Bioverse: Exoplanet Biosignature Detection Under the UV Threshold Hypothesis

Martin Schlecker[1] and et al.

[1] *Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA; schlecker@arizona.edu*

ABSTRACT

...

1. INTRODUCTION

A wide variety of scenarios for the origin of life have been proposed (e.g., Baross & Hoffman 1985; Brasier et al. 2011; Mulkidjanian et al. 2012; Fox & Strasdeit 2013; Deamer & Georgiou 2015; Westall et al. 2018). While we are still far away from testing any of them, new prospects in the search for conditions favorable to life have opened up by thinking of the origin of life as a planetary phenomenon and identifying global-scale environmental properties that might support pathways to life (Sasselov et al. 2020). In particular, specific planetary conditions are needed to create stockpiles of initial compounds for prebiotic chemistry; and planetary processes are required to trigger the prebiotic synthesis. Such planetary conditions can be hypothesized for exoplanets located in the habitable zone of their host star, with persistent liquid water on their surface. For example, the deep-sea hydrothermal (or hydrothermal-sedimentary) context for the origin of life requires a direct contact of an ocean and the planetary mantle/crust (e.g., Baross & Hoffman 1985). The alternate scenario of a surface locally subject to wet-dry cycles requires a planetary exposure to mid-range UV irradiation, as a source of energy and an agent of selection in chemical evolution (e.g., Deamer et al. 2019). This "UV Threshold Hypothesis" states that UV light in a specific wavelength range played a constructive role in getting life started on Earth (Ranjan & Sasselov 2016; Ranjan et al. 2017a; Rimmer et al. 2018; Rapf & Vaida 2016).

The association of chemical pathways to life and planetary environmental conditions offers a new opportunity to test alternate scenarios for life emergence based on planetary-level data collected from the upcoming observations of populations of exoplanets. Deep-sea hydrothermal scenarios require planetary conditions that may not be met on ocean worlds with large amounts of water, where the water pressure on the ocean floor is high enough to form high-pressure ices (Noack et al. 2016; Kite & Ford 2018). In this case, a testable prediction would be that planets with high-pressure ices do not show biosignatures. Likewise, if UV light is required to get life started, then there is a minimum planetary UV flux requirement to have an inhabited world. This requirement is set by competitor thermal processes; if the photoreaction does not move forward at a rate faster than the competitor thermal process(es), then the abiogenesis scenario cannot function. On the other hand, abundant UV light vastly in excess of this threshold does not increase the probability of abiogenesis, since once the UV photochemistry is no longer limiting, some other thermal process in the reaction network will be the rate-limiting process instead. Therefore, a putative dependence of life on UV light is best encoded as a step function (e.g., Ranjan et al. 2017a; Rimmer et al. 2018, 2021a).

The goal of this work is to evaluate the potential of upcoming exoplanet surveys to test the hypothesis that a minimum past NUV flux is required for abiogenesis. We focus on one version of the UV Threshold Hypothesis, the so-called cyanosulfidic scenario, which has been refined to the point where the required threshold flux has been measured to be $F_{\mathrm{NUV,min}} = (6.8 \pm 3.6) \times 10^{10}\,\mathrm{photons\,cm^{-2}\,s^{-1}\,nm^{-1}}$ integrated from 200–280 nm at the surface (Rimmer et al. 2018, 2021b; Rimmer 2023; Ranjan et al. 2023).

We first follow a semi-analytical Bayesian analysis to estimate probabilities of obtaining strong evidence for or against this hypothesis. Under the UV Threshold Hypothesis (H_1), the probability of an exoplanet having detectable biosignatures is zero if the near-ultraviolet (near-Ultraviolet (NUV)) irradiation is less than the threshold, and it is equal to the (unknown) probability of abiogenesis, f_{life} if NUV exceeds the threshold for a sufficiently long period of time. Under the null hypothesis (H_0), that probability simply is f_{life}, that is, it does not correlate with the UV flux. Figure 1 shows these hypotheses as derived from the predictions of the cyanosulfidic scenario. Given a sample of planets, where for some of them we have convincing biosignature detec-

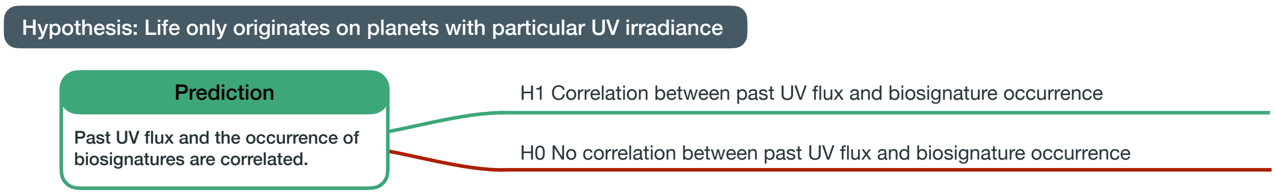

Figure 1. UV Threshold Hypothesis and null hypothesis derived from the cyanosulfidic scenario.

tions but remain agnostic on $f_{\rm life}$, we ask what evidence for H_1 and H_0 we can expect to obtain.

A real exoplanet survey will be subject to observational biases, sample selection effects, and the underlying demographics of the planet sample. To assess the information gain of a realistic exoplanet survey, we employed Bioverse (Bixel & Apai 2021; Hardegree-Ullman et al. 2023; Schlecker et al. 2024; Hardegree-Ullman et al. 2024), a framework that integrates multiple components including statistically realistic simulations of exoplanet populations, a survey simulation module, and a hypothesis testing module to evaluate the statistical power of different observational strategies.

This paper is organized as follows: In Section 2, we introduce both our semi-analytical approach and Bioverse simulations for testing the UV Threshold Hypothesis. Section 3 presents the results of these experiments for a generic survey as well as for a realistic transit survey. In Section 4, we discuss our findings before concluding with a summary in Section 5.

2. METHODS

2.1. *Fraction of inhabited planets with detectable biosignatures*

Here, we conduct a theoretical experiment on the UV Threshold Hypothesis (Sect. **??**) by relating the occurrence of life on an exo-earth candidate with a minimum past quiescent stellar UV flux, focusing on the prebiotically interesting NUV range from 200–280 nm.

SUKRIT: got a good reference for this?

Our concrete hypothesis shall be that life only occurs on planets that at some point in their history have received such radiation at a flux exceeding a threshold $F_{\rm NUV,min}$.

2.2. *Semi-analytical approach*

We first assessed the expected probabilities of obtaining true negative or true positive evidence for the UV Threshold Hypothesis (H_1) above, as well as the probability for misleading or inconclusive evidence, under idealized conditions. This serves as a first-order estimate of the information content of a survey, before we take into account impacts from exoplanet demographics, sample selection, and survey strategy.

Presumably, not all habitable worlds are inhabited and not all inhabited worlds develop detectable biosignatures. The fraction of exo-Earth candidates (EEC) that are both inhabited and harbor detectable biosignatures at the time when we observe them remains speculative; we aggregate them in the unitless parameter $f_{\rm life}$. Let us consider the probability to detect a biosignature $P(L)$, and let our observable be the inferred past NUV flux of the planet $F_{\rm NUV}$. Under Hypothesis H_1 (Equation 8), there exists a special unknown value of $F_{\rm NUV}$, noted $F_{\rm NUV,min}$ such that

$$P(L|F_{\rm NUV}, H_1) = f_{\rm life} \quad \text{if } F_{\rm NUV} > F_{\rm NUV,min} \tag{1}$$

$$P(L|F_{\rm NUV}, H_1) = 0 \quad \text{otherwise} \tag{2}$$

where $f_{\rm life}$ is the unknown probability of abiogenesis. The corresponding null hypothesis is that there exists no such special value of $F_{\rm NUV}$ and that

$$P(L|F_{\rm NUV}, H_{\rm null}) = f_{\rm life}. \tag{3}$$

If we now define a sample of size n as $X = \{F_{\rm NUV,i}, L_i\}_{i\in[1,n]}$ where L_i is equal to 1 if life is detected and 0 otherwise, we can calculate the evidence for hypothesis H_i against H_j through the Bayes factor

$$BF_{H_i,H_j} = \frac{P(X|H_i)}{P(X|H_j)}, \tag{4}$$

with $P(X|H_i)$ and $P(X|H_j)$ likelihoods of obtaining the sample X under either hypothesis.

