On the modelling and statistical analysis of a gamma deteriorating system with imperfect maintenance

Gabriel Salles



UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR - ÉCOLE DOCTORALE 211
Laboratoire de Mathématiques et de leurs Applications de Pau (LMAP)

Gabriel Salles

On the modelling and statistical analysis of a gamma deteriorating system with imperfect maintenance

Thesis committee

Laurent Bordes	Professor, University of Pau	Supervisor
Jean-Yves Dauxois	Professor, INSA Toulouse	Examiner
Olivier Gaudoin	Professor, University of Grenoble	Reviewer
Massimiliano Giorgio	Professor, University of Naples Federico II	Reviewer
Sophie Mercier	Professor, University of Pau	Supervisor
Christian Paroissin	Assistant professor, University of Pau	Examiner

A thesis submitted for the degree of
Doctor of Philosophy in Mathematics

October 2020

Contents

Acronyms

ARA_1	Arithmetic Reduction of Age of order 1
ARD_1	Arithmetic Reduction of Degradation of order 1
ARD_m	Arithmetic Reduction of Degradation of order m
ARD_∞	Arithmetic Reduction of Degradation of order ∞
ARI_1	Arithmetic Reduction of Intensity of order 1
a.s.	almost surely
CBM	Condition-Based Maintenance
CLO	Composite Likelihood based on the Observations
CLI	Composite Likelihood based on the Increments
EB	Empirical Bias
ECR	Exponential Convergence Rate
HDEI	Half Data based on the Even Indexes
HDOI	Half Data based on the Odd Indexes
i.i.d.	independent and identically distributed
MC	Monte Carlo
MLE	Maximum Likelihood Estimation
ME	Moments Estimation
p.d.f	probability density function
PNEE	Proportion of Numerically Exact Estimates
QMC	Quasi Monte Carlo
r.v.	Random Variable
S-ECR	Sub-Exponential Convergence Rate

Part I

Introduction

Safety and dependability are crucial issues in many industries (such as, e.g., railways, aircraft engines or nuclear power plants), which have lead to the development of the so-called reliability theory. For many years, only lifetime data were available and the first reliability studies were focused on lifetime data analysis (see, e.g., [32]), which still remains of interest in many cases. In that context and in case of repairable systems with instantaneous repairs, successive failure (or repair) times appear as the arrival points of a counting process, and failures hence correspond to recurrent events. As for the type of possible repairs, typical classical models are perfect (As-Good-As-New) and minimal (As-Bad-As-Old) repairs, leading to renewal and non homogeneous Poisson processes as underlying counting processes, respectively (see [5]). Regarding the failure intensity, the effects of both these maintenance types are illustrated in **Figure D.1**: a minimal repair does not affect the failure intensity, while a perfect repair reduces it to its initial value. However, the reality often lies in-between, leading to the class of imperfect repairs (also represented in **Figure D.1**). Many models have been envisioned in the literature for their

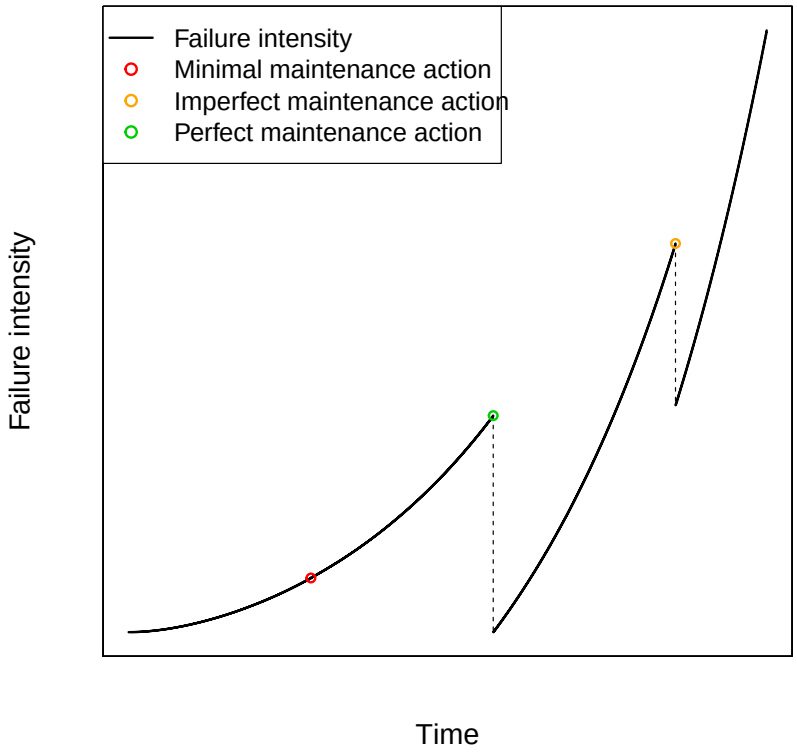

Figure 1

modeling, such as, e.g., virtual age models introduced by Kijima [26], further studied in [13, 16], and extended in [7] where the authors introduce covariates in the virtual age model. Other possible models are geometric processes [27] (extended in [6] and more recently in [12]) or, as already mentioned, models based on reduction of failure intensity [13, 16]. See, e.g., [17] for a recent account and extensions of such models. See also [36] for more references and other models.

Nowadays, the development of online monitoring and the increasing use of sensors for safety assessment make it possible to get specific information on the health of a system and on its effective evolution over time, without waiting for the system failure. This information is often synthesized into a scalar indicator,

which can for instance stand for the length of a crack, the thickness of a cable, the intensity of vibrations, corrosion level, ... This scalar indicator can be considered as a measurement of the deterioration level of the system. The evolution of this deterioration indicator over time is nowadays commonly modeled through a continuous-time and continuous-state stochastic process, which is often considered to have an increasing trend. Classical models include inverse Gaussian [44] or Wiener processes (with trend) [22, 29, 45]. Also, the transformed Wiener process was lately introduced in [20], and further studied in [18], where the degradation increments can be negative. All these stochastic processes are quite common in many other fields out of reliability theory, such as finance, insurance or epidemiology. This thesis focuses on gamma processes, which are widely used since they were introduced in the reliability field mostly simultaneously by Abdel-Hameed [1] and Çinlar [10]. This process is monotonous and is hence well adapted for modeling non decreasing degradation.

Before coming to the definition of a gamma process, let us start with the definition of the gamma distribution, which allows to set up the notations and parametrization used in this document.

A random variable X is said to be gamma distributed with shape parameter $a > 0$ and rate parameter $b > 0$ (denoted $X \sim \Gamma(a, b)$), if its distribution admits the following probability density function (p.d.f.):

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \, \mathbf{1}_{\mathbb{R}^+}(x)$$

with respect to the Lebesgue measure. The corresponding expectation and variance are $\mathbb{E}(X) = a/b$ and $\mathbb{V}(X) = a/b^2$.

Let us now recall some well-known properties of the gamma distribution:
Let X_1 and X_2 be two independent and gamma distributed random variables with respective distributions $\Gamma(a_1\,, b)$ and $\Gamma(a_2\,, b)$, where $a_1, a_2, b > 0$. Then, for all $c > 0$, the random variables cX_1 and $X_1 + X_2$ are gamma distributed $\Gamma\,(a_1\,, b/c)$ and $\Gamma(a_1 + a_2\,, b)$ respectively.

Now, let $a(\cdot) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ be a continuous and non decreasing function such that $a(0) = 0$ and let $b > 0$. Also let $(X_t)_{t\geq 0}$ be a right-continuous stochastic process with left-side limits. Then, the stochastic process $(X_t)_{t\geq 0}$ is a non homogeneous gamma process with shape function $a(\cdot)$ and rate parameter b, as soon as

- $X_0 = 0$ almost surely (a.s.);
- $(X_t)_{t\geq 0}$ has independent increments;
- each increment is gamma distributed: for all $0 \leq s < t$, we have $X_t - X_s \sim \Gamma(a(t) - a(s), b)$,

(see, e.g., [1]). See [40] and its references for a large overview of the gamma processes. Also, an example of simulated trajectories of a gamma process is given in **Figure D.2**.

In order to mitigate the degradation of the system over time and extends its lifetime, preventive maintenance actions can be considered, in addition to corrective repairs which are performed at failure. In the context of deteriorating systems, many preventive maintenance policies from the literature consider condition-based maintenance (CBM) actions, where the preventive repair is triggered by the reaching of a preventive maintenance threshold by the deterioration level. In that context, "most of the existing CBM models have been limited to perfect maintenance actions", as noted by [3] (see also [45]). Some imperfect repair models are however emerging in the latest reliability literature, in this new context of deteriorating systems, see [3] for a recent review. Some models are based on the notion of virtual age previously introduced in the context of recurrent events (see, e.g., [19, 33]), where the system is rejuvenated by a

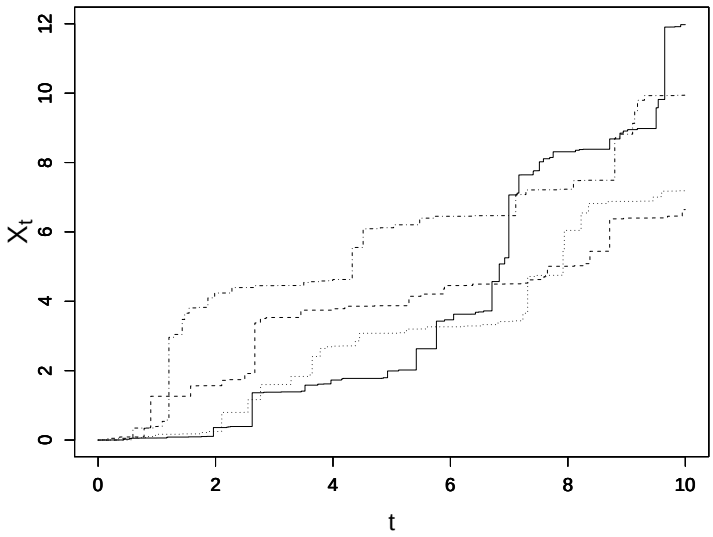

Figure 2: Simulated trajectories of a gamma process with shape function $a(t) = t$ and scale parameter $b = 1$.

maintenance action. Other models consider that an imperfect repair reduces the deterioration level of the system, such as [25, 28, 35, 37, 42], which can be accompanied by some increase in the deterioration rate, as in [15]. Also, some papers consider that the efficiency decreases with the number of repairs (see, e.g., [30, 46]), and further studies, as in [23], deal with imperfect maintenance models such that (i) repairs have a random efficiency (ii) the deterioration rate increases with the number of repairs. In all these papers however, the main point mostly is on the optimization of a maintenance policy, including these imperfect maintenance actions together with perfect repairs (replacements). Up to our knowledge, very few papers from the literature deal with statistical issues concerning imperfect repair models for deteriorating systems, except from [45], where the authors suggest a maximum likelihood method for estimating the parameters of the Wiener process (that describes the deterioration out of repairs) together with an iterative procedure based on a Kalman filter for the different factors implied in successive imperfect repairs.

In the context of systems subject to deterioration and imperfect repairs, the estimates for the parameters of the underlying degradation process and maintenance efficiency are of great use for maintenance policies optimization. Indeed, once the parameters have been estimated, the future behavior of the maintained system can be predicted, which allows to adapt (optimize) the periodicity of the maintenance actions and efficiently plan a general overhaul for instance. From a safety point of view, the principal inquiry is to ensure that the maintenance actions are effective enough to keep with a high probability the degradation level below a fixed threshold (safety level). As long as this safety level is not reached, the maintenance actions can be adjusted, either by adapting their periodicity or by improving their efficiency (if possible). Of course, apart from the previous safety concern, the maintenance costs are another issue. As an example, in [42], the costs minimization is based on the monitoring time and on the imperfect maintenance efficiency. In [23], the author considers a threshold for the degradation level, beyond which an imperfect maintenance action is performed. The optimization is made with respect to this threshold and the inspections periodicity. Finally, in [43], a maintenance policy is proposed, where a replacement

is performed either when the degradation exceeds a given threshold or when a fixed number of imperfect preventive maintenance actions are conducted. See, e.g., the three papers cited above and their reference for an overview on maintenance policies optimization.

This thesis focuses on the development and applications of estimation procedures for three specific imperfect repair models in the context of a gamma deteriorating system. The document is split into four parts, including the present introduction.

Part II deals with the Arithmetic Reduction of Degradation models of order one and infinity (ARD_1 and ARD_∞), where each maintenance action reduces the deterioration level of the system. The ARD_1 model was first introduced in [8] and further studied in [34]. Mimicking the Arithmetic Reduction of Intensity (ARI_1) model of order 1 developed by [16] in the context of recurrent events, the idea of this model is that a maintenance action removes a proportion ρ of the degradation accumulated by the system from the last maintenance action (where $\rho \in [0, 1)$). Based on the same idea, [16] also defined the ARI model of infinite order for recurrent events, that we extend here to the degradation framework in order to introduce the ARD_∞ model. Regarding this imperfect repair model, each maintenance action removes a proportion ρ of the current degradation, that is the degradation accumulated by the system from the initial time $t = 0$. Once these models defined, we place ourselves in a fully parametric framework and the observation scheme is stated. The Moments Estimation (ME) and Maximum Likelihood Estimation (MLE) methods are developed in **Chapter 2** for both models. To be more precise, the identifiability is studied and then we provide expressions for the parameters estimates. Numerical experiments based on simulated data are conducted in **Chapter 3**. In **Chapter 4**, we propose an original estimator for ρ, which does not depend on the underlying gamma process parameters, leading to a semiparametric framework. The idea of this estimator has come from a preliminary study in the framework based on the MLE method, where we observed that for one single trajectory, the minimum of the set of admissible ρ's has quite an interesting behavior when the shape function a of the underlying Gamma deterioration process is concave, getting quickly very close to the unknown efficiency parameter when the number of repairs increases. This semiparametric estimate was first developed in the context of the ARD_1 model, leading to a publication (see [38]) which is exactly reproduced here in **Sections 4.1** to **4.6**. In **Section 4.7** we propose an extension of this work to the ARD_∞ model, which is specific to this thesis.

In **Part III**, we consider an imperfect repair model based on the virtual age introduced by [34]: the Arithmetic Reduction of Age model of order one (ARA_1). Following the idea of [16] in the context of lifetime analysis, this model is such that a maintenance action removes a proportion ρ of the age of the system accumulated since the last maintenance action. Unlike for the ARD_1 model, here the system is rejuvenated, that is it is put back to the exact situation where it was some time before. The first steps of the work are similar to those of the previous part: the model as well as the observation scheme are defined, the ME method is studied from the point of view of the model parameters identifiability, and estimates expressions are provided. Then, the MLE method is developed. However, due to dependency issues, the likelihood function appears to be a product of integrals of large dimension, and thus numerical estimations becomes difficult to compute in a classical way. Hence, the MLE method requires approximating the integrals by the Monte Carlo and randomized Quasi Monte Carlo methods. Also, in order to avoid the computation of high-dimensional integrals, alternatives to the maximum likelihood method are developed: the composite maximum likelihood and the half data methods. The first one is based

on the composite marginal likelihood constructed under independence assumptions, also referred to in the literature as the independence likelihood (see, e.g., [9]). In the present work, this method consists in considering that the observations are independent and thus obtain an approximation of the likelihood function, which allows to develop the maximum composite likelihood estimation method. See [41] for an overview of composite likelihood methods. The half data method is based on sub-sample methods introduced by [21] and commonly used within bootstrap estimation (see, e.g., [4]). Here only one out of two observations are taken into account, which eliminates the dependence issues. Once again, the study of the estimation methods is based on the parameters identifiability, which was very challenging in one case, as well as the expressions of either the estimates or the log-likelihood. However, the identifiability is not treated regarding two out of the six methods we studied since the simulation studies reveal identifiability issues for these specific methods. Lastly in **Chapter 7**, some illustrations of the numerical performances of these methods are provided in two stages. A first study allows to eliminate methods with poor performance while a second study allows to select, among the remaining methods, the most appropriate one as a function of the observations characteristics.

Finally, **Part IV** highlights the conclusions of the thesis and future perspectives.

Part II

Imperfect repairs based on reduction of the degradation level

Chapter 1

Introduction

1.1 Preliminary

A system is considered whose intrinsic deterioration is modelled by a gamma process $(X_t)_{t\geq 0}$ with shape function $a(\cdot)$ and scale parameter b, as defined in the introduction. The system is subject to periodic (period T) and instantaneous imperfect maintenance actions, where each maintenance action removes $\rho\%$ of the deterioration accumulated from the last m maintenance actions, where $m \in N^*$ and ρ in $[0, 1)$ are fixed. This model is called Arithmetic Reduction of Degradation model of order m (ARD_m). The Euclidean parameter ρ is a measure of the efficiency of the imperfect maintenance action.

In the following, two models are considered: the ARD_1 model and the ARD_∞ model. The definitions of both models are first given in the two following sections. Two parametric estimation methods are developed in **Chapter 2**, and numerical experiments based on simulated data are provided in **Chapter 3**. Finally, a semiparametric estimate of the maintenance actions efficiency is developed in **Chapter 4**.

1.2 Arithmetic Reduction of Degradation model of order one

Regarding this model, a maintenance action reduces the degradation accumulated from the last maintenance action only. The repairs have an effect (or efficiency) defined by a parameter $\rho \in [\,0\,,1)$ and thus, at time jT, $j \in \mathbb{N}^*$, the maintenance effect results in a reduction of a proportion ρ of the deterioration accumulated over the time interval $[\,(j-1)T\,, jT\,)$.

Let $(Y_t)_{t\geq 0}$ stands for the random process describing the degradation level evolution of the maintained system according to the ARD_1 model. Let $(X^{(j)})_{j\in\mathbb{N}^*}$ be a sequence of independent copies of $(X_t)_{t\geq 0}$, where $X^{(j)}$ describes the intrinsic deterioration of the system over $((j-1)T, jT)$. The system is assumed to be in the perfect working order at the initial time, that is $Y_0 = X_0^{(1)} = 0$. Over the time interval $[0\,, T)$, the system deteriorates according to $X_t^{(1)}$. Therefore:

$$Y_t = X_t^{(1)} \text{ and } Y_T = (1-\rho)\left(X_T^{(1)} - X_0^{(1)}\right) = (1-\rho)X_T^{(1)}.$$

For $t \in [T\,, 2T)$, the system's degradation level is equal to the sum of the degradation accumulated between times t and T, that is $X_t^{(2)} - X_T^{(2)}$, and the degradation level right after the first maintenance

action, thus:

$$Y_t = Y_T + \left(X_t^{(2)} - X_T^{(2)}\right)$$

and

$$Y_{2T} = Y_{2T^-} + \rho\left(X_{2T^-}^{(2)} - X_T^{(2)}\right) = (1-\rho)\left(\left(X_{2T}^{(2)} - X_T^{(2)}\right) + \left(X_T^{(1)} - X_0^{(1)}\right)\right).$$

Hence, for any t in $[\,nT\,,(n+1)T\,)$, with n in $\mathbb{N}^*$, we have:

$$Y_{nT} = (1-\rho)\sum_{j=1}^{n}\left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right)$$

and

$$Y_t = Y_{nT} + \left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right).$$

Note that the random variables Y_{nT} and $\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right)$ are independent and gamma distributed, $\Gamma\left(a(nT)\,,b/(1-\rho)\right)$ and $\Gamma\left(a(t)-a(nT)\,,b\right)$, respectively.

This definition as well as the following property are already given in [34], and we recall it here for sake of completeness.

Proposition 1. *For each n in $\mathbb{N}$ and $nT \leq t < (n+1)T$, the expectation of the degradation level Y_t is given by $\mathbb{E}\left(Y_t\right) = (a(t) - \rho a(nT))/b$ and its variance by $\mathbb{V}\left(Y_t\right) = (a(t) - \rho\left(2-\rho\right)a(nT))/b^2$.*

Proof. Let us set n in $\mathbb{N}$ and $nT \leq t < (n+1)T$. The expression of Y_t is the following

$$Y_t = Y_{nT} + \left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right),$$

where Y_{nT} and $\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right)$ are independent and gamma distributed $\Gamma\left(a(nT)\,,b/(1-\rho)\right)$ and $\Gamma\left(a(t)-a(nT)\,,b\right)$ respectively. Hence, the expectation as well as the variance of Y_t can be expressed as the sum of both terms, and thus

$$\begin{cases} \mathbb{E}(Y_t) = \mathbb{E}(Y_{nT}) + \mathbb{E}\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right) \\ \mathbb{V}(Y_t) = \mathbb{V}(Y_{nT}) + \mathbb{V}\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right) \end{cases}$$

We now replace the expressions of the mean and variance of the gamma distribution, see **Part I**, which leads to

$$\begin{cases} \mathbb{E}(Y_t) = \dfrac{(1-\rho)a(nT) + a(t) - a(nT)}{b} \\ \mathbb{V}(Y_t) = \dfrac{(1-\rho)^2\,a(nT)}{b^2} + \dfrac{a(t)-a(nT)}{b^2} \end{cases}$$

and finishes the proof.

□

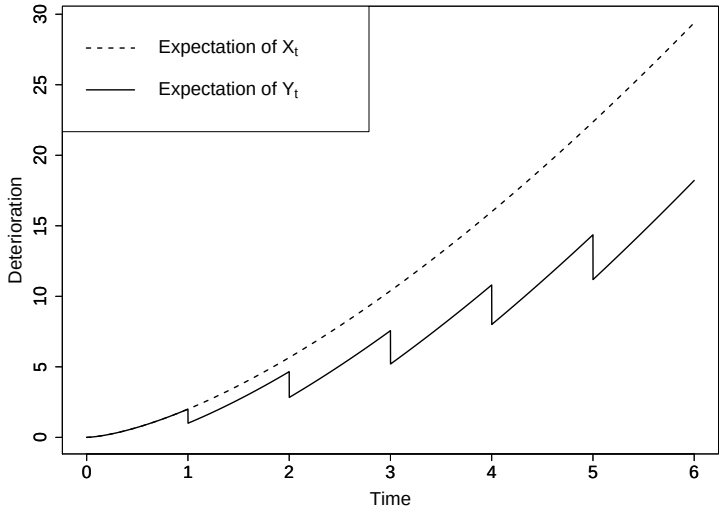

Figure 1.1: Expectations of X_t and Y_t, where the shape function is such that $a : t \mapsto \alpha t^{\beta}$ with $\alpha = 2$ and $\beta = 3/2$, and with $b = 1$ and $\rho = 1/2$.

Example 1. *Let us consider the shape function* $a(t) = \alpha t^{\beta}$*, with* $\alpha = 2$ *and* $\beta = 3/2$. ***Figure 1.1*** *represents the degradation evolution expected of the maintained system, as well as that of the intrinsic deterioration expected (unmaintained system).* ***Figure 1.2*** *represents the variances of the same quantities.*

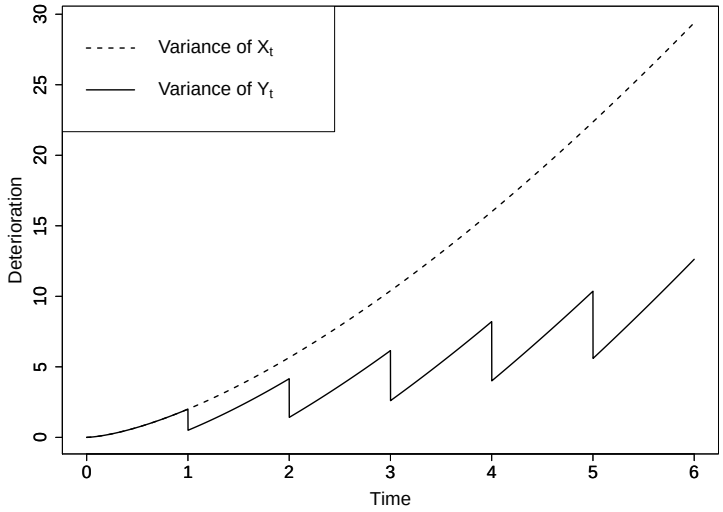

Figure 1.2: Variances of X_t and Y_t, where the shape function is such that $a : t \mapsto \alpha t^\beta$ with $\alpha = 2$ and $\beta = 3/2$, and with $b = 1$ and $\rho = 1/2$.

1.3 Arithmetic Reduction of Degradation model of order infinity

This model is very close to the previous one. The only difference is that a maintenance action affect the degradation accumulated from the initial time $t = 0$. Thus at times jT, for j in $\mathbb{N}^*$, the maintenance action results in a reduction of a proportion ρ of the current degradation level.

Let $(Y_t)_{t\geq 0}$ be the random process describing the degradation level evolution of a system according to the ARD_∞ model. Let $(X^{(j)})_{j\in\mathbb{N}^*}$ be a sequence of independent copies of $(X_t)_{t\geq 0}$. We consider once again that $Y_0 = X_0^{(1)} = 0$. Both ARD_1 and ARD_∞ models have identical behaviour over the time interval $[\,0\,,2T\,)$. Then, for all t in $[\,0\,,T\,)$ we have:

$$Y_t = X_t^{(1)} \text{ and } Y_T = (1-\rho)\left(X_T^{(1)} - X_0^{(1)}\right) = (1-\rho)X_T^{(1)},$$

and for all t in $[\,T\,,2T\,)$

$$Y_t = Y_T + \left(X_t^{(2)} - X_T^{(2)}\right).$$

When the second maintenance action happens, a proportion ρ of the degradation accumulated from the time $t = 0$ is removed, thus $Y_{2T} = (1-\rho)\left(Y_T + \left(X_{2T}^{(2)} - X_T^{(2)}\right)\right)$, which can be written as:

$$Y_{2T} = (1-\rho)^2\left(X_T^{(1)} - X_0^{(1)}\right) + (1-\rho)\left(X_{2T}^{(2)} - X_T^{(2)}\right).$$

From this, for any t in $[nT\,,(n+1)T\,)$ with n in $\mathbb{N}$, we easily derive the expression of Y_t, which is:

$$Y_t = \sum_{j=1}^{n}(1-\rho)^{n-j+1}\left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right) \;+\; \left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right),$$

and where all the increments are mutually independent and gamma distributed. More precisely, for all n

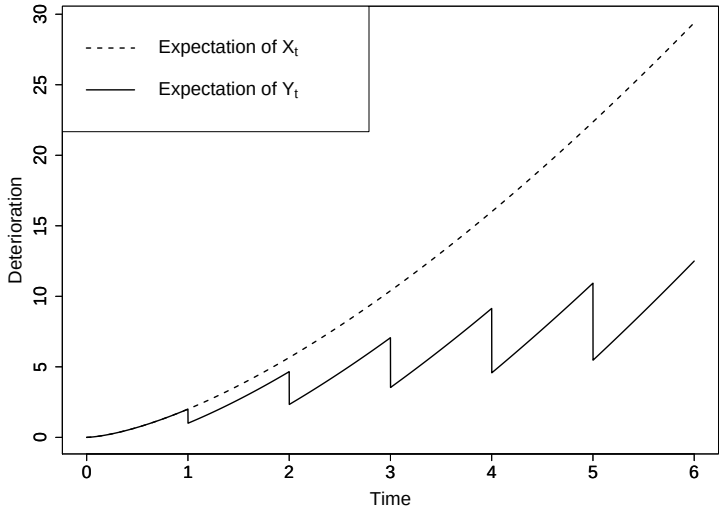

Figure 1.3: Expectations of X_t and Y_t, where the shape function is such that $a : t \mapsto \alpha t^{\beta}$ with $\alpha = 2$ and $\beta = 3/2$, and with $b = 1$ and $\rho = 1/2$.

in $\mathbb{N}^*$ and $1 \leq i \leq n$, we have:

- $\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right) \sim \Gamma\left(a(t) - a(nT), b\right)$;
- $(1-\rho)^{n-j+1}\left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right) \sim \Gamma\left(a(jT) - a((j-1)T), \frac{b}{(1-\rho)^{n-j+1}}\right)$.

Hence, the mean and variance of Y_t are given by

$$\mathbb{E}\left(Y_t\right) = \frac{1}{b}\left(\sum_{j=1}^{n}(1-\rho)^{n-j+1}\left(a(jT) - a((j-1)T)\right) \; + \; \left(a(t) - a(nT)\right)\right)$$

and

$$\mathbb{V}(Y_t) = \frac{1}{b^2}\left(\sum_{j=1}^{n}(1-\rho)^{2(n-j+1)}\left(a(jT) - a((j-1)T)\right) \; + \; \left(a(t) - a(nT)\right)\right)$$

for all $t \in [nT, (n+1)T)$ with $n \in \mathbb{N}$.

Based on the same framework as in **Example 1**, **Figure 1.3** represents the degradation evolution expected of the maintained system, as well as that of the intrinsic deterioration mean (unmaintained system). **Figure 1.4** represents the variances of the same quantities.

For sake of simplicity, from now on, we set $\Delta X^{(j)} = X_{jT}^{(j)} - X_{(j-1)T}^{(j)}$, which represents the increment of the intrinsic deterioration over $[(j-1)T, jT)$.

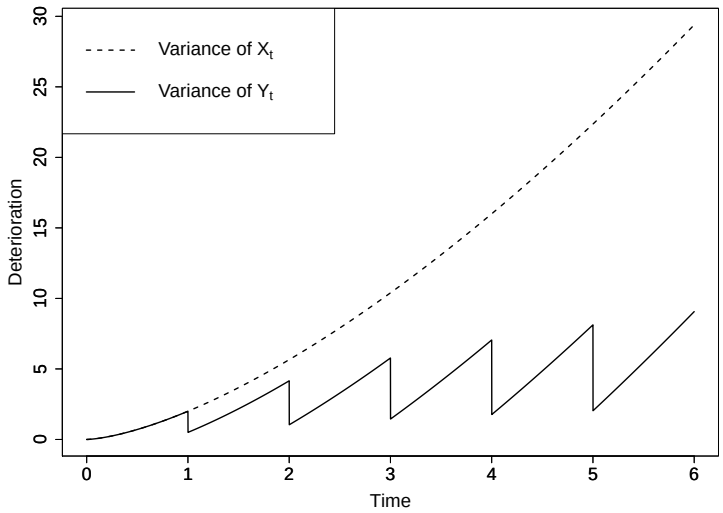

Figure 1.4: Variances of X_t and Y_t, where the shape function is such that $a : t \mapsto \alpha t^{\beta}$ with $\alpha = 2$ and $\beta = 3/2$, and with $b = 1$ and $\rho = 1/2$.

Chapter 2

Parametric inference for the Arithmetic Reduction of Degradation models

2.1 Preliminary

In this chapter, the Moments Estimation (ME) and the Maximum Likelihood Estimation (MLE) methods are developed in the framework of both ARD models defined above, in order to estimate the models parameters. The periodicity T is assumed to be known, thus the parameters of interest are the parameters of the shape function, the scale parameter b and the maintenance efficiency ρ. The estimation methods are then tested in the next chapter, within the framework of a power law shape function. To be more precise, the shape function is defined as $a : t \mapsto \alpha t^{\beta}$ with $\alpha,\ \beta > 0$. We set $\boldsymbol{\theta} = (\alpha, \beta, \rho, b) \in \Theta$ the parameter set, with $\Theta = (0,\infty)^2 \times [0,1), \times (0,\infty)$.

We assume that the degradation level is measured right before the first n maintenance actions for some n in $\mathbb{N}^*$, that is at times T^-, $2T^-$,...,nT^-.. Also, s i.i.d. systems are considered. For sake of readability, let us define the following notations:

- $Y_j = Y_{jT^-}$ for $1 \leq j \leq n$;
- $\boldsymbol{Y} = \left(Y_{jT^-}\right)_{1\leq j\leq n} = (Y_j)_{1\leq j\leq n}$;
- $y_j^{(i)}$ is the observed degradation level of the ith maintained system at times jT^- for $1 \leq j \leq n$ and $1 \leq i \leq s$, which is a realisation of the r.v. Y_j;
- $\boldsymbol{y^{(i)}} = \left(y_j^{(i)}\right)_{1\leq j\leq n}$ is the complete observations set related to the ith maintained system, it is a realisation of $\mathbf{Y}$;
- $\boldsymbol{y}$ is the complete observations set, that is $\boldsymbol{y} = \left(y_j^{(i)}\right)_{1\leq j\leq n,\, 1\leq i\leq s}$, it corresponds to s i.i.d. realisations of $\mathbf{Y}$.

2.2 Moments method estimation

2.2.1 Description of the method

In order to estimate the parameter $\boldsymbol{\theta}$ by the ME method, the distance between the empirical moments and the theoretical moments must be minimized. Here the centered moments are considered, except for

the first moment. Let us consider the distance function D which is defined as

$$D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = \sum_{k=1}^{d}\sum_{j=1}^{n}\left(m_k(\boldsymbol{\theta}\,,jT^-) - m_k(\boldsymbol{\theta_0}\,,jT^-)\right)^2$$

where $\boldsymbol{\theta_0}$ is the true value of the parameter, d the maximal order of the moments, n the number of observations by trajectory and $m_k(\boldsymbol{\theta}\,,jT^-)$ the kth moment (centered if $k>1$) at time jT^-, that is

$$m_1(\boldsymbol{\theta},jT^-) = \mathbb{E}\left(Y_{jT^-}\right) \quad \text{and} \quad \forall k\geq 2,\;\; m_k(\boldsymbol{\theta},jT^-) = \mathbb{E}\left(\left(Y_{jT^-} - \mathbb{E}\left(Y_{jT^-}\right)\right)^k\right).$$

Then, an estimation $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta_0}$ can be obtained through

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min}\,\hat{D}(\boldsymbol{\theta})$$

with $\hat{D}(\boldsymbol{\theta})$ the empirical version of $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0})$, where true unknown moments are replaced by their empirical counterparts. This empirical version of D has the following expression

$$\hat{D}(\boldsymbol{\theta}) = \sum_{k=1}^{d}\sum_{j=1}^{n}\left(m_k(\boldsymbol{\theta}\,,jT^-) - \hat{m}_k(jT^-)\right)^2 \tag{2.1}$$

with $\hat{m}_k(jT^-)$ the kth empirical moment (centered if $k>1$) at time jT^-. These empirical moments are empirical estimates of the moments $m_k(\boldsymbol{\theta_0}\,,jT^-)$ and are defined by

$$\hat{m}_1(jT^-) = \frac{1}{s}\sum_{i=1}^{s} y_j^{(i)} \quad \text{and} \quad \forall k\geq 2,\;\; \hat{m}_k(jT^-) = \frac{1}{s}\sum_{i=1}^{s}\left(y_j^{(i)} - \hat{m}_1(jT^-)\right)^k.$$

This method is based on the classical approach of the Generalized Method of Moments as exposed in [31]. The two following sections are devoted to the application of the ME method as defined above, in the case of the ARD_1 and ARD_∞ models, beginning by the model of order 1.

