









Optimizing Ground-based Observations of O₂ in Earth Analogs

Mercedes López-Morales¹ , Sagi Ben-Ami¹ , Gonzalo Gonzalez-Abad¹ , Juliana García-Mejía^{1,3,4} ,
Jeremy Dietrich^{1,2} , and Andrew Szentgyorgyi¹ 

¹Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA; mlopez-morales@cfa.harvard.edu

²Department of Astronomy and Steward Observatory, University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85719, USA

Received 2019 January 26; revised 2019 April 26; accepted 2019 May 14; published 2019 June 18

Abstract

We present the result of calculations to optimize the search for molecular oxygen, O₂, in Earth analogs transiting around nearby, low-mass stars using ground-based, high-resolution Doppler shift techniques. We investigate a series of parameters, namely spectral resolution, wavelength coverage of the observations, and sky coordinates and systemic velocity of the exoplanetary systems, to find the values that optimize detectability of O₂. We find that increasing the spectral resolution of observations to $R \sim 300,000$ – $400,000$ from the typical $R \sim 100,000$ more than doubles the average depth of O₂ lines in planets with atmospheres similar to Earth's. Resolutions higher than $\sim 500,000$ do not produce significant gains in the depths of the O₂ lines. We confirm that observations in the O₂ A-band are the most efficient except for M9V host stars, for which observations in the O₂ near-infrared (NIR) band are more efficient. Combining observations in the O₂ A, B, and NIR bands can reduce the number of transits needed to produce a detection of O₂ by about one-third in the case of white noise limited observations. However, that advantage disappears in the presence of typical levels of red noise. Therefore, combining observations in more than one band produces no significant gain versus observing only in the A band, unless red noise can be significantly reduced. Blending between the exoplanet's O₂ lines and telluric O₂ lines is a known problem. We find that problem can be alleviated by increasing the resolution of the observations, and by giving preference to targets near the ecliptic.

Key words: methods: observational – planets and satellites: atmospheres – planets and satellites: composition – planets and satellites: detection – planets and satellites: terrestrial planets – techniques: radial velocities

1. Introduction

The appearance of molecular oxygen, O₂, in large amounts in Earth's atmosphere has been linked for decades to the onset of oxygenic photosynthesis by early organisms (see an overview in Catling & Kasting 2017). The *Galileo* probe observations by Sagan et al. (1993), where they observed strong absorption features of telluric O₂ and methane, which departed from local thermodynamical equilibrium, drew further attention to those gases as a strong indicator of life in our planet.

The idea of searching for O₂ in terrestrial planets around nearby stars has gained additional momentum in the past few years. Given the recent discoveries of Proxima b (Anglada-Escudé et al. 2016), the Trappist-1 planets (Gillon et al. 2017), LHS 1140 b (Dittmann et al. 2017), Ross 128 b (Bonfils et al. 2018), and the expected outcome of the *Transiting Exoplanet Survey Satellite* (TESS) mission (Ricker et al. 2016)—predicted to find about 10 planets with radii smaller than two Earth radii within or near the habitable zone of nearby stars (Barclay et al. 2018)—finding the best way to detect O₂ in the atmosphere of those planets is becoming a pressing task. There is also known abiotic pathways to produce O₂ on Earth-like planets (e.g., Domagal-Goldman et al. 2014; Wordsworth & Pierrehumbert 2014; Luger & Barnes 2015). Those potential scenarios enhance the importance of detecting O₂ in Earth analogs, so we can start to understand how diverse the atmospheric characteristics of Earth analogs truly are.

Snellen et al. (2013) first proposed to use high-resolution ($R = 100,000$) Doppler shift observations centered on the O₂ A-band at $0.76 \mu\text{m}$, and cross-correlation techniques to detect O₂ in the atmosphere of transiting Earth twins. They concluded this could be achieved with instrumentation on the upcoming generation of Extremely Large Telescopes (ELTs). Particularly, they concluded that the 39 m European ELT would be able to detect this O₂ signal in the transmission spectrum of an Earth-like planet transiting an M5V star, 12 pc away, by combining the observations of 30 transits. They also pointed out that for planets around M7V stars or later, it would be easier to detect O₂ by observing around the O₂ band centered at $1.27 \mu\text{m}$, where late M dwarfs are brightest.

Rodler & López-Morales (2014, hereafter RLM14) refined the calculations of Snellen et al. (2013) by producing simulated observations that included the effects of Poisson noise and correlated (red) noise, typically introduced by instrumental and atmospheric effects, and by also accounting for atmospheric refraction effects in the exoplanet's transmission spectrum (García Muñoz et al. 2012). In addition, RLM14 explored the expected performance of high-resolution spectrographs planned for the three upcoming ELTs: the Giant Magellan Telescope (GMT), the Thirty Meter Telescope (TMT), and the European Extremely Large Telescope (ELT), including a range of spectral resolutions determined by slit width and a range of stellar spectral types and their velocities with respect to Earth. They concluded that, although still feasible, the number of transits needed to detect O₂ in an Earth twin around a nearby M dwarf was two times larger than predicted by Snellen et al. (2013) after accounting for red noise. They also found that the O₂ A-band at $0.76 \mu\text{m}$ is better than the O₂ near-infrared (NIR) band at $1.27 \mu\text{m}$ to detect the exoplanet's signal for all M-dwarf

³ National Science Foundation Graduate Research Fellow.

⁴ Ford Foundation Fellow.

spectral types, except M9V. This is because instrumental and atmospheric noise levels are typically larger in the NIR. RLM14 also found that blending effects between telluric O₂ lines and the exoplanet's O₂ lines limit the number of targets well suited for this type of observation. This problem appeared most significant in the case of systems with relative velocities of less than $\pm 15 \text{ km s}^{-1}$ and around $\pm 50 \text{ km s}^{-1}$ with respect to Earth, where 50% or more of the exoplanet O₂ lines appeared blended.

More recently Serindag & Snellen (2019) generated more realistic simulations of noise effects by injecting a simulated O₂ signal on to real time series observations of Proxima Centauri with the Ultraviolet and Visual Echelle Spectrograph (UVES) instrument on the VLT. By rescaling the flux of Proxima Centauri collected by the VLT to the flux of a hypothetical M5V star 7 pc away hosting an Earth twin, they find that O₂ can be detected in 20–50 transits. That result is consistent with the predictions by Snellen et al. (2013) and RML14.

RLM14 looked into what would be the optimal spectral resolution of a standard high-resolution slit spectrograph to detect O₂ and found it to be $R = 80,000$ once slit losses and the smearing of the signal as resolution decreases are taken into account. They also suggested that if more starlight could be fed into a narrow slit using pre-slit optics, the efficiency of observations at higher resolutions can greatly improve (see Figure 3 in RLM14). Here we show how the efficiency of the observations can be improved even further by increasing the spectral resolution of the signal to values that optimize the depth of the O₂ spectral lines, making the signal easier to detect.

In this paper we explore what are the best values for a set of instrumental and observational parameters to optimize the detectability of O₂ in Earth-like exoplanets. Those parameters are: spectral resolution (Section 2), combination of absorption bands (Section 3), and the most suitable set of targets to minimize/circumvent the line blending problem between exoplanetary and telluric O₂ lines (Section 4). Sections 5 contain the summary and discussion of our results.

