Draft version February 6, 2025
Typeset using LaTeX twocolumn style in AASTeX631

Bioverse: Potentially Observable Exoplanet Biosignature Patterns Under the UV Threshold Hypothesis for the Origin of Life

Martin Schlecker [1] Dániel Apai [1,2] Antonin Affholder [3] Sukrit Ranjan [2,4] Régis Ferrière [3,5,6]
Kevin K. Hardegree-Ullman [1] Tim Lichtenberg [7] and Stéphane Mazevet [8]

[1] Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA; schlecker@arizona.edu
[2] Lunar and Planetary Laboratory, The University of Arizona, Tucson, AZ 85721, USA
[3] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson AZ, USA
[4] Blue Marble Space Institute of Science, Seattle, 98104, USA
[5] Institut de Biologie de l'École Normale Supérieure, ENS, PSL, Paris, France
[6] International Research Laboratory for Interdisciplinary Global Environmental Studies (iGLOBES), CNRS, ENS, PSL, University of Arizona, Tucson AZ, USA
[7] Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
[8] Observatoire de la Côte d'Azur, Université Côte d'Azur, Nice, France

ABSTRACT

A wide variety of scenarios for the origin of life have been proposed, with many influencing the prevalence and distribution of biosignatures across exoplanet populations. This relationship suggests these scenarios can be tested by predicting biosignature distributions and comparing them with empirical data. Here, we demonstrate this approach by focusing on the cyanosulfidic origins-of-life scenario and investigating the hypothesis that a minimum near-ultraviolet (NUV) flux is necessary for abiogenesis. Using Bayesian modeling and the `Bioverse` survey simulator, we constrain the probability of obtaining strong evidence for or against this "UV Threshold Hypothesis" with future biosignature surveys. Our results indicate that a correlation between past NUV flux and current biosignature occurrence is testable for sample sizes of $\gtrsim$50 planets. The diagnostic power of such tests is critically sensitive to the intrinsic abiogenesis rate and host star properties, particularly maximum past NUV fluxes. Surveys targeting a wide range of fluxes, and planets orbiting M dwarfs enhance the chances of conclusive results, with sample sizes $\gtrsim$100 providing $\gtrsim$80% likelihood of strong evidence if abiogenesis rates are high and the required NUV fluxes are moderate. For required fluxes exceeding a few hundred erg/s/cm^2, both the fraction of inhabited planets and the diagnostic power sharply decrease. Our findings demonstrate the potential of exoplanet surveys to test origins-of-life hypotheses. Beyond specific scenarios, this work underscores the broader value of realistic survey simulations for future observatories (e.g., HWO, LIFE, ELTs, Nautilus) in identifying testable science questions, optimizing mission strategies, and advancing theoretical and experimental studies of abiogenesis.

1. INTRODUCTION

While the probability of abiogenesis is unknown and has primarily been explored through statistical arguments (e.g., Spiegel & Turner 2012; Kipping 2021; Lingam et al. 2024), 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 may still be far from conclusively testing these scenarios, 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 (HZ) of their host star, with persistent liquid water on their surface. For example, if deep-sea or sedimentary hydrothermalism is required for abiogenesis, then the insulation of an ocean from the planetary crust minerals (e.g., due to high-pressure ices) may reduce or eliminate the chances of life emerging (e.g., Baross & Hoffman 1985). The alternate scenario of a surface locally subject to wet-dry cycles requires a planetary exposure to mid-range Ultraviolet (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), and it could provide a probabilistic approach to the interpretation of possible future biosignature detections (e.g., Catling et al. 2018a; Walker et al. 2018).

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 competing 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 described 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 future 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 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 **life emerging and persisting**, f_{life} if NUV exceeds the threshold for a sufficiently long period of time. Under the null hypothesis (H_{null}), 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 detections but remain agnostic on f_{life}, we ask what evidence for H_1 and H_{null} we can expect to obtain.

A real exoplanet survey will be subject to observational biases and sample selection effects, and will be constrained by 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 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 (Ranjan & Sasselov 2016). Our core 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 $F_{\mathrm{NUV,min}}$.

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

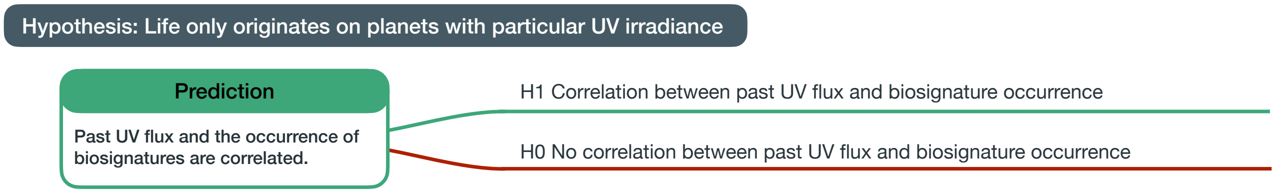

Figure 1. UV Threshold Hypothesis and null hypothesis derived from the cyanosulfidic scenario.

account the effects of 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 exhibit detectable biosignatures at the time of observation is unknown **and is represented by the term f_{life}. This encompasses the probability of life both emerging and persisting to produce detectable biosignatures. Due to our ignorance about its true value or even its order of magnitude, we draw f_{life} from a log-uniform prior probability distribution.** 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_{NUV}. Under Hypothesis H_1, there exists a special unknown value of F_{NUV}, noted $F_{\mathrm{NUV,min}}$ such that

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

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

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

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

In other words, H_{null} states that $P(L)$ is independent of F_{NUV}.

Defining a sample of size n as $X = \{F_{\mathrm{NUV,i}}, L_i\}_{i\in[1,n]}$ where L_i is equal to 1 if life is detected and 0 otherwise, we can calculate the evidence for hypothesis H_i being true 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.

Let $Y = \sum_i^n L_i$ be the random variable counting the number of positive life detections in a sample of size n. Its probability mass function under the null hypothesis H_{null} is that of a binomial distribution:

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

Under H_1, Y also follows a binomial distribution, however it is conditioned by $n_\lambda = n(\{F_{\mathrm{NUV,i}} \text{ if } F_{\mathrm{NUV,i}} > F_{\mathrm{NUV,min}}\}_{i\in[1,n]})$, the number of values of F_{NUV} in the experiment that exceed $F_{\mathrm{NUV,min}}$

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

Hence,

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

and

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

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? The analytical expression for the Bayes factor of this inference problem (Equation 7a) is determined by the unknown variables f_{life} and $F_{\mathrm{NUV,min}}$, as well as by the summary statistic Y (number of biosignature detections). 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.

Under a more realistic scenario, the distribution of n_λ depends on additional planetary properties and their evolution, as well as on observational biases and sample selection effects of the survey. We will address these in the following section.

