Evidence for Black Holes in Green Peas from WISE Colors and Variability

Santosh Harish1, Sangeeta Malhotra1,2, James E. Rhoads1,2, Tianxing Jiang1, Huan Yang3, and Kendrick Knorr1

1 School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA; santosh.harish@asu.edu
2 NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
3 Las Campanas Observatory, Carnegie Institution for Science, Chile

Received 2021 May 7; revised 2022 December 14; accepted 2023 January 22; published 2023 March 16

Abstract

We explore the presence of active galactic nuclei (AGNs)/black holes in Green Pea galaxies (GPs), motivated by the presence of high-ionization emission lines such as He II and [Ne III] in their optical spectra. In order to identify AGN candidates, we used mid-infrared (MIR) photometric observations from the all-sky Wide-field Infrared Survey Explorer (WISE) mission for a sample of 1004 GPs. Considering only >5σ detections with no contamination from neighboring sources in AllWISE, we select 31 GPs out of 134 as candidate AGNs based on a stringent three-band WISE color diagnostic. Using multi-epoch photometry in W1 and W2 bands based on time-resolved unWISE coadded images, we find two sources exhibiting variability in both the WISE bands among 112 GPs with W1 ≤16 mag and no contamination from neighboring sources in unWISE. These two variable sources were selected as AGNs by the WISE three-band color diagnostic as well. Compared to variable AGN fractions observed among low-mass galaxy samples in previous studies, we find a higher fraction (~1.8%) of MIR variable sources among GPs, which demonstrates the uniqueness and importance of studying these extreme objects. Through this work, we demonstrate that MIR diagnostics are promising tools to select AGNs that may be missed by other selection techniques (including optical emission-line ratios and X-ray emission) in star-formation-dominated, low-mass, low-metallicity galaxies.

Unified Astronomy Thesaurus concepts: Emission line galaxies (459); Galaxies (573); Starburst galaxies (1570); AGN host galaxies (2017); Black holes (162)

1. Introduction

Since their discovery more than a decade ago, Green Peas (GPs; Cardamone et al. 2009), a class of low-redshift (z ≲ 0.3) extreme emission-line galaxies, have been studied extensively across the entire electromagnetic spectrum. These galaxies exhibit intense star formation activity, strong nebular lines, low metallicity, low dust content, and high gas pressures (e.g., Izotov et al. 2011; Yang et al. 2016, 2017a; Jiang et al. 2019; Kim et al. 2020). They are also compact in size (Kim et al. 2021) and typically contain low stellar mass. Ultraviolet (UV) studies of GPs have shown the prevalence of Lyα emission lines with an equivalent width distribution matching those of high-redshift (z ≈ 3–6) Lyα emitters (e.g., Henry et al. 2015; Yang et al. 2016, 2017b). They exhibit high [OIII]/[OII] ratios, which are seen in some high-z galaxies, and which may indicate the presence of optically thin ionized regions (e.g., Jaskot & Oey 2013; Nakajima & Ouchi 2014). And there is growing evidence that Lyman continuum radiation escapes more readily in GPs (with escape fractions 2%–72%; Izotov et al. 2016, 2018; Yang et al. 2017b) than in any other known galaxy population, making them the best available low-z analogs of the galaxies that drove cosmological reionization.

GPs are selected on the basis of optical emission lines (generally [OIII]) with extreme equivalent widths. They often exhibit high ionization nebular lines such as HeI 4686 Å and [NeIII] 3869 Å, which suggest a hard ionizing source such as active galactic nuclei (AGNs). However, the relative importance of star formation and accretion in powering GP activity remains poorly known.

Over the last few decades, there has been a growing consensus that most massive galaxies host super-massive black holes (SMBHs) in their central regions (Kormendy & Richstone 1995; Kormendy & Ho 2013). Less is known about BHs among low-mass galaxies, although there has been an increase in low-mass BH candidates in recent times (e.g., Greene & Ho 2007; Reinés et al. 2013). Low-mass galaxies can be crucial in placing constraints on models of BH seed formation and non-merger models of BH evolution (e.g., Volonteri & Natarajan 2009; Van Wassenhove et al. 2010; Greene 2012). Detection of AGN signatures in low-mass galaxies is a significant challenge. A large fraction of these AGNs may be missed by optical selection either because the central engine is obscured, or because emission-line indicators such as [OIII] are contaminated by extreme star formation in the host galaxy (e.g., Goulding & Alexander 2009; Trump et al. 2015). With X-ray selection, the presence of high-mass X-ray binaries and ultra-luminous X-ray sources pose a serious challenge in the detection of low-mass BHs because of their comparable luminosities (e.g., Lehmer et al. 2010; Mineo et al. 2012).

