



An objective Bayesian analysis of life's early start and our late arrival

David Kipping^{a,b,1}

^aDepartment of Astronomy, Columbia University, New York, NY 10027; and ^bCenter for Computational Astrophysics, Flatiron Institute, New York, NY 10010

Edited by Neta A. Bahcall, Princeton University, Princeton, NJ, and approved April 7, 2020 (received for review December 9, 2019)

Life emerged on Earth within the first quintile of its habitable window, but a technological civilization did not blossom until its last. Efforts to infer the rate of abiogenesis, based on its early emergence, are frustrated by the selection effect that if the evolution of intelligence is a slow process, then life's early start may simply be a prerequisite to our existence, rather than useful evidence for optimism. In this work, we interpret the chronology of these two events in a Bayesian framework, extending upon previous work by considering that the evolutionary timescale is itself an unknown that needs to be jointly inferred, rather than fiducially set. We further adopt an objective Bayesian approach, such that our results would be agreed upon even by those using wildly different priors for the rates of abiogenesis and evolution—common points of contention for this problem. It is then shown that the earliest microfossil evidence for life indicates that the rate of abiogenesis is at least 2.8 times more likely to be a typically rapid process, rather than a slow one. This modest limiting Bayes factor rises to 8.7 if we accept the more disputed evidence of ^{13}C -depleted zircon deposits [E. A. Bell, P. Boehnke, T. M. Harrison, W. L. Mao, *Proc. Natl. Acad. Sci. U.S.A.* **112, 14518–14521 (2015)]. For intelligence evolution, it is found that a rare-intelligence scenario is slightly favored at 3:2 betting odds. Thus, if we reran Earth's clock, one should statistically favor life to frequently reemerge, but intelligence may not be as inevitable.**

astrobiology | origin of life | Bayesian statistics

A fundamental question to modern science concerns the prevalence of life, and intelligence, within the Universe. At the time of writing, searches for nonterrestrial life within the Solar System have not yielded any direct evidence for life (1–3), and the remote detection of chemical biomarkers on extrasolar planets remains years ahead of present observational capabilities (4–7). The search for intelligence, through the signatures of their technology, may be detectable under certain assumptions (8–11) and limited observational campaigns have been attempted (12). However, the underlying assumptions make it challenging to use these null results to directly constrain the prevalence of life or intelligence at this time.

Despite having no observational data concerning nonterrestrial life, we are in possession of stronger constraints when it comes to life on Earth. Until this situation changes, inferences concerning the existence of life elsewhere in the Universe must unfortunately rely heavily on this single data point (13). While a single data point is not ideal, it is certainly not devoid of information either (14). This is even true when strong selection biases are in play, such as the fact that our existence is predicated on at least one previously successful abiogenesis event.

Problems such as these lend themselves to Bayesian analysis, where the biases can be encoded into the inference framework. The 2012 seminal paper of Spiegel and Turner (13) applied this to the question of abiogenesis. In that work, the authors treat abiogenesis as a Poisson process, which holds for systems with a discrete number of successes in a finite interval. This is used to define a likelihood function that accounts for the possibility

of multiple successes and the selection effect that a success is demanded for us to observe ourselves. Spiegel and Turner (13) conclude that the priors ultimately dominate their posteriors and that even choosing between three reasonable and diffuse priors leads to greatly different answers. Accordingly, it is very difficult to use the model to say anything definitive about how difficult or easy abiogenesis really is.

Intuitively, the possibility that life typically emerges slowly seems highly improbable given its relatively quick start on Earth (15–18). Indeed, some commentators have remarked that this fact implies that “life is not a fussy, reluctant and unlikely thing” (ref. 19, p. 501). The plausibility of the slow start scenario can be understood to be a consequence of the selection effect, which requires that life emerges fast enough for us to have sufficient time to evolve into complex (“intelligent”) organisms capable of observing ourselves. The early emergence of life on Earth then becomes consistent with a low abiogenesis rate, since worlds where life does not emerge quickly never evolve to the point of an intelligent observer.

This reveals the important role that the evolutionary timescale plays in the inference problem, since it strongly shapes the selection bias effect. In the Bayesian analysis of ref. 13, the evolutionary timescale is not known *a priori* and thus is set to three different fiducial values (1, 2, and 3.1 Gy) to test the sensitivity of their results to this parameter.

In this work, we extend the model to allow the evolutionary timescale to itself now become an inferred parameter. Rather than assume several fiducial values, this parameter can be learned by conditioning upon the time it took for observers to evolve. This enables a joint posterior distribution between the rate of abiogenesis and the rate of intelligence evolution that

Significance

Does life's early emergence mean that it would reappear quickly if we were to rerun Earth's clock? If the timescale for intelligence evolution is very slow, then a quick start to life is actually necessary for our existence—and thus does not necessarily mean it is a generally quick process. Employing objective Bayesianism and a uniform-rate process assumption, we use just the chronology of life's appearance in the fossil record, that of ourselves, and Earth's habitability window to infer the true underlying rates accounting for this subtle selection effect. Our results find betting odds of >3:1 that abiogenesis is indeed a rapid process versus a slow and rare scenario, but 3:2 odds that intelligence may be rare.

Author contributions: D.K. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The author declares no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹Email: d.kipping@columbia.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1921655117/-DCSupplemental>.

First published May 18, 2020.

STATISTICS

EVOLUTION

encodes the covariance between the two. This not only yields a more robust estimate for the abiogenesis rate by marginalizing over the uncertainty in the evolutionary timescale, but also infers the evolutionary timescale.

The Joint Likelihood Function

Distribution of Success Times. Following earlier work (13, 20–22), we describe abiogenesis as a uniform rate (i.e., Poisson) process, defined by a rate parameter λ_L . As with ref. 13, we caution that this does not imply that abiogenesis is a truly single-step, instantaneous event. Rather, we interpret the process to be a model which “integrates out” the likely complex and multistep chemistry which culminates in life. Indeed, it may be that a variety of pathways can lead to abiogenesis, but this ensemble is grouped into a single process where any of them succeeding counts as a success for the ensemble. Further, it is not necessary to strictly define what is meant by “life” here, only that the success of this Poisson process ultimately led to the geological evidence for life and that without it said evidence would not exist.

The assumption of a uniform-rate process over some time interval can at first seem problematic when one considers the stark changes to Earth’s environment over its history. However, much of this change is due to life affecting its environment and thus is a consequence of a success. The abiogenesis rate may indeed be different after life begins, but it is also irrelevant since our model cares only about the first success.

