A More Efficient Search for H2O Megamaser Galaxies: The Power of X-Ray and Mid-infrared Photometry

C. Y. Kuo1, J. Y. Hsiang2, H. H. Chung2, A. Constantin3, Y.-Y. Chang4,5, E. da Cunha6,7,8, D. Pesce9,10, W. T. Chien11, B. Y. Chen2, J. A. Braatz12, Ingyin Zaw13, S. Matsushita1, and J. C. Lin1
1 Physics Department, National Sun Yat-Sen University, No. 70, Lien-Hai Road, Kaosiung City 80424, Taiwan, R.O.C
2 Institute of Astronomy, National Tsing Hua University, General Building II, NTHU, No. 101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan, R.O.C
3 Department of Physics and Astronomy, James Madison University, Harrisonburg, VA 22807, USA
4 Academia Sinica Institute of Astronomy and Astrophysics, P.O. Box 23-141, Taipei 10617, Taiwan, R.O.C.
5 Department of Physics, National Chung Hsing University, 40227, Taichung, Taiwan
6 The Australian National University, Mt Stromlo Observatory, Cotter Road, Weston Creek, ACT 2611, Australia
7 International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
8 ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
9 Center for Astrophysics/Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
10 Black hole Initiative at Harvard University, 20 Garden Street, Cambridge, MA 02138, USA
11 Department of Physics, National Taiwan University, No.1 Sec. 4 Roosevelt Road Taipei 10617, Taiwan
12 National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
13 New York University Abu Dhabi, Abu Dhabi, UAE

Received 2019 November 25; revised 2020 January 21; accepted 2020 February 14; published 2020 March 23

Abstract

We present an investigation of the dependence of H2O maser detection rates and properties on the mid-IR active galactic nucleus (AGN) luminosity, L_{AGN}, and the obscuring column density, N_H, based on mid-IR and hard X-ray photometry. Based on spectral energy distribution fitting that allows for decomposition of the black hole accretion and star formation components in the mid-infrared, we show that the megamaser (disk maser) detection rate increases sharply for galaxies with 12 μm AGN luminosity L_{AGN}^{12 μm} greater than 10^{42} erg s^{-1}, from <3% (<2%) to ~12% (~5%). By using the ratio of the observed X-ray to mid-IR AGN luminosity as an indicator of N_H, we also find that megamaser (disk maser) detection rates are boosted to 15% (7%) and 20% (9%) for galaxies with N_H > 10^{23} cm^{-2} and N_H > 10^{24} cm^{-2}, respectively. Combining these column density cuts with a constraint for high L_{AGN}^{12 μm} (>10^{42} erg s^{-1}) predicts further increases in the megamaser (disk maser) detection rates to 19% (8%) and 27% (14%), revealing unprecedented potential increases of the megamaser and disk maser detection rates by a factor of 7–15 relative to the current rates, depending on the chosen sample selection criteria. A noteworthy aspect of these new predictions is that the completeness rates are only compromised mildly, with the rates remaining at the level of ~95% (~50%) for sources with N_H > 10^{23} cm^{-2} (N_H > 10^{24} cm^{-2}). Applying these selection methods to current X-ray AGN surveys predicts the detection of >15 new megamaser disks.

Unified Astronomy Thesaurus concepts: Megamasers (1023); Astrophysical masers (103); Active galactic nuclei (16); Active galaxies (17); X-ray active galactic nuclei (2035); X-ray surveys (1824); Surveys (1671); Broad band photometry (184); Infrared photometry (792)

Supporting material: machine-readable tables

1. Introduction

H2O megamaser emission at ν ~ 22 GHz (λ ~ 1.3 cm) originating from galactic nuclei at ~0.1–1 pc from the central supermassive black holes (SMBHs) provides, to date, the only known tracer of subparsec structures resolved in both position and velocity that appear to be associated with black hole accretion processes and, consequently, the active galactic nuclei (AGNs) phenomena (e.g., Lo 2005). As work on the prototypical maser galaxy NGC 4258 has demonstrated (Herrnstein et al. 1999), the masing gas in such a system often resides in a subparsec-scale thin disk, with the gas kinematics following nearly perfect Keplerian rotation. As a result, one can relatively easily and accurately measure the mass of an SMBH in a H2O maser disk (e.g., Kuo et al. 2011; Gao et al. 2017; Zhao et al. 2018). Accurate SMBH mass measurements play a crucial role in understanding galaxy formation and evolution processes via the M_{BH}–σ_g relation (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Gütekın et al. 2009, Greene et al. 2016 and references therein).

Moreover, modeling a maser disk in three dimensions (e.g., Herrnstein et al. 1999; Reid et al. 2013) provides an accurate determination of the physical radius of the systemic maser component (i.e., the maser spectral component with velocities closest to the systemic velocity of the host galaxy), which can be used as a standard ruler for measuring accurate angular diameter distances to galaxies beyond the Local Group. In turn, these distance measurements enable a precise determination of the Hubble constant H_0 without inferences about the geometry of the universe (Kuo et al. 2013, 2015; Reid et al. 2013; Gao et al. 2016), thus advancing our understanding of the nature of dark energy when these measurements are used in conjunction with cosmic background radiation constraints (Hu 2005; Olling 2007).

Unfortunately, the chances of finding these golden standards are abysmally low. Of >6000 galaxy nuclei surveyed so far for 22 GHz H2O maser emission, only 180 are detected, with ~30% of those possibly originating in disks. The typical detection rate achieved in the most extensive 22 GHz H2O maser survey to date, the Megamaser Cosmology Project...
rates of H$_2$O megamaser emission by employing in concert the novel investigation of the likelihood of increasing the detection et al. 2018 for a thorough review of these studies targeting of maser galaxies, although without clear predictions scatter, and with new insights toward designing more efficient targeting of maser galaxies, although without clear predictions for a significant increase in the maser detection rate (see Kuo et al. 2018 for a thorough review of these studies).

In Kuo et al. (2018), we followed up on these works with a novel investigation of the likelihood of increasing the detection rates of H$_2$O megamaser emission by employing in concert the optical and mid-infrared photometric properties of the galaxies searched by the MCP with the Green Bank Telescope (hereafter the GBT galaxy sample). We found that galaxies with water megamaser emission tend to be associated with strong emission in all of the mid-infrared bands employed by the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), as well as with previously proposed and newly found indicators of AGN strength in the mid-infrared, such as the red W1 – W2 and W1 – W4 colors of the host galaxy. The correlation with significantly red mid-IR colors also suggests that maser galaxies tend to reside in heavily obscured AGNs, consistent with previous findings (e.g., Zhang et al. 2006, 2010; Greenhill et al. 2008; Castangia et al. 2013; Wagner 2013; Masini et al. 2016). Importantly, these trends predict a potential increase in megamaser detection rates to 6%–15%, depending on the specific sample selection criteria.

While the detection rates predicted by these new survey criteria could be boosted by a factor of 2–8 relative to the current rates, choosing the criteria (e.g., W1 – W2 > 0.5 and W1 – W4 > 7) which give rise to the highest detection rates (e.g., ~18%) often significantly compromises the completeness rate of H$_2$O maser detection (e.g., 32%). Here, the completeness rate refers to the ratio $n_{\text{maser}}/N_{\text{maser}}$, where $N_{\text{maser}}$ is the total number of maser detections in the entire GBT galaxy sample and $n_{\text{maser}}$ indicates the number of detected maser sources showing mid-IR colors that are redder than certain thresholds. A low completeness rate suggests that a only a correspondingly small fraction of the megamasers in the GBT galaxy sample is recovered when applying certain mid-IR color cuts for galaxy selection.

Notably, four of the best-studied disk megamaser systems (i.e., UGC 3789, NGC 5765b, NGC 6264, and NGC 6323)—all of which have well-ordered Keplerian disks suitable for an accurate H$_0$ determination (i.e., Kuo et al. 2013, 2015; Reid et al. 2013; Gao et al. 2016)—exhibit blue mid-IR colors that are essentially dominated by emission from their hosts’ starlight. While the nuclear emission of these systems shows evidence for Compton-thick absorption (i.e., the absorbing column density $N_H \gtrsim 10^{24}$ cm$^{-2}$; Greenhill et al. 2008; Castangia et al. 2013; Masini et al. 2016), suggesting that their nuclear mid-IR colors may be intrinsically red, their AGN might not be luminous enough to dominate over the mid-IR emission from their hosts, making them “WISE blue.”

Furthermore, galaxies with red mid-IR colors that correspond to the high maser detection rates we have found in Kuo et al. (2018) are only a minority (~0.3%) of the entire galaxy population, making it challenging to collect a sizable galaxy sample with these properties to start with. As a consequence, the resulting number of possible H$_2$O megamaser detections from future maser surveys using such selection criteria would be small: among the ~50,000 galaxies with properties cataloged in the 2df (Colless et al. 2001), 6df (Jones et al. 2009), 2MRS (Huchra et al. 2012), RC3 (de Vaucouleurs et al. 1991), and Galaxy Zoo (Lintott et al. 2008) galaxy samples that have not yet been surveyed by the MCP, only 171 galaxies (0.3%) have $W1 – W2 > 0.5$ and $W1 – W4 > 7$, which would in turn yield fewer than 10 expected new disk maser detections.

Given these shortcomings, it follows that aiming at enhancing the detection rate alone is insufficient if one seeks to discover a large number of H$_2$O disk megamasers; to discover the bulk of the megamaser galaxy population, one must also enhance the completeness rate. In the context of using mid-IR emission as a tool to identify megamaser candidates, the key to reaching this goal is to minimize the host galaxy contamination in the mid-IR emission in order to identify intrinsically mid-IR red (dusty) and luminous AGNs.

In this paper, we explore systematic approaches that aim to boost both the maser detection rate and the completeness rate simultaneously by employing the $L_{\text{AGN}}^{12 \mu m}$–$L_{\text{obs}}^{210}$ diagram, where $L_{\text{AGN}}^{12 \mu m}$ and $L_{\text{obs}}^{210}$ refer to the 12 $\mu m$ AGN luminosity (separated from that of the host via spectral energy distribution (SED) decomposition), and the observed 2–10 keV X-ray luminosity, respectively. We show that the use of this particular method provides an effective way to boost the maser detection rates because it allows one to select intrinsically mid-IR luminous AGNs for maser search. In addition, as the ratio of $L_{\text{obs}}^{210}$ and $L_{\text{AGN}}^{12 \mu m}$ has been shown to be an indicator of the absorbing column density of AGNs (e.g., Satyapal et al. 2017), the $L_{\text{AGN}}^{12 \mu m}$–$L_{\text{obs}}^{210}$ diagram provides a potential tool for selecting X-ray-obscured galaxies, including Compton-thick AGN candidates which can be difficult to recognize in X-ray observations probing $\leq$10 keV energy bands (e.g., Bassani et al. 1999; Cappi et al. 2006; Panessa et al. 2006; Goulding et al. 2011; LaMassa et al. 2011; Koulouridis et al. 2016). This helps to increase the completeness rate of maser surveys based on mid-IR galaxy selection because maser galaxies tend to reside in heavily obscured AGNs, which may not be “WISE red.”

This paper is organized as follows: Section 2 presents the galaxy sample and explains the methods used in compiling the X-ray and mid-IR AGN luminosities. Section 3 explores the dependence of the maser detection rates on the 12 $\mu m$ AGN luminosity and the X-ray-obscuring column density, based on the $L_{\text{AGN}}^{12 \mu m}$–$L_{\text{obs}}^{210}$ diagram. The discussions and conclusions of this study are presented and discussed in Sections 4 and 5, respectively.

