al. 2018, *MNRAS*, 481, 3497

Sakamoto, K., Okumura, S. K., Ishizuki, S., & Scoville, N. Z. 1999, *ApJS*, 124, 403

Salim, S., Boquien, M., & Lee, J. C. 2018, *ApJ*, 859, 11

Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, 173, 267

Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, *ApJS*, 227, 2

Salo, H., Laurikainen, E., Laine, J., et al. 2015, *ApJS*, 219, 4

Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014, *A&A*, 563, A49

Sanders, D. B., Mazzarella, J. M., Kim, D., Surace, J. A., & Soifer, B. T. 2003, *AJ*, 126, 1607

Sanders, D. B., & Mirabel, I. F. 1996, *ARA&A*, 34, 749

Schinnerer, E., Meidt, S. E., Pety, J., et al. 2013, *ApJ*, 779, 42

Schinnerer, E., Meidt, S. E., Colombo, D., et al. 2017, *ApJ*, 836, 62

Schinnerer, E., Hughes, A., Leroy, A., et al. 2019, *ApJ*, 887, 49

Schmidt, T. M., Bigiel, F., Klessen, R. S., & de Blok, W. J. G. 2016, *MNRAS*, 457, 2642

Schruba, A., Bialy, S., & Sternberg, A. 2018, *ApJ*, 862, 110

Schruba, A., Kruijssen, J. M. D., & Leroy, A. K. 2019, *ApJ*, 883, 2

Schruba, A., Leroy, A. K., Bolatto, A. D., et al. 2021, *ApJ* to be subm

Schruba, A., Leroy, A. K., Walter, F., Sandstrom, K., & Rosolowsky, E. 2010, *ApJ*, 722, 1699

Schruba, A., Leroy, A. K., Walter, F., et al. 2011, *AJ*, 142, 37

—. 2012, *AJ*, 143, 138

Schruba, A., Leroy, A. K., Kruijssen, J. M. D., et al. 2017, *ApJ*, 835, 278

Schuster, K. F., Kramer, C., Hitschfeld, M., García-Burillo, S., & Mookerjea, B. 2007, *A&A*, 461, 143

Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2018, *ApJ*, 861, 4

Shaya, E. J., Tully, R. B., Hoffman, Y., & Pomarède, D. 2017, *ApJ*, 850, 207

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, *MNRAS*, 343, 978

Sheth, K., Vogel, S. N., Regan, M. W., Thornley, M. D., & Teuben, P. J. 2005, *ApJ*, 632, 217

Sheth, K., Regan, M., Hinz, J. L., et al. 2010, *PASP*, 122, 1397

Silk, J. 1997, *ApJ*, 481, 703

Simonian, G. V., & Martini, P. 2017, *MNRAS*, 464, 3920

Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, *ApJ*, 319, 730

Somerville, R. S., & Davé, R. 2015, *ARA&A*, 53, 51

Sorai, K., Kuno, N., Muraoka, K., et al. 2019, *PASJ*, 125

Sun, J., Leroy, A. K., Schruba, A., et al. 2018, *ApJ*, 860, 172

Sun, J., Leroy, A. K., Ostriker, E. C., et al. 2020a, *ApJ*, 892, 148

Sun, J., Leroy, A. K., Schinnerer, E., et al. 2020b, *ApJL*, 901, L8

Tan, J. C. 2000, *ApJ*, 536, 173

Telford, O. G., Dalcanton, J. J., Williams, B. F., et al. 2020, *ApJ*, 891, 32

Thompson, T. A., & Krumholz, M. R. 2016, *MNRAS*, 455, 334

Thompson, T. A., Quataert, E., & Murray, N. 2005, *ApJ*, 630, 167

Tonry, J. L., Dressler, A., Blakeslee, J. P., et al. 2001, *ApJ*, 546, 681

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898

Tress, R. G., Smith, R. J., Sormani, M. C., et al. 2020a, *MNRAS*, 492, 2973

Tress, R. G., Sormani, M. C., Smith, R. J., et al. 2020b, *arXiv e-prints*, arXiv:2012.05919

Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, *AJ*, 152, 50

Tully, R. B., Rizzi, L., Shaya, E. J., et al. 2009, *AJ*, 138, 323

Utomo, D., Bolatto, A. D., Wong, T., et al. 2017, *ApJ*, 849, 26

Utomo, D., Sun, J., Leroy, A. K., et al. 2018, *ApJL*, 861, L18

Vazdekis, A., Koleva, M., Ricciardelli, E., Röck, B., & Falcón-Barroso, J. 2016, *MNRAS*, 463, 3409

Vazquez-Semadeni, E. 1994, *ApJ*, 423, 681

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261

Vutisalchavakul, N., Evans, II, N. J., & Heyer, M. 2016, *ApJ*, 831, 73

Walch, S., Girichidis, P., Naab, T., et al. 2015, *MNRAS*, 454, 238

Watkins, A. E., Mihos, J. C., & Harding, P. 2016, *ApJ*, 826, 59

Weigel, A. K., Schawinski, K., & Bruderer, C. 2016, *MNRAS*, 459, 2150

Wilson, C. D., Warren, B. E., Israel, F. P., et al. 2012, *MNRAS*, 424, 3050

Wilson, T. L., & Rood, R. 1994a, *ARA&A*, 32, 191

—. 1994b, *ARA&A*, 32, 191

Wong, T., & Blitz, L. 2002, *ApJ*, 569, 157

Wong, T., Hughes, A., Ott, J., et al. 2011, *ApJS*, 197, 16

Wong, T., Xue, R., Bolatto, A. D., et al. 2013, *ApJL*, 777, L4

Wong, T., Hughes, A., Tokuda, K., et al. 2019, *ApJ*, 885, 50

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868

Yim, K., Wong, T., Rand, R. J., & Schinnerer, E. 2020, *MNRAS*, 494, 4558

Young, J. S., Allen, L., Kenney, J. D. P., Lesser, A., & Rownd, B. 1996, *AJ*, 112, 1903

Young, J. S., & Scoville, N. Z. 1991, *ARA&A*, 29, 581

Young, J. S., Xie, S., Tacconi, L., et al. 1995, *ApJS*, 98, 219

Young, L. M., Bureau, M., Davis, T. A., et al. 2011, *MNRAS*, 414, 940

Zucker, C., Battersby, C., & Goodman, A. 2018, *ApJ*, 864, 153

Zuckerman, B., & Evans, N. J., I. 1974, *ApJL*, 192, L149

## APPENDIX

## A. CONTRIBUTIONS

The design, execution, processing, and scientific exploitation of PHANGS-ALMA was a team effort, with major contributions from many people and input from the entire team. This paper also reflects major direct and indirect contributions from many people. We summarize some of the key contributions here.

**Observation Design, Data Processing, and Quality Assurance of the ALMA Data:** Since 2016, observation design, quality assurance, pipeline and algorithm development, and data processing have all been organized through the “PHANGS ALMA Data Reduction” (ADR) working group. J. Pety has led this group since the beginning of the PHANGS collaboration, and the active members participating in almost all key activities and doing the core work on observations and data processing have been C. Faesi, C. Herrera, A. Hughes, D. Liu, A. Leroy, T. Saito, E. Rosolowsky, E. Schinnerer, A. Schruba, and A. Usero. The imaging and postprocessing pipeline for the interferometric data and the software to create derived products has been mostly developed by A. Leroy, D. Liu, E. Rosolowsky, and T. Saito. For the total power data, the processing and software development have been led by C. Faesi, C. Herrera, J. Pety, and A. Usero. A. Hughes led quality assurance efforts and software development. A. Hygate and K. Sliwa also made early key contributions to a wide range of ADR efforts. A broad cross-section of the PHANGS-ALMA team also contributed to quality assurance efforts and gave excellent, frequent feedback. These included: A. Barnes, I. Beslic, M. Chevance, J. den Brok, C. Eibensteiner, C. Faesi, A. García-Rodríguez, C. Herrera, A. Hygate, M. Jimenez Donaire, J. Kim, A. Leroy, D. Liu, J. Pety, J. Puschnig, M. Querejeta, E. Rosolowsky, T. Saito, A. Sardone, E. Schinnerer, A. Schruba, J. Sun, A. Usero, D. Utomo, T. Williams. Please also note that [Leroy et al. \(2021\)](#) includes a more detailed description of contributions to the PHANGS-ALMA data processing and product generation pipeline.