With these points in mind, we can now write that in a time interval t_L , the probability of obtaining at least one successful abiogenesis event ($X_L > 0$) will be

$$\begin{aligned} \Pr(X_L > 0; \lambda_L, t_L) &= 1 - \Pr(X_L = 0; \lambda_L, t_L), \\ &= 1 - e^{-\lambda_L t_L}. \end{aligned} \quad [1]$$

The time interval between successes for a Poisson process follows an exponential distribution. This can be demonstrated by noting that Eq. 1 corresponds to the probability of obtaining at least one success over the interval of time t_L and thus can be understood as the cumulative probability distribution for the achievement of at least one success by that time. Accordingly, the probability density function of the first success with respect to t_L must be given by the derivative of Eq. 1 with respect to time, yielding

$$\Pr(t_L | \lambda_L) = \lambda_L e^{-\lambda_L t_L}. \quad [2]$$

We now deviate from the approach of ref. 13 by considering a second process, labeled “I” for “intelligence,” which can proceed only once the previous process (“L”) is successful. The inverse of this process’s rate parameter, λ_I , describes the characteristic timescale it takes for evolution to develop from the earliest forms of life, to an “intelligent observer” which is carefully defined in *Accounting for Observational Constraints*. We truncate the times, such that both processes have to occur within a finite time T . Since process I can proceed only once process L has occurred, then this requires $t_L + t_I < T$. The joint distribution of times t_L and t_I is therefore given by

$$\Pr(t_L, t_I | \lambda_L, \lambda_I) \propto \begin{cases} \lambda_L \lambda_I e^{-\lambda_L t_L - \lambda_I t_I} & \text{if } t_L + t_I < T, \\ 0 & \text{otherwise.} \end{cases} \quad [3]$$

Imposing the condition of $t_L + t_I < T$ serves to truncate the joint distribution and thus the above is formally a proportionality because it is not yet normalized. After normalization, the expression becomes

$$\Pr(t_L, t_I | \lambda_L, \lambda_I) = \begin{cases} \frac{\lambda_L \lambda_I (\lambda_L - \lambda_I) e^{-\lambda_L t_L - \lambda_I t_I}}{\lambda_L (1 - e^{-\lambda_I T}) - \lambda_I (1 - e^{-\lambda_L T})} & \text{if } t_L + t_I < T, \\ 0 & \text{otherwise.} \end{cases} \quad [4]$$

Accounting for Observational Constraints. Before we discuss how Eq. 4 can be updated to include observational constraints, it is first useful to define exactly what we mean by intelligence in this work. We adopt a functional view of this term and consider that a successful event from process I is defined as some kind of transition—which occurs after abiogenesis—which is fundamentally necessary for analyses such as the one presented here to be possible. In other words, this type of analysis is possible only because process I succeeded and would be impossible if it failed.

Expounding upon this, we can consider that process I results in an observer/entity/society capable of 1) obtaining and dating geological evidence pertaining to the early emergence of life, 2) the ability to model the future climatic conditions of the world such that the habitability window can be estimated, and 3) interpreting the ramifications of this information regarding the underlying rates of abiogenesis and evolution. For the sake of brevity we refer to such outcomes as intelligent observers in what follows. Formally, these three conditions are not equivalent to a technological civilization, but we argue that it is difficult to imagine how these feats would be possible in the absence of one.

In what follows, we attribute process I to correspond to the emergence of human civilization and thus further assume that no previous Earth-dwelling entities/observers/societies have had the capacity to satisfy the three conditions discussed above. This assumption would be invalidated if the “Silurian hypothesis” of ref. 23 were confirmed, which considers the possible existence of industrial civilizations predating humanity (24), in which case we would certainly advocate revisiting the calculations that are described in this paper.

The emergence of human civilization could be defined in a variety of ways. Some possible defining “moments” could be the appearance of hominids, the evolution of *Homo sapiens*, complex language, the Neolithic revolution, or the first radio transmissions into space. Whatever we use, this shifts t_I around only by several million years at most. Since t_I is of order of several gigayears, these disagreements have negligible impact on our final results and thus t_I will be treated as a fixed quantity.

This is not true for the first transition, since it seems to have occurred relatively quickly and the uncertainty associated with it is comparable to the actual timing. Further, it is not possible to accurately date the emergence of life, since any life could (and indeed must) have begun prior to its appearance in the geological record. Accordingly, the true date for the emergence of life, t_L , must predate the actual observation, t'_L ; i.e., $t_L < t'_L$. The probability of this can be calculated through integration:

$$\Pr(t_L < t'_L, t_I | \lambda_L, \lambda_I) = \int_{t_L=0}^{t'_L} \Pr(t_L, t_I | \lambda_L, \lambda_I) dt_L. \quad [5]$$

Since t_I is defined as the time since t_L , then one cannot directly measure this value either. However, we can state that its value is somewhere between t'_I and $t'_I + t'_L$, where t'_I is the observed time difference between the emergence of intelligence and the first evidence for life. This allows us to write our final likelihood function as

$$\begin{aligned} \mathcal{L} &= \Pr(t_L < t'_L, t'_I < t_I < t'_I + t'_L | \lambda_L, \lambda_I) \\ &= \int_{t_I=t'_I}^{t'_I+t'_L} \Pr(t_L < t'_L, t_I | \lambda_L, \lambda_I) dt_I. \end{aligned} \quad [6]$$

Evaluating the above yields a piecewise closed-form likelihood function which is given in *SI Appendix*. The function has two sub-domains, one which applies to the interval $T > 2t'_L + t'_I$ and one which applies to $T < 2t'_L + t'_I$. As shown later, the former case is applicable when $t'_L < 0.904$ Gy and we plot this function in Fig. 1 against λ_L and λ_I , along with the limiting behaviors (*SI Appendix*).

It is interesting to note that the likelihood function is not monotonic and has a global maximum which can be solved for numerically. For example, using the “optimistic” data defined in *Adopted Values for the Observational Data*, it occurs at $\hat{\lambda}_I \rightarrow 0$ and $\hat{\lambda}_L = 21.8$ Gy $^{-1}$ (the full behavior is shown in *SI Appendix*). However, along the $\lambda_I \rightarrow 0$ axis, the likelihood is almost flat beyond this peak. For example, with the same data, the likelihood is 52.7% of the peak when one sets $\lambda_L = \hat{\lambda}_L/10$, but only 99.4% of the peak when one sets $\lambda_L = 10\hat{\lambda}_L$. While maximum-likelihood parameters are instructive, we turn to Bayesian inference to determine the posterior distributions and rigorously compare different model scenarios.

Adopted Values for the Observational Data. To perform any inference with our likelihood function, one first needs to assign values to the observables t'_L and t'_I , as well as T . All three times are relative to some initial time when conditions on Earth became suitable for life to emerge, and so let us first discuss how to define this initial time.

There is of course uncertainty about the conditions on the early Earth and when they became suitable for life (25–27). Earth is generally thought to have been impacted by a Mars-sized body, dubbed “Theia,” 4.51 Gy ago in a cataclysmic event that formed the Moon (28). Such an impact would have been a globally sterilizing event and indeed may have been accompanied by another sterilizing impactor, “Moneta,” 40 My later (29). Mineralogical evidence from zircons indicates that both an

atmosphere and liquid water must have been present on Earth’s surface (4.404 ± 0.008) Gy ago (30). In this work, we consider these to be the necessary basic requirements for abiogenesis to take place and thus adopt $T = 4.408$ Gy throughout. As noted by ref. 13, a step-function-like transition from uninhabitable to habitable is surely too simplistic, but our ignorance of Earth’s early history and indeed the conditions necessary for life mean that we do not at present have a well-motivated complex model to impose in its place.

The earliest evidence for life arrives soon after this time, in the form of ^{13}C -depleted carbon inclusions within 4.1-Gy-old zircon deposits (17). The source of this depletion remains controversial but this would yield an optimistic estimate of $t'_L = 0.304$ Gy. The earliest direct and undisputed evidence for life comes from microfossils discovered in 3.465-Gy-old rocks in western Australia (15, 16, 18), yielding $t'_L = 0.939$ Gy. We highlight that, in both the optimistic and the conservative case, t'_L is much larger than the 76-My uncertainty as to when Earth became habitable and thus is not a dominant source of uncertainty.