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; Hardegree-Ullman et al. 2023). 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 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 $H_{\rm null}$, 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 used 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 $d_{\rm max}$ 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 HZ (see Section 2.3.2). The lower limit was suggested as a minimum planet size to retain an atmosphere (Zahnle & Catling 2017). 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 construct a test the UV Threshold Hypothesis, we required that life occurs only on planets with sufficient past UV irradiation exceeding the origins of life threshold $F_{\rm NUV,min}$. Further, we required this flux to have lasted for a minimum duration $\Delta T_{\rm min}$ to allow for a sufficient "origins timescale" (Rimmer 2023). All commonly investigated origins-of-life scenarios require water as a solvent; we thus considered 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 epoch of observation. To determine HZ occupancy, we took into account the evolution of the host star's luminosity and HZ boundaries.

The HZ describes a region around a star where a planet with Earth's atmospheric composition and climate feedbacks can maintain liquid water on its surface (e.g., Ramirez & Kaltenegger 2017; Ramirez 2018; Mol Lous et al. 2022; Spinelli et al. 2023; Tuchow & Wright 2023). Here, we adopted orbital distance estimates that define the HZ as the region between the runaway greenhouse transition, where the stellar instellation cannot anymore be balanced through infrared cooling to space (Ingersoll 1969), and the maximum greenhouse limit, corresponding to the maximum distance at which surface temperatures allowing liquid water can be maintained through a CO_2 greenhouse effect (Kasting 1991; Kasting et al. 1993; Underwood et al. 2003; Kopparapu et al. 2013, 2014). We used the parametrization in Kopparapu et al. (2014) to derive luminosity and planetary mass-dependent distance limits of the HZ $a_{\rm inner}$ and $a_{\rm outer}$. **The specific boundaries of the habitable zone are known to be sensitive to the star's luminosity, spectral type, the planet's mass, and the planet's atmospheric properties (e.g., Pierrehumbert & Gaidos 2011; Ramirez & Kaltenegger 2014, 2017, 2018; Koll & Cronin 2019; Ramirez 2018, 2020; Bonati & Ramirez 2021; Chaverot et al. 2022; Turbet et al. 2023). The specific habitable zone limits can therefore be variable. Here – to constrain the complexity of this study – we**

adopted a conservative approach that uses the commonly used limits.

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 used 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 $\Delta T_{\mathrm{min}} \geq 1\,\mathrm{Myr}$. We assigned the emergence **and persistence** 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

$$H_1 : P(\mathrm{bio}) = \begin{cases} 0, & F_{\mathrm{NUV}} < F_{\mathrm{NUV,min}} \\ f_{\mathrm{life}}, & F_{\mathrm{NUV}} \geq F_{\mathrm{NUV,min}} \\ & \text{and in HZ for } \Delta t \geq 1\,\mathrm{Myr} \end{cases} \tag{8}$$

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 used `Bioverse`'s survey module to simulate noisy measurements of key observables with a transit survey. We assumed a hypothetical mission that can target a large planet sample with high photometric precision and conduct a biosignature search on these planets (e.g., Apai et al. 2019, 2022). The simulated survey was designed to measure planetary instellation (for HZ occupancy) with a precision of 5% and host star effective temperature with a precision of 50.0 K. 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*

We ought to choose a statistical test that is sensitive to the UV Threshold Hypothesis, and that could be realistically conducted in a future transit survey (which may include auxiliary information from ground-based observations, archived data, or models). Given the available types of data expected from such surveys, our test shall be non-parametric and compare two samples – planets with and without biosignatures – to assess whether they are drawn from the same underlying population in terms of their infered historic maximum NUV flux. Common options include the Kolmogorov-Smirnov test, the Brunner-Munzel test (Brunner & Munzel 2000), and the Mann-Whitney U test (Mann & Whitney 1947). Due to its availability and suitability for large sample sizes, we chose the Mann-Whitney U test, which evaluates if one sample is stochastically greater than the other. Here, we 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 p-value **for each test. To balance the trade-off between Type I and Type II error risks, we set the significance level to the widely adopted threshold $\alpha = 0.05$. To quantify the diagnostic power of the survey, we conducted repeated randomized realizations and calculated the fraction of successful rejections of the null hypothesis, i.e., the statistical power.**

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 distributed for sample sizes 10 and 100, and how likely we are to obtain strong evidence ($BF_{H_i,H_j} > 10$) **in the agnostic case where we draw f_{life} from a log-uniform distribution**. For $n = 10$, strong true evidence for H_1 (H_{null}) can be **expected in** $\sim 3\,\%$ ($\sim 6\,\%$) **of** all random experiments. In the majority of cases, the outcome of the survey will be inconclusive. The situation improves with larger samples: **for $n = 100$,** $14\,\%$

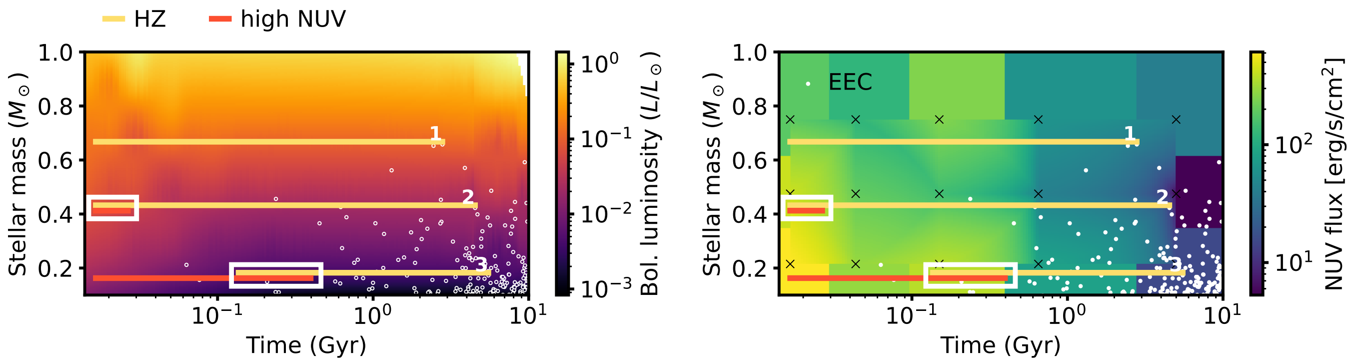

Figure 2. Interpolated stellar luminosity evolution (left) and evolution of the NUV flux in the HZ (right) as a function of host star mass. Scatter points show age and host star mass of the transiting planets in the synthetic planet sample; crosses denote the estimated NUV values in Richey-Yowell et al. (2019). We show three evolutionary tracks for a threshold flux of $F_{\mathrm{NUV,min}} = 300.0\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ that occupy the HZ (yellow sections) and exceed the threshold NUV flux (red sections) at different times. Where these sections overlap (white rectangles), the requirements for abiogenesis are met and we assign a biosignature detection with probability f_{life}. Planet 1 is an EEC orbiting a K dwarf 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.