2. Spectral Resolution

In Earth's atmosphere O₂ extends over a large range of atmospheric heights, from the surface of the planet to altitudes of 85–100 km (see, e.g., Misra et al. 2014), therefore the characteristic shape of the O₂ lines is produced by absorption of stellar photons over a wide range of atmospheric temperatures and pressures. Given our current one-planet sample, we do not know how typical the distribution of O₂ observed in the Earth atmosphere is, but for the purpose of our simulations we assume that Earth-like exoplanets have a similar vertical distribution of O₂ as Earth. Each O₂ spectral line has a finite width and a depth, which are determined by the effect of two broadening mechanisms: pressure broadening in the high-pressure layers of the atmosphere, and Doppler broadening in the lower pressure layers (see, e.g., Chance & Martin 2017). The combination of those two broadening mechanisms defines a minimum width and a maximum depth for the spectral lines, beyond which instrumentation with resolving power larger than the intrinsic broadening of the lines will not improve our capacity to detect those O₂ lines.

We performed a series of simulations to establish at what resolution O₂ lines appear fully resolved in the transmission spectrum of a planet like Earth. We simulated spectra with

resolutions of $R = 10^5$ – 10^6 , in 10^5 steps using the reference forward model (RFM⁵; Dudhia 2017), over a wavelength range between 0.57 and 1.35 μm , which includes four O₂ bands observed in the optical and NIR transmission spectrum of Earth's atmosphere: the α band at 0.63 μm , the B band at 0.69 μm , the A band at 0.76 μm , and the NIR band at 1.27 μm . In our simulations the RFM is driven by the HITRAN 2012 spectral database (Rothman et al. 2013). We used the US standard atmosphere (1976)⁶ to simulate temperature, pressure, and molecular oxygen and water vapor profiles from the surface up to 100 km at 2 km intervals (or layers). To produce the total transmission spectrum of the exoplanet atmosphere, T_T , we integrated the transmission of each layer, T_n , as

$$T_T = \frac{\sum T_n A_n}{A_T}, \quad (1)$$

where A_n is the cross-section area of each layer as seen from Earth and A_T is the total cross-section area of the atmosphere, i.e., $A_T = \sum A_n$. In this integration, we accounted for the effect of refraction in the bottom layers of the atmosphere, as described in RLM14. Specifically, we did not include the first three model layers, between 0 and 6 km from the exoplanet's surface, which are the layers affected by refraction in planets in the habitable zone of early M dwarfs (García Muñoz et al. 2012). We made the same assumption for all host-star spectral types, although refraction might have a lesser effect in late-type M dwarfs (see, e.g., Bétrémieux & Kaltenegger 2014).

We show sample results of these simulations in the top panels of Figures 1 and 2, where we compare transmission spectra in the 0.76 μm A band and the 1.27 μm NIR band at different resolutions. The bottom panels in each figure show how the integrated depth of the O₂ lines increases as a function of resolution for each of those two bands. The integrated line depth for each resolution is calculated by first identifying all the lines in each spectra with depth larger than 5% with respect to the local continuum, and then adding their individual depths. Each line depth is calculated as the minimum of a Gaussian fit to the line.

The vertical axes in the bottom panels of Figures 1 and 2 show the normalized integrated line depths at each resolution, I_R , using the integrated line depth at $R = 1 \times 10^5$, I_{10^5} , as the zero point and normalizing by the integrated line depth at $R = 1 \times 10^6$, I_{10^6} . That is

$$\text{Norm. Integrated Depth} = \frac{(I_R - I_{10^5})}{(I_{10^6} - I_{10^5})}. \quad (2)$$

Those figures show how we gain between 100% and 150% in line depths by increasing the instrumental resolution from $r = 1 \times 10^5$ to $r = 3 \times 10^5$ – 4×10^5 . Given the multiplicative nature of cross-correlation functions (ccfs), an increase in the depth of the signal translates into an increase in the strength of the ccfs peaks, making the signal easier to detect. The O₂ lines appear fully resolved at resolutions larger than about $r = 8.5 \times 10^5$, and increasing instrumental resolution past

⁵ RFM is a GENLN2-based (Edwards 1992), general purpose line-by-line radiative transfer model originally developed at Oxford University under an ESA contract to provide reference spectral calculations for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) instrument (Fischer et al. 2008).

⁶ https://ccmc.gsfc.nasa.gov/modelweb/atmos/us_standard.html

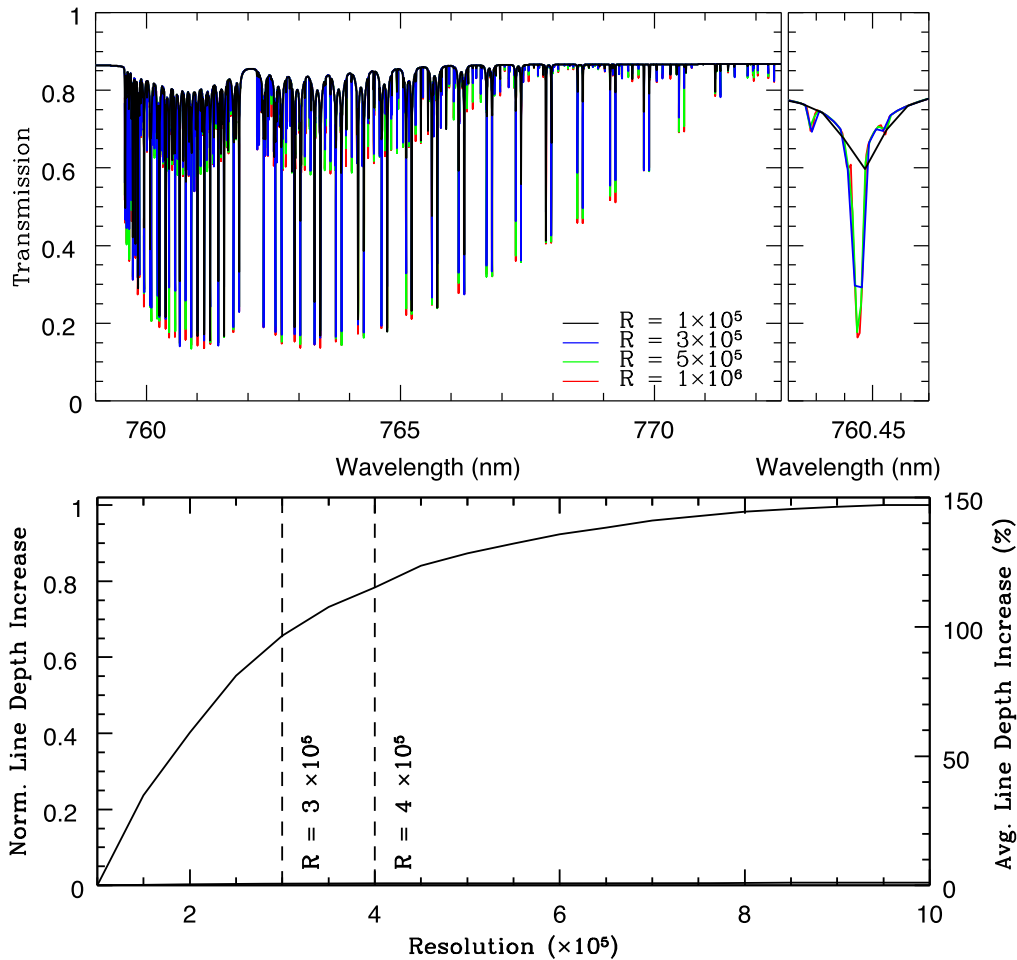


Figure 1. Top: $0.76 \mu\text{m}$ O_2 band at four different resolutions: $R = 100,000$ (black), $300,000$ (blue), $500,000$ (green), and $1,000,000$ (red). The top right panel shows a zoom-in of one of the lines in the band. Bottom: increase in integrated depth of the O_2 lines as a function of resolution, with respect to $R = 100,000$ and normalized to the integrated depth at resolution $R = 1,000,000$.