Mid-infrared (MIR) AGN selection has shown promise in identifying some obscured, optically hidden AGNs in several galaxies. The hard radiation field produced by the AGNs can heat dust to high temperatures, generating strong MIR continuum and an IR spectral energy distribution that is distinct from typical star-forming galaxies. The AGN continuum emission is approximately a power law in the 3–10 μm range, because strong UV and X-ray radiation destroys the molecules responsible for the polycyclic aromatic hydrocarbon (PAH) emission, while heating the surrounding grains to near
dust sublimation temperatures (1000–1500 K). With the advent of the Wide-field Infrared Survey Explorer (WISE), a large number of AGNs have been identified, demonstrating that MIR color diagnostics can select luminous AGNs with a 95% reliability (e.g., Stern et al. 2012).

An alternative way to identify AGNs is based on flux variability. Active nuclei are known to be variable in different spectral regimes with timescales ranging from hours to years (e.g., Sesar et al. 2007; Kozłowski et al. 2016). The variability is thought to be related to thermal fluctuations in the accretion disk driven by a turbulent magnetic field (Kelly et al. 2009) and/or temperature fluctuations in an inhomogenous accretion disk (Ruan et al. 2014). Variability-based selection is a promising means for finding massive BHs in dwarf galaxies. Several studies in the past have shown that such a selection can (1) identify AGNs in the low-mass, low-metallicity regime missed by other selection techniques (e.g., Baldassare et al. 2018), (2) identify low-luminosity AGNs (Trevese et al. 1994), and (3) identify AGNs in dwarf galaxies where host galaxy emission dominates the total luminosity (e.g., Trump et al. 2015).

In this paper, we present findings from the first-ever systematic search for AGNs in GPs using MIR observations. Section 2 describes the GP sample and the relevant MIR WISE photometric data. The MIR color and variability diagnostic used for AGN selection is presented in Section 3. The results and comparison with other selection methods are discussed in Section 4. Our main conclusions are presented in Section 5. All magnitudes presented in this work are in the Vega system, unless stated otherwise.

2. Sample and Data

The sample of GP galaxies was selected following Jiang et al. (2019). In total, there are 1004 objects that are spatially compact ($R_{90} < 3''$), which contain well-detected ($S/N > 5$), strong [OII]$_{5007}$ and/or Hβ emission-lines ($EW([OII]_{5007}) > 300\text{\AA}$ and/or $EW(H\beta) > 100\text{\AA}$), and have a literature spectral classification of “galaxy” with a sub-class of “star-forming,” “starburst,” or “NULL.” The literature classifications are drawn from the galSpecLine (Brinchmann et al. 2004; Tremonti et al. 2004) and emissionLinesPort (Maraston et al. 2013; Thomas et al. 2013) catalogs, based on SDSS Data Release 8 (Aihara et al. 2011) and 12 (Alam et al. 2015), respectively. The object classification is based on the measured optical emission-line ratios. However, as discussed in Section 4.1, some of these sources might still contain an AGN that is indistinguishable from a star-forming object based on optical spectroscopy alone.

2.1. MIR Data

The WISE mission (Wright et al. 2010) conducted a MIR imaging survey of the entire sky using 3.4, 4.6, 12, and 22 μm bandpasses (hereafter, W1, W2, W3, and W4), in 2010. With a field-of-view of $47'' \times 47''$ and angular resolution of ~6'' (W1, W2, W3) and ~12'' (W4), WISE scanned the entire sky once every six months with 12 or more independent, single exposures of each point on the sky. The WISE solid hydrogen cryogen was exhausted in late 2010, and the spacecraft was put into hibernation in 2011 February after surveying the full sky twice. The NEOWISE-R Post-Cryo mission (Mainzer et al. 2011) reactivated the spacecraft and has been surveying the sky from 2013 December until present using the two bluer bandpasses, W1 and W2 (which remain usable after cryogen exhaustion).