2.2.2 Application to the Arithmetic Reduction of Degradation model of order one

Let us recall that the shape function is defined by $a(t) = \alpha t^\beta$ with $\alpha,\ \beta > 0$, and thus the estimation focuses on four parameters $\boldsymbol{\theta} = (\alpha\,,\beta\,,\rho,b) \in \Theta$ with $\Theta = (0\,,+\infty)^2 \times [0\,,1) \times (0\,,+\infty)$. In order to apply the ME method, we first need to see how many moments d and how many data n are necessary to identify the model parameters, which is done in next proposition.

Proposition 2. *The parameters of the ARD_1 model are identifiable from the ME method, that is $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0$ implies that $\boldsymbol{\theta} = \boldsymbol{\theta_0}$ for all $\boldsymbol{\theta},\boldsymbol{\theta_0} \in \Theta$, as soon as $d \geq 2$ and $n \geq 3$. In other words, the identifiability holds from the ME method as soon as the systems are observed at times T^-, $2T^-$ and $3T^-$, and if at least the first two moments (expectation and variance) are used.*

Proof. Here we prove that the assertion $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0 \;\Rightarrow\; \boldsymbol{\theta} = \boldsymbol{\theta_0}$ is true $\forall\,\boldsymbol{\theta},\boldsymbol{\theta_0} \in \Theta$ as soon as $n \geq 3$ and $d \geq 2$. Assume that $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0$, and let $\boldsymbol{\theta} = (\alpha\,,\beta\,,b\,,\rho,b)$ and $\boldsymbol{\theta_0} = (\alpha_0\,,\beta_0\,,\rho_0\,,b_0)$ be in Θ. Based on the definition of the function D, we have for all j in $\{1\,,\ldots\,,n\}$:

$$\begin{cases} m_1(\boldsymbol{\theta}\,, jT^-) = m_1(\boldsymbol{\theta_0}\,, jT^-) \\ m_2(\boldsymbol{\theta}\,, jT^-) = m_2(\boldsymbol{\theta_0}\,, jT^-) \end{cases}$$

When $j = 1$, the first two moments are those of the random variable $X_T^{(1)}$, hence this system is equivalent to

$$\begin{cases} \frac{\alpha}{b} T^\beta = \frac{\alpha_0}{b_0} T^{\beta_0} \\ \frac{\alpha}{b^2} T^\beta = \frac{\alpha_0}{b_0^2} T^{\beta_0} \end{cases}$$

thus

$$\begin{cases} b = b_0 \\ \alpha T^\beta = \alpha_0 T^{\beta_0} \end{cases}$$

Note that only the parameter b is identified. Taking into account the second observation time, that is when $j = 1, 2$ we can write

$$\begin{cases} \frac{\alpha}{b} T^\beta = \frac{\alpha_0}{b_0} T^{\beta_0} \\ \frac{\alpha}{b^2} T^\beta = \frac{\alpha_0}{b_0^2} T^{\beta_0} \\ \frac{\alpha}{b} (2T)^\beta - \frac{\alpha}{b} \rho T^\beta = \frac{\alpha_0}{b_0} (2T)^{\beta_0} - \frac{\alpha_0}{b_0} \rho_0 T^{\beta_0} \\ \frac{\alpha}{b^2} (2T)^\beta - \frac{\alpha}{b^2} \rho(2-\rho) T^\beta = \frac{\alpha_0}{b_0^2} (2T)^{\beta_0} - \frac{\alpha_0}{b_0^2} \rho_0(2-\rho_0) T^{\beta_0} \end{cases}$$

and thus

$$\begin{cases} \frac{\alpha}{b} T^\beta = \frac{\alpha_0}{b_0} T^{\beta_0} \\ \frac{\alpha}{b^2} T^\beta = \frac{\alpha_0}{b_0^2} T^{\beta_0} \\ \frac{\alpha}{b} T^\beta \left(2^\beta - \rho\right) = \frac{\alpha_0}{b_0} T^{\beta_0} \left(2^{\beta_0} - \rho_0\right) \\ \frac{\alpha}{b^2} T^\beta \left(2^\beta - \rho(2-\rho)\right) = \frac{\alpha_0}{b_0^2} T^{\beta_0} \left(2^{\beta_0} - \rho_0(2-\rho_0)\right) \end{cases}$$

which can be reduced to

$$\begin{cases} b = b_0 \\ \alpha T^\beta = \alpha_0 T^{\beta_0} \\ 2^\beta - \rho = 2^{\beta_0} - \rho_0 \\ 2^\beta - \rho(2-\rho) = 2^{\beta_0} - \rho_0(2-\rho_0) \end{cases}$$

Now let us first assume that $\rho_0 \neq 1 - \rho$. The third and fourth equations lead to $\rho(1-\rho) = \rho_0(1-\rho_0)$, hence we have $\rho = \rho_0$. This implies that $2^\beta = 2^{\beta_0}$ and thus $\beta = \beta_0$ from the third equation. Finally, the second equation provides $\alpha = \alpha_0$.

Now let us assume that $\rho_0 = 1 - \rho$. In that case, we need to add the third observation, leading to the following expressions for the expectation and variance when $j = 1, 3$:

$$\begin{cases} \frac{\alpha}{b} T^\beta = \frac{\alpha_0}{b_0} T^{\beta_0} \\ \frac{\alpha}{b^2} T^\beta = \frac{\alpha_0}{b_0^2} T^{\beta_0} \\ \frac{\alpha}{b} T^\beta \left(3^\beta - 2^\beta \rho\right) = \frac{\alpha_0}{b_0} T^{\beta_0} \left(3^{\beta_0} - 2^{\beta_0} \rho_0\right) \\ \frac{\alpha}{b^2} T^\beta \left(3^\beta - 2^\beta \rho(2-\rho)\right) = \frac{\alpha_0}{b_0^2} T^{\beta_0} \left(3^{\beta_0} - 2^{\beta_0} \rho_0(2-\rho_0)\right) \end{cases}$$

This system can be written as

$$\begin{cases} b = b_0 \\ \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ 3^{\beta} - 2^{\beta}\rho = 3^{\beta_0} - 2^{\beta_0}\rho_0 \\ 3^{\beta} - 2^{\beta}\rho(2-\rho) = 3^{\beta_0} - 2^{\beta_0}\rho_0(2-\rho_0) \end{cases}$$

and because $\rho_0 = 1 - \rho$ by assumption, we have

$$\begin{cases} b = b_0 \\ \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ 3^{\beta} - 2^{\beta}\rho = 3^{\beta_0} - 2^{\beta_0} + 2^{\beta_0}\rho \\ 3^{\beta} - 2^{\beta}\rho(2-\rho) = 3^{\beta_0} - 2^{\beta_0} + 2^{\beta_0}\rho^2 \end{cases}$$

The third and fourth equations lead to $2^{\beta}(\rho - \rho^2) = 2^{\beta_0}(\rho - \rho^2)$. If $\rho = 0$ then $\rho_0 = 1 - \rho = 1$ which is impossible since $\rho_0 \in [0, 1)$. Hence $\rho \in (0, 1)$, which implies that $2^{\beta} = 2^{\beta_0}$ and thus $\beta = \beta_0$, and finally $\alpha = \alpha_0$ from the second equation. As a result, from the third equation we deduce that $\rho = \rho_0$. Note that because $\rho_0 = 1 - \rho$, necessarily $\rho = \rho_0 = 0.5$.

Hence, only the first three observations and the first two moments are necessary in order to obtain the identifiability of the model parameters.

□

In the following, we consider $d = 2$, thus Equation (2.1) becomes

$$\hat{D}(\boldsymbol{\theta}) = \sum_{j=1}^{n} \left(\left(m_1(\boldsymbol{\theta}, jT^-) - \hat{m}_1(jT^-) \right)^2 + \left(m_2(\boldsymbol{\theta}, jT^-) - \hat{m}_2(jT^-) \right)^2 \right)$$

which can be written as

$$\hat{D}(\boldsymbol{\theta}) = \sum_{j=1}^{n} \left(\left(\mathbb{E}\left(Y_{jT^-}\right) - \hat{m}_1(jT^-) \right)^2 + \left(\mathbb{V}\left(Y_{jT^-}\right) - \hat{m}_2(jT^-) \right)^2 \right).$$

From now on, another parametrization of the empirical distance function $\hat{D}(\boldsymbol{\theta})$ is considered in order to simplify our calculations. Let us set $\boldsymbol{\theta'} = (\mu, \eta, b, \rho)$, with $\mu = \alpha/b$ and $\eta = \alpha/b^2$. Thus we have for all j in $\{1, \ldots, n\}$,

$$\mathbb{E}(Y_{jT^-}) = \mu T^{\beta} \left(j^{\beta} - \rho\,(j-1)^{\beta} \right) \quad \text{and} \quad \mathbb{V}(Y_{jT^-}) = \eta T^{\beta} \left(j^{\beta} - \rho(2-\rho)\,(j-1)^{\beta} \right),$$

which implies

$$\hat{D}(\boldsymbol{\theta'}) = \mu^2 g_1(\beta, \rho) - 2\mu\, g_2(\beta, \rho) + \eta^2 h_1(\beta, \rho) - 2\eta\, h_2(\beta, \rho) + C \tag{2.2}$$

where the functions g_1, g_2, h_1 and h_2 are independent of μ and η, and C is a constant independent of $\boldsymbol{\theta'}$. To be more precise, the quantity C is given by

$$C = \sum_{j=1}^{n} \left(\hat{m}_1(jT^-)^2 + \hat{m}_2(jT^-)^2 \right)$$

and the functions g_1, g_2, h_1 and h_2 have the following expressions

$$g_1(\beta\,,\rho) = T^{2\beta}\left(\sum_{j=1}^{n} j^{2\beta} - 2\rho\sum_{j=1}^{n}\left(j\,(j-1)\right)^{\beta} + \rho^2\sum_{j=1}^{n}(j-1)^{2\beta}\right)$$

$$g_2(\beta\,,\rho) = T^{\beta}\left(\sum_{j=1}^{n} j^{\beta}\hat{m}_1(jT^-) - \rho\sum_{j=1}^{n}(j-1)^{\beta}\hat{m}_1(jT^-)\right)$$

$$h_1(\beta\,,\rho) = T^{2\beta}\left(\sum_{j=1}^{n} j^{2\beta} - 2\rho(2-\rho)\sum_{j=1}^{n}\left(j\,(j-1)\right)^{\beta} + \rho^2(2-\rho)^2\sum_{j=1}^{n}(j-1)^{2\beta}\right)$$

$$h_2(\beta\,,\rho) = T^{\beta}\left(\sum_{j=1}^{n} j^{\beta}\hat{m}_2(jT^-) - \rho(2-\rho)\sum_{j=1}^{n}(j-1)^{\beta}\hat{m}_2(jT^-)\right)$$

We now look for critical points of $\hat{D}(\boldsymbol{\theta'})$, by looking for the zeros of the partial derivatives of $\hat{D}$ with respect to $\boldsymbol{\theta}$. However, due to the expression of $\hat{D}(\boldsymbol{\theta'})$ given by Equation (2.2), we only write down $\partial_\mu\hat{D}(\boldsymbol{\theta'})$ and $\partial_\eta\hat{D}(\boldsymbol{\theta'})$ as the other partial derivatives are of no use, because of their complexity. This leads to:

$$\begin{cases} \partial_\mu\hat{D}(\boldsymbol{\theta'}) = 0 \\ \partial_\eta\hat{D}(\boldsymbol{\theta'}) = 0 \end{cases}$$

which provides

$$\begin{cases} 2\mu\, g_1(\beta\,,\rho) - 2g_2(\beta\,,\rho) = 0 \\ 2\eta\, h_1(\beta\,,\rho) - 2h_2(\beta\,,\rho) = 0 \end{cases}$$

Hence we can express the parameters μ and η with respect to β and ρ as follows: $\mu = h_\mu(\beta,\rho) = \dfrac{g_2(\beta\,,\rho)}{g_1(\beta\,,\rho)}$ and $\eta = h_\eta(\beta,\rho) = \dfrac{h_2(\beta\,,\rho)}{h_1(\beta\,,\rho)}$. Based on this, Equation (2.2) can be now written as

$$\tilde{D}(\beta,\rho) \equiv \hat{D}\left(h_\mu(\beta,\rho), h_\eta(\beta,\rho), \beta, \rho\right) = C - \left(\frac{g_2^2(\beta\,,\rho)}{g_1(\beta\,,\rho)} + \frac{h_2^2(\beta\,,\rho)}{h_1(\beta\,,\rho)}\right).$$

Then, the function $\tilde{D}(\beta,\rho)$ is numerically minimized to obtain an estimation of β and ρ. Thereafter we estimate μ and η by plugging, setting $\hat{\mu} = h_\mu(\hat{\beta}\,,\hat{\rho})$ and $\hat{\eta} = h_\eta(\hat{\beta}\,,\hat{\rho})$. From this we derive expressions of $\hat{\alpha}$ and $\hat{b}$, and finally we obtain the following estimate

$$\hat{\boldsymbol{\theta}} = \left(\frac{g_2(\hat{\beta}\,,\hat{\rho})h_1(\hat{\beta}\,,\hat{\rho})}{g_1(\hat{\beta}\,,\hat{\rho})h_2(\hat{\beta}\,,\hat{\rho})}\,,\hat{\beta}\,,\hat{\rho},\frac{g_2^2(\hat{\beta}\,,\hat{\rho})h_1(\hat{\beta}\,,\hat{\rho})}{g_1^2(\hat{\beta}\,,\hat{\rho})h_2(\hat{\beta}\,,\hat{\rho})}\right). \tag{2.3}$$

where

$$\left(\hat{\beta}\,,\hat{\rho}\right) = \underset{(\beta,\rho)\in(0,\infty)\times[0,1)}{\arg\min}\ \tilde{D}(\beta,\rho)$$

We now look at the ME method in the case of an ARD_∞ model.

2.2.3 Application to the Arithmetic Reduction of Degradation model of order infinity

This case is similar to the previous one, thus less details are given in this section, beginning with the following proposition about the parameters identifiability.

Proposition 3. *The parameters of the ARD_1 model are identifiable from the ME method, that is $D(\boldsymbol{\theta}, \boldsymbol{\theta_0}) = 0$ implies that $\boldsymbol{\theta} = \boldsymbol{\theta_0}$ for all $\boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$, as soon as $d \geq 2$ and $n \geq 3$.*

Proof. Here we prove that the assertion $D(\boldsymbol{\theta}, \boldsymbol{\theta_0}) = 0 \Rightarrow \boldsymbol{\theta} = \boldsymbol{\theta_0}$ is true $\forall\, \boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$ as soon as $n \geq 3$ and $d \geq 2$. Assume that $D(\boldsymbol{\theta}, \boldsymbol{\theta_0}) = 0$, and let $\boldsymbol{\theta} = (\alpha, \beta, b, \rho, b)$ and $\boldsymbol{\theta_0} = (\alpha_0, \beta_0, \rho_0, b_0)$ be in Θ. So given the function D, we have for all j in $\{1, \dots, n\}$:

$$\begin{cases} m_1(\boldsymbol{\theta}, jT^-) = m_1(\boldsymbol{\theta_0}, jT^-) \\ m_2(\boldsymbol{\theta}, jT^-) = m_2(\boldsymbol{\theta_0}, jT^-) \end{cases}$$

Note that the mean and expectation of the degradation level Y_t over $[0, 2T)$ is equivalent for both ARD_1 and ARD_∞ model. Hence, most of the present proof is similar to this of **Proposition 2**, hence it can be concluded that the identifiability holds if $n = 2$ and $\rho \neq 1 - \rho_0$.

Now assume that $\rho_0 = 1 - \rho$ and let us add the third observation ($j = 3$). Let us recall that for $j = 1$, we have $b = b_0$ and $\alpha T^\beta = \alpha_0 T^{\beta_0}$. Hence, we derive the following equations when $j = 1, 2, 3$:

$$\begin{cases} b = b_0 \\ \alpha T^\beta = \alpha_0 T^{\beta_0} \\ 2^\beta - \rho = 2^{\beta_0} - \rho_0 \\ 2^\beta - \rho(2 - \rho) = 2^{\beta_0} - \rho_0(2 - \rho_0) \\ 3^\beta - 2^\beta + (1 - \rho)(2^\beta - 1) + (1 - \rho)^2 = 3^{\beta_0} - 2^{\beta_0} + (1 - \rho_0)(2^{\beta_0} - 1) + (1 - \rho_0)^2 \\ 3^\beta - 2^\beta + (1 - \rho)^2(2^\beta - 1) + (1 - \rho)^4 = 3^{\beta_0} - 2^{\beta_0} + (1 - \rho_0)^2(2^{\beta_0} - 1) + (1 - \rho_0)^4 \end{cases}$$

We have $\rho_0 = 1 - \rho$ by assumption, thus

$$\begin{cases} b = b_0 \\ \alpha T^\beta = \alpha_0 T^{\beta_0} \\ 2^\beta - \rho = 2^{\beta_0} - (1 - \rho) \\ 2^\beta - \rho(2 - \rho) = 2^{\beta_0} - (1 - \rho)(1 + \rho) \\ 3^\beta - 2^\beta + (1 - \rho)(2^\beta - 1) + (1 - \rho)^2 = 3^{\beta_0} - 2^{\beta_0} + \rho(2^{\beta_0} - 1) + \rho^2 \\ 3^\beta - 2^\beta + (1 - \rho)^2(2^\beta - 1) + (1 - \rho)^4 = 3^{\beta_0} - 2^{\beta_0} + \rho^2(2^{\beta_0} - 1) + \rho^4 \end{cases}$$

The third and fourth equations are identical and provides

$$2\rho - 1 = 2^\beta - 2^{\beta_0} \tag{2.4}$$

while the last two ones leads to

$$\rho(1 - \rho)\left(\rho^2 - 3\rho + 2^\beta + 1\right) = \rho(1 - \rho)\left(\rho^2 + \rho + 2^{\beta_0} - 1\right).$$

Note that if $\rho = 0$, then $\rho_0 = 1$ which is impossible since ρ_0 belongs to $[0, 1)$. Hence $\rho \in (0, 1)$, and the

equation above is equivalent to

$$4\rho - 2 = 2^{\beta} - 2^{\beta_0}. \tag{2.5}$$

Finally, from Equations (2.4) and (2.5), we derive that $4\rho - 2 = 2\rho - 1$. Therefore $\rho = 1/2$ and $\rho_0 = 1 - \rho = 1/2$, and finally $\rho = \rho_0$. This leads to $\beta = \beta_0$ and then $\alpha = \alpha_0$, which finishes the proof.

□

The function $\hat{D}$ (see Equation (2.1)) is also considered here, and an estimation of $\boldsymbol{\theta}$ will be obtained through the minimization of this function. The same parametrisation $\boldsymbol{\theta'} = (\mu\,, \eta\,, b\,, \rho)$ is used, with $\mu = \alpha/b$ and $\eta = \alpha/b^2$, and for j in $\{1\,, \dots\,, n\}$, we have:

$$\mathbb{E}(Y_{jT^-}) = \mu g_1^{(j)}(\beta\,, \rho) \quad \text{and} \quad \mathbb{V}(Y_{jT^-}) = \eta g_2^{(j)}(\beta\,, \rho),$$

with

$$\begin{aligned} g_1^{(j)}(\beta\,, \rho) &= T^{\beta}\left(j^{\beta} - (j-1)^{\beta} + \sum_{i=1}^{j-1}(1-\rho)^{j-i}\left(i^{\beta} - (i-1)^{\beta}\right)\right) \\ g_2^{(j)}(\beta\,, \rho) &= T^{\beta}\left(j^{\beta} - (j-1)^{\beta} + \sum_{i=1}^{j-1}(1-\rho)^{2(j-i)}\left(i^{\beta} - (i-1)^{\beta}\right)\right). \end{aligned}$$

Thus the function $\hat{D}(\boldsymbol{\theta'})$ can be written as

$$\hat{D}(\boldsymbol{\theta'}) = \sum_{j=1}^{n}\left(\mu g_1^{(j)}(\beta\,, \rho) - \hat{m}_1(jT^-)\right)^2 + \sum_{j=1}^{n}\left(\eta g_2^{(j)}(\beta\,, \rho) - \hat{m}_2(jT^-)\right)^2. \tag{2.6}$$

Once again we look for the zeros of the partial derivatives of $\hat{D}$ with respect to μ and η. Here we have

$$\begin{cases} \partial_{\mu}\hat{D}(\boldsymbol{\theta'}) = 0 \\ \partial_{\eta}\hat{D}(\boldsymbol{\theta'}) = 0 \end{cases}$$

which is equivalent to

$$\begin{cases} \mu\sum_{j=1}^{n}\left(g_1^{(j)}(\beta\,, \rho)\right)^2 - \sum_{j=1}^{n} g_1^{(j)}(\beta\,, \rho)\hat{m}_1(jT^-) = 0 \\ \eta\sum_{j=1}^{n}\left(g_2^{(j)}(\beta\,, \rho)\right)^2 - \sum_{j=1}^{n} g_2^{(j)}(\beta\,, \rho)\hat{m}_2(jT^-) = 0 \end{cases}$$

Based on the above equations we can express μ and η as functions of β and ρ that cancel the partial derivatives of $\hat{D}$ with respect to μ and η, leading to

$$\mu = h_{\mu}(\beta\,, \rho) = \frac{\sum_{j=1}^{n} g_1^{(j)}(\beta\,, \rho)\hat{m}_1(jT^-)}{\sum_{j=1}^{n}\left(g_1^{(j)}(\beta\,, \rho)\right)^2}$$

and

$$\eta = h_\eta(\beta\,,\rho) = \frac{\displaystyle\sum_{j=1}^{n} g_2^{(j)}(\beta\,,\rho)\hat{m}_2(jT^-)}{\displaystyle\sum_{j=1}^{n}\left(g_2^{(j)}(\beta\,,\rho)\right)^2}.$$

Then, replacing μ and η by $h_\mu(\beta\,,\rho)$ and $h_\eta(\beta\,,\rho)$ in Equation (2.6), we obtain a function $\tilde{D}$ which only depends on β and ρ, and is given by

$$\tilde{D}(\beta,\rho) \equiv \hat{D}\left(\mu(\beta,\rho),\eta(\beta,\rho),\beta,\rho\right) = C - \left(\frac{\left(\displaystyle\sum_{j=1}^{n} g_1^{(j)}(\beta\,,\rho)\hat{m}_1(jT^-)\right)^2}{\displaystyle\sum_{j=1}^{n}\left(g_1^{(j)}(\beta\,,\rho)\right)^2} + \frac{\left(\displaystyle\sum_{j=1}^{n} g_2^{(j)}(\beta\,,\rho)\hat{m}_2(jT^-)\right)^2}{\displaystyle\sum_{j=1}^{n}\left(g_2^{(j)}(\beta\,,\rho)\right)^2} \right)$$

where C is the same constant as defined in the previous section. Finally, the minimization of this function with respect to β and ρ provides estimations of β and ρ, and thus of the entire parameter set $\boldsymbol{\theta'}$. Then, from the relationship between $\boldsymbol{\theta'}$ and $\boldsymbol{\theta}$ we derive an estimate of $\boldsymbol{\theta}$, which is given by

$$\hat{\boldsymbol{\theta}} = \left(\frac{\displaystyle\sum_{j=1}^{n} g_1^{(j)}(\hat{\beta}\,,\hat{\rho})\hat{m}_1(jT^-)\ \sum_{j=1}^{n}\left(g_2^{(j)}(\hat{\beta}\,,\hat{\rho})\right)^2}{\displaystyle\sum_{j=1}^{n}\left(g_1^{(j)}(\hat{\beta}\,,\hat{\rho})\right)^2\ \sum_{j=1}^{n} g_2^{(j)}(\hat{\beta}\,,\hat{\rho})\hat{m}_2(jT^-)}, \hat{\beta}\,,\hat{\rho}, \frac{\left(\displaystyle\sum_{j=1}^{n} g_1^{(j)}(\hat{\beta}\,,\hat{\rho})\hat{m}_1(jT^-)\right)^2\ \displaystyle\sum_{j=1}^{n}\left(g_2^{(j)}(\hat{\beta}\,,\hat{\rho})\right)^2}{\left(\displaystyle\sum_{j=1}^{n}\left(g_1^{(j)}(\hat{\beta}\,,\hat{\rho})\right)^2\right)^2\ \displaystyle\sum_{j=1}^{n} g_2^{(j)}(\hat{\beta}\,,\hat{\rho})\hat{m}_2(jT^-)} \right).$$

2.3 Maximum likelihood estimation

2.3.1 Application to the Arithmetic Reduction of Degradation model of order one

In this section we deal with the MLE method. We begin with the case where one single system is observed ($s = 1$) and we start by computing the joint p.d.f. of the random vector $\boldsymbol{Y}$, from where the likelihood function is easily expressed for $s > 1$. With that aim, we need the following technical result.

Proposition 4. *The degradation level Y_{jT} right after the maintenance at time jT can be expressed with respect to the random variables $Y_1, \ldots, Y_j$ and the maintenance efficiency parameter ρ as follows:*

$$Y_{jT} = (1-\rho)\sum_{k=1}^{j} \rho^{j-k}\, Y_k, \text{ for all } 1 \le j \le n.$$

Proof. The result is proved by induction on j. When $j = 1$, the result is verified as $(1-\rho)\sum_{k=1}^{j}\rho^{j-k}\,Y_k = (1-\rho)Y_1 = Y_T$. Now assume that

$$Y_{jT} = (1-\rho)\sum_{k=1}^{j}\rho^{j-k}\,Y_k, \text{ for some } 1 \leq j \leq n,$$

we want to prove that

$$Y_{(j+1)T} = (1-\rho)\sum_{k=1}^{j+1}\rho^{j-k+1}\,Y_k,$$

or equivalently that

$$Y_{(j+1)T} = (1-\rho)Y_{j+1} + (1-\rho)\sum_{k=1}^{j}\rho^{j-k+1}\,Y_k. \tag{2.7}$$

Recall that $Y_{(j+1)T} = Y_{jT} + (1-\rho)\Delta X^{(j+1)}$, which can also be expressed as

$$Y_{(j+1)T} = Y_{jT} + (1-\rho)\left(Y_{j+1} - Y_{jT}\right).$$

This leads to

$$Y_{(j+1)T} = (1-\rho)Y_{j+1} + \rho Y_{jT}.$$

By assumption on Y_{jT}, we have

$$Y_{(j+1)T} = (1-\rho)Y_{j+1} + \rho(1-\rho)\sum_{k=1}^{j}\rho^{j-k}\,Y_k$$

which can be written as Equation (2.7) and thus finishes the proof.

□

Lemma 1. *Let $\boldsymbol{y} = (y_1, \ldots, y_n)$ be a realisation of $\boldsymbol{Y}$. The p.d.f. $f_{\boldsymbol{Y}}$ of $\boldsymbol{Y}$ is given by*

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \prod_{j=1}^{n} f_{\Delta X^{(j)}}\left(g_j(\rho, \boldsymbol{y})\right),$$

where $f_{\Delta X^{(j)}}$ is the p.d.f. of the random variable $\Delta X^{(j)}$, and

$$\begin{cases} g_j(\rho, \boldsymbol{y}) = y_j - (1-\rho)\sum_{k=1}^{j-1}\rho^{j-k-1}y_k \\ g_1(\rho, \boldsymbol{y}) = y_1 \end{cases}$$

Proof. The probability multiplication rule provides

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = f_{Y_1}(y_1) \prod_{j=2}^{n} f_{Y_j|(Y_1,\dots,Y_{j-1})}\left(y_j|\left(y_1,\dots,y_{j-1}\right)\right).$$

Then, we have by definition Y_j

$$f_{Y_j|(Y_1,\dots,Y_{j-1})}\left(y_j|\left(y_1,\dots,y_{j-1}\right)\right) = f_{Y_{(j-1)T}+\Delta X^{(j)}|(Y_1,\dots,Y_{j-1})}\left(y_j|\left(y_1,\dots,y_{j-1}\right)\right).$$

Let us now recall that for all $2 \leq j \leq n$ we have $\Delta X^{(j)} \perp (Y_1,\dots,Y_{j-1})$ and

$$Y_{(j-1)T} = (1-\rho)\sum_{k=1}^{j-1} \rho^{j-k-1} Y_k$$

from **Proposition 4**. We derive that,

$$f_{Y_{(j-1)T}+\Delta X^{(j)}|(Y_1,\dots,Y_{j-1})}\left(y_j|\left(y_1,\dots,y_{j-1}\right)\right) = f_{\Delta X^{(j)}}\left(g_j(\rho,\boldsymbol{y})\right).$$

This finishes the proof as $f_{\Delta X^{(1)}}\left(g_1(\rho,\boldsymbol{y})\right) = f_{Y_1}\left(y_1\right)$ by definition of Y_1 and g_1.

□

We now consider s independent and identical systems. The previous lemma allows us to write down the expression of the likelihood function as follows:

$$L\left(\boldsymbol{\theta}|\boldsymbol{y}\right) = \prod_{i=1}^{s} f_{\boldsymbol{Y}}(\boldsymbol{y^{(i)}}) \tag{2.8}$$

which is equivalent to

$$L\left(\boldsymbol{\theta}|\boldsymbol{y}\right) = \prod_{i=1}^{s}\prod_{j=1}^{n} f_{\Delta X^{(j)}}\left(g_j^{(i)}(\rho)\right) \tag{2.9}$$

with

$$\begin{cases} g_j^{(i)}(\rho) = y_j^{(i)} - (1-\rho)\sum_{k=1}^{j} \rho^{j-k-1} y_k^{(i)} \\ g_1^{(i)}(\rho) = y_1^{(i)} \end{cases}$$

Let us recall that for $1 \leq j \leq n$, the random variables $\Delta X^{(j)}$ are gamma distributed with shape and scale parameters $\Delta A^{(j)}$ and b, respectively. As a result, Equation (2.9) becomes

$$L\left(\boldsymbol{\theta}|\boldsymbol{y}\right) = \prod_{i=1}^{s}\prod_{j=1}^{n} \frac{b^{\Delta A^{(j)}}}{\Gamma\left(\Delta A^{(j)}\right)} \left(g_j^{(i)}(\rho)\right)^{\Delta A^{(j)}-1} e^{-b g_j^{(i)}(\rho)} \tag{2.10}$$

where $\Delta A^{(j)} = \alpha T^{\beta}\left(j^{\beta} - (j-1)^{\beta}\right)$.

For similar reasons as for parameters identifiability when applying the ME method, we now study the identifiability in the case of the MLE method.

Proposition 5. *The parameters of the ARD_1 model are identifiable from the likelihood function, that is $L(\boldsymbol{\theta}|\boldsymbol{y}) = L(\boldsymbol{\theta_0}|\boldsymbol{y})$ for all $\boldsymbol{y} = (y_1, \dots, y_n)$ implies that $\boldsymbol{\theta_0} = \boldsymbol{\theta}$ and $\boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$, as soon as $n \geq 2$. In other words, the identifiability holds from the likelihood function as soon as observations are conducted twice, at times T^- and $2T^-$, and whatever s is.*

Proof. Only the case where one single system is observed ($s = 1$), which is sufficient to conclude. Assume that $L(\boldsymbol{\theta}|\boldsymbol{y}) = L(\boldsymbol{\theta_0}|\boldsymbol{y})$ for all $\boldsymbol{y} = (y_1, \dots, y_n)$, the purpose is to show that $\boldsymbol{\theta} = \boldsymbol{\theta_0}$ for all $\boldsymbol{\theta} = (\alpha, \beta, \rho, b)$ and $\boldsymbol{\theta_0} = (\alpha_0, \beta_0, \rho_0, b_0)$ in Θ as soon as $n \geq 2$, whatever s is.