**Management of the PHANGS Collaboration:** E. Schinnerer has served as the leader of the PHANGS collaboration since 2015. G. Blanc, E. Emsellem, A. Leroy, and E. Rosolowsky have acted as the PHANGS steering committee. E. Rosolowsky has served as team manager since 2018. The entire PHANGS core team provides key input and oversight to all major collaboration decisions. The core team includes: F. Bigiel, G. Blanc, E. Emsellem, A. Escala, B. Groves, A. Hughes, K. Kreckel, J.M.D. Kruijssen, J. Lee, A. Leroy, S. Meidt, M. Querejeta, J. Pety, E. Rosolowsky, P. Sanchez-Blazquez, K. Sandstrom, E. Schinnerer, A. Schruba, and A. Usero. Scientific exploitation of PHANGS-ALMA has taken place largely in the context of the “Cold ISM Structure and Its Relation to Star Formation” (Cold ISM) and “Large Scale Dynamical Processes” (Dynamics) science working groups. The “Cold ISM” group was led by A. Hughes and K. Kreckel in 2019 and C. Faesi and A. Hughes since 2020. The “Dynamics” group has been led by S. Meidt and M. Querejeta since 2019.

**Calculation of Galaxy Properties:** G. Anand, J. Lee, and the PHANGS-HST team led compilation of distances for the sample. A. Leroy, S. Meidt, M. Querejeta, K. Sandstrom, J. Sun, and T. Williams played active roles in the discussion and determination of stellar masses and sizes for the sample. I. Ho, A. Leroy, and T. Williams have curated or developed software for curating the master table of PHANGS target properties. E. Behrens and A. Leroy created by-hand masks for artifacts and stars in the multiwavelength data used to estimate the SFR and stellar masses. M. Querejeta played a key role in many aspects of obtaining and processing the IRAC data used for stellar mass estimation. Benchmarks against the MUSE-based galaxy property estimates relied on the hard work of the PHANGS-MUSE team, led by E. Emsellem with major contributions to SFR and  $M_*$  estimates by F. Belfiore, I. Pessa, and P. Sanchez Blazquez. The PHANGS sample definition and the SFR and  $M_*$  estimation built on work on this topic by A. Leroy, K. Sandstrom, J. Chastenet, and I. Chiang, heavily leveraging previous work by S. Salim, M. Boquien, and J. Lee.

**Preparation of this Paper:** A. Leroy led preparation of the text and figures in close collaboration with E. Schinnerer. Frequent heavy edits and early, repeated high level scientific input throughout were contributed by A. Hughes, E. Rosolowsky, K. Sandstrom, and A. Schruba. The survey motivation in Section 2 distills many years of collective scientific discussions and proposals by the team. S. Meidt and M. Querejeta provided major input on the section presenting the estimation of galaxy properties. The entire team provided multiple rounds of careful vetting. C. Faesi and A. Usero compiled the detailed log of observations for the total power data.

**Observatory and Community Support:** Across multiple large and small projects, PHANGS-ALMA has benefited from outstanding support from the Joint ALMA Observatory, the North American ALMA Science Center (NAASC) at the National Radio Astronomy Observatory (NRAO), and the European Southern Observatory (ESO). The Large Program was carried out with the NAASC as supporting ALMA Regional Center, and the NAASC staff, including ARC manager A. Remijan, have been incredibly responsive, helpful, and supportive with a wide range of technical

issues. The ESO ALMA ARC has been similarly supportive whenever issues arose. PHANGS-ALMA built heavily on the high quality ALMA calibration pipeline (E. Humphreys et al. in preparation) and the `CASA` software package. We also heavily leveraged the work by the `astropy` collaboration and the broader scientific `python` community. We also acknowledge the astronomical IDL community, which was instrumental in many early parts of this work.