

Cox Markov models for estimating single cell growth

Federico Bassetti*

Università degli Studi di Pavia, Pavia, Italy

e-mail: federico.bassetti@unipv.it

Ilenia Epifani and Lucia Ladelli

Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

e-mail: ilenia.epifani@polimi.it; lucia.ladelli@polimi.it

Abstract: Recent experimental techniques produce thousands of data of single cell growth, consequently stochastic models of growth can be validated on true data and used to understand the main mechanisms that control the cell cycle. A sequence of growing cells is usually modeled by a suitable Markov chain. In this framework, the most interesting goal is to infer the distribution of the *doubling time* (or of the *added size*) of a cell given its initial size and its elongation rate. In the literature, these distributions are described in terms of the corresponding conditional hazard function, referred as *division hazard rate*. In this work we propose a simple but effective way to estimate the division hazard by using *extended Cox modeling*. We investigate the convergence to the stationary distribution of the Markov chain describing the sequence of growing cells and we prove that, under reasonable conditions, the proposed estimators of the division hazard rates are asymptotically consistent. Finally, we apply our model to study some published datasets of E-Coli cells.

MSC 2010 subject classifications: Primary 60J05, 62N02, 62P10; secondary 62F12, 62M05.

Keywords and phrases: Asymptotic consistency, cell size growth in bacteria, Cox partial likelihood, division hazard rate, extended Cox model, positive Harris recurrent Markov chains.

Received September 2016.

Contents

1	Introduction	2932
2	Extended Cox Markov models for cell size growth data	2935
2.1	Model description	2935
2.2	Timer and adder Cox models	2937
2.3	Inference and consistency results	2938
3	Existence of the stationary distribution	2940
3.1	Sufficient condition for Harris positivity in the time increment representation	2941

*Supported by PRIN2015 (2015SNS29B002) and INDAM-GNAMPA Project 2017

3.2	Sufficient condition for Harris positivity in the space increment representation	2943
4	Applicability of the theoretical results	2944
4.1	Stationarity	2945
4.2	Consistency	2946
5	Statistical analysis of some cell size growth data	2947
5.1	Data description	2948
5.2	Goodness of fit by Cox-Snell residuals	2948
5.3	Comparison between simulated processes and empirical observations	2950
5.4	Pure adder versus concerted control adder	2953
5.5	Modeling the dependence between η and X_F	2956
A	Proofs of Section 2	2960
B	Proofs of Section 3	2968
C	Some definitions and properties of homogenous Markov chains	2973
D	Supplementary Tables	2974
	Acknowledgements	2975
	References	2976

1. Introduction

The understanding of the cell cycle is a classic topic in biology. Most bacteria, such as the *Escherichia coli* and *Bacillus subtilis*, divide when they have approximately doubled their volume, but there is always a range of variability of cell sizes. In order to explain this variability, many models for cell size regulation have been introduced in the literature. The key idea is that cells need a feedback mechanism that controls their size distribution and that leads to size homeostasis.

Most of the models proposed for the control of the cell size can be grouped in three major paradigms: *sizer*, when the cell monitors its size and triggers the cell cycle once it reaches a critical size [14]; *timer*, when the cell attempts to grow for a specific amount of time before the division [14]; *adder*, when the cell attempts to add a constant volume to its newborn size [12]. It is possible also to imagine various combinations of the previous models to describe more complex mechanisms. A proposal in this direction is contained in [9], where a mechanism operating concertedly on the size and the age of the cell is discussed. Other recent proposals that lie between the timer and the adder are contained in [1, 7]. We will refer to these kind of models generically as *concerted control* models.

Using the experimental technique called *mother machine* of [17], it is possible nowadays to follow thousands Gram-negative *Escherichia coli* and Gram-positive *Bacillus subtilis* cells for hundreds of generations under various growth conditions and hence it is now feasible to infer the main mechanism of the growth process in cells.

Experimental evidence (see, for example, [12] II. A. and [17]) suggests that cell volume grows exponentially for various bacteria, that is the size $X_i(s)$ of a cell i at age s is given by

$$X_i(s) := X_{I,i}e^{\alpha_i s}, \quad (1.1)$$

where $X_{I,i}$ is the *initial (newborn) size* of the cell and α_i its *growth (elongation) rate*. Assuming (1.1), the *interdivision (doubling) time* τ_i between two divisions is related to the *final size* $X_{F,i}$ of cell i by $X_{F,i} = X_{I,i}e^{\alpha_i \tau_i}$. The picture is completed by a *division law*, which gives the size of the two newborn cells in term of the final size of the mother cell. Focusing only on one of the two newborn cells, the division law describes the ratio η_i between its size and the size of its mother, i.e. $\eta_i = X_{I,i+1}/X_{F,i}$.

Since the variables $X_{I,i}$, α_i , τ_i and $X_{F,i}$ are related by $X_{F,i} = X_{I,i}e^{\alpha_i \tau_i}$, they can be reduced to three. In this paper, we model the history of a sequence of cells, in which the i -th cell is the ancestor of the $(i+1)$ -th cell, by a Markov chain $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ taking values in \mathbb{R}_+^3 , with $\mathbb{R}_+ := (0, +\infty)$. The dynamic is the following:

- (a) given the final size $X_{F,i}$ of the i -th cell, the size $X_{I,i+1}$ of its daughter cell is given by $X_{I,i+1} = \eta_i X_{F,i}$ where η_i is a random number on $(0, 1)$, whose conditional density given $X_{F,i}$ is $f_\eta(\cdot | X_{F,i})$;
- (b) α_{i+1} is assumed to be drawn from the density f_α independently of all the other variables;
- (c) conditionally on α_{i+1} and $X_{I,i+1}$, the interdivision time τ_{i+1} is drawn from the density $f_\tau(\cdot | \alpha_{i+1}, X_{I,i+1})$.

An alternative way of describing the growth process considers the Markov chain $(X_{I,i}, \alpha_i, \Delta_i)_{i \geq 0}$, where

$$\Delta_i := X_{F,i} - X_{I,i} = X_{I,i}(e^{\alpha_i \tau_i} - 1)$$

is the size increment for the i -th cell. The dynamic of the new triplet is the same as the one described above, where $f_\tau(\cdot | \alpha_{i+1}, X_{I,i+1})$ in (c) is replaced by the conditional density $f_\Delta(\cdot | \alpha_i, X_{I,i+1})$ of Δ_{i+1} given $(X_{I,i+1}, \alpha_{i+1})$.

From a biological point of view the most interesting goal is to understand the doubling time (or the increment size) distribution, that is the distribution of τ_i (or equivalently Δ_i), given the initial size $X_{I,i}$ and the elongation rate α_i , see [1, 7, 9, 12]. These distributions are typically described in terms of the corresponding hazard functions, sometimes referred as *division hazard rate functions*. An important point is to understand the shape of the hazard function and its dependency on the other variables of interest, such as cell instantaneous and initial size, added volume, elapsed time from the previous cell division, and growth rate. Different dependencies correspond to different paradigms: sizer, timer, adder and various forms of concerted control.

For example, [9] considers a concerted control model on τ_i and the inference on the division hazard rate is performed by first choosing a specific parametric form for the conditional hazard h_τ of τ_i , and then estimating the parameters by best fitting of the empirical (conditional) hazard.

Recently, [11] and [12] have shown that data are well reproduced by adder models, where the division rate depends on the added size Δ_i . In particular, [12] suggests a complete independence of the increment Δ_i from any other variables. Specifically, the model proposed in [12] to explain the cell growth is a *perfect adder*, i.e. the conditional hazard of Δ_i has the form

$$h_{\Delta}(\delta|X_{I,i}, \alpha_i) = h_{\Delta,0}(\delta).$$

The empirical distribution of the Δ_i 's is used to infer the shape of the division hazard rate h_{Δ} . However the authors admit that there is no strong evidence against a more complex concerted control mechanism.

The partially contradictory results of the previous studies show the importance of studying efficient and rigorous statistical procedures to infer the shape of the hazard and to validate or reject the proposed models.

The aim of this paper is to propose a new effective method for estimating the division hazard rate using *Cox modeling*. In particular we shall consider a *concerted control timer Cox model* with hazard of the type

$$h_{\tau}(t|X_{I,i}, \alpha_i) = h_0(t)e^{Z(t|X_{I,i}, \alpha_i)\beta_0} \quad (1.2)$$

or, alternatively, a *concerted control adder Cox model* with hazard of the type

$$h_{\Delta}(\delta|X_{I,i}, \alpha_i) = h_0(\delta)e^{Z(\delta|X_{I,i}, \alpha_i)\beta_0}, \quad (1.3)$$

where both $Z(t|X_{I,i}, \alpha_i)$ and $Z(\delta|X_{I,i}, \alpha_i)$ are possibly non constant covariates. Clearly, if β_0 is zero, we recover the *pure timer* model in the former case and the *pure adder* mode of [12] in the latter. Provided $\beta_0 \neq 0$, a correction to the pure timer (pure adder, respectively) is set in, giving rise to particular forms of concerted control models.

We propose to estimate the regression parameter β_0 and the baseline hazard h_0 from data using the classic *Cox partial-likelihood* and *Breslow estimator*. It is worth noticing that the Cox model usually considers conditionally independent observations and exogenous covariates, whereas here we extend it to Markovian observations and endogenous covariates depending on α_i and $X_{I,i}$. The main theoretical contribution of this paper is the asymptotic consistency of the proposed estimators in this Markovian framework, see Propositions 2.1 and 2.3. A crucial assumption for consistency turns out to be the positive Harris recurrence of the growth process chain. For this reason we investigate some conditions for positive Harris recurrence, see Propositions 3.4 and 3.7. These results are of their own interest, since the convergence of the growth process to a stationary distribution ensures size homeostasis.

We applied our estimation methodology to investigate the size regulation mechanism of *Escherichia coli*, analyzing some real data taken from [4] and [17] and we explored what kinds of hypotheses the data support. We found that a Cox model applied to the increments Δ_i (i.e. the concerted control adder Cox model) better explains the data, confirming the fact that the adder paradigm can be a valid choice for modeling the growth mechanism of cells. We compared

a pure adder model (which in our context plays the role of the *null* model) with the concerted control adder Cox model. In almost all the sets of data we analyzed, we found that the model with the hazard depending on the initial size of the cell produces a better fit than the pure incremental model. We also observed that the stationary distribution of the estimated growth process is not very sensitive to the particular choice of the distribution of the elongation rates α_i , while a good estimate of the distribution of the ratios η_i is important to well reproduce the empirical distributions of the data. More precisely, allowing the distribution of η_i to depend on the final size of the mother cell improves the fitting in presence of *filaments* (exceptionally long cells).

Some recent studies claim the need for more refined stochastic models that consider the dynamics in cell-cycle sub-periods related to DNA replication and segregation (see [10] and the reference therein). Until now measures of the length of these sub-periods are available only for a small number of cells. However, it is believed that these data will be more and more accessible in the next years. A natural approach to model a growth processes with sub-periods might consider a concatenation of Markovian processes similar to the ones we are proposing for a single-period model.

The rest of the paper is organized as follows. In Section 2 we detail the Markov chain for the growth process of a sequence of cells. In the same section, we propose a Cox model for the hazard rate of the interdivision times or of the size increments, we construct the estimators of the regression parameter and of the baseline cumulative hazard and, finally, we prove their asymptotic consistency. In Section 3 we studied the Positive Harris recurrence of the growth process chain. Section 4 is devoted to the applicability of our theoretical results to some specific Cox models selected for the data analysis. In Section 5 we develop an accurate statistical analysis of the cell data.

The Appendix contains all the proofs and some additional tables. In more details, Appendix A contains the proofs of the results stated in Section 2, while Appendix B contains the proofs of the results stated in Section 3. Some elementary properties and results on Markov chains are recalled in Appendix C. Finally, Appendix D collects some supplementary tables.

2. Extended Cox Markov models for cell size growth data

In this section we face the problem of using an extended Cox model to estimate the conditional hazard either of the interdivision times τ_i or, alternatively, of the size increments Δ_i , and we establish the asymptotic consistency of the proposed estimators.

2.1. Model description

From now on, we shall briefly refer to $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ as the *time increment representation* of the growth process and to $(X_{I,i}, \alpha_i, \Delta_i)_{i \geq 0}$ as the *space increment representation*.

To simplify the presentation, we start noticing that these two chains share the same probabilistic structure. In order to see this, observe that the dynamic (a)-(c) described in the introduction corresponds to the transition rule

$$\begin{aligned} \mathbb{P}\{X_{I,i} \in dx, \alpha_i \in da, \tau_i \in dt | X_{I,i-1}, \alpha_{i-1}, \tau_{i-1}\} \\ = f_{X_I}(x | X_{I,i-1} e^{\alpha_{i-1} \tau_{i-1}}) f_\alpha(a) f_\tau(t | x, a) dx da dt, \end{aligned} \quad (2.1)$$

where f_{X_I} is the conditional density of $X_{I,i}$ given $X_{I,i-1}$, f_α is as in (b) and f_τ as in (c). As for the initial condition, we assume $(X_{I,0}, \alpha_0, \tau_0) = (x_0, a_0, t_0)$.

Analogously, the transition kernel of the chain $(X_{I,i}, \alpha_i, \Delta_i)_{i \geq 0}$ is

$$\begin{aligned} \mathbb{P}\{X_{I,i} \in dx, \alpha_i \in da, \Delta_i \in d\delta | X_{I,i-1}, \alpha_{i-1}, \Delta_{i-1}\} \\ = f_{X_I}(x | X_{I,i-1} + \Delta_{i-1}) f_\alpha(a) f_\Delta(\delta | x, a) dx da d\delta. \end{aligned} \quad (2.2)$$

Hence, if one sets $Y_i = (X_{I,i}, \alpha_i)$ and T_i is either τ_i in the time increment representation or Δ_i in the space increment representation, it follows that in both cases $(Y_i, T_i)_{i \geq 0}$ is a Markov chain with transition kernel

$$\mathbb{P}\{Y_i \in dy, T_i \in dt | Y_{i-1} = \bar{y}, T_{i-1} = \bar{t}\} = f(t|y)g(y|\bar{y}, \bar{t})dydt \quad (2.3)$$

for suitable transition densities $f(\cdot|\cdot)$ and $g(\cdot|\cdot)$.

In order to introduce the Cox model, it is useful to express the conditional survival function of the positive r.v.'s T_i in terms of the conditional hazard $\lambda_T(u|y)$, i.e., for every $t > 0$, we set

$$S(t|y) := \mathbb{P}\{T_i > t | Y_i = y\} = \int_t^{+\infty} f(s|y)ds = \exp\left\{-\int_0^t \lambda_T(s|y)ds\right\}. \quad (2.4)$$

The hazard function $\lambda_T(\cdot|\cdot) : \mathbb{R}_+ \times \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ is such that $\int_0^{+\infty} \lambda_T(s|y)ds = +\infty$ for every $y \in \mathbb{R}_+^2$. Moreover, we shall denote by $\Lambda_T(t|y) = \int_0^t \lambda_T(s|y)ds$ the corresponding cumulative conditional hazard. Thus, in the case of the time increment representation, one has for $y = (x, a)$

$$\begin{aligned} h_\tau(t|x, a) &= \frac{f_\tau(t|x, a)}{\int_t^{+\infty} f_\tau(s|x, a)ds} = \lambda_T(t|y), \\ H_\tau(t|x, a) &:= \int_0^t h_\tau(s|x, a)ds = \Lambda_T(t|y). \end{aligned} \quad (2.5)$$

Analogously, for the space increment representation, one has

$$\begin{aligned} h_\Delta(\delta|x, a) &= \frac{f_\Delta(\delta|x, a)}{\int_\delta^{+\infty} f_\Delta(u|x, a)du} = \lambda_T(\delta|y), \\ H_\Delta(\delta|x, a) &:= \int_0^\delta h_\Delta(u|x, a)du = \Lambda_T(\delta|y). \end{aligned} \quad (2.6)$$

Our proposal is to model the conditional hazard of the variable T_i with a Cox model with covariate process $\{Z_i(t) = Z(t|Y_i) : 0 \leq t \leq T_i\}$. In other words, we assume that the conditional cumulative hazard of the variable T_i is

$$\Lambda_T(t|y) = \int_0^t \lambda_0(u) e^{\beta_0 Z(u|y)} du,$$

where λ_0 is a suitable unknown baseline hazard function, β_0 an unknown real number and $Z(\cdot|\cdot)$ a given covariate function. If we apply this construction to the time increment representation, this leads to the following form for the survival function:

$$S_\tau(t|x, a) = \exp \left\{ - \int_0^t h_{0,\tau}(s) e^{\beta_0 Z(s|x,a)} ds \right\}. \quad (2.7)$$

Analogously, the survival function for the space increment representation turns out to be

$$S_\Delta(\delta|x, a) = \exp \left\{ - \int_0^\delta h_{0,\Delta}(u) e^{\beta_0 Z(u|x,a)} du \right\}. \quad (2.8)$$

It is worth noticing that specifying either the conditional law of Δ_i or of τ_i by a Cox model leads to non equivalent laws for the growth process.

2.2. Timer and adder Cox models

In our framework the model for the cell growth is specified once f_α , f_η and the hazards H_Δ or H_τ are given. We assume that there are no parameters shared by f_α , f_η and H_Δ , H_τ , thus neither f_α nor f_η are involved in the estimation procedure of the hazards. Hence, we classify the models in terms of the conditional hazard, introducing six different choices of the covariate process Z . This leads to six families of models that we will apply to the cell growth data in Section 5. In all the cases we shall consider, the covariate process Z is built starting from a given function $\mathcal{Z} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$.

Some remarks on f_η and on the shape of \mathcal{Z} are presented in Section 4, with a view to verifying the conditions for stationarity of the chain and the consistency of the estimators proposed in the following.

The model $[\tau|X_I]$ This model is a special case of concerted control depending only on the initial size of the cell. Here, $Z(t|x, a) = \mathcal{Z}(x)$ and hence

$$h_\tau(t|x, a) = h_{0,\tau}(t) \exp\{\mathcal{Z}(x)\beta_0\}.$$

In this case the covariate is independent of t and the model reduces to a Cox model with time independent covariates.

The model $[\tau|X_F]$ Also this model is a special case of concerted control. Here the conditional law of the interdivision time depends on the current size of the cell, but it does not depend directly on the growth rate. More precisely, we choose $Z(t|x, a) = \mathcal{Z}(xe^{at})$ and, in terms of hazard, we get

$$h_\tau(t|x, a) = h_{0,\tau}(t) \exp\{\mathcal{Z}(xe^{at})\beta_0\}. \quad (2.9)$$

The model $[\tau|\Delta]$ In the model $[\tau|\Delta]$, the conditional law of the interdivision time depends on the current increment size of the cell, that is $Z(t|x, a) =$

$\mathcal{Z}(x(e^{at} - 1))$ and hence

$$h_\tau(t|x, a) = h_{0,\tau}(t) \exp\{\mathcal{Z}(x(e^{at} - 1))\beta_0\}. \quad (2.10)$$

The model $[\Delta|X_I]$ The model $[\Delta|X_I]$ is a concerted control model for the size increments, where the conditional law of the increment depends on the initial size of the cell, that is $Z(\delta|x, a) = \mathcal{Z}(x)$ and hence

$$h_\Delta(\delta|x, a) = h_{0,\Delta}(\delta) \exp\{\mathcal{Z}(x)\beta_0\}.$$

The model $[\Delta|\tau]$ The model $[\Delta|\tau]$ is a concerted control model, where the conditional law of the size increment depends on the time variable, i.e. $Z(\delta|x, a) = \mathcal{Z}(\log(1 + \delta/x)/a)$. Note that $\log(1 + \delta/x)/a$ is the time elapsed till an increment δ is attained by a cell of initial size x . In this case

$$h_\Delta(\delta|x, a) = h_{0,\Delta}(\delta) \exp\{\mathcal{Z}(\log(1 + \delta/x)/a)\beta_0\}.$$

The model $[\Delta|\text{null}]$ For the sake of comparison we shall also consider the null model $h_\Delta(\delta|x, a) = h_{0,\Delta}(\delta)$, that corresponds to a pure adder model.