For the purposes of different AGN selection methods employed in this work (see Section 3), we used slightly different versions of the WISE data for each selection. For color-based selection, since we required the deepest available WISE photometry for W1, W2, and W3 bands, we used photometric measurements from the AllWISE Source Catalog, which combines observations from the WISE cryogenic and NEOWISE post-cryogenic phases, with superior sensitivity and photometric accuracy across all four WISE bands. For variability-based selection, we needed multi-epoch observations of WISE data over long time baselines since we were interested in measuring variations in the MIR photometry only over long timescales (~ months to years). unWISE offers such multi-epoch data where multiple exposures in each epoch are stacked together to produce one coadded image per epoch, resulting in two epochal coadds per year (separated by ~ six months) for a typical sky location. Therefore, we used photometry based on unWISE time-resolved coadd images to select variable GPs. Each of these two choices is now described in detail.

The AllWISE Source Catalog contains the most reliable and accurate MIR photometry for over 747 million objects on the sky based on observations from the WISE cryogenic and NEOWISE post-cryo phases of the mission. This catalog is derived from the AllWISE Atlas intensity images ($1''/7$ pixel$^{-1}$) which consists of coadded 7.7 s W1 and W2 single-exposure images from the 4-Band, 3-Band cryogenic and Post-Cryo mission phases, and coadded 8.8 s W3 and W4 single-exposure images from the 4-Band mission phase. In order to find the MIR counterparts of GPs, we cross-matched them, based on their SDSS coordinates, with the AllWISE Source Catalog using a search radius of 2'' and found matches for 516 GPs from our sample.

By stacking together single-exposure images from the same visit, the unWISE team (Lang 2014; Meisner et al. 2017) has produced time-resolved coadds over a multi-year baseline (Meisner et al. 2018) using WISE and NEOWISE-R observations. A typical sky location is imaged every six months by WISE with $\geq 12$ exposures per visit. In unWISE, these individual exposures (with $\sim 1$ day intervals) are stacked to produce one coadd per band per visit, thereby creating one coadd every six months for a given position on the sky in each of the WISE bands W1 and W2. These coadds offer a powerful new tool for detecting variability of relatively faint WISE sources on long timescales (>0.5 yr) with detection of sources $\sim 1.3$ mag fainter than the single-exposure depth (Meisner et al. 2018).

In order to select a reliable sample of variable sources, we restrict our analysis to a subset of bright GP sources by employing the following criteria. Using the unWISE full-depth source catalog (Schlafly et al. 2019), which is based on significantly deeper ($2\times$) imaging than the AllWISE Source Catalog, we cross-matched it with our GP sample using a search radius of 2'' and found 730 MIR counterparts. Considering the completeness and reliability estimates of the catalog from Schlafly et al. (2019), we select only sources with $W1 \leq 16$ mag, $W2 \leq 15.7$ mag for the variability analysis.

4 https://wise2.ipac.caltech.edu/docs/release/allwise/
31 GPs were selected as AGNs using these criteria. Figure 1 shows the location of GPs in the WISE color–color space and the selected AGN candidates based on these criteria. We note that this selection will mostly include AGNs that are significantly brighter than host galaxy emission. For weaker AGNs, whose continuum emission is diluted by the host galaxy emission, their (W1–W2) color can get bluer such that they become inseparable from normal star-forming galaxies (Stern et al. 2012).

3.2. WISE MIR Variability

Compared to optical variability arising from the accretion disk emission of an AGN, the MIR variability due to emission from the hot dust torus is expected to be smoothed out at shorter timescales because IR variations originate from a much bigger region than the optical variations (Kozłowski et al. 2016). Therefore, in this work, we focus on exploring the long-term variability of GPs. For the 112 GPs in our sample, we performed forced-aperture photometry on the time-resolved unWISE coadds (mentioned in Section 2) and constructed light curves using the following procedure.

