Priors for the Bayesian star paradox

Mikael Falconnet

PRIORS FOR THE BAYESIAN STAR PARADOX

MIKAEL FALCONNET

ABSTRACT. We show that the Bayesian star paradox, first proved mathematically by Steel and Matsen for a specific class of prior distributions, occurs in a wider context including less regular, possibly discontinuous, prior distributions.

INTRODUCTION

In phylogenetics, a particular resolved tree can be highly supported even when the data is generated by an unresolved star tree. This unfortunate aspect of the Bayesian approach to phylogeny reconstruction is called the *star paradox*. Recent studies highlight that the paradox can occur in the simplest nontrivial setting, namely for an unresolved rooted tree on three taxa and two states, see Yang and Rannala [7] and Lewis et al. [1]. Kolaczkowski and Thornton [2] presented some simulations and suggested that artifactual high posteriors for a particular resolved tree might disappear for very long sequences. Previous simulations in [7] were plagued by numerical problems, which left unknown the nature of the limiting distribution on posterior probabilities. For an introduction to the Bayesian approach to phylogeny reconstruction we refer to chapter 5 of Yang [5].

The statistical question which supports the star paradox is whether the Bayesian posterior distribution of the resolutions of a star tree becomes uniform when the length of the sequence tends to infinity, that is, in the case of three taxa and two states, whether the posterior distribution of each resolution converges to $1/3$. In a recent paper, Steel and Matsen [3] disprove this, thus ruining Kolaczkowski and Thornton's hope, for a specific class of branch length priors which they call *tame*. More precisely, Steel and Matsen show that, for every tame prior and every fixed $\varepsilon > 0$, the posterior probability of any of the three possible trees stays above $1 - \varepsilon$ with non vanishing probability when the length of the sequence goes to infinity. This result was recognized by Yang [6] and reinforced by theoretical results on the posterior probabilities by Susko [4].

Our main result is that Steel and Matsen's conclusion holds for a wider class of priors, possibly highly irregular, which we call *tempered*. Recall that Steel and Matsen consider smooth priors whose densities satisfy some regularity conditions.

The paper is organized as follows. In Section 1, we describe the Bayesian framework of the star paradox. In Section 2, we define the class of tempered priors on the branch lengths and we state our main result. In Section 3, we state an extension of a technical lemma due to Steel and Matsen, which allows us to extend their result. In Section 4, we prove our main result. Section 5 is devoted to the proofs of intermediate results. In Appendix A, we prove that every tame prior, in Steel and Matsen's sense, is tempered, in the sense of this

Date: November 19, 2010.
2000 *Mathematics Subject Classification.* Primary: 60J28; 92D15. Secondary: 62C10.
Key words and phrases. Phylogenetic trees, Bayesian statistics, star trees.

paper, and we provide examples of tempered, but not tame, prior distributions. Finally, in Appendix B, we prove the extension of Steel and Matsen's technical lemma stated in Section 3.

1. Bayesian framework for rooted trees on three taxa

We consider three taxa, encoded by the set $\tau = \{1,2,3\}$, with two possible states. Phylogenies on τ are supported by one of the four following trees: the star tree R_0 on three taxa and, for every taxon i in τ, the tree R_i such that i is the outlier. Relying on a commonly used notation, this reads as

$$R_1 = (1,(2,3)), \quad R_2 = (2,(1,3)), \quad R_3 = (3,(1,2)).$$

The phylogeny based on R_0 is specified by the common length of its three branches, denoted by t. For each i in τ, the phylogeny based on R_i is specified by a pair of branch lengths $(t_\mathrm{e}, t_\mathrm{i})$, where t_e denotes the external branch length and t_i the internal branch length, see figure 1.

For instance, in the phylogeny based on R_1, the divergence of taxa 2 and 3 occurred t_e units of time ago and the divergence of taxon 1 and a common ancestor of taxa 2 and 3 occurred $t_\mathrm{i} + t_\mathrm{e}$ units of time ago.

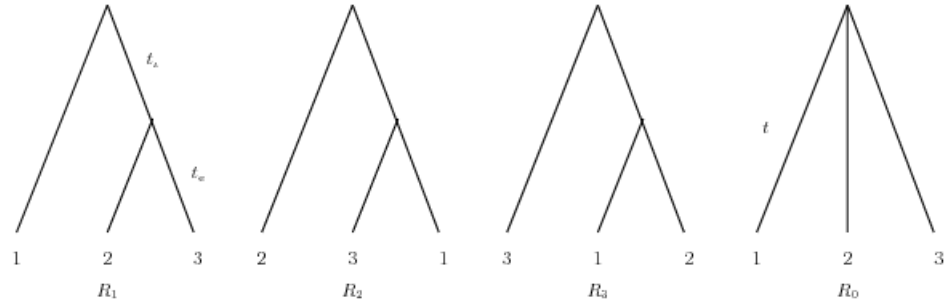

FIGURE 1. The four rooted trees for three species.

We assume that the sequences evolve according to a two-state continuous-time Markov process with equal substitution rates (which we may take to equal 1) between the two character states.

Four site patterns can occur. The first one, denoted by s_0, is such that a given site coincides in the three taxa. The three others, denoted by s_i with i in τ, are such that a given site coincide in two taxa and is different in the third taxon, which is taxon i. In other words, if one writes the site patterns in taxa 1, 2 and 3 in this order and x and y for any two different characters,

$$s_0 = xxx, \quad s_1 = yxx, \quad s_2 = xyx, \quad \text{and} \quad s_3 = xxy.$$

Let $\{s_0, s_1, s_2, s_3\}$ denote the set of site patterns in the specific case described above of three taxa and two states evolving in a two-state symmetric model. Assume that the counting of site pattern s_i is n_i. Then $n = n_0 + n_1 + n_2 + n_3$ is the total length of the sequences and, in the independent two-state symmetric model considered in this paper, the quadruple (n_0, n_1, n_2, n_3) is a sufficient statistics of the sequence data. We use the letter $\mathfrak{n}$ to denote any quadruple (n_0, n_1, n_2, n_3) of nonnegative integers such that $|\mathfrak{n}| = n_0 + n_1 + n_2 + n_3 = n \geqslant 1$.

For every site pattern s_i and every branch lengths $(t_{\mathrm{e}},t_{\mathrm{i}})$, let $p_i(t_{\mathrm{e}},t_{\mathrm{i}})$ denote the probability that s_i occurs on tree R_1 with branch lengths $(t_{\mathrm{e}},t_{\mathrm{i}})$. Standard computations provided by Yang and Rannala [7] show that

$$4p_0(t_{\mathrm{e}},t_{\mathrm{i}}) = 1+\mathrm{e}^{-4t_{\mathrm{e}}}+2\mathrm{e}^{-4(t_{\mathrm{i}}+t_{\mathrm{e}})},$$
$$4p_1(t_{\mathrm{e}},t_{\mathrm{i}}) = 1+\mathrm{e}^{-4t_{\mathrm{e}}}-2\mathrm{e}^{-4(t_{\mathrm{i}}+t_{\mathrm{e}})},$$
$$4p_2(t_{\mathrm{e}},t_{\mathrm{i}}) = 4p_3(t_{\mathrm{e}},t_{\mathrm{i}}) = 1-\mathrm{e}^{-4t_{\mathrm{e}}}.$$

Let $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$ denote a pair of positive random variables representing the branch lengths $(t_{\mathrm{e}},t_{\mathrm{i}})$, and $\mathfrak{N}=(N_0,N_1,N_2,N_3)$ denote a quadruple of integer random variables representing the counts of sites patterns $\mathfrak{n}=(n_0,n_1,n_2,n_3)$.

2. The star tree paradox

Assuming that every taxon evolved from a common ancestor, the aim of phylogeny reconstruction is to compute the most likely tree R_i. To do so, in the Bayesian approach, one places prior distributions on the trees R_i and on their branch lengths $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$.

2.1. **Main result.** Let $\mathbb{P}(\mathfrak{N}=\mathfrak{n}|R_i,\mathfrak{T})$ denote the probability that $\mathfrak{N}=\mathfrak{n}$ assuming that the data is generated along the tree R_i conditionally on the branch lengths $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$. One may consider R_1 only since, for every $\mathfrak{n}=(n_0,n_1,n_2,n_3)$, the symmetries of the setting yield the relations

$$\mathbb{P}(\mathfrak{N}=\mathfrak{n}|R_2,\mathfrak{T}) = \mathbb{P}(\mathfrak{N}=(n_0,n_2,n_3,n_1)|R_1,\mathfrak{T}),$$

and

$$\mathbb{P}(\mathfrak{N}=\mathfrak{n}|R_3,\mathfrak{T}) = \mathbb{P}(\mathfrak{N}=(n_0,n_3,n_1,n_2)|R_1,\mathfrak{T}).$$

Notation 2.1. *For every site pattern s_i, let P_i denote the random variable*

$$P_i = p_i(\mathfrak{T}) = p_i(T_{\mathrm{e}},T_{\mathrm{i}}).$$

For every i in τ and every $\mathfrak{n}$, let $\Pi_i(\mathfrak{n})$ denote the random variable

$$\Pi_i(\mathfrak{n}) = P_0^{n_0}P_1^{n_i}P_2^{n_j+n_k}, \quad \text{with} \quad \{i,j,k\}=\tau.$$