$r = 4 \times 10^5$ only produces about a 20% additional gain in line depths. The same calculations for the $\text{O}_2 \alpha$ band at $0.63 \mu\text{m}$ and the B band at $0.69 \mu\text{m}$ produce similar results.

3. Combination of O_2 Bands

As shown in the original detections of CO in hot Jupiters by Snellen et al. (2010), Brogi et al. (2012), and Rodler et al. (2012), our ability to detect a particular molecular species in an exoplanet atmosphere using high-resolution spectroscopy is enhanced by the capacity of cross-correlation techniques to simultaneously use information from multiple absorption lines, N_l . Work done on hot Jupiters so far, and also the simulation studies done on the detectability of O_2 in Earth analogs (Snellen et al. 2013; Rodler & López-Morales 2014; Serindag & Snellen 2019), have focused on individual bands of specific molecules. However, given that the strength of the ccf signal increases as $\sqrt{N_l}$ (see Equation (1) in Snellen et al. 2015), it might be advantageous to observe multiple bands simultaneously, or to combine observations in different bands. In this section we explore whether multiband observations can improve our ability to detect O_2 . Of the four O_2 bands between $0.6 \mu\text{m}$ and $1.3 \mu\text{m}$, the A band ($0.76 \mu\text{m}$) is the deepest, followed by the B band ($0.69 \mu\text{m}$) and the NIR band ($1.27 \mu\text{m}$) with similar depths. The α band ($0.63 \mu\text{m}$) is significantly weaker than the other three, therefore we do not use that band

in the rest of this study. Using the A , B , and NIR bands, we produce as realistic as possible simulations for different combinations of those bands, assuming exoplanet host stars of different spectral types and typical noise levels. In the simulations, we adopt a spectral resolution of $R = 3 \times 10^5$, based on the results in Section 2, and place the hypothetical planetary system at a fixed distance of 5 pc from Earth.

3.1. Generation of Model Spectra

We generated model spectra following the same methodology described in RLM14 using their Equations (1) and (2) which we reproduce here. The model spectra C are given by

$$C = (a(1 + \epsilon^{-1})^{-1}T) \otimes G, \quad (3)$$

with

$$a = (1 + v_*c^{-1})S + (1 + (v_* + v_{\text{pl}})c^{-1})\epsilon^{-1}P, \quad (4)$$

where S is the spectrum of the host star, P is the transmission spectrum of the planet, T is the telluric spectrum of Earth, and G is the instrumental profile of the instrument used to perform the observations. The other parameters in the equations are the speed of light, c , the instantaneous radial velocity of the planet with respect to its host star, v_{pl} ($\sim 0 \text{ km s}^{-1}$ at primary transit), the relative velocity of the system with respect to Earth, v_* , and

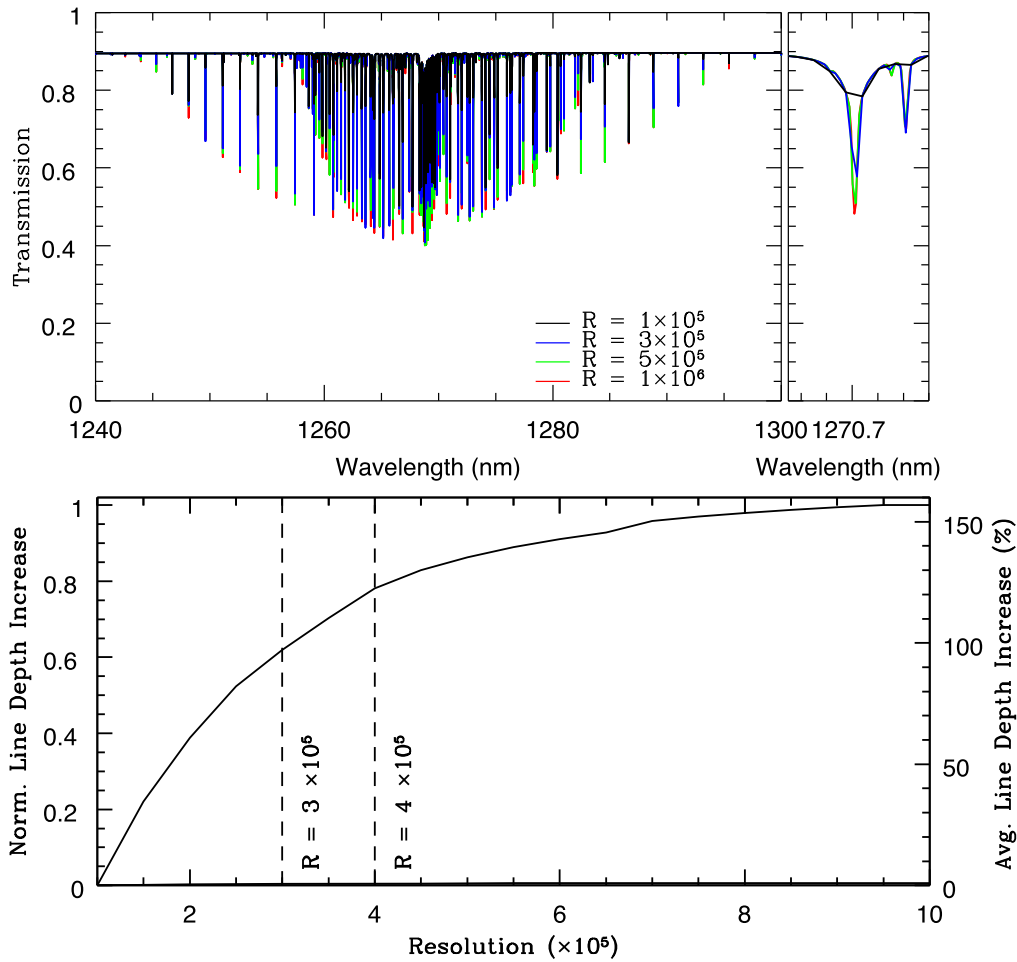


Figure 2. Same as Figure 1, but for the $1.27 \mu\text{m}$ O_2 band.

the ratio of areas between the stellar disk and the atmospheric ring of the planet, ϵ .

For the host-star models, we focused on M-dwarf stars following previous studies (Kaltenegger & Traub 2009; Snellen et al. 2013; Rodler & López-Morales 2014), which show that the exoplanet atmospheric signal is larger for planets around late-type stars. This is because the amplitude of those signals is proportional to the ratio of the areas of the stellar disk and the ring subtended by the atmosphere of the planet. In addition, observing Earth-like planets around M-dwarf stars has the advantage that a large number of transit observations, needed to detect atmospheric features, can be obtained in a relatively short period of time.

We used three stellar emission models with spectral types M1V, M4V, and M9V to explore the effect of relative stellar fluxes between the three O_2 bands. For example, in the case of a planet orbiting a M1V star, the ratio of stellar fluxes between the *B* and NIR O_2 bands will be much smaller than in the case of a planet orbiting a M9V star, as illustrated in Figure 3. More specifically, we used stellar models from the Husser et al. (2013) high-resolution spectral library with effective temperatures of 3600 K (M1V), 3000 K (M4V), and 2300 K (M9V). For all of the models we assumed solar abundances and $\log g = 4.5$ dex. For the telluric spectrum we used the spectrum described in Section 2.1.2 of RLM14.