For each GP, we determine the unWISE coadd tile containing the GP using the coadd_id column in the source catalog from Schlafly et al. (2019). The Epochal Coadd Index Table from Meisner et al. (2018) contains information about the available coadd epochs for a given coadd_id in each of the bands, W1 and W2. Based on this table, we obtain the coadds for all relevant epochs for each GP with a given coadd_id.

Given the WISE survey strategy, Meisner et al. (2018) found that subtracting pairs of consecutive coadd epochs (with the same coadd_id and WISE band) resulted in dipole residuals suggestive of astrometric misalignments. These residuals are due to the asymmetric PSF models adopted by WISE with respect to swapping the scan direction (which can be forward-pointing/backward pointing corresponding to the scan pointing forward/backward along the Earth’s orbit). In order to redress this issue, Meisner et al. provided “recalibrated” astrometric solutions for each epochal coadd where the WCS calibration was performed using flux-weighted centroids. Before performing photometry, we applied the World Coordinate System (WCS) recalibration solutions given in the Epochal Coadd Index Table to each of the relevant coadd images.

For each GP in our sample, since the coadd images are quite large (∼3.5’ × 3.5’), we generated smaller cutout images (∼1.5’ × 1.5’) for all epochal coadd images, centered around the GP coordinates from SDSS. Subsequently, we performed forced-aperture photometry using Source Extractor (SExtractor; Bertin & Arnouts 1996) in dual-mode with a fake detection for all epochal coadd where the WCS calibration was performed using flux-weighted centroids. Before performing photometry, we applied the World Coordinate System (WCS) recalibration solutions given in the Epochal Coadd Index Table to each of the relevant coadd images.

For each GP in our sample, since the coadd images are quite large (∼3.5’ × 3.5’), we generated smaller cutout images (∼1.5’ × 1.5’) for all epochal coadd images, centered around the GP coordinates from SDSS. Subsequently, we performed forced-aperture photometry using Source Extractor (SExtractor; Bertin & Arnouts 1996) in dual-mode with a fake detection image containing a single Gaussian source whose FWHM is 6’ (∼PSF FWHM of the WISE bands). Our aperture has a diameter of ∼12’ which is roughly two times the PSF FWHM of W1 and W2. We construct the individual light curves in W1 and W2 using the measured aperture photometry from multiple epochs for each GP. We did not apply aperture corrections to this aperture photometry, since we are primarily interested in comparing photometry between epochs to detect variability, which does not depend on correcting aperture flux to total flux.

Following Ai et al. (2010) and Kozłowski et al. (2010), we employed the amplitude of variability ($\sigma_v$) and Pearson correlation coefficient ($r_{12}$) metrics, based on photometry from W1 and W2, to select variable candidates among our GP

3. AGN Selection

3.1. WISE MIR Colors

Several MIR color diagnostics exist in the literature for selecting AGNs, which are mostly based on data from either Spitzer (e.g., Stern et al. 2005; Lacy et al. 2007; Donley et al. 2012) or WISE (e.g., Jarrett et al. 2011; Mateos et al. 2012; Stern et al. 2012). Since we are dealing with observations from WISE, we required our selection criteria to be defined based on WISE data that utilize either one- or two-color criteria to select AGNs. The one-color (W1–W2) criterion, however, suffers from contamination from star-forming galaxies (e.g., Hainline et al. 2016). Therefore, we employed the more stringent two-color criteria by Jarrett et al. (2011) to select AGN candidates from our GP sample. As mentioned in Section 2.1, WISE photometry was available for 516 GPs from the AllWISE Source Catalog; however, only 266 of them had S/R ≥ 5 in W1, W2, and W3 bands. For each of these GPs, given the large point-spread function (PSF) size of the WISE bands, we visually inspected corresponding optical imaging from Legacy Survey (Dey et al. 2019; with a median 5σ depth ~22.5 AB mag in the z-band) for contamination within a radius of 15" around the GP coordinates from SDSS. Subsequently, we performed forced-aperture photometry using Source Extractor (SExtractor; Bertin & Arnouts 1996) in dual-mode with a fake detection image containing a single Gaussian source whose FWHM is 6’ (∼PSF FWHM of the WISE bands). Our aperture has a diameter of ∼12’ which is roughly two times the PSF FWHM of W1 and W2. We construct the individual light curves in W1 and W2 using the measured aperture photometry from multiple epochs for each GP. We did not apply aperture corrections to this aperture photometry, since we are primarily interested in comparing photometry between epochs to detect variability, which does not depend on correcting aperture flux to total flux.