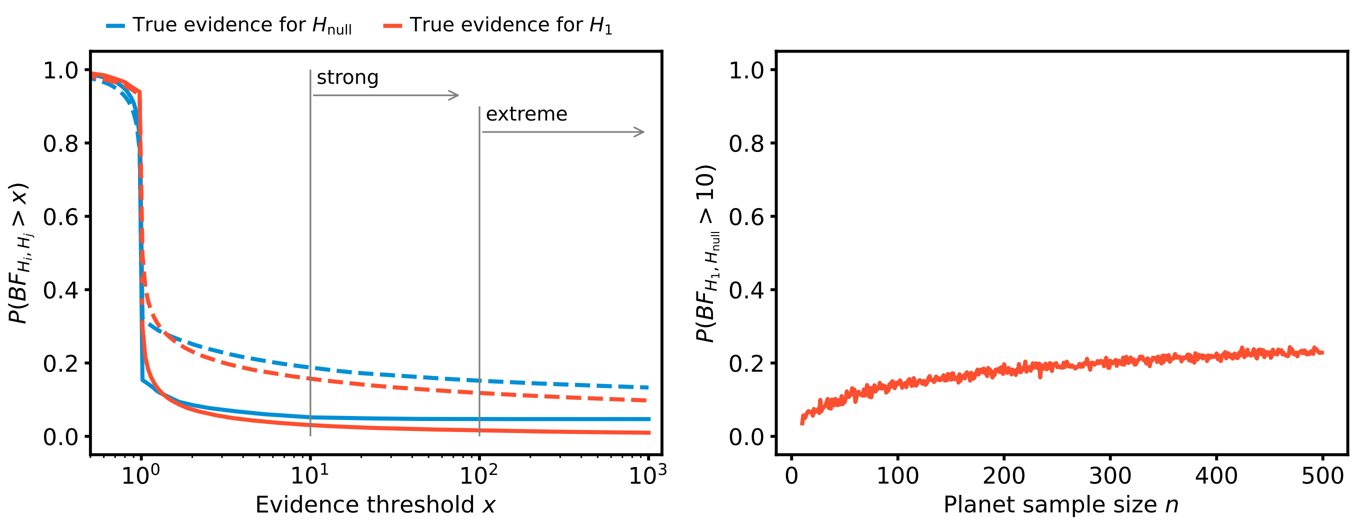

Figure 3. Obtaining true strong evidence with different sample sizes. Left: Probability to reach given evidence levels for H_1 and H_{null} under sample sizes $n = 10$ (solid) and $n = 100$ (dashed). Vertical lines denote thresholds for “strong” evidence, $BF_{H_i,H_j} > 10$, and “extreme” evidence,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

(16 %) **of** random samples permit conclusive inference (strong true evidence) **under** H_1 **(**H_{null}**).**

The expected resulting evidence further depends on the a priori unknown **rate of life's emergence and persistence** 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 **drawn random values from uniform distributions for** $F_{\mathrm{NUV,min}}$, and F_{NUV}**, and from a log-uniform distribution for** f_{life}. A high biosignature detection rate f_{life} increases the evidence (see Equation 7a) but a survey strategy 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 select exoplanets

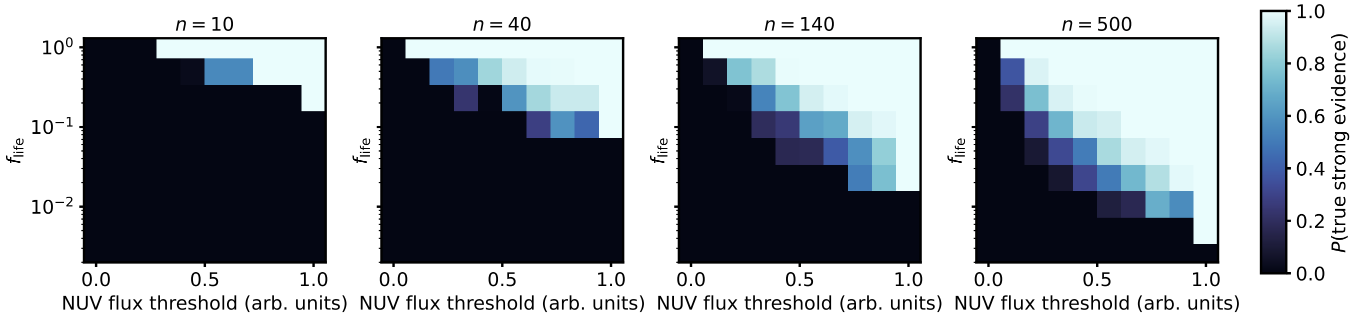

Figure 4. Probability of obtaining true strong evidence for different abiogenesis rates, NUV flux thresholds, and sample sizes. For each of these parameters, higher values increase the probability of yielding strong evidence.

for which a biosignature test is performed based on *a priori* available contextual information (Catling et al. 2018b) in order to maximize the science yield of investing additional resources. For instance, the distribution of $F_{\rm NUV}$ 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 $F_{\rm NUV}$ according to different *Beta* functions and introduce a selectivity parameter $s \in]-1, 1[$ such that $F_{\rm NUV} \sim Beta(1/10^s, 1/10^s)$. 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`

With HZ occupancy as a requirement for abiogenesis, and barring selection biases beyond stellar brightness, the host star distribution of inhabited planets in a simulated transit survey is skewed toward later spectral types. For a fixed planet sample size, the fraction of inhabited planets is highest for planets orbiting M dwarfs 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 $\lesssim 100\,$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 (see Equation 8). Thus – under the UV Threshold Hypothesis – most inhabited transiting planets in the sample orbit M dwarfs.

Here, we are interested in the statistical power of a transit survey with a plausible 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 $F_{\rm NUV,min}$ and the **probability of life emerging and persisting** $f_{\rm life}$.

3.2.1. *Selectivity of simulated transit surveys*

In Section 3.1, we showed 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), which is likely detrimental for statistical hypothesis tests.

3.2.2. *Expected biosignature pattern*

A representative recovery of the injected biosignature pattern is shown in Figure 6. There, we assumed an abiogenesis rate of $f_{\rm life} = 0.8$ 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 the limiting case of **a probability for life's emergence and persistence** of $f_{\rm life} = 1$. 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 inhab-