We recall that $P_2=P_3$ and we note that, if $|\mathfrak{n}|=n_0+n_1+n_2+n_3=n$ with $n\geqslant 1$, then, for every i in τ,

$$\Pi_i(\mathfrak{n}) = P_0^{n_0}P_1^{n_i}P_2^{n-n_0-n_i}.$$

Fix $\mathfrak{n}$ and assume that $|\mathfrak{n}|=n_0+n_1+n_2+n_3=n$ with $n\geqslant 1$. For every i in τ, the posterior probability of R_i conditionally on $\mathfrak{N}=\mathfrak{n}$ is

$$\mathbb{P}(R_i|\mathfrak{N}=\mathfrak{n}) = \frac{n!}{n_0!n_1!n_2!n_3!}\,\frac{1}{\mathbb{P}(\mathfrak{N}=\mathfrak{n})}\,\mathbb{E}(\Pi_i(\mathfrak{n})).$$

Thus, for every i and j in τ,

$$\frac{\mathbb{P}(R_i|\mathfrak{N}=\mathfrak{n})}{\mathbb{P}(R_j|\mathfrak{N}=\mathfrak{n})} = \frac{\mathbb{E}(\Pi_i(\mathfrak{n}))}{\mathbb{E}(\Pi_j(\mathfrak{n}))}.$$

For every $\varepsilon>0$ and every i in τ, let $\mathscr{N}_i^{\varepsilon}$ denote the set of $\mathfrak{n}$ such that, for both indices j in τ such that $j\neq i$,

$$\mathbb{E}(\Pi_i(\mathfrak{n})) \geqslant (2/\varepsilon)\,\mathbb{E}(\Pi_j(\mathfrak{n})).$$

One sees that, for every i in τ and $\mathfrak{n}$ in $\mathscr{N}_i^{\varepsilon}$,

$$\mathbb{P}(R_i|\mathfrak{N}=\mathfrak{n}) \geqslant 1-\varepsilon,$$

which means that the posterior probability of tree R_i among the three possible trees is highly supported.

Recall that, under hypothesis R_0 and for a tame prior distribution on $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$, Steel and Matsen prove that, for every i in τ, $\mathbb{P}(\mathfrak{N}\in\mathscr{N}_i^{\varepsilon})$ does not go to 0 when the sequence length n goes to infinity, and consequently that the posterior probability $\mathbb{P}(R_i|\mathfrak{N})$ can be close to 1 even when the sequence length n is large.

As stated in the introduction, our aim is to prove the same result for tempered prior distributions of $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$, which we now define.

Notation 2.2. *(1) For every $s\in[0,1]$ and $z\in[0,3]$, let*

$$G(z,s)=\mathbb{P}\left(\mathrm{e}^{-4T_{\mathrm{e}}}(1-\mathrm{e}^{-4T_{\mathrm{i}}})\leqslant s\,|\,\mathrm{e}^{-4T_{\mathrm{e}}}(1+2\mathrm{e}^{-4T_{\mathrm{i}}})=z\right).$$

(2) For every positive t and every site pattern s_i, let q_i denote the probability that s_i occurs on tree R_0, hence

$$4q_0=4p_0(0,t)=1+3\mathrm{e}^{-4t},\quad 4q_1=4q_2=4q_3=1-\mathrm{e}^{-4t}.$$

(3) Let ℓ_t denote a positive real number such that $1<4q_0-\ell_t$ and $4q_0+\ell_t<4$, for instance $\ell_t=3\mathrm{e}^{-4t}(1-\mathrm{e}^{-4t})$. Let I and I_t denote the intervals

$$I=[0,3],\quad I_t=[4q_0-1-\ell_t,4q_0-1+\ell_t]\subset]0,3[.$$

(4) For every positive t and integer n, let

$$Q_n(t)=\mathbb{P}\left(T_{\mathrm{i}}\leqslant 1/n,\,t\leqslant T_{\mathrm{e}}\leqslant t+1/n\right).$$

Definition 2.3 (Tempered priors)**.** *The distribution of $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$ is tempered if the following two conditions hold.*

(1) *For every t, there exists a real number s_0 in $]0,1]$, an interval I_t around $4q_0-1$, some bounded functions F_i, some positive numbers α and κ, an integer $k\geqslant 1$ and some real numbers ε_i such that*

$$0=\varepsilon_0<\varepsilon_1<\cdots<\varepsilon_{k-1}\leqslant 2<\varepsilon_k,$$

and such that for every s in $[0,s_0]$ and every z in I_t,

$$\left|G(z,s)-\sum_{i=0}^{k-1}F_i(z)s^{\alpha+\varepsilon_i}\right|\leqslant \kappa s^{\alpha+\varepsilon_k}.$$

(2) *For every positive t, $n^{-1}\log Q_n(t)\to 0$ when $n\to\infty$.*

We detail the properties involved in Definition 2.3 and provide examples of tempered priors in subsection 2.2 below.

We now state our main result, which is an extension of Steel and Matsen's result to our more general setting.

Theorem 2.4. *Consider sequences of length n generated by a star tree R_0 on 3 taxa with strictly positive edge length t. Let $\mathfrak{N}$ be the resulting data, summarized by site pattern counts. Consider any prior on the three resolved trees (R_1,R_2,R_3) which assigns strictly positive probability to each tree, and a tempered prior distribution on their branch lengths $\mathfrak{T}=(T_{\mathrm{e}},T_{\mathrm{i}})$.*

Then, for every i in τ and every positive ε, there exists a positive δ such that, when n is large enough,

$$\mathbb{P}\left(\mathbb{P}(R_i|\mathfrak{N}) \geqslant 1-\varepsilon\right) \geqslant \delta.$$

We prove Theorem 2.4 in Section 4.

2.2. **Motivation and intuitive understanding of Definition 2.3.** In Definition 2.3, condition 2 is easy to describe, to illustrate and to check, while the content of condition 1 might be more difficult to grasp. Condition 1 involves a Taylor expansion around $s=0$ of the conditional cumulative distribution function $s \mapsto G(z,s)$, where the Taylor coefficients depend on z. Such a Taylor expansion roughly describes the prior distribution when $t_{\mathrm{i}} \to 0$ and when t_{e} is roughly constant. The precise definition of $G(z,\cdot)$ and the technical result stated in Proposition 3.2 are both dictated by our approach to the proof of Theorem 2.4. A key hypothesis is that $\varepsilon_0 = 0$ while $\varepsilon_k > 2$, which means that we are given a limited expansion of $s \mapsto G(z,s)$ up to a better order than s^2 when $s \to 0$.

At this point, the reader can wonder how to check if a given prior is tempered or not and if the verification is simply possible in concrete cases, given the convoluted aspect of this definition. Hence we now present some explicit examples of tempered priors. We begin with the following result.

Proposition 2.5. *Assume that $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ has a smooth joint probability density, bounded and everywhere non zero. Then the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is tempered.*

As a consequence, every tame prior fulfills the hypothesis of Proposition 2.5, hence every tame prior is tempered, as claimed in the introduction. This case includes the exponential priors discussed in [7]. We prove Proposition 2.5 in Appendix A.

However some tempered priors are not tame, as illustrated by the following example where Steel and Matsen's condition fails.

Definition 2.6. *Let $a > 0$ and $b > 0$. Let (t_n), (y_n) and (r_n) denote sequences of positive numbers, indexed by $n \geqslant 1$ and defined by the formulas*

$$t_n = n^{-a}, \quad y_n = 1 + 2\mathrm{e}^{-4t_n} \quad r_n = y_n\left(n^{-b} - (n+1)^{-b}\right).$$

Finally, let

$$r = \sum_{n \geqslant 1} r_n.$$

Proposition 2.7. *In the setting of Definition 2.6, assume the following:*

(i) $3a < \min\{1, b\}$.
(ii) *The random variable T_{i} is discrete and such that, for every $n \geqslant 1$,*

$$\mathbb{P}(T_{\mathrm{i}} = t_n) = r_n/r.$$

(iii) *The random variable T_{e} is continuous, independent of T_{i}, with exponential law of parameter 4, that is, with density $4\mathrm{e}^{-4t}$ on $t \geqslant 0$ with respect to the Lebesgue measure.*

Then, the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is not tame but it is tempered, for the parameters

$$k = 3, \quad \alpha = b/a, \quad \varepsilon_1 = 1, \quad \varepsilon_2 = 2, \quad \varepsilon_3 = 3.$$

Since the distribution of T_{i} is an accumulation of Dirac masses, the prior distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ cannot be tame.

Yet, the fact that the prior distribution is tempered does not come only from the fact that the distribution of T_{i} is discrete. For a degenerate example, if $T_{\mathrm{i}} = 0$ almost surely, then $G(z,s) = 1$ for every $s \geqslant 0$, and $G(z,\cdot)$ has no Taylor expansion around zero whose first term is a positive power of s. Note that in this particular case, the Bayesian star paradox does not occur.

However, under the conditions of Proposition 2.7, $G(z,\cdot)$ has a Taylor expansion at 0 fulfilling condition 1 of Definition 2.3. We prove this in Appendix A.

We provide below some examples of less ill-behaved distributions which are tempered but not tame, and an example of a distribution which does not fulfill condition 1, hence is not tempered.

Proposition 2.8. *Assume that T_{e} is a continuous random variable, with exponential law of parameter 4, that is, with density $4\,\mathrm{e}^{-4t}$ on $t \geqslant 0$ with respect to the Lebesgue measure, and that T_{i} is a random variable independent of T_{e}. Then, the following holds.*

(i) *If the distribution of T_{i} is uniform on $[0,\theta]$, with $\theta > 0$, the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is tempered but not tame.*
(ii) *If the distribution of T_{i} has density $\theta t_{\mathrm{i}}^{\theta-1}$ on the interval $[0,1]$, for a given θ in $(0,1)$, the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is tempered but not tame.*
(iii) *If the distribution of T_{i} has density $\log(1/t_{\mathrm{i}})$ on the interval $[0,1]$, the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is not tempered.*
(iv) *If the distribution of T_{i} has density $4t_{\mathrm{i}}\log(1/t_{\mathrm{i}})$ on the interval $[0,1]$, the distribution of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is not tempered.*

Note that in case (iv), the density function of $\mathfrak{T} = (T_{\mathrm{e}}, T_{\mathrm{i}})$ is bounded, non smooth but continuous, but the distribution is not tempered.