2.3. Inference and consistency results

We now deal with the problem of estimating the parameters of the Cox model and we describe the estimation procedure for the process $(Y_i, T_i)_{i \geq 0}$ defined by (2.3) and (2.4).

We need to estimate β_0 and λ_0 as Z is supposed assigned. The classic partial likelihood estimator $\hat{\beta}_n$ of β_0 , based on n observations $(Y_1, T_1), \dots, (Y_n, T_n)$, is defined by

$$\hat{\beta}_n \in \operatorname{argmax}_\beta \prod_{i=1}^n \frac{e^{\beta Z(T_i|Y_i)}}{\sum_{j:T_j \geq T_i} e^{\beta Z(T_i|Y_j)}}$$

and the Breslow estimator of $\Lambda_0(t) := \int_0^t \lambda_0(s) ds$ is

$$\hat{\Lambda}_{0,n}(t) := \sum_{i:T_i \leq t} \frac{1}{\sum_{j:T_j \geq T_i} e^{\hat{\beta}_n Z(T_i|Y_j)}}.$$

Typically the partial likelihood approach is used for independent failures (given the covariates), see e.g. [5] and the references therein. Since here we use the partial likelihood and Breslow estimators for a class of Markovian data, we have to investigate under which assumptions these estimators are consistent. The result is stated in Proposition 2.1 below.

We assume that the reader is familiar with basic notions of the theory of homogeneous Markov chains. However, for the sake of completeness, some definitions and basic results are included in Appendix C. An important hypothesis we need for proving consistency is the Harris positivity of the chain $(Y_i, T_i)_{i \geq 0}$.

In particular, if $(Y_i, T_i)_{i \geq 0}$ is φ -irreducible positive Harris, then it is easy to see that its unique invariant distribution π can be disintegrated as

$$\pi(dydt) = \pi_Y(dy)f(t|y)dt$$

with π_Y the first marginal distribution of π . See Lemma A.1 in Appendix A.

Denote by π_T the second marginal distribution, i.e. $\pi_T(dt) := \int_{\mathbb{R}_+^2} \pi(dydt)$, and define

$$\begin{aligned} c(t|\beta) &:= \int_{\mathbb{R}_+^2} \int_{\{s \geq t\}} e^{\beta Z(t|y)} \pi(dyds), \\ M(\beta) &:= (\beta - \beta_0) \int_{\mathbb{R}_+^2 \times \mathbb{R}_+} Z(t|y) \pi(dydt) - \int_{\mathbb{R}_+} \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \pi_T(dt). \end{aligned} \tag{2.11}$$

Under the assumptions (a)-(d) of the next proposition, one can prove that $c(t|\beta) > 0$ for every t and β (see Lemma A.3 in Appendix A) and hence the function $M(\beta)$ is well defined.

Proposition 2.1. *Assume that:*

- (a) $(Y_i, T_i)_{i \geq 0}$ is φ -irreducible positive Harris with stationary distribution π ;
- (b) $(t, y) \mapsto Z(t|y)$ is a bounded function;
- (c) $t \mapsto Z(t|y)$ is continuous for every y ;
- (d) $\int_0^{\bar{t}} \lambda_0(t) dt < +\infty$ for every $\bar{t} < +\infty$;
- (e) $\frac{d^2}{d\beta^2} M(\beta_0) < 0$.

Then $\hat{\beta}_n$ converges a.s. to β_0 and $\hat{\Lambda}_{0,n}(t)$ converges to $\Lambda_0(t)$ for every $t > 0$ a.s..

In Section 3 we shall discuss reasonable conditions under which assumption (a) holds true. Hypotheses (b)-(d) are, in general, easy to check and the next proposition provides a sufficient condition for the validity of (e).

Proposition 2.2. *If there exists an interval $[a, b] \subset \mathbb{R}_+$ with $\pi_T([a, b]) > 0$, such that $\pi_Y\{y : Z(t|y) = c\} < 1$ for every constant c and almost all $t \in [a, b]$, then $\frac{d^2}{d\beta^2} M(\beta_0) < 0$.*

Finally, the next proposition concerns the asymptotic behavior of the Breslow estimator of the cumulative hazard $\Lambda_T(t|y)$, i.e.

$$\hat{\Lambda}_n(t|y) := \int_{(0,t]} e^{\hat{\beta}_n Z(s|y)} \hat{\Lambda}_{0,n}(ds) = \sum_{i:T_i \leq t} \frac{e^{\hat{\beta}_n Z(T_i|y)}}{\sum_{j:T_j \geq T_i} e^{\hat{\beta}_n Z(T_i|Y_j)}},$$

and the Cox-Snell residuals, i.e.

$$\hat{R}_i := \hat{\Lambda}_n(T_i|Y_i). \tag{2.12}$$

The residuals \hat{R}_i turn out to be very useful for checking the overall fit of the model.

Proposition 2.3. *Under the assumptions of Proposition 2.1, $\hat{\Lambda}_n(t|y)$ converges to $\Lambda_T(t|y)$ for every t and y in \mathbb{R}_+ a.s.. Furthermore, the Cox-Snell residuals \hat{R}_i converge a.s. to $\Lambda_T(T_i|Y_i)$ and are asymptotically independent and exponentially distributed with mean one.*

3. Existence of the stationary distribution

The aim of this section is to provide simple criteria in order to ensure that the growth process is a positive Harris Markov chain, both in the time and in the space increment representation. As we have seen in the previous section, this is a fundamental assumption to prove the consistency of the estimators of β_0 and Λ_0 . Moreover, the convergence of a growth process to a steady state is important also from a biological point of view. Indeed, if one shows that a steady state is reached, then homeostasis is guaranteed.

First of all, let us note that the special form of the transition kernel essentially reduces the study of the asymptotic behavior of the whole growth process to the study of the embedded chain of the initial sizes $(X_{I,i})_{i \geq 0}$. Apropos of this, in the space increment representation, it will be useful to consider the $X_{I,i}$'s as random variables taking values in $[0, +\infty)$. In point of fact, the space increment representation is well-defined even if $X_{I,i} = 0$, provided that the conditional law of Δ_i given $X_{I,i} = 0$ is properly assigned. Clearly, the case $X_{I,i} = 0$ and $\Delta_i > 0$ is not consistent with (1.1), so that the space increment representation and the time increment representation are no longer equivalent. Although the situation in which $X_{I,i} = 0$ has not biological interest, it will be considered below in order to simplify the study of the Harris recurrence. Thus, in this section, we assume the embedded chain $(X_{I,i})_{i \geq 0}$ to be a sequence taking values in $\mathbb{R}_+ = (0, +\infty)$, when we deal with the time increment representation, and in $[0, +\infty)$, when we consider the space increment representation. In the first case, the embedded chain has kernel

$$K(x, dy) = \left(\int_{\mathbb{R}_+^2} \frac{1}{xe^{at}} f_\eta \left(\frac{y}{xe^{at}} \middle| xe^{at} \right) f_\tau(t|x, a) dt f_\alpha(a) da \right) dy,$$

while, in the case of the space increment representation, the kernel of the embedded chain is

$$K^*(x, dy) = \left(\int_{\mathbb{R}_+^2} \frac{1}{x + \delta} f_\eta \left(\frac{y}{x + \delta} \middle| x + \delta \right) f_\Delta(\delta|x, a) d\delta f_\alpha(a) da \right) dy,$$

and $f_\Delta(\cdot|0, a)$ is assumed to be well defined for every $a > 0$.

By direct computations, it is easy to see that, if π_X is an invariant probability measure for the embedded chain $(X_{I,i})_{i \geq 0}$, then π_X is absolutely continuous with respect to the Lebesgue measure and the Markov chain $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ has invariant probability measure

$$\pi(dx dadt) = \pi_X(dx) f_\alpha(a) da f_\tau(t|x, a) dt. \quad (3.1)$$

Analogous considerations hold for the chain $(X_{I,i}, \alpha_i, \Delta_i)_{i \geq 0}$ in the space increment representation. In this case the stationary distribution is

$$\pi^*(dx d\alpha d\delta) = \pi_X(dx) f_\alpha(a) da f_\Delta(\delta|x, a) d\delta.$$

Similar results hold for irreducibility and, under some additional assumptions, also for Harris recurrence, as it is stated in the following proposition.

Proposition 3.1. *If $(X_{I,i})_{i \geq 0}$ is φ -irreducible for some measure φ , then the sequence $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ is an irreducible Markov chain with respect to some suitable (σ -finite non trivial) measure. If moreover $(X_{I,i})_{i \geq 0}$ is positive Harris and irreducible with respect to the Lebesgue measure, then $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ is positive Harris too. The same statement holds in the space increment representation for the chain $(X_{I,i}, \alpha_i, \Delta_i)_{i \geq 0}$.*

Note that the restriction of the kernel K^* to $(0, +\infty)$ is a kernel too and it is absolutely continuous with respect to the restriction of the Lebesgue measure to $(0, +\infty)$. As already noted, if an invariant probability measure exists then it is absolutely continuous with respect to the Lebesgue measure, thus the restriction of the invariant probability measure to $(0, +\infty)$ is still an invariant probability measure for the kernel restricted to $(0, +\infty)$. This shows that also the Harris positivity is preserved by the restricted chain (see Corollary 1 in [13]). From now on, we denote by \mathcal{L} both the Lebesgue measure on $(0, +\infty)$ and on $[0, +\infty)$, according to the considered representation. Moreover, by \mathcal{L} -irreducible chain we mean a chain irreducible with respect to the Lebesgue measure \mathcal{L} .

In the light of the previous considerations, in order to prove that the entire growth process is a positive Harris Markov chain, it is sufficient to prove that the kernel K (or K^*) of the embedded chain $(X_{I,i})_{i \geq 0}$ is positive Harris and \mathcal{L} -irreducible. A classic strategy to prove that a chain $(X_{I,i})_{i \geq 0}$ is positive Harris is to show that: (i) it is φ -irreducible with the support of φ which has non empty interior, (ii) it is weakly Feller, (iii) it satisfies the so called *drift condition*. See Appendix C.

3.1. Sufficient condition for Harris positivity in the time increment representation

The following proposition gives a sufficient condition for \mathcal{L} -irreducibility.

Proposition 3.2. *If there are $u_1(x), u_2(x), t_1(x), t_2(x), a_1(x), a_2(x)$, with $0 \leq u_1(x) < u_2(x) \leq 1, 0 \leq t_1(x) < t_2(x) \leq +\infty, 0 < a_1(x) < a_2(x) \leq +\infty$,*

$$u_1(x) \leq \frac{1 - \epsilon}{e^{a_2(x)t_2(x)}} < \frac{1 + \epsilon}{e^{a_1(x)t_1(x)}} \leq u_2(x),$$

for every $x > 0$ and some $0 < \epsilon < 1$, such that $\mathcal{L} \otimes \mathcal{L}(B_x) > 0$ for

$$B_x := \{(t, a), t \in (t_1(x), t_2(x)), a \in (a_1(x), a_2(x)) \mid f_\tau(t|x, a) f_\alpha(a) > 0\}$$

and $f_\eta(u|x e^{at}) > 0$ for every $u \in (u_1(x), u_2(x)), t \in (t_1(x), t_2(x))$ and $a \in (a_1(x), a_2(x))$, then K is \mathcal{L} -irreducible.

Remark 1. Note that, if $f_\tau(t|x, a) > 0$ for all (t, x, a) , then it is easy to see that the assumptions of Proposition 3.2 are satisfied provided that for some $\tilde{\epsilon} > 0$ and for every $x > 0$ there is $u^* = u^*(x)$ satisfying $f_\eta(u|xe^{at}) > 0$ for every $u \in (u^*(1 - \tilde{\epsilon}), u^*(1 + \tilde{\epsilon}))$.

In the next proposition we consider simple hypotheses yielding that the kernel of the embedded chain is weakly Feller.

Proposition 3.3. Assume that one of the following conditions holds:

- (c1) $x \mapsto f_\eta(yx^{-1}e^{-at})|xe^{at}f_\tau(t|x, a)$ is a continuous function for every (a, t, y) in \mathbb{R}_+^3 ;
- (c2) $x \mapsto f_\eta(u|xe^{at})f_\tau(t|x, a)$ is a continuous function for every (a, t, u) in $\mathbb{R}_+^2 \times (0, 1)$ and for every $\bar{x} > 0$ there is an open set $U_{\bar{x}}$ and a function $c_{\bar{x}}$ such that $f_\eta(u|xe^{at})f_\tau(t|x, a) \leq c_{\bar{x}}(a, t, u)$ for every x in $U_{\bar{x}}$ and $\int_{\mathbb{R}_+^2 \times (0,1)} c_{\bar{x}}(a, t, u) f_\alpha(a) da dt du < +\infty$.

Then, the transition kernel of the embedded chain is weakly Feller.

Finally, an application of the drift condition in Proposition C.1 (Appendix C) gives the next result.

Lemma 3.1. Let the transition kernel of the embedded chain $(X_{I,i})_{i \geq 0}$ be weakly Feller and \mathcal{L} -irreducible. If

$$\begin{aligned} \limsup_{x \rightarrow +\infty} \mathbb{E}[\log^2(X_{I,i+1}) - \log^2(x) | X_{I,i} = x] &< 0 \\ \limsup_{x \rightarrow 0^+} \mathbb{E}[\log^2(X_{I,i+1}) - \log^2(x) | X_{I,i} = x] &< 0 \\ \sup_{x \in (\delta, 1/\delta)} |\mathbb{E}[\log^2(X_{I,i+1}) - \log^2(x) | X_{I,i} = x]| &< +\infty \text{ for every } \delta \in (0, 1), \end{aligned} \tag{3.2}$$

then the chain $(X_{I,i})_{i \geq 0}$ is positive Harris.

Remark 2. Recall that in the time increment representation $X_{I,i+1} = X_{I,i}e^{\alpha_i \tau_i \eta_i}$. It is useful to note that if

$$\begin{aligned} \limsup_{x \rightarrow +\infty} \mathbb{E}[\log(e^{\alpha_i \tau_i \eta_i}) | X_{I,i} = x] &< 0, \quad \liminf_{x \rightarrow 0} \mathbb{E}[\log(e^{\alpha_i \tau_i \eta_i}) | X_{I,i} = x] > 0, \\ \sup_{x > 0} \mathbb{E}[\log^2(e^{\alpha_i \tau_i \eta_i}) | X_{I,i} = x] &< +\infty, \end{aligned} \tag{3.3}$$

then condition (3.2) holds true.

We focus now on time increment models in which the conditional survival function of τ is specified by an extended Cox model, i.e. we assume that $S_\tau(\cdot|\cdot)$ is given by (2.7) with baseline survival function

$$S_{0,\tau}(t) := e^{-\int_0^t h_{0,\tau}(s) ds}.$$

Applying Lemma 3.1 and Remark 2 one obtains the following criterion to check the Harris positivity for the time increment Cox model.

Proposition 3.4. *Let (2.7) hold true with $Z(\cdot|\cdot)$ bounded on $(0, +\infty)^3$. Assume that, for every $t > 0$ and $a > 0$, there exist*

$$\lim_{x \rightarrow 0} Z(t|x, a) = z_0, \quad \lim_{x \rightarrow \infty} Z(t|x, a) = z_\infty. \tag{3.4}$$

Assume also that, for $z_- := \inf_{t,x,a} \beta_0 Z(t|x, a)$,

$$\sup_x \mathbb{E} \left[\log^2(\eta_i) \mid X_{I,i} = x \right] < +\infty, \quad \int_{\mathbb{R}_+} a^2 f_\alpha(a) da \int_{\mathbb{R}_+} t S_{0,\tau}(t)^{e^{z_-}} dt < +\infty, \tag{3.5}$$

$$\limsup_{x \rightarrow +\infty} \mathbb{E} \left[\log(\eta_i) \mid X_{I,i} = x \right] < - \int_{\mathbb{R}_+} a f_\alpha(a) da \int_{\mathbb{R}_+} S_{0,\tau}(t)^{e^{\beta_0 z_\infty}} dt, \tag{3.6}$$

$$\limsup_{x \rightarrow 0^+} \mathbb{E} \left[\log \left(\frac{1}{\eta_i} \right) \mid X_{I,i} = x \right] < \int_{\mathbb{R}_+} a f_\alpha(a) da \int_{\mathbb{R}_+} S_{0,\tau}(t)^{e^{\beta_0 z_0}} dt.$$

If in addition K is weakly Feller and \mathcal{L} -irreducible, then the chain $(X_{I,i})_{i \geq 0}$ is positive Harris recurrent.

3.2. Sufficient condition for Harris positivity in the space increment representation

Let us now consider the growth model in its space increment representation.

Proposition 3.5. *Assume that there is $\epsilon > 0$ such that for every $x \geq 0$ one can determine $u_1(x), u_2(x)$ with $0 \leq u_1(x) < u_2(x) \leq 1$, $d_1(x), d_2(x)$ with $0 \leq d_1(x) < d_2(x) \leq +\infty$ and $A_x \subset \mathbb{R}_+$ such that: if $x = 0$ then $u_1(0) = 0$, $\epsilon < u_2(0)d_1(0)$, $\mathcal{L} \otimes \mathcal{L}(B_0^*) > 0$ for*

$$B_0^* := \{(\delta, a), \delta \in (d_1(0), d_2(0)), a \in A_0 \mid f_\Delta(\delta|0, a)f_\alpha(a) > 0\}$$

and $f_\eta(u|\delta) > 0$ for every $u \in (0, u_2(0))$ and $\delta \in (d_1(0), d_2(0))$; if $x > 0$ then $u_1(x) \leq (1 - \epsilon)/(1 + d_2(x)) < (1 + \epsilon)/(1 + d_1(x)) \leq u_2(x)$, $\mathcal{L} \otimes \mathcal{L}(B_x^*) > 0$ for

$$B_x^* := \{(\delta, a), \delta \in (xd_1(x), xd_2(x)), a \in A_x \mid f_\Delta(\delta|x, a)f_\alpha(a) > 0\}$$

and $f_\eta(u|x+\delta) > 0$ for every $u \in (u_1(x), u_2(x))$ and $\delta \in (xd_1(x), xd_2(x))$. Then, K^* is \mathcal{L} -irreducible.

Remark 3. *Following the same steps of the proof of Proposition 3.5, if, for all (δ, x, a) , $f_\Delta(\delta|x, a) > 0$, it is easy to prove the \mathcal{L} -irreducibility of K^* provided the following request is satisfied: there exists $\epsilon^* > 0$ such that for every $x \geq 0$ one can determine $u^* = u^*(x)$ satisfying $f_\eta(u|x+\delta) > 0$ for every $u \in (u^*(1 - \epsilon^*), u^*(1 + \epsilon^*))$.*

Sufficient conditions in order to ensure that the embedded chain is weakly Feller are contained in the next proposition.