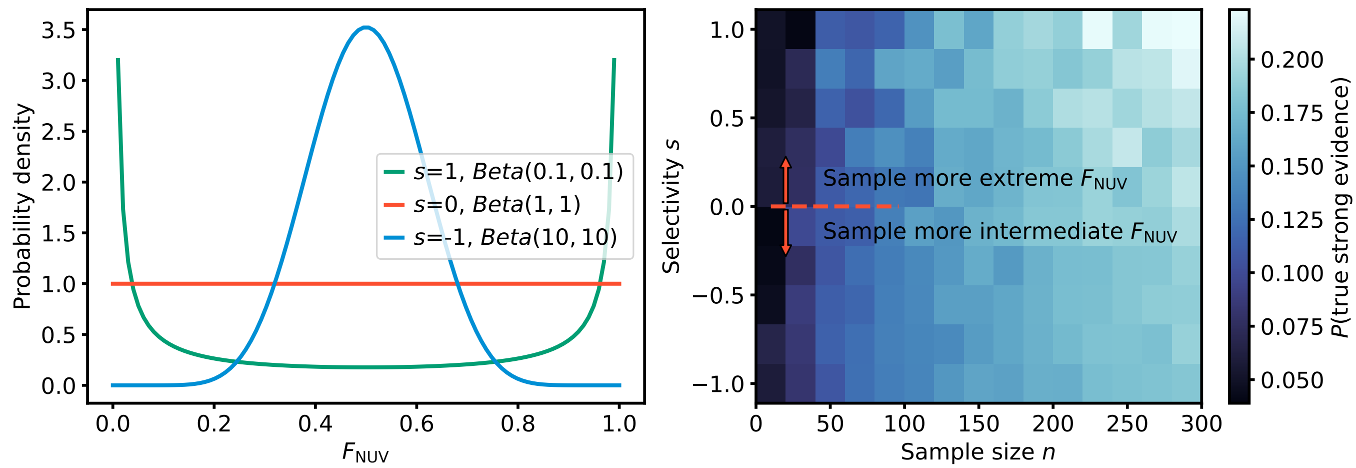

Figure 5. Scaling of the probability of obtaining true strong evidence with sample selectivity. Left: Sampling distribution for different selectivity parameters s. Right: Resulting P(true strong evidence), where f_{life} and $F_{\mathrm{NUV,min}}$ are randomly drawn from **log-uniform and uniform distributions,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

ited planets drops at lower threshold fluxes than for the M dwarf sample.

3.2.3. *Statistical power and sensitivity to 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 $F_{\mathrm{NUV,min}}$ and the abiogenesis **and persistence** rate f_{life}. Figure 8 shows the statistical power as a function of these parameters for a sample size of $N = 250$. Values of $F_{\mathrm{NUV,min}}$ 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, **higher f_{life} increase** the evidence for H_1.

Parameter space regions with statistical power above 90 % lie at $f_{\mathrm{life}} > 0.5$ and mostly at threshold NUV fluxes of ~ 200–$400\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. 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 barely sensitive.

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 (e.g., Malaterre et al. 2022). Using astronomical methods, this question will likely not be testable for individual planets but rather for ensembles of planets. The cyanosulfidic scenario for the origins of life (Patel et al. 2015), in particular its predicted existence of a minimum NUV flux required for prebiotic chemistry, offers 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 NUV flux threshold*

In Sect. 3.1, we show that testing the UV Threshold Hypothesis suffers from 'nuisance' parameters that hamper inference through astronomical observations. 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_{life}. While the inference of a planet's entire UV flux evolution is difficult (e.g., Richey-Yowell et al. 2023), the estimated maximum NUV flux that a planet was exposed to may be used as a proxy, at least if one is interested in a minimum threshold flux and makes the assumption that planetary surfaces offer protection against *too high* UV flux. Indeed, the distribution of the number of planets with detected biosignature in a particular sample of planets with inferred maximum NUV values F_{NUV} depends on both the values of $F_{\mathrm{NUV},min}$ and f_{life} as shown in equation 6.

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 derived 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-

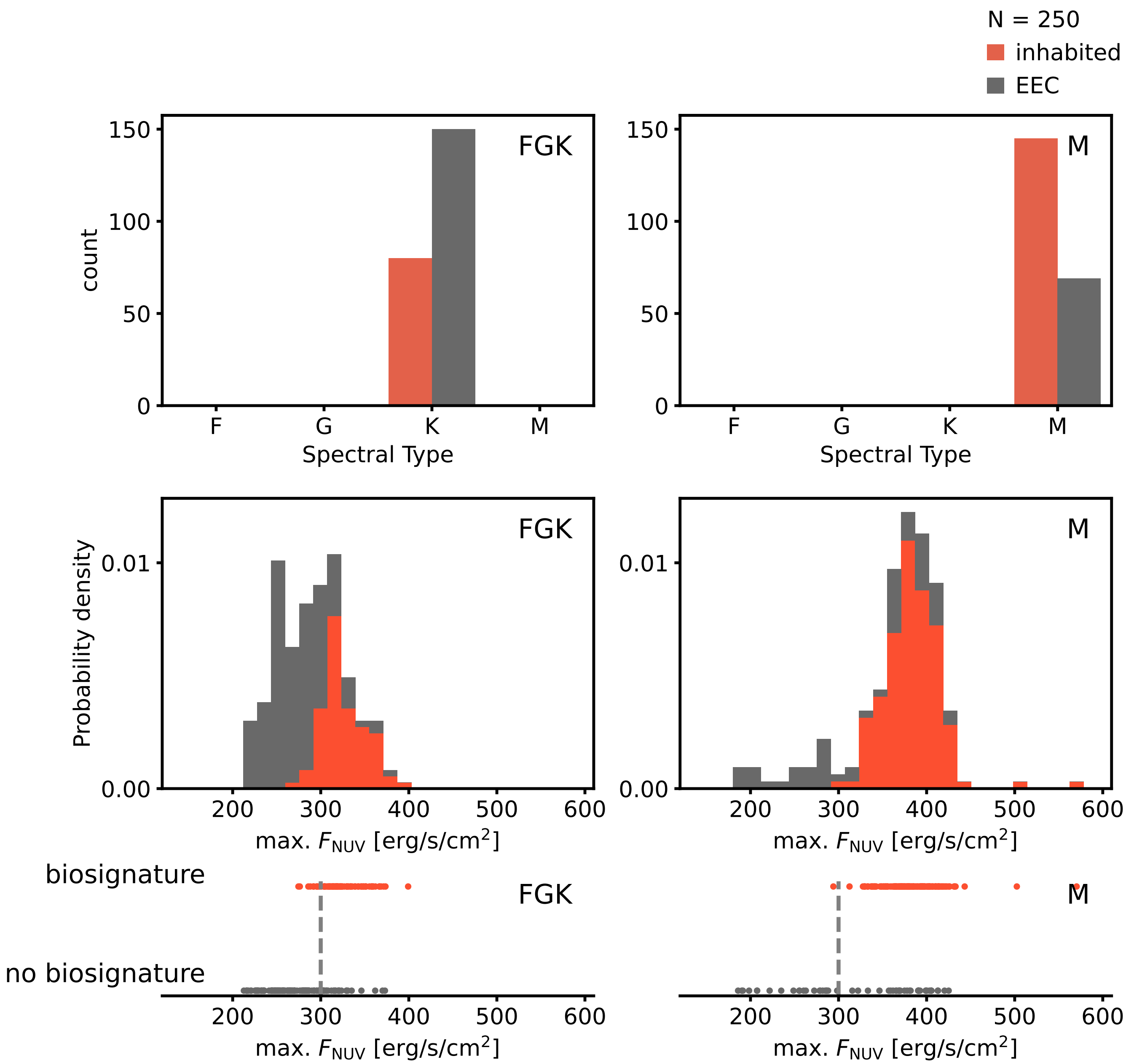

Figure 6. Simulated transit surveys targeting FGK and M stars.
Top: Host stars of all transiting EECs and inhabited planets in a simulated transit survey. In the FGK sample, all EECs and all inhabited planets orbit K dwarfs. In an M dwarf sample of the same size, the fraction of inhabited planets is 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 correspond to selectivities of $s_{\mathrm{FGK}} = -0.47$ and $s_{\mathrm{M}} = -0.02$. Red areas show inhabited planets for an abiogenesis rate of $f_{\mathrm{life}} = 0.8$ 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 $F_{\mathrm{NUV,min}}$.