Figure 1. WISE color–color diagram for Green Peas (GPs). Active galactic nucleus candidates selected by the Jarrett et al. (2011) criteria are shown as red diamonds.

After eliminating duplicate sources which are detected in multiple coadds, our final sample contains 112 sources.
respectively, to the photometric uncertainties respectively. Systematic uncertainties are 0.024 mag and 0.028 mag, \( \sigma_{W1}, \sigma_{W2} \). For sources with light curves in W1 and W2, we calculate Pearson’s \( r_{12} \) as

\[
 r_{12} = \frac{C_{12}}{\Sigma_{W1} \Sigma_{W2}},
\]

where \( C_{12} \) is the covariance between W1 and W2 bands for \( N \)-epoch (\( N_{ep} \)) measurements,

\[
 C_{12} = \frac{1}{N_{ep} - 1} \sum_{i} (m[W1]_{i} - \langle m[W1] \rangle) \times (m[W2]_{i} - \langle m[W2] \rangle),
\]

\( \Sigma_{W1} \) and \( \Sigma_{W2} \) are the standard deviations in W1 and W2, respectively,

\[
 \Sigma_{W1} = \sqrt{\frac{1}{N_{ep} - 1} \sum_{i} (m[W1]_{i} - \langle m[W1] \rangle)^2},
\]

\[
 \Sigma_{W2} = \sqrt{\frac{1}{N_{ep} - 1} \sum_{i} (m[W2]_{i} - \langle m[W2] \rangle)^2},
\]

where \( m[W1]_{i}, m[W2]_{i} \) are the W1, W2 magnitudes in the \( i \)th epoch, and \( \langle m[W1] \rangle, \langle m[W2] \rangle \) are the average magnitudes in W1, W2 respectively, for a given GP source. The amplitude of variability, \( \sigma_{m} \), is derived as

\[
 \sigma_{m} = \begin{cases} 
 \sqrt{\Sigma^2 - \epsilon^2}, & \text{if } \Sigma > \epsilon \\
 0, & \text{otherwise}
\end{cases}
\]

where \( \Sigma \) is the photometric variance as defined in Equations (4) and (5), and \( \epsilon \) is the variance due to measurement errors for \( N \) observations in a given band and is defined as

\[
 \epsilon^2 = \frac{1}{N} \sum_{i} \epsilon_{i}^2,
\]

where \( \epsilon_{i} \) is the photometric error of \( i \)th epoch, based on the scatter among individual measurements within that epoch, and \( \epsilon_{s} \) is the systematic error. For W1 and W2, the reported systematic uncertainties are 0.024 mag and 0.028 mag, respectively (Jarrett et al. 2011), which are added in quadrature to the photometric uncertainties (\( \epsilon_{s} \) in Equation (7)).

True variable objects are those with positive variability amplitudes (\( \sigma_{W1}, \sigma_{W2} > 0 \)) and strong correlation between W1 and W2. Four GP sources were identified as potential AGN candidates fulfilling the following criteria: (1) \( N_{ep} \geq 5 \), (2) \( \sigma_{1}, \sigma_{2} > 0 \), and (3) \( r_{12} > 0.6 \). For these sources, we examined the corresponding optical image in SDSS for possible contaminants/blending and found two of these to be affected by neighboring sources within 15") from the GP coordinates.