Proposition 3.6. *Assume that one of the following conditions holds:*

- (c3) $x \mapsto f_\eta(y(x + \delta)^{-1}|x + \delta)f_\Delta(\delta|x, a)$ is a continuous function on $[0, +\infty)$ for every (a, δ, y) in \mathbb{R}_+^3 ;
- (c4) $x \mapsto f_\eta(u|x + \delta)f_\Delta(\delta|x, a)$ is a continuous function for every (a, δ, u) in $\mathbb{R}_+^2 \times (0, 1)$ and for every $\bar{x} \geq 0$ there is an open set $U_{\bar{x}}$ and a function $c_{\bar{x}}$ such that $f_\eta(u|x + \delta)f_\Delta(\delta|x, a) \leq c_{\bar{x}}(a, \delta, u)$ for every x in $U_{\bar{x}}$ and $\int_{\mathbb{R}_+^2 \times (0, 1)} c_{\bar{x}}(a, \delta, u)f_\alpha(a)dad\delta du < +\infty$.

Then, the transition kernel of the embedded chain is weakly Feller.

As for the drift condition one can use the next

Lemma 3.2. *Let the transition kernel K^* of the embedded chain $(X_{I,i})_{i \geq 0}$ be weakly Feller and \mathcal{L} -irreducible. Assume that for some $\gamma \in (0, 1]$*

$$\sup_{x \geq 0} \mathbb{E}[\Delta_i^\gamma | X_{I,i} = x] < +\infty \tag{3.7}$$

and that

$$\limsup_{x \rightarrow +\infty} \mathbb{E}[\eta_i | X_{I,i} = x] < 1. \tag{3.8}$$

Then, the chain $(X_{I,i})_i$ is positive Harris recurrent.

Let us consider the baseline survival function

$$S_{0,\Delta}(\delta) = e^{-\int_0^\delta h_{0,\Delta}(u)du},$$

then the next proposition summarizes the main result for the space increment Cox model.

Proposition 3.7. *Let (2.8) hold true with $Z(\cdot|\cdot)$ bounded on $[0, +\infty)^3$ and define $z_- = \inf_{\delta > 0, x \geq 0, a > 0} \beta_0 Z(\delta|x, a)$. Assume that for some γ in $(0, 1]$*

$$\int_0^{+\infty} u^{\gamma-1} S_{0,\Delta}(u) e^{z_-} du < +\infty, \tag{3.9}$$

and that (3.8) holds true. If in addition K^* is weakly Feller and \mathcal{L} -irreducible, then the chain $(X_{I,i})_{i \geq 0}$ is positive Harris recurrent.

4. Applicability of the theoretical results

In this section we discuss the applicability of the theoretical results, concerning the existence of a stationary distribution and the asymptotic consistency of the estimators, to the timer and adder Cox models in Subsection 2.2 under the following conditions:

- (H1) $\mathcal{Z}(x) = x^p \mathbb{I}\{x \leq M\} + M^p \mathbb{I}\{x > M\}$ for a suitable $M > 1$ and $p > 0$;
- (H2) the density f_η does not depend on the final size of the cell.

Furthermore, we assume that condition (d) in Proposition 2.1 holds, i.e. for every s , $\int_0^s h_{\tau,0}(u)du < +\infty$ and $\int_0^s h_{\Delta,0}(u)du < +\infty$ for timer and adder models, respectively.

4.1. Stationarity

Let us investigate conditions for recurrence, Feller continuity and Harris positivity.

The models $[\tau|X_I], [\tau|X_F], [\tau|\Delta]$ First observe that, in the case of the Cox model, given $t > 0$, $f_\tau(t|x, a) > 0$ for every (x, a) if and only if $h_{0,\tau}(t) > 0$, and this simplifies the task of verifying the sufficient conditions for the \mathcal{L} -irreducibility stated in Proposition 3.2. As $\int_0^t h_{0,\tau}(s)ds < +\infty$ for every $t > 0$ and $\int_0^{+\infty} h_{0,\tau}(s)ds = +\infty$, one can always find $0 \leq t_1 < t_2 \leq +\infty$, with t_1 and t_2 as big as one likes, and $0 \leq a_1 < a_2 \leq +\infty$ such that

$$\mathcal{L} \otimes \mathcal{L}(B) > 0 \text{ for } B := \{(t, a), t \in (t_1, t_2), a \in (a_1, a_2) | h_{0,\tau}(t)f_\alpha(a) > 0\}.$$

If in addition there exist u_1 and u_2 such that

$$f_\eta(u) > 0 \text{ for } 0 \leq u_1 < u < u_2 \leq 1 \text{ and } u_1 e^{a_2 t_2} < 1 < u_2 e^{a_1 t_1}$$

then it is easy to show that the assumptions of Proposition 3.2 are satisfied and hence the \mathcal{L} -irreducibility is guaranteed for all the timer Cox models.

Moreover, under **(H1)** and **(H2)**, by Proposition 3.3 the embedded chain turns out to be weakly Feller. To see this, for example, let us consider the $[\tau|X_I]$ model. In this case

$$f_\tau(t|x, a) = h_{0,\tau}(t)e^{\beta_0 \mathcal{Z}(x)} \exp \left\{ -e^{\beta_0 \mathcal{Z}(x)} \int_0^t h_{0,\tau}(s)ds \right\}$$

and the first part of condition **(c2)** is satisfied since \mathcal{Z} is continuous. Moreover, for every x , we may take

$$c_x(a, t, u) = C h_{0,\tau}(t) e^{z_-} \exp \left\{ -e^{z_-} \int_0^t h_{0,\tau}(s)ds \right\}$$

for a suitable positive constant C and $z_- = \min\{\beta_0 M^p, 0\}$. Hence also the second part of **(c2)** is satisfied and the chain is weakly Feller.

Finally the assumptions of Proposition 3.4 simplify as follows. One has $z_0 = 0$ and $z_\infty = M^p$, hence the conditions in (3.6) become

$$-\int_{\mathbb{R}_+} S_{0,\tau}(t)dt < \frac{\mathbb{E}[\log(\eta_i)]}{\mathbb{E}[\alpha_i]} < -\int_{\mathbb{R}_+} S_{0,\tau}^{e^{\beta_0 M^p}}(t)dt,$$

which is possible only if $\beta_0 > 0$. On the other hand, if $\beta_0 > 0$, then (3.5) is satisfied when

$$\mathbb{E}[\log^2(\eta_i)] < +\infty, \quad \mathbb{E}[\alpha_i^2] < +\infty, \quad \int_{\mathbb{R}_+} t S_{0,\tau}(t)dt < +\infty.$$

The models $[\Delta|X_I], [\Delta|\tau]$ Let us now consider the space increment representation. Thanks to **(H2)**, reasoning in a way analogous to the time represen-

tation, one can show that the sufficient conditions of Proposition 3.5 for the \mathcal{L} -irreducibility are satisfied if

$$\exists a > 0 \text{ such that } f_\eta(u) > 0 \quad \forall 0 < u < a. \quad (4.1)$$

Again $z_- = \min\{\beta_0 M^p, 0\}$ and one easily checks that condition **(c3)** is satisfied and hence by Proposition 3.6 the weakly Feller property always holds true. Furthermore, notice that condition (3.8) is trivially satisfied. Thus, in order to apply Proposition 3.7, it remains to verify condition (3.9).

The null models Finally we treat separately the null model in time $[\tau|\text{null}]$ and in space $[\Delta|\text{null}]$, i.e. the limit case in a timer and adder Cox model when the parameter $\beta_0 = 0$.

In the model $[\tau|\text{null}]$ the kernel is

$$K(x, dy) = g(yx^{-1})x^{-1}dy$$

with

$$g(u) = \int_{\mathbb{R}_+^2} e^{-at} f_\eta(ue^{-at}) f_\tau(t) f_\alpha(a) dt da.$$

Hence an invariant probability measure π_X for K should solve the distributional equation

$$X \stackrel{\mathcal{D}}{=} XY$$

where X and Y are independent, X has law π_X and Y has density g . We have already noticed that, if a stationary distribution π_X exists, it must be absolutely continuous with respect to \mathcal{L} . In this case it is easy to see that this equation do not have solutions. Thus, the chain $X_{I,i}$ has no invariant probability measure. For this reason, the model $[\tau|\text{null}]$ has not been included in Subsection 2.2 and will not be considered in the data analysis.

In the case of the model $[\Delta|\text{null}]$, if a stationary distribution of the process $(X_{I,i})_{i \geq 0}$ exists, then it is a perpetuity, see e.g. [3]. More precisely, it is a solution of the distributional equation

$$X_\infty \stackrel{\mathcal{D}}{=} \eta(\Delta + X_\infty)$$

where X_∞ and (η, Δ) are independent and (η, Δ) has the same distribution of each (η_i, Δ_i) . By Theorem 3.1 and Corollary 4.1 in [3], if $\mathbb{E}[\log^+(\eta\Delta)] < +\infty$, the previous equation has a unique solution (in law) and hence K^* has a stationary distribution. Under condition (4.1), the kernel K^* is \mathcal{L} -irreducible and, since for every x , $K^*(x, \cdot)$ is absolutely continuous with respect to \mathcal{L} , it follows from Proposition C.2 that K^* is positive Harris recurrent.

4.2. Consistency

We now deal with the applicability of Proposition 2.1 in order to obtain the consistency of the estimators. Assumptions (b) and (c) are an immediate consequence of **(H1)**. If in addition, we assume **(H2)** and (d), the properties of

Harris positivity and \mathcal{L} -irreducibility of the chain $(X_{I,i})_i$, which imply (a), are discussed in Subsection 4.1. Below, we study the validity of (e) by means of Proposition 2.2.

The models $[\tau|X_I]$, $[\tau|X_F]$, $[\tau|\Delta]$ Under **(H1)** it follows immediately that $\pi_Y\{y : Z(t|y) = c\} = \pi_Y\{\emptyset\} = 0$ if $c < 0$ or $c > M^p$. Suppose now $0 \leq c \leq M^p$, and, for $t > 0$, set $\gamma_t(a) = 1$ under the model $[\tau|X_I]$, $\gamma_t(a) = e^{at}$ under $[\tau|X_F]$ and $\gamma_t(a) = e^{at} - 1$ under $[\tau|\Delta]$. With this position,

$$\{y : Z(t|y) = c\} = \begin{cases} \{y = (x, a) \in \mathbb{R}_+^2 : x\gamma_t(a) = c^{1/p}\} & \text{if } 0 \leq c < M^p \\ \{y = (x, a) \in \mathbb{R}_+^2 : x\gamma_t(a) \geq c^{1/p}\} & \text{if } c = M^p. \end{cases}$$

Now choose $\bar{\gamma}$ such that $\int_{\{a:\gamma_t(a) < \bar{\gamma}\}} f_\alpha(a)da > 0$ and note that

$$\left\{y = (x, a) \in \mathbb{R}_+^2 : x < \frac{c^{1/p}}{\bar{\gamma}}, \gamma_t(a) < \bar{\gamma}\right\} \subseteq \{y : Z(t|y) \neq c\}.$$

Hence, in order to show that $\pi_Y\{y : Z(t|y) = c\} < 1$, it suffices to prove that

$$\pi_Y\left\{y = (x, a) \in \mathbb{R}_+^2 : x < \frac{c^{1/p}}{\bar{\gamma}}, \gamma_t(a) < \bar{\gamma}\right\} > 0.$$

The last claim follows easily since by (3.1) one can write $\pi_Y(dxda) = \pi_X(dx)f_\alpha(a)da$ and then

$$\begin{aligned} \pi_Y\left\{y = (x, a) \in \mathbb{R}_+^2 : x < \frac{c^{1/p}}{\bar{\gamma}}, \gamma_t(a) < \bar{\gamma}\right\} \\ = \pi_X\left\{0 < x < \frac{c^{1/p}}{\bar{\gamma}}\right\} \int_{\{a:\gamma_t(a) < \bar{\gamma}\}} f_\alpha(a)da. \end{aligned}$$

To conclude, note that if the assumptions discussed in the previous subsection hold, the chain $(X_{I,i})_i$ is \mathcal{L} -irreducible and hence \mathcal{L} is absolutely continuous with respect to π_X (see Appendix C), which yields that $\pi_X\{0 < x < c^{1/p}/\bar{\gamma}\} > 0$.

The models $[\Delta|X_I]$, $[\Delta|\tau]$, $[\Delta|\text{null}]$ As for the models $[\Delta|X_I]$ and $[\Delta|\tau]$, the study of the validity of (e) follows the same lines of the previous paragraph.

In the case of the model $[\Delta|\text{null}]$, conditions (a) and (d) are sufficient to guarantee the consistency of the estimator

$$\hat{\Lambda}_n(t|y) = \hat{\Lambda}_{0,n}(t) = \sum_{i:T_i \leq t} \frac{1}{\sum_j \mathbb{I}_{\{T_j \geq T_i\}}}.$$

5. Statistical analysis of some cell size growth data

In this section we apply the proposed extended Cox models to analyse some cell growth data taken from [4], [9] and [17].

5.1. Data description

The data in [9] have been elaborated from the publicly available dataset [17]. The original crude sample consists of measurements of size (length, width and projected area) at each time frame of 1 min from segmented images of cells growing in steady exponential conditions in a microfluidic device made of micron-sized channels. All the data correspond to the strain E. coli MG1655. To reduce the noise in the size measurements, the Authors in [9] extracted the cell size at a given time point from the fit of exponential cell growth, instead of using the value obtained for cell size at that frame. To this end, a linear fit was performed on the log of cell size against the time, by the least-squares method. This procedure uniquely associates to each cell its initial size X_0 (expressed in μm), final size X_F (in μm), doubling time τ (in minutes) and growth rate α . For more details the reader is referred to the supplementary material of [9]. We perform our analysis on three datasets, each of 10000 cells, from now on denoted by MG-A, MG-B, MG-C. All the data refer to the so called old-pole cells, i.e. the cells at the closed end of each channel in the mother machine.

The second group of data are taken from [4] and have been obtained by using agarose pad microscopy under varying nutrient quality. We consider four datasets corresponding to two different strains (referred to as P5-ori and MRR), and three media (denoted by LB, CAA and Glc). For details on the strains and the media see [4]. In the following the four datasets are denoted by CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR and contain 5905, 979, 2073, 2296 observations, respectively. As a first approximation, we treated each dataset as a unique sequence of mother-daughter cells.

5.2. Goodness of fit by Cox-Snell residuals

Let us recall that in all the models we assume that the growth rates α_i are independent and identically distributed and that there are no parameters shared by the densities of α and η and the hazards of Δ and τ . Hence, as noted in the previous section, neither f_α nor f_η are involved in the estimation procedure of the hazards H_Δ and H_τ , which we are mainly interested in. For this reason, at a first stage, we may not specify f_α and f_η . As for the function \mathcal{Z} , in line with hypothesis **(H1)** of Section 4, in order to satisfy assumption (b) of Proposition 2.1 we choose $\mathcal{Z}(x) = \min(x, M)$, with $M = 350$. As a matter of fact, the truncation at 350 is fictitious, because both the final size and the doubling time of every cell are less than 80. Moreover, other functions have been considered, in particular polynomial and logarithmic \mathcal{Z} , and the results are very similar to those obtained under $\mathcal{Z}(x) = \min(x, M)$.

In order to assess the fit of each alternate model listed in Subsection 2.2, we first estimated the hazard functions and then we computed the Cox-Snell residuals \hat{R}_i as in (2.12). Then, we checked whether the residuals \hat{R}_i 's behave as a sample from a unit exponential distribution, by means of the Kolmogorov-Smirnoff test. Indeed, thanks to Proposition 2.3, if a Cox model fits the data,

then the one-sample Kolmogorov-Smirnov test should fail to reject the null hypothesis that the residuals are exponentially distributed with mean one and the corresponding p -value should be high. Tables 1 and 2 report p -values and Kolmogorov-Smirnoff statistics for the datasets MG-A, MG-B, MG-C and CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR, respectively.

Concerning the datasets MG-A, MG-B and MG-C, the summaries in Table 1 suggest that both models $[\Delta|X_I]$ and $[\Delta|\tau]$ fit the data well, whereas all the Cox models for the doubling time τ , i.e. $[\tau|\Delta]$, $[\tau|X_I]$ and $[\tau|X_F]$, fit the data badly (p -value < 0.05). This supports the idea that the Cox model behaves quite well when applied to the increments Δ_i . We also computed the Nelson-Aalen estimator of the cumulative hazard rate for each of the candidate Cox models and compared them to the unit exponential cumulative hazard rate. Results give similar evidence (plots not shown).

TABLE 1
Kolmogorov-Smirnoff test to check if the estimated Cox-Snell residuals of each model listed in the first row are exponential with rate one for the three datasets: MG-A, MG-B and MG-C.

dataset		$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
MG-A	D -statistic	0.0099	0.0100	0.0152	0.0173	0.0160
	p -value	0.2823	0.2676	0.0201	0.0050	0.0118
MG-B	D -statistic	0.0075	0.0093	0.0168	0.0201	0.0177
	p -value	0.6312	0.3576	0.0073	0.0006	0.0037
MG-C	D -statistic	0.0096	0.0071	0.0267	0.0276	0.0238
	p -value	0.3158	0.6893	0.0000	0.0000	0.0000

TABLE 2
Kolmogorov-Smirnoff test to check if the estimated Cox-Snell residuals of each model listed in the first row are exponential with rate one for CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR datasets.

dataset	sample size		$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
CAA-P5ori	5905	D -statistic	0.0089	0.0333	0.0334	0.0740	0.0578
		p -value	0.7413	0.0000	0.0000	0.0000	0.0000
Glc-MRR	2703	D -statistic	0.0141	0.0395	0.0253	0.0456	0.0396
		p -value	0.6538	0.0004	0.0635	0.0000	0.0004
Glc-P5ori	979	D -statistic	0.0042	0.0096	0.0186	0.0503	0.0547
		p -value	1.0000	1.0000	0.8886	0.0141	0.0057
LB-MRR	2296	D -statistic	0.0090	0.0080	0.0190	0.0197	0.0180
		p -value	0.9922	0.9987	0.3775	0.3366	0.4463

As regards Cox-Snell residuals for the data CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR reported in Table 2, it is worth noticing that their sample sizes are smaller than in the previous datasets and that they are more noisy. The Cox-Snell residuals of this second group of data do not lead to an interpretation as clean as for the first group. Only the model $[\Delta|X_I]$ adequately fits all datasets. Conversely, the Cox models $[\tau|X_I]$ and $[\tau|X_F]$ badly fit all datasets with the exception of LB-MRR. Lastly, the models $[\tau|\Delta]$ and $[\Delta|\tau]$ are essentially equivalent in term of Cox-Snell residuals, although none of them exhibits a good fit uniformly on this group of data.

To summarize, if one looks for a Cox model that works for all datasets, then the model $[\Delta|X_I]$ is the only one that we do never reject by the Cox-Snell residual test.

5.3. Comparison between simulated processes and empirical observations

In the perspective of deselecting poor models, as a further measure of fit, we compare the actual marginal empirical distributions of the interdivision times τ_i , the increments Δ_i , the initial sizes $X_{I,i}$ and the final sizes $X_{F,i}$ with the corresponding empirical distributions of an artificial dataset generated from the estimated models $[\Delta|X_I]$, $[\Delta|\tau]$, $[\tau|X_I]$, $[\tau|X_F]$ and $[\tau|\Delta]$. As a competing model, we also consider the pure adder model $[\Delta|\text{null}]$. These comparisons rely on the stationarity of both the actual and simulated data.