Throughout this paper we adopt a $\Lambda$CDM cosmology, where $\Lambda$ denotes the cosmological constant that accounts for dark energy, and a universe that contains cold dark matter (CDM). We adopt $\Omega_{0.3}$, $\Omega_{0.6}$, $\Omega_{0.6}$, and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$ for the cosmological parameters.

2. Galaxy Sample and Photometric Data Collection

2.1. Galaxy Sample and the X-Ray Data

The galaxies we study in this paper are drawn from the GBT sample studied by Kuo et al. (2018), which represents the
largest and most comprehensive catalog\textsuperscript{14} of galaxies surveyed for water maser emission at 22 GHz. This sample contains 4836 galaxies surveyed by the GBT prior to 2016 September. These galaxies are primarily narrow-emission-line AGNs selected from galaxy catalogs including the Sloan Digital Sky Survey (SDSS; York et al. 2000; Abazajian et al. 2009), Two Micron All Sky Survey (2MASS) Redshift Survey (Huchra et al. 2012), 2dF Survey (Colless et al. 2001), and 6dF Survey (Jones et al. 2009) via their optical emission-line ratios using the BPT diagram method (Baldwin et al. 1981), with no additional criteria such as galaxy colors or magnitudes.

The redshifts of the GBT sample galaxies range from 0.0 to 0.07, and the overall maser detection rate is \(\sim 3\%\). Hereafter, we refer to galaxies with confirmed H\textsubscript{2}O maser emission as maser galaxies or masers.\textsuperscript{15} and we refer to galaxies with no maser detection as nonmaser galaxies or nonmasers. Within the category of masers, we further define three subsamples called kilomasers, megamasers, and disk masers. Kilomasers and megamasers refer to masers having an isotropic H\textsubscript{2}O luminosity \(L_{\text{H}_2O} < 10^9\) L\(_{\odot}\) and \(L_{\text{H}_2O} > 10^9\) L\(_{\odot}\), respectively. Disk masers are megamaser systems whose maser spectra display characteristic spectral structures that suggest the presence of subparsec-scale maser disks; thus, megamasers indicate a maser population that includes both disk masers and nondisk megamasers (see Section 2 in Kuo et al. 2018 for a more detailed description of these types of masers). The isotropic H\textsubscript{2}O luminosities for all of the maser galaxies can be found in Table 1 of Kuo et al. (2018).

To obtain literature values for the observed 2–10 keV X-ray luminosities \(L_{\text{X}}\) for galaxies in the GBT sample, we prioritize those X-ray surveys which provide X-ray properties based on X-ray spectral fitting. X-ray spectral analysis is important for our work because our investigation primarily relies on the intrinsic X-ray luminosities of AGNs whose nonthermal continuum emission can be described by power-law distributions in their X-ray spectra. Spectral decomposition is necessary to separate the power-law component from the thermal emission (e.g., soft excess). When the signal to noise of an X-ray spectrum is high enough (e.g., spectra from the Swift/BAT survey described below), we adopt data in which any existing Fe K-alpha line components are removed.

After assembling a set of X-ray surveys which provide 2–10 keV X-ray luminosities based on X-ray spectral analysis, we cross-matched the GBT sample with the four largest X-ray source catalogs in our collection. These X-ray catalogs include (1) the Swift/BAT 70 month AGN catalog (Ricci et al. 2017), (2) the XMM-Newton spectral-fit database (Corral et al. 2015), (3) the Chandra Survey of Nearby Galaxies (She et al. 2017), and (4) the INTEGRAL/IBIS AGN catalog (Malizia et al. 2012, 2016). We also searched the 40 month catalog of the NuSTAR Serendipitous Survey (Lansbury et al. 2017) for counterparts of the GBT galaxy sample, but we did not find any matches.

In addition to the above large X-ray source catalogs, we also cross-matched the GBT sample with smaller X-ray catalogs reported in the NASA/IPAC Extragalactic Database (NED). Upon extensively searching for all references of X-ray photometric data for each galaxy in the GBT sample, we identified a set of measurements that we collectively refer to here as the NED catalog. The number of GBT galaxies found in each X-ray catalog, including a split per maser type, are listed in Table 1.

In cross-matching sources between the X-ray catalogs, we use a matching radius of 10″, from which we identify 691 X-ray counterparts of the GBT sample. To ensure reliable cross-matching and compatibility with additional optical and near-IR photometric data used in SED fitting, we further require the matched X-ray counterparts to be within the effective (i.e., half-light) radii \(r_{\text{eff}}\) of the GBT galaxies, with \(r_{\text{eff}}\) obtained from the SDSS or the 2MASS galaxy surveys (see Section 2.2.2 for details). This multiwavelength, multicatalog search leads to successful cross-matching of 642 galaxies (hereafter the X-ray sample) that have all the necessary broadband rest-frame UV-to-mid-IR photometry for measuring an intrinsic mid-IR AGN luminosity based on SED decomposition (see Section 2.2). The majority (93%) of the X-ray sample sources are included in the Swift/BAT catalog (44%) and the XMM-Newton spectral-fit database (49%). The total numbers of kilomasers, megamasers, and disk masers in the X-ray sample are 15 (2.3%), 53 (8.3%), and 24 (3.7%), respectively.

Table 2 lists the coordinates, redshifts, distances, and the mid-IR (at 12 \(\mu\)m) and optical (in [O III] emission) luminosities of the X-ray sample sources. The X-ray properties of these sources, which include the 2–10 keV luminosities, neutral hydrogen absorbing column densities \(N_{\text{HI}}\), and the photon index values \(\Gamma\) (from fitting power laws to the X-ray spectra) are shown in Table 3. The distance distribution of these X-ray sources and the maser detection rates as a function of distance are shown in the left and right panels of Figure 1, respectively. In Section 4.2, we explore why the detection rate of megamasers in the X-ray sample has a different distance distribution with respect to the GBT sample. The following subsections present in detail the considerations for retrieving the necessary data from the above-listed X-ray catalogs.

\textsuperscript{14} The catalog can be accessed via the weblink https://safe.nrao.edu/wiki/bin/view/Main/MegamaserProjectSurvey.
\textsuperscript{15} Of all the 180 megamaser detections mentioned in the introduction, two are found associated with the high-redshift quasars MG J0414+0534 (\(z = 2.64\); Impellizzeri et al. 2008) and SDSS J080430.99+360718.1 (\(z = 0.66\); Barvainis & Antonucci 2005). These two maser sources are not discovered as part of the GBT maser survey, and we do not include them in the analysis presented in this paper.
probes X-ray sources in the 14–195 keV energy range (Ricci et al. 2017). The spectral analysis used in this catalog combines X-ray data from XMM-Newton, Swift/XRT, ASCA, Chandra, and Suzaku observations in the X-ray band covering 0.3–10 keV. Because of the broadband nature (0.3–150 keV) of the analysis, the Swift/BAT AGN catalog provides the most accurate measurements of the X-ray properties for GBT sample sources included in this catalog. In particular, as Swift/BAT can detect X-ray photons with energies beyond 10 keV (which can traverse heavy obscuration), measurements of $N_H$ from this catalog are far less susceptible to systematic biases caused by large absorbing columns in Compton-thick AGNs than observations that only probe the $< 10$ keV energy bands (e.g., Chandra and XMM-Newton).

The X-ray spectral analysis in the Swift/BAT catalog was carried out for two categories of sources (i.e., nonblazar AGN and blazars) using a series of models of successive complexity. For nonblazar AGNs, which are further subdivided into obscured and unobscured sources, 17 different spectral decomposition models were fit to the 0.3–150 keV broadband data. For blazars, seven spectral models were adopted to fit the data. For all of the Swift/BAT sources, the broadband data enable the X-ray spectral analysis to remove the Fe Kα component and various thermal components in the observed X-ray luminosities $L_{2–10}^{\text{obs}}$, ensuring that the AGN luminosities listed in Table 3 only include the power-law component of the AGN X-ray continuum emission.

Because of the reliability of its X-ray parameter estimation, we give higher priority to the Swift/BAT data when a GBT sample galaxy appears in multiple X-ray catalogs including the Swift/BAT. Among the 4836 galaxies in the GBT sample, 307 have reliable matches to the Swift/BAT 70 month catalog. For a few sources in our sample which have X-ray counterparts in the Swift/BAT survey, instead of using the $L_{2–10}^{\text{obs}}$ provided by the Swift/BAT catalog, we adopt the estimated 2–10 keV luminosities $L_{2–10}^{\text{est}}$ obtained by transforming the observed 14–195 keV luminosities from the Swift/BAT survey ($L_{2–10}^{\text{obs}}$) into 2–10 keV luminosities using a correction factor of 0.37 (i.e., $L_{2–10}^{\text{est}} \equiv 0.37 \times L_{2–10}^{\text{obs}}$; see Koss et al. 2017). As indicated in Figure 15 in Ricci et al. (2017), $L_{2–10}^{\text{est}}$ shows a strong correlation with $L_{2–10}^{\text{obs}}$ for objects with log $N_H \lesssim 23.7$, with a scatter of ~0.3 dex which is likely caused by differences in the shapes of the X-ray continuum spectra and by intrinsic flux variability of the X-ray emission (Ricci et al. 2017). However, a few sources have absorption-corrected 2–10 keV fluxes that are considerably higher than those expected from the observed 14–150 keV flux. According to Ricci et al. (2017), their $L_{2–10}^{\text{obs}}$ might be overestimated. Given the possibility of

16. The conversion factor 0.37 corresponds to an X-ray spectrum photon index $\Gamma$ of 1.8.

17. The scatter in log $L_{2–10}^{\text{est}}$ will be 0.32 if the X-ray continuum spectra of a sample of X-ray AGNs are fully described by power-law distributions with the photon indices $\Gamma$ following a Gaussian distribution with a mean of 1.76 and a standard deviation of 0.29 (i.e., the mean and the standard deviation of $\Gamma$ for Swift/BAT sources in the X-ray sample).
significant overestimation, we adopt $L_{\text{2-10 keV}}^{\text{estimate}}$ as the observed 2–10 keV luminosities for sources in our sample with $L_{\text{2-10}}^{\text{obs}} > 5 \times L_{\text{2-10}}^{\text{estimate}}$.

2.1.2. The XMM-Newton Spectral-fit Database

The XMM-Newton spectral-fit database (XMMFITCAT; Corral et al. 2015) provides X-ray spectral-fitting results for sources detected in the XMM-Newton serendipitous survey (Rosen et al. 2016), which probes the 0.2–12 keV energy band. In its most recent release (3MM-DR6), XMMFITCAT includes fitting results for 146,825 detections. Among all unique XMM-Newton detections, we found 505 matches between XMMFITCAT and the GBT sample, with 168 sources also belonging to the Swift/BAT AGN catalog. Therefore, XMMFITCAT provides X-ray information for 337 independent X-ray sources for our study.