If we define $k = \sum L_i$ and denote Y the random variable that describes it, $H_{\rm null}$ represents the likelihood that the number of planets with life in the sample follows the binomial distribution

$$P(Y = k|H_{\rm null}) = \binom{n}{k} f_{\rm life}^k (1 - f_{\rm life})^{n-k}. \tag{5}$$

Under H_1, Y also follows a binomial distribution, however it is conditioned by $n_\lambda = Card(\{F_{\rm NUV,i} \text{ if } F_{\rm NUV,i} >$

$F_{\mathrm{NUV,min}}\}_{i\in[1,n]})$ the number of values of F_{NUV} in the experiment that exceed $F_{\mathrm{NUV,min}}$

$$P(Y=k|H_1)=\binom{n_\lambda}{k} f_{\mathrm{life}}^k (1-f_{\mathrm{life}})^{n_\lambda-k}. \quad (6)$$

Hence,

$$BF_{H_1,H_{\mathrm{null}}}=\frac{P(Y=k|H_1)}{P(Y=k|H_{\mathrm{null}})}=\frac{\binom{n_\lambda}{k}}{\binom{n}{k}}(1-f_{\mathrm{life}})^{n_\lambda-n}, \quad (7a)$$

and

$$BF_{H_{\mathrm{null}},H_1}=\frac{P(Y=k|H_{\mathrm{null}})}{P(Y=k|H_1)}=\frac{\binom{n}{k}}{\binom{n_\lambda}{k}}(1-f_{\mathrm{life}})^{n-n_\lambda}. \quad (7b)$$

Given a sample of planets, where for some of them we have convincing biosignature detections but remaining agnostic on f_{life}: What evidence for H_1 and H_{null} can we expect to get? Our Bayes factor (Equation 7a) is determined by the unknown variables f_{life} and $F_{\mathrm{NUV,min}}$, as well as the number of planets with biosignature detections in the sample k. To compute the distribution of evidences, we repeatedly generated samples under H_1 and H_{null} and computed the Bayes factors $BF_{H_1,H_{\mathrm{null}}}$ and $BF_{H_{\mathrm{null}},H_1}$. We then evaluated the fraction of Monte Carlo runs in which certain evidence thresholds (Jeffreys 1939) were exceeded.

2.3. *Exoplanet survey simulations with Bioverse*

To assess the diagnostic power of realistic exoplanet surveys, we employed our survey simulator and hypothesis testing framework Bioverse (Bixel & Apai 2021). The general approach is as follows:

1. **Exoplanet population synthesis:** We populate the Gaia Catalogue of Nearby Stars (Smart et al. 2021) with synthetic exoplanets whose orbital parameters and planetary properties reflect our current understanding of exoplanet demographics (Bergsten et al. 2022). Here, we also inject the demographic trend in question - in this case we assign biosignatures according to H_1, i.e., to planets in the habitable zone (HZ) that have received NUV fluxes above a certain threshold.

2. **Survey simulation:** We simulate the detection and characterization of these exoplanets with a hypothetical survey, taking into account the survey's sensitivity, target selection, and observational biases. To model the sensitivity of the information gain of a proposed mission to sample selection and survey strategy, we conduct survey simulations with Bioverse using different sample sizes and survey strategies.

3. **Hypothesis testing:** We evaluate the likelihood that a given survey would detect a specified demographic trend in the exoplanet population and estimate the precision with which the survey could constrain the parameters of that trend. A common definition of the null hypothesis H_0, which is also applied here, is that there is no relationship between the independent variable (here: maximum NUV flux) and the dependent variable (here: biosignature occurrence). The alternative hypothesis H_1 proposes a specific relationship between the independent and dependent variables. Bioverse offers either Bayesian model comparison or non-parametric tests to evaluate the evidence for or against the null hypothesis.

To determine the diagnostic capability of a given survey, Bioverse runs multiple iterations of the simulated survey and calculates the fraction of realizations that successfully reject the null hypothesis. We use this metric, known as the statistical power, to quantify the potential information content of the survey, identify critical design trades, and find strategies that maximize the survey's scientific return.

2.3.1. *Simulated star and planet sample*

We generated two sets of synthetic exoplanet populations, one for FGK-type stars and one for M-type stars. The stellar samples are drawn from the Gaia Catalogue of Nearby Stars (Smart et al. 2021) with a maximum Gaia magnitude of 16 and a maximum stellar mass of 1.5 $M_\odot$. We included stars out to a maximum distance d_{max} that depends on the required planet sample size. Planets were generated and assigned to the synthetic stars following the occurrence rates and size/orbit distributions of Bergsten et al. (2022). Following Bixel & Apai (2021), we considered only transiting EECs with radii $0.8\,S^{0.25} < R < 1.4$ that are within the habitable zone (see Section 2.3.2). The lower limit was suggested as a minimum planet size to retain an atmosphere (Zahnle & Catling 2017). To generate planet samples larger than what the stellar catalog in combination with these occurrence rates yields, we scaled up the occurrence rates by a constant factor that yields the desired number of planets. This was in particular necessary for the FGK sample, where the rate of transiting planets that occupy the habitable zone is low.

For all survey simulations and hypothesis tests, we repeated the above in a Monte Carlo fashion to generate randomized ensembles of synthetic star and planet populations (Bixel & Apai 2021).

2.3.2. *Habitable zone occupancy and UV flux*

To test the UV Threshold Hypothesis, we require that life occurs only on planets with sufficient past UV irradiation exceeding the origins of life threshold $F_{\mathrm{NUV,min}}$. Further, we require this flux to have lasted for a minimum duration $\Delta T_{\min}$ to allow for a sufficient "origins timescale" (Rimmer 2023). All common Origin of Life scenarios require water as a solvent; we thus consider only rocky planets that may sustain liquid water on their surface, i.e., that occupy their host star's momentary HZ during the above period, as well as at the time of observation.

Does this need more detailed explanation?

This takes into account the evolution of the host star's luminosity and HZ boundaries.

Different formulations of habitable zones as regions around a star where a planet with Earth's atmospheric composition can maintain liquid water on its surface exist (e.g., Mol Lous et al. 2022; Spinelli et al. 2023; Tuchow & Wright 2023). CITE! Ramirez & Kaltenegger 2017, 2018 Here, we adopt the popular estimates of Kasting et al. (1993) and Kopparapu et al. (2013, 2014) that define a temperate zone between the runaway greenhouse transition CITE! and the maximum greenhouse limit CITE!. We use the parametrization in Kopparapu et al. (2014) to derive luminosity and planetary mass-dependent edges of the HZ a_{inner} and a_{outer}.

To determine HZ occupancy, we interpolated the stellar luminosity evolution grid of Baraffe et al. (1998) using a Clough Tocher interpolant (Nielson 1983; Alfeld 1984, see left panel of Figure 2) to compute the evolution of the inner (runaway greenhouse) and outer (maximum greenhouse) edges as a function of planet mass and stellar spectral type (Kopparapu et al. 2014). Being a local interpolation method, Clough Tocher enables rapid processing while producing a smooth interpolating surface that highlights local trends. From this, we get each planet's epochs within and outside the HZ.

For the NUV flux, we use the age- and stellar mass-dependent NUV fluxes in the HZ obtained by Richey-Yowell et al. (2023), which considers GALEX UV data in the wavelength range of 177–283 nm. We linearly interpolate in their measured grid, where we convert spectral type to stellar mass using the midpoints of their mass ranges ($0.75\,\mathrm{M}_\odot$ for K stars, $0.475\,\mathrm{M}_\odot$ for early-type M stars, and $0.215\,\mathrm{M}_\odot$ for late-type M stars). Outside the age and stellar mass range covered in Richey-Yowell et al. (2023), we extrapolate using nearest simplex (see right panel of Figure 2).