It is important to note that in order to get a sample from an estimated model one needs to take into account the distributions of α and η . According to **(H2)** in Section 4, the simplest choice assumes η independent of the final size X_F . In this way one can estimate the laws of α and η by the corresponding empirical distributions. As mentioned in the introduction, the law of α is quite peaked, hence another possibility is to estimate it by a (truncated) gaussian distribution with the sample mean and variance as parameters. Apropos of this, let us recall that [9] assumes a gaussian model for α . Indeed, an exploratory analysis of the growth rates α_i shows that, if one excludes the extreme values (around 0.1% of the largest and 0.1% of the smallest ones), then a gaussian model fits the data. Once f_α and f_η has been estimated, we generated the artificial dataset and computed the two-sample Kolmogorov-Smirnoff distance between the actual and the simulated data. As the actual and simulated samples are Markovian, we do not perform a homogeneity Kolmogorov-Smirnoff test.

The results in Table 3 confirm that the incremental Cox models fit better the datasets MG-A, MG-B and MG-C than the interdivision time Cox models, for α and η generated from the empirical distributions. Note that the models $[\tau|X_I]$, $[\tau|X_F]$ and $[\tau|\Delta]$ give small values of the Kolmogorov-Smirnoff statistics only for the variable τ . In all the three datasets, the concerted control model $[\Delta|X_I]$ is slightly better than both the pure adder model $[\Delta|\text{null}]$ and the model $[\Delta|\tau]$. All these findings are confirmed by the plots of the histograms of the actual data compared with the ones of the artificial data. Figures 1(a), 1(b) and 3(a) show the plots for MG-A, MG-B and MG-C datasets, respectively.

It is worth mentioning that the comparison between artificial and actual data are not very sensitive to the choice of the law of α . For example, Table 8 in Appendix D reports similar values for Kolmogorov-Smirnoff distances under a gaussian fit of α . Conversely, the results are more affected by the choice of the law of η . For instance, the fit is considerably worse if one generates η from a (truncated) gaussian distribution with sample mean and variance as parameters.

As regards the datasets CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR from [4], the comparison between simulated and true observations is less easy to interpret.

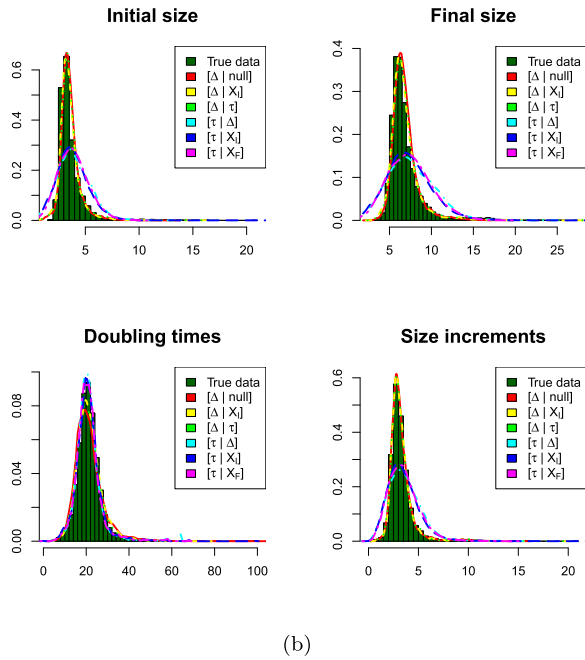
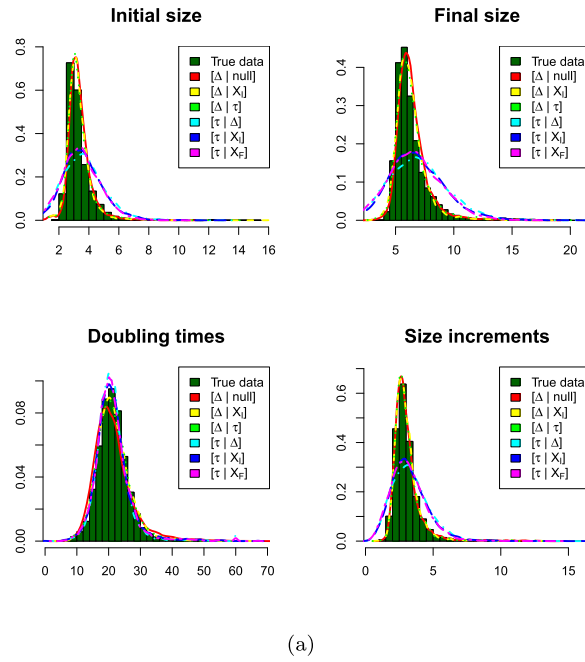


FIG 1. Histograms of the actual and simulated MG-A datasets in Figure 1(a) and of the actual and simulated MG-B datasets in Figure 1(b). (Artificial α and η sampled from the corresponding empirical distributions).

TABLE 3

Values of the Kolmogorov-Smirnoff statistic for comparing empirical versus simulated data in MG-A, MG-B and MG-C cases. (Artificial α and η sampled from the corresponding empirical distributions).

dataset	variable	$[\Delta \text{null}]$	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
MG-A	X_I	0.108	0.073	0.073	0.252	0.220	0.200
	X_F	0.095	0.053	0.058	0.262	0.228	0.212
	τ	0.118	0.073	0.076	0.048	0.041	0.046
	Δ	0.010	0.011	0.011	0.205	0.176	0.163
MG-B	X_I	0.103	0.067	0.111	0.245	0.220	0.239
	X_F	0.086	0.043	0.081	0.254	0.223	0.250
	τ	0.113	0.076	0.100	0.052	0.039	0.046
	Δ	0.010	0.010	0.006	0.217	0.175	0.190
MG-C	X_I	0.125	0.101	0.170	0.363	0.323	0.362
	X_F	0.107	0.078	0.130	0.342	0.296	0.335
	τ	0.104	0.080	0.120	0.082	0.086	0.088
	Δ	0.011	0.011	0.017	0.244	0.185	0.224

First of all, it turns out that the fitting is strongly affected by the variance of the simulated η . In particular, for all models listed in Subsection 2.2, we get better results under a constant η equal to the empirical mean, than for η simulated from the empirical distribution. This may be due to the unexpected high value of η 's variance, probably caused by the presence of outliers and errors in the reconstruction of the genealogy of the cells.

In Table 4 we report the two-sample Kolmogorov-Smirnoff statistics, when the artificial η_i 's are equal to the empirical mean and the artificial α_i 's are sampled from the empirical distribution. The model $[\Delta|X_I]$ confirms a good fit for all four datasets.

TABLE 4

Values of the Kolmogorov-Smirnoff statistic for comparing empirical versus simulated data in CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR cases. (Artificial α sampled from the corresponding empirical distributions, η constant and equal to the corresponding empirical mean).

dataset	variable	$[\Delta \text{null}]$	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
CAA-P5ori	X_I	0.114	0.101	0.091	0.211	0.257	0.146
	X_F	0.086	0.073	0.060	0.195	0.238	0.124
	τ	0.134	0.133	0.072	0.081	0.140	0.091
	Δ	0.014	0.013	0.024	0.134	0.142	0.090
Glc-MRR	X_I	0.085	0.060	0.066	0.231	0.228	0.214
	X_F	0.096	0.070	0.064	0.202	0.198	0.183
	τ	0.115	0.101	0.085	0.068	0.093	0.082
	Δ	0.009	0.010	0.027	0.151	0.132	0.132
Glc-P5ori	X_I	0.090	0.090	0.085	0.153	0.451	0.174
	X_F	0.051	0.051	0.045	0.139	0.451	0.164
	τ	0.067	0.061	0.069	0.065	0.143	0.086
	Δ	0.007	0.007	0.015	0.099	0.287	0.122
LB-MRR	X_I	0.087	0.093	0.228	0.567	0.487	0.518
	X_F	0.082	0.074	0.167	0.512	0.428	0.458
	τ	0.099	0.085	0.154	0.105	0.120	0.115
	Δ	0.016	0.007	0.018	0.366	0.276	0.311

For the CAA-P5ori and Glc-MRR data, the two-sample Kolmogorov-Smirnoff statistics under the models $[\Delta|\tau]$ and $[\Delta|X_I]$ are very similar, even if the model $[\Delta|\tau]$ is rejected by the Cox-Snell test (see Table 2). Thus, also for these second group of datasets the Cox model seems to behave quite well when applied to the increments Δ_i , while it does not work for the interdivision time τ_i .

Table 9 in Appendix D gives the two-sample Kolmogorov-Smirnoff statistics when both α and η are generated from the empirical distributions. Similar conclusions can be drawn, even if the Kolmogorov-Smirnoff values are in general higher, indicating a worse fit.

As for the null incremental model $[\Delta|\text{null}]$, we observe that the two-sample Kolmogorov-Smirnoff statistics are very similar to those obtained with the non-null models $[\Delta|X_I]$ and $[\Delta|\tau]$, regardless of the choice of η . In particular, $[\Delta|\text{null}]$ and $[\Delta|X_I]$ essentially yield identical results for Glc-P5ori dataset. As one can see in Table 10 in Appendix D, this fact is consistent with the $\hat{\beta}_n$ estimate that is very close to zero, one order of magnitude less than the $\hat{\beta}_n$ estimates for the remaining datasets.

5.4. Pure adder versus concerted control adder

As mentioned in the introduction, from a biological point of view it is interesting to understand if a pure adder mechanism explains the growth process or if a more complex mechanism, as a generalized Cox model, is needed. On the basis of the results of Subsection 5.3, it is not clear if the null model can be rejected. Indeed, the simulated samples from a pure adder model reasonably reproduce the marginal distributions of all the variables of interest. Hence in this section we compare the null model $[\Delta|\text{null}]$ with the competing $[\Delta|X_I]$ one by means of suitable statistics which take into account the dependence between $X_{I,i}$ and Δ_i . To this aim, we consider the conditional mean of Δ_i given $X_{I,i}$. We focus on the $[\Delta|X_I]$ model because, on the one hand, it gives both good Cox-Snell residuals and good fit for the simulated data in all the datasets, on the other hand it is the simplest model which is not the null one.

Let $M_{[\Delta|\text{null}]}(x)$ denote the conditional mean $\mathbb{E}[\Delta_i|X_{I,i} = x]$ under the null model and $M_{[\Delta|X_I]}(x)$ the conditional mean under the $[\Delta|X_I]$ model. Therefore

$$M_{[\Delta|\text{null}]}(x) = \mathbb{E}[\Delta_i] \quad \text{and} \quad M_{[\Delta|X_{I,i}]}(x) = \int_0^{+\infty} S_{0,\Delta}(\delta) e^{\beta_0 z(x)} d\delta,$$

for almost all x . Then we estimate the L^1 error

$$L_{model} := \mathbb{E}_\pi [|M_{model}(X_{I,i}) - \mathbb{E}[\Delta_i|X_{I,i}]|],$$

where *model* stands for $[\Delta|\text{null}]$ or $[\Delta|X_I]$ and \mathbb{E}_π denotes the expectation with respect to the stationary distribution of the process.

We estimated $M_{[\Delta|\text{null}]}$ by the empirical mean $\hat{M}_{[\Delta|\text{null}]} = n^{-1} \sum_{i=1}^n \Delta_i$ and $M_{[\Delta|X_I]}(x)$ by

$$\hat{M}_{[\Delta|X_I]}(x) = \int_0^{+\infty} \hat{S}_{0,\Delta,n}(\delta) e^{\hat{\beta}_n z(x)} d\delta,$$

where $\hat{S}_{0,\Delta,n}(\delta) = \exp\{-\hat{H}_{0,\Delta,n}(\delta)\}$, $\hat{H}_{0,\Delta,n}$ is the Breslow estimator of $H_{0,\Delta}$ and $\hat{\beta}_n$ is the estimator of β_0 (see [19]). Lastly, we estimated the “true” conditional mean $\mathbb{E}[\Delta_i|X_{I,i} = x]$ non-parametrically by the Nadaraya-Watson estimator

$$\hat{M}_{ker}(x) = \sum_{i=1}^n \Delta_i G\left(\frac{x - X_{I,i}}{b}\right) / \sum_{i=1}^n G\left(\frac{x - X_{I,i}}{b}\right),$$

with G a bounded symmetric kernel density with bandwidth b (see [18]). Here we assumed the Epanechnikov kernel $G(\nu) = (3/4)(1 - \nu^2)^+$ and we chose the bandwidth b given in Equation (5.5) page 127 in [16]. In this way we obtained the following plug-in estimator

$$\hat{L}_{model} = \frac{1}{n} \sum_{i=1}^n |\hat{M}_{ker}(X_{I,i}) - \hat{M}_{model}(X_{I,i})|.$$

The estimates $\hat{M}_{[\Delta|X_I]}(x)$ and $\hat{M}_{ker}(x)$ may be very poor in the tail of the empirical distribution of the $X_{I,i}$'s and, as a consequence, the empirical estimate of L_{model} turns out to be sensitive to few extreme $X_{I,i}$ values. To alleviate this problem, we considered the following truncated version

$$\hat{L}_{model}(p) = \frac{1}{n} \sum_{i=1}^{\lceil np \rceil} |\hat{M}_{ker}(X_{(i)}) - \hat{M}_{model}(X_{(i)})|,$$

where $X_{(i)}$ is the i -th order statistic of the $X_{I,i}$'s, $p \in (0, 1)$ and $\lceil np \rceil$ is the integer part of np . Note that $\hat{L}_{model}(p)$ is a plug-in estimator of $L_{model}(p) = \mathbb{E}_\pi[|M_{model}(X_{I,i}) - \mathbb{E}[\Delta_i|X_i]| \mathbb{I}\{X_{I,i} \leq q_p\}]$, where q_p is the p -th quantile of the stationary distribution of the $X_{I,i}$'s.

Table 5 collects the results for the MG-A, MG-B, MG-C datasets with $p = 0.985, 0.990, 0.995$ and under different choices of \mathcal{Z} . Analogous results for CAA-P5ori, Glc-MRR, LB-MRR, Glc-P5ori datasets are reported in Table 6. In particular, the third column in both tables shows $\hat{L}_{[\Delta|\text{null}]}$, while the fourth one reports $\hat{L}_{[\Delta|X_I]}$ in the case $\mathcal{Z}(x) = \min(x, 350) =: \mathcal{Z}_1(x)$. Comparing these two columns, we conclude that the model $[\Delta|X_I]$ works better than the null model, except for Glc-P5ori. It seems that the values of \hat{L}_{model} for Glc-P5ori are not affected by the choice of any model. This matches the fact that the $\hat{\beta}_n$ estimate under $[\Delta|X_I]$ is very close to zero, suggesting that $[\Delta|\text{null}]$ is the true model for Glc-P5ori.

We performed the same analysis with $\mathcal{Z}(x) = \min(x^\gamma, 350^\gamma)$ for different choices of γ , without observing significant differences. For the MG datasets all these choices give an estimated conditional mean $\hat{M}_{[\Delta|X_I]}$ that is close to $\hat{M}_{ker}(x)$ for x smaller than a suitable threshold (between 4 and 5), while they are not completely satisfactory for larger values. See, e.g., black and blue triangle-lines in Figure 2(a). For the group of data CAA-P5ori, Glc-MRR, LB-MRR and Glc-P5ori, the behavior of $\hat{M}_{ker}(x)$ is more irregular. The estimated conditional mean of Glc-MRR exhibits a pattern similar to the MG datasets. The

TABLE 5

Empirical estimates of the truncated L^1 errors for MG-A, MG-B and MG-C datasets: $\hat{L}_{[\Delta|\text{null}]}$ values are shown in the third column, $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = \min(x, 350) =: \mathcal{Z}_1(x)$ in the fourth column, $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = \hat{q}_{0.015}\mathbb{I}_{[0, \hat{q}_{0.015}]}(x) + x\mathbb{I}_{(\hat{q}_{0.015}, \hat{q}_{0.985})}(x) + \hat{q}_{0.985}\mathbb{I}_{[\hat{q}_{0.985}, +\infty)}(x) =: \mathcal{Z}_2(x)$ in the fifth column and $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = x\mathbb{I}_{(0, \hat{q}_{0.98})}(x) + \hat{q}_{0.98}\mathbb{I}_{[\hat{q}_{0.98}, +\infty)}(x) =: \mathcal{Z}_3(x)$ in the sixth column.

dataset	p	$\hat{L}_{[\Delta \text{null}]}$	$\hat{L}_{[\Delta X_I]}$		
			\mathcal{Z}_1	\mathcal{Z}_2	\mathcal{Z}_3
MG-A	0.985	0.2669	0.0760	0.0505	0.0493
	0.990	0.2742	0.0776	0.0538	0.0514
	0.995	0.2820	0.0968	0.0572	0.0545
MG-B	0.985	0.3160	0.1127	0.0849	0.0735
	0.990	0.3225	0.1304	0.1005	0.0788
	0.995	0.3289	0.1601	0.1159	0.0842
MG-C	0.985	0.2561	0.0937	0.0926	0.0918
	0.990	0.2688	0.0947	0.0933	0.0927
	0.995	0.2763	0.1007	0.0977	0.0966

TABLE 6

Empirical estimates of the truncated L^1 errors for CAA-P5ori, Glc-MRR, LB-MRR and Glc-P5ori datasets: $\hat{L}_{[\Delta|\text{null}]}$ values are shown in the third column, $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = \min(x, 350) =: \mathcal{Z}_1(x)$ in the fourth column, $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = \hat{q}_{0.015}\mathbb{I}_{[0, \hat{q}_{0.015}]}(x) + x\mathbb{I}_{(\hat{q}_{0.015}, \hat{q}_{0.985})}(x) + \hat{q}_{0.985}\mathbb{I}_{[\hat{q}_{0.985}, +\infty)}(x) =: \mathcal{Z}_2(x)$ in the fifth column and $\hat{L}_{[\Delta|X_I]}$'s with $\mathcal{Z}(x) = x\mathbb{I}_{(0, \hat{q}_{0.98})}(x) + \hat{q}_{0.98}\mathbb{I}_{[\hat{q}_{0.98}, +\infty)}(x) =: \mathcal{Z}_3(x)$ in the sixth column.

dataset	p	$\hat{L}_{[\Delta \text{null}]}$	$\hat{L}_{[\Delta X_I]}$		
			\mathcal{Z}_1	\mathcal{Z}_2	\mathcal{Z}_3
CAA-P5ori	0.985	0.1473	0.1029	0.0969	0.0971
	0.990	0.1475	0.1041	0.0983	0.0981
	0.995	0.1477	0.1057	0.0999	0.0993
Glc-MRR	0.985	0.3074	0.1660	0.1360	0.1340
	0.990	0.3101	0.1757	0.1452	0.1403
	0.995	0.3104	0.1951	0.1560	0.1484
LB-MRR	0.985	0.1940	0.1756	0.1892	0.1881
	0.990	0.2002	0.1819	0.1954	0.1942
	0.995	0.2069	0.1875	0.2010	0.1999
Glc-P5ori	0.985	0.0636	0.0637	0.0636	0.0649
	0.990	0.0713	0.0714	0.0713	0.0727
	0.995	0.0748	0.0748	0.0747	0.0762

estimated conditional mean of Glc-P5ori is essentially constant, while the patterns of CAA-P5ori and LB-MRR are quite oscillatory. In order to save space, in Figure 2(b) we display the parametric and non-parametric estimates of the conditional mean of Glc-MRR dataset only. Summarizing, at least for the MG group of data, \hat{M}_{ker} is monotonic in the range corresponding to the 90-95% smallest $X_{I,i}$'s (the percentage depending on the dataset), while for the remaining values of x the behavior is more irregular. Although the estimates $\hat{M}_{ker}(x)$ are less reliable for large x , they suggest that the “true” conditional mean $\mathbb{E}[\Delta|X_I](x)$ is bounded. On the contrary, one can easily see that if $\beta_0 < 0$ and \mathcal{Z} is strictly increasing, then the conditional mean is strictly increasing with x ; the opposite

is true if $\beta_0 > 0$. Here all the $\hat{\beta}_n$ estimates are negative and coherently the corresponding estimates $\hat{M}_{[\Delta|X_I]}(x)$ are strictly increasing. Unfortunately, this fact is supported neither by the data nor by biological considerations. In the light of these findings, we explored the impact of a “real” truncation in \mathcal{Z} , and we considered $\mathcal{Z}(x) = m\mathbb{I}_{[0,m]}(x) + x\mathbb{I}_{(m,M)}(x) + M\mathbb{I}_{[M,+\infty)}(x)$. As a rule of thumb, we choosed m and M based on some suitable quantiles of the $X_{I,i}$ ’s. For instance, the results shown on the fifth column in Tables 5 and Table 6 are based on a truncation at m equal to the 0.015 empirical quantile of the $X_{I,i}$ ’s and at M equal to the 0.985 quantile, i.e. $\mathcal{Z}(x) = \hat{q}_{0.015}\mathbb{I}_{[0,\hat{q}_{0.015}]}(x) + x\mathbb{I}_{(\hat{q}_{0.015},\hat{q}_{0.985})}(x) + \hat{q}_{0.985}\mathbb{I}_{[\hat{q}_{0.985},+\infty)}(x) =: \mathcal{Z}_2(x)$. In addition, the sixth column of each table contains the results for $\mathcal{Z} = x\mathbb{I}_{(0,\hat{q}_{0.98})}(x) + \hat{q}_{0.98}\mathbb{I}_{[\hat{q}_{0.98},+\infty)}(x) =: \mathcal{Z}_3(x)$. Actually, we observe that the corresponding estimated errors $\hat{L}_{[\Delta|X_I]}$ decrease; even a visual inspection of the plots of the estimated conditional means in Figure 2 confirms these findings. We computed the Cox-Snell residuals of the model $[\Delta|X_I]$ with such truncated \mathcal{Z} ’s functions and we found that the corresponding p -values are smaller than those from the non truncated case, although they are bigger than 0.05. See Table 11 in Appendix D.