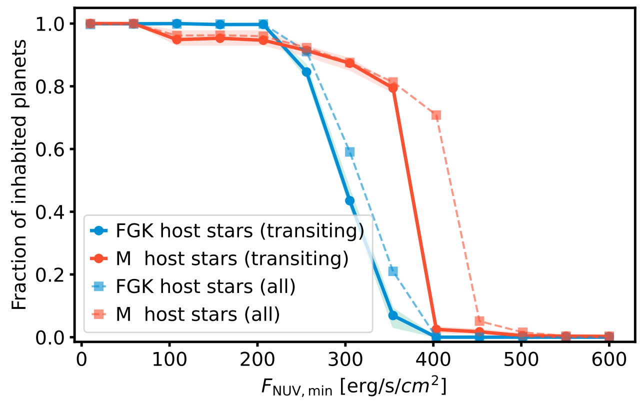

Figure 7. Fraction of inhabited planets for different threshold NUV fluxes under the UV Threshold Hypothesis if the abiogenesis rate $f_{\mathrm{life}} = 1$. For all samples, the fraction of inhabited planets drops sharply with increasing threshold NUV flux due to the combined effects of never receiving 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 actually choose to 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 likely 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 may be 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 – barring 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. *How planetary context may constrain the UV Threshold Hypothesis*

It comes to no surprise that the success rate for testing the UV Threshold Hypothesis 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 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 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 – depending on the (unknown) threshold NUV flux – limit the diagnostic power of a survey. In the case of a pure FGK sample, it will essentially only be sensitive to NUV flux thresholds $\sim$ 200–400 $\mathrm{erg\,s^{-1}\,cm^{-2}}$, and the chance of detecting biosignatures diminishes rapidly for higher thresholds within this range (see Figure 7). Detected biosignatures in an FGK sample would have little constraining power on testing the UV Threshold Hypothesis; they would either suggest that low NUV fluxes are sufficient for abiogenesis or indicate a different abiogenic pathway (e.g., Westall et al. 2018). A lack of biosignatures in a larger FGK sample would support the UV Threshold Hypothesis.

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 significant fraction of them emit high NUV fluxes into

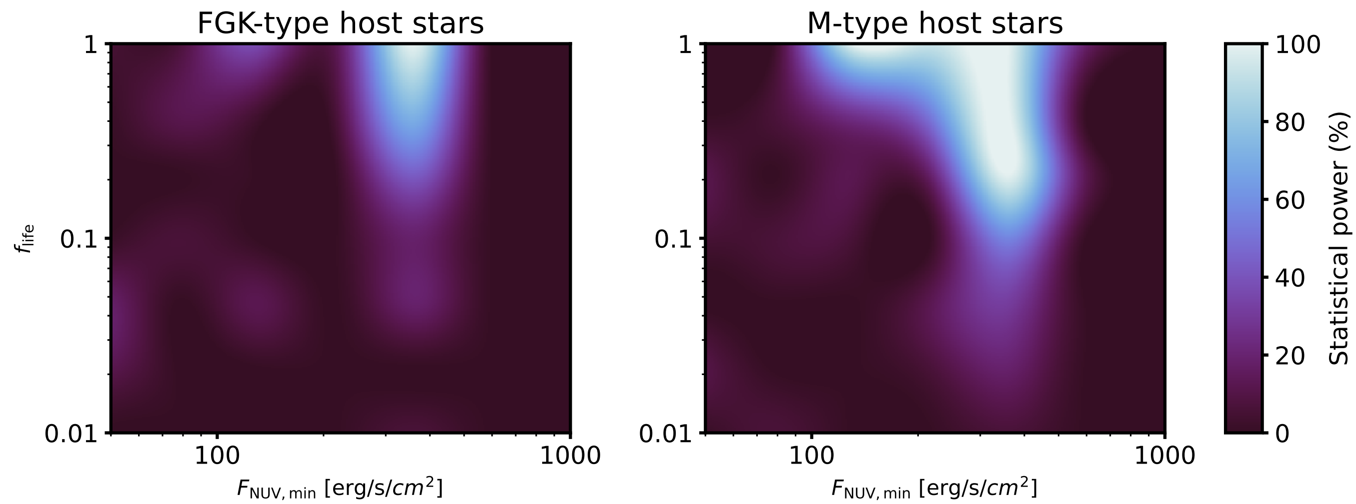

Figure 8. Statistical power as a function of threshold NUV flux and abiogenesis rate. Even for a large sample (here: $N = 250$), a high statistical power of the transit survey requires high **rates of life emerging and persisting** f_{life}. Intermediate values of $F_{\mathrm{NUV,min}}$ are more likely to yield strong evidence than extreme values. **For** $f_{\mathrm{life}} \gtrsim 80\,\%$, the sensitivity of the M dwarf sample extends into the low NUV flux end.

their HZ during their early stages, in particular later subtypes (Richey-Yowell et al. 2023). This will 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 NUV flux is favorable or necessary for life. At the same time, a fraction of host stars in our M dwarf sample extends it to lower maximum past NUV fluxes, enabling tests of the low NUV flux end of the hypothesis. The higher and more variable NUV fluxes in M dwarfs thus increase the likelihood of obtaining strong evidence for or against the UV Threshold 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. We note that relevant mission concepts, such as the Large Interferometer for Exoplanets (*LIFE*, Quanz et al. 2022; Glauser et al. 2024), include M-dwarf systems among their primary targets (Kammerer & Quanz 2018; Carrión-González et al. 2023). Our 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 **rate of life's emergence and persistence** is very high, say $f_{\mathrm{life}} = 1$, under the condition that prebiotic chemistry requires a minimum NUV flux *and* liquid water, if the threshold flux turns out to be high 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 inferred 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 preferring 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.

Our simulations find no 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 potential future 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.