We then determined which planets were both in the HZ and had NUV fluxes above $F_{\mathrm{NUV,min}}$. To avoid considering short transitional phases, we require this situation to last for a minimum duration $\Delta T_{\min} \geq 10\,\mathrm{Myr}$. We assigned the development of life to a random fraction f_{life} of all temperate planets fulfilling these requirements. For the probability of a planet having detectable biosignatures, $P(\mathrm{bio})$, the UV Threshold Hypothesis then states

$$H_1 : P(\mathrm{bio}) = \begin{cases} 0, & F_{\mathrm{NUV}} < F_{\mathrm{NUV,min}} \\ f_{\mathrm{life}}, & F_{\mathrm{NUV}} \geq F_{\mathrm{NUV,min}} \text{ and in HZ for } \Delta t \geq 10\,\mathrm{Myr} \end{cases} \tag{8}$$

and the corresponding null hypothesis $H_{\mathrm{null}} : P(\mathrm{bio}) = f_{\mathrm{life}}$, i.e., no correlation with UV flux.

2.3.3. *Transit survey simulations*

With the synthetic star and planet samples generated, we simulated observations of these planets with a hypothetical transit survey. Using Bioverse's survey module, we simulated noisy measurements of key observables assuming a capable transit survey that can characterize a large sample with high photometric precision. We roughly followed the mission parameters of the Nautilus mission concept (Apai et al. 2019, 2022) and measured planetary instellation (for HZ occupancy) with a precision of 5% and host star effective temperature with a precision of 50.0 K. We assumed that the maximum past NUV flux a planet received can be determined within a precision of 5%. To marginalize over choices of biosignatures and their detectability, which are beyond the scope of this study, we assumed that any inhabited planet would show a biosignature detectable by the survey.

2.3.4. *Hypothesis testing*

To evaluate the evidence for correctly rejecting the null hypothesis, we employed the Mann-Whitney U test (?). This is a non-parametric test to determine whether two independent samples were drawn from a population with the same distribution and is in particular sensitive to one sample being stochastically greater than the other. We used the Mann-Whitney U test to compare the distributions of NUV fluxes of planets with and without biosignatures. The implementation in Bioverse relies on the `scipy.stats.mannwhitneyu` function (Virtanen et al. 2020) and returns a two-tailed p-value, for which we set a significance level of $\alpha = 0.05$ to reject the null hypothesis. For every hypothesis test, we repeated randomized survey realizations to estimate the fraction of successful rejections of the null hypothesis, i.e., the statistical power of the survey.

3. RESULTS

3.1. *Semi-analytical assessment*

In Section 2.2 we computed the probability for true positive evidence for H_1 and H_{null}, respectively (Equations 7a, 7b). Figure 3 shows how these evidences are

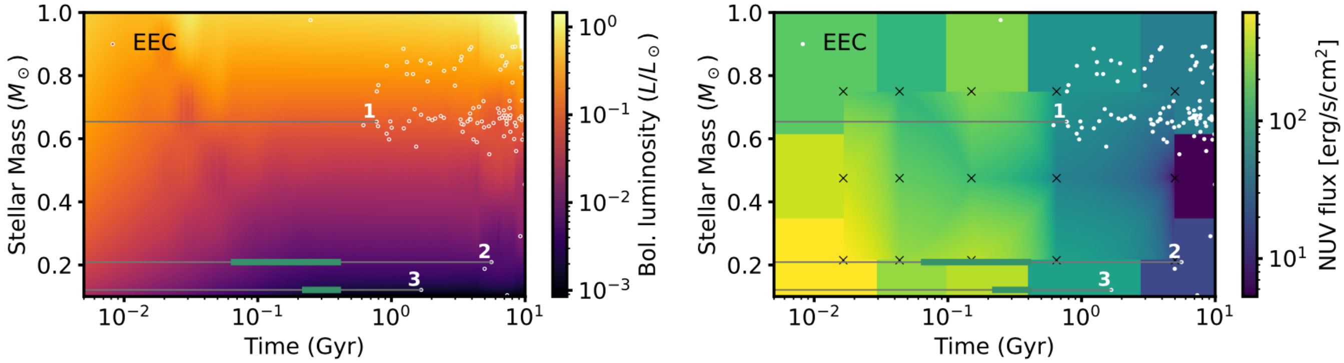

Figure 2. Interpolated stellar luminosity evolution (left) and evolution of the NUV flux in the HZ (right) as a function of host star mass. The scatter plots show age and host star mass of the transiting planets in the synthetic FGK sample; crosses denote the estimated NUV values in Richey-Yowell et al. (2019). A few example tracks for an example threshold flux of $F_{\mathrm{NUV,min}} = 300.0\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ are shown; extended overlap of HZ occupancy and high NUV flux (green sections) fulfills our requirement for abiogenesis. Planet 1 is an EEC that never receives sufficient NUV flux for abiogenesis. Planet 2 and Planet 3 enter the HZ at different times and receive sufficient NUV flux for different durations until their respective host star evolves below the threshold.

distributed for sample sizes 10 and 100, and how likely we are to obtain strong evidence ($BF_{H_i,H_j} > 10$). For $n = 10$, strong true evidence for H_1 (H_{null}) can be expected in $\sim 30\,\%$ ($\sim 40\,\%$) of all random experiments. In the majority of cases, the outcome of the survey will be inconclusive. The situation improves with larger samples: for $n = 100$, $80\,\%$ of random samples permit conclusive inference (strong true evidence) under either H_1 or H_{null}.

The expected resulting evidence further depends on the a priori unknown abiogenesis rate f_{life} and on the NUV flux threshold. Figure 4 illustrates this dependency: For very low values of either parameter, samples drawn under the null or alternative hypotheses are indistinguishable and the Bayesian evidence is always low. Both higher f_{life} and higher NUV flux thresholds increase the probability of obtaining strong evidence. Larger sample sizes enable this at lower values of these parameters.

So far, we have assumed random, uniform distributions of f_{life}, $F_{\mathrm{NUV,min}}$, and F_{NUV}. A high biosignature detection rate f_{life} increases the evidence (cmp. Equation 7a) but we cannot influence it. The same is true for $F_{\mathrm{NUV,min}}$, where again higher values increase the evidence as the binomial distribution for H_1 gets increasingly skewed and shifted away from the one for H_{null}. However, one might 'cherry-pick' exoplanets for which a biosignature test is performed based on *a priori* available contextual information (Catling et al. 2018) in order to maximize the science yield of investing additional resources. For instance, the distribution of F_{NUV} in the planet sample can be influenced by the survey strategy, and a targeted sampling approach could favor extreme values. We model this by distributing F_{NUV} according to different *Beta* functions and introduce a selectivity parameter $s \in]-1, 1[$ such that $F_{\mathrm{NUV}} \sim Beta(1/10^s, 1/10^s)$. Figure 5 shows how the probability of obtaining true strong evidence for H_1 scales with selectivity s. For large samples, a high selectivity ($s \sim 1$) can increase the probability of obtaining true strong evidence from $\sim 70\,\%$ for $s = 0$ (random uniform distribution) to $> 90\,\%$.

3.2. *Survey simulations with Bioverse*

In a magnitude- and volume-limited sample of a transit survey, the host star distribution for EECs will be skewed toward later spectral types and dominated by M dwarfs due to how the HZ scales with spectral type. For a fixed planet sample size, the fraction of inhabited planets is highest in the M dwarf sample due to the higher NUV fluxes in the HZ of these stars (see Figures 2, 6). Their NUV fluxes are generally highest at early times $\lesssim 100\,\mathrm{Myr}$. These host stars, in particular late subtypes, also provide extended periods of increased NUV emission that overlap with times when some of these planets occupy the HZ (see Figure 2), our requirement for abiogenesis (compare Equation 8). Thus – under the UV Threshold Hypothesis, most inhabited transiting planets in a magnitude- and volume-limited sample orbit M dwarfs.

Here, we are interested in the statistical power of a transit survey with a realistic sample selection and size. In the following, we fix the sample size to 250 and consider two different survey strategies targeting FGK and M dwarfs, respectively. We further investigate the sensitivity of the survey to the a priori unknown threshold NUV flux $F_{\mathrm{NUV,min}}$ and the abiogenesis rate f_{life}.