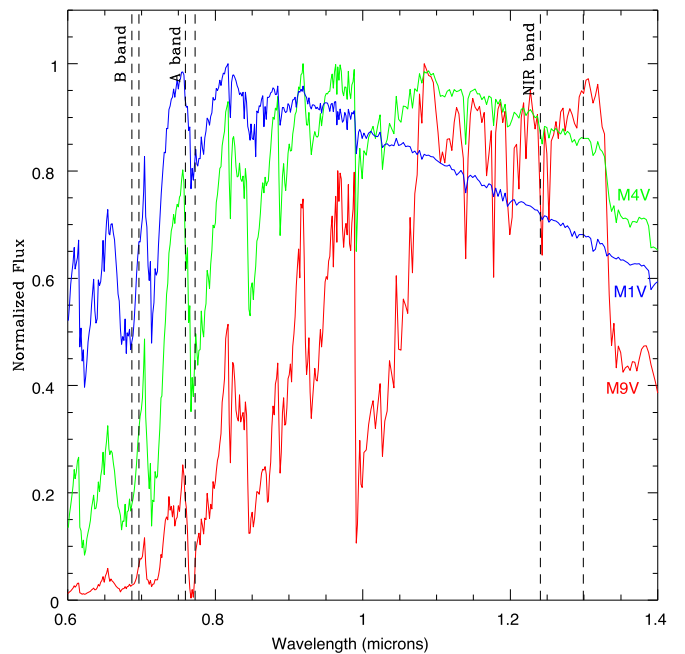


Figure 3. Husser et al. (2013) model emission spectra for M1V, M4V, and M9V stars. The resolution of the models has been binned down to $R = 3000$ for a clearer display. The vertical dashed lines highlight the location of the O_2 *B*, *A*, and NIR bands.

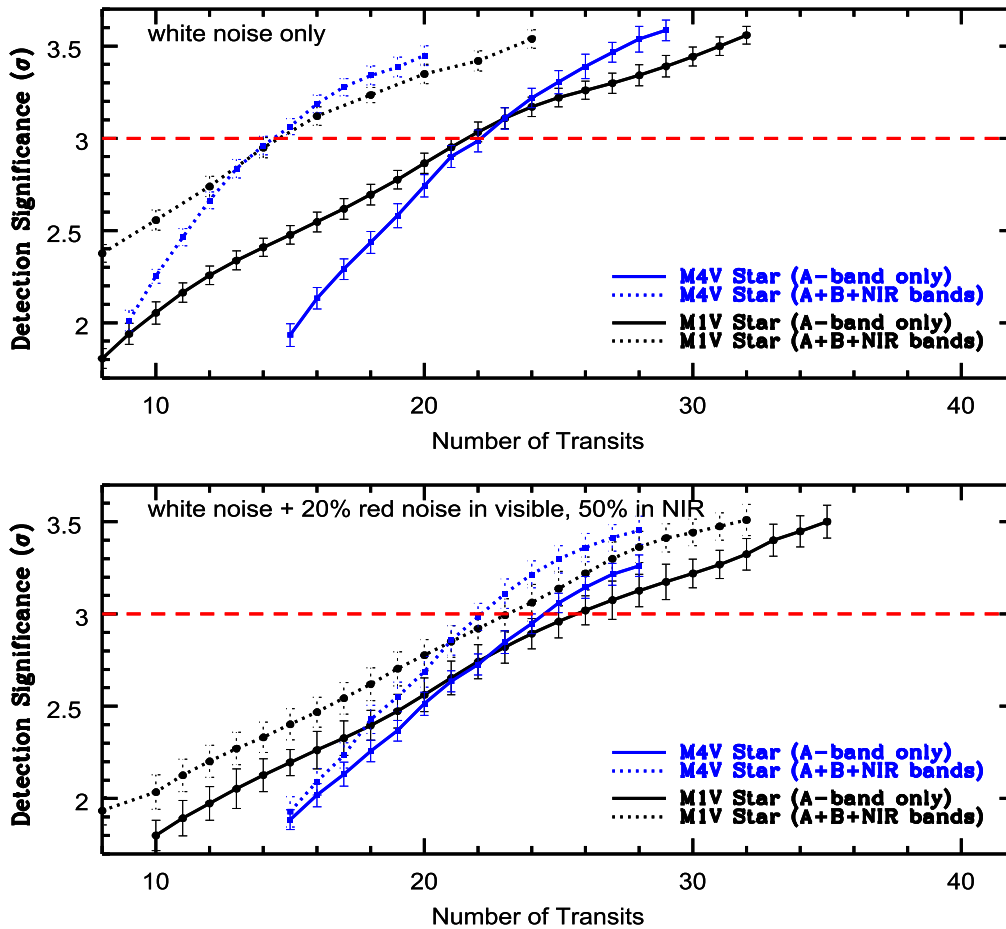


Figure 4. Number of transits needed to reach a 3σ detection using only the O_2 A-band (solid lines) vs. combining the A, B, and NIR bands (dashed lines). Blue is for a M4V host star. Black is for a M1V host star. The top panel only includes white noise. The bottom panel includes 20% red noise in the optical bands and 50% in the NIR band.

The exoplanet spectrum was generated as described in Section 2. For the instrumental profile G we assumed a Gaussian with a FWHM equal to the instrument’s spectral resolution (in this case $R = 3 \times 10^5$). We also assumed a fixed system velocity of $v_* = 20 \text{ km s}^{-1}$, and ϵ values of 125,000, 35,000, and 4000, respectively, for the M1V, M4V, and M9V host-star cases as listed in Table 1 of RLM14. Finally, we explored the ideal case in which only white (Poisson) photon noise is present in the observations, and the more realistic case in which the observations contain some level of red (correlated) noise in addition to white noise. Specifically, we used typically measured red noise levels of 20% in the visible (Pont et al. 2006) and 50% in the NIR (Rogers et al. 2009).

3.2. Simulations Setup

We ran simulations aimed at answering the following two questions. (1) Does combining observations in the A, B, and NIR bands significantly reduce the number of transits needed to produce a O_2 detection? (2) Does the optimal band to detect O_2 depend on the spectral type of the host star? The answers to those questions are illustrated in Figures 4 and 5.

To generate Figures 4 and 5, we estimated the photon flux per transit assuming transit durations of 4.0 hr, 2.1 hr, and 0.43 hr for an Earth-like planet in the middle of the habitable zone of a M1V, M4V, and M9V star. To compare our results with those in RLM14, we assumed that the observations are performed with

an instrument similar to G-CLEF on the GMT (Szentgyorgyi et al. 2012), and assumed duty cycles of 0.93 for the M1V star, and 0.98 for the M4V and M9V stars. These duty cycles were computed using the most recent exposure time calculator (ETC) for G-CLEF. In the new ETC, observations of the M1V star system have a higher duty cycle than the 0.86 value used in RLM14. The duty cycles for the M4V and M9V stars remained the same. The calculations also assume that the same amount of flux is collected from the star at $R = 3 \times 10^5$ and $R = 1 \times 10^5$,⁷ and use a readout (dead) time of 7 s and a readout noise of 5 electrons per exposure, based on the most up-to-date values for those parameters. Like in RLM14, we assumed a dead time of 10 s and a readout noise of 10 electrons for a hypothetical detector in the NIR. We also set a maximum exposure time of 600 s (200 s in the case of the M9V host star), so the radial velocity shift of the remote planet does not smear out the absorption lines in its atmospheric spectrum.

Finally, to estimate uncertainties in the achieved detection levels (or equivalently, in the number of transits needed), we generated 1000 model spectra with randomized noise values and used those as a pool of transit spectra for a bootstrap analysis. In that bootstrap analysis we simulated transits by

⁷ In other words, we assume no additional slit losses following the logic that this can be achieved using pre-slit optics as discussed in RLM14. This assumption allows us to compare how the efficiency of O_2 observations would improve with resolution.