We can conclude that the non-null model $[\Delta|X_I]$ behaves better than the null model $[\Delta|\text{null}]$ in all considered datasets with the exception of Glc-P5ori.

5.5. Modeling the dependence between η and X_F

In some growth conditions and for some strains, exceptionally long cells (filaments) can be observed. These cells produce a second minor mode in the distribution of cell size, as can be clearly seen for the MG-C dataset in Figure 3. But this second peak is definitely not reproduced by the models of the previous subsections, as in those applications we assumed the independence of η from the final size of the cell. Already in [17] it has been observed that filaments are probably related to aging, an effect that cannot be captured by any homogenous Markov model. Nevertheless, there is another interesting feature that one can try and take into account in the models we are considering: there is a striking difference in the division patterns of the cells depending on their final size. In the note accompanying the datasets of [17], it is pointed out that *filaments longer than 15 μm never divide at the midcell but at the quarter positions*. This is evident for all the MG datasets, see for example Figure 4 for MG-C. This fact suggests that, in modeling the law of η , one should account for some dependence on X_F .

Roughly speaking, the η ’s behave in different manners as the final size X_F increases and, in particular, for final size bigger than 13 μm we observe a bimodal-shaped distribution of η with modes around 0.25 and 0.75. In order to bring this feature into our model as simply as possible, we assume that

$$f_\eta(u|x) = \sum_{j=1}^K \mathbb{I}\{\tilde{x}_{j-1} \leq x < \tilde{x}_j\} f_{\eta,j}(u)$$

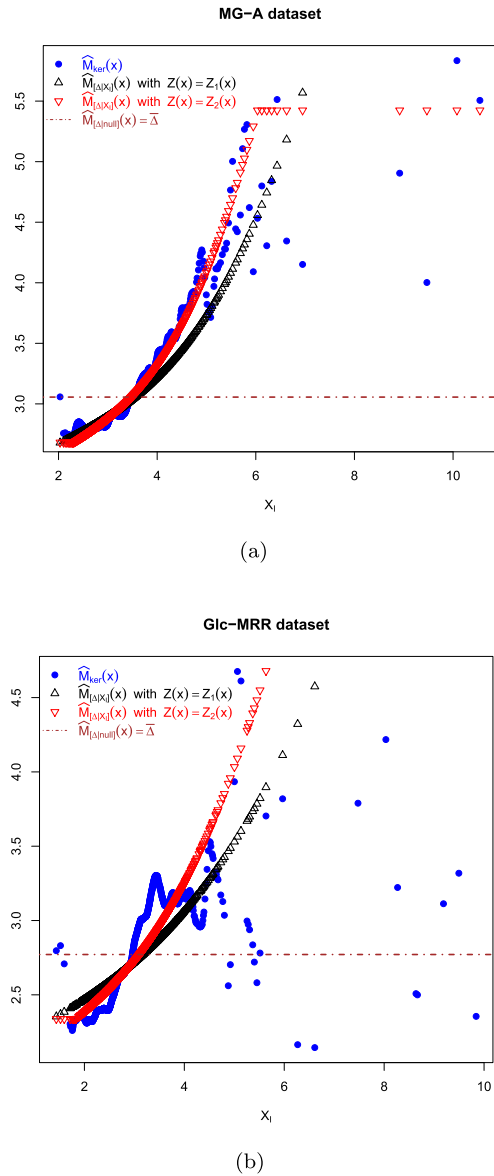
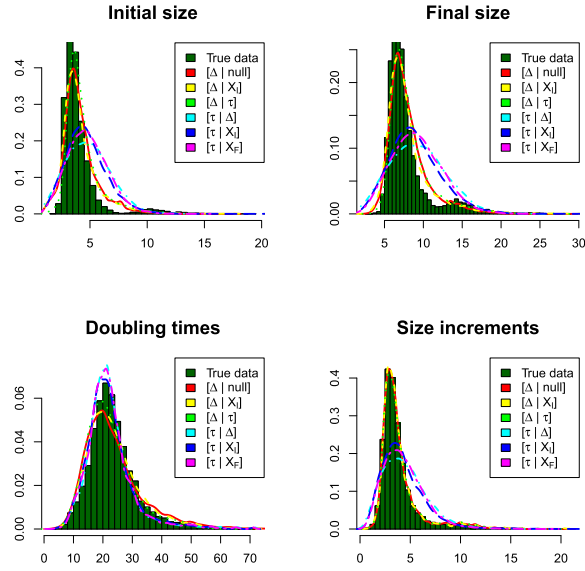
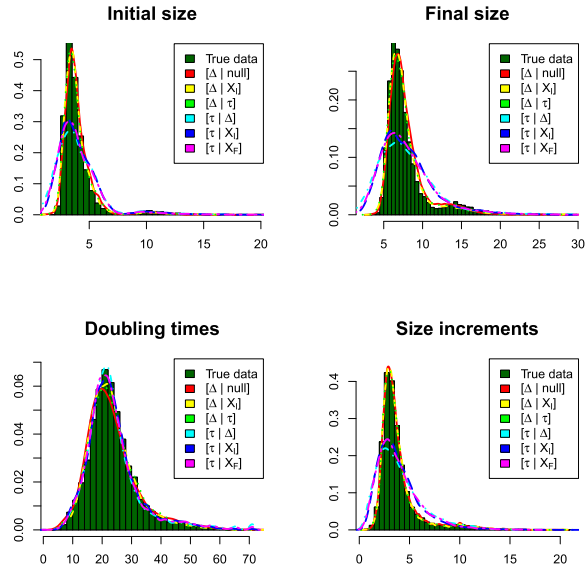


FIG 2. Estimates of the conditional means for MG-A dataset in Figure 2(a) and for Glc-MRR dataset in Figure 2(b). (Black) up-pointing triangle-line corresponds to the estimate $\hat{M}_{[\Delta|X_I]}(x)$ with $Z(x) = \min(x, 350) =: Z_1(x)$, (red) down-pointing triangle-line to $\hat{M}_{[\Delta|X_I]}(x)$ with $Z(x) = \hat{q}_{0.015} \mathbb{1}_{[0, \hat{q}_{0.015}]}(x) + x \mathbb{1}_{(\hat{q}_{0.015}, \hat{q}_{0.985})}(x) + \hat{q}_{0.985} \mathbb{1}_{[\hat{q}_{0.985}, +\infty)}(x) =: Z_2(x)$, (blue) solid dotted line to the non-parametric Nadaraya-Watson estimator $\hat{M}_{ker}(x)$ and, (brown) horizontal line to the empirical mean of the increments Δ .



(a)



(b)

FIG 3. Histograms of the actual and simulated MG-C datasets in Figure 3(a) with both artificial α and η sampled from the corresponding empirical distributions, and of the actual and simulated MG-C datasets in Figure 3(b) with artificial α sampled from the corresponding empirical distributions and artificial η simulated from the X_F -dependent density f_η .

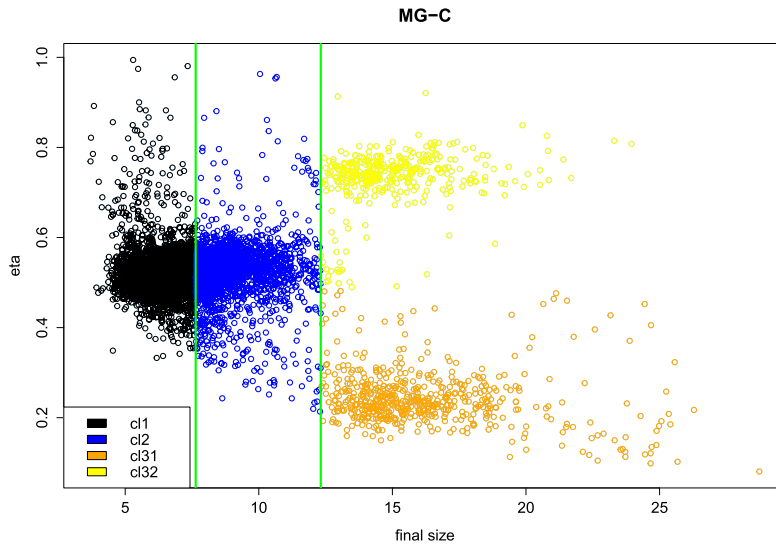


FIG 4. Clustering procedure applied to MG-C dataset.

where $\tilde{x}_0 = 0 < \tilde{x}_1 < \dots < \tilde{x}_K = +\infty$ and $f_{\eta,j}(u)$ is a density on the unit interval. In what follows $K = 3$, $f_{\eta,j}(u) = \mathcal{N}_{[0,1],\mu_j,\sigma_j}(u)$ for $j = 1, 2$ and $f_{\eta,3}(u) = p_3 \mathcal{N}_{[0,1],\mu_{31},\sigma_{31}}(u) + (1 - p_3) \mathcal{N}_{[0,1],\mu_{32},\sigma_{32}}(u)$, where $\mathcal{N}_{[0,1],a,b}$ denotes a gaussian density with mean a and standard deviation b truncated on $[0, 1]$ and $0 < p_3 < 1$.

The estimation of the parameters $(\tilde{x}_1, \tilde{x}_2, \mu_1, \mu_2, \mu_{31}, \mu_{32}, \sigma_1, \sigma_2, \sigma_{31}, \sigma_{32}, p_3)$ is obtained in the following way. Given a set of data $(X_{F,i}, \eta_i)_i$, we start by identifying four clusters as follows.

Step 1 Using the *k-means* algorithm, we first split the data into three clusters: cl_1 corresponding to small values of $X_{F,i}$, cl_2 corresponding to the medium ones and cl_3 corresponding to the large ones.

Step 2 As the third cluster cl_3 highlights two point-clouds, apparently based on the values of η_i , we perform another *k-means* algorithm with $k = 2$ to the only η_i 's of the data in cl_3 and denote by cl_{31}, cl_{32} the clusters induced in cl_3 by the last 2-means classification of the η_i 's.

The result of the above clustering procedure applied to MG-C dataset is shown in Figure 4. Hence we take \tilde{x}_1 and \tilde{x}_2 equal to the maximum values attained by the $X_{F,i}$'s in cl_1 and cl_2 , respectively, and we set p_3 equal to the proportion of the data in cl_3 belonging to cl_{31} . Finally, we estimate μ_j by the sample mean and σ_j by the sample variance of the η_i 's such that the couple $(\eta_i, X_{F,i})$ belongs to cl_j , whereas μ_{3j} and σ_{3j} are estimated by the sample mean and variance of η_i 's such that $(\eta_i, X_{F,i})$ is in cl_{3j} , for $j = 1, 2$.

We performed the validation procedure described in Subsection 5.3 using this estimate of the conditional density f_η , for the MG datasets. The corresponding

two-sample Kolmogorov-Smirnoff distances between actual and simulated data are collected in Table 7. Even with this choice of f_η , Cox models applied to the

TABLE 7
Values of the Kolmogorov-Smirnoff statistic for comparing empirical versus simulated data in MG-A, MG-B and MG-C cases. (Artificial α sampled from the corresponding empirical distributions, artificial η simulated from the X_F -dependent density f_η)

dataset	variable	$[\Delta \text{null}]$	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
MG-A	X_I	0.091	0.041	0.046	0.196	0.208	0.188
	X_F	0.096	0.040	0.043	0.203	0.228	0.206
	τ	0.100	0.065	0.061	0.087	0.087	0.091
	Δ	0.013	0.013	0.009	0.160	0.179	0.159
MG-B	X_I	0.093	0.034	0.018	0.185	0.174	0.170
	X_F	0.091	0.031	0.021	0.188	0.177	0.185
	τ	0.107	0.067	0.065	0.078	0.066	0.082
	Δ	0.013	0.010	0.017	0.160	0.142	0.157
MG-C	X_I	0.066	0.038	0.029	0.194	0.166	0.174
	X_F	0.073	0.033	0.038	0.196	0.171	0.174
	τ	0.076	0.053	0.050	0.059	0.050	0.055
	Δ	0.005	0.012	0.008	0.154	0.133	0.136

increments Δ_i fit better than Cox models for times τ_i , and $[\Delta|X_I]$ and $[\Delta|\tau]$ are preferable to $[\Delta|\text{null}]$. Moreover, a comparison between Table 3 and Table 7 points out that the introduction of a dependence between η and X_F improves the fitting of all the marginal distributions for all the MG datasets. In particular, the strongest improvement is obtained for the MG-C dataset, which contains the highest percentage of filaments (8.5% of cells longer than $13\mu\text{m}$ versus 1.7% in MG-A and 2.8% in MG-B).

The histograms of actual and simulated MG-C data in Figure 3(a) (for η sampled from the empirical distribution) and in Figure 3(b) (for η sampled from the distribution described above) reveal that the Δ -models with η dependent on X_F reproduce more accurately the tails of the true data. In particular, the resulting distributions for X_I and τ turn out to be more peaked and the marginal distribution of X_F better mimics the second minor mode of the true data. We did not apply this statistical analysis to the datasets CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR, since they are not large enough to contain a non-negligible number of filaments.

Appendix A: Proofs of Section 2

The proof of Proposition 2.1 follows the same line of the analogous results in [2] and it will be carried out in a series of lemmata. The proofs of Proposition 2.2 and Proposition 2.3 are given at the end of this section.

Proof of Proposition 2.1. Let

$$M_n(\beta) := A_n(\beta) - B_n(\beta)$$

where

$$A_n(\beta) = \frac{1}{n} \sum_{i=1}^n Z(T_i|Y_i)(\beta - \beta_0)$$

$$B_n(\beta) = \frac{1}{n} \sum_{i=1}^n \log \left(\frac{\frac{1}{n} \sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} e^{\beta Z(T_i|Y_j)}}{\frac{1}{n} \sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} e^{\beta_0 Z(T_i|Y_j)}} \right).$$

In order to show that $\hat{\beta}_n$ converges to β_0 \mathbb{P} -a.s., one observes that that M_n is a concave function and $\hat{\beta}_n \in \operatorname{argmax} M_n(\beta)$ (see Lemma A.2). Then one shows that $M_n(\beta)$ converges \mathbb{P} -a.s. to $M(\beta)$ for all β (see Lemma A.6) and that $\beta \mapsto M(\beta)$ has a unique maximum in β_0 (see Lemma A.4). At this stage the first part of the thesis follows by Thm. 7.7 in [6]. The second part of the thesis is proved in Lemma A.7. \square

Lemma A.1. *If $(Y_i, T_i)_{i \geq 0}$ is φ -irreducible positive Harris, then there exists a probability measure π_Y on $\mathcal{B}(\mathbb{R}_+^2)$ such that*

$$\pi(dydt) = \pi_Y(dy) f(t|y) dt \tag{A.1}$$

is the unique invariant distribution for $(Y_i, T_i)_{i \geq 0}$ and

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n \psi(Y_i, T_i) = \int_{\mathbb{R}_+^2 \times \mathbb{R}_+} \psi(y, t) \pi(dydt)$$

almost surely for every ψ such that $\int_{\mathbb{R}_+^2 \times \mathbb{R}_+} |\psi(y, t)| \pi(dydt) < +\infty$ and for every initial condition.

Proof. If $(Y_i, T_i)_{i \geq 0}$ is φ -irreducible and positive Harris, this means that there is a (unique) invariant probability measure π , that is

$$\pi(dydt) = \int f(t|y) dt g(dy|\bar{y}, \bar{t}) \pi(d\bar{y}d\bar{t}) = f(t|y) dt \int g(dy|\bar{y}, \bar{t}) \pi(d\bar{y}d\bar{t}).$$

Hence (A.1) is true with

$$\pi_Y(dy) = \int g(dy|\bar{y}, \bar{t}) \pi(d\bar{y}d\bar{t}). \tag{A.2}$$

The second part of the thesis is the law of large number for a (φ -irreducible and) positive Harris Markov chain. See Appendix C. \square

Lemma A.2. *M_n is a concave (random) function and $\hat{\beta}_n \in \operatorname{argmax} M_n(\beta)$.*

Proof. First observe that $M_n(\beta) = n^{-1}(\log(PL_n(\beta)) - \log(PL_n(\beta_0)))$ and hence $\hat{\beta}_n \in \operatorname{argmax} M_n(\beta)$. The concavity of M_n follows by showing that $\frac{d^2}{d\beta^2} M_n(\beta) \leq 0$. Since $A_n(\beta)$ is a linear function of β , then $\frac{d^2}{d\beta^2} M_n(\beta) = -\frac{d^2}{d\beta^2} B_n(\beta)$ and an easy computation gives

$$-\frac{d^2}{d\beta^2}B_n(\beta) = -\frac{1}{n} \sum_{i=1}^n \left(\frac{\sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} Z^2(T_i|Y_j) e^{\beta Z(T_i|Y_j)}}{\sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} e^{\beta Z(T_i|Y_j)}} - \left(\frac{\sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} Z(T_i|Y_j) e^{\beta Z(T_i|Y_j)}}{\sum_{j=1}^n \mathbb{I}\{T_j \geq T_i\} e^{\beta Z(T_i|Y_j)}} \right)^2 \right)$$

Finally, each term of the first sum is positive in view of the Cauchy-Schwarz Inequality. \square

Setting

$$c_n(t|\beta) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}\{T_j \geq t\} e^{\beta Z(t|Y_j)}$$

and $N_n(dt) := n^{-1} \sum_{i=1}^n \delta_{T_i}(dt)$ we can write

$$B_n(\beta) = \int_{\mathbb{R}_+} \log \left(\frac{c_n(t|\beta)}{c_n(t|\beta_0)} \right) N_n(dt).$$

Notice that the above integral is well defined since $c_n(T_i|\beta) > 0$ for all β and $i = 1, \dots, n$.