For t'_I , for reasons discussed in *Accounting for Observational Constraints* we can simply attribute this time to be the modern era. We therefore adopt intelligent observing as arriving “now” such that $t'_I = 4.404$ Gy – t'_L in what follows.

Finally, we turn to T , which defines the interval over which Earth is expected to persist as habitable. It is important that we refine this definition to be habitable for intelligent beings, such as ourselves. If Earth evolves into a state where only simple microbial life is possible, then a success can no longer occur for process I . As the Sun evolves, its luminosity will increase, which in turn increases the rate of weathering of silicate rocks on Earth (31). This increased weathering draws down carbon from the atmosphere, thus gradually depleting the atmospheric content of carbon dioxide. Once levels drop below ~ 10 ppm, plant C4 photosynthesis will no longer be viable (32), leading to their imminent demise. It is also possible that higher temperatures and the progressive loss of Earth’s oceans could trigger an earlier die-off (33).

The end of plant life leads to a collapse in both the food chain and Earth’s oxygen productivity, upon which animal life is critically dependent. Large endotherms, such as mammals and birds, will be the first to become extinct as a result of their higher oxygen requirements (31, 34). Thus, one can reasonably consider that the habitable window for intelligence decidedly ends once the reign of plant life comes to a close. The timing for this is predicted to be 0.9 Gy by ref. 33, a value which is adopted in what follows to give $T = 5.304$ Gy.

In summary, we set $T = 5.304$ Gy but consider two values for t'_L of $t'_L = 0.304$ Gy (“optimistic”) and $t'_L = 0.939$ Gy (“conservative”). This in turn gives two values of t'_I of $t'_I = 4.404$ Gy – t'_L .

The Nonobjective λ Power-Law Prior

The Role of the Prior. Equipped with a likelihood function, one may infer the a posteriori distribution of λ_L and λ_I using Bayes’ theorem:

$$\Pr(\lambda_L, \lambda_I | t'_L, t'_I) = \frac{\Pr(t'_L, t'_I | \lambda_L, \lambda_I) \Pr(\lambda_L, \lambda_I)}{\int \int \Pr(t'_L, t'_I | \lambda_L, \lambda_I) \Pr(\lambda_L, \lambda_I) d\lambda_L d\lambda_I}. \quad [7]$$

In any Bayesian inference problem, the posterior is a product of the likelihood and the prior and thus is affected by both. In cases where one possesses little or no information about the target parameters in advance, such as here, the ideal prior should be minimally informative (“diffuse”) such that it does not strongly influence the result (35). In objective Bayesianism, the resulting posterior should be expected to be universally agreed

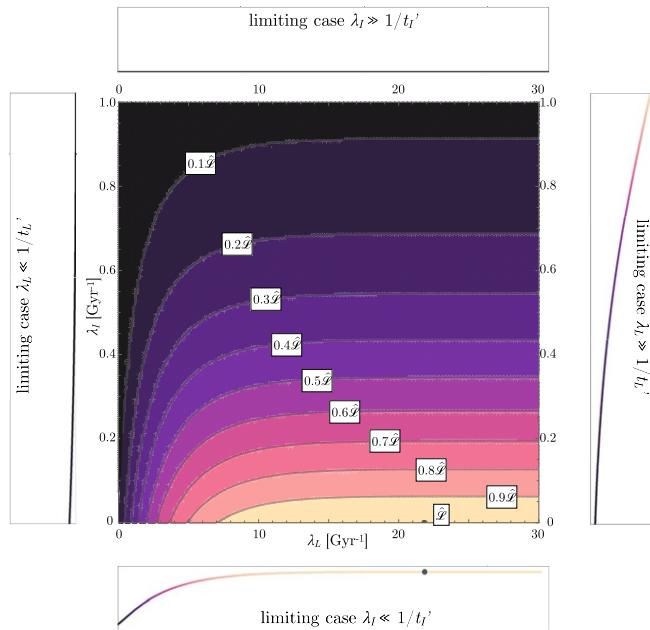


Fig. 1. The joint-likelihood function for the rate of abiogenesis, λ_L , and the rate of intelligence emergence, λ_I (for cases where $T > 2t'_L + t'_I$). Contours relative to the maximum-likelihood position (gray circle) are depicted by gray lines, where one can see a preference against low λ_L values. The limiting behaviors of the likelihood function are shown along *Top*, *Bottom*, *Left*, and *Right* edges.

upon by everyone (36)—whereas a subjective Bayesian would argue that probability corresponds to the degree of personal belief (37).

When equipped with strongly constraining data, even those using different diffuse priors will generally find convergent solutions, since the likelihood overwhelms the prior, thus naturally leading to objective Bayesianism. This is certainly not the case for our problem, since it has already been established that the posterior for λ_L is very sensitive to the priors (13). In such a case, one should tread carefully and seek a prior which can be objectively defined, such that other parties could agree upon the choice of prior and thus the resulting posterior.

We therefore proceed by considering how to define a prior distribution which is minimally informative and also not dependent upon subjective choices of the prior distribution parameters. However, we will later show that several important inference statements can be made independent of the prior.

Power Law in λ . In previous work (13, 22), the a priori distribution for λ_L was assumed to be of the form λ_L^n —a power law. In extending the likelihood to include λ_I , one may similarly extend this power-law prior to encompass λ_I by writing that $\Pr(\lambda_L, \lambda_I) = \Pr(\lambda_L)\Pr(\lambda_I)$ (i.e., assuming independence), where

$$\Pr(\lambda) = \begin{cases} \frac{\lambda^{-1}}{\log(\lambda_{\max}) - \log(\lambda_{\min})} & \text{if } n = -1, \\ \frac{\lambda_{\max}^{n+1} - \lambda_{\min}^{n+1}}{(n+1)\lambda_{\max}^n} & \text{otherwise.} \end{cases} \quad [8]$$

For $n = 0$, this returns a uniform in λ prior, for $n = -1$ a uniform in $\log \lambda$ prior, and for $n = -2$ a uniform in λ^{-1} prior: the three priors considered by ref. 13. In adopting a prior of this form, one must choose values for three shape parameters: the index, n and the prior bounds λ_{\min} and λ_{\max} .

Assigning the Prior Shape Parameters. In ref. 13, the favored index was $n = -1$ on the basis that a log-uniform prior exhibits scale-invariance ignorance for λ . For a real-valued parameter constrained only by a minimum and a maximum threshold, $n = -1$ also corresponds to the Jeffreys prior—a standard approach to defining objective priors (35). However, in this case, the $n = -1$ power law is not actually objective since one cannot objectively define a minimum and a maximum threshold. Because a power-law prior does not have semiinfinite support from $\lambda = 0$ to $\lambda = \infty$, then these prior bounds have to be subjectively chosen.

For these bounds, ref. 13 set $\lambda_{\max} = 10^3 \text{ Gy}^{-1}$ somewhat arbitrarily and a range of plausible values were offered for λ_{\min} . While useful as an exercise to test the sensitivity of the posterior to the prior, this approach does not enable an objectively defined solution. In an effort to objectively assign a power-law prior, we consider here imposing the condition that the prior should be fair and unbiased, which we define in what follows.