Unlike the Swift/BAT catalog, XMMFITCAT does not directly provide a measurement of $L_{\text{2-10}}^{\text{obs}}$. Instead, it provides best-fit parameters for six spectral models (Corral et al. 2015). To obtain $L_{\text{2-10}}^{\text{obs}}$, we adopt the model parameters from the “preferred model” (Corral et al. 2015) recommended by XMMFITCAT, which is selected according to the goodness of each fit, which we then use to estimate $L_{\text{2-10}}^{\text{obs}}$ from the power-law component of the X-ray emission following the equation

$$L_{\text{2-10}}^{\text{obs}} = 4\pi D_L^2 \int_{2 \text{ keV}}^{10 \text{ keV}} e^{-N_H \sigma(E)} N(E)^{1-\Gamma} dE,$$

where $D_L$ is the luminosity distance, $\sigma(E)$ is the photoelectric cross-section from Morrison & McCammon (1983), $E$ is the photon energy in units of keV, $N_H$, $N$, and $\Gamma$ are the best-fit absorbing column density, normalization factor, and photon index given by XMMFITCAT, respectively. When the double power-law model (i.e., model 5) is selected as the preferred model, we evaluate $L_{\text{2-10}}^{\text{obs}}$ by summing the luminosities of the
two power-law components. When the preferred model has no power-law component and purely includes thermal emission (i.e., model 1), we use the 2–10 keV X-ray luminosity of the thermal component as an upper limit on $L_{2-10}^{\text{int}}$. There are nine sources in this last category, and they are indicated as upper limits for both $L_{2-10}^{\text{obs}}$ and $L_{2-10}^{\text{int}}$ in Table 3.

Given the significant overlap between the XMMFITCAT and Swift/BAT AGN catalogs, we are able to check for consistency in $L_{2-10}^{\text{obs}}$ measurements. Figure 2 compares the luminosities obtained from data appearing in both catalogs, where it is readily apparent that with the exception of a few outliers, there is good general consistency between the two sources of $L_{2-10}^{\text{obs}}$ measurements (the two main outliers for which $L_{2-10}^{\text{obs}}$/Swift/BAT are substantially higher than $L_{2-10}^{\text{obs}}$/XMM-Newton are 3C 84 and 2MASX J05580206–3820043, which are nonmasers). The mean scatter between the two different measurements is 0.16 dex, well within the scatter (∼0.3 dex) seen in the correlation between $L_{2-10}^{\text{est}}$ and $L_{2-10}^{\text{obs}}$ (see Section 2.1.1), which is partly caused by the intrinsic flux variability (Ricci et al. 2017).

2.1.3. INTEGRAL/IBIS Survey

The latest version of the INTEGRAL/IBIS AGN catalog (Malizia et al. 2012, 2016) consists of 363 high-energy emitters confirmed to be AGNs. This catalog provides 2–10 keV fluxes and absorbing column densities for these sources based on spectral fitting. We found 147 counterparts in the GBT sample, with 143 of them being either included in the Swift/BAT AGN catalog or XMRFITCAT. Thus, cross-matching with the INTEGRAL/IBIS AGN catalog only provides X-ray information for four additional sources.

2.1.4. The Chandra Survey of Nearby Galaxies

This X-ray catalog is derived from the Chandra survey of nearby galaxies (She et al. 2017), and it contains 314 AGNs within 50 Mpc. The majority of these sources are low-luminosity AGNs. X-ray spectra were extracted for 154 AGNs having photon counts $\geq$100 in the energy band 0.3–8 keV. For these sources, She et al. (2017) fit either a simple absorbed power-law model, a power-law model plus a thermal component, or a partial absorption model to the spectral data to determine the X-ray properties. For the 160 sources without sufficient photon counts (i.e., $<100$), the authors estimate the 2–10 keV flux and luminosity with their source count rate assuming an absorbed power-law model, and they infer the absorbing column $N_{\text{H}}$ based on the observed hardness ratio.
Here, we consider only the 154 AGNs with available spectral-fitting results. After cross-matching with the GBT sample and removing sources already included in the previous catalogs, the Chandra AGN catalog provides 27 additional AGNs for our current study. For most of these sources (24 out of 27), the X-ray spectra can be fit with either a simple absorbed power-law or with a power-law model plus a thermal component; for these sources, we evaluate $L_{2-10}^{\text{obs}}$ using Equation (1), based on the spectral-fitting results provided by Table 4 in She et al. (2017). For the remaining (3 out of 27) sources that are best fit by the partial absorption model, we requested the partial covering factor from R. She (2020, private communication). We use this additional information to evaluate the $2-10$ keV luminosity based on

$$L_{2-10}^{\text{obs}} = 4\pi D_L^2 \int^{10^{10}\text{keV}}_{2\text{keV}} [\sigma N_{e}(E) + (1 - \eta)] NE^{-1}dE,$$

where $\eta$ is the dimensionless covering fraction ($0 < \eta < 1$).

2.1.5. NED Catalog

To identify additional X-ray measurements, particularly estimates of the absorbing column density for GBT sample galaxies not included in the four large catalogs discussed above, we searched the references provided by NED. Our extensive literature search resulted in X-ray photometric data for an additional 39 sources in the GBT sample, of which 16 galaxies have $L_{2-10}^{\text{obs}}$ and $N_H$ derived from spectral fitting (see the references in Table 2). For the 23 sources without spectral fits, the references provide only the total $2-10$ keV luminosity, which may contain thermal emission in this energy band. As a result, we do not use these measurements to infer $N_H$ from the $L_{12\mu m} - L_{2-10}^{\text{obs}}$ diagram so as not to contaminate it with underestimates of the column density (upper limits are not necessarily relevant for this study).

2.2. Mid-infrared AGN Luminosity

2.2.1. The Methods

Studies of mid-IR–X-ray relations in AGNs often determine mid-IR AGN luminosities ($L_{\text{MIR}}$) by performing high-resolution (i.e., subarcsecond) imaging of the target sources in these wavelengths (e.g., Fiore et al. 2009; Gandhi et al. 2009; Goulding et al. 2011; Sazonov et al. 2012; Asmus et al. 2014, 2015, 2016). High-resolution imaging permits high contrast with the host galaxy light, thus ensuring that the AGN dominates the mid-IR emission in the central region of the source. This generally leads to an accurate measurement of $L_{\text{MIR}}$ (or simply the $6\mu m$ or $12\mu m$ luminosity), which has been shown to be tightly correlated with the absorption-corrected intrinsic $2-10$ keV X-ray luminosity $L_{2-10}^{\text{int}}$ of the central AGN (the $L_{\text{MIR}} - L_{2-10}^{\text{int}}$ relations; Lutz et al. 2004; Fiore et al. 2009; Gandhi et al. 2009; Sazonov et al. 2012; Asmus et al. 2015). One of the most well-known relations was found by Gandhi et al. (2009), the Gandhi relation hereafter), who discovered a tight correlation between the $12\mu m$ AGN luminosity $L_{12\mu m}^{\text{AGN}} \equiv \nu L_{\nu}(12\mu m)$ and the intrinsic $2-10$ keV luminosity $L_{2-10}^{\text{int}}$.

An alternative way to reliably measure the $L_{\text{MIR}}$ of an AGN is to fit SED templates to broadband photometric observations of the source, which allow for a decomposition of the total observed mid-IR emission into components associated with the AGN and with the host galaxy (e.g., Georgantopoulos et al. 2011; Mateos et al. 2015; Kouloridis et al. 2016). This approach is particularly useful for galaxy samples where high-resolution mid-IR imaging is not available, as is the case for our GBT sample. To follow this approach, we perform SED decomposition for our sources using a recent version of the SED-fitting code MAGPHYS (da Cunha et al. 2008, 2015; Chang et al. 2017) applied to broadband UV-to-mid-IR photometric data.

2.2.2. The SED Decomposition

In MAGPHYS, the AGN emission is reproduced from a set of empirical templates (see Figure 2 in Chang et al. 2017; from Mullaney et al. 2011, templates 1 and 2, corresponding to low- and high-luminosity Seyfert 2s; Richards et al. 2006, template 3, for the empirical QSO template; and Polletta et al. 2007, template 4, for the average Seyfert 1 galaxies), which span in a representative way the global range of known AGN SEDs. To avoid degeneracies in the SED fitting, we only use four typical templates.

The broadband photometric data used in our SED fitting includes data taken in the UV, optical, near-IR, and mid-IR wavelengths, with the UV-to-optical data obtained from Data Release 14 (DR14) of the SDSS (e.g., Abolfathi et al. 2018), the near-IR data from the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), and the mid-IR data from the Widefield Infrared Survey Explorer (WISE; Wright et al. 2010). We obtain the SDSS data via Skyserver DR1418 and access the 2MASS and WISE data via the NASA/IPAC Infrared Science Archive using the IRS/7 catalog tool.

The final product of the MAGPHYS SED fitting is a set of best-fit parameters, including the infrared AGN luminosity $L_{\text{IR, AGN}}$ in the wavelength range 3–2000 $\mu m$. Because we are mainly interested in the $12\mu m$ AGN luminosity $L_{12\mu m}^{\text{AGN}}$, we measure this value by using the conversion factors $\epsilon \equiv L_{\text{IR, AGN}} / L_{12\mu m}^{\text{AGN}}$ of 3.07, 2.49, 2.62, and 2.50 for templates 1, 2, 3, and 4, respectively, which we have calculated directly from the four SED templates.

The following section provides details on the data collection and strategies for minimizing host galaxy contamination in the $12\mu m$ AGN luminosity measurements.

2.2.3. The Broadband UV-to-mid-IR Photometry

To obtain the broadband photometry, we cross-match the X-ray sample with the SDSS, 2MASS, and WISE catalogs using a matching radius of $6\arcsec$ (i.e., the resolution of WISE at $3.4\mu m$ and $4.6\mu m$). We use a smaller matching radius here because we assume that the position accuracies of galaxies in the optical and infrared catalogs are better than those in the X-ray catalogs. We note that increasing the matching radius to $10\arcsec$ does not increase the number of successful cross-matches and therefore does not affect our results.

Among the source catalogs provided by 2MASS, we adopt the extended source catalog (XSC) for cross-matching. To minimize the chance of mismatch, we further require the X-ray sample galaxies to be within the effective radii $r_{\text{eff}}$ of the matched sources, with $r_{\text{eff}}$ determined primarily from the 2MASS catalog (88% galaxies). When $r_{\text{eff}}$ is not available in

---

18 http://cas.sdss.org/dr14
2MASS XSC for a particular source (12% of the X-ray sample), we adopt the value from the SDSS catalog for making the comparison. Based on this cross-matching procedure, we find that there are 291 galaxies (45%) in the X-ray sample that do not have UV–optical photometry available from the SDSS catalog. For these sources, we perform the SED fitting using only the photometry from the near-IR and mid-IR bands.

Following the procedure described in Mateos et al. (2015), we adopt the SDSS model magnitudes in the $u$, $g$, $r$, $i$, $z$ (i.e., UV to optical) bands, while for the 2MASS $JHK$ bands (at 1.25, 1.65, and 2.17 μm), we use the fiducial Kron elliptical aperture magnitudes. For both the SDSS and 2MASS data, we applied a Galactic extinction correction according to Schlafly & Finkbeiner (2011). When evaluating the global galaxy flux densities in the WISE bands (i.e., $W_1$, $W_2$, $W_3$, $W_4$, at 3.4, 4.6, 12, 22 μm, respectively), we use the elliptical aperture magnitudes ($w_{gmag}$) for the X-ray sample sources that are flagged as extended (constituting 99% of the sample), and we adopt profile-fit photometric magnitudes for the remaining 1% of sources that are spatially unresolved in all of the WISE bands.

To account for possible systematic errors in the photometric data, we adopt the systematic uncertainties of 0.1 mag used by Chang et al. (2015) for the SDSS and WISE bands, and we add 0.05 mag of systematic uncertainty in quadrature with the photometric errors for the 2MASS bands. For the WISE data, we further correct for the systematic offsets in $w_{gmag}$ that lead to underestimation of the global galaxy fluxes based on the information provided by the WISE Explanatory Supplement.\(^{19}\)

### 2.2.4. Host Galaxy Contamination in the Mid-IR AGN Luminosity

Before using $L_{12,\mu m}^{AGN}$ as measured from the SED fitting to infer the absorbing column density using comparisons with $L_{2-10}^{obs}$, it is important to check whether the residual mid-IR emission from star-forming activity (hereafter referred to as host galaxy contamination) contributes substantially to $L_{12,\mu m}^{AGN}$. Significant host galaxy contamination in the mid-IR AGN luminosity will make the ratio $L_{2-10}^{obs}/L_{12,\mu m}^{AGN}$ artificially low and result in an overestimate of the absorbing column density.