We prove Proposition 2.8 in Appendix A.

3. Extension of Steel and Matsen's lemma

The Bayesian star paradox due to Steel and Matsen relies on a technical result which we slightly rephrase as follows. For every nonnegative real t and every $[0,1]$ valued random variable V, introduce

$$M_t = \mathbb{E}(V^t), \quad R_t = 1 - \frac{M_{t+1}}{M_t} = \frac{\mathbb{E}(V^t(1-V))}{\mathbb{E}(V^t)}.$$

Proposition 3.1 (Steel and Matsen's lemma)**.** *Let $0 \leqslant \eta < 1$ and $B > 0$. There exists a finite K, which depends on η and B only, such that the following holds. For every $[0,1]$ valued random variable V with a smooth probability density function f such that $f(1) > 0$ and $|f'(v)| \leqslant Bf(1)$ for every $\eta \leqslant v \leqslant 1$, and for every integer $k \geqslant K$,*

$$2kR_k \geqslant 1.$$

Indeed the asymptotics of R_k when k is large depends on the behaviour of the distribution of V around 1.

Our next proposition proves that the conclusion of Steel and Matsen's lemma above holds for a wider class of random variables.

Proposition 3.2. *Let V a random variable on $[0,1]$. Suppose that there exists an integer $n \geqslant 1$ and real numbers $0 \leqslant v_0 < 1$, $\alpha > 0$, ε_i and γ_i, such that*

$$0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_{n-1} \leqslant 1 < \varepsilon_n,$$

and, for every $v_0 \leqslant v \leqslant 1$,

$$\left| \mathbb{P}(V \geqslant v) - \sum_{i=0}^{n-1} \gamma_i (1-v)^{\alpha+\varepsilon_i} \right| \leqslant \gamma_n (1-v)^{\alpha+\varepsilon_n}.$$

Then there exists a finite $\theta(\gamma)$, which depends continuously on $\gamma = (\gamma_0, \ldots, \gamma_n)$, such that for every $t \geqslant \theta(\gamma)$,

$$2tR_t \geqslant \alpha.$$

Remark 1. *We insist on the fact that $\theta(\gamma)$ depends continuously on the multiparameter $\gamma = (\gamma_0, \ldots, \gamma_n)$. To wit, in the proof of Proposition 5.6, we apply Proposition 3.2 with bounded functions of z. This means that for every z in I_t, one gets a number θ which depends on z through the bounded functions such that the control on the distribution of V holds. The continuity of θ ensures that there exists a number independent of z such that Proposition 5.6 holds.*

Remark 2. *If one computes a Taylor expansion of the function $v \mapsto \mathbb{P}(V \geqslant v)$ at $v = 1^-$ under the conditions of Steel and Matsen's lemma, one sees that conditions of Proposition 3.2 hold. Hence Proposition 3.2 is an extension of Steel and Matsen's lemma.*

We prove Proposition 3.2 in Appendix B. The proof of Theorem 2.4 relies on it.

4. Synopsis of the proof of Theorem 2.4

This section is devoted to a sketch of the proof of Theorem 2.4. We use the definitions below. Note that the set $F_c^{(n)}$ defined below is not the set introduced by Steel and Matsen. For a technical reason in the proof of Proposition 4.2 stated below, we had to modify their definition. Note however that Propositions 4.2 and 4.3 below are adaptations of ideas in Steel and Matsen's paper.

Notation 4.1. *Define functions Δ_i as follows. For every nonnegative integers $\mathfrak{n} = (n_0, n_1, n_2, n_3)$ such that $|\mathfrak{n}| = n_0 + n_1 + n_2 + n_3 = n$ with $n \geqslant 1$,*

$$\Delta_0(\mathfrak{n}) = \frac{n_0 - q_0 n}{\sqrt{n}},$$

and, for every i in τ,

$$\Delta_i(\mathfrak{n}) = \frac{n_i - 1/3(n - n_0)}{\sqrt{n}}.$$

For every $c > 1$, introduce

$$F_c^{(n)} = \{\mathfrak{n}\,;\, |\mathfrak{n}| = n,\, -2c \leqslant \Delta_2(\mathfrak{n}) \leqslant -c,\, -2c \leqslant \Delta_3(\mathfrak{n}) \leqslant -c,\, -c \leqslant \Delta_0(\mathfrak{n}) \leqslant 0\}.$$

For every i in τ and every positive η, let A_η^i denote the event

$$A_\eta^i = \left\{ \forall j \in \tau,\, j \neq i,\, \mathbb{E}(\Pi_i(N) \,|\, N) \geqslant \eta \mathbb{E}(\Pi_j(N) \,|\, N) \right\}.$$

Since $\Delta_1+\Delta_2+\Delta_3=0$, every $\mathfrak{n}$ in $F_c^{(n)}$ is such that $2c \leqslant \Delta_1(\mathfrak{n}) \leqslant 4c$. We note that $F_c^{(n)}$ is not symmetric about τ and gives a preference to 1. That is why we only deal with A_η^1 in the following proof. To deal with A_η^i, one would change the definition $F_c^{(n)}$ accordingly.

From the reasoning in Section 2, it suffices to prove that for every positive η, there exists a positive δ such that, when n is large enough,

$$\mathbb{P}\left(A_\eta^1\right) \geqslant \delta.$$

Suppose that one generates $n \geqslant 1$ sites on the star tree R_0 with given branch length t and let $\mathfrak{N}=(N_0,N_1,N_2,N_3)$ denote the counts of site patterns defined in Section 1, hence $N_0+N_1+N_2+N_3=n$.

From central limit estimates, the probability of the event $\left\{\mathfrak{N} \in F_c^{(n)}\right\}$ is uniformly bounded from below, say by $\delta>0$, when n is large enough. Hence,

$$\mathbb{P}\left(A_\eta^1\right) \geqslant \delta \mathbb{P}\left(A_\eta^1 \,|\, \mathfrak{N} \in F_c^{(n)}\right)$$

We wish to prove that there exists a positive α independent of c such that for n large enough and for every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j=2$ and $j=3$,

$$\mathbb{E}(\Pi_1(\mathfrak{n})) \geqslant c^2\alpha\, \mathbb{E}(\Pi_j(\mathfrak{n})).$$

This follows from the two results below, adapted from Steel and Matsen's paper.

Proposition 4.2. *Fix t and assume that $\mathfrak{n}$ is in $F_c^{(n)}$. Then, when n is large enough, for $j=2$ and $j=3$,*

$$\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0-1 \in I_t) \geqslant \mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0-1 \notin I_t).$$

Proposition 4.3. *Fix t and assume that $\mathfrak{n}$ is in $F_c^{(n)}$. Then, there exists a positive α, independent of c, such that for every z in I_t, and for $j=2$ and $j=3$,*

$$\mathbb{E}(\Pi_1(\mathfrak{n}) \,|\, 4P_0-1=z) \geqslant c^2\alpha \mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0-1=z).$$

We prove Propositions 4.2 and 4.3 in Section 5.

From these two results, for $j=2$ and $j=3$,

$$\mathbb{E}(\Pi_1(\mathfrak{n})) \geqslant c^2\alpha \mathbb{P}(4P_0-1 \in I_t)\, \mathbb{E}(\Pi_j(\mathfrak{n})).$$

Assume that c is so large that $c^2\alpha\mathbb{P}(4P_0-1\in I_t) \geqslant \eta$. Then, for every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j=2$ and $j=3$,

$$\mathbb{E}(\Pi_1(\mathfrak{n})) \geqslant \eta\, \mathbb{E}(\Pi_j(\mathfrak{n})).$$

This implies that $\mathbb{P}\left(A_\eta^1 \,|\, \mathfrak{N} \in F_c^{(n)}\right)=1$, which yields the theorem.

5. Proofs of Propositions 4.2 and 4.3

5.1. **Proof of Proposition 4.2.** The proof is decomposed into two intermediate results, stated as lemmata below and using estimates on auxiliary random variables introduced below.

Notation 5.1. *For every $n \geqslant 1$ and $t > 0$, let $\Gamma_t(n) = [0,1/n] \times [t,t+1/n]$.*
For every $t > 0$, let $\mu_t = q_0^{q_0} q_1^{q_1} q_2^{q_2} q_3^{q_3} = q_0^{q_0} q_1^{3q_1}$ and U_t denote the random variable

$$U_t = \prod_{i=0}^{3} (P_i/q_i)^{q_i}.$$

For every $\mathfrak{n}$ and for $j = 2$ and $j = 3$, let $W_j(\mathfrak{n})$ denote the random variable

$$W_j(\mathfrak{n}) = P_0^{\Delta_0(\mathfrak{n})} P_1^{(\Delta_j - \Delta_0/3)(\mathfrak{n})} P_2^{(\Delta_1 + \Delta_k - 2\Delta_0/3)(\mathfrak{n})}, \quad \text{with} \quad \{j,k\} = \{2,3\}.$$

One sees that

$$U_t = P_0^{q_0} P_1^{q_1} P_2^{2q_1}/\mu_t, \qquad Q_n(t) = \mathbb{P}(\mathfrak{T} \in \Gamma_t(n)),$$

and, for $j = 2$ and $j = 3$,

$$W_j = (P_0/P_2)^{\Delta_0} (P_1/P_2)^{\Delta_j - \Delta_0/3}.$$

Lemma 5.2. *(1) For every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j = 2$ and $j = 3$, $W_j(\mathfrak{n}) \leqslant 1$.*
(2) For every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j = 2$ and $j = 3$, $W_j(\mathfrak{n}) \geqslant (q_1)^c$ on the event $\{\mathfrak{T} \in \Gamma_t(n)\}$.
(3) There exists a finite constant κ such that $U_t^n \geqslant \mathrm{e}^{-\kappa}$ uniformly on the integer $n \geqslant 1$ and on the event $\{\mathfrak{T} \in \Gamma_t(n)\}$.