As currently stated, the prior in Eq. 8 appears reasonably diffuse for $n = 0, -1$, and -2 , and the bounds could essentially be anything. However, it is worth recalling that the Poisson model is used as a vehicle to describe the Bernoulli probability of one or more successful events occurring. Accordingly, a natural alternative parameterization for this problem is to consider the fraction of experiments in which the Poisson processes culminate in at least one success, f_L and f_I . Although these terms are similar to the fractions defined in the Drake equation (38), here an “experiment” really refers to rerunning Earth’s history back and observing how the stochastic processes play out each time (rather than some other world). Since $f = 1 - e^{-\lambda T}$, then the power-law prior in λ is transformed into f space as

$$\Pr(f) = \frac{n+1}{1-f} \frac{(-\log(1-f))^n}{(-\log(1-f_{\max}))^{n+1} - (-\log(1-f_{\min}))^{n+1}}. \quad [9]$$

Recall that we seek to define a prior which is both fair and unbiased. For a Bernoulli process, such as a coin toss, a “fair” prior can be defined as one for which the chance of a positively loaded coin is no more or less likely than that of a negatively loaded one. Accordingly, in our problem, we define a fair prior as one which does not a priori favor either an optimistic ($f > 1/2$) or a pessimistic ($f < 1/2$) worldview. This can be quantified by defining the prior odds ratio between the two scenarios using \mathbb{F} :

$$\mathbb{F} \equiv \frac{\int_{f=1/2}^{f_{\max}} \Pr(f) df}{\int_{f=f_{\min}}^{1/2} \Pr(f) df}. \quad [10]$$

Setting $\mathbb{F} = 1$, one may solve for f_{\max} (which corresponds to λ_{\max}) as a function of f_{\min} (corresponding to λ_{\min}):

$$\lim_{\mathbb{F} \rightarrow 1} \lambda_{\max} = \begin{cases} \frac{(\log 2)^2}{T^2 \lambda_{\min}} & \text{if } n = -1, \\ \frac{((2 \log 2)^{n+1} - (\lambda_{\min} T)^{n+1})^{1/(n+1)}}{T} & \text{otherwise.} \end{cases} \quad [11]$$

Although this ensures a fair prior (subject to our definition), it does not necessarily ensure an unbiased one. An unbiased prior is defined here as one for which the a priori expectation value of f , given by $E[f] \equiv \int_{f_{\min}}^{f_{\max}} f \Pr(f) df$, equals one-half. After imposing the $\mathbb{F} = 1$ constraint enabled by Eq. 11, we evaluate $E[f]$ in the limit of $f_{\min} \rightarrow 0$ as a function of n . This reveals for $n = -1$ that the expectation value converges to one-half, as desired for an unbiased prior (SI Appendix). The only other value of n in the range $-1 < n < 2$ that yields a fair and unbiased distribution is $n = -0.709\dots$, but this is shown in SI Appendix to require an overly restrictive prior bound limit and thus is not used.

Using these two constraints, our fair and unbiased prior for λ takes the form λ^{-1} , with bounds following the relationship given by Eq. 11. Unfortunately, our two constraints applied to three parameters are insufficient to uniquely define the prior—it is still necessary to choose λ_{\min} subjectively. The posterior could still be argued to be objective if it were found to be broadly insensitive to the choice of λ_{\min} —however, this unfortunately turns out to be false, as shown in what follows.

Resulting Posteriors

To compute marginalized posteriors, we initially tried sampling using Markov chain Monte Carlo and nested sampling techniques, but found that the resulting posteriors were too poorly sampled in the tails. Instead, we directly integrate the posterior density in each parameter. For example, for the λ_L marginalized posterior, we slide along in a fine grid of λ_L values and numerically integrate (using the Gauss-Kronrod rule) the joint posterior density over the limits $\lambda_I = \lambda_{\min}$ to $\lambda_I = \lambda_{\max}$.

The resulting marginalized posteriors are shown in Fig. 2 for two arbitrary choices of λ_{\min} : 10^{-3} Gy^{-1} and 10^{-6} Gy^{-1} . As can be seen from Fig. 2, and perhaps not surprisingly, the resulting distributions are certainly sensitive to the choice of λ_{\min} . In conclusion, we argue here that a λ power-law prior is simply unacceptable as an objective prior for this problem.

The Objective Bernoulli Prior

Fair and Unbiased Priors with Semiinfinite Support. Since we have no way of objectively choosing λ_{\min} , then a fair and unbiased (subject to our definitions) power-law prior in λ does not provide a viable path to defining an objective posterior.

Rather than prescribing a prior in λ space and then evaluating its fairness and bias in f space, we consider here simply writing

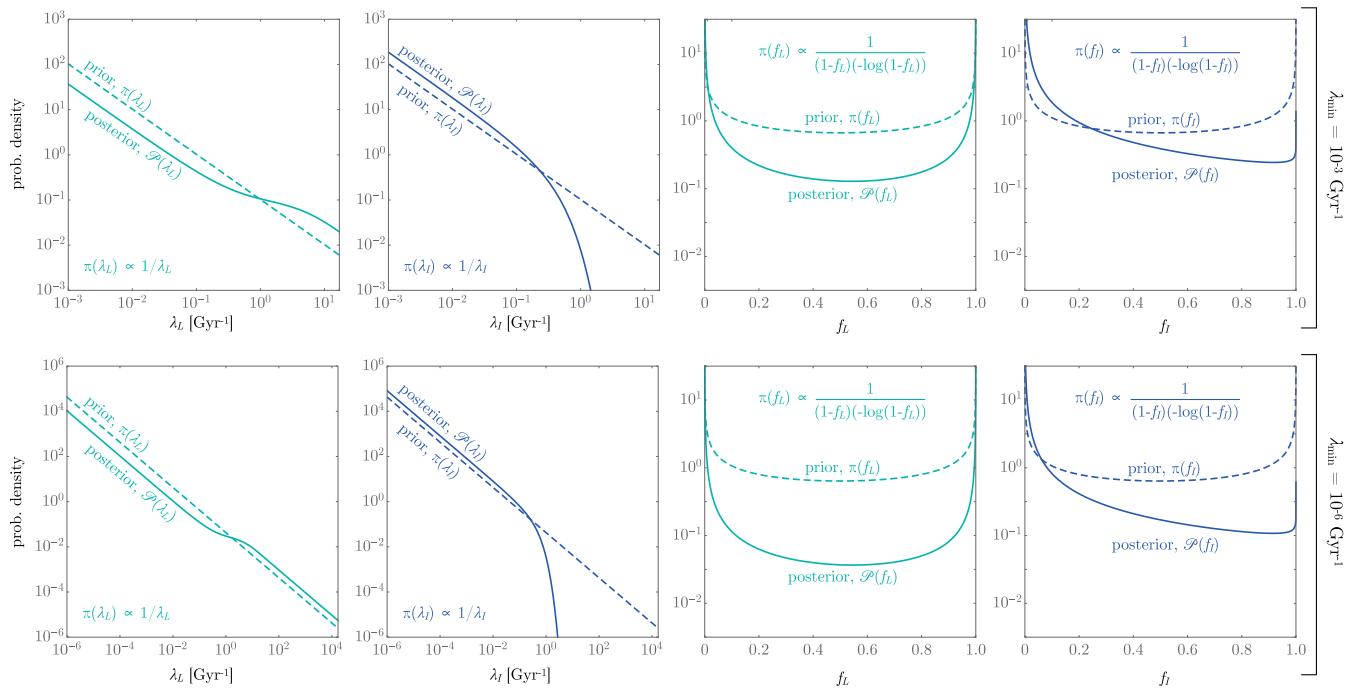


Fig. 2. Marginalized posterior distribution (solid lines) for λ_L (Left), λ_I (Center Left), f_L (Center Right), and f_I (Right) using the log-uniform λ prior (dashed lines) and the optimistic data. Top row assumes $\lambda_{\min} = 10^{-3}$ Gyr $^{-1}$ and Bottom row assumes $\lambda_{\min} = 10^{-6}$ Gyr $^{-1}$. The y-axis scale is chosen to highlight the dynamic range.

down a fair and unbiased prior in f space directly. Since f represents a fraction, it can be interpreted as the Bernoulli probability of a success over the interval T . For a Bernoulli process, a so-called Haldane prior (39) of the form $\propto (f(1-f))^{-1}$ was argued by the objective Bayesian Jaynes (36) to represent the least informative prior. Such a prior is fair and unbiased by construction and places extreme weight on the solutions $f=1$ and $f=0$ at the expense of intermediate values. The intuition behind this is that either a very small fraction of planets will be successful or almost all of them will be, but it is unlikely the laws of nature are tuned such that approximately half of the planets are successful.