In addition to the above, we measured the \( \chi^2 \) in both bands (W1 and W2) relative to the best-fitting flux, to test the likelihood of the data under the null hypothesis of a constant source. We found high \( \chi^2 \) and low probabilities (\( p < 0.01 \)) of the null hypothesis in three of the four GPs. With a \( p \)-value of >0.1 in both bands, we were unable to reject the possibility of GP J143202.85+515252.2 being a constant source. We also performed a simple Monte Carlo simulation to determine whether the observed variability amplitudes (\( \sigma_{W1} \) and \( \sigma_{W2} \)) are obtained by chance. For this simulation, we construct the non-varying light curves for each GP as follows: using the mean GP flux and the corresponding observed errors in each epoch, we generate fake WISE photometry across all epochs in both W1 and W2 bands and measure their variability amplitudes. We repeat this procedure for 100,000 realizations. Based on the observed distribution, we calculated the probability of randomly obtaining a variability amplitude (in each band) that is equal to or greater than the observed value and found them to be highly insignificant (\( \gtrsim 1\% \)) except for the GP J143202.85+515252.2.

Eventually, excluding J143202.85+515252.2, we conclude that there are three GP objects that are MIR variable in W1 and W2. Light-curves of GP J024052.20-082827.4 and J120503.54+455150.9 are shown in Figure 2 and their properties are given in Table 1. Additionally, GP J132738.16+132444.5 passed our tests of variability but is among the sources whose WISE photometry is affected by close neighbors.

Recently, Secrest & Satyapal (2020) performed a search for variable AGNs in the MIR using WISE/NEOWISE-R for a sample of dwarf galaxies and found significant variability in 0.09% of their sample. Another recent study (Ward et al. 2022) searched for intermediate-mass BH candidates based on variability, using WISE, in a bigger dwarf galaxy sample and found 0.2% of their sample to contain variable AGNs.
Compared to these previous studies, we find a higher fraction (~1.8%) of variable AGNs in our GP sample. If the parent population of GPs had an 0.2% variability fraction, assuming estimates from Ward et al. (2022), our expectation value would be 0.22 variable objects, and the likelihood of our observing two variable sources would be \( \exp(-0.22) \times 0.22^2/2 \approx 2\% \). The corresponding number assuming the 0.09% variability fraction from Secrest & Satyapal 2020 is smaller, \( \approx 0.5\% \). This suggests (at the \( \sim 2\sigma \) level) that fGPs host AGNs more often than the parent samples of those two studies.

### 4. Discussion

#### 4.1. Limitations of BPT Selection

The classical Baldwin–Phillips–Terlevich (BPT) diagram (Baldwin et al. 1981), \( \log([\text{N}II]/\text{H}alpha) \) versus \( \log([\text{O}III]/\text{H}beta) \), is a widely employed diagnostic to differentiate AGNs from galaxies. Most strong AGNs with solar/super-solar metallicity tend to occupy the “rising branch” on the right-hand side of the diagram (Figure 3). However, low-metallicity (sub-solar or less) AGNs occupy a region on the left-edge of the diagram where they cannot be easily distinguished from star-forming galaxies. Based on chemical evolution estimates from cosmological hydrodynamic simulations, Kewley et al. (2013) showed that low-metallicity galaxies \( (12+\log(O/H) \lesssim 8.4) \) containing AGNs will be impossible to distinguish from low-metallicity star-forming galaxies using the BPT diagram. Around 98% of GP objects have metallicities below this limit. Therefore, the BPT diagram is inefficient at selecting AGN candidates among the GP sample.

#### 4.2. MIR Color-selected AGN Candidates

Different methods of AGN selection can be sensitive to different types of AGNs with varying luminosities, redshifts, line-of-sight angles, obscuration, and so on. Although X-ray detection is considered to be the most robust diagnostic of AGN presence, a significant fraction of AGNs that are selected using IR diagnostics are not detected in X-rays (e.g., Simmonds et al. 2016). Such sources are expected to be significantly obscured and/or low-luminosity AGNs.

MIR color-based selection relies on the fact that hard radiation from an AGN heats the surrounding dusty torus to high temperatures, limited only by the grain sublimation temperature. This produces an approximate power-law continuum in the MIR that is easily distinguishable from stars and typical star-forming galaxies (e.g., Stern et al. 2012; Assef et al. 2013). Since MIR is insensitive to extinction, this selection is able to identify AGNs, especially obscured ones that are optically hidden (e.g., Assef et al. 2013; Stern et al. 2014).