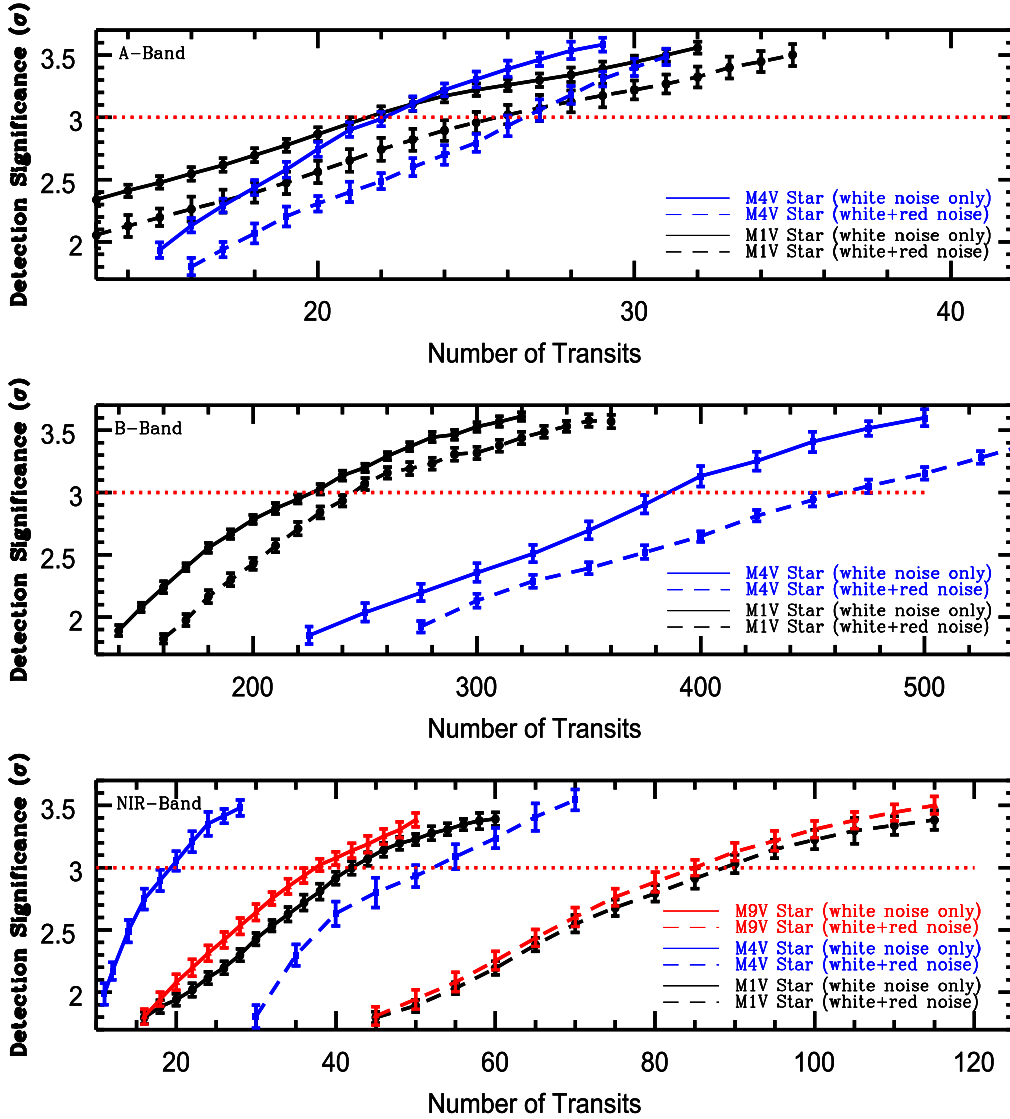


Figure 5. Number of transits needed to reach a 3σ detection using only the O_2 A-band (top panel), B band (middle panel), and NIR band (bottom panel), assuming host stars of spectral types M1V, M4V, and M9V. The solid lines show results when we only include white noise in the simulations. The dashed lines show results when we include red noise. The simulations using a M9V host star produce very large numbers of required transits in the case of the A and B bands, and are therefore not included in those panels.

randomly selecting N spectra from the 1000 models pool, where N is the number of in-transit observations, and produced an observed planet spectrum as $P_{\text{obs}} = C_{\text{in}}/C_{\text{out}}$, where C_{in} is the integrated in-transit spectra and C_{out} is the integrated spectra observed out of transit. C_{out} is produced by integrating N random spectra of the star and tellurics alone, and also with randomized noise values. P_{obs} is then cross-correlated with a noiseless, template exoplanet transmission spectrum. We determine the detection significance level σ by examining the significance of the peak in the ccf at $v_* = 20 \text{ km s}^{-1}$, which is the input velocity of the system in the models. We repeated this process 500 times for each set of number of transits in Figures 4 and 5. The error bars of each data point in those figures correspond to the 1σ standard deviations of those 500 iterations.

3.3. Simulations Results

Figure 4 illustrates the number of transits necessary to produce a 3σ detection of oxygen for a planet around a M1V

and a M4V star when only using the O_2 A-band versus using the A, B, and NIR bands combined. The case of a planet around a M9V star requires over 200 transits to produce a 3σ detection and therefore has not been included in the plot. In the white noise only case, we find that we need 22 ± 1 transits to produce a 3σ detection of oxygen when only using the A band. We obtain the same number of transits for both stars, consistently with the same simulation in RLM14. However, the total number of transits needed is 34% less than in RLM14 because of the higher spectral resolution of our simulations. We find that combining the A, B, and NIR bands is advantageous in the case where the observations are only affected by white noise, with the required number of transits dropping to only 14 ± 1 when combining the three bands. That represents a $36\% \pm 5\%$ improvement in the number of transits compared to the results using only the A band. When we include typically observed levels of red noise in the simulations, we find that the number of transits needed to produce a 3σ detection of oxygen using only the A band is 24–26 transits, slightly larger than in

the white noise only case, which is expected. However, in this case the number of transits needed to produce a 3σ detection of oxygen does not significantly decrease when combining observations in the three O₂ bands. In fact, doing so only reduces the number of transits by $10\% \pm 8\%$. Therefore, there is no significant advantage in combining observations in the A, B, and NIR bands versus observing only in the A band when typical levels of red noise are present. It is worth mentioning that we also find that the contribution of the B band is very small (as can be also inferred from Figure 5), and including it in combination with the A and NIR bands only improves the number of transits by 0.5 versus the results from only combining the A and NIR bands. We show the three bands in Figure 4 for completeness.

Figure 5 shows the number of transits necessary to produce a 3σ detection using individual O₂ bands for planets orbiting M1V, M4V, and M9V host stars. The simulation results for the M9V host-star case have not been included in the A-band and B-band panels because they yield impracticably large numbers of transits. The reason for this is the very low amounts of flux M9V stars emit at those wavelengths, as illustrated in Figure 3.

The main conclusion from Figure 5 is that, in the most realistic case of data sets with typical levels of red noise, observations in the A band will be the most efficient for all low-mass host stars except for M9V stars, since for those stars the flux emitted in the A band is practically zero (see Figure 3). For planets orbiting around M9V stars, O₂ searches are only feasible using the NIR band, but will require scores of transits, as illustrated in the bottom panel of Figure 5. The middle panel in the figure shows how it will be necessary to observe several hundreds of transits to produce a 3σ detection of O₂ in the B band. This is both because the B band is shallower than the A and NIR bands, and also because M dwarfs emit less flux in that band than at the wavelengths where the A and NIR bands are located (see Figure 3). Therefore, the B band is not well suited for observations of O₂ in planets around M-dwarf stars.