Proof of Lemma 5.2. (1) For every $\mathfrak{T}$, $P_0 \geqslant P_1 \geqslant P_2$. On $F_c^{(n)}$, $\Delta_0 \leqslant 0$ and for $j = 2$ and $j = 3$, $\Delta_j - \Delta_0/3 \leqslant 0$ hence

$$(P_0/P_1)^{\Delta_0} \leqslant 1, \quad (P_0/P_2)^{\Delta_j - \Delta_0/3} \leqslant 1.$$

This proves the claim.

(2) One has $P_0 \leqslant 1$ everywhere and $P_1 \geqslant q_1$ and $P_2 \geqslant q_1$ on the event $\{\mathfrak{T} \in \Gamma_t(n)\}$. On $F_c^{(n)}$, $\Delta_0 \leqslant 0$ and for $j = 2$ and $j = 3$, $\Delta_j - \Delta_0/3 \leqslant 0$ hence $W_j \geqslant q_2^{-\Delta_j - 2\Delta_0/3}$. Finally, on $F_c^{(n)}$, $\Delta_j + 2\Delta_0/3 \leqslant -c$. This proves the claim.

(3) For every $\mathfrak{T}$ in $\Gamma_t(n)$, one has $T_{\mathrm{i}} \geqslant 0$ and $T_{\mathrm{e}} \geqslant t$, hence $P_1 \geqslant q_1$ and $P_2 \geqslant q_2 = q_1$. Likewise, $T_{\mathrm{i}} \leqslant 1/n$ and $T_{\mathrm{e}} \leqslant t + 1/n$ hence

$$P_0 \geqslant p_0(1/n, t+1/n) \geqslant q_0 - 5\mathrm{e}^{-4t}(1 - \mathrm{e}^{-4/n})/4.$$

This yields that, for every $n \geqslant 1$ and every $\mathfrak{T}$ in $\Gamma_t(n)$,

$$U_t^n \geqslant (1 - 5\mathrm{e}^{-4t}/(q_0 n))^n \to \exp(-5\mathrm{e}^{-4t}/q_0) > 0,$$

which implies the desired lower bound. □

Lemma 5.3. *For every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j = 2$ and $j = 3$,*

$$\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0 - 1 \in I_t) \geqslant \mu_t^n Q_n(t) \exp(-O(\sqrt{n})),$$

and

$$\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0 - 1 \notin I_t) \leqslant \mu_t^n \exp(-n\ell_t^2/32).$$

Proof of Lemma 5.3. Since $P_0 = p_0(\mathfrak{T})$, for every $\mathfrak{T}$ in $\Gamma_t(n)$, when n is large, $4P_0 - 1$ is in the interval I_t. Consequently,

$$\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_1 - 1 \in I_t) \geqslant Q_n(t) \mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, \mathfrak{T} \in \Gamma_t(n)).$$

On the event $\{\mathfrak{T} \in \Gamma_t(n)\}$,

$$\Pi_j(\mathfrak{n}) = \mu_t^n U_t^n W_j(\mathfrak{n})^{\sqrt{n}} \geqslant \mu_t^n \mathrm{e}^{-\kappa} (q_1)^{c\sqrt{n}},$$

from parts (2) and (3) of Lemma 5.2, which proves the first part of the lemma.

Turning to the second part, let d_{KL} denote the Kullback-Leibler distance between discrete probability measures. When $4P_0 - 1$ is not in I_t,

$$\mathrm{d}_{KL}(q,P) \geqslant (1/2)\|q-P\|_1^2 \geqslant (1/2)(q_0-P_0)^2 \geqslant \ell_t^2/32.$$

Note that

$$\Pi_j(\mathfrak{n}) = \mu_t^n W_j(\mathfrak{n})^{\sqrt{n}} \exp(-n\mathrm{d}_{KL}(q,P)),$$

hence the estimate on $\mathrm{d}_{KL}(q,P)$, and part (1) of Lemma 5.2, imply the second part of the lemma. □

Turning finally to the proof of Proposition 4.2, we note that $Q_n(t) = \mathrm{e}^{o(n)}$ because the distribution of $\mathfrak{T}$ is tempered. Furthermore, Lemma 5.3 shows that, when n is large enough,

$$\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0 - 1 \in I_t) \geqslant \mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, 4P_0 - 1 \notin I_t),$$

and this concludes the proof of Proposition 4.2.

5.2. **Proof of Proposition 4.3.** Our proof of Proposition 4.3 is based on Lemma 5.5 and Proposition 5.6 below.

Notation 5.4. *For every u in $[0,1]$, let $\zeta(u) = (1+2u)(1-u)^2$. Let U and V denote the random variables defined as*

$$U = (P_1 - P_2)/(1-P_0), \qquad V = \zeta(U).$$

Lemma 5.5. *For every $\mathfrak{n}$ in $F_c^{(n)}$ and for $j = 2$ and $j = 3$,*

$$\frac{\mathbb{E}(\Pi_1(\mathfrak{n}) \,|\, P_0)}{\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, P_0)} \geqslant 4c^2 n \frac{\mathbb{E}\,(V^s(1-V) \,|\, P_0)}{\mathbb{E}\,(V^s \,|\, P_0)}, \quad \text{where } s = (n-n_0)/3.$$

Proof of Lemma 5.5. Recall that, for every $c > 1$, $F_c^{(n)}$ is

$$F_c^{(n)} = \{\mathfrak{n} : |\mathfrak{n}| = n, -2c \leqslant \Delta_2(\mathfrak{n}) \leqslant c, -2c \leqslant \Delta_3(\mathfrak{n}) \leqslant c, -c \leqslant \Delta_0(\mathfrak{n}) \leqslant 0\}.$$

Using the Δ variables, one can rewrite Π_1, Π_2 and Π_3 as

$$\Pi_i(\mathfrak{n}) = P_0^{n_0}(P_1P_2^2)^s (P_1/P_2)^{\Delta_i(\mathfrak{n})\sqrt{n}}, \quad i = 1,2,3,\ s = (n-n_0)/3.$$

Assume that $\mathfrak{n}$ is in $F_c^{(n)}$. Then, $\Delta_1(\mathfrak{n}) \geqslant 2c$, $\Delta_j(\mathfrak{n}) \leqslant 0$ for $j = 2$ and $j = 3$, and $P_1 \geqslant P_2$. Hence

$$\Pi_1(\mathfrak{n}) \geqslant P_0^{n_0}(P_1P_2^2)^s (P_1/P_2)^{2c\sqrt{n}}, \qquad \Pi_j(\mathfrak{n}) \leqslant P_0^{n_0}(P_1P_2^2)^s.$$

Furthermore,

$$P_1P_2^2 = (1/27)V(1-P_0)^3, \quad P_1/P_2 = (1+2U)/(1-U),$$

hence for $j = 2$ and $j = 3$,

$$\frac{\mathbb{E}(\Pi_1(\mathfrak{n}) \,|\, P_0)}{\mathbb{E}(\Pi_j(\mathfrak{n}) \,|\, P_0)} \geqslant \frac{\mathbb{E}\left(V^s((1+2U)/(1-U))^{2c\sqrt{n}} \,|\, P_0\right)}{\mathbb{E}\,(V^s \,|\, P_0)}.$$

Direct computations (or Lemma 3.2 in Steel and Matsen [3]) show that, for every u in $[0,1)$ and every $m \geqslant 3$,

$$((1+2u)/(1-u))^m \geqslant m^2(1-\zeta(u)),$$

hence

$$((1+2U)/(1-U))^{2c\sqrt{n}} \geqslant 4c^2 n(1-V).$$

The conclusion of Lemma 5.5 follows. □

Proposition 5.6. *Assume that the distribution of $\mathfrak{T}$ is tempered. Then there exists θ and α, both positive and independent of c, such that for every $s \geqslant \theta$, on the event $\{4P_0 - 1 \in I_t\}$,*

$$4s\,\mathbb{E}(V^s(1-V)\,|\,P_0) \geqslant \alpha\,\mathbb{E}(V^s\,|\,P_0).$$

Proof of Proposition 5.6. We recall that U and V denote random variables defined as

$$U = (P_1 - P_2)/(1-P_0), \quad V = \zeta(U), \quad \zeta(u) = (1+2u)(1-u)^2.$$

To use Proposition 3.2, one must compute a Taylor expansion at $v = 1^-$ or, equivalently, at $u = 0^+$, of the conditional probability

$$\mathbb{P}(V \geqslant v\,|\,P_0) = \mathbb{P}(U \leqslant u\,|\,P_0),$$

where $u = \zeta^{-1}(v)$. Besides, for v close to 1,

$$u = \zeta^{-1}(v) = w/\sqrt{3} + w^2/9 + 5w^3/54\sqrt{3} + O(w^4), \quad \text{with } w = \sqrt{1-v}.$$

Since $U = (P_1 - P_2)/(1-P_0)$,

$$\mathbb{P}(U \leqslant u\,|\,4P_0 - 1 = z) = \mathbb{P}\left(S_{\mathrm{e}}(3 - S_{\mathrm{i}}) \leqslant 2s\,|\,S_{\mathrm{e}}S_{\mathrm{i}} = z\right),$$

where we used the notations

$$S_{\mathrm{e}} = \mathrm{e}^{-4T_{\mathrm{e}}}, \quad S_{\mathrm{i}} = 1 + 2\mathrm{e}^{-4T_{\mathrm{i}}}, \quad 2s = u(3-z).$$