To determine the level of any host galaxy contamination, we compare $L_{12,\mu m}^{AGN}$ as measured from the SED fitting to the prediction based on the Asmus relation (Asmus et al. 2015), which is the extension of the Gandhi relation to AGNs having a low luminosity of $L_{2-10} \sim 10^{40}$ erg s\(^{-1}\). In order to make predictions with the Asmus relation, we require reliable measurements of $L_{2-10}^{int}$ (i.e., those that are not affected by the systematic biases caused by heavy absorbing columns in Compton-thick AGNs). For this comparison, we therefore primarily use galaxies from the Swift/BAT catalog. We also include, where appropriate, Compton-thin sources from the XMM-Newton catalog, for which $L_{2-10}^{int}$ can be reliably inferred from the best-fit parameters of the X-ray spectral analysis in XMMFITCAT.

To differentiate between Compton-thin and Compton-thick sources, we first collect the XMM-Newton galaxies that have X-ray-to-[O III] luminosity ratios (i.e., the $T$ ratio; $T ≡ L_{2-10}^{obs}/L_{[O III]}$) available, followed by examining these sources with the $T$-ratio method from Bassani et al. (1999). The $T$-ratio method provides a convenient way to separate Compton-thin AGNs ($T > 1$) from Compton-thick sources ($T < 1$; Bassani et al. 1999; Cappi et al. 2006; Panessa et al. 2006). Considering the effect of X-ray variability and the uncertainty in the reddening correction to $L_{[O III]}$ (Brightman & Nandra 2011), we set $L_{2-10}^{int}/L_{[O III]} > 5$ to better select Compton-thin AGNs in our comparison for XMM-Newton sources. Note that the [O III] luminosities used in our examination were gathered via an extensive literature search (see references in Table 1); of all the reported $L_{[O III]}$ values, we only choose those measurements having a Balmer decrement available for reddening correction, for which we assume an intrinsic ratio of (Hβ/Hδ) of 3.0 and a Hβ/Hα color index for extinction of 2.94 (Bassani et al. 1999).

Figure 3 presents the relationship between $L_{12,\mu m}^{AGN}$ and $L_{12,\mu m}^{AGN} = \nu L_{\nu}(12 \mu m)$ based on two different sets of flux measurements, along with the Asmus relation. The first set of measurements (left panel) employs global galaxy fluxes for the SED fitting. While these measurements are generally consistent with the Asmus relation for bright X-ray AGNs (log $L_{2-10}^{int} > 42.5$), they do not exhibit any apparent correlation below this threshold. This behavior matches the findings of Mullaney et al. (2011), who showed that $L_{MBR}$ derived from SED fitting might include significant host galaxy contamination for lower luminosity AGNs. Therefore, in order to infer reliable absorbing column densities for the lower luminosity sources, we need to find ways to obtain more accurate $L_{12,\mu m}^{AGN}$ estimates so as to substantially reduce the scatter in the correlation with $L_{2-10}^{int}$.

### 2.2.5. The Approach to Minimizing the Host Galaxy Contamination

To minimize the host galaxy contamination in lower luminosity AGNs while retaining the total AGN flux, we find—for all WISE bands—that it is helpful for the SED fitting to use the profile-fit magnitudes ($w_{gmag}$) rather than using $w_{gmag}$. Because AGNs are point sources at mid-IR wavelengths, we add photometric errors for the 2MASS bands, and we add 0.05 mag of systematic uncertainty in quadrature with the photometric errors for the 2MASS bands. For the WISE data, we further correct for the systematic offsets in $w_{gmag}$ that lead to underestimation of the global galaxy fluxes based on the information provided by the WISE Explanatory Supplement.

http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec6_3e.html

The right panel of Figure 3 presents the $L_{12,\mu m}^{AGN}$ values obtained from fitting these host-galaxy-minimized SED measurements. We see that these $L_{MBR}$ match the prediction of the Asmus relation across the six orders of magnitude spanned by our sample. We also see that the scatter in the $L_{12,\mu m}^{AGN}/L_{2-10}^{int}$ relation has been substantially reduced for the data with log $L_{2-10}^{int}$ less than 42.5. This behavior suggests that the 12 μm AGN luminosity derived from the aperture photometry SED fitting is reliable and can be used to build the $L_{12,\mu m}^{AGN}/L_{2-10}^{obs}$ diagram for investigating the maser detection...
problem. Note that despite the relatively large scatter (0.82 dex) still remaining for the log $L_{210}^{\text{int}}$ data, the impact on our conclusions will remain minimal as nearly all maser galaxies have $L_{210}^{\text{int}}$ greater than $10^{41}$ erg s$^{-1}$.

3. Results: Consequences for the Megamaser Detection Rates

Before looking for ways to enhance the maser detection rates using the mid-IR and X-ray photometry simultaneously, it is useful to first study the dependence of maser detection rates on the mid-IR AGN luminosity ($L_{12\mu \text{m}}^{\text{AGN}}$) and absorbing column density ($N_{\text{H}}$), both separately and in concert. Evaluating the relative importance of these two factors may help to identify sample selection criteria that simultaneously maximize both the detection and completion rates for maser galaxies.

3.1. The Effect of Mid-IR AGN Luminosity

In the left panel of Figure 4, we plot the detection rates of kilomasers (dotted yellow), megamasers (solid cyan), and disk masers (dashed magenta) as functions of $L_{12\mu \text{m}}^{\text{AGN}}$. Right panel: the fraction of galaxies in the X-ray sample as a function of 12 $\mu$m AGN luminosity, $L_{12\mu \text{m}}^{\text{AGN}}$. $\Delta N$ and $N$ refer to the number of sources in each luminosity bin and the total number of galaxies in the X-ray sample, respectively.
10^{38} \text{ to } 10^{42} \text{ erg s}^{-1} \text{ and then get boosted substantially by a factor of } \sim 2.7 \text{ when } L_{12 \mu m}^{AGN} \text{ is greater than } 10^{42} \text{ erg s}^{-1} \text{ (i.e., } L_{12 \mu m}^{int} \geq 5.0 \times 10^{41} \text{ erg s}^{-1}, \text{ according to the Asmus relation). This behavior suggests that, even without including the effect of } N_{H}, \text{ one could immediately increase the megamaser detection rates by a factor of } \sim 3 \text{ relative to the GBT sample}^{20} \text{ by making a mid-IR AGN luminosity cut of } L_{12 \mu m}^{AGN} > 10^{42} \text{ erg s}^{-1} \text{ when selecting Type II AGNs for maser surveys.}

However, as shown in the right panel of Figure 4, the majority (67\%) of galaxies in the X-ray sample already have $L_{12 \mu m}^{AGN} > 10^{42} \text{ erg s}^{-1}$. The galaxy distribution is biased toward higher mid-IR AGN luminosity through the inclusion of the Swift/BAT sample, which primarily consists of sources with $L_{12 \mu m}^{AGN} > 10^{42} \text{ erg s}^{-1}$. Because the X-ray sample is dominated by these higher luminosity AGNs, making the mid-IR AGN luminosity cut of $L_{12 \mu m}^{AGN} > 10^{42} \text{ erg s}^{-1}$ in the X-ray sample only leads to an increase in megamaser detection rates by a factor of $\sim 1.4$ with respect to the X-ray sample itself.

3.2. The Effect of the Obscuring Column Density

To explore how the maser detection rates depend on obscuring column density, we first infer $N_{H}$ for the X-ray sample from the ratio of the observed X-ray to the mid-IR AGN luminosity based on the $L_{12 \mu m}^{AGN}/L_{2-10}^{obs}$ diagram (Satyalap et al. 2017). This diagram allows one to estimate the $N_{H}$ value of an AGN because the more obscured an AGN is, the stronger the X-ray absorption will be, suppressing the observed X-ray luminosity $L_{2-10}^{obs}$. Because the reprocessed mid-IR continuum emission in AGNs is relatively unaffected by the X-ray-absorbing material and thus gives a reliable estimate of the intrinsic AGN luminosity (even for Compton-thick sources; Goulding et al. 2011), the ratio of the observed X-ray to mid-IR AGN luminosity allows one to evaluate the magnitude of suppression in $L_{2-10}^{obs}$ and thus provides an estimate of $N_{H}$, especially at $N_{H} > 10^{23} \text{ cm}^{-2}$ (Asmus et al. 2015).

In Figure 5, we plot $L_{12 \mu m}^{AGN}$ against $L_{2-10}^{obs}$ for the sources in the X-ray sample. In this diagram, the yellow triangles, cyan squares, and magenta filled circles represent the kilomasers, megamasers, and disk masers, respectively; the sample of X-ray galaxies that have been searched for 22 GHz emission with the GBT but have no detection (i.e., the nonmasers) is shown as gray circles.

To estimate the $N_{H}$ value of each AGN, we compare $L_{2-10}^{obs}$ with $L_{12 \mu m}^{int}$ evaluated from the Asmus relation (the purple line) at each given $L_{12 \mu m}^{AGN}$. The cyan, green, orange, and brown lines indicate the $L_{2-10}^{obs}$ expected for sources with log $N_{H} = 22, 23, 24, \text{ and } 24.3 \text{ cm}^{-2}$, respectively, based on Equation (1) and the assumption that the photon index $\Gamma = 1.8$ (Nandra & Pounds 1994; Dadina 2008). We note that while we do not include a reflected X-ray component from the AGN (which becomes important only when log $N_{H} > 24.3$), our prediction is nevertheless consistent with that made by the MYTORUS model$^{22}$ (Murphy & Yaqoob 2009; Satyalap et al. 2017), which considers both intrinsic absorption and reprocessed X-ray emission from a toroidal absorber for the full range of column density distributions. Therefore, replacing our prediction with that from the MYTORUS model does not change our conclusion.

The comparison shown in Figure 5 indicates that the majority (93\%) of maser sources have $N_{H} \geq 10^{23} \text{ cm}^{-2}$, consistent with the findings from Greenhill et al. (2008) and Castangia et al. (2013), who derived the $N_{H}$ of maser sources based on X-ray spectral fitting. This behavior suggests that the maser detection rates are strong functions of absorbing column density.

In the first four rows of Table 4, we list the detection rates for all masers, megamasers, and disk masers for sources having $N_{H} < 10^{23} \text{ cm}^{-2}$, $10^{23} < N_{H} < 10^{24} \text{ cm}^{-2}$, $N_{H} \geq 10^{24} \text{ cm}^{-2}$, and $N_{H} > 10^{24} \text{ cm}^{-2}$, respectively. The table shows that the maser detection rates get boosted by more than an order of magnitude when the column density of an AGN increases from $N_{H} < 10^{23} \text{ cm}^{-2}$ to $N_{H} \geq 10^{23} \text{ cm}^{-2}$. Moreover, with respect to the maser detection rates of the GBT sample, this table also shows that one can increase the detection rate of megamasers (disk masers) by a factor of $\sim 6 (7)$ and $\sim 7 (10)$ by choosing galaxies with $N_{H} > 10^{23} \text{ cm}^{-2}$ and $N_{H} > 10^{24} \text{ cm}^{-2}$, respectively.