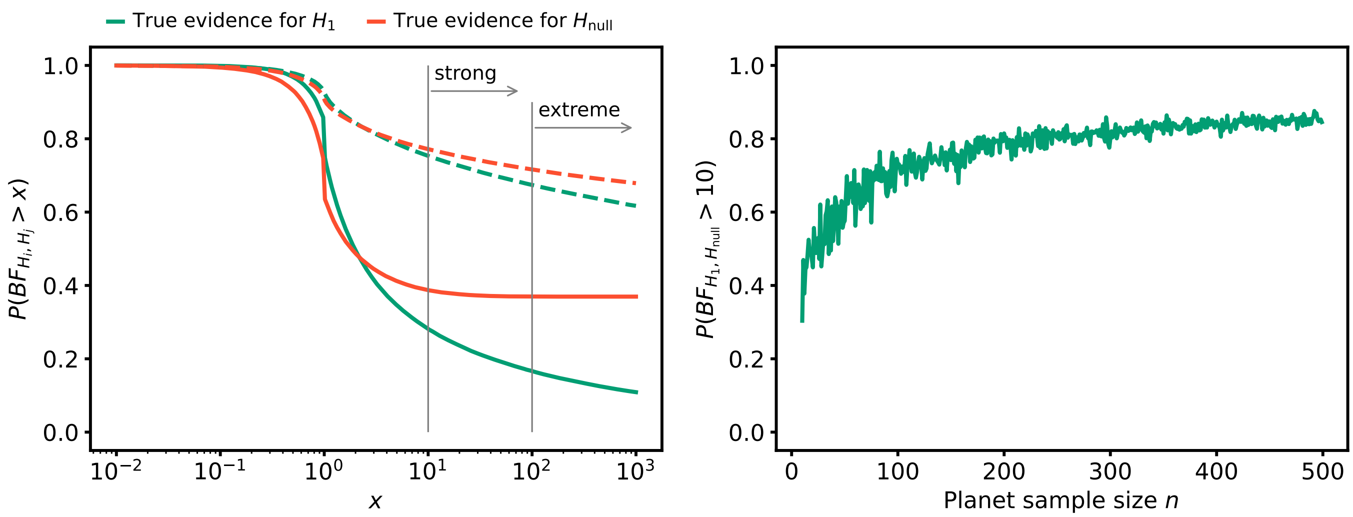

Figure 3. Obtaining true strong evidence with different sample sizes. Left: Probability to reach given evidence levels for H_1 and $H_{\rm null}$ under sample sizes $n = 10$ (solid) and $n = 100$ (dashed). Vertical lines denote thresholds for "strong" evidence, $BF_{H_i,H_j} > 10$, and "extreme" evidence, $BF_{H_i,H_j} > 100$. Right: Probability of obtaining true strong evidence for H_1 as a function of sample size n.

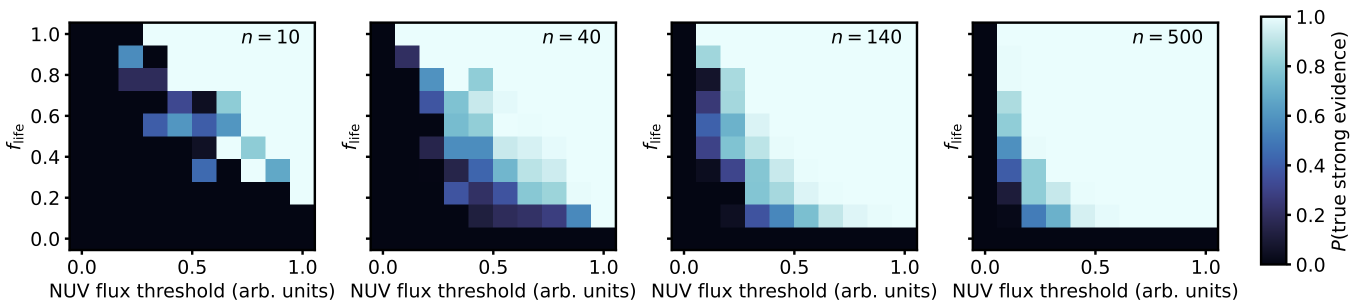

Figure 4. Probability of obtaining true strong evidence for different abiogenesis rates, NUV flux thresholds, and sample sizes. For each of these parameters, higher values increase the probability of yielding strong evidence.

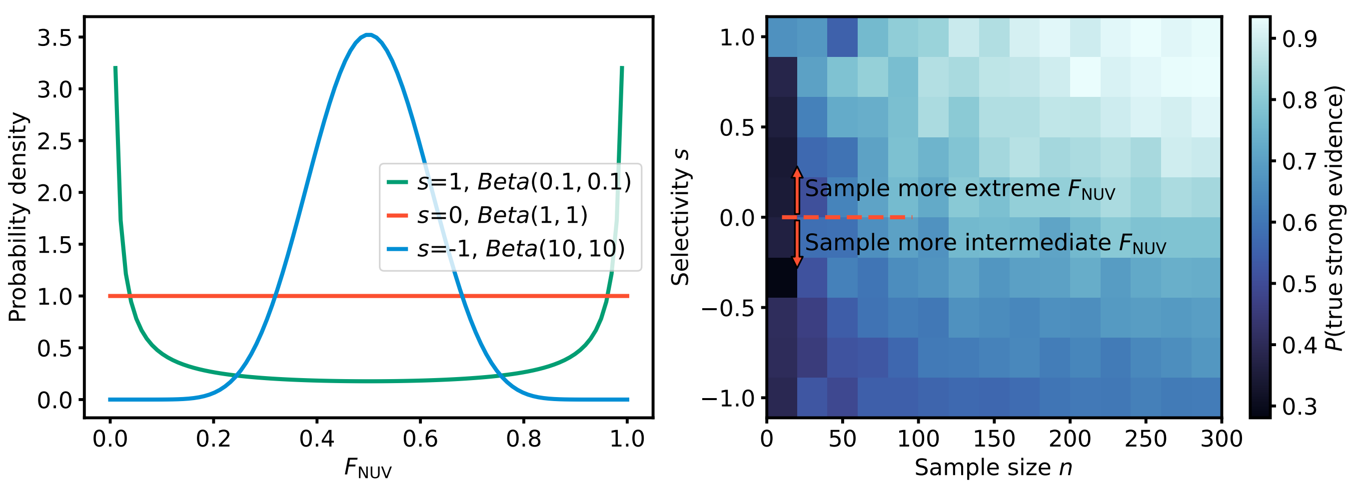

Figure 5. Scaling of the probability of obtaining true strong evidence with sample selectivity. Left: Sampling distribution for different selectivity parameters s. Right: Resulting P(true strong evidence), where $f_{\rm life}$ and $F_{\rm NUV,min}$ are randomly drawn from a uniform distribution. Sampling more extreme values of $F_{\rm NUV}$ is more likely to yield strong evidence.

3.2.1. *Selectivity of simulated transit surveys*

In Section 3.1, we demonstrated that the probability of obtaining true strong evidence for the hypothesis that life only originates on planets with a minimum past NUV flux is sensitive to the distribution of sampled past NUV fluxes, i.e., the selectivity of the survey (compare Figure 5). For both surveys targeting M dwarfs and those targeting FGK dwarfs, the maximum NUV distribution is rather unimodal. Applying the approach from Sect. 3.1 of fitting a Beta function to the distribution, we find rather low selectivities (see Figure 6).

3.2.2. *Expected biosignature pattern in a transit survey*

A representative recovery of the injected biosignature pattern is shown in Figure 6. There, we assumed an abiogenesis rate of $f_{\rm life} = 1$ and a minimum NUV flux of $F_{\rm NUV,min} = 300.0\,{\rm erg\,s^{-1}\,cm^{-2}}$. All injected biosignatures are assumed to be detected without false positive ambiguity, and the maximum NUV flux is estimated from the host star's spectral type and age with an uncertainty corresponding to the intrinsic scatter in the NUV fluxes in Richey-Yowell et al. (2023). This leads to a distribution of biosignature detections with detections increasingly occurring above a threshold inferred NUV flux. In this example case, the few biosignature detections in the FGK sample lead to a higher evidence than in the M dwarf sample, where the majority of planets are above the threshold NUV flux.