Using Definition 2.3, one has

$$G(z,s) = \mathbb{P}\left(S_{\mathrm{e}}(3 - S_{\mathrm{i}}) \leqslant 2s\,|\,S_{\mathrm{e}}S_{\mathrm{i}} = z\right).$$

Since the distribution of $\mathfrak{T}$ is tempered, there exists some bounded functions F_i defined on I_t, a positive number α, $n+1$ real numbers

$$0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_{n-1} \leqslant 2 < \varepsilon_n,$$

and two positive numbers κ and s_0 such that for every $0 \leqslant s \leqslant s_0$ and every z in I_t,

$$\left| G(z,s) - \sum_{i=0}^{n-1} F_i(z) s^{\alpha+\varepsilon_i} \right| \leqslant \kappa s^{\alpha+\varepsilon_n}.$$

Combining this with the relation $2s = u(3-z)$ and the expansion of $u = \zeta^{-1}(v)$ along the powers of w, one sees that there exists some bounded functions f_i on I_t, a positive number κ' and $0 \leqslant v_0 < 1$ such that for every $v_0 \leqslant v \leqslant 1$ and every $z \in I_t$,

$$\left| \mathbb{P}(V \geqslant v\,|\,4P_0 - 1 = z) - \sum_{i=0}^{n-1} f_i(z)(1-v)^{\alpha/2+\varepsilon_i/2} \right| \leqslant \kappa'(1-v)^{\alpha/2+\varepsilon_n/2}.$$

Since the functions f_i are bounded and positive on I_t, Proposition 3.2 implies that there exists a positive number θ such that for every z in I_t and every $s \geqslant \theta$, the conclusion of Proposition 5.6 holds. □

Assuming this, the proof of Proposition 4.3 is as follows. Let s, θ and α be as in Lemma 5.5 and Proposition 5.6. Since $n - n_0 = (1-q_0)n - \Delta_0\sqrt{n} \geqslant (1-q_0)n$ for every $\mathfrak{n}$ in $F_c^{(n)}$, one knows that $s = (n-n_0)/3 \geqslant \theta$ when n is large enough. Furthermore, $s \leqslant n/3$. Finally, for every $\mathfrak{n}$ in $F_c^{(n)}$ with n large enough, on the event $\{4P_0 - 1 \in I_t\}$ and for $j = 2$ and $j = 3$,

$$\mathbb{E}(\Pi_1(\mathfrak{n})\,|\,P_0) \geqslant 3c^2\alpha\,\mathbb{E}(\Pi_j(\mathfrak{n})\,|\,P_0).$$

This concludes the proof of Proposition 4.3.

ACKNOWLEDGMENTS

I would like to thank Mike Steel and an anonymous referee for some helpful comments.

REFERENCES

[1] M.T. Holder, P. Lewis, and K.E. Holsinger. Polytomies and bayesian phylogenetic inference. *Systematic Biology*, 54(2):241–253, 2005.
[2] B. Kolaczkowski and J.W. Thornton. Is there a star tree paradox? *Mol. Biol. Evol.*, 23:1819–1823, 2006.
[3] M. Steel and F. A. Matsen. The Bayesian 'star paradox' persists for long finite sequences. *Molecular Biology and Evolution*, 24:1075–1079, 2007.
[4] E. Susko. On the distributions of bootstrap support and posterior distributions for a star tree. *Systematic Biology*, 57(4):602–612, 2008.
[5] Z. Yang. *Computational Molecular Evolution*. Oxford Series in Ecology and Evolution, 2006.
[6] Z. Yang. Fair-Balance paradox, star-tree paradox, and bayesian phylogenetics. *Molecular Biology and Evolution*, 24:1639–1655, 2007.
[7] Z. Yang and B. Rannala. Branch-length prior influences Bayesian posterior probability of phylogeny. *Syst. Biol.*, 54(3):455–470, 2005.

APPENDIX A. PROOF OF PROPOSITIONS 2.5, 2.7, AND 2.8

Notation A.1. *Introduce the random variables*

$$(S_{\mathrm{e}}, S_{\mathrm{i}}) = \varsigma(T_{\mathrm{e}}, T_{\mathrm{i}}), \qquad \textit{where } \varsigma(t_{\mathrm{e}}, t_{\mathrm{i}}) = (\mathrm{e}^{-4t_{\mathrm{e}}}, 1 + 2\mathrm{e}^{-4t_{\mathrm{i}}}),$$

that is,

$$S_{\mathrm{e}} = \mathrm{e}^{-4T_{\mathrm{e}}}, \qquad S_{\mathrm{i}} = 1 + 2\mathrm{e}^{-4T_{\mathrm{i}}}.$$

Hence, $G(z,\cdot)$ is defined by

$$G(z,s) = \mathbb{P}\left(3S_{\mathrm{e}} \leqslant 2s + z \,|\, S_{\mathrm{e}}S_{\mathrm{i}} = z\right),$$

A.1. **Proof of Proposition 2.5.** The distribution of $(S_{\mathrm{e}}, S_{\mathrm{i}})$ has a smooth joint probability density, say ϖ, defined on the set $0 < x \leqslant 1 < y \leqslant 3$ by

$$\varpi(x,y) = \frac{\omega \circ \varsigma^{-1}(x,y)}{16x(y-1)}.$$

For tame priors, the probability $Q_n(t)$ introduced in condition 2 of Definition 2.3 is of order $1/n^2$, hence this condition holds.

The definition of $G(z,s)$ as a conditional expectation can be rewritten as

$$G(z,s) = \mathbb{P}\left(3S_{\mathrm{e}} \leqslant 2s + S_{\mathrm{e}}S_{\mathrm{i}} \,|\, S_{\mathrm{e}}S_{\mathrm{i}} = z\right).$$

Hence, for every measurable bounded function H,

$$\mathbb{E}\left(H(S_{\mathrm{e}}S_{\mathrm{i}})\,;\, 3S_{\mathrm{e}} \leqslant 2s + S_{\mathrm{e}}S_{\mathrm{i}}\right) = \mathbb{E}\left(H(S_{\mathrm{e}}S_{\mathrm{i}})G(S_{\mathrm{e}}S_{\mathrm{i}}, s)\right),$$

that is,

$$\iint H(xy)\mathbf{1}\{3x \leqslant 2s + xy\}\varpi(x,y)\mathrm{d}x\mathrm{d}y = \iint H(xy)G(xy,s)\varpi(x,y)\mathrm{d}x\mathrm{d}y.$$

The change of variable $z = xy$ yields

$$\iint H(z)\mathbf{1}\{3x \leqslant 2s + z\}\varpi\left(x, z/x\right)\mathrm{d}z\mathrm{d}x/x = \iint H(z)G(z,s)\varpi\left(x, z/x\right)\mathrm{d}z\mathrm{d}x/x.$$

This must hold for every measurable bounded function H, hence one can choose

$$G(z,s) = H(z,s)/H(z,\infty),$$

with

$$H(z,s) = \int \mathbf{1}\{3x \leqslant 2s+z\}\varpi(x,z/x)\,\mathrm{d}x/x.$$

Since $0 \leqslant S_\mathrm{e} \leqslant 1 \leqslant S_\mathrm{i} \leqslant 3$ almost surely, the integral defining $H(z,s)$ may be further restricted to the range $0 \leqslant x \leqslant 1$ and $z/3 \leqslant x \leqslant z$. Finally, for every $s \geqslant 0$ and $z \in [0,3]$,

$$G(z,s) = H(z,s)/H(z,1),$$

where

$$H(z,s) = \int_{m(0,z)}^{m(s,z)} \varpi(x,z/x)\,\mathrm{d}x/x, \quad \text{with } m(s,z) = \min\{1,z,(2s+z)/3\}.$$

Hence, $m(0,z) = z/3$ and, for small positive values of s, $m(s,z) = m(0,z) + 2s/3$. When $0 \leqslant z \leqslant 1$, $m(s,z) \to m(\infty,z) = z$ when $s \to \infty$ and this limit is reached for $s = z$. When $1 \leqslant z \leqslant 3$, $m(s,z) \to m(\infty,z) = 1$ when $s \to \infty$ and this limit is reached for $s = (3-z)/2$. In both cases, $m(\infty,z) = m(1,z)$ hence $H(z,\infty) = H(z,1)$.

Because ω and ς^{-1} are smooth, the Taylor-Lagrange formula shows that, for every $s \geqslant 0$ and every fixed z,

$$H(z,s) = H(z,0) + H'(z,0)s + H''(z,0)\frac{s^2}{2} + H^{(3)}(z,0)\frac{s^3}{6} + \int_0^s (x-s)^3 H^{(4)}(z,s)\frac{\mathrm{d}x}{24},$$

where all the derivatives are partial derivatives with respect to the second argument s.

Simple computations yield $H(z,0) = 0$ and the values of the three derivatives $H'(z,0)$, $H''(z,0)$ and $H^{(3)}(z,0)$ as combinations of ω and of partial derivatives of ω, evaluated at the point $(\vartheta,0)$, where $3\mathrm{e}^{-4\vartheta} = z$.