Finally, we elaborate on the results in the bottom panel of Figure 5 for the NIR band. Because our simulations assume a fixed distance of 5 pc for all the systems, the amount of flux received from each system correlates with the spectral type of the host star. This, combined with the emission spectrum of each star (Figure 3), yields some interesting results. We find that, for the case of an M4V host star, observations in the NIR could be slightly more efficient than observations in the A band if it were possible to obtain observations free of red noise (see the solid blue lines in the top panel of Figure 4 and the bottom panel of Figure 5). We also find that, if two systems hosted by a M1V and a M9V star were found at the same distance from Earth, both would require about the same number of transits in the NIR band to produce an O₂ detection. This is because the extra flux of the M9V host star at those wavelengths would get compensated by the higher intrinsic brightness of the M1V host star. However, it would still be about 3.3 times more efficient for the system around a M1V host star to observe it with an instrument in the visible covering the O₂ A-band, as illustrated in the top panel of Figure 5.

4. Relative Velocity and Sky Coordinates

High-resolution, Doppler shift searches for O₂ (or any other molecule) in Earth-like planets take advantage of the relative velocity of the exoplanet system with respect to Earth, v_* . Given high enough resolution, the O₂ lines from the exo-Earth

can be separated from the telluric O₂ lines if they appear sufficiently Doppler shifted. One obvious conclusion from this method is that it will not work for exoplanetary systems with $v_* = 0$, because the exoplanet's O₂ lines will fully overlap with telluric O₂ lines. However, as first pointed out by RLM14, this line blending problem is not unique to $v_* = 0$ km s⁻¹ and there is other relative velocities at which blending effects will be significant, given the high density of the O₂ lines.

In this section we refine the blending analysis by RLM14 for the O₂ A-band and NIR band using a finer grid of relative velocities ($\Delta v_* = 0.5$ km s⁻¹) and also considering a larger number of O₂ lines. In addition, we update the blending criteria used in RLM14 by counting all of the exoplanet lines with a separation from a telluric line of less than one resolution element as blended. The reason for excluding those lines is that, even though some of them would still appear in the exoplanetary spectrum after subtracting the telluric lines, the signal-to-noise level in those portions of the exoplanet spectrum would be too low to detect those lines. Finally, we extend the line blending analysis here to resolutions of 3×10^5 and 5×10^5 , following the conclusions in Sections 2 and 3 that higher resolutions increase the depth of the O₂ lines and therefore their detectability.

Figure 6 shows the normalized fraction of blended lines as a function of the relative velocity of the exoplanetary system with respect to Earth, for the O₂ A-band and NIR band. The figure, which shows blending fractions for resolutions of 1×10^5 , 3×10^5 , and 5×10^5 , illustrates how there will always be a non-zero level of blending between the exoplanet lines and the telluric lines. Blending is higher in the A band than in the NIR band, because of the higher density of lines. However, as shown in Sections 2 and 3, the A band is still the best suited for O₂ observations for the majority of systems, because of deeper absorption lines, and higher stellar emission levels for all host stars except M9Vs.

Focusing on the A band, our results agree overall with RLM14. Systems with relative velocities between about ± 15 – 30 km s⁻¹ are the least affected by line blending. The figure also shows how increasing the spectral resolution reduces the level of blending. In addition, systems with negative relative velocities with respect to Earth will be on average less affected by blending than systems with positive relative velocities. This is because of the asymmetric shape of the O₂ A-band, and also of the individual lines. Also this effect increases with resolution. To quantify this effect we calculated the average blending between -150 km s⁻¹ and -15 km s⁻¹, and between $+15$ and $+150$ km s⁻¹ for the three blending functions shown in the top panel of Figure 6. The results show that, for $R = 5 \times 10^5$, systems with negative relative velocities have on average 5.95% less blending of lines than systems with positive relative velocities. In the case of $R = 3 \times 10^5$ the blending difference is 5.46%, while for $R = 1 \times 10^5$ the blending difference is 1.95%. In the case of the NIR band, which has a more symmetric shape, this effect is not noticeable.

Although blending is unavoidable in the search for O₂ in an exoplanetary atmosphere when observing from the ground, it can be reduced by combining information about the systemic velocity of the exoplanetary system with their sky coordinates and with the position of Earth in its orbit during the observations. Here we produce a quantitative estimation of which exoplanetary systems will be best suited to search for O₂ in their atmospheres based on their sky coordinates and

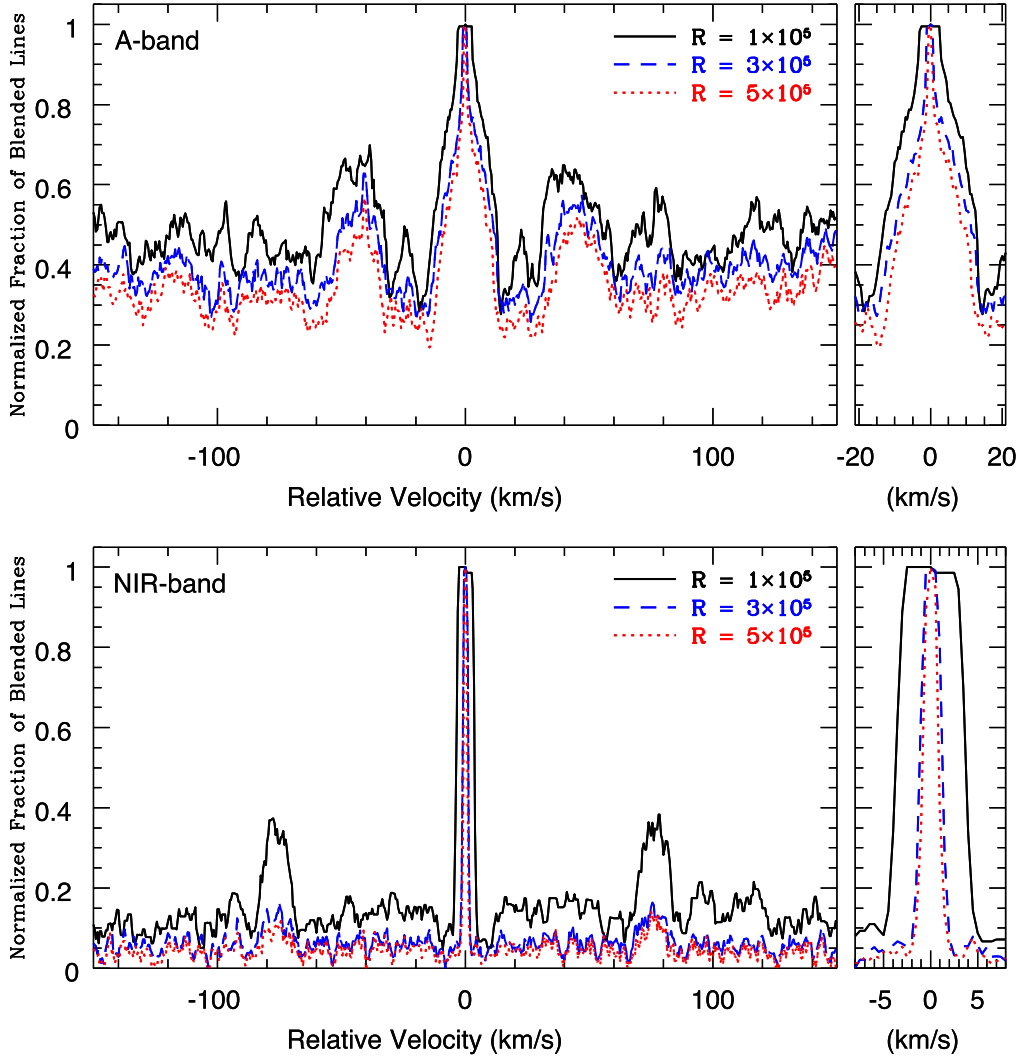


Figure 6. Top: normalized fraction of blended lines in the O_2 A-band as a function of the relative velocity of the exoplanet system with respect to Earth for three spectral resolution values ($R = 1 \times 10^5$, 3×10^5 , and 5×10^5). The top right panel shows a zoom-in around $v_* = 0$ km s^{-1} . Bottom: same as the top panel, but for the O_2 NIR band.