Figure 7 shows the fraction of inhabited planets under the UV Threshold Hypothesis for different threshold NUV fluxes and for a high abiogenesis rate of 1. This fraction decreases sharply with increasing threshold flux, as fewer planets receive sufficient NUV flux for abiogenesis. Another effect responsible for this drop is that some planets receive the required NUV flux only before entering the HZ – this is especially likely for M dwarfs. For the FGK sample, the fraction of inhabited planets drops at lower threshold fluxes than for the M dwarf sample.

3.2.3. *Statistical power for a transit survey and sensitivity on astrophysical parameters*

We now investigate the sensitivity of the achieved statistical power of our default transit survey to the a priori unconstrained threshold NUV flux $F_{\rm NUV,min}$ and the abiogenesis rate $f_{\rm life}$. Figure 8 shows the statistical power as a function of these parameters for a sample size of $N = 250$. Values of $F_{\rm NUV,min}$ that lie between the extrema of the inferred maximum NUV flux increase the achieved statistical power of the survey, as in this case the dataset under the alternative hypothesis H_1 differs more from the null hypothesis. Furthermore, a higher abiogenesis rate $f_{\rm life}$ increases the evidence for H_1.

Notably, the sensitivity of the M dwarf sample extends into the low NUV flux end due to the broader distribution of maximum past NUV fluxes in this sample. Here, the FGK sample is less sensitive. The only parameter space region with statistical power above 90 % lies at abiogenesis rates $f_{\rm life} > 0.9$ and threshold NUV fluxes of ~ 30–$70\,{\rm erg\,s^{-1}\,cm^{-2}}$.

Update this with final results from stat. power grid

4. DISCUSSION

A key question in the quest to understand the origins of life is which natural processes best explain how living matter spontaneously appears from nonliving matter (Malaterre et al. 2022). The cyanosulfidic scenario (Patel et al. 2015), in particular its predicted existence of a minimum NUV flux required for prebiotic chemistry, seems to offer an opportunity to test an origins of life hypothesis with a statistical transit survey sampling planets with varying NUV flux histories. In the following, we discuss the prospects of testing the UV Threshold Hypothesis in light of our results.

4.1. *Sampling strategy for testing a predicted minimum NUV flux*

In Sect. 3.1, we show that testing the hypothesis of a minimum past NUV flux required for abiogenesis suffers from 'nuisance' parameters that render inference through astronomical observations difficult. Here, these parameters are the unspecified value of the NUV threshold hypothesized to exist under H_1, and the unknown probability of detectable life emerging on a habitable planet $f_{\rm life}$. While the history of a planet's received UV flux is difficult to infer (e.g., Richey-Yowell et al. 2023), the estimated maximum NUV flux that a planet was exposed to may be used as a proxy. Indeed, the distribution of the number of planets with detected biosignature in a particular sample of planets with determined maximum NUV values $F_{\rm NUV}$ depends on both the values of $F_{{\rm NUV},min}$ and $f_{\rm life}$ as shown in equation 6. Avoiding a loss in statistical power due to nuisance parameters is a common difficulty in statistical testing, and is the motivation behind elaborate test statistics.

In our semi-analytical analysis (Section 3.1), we project a possible test performed by a future observer equipped with a sample of exoplanets with known past maximum NUV exposure for which biosignature detection has been attempted. This is necessarily reductive as this observer will have more knowledge about experimental conditions and will therefore be able to use this information to guide hypothesis testing. For instance, we have made the choice to consider the total number of detected biosignatures as our summary statistic (Equa-

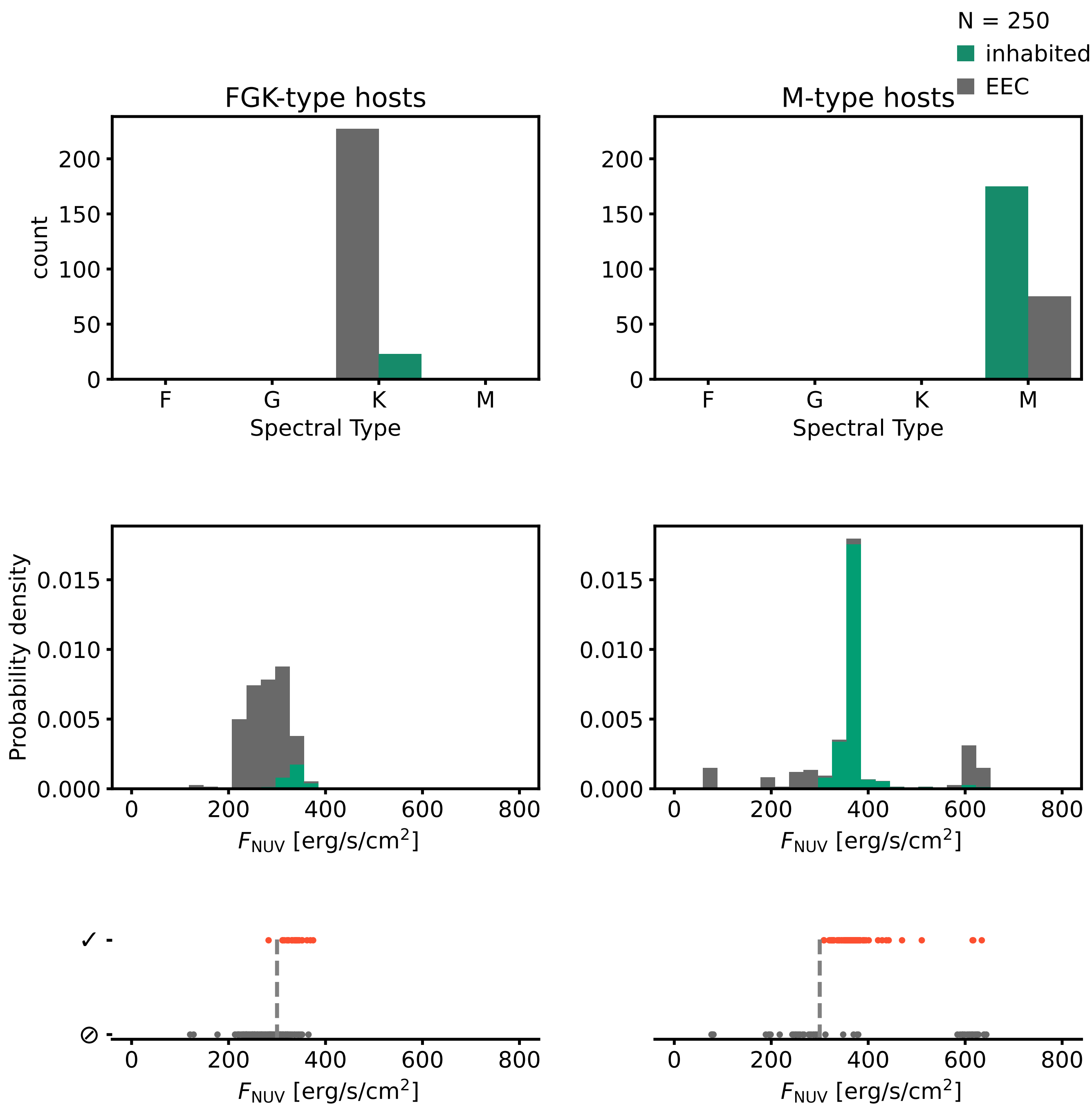

Figure 6. Simulated transit surveys targeting FGK and M stars.
Top: Host stars of all transiting EECs and inhabited planets in a simulated transit survey. In the FGK sample, most EECs and all inhabited planets orbit K dwarfs. In an M dwarf sample of the same size, the fraction of inhabited planets is larger.
Center: Distribution of inferred maximum past NUV flux in transit surveys targeting EECs around FGK and M stars, respectively. The best-fit beta distributions (gray) correspond to selectivities of $s_{\mathrm{FGK}} = -2.3$ and $s_{\mathrm{M}} = -1.3$. Green areas show inhabited planets for an abiogenesis rate of $f_{\mathrm{life}} = 1$ and a generic threshold NUV flux $F_{\mathrm{NUV,min}} = 300.0\,\mathrm{erg\,s^{-1}\,cm^{-2}}$.
Bottom: Recovered biosignature detections (✓) and non-detections (⊘) of simulated transit surveys. The dashed line denotes $F_{\mathrm{NUV,min}}$.