Furthermore, as $f(t|y) = S(t|y, \beta_0) e^{\beta_0 Z(t|y)} \lambda_0(t)$, one gets from (A.2) and (2.11) that

$$\begin{aligned} c(t|\beta) &= \int_{\mathbb{R}_+^2} S(t|y, \beta_0) e^{\beta Z(t|y)} \pi_Y(dy) \quad \text{and} \\ \pi_T(dt) &= \left(\int_{\mathbb{R}_+^2} S(t|y, \beta_0) e^{\beta_0 Z(t|y)} \pi_Y(dy) \right) \lambda_0(t) dt = c(t|\beta_0) \lambda_0(t) dt =: f_T(t) dt. \end{aligned} \tag{A.3}$$

Lemma A.3. *Under assumptions (a)-(d) of Proposition 2.1, $\pi_T[\bar{t}, +\infty) > 0$ for every $\bar{t} < +\infty$. Moreover, for every β , $0 < k^-(\beta) \pi_T[\bar{t}, +\infty) \leq c(t|\beta) \leq k^+(\beta) < +\infty$ for every $t \in [0, \bar{t}]$ and for suitable functions $k^-(\beta)$ and $k^+(\beta)$.*

Proof. Let $k^+(\beta) := \sup_{(t,y)} e^{Z(t|y)\beta} < +\infty$ and $k^-(\beta) := \inf_{(t,y)} e^{Z(t|y)\beta} > 0$. Clearly $1 \geq S(t|y, \beta_0) \geq \exp\{-k^+(\beta_0) \int_0^t \lambda_0(s) ds\}$ and hence, recalling (A.3):

$$\begin{aligned} f_T(t) &\geq \lambda_0(t) k^-(\beta_0) \int_{(\mathbb{R}_+^2)^2} e^{-k^+(\beta_0) \int_0^t \lambda_0(s) ds} \pi_Y(dy) \\ &= \lambda_0(t) k^-(\beta_0) e^{-k^+(\beta_0) \int_0^t \lambda_0(s) ds}. \end{aligned}$$

But, since $\int_0^t \lambda_0(s) ds < +\infty$ for every t and $\int_0^{+\infty} \lambda_0(s) ds = +\infty$, then it follows immediately that $\pi_T[\bar{t}, +\infty) = \int_{\bar{t}}^{+\infty} f_T(t) dt > 0$ for every $\bar{t} < +\infty$. Finally, $c(t|\beta) \leq k^+(\beta)$ and, recalling (A.3), $c(t|\beta) \geq k^-(\beta) \pi_T[\bar{t}, +\infty)$ for every $t \leq \bar{t}$. \square

Lemma A.4. *Under assumptions (b)-(e) of Proposition 2.1, the function M has a unique minimum in β_0 .*

Proof. First, in view of Lemma A.3, $c(t|\beta) > 0$ for every t, β and, by (b), one can differentiate under the integral sign to get

$$\begin{aligned} \frac{d}{d\beta} M(\beta) &= \int Z(t|y)\pi(dt dy) - \int \frac{\frac{\partial}{\partial\beta} c(t|\beta)}{c(t|\beta)} f_T(t) dt \\ &= \int Z(t|y)\pi(dt dy) - \int \frac{\int Z(t|y)S(t|y, \beta_0)e^{\beta Z(t|y)}\pi_Y(dy)}{c(t|\beta)} f_T(t) dt \end{aligned}$$

and

$$\begin{aligned} -\frac{d^2}{d\beta^2} M(\beta) &= \int \left(\frac{\int Z^2(t|y)S(t|y, \beta_0)e^{\beta Z(t|y)}\pi_Y(dy)}{c(t|\beta)} \right. \\ &\quad \left. - \left(\frac{\int Z(t|y)S(t|y, \beta_0)e^{\beta Z(t|y)}\pi_Y(dy)}{c(t|\beta)} \right)^2 \right) f_T(t) dt. \end{aligned}$$

For $\beta = \beta_0$ one has $f_T(t)/c(t|\beta) = \lambda_0(t)$, so that

$$\begin{aligned} &\int \frac{\int Z(t|y)S(t|y, \beta_0)e^{\beta_0 Z(t|y)}\pi_Y(dy)}{c(t|\beta)} f_T(t) dt \\ &= \int \int Z(t|y)S(t|y, \beta_0)e^{\beta_0 Z(t|y)}\pi_Y(dy)\lambda_0(t) dt = \int Z(t|y)\pi(dt dy), \end{aligned}$$

hence $\frac{d}{d\beta} M(\beta)|_{\beta=\beta_0} = 0$. Moreover, as

$$\frac{\int S(t|y, \beta_0)e^{\beta Z(t|y)}\pi_Y(dy)}{c(t|\beta)} = 1,$$

then the Cauchy-Schwarz inequality gives $\frac{d^2}{d\beta^2} M(\beta) \leq 0$ and, by assumption (e), $\frac{d^2}{d\beta^2} M(\beta_0) < 0$, which yields that β_0 is the unique maximum of M . \square

Lemma A.5. *If assumption (a)-(d) of Proposition 2.1 are in force, then*

$$\sup_{0 \leq t \leq \bar{t}} |c_n(t|\beta) - c(t|\beta)| \rightarrow 0 \quad \mathbb{P}\text{-a.s.} \tag{A.4}$$

for every $0 < \bar{t} < +\infty$ and every β .

Proof. Set $\zeta_n(t) := n^{-1} \sum_{j=1}^n \mathbb{I}\{T_j \geq t\}$ and define

$$\begin{aligned} \zeta(t) &:= \pi_T[t, +\infty), \\ \varpi(\delta, y, \beta) &:= \sup_{s, t \leq \bar{t}, |s-t| \leq \delta} |e^{Z(t|y)\beta} - e^{Z(s|y)\beta}|, \\ H(\delta, \beta) &:= \int \varpi(\delta, y, \beta)\pi_Y(dy). \end{aligned}$$

Now observe that $t \mapsto \zeta(t)$ and $t \mapsto c(t|\beta)$ are continuous functions, see (A.3), and hence they are uniformly continuous on $[0, \bar{t}]$. Moreover, it follows by (a)

and (c) that $\varpi(\delta, y, \beta) \leq 2e^{\beta\|Z\|_\infty}$ and that $\lim_{\delta \rightarrow 0} \varpi(\delta, y, \beta) = 0$ for every y, β . Hence, $\lim_{\delta \rightarrow 0} H(\delta, \beta) = 0$. Putting all together, for every ϵ there is δ such that

$$\sup_{s, t \leq \bar{t}, |s-t| \leq \delta} |c(t|\beta) - c(s|\beta)| \leq \epsilon \tag{A.5}$$

$$\sup_{s, t \leq \bar{t}, |s-t| \leq \delta} |\zeta(t) - \zeta(s)| \leq \epsilon \tag{A.6}$$

$$H(\delta, \beta) \leq \epsilon. \tag{A.7}$$

Now choose t_0, \dots, t_m such that $0 = t_0 < \dots < t_m = \bar{t}$ and $t_i - t_{i-1} \leq \delta$. Hence,

$$\begin{aligned} & \sup_{0 \leq t \leq \bar{t}} |c_n(t|\beta) - c(t|\beta)| \\ & \leq \max_i \sup_{t_{i-1} \leq t \leq t_i} \left\{ |c_n(t|\beta) - c_n(t_i|\beta)| + |c_n(t_i|\beta) - c(t_i|\beta)| + |c(t_i|\beta) - c(t|\beta)| \right\}. \end{aligned}$$

By (A.5)

$$\sup_{0 \leq t \leq \bar{t}} |c_n(t|\beta) - c(t|\beta)| \leq \max_i \sup_{t_{i-1} \leq t \leq t_i} |c_n(t|\beta) - c_n(t_i|\beta)| + J_{1,n}(\delta) + \epsilon$$

with $J_{1,n}(\delta) = \max_{i=1, \dots, m} |c_n(t_i|\beta) - c(t_i|\beta)|$. Now consider $t_{i-1} \leq t \leq t_i$. Clearly $\mathbb{I}\{T_j \geq t\} \geq \mathbb{I}\{T_j \geq t_i\}$, so that

$$\begin{aligned} |c_n(t|\beta) - c_n(t_i|\beta)| & \leq \left| \frac{1}{n} \sum_{j=1}^n (\mathbb{I}\{T_j \geq t\} - \mathbb{I}\{T_j \geq t_i\}) e^{Z(t|Y_j)\beta} \right| \\ & \quad + \frac{1}{n} \sum_{j=1}^n |e^{Z(t|Y_j)\beta} - e^{Z(t_i|Y_j)\beta}| \\ & \leq k^+(\beta) |\zeta_n(t) - \zeta_n(t_i)| + \frac{1}{n} \sum_{j=1}^n \varpi(\delta, Y_j, \beta) \\ & \leq k^+(\beta) |\zeta_n(t_{i-1}) - \zeta_n(t_i)| + \frac{1}{n} \sum_{j=1}^n \varpi(\delta, Y_j, \beta) \\ & \leq k^+(\beta) (|\zeta_n(t_{i-1}) - \zeta(t_{i-1})| + |\zeta(t_{i-1}) - \zeta(t_i)| + |\zeta(t_i) - \zeta_n(t_i)|) \\ & \quad + \frac{1}{n} \sum_{j=1}^n \varpi(\delta, Y_j, \beta) \end{aligned}$$

for $k^+(\beta) = \sup_{(t,y)} e^{Z(t|y)\beta} < +\infty$. Therefore, using (A.6),

$$\max_i \sup_{t_{i-1} \leq t \leq t_i} |c_n(t|\beta) - c_n(t_i|\beta)| \leq 2k^+(\beta) J_{2,n}(\delta) + k^+(\beta)\epsilon + \frac{1}{n} \sum_{j=1}^n \varpi(\delta, Y_j, \beta)$$

with $J_{2,n}(\delta) = \max_{i=1, \dots, m} |\zeta_n(t_i) - \zeta(t_i)|$. We have proved that

$$\sup_{0 \leq t \leq \bar{t}} |c_n(t|\beta) - c(t|\beta)| \leq J_n(\delta) + (k^+(\beta) + 1)\epsilon + H_n(\delta, \beta)$$

with $J_n(\delta) = J_{1,n}(\delta) + 2k^+(\beta)J_{2,n}(\delta)$ and $H_n(\delta, \beta) = n^{-1} \sum_{j=1}^n \varpi(\delta, Y_j, \beta)$. Now by Lemma A.1, $J_n(\delta)$ converges almost surely to zero and $H_n(\delta, \beta)$ converges almost surely to $H(\delta, \beta)$ with $H(\delta, \beta) \leq \epsilon$ by (A.7). This gives (A.4). \square

Lemma A.6. *Assume (a)-(d) of Proposition 2.1. Then, for every β , $M_n(\beta)$ converges \mathbb{P} -a.s. to $M(\beta)$.*

Proof. Let us fix β . Using Lemma A.2 it follows immediately that $A_n(\beta)$ converges a.s. to $(\beta - \beta_0) \int_{\mathbb{R}_+^2 \times \mathbb{R}_+} Z(t|y)\pi(dydt)$. Using assumption (b) it follows that

$$0 < c_-(\beta) := \frac{k^-(\beta)}{k^+(\beta_0)} \leq \frac{c_n(t|\beta)}{c_n(t|\beta_0)} \leq c_+(\beta) := \frac{k^+(\beta)}{k^-(\beta_0)} < +\infty$$

for all t, n, ω such that $\zeta_n(t) = n^{-1} \sum_{j=1}^n \mathbb{I}\{T_j \geq t\} > 0$. Moreover, by Lemma A.3 one has $0 < k_-(\beta)\pi_T[\bar{t}, +\infty) \leq c(t|\beta) \leq k_+(\beta) < +\infty$ for every $t \in [0, \bar{t}]$. In the light of these facts and (A.4), one can show that for every $\bar{t} < +\infty$ and every β

$$\sup_{0 \leq t \leq \bar{t}} |\log(c_n(t|\beta)) - \log(c(t|\beta))| \rightarrow 0 \quad \mathbb{P}\text{-a.s.}$$

Indeed, since $\inf_{t \in [0, \bar{t}]} c(t|\beta) \geq \delta > 0$, then $c_n(t|\beta) \geq \delta/2$ for every $t \in [0, \bar{t}]$ and every $n \geq \bar{n}(\omega, \bar{t})$, on a set of probability one. As $|\log(x) - \log(y)| = |\int_y^x t^{-1} dt| \leq |x - y|L$ for $x, y \geq L^{-1}$, it follows that

$$\sup_{0 \leq t \leq \bar{t}} |\log(c_n(t|\beta)) - \log(c(t|\beta))| \leq \frac{2}{\delta} \sup_{0 \leq t \leq \bar{t}} |c_n(t|\beta) - c(t|\beta)|$$

for every $n \geq \bar{n}(\omega, \bar{t})$. The claim follows.

To summarize, for every $\bar{t} < +\infty$ and every β , we proved that

$$\sup_{0 \leq t \leq \bar{t}} \left| \log \left(\frac{c_n(t|\beta)}{c_n(t|\beta_0)} \right) - \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \right| \rightarrow 0 \quad \mathbb{P}\text{-a.s.} \tag{A.8}$$

Recall that $N_n(dt) := n^{-1} \sum_{i=1}^n \delta_{T_i}(dt)$ and observe that $\zeta_n(T_i) > 0$ for every n, i . Hence, setting $\bar{c}(\beta) := \max(|\log(c_-(\beta))|, |\log(c_+(\beta))|)$, one can write, for every $\bar{t} < +\infty$,

$$\begin{aligned} |B_n(\beta) - \int \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \pi_T(dt)| &\leq \bar{c}(\beta)[N_n([\bar{t}, +\infty)) + \pi_T([\bar{t}, +\infty))] \\ &\quad + \left| \int_{[0, \bar{t}]} \log \left(\frac{c_n(t|\beta)}{c_n(t|\beta_0)} \right) N_n(dt) - \int_{[0, \bar{t}]} \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \pi_T(dt) \right| \\ &\leq \bar{c}(\beta)[N_n([\bar{t}, +\infty)) + \pi_T([\bar{t}, +\infty))] + \int_{[0, \bar{t}]} \left| \log \left(\frac{c_n(t|\beta)}{c_n(t|\beta_0)} \right) - \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \right| N_n(dt) \\ &\quad + \left| \int_{[0, \bar{t}]} \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) [N_n(dt) - \pi_T(dt)] \right| \\ &\leq \bar{c}(\beta)[N_n([\bar{t}, +\infty)) + \pi_T([\bar{t}, +\infty))] + \sup_{0 \leq t \leq \bar{t}} \left| \log \left(\frac{c_n(t|\beta)}{c_n(t|\beta_0)} \right) - \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \right| \end{aligned}$$

$$+ \left| \int_{[0, \bar{t})} \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) [N_n(dt) - \pi_T(dt)] \right| =: R_{1,n}(\bar{t}) + R_{2,n}(\bar{t}) + R_{3,n}(\bar{t}).$$

Now let \bar{t} such that $\pi_T([\bar{t}, +\infty)) \leq \epsilon(2\bar{c}(\beta))^{-1}$. Using Lemma A.1 one gets that $R_{1,n}(\bar{t})$ converges a.s. to $2\bar{c}(\beta)\pi_T([\bar{t}, +\infty)) \leq \epsilon$. Moreover, by (A.8), $R_{2,n}(\bar{t})$ converges a.s. to 0 and, by Lemma A.1, $R_{3,n}(\bar{t})$ converges a.s. to 0. This shows that $B_n(\beta)$ converges a.s. to $\int \log \left(\frac{c(t|\beta)}{c(t|\beta_0)} \right) \pi_T(dt)$ and the thesis follows. \square

Lemma A.7. *If assumptions (a)-(e) of Proposition 2.1 are in force then $\hat{\beta}_n \rightarrow \beta_0$ \mathbb{P} -a.s. and $\hat{\Lambda}_{0,n}(t)$ converges to $\int_{(0,t]} \lambda_0(s) ds$ for every $t > 0$ \mathbb{P} -a.s..*

Proof. Combining Thm. 7.7 in [6] with Lemmata A.2, A.4 and A.6, one gets that $\hat{\beta}_n \rightarrow \beta_0$ a.s. Note that this proves the first part of Proposition 2.1. Now observe that

$$\hat{\Lambda}_{0,n}(t) = \int_{(0,t]} \frac{1}{c_n(s|\hat{\beta}_n)} N_n(ds).$$

Once again, by Lemma A.3, for every $0 < t < +\infty$, $\inf_{s \in [0,t]} c(s|\beta_0) \geq \delta = \delta(t) > 0$ and hence $\inf_{s \in [0,t]} c_n(s|\beta_0) \geq \delta/2$ for every $n \geq \bar{n}(\omega, t)$ on a set of probability one. So that, for every $n \geq \bar{n}(\omega, t)$

$$\begin{aligned} \sup_{s \in (0,t]} \left| \frac{1}{c_n(s|\beta_0)} - \frac{1}{c(s|\beta_0)} \right| &= \sup_{s \in (0,t]} \left| \frac{c_n(s|\beta_0) - c(s|\beta_0)}{c_n(s|\beta_0)c(s|\beta_0)} \right| \\ &\leq \frac{2}{\delta^2} \sup_{s \in (0,t]} |c_n(s|\beta_0) - c(s|\beta_0)|. \end{aligned} \tag{A.9}$$

Now

$$\begin{aligned} &\left| \int_{(0,t]} \frac{1}{c_n(s|\hat{\beta}_n)} N_n(ds) - \int_{(0,t]} \frac{1}{c(s|\beta_0)} \pi_T(ds) \right| \\ &\leq \left| \int_{(0,t]} \frac{1}{c_n(s|\hat{\beta}_n)} N_n(ds) - \int_{(0,t]} \frac{1}{c_n(s|\beta_0)} N_n(ds) \right| \\ &\quad + \int_{(0,t]} \left| \frac{1}{c_n(s|\beta_0)} - \frac{1}{c(s|\beta_0)} \right| N_n(ds) \\ &\quad + \left| \int_{(0,t]} \frac{1}{c(s|\beta_0)} N_n(ds) - \int_{(0,t]} \frac{1}{c(s|\beta_0)} \pi_T(ds) \right| \\ &=: I_{1,n} + I_{2,n} + I_{3,n}. \end{aligned}$$

By Taylor expansion of $\beta \mapsto \int_{(0,t]} c_n^{-1}(s|\beta) N_n(ds)$, one gets

$$I_{1,n} = |\hat{\beta}_n - \beta_0| \left| \int_{(0,t]} \frac{\frac{\partial}{\partial \beta} c_n(s|\beta_n^*)}{c_n^2(s|\beta_n^*)} N_n(ds) \right|$$

for some β_n^* in the segment connecting $\hat{\beta}_n$ and β_0 , i.e. $|\beta_n^* - \beta_0| \leq |\beta_n - \beta_0|$. Since $\beta_n^* = \beta_n^* - \beta_0 + \beta_0 \geq -|\beta_n^* - \beta_0| + \beta_0$, then

$$c_n(t|\beta_n^*) \geq c_n(t|\beta_0) e^{-|\beta_n^* - \beta_0| \|Z\|_\infty}$$

and hence

$$\frac{1}{c_n(t|\beta_n^*)} \leq \frac{e^{|\beta_n^* - \beta_0| \|Z\|_\infty}}{c_n(t|\beta_0)}. \tag{A.10}$$

Therefore

$$\frac{\frac{\partial}{\partial \beta} c_n(s|\beta_n^*)}{c_n^2(s|\beta_n^*)} \leq \frac{e^{2|\beta_n^* - \beta_0| \|Z\|_\infty}}{c_n^2(t|\beta_0)} \|Z\|_\infty e^{|\beta_n^*| \|Z\|_\infty} \leq \frac{4}{\delta^2} \|Z\|_\infty e^{|\beta_n^*| \|Z\|_\infty} e^{2|\beta_n^* - \beta_0| \|Z\|_\infty}.$$

Since $\hat{\beta}_n \rightarrow \beta_0$ a.s. and hence $\beta_n^* \rightarrow \beta_0$, it follows that $I_{1,n}$ converges a.s. to 0.