Furthermore, the hypothesis on ω ensures that $H^{(4)}(z,\cdot)$ is bounded, in the following sense: there exist positive numbers s_0 and κ_0 such that for every s in $[0,s_0]$ and every z in I_t,

$$H^{(4)}(z,s) \leqslant 24\kappa_0.$$

Hence, $\mathfrak{T} = (T_\mathrm{e}, T_\mathrm{i})$ fulfills the first condition to be tempered, with

$$k = 3, \quad \alpha = 1, \quad \varepsilon_1 = 1, \quad \varepsilon_2 = 2, \quad \varepsilon_3 = 3, \quad \kappa = \kappa_0,$$

and, for every $0 \leqslant i \leqslant 2$,

$$F_i(z) = H^{(i+1)}(z,0)/H(z,1).$$

Finally, since ω is smooth, the functions F_i are bounded on I_t.

A.2. **Proof of Proposition 2.7.** Recall that, using the random variables $S_\mathrm{e} = \mathrm{e}^{-4T_\mathrm{e}}$ and $S_\mathrm{i} = 1 + 2\mathrm{e}^{-4T_\mathrm{i}}$, the function G is characterized by the fact that, for every measurable bounded function H,

$$\mathbb{E}\left(H(S_\mathrm{e}S_\mathrm{i}) : S_\mathrm{e}(3-S_\mathrm{i}) \leqslant 2s\right) = \mathbb{E}\left(H(S_\mathrm{e}S_\mathrm{i})G(S_\mathrm{e}S_\mathrm{i},s)\right).$$

Here, S_e and S_i are independent, the distribution of S_e is uniform on $[0,1]$ and the distribution of S_i is discrete with

$$\mathbb{P}(S_\mathrm{i} = y_n) = r_n/r.$$

Thus,

$$\sum_n r_n \int_0^1 H(xy_n)\,\mathbf{1}\{x(3-y_n) \leqslant 2s\}\,\mathrm{d}x = \sum_n r_n \int_0^1 H(xy_n)\,G(xy_n,s)\,\mathrm{d}x.$$

The changes of variable $z = y_n x$ in each integral yield

$$\sum_n \frac{r_n}{y_n} \int H(z)\,\mathbf{1}\{z \leqslant y_n\}\mathbf{1}\{3z \leqslant (2s+z)y_n\}\,\mathrm{d}z = \sum_n \frac{r_n}{y_n} \int H(z)\,\mathbf{1}\{z \leqslant y_n\}\,G(z,s)\,\mathrm{d}z.$$

This must hold for every measurable bounded function H, hence

$$G(z,s) = H(z,s)/H(z,\infty), \quad H(z,s) = \sum_n (r_n/y_n)\,\mathbf{1}\{z \leqslant y_n\}\mathbf{1}\{3z \leqslant (2s+z)y_n\}.$$

Since $r_n/y_n = n^{-b} - (n+1)^{-b}$ for $n \geqslant 1$, $H(z,s) = n(z,s)^{-b}$ where

$$n(z,s) = \inf\{n \geqslant 1 \,|\, z \leqslant y_n,\, 3z \leqslant (2s+z)y_n\}.$$

Since $y_n \to 3$ when $n \to \infty$, $n(z,s)$ is finite for every $z < 3$ and $s > 0$.

For every $z > 0$, when s is large enough, namely $s \geqslant (3-z)/2$, the condition that $3z \leqslant (2s+z)y_n$ becomes useless and

$$n(z,s) = \inf\{n \geqslant 1 \,|\, z \leqslant y_n\},$$

hence $n(z,s)$ and $H(z,s)$ are independent of s. If $z \geqslant 1$, this implies that $n(z,s)$ and $H(z,s)$ are independent of $s \geqslant 1$. If $z < 1$ and $s \geqslant 1$, the conditions that $z \leqslant y_n$ and $3z \leqslant (2s+z)y_n$ both hold for every $n \geqslant 1$ hence $n(z,s) = 1$ and $H(z,s) = 1$. In both cases, $H(z,\infty) = H(z,1)$.

We are interested in small positive values of s. For every $z < 3$, when s is small enough, namely $s \leqslant (3-z)/2$, the condition $z \leqslant y_n$ becomes useless and

$$n(z,s) = \inf\{n \geqslant 1 \,|\, 3z \leqslant (2s+z)y_n\},$$

When furthermore $s < z$, $n \geqslant n(z,s)$ is equivalent to the condition

$$n^{-a} \leqslant h(s/z), \quad \text{with} \quad h(u) = -\frac{1}{4}\ln\left(1 - \frac{3u}{1+2u}\right),\ 0 \leqslant u < 1.$$

Finally, for every $s < \min\{z, (3-z)/2\}$, $n(z,s)$ is the unique integer such that

$$n(z,s) - 1 < h(s/z)^{-1/a} \leqslant n(z,s).$$

This reads as

$$h(u)^{b/a}[1+h(u)^{1/a}]^{-b} < H(z,1)\,G(z,s) \leqslant h(u)^{b/a}, \quad u = s/z.$$

One sees that the function h is analytic and that $h(u) = (3u/4) + o(u)$ when $u \to 0$, hence,

$$h(u)^{b/a} = (3u/4)^{b/a}(1 + a_1 u + a_2 u^2 + a_3 u^3 + o(u^3)),$$

when $u \to 0$, for given coefficients a_1, a_2 and a_3. Likewise, since $1/a > 3$, $h(u)^{1/a} = o(u^3)$ when $u \to 0$. This implies that

$$\left(1 + h(u)^{1/a}\right)^{-b} = 1 + o(u^3),$$

hence

$$H(z,1)\,G(z,s) = (3u/4)^{b/a}(1 + a_1 u + a_2 u^2 + a_3 u^3 + o(u^3)).$$

This yields the first part of Definition 2.3, with

$$k = 3, \quad \alpha = b/a, \quad (\varepsilon_1, \varepsilon_2, \varepsilon_3) = (1,2,3),$$

and

$$F_0(z) = (3/4z)^{b/a}/H(z,1), \quad F_1(z) = a_1 F_0(z)/z, \quad F_1(z) = a_2 F_0(z)/z^2.$$

The remaining step is to get rid of the dependencies over z of our upper bounds. For instance, the reasoning above provides as an error term a multiple of

$$u^{\alpha+3}/H(z,1) = s^{\alpha+3}/(z^{\alpha+3}H(z,1)),$$

instead of a constant multiple of $s^{\alpha+3}$. But $\inf I_t > 0$, hence the $1/z^{\alpha+3}$ contribution is uniformly bounded.

As regards $H(z,1)$, we first note that $H(z,1)=1$ if $z\leqslant 1$. If $z\geqslant 1$, elementary computations show that $H(z,1)\geqslant c$ if and only if $n(z,1)\leqslant c^{-1/b}$ if and only if $\exp(-c^{a/b})\geqslant (z-1)/2$, which is implied by the fact that $1-c^{a/b}\geqslant (z-1)/2$, which is equivalent to the upper bound $c^{a/b}\leqslant (3-z)/2$. Since $\sup I_t<3$, this can be achieved uniformly over z in I_t and $1/H(z,1)$ is uniformly bounded as well.

Finally, we asked for an expansion valid on $s\leqslant s_0$, for a fixed s_0, and we proved an expansion valid over $s/z\leqslant u_0$, for a fixed u_0. But one can choose $s_0=u_0\inf I_t$. This concludes the proof that the conditions in the first part of Definition 2.3 hold.

We now prove that the second part of Definition 2.3 holds. Since T_{i} and T_{e} are independent, for every positive integer n,

$$Q_n(t)=\mathbb{P}(T_{\mathrm{i}}\leqslant 1/n)\,\mathbb{P}(t\leqslant T_{\mathrm{e}}\leqslant t+1/n).$$

One has

$$n\mathbb{P}(t\leqslant T_{\mathrm{e}}\leqslant t+1/n)\to 4\mathrm{e}^{-4t}\quad\text{when}\quad n\to+\infty,$$

and

$$\frac{1}{r(n^{1/a}+1)^b}\leqslant\mathbb{P}(T_{\mathrm{i}}\leqslant 1/n)\leqslant\frac{3}{rn^{b/a}}.$$

Since $Q_n(t)$ is bounded from below by a multiple of $1/n^{1+b/a}$, the second point of Definition 2.3 holds.

A.3. **Proof of Proposition 2.8.** Recall once again that, using the random variables $S_{\mathrm{e}}=\mathrm{e}^{-4T_{\mathrm{e}}}$ and $S_{\mathrm{i}}=1+2\mathrm{e}^{-4T_{\mathrm{i}}}$, the function G is characterized by the fact that, for every measurable bounded function H,

$$\mathbb{E}\left(H(S_{\mathrm{e}}S_{\mathrm{i}}):S_{\mathrm{e}}(3-S_{\mathrm{i}})\leqslant 2s\right)=\mathbb{E}\left(H(S_{\mathrm{e}}S_{\mathrm{i}})G(S_{\mathrm{e}}S_{\mathrm{i}},s)\right).$$

Case (**i**). Here, S_{e} and S_{i} are independent, the distribution of S_{e} is uniform on $[0,1]$ and S_{i} is a continuous random variable with density

$$\frac{1}{4\theta(s_i-1)}\mathbf{1}\{1+2\mathrm{e}^{-4\theta}\leqslant s_i\leqslant 3\}$$

with respect to the Lebesgue measure. Let ϖ denote the joint probability density defined as

$$\varpi(x,y)=\mathbf{1}\{0\leqslant x\leqslant 1\}\mathbf{1}\{1+2\mathrm{e}^{-4\theta}\leqslant y\leqslant 3\}\frac{1}{4\theta(y-1)}.$$

Thus,

$$\iint H(xy)\mathbf{1}\{3x\leqslant 2s+xy\}\varpi(x,y)\mathrm{d}x\mathrm{d}y=\iint H(xy)G(xy,s)\varpi(x,y)\mathrm{d}x\mathrm{d}y.$$