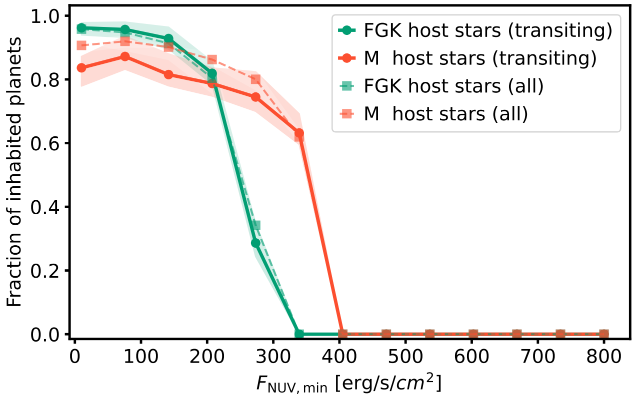

Figure 7. Fraction of inhabited planets for different threshold NUV fluxes under the UV Threshold Hypothesis if the abiogenesis rate is 1. Shaded regions denote 90 percent confidence intervals of randomized sample generations, and dashed lines correspond to samples including non-transiting planets. For all samples, the fraction of inhabited planets drops sharply with increasing threshold NUV flux due to the combined effects of never receiving sufficient NUV flux for abiogenesis or receiving it before entering the HZ.

tion 6), which is not 'sufficient' to infer $F_{\mathrm{NUV},min}$ and f_{life} separately. However, by conditioning the Bayes factor to these variables (Equation 7a), we calculate the probability distribution of the Bayesian evidence in favor of H_1. In doing so, we may evaluate how evidence depends on the uncertainty over these unknown parameters in general terms, without assuming which particular test a future observer might choose to actually perform over real data when available. From this, we can see that target selection can strongly affect the conclusiveness of a future test of the UV Threshold Hypothesis.

The particular finding that prioritizing extreme values of past NUV flux can enhance statistical power clashes with observational constraints, as the composition of the subset of planets that we can observe and for which detection of biosignature can be attempted is not independent from their NUV flux history. Hence, for our future observer, selectivity and sample size are in conflict. This trade-off can be quantified in terms of expected evidence yield, which we have done in Section 3.1. Our analysis shows that regardless of selectivity, sample sizes smaller than 50 likely result in inconclusive tests, and that increasing selectivity towards extreme F_{NUV} offers limited inference gains compared to the uniform case ($s = 0$; Figure 5). For larger samples, however, a narrow distribution of F_{NUV} may prevent inference entirely. We thus argue that selecting a sample with F_{NUV} distributed uniformly or emphasizing extreme values should – barren any practical counterarguments – be considered in any future attempt at testing the UV Threshold Hypothesis. Since the practical implementation of an exoplanet survey can stand in the way of such a selection, the following discussion focuses on the results of our transit survey simulations with Bioverse.

4.2. *Constraining power for the origins of life as a function of biosignature location*

It comes to no surprise that the success rate of such a test is sensitive to the sample size of the survey and to the occurrence of life on temperate exoplanets. As we have shown, the statistical power of this test also depends on the distribution of past NUV fluxes in the sample and on the required threshold flux. Optimizing the survey to sample a wide range of NUV flux values, particularly at the extremes, can enhance the likelihood of obtaining strong evidence for or against the hypothesis. Intermediate values of the threshold NUV flux are more likely to yield strong evidence than extreme values, as the dataset under the alternative hypothesis H_1 differs more from the null hypothesis in this case while still being sufficiently populated. The required threshold flux is, of course, a priori unknown and we cannot influence it. If, however, better theoretical predictions for the required NUV flux for abiogenesis become available (Rimmer et al. 2021a), the survey strategy can be further optimized, for instance by targeting planets that are estimated to have received a NUV flux slightly below and above this threshold or by applying a bisection algorithm in a sequential survey (Fields et al. 2023).

4.3. *An M dwarf opportunity*

An interesting aspect lies in the distribution of host star properties, as different spectral types probe different past NUV flux regimes. FGK stars show a narrow distribution of maximum past NUV fluxes in the HZ, which may limit the constraining power of a survey depending on the (unknown) threshold NUV flux. With a pure FGK sample, a survey will be more sensitive to the low NUV flux end of the hypothesis. A lack of biosignatures on these planets would support the UV Threshold Hypothesis. Their presence would have little constraining power; it would either suggest that low NUV fluxes are sufficient for abiogenesis or indicate a different abiogenic pathway (e.g., Westall et al. 2018).

On the other hand, M dwarfs show a wider distribution of maximum past NUV fluxes in their HZs. While old M dwarfs can be considered low-UV environments, a large fraction of them emit high NUV fluxes into their HZ during their early stages, in particular later subtypes (Richey-Yowell et al. 2023). an M dwarf sample will thus help to test the high NUV flux end of the UV Threshold Hypothesis; a higher occurrence of biosignatures here would support the hypothesis that a higher

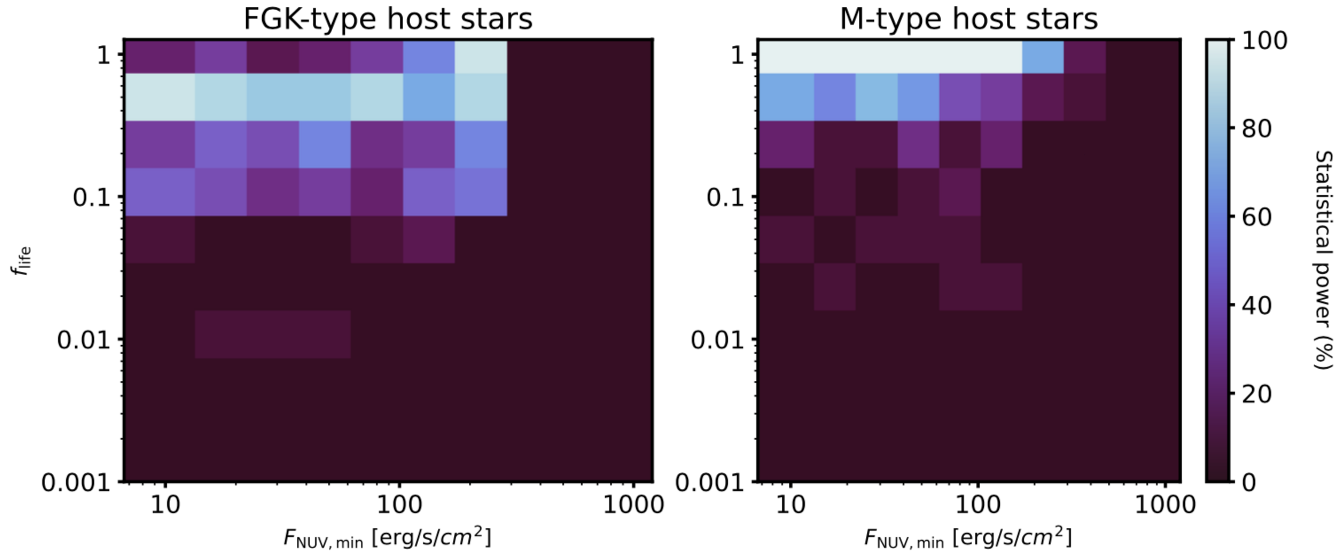

Figure 8. Statistical power as a function of threshold NUV flux and abiogenesis rate. Even for a large sample (here: $N = 250$), a high statistical power of the transit survey requires high abiogenesis rates $f_{\rm life}$. Intermediate values of $F_{\rm NUV,min}$ are more likely to yield strong evidence than extreme values; the sensitivity of the M dwarf sample extends into the low NUV flux end.