Using the Fisher information matrix, one can define the Jeffreys prior for a Bernoulli distribution. The solution is not the Haldane prior, but rather a softer variant of the form $\propto (f(1-f))^{-1/2}$. This translates to a prior in λ space given by

$$\Pr(\lambda) = \frac{T}{\pi \sqrt{e^{\lambda T} - 1}}. \quad [12]$$

Both the Haldane and Jeffreys priors are fair and unbiased with respect to f and indeed any prior of the more general form $\propto (f(1-f))^n$ satisfies these conditions. However, the $n=-1$ case, corresponding to the Haldane prior, is improper and indeed leads to an improper posterior too, but for all other $n > -1$ the prior can be normalized to a finite quantity. Accordingly, the Jeffreys prior is fair, unbiased, objectively defined, and proper over the interval $f = [0, 1]$. Since $f=0$ corresponds to $\lambda=0$ and $f=1$ corresponds to $\lambda=\infty$, this naturally yields a proper prior with semiinfinite support in λ , something which was not possible with the power-law case discussed earlier. Together, these properties make the distribution well suited for our problem and we argue it solves the dilemma faced in earlier work (13).

Although we consider the Jeffreys prior to be the ideal objective prior for our problem, it is instructive to consider posteriors

with $n=0$ (a uniform prior in f) as well, which has a λ -space form of

$$\Pr(\lambda) = Te^{-\lambda T}. \quad [13]$$

Bayes Factors Independent of the λ Prior. Equipped with our likelihood function and prior, one may now sample/integrate the posterior probability distribution to compute marginalized distributions. Marginalization irreversibly bakes the prior into the resulting collapsed posteriors, but specific probabilistic statements can be made in a Bayesian framework without marginalizing. In particular, we consider here an exercise in Bayesian model comparison, where we seek to compare four models, \mathcal{M} , defined as the unique corners of the parameter volume:

- \mathcal{M}_{00} : $\lambda_L \ll 1/t'_L$ and $\lambda_I \ll 1/t'_I$
- \mathcal{M}_{01} : $\lambda_L \ll 1/t'_L$ and $\lambda_I \gg 1/t'_I$
- \mathcal{M}_{10} : $\lambda_L \gg 1/t'_L$ and $\lambda_I \ll 1/t'_I$
- \mathcal{M}_{11} : $\lambda_L \gg 1/t'_L$ and $\lambda_I \gg 1/t'_I$.

Binarizing the parameter volume into these four camps may at first seem arbitrary—What about intermediate values? However, this partitioning is consistent with objective Bayesianism. The objective Bernoulli prior treats life/intelligence as being either very rare or very common, but unlikely to be finely tuned such that it approaches the intermediate value of one-half—thus motivating the models above.

Conditioned upon some available data, \mathcal{D} , one may express the odds ratio between two models as $\Pr(\mathcal{M}_1|\mathcal{D})/\Pr(\mathcal{M}_2|\mathcal{D}) = [\Pr(\mathcal{D}|\mathcal{M}_1)/\Pr(\mathcal{D}|\mathcal{M}_2)][\Pr(\mathcal{M}_1)/\Pr(\mathcal{M}_2)]$. The terms inside the first square bracket are known as the Bayes factor, which equals the odds factor under the simple assumption that no model is a priori preferred over any other. The Bayes factor is the ratio of two “evidences” given by $\mathcal{Z} \equiv \Pr(\mathcal{D}|\mathcal{M})$, and for the four models defined above, \mathcal{Z} can be expressed analytically and independent of the prior $\pi(\lambda_L, \lambda_I)$. This can be seen by noting that, for example with model \mathcal{M}_{00} :

$$\begin{aligned} \mathcal{Z}_{00} &= \Pr(\underbrace{t'_L, t'_I}_{\mathcal{D}} \mid \lambda_L \ll 1/t'_L, \lambda_I \ll 1/t'_I) \overbrace{\Pr(t'_L, t'_I \mid \lambda_L, \lambda_I)}^{\mathcal{M}_{00}} \\ &= \lim_{\lambda_L \ll 1/t'_L} \lim_{\lambda_I \ll 1/t'_I} \Pr(t'_L, t'_I \mid \lambda_L, \lambda_I). \end{aligned} \quad [14]$$

Thus, the evidences of these four corner models are independent of the prior of $\pi(\lambda_L, \lambda_I)$, meaning that even those adopting different priors would consistently agree on the Bayes factors. We note that ref. 13 used a similar strategy and we provide an alternative explanation of this prior-free model comparison in *SI Appendix*, in terms of the Savage–Dickey ratio (40).

In practice, Bayes factors for the two corners with rapid intelligence emergence ($\lambda_I \gg 1/t'_I$) tend to zero, since this is the behavior of the likelihood function (*SI Appendix*). This can be understood by the fact that if the rate of intelligence emergence

were extremely fast, then it would be incompatible with taking as long as it did here on Earth. In contrast, life could emerge much faster than t'_L because t'_L represents only the first appearance of life in the geological record, not the actual date of abiogenesis.

Since these two corners are zero, we instead compare models \mathcal{M}_{10} to \mathcal{M}_{00} . This represents the Bayes factor between a scenario where abiogenesis is a fast versus a slow process conditional upon the premise that intelligence emergence is itself a slow process. In this case, we find

$$\frac{\mathcal{Z}_{10}}{\mathcal{Z}_{00}} = \begin{cases} \frac{T}{2t'_L} & \text{if } T > 2t'_L + t_I, \\ \frac{t'_L}{4(T-t'_I)t'_L - 2t'_L - (T-t'_I)^2} & \text{if } T < 2t'_L + t_I, \end{cases} \quad [15]$$

which evaluates to $(\mathcal{Z}_{10}/\mathcal{Z}_{00}) = 8.73$ and 2.83 for the optimistic and conservative data inputs, respectively. The above also