Even though red MIR colors are a strong indication of AGN activity in such galaxies, intense star formation activity in low-metallicity starburst galaxies (such as blue compact dwarfs, which are similar to GPs in many ways) can in extreme cases heat dust to high temperatures that produce extremely red MIR colors (e.g., Griffith et al. 2011; Izotov et al. 2014). The two-color selection that we use for AGN candidate selection is more robust to such cases than are single-color criteria, which are more vulnerable to sample contamination by star-forming galaxies (e.g., Hainline et al. 2016).

While we cannot entirely rule out extreme star formation on the basis of MIR colors alone, it is statistically likely that many of the 31 GPs that meet the Jarrett et al. (2011) color selection criteria are truly AGNs. First, Assef et al. (2013) found a >75% reliability for MIR selection of AGNs with the criteria of Jarrett et al. in a brightness range similar to our GP sample (W2 < 15.73), which suggests \( \sim 24 \) AGN and \( \sim 8 \) star-forming objects among our color-selected candidates.

Second, based on the variability properties of quasars, we expect that our detection of two variable sources implies the presence of additional AGNs that happened not to vary detectably during the monitoring period. Given sensitivity to \( \sim 0.1 \) mag variations over a monitoring timescale of 300–3000 days, the observed quasar variability estimates from MacLeod et al. (2012) imply that \( \sim 30\%–50\% \) of AGNs should show detectable variation in the WISE data we analyzed. Using the Spitzer Deep Wide-Field Survey, Kozlowski et al. (2010) found that \( \sim 15\% \) of AGNs in their sample were sufficiently variable to be selected as a variable AGN. The two AGNs observed to vary in our sample then imply \( \sim 4–15 \) AGNs in total.

### Table 1

| ID          | R.A. (J2000) | Decl. (J2000) | \( \tau_{\text{spec}} \) | \( \sigma_{W1} \) (mag) | \( \sigma_{W2} \) (mag) | Pearson \( r \) |
|-------------|--------------|---------------|--------------------------|-------------------------|-------------------------|-----------------|
| J024052.20-082827.4 | 02:40:52.20   | -08:28:27.4   | 0.0822                   | 0.03                    | 0.07                    | 0.85            |
| J120503.54+455150.9  | 12:05:03.54   | +45:51:50.9   | 0.0654                   | 0.14                    | 0.09                    | 0.96            |

**Note.** Column 1: GP ID, column 2: R.A. (J2000), column 3: decl. (J2000), column 4: spectroscopic redshift from SDSS, column 5: variability amplitude in W1, column 6: variability amplitude in W2, column 7: Pearson \( r \) correlation coefficient for W1 and W2.
that these two variable GPs most likely contain a low-mass, could arise from the gas accretion onto a black hole in an AGN.

4.3. MIR Variable AGN Candidates

Both variable AGN candidates identified in this work are also identified as AGNs by MIR color selection (Figure 4). Considering the location of various classes of objects in the WISE color–color space (as shown in Figure 12 of Wright et al. 2010), all our AGN candidates exhibit colors that coincide with colors typically observed among QSOs/Seyfert objects. Of all the AGN candidates, the variable GPs have the highest [OIII]/[OII] ratio (O32 > 18) and are also among the most metal-poor objects with gas-phase metallicities, 12+log (O/H) < 7.9, as shown in Figure 5.

In addition to the MIR AGN signatures, several other characteristics indicative of AGNs have been observed in the optical spectra of these two variable GPs. Recently, Izotov et al. (2017, 2021) have detected high-ionization emission lines such as [Ne v] λ3426 and He II λ4686, indicating the presence of hard ionizing radiation in these galaxies. In order to determine the physical processes responsible for the hard radiation, they compare observations with models of photoionized H II regions using CLOUDY (Ferland et al. 2013), in combination with ionizing radiation derived from BPASS stellar population models (Eldridge et al. 2017) as well as a combination of stellar radiation from STARBURST99 models (Leitherer et al. 1999) with ionizing radiation produced by shocks or by non-thermal radiation from an AGN. They find that pure stellar ionizing radiation is unlikely to produce the high-ionization emission in those galaxies and attribute the hard radiation to fast radiative shocks with velocities up to ~500 km s^{-1}. However, they also conclude that their observations are consistent with the presence of an AGN source, contributing ~10%–20% to the total luminosity of ionizing radiation in all those galaxies. In addition to the high ionization lines, they also find low-intensity broad components of Hα in the optical spectra for both these variable GPs and this could arise from the gas accretion onto a black hole in an AGN.

