A Multi-Parameter Statistical Analysis of the Connection Between $\text{H}_2\text{O}$ Maser Emission and Nuclear Galactic Activity

– or –

How to Hunt Masers with Statistics

Emil Christensen

Department of Physics and Astronomy
James Madison University

This work benefitted from collaboration with James Corcoran, Nathan DiDomenico, Robert Roten, Emily Dick and members of the Mega-Maser Cosmology Project (MCP) and from the generous advice and support of Dr. Anca Constantin

Article published on www.jmu.edu/stories (March 2013)
Motivation

> 3500 objects surveyed

150 masers found!

≈ 4% maser detection rate

≈ 2% mega-maser ($L_{H2O} > 10 L_{SUN}$)

< 1% in crucial disk like configuration
### The Literature

We have completed an extensive optical survey of the central region of a total sample of 3231 nearby galaxies with characteristics similar to those of the 486 galaxies contained in the complete spectral atlas of the 486 galaxies contained in the Palomar and SDSS galaxy catalogs (Osterbrock & De Robertis 1985). The main parameters of the emission lines (intensity ratios, fluxes, line widths, and equivalent widths) are measured and tabulated, as are several stellar absorption-line measurements of several stellar absorption-line and continuum lines. We present the results of applying it to our survey, while presents the detailed & Sargent hereafter.

### Other Palomar & SDSS

#### Table 6: Maser sample information

<table>
<thead>
<tr>
<th>Source</th>
<th>Control</th>
<th>Masers</th>
<th>Mega-masers</th>
<th>Disks</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDSS</td>
<td>1181</td>
<td>46</td>
<td>34</td>
<td>7</td>
</tr>
<tr>
<td>Palomar</td>
<td>183</td>
<td>26</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Palomar &amp; SDSS$^a$</td>
<td>25</td>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Other$^b$</td>
<td>0</td>
<td>27</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1339</strong></td>
<td><strong>92</strong></td>
<td><strong>65</strong></td>
<td><strong>15</strong></td>
</tr>
</tbody>
</table>

$^a$ Galaxies with both Palomar and SDSS spectra; in these cases, we adopt the Palomar spectra and measurements.

So what are we looking at again?

- **Line Fluxes**
- **Line Ratios**
  - [O III]/Hβ
  - [N II]/Hα
  - [S II]/Hα
  - [O I]/Hα
- **BPT Diagrams**
- **Stellar Properties**
  - \( D_{4000} \)
  - (Age of stellar population)
  - \( M_{\text{star}} \)
  - (Total stellar mass)
  - \( \sigma^* \)
  - (Stellar Velocity Dispersion)
- **Emitting Gas Parameters**
  - [S II] Ratio (\( N_e \))
    - (\( \lambda \lambda 6716/6731 \) Å)
  - Hα / Hβ
    - (Balmer Decrement)

---

The forbidden lines of the doublet [O I] \( \lambda \lambda 6300, 6364 \) arise from collisional excitation of O\( ^0 \) by hot electrons. Because the ionization potential of O\( ^0 \) (13.6 eV) is nearly identical to that of hydrogen, in an ionization-bounded nebula [O I] is produced predominantly in the "partially ionized zone," wherein both neutral oxygen and free electrons coexist. In addition to O\( ^0 \), the conditions of the partially ionized zone are also favorable for S\( ^+ \) and N\( ^+ \), whose ionization potentials are 23.3 eV and 29.6 eV, respectively. Hence, in the absence of abundance anomalies, [N II] \( \lambda \lambda 6548, 6583 \) and [S II] \( \lambda \lambda 6716, 6731 \) are strong (relative to, say, Hα) whenever [O I] is strong, and vice versa.

In a nebula photoionized by young, massive stars, the partially ionized zone is very thin because the ionizing spectrum of OB stars contains few photons with energies greater than 13.6 eV. Hence, in the optical spectra of H II regions and starburst nuclei, hereinafter H II nuclei, the low-ionization transitions [N II], [S II], and especially [O I] are very weak. [As originally defined by Weedman et al. (1981), a starburst nucleus is one whose current star-formation rate is much higher than its past average rate. Because in general we do not know the star-formation history of any individual object, I adopt the more general designation of H II nucleus.] By contrast, a harder radiation field, such as that of an AGN power-law continuum that extends into the extreme ultraviolet (UV) and X-rays, penetrates much deeper into an optically thick cloud, creating an extensive partially ionized zone and hence strong low-ionization forbidden lines. A hard AGN radiation field also boosts the production of collisionally excited forbidden line emission because its high thermal energy deposition rate enhances the gas temperature.

2.2. Sample Spectra

The spectra shown in Figure 1 illustrate the empirical distinction between AGNs and H II nuclei. In NGC 7714, which has a well-known starburst nucleus (Weedman et al. 1981), [O I], [N II], and [S II] are weak relative to Hα. The [O III] \( \lambda \lambda 4959, 5007 \) doublet is quite strong compared to [O II], [N II], and [S II].
Line Diagnostic Diagram

or

“What’s Under the Hood”

Not all masers are Seyferts

Potential observational bias
Correlation Analysis

Luminosities are highly correlated

Significant maser/non-maser separation
Individual Parameter Comparisons

Hα, [O III] and [O I] luminosity distributions clearly different as we would expect!

High detection rates for certain “goldilocks regions.”
Ok, they’re different. So what?