As regards $I_{2,n}$, by (A.9)

$$I_{2,n} \leq \frac{2}{\delta^2} \sup_{s \in (0,t]} |c_n(s|\beta_0) - c(s|\beta_0)|.$$

By Lemma A.5 we know that (A.4) holds and then also $I_{2,n}$ converges a.s. to zero. Finally, using again $\inf_{s \in [0,t]} c(s|\beta_0) \geq \delta(t) > 0$ and the strong law of large numbers of Lemma A.1, $I_{3,n}$ converges a.s. to zero. We have proved that

$$\int_{(0,t]} \frac{1}{c_n(s|\hat{\beta}_n)} N_n(ds) \rightarrow \int_{(0,t]} \frac{1}{c(s|\beta_0)} \pi_T(ds) = \int_{(0,t]} \lambda_0(s) ds = \Lambda_0(t) \quad \mathbb{P}\text{-a.s.}$$

As $\hat{\Lambda}_{0,n}$ are increasing processes, the thesis follows. □

Proof of Proposition 2.2. From the expression of $\frac{d^2}{d\beta^2} M(\beta_0)$ one gets

$$\frac{d^2}{d\beta^2} M(\beta_0) = - \int \text{Var}(Z(t|Y_t)) f_T(t) dt$$

where, for any fixed $t > 0$, Y_t is a random variable with distribution

$$\frac{S(t|y, \beta_0) e^{\beta_0 Z(t|y)} \pi_Y(dy)}{c(t|\beta_0)}.$$

Hence

$$\begin{aligned} P(Z(t|Y_t) \neq c) &= \int_{\{y: Z(t|y) \neq c\}} \frac{S(t|y, \beta_0)}{c(t|\beta_0)} e^{\beta_0 Z(t|y)} \pi_Y(dy) \\ &\geq \frac{k^-(\beta_0)}{k^+(\beta_0)} e^{-k^+(\beta_0) \int_0^t \lambda_0(s) ds} \pi_Y(\{y : Z(t|y) \neq c\}), \end{aligned}$$

for every t and c , and the last term is greater than zero if and only if

$$\pi_Y(\{y : Z(t|y) = c\}) < 1.$$

Thus, if the last condition is true for every c and for almost all $t \in [a, b]$ with $\pi_T([a, b]) = \int_a^b f_T(u) du > 0$, then

$$\frac{d^2}{d\beta^2} M(\beta_0) < 0. \tag{□}$$

Proof of Proposition 2.3. Thanks to Proposition 2.1 there exists Ω_0 such that $\mathbb{P}(\Omega_0) = 0$ and for a fixed $\omega \in \Omega_0^c$ it follows that $\hat{\beta}_n \rightarrow \beta_0$ and $\hat{\Lambda}_{0,n}(t) \rightarrow \int_{(0,t]} \lambda_0(s)ds$ for every $t > 0$. Then, for every $K > 0$ and $t \leq K$,

$$\begin{aligned} |\hat{\Lambda}_n(t|y) - \int_{(0,t)} e^{\beta_0 Z(s|y)} \lambda_0(s)ds| &= \left| \int_{(0,t]} e^{\hat{\beta}_n Z(s|y)} \hat{\Lambda}_{0,n}(ds) - \int_{(0,t)} e^{\beta_0 Z(s|y)} \lambda_0(s)ds \right| \\ &\leq \int_{(0,t]} |e^{\hat{\beta}_n Z(s|y)} - e^{\beta_0 Z(s|y)}| \hat{\Lambda}_{0,n}(ds) \\ &\quad + \left| \int_{(0,t]} e^{\beta_0 Z(s|y)} \hat{\Lambda}_{0,n}(ds) - \int_{(0,t)} e^{\beta_0 Z(s|y)} \lambda_0(s)ds \right| \\ &\leq C(|\hat{\beta}_n - \beta_0| \hat{\Lambda}_{0,n}(K) + |\hat{\Lambda}_{0,n}(K) - \int_{(0,K)} \lambda_0(s)ds|) \end{aligned}$$

for $n \geq n_0(\omega)$ and a suitable constant C , since Z is a bounded function of s and y . By Lemma A.7 the last term of the inequality tends to zero and hence the first part of the corollary is proved. As a consequence, $\hat{R}_i = \hat{\Lambda}_n(T_i|Y_i) \rightarrow \Lambda_T(T_i|Y_i) = \int_{(0,T_i)} e^{\beta_0 Z(s|Y_i)} \lambda_0(s)ds$ a.s.. If one denotes by $\Lambda_T^{-1}(\cdot|y)$ a (generalized) inverse function of $\Lambda_T(\cdot|y)$, then, for every positive u , one has

$$\mathbb{P}\left\{T_i > \Lambda_T^{-1}(u|Y_i) \mid Y_i = y\right\} = \exp\{-\Lambda_T(\Lambda_T^{-1}(u|y)|y)\} = e^{-u}. \tag{A.11}$$

Using this fact, one can easily prove that the $\Lambda_T(T_i|Y_i)$'s are independent and exponentially distributed with mean 1 and the second part of the proposition follows. □

Appendix B: Proofs of Section 3

Proof of Proposition 3.1. Define

$$\begin{aligned} \rho(\bar{x}; dxdadt) &:= \left(\int_{\mathbb{R}^2} f_{X_I}(x|\bar{x}e^{\bar{a}\bar{t}}) f_\tau(\bar{t}|\bar{x}, \bar{a}) f_\alpha(\bar{a}) d\bar{a}d\bar{t} \right) f_\alpha(a) f_\tau(t|x, a) dxdadt \\ &= \mathbb{P}\{X_{I,n} \in dx, \alpha_n \in da, \tau_n \in dt | X_{I,n-1} = \bar{x}\}, \end{aligned}$$

for $n = 1, \dots$ and

$$\tilde{\varphi}(dxdadt) := \int_{\mathbb{R}_+} \varphi(d\bar{x}) \rho(\bar{x}; dxdadt).$$

Then for any measurable subset C of \mathbb{R}_+^3 such that $\tilde{\varphi}(C) > 0$ there exists $\mathbb{X}_0 \subset \mathbb{R}_+$, with $\mathbb{X}_0 = \mathbb{X}_0(C)$ measurable set satisfying $\varphi(\mathbb{X}_0) > 0$ and $\rho(\bar{x}; C) > 0$ for every $\bar{x} \in \mathbb{X}_0$. On the other hand, the φ -irreducibility of the embedded Markov Chain $(X_{I,n})_{n \geq 0}$ yields

$$\sum_{n \geq 1} \mathbb{P}\{X_{I,n} \in \mathbb{X}_0 | X_{I,0} = x\} \frac{1}{2^n} > 0 \tag{B.1}$$

for every $x \in \mathbb{R}_+$ see e.g. Proposition 4.2.1 in [8]. We now claim that (B.1) implies

$$\sum_{n \geq 1} \mathbb{P}\{(X_{I,n+j}, \alpha_{n+j}, \tau_{n+j}) \in C | X_{I,j} = x\} \frac{1}{2^n} > 0 \tag{B.2}$$

for every $x \in \mathbb{R}_+$ and $j = 0, 1, \dots$. Because of the homogeneity of the chain $(X_{I,n}, \alpha_n, \tau_n)$, it is sufficient to prove that (B.1) implies (B.2) when $j = 0$. To see this, note that

$$\begin{aligned} \sum_{n \geq 1} \mathbb{P}\{(X_{I,n}, \alpha_n, \tau_n) \in C | X_{I,0} = x\} \frac{1}{2^n} &= \frac{\rho(x; C)}{2} + \\ &+ \frac{1}{2} \int_{\mathbb{R}_+} \sum_{n \geq 1} \mathbb{P}\{(X_{I,n+1}, \alpha_{n+1}, \tau_{n+1}) \in C | X_{I,n} = \bar{x}\} \mathbb{P}\{X_{I,n} \in d\bar{x} | X_{I,0} = x\} \frac{1}{2^n} \\ &\geq \frac{1}{2} \int_{\mathbb{X}_0} \sum_{n \geq 1} \mathbb{P}\{(X_{I,n+1}, \alpha_{n+1}, \tau_{n+1}) \in C | X_{I,n} = \bar{x}\} \mathbb{P}\{X_{I,n} \in d\bar{x} | X_{I,0} = x\} \frac{1}{2^n} \\ &= \frac{1}{2} \int_{\mathbb{X}_0} \rho(\bar{x}; C) \sum_{n \geq 1} \mathbb{P}\{X_{I,n} \in d\bar{x} | X_{I,0} = x\} \frac{1}{2^n} > 0 \end{aligned}$$

by (B.1). Finally,

$$\begin{aligned} &\sum_{n \geq 1} \mathbb{P}\{(X_{I,n}, \alpha_n, \tau_n) \in C | (X_{I,0}, \alpha_0, \tau_0) = (x_0, a_0, t_0)\} \frac{1}{2^n} \\ &\geq \frac{1}{2} \sum_{n \geq 1} \mathbb{P}\{(X_{I,n+1}, \alpha_{n+1}, \tau_{n+1}) \in C | (X_{I,0}, \alpha_0, \tau_0) = (x_0, a_0, t_0)\} \frac{1}{2^n} \\ &= \frac{1}{2} \int_{\mathbb{R}_+} \left[\sum_{n \geq 1} \mathbb{P}\{(X_{I,n+1}, \alpha_{n+1}, \tau_{n+1}) \in C | X_{I,1} = x\} \frac{1}{2^n} \right] \\ &\quad \times \mathbb{P}\{X_{I,1} \in dx | (X_{I,0}, \alpha_0, \tau_0) = (x_0, a_0, t_0)\} \\ &\geq \frac{1}{2} \int_{\mathbb{R}_+} \left[\sum_{n \geq 1} \mathbb{P}\{(X_{I,n+1}, \alpha_{n+1}, \tau_{n+1}) \in C | X_{I,1} = x\} \frac{1}{2^n} \right] f_{X_I}(x | x_0 e^{a_0 t_0}) dx > 0 \end{aligned}$$

by (B.2) and the $\tilde{\varphi}$ -irreducibility of $(X_{I,n}, \alpha_n, \tau_n)_{n \geq 0}$ follows.

Let now $(X_{I,i})_{i \geq 0}$ be φ -irreducible and positive Harris. Then there exists a (unique) invariant probability measure π_X , and it follows from the first part of the proof that $(X_{I,i}, \alpha_i, \tau_i)_{i \geq 0}$ is $\tilde{\varphi}$ -irreducible with invariant probability measure π given by (3.1). Hence by Proposition C.2 it suffices to show that for all x $K(x, \cdot)$ is absolutely continuous with respect to π . To this aim consider $A \in \mathcal{B}(\mathbb{R}_+)$ such that

$$0 = \pi(A) = \int_A f_\alpha(a) f_\tau(t | x, a) \pi_X(dx) dadt = \int I(x) \pi_X(dx)$$

where $I(x) = \int_{\{(a,t) \in \mathbb{R}_+^2 \mid (x,a,t) \in A\}} f_\alpha(a) f_\tau(t|x,a) da dt$. Then $I(x) = 0$ π_X -a.s.. Since $(X_{I,i})_{i \geq 0}$ is \mathcal{L} -irreducible, then \mathcal{L} is absolutely continuous with respect to π_X (see Appendix C) and then $I(x) = 0$ a.s.. It follows that

$$\begin{aligned} K((\bar{x}, \bar{a}, \bar{t}), A) &= \int_A f_{X_I}(x|\bar{x}e^{\bar{a}\bar{t}}) f_\alpha(a) f_\tau(t|x,a) dx da dt \\ &= \int f_{X_I}(x|\bar{x}e^{\bar{a}\bar{t}}) I(x) dx = 0, \quad \forall (\bar{x}, \bar{a}, \bar{t}), \end{aligned}$$

which completes the proof.

The statement for the chain $(X_{I,n}, \alpha_n, \Delta_n)_{n \geq 0}$ can be proved in a similar way. □

Lemma B.1. *Let $k(x, y)$ be the density kernel of $K(x, dy)$ with respect to the Lebesgue measure \mathcal{L} . If there exist $\epsilon > 0$ and $N \subset \mathbb{R}_+$ such that $\mathcal{L}(N) = 0$ and*

$$k(x, y) > 0 \quad \text{for all } x \in \mathbb{R}_+ \text{ and every } y \in (x(1 - \epsilon), x(1 + \epsilon)) \cap N^c \quad (\text{B.3})$$

then K is \mathcal{L} -irreducible. Analogously, if the density kernel k^ of K^* is such that for some $\epsilon' > \epsilon > 0$ and $N \subset [0, +\infty)$ with $\mathcal{L}(N) = 0$*

$$k^*(0, y) > 0 \quad \text{for every } y \in (0, \epsilon') \cap N^c$$

$$k^*(x, y) > 0 \quad \text{for all } x \in \mathbb{R}_+ \text{ and every } y \in (x(1 - \epsilon), x(1 + \epsilon)) \cap N^c$$

then K^ is \mathcal{L} -irreducible.*

Proof. By induction it is easy to show that for every $x \in N^c$ and every $y \in (x(1 - \epsilon)^n, x(1 + \epsilon)^n) \cap N^c$, one has $k^n(x, y) > 0$. Indeed, assume $x \in N^c$ and $k^n(x, y) > 0$ for every $y \in (x(1 - \epsilon)^n, x(1 + \epsilon)^n) \cap N^c$ and, for any $y \in (x(1 - \epsilon)^{n+1}, x(1 + \epsilon)^{n+1}) \cap N^c$, set $A_y := (x(1 - \epsilon)^n, x(1 + \epsilon)^n) \cap (y(1 + \epsilon)^{-1}, y(1 - \epsilon)^{-1}) \cap N^c$. Note that, since $y(1 + \epsilon)^{-1} < x(1 + \epsilon)^n$ and $x(1 - \epsilon)^n < y(1 - \epsilon)^{-1}$, then $\mathcal{L}(A_y) > 0$. Now if $z \in A_y$ then $k^n(x, z) > 0$ and $k(z, y) > 0$, hence

$$k^{n+1}(x, y) \geq \int_{A_y} k^n(x, z) k(z, y) dz > 0.$$

This shows that, for every $x > 0, y > 0$ and $x, y \notin N$, there is $n = n(x, y)$ such that $k^n(x, y) > 0$, so that

$$\sum_{n \geq 1} k^n(x, y) > 0 \quad \forall x, y > 0,$$

which gives the thesis of the first part of the lemma. The second part can be proved in a similar way. □

Proof of Proposition 3.2. Recall that the kernel K of the imbedded chain $(X_{I,i})_{i \geq 0}$ in the time increment representation is

$$K(x, dy) = \left(\int_{\mathbb{R}_+^2} \frac{1}{xe^{at}} f_\eta\left(\frac{y}{xe^{at}} \mid xe^{at}\right) f_\tau(t|x, a) dt f_\alpha(a) da \right) dy$$

and hence, for every x and almost all y , its kernel density k satisfies

$$k(x, y) \geq \int_{B_x} \frac{1}{xe^{at}} f_\eta\left(\frac{y}{xe^{at}} \middle| xe^{at}\right) f_\tau(t|x, a) dt f_\alpha(a) da.$$

Thanks to the assumptions and the definition of B_x , if $y \in (x(1 - \epsilon), x(1 + \epsilon))$ the integrand is positive on B_x and so does $k(x, y)$. The \mathcal{L} -irreducibility of the imbedded chain follows by Lemma B.1. \square

Proof of Proposition 3.3. Using the continuity assumption (c1) and the Scheffé’s theorem, one gets that for every measurable set A the function $x \mapsto \mathbb{P}\{X_{I,i+1} \in A | X_{I,i} = x\}$ is continuous on \mathbb{R}_+ and hence the chain is strongly (and a fortiori weakly) Feller. Alternatively, if (c2) holds, the dominated convergence yields that $\lim_{x \rightarrow \bar{x}} \int g(y) K(x, dy) = \int g(y) K(\bar{x}, dy)$, for every $\bar{x} > 0$ and every bounded continuous function g . Thus the chain is weakly Feller. \square

Proof of Lemma 3.1. Condition (3.2) yields the drift condition (C.1) of Proposition C.1 in Appendix C for the chain $W_n = X_{I,n}$, for suitable positive constants ϵ and C , a compact set $A = [a, b]$ in \mathbb{R}_+ and the test function $V(x) = \log^2(x)$. As the Lebesgue measure \mathcal{L} on \mathbb{R}_+ has support with non-empty interior, $\mathcal{L}[a, b] > 0$ if $a < b$, and the sub-levels of V are clearly compact sets, then the thesis follows. \square

Proof of Proposition 3.4. Set $c^* := e^{z^-}$. Noting that $S_\tau(t|x, a) \leq S_{0,\tau}(t)^{c^*}$, one has

$$\begin{aligned} \mathbb{E}[\alpha_i^2 \tau_i^2 | X_{I,i} = x] &= \int_{\mathbb{R}_+ \times \mathbb{R}_+} a^2 2t S_\tau(t|x, a) dt f_\alpha(a) da \\ &\leq 2 \int_{\mathbb{R}_+} a^2 f_\alpha(a) da \int_{\mathbb{R}_+} t S_{0,\tau}(t)^{c^*} dt. \end{aligned}$$

Hence, since $\log(\eta_i e^{\alpha_i \tau_i}) = \log(\eta_i) + \alpha_i \tau_i$, assumption (3.5) yields that

$$\sup_{x>0} \mathbb{E}[\log^2(\eta_i e^{\alpha_i \tau_i}) | X_{I,i} = x] < +\infty.$$

Now

$$\begin{aligned} \limsup_{x \rightarrow +\infty} \mathbb{E}[\log(\eta_i e^{\alpha_i \tau_i}) | X_{I,i} = x] &\leq \limsup_{x \rightarrow +\infty} \mathbb{E}[\log(\eta_i) | X_{I,i} = x] \\ &\quad + \limsup_{x \rightarrow +\infty} \mathbb{E}[\alpha_i \tau_i | X_{I,i} = x]. \end{aligned}$$