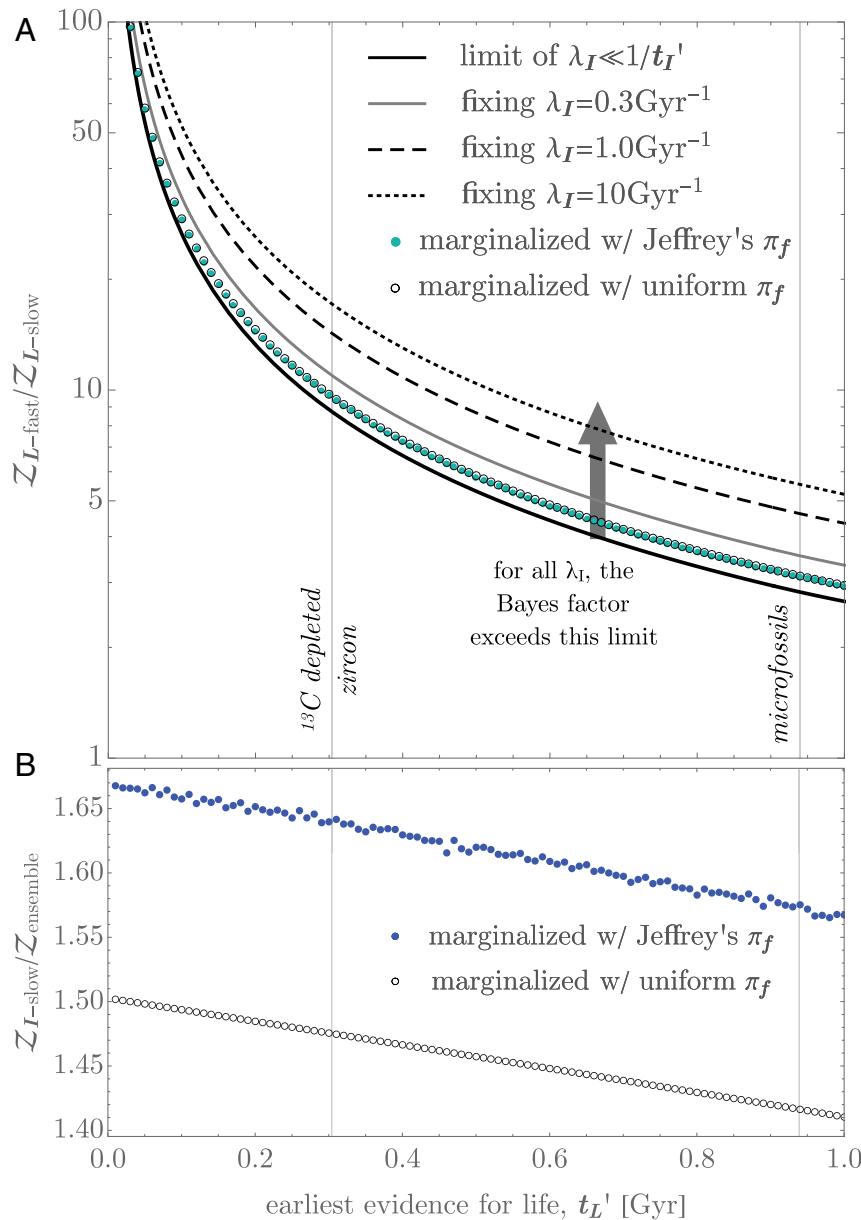


Fig. 3. (A) Bayes factor for a model where life emerges rapidly ($\lambda_L \gg 1/t'_L$) versus slowly ($\lambda_L \ll 1/t'_L$) on Earth. (A) A quick start is favored by at least a factor of 3 conditioned upon early microfossil evidence, independent of our assumptions regarding the evolutionary timescale of intelligent observers and priors on the abiogenesis rate. (B) Bayes factor of a scenario where intelligent observers typically emerge on a much longer timescale than occurred on Earth, versus the ensemble of possibilities. There is a weak preference for a rare intelligence scenario.

reveals the dependency on T is nearly linear; if T is revised significantly up, then optimism for life would also increase. We highlight that $\mathcal{Z}_{10} < \mathcal{Z}_{00}$ if $t'_L > 3.72$ Gy—i.e., if the earliest evidence for life were from no earlier than 680 My ago, we would conclude that abiogenesis was an improbable event.

If we relax the assumption that $\lambda_I \ll 1/t'_I$ and let intelligence become faster, then the Bayes factor monotonically rises, as shown in Fig. 3. This therefore means that the Bayes factor of a quick versus slow abiogenesis scenario must be greater than the limiting case of $\mathcal{Z}_{10}/\mathcal{Z}_{00}$, irrespective of whatever value λ_I takes (or indeed whatever the prior is).

On this basis, we can conclude that even with the most conservative date for the emergence of life, a scenario where abiogenesis occurs rapidly is at least three times more likely than a slow emergence, independent of the priors and even the timescale it takes for intelligence to emerge. If the more ambiguous evidence for an earlier start to life is confirmed (17), then this would increase the odds to a factor of 9, representing relatively strong preference for a model where life would consistently emerge rapidly on Earth, if time were replayed.

Bayes Factors after Marginalization. Thus far we have avoided using the marginalized posteriors, which has the benefit of enabling model comparison independent of the prior $\pi(\lambda_L, \lambda_I)$. However, it also has the disadvantage that we can compare only conditional scenarios. For example, our result for $(\mathcal{Z}_{10}/\mathcal{Z}_{00})$ is a Bayes factor conditional upon the assumption of a slow intelligence emergence. While it turns out this can be interpreted as a lower limit on a fast versus slow abiogenesis scenario, marginalization allows for a calculation which integrates over the uncertainty in λ_I . For example, one may write that the evidence for a model where $\lambda_L \ll 1/t'_L$ (slow abiogenesis) marginalized over λ_I is given by

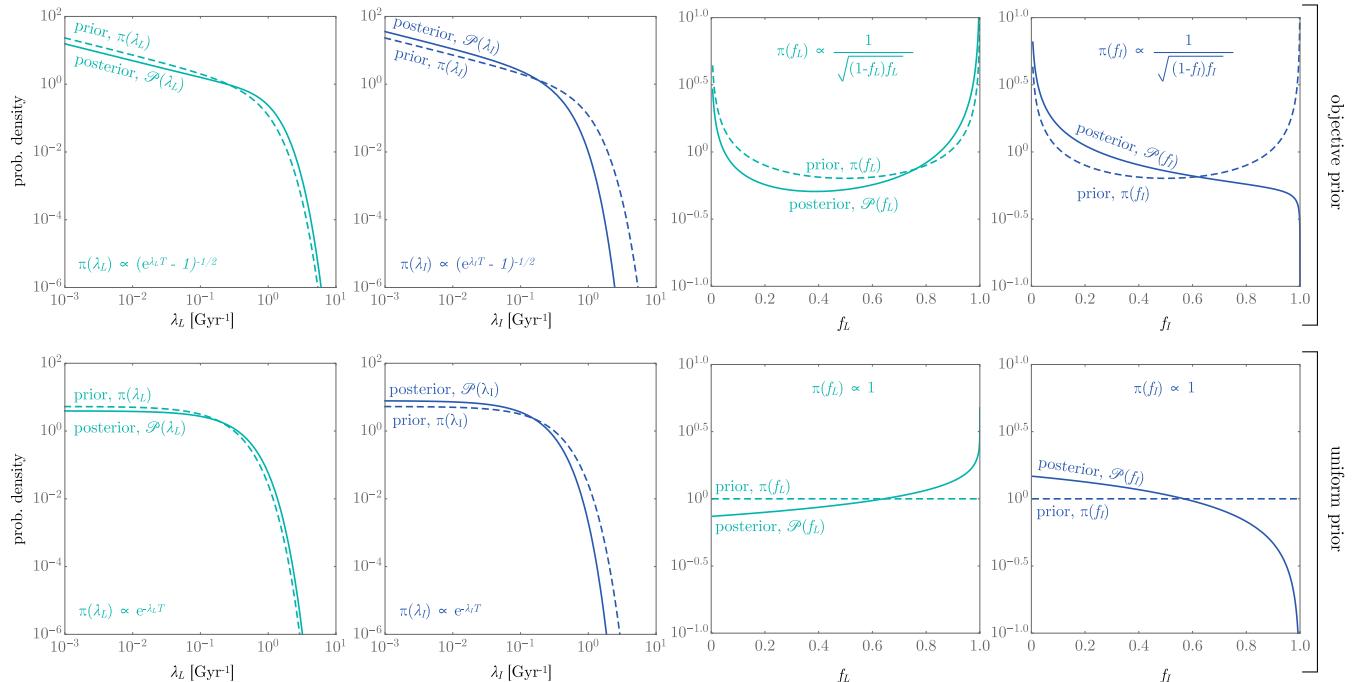


Fig. 4. (Top row) Marginalized posteriors (solid lines) for λ_L (Left), λ_I (Center Left), f_L (Center Right), and f_I (Right) using the objective Bernoulli prior (dashed lines), using the optimistic data. (Bottom row) Same as Top row but using the uniform prior in f for comparison. The y-axis scale is chosen to highlight the dynamic range.