First, when $n = 1$, the observation is a realization of $\Delta X^{(1)}$, which does not depend on rho, so that the model obviously is non identifiable. Assume that $n = 2$. We have by assumption:

$$\ell(\boldsymbol{\theta}|\,\boldsymbol{y}) - \ell(\boldsymbol{\theta_0}|\,\boldsymbol{y}) = 0,$$

where ℓ is the log-likelihood function, and thus by definition this is equivalent to

$$\begin{aligned} &\log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta}) - \log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta_0}) \\ &+ \log\left(\frac{f_{\Delta X^{(2)}}(y_2 - (1-\rho)y_1|\boldsymbol{\theta})}{f_{\Delta X^{(2)}}(y_2 - (1-\rho_0)y_1|\boldsymbol{\theta_0})}\right) = 0. \end{aligned} \tag{2.11}$$

for all y_1 in $(0, \infty)$ and y_2 in $\big(\max\big((1-\rho)y_1, (1-\rho_0)y_1\big), \infty\big)$. This equation is assumed to be true, regardless of the observations set (y_1, y_2). Considering that y_2 tends towards ∞, the quotient of the p.d.f. tends towards 1 and we obtain that

$$\log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta}) - \log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta_0}) = 0.$$

Thus, Equation (2.11) entails that

$$\begin{cases} \log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta}) - \log f_{\Delta X^{(1)}}(y_1|\boldsymbol{\theta_0}) = 0 \\ \log f_{\Delta X^{(2)}}(y_2 - (1-\rho)y_1|\boldsymbol{\theta}) - \log f_{\Delta X^{(2)}}(y_2 - (1-\rho_0)y_1|\boldsymbol{\theta_0}) = 0 \end{cases} \tag{2.12}$$

for all y_1, y_2, and where the first equation depends on y_1 only while the second one depends on both y_1 and y_2. Using the expression of $f_{\Delta X^{(1)}}$, the first equation in (2.12) becomes

$$\log \frac{\Gamma\left(\alpha_0 T^{\beta_0}\right) b^{\alpha T^\beta}}{\Gamma\left(\alpha T^\beta\right) b_0^{\alpha_0 T^{\beta_0}}} + \left(\alpha T^\beta - \alpha_0 T^{\beta_0}\right) \log y_1 + (b_0 - b) y_1 = 0, \tag{2.13}$$

which provides

$$\begin{cases} \log \frac{\Gamma\left(\alpha_0 T^{\beta_0}\right) b^{\alpha T^\beta}}{\Gamma(\alpha T^\beta) b_0^{\alpha_0 T^{\beta_0}}} = 0 \\ \alpha T^\beta - \alpha_0 T^{\beta_0} = 0 \\ b_0 - b = 0 \end{cases} \tag{2.14}$$

based on the fact that the three functions $y_1 \mapsto 1$, $y_1 \mapsto y_1$ and $y_1 \mapsto \log y_1$ involved in (2.13) are linearly independent. Then, $b = b_0$ and $\alpha T^\beta = \alpha_0 T^{\beta_0}$. Now, in the same way, we can write the second equation in (2.12) as

$$\log \frac{\Gamma\left(\alpha_0 T^{\beta_0}(2^{\beta_0}-1)\right) b^{\alpha T^\beta (2^\beta - 1)}}{\Gamma\left(\alpha T^\beta (2^\beta - 1)\right) b_0^{\alpha_0 T^{\beta_0}(2^{\beta_0}-1)}} + \left(\alpha T^\beta (2^\beta - 1) - 1\right) \log(y_2 - (1-\rho)y_1)$$
$$- \left(\alpha_0 T^{\beta_0}(2^{\beta_0} - 1) - 1\right) \log(y_2 - (1-\rho_0)y_1) + b_0(y_2 - (1-\rho_0)y_1) - b(y_2 - (1-\rho)y_1) = 0.$$

Then, replacing b by b_0 and αT^β by $\alpha_0 T^{\beta_0}$, we obtain

$$\log \frac{\Gamma\left(\alpha T^\beta (2^{\beta_0}-1)\right) b^{\alpha T^\beta (2^\beta - 2^{\beta_0})}}{\Gamma\left(\alpha T^\beta (2^\beta - 1)\right)} + \left(\alpha T^\beta (2^\beta - 1) - 1\right) \log(y_2 - (1-\rho)y_1) \tag{2.15}$$
$$- \left(\alpha T^\beta (2^{\beta_0} - 1) - 1\right) \log(y_2 - (1-\rho_0)y_1) + b y_1(\rho_0 - \rho) = 0.$$

We consider that y_1 tends towards 0, which leads to

$$\log \frac{\Gamma\left(\alpha T^\beta (2^{\beta_0}-1)\right) b^{\alpha T^\beta (2^\beta - 2^{\beta_0})}}{\Gamma\left(\alpha T^\beta (2^\beta - 1)\right)} + \alpha T^\beta (2^\beta - 2^{\beta_0}) \log(y_2) = 0$$

and because the functions $y_2 \mapsto 1$ and $y_2 \mapsto \log y_2$ are linearly independent, we have

$$\alpha T^\beta (2^\beta - 2^{\beta_0}) \log(y_2) = 0$$

for all $y_2 > 0$, and thus $\beta = \beta_0$. As a result, Equation (2.15) can be written as

$$\left(\alpha T^\beta (2^\beta - 1) - 1\right) \log\left(\frac{y_2 - (1-\rho)y_1}{y_2 - (1-\rho_0)y_1}\right) + b y_1(\rho_0 - \rho) = 0$$

which is equivalent to

$$\left(\alpha T^\beta (2^\beta - 1) - 1\right) \log\left(1 + \frac{\rho - \rho_0}{y_2/y_1 - (1-\rho_0)}\right) + b y_1(\rho_0 - \rho) = 0.$$

Now, when y_2 tends towards ∞, we derive that $b y_1(\rho_0 - \rho) = 0$ for all $y_1 > 0$, hence $\rho = \rho_0$. Merging this result with the previous, we have

$$\begin{cases} \rho = \rho_0 \\ \beta = \beta_0 \\ b = b_0 \\ \alpha T^\beta = \alpha_0 T^{\beta_0} \end{cases}$$

and finally $\boldsymbol{\theta} = \boldsymbol{\theta_0}$.

□

The maximum likelihood estimate can be computed numerically by maximizing the log-likelihood function $l(\boldsymbol{\theta}|\boldsymbol{y}) = \log L(\boldsymbol{\theta}|\boldsymbol{y})$ whose expression is

$$\ell(\boldsymbol{\theta}|\boldsymbol{y}) = s\left(\alpha (nT)^\beta \log b - \sum_{j=1}^{n} \log \Gamma\left(\Delta A^{(j)}\right)\right) \tag{2.16}$$

$$+\sum_{i=1}^{s}\sum_{j=1}^{n}\left(\Delta A^{(j)}-1\right)\log g_j^{(i)}(\rho)\;-\;b\sum_{i=1}^{s}\sum_{j=1}^{n}g_j^{(i)}(\rho).$$

We now look for the zeros of the partial derivatives of l, and the equation $\partial_b\, l\,(\boldsymbol{\theta}|\boldsymbol{y})=0$ implies that

$$b=\frac{s\alpha(nT)^{\beta}}{\sum_{i=1}^{s}\sum_{j=1}^{n}g_j^{(i)}(\rho)}.$$

Now b is substituted by this expression into Equation (2.16), which provides the profile likelihood function $\tilde{l}$ defined by

$$\begin{aligned}\tilde{l}\,((\alpha,\beta,\rho)|\boldsymbol{y}) &= \sum_{i=1}^{s}\sum_{j=1}^{n}\left(\Delta A^{(j)}-1\right)\log g_j^{(i)}(\rho) \\ &- s\left(\alpha(nT)^{\beta}\left(1-\log\frac{s\alpha(nT)^{\beta}}{\sum_{i=1}^{s}\sum_{j=1}^{n}g_j^{(i)}(\rho)}\right)-\sum_{j=1}^{n}\log\Gamma\left(\Delta A^{(j)}\right)\right).\end{aligned} \tag{2.17}$$

Finally, we obtain an estimate $\hat{\boldsymbol{\theta}}=(\hat{\alpha},\hat{\beta},\hat{\rho},\hat{b})$ of $\boldsymbol{\theta}$ given by:

$$\hat{\boldsymbol{\theta}}=\left(\underset{(\alpha,\beta,\rho)}{\arg\max}\,\tilde{l}\,((\alpha,\beta,\rho)|\boldsymbol{y})\;,\;\frac{s\hat{\alpha}(nT)^{\hat{\beta}}}{\sum_{i=1}^{s}\sum_{j=1}^{n}g_j^{(i)}(\rho)}\right). \tag{2.18}$$

2.3.2 Application to the Arithmetic Reduction of Degradation model of order infinity

Here the MLE method is studied in the case of a ARD_∞ model. However, the development of this method is exactly the same as in the previous section, the only difference is the expression of Y_{jT} with respect to ρ and the observations, and consequently the definition of the function $g_j^{(i)}$.

For the ARD_∞ model, the degradation level Y_{jT} right after the maintenance at time jT can be easily expressed with respect to Y_j and the maintenance efficiency parameter ρ, as we have $Y_{jT}=(1-\rho)Y_j$. Then, we define $g_j^{(i)}$ as

$$g_j^{(i)}(\rho)=Y_j-(1-\rho)Y_{j-1}.$$

With such an expression of $g_j^{(i)}$, **Lemma 1** holds for the ARD_∞ model and the likelihood function has the same expression provided by (2.16). As a result, the likelihood expression is the same here. Note that both models are identical over the time interval $[0,2T)$, during which the first two observations are conducted. Hence, the parameters of the ARD_∞ model are identifiable as before. As a conclusion, Equations (2.17) and (2.18) provide an estimate $\hat{\boldsymbol{\theta}}$ of the model parameters.

The aim of the next chapter is to investigate numerically the estimation methods performances for each model, starting with the ARD_1 model.

Chapter 3

Simulation study

3.1 Arithmetic Reduction of Degradation model of order one

We place ourselves in the following framework for the model parameters and the observations characteristics:

- Shape function parameter: $\alpha = 1$ and $\beta \in \{1.2, 1.6\}$;
- Scale parameter: $b = 1$;
- Maintenance actions efficiency: $\rho \in \{0.2, 0.5, 0.8\}$;
- Period of repairs: $T = 1$;
- Observations times: $\{jT^-; 1 \leq j \leq n\}$ with $n \in \{5, 20\}$;
- Number of observed i.i.d. systems: $s \in \{50, 200\}$.

For each parameters combination, 2000 data sets are generated, and an estimation of $\boldsymbol{\theta}$ is computed for each data set. The ME method is here based on the minimization of a bivariate function, resulting in an estimation of β and ρ. The optimization is performed through a constrained gradient method: β is sought over the interval $[0.1, 5]$ and ρ in $[0.01, 0.99]$. This iterative algorithm is initialized at point $(2.5, 0.5)$ when $\rho \neq 0.5$, and at point $(2.5, 0.8)$ otherwise. Regarding the MLE method, it is based on the maximization of a function depending on three parameters: α, β and ρ. Once again the optimization is performed through a constrained gradient method in the same way as before for β and ρ, and the parameter α is sought in the same interval as β. Some investigations lead us to conclude that the initialization point of the iterative optimization procedure has no influence on the estimation results, and it only has very few impact on the CPU times. Moreover, the CPU time is negligible as the mean CPU time for one single estimation is less than 0.1 second for the ME method and 0.25 second for the MLE method.
The results are analysed through two indicators, the Relative Bias (RB) in percentage, defined as

$$RB\left(\hat{\boldsymbol{\theta}}\right) = 100 \times \frac{\mid \bar{\boldsymbol{\theta}} - \boldsymbol{\theta_0} \mid}{\boldsymbol{\theta_0}},$$

where $\bar{\boldsymbol{\theta}}$ is the empirical mean of the estimates, and the empirical variances of the estimates. For each possible combination of (β, ρ), the RB and the variances of the estimations are summarized in **Appendix A**. The figures deal either with the RB or with the variance, and are composed of four graphs: one for

each parameter of the model. These graphs show four curves each, representing either the relative bias or the variance evolution with s for both the ME and the MLE methods when $n = 5$ and $n = 20$.
To be able to compare the results from one parameter to another, as well as the parameters sets, the y-axis range is the same for all plots. However, both the RB and the variance can be higher than the maximum value displayed in the y-axis. This signifies that if a curve does not appear on a graph, as the black one in the top right-hand graph as well as the bottom left-hand one of **Figure A.1**, it means that all the points of the curve are out of range.

Analysis of the results

When $n = 5$, the ME method provides good estimations in terms of relative bias and variance of:

- α and b when $\rho = 0.2$;
- β when $\rho = 0.5$;
- b when $\rho = 0.8$.

The value of β does not affect significantly the estimations quality, but note that ρ is never well estimated here, even if the RB is sometimes low, the variance is always fairly high. As a result, this method provides poor quality estimations regarding the four parameters simultaneously.
If $n = 20$, once again the value of β does not affect significantly the results. In this case, when $\rho = 0.8$, the ME method provides good estimations of the parameters. However, if $\rho = 0.5$ then the estimation quality of b is poor, it is also true for the estimation of ρ when its true value is set to 0.2.

Regarding the MLE method, it provides good quality estimations, better than those from the ME method. The RB of all the parameters estimators are less than 2%, except when $\rho = 0.2$, $n = 5$ and $s = 50$ where the RB related to ρ approaches 5%. Also, the variances are low, especially regarding β and ρ: they are between 5.4×10^{-3} and 9.1×10^{-5} for β, and between 2.2×10^{-3} and 1.4×10^{-6} for ρ. Here the value of β does not affect the results as well. Finally, the MLE method is more efficient in the case where $n = 20$ than when $n = 5$, except sometimes regarding the estimation of the scale parameter.

Remark 1. *The case where $\beta < 1$ is not studied here due to numerical issues with the computation of the likelihood function given in Equation (2.10). Actually, if $\beta < 1$ then the shape parameter $\Delta A^{(j)}$ of the degradation increments $\Delta X^{(j)}$ is less than 1 as well. Hence the quantity $\left(g_j^{(i)}(\rho)\right)^{\Delta A^{(j)}-1}$ might tend towards infinity. To be more precise, there exists in some case a value $\tilde{\rho}$ in $(0, 1)$ such that the likelihood is only defined over $(\tilde{\rho}, 1)$ regarding the parameter ρ, with $g_j^{(i)}(\tilde{\rho}) = 0$ and $\lim_{\rho \to \tilde{\rho}^+} L\left(\boldsymbol{\theta} | \left(\boldsymbol{y^{(1)}}, \ldots, \boldsymbol{y^{(s)}}\right)\right) = \infty$. Although this particularity allows to have a good estimation of the maintenance efficiency (see **Chapter 4**), the sensitivity of the likelihood in the neighbourhood of $\tilde{\rho}$ prevents the optimization from estimating properly the other parameters.*

3.2 Arithmetic Reduction of Degradation model of order infinity

Here the same analysis as in the previous section is done, with an identical framework for the model parameters, the observations characteristics, the optimization procedure as well as for the results summary which appears in **Appendix B**. Also, the case $\beta < 1$ is not studied for the reasons already mentioned

above.

Regarding the ME method, the parameter β is here well estimated when $\rho = 0.2$, associated with a low variance, and that whatever the values of n, s or β. When $n = 5$, this method is not reliable if $\rho = 0.5$ as none of the parameters is properly estimated, while α and b are rather well estimated if $\rho = 0.8$. When $n = 20$, the method provides good estimations of β and b if $\rho = 0.5$. Finally, as for the ARD_1 model, the ME method is efficient when $\rho = 0.8$, but the variance of this parameter is rather high. Once again the true value of β does not affect the results in a significant way.

About the MLE method, the results are close to those of the previous section:

- globally very good estimations with RB values smaller than 2%, except when $\rho = 0.2$;
- when $\rho = 0.2$, the RB related to this parameter is higher (it approaches 3%);
- the variances of the estimations of β and ρ are low, with similar boundaries as before;
- this method is better as n increases, except sometimes for α.

As a result, for both ARD_1 and ARD_∞ models, the MLE method is the best.

Chapter 4

Semiparametric estimate of the maintenance actions efficiency

Sections 4.1 to 4.6 are devoted to a semiparametric estimation procedure in the context of an ARD1 model. These sections are the exact reproduction of the following paper: [38] G. Salles, S. Mercier, and L. Bordes. Semiparametric estimate of the efficiency of imperfect maintenance actions for a gamma deteriorating system. *Journal of Statistical Planning and Inference*, 206:278-297, 2020.

Section 4.7 has been added to the previous paper, where the same procedure is explored in the context of an ARD_∞ model.[1]

4.1 Introduction

Safety and dependability are crucial issues in many industries (such as, e.g., railways, aircraft engines or nuclear power plants), which have lead to the development of the so-called reliability theory. For many years, only lifetime data were available and the first reliability studies were focused on lifetime data analysis (see, e.g., [32]), which still remains of interest in many cases. In that context and in case of repairable systems with instantaneous repairs, successive failure (or repair) times appear as the arrival points of a counting process, and failures hence correspond to recurrent events. As for the type of possible repairs, typical classical models are perfect (As-Good-As-New) and minimal (As-Bad-As-Old) repairs, leading to renewal and non homogeneous Poisson processes as underlying counting processes, respectively (see [5]). The reality often lies in-between, leading to the class of imperfect repairs. Many models have been envisioned in the literature for their modeling, such as, e.g., virtual age models introduced by Kijima [26] and further studied in [13, 16], geometric processes [27] (extended in [6]) or models based on reduction of failure intensity [13, 16]. See, e.g., [17] for a recent account and extensions of such models. See also [36] for more references and other models.

Nowadays, the development of online monitoring and the increasing use of sensors for safety assessment make it possible to get specific information on the health of a system and on its effective evolution over time, without waiting for the system failure. This information is often synthesized into a scalar indicator, which can for instance stand for the length of a crack, the thickness of a cable, the intensity of vibrations, corrosion level, ... This scalar indicator can be considered as a measurement of the deterioration level

[1] Due to the introduction of a published paper within this chapter, some notations differ from those used in the remainder of the document.

of the system. The evolution of this deterioration indicator over time is nowadays commonly modeled through a continuous-time and continuous-state stochastic process, which is often considered to have an increasing trend. Classical models include inverse Gaussian [44] or Wiener processes with trend [22, 29, 45], which are also quite common in many other fields out of reliability theory, such as finance, insurance or epidemiology. This paper focuses on gamma processes, which are widely used since they were introduced in the reliability field by Çinlar [10]. See [40] and its references for a large overview.

In order to mitigate the degradation of the system over time and extends its lifetime, preventive maintenance actions can be considered, in addition to corrective repairs which are performed at failure. In the context of deteriorating systems, many preventive maintenance policies from the literature consider condition-based maintenance (CBM) actions, where the preventive repair is triggered by the reaching of a preventive maintenance threshold by the deterioration level. In that context, "most of the existing CBM models have been limited to perfect maintenance actions", as noted by [3] (see also [45]). Some imperfect repair models are however emerging in the latest reliability literature, in this new context of deteriorating systems, see [3] for a recent review. Some models are based on the notion of virtual age previously introduced in the context of recurrent events (see, e.g., [19, 33]), where the system is rejuvenated by a maintenance action. Other models consider that an imperfect repair reduces the deterioration level of the system, such as [25, 28, 35, 37, 42], which can be accompanied by some increase in the deterioration rate, as in [15]. Also, some papers consider that the efficiency decreases with the number of repair (see, e.g., [30, 46]), and further studies, as in [23], deal with imperfect maintenance models such that (i) repairs have a random efficiency (ii) the deterioration rate increases with the number of repairs. In all these papers however, the main point mostly is on the optimization of a maintenance policy, including these imperfect maintenance actions together with perfect repairs (replacements). Up to our knowledge, very few papers from the literature deal with statistical issues concerning imperfect repair models for deteriorating systems, except from [45], where the authors suggest a maximum likelihood method for estimating the parameters of the Wiener process together with an iterative procedure based on a Kalman filter for the different factors implied in successive imperfect repairs. This estimation procedure is developed in a fully parametric context and validated on simulated data, without any study of the asymptotic properties of the estimators.

The evaluation of the maintenance actions efficiency is mainly used for maintenance policies optimization. Once the repair efficiency has been estimated, the future behavior of the maintained system can be predicted, which allows to adapt (optimize) the periodicity of the maintenance actions and efficiently plan a general overhaul. From a safety point of view, the principal inquiry is to ensure that the maintenance actions are effective enough to keep with a high probability the degradation level below a fixed threshold (safety level). As long as this safety level is not reached, the maintenance actions must be adjusted, either by adapting their periodicity or by improving their efficiency (if possible). Of course, apart from the previous safety concern, the maintenance costs are another issue. As an example, in [42], the costs minimization is based on the monitoring time and on the imperfect maintenance efficiency. In [23], the author considers a threshold for the degradation, beyond which an imperfect maintenance is performed. The optimization is made with respect to this threshold, the inspections periodicity and the repairs efficiency. See, e.g., both papers cited above and their reference for an overview on maintenance policies optimization.

This paper focuses on a specific imperfect repair model, where each maintenance action reduces the deterioration level of the system. The model was first introduced in [8] and further studied in [34], where it was called Arithmetic Reduction of Degradation model of order 1 (ARD_1). Mimicking Arithmetic Reduction of Age (ARA_1) and Arithmetic Reduction of Intensity (ARI_1) models of order 1 developed by

[16] in the context of recurrent events, the idea of an ARD_1 model is that a maintenance action removes a proportion ρ of the degradation accumulated by the system from the last maintenance action (where $\rho \in [0,1)$). The parameter ρ appears as a measure of the maintenance efficiency, which lies between As-Good-As-New when $\rho \to 1^-$ and As-Bad-As-Old when $\rho = 0$. Along the same lines as [8, 16, 34], the maintenance actions efficiency is here assumed to be fixed and independent of the intrinsic degradation.

This paper is concerned with the development and study of an estimation procedure for the maintenance efficiency parameter ρ, in the context of a gamma deteriorating system subject to periodic ARD_1 imperfect repairs. Observations are lead on just before each maintenance action. Considering n successive repairs, this leads to multivariate data, from where an estimator of ρ is proposed. The idea of this estimator has come from a preliminary study in a parametric framework based on the maximum likelihood method, where we have observed that the minimum of admissible ρ's has quite an interesting behavior, getting quickly very close to the unknown efficiency parameter when n increases. This has lead to the proposition of an original estimator for ρ, which depends only on the data, and not on the shape function and rate parameter of the gamma process, leading to a semiparametric framework. Under technical assumptions, the strong consistency of this new estimator is shown, as the number n of repairs tends towards infinity. Also, the convergence rate is proved to be surprisingly high, and can even reach an exponential speed in some cases. This estimator hence appears to be super consistent (under specific conditions). This is illustrated on simulated data at the end of the paper, where we provide two examples for which we observe that roughly 95% of the estimates are exact at the machine precision level (6.10^{-17}) as soon as $n > 40$ and $n > 88$, respectively, with a mean error below 10^{-15} in both cases. The study is next extended to the case where s independent and identical systems are observed (n times each). A similar semiparametric estimator is proposed for the (common) maintenance efficiency and the strong consistency is proved to hold as s tends towards infinity, no matter the fixed value of n and out of any technical condition requirement. The convergence rate is studied, which is shown to depend on the shape function of the gamma process and on the maintenance period, leading to a speed that can be either slower or faster than $\sqrt{s}$, according to the case.

The outline of this paper is as follows. The framework is specified in Section 4.2, which covers the gamma deterioration process, the ARD_1 imperfect repair model and the observation scheme. Section 4.3 is devoted to the study of the semiparametric estimator in the case where one single system is observed, which includes its asymptotic properties when the number of repairs tends towards infinity. Section 4.4 deals with the extension to several systems and considers the asymptotic properties with respect to the number of observed systems. Some illustrations of the estimator performances are provided in Section 4.5 and conclusions are formulated in Section 4.6.

4.2 Framework

4.2.1 Intrinsic deterioration

Let us first recall that a random variable X is said to be gamma distributed with a and b as shape and rate parameters, respectively ($X \sim \Gamma(a,b)$ with $a,b > 0$), if its distribution admits the following probability density function (p.d.f.):

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \, \mathbf{1}_{\mathbb{R}^+}(x)$$

with respect to Lebesgue measure. Its mean, variance and Laplace transform are provided by

$$\mathbb{E}(X) = \frac{a}{b},\ \mathbb{V}(X) = \frac{a}{b^2},\ \mathcal{L}_X(t) = \mathbb{E}\left(e^{-tX}\right) = \left(\frac{b}{b+t}\right)^a,\ \forall t \geq 0,$$

respectively. Moreover, $cX \sim \Gamma(a, b/c)$ for any $c > 0$, and the sum of n independent random variables $X_i \sim \Gamma(a_i, b)$ (with $1 \leq i \leq n$) is also gamma distributed with $X_1 + \cdots + X_n \sim \Gamma(a_1 + \cdots + a_n, b)$.

Now, let $a(\cdot) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ be a continuous and non decreasing function such that $a(0) = 0$ and let $b > 0$. Also let $(X_t)_{t\geq 0}$ be a right-continuous process with left-side limits. Then, we recall that $(X_t)_{t\geq 0}$ is a non homogeneous gamma process with shape function $a(\cdot)$ and rate parameter b, as soon as

- $X_0 = 0$ almost surely (a.s.),
- $(X_t)_{t\geq 0}$ has independent increments,
- each increment is gamma distributed: for all $0 \leq s < t$, we have $X_t - X_s \sim \Gamma(a(t) - a(s), b)$,

(see, e.g., [1]).

In all the sequel, the intrinsic deterioration of the system (that is out of repairs) is assumed to be modeled by a non homogeneous gamma process $(X_t)_{t\geq 0}$ with shape function $a(\cdot)$ and rate parameter b.

4.2.2 The imperfect repair model

In order to lower the deterioration level, instantaneous and periodic imperfect repairs are carried out on the system every T units of time ($T > 0$). Following [8, 34], an Arithmetic Reduction Degradation model of order 1 (ARD_1) is considered, where a maintenance action removes a proportion $\rho \in [0, 1)$ of the deterioration accumulated since the last maintenance action (or from time $t = 0$). The model used in the present paper is just the same as that used in [34], which we now recall, for sake of completeness.

Between repairs, the system is assumed to evolve according to independent and identically distributed (i.i.d.) copies $\left(X_t^{(i)}\right)_{t\geq 0}$, $i = 1, 2, \ldots$ of the gamma process $(X_t)_{t\geq 0}$, where exponent (i) refers to the i-th between-repair period $[(i-1)T, iT)$ (where time 0 is considered as a repair time). We set $(Y_t)_{t\geq 0}$ to describe the overall deterioration level of the maintained system, as a result of the intrinsic deterioration and of the imperfect periodic repairs.

On the first time interval $[0, T)$, there is no repair and

$$Y_t = X_t^{(1)} \text{ for all } t \in [0, T).$$

This implies that $Y_{T^-} = X_{T^-}^{(1)} = X_T^{(1)}$ a.s., based on the almost sure continuity of a gamma process.

At time T, the deterioration level is reduced of $\rho X_T^{(1)}$ so that $Y_T = (1-\rho)X_T^{(1)}$ a.s.

On the second time interval $[T, 2T)$, we now have:

$$Y_t = Y_T + X_t^{(2)} - X_T^{(2)} \text{ for all } t \in [T, 2T),$$

which leads to

$$Y_{2T^-} = Y_T + X_{2T^-}^{(2)} - X_T^{(2)} = (1-\rho)X_T^{(1)} + X_{2T}^{(2)} - X_T^{(2)} \text{ a.s.}$$

and

$$Y_{2T} = (1-\rho)X_T^{(1)} + (1-\rho)\left(X_{2T}^{(2)} - X_T^{(2)}\right) \text{ a.s.}$$

More generally, on the j-th time interval (with $j \in \mathbb{N}$), the effective degradation level can be expressed as

$$Y_t = Y_{jT} + \left(X_t^{(j+1)} - X_{jT}^{(j+1)}\right) \text{ for all } t \in [jT, (j+1)T),$$

which leads to

$$Y_{jT^-} = (1-\rho)\sum_{p=1}^{j-1}\left(X_{pT}^{(p)} - X_{(p-1)T}^{(p)}\right) + \left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right) \tag{4.1}$$

and

$$Y_{jT} = (1-\rho)\sum_{p=1}^{j}\left(X_{pT}^{(p)} - X_{(p-1)T}^{(p)}\right)$$

(with the convention that an empty sum is zero).

An example of trajectory of $(Y_t)_{t\geq 0}$ is given in Figure 4.1 for $a(t) = t^{1.5}$, $b = 1$, $T = 1$ and $\rho = 0.5$, together with the corresponding trajectories of the $(X_t^{(j)})_{t\geq 0}, j = 1, 2, \ldots$

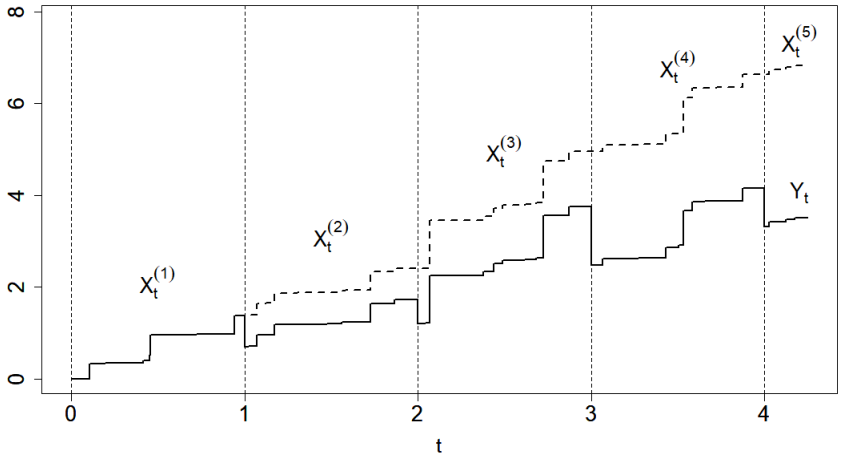

Figure 4.1: An example of simulated trajectories of $\left(X_t^{(j)}\right)_{t\geq 0}, j = 1, \ldots, 5$ and of $(Y_t)_{t\geq 0}$ (intrinsic and overall degradation levels, respectively) with parameters $a(t) = t^{1.5}$, $b = 1$, $T = 1$ and $\rho = 0.5$.

Note that out of maintenance times ($t \notin \{jT, j = 1, 2, \ldots\}$), the random variable Y_t is the sum of two gamma random variables which do not share the same rate parameter (except if $\rho = 0$). Hence, it is not gamma distributed. Please see [34] for more details on this model.

The periodicity T is assumed to be known and the previous model is called ARD_1 model with parameter $(a(\cdot), b, \rho)$ in the following (with T omitted).

As known from the introduction, our focus is on the development of an estimation procedure for the maintenance efficiency parameter ρ. We now specify the observation scheme and derive some first consequences.

4.2.3 Observation scheme and first consequences

The deterioration level of the maintained system is assumed to be (perfectly) measured n times ($n \in \mathbb{N}^*$), right before the n first maintenance actions, that is at times $T^-, \ldots, nT^-$. The data hence is an

observation of $(Y_{T^-}, \ldots, Y_{nT^-})$, where Y_{jT^-} is provided by (4.1).

For the sake of simplicity, we set

$$\begin{aligned} Y_j &= Y_{jT^-}, \\ U_j &= X_{jT}^{(j)} - X_{(j-1)T,}^{(j)} \\ a_j &= a(jT) - a\big((j-1)T\big) \end{aligned}$$

for all $j = 1, \ldots, n$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$.

With the previous notation and based on the independent increments of a gamma process, the random variables U_j's can be seen to be independent with $U_j \sim \Gamma(a_j, b)$ for all $j = 1, \ldots, n$. Also, for an ARD_1 model with parameter $(a(\cdot), b, \rho)$, Equation (4.1) can now be written as:

$$Y_j = (1-\rho)\sum_{p=1}^{j-1} U_p + U_j \tag{4.2}$$

for all $1 \leq j \leq n$.

We first check the theoretical identifiability of the model, considering the parameters of the underlying gamma process as nuisance parameters (and T fixed).

Proposition 6. (Identifiability) *Let* $\mathbf{Y} = \left(Y_1, \ldots, Y_n\right)$ *and* $\tilde{\mathbf{Y}} = \left(\tilde{Y}_1, \ldots, \tilde{Y}_n\right)$ *be two random vectors based on the* ARD_1 *repair model with parameters* $(a(\cdot), b, \rho)$ *and* $\left(\tilde{a}(\cdot), \tilde{b}, \tilde{\rho}\right)$*, respectively (and the same period* T*). Assume that there are at least two observations (*$n \geq 2$*). Then if* $\mathbf{Y}$ *and* $\tilde{\mathbf{Y}}$ *are identically distributed (denoted by* $\mathbf{Y} \overset{\mathcal{D}}{=} \tilde{\mathbf{Y}}$ *), necessarily* $\rho = \tilde{\rho}$*.*

Proof. Assume that $\mathbf{Y} \overset{\mathcal{D}}{=} \tilde{\mathbf{Y}}$ and $n \geq 2$. Then Y_1 and $\tilde{Y}_1$ are identically distributed, with $Y_1 = U_1 \sim \Gamma(a_1, b)$ and $\tilde{Y}_1 = \tilde{U}_1 \sim \Gamma\left(\tilde{a_1}, \tilde{b}\right)$. This implies that $a_1 = \tilde{a}_1$ and $b = \tilde{b}$.

Also, $Y_2 = (1-\rho)U_1 + U_2$ and $\tilde{Y}_2 = (1-\tilde{\rho})\tilde{U}_1 + \tilde{U}_2$ must share the same distribution, and hence the same Laplace transform.

Based on the independence between U_1 and U_2, the Laplace transform of Y_2 is

$$\mathcal{L}_{Y_2}(t) = \mathcal{L}_{(1-\rho)U_1}(t)\,\mathcal{L}_{U_2}(t) = \left(\frac{b}{b+(1-\rho)t}\right)^{a_1}\left(\frac{b}{b+t}\right)^{a_2},$$

with a similar expression for $\tilde{Y}_2$. Remembering that $a_1 = \tilde{a}_1$ and $b = \tilde{b}$, this leads to

$$\left(\frac{b}{b+(1-\rho)t}\right)^{a_1}\left(\frac{b}{b+t}\right)^{a_2} = \left(\frac{b}{b+(1-\tilde{\rho})t}\right)^{a_1}\left(\frac{b}{b+t}\right)^{\tilde{a}_2},$$

for all $t \geq 0$, which can be simplified into

$$\left(\frac{1+(1-\tilde{\rho})u}{1+(1-\rho)u}\right)^{a_1} = \frac{1}{(1+u)^{\tilde{a}_2 - a_2}}$$

for all $u \geq 0$, setting $u = t/b$. A first order series expansion at point 0 induces that $a_1(\tilde{\rho} - \rho) = \tilde{a_2} - a_2$ and next that

$$\frac{1+(1-\tilde{\rho})u}{1+(1-\rho)u} = \frac{1}{(1+u)^{\tilde{\rho}-\rho}}$$

for all $u \geq 0$. Taking the limit when $u \to +\infty$ in the previous equation, we get

$$\frac{1-\tilde{\rho}}{1-\rho} = \lim_{u\to+\infty} \frac{1}{(1+u)^{\tilde{\rho}-\rho}} = \begin{cases} 0 \text{ if } \tilde{\rho} > \rho, \\ 1 \text{ if } \tilde{\rho} = \rho, \\ \infty \text{ if } \tilde{\rho} < \rho. \end{cases}$$

This is possible only if $\tilde{\rho} = \rho$ since ρ and $\tilde{\rho}$ belong to $[0, 1)$, which achieves the proof. □

The identifiability hence holds as soon as two observations are available.