NUV flux is beneficial or necessary for life. As our simulations show, the higher and more variable NUV fluxes in M dwarfs increase the likelihood of obtaining strong evidence for or against the hypothesis. The combination of a lack of UV radiation today, which makes biosignature gases more detectable (Segura et al. 2005), and a UV-rich past that may have enabled abiogenesis could make M dwarfs the preferred targets for biosignature searches.

These findings underscore the importance of constraining the UV emission profiles of EEC host stars throughout their evolutionary stages to assess the viability of M-dwarf planets as testbeds for theories on the origins of life (Rimmer et al. 2021a; Ranjan et al. 2023).

4.4. *Sensitivity to astrophysical parameters*

Our Bioverse simulations that take into account exoplanet demographics, the evolution of habitability and NUV fluxes, and observational biases show that not only the likelihood of a conclusive test of the UV Threshold Hypothesis, but also the likelihood of successful biosignature detection itself is extremely sensitive to the threshold NUV flux if the hypothesis is true. Even if all biosignatures can be detected and the nominal abiogenesis rate is very high, say $f_{\rm life} = 1$, under the condition that prebiotic chemistry requires a minimum NUV flux *and* liquid water, the probability of finding life on a randomly selected planet may be very low. As we showed, for high required fluxes the two requirements of simultaneous HZ occupancy and sufficient NUV flux conspire to diminish the fraction of inhabited planets in the sample. Taking the infered fluxes from Richey-Yowell et al. (2023) at face value (but taking into account intrinsic scatter), a minimum required NUV flux of $\gtrsim 400\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ reduces the fraction of inhabited planets to below $\sim 1\,\%$. This not only calls for a large sample size and a targeted sample selection prefering high expected past NUV fluxes, but also highlight the necessity of continued theoretical and experimental research into the role of UV radiation in prebiotic chemistry (Ranjan et al. 2017b; Rimmer et al. 2018, 2021a).

4.5. *Contextual support for potential biosignature detections*

The predicted interplay of NUV flux and HZ occupancy in enabling abiogenesis via the cyanosulfidic scenario could in principle be used to add or remove credibility from a tentative biosignature detection. For example, with a strong belief that this scenario is the only viable one for the origins of life, a biosignature detection on a planet orbiting a strongly UV-radiating star may add credibility to the detection. Conversely, a biosignature detection on a planet estimated to have received very little UV radiation would increase the likelihood of a false positive detection. On the other hand, should the detection in the latter case be confirmed, it could be used to falsify the UV Threshold Hypothesis.

Unfortunately, our simulations do not yield a clear criterion for the credibility of a biosignature detection based on spectral type of the host star, as both FGK and M dwarf samples show similar maximum past NUV flux distributions. The abiogenesis rate in both samples show similar trends with the threshold NUV flux, and the fraction of inhabited planets drops at similar threshold fluxes (see Section 3.2). A potentially inhabited planet's host star spectral type may thus not be a strong indicator for the credibility of a biosignature detection in the context of the UV Threshold Hypothesis.

4.6. *Overall prospects for testing the UV Threshold Hypothesis*

Our results show that the UV Threshold Hypothesis is testable with upcoming exoplanet surveys, but that the success of such a test depends on the sample size, the distribution of past NUV fluxes, and several unknown astrophysical nuisance parameters. Even under idealized conditions, obtaining strong evidence for or against the hypothesis likely requires sample sizes on the order of 100 (see Section 3.1). This is true for a future transit survey, the specifics of which we have reflected in our Bioverse simulations (see Section 3.2). However, we have shown that the impacts from the combined requirements of the UV Threshold Hypothesis on the fraction of inhabited planets in a sample are comparable in the non-transiting case. The likely required sample sizes are larger than the projected sample sizes of most planned missions; for example the Habitable Worlds Observatory (HWO) expects to characterize a sample of only ~ 25 Earth analogs (Mamajek & Stapelfeldt 2023; Tuchow et al. 2024).

Meaningful tests of the UV Threshold Hypothesis may thus require a different survey strategy that increases the sample of characterized planets. Such concepts, like the Nautilus mission (Apai et al. 2019, 2022), which aims to characterize up to ~ 1000 rocky planets via utilizing an innovative telescope design, are only appearing on the horizon. Constraints on the origins of life from exoplanet surveys may thus be a long-term endeavor, and the UV Threshold Hypothesis may not be testable yet with the current and next generation of exoplanet surveys. In the meantime, it is important to refine predictions on the role of UV radiation in prebiotic chemistry with both theoretical and experimental work.

4.7. *Caveats*

Our work is based on a number of assumptions and simplifications that may affect the results and conclusions. We discuss some of these caveats here.

4.7.1. *Existence of an atmosphere-crust interface*

By its nature, cyanosulfidic scenario relies on rock surfaces exposed to the planetary atmosphere. Water worlds that have their entire planeary surface covered with water contradict this requireent and do not allow for the wet-dry cycling inherent to this origin of life scenario. The competition of tectonic stress with gravitational crustal spreading (?) sets the maximum possible height of mountains, which in the solar system does not exceed ~20 km. Such mountains will be permanently underwater on water worlds. Another impediment to wet-dry cycles may be tidal locking of the planet as it stalls stellar tide-induced water movement and diurnal irradiation variability.

4.7.2. *Stellar flares*

Our assumptions on past UV flux neglect the contribution of stellar flares, which may be hypothesized as an alternative source of UV light (Buccino et al. 2007; Ranjan et al. 2017a). This concerns mainly ultracool dwarfs, due to their low quiescent emission and high pre-main sequence stellar activity (Buccino et al. 2007; ?). However, recent work indicates that the majority of stars show inadequate activity levels for a sufficient contribution through flares (Glazier et al. 2020; Ducrot et al. 2020; Günther et al. 2020).

4.7.3. *Atmosphere transmission*

We do not take into account absorption of UV radiation by the planetary atmosphere. Theoretical work suggests that the atmosphere of prebiotic Earth was largely transparent at NUV wavelengths with the only known source of attenuation being Rayleigh scattering (Ranjan & Sasselov 2017; Ranjan et al. 2017a). We thus approximated surface UV flux using top-of-atmosphere fluxes. If there are planets in a sample that do not have a transparent atmosphere at NUV wavelengths and require higher fluxes for abiogenesis, the fraction of inhabited planets in the sample will be lower. However, these planets will not pollute the below-threshold subsample, as they will not be able to host life under the UV Threshold Hypothesis.

5. CONCLUSIONS

We have investigated the potential of upcoming exoplanet surveys to test the hypothesis – motivated by the cyanosulfidic origins-of-life scenario – that a minimum past NUV flux is required for abiogenesis. To this end, we first employed a semi-analytical Bayesian analysis to estimate probabilities of obtaining strong evidence for or against this hypothesis. We then used the Bioverse framework to assess the diagnostic power of realistic transit surveys, taking into account exoplanet demographics, time-dependency of habitability and NUV fluxes, observational biases, and target selection.

Our main findings are:

1. The UV Threshold Hypothesis of the cyanosulfidic scenario for the origins of life should lead to a correlation between past NUV flux and current occurrence of biosignatures that may be observationally testable.

2. The required sample size for detecting this correlation depends on the abiogenesis rate on temperate exoplanets and the distribution of host star properties in the sample; in particular their maximum past NUV fluxes.

3. Under the UV Threshold Hypothesis, the fraction of inhabited planets in a transit survey is sensitive to the threshold NUV flux and is expected to drop sharply for required fluxes above a few hundred $\mathrm{erg\,s^{-1}\,cm^{-2}}$.

4. If the predicted UV correlation exists, obtaining strong evidence for the hypothesis is likely ($\gtrsim 80\,\%$) for sample sizes ≥ 100 if the abiogenesis rate is high ($\gtrsim 50\,\%$) and if no very high NUV fluxes are required. A survey strategy that targets extreme values of inferred past NUV irradiation increases the diagnostic power.

5. Samples of planets orbiting M dwarfs overall yield higher chances of successfully testing the UV Threshold Hypothesis. They may also be more likely to yield biosignature detections under this hypothesis.