changes in v_* . As shown in the top panel of Figure 7, the relative velocity of a system with respect to Earth can be written as

$$v_* = v_{\text{Sun},*} + v_{\text{Earth}} \cdot \sin \theta \cdot \cos \phi, \quad (5)$$

where $v_{\text{Sun},*}$ is the relative velocity between the exoplanet system and the Sun, v_{Earth} is the velocity of Earth around the Sun, ϕ is the angle between the position of the exoplanet host star in the sky and the ecliptic, and θ is the angle between the location of Earth on the ecliptic at a given epoch and the projected line of sight between the location in the sky of the exoplanet host star and the Sun.

The bottom panel of Figure 7 shows how the relative velocity of the exoplanetary system changes relative to Earth as Earth moves in its orbit around the Sun for different values of ϕ . If the exoplanet host star is located perpendicular to the ecliptic ($\phi = 90$ deg), its velocity relative to Earth will remain constant. In this case, systems with v_* values that coincide with high blending peaks in Figure 6 are not optimal, and priority should be given to systems with v_* values with low blending levels. If the exoplanet host star is located along the ecliptic

($\phi = 0$ deg), its velocity relative to Earth will vary by 59.6 km s^{-1} periodically through the year as Earth moves along its orbit. In this case, the blending levels of a given system will change periodically, so observers can optimize observations by targeting the system during epochs when blending is reduced. The same is true for systems with intermediate ϕ values, for which the amplitude of their v_* variation will depend on their actual sky coordinates.

5. Summary and Discussion

We have generated detailed simulations of observations of O_2 in the atmosphere of a hypothetical nearby transiting Earth twin using large ground-based telescopes. Our goal is to establish informed guidelines for the design of instrumentation and observational approaches best suited to search for O_2 in the atmospheres of nearby potentially Earth-like planets. We find that performing the observations at resolutions of $R = 3\text{--}4 \times 10^5$ instead of the standard $R \sim 1 \times 10^5$ of most current and planned high-resolution spectrograph designs more than doubles the depth of the exoplanet's O_2 lines. That increase in line depth translates into a reduction of over 30% in the number of transits needed to produce

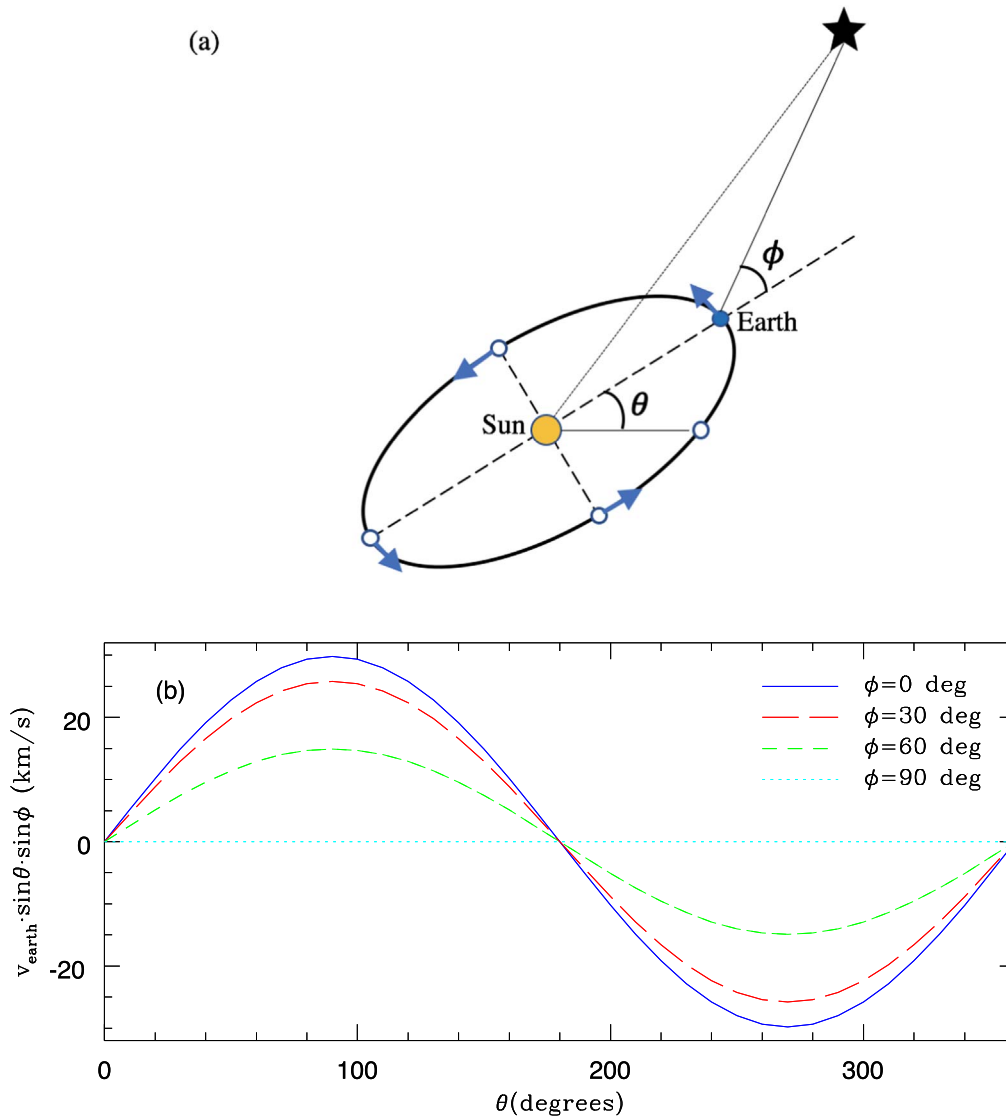


Figure 7. Top: schematic diagram to illustrate the derivation of Equation (5) in the text (see the description of the parameters shown in the figure in Section 4). Bottom: variation in the projected velocity of Earth during its orbit around the Sun for different position angles, ϕ with respect to the ecliptic.

a detection. For our calculations we have assumed that the atmosphere of the simulated exoplanet has a pressure–temperature (P – T) profile and an O_2 atmospheric vertical distribution identical to Earth. In reality this might not be the case: we might find planets with different P – T profiles or with different atmospheric vertical distributions of O_2 . In those cases the O_2 lines will appear intrinsically broader or still not fully resolved at $R = 3\text{--}4 \times 10^5$, but given the current absence of other examples, using Earth as a proxy seems justified. Also, observing at higher resolutions than the currently planned 1×10^5 might yield additional information about O_2 in exoplanetary atmospheres by enabling more detailed sampling of the absorption lines.