From now on, we assume that the true maintenance efficiency parameter is $\rho_0 \in [0, 1)$. The Y_j's and the U_j's hence correspond to an ARD_1 model with parameters $(a(\cdot), b, \rho_0)$. A first link between the Y_j's and the U_j's $(j = 1, \cdots, n)$ has been provided in Equation (4.2). We now invert this system of equations, thus providing an expression of the U_j's with respect to the Y_j's, that will be used in the sequel.

Lemma 2. *For each $j \in \{1, \cdots, n\}$, the increment U_j can be expressed with respect to the observations $Y_1, \ldots, Y_j$ and to the maintenance efficiency parameter ρ_0 as follows:*

$$U_j = \sum_{p=1}^{j} \rho_0^{j-p} (Y_p - Y_{p-1}), \tag{4.3}$$

where we set $Y_0 = 0$.

Proof. This result is proved by induction on j. For $j = 1$, the ARD_1 model definition provides $Y_1 = U_1$.

Now, assume that (4.3) is true for some fixed $1 \leq j \leq n-1$. Observe from (4.2) that

$$\begin{aligned} Y_{j+1} - Y_j &= (1-\rho_0) \sum_{p=1}^{j} U_p + U_{j+1} - \left[(1-\rho_0) \sum_{p=1}^{j-1} U_p + U_j \right] \\ &= U_{j+1} - \rho_0 U_j, \end{aligned}$$

or equivalently that $U_{j+1} = Y_{j+1} - Y_j + \rho_0 U_j$.

Using the induction assumption, we easily derive that

$$\begin{aligned} U_{j+1} &= Y_{j+1} - Y_j + \rho_0 \sum_{p=1}^{j} \rho_0^{j-p} (Y_p - Y_{p-1}) \\ &= \sum_{p=1}^{j+1} \rho_0^{j+1-p} (Y_p - Y_{p-1}). \end{aligned}$$

Hence, Equation (4.3) holds for $j+1$, which achieves the proof. □

For each $1 \leq j \leq n$, let us now define the function $g_j(\rho, \mathbf{Y})$ by

$$g_j(\rho, \mathbf{Y}) = \sum_{p=1}^{j} \rho^{j-p} (Y_p - Y_{p-1}), \ \forall \rho \in [0, 1), \tag{4.4}$$

where we recall that the Y_j's refer to the true maintenance efficiency parameter ρ_0.

Lemma 2 ensures that $g_j(\rho_0, \mathbf{Y})$ matches with the increment U_j, that is

$$g_j(\rho_0, \mathbf{Y}) = U_j, \tag{4.5}$$

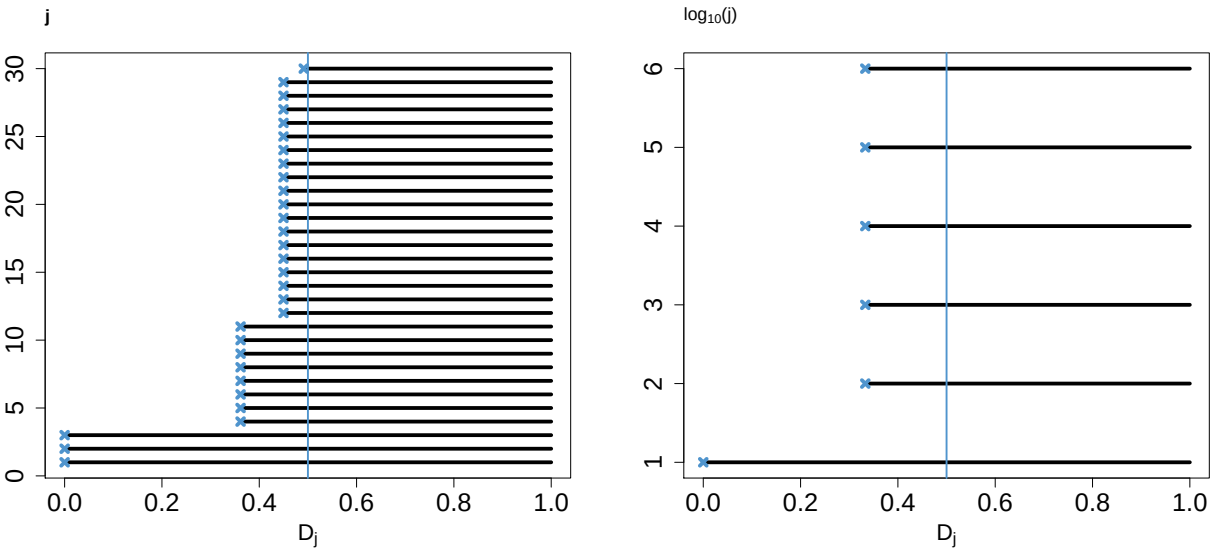

Figure 4.2: Representation of the sets $D_j, j = 1, \ldots, n$ (black horizontal segments) for $\rho_0 = 0.5$ (vertical blue line) and $b = 1$, with $a(t) = \sqrt{t}$ (concave function), $n = 30$ for the left plot, and $a(t) = t^{1.5}$ (convex function), $n = 10^6$ for the right one, where the lower bounds M_j's of the D_j's are highlighted by blue crosses, Example 2.

for each $1 \leq j \leq n$. As these increments are gamma distributed, they necessarily are non negative. Hence the true parameter ρ_0 fulfils the condition $g_j(\rho_0, \mathbf{Y}) \geq 0$ for each $j \in \{1, \ldots, n\}$. An important consequence is that the range for the admissible ρ's can be restricted to the set

$$D_n = \{\rho \in [0, 1) : g_j(\rho, \mathbf{Y}) \geq 0 \text{ for all } j \in \{1, \ldots, n\}\}.$$

For a better understanding of what the D_n's are, we now look at an example, based on simulated data, where we consider the successive $D_1, \ldots, D_n$ (where the D_j's, $j = 1, \ldots, n$, are defined in a similar way as D_n).

Example 2. *Two sets of parameters are considered, with $\rho_0 = 0.5$, $T = 1$ and $b = 1$ for both, $a(t) = \sqrt{t}$ (concave function) for the first set and $a(t) = t^{1.5}$ (convex function) for the second one. An observation of $\mathbf{Y}$ is generated for each of the two parameter sets, and the corresponding observations of the D_j's are next computed. They are plotted in the left and right plots of Figure 4.2 for the concave and convex cases, respectively. The range for n is $\{1, \cdots, 30\}$ for the left plot (concave case) and $\{1, \cdots, 10^6\}$ for the right plot (convex case). Also, the parameter ρ_0 is highlighted by a vertical blue line on each plot. We can observe that in both cases, the sets D_j's are intervals of the shape $[M_j, 1)$ and that $M_1 \leq M_2 \leq \cdots \leq M_n \leq \rho_0$. (Please note that the M_j's are indicated by blue crosses on the graphs). As can be seen on the left plot, it seems that, in case of a concave shape function, the sequence (M_j) converges very quickly towards ρ_0 when j increases. When the shape function is convex, it might also be convergent towards ρ_0 (?), but if so, it can only be at a very slow rate.*

From the previous example, it seems that M_n could be a very good estimator for ρ_0 in the concave case. However, in the convex case, even if the sequence (M_j) happened to converge towards ρ_0 when j increases (which we do not know), the rate of convergence would apparently be far below the classical square-root speed that could be obtained, e.g., with a maximum likelihood estimator. There hence is no interest in pursuing on this way in the convex case.

As a summary, from the previous observations, we suggest to use M_n as an estimator of the maintenance efficiency parameter ρ_0, that we hope to be convergent at a very high speed in the case of a concave shape function. Note that it is a semiparametric estimator of ρ_0 since the parameters of the gamma process are unknown and not restricted to a parametric family (the shape function $a(\cdot)$ is unknown).

4.3 The semiparametric estimator and its asymptotic properties

This section is devoted to the formal definition of the semiparametric estimator (Subsection 4.3.1), together with the study of its asymptotic properties (Subsection 4.3.3), when the number of imperfect repairs n tends to infinity, with one single system observed.

4.3.1 Definition and first properties

Let us first recall from the previous section that

$$D_j = \{\rho \in [0,1) : g_k(\rho, \mathbf{Y}) \geq 0 \text{ for all } k \in \{1,\dots,j\}\}$$

for all $1 \leq j \leq n$. Note that each set D_j is non empty, because $g_k(1^-, \mathbf{Y}) = Y_k \geq 0$ for all $k \in \{1,\dots,j\}$.

Also, $g_1(\rho, \mathbf{Y}) = Y_1 \geq 0$ for all $\rho \in [0,1)$, which implies that $D_1 = [0,1)$. Finally, it is readily seen that

$$D_{j+1} = \{\rho \in D_j : g_{j+1}(\rho, \mathbf{Y}) \geq 0\}$$

and that $D_{j+1} \subset D_j$ for all $1 \leq j \leq n-1$.

Let us set

$$M_j = \inf(D_j)$$

for all $1 \leq j \leq n$.

Proposition 7. *The function $g_j(\rho, \mathbf{Y})$ and the sequence $(M_j)_{1\leq j\leq n}$ (almost surely) satisfy the following properties:*

1. *$M_j \leq M_{j+1}$ for all $1 \leq j \leq n-1$;*
2. *$\rho \mapsto g_j(\rho, \mathbf{Y})$ is non negative on D_j for all $1 \leq j \leq n$;*
3. *$\rho \mapsto g_j(\rho, \mathbf{Y})$ is non decreasing on D_{j-1} for all $1 \leq j \leq n$ (where we set $D_0 = D_1 = [0,1)$ for convenience);*
4. *$D_j = [M_j, 1)$ for all $1 \leq j \leq n$;*
5. *$M_j \leq \rho_0$ for all $1 \leq j \leq n$.*

Proof. Points 1 and 2 are clear due to $D_{j+1} \subset D_j$ and to the definition of D_j.

Let us show the three following points (Points 3-5) all together by induction on j.

At first, we have $D_0 = D_1 = [0,1)$, $M_1 = 0 \leq \rho_0$ and $g_1(\rho, \mathbf{Y}) = Y_1$ for all $\rho \in [0,1)$. Hence Points 3-5 are true for $j = 1$.

Now, assume Points 3-5 to be true for some $j \in \{1, \cdots, n-1\}$ and, to begin with, let us note that Equation (4.4) implies that

$$\begin{aligned} g_{j+1}(\rho, \mathbf{Y}) &= Y_{j+1} - Y_j + \rho \sum_{p=1}^{j} \rho^{j-p} (Y_p - Y_{p-1}) \\ &= Y_{j+1} - Y_j + \rho\, g_j(\rho, \mathbf{Y}) \end{aligned} \tag{4.6}$$

for all $1 \leq j \leq n-1$ (where $Y_{j+1} - Y_j$ might be negative).

By the induction assumption, $g_j(\rho, \mathbf{Y})$ is non decreasing on D_{j-1}, and hence also on D_j (as $D_j \subset D_{j-1}$). Based on the previous recursion formula (4.6) and as $g_j(\rho, \mathbf{Y})$ also is non negative on D_j, this implies that $g_{j+1}(\rho, \mathbf{Y})$ is non decreasing on D_j.

As $D_j = [M_j, 1)$ by the induction assumption, we now have

$$M_{j+1} = \inf \{\rho \in [M_j, 1) : g_{j+1}(\rho, \mathbf{Y}) \geq 0\},$$

where g_{j+1} is non decreasing and continuous on $[M_j, 1)$, with $g_{j+1}(1^-, \mathbf{Y}) = Y_{j+1} > 0$ (almost surely). This implies that D_{j+1} is an interval and $D_{j+1} = [M_{j+1}, 1)$.

Finally, from Equation (4.5), we have $g_{j+1}(\rho_0, \mathbf{Y}) = U_{j+1} \geq 0$. As $M_j \leq \rho_0$ by the induction assumption and $g_{j+1}(\rho, \mathbf{Y})$ is known to be non decreasing on $[M_j, 1)$, we necessarily have $M_{j+1} \leq \rho_0$. Hence Points 2-5 are true for $j+1$, and this achieves the proof. □

Based on the previous results, we can see that the sequence $(M_j)_{1 \leq j \leq n}$ can be alternatively defined through

$$\begin{cases} M_1 = 0, \\ M_{j+1} = \inf \{\rho \in [M_j, 1) \,:\, g_{j+1}(\rho, \mathbf{Y}) \geq 0\}, \quad \text{for all } 1 \leq j \leq n-1. \end{cases} \tag{4.7}$$

Also, considering a possibly infinite number of observations, the sequence $(M_n)_{n \geq 1}$ is non decreasing and upperly bounded by ρ_0. It hence is an almost sure convergent sequence. It remains to prove that it converges towards ρ_0, which is done in Subsection 4.3.3, under specific technical assumptions (among with, concavity of the shape function of the gamma process). With that aim, some technical results have first to be established, which is done in the next subsection.

4.3.2 Technical results

Lemma 3. *Let $\rho \in [0, 1)$. Then, $g_j(\rho, \mathbf{Y}) \geq 0$ implies that $\rho_0 - \rho \leq \dfrac{U_j}{U_{j-1}}$ for each $2 \leq j \leq n$.*

Proof. Let us first prove by induction that

$$g_j(\rho, \mathbf{Y}) = (\rho - \rho_0) \sum_{p=1}^{j-1} \rho^{j-1-p} U_p + U_j.$$

For $j = 1$, the result is clear because $g_1(\rho, \mathbf{Y}) = Y_1 = U_1$. Assume that the result holds for some $j \in \{1, \cdots, n-1\}$. Based on (4.6) and (4.5), we know that

$$\begin{aligned} g_{j+1}(\rho, \mathbf{Y}) &= Y_{j+1} - Y_j + \rho\, g_j(\rho, \mathbf{Y}), \\ U_{j+1} &= Y_{j+1} - Y_j + \rho_0\, U_j, \end{aligned}$$

(taking $\rho = \rho_0$ in the first line to derive the second one), which provides

$$g_{j+1}(\rho, \mathbf{Y}) = U_{j+1} - \rho_0\, U_j + \rho\, g_j(\rho, \mathbf{Y}).$$

Using the induction assumption, $g_{j+1}(\rho, \mathbf{Y})$ can now be expressed as follows:

$$\begin{aligned} g_{j+1}(\rho, \mathbf{Y}) &= U_{j+1} - \rho_0 U_j + \rho \left((\rho - \rho_0) \sum_{p=1}^{j-1} \rho^{j-1-p} U_p + U_j \right) \\ &= (\rho - \rho_0) \sum_{p=1}^{j} \rho^{j-p} U_p + U_{j+1} \end{aligned}$$

where the last equality results from straightforward calculations. Thus we obtain the first result.

Next we note that $g_j(\rho, \mathbf{Y}) \geq 0$ is true as soon as

$$(\rho - \rho_0) \sum_{p=1}^{j-1} \rho^{j-1-p} U_p + U_j \geq 0,$$

or equivalently

$$\rho_0 - \rho \leq \frac{U_j}{U_{j-1} + \sum_{p=1}^{j-2} \rho^{j-1-p} U_p}.$$

This implies the result since $\sum_{p=1}^{j-2} \rho^{j-1-p} U_p \geq 0$. □

In the following, we will have to control quantities of the shape $\mathbb{P}(\rho_0 - \rho > \varepsilon)$, which will be done by controlling quantities of the shape $\mathbb{P}(U_j/U_{j-1} > \varepsilon)$, using arguments based on the previous lemma. This will be achieved through the use of the following Remark and Lemma.

Remark 2. *For each $j \geq 2$, the random variables U_{j-1} and U_j are known to be independent and gamma distributed $\Gamma(a_{j-1}, b)$ and $\Gamma(a_j, b)$, respectively. It follows that, for all $\varepsilon \geq 0$,*

$$\mathbb{P}\left(\frac{U_j}{U_{j-1}} > \varepsilon \right) = \mathbb{P}\left(\frac{U_j}{U_{j-1} + U_j} > \frac{\varepsilon}{1+\varepsilon} \right)$$

where the random variable $U_j/(U_{j-1} + U_j)$ is beta distributed $\mathcal{B}(a_j, a_{j-1})$ (standard property of gamma distributions), with p.d.f.

$$f_{a_j, a_{j-1}}(t) = \frac{1}{\mathcal{B}(a_j, a_{j-1})} t^{a_j - 1} (1-t)^{a_{j-1} - 1}, \ \forall t \in [0,1]. \tag{4.8}$$

Lemma 4. *Let us denote by $I_x(\alpha_1, \alpha_2)$ the cumulative density function of the beta distribution $\mathcal{B}(\alpha_1, \alpha_2)$ with positive parameters α_1 and α_2 (which is also called the regularized incomplete beta function). For all $x \in [0,1]$,*

$$I_x(\alpha_1, \alpha_2) \geq \frac{x^{\alpha_1} (1-x)^{\alpha_2}}{1 + \frac{\alpha_1}{\alpha_2}}. \tag{4.9}$$

Proof. Let us first show that

$$I_x(\alpha_1,\alpha_2) \geq \frac{x^{\alpha_1}(1-x)^{\alpha_2}}{\alpha_1 \mathcal{B}(\alpha_1,\alpha_2)} \tag{4.10}$$

for all $x \in [0,1]$.

Note that Inequality (4.10) can be seen as a direct consequence of [14, Eq. 8.17.20]. As it is stated without proof in the quoted reference, we prefer to propose some details here.

For (α_1,α_2) fixed, let us set

$$g(x) = I_x(\alpha_1,\alpha_2) - \frac{x^{\alpha_1}(1-x)^{\alpha_2}}{\alpha_1 \mathcal{B}(\alpha_1,\alpha_2)}, \forall x \in [0,1].$$

Based on the p.d.f. of a beta distribution recalled in (4.8), it is easy to check that

$$g'(x) = \frac{x^{\alpha_1}(1-x)^{\alpha_2-1}}{\mathcal{B}(\alpha_1,\alpha_2)}\left(1+\frac{\alpha_2}{\alpha_1}\right) \geq 0.$$

Thus $g(x)$ is non decreasing with respect to x. As $g(0)=0$, we derive that $g(x) \geq 0$ for all $x \in [0,1]$ and Inequality (4.10) is true.

It remains to show that

$$\frac{1}{\alpha_1 \mathcal{B}(\alpha_1,\alpha_2)} \geq \frac{1}{1+\frac{\alpha_1}{\alpha_2}} \tag{4.11}$$

to derive (4.9). Now, [2, Eq. 6.1.3 p. 255] states that for all positive real number α, the inverse of $\Gamma(\alpha)$ can be expressed as

$$\Gamma(\alpha)^{-1} = \alpha \exp(\gamma\alpha) \prod_{m\geq 1}\left[\left(1+\frac{\alpha}{m}\right)\exp\left(-\frac{\alpha}{m}\right)\right],$$

where γ is Euler's constant. By definition of the Beta function, we hence have

$$\begin{aligned}\mathcal{B}(\alpha_1,\alpha_2) &= \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} \\ &= \frac{\alpha_1+\alpha_2}{\alpha_1\alpha_2}\left[\frac{\prod_{m\geq 1}\left[\left(1+\frac{\alpha_1+\alpha_2}{m}\right)\exp\left(-\frac{\alpha_1+\alpha_2}{m}\right)\right]}{\prod_{m\geq 1}\left[\left(1+\frac{\alpha_1}{m}\right)\exp\left(-\frac{\alpha_1}{m}\right)\right]\prod_{m\geq 1}\left[\left(1+\frac{\alpha_2}{m}\right)\exp\left(-\frac{\alpha_2}{m}\right)\right]}\right].\end{aligned}$$

As the products are convergent, this can be simplified into

$$\begin{aligned}\mathcal{B}(\alpha_1,\alpha_2) &= \frac{\alpha_1+\alpha_2}{\alpha_1\alpha_2}\prod_{m\geq 1}\left[\frac{\left(1+\frac{\alpha_1+\alpha_2}{m}\right)}{\left(1+\frac{\alpha_1}{m}\right)\left(1+\frac{\alpha_2}{m}\right)}\right] \\ &\leq \frac{1}{\alpha_1}+\frac{1}{\alpha_2},\end{aligned}$$

which provides (4.11) and the result. □

Corollary 1. *Let $2 \leq j \leq n$. Then $\rho_0 - M_j \leq U_j/U_{j-1}$ and*

$$\mathbb{P}(\rho_0 - M_j > \varepsilon) \leq \mathbb{P}\left(\frac{U_j}{U_{j-1}} > \varepsilon\right) \leq 1 - \frac{\tilde{\varepsilon}^{a_j}(1-\tilde{\varepsilon})^{a_{j-1}}}{1+\frac{a_j}{a_{j-1}}}$$

for all $\varepsilon \in (0,1)$, *with* $\tilde{\varepsilon} = \varepsilon/(1+\varepsilon)$.

Proof. By definition of M_j, we know that $g_j(M_j, \mathbf{Y}) \geq 0$. Based on Lemma 3, we derive that $\rho_0 - M_j \leq U_j/U_{j-1}$. Hence:

$$\mathbb{P}(\rho_0 - M_j > \varepsilon) \leq \mathbb{P}\left(\frac{U_j}{U_{j-1}} > \varepsilon\right).$$

Now, a direct consequence from Remark 2 and Lemma 4 can be written as follows:

$$\mathbb{P}\left(\frac{U_j}{U_{j-1}} > \varepsilon\right) = 1 - I_{\tilde{\varepsilon}}(a_j, a_{j-1}) \leq 1 - \frac{\tilde{\varepsilon}^{a_j}(1-\tilde{\varepsilon})^{a_{j-1}}}{1 + \frac{a_j}{a_{j-1}}}$$

with $\tilde{\varepsilon} = \varepsilon/(1+\varepsilon)$, which allows to conclude. □

All the previous results are valid without any assumption on the shape function $a(\cdot)$. We now come to specific technical results, which requires some concavity assumption for $a(\cdot)$ to hold. To be more precise, our main assumption states as follows.

Assumption (H) The shape function $a(\cdot)$ of the gamma process is concave and differentiable on $\mathbb{R}^+$, and such that $\lim_{t\to\infty} a'(t) = 0$.

Remark 3. *All the asymptotic results of the paper remain valid if the concavity and differentiability properties for* $a(\cdot)$ *only hold from one point* t_0 *(that is on a set* $[t_0, +\infty)$*), and not on the whole half-line* $\mathbb{R}_+$.

Remark 4. *Many classical concave shape functions fulfills Assumption (H), such as*

$$a_1(t) = \alpha t^\beta,\ a_2(t) = \log\left(1+\alpha t^\beta\right) \quad or \quad a_3(t) = 1 - \exp\left(-\alpha t^\beta\right),$$

with $\alpha > 0$ *and* $0 < \beta < 1$ *for the first case and* $0 < \beta \leq 1$ *in the other two cases.*

When $\beta > 1$, *both shape functions* $a_2(\cdot)$ *and* $a_3(\cdot)$ *are concave only from the point* t_0, *with* $t_0 = [(\beta-1)/\alpha]^{1/\beta}$ *and* $t_0 = [(\beta-1)/(\alpha\beta)]^{1/\beta}$, *respectively. Hence, as stated in Remark 3, all the asymptotic results remain valid for* $a(\cdot) = a_2(\cdot)$ *or* $a_3(\cdot)$ *and for all* $\beta > 0$.

Nevertheless, for the sake of simplicity, the shape function $a(\cdot)$ *is assumed to be concave and differentiable from the initial time, in all the results requiring Assumption (H) to hold.*

Lemma 5. *Suppose Assumption (H) to hold. Then the sequence* $(a_n)_{n\in\mathbb{N}^*}$ *is non increasing, and tends to 0 when* n *tends to* ∞.

Proof. By the mean value theorem, there exists $c(n)$ in $((n-1)T, nT)$ such that

$$a_n = a(nT) - a((n-1)T) = T\,a'(c(n)).$$

As $\lim_{t\to\infty} a'(t) = 0$ by assumption and $\lim_{n\to\infty} c(n) = \infty$, it follows that $\lim_{t\to\infty} a'(c(n)) = 0$, which induces the convergence of $(a_n)_{n\in\mathbb{N}^*}$ towards 0. Finally, $a_n \geq a_{n+1}$ for all $n \geq 0$ is a direct consequence of the concavity of $a(\cdot)$. □

Lemma 6. *Suppose Assumption (H) to hold. Then,*

$$\mathbb{P}(\rho_0 - M_{4n} > \varepsilon) \leq \left(1 - \frac{\varepsilon^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right)^n$$

for all $n \geq 1$ *and all* $\varepsilon \in (0,1)$.

Proof. Let $2 \leq k \leq n$. Based on Corollary 1, we know that $\rho_0 - M_k \leq U_k/U_{k-1}$. As $(M_k)_{2\leq k\leq n}$ is non decreasing, we get that

$$\rho_0 - M_n = \min_{2\leq k\leq n} (\rho_0 - M_k) \leq \min_{2\leq k\leq n} \frac{U_k}{U_{k-1}}.$$

Now, in order to boil down to independent random variables U_k/U_{k-1}, let us consider only the even terms $k = 2j$. Also, for sake of simplification, let us substitute n by $4n$. We get

$$\rho_0 - M_{4n} \leq \min_{2\leq k\leq 4n} \frac{U_k}{U_{k-1}} \leq \min_{1\leq j\leq 2n} \frac{U_{2j}}{U_{2j-1}}.$$

This induces

$$\mathbb{P}\left(\rho_0 - M_{4n} > \varepsilon\right) \leq \mathbb{P}\left(\min_{1\leq j\leq 2n} \frac{U_{2j}}{U_{2j-1}} > \varepsilon\right) = \prod_{j=1}^{2n} \mathbb{P}\left(\frac{U_{2j}}{U_{2j-1}} > \varepsilon\right)$$

since the ratios are independent random variables. Based on Corollary 1 again, we derive that

$$\mathbb{P}\left(\rho_0 - M_{4n} > \varepsilon\right) \leq \prod_{j=1}^{2n} \left[1 - \frac{\tilde{\varepsilon}^{\,a_{2j}} (1-\tilde{\varepsilon})^{a_{2j-1}}}{1 + \frac{a_{2j}}{a_{2j-1}}}\right] \tag{4.12}$$

where $\tilde{\varepsilon} = \varepsilon/(1+\varepsilon)$. Under Assumption (H), we know from Lemma 5 that $a_{2j} \leq a_{2j-1}$ for all $1 \leq j \leq 2n$. Thus

$$\frac{1}{1 + \frac{a_{2j}}{a_{2j-1}}} \geq \frac{1}{2}$$

and

$$\begin{aligned}\prod_{j=1}^{2n} \left[1 - \frac{\tilde{\varepsilon}^{\,a_{2j}} (1-\tilde{\varepsilon})^{a_{2j-1}}}{1 + \frac{a_{2j}}{a_{2j-1}}}\right] &\leq \prod_{j=1}^{2n} \left[1 - \frac{1}{2}\tilde{\varepsilon}^{\,a_{2j}} (1-\tilde{\varepsilon})^{a_{2j-1}}\right] \\ &\leq \prod_{j=n+1}^{2n} \left[1 - \frac{1}{2}\tilde{\varepsilon}^{\,a_{2j}} (1-\tilde{\varepsilon})^{a_{2j-1}}\right] \end{aligned} \tag{4.13}$$

(as each term is smaller than 1 in the product).

Using again the non increasingness of $(a_j)_{n+1\leq j\leq 2n}$, we get that $\tilde{\varepsilon}^{a_{2j}} \geq \tilde{\varepsilon}^{a_{2(n+1)}}$ and $(1-\tilde{\varepsilon})^{a_{2j-1}} \geq (1-\tilde{\varepsilon})^{a_{2n+1}}$ for all $j \in \{n+1, n+2, \ldots, 2n\}$, since $\tilde{\varepsilon} \in (0,1)$. This implies

$$\prod_{j=n+1}^{2n} \left[1 - \frac{1}{2}\tilde{\varepsilon}^{\,a_{2j}} (1-\tilde{\varepsilon})^{\,a_{2j-1}}\right] \leq \left(1 - \frac{1}{2}\tilde{\varepsilon}^{\,a_{2(n+1)}} (1-\tilde{\varepsilon})^{\,a_{2n+1}}\right)^n. \tag{4.14}$$

Putting together (4.12), (4.13) and (4.14) leads to

$$\begin{aligned}\mathbb{P}\left(\rho_0 - M_{4n} > \varepsilon\right) &\leq \left(1 - \frac{1}{2}\tilde{\varepsilon}^{\,a_{2(n+1)}} (1-\tilde{\varepsilon})^{\,a_{2n+1}}\right)^n \\ &= \left(1 - \frac{\varepsilon^{\,a_{2(n+1)}}}{2\,(1+\varepsilon)^{\,a_{2n+1}+a_{2(n+1)}}}\right)^n\end{aligned}$$

by definition of $\tilde{\varepsilon}$. Finally, because $1/(1+\varepsilon) > 1/2$ (as $\varepsilon < 1$), we have

$$\mathbb{P}(\rho_0 - M_{4n} > \varepsilon) \leq \left(1 - \frac{\varepsilon^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right)^n,$$

which finishes the proof. □

We are now ready to state our main results, which is done in next subsection.

4.3.3 Consistency and convergence rates

Theorem 1. *Suppose Assumption (H) to hold. Then M_n is a strongly consistent estimator of ρ_0 ($M_n \longrightarrow \rho_0$ almost surely) as the number of repairs n tends to infinity.*

Proof. First, because $(M_n)_{n\geq 1}$ is non decreasing, it is sufficient to prove the almost sure convergence of the subsequence $(M_{4n})_{n\geq 1}$ towards ρ_0. (The same remark is valid for the convergence rate, hereafter). Let $\varepsilon \in (0,1)$. From Lemma 6, we know that

$$\sum_{n\geq 1} \mathbb{P}(\rho_0 - M_{4n} > \varepsilon) \leq \sum_{n\geq 1} \left(1 - \frac{\varepsilon^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right)^n. \tag{4.15}$$

Hence, it is enough to show the convergence of the right-side series in the previous inequality, to show the strong consistency.

Under Assumption (H) and by Lemma 5, we have $\lim_{n\to+\infty} a_{2(n+1)} = \lim_{n\to+\infty} a_{2n+1} = 0$. Then

$$\lim_{n\to+\infty} \sqrt[n]{\left(1 - \frac{\varepsilon^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right)^n} = \lim_{n\to+\infty} \left(1 - \frac{\varepsilon^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right) = \frac{1}{2} < 1.$$

The root test ensures the convergence of the right-side series in (4.15), which allows to conclude. □

We now look at the convergence rate, which reveals itself to be very high (at least sub-exponential, or even exponential).

Theorem 2. *Suppose Assumption (H) to hold. Then we have:*

1. *The almost sure convergence rate of the estimator M_n is at least sub-exponential (that is at least polynomial of order k, for any $k > 0$) as soon as $a_{2n} = O\left((\log n)^{-1}\right)$.*
2. *The almost sure convergence rate is at least exponential as soon as $a_{2n} = O\left(n^{-1}\right)$.*
3. *The convergence rate in probability is at least exponential as soon as $a_{2n} = o\left(n^{-1}\log n\right)$.*

Proof. Let $\varepsilon_n \in (0,1)$ for all $n \in \mathbb{N}^*$. Based on Lemma 6, we have

$$\mathbb{P}(\rho_0 - M_{4n} > \varepsilon_{4n}) \leq \left(1 - \frac{\varepsilon_{4n}^{a_{2(n+1)}}}{2^{1+a_{2n+1}+a_{2(n+1)}}}\right)^n.$$

As $a_{2n+1} \leq a_{2(n+1)} \leq a_1$ from Lemma 5, we get that $2^{1+a_{2n+1}+a_{2(n+1)}} \leq 2^{1+2a_1}$ and hence

$$\mathbb{P}(\rho_0 - M_{4n} > \varepsilon_{4n}) \leq (1-u_n)^n \tag{4.16}$$

with $u_n = C\ \varepsilon_{4n}^{a_{2(n+1)}}$ and $C = 1/\left(2^{1+2a_1}\right)$. From the root test, we know that the series with generic term $(1-u_n)^n$ is convergent as soon as $\limsup_{n\to+\infty}(1-u_n) < 1$, or equivalently as soon as $\liminf_{n\to\infty} u_n > 0$. Hence, the series with generic term the left-side expression in (4.16) is convergent as soon as $\liminf_{n\to\infty} u_n > 0$.

Let us now look at the three different points of the theorem. Assume first that $\varepsilon_n = \varepsilon/n^k$ with $k > 0$ and $\varepsilon \in (0,1)$. This provides

$$u_n = C\ \left(\frac{\varepsilon}{(4n)^k}\right)^{a_{2(n+1)}} = C\ \exp\left\{a_{2(n+1)}\left[\log(\varepsilon) - k\log(4) - k\log(n)\right]\right\}.$$

Assume further that $a_{2n} = O\left((\log n)^{-1}\right)$, or equivalently that $a_{2(n+1)} = O\left((\log n)^{-1}\right)$. Then, there exists $K > 0$ such that $a_{2(n+1)}\log(n) < K$, from where we derive that

$$u_n > C\ \exp\left\{a_{2(n+1)}\left(\log(\varepsilon) - k\log(4)\right) - k\ K\right\}.$$

Hence

$$\liminf_{n\to\infty} u_n \geq C \liminf_{n\to\infty}\ \exp\left\{-k\ K\right\} > 0$$

because $a_{2(n+1)}$ converges towards 0 (see Lemma 5).