Interestingly, unlike sample selections based solely on the mid-IR colors (Kuo et al. 2018), these dramatic increases in maser detection rates do not come at the cost of the completeness rates ($C_{\text{maser}}$) when X-ray detection and information are included. Here, $C_{\text{maser}}$ is defined as $N_{\text{cut}}^{\text{maser}}/N_{\text{tot}}^{\text{maser}}$, where $N_{\text{tot}}^{\text{maser}}$ indicates the total number of maser detections in the X-ray sample and $N_{\text{cut}}^{\text{maser}}$ refers to the number of detected masers which satisfy certain cuts in $L_{12 \mu m}^{AGN}$ or $N_{H}$. When evaluating the completeness rates for megamasers ($C_{\text{megamasers}}$) or disk masers ($C_{\text{disk}}$), we replace the $N_{\text{cut}}^{\text{maser}}$ in the definition for

$^{20}$ The detection rates of all masers, megamasers, and disk masers in the GBT sample are $3.3\% \pm 0.1\%$, $2.7\% \pm 1.1\%$, and $0.9\% \pm 0.8\%$, respectively (see Table 3 in Kuo et al. 2018).

$^{21}$ log $L_{12 \mu m}^{obs} = log L_{12 \mu m}^{int} - \epsilon$, where $\epsilon = 0.0414, 0.2810, 1.3556$, and $2.2900$ for log $N_{H} = 22, 23, 24, \text{ and } 24.3 \text{ cm}^{-2}$, respectively.

$^{22}$ log $L_{12 \mu m}^{obs} = log L_{12 \mu m}^{int} - \eta$, where $\eta = 0.2929, 1.3673$, and $2.20703$ for log $N_{H} = 23, 24, \text{ and } 24.3 \text{ cm}^{-2}$, respectively.
C_maser by N_{cut}^{Mmaser} and N_{cut}^{disk}, where N_{cut}^{Mmaser} and N_{cut}^{disk} indicates the number of detected megamasers and disk masers that meet the chosen selection criteria, respectively.

From Table 1, it can be seen that the X-ray sample completeness rate for sources with N_H ≳ 10^{23} cm^{-2} remains close to 100% for megamasers (C_{Mmaser} = 93.3%) and disk masers (C_{disk} = 94.7%), even when the detection rates are as high as 18.8% ± 2.6% and 6.4% ± 1.5%, respectively. Even if one wants to achieve the highest maser detection rates (i.e., 25.0% ± 4.5%) by selecting Compton-thick candidates (i.e., N_H ≳ 10^{24} cm^{-2}), the completeness rates (~50%) remain substantially higher than the rates (~30%) achieved by the mid-IR selection criterion of W1 − W2 > 0.5 and W1 − W4 > 7 that led to the highest detection rate (18%) in Kuo et al. (2018). This result indicates that selecting sources based on the obscuring column density allows one to achieve both high detection rates and high completeness rates simultaneously.

### 3.3. The Joint Effect of L_{12 μm} and N_H

Given that both the mid-IR AGN luminosity and the absorbing column density can play substantial roles for maser detection when examined individually, it is natural to explore whether combining the two factors could provide an even more effective optimization. In the last four rows of Table 4, we show the maser detection rates and completeness rates as a function of log N_H for all sources having L_{12 μm}^{AGN} \geq 10^{42} erg s^{-1}. We see that the detection rates of megamasers and disk masers in highly obscured and Compton-thick AGNs increase by another factor of 1.3–1.5 by enforcing that the 12 μm AGN luminosity be greater than 10^{42} erg s^{-1}. While the mid-IR AGN luminosity appears to play a secondary role in boosting the megamaser detection rates, the highest detection rates achieved by making a cut in L_{12 μm}^{AGN} become 27.0% ± 6.2% and 13.5% ± 4.3% for megamasers and disk masers, respectively, corresponding to a factor of 10 and 15 boost in the detection rates with respect to the GBT sample. Meanwhile, the completeness rates remain at the level of ~80% (~50%) for sources with N_H \geq 10^{22} cm^{-2} (N_H \geq 10^{21} cm^{-2}), indicating that selecting AGNs based on L_{AGN} and N_H can lead to high maser detection rates and completeness rates simultaneously.

### 3.4. The L_{2−10 keV}^{obs}−L_{O III}^{AGN} Diagram

To examine the robustness of our conclusion made in the previous section, which is drawn from an indirect estimate of absorbing column density based on the L_{2−10 keV}^{obs}−L_{12 μm}^{AGN} diagram, it would be useful to see if we can arrive at the same conclusion by using independent estimates or direct measurements of N_H, which are available for subsets of the X-ray sample.

Among the 642 sources in the X-ray sample, 342 (53%) of them have reddenning-corrected [O III] λ5007 luminosities L_{O III}^{AGN} available in the literature (see Table 2). Similar to the reprocessed mid-IR continuum emission, L_{O III}^{AGN} is often used as an indicator of isotropic AGN luminosity because it is not affected by the optically thick X-ray-absorbing material (Heckman et al. 2005; Panessa et al. 2006; Goulding et al. 2011; Koulouridis et al. 2016). Through the examination of a local sample of Seyfert 2 galaxies, Bassani et al. (1999) demonstrated that the X-ray−[O III] flux ratio (i.e., the T ratio; T ≡ L_{2−10 keV}^{obs} / L_{O III}^{AGN}) of T < 1 provides a reliable diagnostic for Compton-thick AGNs. Therefore, for the subset of the X-ray sample with L_{O III}^{AGN} available, we are able to explore the maser detection rates for Compton-thick sources by identifying these sources based on the T-ratio method.

In the left panel of Figure 6, we plot L_{O III}^{AGN} against the observed 2−10 keV AGN luminosity. The dashed line indicates the region where the T ratio equals 1. All of the sources lying below the dashed lines (i.e., T < 1) are candidates for Compton-thick AGNs. Based on the distribution seen in the plot, we can infer that the detection rates of all masers, megamasers, and disk masers in the Compton-thick candidates are 25.5% ± 5.2%, 21.3% ± 4.8%, 13.8% ± 3.8%, respectively. These values are consistent with the maser detection rates inferred from the L_{12 μm}^{AGN}−L_{O III}^{AGN} diagram for sources with log N_H \geq 24 (see the fourth row of Table 4), supporting the conclusions from Section 3.2 for Compton-thick sources.

Note that although L_{O III}^{AGN} provides an independent measurement of the intrinsic AGN luminosity, we do not use the the ratio of the observed X-ray to [O III] luminosity to infer the absorbing column density as we did in Section 3.2 for Compton-thin sources (i.e., log N_H < 24). This is because the intrinsic 2−10 keV luminosity and L_{O III}^{AGN} do not show a tight correlation (Heckman et al. 2005; Georgantopoulos & Akylas 2010; Berney et al. 2015), particularly for Seyfert 2 galaxies. Nevertheless, one can see from the left panel of Figure 6 that some maser galaxies do have T > 1. This suggests that maser galaxies reside in Compton-thin AGNs as well, consistent with what we see in Figure 5. To robustly confirm the maser detection rates in Compton-thin sources, it would be best to use a direct measurement of N_H from X-ray spectral fitting.

### Table 4: Effective Search Criteria for Megamasers and Disks

| log N_H (cm^{-2}) | L_{12 μm}^{AGN} (erg s^{-1}) | % All MCP Galaxies | R_{maser} | C_{maser} | R_{Mmaser} | C_{Mmaser} | R_{Disk} | C_{Disk} | N_{Disk} |
|-------------------|-----------------------------|-------------------|-----------|----------|----------|----------|---------|---------|---------|
| log N_H < 23      | > 10^{18}                   | 48.7              | 1.1 ± 0.7 | 5.4      | 1.1 ± 0.7 | 6.7      | 0.4 ± 0.4 | 5.3      | 1 ± 1   |
| 23 ≤ log N_H < 24 | > 10^{18}                   | 31.2              | 15.4 ± 3.0| 46.1     | 11.8 ± 2.6| 44.4     | 4.7 ± 1.7 | 42.1     | 6 ± 2   |
| log N_H ≥ 24      | > 10^{18}                   | 51.2              | 19.1 ± 2.6| 94.6     | 15.1 ± 2.3| 93.3     | 6.5 ± 1.5 | 94.7     | 15 ± 2  |
| log N_H < 23      | > 10^{18}                   | 20.0              | 25.0 ± 4.8| 47.4     | 20.3 ± 4.3| 48.9     | 9.3 ± 2.9 | 52.6     | 9 ± 3   |
| 23 ≤ log N_H < 24 | > 10^{18}                   | 32.5              | 1.1 ± 0.8 | 3.6      | 1.1 ± 0.8 | 4.4      | 0.0 ± 0.0 | 0.0      | 0       |
| log N_H ≥ 24      | > 10^{42}                   | 23.1              | 16.8 ± 3.7| 37.5     | 13.6 ± 3.3| 37.8     | 4.8 ± 2.0 | 31.6     | 3 ± 1   |
| log N_H < 23      | > 10^{42}                   | 36.8              | 22.1 ± 3.3| 78.6     | 19.1 ± 3.1| 84.4     | 8.0 ± 2.0 | 84.2     | 10 ± 3  |
| 23 ≤ log N_H < 24 | > 10^{42}                   | 13.7              | 31.1 ± 6.5| 41.1     | 27.0 ± 6.2| 46.7     | 13.5 ± 4.3| 52.6     | 8 ± 3   |

Note: R_{maser}, R_{Mmaser}, and R_{Disk} are the detection rates of all masers, megamasers, and disks, respectively. C is the completeness rate, listed separately for all masers, megamasers, and disks. The criteria that give the highest detection and completion rates for megamasers and disks are highlighted in bold. N_{Disk} shows the predicted number of new disk masers that could be detected from sources in the Swift/BAT AGN catalog and the XMM-Newton spectral-fitting database, which are not yet searched for masers with the GBT.
3.5. The $L_{21\,\text{cm}}^{\text{int}}-N_{\text{H}}$ Diagram

The spectral analyses from the X-ray surveys provide the absorbing column densities $N_{\text{H}}$ for all sources in the X-ray sample (see Table 2). However, out of the 314 X-ray sources from the XMM-Newton catalog, there are 132 sources for which the column densities were fixed in the spectral-fitting process. This happens when the quality of the spectrum is not high enough (e.g., low signal-to-noise ratio) to fit $N_{\text{H}}$ robustly. These $N_{\text{H}}$ values tend to be subject to greater uncertainties, and we do not use them in our present comparison.

Furthermore, for sources drawn from X-ray surveys probing $\lesssim 10\,\text{keV}$ bands (e.g., XMM-Newton and Chandra), the $N_{\text{H}}$ values can be significantly underestimated if the AGNs are Compton-thick (e.g., Cappi et al. 2006; Panessa et al. 2006; Singh et al. 2011; Castangia et al. 2013). Therefore, to reliably measure the maser detection rates as a function of $N_{\text{H}}$, we primarily use galaxies from the Swift/BAT AGN catalog, which provides reliable $N_{\text{H}}$ even for Compton-thick sources. In addition, we also use XMM-Newton sources that have $L_{\text{O III}}$ available so that we can infer a more reliable $N_{\text{H}}$ for Compton-thick AGNs based on the $T$-ratio method. In total, this sample of more-reliable $N_{\text{H}}$ measurements includes 417 sources, with 307 from the Swift/BAT AGN catalog and 110 from XMMFITCAT.