The change of variable $z=xy$ yields

$$\iint H(z)\mathbf{1}\{3x\leqslant 2s+z\}\varpi(x,z/x)\,\mathrm{d}z\mathrm{d}x/x=\iint H(z)G(z,s)\varpi(x,z/x)\,\mathrm{d}z\mathrm{d}x/x.$$

This must hold for every measurable bounded function H, one can choose

$$G(z,s)=H(z,s)/H(z,\infty),$$

with

$$H(z,s)=\int\mathbf{1}\{3x\leqslant 2s+z\}\mathbf{1}\{0\leqslant x\leqslant 1\}\mathbf{1}\{1+2\mathrm{e}^{-4\theta}\leqslant z/x\leqslant 3\}\mathrm{d}x/(z-x).$$

Finally, for every $s\geqslant 0$ and z in $[0,3]$,

$$G(z,s)=H(z,s)/H(z,1+\mathrm{e}^{-4\theta}),$$

where

$$H(z,s)=\int_{m(0,z)}^{m(s,z)}\frac{\mathrm{d}x}{z-x},\quad \text{with } m(s,z)=\min\{1,z/(1+2\mathrm{e}^{-4\theta}),(2s+z)/3\}.$$

Hence, $m(0,z)=z/3$ and, for small positive values of s, $m(s,z)=m(0,z)+2s/3$. When $0\leqslant z\leqslant 1+2\mathrm{e}^{-4\theta}$, $m(s,z)\to m(\infty,z)=z/(1+2\mathrm{e}^{-4\theta})$ when $s\to\infty$ and this limit is reached for $s=\frac{1+\mathrm{e}^{-4\theta}}{1+2\mathrm{e}^{-4\theta}}z$. When $1+2\mathrm{e}^{-4\theta}\leqslant z\leqslant 3$, $m(s,z)\to m(\infty,z)=1$ when $s\to\infty$ and this limit is reached for $s=(3-z)/2$. In both cases, $m(\infty,z)=m(1+\mathrm{e}^{-4\theta},z)$ hence $H(z,\infty)=H(z,1+\mathrm{e}^{-4\theta})$.

For every fixed $0\leqslant z\leqslant 1+2\mathrm{e}^{-4\theta}$ and every $0\leqslant s\leqslant\dfrac{1+\mathrm{e}^{-4\theta}}{1+2\mathrm{e}^{-4\theta}}z$,

$$H(z,s)=\log\left(\frac{z}{z-s}\right).$$

For every fixed $1+2\mathrm{e}^{-4\theta}\leqslant z\leqslant 3$ and every $0\leqslant s\leqslant\dfrac{3-z}{2}$,

$$H(z,s)=\log\left(\frac{z}{z-s}\right).$$

Hence, there exists a positive s_0 such that for every z in I_t and every s in $[0,s_0]$,

$$H(z,s)=\log\left(\frac{z}{z-s}\right)=\log\left(1-\frac{s}{z}\right).$$

Such a function has a Taylor expansion around $s=0$ with uniformly bounded coefficient over z in I_t. Hence, $\mathfrak{T}=(T_\mathrm{e},T_\mathrm{i})$ fulfills the first condition to be tempered.

We now prove that the second part of Definition 2.3 holds. Since T_i and T_e are independent, for every positive integer n,

$$Q_n(t)=\mathbb{P}(T_\mathrm{i}\leqslant 1/n)\,\mathbb{P}(t\leqslant T_\mathrm{e}\leqslant t+1/n).$$

One has

$$n\mathbb{P}(t\leqslant T_\mathrm{e}\leqslant t+1/n)\to 4\mathrm{e}^{-4t}\quad\text{when}\quad n\to+\infty,$$

and

$$\mathbb{P}(T_\mathrm{i}\leqslant 1/n)=\frac{1}{\theta n},\quad\text{when}\quad n\text{ is large enough.}$$

Since $Q_n(t)$ is bounded from below by a multiple of $1/n^2$, the second point of definition 2.3 holds.

Case (**ii**). Here, S_e and S_i are independent, the distribution of S_e is uniform on $[0,1]$ and S_i is a continuous random variable with density

$$\frac{\theta}{4(s_i-1)}\left[-\frac{1}{4}\log\left(\frac{s_i-1}{2}\right)\right]^{\theta-1}\mathbf{1}\{1+2\mathrm{e}^{-4}\leqslant s_i<3\}$$

with respect to the Lebesgue measure. One can choose

$$G(z,s)=H(z,s)/H(z,\infty),$$

where

$$H(z,s)=\int_{m(0,z)}^{m(s,z)}\left[\frac{-1}{4}\log\left(\frac{z-x}{2x}\right)\right]^{\theta-1}\frac{\mathrm{d}x}{z-x},$$

with

$$m(s,z)=\min\{1,z/(1+2\mathrm{e}^{-4}),(2s+z)/3\}.$$

Hence, $m(0,z) = z/3$ and, for small positive values of s, $m(s,z) = m(0,z) + 2s/3$. When $0 \leqslant z \leqslant 1 + 2\mathrm{e}^{-4}$, $m(s,z) \to m(\infty,z) = z/(1+2\mathrm{e}^{-4})$ when $s \to \infty$ and this limit is reached for $s = \frac{1+\mathrm{e}^{-4}}{1+2\mathrm{e}^{-4}} z$. When $1 + 2\mathrm{e}^{-4} \leqslant z \leqslant 3$, $m(s,z) \to m(\infty,z) = 1$ when $s \to \infty$ and this limit is reached for $s = (3-z)/2$. In both cases, $m(\infty,z) = m(1+\mathrm{e}^{-4},z)$ hence $H(z,\infty) = H(z,1+\mathrm{e}^{-4})$.

Hence, there exists a positive s_0 such that for every z in I_t and every s in $[0,s_0]$,

$$H(z,s) = \int_0^s \frac{1}{(z-x)} \left[\frac{-1}{4} \log\left(1 - \frac{3x}{2x+z}\right) \right]^{\theta-1} \mathrm{d}x.$$

Such a function has a Taylor expansion around $s = 0$ with uniformly bounded coefficient over z in I_t. For instance, when $\theta = 1/2$,

$$H(z,s) = \frac{4}{\sqrt{3z}} \sqrt{s} + \frac{5}{(3z)^{3/2}} s^{3/2} + \frac{9\sqrt{3}}{40 z^{5/2}} s^{5/2} + O(s^{7/2}),$$

where $O(s^{7/2})$ is uniformly bounded over z in I_t. Hence, $\mathfrak{T} = (T_\mathrm{e}, T_\mathrm{i})$ fulfills the first condition to be tempered.

We now prove that the second part of Definition 2.3 holds. Since T_i and T_e are independent, for every positive integer n,

$$Q_n(t) = \mathbb{P}(T_\mathrm{i} \leqslant 1/n)\, \mathbb{P}(t \leqslant T_\mathrm{e} \leqslant t + 1/n).$$

One has

$$n\mathbb{P}(t \leqslant T_\mathrm{e} \leqslant t + 1/n) \to 4\mathrm{e}^{-4t} \quad \text{when} \quad n \to +\infty,$$

and

$$\mathbb{P}(T_\mathrm{i} \leqslant 1/n) = \frac{1}{n^\theta}, \quad \text{when} \quad n \text{ is large enough.}$$

Since $Q_n(t)$ is bounded from below by a multiple of $1/n^{1+\theta}$, the second point of definition 2.3 holds.

Case (**iii**). Here, S_e and S_i are independent, the distribution of S_e is uniform on $[0,1]$ and S_i is a continuous random variable with density

$$-\frac{1}{4(s_i-1)} \log\left[-\frac{1}{4} \log\left(\frac{s_i - 1}{2} \right) \right] \mathbf{1}\{1 + 2\mathrm{e}^{-4} \leqslant s_i < 3\}$$

with respect to the Lebesgue measure.

One can choose

$$G(z,s) = H(z,s)/H(z,\infty),$$

where

$$H(z,s) = \int_{m(0,z)}^{m(s,z)} \log\left[\frac{-1}{4} \log\left(\frac{z-x}{2x} \right) \right] \frac{\mathrm{d}x}{z-x},$$

with

$$m(s,z) = \min\{1, z/(1+2\mathrm{e}^{-4}), (2s+z)/3\}.$$

Hence, there exists a positive s_0 such that for every z in I_t and every s in $[0,s_0]$,

$$H(z,s) = -\int_0^s \frac{1}{(z-x)} \log\left[\frac{-1}{4} \log\left(1 - \frac{3x}{2x+z}\right) \right] \mathrm{d}x.$$

The Taylor expansion around zero of $H(z,s)$ reads as

$$zH(z,s) = (1 - \log(3/(4z)))\, s - s\log(s) + o(s\log(s)),$$

hence $\mathfrak{T} = (T_\mathrm{e}, T_\mathrm{i})$ does not fulfill the first condition to be tempered.

Case (**iv**). Here, S_e and S_i are independent, the distribution of S_e is uniform on $[0,1]$ and S_i is a continuous random variable with density

$$\frac{1}{16(s_i-1)} \log\left(\frac{s_i-1}{2}\right) \log\left[-\frac{1}{4}\log\left(\frac{s_i-1}{2}\right)\right] \mathbf{1}\{1+2\mathrm{e}^{-4} \leqslant s_i < 3\}$$

with respect to the Lebesgue measure.