This shows that the series with generic term the left-side expression in (4.16) is convergent for $\varepsilon_n = \varepsilon/n^k$ and any (k,ε), which means that M_n almost surely converges towards ρ_0 at speed at least n^{-k} for any $k > 0$, namely the convergence rate is at least sub-exponential, which proves the first point.

Now let us set $\varepsilon_n = \varepsilon\exp(-kn)$ with $k > 0$ and $\varepsilon \in (0,1)$.

We have

$$\begin{aligned} u_n &= C\ \varepsilon^{\,a_{2(n+1)}}\exp\left(-4kna_{2(n+1)}\right) \\ &= C\ \exp\left(na_{2(n+1)}\left(\frac{\log(\varepsilon)}{n} - 4k\right)\right). \end{aligned}$$

Assume that $a_{2n} = O\left(n^{-1}\right)$. Then, there exists $K > 0$ such that $na_{2(n+1)} < K$, from where we derive that

$$u_n > C\ \exp\left(K\left(\frac{\log(\varepsilon)}{n} - 4k\right)\right).$$

Hence

$$\liminf_{n\to\infty} u_n \geq C\ \exp(-4Kk) > 0,$$

which allows to conclude for the second point.

Finally, assume that $a_{2n} = o\left(n^{-1}\log n\right)$. The point here is to show the convergence in probability. Based on (4.16), it is sufficient to show that $\lim_{n\to+\infty}(1-u_n)^n = 0$.

We have

$$(1-u_n)^n = \exp\left\{n\log\left[1 - C\ \varepsilon^{\,a_{2(n+1)}}\exp\left(-4kna_{2(n+1)}\right)\right]\right\}.$$

As $\log(1-x) \leq -x$ for all $x \in (0,1)$, we get that

$$\begin{aligned}(1-u_n)^n &\leq \exp\left\{-C\ n\varepsilon^{\,a_{2(n+1)}} \exp\left(-4kna_{2(n+1)}\right)\right\} \\ &= \exp\left(-C\ v_n\right) \end{aligned} \tag{4.17}$$

with

$$v_n = \exp\left\{\log(n)\left[1 + \frac{na_{2(n+1)}}{\log(n)}\left(\frac{\log(\varepsilon)}{n} - 4k\right)\right]\right\}$$

Based on $a_{2n} = o\left(n^{-1}\log n\right)$, we have $\lim_{n\to+\infty} na_{2n}/\log(n) = 0$, which implies

$$\lim_{n\to+\infty}\left[1 + \frac{na_{2(n+1)}}{\log(n)}\left(\frac{\log(\varepsilon)}{n} - 4k\right)\right] = 1$$

and hence v_n tends to ∞. We derive from (4.17) that $(1-u_n)^n$ converges towards 0, which allows to conclude. □

Example 3. *Let $a(t) = \alpha t^\beta$ with $0 < \alpha, \beta < 1$, which is already known to fulfill Assumption (H) from Remark 4. Also, we have*

$$\begin{aligned} a_{2n} &= \alpha T^\beta\left((2n)^\beta - (2n-1)^\beta\right) \\ &= \alpha(2nT)^\beta\left(1 - \left(1 - \frac{1}{2n}\right)^\beta\right) \\ &\underset{n\to+\infty}{\sim} C\frac{1}{n^{1-\beta}} \end{aligned}$$

where $C = \alpha\beta 2^{\beta-1}T^\beta$. It is easy to check that the condition $a_{2n} = O\left((\log n)^{-1}\right)$ from Point 1 in Theorem 2 is satisfied (but not the conditions for the other points). Hence, we can conclude that the almost sure convergence holds with an at least sub-exponential rate.

Example 4. *Let $a(t) = \log\left(1+\alpha t^\beta\right)$ with $\alpha > 0$, $0 < \beta \leq 1$, which is already known to fulfill Assumption (H) from Remark 4. Also, based on $\log(x) \sim x - 1$ when $x \to 1$ for the second line, we have*

$$\begin{aligned} a_{2n} &= \log\left(\frac{1+\alpha T^\beta(2n)^\beta}{1+\alpha T^\beta(2n-1)^\beta}\right) \\ &\underset{n\to+\infty}{\sim} \alpha T^\beta\frac{(2n)^\beta - (2n-1)^\beta}{1+\alpha T^\beta(2n-1)^\beta} \\ &\underset{n\to+\infty}{\sim} \frac{1}{(1-\frac{1}{2n})^\beta} - 1 \\ &\underset{n\to+\infty}{\sim} \frac{\beta}{2n}. \end{aligned}$$

Hence, the condition $a_{2n} = O\left(n^{-1}\right)$ is satisfied (strongest condition in Theorem 2), and the almost sure convergence holds with an at least exponential rate. Note that this result would remain valid for $\beta > 1$ as the shape function is concave from point $t_0 = [(\beta-1)/\alpha]^{1/\beta}$ (see Remark 4).

Example 5. *Let $a(t) = 1 - \exp\left(-\alpha t^\beta\right)$ with $\alpha > 0$, $0 < \beta \leq 1$, which is already known to fulfill Assumption (H) from Remark 4. We have*

$$a_{2n} = \exp\left(-\alpha\left(2nT\right)^\beta\right) - \exp\left(-\alpha\left(2n-1\right)T^\beta\right),$$

which clearly implies $a_{2n} = O\left(n^{-1}\right)$. *Hence the almost sure convergence holds with an at least exponential rate. Note that, here again, the result would remain valid for* $\beta > 1$ *as the shape function is concave from point* $t_0 = [(\beta - 1)/(\alpha\beta)]^{1/\beta}$ *(see Remark 4).*

Up to here, it was supposed that one single system is observed. In the next section, we now envision the possibility of observing several systems.

4.4 Extension to the case where several systems are observed

4.4.1 Extended semiparametric estimator

In this section, s identical and independent systems are considered. They share the same intrinsic deterioration and ARD_1 repair model with parameter $(a(\cdot), b, \rho_0)$ and they are all observed at times T^-, $2T^-$, ..., nT^-, as described in Section 4.2. For each $i \in \{1, \ldots, s\}$, we add exponent (i) to each quantity referring to the i-th system. For instance, $\mathbf{Y^{(i)}} = \left(Y_1^{(i)}, \ldots, Y_n^{(i)}\right)$ stands for the multivariate observation of the i-th system at times T^-, $2T^-$, ..., nT^-. Also, the sequence $\left(M_j^{(i)}\right)_{1 \leq j \leq n}$ is defined by

$$\begin{cases} M_1^{(i)} = 0 \\ M_j^{(i)} = \inf\left\{\rho \in \left[M_{j-1}^{(i)}, 1\right) \,:\, g_j\left(\rho, \mathbf{Y^{(i)}}\right) \geq 0\right\} & \text{for all } 2 \leq j \leq n \end{cases}$$

in a similar way as in (4.7).

The extended semiparametric estimator is defined as

$$M_{s,n} = \max_{1 \leq i \leq s}\left(M_n^{(i)}\right)$$

for all $n \leq 1$ and $s \geq 1$.

The asymptotic properties of each sequence $\left(M_n^{(i)}\right)_{n \in \mathbb{N}^*}$ (with i fixed) has been studied in the previous section. Clearly, similar results are valid for the sequence $(M_{s,n})_{n \in \mathbb{N}^*}$ with s fixed (with an even higher rate of convergence as $M_n^{(i)} \leq M_{s,n} \leq \rho_0$ for each i). We hence focus on the asymptotic properties of $(M_{s,n})_{s \in \mathbb{N}^*}$ with n fixed in the sequel of this section. We take $n \geq 2$, which ensures the identifiability, based on Proposition 6.

4.4.2 Consistency and convergence rates according to the number of observed systems

Theorem 3. *Let* $n \geq 2$. *Then* $M_{s,n}$ *is a strongly consistent estimator of* ρ_0 *as the number of observed systems* s *tends to infinity.*

Proof. From Proposition 7, we know that $M_2^{(i)} \leq M_n^{(i)} \leq \rho_0$ for each $1 \leq i \leq s$, which implies that $M_{s,2} \leq M_{s,n} \leq \rho_0$. Hence, it is enough to prove that $M_{s,2}$ is a strongly consistent estimator of ρ_0. From the definition of $M_{s,2}$, we have

$$\rho_0 - M_{s,2} = \min_{1 \leq i \leq s}\left(\rho_0 - M_2^{(i)}\right)$$

and as the systems are i.i.d., this provides

$$\mathbb{P}\left(\rho_0 - M_{s,2} > \varepsilon\right) = \mathbb{P}\left(\rho_0 - M_2^{(1)} > \varepsilon\right)^s \tag{4.18}$$

for all $\varepsilon > 0$. Now, Corollary 1 leads to

$$\mathbb{P}\left(\rho_0 - M_{s,2} > \varepsilon\right) \leq \mathbb{P}\left(\frac{U_2^{(1)}}{U_1^{(1)}} > \varepsilon\right)^s$$

and because $\mathbb{P}\left(U_2^{(1)}/U_1^{(1)} > \varepsilon\right) < 1$, we get

$$\sum_{s\geq 1} \mathbb{P}\left(\frac{U_2^{(1)}}{U_1^{(1)}} > \varepsilon\right)^s < \infty.$$

Thus, $M_{s,2}$ tends towards ρ_0 almost surely, which proves the result. □

Theorem 4. *Let $n \geq 2$. The almost sure convergence rate of the estimator $M_{s,n}$ (with respect to s) is at least s^{-k}, for all positive real number k such that*

$$k < \frac{1}{\min\limits_{2\leq p\leq n} a_p}.$$

Proof. Let us set $\varepsilon_s = \varepsilon/s^k$, with $k > 0$ and $\varepsilon \in (0,1)$. The point is to show that the series with generic term $\mathbb{P}\left(\rho_0 - M_{s,n} > \varepsilon_s\right)$ converges for all $k < 1/\min_{2\leq p\leq n} a_p$. Using a similar procedure as for Equation (4.18) and based on Corollary 1, we have[2]

$$\mathbb{P}\left(\rho_0 - M_{s,n} > \varepsilon_s\right) = \mathbb{P}\left[\min_{1\leq i\leq s}\left(\rho_0 - M_n^{(i)}\right) > \varepsilon_s\right] \leq \left(\prod_{p=2}^{n} P_p(s)\right)^s \tag{4.19}$$

where

$$P_p(s) = 1 - \frac{\varepsilon_s^{\,a_p}}{\left(1+\frac{a_p}{a_{p-1}}\right)(1+\varepsilon_s)^{a_{p-1}+a_p}}.$$

Note that $P_p(s) \in (0,1)$ for any $s \geq 1$ and $2 \leq p \leq n$ so that the product in (4.19) is smaller than each of its term. Hence, keeping only the $2p$-th term, we get

$$\mathbb{P}\left(\rho_0 - M_{s,2n} > \varepsilon_s\right) \leq \left(P_{2p}(s)\right)^s$$

for all $p \geq 1$ such that $2p \leq n$. Then, using that $\varepsilon_s = \varepsilon s^{-k}$ and $1/(1+\varepsilon_s) > 1/2$, we obtain

$$\left(P_{2p}(s)\right)^s \leq \left(1 - \frac{s^{-ka_{2p}}}{\left(1+\frac{a_{2p}}{a_{2p-1}}\right)2^{\,a_{2p-1}+a_{2p}}}\varepsilon^{\,a_{2p}}\right)^s.$$

Now, using that $\log(1-x) \leq -x$ for all x in $[0,1)$, it follows that $\left(P_{2p}(s)\right)^s \leq u_s$, where

$$u_s = \exp\left(-C_p\, s^{1-ka_{2p}}\right) \text{ and } C_p = \varepsilon^{\,a_{2p}}\left(\left(1+\frac{a_{2p}}{a_{2p-1}}\right)2^{\,a_{2p-1}+a_{2p}}\right)^{-1}.$$

Gathering the previous inequalities, we now have $\mathbb{P}\left(\rho_0 - M_{s,n} > \varepsilon_s\right) \leq u_s$ and the point is to study the

[2] To be more precise, we have $\mathbb{P}\left[\min\limits_{1\leq i\leq s}\left(\rho_0 - M_n^{(i)}\right) > \varepsilon_s\right] = \mathbb{P}\left(\rho_0 - M_n^{(1)} > \varepsilon_s\right)^s \leq \mathbb{P}\left(\min\limits_{2\leq p\leq n}\frac{U_p}{U_{p-1}} > \varepsilon_s\right)^s$, which leads to the result from Corollary 1.

convergence of the series with generic term u_s. If $1 - ka_{2p} \leq 0$, then u_s converges towards 1 or $\exp(-C_p)$ (if $ka_{2p} = 1$), and the series is divergent. If $1 - ka_{2p} > 0$, then

$$\lim_{s \to +\infty} s^2 u_s = 0$$

and $u_s \underset{s \to +\infty}{=} o\left(\frac{1}{s^2}\right)$, which entails that the series with generic term u_s is convergent, and hence also the series with generic term $\mathbb{P}\left(\rho_0 - M_{s,n} > \varepsilon_s\right)$. This allows to conclude that the almost sure convergence rate of the estimator $M_{s,n}$ is at least s^{-k}, for any $k < a_{2p}^{-1}$ and any $p \geq 1$ such that $2p \leq n$.

Now, keeping only the $2p+1$-th term in the product in (4.19) provides

$$\mathbb{P}\left(\rho_0 - M_{s,2n} > \varepsilon_s\right) \leq \left(P_{2p+1}(s)\right)^s$$

for any p such that $2p + 1 \leq n$. Similar arguments as above allow to derive that the convergence rate it at least s^{-k}, for any $k < a_{2p+1}^{-1}$ and any $p \geq 1$ such that $2p + 1 \leq n$. Finally, it is hence true for any $k < a_p^{-1}$ and any p such that $2 \leq p \leq n$, and consequently for any $k < \max_{2 \leq p \leq n}\left(a_p^{-1}\right)$, which allows to conclude. □

The real number a_p corresponds to the increment of the shape function over the time interval $[(p-1)T, pT)$. The overall convergence rate obtained in the previous theorem corresponds to the smallest increment. Hence the smaller this increment, the slower the degradation and the higher the convergence rate. More precisely, when the shape function is concave, the increments decrease over time and the convergence rate is at least s^{-1/a_n} because the smallest increment is the last one. On the other hand, the increments increase over time when the shape function is convex, hence the smallest increment is the second one and the convergence rate is at least s^{-1/a_2}. Then the convergence rate is higher than the standard square-root speed as soon as the smallest increment is less than 2. Note that this condition depends on both the shape function and the period T, as illustrated in the next example.

Example 6. *Let $a(t) = \alpha t$ with $\alpha > 0$, hence $a_p = \alpha T$ for each $2 \leq p \leq n$ and the almost sure convergence rate of $M_{s,n}$ with respect to s is at least $s^{-1/\alpha T}$. In comparison with the classical rate $s^{-1/2}$, it is higher if $\alpha T < 2$ and lower if $\alpha T > 2$.*

4.5 Empirical illustration based on simulated data

The aim of this section is to illustrate our most significant results, that is the fast convergence rates obtained in Section 4.3 in the case where a single system is observed. In that case asymptotic results are obtained with respect to an increasing number of repairs n. The point hence is to observe the empirical behavior of the semiparametric estimator M_n of ρ_0, which from Theorems 1 and 2 is known to be strongly consistent, with a convergence rate that can be either exponential (ECR) or sub-exponential (S-ECR) with respect to n (considering either almost sure convergence or convergence in probability).

These illustrations are based on two simulated datasets, considering $a(t) = \log(1+t)$ (first case) and $a(t) = \sqrt{t}$ (second case) as shape functions, which from Examples 4 and 3 provide exponential and sub-exponential almost sure convergence rates respectively. The first (resp. second) case will hence be referred to as ECR (resp. S-ECR) in the sequel.

Data simulation and empirical bias

To be able to compare results, we place ourselves within the same framework for both cases. First, the model parameters as well as the observation characteristics of a maintained system are the following:

- Shape function: $a(t) = \log(1+t)$ (ECR) and $a(t) = \sqrt{t}$ (S-ECR);
- Scale parameter: $b = 1$;
- Maintenance efficiency parameter: $\rho_0 = 0.5$;
- Period for repairs: $T = 1$;
- Observation times: $\{nT^-; 1 \leq n \leq 250\}$ because the system is maintained 250 times.

Thus for a single maintained system simulated over the time interval [0,250], that is a degradation trajectory over the time interval [0,250], M_j is computed for each observation time (right before the repair time) providing a realization $(m_1, \ldots, m_{250})$ of the random vector $(M_1, \ldots, M_{250})$. We generate 250 000 i.i.d trajectories, which leads to 250 000 i.i.d. realizations $\left\{\left(m_1^{(i)}, \ldots, m_{250}^{(i)}\right); 1 \leq i \leq 250\,000\right\}$ of $(M_1, \ldots, M_{250})$. In other words, we have 250 000 estimations of ρ_0 at each observation time $T^-, 2T^-, \ldots, 250T^-$. Then the Empirical Bias (EB) is computed at each observation time nT^-as follows

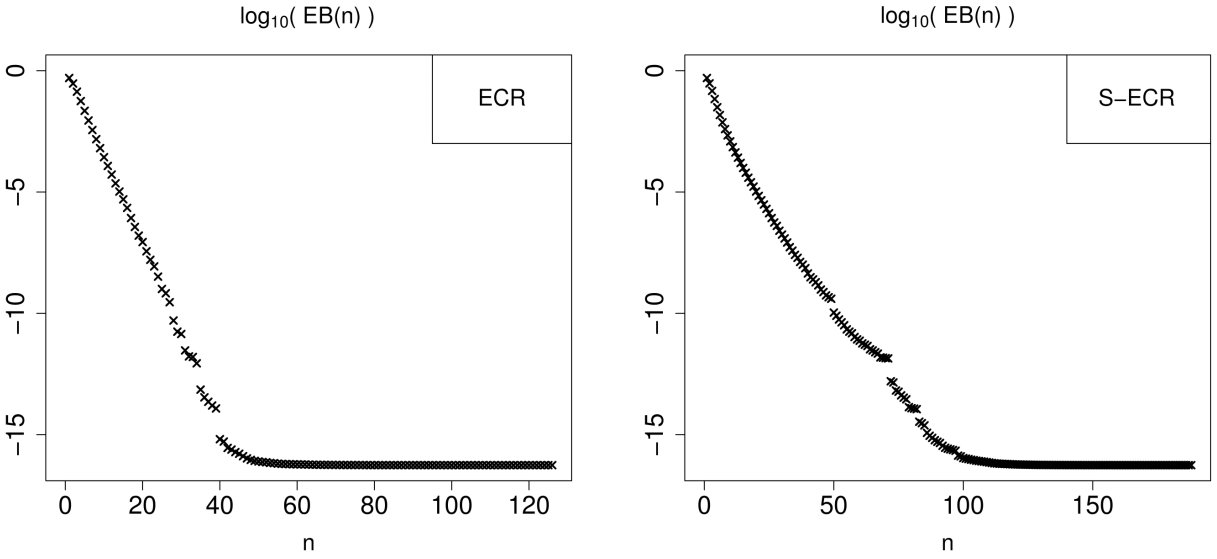

Figure 4.3: Plots of the common logarithm of the Empirical Bias versus n, for the ECR case on the left and the S-ECR case on the right.

$$\mathrm{EB}(n) = \frac{1}{250\,000} \sum_{i=1}^{250\,000} \left(\rho_0 - m_n^{(i)}\right),$$

and the common logarithm of EB(n) for both the ECR and S-ECR cases are reported on Figure 4.3. In both plots, we can observe that

- the empirical bias tends towards a non-zero constant as n increases;
- viewed as a function of n, the empirical bias has jumps.

Our point now is to explain these two particularities, which are induced by computing limitations, as is discussed hereafter.

Let M_{num} be the largest positive real number such that $1 - M_{\text{num}} = 1$ (numerically negligible), which is roughly equal to 1.11×10^{-16} in our case. This entails that $(1 - M_{\text{num}})/2 = 1/2$, and hence

$$\rho_0 - \frac{M_{\text{num}}}{2} = \rho_0,$$

based on $\rho_0 = 1/2$. Then, when we obtain an estimate $m_n^{(i)} = 0.5$ for some $1 \leq i \leq 250\ 000$ and $1 \leq n \leq 250$, it only means that $\rho_0 - m_n^{(i)} \leq M_{\text{num}}/2$ and not that $m_n^{(i)} = \rho_0$. Thus, the numerical estimate of the bias is correct only if $\rho_0 - m_n^{(i)} \geq M_{\text{num}}/2$, otherwise it is underestimated.

This leads us to introduce the Proportion of Numerically Exact Estimates (PNEE) as a function of n

$$\text{PNEE}(n) = \frac{\#\left\{1 \leq i \leq 250\ 000 : \ \rho_0 - m_n^{(i)} = \frac{M_{\text{num}}}{2}\right\}}{250\ 000}.$$

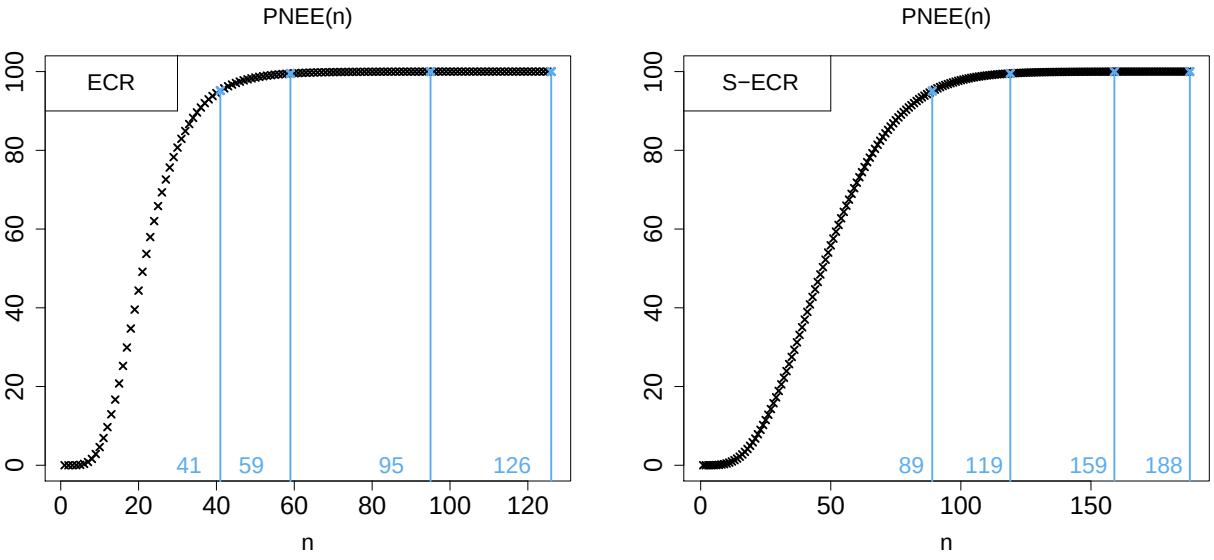

Figure 4.4: Plots of the PNEE as a function of n, for the ECR case on the left and the S-ECR case on the right. Blue vertical lines point to the smallest repair numbers n for which PNEE(n) is larger or equal than 95%, 99.5%, 99.99% and 100% from left to right.

This figure shows that the larger n, the more there are estimates equal to $M_{\text{num}}/2$, which explains why EB(n) tends towards $M_{\text{num}}/2$ instead of 0. Moreover we have $M_{\text{num}}/2 \approx 5.55 \times 10^{-17}$, and so $\log_{10}(M_{\text{num}}/2) \approx -16.25$, which matches the numerical results of Figure 4.3. Note that in both ECR and S-ECR cases, all the estimates are numerically exact before the last observation time $250T^-$, which illustrates a very fast convergence rate of the estimator. Looking carefully at the distribution of the estimates $m_n^{(i)}, i = 1, \ldots, 250\ 000$, for each n, we have observed that there are a (very) few extreme values, which correspond to trajectories for which the convergence is slow when n increases. For each trajectory, the sequence of estimates $m_n^{(i)}, n = 1, \ldots, 250$, is piecewise constant and approximates ρ_0 by below. When a large proportion of sequences has already reached the numerical precision $M_{\text{num}}/2$ (and hence remains constant when n increases), each jump in an extreme trajectory entails a negative jump in the empirical bias, which is not counterbalanced by any positive jump. This explains the negative jumps

observed in Figure 4.3.

As a summary, both jumps and convergence to a non-zero constant in Figure 4.3 are due to numerical limitations and the corresponding points on the plots should not be considered for the further study of the empirical bias. Thus, in the sequel, we only focus on the first values of n for which there is no jump, that is on $1 \leq n \leq 26$ for the ECR case and $1 \leq n \leq 49$ for the S-ECR case.

Our aim now is to explore the convergence rate from an empirical point of view. This can be done through the study of the empirical bias, as is now explained.

Link between the bias and the exponential convergence rate

Let us show that the convergence rate of M_n is at least exponential whenever $\log_{10}\left(\rho_0 - \mathbb{E}\left(M_n\right)\right)$ decreases linearly. Indeed, assuming that $\log_{10}\left(\rho_0 - \mathbb{E}\left(M_n\right)\right)$ decreases linearly, there exist $\tilde{k} > 0$ and $C \in \mathbb{R}$ such that

$$\log_{10}\left(\rho_0 - \mathbb{E}\left(M_n\right)\right) = C - \tilde{k}n.$$

Because $M_n \in [0, \rho_0]$ for any $n \geq 1$ with probability one, the random variable $\rho_0 - M_n$ is non negative and by the Markov's inequality we have

$$\mathbb{P}\left(\rho_0 - M_n > \varepsilon_n\right) \leq \frac{\rho_0 - \mathbb{E}\left(M_n\right)}{\varepsilon_n}.$$

Setting $\varepsilon_n = \varepsilon \exp(-kn)$ with $k \in (0, \tilde{k})$ and $\varepsilon > 0$, we have

$$\mathbb{P}\left(\rho_0 - M_n > \varepsilon_n\right) \leq \frac{\exp(C)}{\varepsilon} \exp\left(-\left(\tilde{k} - k\right) n\right) \underset{n\to\infty}{\longrightarrow} 0,$$

and

$$\sum_{n\geq 1} \mathbb{P}\left(\rho_0 - M_n > \varepsilon_n\right) \leq \sum_{n\geq 1} \frac{\exp(C)}{\varepsilon} \exp\left(-\left(\tilde{k} - k\right) n\right) < +\infty.$$

Therefore the two last results allow to conclude that the rate of convergence of M_n is at least exponential for the convergence in probability as well as for the almost sure convergence. If $\log_{10}\left(\rho_0 - \mathbb{E}\left(M_n\right)\right)$ decreases at a slower rate than the linear rate, we can not conclude to a sub-exponential convergence rate, however we observe that the empirical evidence of a slower convergence rate for $\log_{10}\left(\rho_0 - \mathbb{E}\left(M_n\right)\right)$ coincides with a slower convergence rate for $\rho_0 - M_n$ in Theorem 2.

Link between theoretical and empirical results

Because of the previous explanations on the behavior of $\mathrm{EB}(n)$, linear regressions are performed on $\{\mathrm{EB}(n); 1 \leq n \leq 26\}$ for the ECR case and $\{\mathrm{EB}(n); 1 \leq n \leq 49\}$ for the S-ECR case. We assume that the relationship between $\mathrm{EB}(n)$ and n is modelled by a linear regression either simple or quadratic. The results are summarized on Figure 4.5, Table 4.1 and Table 4.2.

The tables provide the linear regressions outcomes, that is the coefficients of the first and second degree polynomial regression (and their related p-values and indicators about residual error (minimum, maximum, first, second and third quartiles). For the ECR case, both first (simple) and second (quadratic) degree polynomial regression fit well $\log_{10}\left(\mathrm{EB}(n)\right)$, with adjusted R^2 of 0.9994 and 0.9996, respectively. However, Table 4.1 shows that the second degree term is not significant neither in comparison with

Table 4.1: Linear regressions summary for the ECR case

(value / p-value)	Intercept coefficient	First degree coefficient	Second degree coefficient
1^{st} degree	0.10 / 6.12×10^{-4}	-0.36 / $< 2 \times 10^{-16}$	
2^{nd} degree	0.19 / 1.70×10^{-5}	-0.38 / $< 2 \times 10^{-16}$	6.9×10^{-4} / 4×10^{-3}

(Residual error)	Minimum	1^{st} quartile	Median	3^{rd} quartile	Maximum
1^{st} degree	-0.09	-0.05	-0.01	0.04	0.12
2^{nd} degree	-0.14	-0.03	3×10^{-3}	0.04	0.10

Table 4.2: Linear regressions summary for the S-ECR case

(value / p-value)	Intercept coefficient	First degree coefficient	Second degree coefficient
1^{st} degree	-0.95 / 4.79×10^{-13}	-0.19 / $< 2 \times 10^{-16}$	
2^{nd} degree	-0.21 / 4.90×10^{-6}	-0.27 / $< 2 \times 10^{-16}$	1.76×10^{-4} / $< 2 \times 10^{-16}$

(Residual error)	Minimum	1^{st} quartile	Median	3^{rd} quartile	Maximum
1^{st} degree	-0.31	-0.27	-0.14	0.21	0.84
2^{nd} degree	-0.15	-0.06	-5×10^{-3}	0.06	0.22

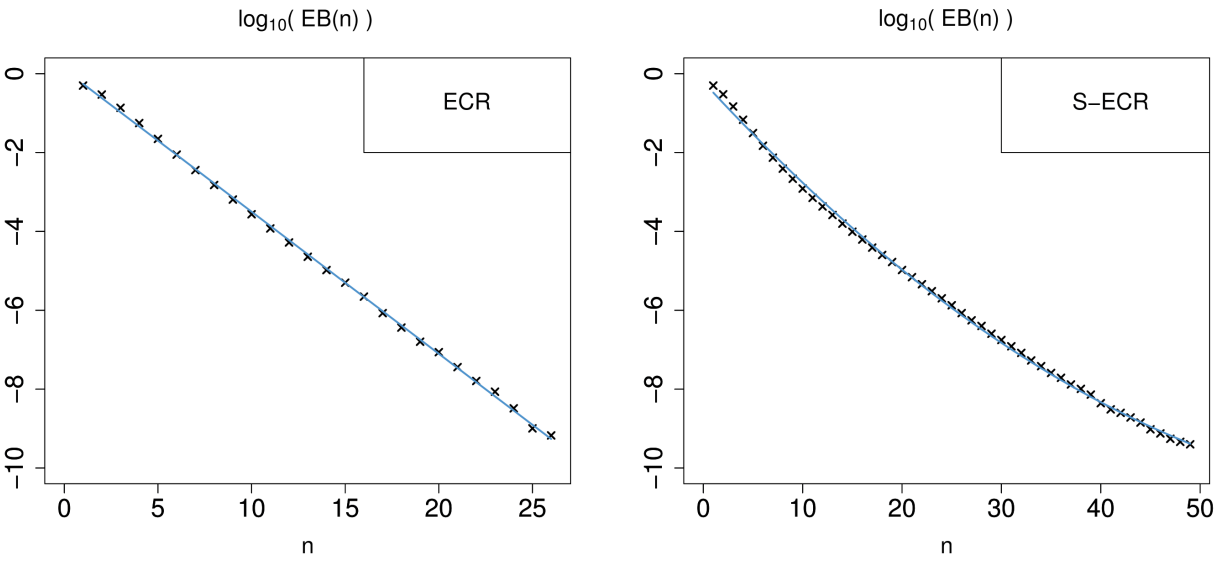

Figure 4.5: Partial plots of the common logarithm of the Empirical Bias versus n, for the ECR case on the left and the S-ECR case on the right. The blue lines correspond to the first degree polynomial regression for the ECR case and to the second degree polynomial regression for the S-ECR case.

the other terms nor for improving the model quality (the adjusted R^2 increases and the residual error decreases). We conclude to a linear decrease of $\log_{10}(\mathrm{EB}(n))$ which induces an at least exponential convergence rate. It is thus consistent with results of Theorem 2 as well as with the related plot in Figure 4.5. Concerning the S-ECR case (see Table 4.2), the addition of the second degree coefficient improves the model quality, especially the adjusted R^2, which goes from 0.9845 for a linear polynomial to 0.9989 for a quadratic one. Hence the quadratic linear regression model is more relevant than the simple lin-

ear regression model, which coincides with the (at least) sub-exponential convergence rate mentioned in Theorem 2.
Regarding the optimality of our results, we recall that the condition of Theorem 2 to obtain an at least exponential convergence rate is $a_{2n} = o\left(n^{-1}\log n\right)$. Now repeating the study with $a_{2n} = n^{-1}\log n$ we see in Figure 4.6 that again an exponential convergence rate is expected because the bias decreases linearly. We conclude that the condition on a_{2n} in Theorem 2 is probably sufficient but not necessary.