$$\mathcal{Z}_{L-\text{slow}} = \int_{\lambda_I=0}^{\infty} \left(\lim_{\lambda_L \ll 1/t'_L} \mathcal{L} \right) \pi(\lambda_I) d\lambda_I. \quad [16]$$

We numerically evaluated the evidences for $\mathcal{Z}_{L-\text{slow}}$ using the above, $\mathcal{Z}_{L-\text{slow}}$ ($\lambda_L \gg 1/t'_L$) as well as $\mathcal{Z}_{L-\text{slow}}$ ($\lambda_I \ll 1/t'_I$) and the ensemble evidence over all possibilities, $\mathcal{Z}_{\text{ensemble}}$. As before, the fast intelligence emergence scenario has zero evidence since the likelihood tends to zero in this regime, for reasons discussed earlier. The resulting Bayes factors for $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}})$ are shown in Fig. 3A by the solid (Jeffreys prior) and open (uniform prior) circles.

The two sets of points are almost indistinguishable and consistently lie above our previously derived lower limit on the Bayes factor, as expected. Using the optimistic data, we find that $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) = 9.538$ and $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) = 9.648$ for the Jeffreys and uniform priors, respectively, both of which satisfy $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) > (\mathcal{Z}_{10}/\mathcal{Z}_{00}) = 8.73$. For the conservative data, these numbers become $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) = 3.110$ and $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) = 3.137$, again satisfying $(\mathcal{Z}_{L-\text{fast}}/\mathcal{Z}_{L-\text{slow}}) > (\mathcal{Z}_{10}/\mathcal{Z}_{00}) = 2.83$.

Further understanding of these Bayes factors can be gained by evaluating the marginalized posteriors. We numerically marginalize the posteriors in the case of the optimistic data and show the resulting distributions in Fig. 4. From these, one can see that the $f_L \rightarrow 0$ limit drops below the prior, whereas the $f_L \rightarrow 1$ limit rises above it. Together, these results paint a consistent picture that the timing of life's emergence and that of intelligent observers favor the hypothesis that life would likely reemerge rapidly on Earth were the clock to be rerun.

And What of Intelligence? Thus far, we have calculated Bayes factors concerning fast versus slow abiogenesis rates. For the rate of emergence of intelligent observers, Bayes factors against a fast emergence scenario tend to infinity, since $\lim_{\lambda_I \gg 1/t'_I} \mathcal{L} \rightarrow 0$ (as shown in *SI Appendix*). Instead, it is more useful to compare the slow intelligence scenario ($\lambda_I \ll 1/t'_I$) against the ensemble of

models where λ_I can take any value, for which the Bayes factor is

$$\frac{\mathcal{Z}_{\text{I-slow}}}{\mathcal{Z}_{\text{ensemble}}} = \frac{\int_{\lambda_L=0}^{\infty} \left(\lim_{\lambda_I \ll 1/t'_L} \mathcal{L}(\lambda_L, \lambda_I) \right) \pi(\lambda_L) d\lambda_L}{\int_{\lambda_L=0}^{\infty} \int_{\lambda_I=0}^{\infty} \mathcal{L}(\lambda_L, \lambda_I) \pi(\lambda_L) \pi(\lambda_I) d\lambda_I d\lambda_L}. \quad [17]$$

We evaluated the above numerically using optimistic data to yield 1.638 and 1.474 using the Jeffreys and uniform priors, respectively. Switching to the conservative data points barely affects these numbers with 1.572 and 1.416 (the full range of possibilities is depicted in Fig. 3B). This suggests a slight preference, 3:2 betting odds, that intelligent observers would rarely reemerge—a value which is broadly robust against the two priors considered and the range of possible t'_L values.

Conclusions

In this work, we have attempted to build upon the seminal paper of ref. 13 which devised a Bayesian formalism for interpreting life's early emergence on Earth. Unlike that work, we do not treat the timescale for life to develop into intelligent observers as a fixed quantity, but rather infer it jointly as a free parameter. This important difference feeds into an overall theme of the analysis presented here—to present an objective Bayesian analysis of life's early emergence on Earth and our relatively late arrival within the context of Earth's habitable window.

In this vein, we have demonstrated that the commonly used power-law prior for this problem is not objective as the results strongly depend on arbitrary choices on the prior's domain. We show that priors on the Bernoulli probability of life/intelligence emerging naturally provide semiinfinite support and yield distribution which can be seen to be fair and unbiased for this problem. Even so, it is possible to derive numerous model comparison results which are fully independent of these priors—meaning that even those using wildly different priors would consistently agree on the results.

The early emergence of life on Earth is naively interpreted as meaning that if we reran the tape, life would generally reappear quickly. But if the timescale for intelligence evolution is long, then a quick start to life is simply a necessary byproduct of our existence—not evidence for a general rapid abiogenesis rate. Using our objective Bayesian framework, we show that the Bayes factor between a fast versus a slow abiogenesis scenario is at least a factor of 3—irrespective of the prior or the timescale for intelligence evolution. This factor is boosted to 9 when we replace the

earliest microfossil evidence (15, 16, 18) with the more disputed ^{13}C -depleted zircon deposits reported by ref. 17. These results are also supported by marginalizing over our objective priors. An additional objective result concerning abiogenesis is that the maximum-likelihood timescale for life to first appear is 190 My after conditions became habitable (4.21 Gy ago) using the microfossil evidence or even just 46 My using the more disputed data. It is emphasized that these results are conditioned solely upon the chronology data concerning life.

We find that the rate at which intelligent observers evolve is less well constrained. Certainly, the possibility that the rate of intelligence emergence is rapid (much less than gigayears) is strongly excluded, which is not surprising given that it took several gigayears here on Earth. But the possibility that intelligence is extremely rare and Earth “lucked out” remains quite viable. Overall, we find a weak preference, 3:2 betting odds, that intelligence rarely emerges given our late arrival.

It is tempting to apply these numbers to potentially habitable exoplanets being discovered. However, we caution that our analysis purely concerns the Earth, treating abiogenesis as a stochastic process against a backdrop of events and conditions which might be plausibly unique to Earth. If conditions sufficiently similar to the early conditions exist and sustain on other worlds for 1 Gy or more, then our analysis would then favor the hypothesis that life is common, by a factor of $K > 3$. However, the alternative is clearly not discounted and our Bayes factor does not cross the threshold to which it would be conventionally described as “strong” ($K > 10$) or “decisive” ($K > 100$) evidence (41). Yet, future revision regarding the earliest evidence for life could plausibly trigger this.