Since $S_\tau(t|x, a) \leq S_{0,\tau}(t)^{c^*}$ one can use (3.5) and dominated convergence to write

$$\begin{aligned} \lim_{x \rightarrow +\infty} \mathbb{E}[\alpha_i \tau_i | X_{I,i} = x] &= \lim_{x \rightarrow +\infty} \int a S_\tau(t|x, a) dt f_\alpha(a) da \\ &= \int a \lim_{x \rightarrow +\infty} S_\tau(t|x, a) dt f_\alpha(a) da \end{aligned}$$

$$= \int_{\mathbb{R}_+} a f_\alpha(a) da \int_{\mathbb{R}_+} S_{0,\tau}(t) e^{\beta_0 z_\infty} dt.$$

The first of the two conditions (3.6) yields $\limsup_{x \rightarrow +\infty} \mathbb{E}[a_i \tau_i | X_{I,i} = x] < 0$. Analogously, by dominated convergence

$$\begin{aligned} \lim_{x \rightarrow 0^+} \mathbb{E}[\alpha_i \tau_i | X_{I,i} = x] &= \lim_{x \rightarrow 0^+} \int_{\mathbb{R}_+ \times \mathbb{R}_+} a f_\alpha(a) da S_\tau(t|x, a) dt da \\ &= \int_{\mathbb{R}_+ \times \mathbb{R}_+} a f_\alpha(a) \lim_{x \rightarrow 0^+} S_\tau(t|x, a) dt da \\ &= \int_{\mathbb{R}_+} a f_\alpha(a) da \int_{\mathbb{R}_+} S_{0,\tau}(t) e^{\beta_0 z_0} dt. \end{aligned}$$

Hence the second of the two conditions (3.6) yields $\liminf_{x \rightarrow 0^+} \mathbb{E}[a_i \tau_i | X_{I,i} = x] > 0$. This proves (3.3) and, in its turn, (3.2). The thesis follows by Lemma 3.1. \square

Proof of Proposition 3.5. Recall that the kernel K^* of the imbedded chain $(X_{I,i})_{i \geq 0}$ in the space increment representation is

$$K^*(x, dy) = \left(\int_{\mathbb{R}_+^2} \frac{1}{x + \delta} f_\eta\left(\frac{y}{x + \delta} \middle| x + \delta\right) f_\Delta(\delta|x, a) d\delta f_\alpha(a) da \right) dy$$

and hence, for every x and almost all y , its kernel density k^* satisfies

$$k^*(x, y) \geq \int_{B_x^*} \frac{1}{x + \delta} f_\eta\left(\frac{y}{x + \delta} \middle| x + \delta\right) f_\Delta(\delta|x, a) d\delta f_\alpha(a) da.$$

Thanks to the assumptions and the definition of B_x , if $y \in (0, u_2(0)d_1(0))$ and $x = 0$ or if $y \in (x(1 - \epsilon), x(1 + \epsilon))$ and $x > 0$, the integrand is positive on B_x^* and so does $k^*(x, y)$. The \mathcal{L} -irreducibility of the imbedded chain follows again by Lemma B.1. \square

Proof of Proposition 3.6. The proof follows the same lines of the proof of Proposition 3.3. \square

Proof of Lemma 3.2. We shall apply Proposition C.1 in Appendix C to the chain $W_n = X_{I,n}$ for $V(x) = x^\gamma$. Since the Lebesgue measure \mathcal{L} on $[0, +\infty)$ has support with non-empty interior and the sub-levels of V are compact sets, it is enough to verify (C.1) for a suitable closed interval A . As $0 < \gamma < 1$ and $0 \leq \eta_n \leq 1$ a.s., one has

$$\begin{aligned} \mathbb{E}[X_{I,n+1}^\gamma | X_{I,n} = x] - x^\gamma &= \mathbb{E}[(x + \Delta_n) \eta_n]^\gamma | X_{I,n} = x] - x^\gamma \\ &\leq x^\gamma (\mathbb{E}[\eta_n^\gamma | X_{I,n} = x] - 1) + C \end{aligned}$$

with $C := \sup_{x \geq 0} \mathbb{E}[\Delta_n | X_{I,n} = x] < +\infty$ by (3.7). Now $\mathbb{E}[\eta_n^\gamma | X_{I,n} = x] - 1 \leq 0$ for every $x \geq 0$ and $\mathbb{E}[\eta_n^\gamma | X_{I,n} = x] - 1 \leq -\epsilon$ for x large enough by (3.8). \square

Proof of Proposition 3.7. It is enough to note that

$$\begin{aligned} \mathbb{E}[\Delta_i^\gamma | X_{I,i} = x] &= \int_0^{+\infty} \int_0^{+\infty} \gamma u^{\gamma-1} S_\Delta(\delta|x, a) d\delta f_\alpha(a) da \\ &\leq \gamma \int_0^{+\infty} u^{\gamma-1} S_{0,\Delta}(\delta)^{e^{z-}} d\delta \end{aligned}$$

and hence condition (3.8) is satisfied. The thesis follows by Lemma 3.2. \square

Appendix C: Some definitions and properties of homogenous Markov chains

A homogeneous Markov chain $(W_n)_{n \geq 0}$ taking values in a measurable space $(\mathbb{U}, \mathcal{U})$ with transition kernel $\mathcal{K}(u_0, du) = \mathbb{P}\{W_{n+1} \in du | W_n = u_0\}$ is said to be φ -irreducible if there is a $(\sigma$ -finite non trivial) measure φ such that $\sum_{n \geq 1} \mathcal{K}^n(u_0, A) > 0$ for every $u_0 \in \mathbb{U}$ and $A \in \mathcal{U}$ with $\varphi(A) > 0$.

A φ -irreducible Markov chain with kernel \mathcal{K} is said to be *Harris recurrent* if, for every A such that $\psi(A) = \int \sum_{n \geq 0} \mathcal{K}^n(x, A) 2^{-(n+1)} \varphi(dx) > 0$ and every $u_0 \in A$, $\mathbb{P}(W_n \in A \text{ i.o.} | W_0 = u_0) = 1$. Finally a φ -irreducible Markov chain is said to be *positive Harris* if it is Harris recurrent and it admits an invariant probability measure, that is if there is a probability measure π_W such that $\pi_W(A) = \int \mathcal{K}(u, A) \pi_W(du)$ for every $A \in \mathcal{U}$. In point of fact, if \mathcal{K} is positive Harris, the invariant probability measure π_W is unique and every irreducibility measure φ is absolutely continuous with respect to π . See e.g. Thm. 10.4.4 and Thm. 10.4.9 in [8].

If a φ -irreducible Markov chain $(W_n)_{n \geq 0}$ is positive Harris, then it satisfies the strong law of large numbers, i.e. for every initial condition $W_0 = u_0$

$$\frac{1}{n} \sum_{i=1}^n g(W_i) \rightarrow \int g(u) \pi_W(du) \quad \mathbb{P}_{u_0}\text{-a.s.}$$

whenever $\int |g(u)| \pi_W(du) < +\infty$. See, e.g. Thm. 17.0.1 in [8].

Furthermore, recall that if $(\mathbb{U}, \mathcal{U})$ is a topological space with the Borel σ -field, the chain $(W_n)_{n \geq 0}$, or equivalently its transition kernel \mathcal{K} , is said to be weakly Feller if \mathcal{K} maps $C_b(\mathbb{U})$ in $C_b(\mathbb{U})$ that is $w \rightarrow \int \mathcal{K}(w, du) g(u)$ is continuous for every bounded and continuous function g . Analogously, $(W_n)_{n \geq 0}$, or equivalently its transition kernel \mathcal{K} , is said to be strongly Feller if $w \rightarrow \int \mathcal{K}(w, du) g(u)$ is continuous for every bounded and measurable function g .

Let now suppose that the homogenous Markov chain $(W_n)_{n \geq 0}$ takes values in a Borel subset of \mathbb{R}^n and let \mathcal{K} denote its transition kernel as above. A sufficient condition for $(W_n)_{n \geq 0}$ to be Harris positive is the following version of the so-called “drift” criteria.

Proposition C.1. *Let the chain $(W_n)_{n \geq 0}$ be weakly Feller and φ -irreducible with φ such that the support of φ has non-empty interior. If the drift conditions*

$$\begin{aligned} \mathbb{E}[V(W_{n+1}) | W_n = w] - V(w) &\leq -\epsilon \quad \forall w \in A^c \\ \mathbb{E}[V(W_{n+1}) | W_n = w] - V(w) &\leq C \quad \forall w \in A \end{aligned} \tag{C.1}$$

hold for some compact set A with $\varphi(A) > 0$, some constants $\epsilon > 0$ and $C < +\infty$, and some non-negative, measurable function V such that the sets $\{w : V(w) \leq \lambda\}$ are compact sets for every $\lambda > 0$, then $(W_n)_{n \geq 1}$ (or equivalently \mathcal{K}) is positive Harris.

Proof. Since the chain is weakly Feller and in view of the property of the irreducibility measure φ , it follows from Thm. 6.0.1 and Thm. 6.2.7 in [8] that every compact set is petite. Hence, the function V satisfies that $\{w : V(w) \leq \lambda\}$ is petite for every $\lambda > 0$, i.e. V are unbounded off petite sets (see p. 191 in [8]). As (C.1) implies that

$$\Delta V(w) := \int V(u)\mathcal{K}(w, du) - V(w) \leq 0 \quad \forall w \in A^c \tag{C.2}$$

with A compact and hence petite set, then by Thm. 8.4.3 and Prop. 9.1.7 (ii) in [8] we obtain that the chain $(W_n)_{n \geq 1}$ is Harris recurrent. Finally, by Thm. 11.0.1 in [8], the drift conditions (C.1) imply that the invariant measure has finite total mass, i.e. up to a constant it is a probability measure. That is the chain $(W_n)_{n \geq 1}$ is positive Harris. \square

Finally, let us recall Corollary 1 in [13], which gives a sufficient condition for Harris recurrence of the kernel \mathcal{K} when there exists an invariant σ -finite measure.

Proposition C.2. *Suppose \mathcal{K} is π -irreducible with $\pi\mathcal{K} = \pi$. If $\mathcal{K}(x, \cdot)$ is absolutely continuous with respect to π for all x , then \mathcal{K} is Harris recurrent.*

Appendix D: Supplementary Tables

TABLE 8

Values of the Kolmogorov-Smirnoff statistic for comparing empirical versus simulated data in MG-A, MG-B and MG-C cases. (Artificial α sampled from gaussian distributions, artificial η from the corresponding empirical distributions).

dataset	variable	$[\Delta \text{null}]$	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
MG-A	X_I	0.126	0.071	0.076	0.234	0.216	0.208
	X_F	0.112	0.059	0.061	0.244	0.226	0.221
	τ	0.113	0.071	0.080	0.048	0.035	0.032
	Δ	0.009	0.012	0.011	0.193	0.174	0.171
MG-B	X_I	0.087	0.059	0.114	0.165	0.244	0.250
	X_F	0.077	0.039	0.086	0.172	0.249	0.256
	τ	0.113	0.076	0.094	0.042	0.048	0.051
	Δ	0.008	0.013	0.013	0.149	0.199	0.197
MG-C	X_I	0.119	0.094	0.190	0.370	0.323	0.319
	X_F	0.100	0.070	0.152	0.347	0.297	0.291
	τ	0.097	0.081	0.122	0.088	0.084	0.086
	Δ	0.011	0.009	0.020	0.246	0.194	0.188

TABLE 9

Values of the Kolmogorov-Smirnoff statistic for comparing empirical versus simulated data in CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR cases. (Artificial α and η sampled from the corresponding empirical distributions).

dataset	variable	$[\Delta \text{null}]$	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
CAA-P5ori	X_I	0.120	0.140	0.151	0.251	0.246	0.177
	X_F	0.055	0.089	0.095	0.235	0.243	0.146
	τ	0.145	0.130	0.103	0.044	0.119	0.065
	Δ	0.008	0.021	0.041	0.179	0.142	0.118
Glc-MRR	X_I	0.098	0.128	0.157	0.252	0.222	0.212
	X_F	0.056	0.065	0.092	0.228	0.194	0.176
	τ	0.122	0.110	0.113	0.040	0.058	0.040
	Δ	0.011	0.016	0.051	0.168	0.143	0.125
Glc-P5ori	X_I	0.147	0.149	0.148	0.191	0.443	0.233
	X_F	0.091	0.092	0.089	0.172	0.459	0.231
	τ	0.082	0.085	0.084	0.063	0.157	0.112
	Δ	0.007	0.009	0.015	0.109	0.301	0.137
LB-MRR	X_I	0.167	0.184	0.202	0.321	0.304	0.276
	X_F	0.110	0.119	0.131	0.299	0.283	0.260
	τ	0.113	0.103	0.105	0.045	0.045	0.055
	Δ	0.008	0.017	0.019	0.223	0.207	0.180

TABLE 10

$\hat{\beta}_n$ estimates of the regression parameter β_0 under the Cox models listed in Section 2.2 with $\mathcal{Z}(x) = \min(x, 350)$, for the seven datasets under study.

dataset	size	$[\Delta X_I]$	$[\Delta \tau]$	$[\tau \Delta]$	$[\tau X_I]$	$[\tau X_F]$
MG-A	10000	-0.3532	0.0750	0.3047	0.3129	0.1657
MG-B	10000	-0.2657	0.0620	0.2823	0.2633	0.1530
MG-C	10000	-0.108	0.0292	0.2353	0.2633	0.1387
CAA-P5ori	5905	-0.2540	0.0445	0.6884	1.0694	0.5773
Glc-MRR	2703	-0.2679	0.0444	0.3531	0.4547	0.2527
Glc-P5ori	979	-0.0281	-0.0039	1.7775	0.6192	1.0320
LB-MRR	2296	-0.1076	0.0313	0.1807	0.2246	0.11237

TABLE 11

p -values of the Kolmogorov-Smirnoff test to check if the estimated Cox-Snell residuals of the Cox model $[\Delta|X_I]$ are exponential with rate one for MG-A, MG-B, MG-C, CAA-P5ori, Glc-P5ori, Glc-MRR, LB-MRR datasets under $\mathcal{Z}(x) = \min(x, 350) = \mathcal{Z}_1(x)$ on the first row, $\mathcal{Z}(x) = \hat{q}_{0.015}\mathbb{I}_{[0, \hat{q}_{0.015}]}(x) + x\mathbb{I}_{(\hat{q}_{0.015}, \hat{q}_{0.985}]}(x) + \hat{q}_{0.985}\mathbb{I}_{[\hat{q}_{0.985}, +\infty)}(x) = \mathcal{Z}_2(x)$ on the second row and $\mathcal{Z}(x) = x\mathbb{I}_{(0, \hat{q}_{0.98}]}(x) + \hat{q}_{0.98}\mathbb{I}_{[\hat{q}_{0.98}, +\infty)}(x) = \mathcal{Z}_3(x)$ on the third row.

	MG-A	MG-B	MG-C	CAA-P5ori	Glc-MRR	Glc-P5ori	LB-MRR
\mathcal{Z}_1	0.2823	0.6312	0.3158	0.7413	0.6538	1.0000	0.9922
\mathcal{Z}_2	0.0724	0.1393	0.2132	0.5611	0.1571	1.0000	0.9802
\mathcal{Z}_3	0.0752	0.0778	0.2063	0.5686	0.1481	1.0000	0.9726

Acknowledgements

We are grateful to M. Cosentino Lagomarsino, M. Osella and A. S. Kennard for many useful discussions and for sharing with us the data from the papers [4] and [9]. We would like to thank the anonymous reviewers and an associate editor for their suggestions and remarks.

References

- [1] AMIR, A. 2014. Cell size regulation in bacteria. *Physical Review Letters*, **112**, 208102_1-208102_5.
- [2] ANDERSEN, P., & GILL R. 1982. Cox's regression model for counting processes: a large sample study. *Annals of Statistics*, **10**, 1100–1120. [MR0673646](#)
- [3] GOLDIE, C. M., & MALLER, R. A. 2000. Stability of perpetuities. *Annals of Probability*, **28**, 1195–1218. [MR1797309](#)
- [4] KENNARD, A. S., OSELLA, M. C., JAVER, A., GRILLI, J., NGHE, P., TANS, S., CICUTA, P., & COSENTINO LAGOMARSINO, M. C. 2015. Individuality and universality in the growth-division laws of single E. coli cells. *Physical Review E*, **93**, 012408
- [5] KLEIN, J. P., & MOESCHBERGER, M. L. 2003. *Survival analysis: Techniques for censored and truncated data*. Springer, New York.
- [6] LIESE, F., & MIESCKE, K. J. 2008. *Statistical decision theory*. Springer, New York. [MR2421720](#)
- [7] MARANTAN, A. & AMIR, A. 2016. Stochastic modeling of cell growth with symmetric or asymmetric division. *arXiv Preprint arXiv:1602.01848 [q-bio]*.
- [8] MEYN, S., & TWEEDIE, R. L. 1993. *Markov Chains and Stochastic Stability*. Springer-Verlag, London. [MR1287609](#)
- [9] OSELLA, M., NUGENT, E., & COSENTINO LAGOMARSINO, M. 2014. Concerted control of Escherichia coli cell division. *PNAS* DOI: 10.1073/pnas.1313715111.
- [10] OSELLA M., TANS, S. J. & COSENTINO LAGOMARSINO, M. 2017. Step by Step, Cell by Cell: Quantification of the Bacterial Cell Cycle. *Trends in Microbiology*, **25**, 250–256.
- [11] SOIFER, I., ROBERT, L., & AMIR, A. 2016. Single-Cell Analysis of Growth in Budding Yeast and Bacteria Reveals a Common Size Regulation Strategy. *Current Biology*, **26**, 356–361.
- [12] TAHERI-ARAGHI, S., BRADDE, S., SAULS, J., HILL, N. S., LEVIN, P. A., PAULSSON, J., VERGASSOLA, M., & JUN, S. 2015. Cell-size control and homeostasis in bacteria. *Current Biology*, **25**, 385–391.
- [13] TIERNEY, L. 1994. Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, **22**, 1701–1762. [MR1329166](#)
- [14] TURNER, J. J., EWALD, J. C., & SKOTHEIM, J. M. 2012. Cell size control in yeast. *Current Biology*, **22**, R350–R359.
- [15] TYSON, J. J., & DIEKMANN, O. 1986. Sloppy size control of the cell division cycle. *Journal of Theoretical Biology*, **118**, 405–426. [MR0829880](#)
- [16] VENABLES, W. N., & RIPLEY, B. D. 2002. *Modern Applied Statistics with S, Fourth Edition* Springer, New York. [MR1337030](#)
- [17] WANG, P., ROBERT, L., PELLETIER, J., DANG, W. L., TADDEI, F., WRIGHT, A. & JUN, S. 2010. Robust growth of Escherichia coli. *Current Biology*, **20**, 1099–1103.

- [18] WILCOX, R. R. 2009. Nonparametric Estimation in *Handbook of Computational Econometrics* (eds D. A. Belsley and E. J. Kontoghiorghes), John Wiley & Sons, Ltd, Chichester. [MR1973663](#)
- [19] ZUCKER, D. M. 1998. Restricted mean life with covariates: modification and extension of a useful survival analysis method. *Journal of the American Statistical Association*, **93**, 702–709. [MR1631365](#)