Detection Rate Inside: 14.0%
Detection Rate Outside: 2.6%
Detection Rate Overall: 6.0%

More than double the detection rate for a simple 2 parameter constraint!
N-parameters Part I: Principal Component Analysis

- A transformation of the variables into new linearly uncorrelated variables called principal components (PCs).
- Each PC attempts to account for as much of the variance as possible.
- Reduces dimensionality and information redundancy.

![Diagram of Principal Component Analysis](https://onlinecourses.science.psu.edu/stat857/book/export/html/11)

![Graphs showing redundancy](http://www.cs.cmu.edu/~elaw/papers/pca.pdf)
PCA Results

**EV’s provide good separation between masers and non-masers**

<table>
<thead>
<tr>
<th>Variable</th>
<th>EV-1</th>
<th>EV-2</th>
<th>EV-3</th>
<th>EV-4</th>
<th>EV-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z$</td>
<td>0.218</td>
<td>-0.448</td>
<td>-0.164</td>
<td>-0.404</td>
<td>-0.320</td>
</tr>
<tr>
<td>$\sigma^*$</td>
<td>0.276</td>
<td>0.182</td>
<td>0.254</td>
<td>-0.008</td>
<td>0.062</td>
</tr>
<tr>
<td>$H\alpha/H\beta$</td>
<td>0.002</td>
<td>0.111</td>
<td>0.664</td>
<td>-0.238</td>
<td>-0.625</td>
</tr>
<tr>
<td>$\lambda 6716/6731$</td>
<td>-0.090</td>
<td>0.146</td>
<td>-0.357</td>
<td>-0.792</td>
<td>0.043</td>
</tr>
<tr>
<td>$\log L \left[ {\text{O} III} \right]$</td>
<td>0.275</td>
<td>-0.620</td>
<td>0.025</td>
<td>0.021</td>
<td>-0.012</td>
</tr>
<tr>
<td>$\log \left[ {\text{O} III}/H\beta \right]$</td>
<td>0.389</td>
<td>-0.337</td>
<td>-0.001</td>
<td>0.147</td>
<td>0.044</td>
</tr>
<tr>
<td>$\log \left[ {\text{N} II}/H\alpha \right]$</td>
<td>0.450</td>
<td>0.282</td>
<td>0.057</td>
<td>-0.022</td>
<td>-0.083</td>
</tr>
<tr>
<td>$\log \left[ {\text{S} II}/H\alpha \right]$</td>
<td>0.442</td>
<td>0.309</td>
<td>-0.178</td>
<td>-0.037</td>
<td>0.034</td>
</tr>
<tr>
<td>$\log \left[ {\text{O} I}/H\alpha \right]$</td>
<td>0.489</td>
<td>0.217</td>
<td>-0.119</td>
<td>0.013</td>
<td>0.036</td>
</tr>
<tr>
<td>$L \left( H_2O \right)$</td>
<td>0.070</td>
<td>-0.113</td>
<td>0.540</td>
<td>-0.358</td>
<td>0.700</td>
</tr>
</tbody>
</table>

**PC-1** ➔ Correlation of the four line ratios confirming the Line Diagnostic Diagram.

**PC-2** ➔ Correlation of redshift and $[OIII]$ luminosity (distant objects are more intrinsically luminous)

**PC-3** ➔ Correlation of $H\alpha/H\beta$ (reddening due to dust) and $H_2O$ luminosity (maser strength)
N-parameters Part II:
Discriminant Analysis

• Finds a linear combination of variables that optimizes the separation of two or more groups

• We desire to construct a Discriminant function which can then be applied to new data

source: Cooley & Lohnes ((1971)
### DA Preliminary Results

#### Table 1. Predictions of the Discriminant Analysis

<table>
<thead>
<tr>
<th></th>
<th>Classified as maser</th>
<th>Classified as non-maser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual maser</td>
<td>17</td>
<td>51</td>
<td>68</td>
</tr>
<tr>
<td>Actual non-maser</td>
<td>11</td>
<td>1297</td>
<td>1308</td>
</tr>
<tr>
<td>Classification Accuracy</td>
<td>25%</td>
<td>99.2%</td>
<td></td>
</tr>
</tbody>
</table>

**Optimal maser detection rate: 60.7%**  
**Maser misclassification rate: 75%**
Conclusions

- Clear separation of maser and non-maser distributions
- Should be able to double or triple the detection rate
- Still fine tuning the DA
- Future work will expand the parameter space into other wavelengths
Thank You!

Questions?
Why We Care About Mega-Masers

The ONLY direct distance measurements

The MOST accurate measurement of black hole mass

Five Steps in the Distance Ladder

1. Radar 1 AU
2. Parallax 3000 ly
3. Cepheids 100 M-ly
4. Tully-Fisher 500 M-ly
5. Supernovae 10 G-ly

\[ M_* = \left( \frac{|v_K|^2 \theta}{G} \right) D_A = \left( \frac{\pi v_1^2}{6.48 \times 10^8 G} \right) D_A, \]

\[ D_a = r/\theta = v_r^2/(a\theta). \]
Attempt to fill the gaps
Co-Investigator on optical spectroscopy proposals for 75 masers

ESO New Technology Telescope (NTT)
La Silla, Chile

Southern Astrophysical Research (SOAR) Telescope
Cerro Pachón, Chile
Good match

Bad match
SDSS Only

![Graph showing distributions of Stellar Age and Stellar Mass for Maser and Control Galaxies.]

- Maser Galaxies:
  - Stellar Age Average: 1.42
  - Stellar Mass Average: 10.1

- Control Galaxies:
  - Stellar Age Average: 1.53
  - Stellar Mass Average: 10.3