In the right panel of Figure 6, we plot $N_{\text{H}}$ as a function of $L_{21\,\text{cm}}^{\text{int}}$. In this diagram, for those sources drawn from the XMM-Newton catalog with $T \lesssim 1$, we follow Cappi et al. (2006) in setting $\log N_{\text{H}} = 24.3$ to indicate that these sources are Compton-thick.\footnote{If $\log N_{\text{H}}$ equals 24.3, it indicates that 99.5% of the AGN X-ray emission is blocked by the obscuring material, assuming the photon index of the X-ray spectrum to be $\Gamma = 1.8$.} We do not use the $N_{\text{H}}$ provided by XMMFITCAT for these sources because their $N_{\text{H}}$ are likely to be underestimated. To avoid significant clustering and overlapping of points in the figure, we randomize the $N_{\text{H}}$ values uniformly between 24.1 and 24.5. The inferred maser detection rates as a function of $\log N_{\text{H}}$ are shown in the left panel of Figure 7. From this plot, we can see that the overall distribution is consistent with what we show in Table 4. To further compare the detection rate distributions for more luminous sources, we show in the right panel of Figure 7 the maser detection rates as a function of $\log N_{\text{H}}$ for sources with $\log L_{21\,\text{cm}}^{\text{int}} > 41.7$ (i.e., $L_{\text{AGN}}^{21\,\text{cm}} > 10^{42}\,\text{erg s}^{-1}$).

While the distribution remains unchanged for $\log N_{\text{H}} < 24\)\), the detection rates increase from 30.4% ± 6.6% to 38.5% ± 8.6% for megamasers and from 15.9% ± 4.8% to 19.2% ± 6.2% for disk masers. Although there are 11.5% and 5.7% increases relative to the detection rates for Compton-thick megamasers and disk masers shown in Table 4, respectively, the rates are consistent at the $\sim 1\sigma$ level.

3.6. Photon Index and Eddington Ratio

We have shown that the maser detection rates depend substantially on two fundamental properties: the mid-IR intrinsic luminosity and the obscuring column density. Given the tight relationship between the maser phenomena and AGN activities, one may wonder whether the maser detection rates also correlate with another fundamental AGN property: the Eddington ratio $\lambda_{\text{Edd}} \equiv L_{\text{bol}} / L_{\text{Edd}}$ (where $L_{\text{bol}}$ and $L_{\text{Edd}}$ refer to the bolometric and Eddington luminosity, respectively). Because $\lambda_{\text{Edd}}$ is known to be correlated with the photon index $\Gamma$ of the AGN X-ray spectrum (e.g., Shemmer et al. 2006; Constantin et al. 2009; Brightman et al. 2013), it would be interesting to see whether masers are correlated with $\Gamma$ as well. We explore these possibilities in Figure 8.
The left and middle panels of Figure 8 show the $\Gamma - \lambda_{\text{Edd}}$ diagrams for sources from the 
$\text{Swift}/\text{BAT}$ catalog and the non-$\text{Swift}/\text{BAT}$ catalogs, respectively. We make the plots for the 
two different subsets of the X-ray sample separately because 
the photon indices provided by the $\text{Swift}/\text{BAT}$ survey measure 
the slopes of the intrinsic power-law spectra, whereas 
$\Gamma$ values 
in the non-$\text{Swift}/\text{BAT}$ catalog only indicate the apparent slopes 
of the X-ray spectra (which could contain significant reflection 
components in Compton-thick sources).

When evaluating $\lambda_{\text{Edd}}$, we obtain $L_{\text{bol}}$ from the 12 $\mu$m AGN 
luminosity by adopting the luminosity-dependent bolometric 
correction from Gandhi et al. (2009). The Eddington luminosity 
is calculated from the black hole mass as

$$L_{\text{Edd}} = 1.25 \times 10^{38} \left( \frac{M_{\text{BH}}}{M_{\odot}} \right) \text{erg s}^{-1}. \quad (3)$$

Following Pons & Watson (2014), we infer the black hole mass 
$M_{\text{BH}}$ from the rest-frame total $K$-band luminosity $L_{K,\text{tot}}$ of the host 
galaxy using the $M_{\text{BH}} - L_{K,\text{tot}}$ relation from Läsker et al. (2014).

Given the intrinsic scatter seen in the $M_{\text{BH}} - L_{K,\text{tot}}$ relation, the 
expected scatter in log $L_{\text{Edd}}$ is $\sim 0.5$ dex.

From the plot showing the $\text{Swift}/\text{BAT}$ sample, we see that there is no clear correlation between the photon indices and the 
maser sources, suggesting that the slopes of the intrinsic AGN X-ray spectra for the maser galaxies are statistically similar. In 
contrast, for sources in the non-$\text{Swift}/\text{BAT}$ surveys, represented mainly by the $\text{XMM-Newton}$ sources, the middle panel 
of Figure 8 shows that megamasers and disk masers 
preferentially have lower $\Gamma$ than the average value of the 
$\text{Swift}/\text{BAT}$ sample (i.e., $\Gamma_{\text{avg}} = 1.76$). This is consistent with 
the fact that $\sim 50\%$ of megamaser galaxies reside in AGNs with 
$N_{\text{H}} \gtrsim 10^{24}$ cm$^{-2}$, in which strong reflection components flatten 
the X-ray spectra at higher energies and give rise to smaller

---

24 Here, the best fit is obtained by the double power-law model, and we adopt 
the smaller of the two $\Gamma$ values provided by XMMFITCAT. The power-law 
spectrum associated with this photon index may reflect the scattered component 
of the hard X-ray emission (Corral et al. 2015).
By inspecting the maser distribution in the domain of the Eddington ratio in the right panel of Figure 8, we see that the maser detection rates for sources having $\lambda_{\text{Edd}}$ below $10^{-4}$ are zero, suggesting that $\lambda_{\text{Edd}} \sim 10^{-4}$ is an accretion efficiency threshold for maser excitation. We note, however, that this interpretation is based on a small fraction (11%) of the X-ray sample sources with $\lambda_{\text{Edd}} < 10^{-4}$, which may not be statistically significant enough to draw firm conclusions. In the X-ray sample we employ in this analysis, the sources with $\lambda_{\text{Edd}} < 10^{-4}$ mainly come from the XMM-Newton catalog, while such low-$\lambda_{\text{Edd}}$ AGNs are nearly absent in the Swift/BAT sample because the spectral analysis in the Swift/BAT catalog requires sufficient photon counts $N$ up to 150 keV, where $N$ can become significantly lower than the photon counts at $\lesssim 10$ keV bands; therefore, the Swift/BAT sample tends to include more luminous and higher $\lambda_{\text{Edd}}$ AGNs than the XMM-Newton sample in order to achieve sufficient sensitivity across the wide energy bands. Thus, the robustness of this interpretation will need to be tested by including more low-$\lambda_{\text{Edd}}$ sources in the analysis in the future.

For sources with $\lambda_{\text{Edd}} > 10^{-4}$, the detection rates of megamasers and disk masers are consistent with being flat functions of $\lambda_{\text{Edd}}$. This may imply that, as long as $\lambda_{\text{Edd}}$ is above the threshold, the accretion efficiency may not play an important role for exciting disk maser emission (at least for X-ray-selected AGNs). It is likely that the maser detection rates may be different functions of $\lambda_{\text{Edd}}$ for optically selected AGNs because these AGNs could contain a substantial population of X-ray-weak, unobscured Seyfert 2 galaxies (Georgantopoulos & Akylas 2010) at low luminosity and accretion rates (i.e., $\lambda_{\text{Edd}} < 10^{-5}$), for which the maser detection rates are expected to be low.

4. Discussion

4.1. Why Do Megamasers Prefer Mid-IR-luminous and X-ray-obsured AGNs?

The analysis presented here reveals that the H$_2$O megamaser detection rate ($R_{\text{Mmaser}}$) is a strong function of the column density $N_{\text{H}}$ of AGNs. In addition, the detection rate also increases with the 12 $\mu$m AGN luminosity $L_{12}\mu$m. Given the expectation that it is the dusty torus conceived in the unification paradigm of AGNs (Antonucci & Miller 1985) which blocks the X-ray photons and provides the reprocessed mid-IR emission in obscured AGNs, the dependence of $R_{\text{Mmaser}}$ on $N_{\text{H}}$ and $L_{12}\mu$m suggests that H$_2$O megamaser emission is closely associated with AGN circumnuclear obscuration, as previously expected (see also Kuo et al. 2018).

These results support previous findings regarding links between the properties of the obscuring material and those of the water megamasers. Masini et al. (2016) showed that the maser disk emission very likely originates from a geometrically thin disk composed of a large number of molecular clouds that connect the inner edge of the torus and the outer part of the accretion disk. Because maser emission is beamed and long path lengths are required for strong maser amplification, the detection rate of megamaser disks is then expected to be the highest when maser disks are close to being edge-on.

Assuming that the disk plane for maser emission is aligned with the equatorial plane of the circumnuclear dusty torus, one can perceive that only for an edge-on maser disk would the absorbing column density $N_{\text{H}}$ become especially high because one would encounter the greatest number of obscuring clouds along the line of sight. According to the standard cloud distribution of the clumpy torus model proposed by Elitzur & Shlosman (2006), the number of torus clouds intercepted by the line of sight can be expressed as $N_{\text{cl}} \propto \exp(-\beta^2/\sigma^2)$, where $\beta$ is the angular offset between the observing line of sight and the equatorial plane, and $\sigma$ ($\sim 45^\circ$) measures the angular width of the cloud distribution, which supports a correlation between the inclination angle and $N_{\text{cl}}$ (dropping quickly when $\beta > 0^\circ$), and thus $N_{\text{H}}$. Therefore, it can be expected that the further a maser disk is from being edge-on, the shorter the amplification path length would be, corresponding to weaker maser emission, which will be harder to detect. As the maser disk plane has a greater departure from being edge-on, i.e., a larger viewing angle $\beta$, $N_{\text{cl}}$ and thus $N_{\text{H}}$ would decrease.

A similar argument could also be applied to megamaser emission arising from a dusty torus (e.g., NGC 3079; Elitzur & Shlosman 2006). In such a scenario, strong maser emission can occur when a seed photon of maser emission from a torus cloud is amplified by other clouds which lie across the line of sight. Based on the cloud distribution described above, the chance for such alignments would be greatest when $\beta$ is small. Therefore, the strength of the megamaser emissions from torus clouds, and the associated detection rate, will again be strongly dependent on the viewing angle of the torus, which is well correlated with the absorbing column density $N_{\text{H}}$.

On the other hand, a large number of clouds along the line of sight is clearly not sufficient for generating strong maser emission. This is because maser excitation can become significant only when the gas temperature $T_{\text{gas}}$ is greater than $\sim 400$ K, which is the temperature above which the reaction network for generating H$_2$O molecules becomes very efficient and can lead to a water abundance of $10^{-5}$ relative to the density of H nuclei (Neufeld et al. 1994; Neufeld & Maloney 1995). The key to raising the gas temperature above $\sim 400$ K is the X-ray heating rate, which depends on both the illuminating hard X-ray flux $F_X$ and the column density $N_{\text{H}}$ of X-ray-shielding material for a maser-emitting cloud. Based on the relationship between $T_{\text{gas}}$, $F_X$, and $N_{\text{H}}$ obtained by Neufeld et al. (1994), it can be inferred that for a population of disk (or torus) clouds with a given $N_{\text{H}}$ distribution, the higher the X-ray flux is, the higher the fraction of the clouds that will be heated to sufficient $T_{\text{gas}}$ for maser emission will be. Therefore, if an AGN has a higher X-ray luminosity $L_X$ (i.e., higher 12 $\mu$m AGN luminosity), the megamaser emission is also more likely to become stronger thanks to the higher number of clouds at suitable $T_{\text{gas}}$ for maser emission and amplification, leading to higher megamaser detection rates.