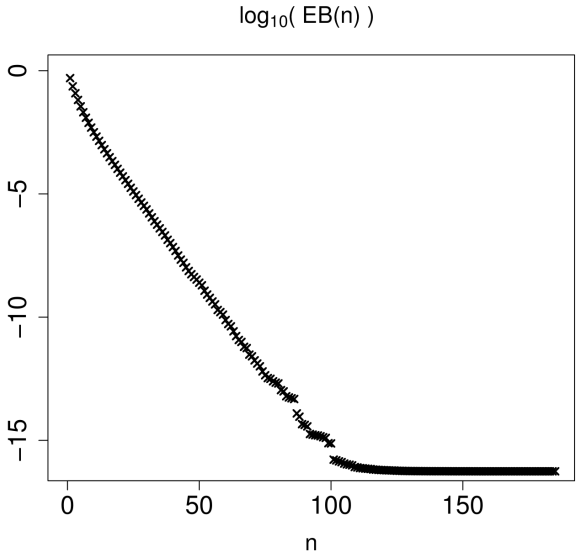

Figure 4.6: Plots of the common logarithm of the Empirical Bias versus n, with $a_{2n} = n^{-1}\log n$

4.6 Concluding remarks

In this paper we propose a semiparametric inference method for the maintenance efficiency parameter involved in the ARD_1 repair model for a Gamma deteriorating system. For a single system the main condition that insures the strong consistency of our semiparametric estimator of the maintenance efficiency parameter is the concavity of the shape function of the underlying Gamma deteriorating process. Two types of asymptotic results are obtained: either a single system is observed with the number of repairs tending to infinity, or it is the number of systems that tends to infinity. In the case of a single system the almost sure convergence rate of the estimator can be particularly fast, at least exponential for some particular cases. The simulation study illustrates the convergence rates obtained in case of a single trajectory. We observe that the theoretical convergence rates are consistent with the numerical simulation results. However it seems that it is still possible to refine the mathematical conditions under which the convergence rates are obtained. Thus improving the assumptions accuracy may constitute further work. Note that when several systems are considered the convergence rate of the estimator is slower but its strong consistency holds whatever the shape function is. Depending on the shape function, the convergence rate of the estimator may overcome the usual square root rate.

In the wake of this study we want to mention several lines of work that we consider as are important. First, the observation scheme could be decoupled from the scheme of maintenance actions. Indeed, it would be interesting, for instance for an ARD_1 model, to consider a system with scheduled times of maintenance actions for which the observation times are independent of the maintenance schedule. Also,

there exist many other models that extend the ARD_1 model, such as for instance the ARD_m (resp. ARD_∞) for which the basic idea is that a maintenance action removes a proportion of the degradation accumulated by the system from the last m maintenance actions (resp. since the system was put into operation).

As explained in the introduction, there exist also alternatives to arithmetic reduction of degradation models such as those based on arithmetic reduction of age. For such models, instead of reducing the degradation level of the system, the maintenance action consists in reducing the age of the system. The use of these models is not restricted to gamma processes, and can be generalized to any non homogeneous Lévy process. As an example, [24] deals with both the ARD_1 and ARD_∞ models, as well as two arithmetic reduction of age models, considering a Wiener process based degradation. Nevertheless, the estimation procedure we developed highly relies on the non negativity of the gamma process and hence could not be adapted to a general non monotonous Lévy process. The adaptation of the estimation procedure of the present paper to another monotonous Lévy process than the gamma process would be interesting to study.

Hence there remain many estimation procedures to be developed for all these imperfect repair models for deteriorating systems, but the semiparametric estimation of the maintenance efficiency for the models mentioned above with various observation schemes, is probably the most challenging problems we aim at investigating in near future.

4.7 Extension of the semiparametric estimation method to the Arithmetic Reduction of Degradation model of order infinity

The purpose of this section is to prove that the semiparametric estimation method studied previously remains valid in the case where the ARD_∞ model is considered.

First of all, **Proposition 6** also holds for the ARD_∞ model, because only the first two observations are necessary to obtain the model identifiability, and both models are the same over $[0, 2T)$. Now let us recall that the degradation level Y_t at any time t such that $(j-1)T \leq t < jT$ can be expressed as

$$Y_t = X_t^{(j)} - X_{(j-1)T}^{(j)} + \sum_{k=1}^{j-1} (1-\rho_0)^{j-k} U_k$$

where $U_k = X_{kT}^{(k)} - X_{(k-1)T}^{(k)}$ for $1 \leq k \leq j-1$, and ρ_0 is the true value of the maintenance actions efficiency. The observation times are still jT^- for $1 \leq j \leq n$, and let us set $Y_j = Y_{jT^-}$. Therefore, we have

$$Y_j = U_j + \sum_{k=1}^{j-1} (1-\rho_0)^{j-k} U_k. \tag{4.20}$$

Note that the degradation level right after the jth maintenance action is $(1-\rho_0)Y_j$, hence

$$Y_j = U_j + (1-\rho_0) Y_{j-1}$$

As a result, for each $1 \leq j \leq n$, the increments U_j can be expressed with respect to the observations

Y_{j-1} and Y_j as well as the maintenance efficiency parameter ρ_0 as follows:

$$U_j = Y_j - (1-\rho_0)Y_{j-1}.$$

Similarly as in **Subsection 4.3.1**, we define here the function g_j as well as the estimator sequence $(M_j)_{1\leq j\leq n}$. It is not necessary to define the set D_j of the possible values for ρ_0 as the sequence $(M_j)_{1\leq j\leq n}$ has an explicit expression here.
Let us set $g_j(\rho, \boldsymbol{Y}) = Y_j - (1-\rho)Y_{j-1}$, hence we have $g_j(\rho_0, \boldsymbol{Y}) = U_j > 0$ a.s. Therefore, the parameter ρ_0 fulfils the condition $g_j(\rho_0, \boldsymbol{Y}) \geq 0$ for all $1 \leq j \leq n$. Let us recall that $g_1(\rho_0, \boldsymbol{Y}) = Y_1 \geq 0$, and for $j = 2, \ldots, n$ the condition $g_j(\rho_0, \boldsymbol{Y}) \geq 0$ is equivalent to $\rho_0 \geq 1 - Y_j/Y_{j-1}$, thus

$$\rho_0 \geq \max_{2\leq j\leq n} \left(1 - \frac{Y_j}{Y_{j-1}}\right).$$

Finally, we define the sequence $(M_j)_{1\leq j\leq n}$ as

$$\begin{cases} M_1 = 0 \\ M_j = \max\left(M_{j-1}\,,\, 1 - \frac{Y_j}{Y_{j-1}}\right) \quad \text{for all } 2 \leq j \leq n \end{cases}$$

which is non-decreasing and bounded by ρ_0 because $\rho_0 \in [M_j, 1)$ for all $1 \leq j \leq n$.
The last step is to prove that **Lemma 3** still holds here. As a result, all the technical results displayed in **Subsection 4.3.2** hold regardless if the order of the ARD model is 1 of ∞, because they only depend on the sequence properties as well as this lemma. Let us now recall the lemma and prove it within the ARD_∞ context.

Lemma 3. *Let us set $\rho_0 \in [0, 1)$, then $g_j(\rho, \boldsymbol{Y}) \geq 0$ implies that $\rho_0 - \rho \leq \frac{U_j}{U_{j-1}}$ for all $2 \leq j \leq n$.*

Proof. Assume $j \geq 2$. We derive from Equation (4.20) that

$$\begin{aligned} 1 - \frac{Y_j}{Y_{j-1}} &= 1 - \frac{U_j + (1-\rho_0)Y_{j-1}}{Y_{j-1}} \\ &= 1 - \frac{U_j}{Y_{j-1}} - (1-\rho_0) \end{aligned}$$

Hence, $\rho_0 - \rho \leq U_j/Y_{j-1}$, and replacing Y_{j-1} by its expression allows to write that

$$\rho_0 - \rho \leq \frac{U_j}{U_{j-1} + (1-\rho_0)Y_{j-2}}$$

which leads to the result since $(1-\rho_0)Y_{j-2} > 0$ a.s.

□

From this, it can be concluded that the main results displayed in **Theorems 1**, **2**, **3** and **4** are verified whether the underlying model is the ARD_1 or the ARD_∞.

Part III

Imperfect repairs based on reduction of age

Chapter 5

Introduction

A system is considered whose intrinsic deterioration is modelled by a gamma process $(X_t)_{t\geq 0}$ with shape function a and scale parameter b. The system is subject to periodic (period T) and instantaneous imperfect maintenance actions, and we place ourselves into the framework of a Arithmetic Reduction of Age model of order one (ARA_1). Within the context of virtual age models, an imperfect repair puts back the system to a similar state as it was before, namely the system is rejuvenated. In the case of the ARA_1 model, a repair removes a proportion $\rho \in [0,1)$ of the age of the system accumulated since the last maintenance action. This model is defined below.

In **Chapter 6**, several estimation methods are developed, starting with the most usual ones, namely the methods of moments and the Maximum Likelihood Estimation (MLE). However, the expression of the joint density of the observations, and consequently the likelihood, is a product of integrals of large dimension, and thus numerical estimations becomes difficult to compute in a classical way. At first, an expectation maximization algorithm, as well as a differential evolution maximization (see [39]) associated with quasi Monte Carlo approximations for the integrals, were developed and tested. However, none of these methods were suitable due to numerical issues and high CPU times. Hence, the MLE is treated here by approximating the integrals by the Monte Carlo and randomized Quasi Monte Carlo methods.

Besides these methods, in order to avoid the problem of high-dimensional integrals, alternative maximum likelihood methods are developed: the composite maximum likelihood and the half data method. The first method consists in assuming that the observations are independent, while the second one only takes into account one out of two observations. This allows to reduce the dimension of the integrals to 1.

After the methods study, some illustrations of their performances are provided in **Chapter 7**.

Let us first define the ARA_1 model. Let $\left(X^{(j)}\right)_{j\in\mathbb{N}^*}$ be a sequence of independent copies of the gamma process $X = (X_t)_{t\geq 0}$, which corresponds to the intrinsic degradation of the system. Let $(Z_t)_{t\geq 0}$ stands for the effective degradation level of the system. In the framework of the ARA_1 model, each maintenance action remove ρT units of time to the virtual age of the system, which can be expressed as follows:

$$V(t,\rho) = t - \rho n T \text{ for } nT \leq t < (n+1)T$$

where $V(t,\rho)$ is the virtual age of the system at real time t. Note that $V(nT,\rho) = (1-\rho)nT$ and $V(nT^-,\rho) = [(1-\rho)n+\rho]T$. Also, the nth maintenance action puts the system back to the state it was ρT units of time before, that is at real time $nT - \rho T = (n-\rho)T$. The related virtual age is

$$V((n-\rho)T,\rho) = (n-\rho)T - \rho(n-1)T = (1-\rho)nT = V(nT,\rho).$$

This is consistent because at real time nT, the system is put back to the exact state it was at real time $(n-\rho)\,T$. This leads to the following expression for the degradation level right after the nth maintenance action:

$$Z_{nT} = Z_{(n-\rho)T} \tag{5.1}$$

and finally the general expression of the virtual age of the system for all $t \geq 0$ is

$$V(t,\rho) = \sum_{j\geq 0} (t - \rho j T)\, \mathbf{1}_{[jT,(j+1)T)}(t).$$

Hence, the degradation level Z_t for $nT < t \leq (n+1)T$, with n in $\mathbb{N}^*$, can be expressed as follows

$$Z_t = Z_{nT} + \left(X^{(n+1)}_{V(t,\rho)} - X^{(n+1)}_{V(nT,\rho)}\right) = Z_{nT} + \left(X^{(n+1)}_{t-\rho nT} - X^{(n+1)}_{(1-\rho)nT}\right)$$

with

$$Z_{nT} = Z_{(n-1)T} + \left(X^{(n)}_{V(nT,\rho)} - X^{(n)}_{V((n-1)T,\rho)}\right) = Z_{(n-1)T} + \left(X^{(n)}_{(1-\rho)nT} - X^{(n)}_{(1-\rho)(n-1)T}\right),$$

and where $\left(X^{(n+1)}_{t-\rho nT} - X^{(n+1)}_{(1-\rho)nT}\right)$ is gamma distributed $\Gamma\big(a(t-\rho nT) - a((1-\rho)nT), b\big)$. Note that Z_{nT} can be written as

$$Z_{nT} = \sum_{j=1}^{n} \left(X^{(j)}_{(1-\rho)jT} - X^{(j)}_{(1-\rho)(j-1)T}\right),$$

which is also gamma distributed with shape and scale parameters $a((1-\rho)nT)$ and b respectively. Thus Z_t is the sum of two gamma distributed random variables with the same scale parameter, which entails that Z_t is gamma distributed $\Gamma\big(a(t-\rho nT), b\big)$ for all $nT \leq t < (n+1)T$, and its mean and variance are given by

$$\mathbb{E}(Z_t) = \frac{a(t-\rho nT)}{b}$$

and

$$\mathbb{V}(Z_t) = \frac{a(t-\rho nT)}{b^2}.$$

Based on the same framework as in **Example 1**, **Figure 1.3** represents both the degradation evolution mean and variance of the maintained system, as well as those of the intrinsic deterioration mean (unmaintained system). The scale parameter is here equal to one, hence the mean and the variance are equal as well due to their similar expressions.

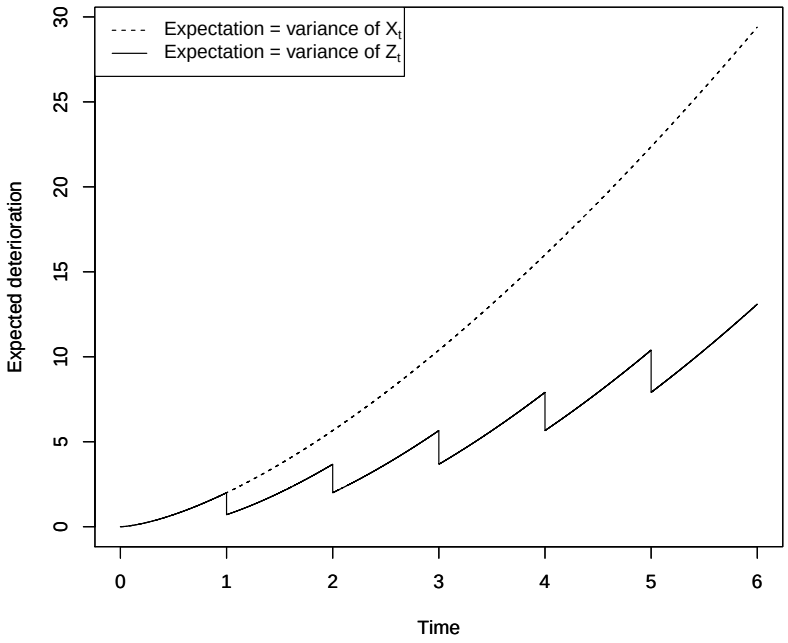

Figure 5.1: Expectations (equal to variances) of X_t and Z_t, where the shape function is such that $a : t \mapsto \alpha t^{\beta}$ with $\alpha = 2$ and $\beta = 3/2$, and with $b = 1$ and $\rho = 1/2$.

Chapter 6

Parametric inference for the Arithmetic Reduction of Age model of order one

6.1 Preliminary

As mentioned in the introduction of this part, several estimation methods are developed in this chapter, in order to estimate the ARA_1 model parameters. The periodicity T is assumed to be known, thus the parameters of interest are the parameters of the shape function, the scale parameter b and the maintenance efficiency ρ. The estimation methods are developed, and then tested in the next chapter, within the framework of a power law shape function. To be more precise, the shape function is defined as $a : t \mapsto \alpha t^{\beta}$ with $\alpha,\ \beta > 0$. We set $\boldsymbol{\theta} = (\alpha, \beta, \rho, b) \in \Theta$ the parameter set, with the parameter space $\Theta = (0,\infty)^2 \times [0,1), \times (0,\infty)$.

We assume that the degradation level is measured right before the first n maintenance actions for some n in $\mathbb{N}^*$, that is at times T^-, $2T^-$,...,nT^-. Also, s i.i.d. systems are considered. For sake of readability, let us define the following notations:

- $Z_j = Z_{jT^-}$ for $1 \leq j \leq n$;
- $\boldsymbol{Z} = \left(Z_{jT^-}\right)_{1\leq j\leq n} = (Z_j)_{1\leq j\leq n}$;
- $z_j^{(i)}$ is the observed degradation level of the ith maintained system at times jT^- for $1 \leq j \leq n$ and $1 \leq i \leq s$, which is a realisation of the r.v. Z_j;
- $\boldsymbol{z^{(i)}} = \left(z_j^{(i)}\right)_{1\leq j\leq n}$ is the complete observations set related to the ith maintained system;
- $\boldsymbol{z}$ is the complete observations set, that is $\boldsymbol{z} = \left(z_j^{(i)}\right)_{1\leq j\leq n,\, 1\leq i\leq s}$;
- $\boldsymbol{\xi}$ is the parameter set excluding b, that is $\boldsymbol{\xi} = (\alpha, \beta, \rho)$.

Note that the Z_j's are mutually dependent. Hence, in the next sections, some methods will be developed considering the observations difference (increments), that is $\boldsymbol{\Delta Z} = (Z_j - Z_{j-1})_{1\leq j\leq n}$, with $Z_0 = 0$ by assumption on the initial deterioration level, and these random variables $Z_j - Z_{j-1}$ are only pairwise dependent.

In the following, for sake of simplicity, the same notations are kept (Δ is omitted), that is

- $Z_j = Z_{jT^-} - Z_{(j-1)T^-}$ for $1 \leq j \leq n$;
- $\boldsymbol{Z} = \left(Z_{jT^-} - Z_{(j-1)T^-}\right)_{1\leq j\leq n} = (Z_j)_{1\leq j\leq n}$;
- $z_j^{(i)}$ is the observed increments of the degradation of the ith maintained system between the instants right before the $j-1$th and the jth maintenance for $1 \leq j \leq n$ and $1 \leq i \leq s$, which is a realisation of the r.v. Z_j;
- $\boldsymbol{z^{(i)}} = \left(z_j^{(i)}\right)_{1\leq j\leq n}$;
- $\boldsymbol{z} = \left(z_j^{(i)}\right)_{1\leq j\leq n,\, 1\leq i\leq s}$.

We provide another representation of such transformed observations in the lemma below.

Lemma 7. *Let Z_j be the degradation increment between the instants right before the $j-1$th and the jth maintenance. For all $1 \leq j \leq n$, the r.v. Z_j can be expressed with respect to independent and gamma distributed random variables as follows:*

$$Z_j = U_j + V_j - V_{j-1}$$

with the convention $V_0 = 0$ and where

- *$U_1, \ldots, U_n, V_1, \ldots, V_n$ are independent;*
- *$U_j \sim \Gamma\left(\mu_j\left(\boldsymbol{\xi}\right), b\right)$ for $1 \leq j \leq n$;*
- *$V_j \sim \Gamma\left(\nu_j\left(\boldsymbol{\xi}\right), b\right)$ for $1 \leq j \leq n$,*

and the shape functions μ_j and ν_j are such that

$$\mu_j\left(\boldsymbol{\xi}\right) = \left[\left(j(1-\rho)\right)^{\beta} - \left((j-1)(1-\rho)\right)^{\beta}\right] \alpha T^{\beta}$$

and

$$\nu_j\left(\boldsymbol{\xi}\right) = \left[\left(j-(j-1)\rho\right)^{\beta} - \left(j(1-\rho)\right)^{\beta}\right] \alpha T^{\beta}.$$

Moreover, because of the independence of U_j and V_j, we also can write Z_j as

$$Z_j = W_j - V_{j-1}$$

where W_j and V_{j-1} are independent, and W_j is gamma distributed $\Gamma\left(\omega_j\left(\boldsymbol{\xi}\right), b\right)$ with

$$\omega_j\left(\boldsymbol{\xi}\right) = \left[\left(j-(j-1)\rho\right)^{\beta} - \left((j-1)(1-\rho)\right)^{\beta}\right] \alpha T^{\beta}.$$

Proof. Let U_j be the increment of the underlying degradation process over the real time interval $\left((j-1)T, (j-\rho)T\right]$. Therefore, we have

$$U_j = Z_{(j-\rho)T} - Z_{(j-1)T}$$

which can be written as

$$U_j = X^{(j)}_{V((j-\rho)T,\rho)} - X^{(j)}_{V((j-1)T,\rho)}.$$

Thus, we derive that U_j is gamma distributed $\Gamma\left(a\left[V\left(\left(j-\rho\right)T,\rho\right)\right] - a\left[V\left(\left(j-1\right)T,\rho\right)\right], b\right)$, which matches the definition of U_j as

$$\begin{aligned} a\left[V\left(\left(j-\rho\right)T,\rho\right)\right] - a\left[V\left(\left(j-1\right)T,\rho\right)\right] &= a\left[\left(1-\rho\right)jT\right] - a\left[\left(1-\rho\right)\left(j-1\right)T\right] \\ &= \mu(\boldsymbol{\xi}). \end{aligned}$$

Note that from Equation (5.1), we have

$$U_j = Z_{jT} - Z_{(j-1)T}. \tag{6.1}$$

Now, let V_j be the increment of the underlying degradation process over $\left(\left(j-\rho\right)T, jT\right)$, with

$$\begin{aligned} V_j &= Z_{jT^-} - Z_{(j-\rho)T} \\ &= X^{(j)}_{V(jT^-,\rho)} - X^{(j)}_{V((j-\rho)T,\rho)} \end{aligned}$$

and so V_j is gamma distributed $\Gamma\left(\nu(\boldsymbol{\xi}), b\right)$. Once again from Equation (5.1), we have

$$V_j = Z_{jT^-} - Z_{jT}. \tag{6.2}$$

Therefore, Equations (6.1) and (6.2), provide the following expression for the observations increments:

$$\begin{aligned} Z_{jT^-} - Z_{(j-1)T^-} &= \left(Z_{jT^-} - Z_{jT}\right) + \left(Z_{jT} - Z_{(j-1)T}\right) - \left(Z_{(j-1)T^-} - Z_{(j-1)T}\right) \\ &= U_j + V_j - V_{j-1} \\ &= W_j - V_{j-1} \end{aligned}$$

where $W_j = U_j + V_j$ and where the random variables $U_1, \ldots, U_n, V_1, \ldots, V_n$ are independent.

□

Now the model identifiability must be verified. To this aim, we first need the result given in the lemma below.

Lemma 8. *Let X_1, X_2, Y_1 and Y_2 be independent and gamma distributed random variables with parameters (a_1, b_1), (a_2, b_1), (c_1, b_2) and (c_2, b_2) respectively. If $X_1 - X_2 \overset{\mathcal{D}}{=} Y_1 - Y_2$, then*

$$\begin{cases} b_1 = b_2 \\ a_1 = c_1 \\ a_2 = c_2 \end{cases}$$

Proof. Assume that $X_1 - X_2 \overset{\mathcal{D}}{=} Y_1 - Y_2$. We first look for an expression of the third centered moment of X_1. Let us recall that the third moment of the difference of two random variables is equal to the

difference of the moments. Let μ be the Lévy's measure associated to the random variable X_1. By [11, Eq. 2.32] and [11, Prop. 3.13], we have

$$\mathbb{E}\left(\left(X_t - \mathbb{E}\left(X_t\right)\right)^3\right) = \int_{\mathbb{R}} x^3 \mu(dx) = \int_0^\infty x^3 a_1 \frac{e^{-b_1 x}}{x} dx$$

and thus

$$\mathbb{E}\left(\left(X_t - \mathbb{E}\left(X_t\right)\right)^3\right) = \frac{2a_1}{b_1^3}$$

The expectations, variances and third centered moments of $X_1 - X_2$ and $Y_1 - Y_2$ are equal, which leads to

$$\begin{cases} \dfrac{a_1}{b_1} - \dfrac{a_2}{b_1} = \dfrac{c_1}{b_2} - \dfrac{c_2}{b_2} \\ \dfrac{a_1}{b_1^2} + \dfrac{a_2}{b_1^2} = \dfrac{c_1}{b_2^2} + \dfrac{c_2}{b_2^2} \\ \dfrac{2a_1}{b_1^3} - \dfrac{2a_2}{b_1^3} = \dfrac{2c_1}{b_2^3} - \dfrac{2c_2}{b_2^3} \end{cases}$$

The first and third equation provide $b_1 = b_2$, and then from the first and second equation we have

$$\begin{cases} a_1 = c_1 \\ a_2 = c_2 \end{cases}$$

which finishes the proof.

□

Theorem 5. *Let $\boldsymbol{Z}$ and $\tilde{\boldsymbol{Z}}$ be two random vectors based on the ARA_1 repair model with parameter $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{\theta}}$ respectively, with $\boldsymbol{\theta}$, $\tilde{\boldsymbol{\theta}} \in \Theta = (0,\infty)^2 \times (0,1) \times (0,\infty)$. If $\boldsymbol{Z} \overset{\mathcal{D}}{=} \tilde{\boldsymbol{Z}}$ and $n \geq 2$, then $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$.*

Proof. Assume that $\boldsymbol{Z} \overset{\mathcal{D}}{=} \tilde{\boldsymbol{Z}}$ and $n = 2$. Then $W_1 \overset{\mathcal{D}}{=} \tilde{W}_1$ and $W_2 - V_1 \overset{\mathcal{D}}{=} \tilde{W}_2 - \tilde{V}_1$, and from **Lemma 8** we deduce

$$\begin{cases} b = \tilde{b} \\ \omega_1(\boldsymbol{\xi}) = \omega_1(\tilde{\boldsymbol{\xi}}) \\ \nu_1(\boldsymbol{\xi}) = \nu_1(\tilde{\boldsymbol{\xi}}) \\ \omega_2(\boldsymbol{\xi}) = \omega_2(\tilde{\boldsymbol{\xi}}) \end{cases}$$

with $\boldsymbol{\xi} = (\alpha, \beta, \rho)$ and $\tilde{\boldsymbol{\xi}} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\rho})$. The expressions of ω_1, ω_2 and ν_1 lead to

$$\begin{cases} b = \tilde{b} \\ \alpha T^{\beta} = \tilde{\alpha} T^{\tilde{\beta}} \\ \left[1 - (1-\rho)^{\beta}\right] \alpha T^{\beta} = \left[1 - (1-\tilde{\rho})^{\tilde{\beta}}\right] \alpha T^{\tilde{\beta}} \\ \left[(2-\rho)^{\beta} - (1-\rho)^{\beta}\right] \alpha T^{\beta} = \left[(2-\tilde{\rho})^{\tilde{\beta}} - (1-\tilde{\rho})^{\tilde{\beta}}\right] \alpha T^{\tilde{\beta}} \end{cases}$$

which can be written as

$$\begin{cases} \alpha T^{\beta} = \tilde{\alpha} T^{\tilde{\beta}} \\ b = \tilde{b} \\ (1-\rho)^{\beta} = (1-\tilde{\rho})^{\tilde{\beta}} \\ (2-\rho)^{\beta} = (2-\tilde{\rho})^{\tilde{\beta}} \end{cases}$$

Let us set $r = 1 - \rho \in (0,1)$, the last two equations become:

$$\begin{cases} r^{\beta} = \tilde{r}^{\tilde{\beta}} \\ (1+r)^{\beta} = (1+\tilde{r})^{\tilde{\beta}} \end{cases}$$

and because $r \in (0,1)$, we have

$$\begin{cases} \beta \log(r) = \tilde{\beta} \log(\tilde{r}) \\ \beta \log(1+r) = \tilde{\beta} \log(1+\tilde{r}) \end{cases}$$

and finally

$$\frac{\log(1+r)}{\log(r)} = \frac{\log(1+\tilde{r})}{\log(\tilde{r})}.$$

However, the function $r \mapsto \log(1+r)/\log(r)$ is injective over $(0,1)$, hence necessarily $r = \tilde{r}$. From this we deduce that $\rho = \tilde{\rho}$, which leads to $\beta = \tilde{\beta}$ and then $\alpha = \tilde{\alpha}$. Therefore $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$.

□

In other words, the model parameters are identifiable as soon as the observations are conducted twice, at times T^{-} and $2T^{-}$. Now, we go on with the development of parametric estimation methods.

6.2 Method of moments estimators

Here we deal with the Moments Estimation (ME) method, which is the exact copy of the method applied to the ARD models. Hence, very few details are given in the following. See **Section 2.2** for the detailed definition of the method as well as its application. First, let us recall that this method is based on the minimization of the distance function D defined as

$$D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = \sum_{i=1}^{d}\sum_{j=1}^{n}\left(m_i(\boldsymbol{\theta}\,,jT^-) - m_i(\boldsymbol{\theta_0}\,,jT^-)\right)^2$$

and we first check the parameters identifiability in this framework through the following proposition.

Proposition 8. *The parameters of the ARA_1 model are identifiable from the ME method, that is $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0$ implies that $\boldsymbol{\theta} = \boldsymbol{\theta_0}$ for all $\boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$, as soon as $d \geq 2$ and $n \geq 3$.*

Proof. Here we prove that the assertion $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0 \;\Rightarrow\; \boldsymbol{\theta} = \boldsymbol{\theta_0}$ is true $\forall\, \boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$ as soon as $n \geq 3$ and $d \geq 2$. Assume that $D(\boldsymbol{\theta}\,,\boldsymbol{\theta_0}) = 0$, and let $\boldsymbol{\theta} = (\alpha\,,\beta\,,b\,,\rho)$ and $\boldsymbol{\theta_0} = (\alpha_0\,,\beta_0\,,b_0\,,\rho_0)$ be in Θ. So given the function D, we have for all j in $\{1\,,\ldots,n\}$:

$$\begin{cases} m_1(\boldsymbol{\theta}\,,jT^-) = m_1(\boldsymbol{\theta_0}\,,jT^-) \\ m_2(\boldsymbol{\theta}\,,jT^-) = m_2(\boldsymbol{\theta_0}\,,jT^-) \end{cases}$$

which is equivalent to

$$\begin{cases} \dfrac{\alpha\,(j-\rho(j-1))^{\beta}\,T^{\beta}}{b} = \dfrac{\alpha_0\,(j-\rho_0(j-1))^{\beta_0}\,T^{\beta_0}}{b_0} \\[2ex] \dfrac{\alpha\,(j-\rho(j-1))^{\beta}\,T^{\beta}}{b^2} = \dfrac{\alpha_0\,(j-\rho_0(j-1))^{\beta_0}\,T^{\beta_0}}{b_0^2} \end{cases}$$

When $j = 1$, we have

$$\begin{cases} \dfrac{\alpha T^{\beta}}{b} = \dfrac{\alpha_0 T^{\beta_0}}{b_0} \\[2ex] \dfrac{\alpha T^{\beta}}{b^2} = \dfrac{\alpha_0 T^{\beta_0}}{b_0^2} \end{cases}$$

which leads to

$$\begin{cases} \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ b = b_0 \end{cases} \tag{6.3}$$

Since the expectations and the variances have similar expression, considering both of them only allows the identifiability of the scale parameter and the quantity αT^{β}. Hence only the expectation is considered in the following, and taking into account the second and third observation times leads to

$$\begin{cases} \dfrac{\alpha\,(2-\rho)^{\beta}\,T^{\beta}}{b} = \dfrac{\alpha_0\,(2-\rho_0)^{\beta_0}\,T^{\beta_0}}{b_0} \\[2ex] \dfrac{\alpha\,(3-2\rho)^{\beta}\,T^{\beta}}{b} = \dfrac{\alpha_0\,(3-2\rho_0)^{\beta_0}\,T^{\beta_0}}{b_0} \end{cases}$$

and from System (6.3) we deduce

$$\begin{cases} (2-\rho)^{\beta} = (2-\rho_0)^{\beta_0} \\ (3-2\rho)^{\beta} = (3-2\rho_0)^{\beta_0} \end{cases}$$

The parameters $\rho,\ \rho_0 \in [0,1)$, thus the quantities $3-2\rho$ and $3-2\rho_0$ are strictly positive and we have

$$\frac{\log(2-\rho)}{\log(3-2\rho)} = \frac{\log(2-\rho_0)}{\log(3-2\rho_0)}$$

and because the function

$$f :\mapsto \frac{\log(2-x)}{\log(3-2x)}$$

is injective over $[0,1)$, we conclude that $\boldsymbol{\theta} = \boldsymbol{\theta_0}$.