One can choose

$$G(z,s) = H(z,s)/H(z,\infty),$$

where

$$H(z,s) = \int_{m(0,z)}^{m(s,z)} \log\left(\frac{z-x}{2x}\right) \log\left[\frac{-1}{4}\log\left(\frac{z-x}{2x}\right)\right] \frac{\mathrm{d}x}{z-x},$$

with

$$m(s,z) = \min\{1, z/(1+2\mathrm{e}^{-4}), (2s+z)/3\}.$$

Hence, there exists a positive s_0 such that for every z in I_t and every s in $[0,s_0]$,

$$H(z,s) = -\int_0^s \frac{1}{(z-x)} \log\left(1-\frac{3x}{2x+z}\right) \log\left[\frac{-1}{4}\log\left(1-\frac{3x}{2x+z}\right)\right] \mathrm{d}x.$$

The Taylor expansion around zero of $H(z,s)$ reads as

$$2z^2 H(z,s) = (3/2 - 3\log(3) + 3\log(z) + 6\log(2))\, s^2 - 3s^2 \log(s) + o(s^2\log(s)),$$

hence $\mathfrak{T} = (T_\mathrm{e}, T_\mathrm{i})$ does not fulfill the first condition to be tempered.

Appendix B. Proof of Proposition 3.2

Notation B.1. *Recall that* Γ *denotes the Gamma function defined for every positive number* x *by*

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\mathrm{d}t.$$

For every real number t*, let* $[t]$ *denote the integer part of* t*, that is, the largest integer not greater than* t*, and let* $\{t\}$ *denote the fractional part of* t*, hence* $t = \{t\} + [t]$*,* $[t]$ *is an integer and* $\{t\}$ *belongs to the interval* $[0,1)$*.*

For fixed values of the coefficients α, γ_i and ε_i, introduce, for every $t > 0$,

$$M_t^{\pm} = \int_0^1 t v^{t-1} F_{\pm}(v)\,\mathrm{d}v, \quad \text{where } F_{\pm}(v) = \sum_{i=0}^{n-1} \gamma_i (1-v)^{\alpha+\varepsilon_i} \pm \gamma_n (1-v)^{\alpha+\varepsilon_n}.$$

Hence,

$$M_t = \int_0^1 t v^{t-1} \mathbb{P}(V \geqslant v)\,\mathrm{d}v = M_t^{\pm} + \int_0^1 t v^{t-1}[\mathbb{P}(V \geqslant v) - F_{\pm}(v)]\,\mathrm{d}v,$$

and

$$M_t^{\pm} = tB(t, \alpha+1)\left(\sum_{i=1}^{n-1} \gamma_i \Lambda(\varepsilon_i, t) P(\varepsilon_i, t) \pm \gamma_n \Lambda(\varepsilon_n, t) P(\varepsilon_n, t)\right),$$

where

$$\Lambda(\varepsilon,t) = \frac{\Gamma(\{t\}+\alpha+1)}{\Gamma(\{t\}+\alpha+\varepsilon+1)}, \quad P(\varepsilon,t) = \prod_{\ell=1}^{[t]+1}\left(1-\frac{\varepsilon}{\alpha+\varepsilon+\{t\}+\ell}\right),$$

and B denotes the beta function

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

From the control of the distribution of V,

$$M_t^- - \gamma v_0^t \leqslant M_t \leqslant M_t^+ + v_0^t \quad \text{where } \gamma = \sum_{i=0}^{n} |\gamma_i|.$$

Combining this with the general expression of $M_t^{\pm}$ given above, one gets

$$\frac{M_{t+1}}{M_t} \leqslant \frac{(t+1)B(t+1,\alpha+1)\chi_+(t+1) + v_0^{t+1}}{tB(t,\alpha+1)\chi_-(t) - \gamma v_0^t},$$

where

$$\chi_\pm(t) = \sum_{i=0}^{n-1} \gamma_i \Lambda(\varepsilon_i,t)P(\varepsilon_i,t) \pm \gamma_n \Lambda(\varepsilon_n,t)P(\varepsilon_n,t).$$

Using the fact that

$$\frac{(t+1)B(t+1,\alpha+1)}{tB(t,\alpha+1)} = \frac{t+1}{t+\alpha+1},$$

and that

$$tB(t,\alpha+1)Q_\alpha(t) \geqslant 1, \quad \text{where } Q_\alpha(t) = \frac{(t+\alpha)(t+\alpha-1)\dots(t+\{\alpha\})}{\Gamma(\alpha+1)},$$

one sees that

$$\frac{M_{t+1}}{M_t} \leqslant \frac{t+1}{t+\alpha+1} \frac{\gamma_0 + \chi_+(t+1) + Q_\alpha(t+1)v_0^{t+1}}{\gamma_0 + \chi_-(t) - \gamma Q_\alpha(t) v_0^t}.$$

Furthermore,

$$\frac{\gamma_0 + \chi_+(t+1) + Q_\alpha(t+1)v_0^{t+1}}{\gamma_0 + \chi_-(t) - Q_\alpha(t)v_0^t} = 1 + \frac{\chi_+(t+1) - \chi_-(t) + \kappa(t)v_0^t}{\gamma_0 + \chi_-(t) - \gamma Q_\alpha(t) v_0^t}.$$

where $\kappa(t) = v_0 Q_\alpha(t+1) + \gamma Q_\alpha(t)$ is a polynomial function in t.

From Lemma B.2 below, there exists a positive number C which depend on the exponents α and ε_i, $0 \leqslant i \leqslant n$, only, such that

$$\chi_+(t+1) - \chi_-(t) \leqslant [2\gamma_n + \varepsilon_n \gamma] C t^{-\beta}, \quad \chi_-(t) \geqslant -C\gamma t^{-\varepsilon_1}.$$

where

$$\beta = \min\{\varepsilon_n, 1+\varepsilon_1\}, \qquad 1 < \beta \leqslant 2.$$

Combining these estimates on $\chi_+(t+1)$ and $\chi_-(t)$, one sees that there exists finite continuous functions θ_1 and A of the exponents γ_i, α, and ε_i, such that, for every $t \geqslant \theta_1$,

$$R_t \geqslant \alpha/t - A/t^\beta.$$

Since $\beta > 1$, there exists θ_2 such that $2At \leqslant \alpha t^\beta$ for every $t \geqslant \theta_2$. Choosing finally $\theta = \max(\theta_1, \theta_2)$ yields Proposition 3.2.

Lemma B.2. *Let $\beta = \min\{\varepsilon_n, 1+\varepsilon_1\}$. There exists a positive number C, which depends on the exponents α and ε_i only, such that*

$$\chi_+(t+1) - \chi_-(t) \leqslant [2\gamma_n + \varepsilon_n \gamma] C t^{-\beta}, \quad \chi_-(t) \geqslant -C\gamma t^{-\varepsilon_1}.$$

Proof of Lemma B.2. For every real number $t \geqslant 1$ and every $1 \leqslant i \leqslant n$,

$$\mathrm{e}^{-S(\varepsilon_i,t)-T(\varepsilon_i,t)} \leqslant P(\varepsilon_i,t) \leqslant \mathrm{e}^{-S(\varepsilon_i,t)},$$

where

$$S(\varepsilon,t) = \sum_{\ell=1}^{[t]+1} \frac{\varepsilon}{\alpha+\varepsilon+\{t\}+\ell} \quad \text{and} \quad T(\varepsilon,t) = \sum_{\ell=1}^{[t]+1} \frac{\varepsilon^2}{(\alpha+\varepsilon+\{t\}+\ell)^2}.$$

Thus, there exists two positive real numbers C_i^- and C_i^+ such that for every real number $t \geqslant 1$, $C_i^- \leqslant t^{\varepsilon_i} P(\varepsilon_i,t) \leqslant C_i^+$, and one can choose $C_i^+ = (\alpha+\varepsilon_i+3)^{\varepsilon_i}$.

Let $C = \max\{C_i^+ \,;\, 1 \leqslant i \leqslant n\}$. Using the two relations

$$P(\varepsilon_i,t) - P(\varepsilon_i,t+1) = P(\varepsilon_i,t)\frac{\varepsilon_i}{\alpha+\varepsilon_i+t+2},$$

and

$$P(\varepsilon_n,t) + P(\varepsilon_n,t+1) = P(\varepsilon_n,t)\left(2 - \frac{\varepsilon_n}{\alpha+\varepsilon_n+t+2}\right),$$

one sees that

$$\chi_+(t+1) - \chi_-(t) = 2\gamma_n \Lambda(\varepsilon_n,t)P(\varepsilon_n,t) - \sum_{i=1}^{n} \gamma_i \Lambda(\varepsilon_i,t)P(\varepsilon_i,t)\frac{\varepsilon_i}{\alpha+\varepsilon_i+t+2}.$$

For every $1 \leqslant i \leqslant n$, the function $\Lambda(\varepsilon_i,\cdot)$ is positive and bounded by 1. Hence,

$$\begin{aligned}\chi_+(t+1) - \chi_-(t) &\leqslant 2\gamma_n P(\varepsilon_n,t) + \sum_{i=1}^{n} |\gamma_i| P(\varepsilon_i,t)\frac{\varepsilon_i}{\alpha+\varepsilon_i+t+2} \\ &\leqslant C\left(2\gamma_n t^{-\varepsilon_n} + \gamma\varepsilon_n t^{-(1+\varepsilon_1)}\right),\end{aligned}$$

and the first inequality in the statement of the lemma holds. The same kind of estimates yields

$$\chi_-(t) \geqslant -\sum_{i=0}^{n-1} |\gamma_i| \Lambda(\varepsilon_i,t)P(\varepsilon_i,t) - \gamma_n \Lambda(\varepsilon_n,t)P(\varepsilon_n,t),$$

hence the second inequality holds. This concludes the proof of Lemma B.2. □

UNIVERSITÉ JOSEPH FOURIER GRENOBLE 1, INSTITUT FOURIER UMR 5582 UJF-CNRS, 100 RUE DES MATHS, BP 74, 38402 SAINT MARTIN D'HÈRES, FRANCE