We also find that given the relative depths of the four O_2 absorption bands in the optical to NIR wavelength range, the relative emission of M0V to M9V host stars in those bands, and typical levels of correlated noise for current ground-based instrumentation/observations, there is no advantage to performing combined observations in multiple O_2 bands and the A band remains the best suited band to search for O_2 . Therefore, based on our current knowledge, it appears that future efforts to detect O_2 in the atmospheres of transiting exoplanets are best

focused on optimizing instrumentation and observational approaches for the A band. Designing new instrumentation which optimizes O_2 A -band observations (see Ben-Ami et al. 2018), and maybe also building dedicated spectrographs with optimize throughput in the O_2 bands, will be key. The recommendation to use only the O_2 A -band, however, should be revised if/when we manage to further reduce red noise levels in the observations (especially in the NIR), since our simulations show how combining observations in different bands will be advantageous if the red noise levels drop. One caveat of our simulations is that they use red noise levels found in optical and NIR time series photometry. However, it is not known whether high-resolution time series observations at wavelengths covering O_2 bands present similar levels of red noise. Such observations would be useful. In addition, improvements in detector noise will have a great impact in this particular science case and also others.

Because of the high density of lines in the O_2 bands, ground-based observations will always suffer from some level of blending with telluric O_2 lines. The problem is most pronounced for systems with relative velocities less than about $\pm 15 \text{ km s}^{-1}$







and between ± 35 and 55 km s^{-1} with respect to Earth. We find that, as expected, increasing the spectral resolution of the observations decreases the amount of blending, and therefore increases the number of systems that can be observed. We also find that it will be possible to observe some systems with high blending if their sky coordinates and epoch of the year are taken into account when planning the observations. It is important to take these limitations introduced by blending with telluric lines into account, especially in the case of NASA *TESS* mission candidates. The continuing viewing zone (CVZ) of *TESS* is centered around the ecliptic poles. Therefore we expect to find most of the best primary mission planet candidates in those locations as illustrated, for example, in Figure 4 of Huang et al. (2018). Also, regions within ± 6 deg from the ecliptic plane will not be observed during *TESS*'s primary mission. Furthermore, the CVZ of the *James Webb Space Telescope* also points toward the ecliptic pole, so the same best targets from *TESS* will be the focus of follow-up observations in search for potential biomarkers, such as H_2O or CH_4 from space. We suggest that the observability of O_2 should be included as a key criterion when ranking those candidates for atmospheric follow-up. Blending with telluric lines can be eliminated by observing from space, but the need for very large photon collecting areas currently limits O_2 searches to ground-based ELTs.

Another factor that needs to be determined is the stability of telluric lines, in particular how stable are the wings of telluric lines at the resolution and timescales required for these observations. Like in the case of the red noise levels for the high-resolution observations mentioned above, observations to establish how stable telluric lines are will be useful in the design of future observations of O_2 .

Finally, this and previous studies show how the expected exoplanet O_2 signals are very small. Even for systems just a few parsecs away from Earth, and with the large photon collecting areas of ELTs, we will still need to co-add observations from many transit events of an individual system to produce a significant O_2 detection. This problem could be improved by observing, whenever possible, the same targets with more than one ELT. Coordinating and combining observations of the same targets with different ELTs will not only lead to a faster collection of the necessary amount of photons to produce a detection, but it will also serve as a cross-check for spurious signals or systematics.

The authors thank the Brinson Foundation and the Smithsonian Institution for providing funding to support this project. We also thank the anonymous referee for helpful comments and suggestions. J.G.M. acknowledges support by the National Science Foundation through a Graduate Research Fellowship under grant No. DGE1745303 and by the Ford Foundation through a Ford Foundation Predoctoral Fellowship, administered by the National Academies of Sciences, Engineering, and Medicine.

ORCID iDs

Mercedes López-Morales  <https://orcid.org/0000-0003-3204-8183>
 Sagi Ben-Ami  <https://orcid.org/0000-0001-6760-3074>
 Gonzalo Gonzalez-Abad  <https://orcid.org/0000-0002-8090-6480>
 Juliana García-Mejía  <https://orcid.org/0000-0003-1361-985X>
 Jeremy Dietrich  <https://orcid.org/0000-0001-6320-7410>
 Andrew Szentgyorgyi  <https://orcid.org/0000-0002-0255-2525>

References

- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, *Natur*, **536**, 437
 Barclay, T., Pepper, J., & Quintana, E. V. 2018, *ApJS*, **239**, 2
 Ben-Ami, S., López-Morales, M., García-Mejía, J., Gonzalez Abad, G., & Szentgyorgyi, A. 2018, *ApJ*, **861**, 79
 Bétrémieux, Y., & Kaltenegger, L. 2014, *ApJ*, **791**, 7
 Bonfils, X., Astudillo-Defru, N., Díaz, R., et al. 2018, *A&A*, **613**, A25
 Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, *Natur*, **486**, 502
 Catling, D. C., & Kasting, J. F. 2017, *Atmospheric Evolution on Inhabited and Lifeless Worlds* (Cambridge: Cambridge Univ. Press)
 Chance, K., & Martin, R. V. 2017, *Spectroscopy and Radiative Transfer of Planetary Atmospheres* (Oxford: Oxford Univ. Press)
 Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, *Natur*, **544**, 333
 Domagal-Goldman, S. D., Segura, A., Claire, M. W., Robinson, T. D., & Meadows, V. S. 2014, *ApJ*, **792**, 90
 Dudhia, A. 2017, *JQSRT*, **186**, 243
 Edwards, D. P. 1992, GENLN2: A General Line-by-line Atmospheric Transmittance and Radiance Model. Version 3.0: Description and Users Guide, Tech. Rep. NCAR/TN-367+STR
 Fischer, H., Birk, M., Blom, C., et al. 2008, *Atmospheric Chemistry and Physics*, **8**, 2151
 García Muñoz, A., Zapatero Osorio, M. R., Barrena, R., et al. 2012, *ApJ*, **755**, 103
 Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, *Natur*, **542**, 456
 Huang, C. X., Shporer, A., Dragomir, D., et al. 2018, arXiv:1807.11129
 Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, **553**, A6
 Kaltenegger, L., & Traub, W. A. 2009, *ApJ*, **698**, 519
 Luger, R., & Barnes, R. 2015, *AsBio*, **15**, 119
 Misra, A., Meadows, V., & Crisp, D. 2014, *ApJ*, **792**, 61
 Pont, F., Zucker, S., & Queloz, D. 2006, *MNRAS*, **373**, 231
 Ricker, G. R., Vanderspek, R., Winn, J., et al. 2016, *Proc. SPIE*, **9904**, 99042B
 Rodler, F., & López-Morales, M. 2014, *ApJ*, **781**, 54
 Rodler, F., López-Morales, M., & Ribas, I. 2012, *ApJL*, **753**, L25
 Rogers, J. C., Apai, D., López-Morales, M., Sing, D. K., & Burrows, A. 2009, *ApJ*, **707**, 1707
 Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, *JQSRT*, **130**, 4
 Sagan, C., Thompson, W. R., Carlson, R., Gurnett, D., & Hord, C. 1993, *Natur*, **365**, 715
 Serindag, D. B., & Snellen, I. A. G. 2019, *ApJL*, **871**, L7
 Snellen, I., de Kok, R., Birkby, J. L., et al. 2015, *A&A*, **576**, A59
 Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, *Natur*, **465**, 1049
 Snellen, I. A. G., de Kok, R. J., le Poole, R., Brogi, M., & Birkby, J. 2013, *ApJ*, **764**, 182
 Szentgyorgyi, A., Frebel, A., Furesz, G., et al. 2012, *Proc. SPIE*, **8446**, 84461H
 Wordsworth, R., & Pierrehumbert, R. 2014, *ApJL*, **785**, L20