Overall, our work demonstrates that future exoplanet surveys have the potential to test the hypothesis that a minimum past NUV flux is required for abiogenesis. More generally, we found that models of the origins of life provide hypotheses that may be testable with these surveys. Conducting realistic survey simulations with representative samples is key to making accurate predictions, developing the optimal survey strategy, and increasing science return. Our work highlights the importance of understanding the context in which a biosignature detection is made, which can not only help to assess the credibility of the detection but also to test competing theories of the origins of life on Earth and beyond.

REPRODUCIBILITY

REFERENCES

Alfeld, P. 1984, Computer Aided Geometric Design, 1, 169, doi: 10.1016/0167-8396(84)90029-3

Apai, D., Milster, T. D., Kim, D. W., et al. 2019, AJ, 158, 83, doi: 10.3847/1538-3881/ab2631

Apai, D., Milster, T. D., Kim, D., et al. 2022, in Optical Manufacturing and Testing XIV, Vol. 12221 (SPIE), 59–71, doi: 10.1117/12.2633184

Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, Astronomy and Astrophysics, v.337, p.403-412 (1998), 337, 403

Baross, J. A., & Hoffman, S. E. 1985, Origins Life Evol Biosphere, 15, 327, doi: 10.1007/BF01808177

Bergsten, G. J., Pascucci, I., Mulders, G. D., Fernandes, R. B., & Koskinen, T. T. 2022, AJ, 164, 190, doi: 10.3847/1538-3881/ac8fea

Bixel, A., & Apai, D. 2021, The Astronomical Journal, 161, 228, doi: 10.3847/1538-3881/abe042

Brasier, M. D., Matthewman, R., McMahon, S., & Wacey, D. 2011, Astrobiology, 11, 725, doi: 10.1089/ast.2010.0546

Buccino, A. P., Lemarchand, G. A., & Mauas, P. J. D. 2007, Icarus, 192, 582, doi: 10.1016/j.icarus.2007.08.012

Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., et al. 2018, Astrobiology, 18, 709, doi: 10.1089/ast.2017.1737

Deamer, D., Damer, B., & Kompanichenko, V. 2019, Astrobiology, 19, 1523, doi: 10.1089/ast.2018.1979

Deamer, D. W., & Georgiou, C. D. 2015, Astrobiology, 15, 1091, doi: 10.1089/ast.2015.1338

Ducrot, E., Gillon, M., Delrez, L., et al. 2020, A&A, 640, A112, doi: 10.1051/0004-6361/201937392

Fields, B., Gupta, S., & Sandora, M. 2023, International Journal of Astrobiology, 22, 583, doi: 10.1017/S1473550423000150

Fox, S., & Strasdeit, H. 2013, Astrobiology, 13, 578, doi: 10.1089/ast.2012.0934

Glazier, A. L., Howard, W. S., Corbett, H., et al. 2020, ApJ, 900, 27, doi: 10.3847/1538-4357/aba4a6

Günther, M. N., Zhan, Z., Seager, S., et al. 2020, AJ, 159, 60, doi: 10.3847/1538-3881/ab5d3a

Hardegree-Ullman, K. K., Apai, D., Bergsten, G. J., Pascucci, I., & López-Morales, M. 2023, AJ, 165, 267, doi: 10.3847/1538-3881/acd1ec

Hardegree-Ullman, K. K., Apai, D., Haffert, S. Y., et al. 2024, Bioverse: GMT and ELT Direct Imaging and High-Resolution Spectroscopy Assessment $\unicode{x2013}$ Surveying Exo-Earth O$_{\mathrm{2}}$ and Testing the Habitable Zone Oxygen Hypothesis, https://arxiv.org/abs/2405.11423v1

Jeffreys, H. 1939, Theory of Probability

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108, doi: 10.1006/icar.1993.1010

Kite, E. S., & Ford, E. B. 2018, ApJ, 864, 75, doi: 10.3847/1538-4357/aad6e0

Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, ApJL, 787, L29, doi: 10.1088/2041-8205/787/2/L29

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131, doi: 10.1088/0004-637X/765/2/131

Malaterre, C., Jeancolas, C., & Nghe, P. 2022, Astrobiology, 22, 851, doi: 10.1089/ast.2021.0162

Mamajek, E., & Stapelfeldt, K. 2023, 55, 116.07

Mol Lous, M., Helled, R., & Mordasini, C. 2022, Nat Astron, 6, 819, doi: 10.1038/s41550-022-01699-8

Mulkidjanian, A. Y., Bychkov, A. Y., Dibrova, D. V., Galperin, M. Y., & Koonin, E. V. 2012, Proceedings of the National Academy of Sciences, 109, E821, doi: 10.1073/pnas.1117774109

Nielson, G. M. 1983, Math. Comp., 40, 253, doi: 10.1090/S0025-5718-1983-0679444-7

Noack, L., Höning, D., Rivoldini, A., et al. 2016, Icarus, 277, 215, doi: 10.1016/j.icarus.2016.05.009

Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D., & Sutherland, J. D. 2015, Nature Chem, 7, 301, doi: 10.1038/nchem.2202

Ranjan, S., Nayak, P. K., Pineda, J. S., & Narang, M. 2023, The Astronomical Journal, 166, 70, doi: 10.3847/1538-3881/ace32d

Ranjan, S., & Sasselov, D. D. 2016, Astrobiology, 16, 68, doi: 10.1089/ast.2015.1359

—. 2017, Astrobiology, 17, 169, doi: 10.1089/ast.2016.1519

Ranjan, S., Wordsworth, R., & Sasselov, D. D. 2017a, Astrobiology, 17, 687, doi: 10.1089/ast.2016.1596

—. 2017b, ApJ, 843, 110, doi: 10.3847/1538-4357/aa773e

Rapf, R. J., & Vaida, V. 2016, Phys. Chem. Chem. Phys., 18, 20067, doi: 10.1039/C6CP00980H

Richey-Yowell, T., Shkolnik, E. L., Schneider, A. C., et al. 2019, ApJ, 872, 17, doi: 10.3847/1538-4357/aafa74

—. 2023, ApJ, 951, 44, doi: 10.3847/1538-4357/acd2dc

Rimmer, P. B. 2023, in Conflicting Models for the Origin of Life (John Wiley & Sons, Ltd), 407–424, doi: 10.1002/9781119555568.ch16

Rimmer, P. B., Ranjan, S., & Rugheimer, S. 2021a, Elements, 17, 265, doi: 10.2138/gselements.17.4.265

Rimmer, P. B., Thompson, S. J., Xu, J., et al. 2021b, Astrobiology, 21, 1099, doi: 10.1089/ast.2020.2335

Rimmer, P. B., Xu, J., Thompson, S. J., et al. 2018, Science Advances, 4, eaar3302, doi: 10.1126/sciadv.aar3302

Sasselov, D. D., Grotzinger, J. P., & Sutherland, J. D. 2020, Science Advances, 6, eaax3419, doi: 10.1126/sciadv.aax3419

Schlecker, M., Apai, D., Lichtenberg, T., et al. 2024, Planet. Sci. J., 5, 3, doi: 10.3847/PSJ/acf57f

Segura, A., Kasting, J. F., Meadows, V., et al. 2005, Astrobiology, 5, 706, doi: 10.1089/ast.2005.5.706

Smart, R. L., Sarro, L. M., Rybizki, J., et al. 2021, A&A, 649, A6, doi: 10.1051/0004-6361/202039498

Spinelli, R., Borsa, F., Ghirlanda, G., Ghisellini, G., & Haardt, F. 2023, The Ultraviolet Habitable Zone of Exoplanets, doi: 10.48550/arXiv.2303.16229

Tuchow, N., Stark, C., & Mamajek, E. 2024, HPIC: The Habitable Worlds Observatory Preliminary Input Catalog, arXiv, doi: 10.48550/arXiv.2402.08038

Tuchow, N. W., & Wright, J. T. 2023, ApJ, 944, 71, doi: 10.3847/1538-4357/acb054

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Westall, F., Hickman-Lewis, K., Hinman, N., et al. 2018, Astrobiology, 18, 259, doi: 10.1089/ast.2017.1680

Zahnle, K. J., & Catling, D. C. 2017, ApJ, 843, 122, doi: 10.3847/1538-4357/aa7846