4.2. The Effect of Galaxy Distance on Maser Detection

4.2.1. The Role of the Sensitivity Limit

Galaxy distance has a subtle effect on maser detection, and this effect may play different roles for optically selected and X-ray-selected AGNs. From the right panel of Figure 1, we can see that the maser detection rates of the X-ray sample drop...
quickly as a function of distance $D$. In addition, the detection rates of megamasers and disk masers increase by a factor of a few relative to the GBT sample at $D \sim 50$ Mpc. The rapidly decreasing trend in the megamaser detection rates—which are substantially enhanced at short distances—is in sharp contrast to the flat detection rate distribution for the GBT sample (see Figure 1 in Kuo et al. 2018), which shows that the detection rates of megamasers and disk masers are nearly constant ($\sim 2.7\%$ for megamasers and $\sim 0.9\%$ for disk masers) from $D = 0 \sim 180$ Mpc. Given that the X-ray sample is a subset of the GBT sample, we expect that the dramatically different appearances in the distributions of the megamaser detection rates are the result of selection effects. It is likely that the X-ray selection gives rise to a significant distance effect in the search for H$_2$O megamasers.

When looking for the causes of this difference in the detection rate distributions, we notice that the sensitivity limit for megamaser/disk maser detection in an H$_2$O maser survey cannot fully account for what we see. In the left panel of Figure 9, we plot the H$_2$O maser luminosities of all maser sources in the GBT sample as a function of $D$. The dashed and dotted–dashed lines show the $3\sigma$ and $5\sigma$ detection limits of the GBT maser survey presented in Kuo et al. (2018). Here, we define the detection limits as the sensitivity needed for detecting a single maser line with a line width broader than 1 km s$^{-1}$ (i.e., the typical width of a H$_2$O maser line) at $3\sigma$ or $5\sigma$ significance.

From the trends in the detection limits, one can infer that sensitivity starts to substantially affect the detection rate of megamasers (i.e., $L_{\text{H}_2\text{O}} \geq 10^7 L_{\odot}$, the sources above the horizontal blue line in the plot) only when $D$ is greater than $\sim 170$ Mpc, suggesting that one could in principle detect nearly all megamasers up to $D \sim 170$ Mpc, given the sensitivity level achieved in the GBT maser survey. This would not only explain why the megamaser detection rate is nearly constant for the GBT sample within $D \sim 180$ Mpc, but it would also suggest that the rapidly decreasing trend in the maser detection rates for the X-ray sample results from a different cause.

This conclusion would not change dramatically if one imposed a stricter definition for the detection limit by taking into account the fact that a maser detection usually appears in the form of line complexes. A line complex tends to be more difficult to detect for a given total maser luminosity because the total maser flux gets diluted when spread across multiple lines, resulting in each line having a smaller signal-to-noise ratio. In Figure 9, the dotted line represents the detection limit for a maser line complex consisting of at least three individual maser lines, with each line detected at the $3\sigma$ level. Based on this definition, one would start to lose detections of megamasers only when $D$ is greater than 100 Mpc. The megamaser detection rate will still be a flat function of $D$ within 100 Mpc, inconsistent with the quick drop in the detection rates seen in Figure 1.

Given the close relationship between megamasers/disk masers and the highly obscured/Compton-thick AGNs, we argue that the reason why the detection rates of megamasers and disk masers in the X-ray sample are strongly decreasing functions of $D$ is because the fraction of high column density AGNs decreases rapidly with $D$ for a sample of X-ray-selected AGNs. Due to strong photoelectric absorption, it is expected that the X-ray emission from highly obscured and Compton-thick AGNs will get significantly suppressed (e.g., Treister et al. 2009). These suppressed X-ray sources become harder to detect at greater distances in a sensitivity-limited X-ray survey, leading to a decreasing fraction of heavily obscured AGNs as $D$ increases. As a result, if one searches for masers from a sample of X-ray-selected sources in which the fraction of highly obscured AGNs becomes smaller because of the X-ray...
sensitivity limit, the expected maser detection yields will also be lower because the search will be conducted primarily among lower column density AGNs.

Indeed, when we examine the Compton-thick candidates seen in the \( L_{\text{AGN}}^{12 \mu m} - L_{\text{obs}}^{2-10} \) diagram, we do see that the fraction of these Compton-thick AGNs drops quickly with \( D \) (e.g., right panel of Figure 9), and this trend is clearly correlated with the detection rate distributions seen in Figure 1. This correlation suggests that the detectability of heavily obscured X-ray sources as a function of distance is likely to be the dominant cause for the decreasing trend of the maser detection rates for the overall samples of X-ray-selected AGNs.

4.2.2. The Role of Mid-IR AGN Luminosity

While the decreasing fraction of Compton-thick AGN in the X-ray sample provides an explanation for the declining trend of the megamaser detection rate, it cannot easily explain the enhanced peak in the megamaser detection rate distribution at \( D = 25\text{--}50 \text{Mpc} \), suggesting that the X-ray selection may introduce a secondary distance effect in the search for H2O megamasers. After examining factors that might enhance megamaser detection rates as a function of distance, we note that this peak could be associated with the dependence of the megamaser detection rate on the mid-IR AGN luminosity.

Figure 10 shows the 12 \( \mu \)m AGN luminosities \( L_{\text{AGN}}^{12 \mu m} \) as a function of \( D \) for the sources in the GBT X-ray sample (the magenta dots) and the whole GBT sample regardless of the X-ray detection (the gray dots). The mid-IR AGN luminosities for the GBT sample are obtained in exactly the same way as described in Section 2.2. As expected, the AGN luminosities for both samples increase as a function of distance. However, because of the sensitivity limits, the X-ray surveys tend to pick the most-luminous fraction of AGNs in the GBT sample when the distance is greater than \( D > 25\text{--}50 \text{Mpc} \). Note also that the X-ray surveys beyond 25 Mpc are gradually dominated by sources with \( L_{\text{AGN}}^{12 \mu m} \gtrsim 10^{42} \text{erg s}^{-1} \), whereas the majority of the X-ray sources within 25 Mpc have \( L_{\text{AGN}}^{12 \mu m} < 10^{42} \text{erg s}^{-1} \).

As described in Section 3.1, the megamaser detection rate is an increasing function of \( L_{\text{AGN}}^{12 \mu m} \) when \( L_{\text{AGN}}^{12 \mu m} < 10^{42} \text{erg s}^{-1} \), and it increases by a factor of a few when \( L_{\text{AGN}}^{12 \mu m} > 10^{42} \text{erg s}^{-1} \). Therefore, one can expect that when the fraction of Compton-thick sources is nearly constant within \( D = 50 \text{Mpc} \), the distance distribution of the megamaser detection rate is mainly affected by the mid-IR AGN luminosity, and it is expected to increase by a factor of a few when the distance is greater than \( \sim 25 \text{Mpc} \) because the X-ray sources gradually have \( L_{\text{AGN}}^{12 \mu m} \gtrsim 10^{42} \text{erg s}^{-1} \). Beyond 50 Mpc, the rapidly decreasing fraction of Compton-thick sources starts to affect the detection rate distribution significantly and causes the megamaser detection rate to drop. In the end, as a result of the X-ray detection limit, the X-ray selection gives rise to the quickly increasing trend of mid-IR AGN luminosity within 50 Mpc and the rapidly decreasing Compton-thick fraction beyond 50 Mpc, leading to an enhanced peak at \( D \sim 25\text{--}50 \text{Mpc} \) in the megamaser detection rate distribution and a quick decrease when \( D > 50 \text{Mpc} \).

Finally, the effect of AGN luminosity discussed here could also explain why an early maser survey conducted by Braatz et al. (1996) achieved a relatively high maser detection rate (i.e., \( 11\% \pm 5\% \)) compared with the whole GBT sample for nearby AGNs located within \( \sim 30 \text{Mpc} \). While the majority of the galaxies in the GBT sample within 30 Mpc have

**Figure 10.** The 12 \( \mu \)m AGN luminosity as a function of distance. The magenta and gray dots show galaxies in the GBT X-ray sample and the whole GBT sample of galaxies without X-ray information, respectively.

\[ L_{\text{AGN}}^{12 \mu m} < 10^{41} \text{erg s}^{-1} \] is the majority of AGNs included in Braatz et al. (1996) have \( L_{\text{obs}}^{2-10} > 10^{41} \text{erg s}^{-1} \) (i.e., \( L_{\text{AGN}}^{12 \mu m} > 2.5 \times 10^{41} \text{erg s}^{-1} \), see Falocco et al. 2014). Given that the megamaser detection rate ranges between 4\% and 14\% when \( L_{\text{obs}}^{2-10} \) is greater than \( 10^{41} \text{erg s}^{-1} \) (see Figure 4), it is conceivable that an early maser survey that selected primarily luminous AGNs could achieve a higher detection rate than the general GBT survey for nearby optically active sources.

4.3. Prediction

In Section 3, we showed that the \( L_{\text{AGN}}^{12 \mu m} - L_{\text{obs}}^{2-10} \) diagram provides a convenient way to select AGNs that are simultaneously heavily obscured and mid-IR luminous, for boosting the maser detection rates. This method can be easily applied to a new sample of X-ray sources, and it allows us to predict how many megamasers and disk masers we could discover in the near future.

To establish the new X-ray sample for making such predictions, we first collect all X-ray AGNs from the largest pools of X-ray sources with spectral-fitting results available—the Swift/BAT AGN catalog and the XMM-Newton spectral-fit database—followed by removing all sources that have already been observed by the GBT as part of a maser search. Because of the sensitivity limit of VLBI observations for black hole mass measurements, we only include galaxies within \( z < 0.07 \) (\( \sim 300 \text{Mpc} \)) in our new X-ray sample of galaxies that have never been surveyed for emission in 22 GHz. Because the XMM-Newton catalog does not provide spectroscopic redshifts,

26 The AGN sample in Braatz et al. (1996) was drawn either from the Véron-Cetty & Véron (1991) catalog or from Huchra’s catalog of AGNs (J. P. Huchra 1993, private communication). While the range of \( L_{\text{AGN}}^{12 \mu m} \) for sources in the Véron catalog can be found in Falocco et al. (2014), there are no published X-ray luminosities for Huchra’s AGN sample obtained via private communication. Therefore, for the argument presented in this paragraph, we assume that an early AGN sample such as that of J. P. Huchra (1993, private communication) mainly consists of AGNs as luminous as the sources included in the Véron-Cetty & Véron catalog.
prior to selecting X-ray sources within $z < 0.07$, we cross-match the XMM-Newton catalog with 12 galaxy redshift surveys$^{27}$ to obtain galaxy redshifts and to maximize the number of X-ray sources with associated spectroscopic redshifts.

After assembling all X-ray sources within $z = 0.07$, we collect the broadband UV-to-mid-IR photometry in exactly the same way as described in Section 2. In addition, we perform SED fitting with MAGPHYS to measure the 12 $\mu$m AGN luminosity. The total number of X-ray sources having both spectroscopic redshifts and the photometric data necessary for the SED fitting is 628, with 441 (70%) of them lying within $z = 0.04$ ($\sim$170 Mpc).

In Figure 11, we plot $L_{12}\mu m$ against $L_{60-100}$ for the new sample of non-GBT X-ray sources. The gray and dark gray symbols represent X-ray sources within and beyond $z = 0.04$, respectively. From this plot, we can see that the number of X-ray sources having $N_{H} \gtrsim 10^{24}$ cm$^{-2}$ is substantially smaller for galaxies beyond $z = 0.04$, consistent with the distance effect discussed in Section 4.2.