□

In other words, the identifiability holds from the ME method as soon as the systems are observed three times, at times T^-, $2T^-$ and $3T^-$, and if at least the first two moments (expectation and variance) are used. Then, an estimation $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta_0}$ can be obtained through

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min}\, \hat{D}(\boldsymbol{\theta})$$

with $\hat{D}(\boldsymbol{\theta})$ the empirical version of $D(\boldsymbol{\theta}, \boldsymbol{\theta_0})$ (see Equation (2.1)). Let us set $n \geq 3$ and $d = 2$, and recall the used parametrization for the ME method: $\tilde{\boldsymbol{\theta}} = (\mu, \eta, \beta, \rho)$ with $\mu = \alpha/b$ and $\eta = \alpha/b^2$, that is $\alpha = \mu^2/\eta$ and $b = \mu/\eta$). Thus, we have

$$\begin{aligned} \hat{D}(\boldsymbol{\theta}) &= \hat{D}(\tilde{\boldsymbol{\theta}}) \\ &= \left(\mu^2+\eta^2\right)\left[\sum_{j=1}^{n}\left(\left(j-(j-1)\rho\right)^{\beta}T^{\beta}\right)^2\right] - 2\mu\left[\sum_{j=1}^{n}\left(j-(j-1)\rho\right)^{\beta}T^{\beta}\hat{m}_1\left(jT^-\right)\right] \\ &- \eta\left[\sum_{j=1}^{n}\left(j-(j-1)\rho\right)^{\beta}T^{\beta}\hat{m}_2\left(jT^-\right)\right] + \sum_{j=1}^{n}\left(\hat{m}_1^2\left(jT^-\right) + \hat{m}_2^2\left(jT^-\right)\right) \end{aligned}$$

where $\hat{m}_1\left(jT^-\right)$ and $\hat{m}_2\left(jT^-\right)$ are the empirical expectations and variances of Z_{jT^-}. Now we look for zeros of the partial derivatives of $\hat{D}$ with respect to μ and η:

$$\begin{cases} \partial_{\mu}\hat{D}(\tilde{\boldsymbol{\theta}}) = 0 \\ \partial_{\eta}\hat{D}(\tilde{\boldsymbol{\theta}}) = 0 \end{cases}$$

which leads to

$$\begin{cases} \mu = \dfrac{\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_1\left(jT^-\right)}{\sum_{j=1}^{n} \left(\left(j-(j-1)\rho\right)^{\beta} T^{\beta}\right)^2} \\ \\ \rho = \dfrac{\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_2\left(jT^-\right)}{\sum_{j=1}^{n} \left(\left(j-(j-1)\rho\right)^{\beta} T^{\beta}\right)^2} \end{cases}$$

Given these expressions for μ and η, the function $\hat{D}(\tilde{\boldsymbol{\theta}})$ can be written as

$$\left[\sum_{j=1}^{n} \left(\hat{m}_1^2\left(jT^-\right) + \hat{m}_2^2\left(jT^-\right)\right)\right] - \frac{\left[\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_1\left(jT^-\right)\right]^2 + \left[\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_2\left(jT^-\right)\right]^2}{\sum_{j=1}^{n} \left(\left(j-(j-1)\rho\right)^{\beta} T^{\beta}\right)^2}$$

which becomes a function only depending on β and ρ. In order to obtain estimates $\hat{\beta}$ and $\hat{\rho}$ of β and ρ, it remains to minimize numerically this function, which is equivalent to maximize the quantity

$$\frac{\left[\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_1\left(jT^-\right)\right]^2 + \left[\sum_{j=1}^{n} \left(j-(j-1)\rho\right)^{\beta} T^{\beta} \hat{m}_2\left(jT^-\right)\right]^2}{\sum_{j=1}^{n} \left(\left(j-(j-1)\rho\right)^{\beta} T^{\beta}\right)^2} \tag{6.4}$$

Finally, the estimator $\hat{\boldsymbol{\theta}}$ of θ has the following expression

$$\left(\frac{\left[\sum_{j=1}^{n} \left(j-(j-1)\hat{\rho}\right)^{\hat{\beta}} T^{\hat{\beta}} \hat{m}_1\left(jT^-\right)\right]^2}{\left[\sum_{j=1}^{n} \left(j-(j-1)\hat{\rho}\right)^{\hat{\beta}} T^{\hat{\beta}} \hat{m}_2\left(jT^-\right)\right] \left[\sum_{j=1}^{n} \left(\left(j-(j-1)\hat{\rho}\right)^{\hat{\beta}} T^{\hat{\beta}}\right)^2\right]} \;,\; \hat{\beta} \;,\; \hat{\rho} \;,\; \frac{\sum_{j=1}^{n} \left(j-(j-1)\hat{\rho}\right)^{\hat{\beta}} T^{\hat{\beta}} \hat{m}_1\left(jT^-\right)}{\sum_{j=1}^{n} \left(j-(j-1)\hat{\rho}\right)^{\hat{\beta}} T^{\hat{\beta}} \hat{m}_2\left(jT^-\right)} \right)$$

6.3 Maximum Likelihood Estimation

In this section we deal with the MLE method. We begin with the case where one single system is observed ($s = 1$) and we start by computing the joint p.d.f. of the random vector $\boldsymbol{Z}$, from where the likelihood function is easily derived when $s > 1$.

We know from **Theorem 5** that the model identifiability holds only if $n \geq 2$. Then, let us set $n \geq 2$, the joint p.d.f. $f_{\boldsymbol{Z}}$ of $\boldsymbol{Z}$ is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{Z}) = \int_{\mathbb{R}^{n-1}} f_{\boldsymbol{Z}|\boldsymbol{V_{n-1}}}(\boldsymbol{Z} \mid v_1, \dots, v_{n-1})\, f_{\boldsymbol{V_{n-1}}}(v_1, \dots, v_{n-1})\ dv_1 \dots dv_{n-1}$$

where $\boldsymbol{Z} = (z_1, z_2, \cdots, z_n) \in \mathbb{R}^+ \times \mathbb{R}^{n-1}$. The r.v. V_j are independent, then

$$f_{\boldsymbol{V_{n-1}}}(v_1, \dots, v_{n-1}) = \prod_{j=1}^{n-1} f_{V_j}(v_j)$$

and because $Z_j = U_j + V_j - V_{j-1}$, we can express the p.d.f. $f_{\boldsymbol{Z}|\boldsymbol{V_{n-1}}}$ as

$$f_{Z_1|\boldsymbol{V_{n-1}}}(z_1 \mid v_1, \dots, v_{n-1}) = f_{U_1+V_1|V_1}(z_1 \mid v_1) = f_{U_1}(z_1 - v_1),$$

$$f_{Z_j|\boldsymbol{V_{n-1}}}(z_j \mid v_1, \dots, v_{n-1}) = f_{U_j+V_j-V_{j-1}|V_{j-1},V_j}(z_j \mid v_{j-1}, v_j) = f_{U_j}(z_j + v_{j-1} - v_j)$$

for $1 < j < n$, and

$$f_{Z_n|\boldsymbol{V_{n-1}}}(z_n \mid v_1, \dots, v_{n-1}) = f_{U_n+V_n-V_{n-1}|V_{n-1}}(z_n \mid v_{n-1}) = f_{W_n}(z_n + v_{n-1}).$$

Therefore, with the convention $v_0 = 0$, the joint p.d.f. $f_{\boldsymbol{Z}}$ of $\boldsymbol{Z}$ can be expressed as follows:

$$f_{\boldsymbol{Z}}(\boldsymbol{Z}) = \int_{\mathbb{R}^{n-1}} \left[\prod_{j=1}^{n-1} f_{V_j}(v_j)\ f_{U_j}(z_j + v_{j-1} - v_j) \right] \times f_{W_n}(z_n + v_{n-1})\ dv_1 \dots dv_{n-1}. \tag{6.5}$$

We now reduce the integration domain to $[0,1]^{n-1}$ in order to be able to apply either Monte Carlo (MC) or Quasi Monte Carlo (QMC) methods for the integrals computations. This will prevent issues within the implementation of this method, since the computation of an integral over $\mathbb{R}^{n-1}$ might be an issue when n is large.

Proposition 9. *The joint p.d.f. of the random vector $\boldsymbol{Z}$ can be expressed as*

$$f_{\boldsymbol{Z}}(\boldsymbol{Z}) = \prod_{j=1}^{n-1} \Big(C_j - m_{j+1}\Big) \int_{[0,1]^{n-1}} G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{z})\, dx_1 \dots dx_{n-1}$$

where $\boldsymbol{x} = (x_1\ \dots\ x_{n-1})$, $m_j = \max(0, -z_j)$, $C_j = \sum_{i=1}^{j} z_i$, *and*

$$G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{z}) = \prod_{j=1}^{n-1} f_{V_j}\left((C_j - m_{j+1})\ x_j + m_{j+1}\right)$$

$$\times \prod_{j=1}^{n-1} f_{U_j}\left(z_j + (C_{j-1} - m_j)\, x_{j-1} - (C_j - m_{j+1})\, x_j + m_j - m_{j+1}\right)$$
$$\times f_{W_n}\left(z_n + x_{n-1}\big(C_{n-1} - m_n\big) + m_n\right)$$

Proof. The proof stands in two steps. First, we prove that the integral domain can be reduced to the set $\prod_{j=1}^{n-1} \left(\max(0, -z_{j+1}), C_j\right)$ for all $1 \leq j \leq n-1$. Then, a variable change will ensure the result.

The function to integrate in Equation (6.5) is a product of p.d.f., which can be zero over $\mathbb{R}^{n-1}$. Hence, in order to reduce the integral domain, we look for a subset of $\mathbb{R}^{n-1}$ containing all the value of the v_j for which the p.d.f. are non zero. Here, the functions f_{U_j}, f_{V_j} and f_{W_n} are all non zero if

- $0 < v_j$ for $1 \leq j \leq n-1$;
- $0 < z_j + v_{j-1} - v_j$ for $1 \leq j \leq n-1$;
- $\max(0, -z_n) < v_{n-1}$.

Let us set $m_j = \max(0, -z_j)$ and $C_j = \sum_{i=1}^{j} z_i > 0$ for $1 \leq j \leq n-1$, and show by induction that $m_{j+1} < v_j < C_j$ for all $1 \leq j \leq n-1$.

If $j = 1$, then $0 < z_1 + v_0 - v_1$ and $v_1 < z_1 = C_1$. Similarly, $0 < z_2 + v_1 - v_2$ hence $v_2 - z_2 < v_1$, and because $v_2 > 0$ we have $-z_2 < v_1$, and $0 < v_1$ implies that $m_2 < v_1$. Hence $m_2 < v_1 < C_1$.

Now assume that $m_j < v_{j-1} < C_{j-1}$ for $2 \leq j \leq n-2$, the aim is to show that $m_{j+1} < v_j < C_j$. We know by assumption that

- v_j, $v_{j+1} > 0$
- $0 < z_j + v_{j-1} - v_j$
- $0 < z_{j+1} + v_j - v_{j+1}$

and $v_j < z_j + C_{j-1} = C_j$ because $v_j < z_j + v_{j-1}$. Also, $v_{j+1} - z_{j+1} < v_j$, $v_j > 0$ and $v_{j-1} > 0$, which implies that $-z_{j+1} < v_j$ and finally $m_{j+1} < v_j$ because of the positivity of v_j.

Finally, if $j = n-1$ we know that $m_n < v_{n-1}$. Moreover, $v_{n-1} < z_{n-1} + v_{n-2} < z_{n-1} + C_{n-2} = C_{n-1}$, which leads to $m_n < v_{n-1} < C_{n-1}$. We conclude that the integral domain can be written as

$$\prod_{j=1}^{n-1} \left(\max(0, -z_{j+1}), C_j\right).$$

Note that the p.d.f. might be zero within this domain, but cannot be non zero outside. Now, the following change of variables can be applied. Let us set

$$x_j = \frac{v_j - m_{j+1}}{C_j - m_{j+1}} \text{ for all } 1 \leq j \leq n-1$$

hence

$$dv_j = (C_j - m_{j+1})\, dx_j$$

for $1 \leq j \leq n-1$, and the integration domain becomes $[0,1]^{n-1}$. Finally, substituting the v_j and the dv_j by x_j and dx_j leads to the result.

□

We now consider s observed systems. Let us set $S(\boldsymbol{z_{i,n}})$ such that

$$S(\boldsymbol{z_{i,n}}) = \prod_{j=1}^{n-1} \left(C_{i,j} - m_{i,j+1} \right)$$

with $C_{i,j} = \sum_{k=1}^{j} z_{i,k}$ and $m_{i,j+1} = \max(0, -z_{i,j+1})$. The log-likelihood is easily derived from the previous result, which has the following expression

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{z}) = \sum_{i=1}^{s} \log S(\boldsymbol{z_{i,n}}) + \sum_{i=1}^{s} \log \int_{[0,1]^{n-1}} G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{z})\, dx_1 \dots dx_{n-1}. \tag{6.6}$$

6.4 Half data method

We recall that the random variables $Z_j = U_j + V_j - V_{j-1}$ for $1 \leq j \leq n$ are not independent, which leads to a cumbersome expression for the joint p.d.f., which is difficult to assess from a numerical point of view, as we have seen in the previous section. However, the Z_j's with non consecutive indices j's can be seen to be independent, as these variables are pairwise dependent. We hence follow the idea of the sub-sample methods (see [4, 21]), and we consider only one-out-of two Z_j's, restricting the information to either odd indices, or even indices, and letting down the other ones.

6.4.1 Half data based on the odd indexes

For sake of simplicity, n is assumed here to be odd, and we set $p = (n+1)/2$. The half data vector $\tilde{\boldsymbol{Z}}$ is composed by the r.v.s. Z_j whose index is odd, that is

$$\tilde{\boldsymbol{Z}} = \left(Z_{2j-1} \right)_{1 \leq j \leq p} = \left(W_{2j-1} - V_{2j-2} \right)_{1 \leq j \leq p}$$

where the Z_{2j-1} are independent. Based on **Lemma 8**, the identifiability holds if the following system of equations

$$\begin{cases} \omega_{2j-1}(\boldsymbol{\xi}) = \omega_{2j-1}(\boldsymbol{\xi_0}) \\ \\ \nu_{2j-2}(\boldsymbol{\xi}) = \nu_{2j-2}(\boldsymbol{\xi_0}) \end{cases} \tag{6.7}$$

for $j = 1, \dots, p$ has a unique solution in $(\mathbb{R}^+)^2 \times (0,1)$, that is $\boldsymbol{\xi} = \boldsymbol{\xi_0}$. We first consider the case $n = 3$ (where two observations per trajectory are used), which will be seen to be non identifiable on the whole domain. We next go the the case n=5 (where three observations per trajectory are used), which will be shown to provide identifiability on the whole domain.

Case $n = 3$

System (6.7) becomes :

$$\begin{cases} \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ \left[(3-2\rho)^{\beta} - (2-2\rho)^{\beta}\right] \alpha T^{\beta} = \left[(3-2\rho_0)^{\beta_0} - (2-2\rho_0)^{\beta_0}\right] \alpha_0 T^{\beta_0} \\ \left[(2-\rho)^{\beta} - (2-2\rho)^{\beta}\right] \alpha T^{\beta} = \left[(2-\rho_0)^{\beta_0} - (2-2\rho_0)^{\beta_0}\right] \alpha_0 T^{\beta_0} \end{cases}$$

which is equivalent to

$$\begin{cases} \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ (1+2r)^{\beta} - (2r)^{\beta} = (1+2r_0)^{\beta_0} - (2r_0)^{\beta_0} \\ (1+r)^{\beta} - (2r)^{\beta} = (1+r_0)^{\beta_0} - (2r_0)^{\beta_0} \end{cases}$$

with $r = 1-\rho$ and $r_0 = 1-\rho_0$. We concentrate on the last two equations and try to identify the cases where they imply that $(r,\beta) = (r_0,\beta_0)$. In such cases, the first equation leads to $\alpha = \alpha_0$, and identifiability is obtained.
Let us set

$$\begin{aligned} h_1(r,\beta) &= (1+2r)^{\beta} - (2r)^{\beta} \\ h_2(r,\beta) &= (1+r)^{\beta} - (2r)^{\beta} \end{aligned}$$

with $\beta > 0$ and $r \in (0,1)$.

Lemma 9. *Let (r,β) and (r_0,β_0) be such that $h_i(r,\beta) = h_i(r_0,\beta_0)$ for $i = 1,2$. Then,*

1. *if $\beta_0 = \beta$ or $r = r_0$ then $(r,\beta) = (r_0,\beta_0)$;*
2. *if $\beta_0 \leq 1$ then $(r,\beta) = (r_0,\beta_0)$;*
3. *if $\beta_0 > 1$ then $\beta > 1$;*
4. *if β_0, $\beta > 1$, then $\beta_0 > \beta$ implies $r > r_0$ and $\beta_0 < \beta$ implies $r < r_0$;*
5. *if (r_0,β_0) is such that $h_2(r_0,\beta_0) < 1$, then $(r,\beta) = (r_0,\beta_0)$.*

Proof. The partial derivatives of h_1 and h_2 are given by

$$\begin{aligned} \partial_r h_1(r,\beta) &= 2\beta\left[(2r+1)^{\beta-1} - (2r)^{\beta-1}\right] \\ \partial_\beta h_1(r,\beta) &= (2r+1)^{\beta}\ln(2r+1) - (2r)^{\beta}\ln(2r) \end{aligned}$$

$$\begin{aligned} \partial_r h_2(r,\beta) &= \beta\left((r+1)^{\beta-1} - 2(2r)^{\beta-1}\right) \\ \partial_\beta h_2(r,\beta) &= (r+1)^{\beta}\ln(r+1) - (2r)^{\beta}\ln(2r) \end{aligned}$$

Note that both partial derivatives with respect to β are strictly positive. Moreover, $\partial_r h_1(r,1) = 0$ and the quantity $\partial_r h_1(r,\beta)$ has the same sign as $\beta - 1$, and if $\beta \leq 1$ then $\partial_r h_2(r,\beta) < 0$.
Also, when $\beta > 1$, the partial derivative $\partial_r h_2(r,\beta)$ changes sign at $g(\beta)$ from positive to negative, where

$$g(\beta) = \frac{1}{2^{\frac{\beta}{\beta-1}} - 1}$$

and $\beta \mapsto g(\beta)$ is non-decreasing from $(1,\infty)$ to $(0,1)$.

1. Assume $\beta = \beta_0$. If $\beta \leq 1$, then $r \longmapsto h_2(r,\beta)$ is injective and thus $h_2(r,\beta) = h_2(r_0,\beta_0)$ implies $r = r_0$. Similarly, if $\beta > 1$ then $r \longmapsto h_1(r,\beta)$ is injective and $r = r_0$.
   Assume now that $r = r_0$, hence $\beta = \beta_0$ because the function $\beta \longmapsto h_1(r,\beta)$ is injective.

2. Assume $\beta_0 = 1$, thus $h_1(r_0,1) = 1$ and by assumption $h_1(r,\beta) = 1$. However if $\beta > 1$, then $h_1(r,\beta) > h_1(0,\beta) = 1$ and in the same way, if $\beta < 1$, then $h_1(r,\beta) < 1$. This implies $\beta = \beta_0 = 1$, and $r = r_0$ is directly obtained from the previous case.

3. Assume $\beta_0 < 1$, thus $h_1(r_0,\beta_0) < h_1(r_0,1) = 1 = h_1(r,1)$ and so $h_1(r,\beta) < 1$, hence $\beta < 1$.
   If $\beta > \beta_0$, then $h_1(r,\beta) > h_1(r,\beta_0)$. Moreover $h_1(r,\beta) = h_1(r_0,\beta_0)$ hence

   $$\begin{aligned} 0 &= h_1(r,\beta) - h_1(r_0,\beta_0) \\ &= h_1(r,\beta) - h_1(r,\beta_0) + h_1(r,\beta_0) - h_1(r_0,\beta_0) \end{aligned}$$

   and so $h_1(r,\beta_0) - h_1(r_0,\beta_0) < 0$ because $h_1(r,\beta) > h_1(r,\beta_0)$, which leads to $r > r_0$.
   Let us now define h_3 such that $h_3(r,\beta) = h_1(r,\beta) - h_2(r,\beta)$. We have $h_3(r,\beta) = (1+2r)^\beta - (1+r)^\beta$ and

   $$\begin{aligned} \partial_r h_3(r,\beta) &= \beta\left(2(2r+1)^{\beta-1} - (1+r)^{\beta-1}\right) > 0; \\ \partial_\beta h_3(r,\beta) &= (2r+1)^\beta \ln(2r+1) - (1+r)^\beta \ln(1+r) > 0. \end{aligned}$$

   Based on the fact that $\beta > \beta_0$ and $r > r_0$, it can be stated that

   $$h_3(r,\beta) > h_3(r_0,\beta_0),$$

   which is impossible because $h_3(r,\beta) = h_3(r_0,\beta_0)$ by definition, hence $(r,\beta) = (r_0,\beta_0)$. In the same way, this result holds when $\beta < \beta_0$.

4. Assume $\beta_0 > 1$, then $h_1(r,\beta) = h_1(r_0,\beta_0) > h_1(r_0,1) = 1 = h_1(r,1)$ hence $\beta > 1$.
   With similar arguments as in the previous case, if $\beta < \beta_0$, then $r > r_0$. Finally if $\beta > \beta_0$, then $r < r_0$ by symmetry.

5. If $\beta_0 \leq 1$ then $(r,\beta) = (r_0,\beta_0)$ from the second case.
   Assume $\beta_0 > 1$ and $h_2(r_0,\beta_0) < 1$. Then, we know from the third case that $\beta > 1$, which implies that $r \mapsto h_2(r,\beta)$ is increasing over $(0, g(\beta)]$. Also, note that $h_2(0,\beta) = 1$ for all $\beta > 0$, and

because $r \mapsto h_2(r,\beta)$ is increasing over $(0, g(\beta)]$, necessarily $h_2(r,\beta) > 1$, which is impossible as $h_2(r_0,\beta_0) = h_2(r,\beta)$. Hence $r_0 > g(\beta_0)$ and $r > g(\beta)$ with similar arguments.
Also, either $\beta_0 > \beta$ and $r > r_0$, or $\beta_0 < \beta$ and $r < r_0$. If $\beta_0 > \beta$ and $r > r_0$ then $h_2(r_0,\beta_0) > h_2(r_0,\beta)$ because $\partial_\beta h_2(r,\beta) > 0$. Moreover, $r > r_0 > g(\beta_0) > g(\beta)$ and the function $r \mapsto h_2(r,\beta)$ is decreasing over $[g(\beta), 1)$, hence $h_2(r_0,\beta) > h_2(r,\beta)$. This leads to $h_2(r_0,\beta_0) > h_2(r,\beta)$ which is impossible since $h_2(r_0,\beta_0) = h_2(r,\beta)$. In the same way, the result holds when $\beta_0 < \beta$ and $r < r_0$.

□

Corollary 2. *Let $(\alpha,\beta,r,\alpha_0,\beta_0,r_0)$ be such that (6.7) is satisfied. If $\beta_0 \leq 1$ or $h_2(r_0,\beta_0) < 1$ then $\boldsymbol{\xi} = \boldsymbol{\xi_0}$ as soon as $n \geq 3$.*

Proof. From **Lemma 9**, if $\beta_0 \leq 1$ or $h_2(r_0,\beta_0) < 1$ then $(\beta, r) = (\beta_0, r_0)$, and the first line of (6.7) ensures that $\alpha = \alpha_0$. Let us recall that $r = 1-\rho$ and $r_0 = 1-\rho_0$, this leads to the conclusion $\boldsymbol{\xi} = \boldsymbol{\xi_0}$. □

The remaining cases now are $1 < \beta < \beta_0$ and $r < r_0$, or $1 < \beta_0 < \beta$ and $r > r_0$.

Let (r,β) and (r_0,β_0) be such that: (*)

- $h_i(r,\beta) = h_i(r_0,\beta_0)$ *for* $i = 1, 2$,
- $1 < \beta < \beta_0$ *and* $r_0 < r$.

Note that the case where $1 < \beta_0 < \beta$ and $r > r_0$ is identical to this one by symmetry, hence the study of one of the cases is enough.
Now, we investigate whether Conditions (*) are compatible. To this end, let us now define the function u_2 by

$$u_2(\gamma) = (1+\gamma r)^\beta - (2r)^\beta - (1+\gamma r_0)^{\beta_0} + (2r_0)^{\beta_0}$$

for $\gamma > 0$, where the parameters (r,β,r_0,β_0) are fixed and omitted in the notation $u_2(\gamma)$, for sake of simplification. From Conditions (*) we have $u_2(1) = u_2(2) = 0$, which is possible only if u_2 is non-monotonic over $[1,2]$. Therefore, in the following, the monotonicity of u_2 is studied.

Lemma 10. *Suppose Conditions (*) to hold. Then $u_2'(\gamma) > 0$ if and only if $\frac{r\beta}{r_0\beta_0} > w(\gamma)$, with*

$$w(\gamma) = \frac{(1+\gamma r_0)^{\beta_0-1}}{(1+\gamma r)^{\beta-1}} \quad \text{for all } \gamma > 0.$$

Moreover, $w(\gamma)$ reaches its minimum at γ_0 given by

$$\gamma_0 = \frac{r_0 - r - r_0\beta_0 + r\beta}{r r_0 (\beta_0 - \beta)}.$$

Proof. The derivative of u_2 is given by

$$u_2'(\gamma) = r\beta(1+\gamma r)^{\beta-1} - r_0\beta_0(1+\gamma r_0)^{\beta_0-1},$$

thus $u_2'(\gamma) > 0$ if and only if $\frac{r\beta}{r_0\beta_0} > w(\gamma)$. Then, $w'(\gamma)$ has the same sign as the quantity

$$(\beta_0 - 1) r_0 (1+\gamma r) - (\beta - 1) r (1+\gamma r_0)$$

which is positive if and only if $\gamma > \gamma_0$, and so $w(\gamma) > w(\gamma_0)$ for all $\gamma \neq \gamma_0$. □

Corollary 3. *Suppose Conditions (*) to hold. If the function u_2 is non-monotonic over $[1,2]$, then one of the following assertions holds :*

- $\gamma_0 \leq 1$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$ *with* $w(1) < w(2)$;
- $\gamma_0 \geq 2$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(2), w(1))$ *with* $w(2) < w(1)$;
- $\gamma_0 \in (1,2)$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), \max(w(1), w(2)))$.

Proof. Remember from **Lemma 10** that the function w is first decreasing on $[0, \gamma_0]$ and next increasing on $[\gamma_0, \infty)$, and note that there are three possible cases, according to whether $\gamma_0 \leq 1$, $\gamma_0 \geq 2$ or $\gamma_0 \in (1,2)$.

Assume first that $\gamma_0 \leq 1$. Then w is increasing on $[1,2]$, and hence $w(1) < w(2)$. Based on **Lemma 10** again, if the function u_2 is non monotonic, it means that $\frac{r\beta}{r_0\beta_0}$ must belong to $(w(1), w(2))$, otherwise, the sign of u_2' would remain a constant on $[1,2]$.

In the same way, if $\gamma_0 \geq 2$, then w is decreasing on $[1,2]$, and hence $w(2) < w(1)$. Also, $\frac{r\beta}{r_0\beta_0}$ must belong to $(w(2), w(1))$, with similar arguments.

Finally, assume that $\gamma_0 \in (1,2)$. Then, the function w is decreasing on $[1, \gamma_0]$ and increasing on $[\gamma_0, 2]$, so that

$$w([1,2]) = [w(\gamma_0), \max(w(1), w(2))].$$

If the function u_2 is non monotonic, it now means that $\frac{r\beta}{r_0\beta_0}$ must belong to $(w(\gamma_0), \max(w(1), w(2)))$. □

Lemma 11. *Suppose Conditions (*) to hold. Then $\gamma_0 > 1$ implies $\frac{r\beta}{r_0\beta_0} \geq w(1)$.*

Proof. Assume that $\gamma_0 > 1$. Our aim is to show that $\frac{r\beta}{r_0\beta_0} \geq w(1)$ which is equivalent to prove that

$$g(r,\beta) \leq 0 \tag{6.8}$$

with $g(r,\beta) = r_0\beta_0(1+r_0)^{\beta_0-1} - r\beta(1+r)^{\beta-1}$, for all r in $(0,1)$ and $\beta > 0$ such that $\gamma_0 > 1$.
As a first step, let us show that $\gamma_0 > 1$ implies that

$$r > v_1(\beta) = r_0 \frac{\beta_0 - 1}{\beta - 1 - r_0(\beta_0 - \beta)}.$$

Let us note that $\gamma_0 > 1$ can be written as

$$\frac{r}{1+r} > \frac{r_0}{1+r_0}\frac{\beta_0 - 1}{\beta - 1}, \tag{6.9}$$

or equivalently as

$$r\left(1-\frac{r_0}{1+r_0}\frac{\beta_0-1}{\beta-1}\right)>1. \tag{6.10}$$

Let us check that the factor between brackets is positive in the previous expression. As $r/(r+1)<1$, Inequality (6.9) entails that

$$1>\frac{r_0}{1+r_0}\frac{\beta_0-1}{\beta-1} \tag{6.11}$$

which can be written as

$$\beta>\beta_1\equiv 1+\frac{r_0}{1+r_0}\left(\beta_0-1\right).$$

Now it is easy to check that based on $\beta>\beta_1$ that Inequality (6.11) is true, and we now obtain from (6.10) that

$$r>\frac{1}{1-\frac{r_0}{1+r_0}\frac{\beta_0-1}{\beta-1}}-1=v_1\left(\beta\right).$$

Based on the fact that g is decreasing with respect to r, we now obtain that

$$g\left(r,\beta\right)<g\left(v_1\left(\beta\right),\beta\right).$$

In order to show (6.8), it is now sufficient to prove that $g\left(v_1\left(\beta\right),\beta\right)<0$. We have

$$\begin{aligned}
g\left(v_1\left(\beta\right),\beta\right) &= r_0\beta_0\left(1+r_0\right)^{\beta_0-1}-\frac{r_0\left(\beta_0-1\right)}{\beta-1-r_0\left(\beta_0-\beta\right)}\beta\left(1+\frac{r_0\left(\beta_0-1\right)}{\beta-1-r_0\left(\beta_0-\beta\right)}\right)^{\beta-1}\\
&= r_0\beta_0\left(1+r_0\right)^{\beta_0-1}-\frac{r_0\left(\beta_0-1\right)}{\beta-1-r_0\left(\beta_0-\beta\right)}\beta\frac{\left(r_0+1\right)^{\beta-1}\left(\beta-1\right)^{\beta-1}}{\left(\beta+\beta r_0-r_0\beta_0-1\right)^{\beta-1}}\\
&= r_0\left(r_0+1\right)^{\beta-1}\left[\beta_0\left(1+r_0\right)^{\beta_0-\beta}-\frac{\beta_0-1}{\beta-1}\beta\frac{\left(\beta-1\right)^{\beta}}{\left(\beta+\beta r_0-r_0\beta_0-1\right)^{\beta}}\right]\\
&= r_0\left(r_0+1\right)^{\beta-1}\phi\left(\beta\right)
\end{aligned}$$

with

$$\phi\left(\beta\right)=\beta_0\left(1+r_0\right)^{\beta_0-\beta}-\left(\beta_0-1\right)\frac{\beta}{\beta-1}\frac{1}{\left(1-r_0\frac{\beta_0-\beta}{\beta-1}\right)^{\beta}}.$$

The quantity $g\left(v_1\left(\beta\right),\beta\right)$ has the same sign as $\phi\left(\beta\right)$, hence it remains to prove that $\phi\left(\beta\right)\leq 0$. First, we have $\beta/(\beta-1)>\beta_0/(\beta_0-1)$ because $\beta<\beta_0$. This implies

$$\phi\left(\beta\right)<\beta_0\left(\left(1+r_0\right)^{\beta_0-\beta}-\frac{1}{\left(1-r_0\frac{\beta_0-\beta}{\beta-1}\right)^{\beta}}\right)$$

and consequently it is sufficient to show that

$$\beta \ln\left(\frac{1-r_0\frac{\beta_0-\beta}{\beta-1}}{1+r_0}\right) < -\beta_0 \ln\left(1+r_0\right). \tag{6.12}$$

Let us set z such that

$$z = \frac{\beta_0-\beta}{\beta-1}.$$

Inequation (6.12) is equivalent to $\tau(z) < -\beta_0 \ln(1+r_0)$ where

$$\tau(z) = \frac{z+\beta_0}{z+1}\ln\left(\frac{1-r_0 z}{1+r_0}\right) \quad \text{with} \quad 0 < z < \frac{\beta_0-\beta_1}{\beta_1-1} = \frac{1}{r_0}.$$

The derivative of τ is given by

$$\tau'(z) = -\frac{\beta_0-1}{(z+1)^2}\ln\left(\frac{1-r_0 z}{1+r_0}\right) + \frac{z+\beta_0}{z+1}\frac{-r_0}{1-r_0 z}$$

and has the same sign as

$$\sigma(z) = -(\beta_0-1)\ln\left(\frac{1-r_0 z}{1+r_0}\right) - r_0\frac{(z+\beta_0)(z+1)}{1-r_0 z}.$$

Also

$$\begin{aligned}\sigma'(z) &= -(\beta_0-1)\frac{-r_0}{1-r_0 z} - r_0\frac{1}{(1-r_0 z)^2}\left(-r_0 z^2 + 2z + \beta_0 + r_0\beta_0 + 1\right)\\ &= r_0(z+1)\frac{zr_0 - r_0\beta_0 - 2}{(zr_0-1)^2}\end{aligned}$$

and $\sigma'(z) < 0$ because $zr_0 < 1$. Thus the function σ is non-increasing, with $\sigma(0) < 0$, therefore $\sigma(z) < 0$. This leads to $\tau'(z) < 0$ for all $0 < z < \frac{1}{r_0}$, which implies that τ is also non-increasing, with $\tau(0) = -\beta_0\ln(1+r_0)$. Therefore, Inequation (6.12) is satisfied because

$$\tau(z) < -\beta_0 \ln(1+r_0)$$

for all $0 < z < \frac{1}{r_0}$, which implies $g(r,\beta) \leq 0$ and finally $\frac{r\beta}{r_0\beta_0} \geq w(1)$. □

Corollary 4. *Suppose Conditions (*) to hold. If the function u_2 is non-monotonic, then $\gamma_0 < 2$ and $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$, with $w(1) < w(2)$.*

Proof. From the second point of **Corollary 3**, if u_2 is non-monotonic and $\gamma_0 \geq 2$, then $\frac{r\beta}{r_0\beta_0} \in (w(2), w(1))$. This is inconsistent with the results of **Lemma 11** because $\gamma_0 > 1$ implies $\frac{r\beta}{r_0\beta_0} \geq w(1)$, hence this point cannot be verified. Moreover, in the case where $\gamma_0 \in (1,2)$, we have $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), \max(w(1), w(2)))$, that is either $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), w(2))$ or $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), w(1))$. Once again because $\gamma_0 > 1$ implies $\frac{r\beta}{r_0\beta_0} \geq w(1)$, the second case is impossible. It remains the case $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), w(2))$, but $w(1) > w(\gamma_0)$ by definition of γ_0, hence necessarily $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$ with $w(1) < w(2)$. □

In summary, if $h_i(r,\beta) = h_i(r_0,\beta_0)$ for $i = 1,2$, then the identifiability holds when $\beta_0 \leq 1$ or $h_2(r_0,\beta_0) < 1$, and in the case where $\beta_0 > 1$ we know that $\beta > 1$. Moreover, if $\beta < \beta_0$ and $r > r_0$ we have from **Corollary 4**:

- $\gamma_0 < 2$;
- $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$, with $w(1) < w(2)$.

We now look at the possible zone for (r, β) with a fixed (r_0, β_0) based on the previous results, from a numerical point of view.