Given the challenging nature of detecting and characterizing small (Earth-sized) exoplanets, most exoplanet mission concepts currently under development or considered lack the potential for characterizing large enough samples. Ground-based 25–40-meter class extremely large telescopes are expected to have the capabilities to detect biosignatures on exoplanets like Proxima Centauri b (e.g., Wang et al. 2017; Hawker & Parry 2019; Zhang et al. 2024; Vaughan et al. 2024). Hardegree-Ullman et al. (2024) used `Bioverse` to determine potential yields for a 10-year direct imaging and high-resolution spectroscopy survey of O_2 on the Giant Magellan Telescope (GMT) and on the Extremely Large Telescope (ELT) and found that between 7 and 19 habitable zone Earth-sized planets could be probed for Earth-like oxygen levels. Such a sample is too small to test the UV Threshold Hypothesis, but it may be synergistic with other detection methods.

The Habitable Worlds Observatory (HWO) is expected to characterize a sample of $\sim$ 25 Earth analogs (Mamajek & Stapelfeldt 2023; Tuchow et al. 2024). Depending on the technical design, *LIFE* is expected to target 25–80 EECs (Kammerer & Quanz 2018; Quanz et al. 2022), which could be just sufficient to constrain the UV Threshold Hypothesis. One current exception is the Nautilus Space Observatory concept (Apai et al. 2019, 2022). Nautilus aims to characterize up to $\sim$ 1000 EEC via transmission spectroscopy, building on an innovative optical technology. To guide the definition of future biosignature surveys, 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. ***The UV Threshold Hypothesis as a narrow step function***

A key aspect of the UV Threshold Hypothesis is the proposed step-function dependence of abiogenesis likelihood on UV flux. This approach stems from the constraints governing photochemical pathways, which exhibit a threshold behavior: below a certain flux, competing thermal reactions dominate, preventing abiogenesis, while above the threshold, UV photochemistry proceeds at sufficient rates, and other stochastic processes become rate-limiting.

Ranjan et al. (2017b) speculated that UV photochemistry might be rate-limiting for abiogenesis, particularly on planets orbiting M dwarfs, due to their lower baseline UV fluxes. This could delay abiogenesis by orders of magnitude, resulting in a continuous dependence of abiogenesis likelihood on UV flux. However, recent studies challenge this view for the cyanosulfidic scenario. Rimmer et al. (2021a) calculated photochemical timescales on early Earth at 180-300 hours (7.5-12.5 days), significantly shorter than the timescale for stochastic geological events (Rimmer 2023). Even with 1000x slower photochemistry on M-dwarf planets, abiogenesis would occur within 20-30 years — geologically negligible compared to stochastic processes — supporting the step-function model.

Nonetheless, alternative abiogenesis pathways or combinations of pathways may exhibit contin-

uous or mixed dependencies on UV flux. While the step-function formalism is justified for the cyanosulfidic scenario, future work should explore UV dependencies across other scenarios to refine predictions for biosignature distributions and testable hypotheses.

4.7.2. *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 planetary surface covered by oceans contradict this requirement 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 (Melosh 2011) sets the maximum possible height of mountains, which in the solar system does not exceed $\sim$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 (e.g., Ranjan et al. 2017b). However, recent dynamical models suggest tidally locked planets to undergo rapid drift of their sub-stellar point (Revol et al. 2024).

4.7.3. *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; West et al. 2008). 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).

4.7.4. *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. Exoplanet surveys focusing on highly irradiated planets offer an opportunity to constrain the typical oxidation state of rocky exoplanets, providing insights into the average composition of their secondary atmospheres (Lichtenberg & Miguel 2024). This is particularly relevant for prebiotic worlds, as varying oxidation states significantly perturb the classical habitable zone concept (Nicholls et al. 2024) and also influence surface UV levels through changing atmospheric transmission. Optimally, the atmospheric composition of young rocky protoplanets will be probed to constrain the possible range of atmospheric and mantle oxidation states during early planetary evolution by future direct imaging concepts (Cesario et al. 2024).

4.7.5. ***Other mechanisms regulating habitability and abiogenesis***

Our study focuses primarily on two factors that may regulate the emergence and persistence of life on a planet: the NUV flux and the viability of liquid water, which provide a testable framework for assessing biosignature distributions. Of course, planetary habitability and abiogenesis depend on a broader range of physical and chemical conditions. For one, a planet must retain an atmosphere capable of supporting liquid water and surface chemistry. Atmospheric escape processes have been studied in detail, and their occurrence is supported by planet formation models (e.g., Owen & Wu 2013; Schlichting et al. 2014; Ginzburg et al. 2016; Mordasini 2020; Burn et al. 2024) as well as exoplanet demographics (e.g., Owen & Estrada 2019; Bergsten et al. 2022; Rogers et al. 2021). Population synthesis studies demonstrate that models of atmospheric escape can explain key statistical features of the observed exoplanet population (Rogers & Owen 2020; Emsenhuber et al. 2021; Schlecker et al. 2021; Burn et al. 2024).

Planets around M dwarfs, in particular, may be subject to significant atmospheric loss due to the extended high-luminosity pre-main-sequence phase of their host stars (e.g., Luger & Barnes 2015). Recent works (e.g., Coy et al. 2024; Luque et al. 2024) interpret JWST eclipse measurements as growing evidence for the absence of substantial atmospheres on M-dwarf rocky exoplanets. If these planets indeed lack atmospheres, their habitability would be primarily dictated by the rate of atmospheric escape (e.g., Owen & Campos Estrada 2020). However, alternative interpretations remain viable: Ducrot et al. (2024) and Hammond et al. (2025) recently demonstrated that current JWST data cannot defini-

tively distinguish between a bare-rock scenario and an atmosphere composed primarily of N_2-CO_2-H_2O. This underscores the need for future observations and modeling efforts to constrain atmospheric retention and escape processes more robustly.

Beyond atmospheric escape, internal heating from tidal forces or radioactive decay can extend or constrain the limits of planetary surface habitability (e.g., Barnes et al. 2013; Oosterloo et al. 2021). Tidal effects are especially relevant for HZ M dwarf planets, where strong stellar interactions can drive internal heating, loss of water or an atmosphere, or runaway greenhouse conditions. Tidal locking may also create extreme climate zones, challenging habitability. While not modeled here, the factors outlined above may offer directions for future work.

5. CONCLUSIONS

We propose that specific origins-of-life scenarios may leave a detectable imprint on the distribution of biosignatures in exoplanet populations. 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. Samples smaller than 50 planets are unlikely to yield conclusive results.

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 important to identify testable science questions, support trade studies, help define science cases for future missions, and guide further theoretical and experimental work on the origins of life. 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 hypotheses on the origins of life on Earth and beyond.

ACKNOWLEDGMENTS

The authors thank Kevin Heng, Dominik Hintz, and Rhys Seeburger for insightful discussions. This material is based upon work supported by the National Aeronautics and Space Administration under Agreement No. 80NSSC21K0593 for the program "Alien Earths". The results reported herein benefited from collaborations and/or information exchange within NASA's Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by NASA's Science Mission Directorate. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. T.L. was supported by the Branco Weiss Foundation, the Netherlands eScience Center (PROTEUS project, NLESC.OEC.2023.017), and the Alfred P. Sloan Foundation (AEThER project, G202114194).