Given the observed variability and redder colors in the MIR and the aforementioned optical characteristics, we conclude that these two variable GPs most likely contain a low-mass, low-luminosity AGN, with gas-phase metallicities lower than that of a typical AGN.

5. Conclusions

Using a sample of ~1000 GPs selected from SDSS, based on strong [OIII]λ5007 and/or Hβ emission, we searched for AGN presence based on MIR colors and variability using WISE/NEOWISE-R observations. Considering only >5σ detections in the shortest three WISE bands with no contamination from neighboring sources, 134 GPs were used for AGN selection based on MIR colors. For variability-based AGN selection, 112 GPs with W1 ≤16 and no contamination from neighboring sources were used. Our main findings are as follows.

1. Based on photometry from the shortest three WISE bands, we selected a sample of 31 AGN candidates (out of 134 GPs) using the color–color criteria from Jarrett et al. (2011).
2. Using time-resolved unWISE coadd images, we derived multi-year light curves in W1 and W2 bands for a subset of bright GPs in the MIR. Based on variability amplitude and Pearson’s correlation coefficient (r) metrics, we identified two GPs (out of 112) that exhibit notable variability in the MIR.
3. Both variable GPs are selected as AGNs also based on MIR colors. In addition to the MIR signatures, optical spectra of these variable GPs from previous studies (Izotov et al. 2017, 2021) show evidence of hard ionizing radiation based on the presence of high ionization lines such as [Ne v] and HeII. They also find low-intensity broad component emission in the Hα line in these GPs. Given the variability we observe, these optical signatures are best interpreted as additional indications of gas accretion onto a central BH.

Our findings suggest MIR colors and variability can select AGN candidates in low-mass galaxies that elude detection via other selection methods. Since GPs are extreme emission-line galaxies with high specific star formation rates and low typical metallicities, optical selection methods such as those based on emission-line ratios may be unable to detect the presence of AGNs in GPs, both because optical emission lines may be
dominated by star formation activity, and because the optical line ratio signature of AGNs becomes less distinct from star formation at low metallicity. Multi-wavelength follow-up observations of these sources will be crucial in validating the AGN presence and constraining their properties, including BH mass and accretion rates. If any of these GP are shown not to host an AGN, then an investigation regarding the origin of hot dust and flux variability will be of great interest in itself.

We thank an anonymous referee and Nathan Secrest for their helpful comments and suggestions on an earlier draft, which significantly improved the manuscript. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. WISE and NEOWISE are funded by the National Aeronautics and Space Administration.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics|Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofisica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

This work has made use of the following open-source softwares: SciPy (Virtanen et al. 2020), NumPy (Harris et al. 2020), Matplotlib (Hunter 2007), Astropy (Astropy Collaboration et al. 2013, 2018), DS9 (Joye & Mandel 2003) and TOPCAT (Taylor 2005).
Taylor, M. B. 2005, in ASP Conf. Ser. 347, Astronomical Data Analysis Software, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29
Thomas, D., Steele, O., Maraston, C., et al. 2013, MNRAS, 431, 1383
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Trevese, D., Kron, R. G., Majewski, S. R., Bershady, M. A., & Koo, D. C. 1994, ApJ, 433, 494
Trump, J. R., Sun, M., Zeimann, G. R., et al. 2015, ApJ, 811, 26
Van Wassenhove, S., Volonteri, M., Walker, M. G., & Gair, J. R. 2010, MNRAS, 408, 1139
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Volonteri, M., & Natarajan, P. 2009, MNRAS, 400, 1911
Ward, C., Gezari, S., Nugent, P., et al. 2022, ApJ, 936, 104
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Yang, H., Malhotra, S., Gronke, M., et al. 2016, ApJ, 820, 130
Yang, H., Malhotra, S., Gronke, M., et al. 2017b, ApJ, 844, 171
Yang, H., Malhotra, S., Rhoads, J. E., et al. 2017a, ApJ, 838, 4