To achieve high maser detection rates and completeness rates simultaneously for this new sample, we adopt the criterion $N_{H} \gtrsim 10^{23}$ cm$^{-2}$ for sample selection. The number of X-ray sources satisfying this condition is 233, which then corresponds to an expected number of new megamaser and disk maser detections of $35 \pm 5$ and $15 \pm 3$, respectively. Alternatively, we can also adopt the criterion $N_{H} \gtrsim 10^{24}$ cm$^{-2}$ to select galaxies, and this leads to a total number of 96 sources satisfying the condition, corresponding to new detections of $19 \pm 4$ megamasers and $9 \pm 3$ disk masers. For the rest of the selection criteria presented in this paper, we list the predicted numbers of disk masers in Column 10 of Table 4.

We note that in the above predictions we make the assumption that all of the megamasers and disk masers can be detected up to $z = 0.07$. For this assumption to be valid, we require that the integration time for sources between $z = 0.04$ to $\sim 0.07$ be increased by a factor of 4 on average (i.e., $\sim 40$ minutes per source) in order to detect all megamasers above the $3\sigma$ level. Fortunately, the number of heavily obscured AGNs between $z = 0.04$ to $\sim 0.07$ is $\lesssim 48$ ($\sim 21\%$ of the targets), and even, thus if one increases the integration time for these higher redshift sources, the total amount of observing time needed for the 233 X-ray AGNs amounts to only $\sim 63$ hours with the GBT. This exposure time requirement is about a factor of 3 smaller than the typical GBT time awarded to the MCP per year before $\sim 2013$ (200 hr), while the maser detection is at least five times more efficient in terms of the detection rates.

5. Conclusion

To reach the ultimate potential of H$_2$O megamasers for measuring a percent-level $H_0$, as well as for a significant increase in the number of accurately measured SMBH masses, it is important to enhance the efficiency with which we can discover new megamaser disk systems. While the mid-IR color cuts for selecting red/dusty AGNs presented in Kuo et al. (2018) reveal a dramatic increase (at least a factor of 4) in the maser detection rates, this method is compromised by a small maser completeness rate ($\sim 30\%$ for the criteria leading to the highest detection rates) and a small fraction ($< 10\%$) of AGNs complying with the criteria. The low completeness rate results from the fact that the mid-IR emission in the majority ($\sim 60\%$–70\%) of maser galaxies is dominated by the host galaxy component, making these “WISE blue” galaxies easily missed by our mid-IR selection criteria.

In this paper, we demonstrate that by incorporating hard X-ray information in the analysis and performing SED fitting, we can minimize the impact of the host galaxy emission on identifying hosts of water megamasers. Instead of selecting dusty (heavily obscured) AGNs based on mid-IR colors, we use the ratio of the observed X-ray to mid-IR AGN luminosity to infer the obscuring column density $N_{H}$, which can be successfully used as a proxy for the highly sought megamasers systems. This method is particularly useful for discovering heavily obscured AGN candidates from X-ray surveys probing $\lesssim 10$ keV bands. Our analysis shows that the detection rates of megamasers and disk masers primarily depend on $N_{H}$, with the majority (93\%) of them having $N_{H} \gtrsim 10^{23}$ cm$^{-2}$.

Because of the obscured nature of the megamaser systems, selecting galaxies with $N_{H} \gtrsim 10^{23}$ cm$^{-2}$ and $N_{H} \gtrsim 10^{24}$ cm$^{-2}$ leads to impressively high detection rates: $\sim 15\%$–$28\%$ for megamasers and $\sim 6\%$–$13\%$ for disk masers, which are $\sim 5$–10 times higher than the typical rates achieved in the MCP survey, and the compromise in the completeness rates is mild. With respect to the X-ray sample, the maser completeness rates are as high as $\sim 95\%$ and $\sim 50\%$ if one chooses galaxies with $N_{H} \gtrsim 10^{23}$ cm$^{-2}$ and $N_{H} \gtrsim 10^{24}$ cm$^{-2}$, respectively. Making a cut of $L_{12}\mu m \gtrsim 10^{12}$ erg s$^{-1}$ would lead to only a mild decrease in the completeness rates. This shows that X-ray-plus-mid-IR selection is not only effective for significantly boosting the

$^{27}$ These redshift surveys include 1. Galaxy Zoo (Lintott et al. 2008), 2. 2MASS Redshift Survey (Huchra et al. 2012), 3. 2dF Survey (Colless et al. 2001), 4. 6dF Survey (Jones et al. 2009), 5. RC3 catalog (de Vaucouleurs et al. 1991), 6. Galaxy Zoo 2 (Willett et al. 2013), 7. CFA Redshift Catalog (ZCAT; Huchra et al. 1995); 8. CFA25 (Huchra et al. 1999), 9. Dark Energy Survey (Abbott et al. 2018), 10. Point Source Catalog Survey (PSCz; Saunders et al. 2000), 11. Updated Zwicky Catalog (UZC; Falco et al. 1999), 12. SDSS DR 14 (Abolfathi et al. 2018).
The maser galaxy selection methods established in this paper can be easily applied to current and future X-ray AGN surveys for discovering more disk maser systems. When applied to the 628 X-ray sources available in the Swift/BAT AGN catalog and the XMM-Newton spectral-fit database within $z = 0.07$, we predict the detection of $\sim 15$ new disk megamaser systems by new 22 GHz surveys. Although this number is only about 23% of the number of new disk masers ($\sim 70$) needed to achieve a 1% Hubble constant measurement with current instrumentation (e.g., Kuo et al. 2018), it is certain that the number of disk maser candidates will continue to increase as both Swift/BAT and XMM-Newton AGN catalogs keep growing, while future X-ray observatories such as STROBE-X (Ray et al. 2018) will discover more heavily obscured AGN.

Therefore, along with the detection rate enhancement methods proposed by Kuo et al. (2018), it is promising that we will gather enough disk maser candidates to make a 1% $H_0$ determination, along with the very desirable black hole masses at percent-level accuracy, when the Next Generation Very Large Array (ngVLA) is well established and included in the VLBI, which is the main tool used to provide the necessary highly accurate astrometric measurements for these goals.

We gratefully thank the anonymous referee for a thorough review that improved this manuscript. This publication is supported by the Ministry of Science and Technology, R.O.C., under the project 108-2112-M-110-002. A.C. acknowledges support from the National Science Foundation under grant No. AST 1814594. This work has made use of data products from the Wide-field Infrared Survey Explorer (WISE) and the SDSS, and the 2MASS. WISE is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

The Two Micron All Sky Survey (2MASS) is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Funding for SDSS has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS website is http://www.sdss.org/.

**ORCID iDs**

C. Y. Kuo https://orcid.org/0000-0001-6211-5581
D. Pesce https://orcid.org/0000-0002-5278-9221
J. A. Braatz https://orcid.org/0000-0002-1468-9203
Ingyin Zaw https://orcid.org/0000-0002-5208-1426

**References**

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, 
ApJS, 182, 543
Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18
Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42
Akylas, A., & Georgantopoulos, I. 2009, A&A, 500, 999
Antonucci, R. R. J., & Miller, J. S. 1985, ApJ, 297, 621
Asmus, D., Gandhi, P., Höning, S. F., Smette, A., & Duschl, W. J. 2015, 
MNRAS, 454, 766
Asmus, D., Höning, S. F., & Gandhi, P. 2016, ApJ, 822, 109
Asmus, D., Höning, S. F., Gandhi, P., Smette, A., & Duschl, W. J. 2014, 
MNRAS, 439, 1648
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Barvainis, R., & Antonucci, R. 2005, ApJ, 628, 89
Bassani, L., Dadina, M., Maiolino, R., et al. 1999, ApJS, 121, 473
Berney, S., Koss, M., Trakhtenbrot, B., et al. 2015, MNRAS, 454, 3622
Braatz, J. A., Reid, M. J., Humphreys, E. M. L., et al. 2010, ApJ, 718, 657
Braatz, J. A., Wilson, A. S., & Henkel, C. 1996, ApJ, 106, 51
Brightman, M., & Nandra, K. 2011, MNRAS, 414, 3084
Brightman, M., Silverman, J. D., Mainieri, V., et al. 2013, MNRAS, 433, 2485
Cappi, M., Panessa, F., Bassani, L., et al. 2006, A&A, 446, 459
Castanga, P., Panessa, F., Henkel, C., Kadler, M., & Tarchi, A. 2013, MNRAS, 436, 3388
Castanga, P., Surcis, G., Tarchi, A., et al. 2019, A&A, 629, 25
Chang, Y.-Y., Le Floc’h, E., Juncau, S., et al. 2017, ApJS, 233, 19
Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, ApJS, 219, 8
Chitnis, V. R., Pendharkar, J. K., Bose, D., et al. 2009, ApJ, 698, 1207
Colless, M., Dalton, G., & Maddox, S. 2001, MNRAS, 328, 1039
Constantin, A. 2012, JPhCS, 372, 012047
Constantin, A., Green, P., Aldcroft, T., et al. 2009, ApJ, 705, 1336
Corral, A., Georgantopoulos, I., Watson, M. G., et al. 2015, A&A, 576, 71
da Cunha, E., Walter, F., & Smail, I. R. 2015, ApJ, 806, 110
Darling, S. 2000, MNRAS, 318, L71
Dahari, O., & De Robertis, M. R. 1998, ApJS, 67, 249
da Vaucouleurs, G., de Vaucouleurs, A., & Corwin, H. G., Jr. 1991, Third Reference Catalogue of Bright Galaxies, Vol. I, II, and III (New York: Springer)
Elitzur, M., & Shlosman, I. 2000, ApJ, 539, L9
Falco, E. E., Kurtz, M. J., Geller, M. J., et al. 1999, PASP, 111, 438
Falocco, S., Carrera, F. J., Barcons, X., & Miquel, R. 2014, A&A, 568, 15
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, 9
Fiore, F., Puccetti, S., Brusa, M., et al. 2009, ApJ, 693, 447
Gandhi, P., Horst, H., Smette, A., et al. 2009, A&A, 502, 457
Gao, F., Braatz, J. A., Reid, M. J., et al. 2016, ApJL, 817, 128
Gao, F., Braatz, J. A., Reid, M. J., et al. 2017, ApJ, 834, 52
Gebran, J., Bower, G., & Gilbert, K. 2000, A&A, 539, 13
Giovanantonio, I., & Akylas, A. 2010, A&A, 509, 38
Girogantopoulos, I., Rovilos, E., Akylas, A., et al. 2011, A&A, 534, 23
Giozzi, M., Foschini, L., Sambruna, R. M., & Tavecchio, F. 2008, A&A, 478, 723
González-Martín, O., Masegosa, J., Márquez, I., Guerrero, M. A., & Dultzin-Hacyan, D. 2006, A&A, 460, 45
Gonçalves, A. D., Alexander, D. M., Mullaney, J. R., et al. 2011, MNRAS, 411, 1231
Greene, J. E., Seth, A., Kim, M., et al. 2016, ApJ, 826, 32
Greenhill, L. J., Tilak, A., & Madejski, G. 2008, ApJ, 686, 13
Guainazzi, M., Matt, G., & Perola, G. C. 2005, A&A, 444, 119
Gützkow, K., Richstone, D. O., Gehrdt, K., et al. 2009, ApJ, 698, 198
Heckman, T. M., Ptak, A., Hornschemeier, A., & Kauffmann, G. 2005, ApJ, 634, 161