AUTHOR CONTRIBUTIONS

M.S., D.A., and S.R. conceived the project, planned its implementation, and interpreted the results. M.S. developed the planetary evolution component to `Bioverse`, carried out the hypothesis tests and statistical analyses, and wrote the manuscript. D.A. leads the "Alien Earths" program through which this project is funded, helped to guide the strategy of the project, and provided text contributions. A.A. carried out the semi-analytical computations regarding the correlation of past UV flux and biosignature occurrence. S.R. advised on planetary NUV flux evolution and the cyanosulfidic scenario of the origins of life. R.F. wrote the initial draft of the Introduction and advised on the evolutionary biology aspects of the project. K.H.-U. contributed to the `Bioverse` software development and simulations. T.L. supported the selection of testable hypotheses and provided text contributions to the initial draft. S.M. advised on the scope of the project and supported the selection of testable hypotheses. All authors provided comments and suggestions on the manuscript.

REPRODUCIBILITY

All code required to reproduce our results, figures, and this article itself is available at https://github.com/matiscke/originsoflife.

APPENDIX

A. ALTERNATIVE PRIORS FOR THE ABIOGENESIS RATE

In the main text, we used a log-uniform prior for the probability of life emerging and persisting f_{life} in order to reflect our ignorance about this parameter and its order of magnitude. We consider this a reasonable choice, as it is agnostic about the scale of f_{life} and assigns equal prior probability to all orders of magnitude. Nevertheless, we repeated our semi-analytical analysis for the more optimistic choice of a prior that is *uniform* in f_{life}. Figure 9 shows the resulting probabilities of obtaining strong evidence for or against the UV Threshold Hypothesis under this prior. While the resulting trends are qualitatively similar, the uniform prior yields overall higher probabilities of such a conclusive test for all sample sizes.

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, The Astronomical Journal, 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

Barnes, R., Mullins, K., Goldblatt, C., et al. 2013, Astrobiology, 13, 225, doi: 10.1089/ast.2012.0851

Baross, J. A., & Hoffman, S. E. 1985, Origins of life and evolution of the biosphere, 15, 327, doi: 10.1007/BF01808177

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

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

Bonati, I., & Ramirez, R. M. 2021, Monthly Notices of the Royal Astronomical Society, 504, 1029, doi: 10.1093/mnras/stab891

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

Brunner, E., & Munzel, U. 2000, Biometrical Journal, 42, 17, doi: 10.1002/(SICI)1521-4036(200001)42:1<17::AID-BIMJ17>3.0.CO;2-U

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

Burn, R., Mordasini, C., Mishra, L., et al. 2024, Nature Astronomy, 8, 463, doi: 10.1038/s41550-023-02183-7

Carrión-González, Ó., Kammerer, J., Angerhausen, D., et al. 2023, Astronomy & Astrophysics, 678, A96, doi: 10.1051/0004-6361/202347027

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

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

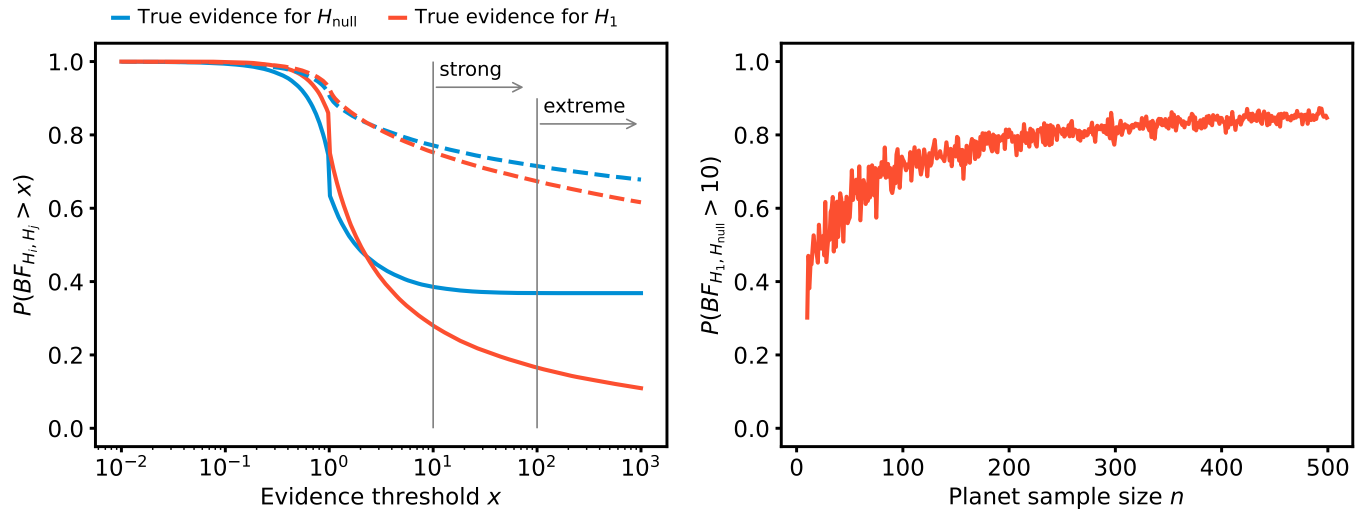

Figure 9. **As Figure 3, but using a uniform prior for the abiogenesis rate f_{life}. The probabilities of obtaining strong evidence for or against the UV Threshold Hypothesis are overall higher than for the log-uniform prior.**

Cesario, L., Lichtenberg, T., Alei, E., et al. 2024, Large Interferometer For Exoplanets (LIFE). XIV. Finding Terrestrial Protoplanets in the Galactic Neighborhood, arXiv, doi: 10.48550/arXiv.2410.13457

Chaverot, G., Turbet, M., Bolmont, E., & Leconte, J. 2022, Astronomy & Astrophysics, 658, A40, doi: 10.1051/0004-6361/202142286

Coy, B. P., Ih, J., Kite, E. S., et al. 2024, Population-Level Hypothesis Testing with Rocky Planet Emission Data: A Tentative Trend in the Brightness Temperatures of M-Earths, arXiv, doi: 10.48550/arXiv.2412.06573

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, Astronomy & Astrophysics, 640, A112, doi: 10.1051/0004-6361/201937392

Ducrot, E., Lagage, P.-O., Min, M., et al. 2024, Nature Astronomy, 1, doi: 10.1038/s41550-024-02428-z

Emsenhuber, A., Mordasini, C., Burn, R., et al. 2021, Astronomy & Astrophysics, 656, A70, doi: 10.1051/0004-6361/202038863

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

Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, The Astrophysical Journal, 825, 29, doi: 10.3847/0004-637x/825/1/29

Glauser, A. M., Quanz, S. P., Hansen, J., et al. 2024, in Optical and Infrared Interferometry and Imaging IX, Vol. 13095 (SPIE), 354–374, doi: 10.1117/12.3019090