Overall, our work supports an optimistic outlook for future searches for biosignatures (4–7). The slight preference for a rare intelligence scenario is consistent with a straightforward resolution to the Fermi paradox. However, our work says nothing about the lifetime of civilizations, and indeed the weight of evidence in favor of this scenario is sufficiently weak that searches for technosignatures should certainly be a component in observational campaigns seeking to resolve this grand mystery.

Data Availability. All data used in this work are fully stated in the text of this paper.

ACKNOWLEDGMENTS. D.K. acknowledges support from the Alfred P. Sloan Foundation. I thank Jason Wright for helpful comments on an early draft. I especially thank the anonymous reviewers for their constructive feedback.

1. D. S. McKay *et al.*, Search for past life on Mars: Possible relic biogenic activity in Martian meteorite ALH84001. *Science* **273**, 924–930 (1996).
2. R. Navarro-González, E. Vargas, J. de la Rosa, A. C. Raga, C. P. McKay, Reanalysis of the Viking results suggests perchlorate and organics at midlatitudes on Mars. *J. Geophys. Res.* **115**, E12010 (2010).
3. J. H. Waite *et al.*, Cassini finds molecular hydrogen in the Enceladus plume: Evidence for hydrothermal processes. *Science* **356**, 155–159 (2017).
4. D. J. Des Marais *et al.*, Remote sensing of planetary properties and biosignatures on extrasolar terrestrial planets. *Astrobiology* **2**, 153–181 (2002).
5. E. W. Schwambier *et al.*, Exoplanet biosignatures: A review of remotely detectable signs of life. *Astrobiology* **18**, 663–708 (2018).
6. J. Krissansen-Totton, S. Olson, D. C. Catling, Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life. *Sci. Adv.* **4**, eaao5747 (2018).
7. Y. Fujii *et al.*, Exoplanet biosignatures: Observational prospects. *Astrobiology* **18**, 739–778 (2018).
8. G. Cocconi, P. Morrison, Searching for interstellar communications. *Nature* **184**, 844–846 (1959).
9. F. J. Dyson, Search for artificial stellar sources of infrared radiation. *Science* **131**, 1667–1668 (1960).
10. J. T. Wright, K. M. S. Cartier, M. Zhao, D. Jontof-Hutter, E. B. Ford, The search for extraterrestrial civilizations with large energy supplies. IV. The signatures and information content of transiting megastructures. *Astrophys. J.* **816**, 17 (2016).
11. D. Kipping, Transiting quasites as a possible technosignature. *Res. Notes Am. Astronom. Soc.* **3**, 91 (2019).
12. J. T. Wright, S. Kanodia, E. Lubar, How much SETI has been done? Finding needles in the n-dimensional cosmic haystack. *AJNR* **156**, 260 (2018).
13. D. S. Spiegel, E. L. Turner, Bayesian analysis of the astrobiological implications of life's early emergence on Earth. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 395–400 (2012).
14. F. Simpson, The size distribution of inhabited planets. *MNRAS* **456**, L59–L63 (2016).
15. J. W. Schopf, Fossil evidence of Archean life. *Philos. Trans. Biol. Sci.* **361**, 869–885 (2006).
16. J. W. Schopf, A. B. Kudryavtsev, A. D. Czaja, A. B. Tripathi, Evidence of Archean life: Stromatolites and microfossils. *Precambrian Res.* **158**, 141–155 (2007).
17. E. A. Bell, P. Boehnke, T. M. Harrison, W. L. Mao, Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14518–14521 (2015).
18. J. W. Schopf, K. Kitajima, M. J. Spicuzza, A. B. Kudryavtsev, J. W. Valley, SIMS analyses of the oldest known assemblage of microfossils document their taxon-correlated carbon isotope compositions. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 53–58 (2018).
19. A. C. Allwood, Evidence of life in Earth's oldest rocks. *Nature* **537**, 500–501 (2016).
20. B. Carter, Five- or six-step scenario for evolution? *Int. J. Astrobiol.* **7**, 177–182 (2008).
21. C. Scharf, L. Cronin, Quantifying the origins of life on a planetary scale. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8127–8132 (2016).
22. J. Chen, D. Kipping, On the rate of abiogenesis from a Bayesian informatics perspective. *Astrobiology* **18**, 1574–1584 (2018).
23. G. A. Schmidt, A. Frank, The Silurian hypothesis: Would it be possible to detect an industrial civilization in the geological record? *Int. J. Astrobiol.* **18**, 142–150 (2019).
24. J. T. Wright, Prior indigenous technological species. *Int. J. Astrobiol.* **17**, 96–100 (2018).
25. N. H. Sleep, K. Zahnle, Carbon dioxide cycling and implications for climate on ancient Earth. *J. Geophys. Res. Plan.* **106**, 1373–1399 (2001).
26. N. H. Sleep, A. M. Hessler, Weathering of quartz as an Archean climatic indicator. *Earth Planet Sci. Lett.* **241**, 594–602 (2006).

27. N. T. Arndt, E. G. Nisbet, Processes on the young Earth and the habitats of early life. *Annu. Rev. Earth Planet Sci.* **40**, 521–549 (2012).
28. W. K. Hartmann, D. R. Davis, Satellite-sized planetesimals and lunar origin. *Icarus* **24**, 504–515 (1975).
29. S. A. Benner *et al.*, When did life likely emerge on Earth in an RNA-first process? *ChemSystemsChem* **2**, e1900035 (2020).
30. S. A. Wilde, J. W. Valley, W. H. Peck, C. M. Graham, Evidence from detrital zircons for the existence of continental crust and oceans on the earth 4.4 Gyr ago. *Nature* **409**, 175–178 (2001).
31. J. T. O’Malley-James, J. S. Greaves, J. A. Raven, C. S. Cockell, Swansong biospheres: Refuges for life and novel microbial biospheres on terrestrial planets near the end of their habitable lifetimes. *Int. J. Astrobiol.* **12**, 99–112 (2013).
32. R. W. Pearcy, J. Ehleringer, Comparative ecophysiology of C3 and C4 plants. *Plant Cell Environ.* **7**, 1–13 (1984).
33. K. Caldeira, J. F. Kasting, The life span of the biosphere revisited. *Nature* **360**, 721–723 (1992).
34. P. G. Falkowski *et al.*, The rise of oxygen over the past 205 million years and the evolution of large placental mammals. *Science* **309**, 2202–2204 (2005).
35. H. Jeffreys, An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. Math. Phys. Sci.* **186**, 453–461 (1946).
36. E. T. Jaynes, Prior probabilities. *IEEE Trans. Syst. Sci. Cybern.* **4**, 227–241 (1968).
37. B. d. Finetti, *Theory of Probability: A Critical Introductory Treatment* (John Wiley & Sons, Ltd, 1975).
38. F. D. Drake, *The Radio Search for Intelligent Extraterrestrial Life*, G. Mamikunian, M. H. Briggs, Eds. (Wiley, Hoboken, NJ, 1965), chap. 2, pp. 323–345.
39. J. B. S. Haldane, A note on inverse probability. *Proc. Camb. Philos. Soc.* **28**, 55–61 (1932).
40. J. M.Dickey, The weighted likelihood ratio, linear hypotheses on normal location parameters. *Ann. Math. Stat.* **42**, 204–223 (1971).
41. R. E. Kass, A. E. Raftery, Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).

STATISTICS

EVOLUTION