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

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

Hammond, M., Guimond, C. M., Lichtenberg, T., et al. 2025, The Astrophysical Journal Letters, 978, L40, doi: 10.3847/2041-8213/ada0bc

Hardegree-Ullman, K. K., Apai, D., Bergsten, G. J., Pascucci, I., & López-Morales, M. 2023, The Astronomical Journal, 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

Hawker, G. A., & Parry, I. R. 2019, MNRAS, 484, 4855, doi: 10.1093/mnras/stz323

Ingersoll, A. P. 1969, Journal of Atmospheric Sciences, 26, 1191, doi: 10.1175/1520-0469(1969)026<1191:TRGAHO>2.0.CO;2

Jeffreys, H. 1939, Theory of Probability

Kammerer, J., & Quanz, S. P. 2018, Astronomy & Astrophysics, 609, A4, doi: 10.1051/0004-6361/201731254

Kasting, J. F. 1991, Icarus, 94, 1, doi: 10.1016/0019-1035(91)90137-I

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

Kipping, D. 2021, Proceedings of the National Academy of Sciences, 118, e2026808118, doi: 10.1073/pnas.2026808118

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

Koll, D. D. B., & Cronin, T. W. 2019, The Astrophysical Journal, 881, 120, doi: 10.3847/1538-4357/ab30c4

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

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

Lichtenberg, T., & Miguel, Y. 2024, Super-Earths and Earth-like Exoplanets, arXiv, doi: 10.48550/arXiv.2405.04057

Lingam, M., Nichols, R., & Balbi, A. 2024, Astrobiology, 24, 813, doi: 10.1089/ast.2024.0037

Luger, R., & Barnes, R. 2015, Astrobiology, 15, 119, doi: 10.1089/ast.2014.1231

Luque, R., Coy, B. P., Xue, Q., et al. 2024, A Dark, Bare Rock for TOI-1685 b from a JWST NIRSpec G395H Phase Curve, arXiv, doi: 10.48550/arXiv.2412.03411

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

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

Mann, H. B., & Whitney, D. R. 1947, The Annals of Mathematical Statistics, 18, 50

Melosh, H. J. 2011, Planetary Surface Processes, Cambridge Planetary Science (Cambridge: Cambridge University Press), doi: 10.1017/CBO9780511977848

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

Mordasini, C. 2020, Astronomy and Astrophysics, 638, 1, doi: 10.1051/0004-6361/201935541

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

Nicholls, H., Lichtenberg, T., Bower, D. J., & Pierrehumbert, R. 2024, Magma Ocean Evolution at Arbitrary Redox State, arXiv, doi: 10.48550/arXiv.2411.19137

Nielson, G. M. 1983, Mathematics of Computation, 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

Oosterloo, M., Höning, D., Kamp, I. E. E., & van der Tak, F. F. S. 2021, doi: 10.1051/0004-6361/202039664

Owen, J., & Estrada, B. C. 2019, 12, 1

Owen, J. E., & Campos Estrada, B. 2020, Monthly Notices of the Royal Astronomical Society, 491, 5287, doi: 10.1093/mnras/stz3435

Owen, J. E., & Wu, Y. 2013, Astrophysical Journal, 775, 1, doi: 10.1088/0004-637X/775/2/105

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

Pierrehumbert, R., & Gaidos, E. 2011, The Astrophysical Journal Letters, 734, L13, doi: 10.1088/2041-8205/734/1/L13

Quanz, S. P., Ottiger, M., Fontanet, E., et al. 2022, Astronomy & Astrophysics, 664, A21, doi: 10.1051/0004-6361/202140366

Ramirez, R. M. 2018, Geosciences, 8, 280, doi: 10.3390/geosciences8080280

—. 2020, Monthly Notices of the Royal Astronomical Society, 494, 259, doi: 10.1093/mnras/staa603

Ramirez, R. M., & Kaltenegger, L. 2014, The Astrophysical Journal Letters, 797, L25, doi: 10.1088/2041-8205/797/2/L25

—. 2017, The Astrophysical Journal Letters, 837, L4, doi: 10.3847/2041-8213/aa60c8

—. 2018, The Astrophysical Journal, 858, 72, doi: 10.3847/1538-4357/aab8fa

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, The Astrophysical Journal, 843, 110, doi: 10.3847/1538-4357/aa773e

Rapf, R. J., & Vaida, V. 2016, Physical Chemistry Chemical Physics, 18, 20067, doi: 10.1039/C6CP00980H

Revol, A., Bolmont, É., Sastre, M., et al. 2024, Astronomy & Astrophysics, 691, L3, doi: 10.1051/0004-6361/202451532

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

—. 2023, The Astrophysical Journal, 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

Rogers, J. G., Gupta, A., Owen, J. E., & Schlichting, H. E. 2021, 17, 1

Rogers, J. G., & Owen, J. E. 2020, 18, 1

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, The Planetary Science Journal, 5, 3, doi: 10.3847/PSJ/acf57f

Schlecker, M., Pham, D., Burn, R., et al. 2021, Astronomy & Astrophysics, 656, A73, doi: 10.1051/0004-6361/202140551

Schlichting, H., Sari, R., & Yalinewich, A. 2014, eprint arXiv:1406.6435, 1. https://arxiv.org/abs/1406.6435

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, Astronomy & Astrophysics, 649, A6, doi: 10.1051/0004-6361/202039498

Spiegel, D. S., & Turner, E. L. 2012, Proceedings of the National Academy of Sciences, 109, 395, doi: 10.1073/pnas.1111694108

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, The Astrophysical Journal, 944, 71, doi: 10.3847/1538-4357/acb054

Turbet, M., Fauchez, T. J., Leconte, J., et al. 2023, Water Condensation Zones around Main Sequence Stars, https://arxiv.org/abs/2308.15110v1

Underwood, D. R., Jones, B. W., & Sleep, P. N. 2003, International Journal of Astrobiology, 2, 289, doi: 10.1017/S1473550404001715

Vaughan, S. R., Birkby, J. L., Thatte, N., et al. 2024, MNRAS, 528, 3509, doi: 10.1093/mnras/stae242

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

Walker, S. I., Bains, W., Cronin, L., et al. 2018, Astrobiology, 18, 779, doi: 10.1089/ast.2017.1738

Wang, J., Mawet, D., Ruane, G., Hu, R., & Benneke, B. 2017, The Astronomical Journal, 153, 183, doi: 10.3847/1538-3881/aa6474

West, A. A., Hawley, S. L., Bochanski, J. J., et al. 2008, The Astronomical Journal, 135, 785, doi: 10.1088/0004-6256/135/3/785

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, The Astrophysical Journal, 843, 122, doi: 10.3847/1538-4357/aa7846

Zhang, H., Wang, J., & Plummer, M. K. 2024, AJ, 167, 37, doi: 10.3847/1538-3881/ad109e