Example 7. *In **Figure 6.1** are illustrated the possible zones for (r, β) based on **Corollary 4** with $(r_0, \beta_0) = (0.01, 5)$ and $(r_0, \beta_0) = (0.05, 4)$. The blue hatched area corresponds to the (r, β) such that $\frac{r\beta}{r_0\beta_0} > w(1)$, while the grey hatched area corresponds to the (r, β) such that $\frac{r\beta}{r_0\beta_0} < w(2)$, and in both zones $\gamma_0 < 2$. Therefore, from **Corollary 4**, if there is a solution (r, β) satisfying Conditions (*) then (r, β) belongs to the intersection of the hatched zones, which can be very small as seen in the right-hand plot.*

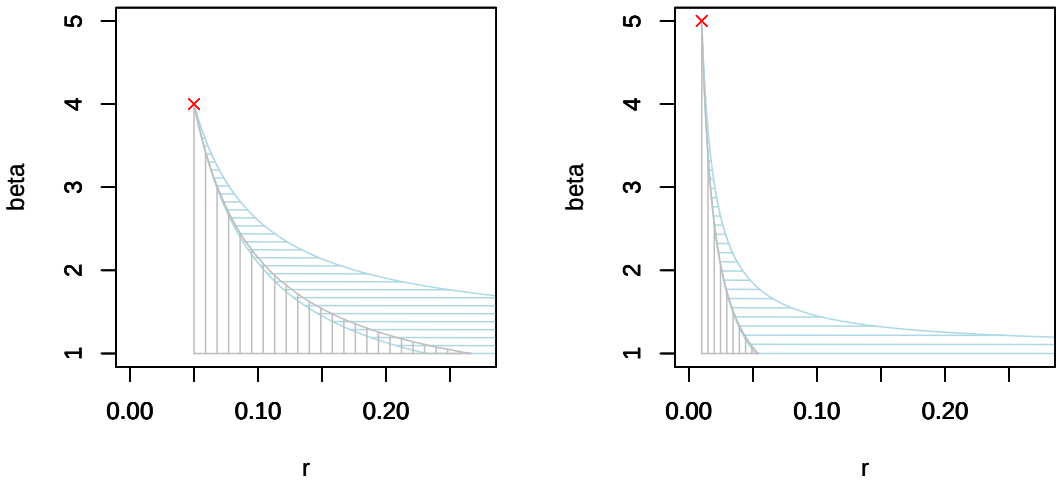

Figure 6.1: Both graphs highlight the possible areas for (r, β) fulfilling Conditions (*) given a fixed (r_0, β_0), based on **Corollary 4**. The blue hatched area corresponds to the (r, β) such that $\gamma_0 < 2$ and $\frac{r\beta}{r_0\beta_0} > w(1)$, while the grey hatched area corresponds to the (r, β) such that $\gamma_0 < 2$ and $\frac{r\beta}{r_0\beta_0} < w(2)$. The red crosses represent the values of (r_0, β_0), with $(r_0, \beta_0) = (0.01, 5)$ on the left-hand plot and $(r_0, \beta_0) = (0.05, 4)$ on the right-hand one.

Up to now, we only know that, if $h_i(r, \beta) = h_i(r_0, \beta_0)$ for $i = 1, 2$ with $1 < \beta < \beta_0$ and $r_0 < r$, then (r, β) should be in the intersection of the two hatched zones in **Figure 6.1**, obtained through **Corollary 4**. Based on some numerical investigations, it seems that such an intersection is never empty whatever (r_0, β_0) is. Hence, we are not able to conclude about possible identifiability, based on the previous theoretical results.

We next investigate identifiability from a numerical point of view. For each (r_0, β_0) in $(0, 1) \times (1, \infty)$, let us define the function

$$g_{(r_0,\beta_0)}(r, \beta) = \Big(h_1(r, \beta) - h_1(r_0, \beta_0)\Big)^2 + \Big(h_2(r, \beta) - h_2(r_0, \beta_0)\Big)^2,$$

which is zero as soon as $h_i(r,\beta) = h_i(r_0,\beta_0)$ for $i = 1, 2$. We begin with a first numerical illustration of the behavior of this function.

Example 8. *Let us set* $(r_0,\beta_0) \approx (0.07, 4.14)$. *The function* $g_{(r_0,\beta_0)}(r,\beta)$ *is plotted in* ***Figure 6.2****. It can be seen that besides* (r_0,β_0) *where the function* $g_{(r_0,\beta_0)}(r,\beta)$ *is zero, there is another point* $(r,\beta) = (0.1, 3)$ *for which this function is very close to zero, and there is one single such point. From a numerical point of view, there hence seems to exist a solution to Conditions* (*)*, and identifiability seems not to hold. The possible area for* (r,β) *given by* ***Corollary 4*** *is plotted in* ***Figure 6.3****, where* $(r,\beta) = (0.1, 3)$ *is indicated by a red cross. We observe that* $(r,\beta) = (0.1, 3)$ *is in the possible area that was theoretically obtained, which is coherent. More specifically, we have:*

- $h_i(r,\beta) \approx h_i(r_0,\beta_0)$ *for* $i = 1, 2$*;*
- $1 < \beta < \beta_0$ *and* $r_0 < r$*;*
- $\gamma_0 \approx -2.48 < 2$*;*
- $\frac{r\beta}{r_0\beta_0} \approx 1.03 \in \big(w(1), w(2)\big)$ *where* $w(1) \approx 1.02$ *and* $w(2) \approx 1.05$.

Similar results were found dealing with both points from ***Example 7****. In each case, a unique* (r,β) *was found:* $(r,\beta) \approx (0.0213, 2.385)$ *when* $(r_0,\beta_0) = (0.01, 5)$ *and* $(r,\beta) \approx (0.072, 2.853)$ *when* $(r_0,\beta_0) = (0.05, 4)$.
These numerical counter-examples lead us to conjecture that the identifiability does not hold over the entire set $(0,1) \times (0,\infty)$ *when* $n = 3$.

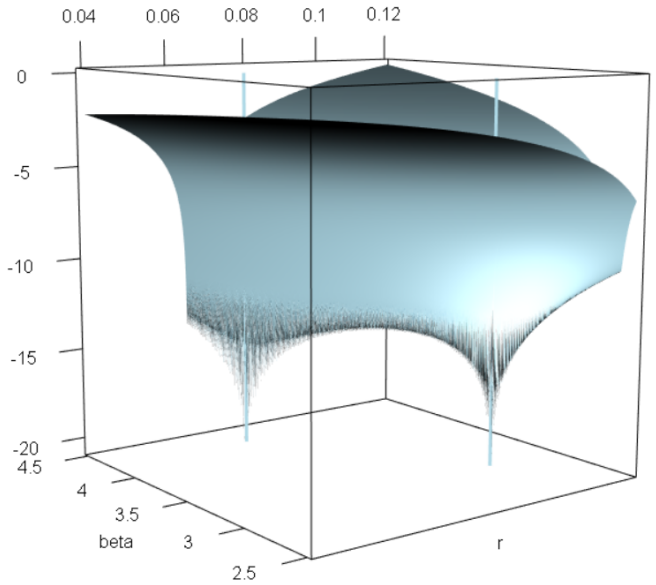

Figure 6.2: The surface represents the common logarithm of $g_{(r_0,\beta_0)}(r,\beta)$ depending on (r,β), with $(r_0,\beta_0) \approx (0.07, 4.14)$. Both blue lines highlight the two points (r_0,β_0) and $(r,\beta) = (0.1, 3)$ for which $g_{(r_0,\beta_0)}(r,\beta) \approx 0$.

Now, we are looking for the zone where identifiability does not hold, from a numerical point of view. For some $(r_0,\beta_0) \in (0,1) \times (1,50)$, approximate solutions (r,β) such that $g_{(r_0,\beta_0)}(r,\beta) = 0$ are searched. To be more specific, for each (r_0,β_0), we look for (r,β) such that

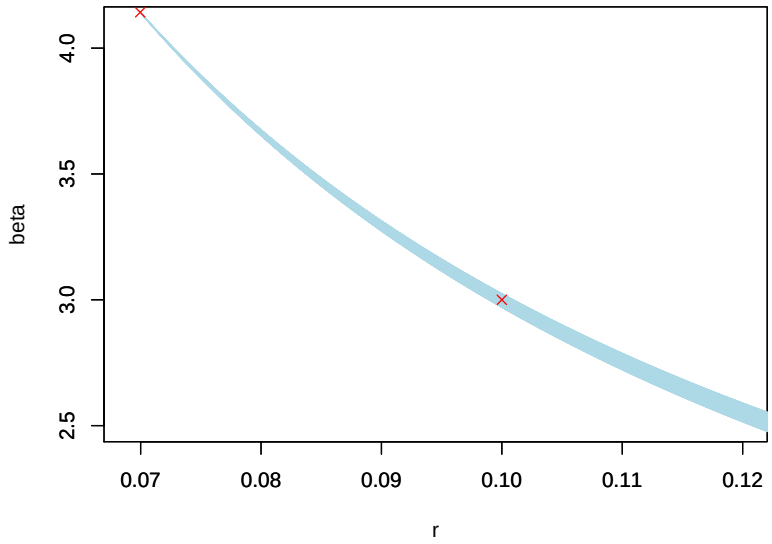

Figure 6.3: The possible area for (r, β) fulfilling Conditions (*) given $(r_0, \beta_0) \approx (0.07, 4.14)$, based on **Corollary 4**, is plotted in blue. Between the curves, we have $\gamma_0 < 2$ and $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$, with $w(1) < w(2)$. The red crosses represent (r, β) and (r_0, β_0).

- $(r, \beta) \neq (r_0, \beta_0)$;
- (r, β) is not in a neighbourhood of (r_0, β_0), that is $\|(r_0, \beta_0) - (r, \beta)\|_1 > 0.01$;
- $g_{(r_0,\beta_0)}(r, \beta) < 10^{-9}$.

For each (r_0, β_0) that has been considered in the study, either there was no solution for (r, β), so that it seems that such a (r_0, β_0) belongs to the identifiability zone, either there was one single solution (r, β) (as in the previous examples), and identifiability does not seem to hold for such a (r_0, β_0). The points where identifiability does not seem to hold are indicated in **Figure 6.4** with a black cross. The red line corresponds to the boundary of the zone containing such non identifiable points, and no solution has been found outside this area. The area where the identifiability were proved to hold is also plotted in blue. **Besides this zone, the previous numerical results lead us to the conjecture that the identifiability does not hold within the red area, while it holds outside.**

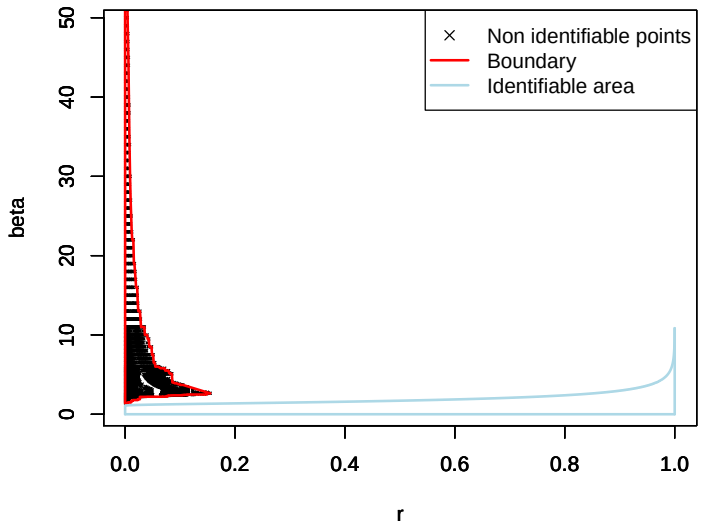

Figure 6.4: The black points are the values for (r_0, β_0) for which there exists (r, β) such that $g_{(r_0,\beta_0)}(r, \beta) < 10^{-9}$ and (r, β) is not in a neighbourhood of (r_0, β_0). The red lines correspond to the boundary of such points, and the area where theoretical identifiability holds is plotted in blue.

Case $n = 5$

The system (6.7) becomes :

$$\begin{cases} \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ h_1(r, \beta) = h_1(r_0, \beta_0) \\ h_2(r, \beta) = h_2(r_0, \beta_0) \\ h_3(r, \beta) = h_3(r_0, \beta_0) \\ h_4(r, \beta) = h_4(r_0, \beta_0) \end{cases} \tag{6.13}$$

where

$$h_3(r, \beta) = (1 + 4r)^{\beta} - (4r)^{\beta}$$
$$h_4(r, \beta) = (1 + 3r)^{\beta} - (4r)^{\beta}$$

Here Conditions (*) have to be redefined as two equations were added to the previous system.

Let (r, β) *and* (r_0, β_0) *be such that:* (**)

- $h_i(r, \beta) = h_i(r_0, \beta_0)$ *for* $i \in \{1, 2, 3, 4\}$

- $1 < \beta < \beta_0$ *and* $r_0 < r$.

As before for $n = 3$, let us define u_4 such that

$$u_4(\gamma) = (1 + \gamma r)^{\beta} - (4r)^{\beta} - (1 + \gamma r_0)^{\beta_0} + (4r_0)^{\beta_0}$$

for $\gamma \in [3, 4]$. Hence from Conditions (**) we have

$$u_2(1) = u_2(2) = 0$$
$$u_4(3) = u_4(4) = 0$$

Using similar arguments as those in **Corollary 3** leads to the following corollary for u_4.

Corollary 5. *Suppose Conditions* (**) *to hold. If the function* u_4 *is non-monotonic over* $[3, 4]$*, then one of the following assertions holds :*

- $\gamma_0 \leq 3$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(3), w(4))$;
- $\gamma_0 \geq 4$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(4), w(3))$;
- $\gamma_0 \in (3, 4)$ *and* $\frac{r\beta}{r_0\beta_0} \in (w(\gamma_0), \max(w(3), w(4)))$.

In **Figure 6.5**, the same area as in **Figure 6.3** is plotted in blue, as well as the area where the results of **Corollary 5** are satisfied, for $(r_0, \beta_0) \approx (0.07, 4.14)$. The intersection of the two areas seems empty, which indicates that identifiability holds in this case.

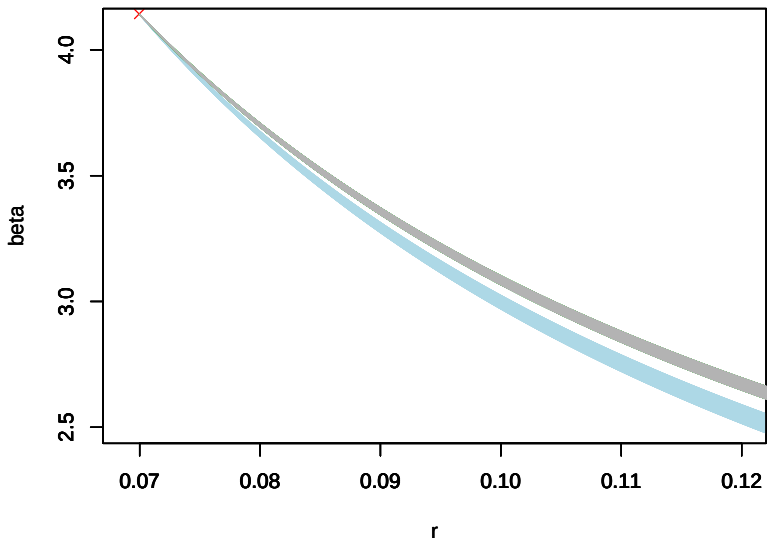

Figure 6.5: Both areas match the points (r, β) for which either the results of **Corollary 4** (blue area) or the results of **Corollary 5** (grey area) are satisfied, with $\beta_0 > \beta$, $r > r_0$ and $(r_0, \beta_0) \approx (0.07, 4.14)$. The red cross represents (r_0, β_0).

Theorem 6. ***(Identifiability)*** *Let* $(\boldsymbol{\xi}, \boldsymbol{\xi_0})$ *be such that* (6.13) *is satisfied, then* $\boldsymbol{\xi} = \boldsymbol{\xi_0}$ *as soon as* $n \geq 5$.

Proof. From **Lemma 2**, if $\beta_0 \leq 1$ then the identifiability holds. Now assume Conditions (**) to hold. Necessarily $\gamma_0 < 2$ and $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$ based on **Corollary 4**.
If $\gamma_0 \leq 1$, then $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$ and $\frac{r\beta}{r_0\beta_0} \in (w(3), w(4))$ from **Corollary 5**. However, $w(\gamma)$ is non-decreasing over $[1, 4]$, which is inconsistent. Hence $\gamma_0 \in (1, 2)$.
Then, $\frac{r\beta}{r_0\beta_0} \in (w(1), w(2))$ and $\frac{r\beta}{r_0\beta_0} \in (w(3), w(4))$ and thus $\frac{r\beta}{r_0\beta_0} \in (\max(w(1), w(3)), \min(w(2), w(4)))$.
The function w is non-decreasing over $(\gamma_0, 4)$, which leads to $w(4) > w(3) > w(2)$ as $\gamma_0 < 2$, and we know that $w(1) < w(2)$. Hence $\max(w(1), w(3)) = w(3)$ and $\min(w(2), w(4)) = w(2)$. Therefore $\frac{r\beta}{r_0\beta_0} \in (w(3), w(2))$, which is impossible because $w(3) > w(2)$.
As a conclusion, the case where $1 < \beta < \beta_0$ and $r > r_0$ is impossible, as well as the case $\beta > \beta_0 > 1$ and $r < r_0$ by symmetry, thus $(\beta, r) = (\beta_0, r_0)$ and finally $\boldsymbol{\xi} = \boldsymbol{\xi_0}$.

□

6.4.1.0.1 Likelihood expression

By definition, these random variables Z_{2j-1}, for $1 \leq j \leq p$, are independent. Hence the joint density function of $\tilde{\boldsymbol{Z}}$ is given by

$$f_{\tilde{\boldsymbol{Z}}}(\boldsymbol{z_n}) = \prod_{j=1}^{p} f_{Z_{2j-1}}(z_{2j-1}) = f_{W_1}(z_1) \prod_{j=2}^{p} f_{W_{2j-1}-V_{2j-2}}(z_{2j-1}). \tag{6.14}$$

For $2 \leq j \leq p$, the p.d.f. of $W_{2j-1} - V_{2j-2}$ can be developed as

$$f_{W_{2j-1}-V_{2j-2}}(z_{2j-1}) = \int_{\mathbb{R}} f_{W_{2j-1}}(z_{2j-1} + x) f_{V_{2j-2}}(x)\, dx.$$

Finally, replacing the p.d.f.'s by their expressions in Equation (6.14), we obtain the following expression for the joint density of $\tilde{\boldsymbol{Z}}$

$$\begin{aligned} f_{\tilde{\boldsymbol{Z}}}(\boldsymbol{z}) = & \frac{z_1^{\omega_1 - 1} b^{\Omega_p} \exp\left(-b \sum_{j=1}^{p} z_{2j-1}\right)}{\Gamma(\omega_1) \prod_{j=2}^{p} \Gamma(\omega_{2j-1})\Gamma(\nu_{2j-2})} \\ & \times \left[\prod_{j=2}^{p} \int_{\max(0, -z_{2j-1})}^{\infty} x^{\nu_{2j-2}-1} (z_{2j-1} + x)^{\omega_{2j-1}-1} \exp(-2bx)\, dx\right] \mathbf{1}_{\mathbb{R}^+}(z_1) \end{aligned}$$

with $\nu_0 = 0$ and $\Omega_p = \sum_{j=1}^{p} (\omega_{2j-1} + \nu_{2j-2})$.
Now s identical systems are considered. Hence an index i is added to each quantity, referring to the ith system, and the observation set is given by

$$\tilde{\boldsymbol{Z}} = \left(Z_{2j-1}^{(i)}\right)_{1 \leq i \leq s\ ,\ 1 \leq j \leq p}$$

whose related density function is given by

$$f_{\tilde{\boldsymbol{Z}}}(\boldsymbol{z}) = \frac{b^{s\tilde{\Omega}_p}\exp\left(-b\sum_{i=1}^{s}\sum_{j=1}^{p} z_{2j-1}^{(i)}\right)\prod_{i=1}^{s}\left(z_1^{(i)}\right)^{\omega_1-1}}{\Gamma(\omega_1)^s\prod_{j=2}^{p}\Gamma(\omega_{2j-1})^s\Gamma(\nu_{2j-2})^s} \times\left[\prod_{i=1}^{s}\prod_{j=2}^{p}\int_{\max(0,-z_{2j-1}^{(i)})}^{\infty} x^{\nu_{2j-2}-1}(z_{2j-1}^{(i)}+x)^{\omega_{2j-1}-1}\exp(-2bx)\,dx\right]\mathbf{1}_{\mathbb{R}^+}(z_1).$$

Finally, the log-likelihood has the following expression

$$\ell(\boldsymbol{\theta}\mid\boldsymbol{z}) = s\tilde{\Omega}_p\log b - b\sum_{i=1}^{s}\sum_{j=1}^{p} z_{2j-1}^{(i)} + (\omega_1-1)\sum_{i=1}^{s}\log z_1^{(i)} - s\log\Gamma(\omega_1) - s\sum_{j=2}^{p}\log\Gamma(\omega_{2j-1}) - s\sum_{j=2}^{p}\log\Gamma(\nu_{2j-2}) + \sum_{i=1}^{s}\sum_{j=2}^{p}\log I\left(\boldsymbol{\theta}\mid z_{2j-1}^{(i)}\right) \tag{6.15}$$

where

$$I\left(\boldsymbol{\theta}\mid z_{2j-1}^{(i)}\right) = \int_{\max(0,-z_{2j-1}^{(i)})}^{\infty} x^{\nu_{2j-2}-1}(z_{2j-1}^{(i)}+x)^{\omega_{2j-1}-1}\exp(-2bx)\,dx.$$

Note that all the results hold if n is an even integer, but in this case p must be such that $p = n/2$.

6.4.2 Half data based on the even indexes

In a manner similar to that used in the previous section, we only consider the Z_j's whose index is even, in order for these random variables to be independent. The identifiability is not studied here. Indeed, in **Chapter 7** dealing with numerical analysis, we will see that this method may have identifiability issues.

Here n is assumed to be even, and we set $p = n/2$. The considered observations set is

$$\tilde{\boldsymbol{Z}} = (Z_j)_{1\le j\le p} = (W_j - V_{j-1})_{1\le j\le p}$$

The joint p.d.f. of $\tilde{\boldsymbol{Z}}$ is here given by

$$f_{\tilde{\boldsymbol{Z}}}(\tilde{\boldsymbol{z}}_{\boldsymbol{n}}) = \prod_{j=1}^{p} f_{Z_{2j}}(z_{2j}) = \prod_{j=1}^{p} f_{W_{2j}-V_{2j-1}}(z_{2j})$$

Therefore, in a similar way than for the previous method, we derive the joint density expression below

$$f_{\tilde{\boldsymbol{Z}}}(\tilde{\boldsymbol{z}}) = \frac{b_p^{\Omega}\exp\left(-b\sum_{j=1}^{p} z_{2j}\right)}{\prod_{j=1}^{p}\Gamma(\omega_{2j})\Gamma(\nu_{2j-1})}\left[\prod_{j=1}^{p}\int_{\max(0,-z_{2j})}^{\infty} x^{\nu_{2j-1}-1}(z_{2j}+x)^{\omega_{2j}-1}\exp(-2bx)\,dx\right] \tag{6.16}$$

where

$$\Omega_p = \sum_{j=1}^{p} \left(\omega_{2j} + \nu_{2j-1}\right).$$

Now, considering s i.i.d. systems, we easily derive the log-likelihood expression from Equations (6.15) and (6.16), which is given by

$$\begin{aligned} \ell\left(\boldsymbol{\theta} \mid \boldsymbol{z}\right) = s\tilde{\Omega}_p \log b - b \sum_{i=1}^{s}\sum_{j=1}^{p} z_{2j}^{(i)} - s\sum_{j=1}^{p} \log\Gamma(\omega_{2j}) \\ - s\sum_{j=1}^{p} \log\Gamma(\nu_{2j-1}) + \sum_{i=1}^{s}\sum_{j=1}^{p} \log I\left(\boldsymbol{\theta} \mid z_{2j}^{(i)}\right) \end{aligned} \tag{6.17}$$

where

$$I\left(\boldsymbol{\theta} \mid z_{i,2j}\right) = \int_{\max(0,-z_{2j}^{(i)})}^{\infty} x^{\nu_{2j-1}-1}(z_{2j-1}^{(i)} + x)^{\omega_{2j}-1} \exp\left(-2bx\right)\, dx.$$

Note that if n is odd, the previous results hold with $p = (n-1)/2$.

6.5 Maximum composite likelihood estimation

Another way to avoid dependency issues is to assume that the variables are independent, even if they are not. It is the purpose in this section, where two MLE methods are developed considering composite likelihood. To be more precise, the first one deal with the observations (assumed mutually independent), while the second one is based on the increments (also assumed mutually independent), according to the definition provided by [41].

6.5.1 Composite likelihood based on the observations

For all $j \in \{1, \ldots, n\}$, the r.v. $Z_j = Z_{jT^-}$ are gamma distributed $\Gamma\big(a\left(jT - \rho(j-1)T\right), b\big)$. Let us set $a_j = a\left(jT - \rho(j-1)T\right)$ the shape parameters of these distributions, we have the following expression for the p.d.f. of Z_j

$$f_{Z_j}(z) = \frac{b^{a_j}}{\Gamma(a_j)}\, z^{a_j-1} \exp\left(-bz\right) \mathbf{1}_{\mathbb{R}^+}(z)$$

Assuming that these r.v.s are independent, the likelihood is given by

$$L\left(\mathbf{z} \mid \boldsymbol{\theta}\right) = \frac{b^{\sum_{j=1}^{n} a_j}}{\prod_{j=1}^{n} \Gamma(a_j)} \exp\left(-b\sum_{j=1}^{n} z_j\right) \prod_{j=1}^{n} z_j^{a_j-1}.$$

Now we look at the parameters identifiability in this case, as it might not hold under the independence assumption.

Proposition 10. *The parameters of the ARA_1 model are identifiable from the composite likelihood function, that is $L(\boldsymbol{\theta}|\boldsymbol{z}) = L(\boldsymbol{\theta_0}|\boldsymbol{z})$ for all $\boldsymbol{y} = (z_1, \ldots, z_n)$ implies that $\boldsymbol{\theta_0} = \boldsymbol{\theta}$ and $\boldsymbol{\theta}, \boldsymbol{\theta_0} \in \Theta$, as soon as $n \geq 3$. In other words, the identifiability holds from the likelihood function as soon as observations are conducted three times, and whatever s is.*

Proof. Given the expression of the likelihood, we know that the gamma distribution parameters are identifiable. Hence, for $j = 1, 2, 3$, we deduce that

$$\begin{cases} \alpha T^{\beta} = \alpha_0 T^{\beta_0} \\ b = b_0 \\ (2-\rho)^{\beta} = (2-\rho_0)^{\beta_0} \\ (3-2\rho)^{\beta} = (3-2\rho_0)^{\beta_0} \end{cases}$$

This system is the same one as this in **Proposition 8**, which was proved to have a unique solution, and so the identifiability holds.

□

Now, the log-likelihood is given by

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{z}) = \log(b) \sum_{j=1}^{n} a_j - b \sum_{j=1}^{n} z_j - \sum_{j=1}^{n} \log\Gamma(a_j) + \sum_{j=1}^{n} (a_j - 1) \log z_j.$$

The parameter b can be written as a function of the other ones by solving the equation $\partial_b \ell(\boldsymbol{\theta} \mid \boldsymbol{z}) = 0$, which leads to

$$b = \frac{\sum_{j=1}^{n} a_j}{\sum_{j=1}^{n} z_j}.$$

Then, substituting b by its expression above in the log-likelihood leads to

$$\ell\left(\tilde{\boldsymbol{\theta}} \mid \boldsymbol{z_n}\right) = \left(\log\left(\sum_{j=1}^{n} a_j\right) - 1 - \log\left(\sum_{j=1}^{n} z_j\right)\right) \sum_{j=1}^{n} a_j - \sum_{j=1}^{n} \log\Gamma(a_j) + \sum_{j=1}^{n} (a_j - 1) \log z_j$$

where $\tilde{\boldsymbol{\theta}} = (\alpha, \beta, \rho)$.

We now deal with the extension of this method considering s i.i.d. systems. The log-likelihood has the following expression

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{z}) = s \log b \sum_{j=1}^{n} a_j - s \sum_{j=1}^{n} \log\Gamma(a_j) + \sum_{i=1}^{s} \log\left[\exp\left(-b \sum_{j=1}^{n} z_j^{(i)}\right) \prod_{j=1}^{n} \left(z_j^{(i)}\right)^{a_j - 1}\right]$$

$$= s \log b \sum_{j=1}^{n} a_j - s \sum_{j=1}^{n} \log \Gamma(a_j) - b \sum_{i=1}^{s} \sum_{j=1}^{n} z_j^{(i)} + \sum_{i=1}^{s} \sum_{j=1}^{n} (a_j - 1) \log z_j^{(i)}.$$

Once again the parameter b can be expressed with respect to the other ones, leading to

$$b = \frac{s \sum_{j=1}^{n} a_j}{\sum_{i=1}^{s} \sum_{j=1}^{n} z_j^{(i)}}$$

and injecting this expression in ℓ entails the following expression for the profile composite likelihood function:

$$\ell\left(\tilde{\boldsymbol{\theta}} \mid \boldsymbol{z_n}\right) = \log \left(\frac{s \sum_{j=1}^{n} a_j}{\sum_{i=1}^{s} \sum_{j=1}^{n} z_{i,j}} \right) s \sum_{j=1}^{n} a_j - s \sum_{j=1}^{n} \log \Gamma(a_j) - \frac{s \sum_{j=1}^{n} a_j}{\sum_{i=1}^{s} \sum_{j=1}^{n} z_{i,j}} \sum_{i=1}^{s} \sum_{j=1}^{n} z_{i,j} + \sum_{i=1}^{s} \sum_{j=1}^{n} (a_j - 1) \log z_{i,j}$$

which can be reduced to

$$\ell\left(\tilde{\boldsymbol{\theta}} \mid \boldsymbol{z_n}\right) = \left(s \sum_{j=1}^{n} a_j \right) \left(\log s - 1 + \log \sum_{j=1}^{n} a_j - \log \sum_{i=1}^{s} \sum_{j=1}^{n} z_{i,j} \right) - s \sum_{j=1}^{n} \log \Gamma(a_j) + \sum_{i=1}^{s} \sum_{j=1}^{n} (a_j - 1) \log z_{i,j}. \tag{6.18}$$

6.5.2 Composite likelihood based on the increments

As in the previous section, the aim here is to provide the expression of the composite log-likelihood, but dealing with the difference of the observations $Z_j = Z_{jT^-} - Z_{(j-1)T^-}$. Let us recall that $Z_j = W_j - V_{j-1}$. Thus, the p.d.f. of the Z_j's for $1 \leq j \leq n$ is given by $f_{W_j - V_{j-1}}(z)$, whose expression is close to that in Equation (6.16), hence we can write

$$f_{Z_j}(z) = \frac{b^{w_j + v_{j-1}}}{\Gamma(w_j)\Gamma(v_{j-1})} \exp(-bz) \int_{max(0,-z)}^{\infty} (z+x)^{w_j - 1} x^{v_{j-1} - 1} \exp(-2bx) \, dx.$$

Similarly as in the previous case, assuming that the Z_j's are independent allows us to write the joint density of $\boldsymbol{Z}$, whose expression is given below

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{b^{\sum_{j=1}^{n} w_j + v_{j-1}}}{\prod_{j=1}^{n} \Gamma(w_j)\Gamma(v_{j-1})} \exp\left(-b\sum_{j=1}^{n} z_j\right) \prod_{j=1}^{n} \left[\int_{max(0,-z_j)}^{\infty} (z_j + x)^{w_j - 1}\, x^{v_{j-1}-1} \exp(-2bx)\, dx\right]$$

and thus the composite log-likelihood has the following expression

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{Z}) = \log(b)\sum_{j=1}^{n}(w_j + v_{j-1}) - b\sum_{j=1}^{n} z_j - \sum_{j=1}^{n}\log\Gamma(w_j) - \sum_{j=1}^{n}\log\Gamma(v_{j-1})$$

$$+ \sum_{j=1}^{n}\left[\log\int_{max(0,-z_j)}^{\infty} (z_j + x)^{w_j - 1}\, x^{v_{j-1}-1} \exp(-2bx)\, dx\right].$$

Now we extend the log-likelihood expression to the case where s i.i.d. systems are observed. The joint density of $\boldsymbol{Z}$ is

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \prod_{i=1}^{s}\left[\frac{b^{\sum_{j=1}^{n} w_j + v_{j-1}}}{\prod_{j=1}^{n} \Gamma(w_j)\Gamma(v_{j-1})} \exp\left(-b\sum_{j=1}^{n} z_j^{(i)}\right) \times \prod_{j=1}^{n}\left[\int_{max(0,-z_j^{(i)})}^{\infty} (z_j^{(i)} + x)^{w_j - 1}\, x^{v_{j-1}-1} \exp(-2bx)\, dx\right]\right]$$

or equivalent

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{b^{s\sum_{j=1}^{n} w_j + v_{j-1}}}{\prod_{j=1}^{n} \Gamma(w_j)^s\Gamma(v_{j-1})^s} \exp\left(-b\sum_{i=1}^{s}\sum_{j=1}^{n} z_j^{(i)}\right) \prod_{i=1}^{s}\prod_{j=1}^{n} I_{i,j}(\boldsymbol{z}, \boldsymbol{\theta})$$

where

$$I_{i,j}(\boldsymbol{z}, \boldsymbol{\theta}) = \int_{max(0,-z_j^{(i)})}^{\infty} (z_j^{(i)} + x)^{w_j - 1}\, x^{v_{j-1}-1} \exp(-2bx)\, dx$$

Finally, the composite log-likelihood is given by

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{Z}) = s\log b\sum_{j=1}^{n}(w_j + v_{j-1}) - s\sum_{j=1}^{n}(\log\Gamma(w_j) + \log\Gamma(v_{j-1})) \tag{6.19}$$
$$- b\sum_{i=1}^{s}\sum_{j=1}^{n} z_j^{(i)} + \sum_{i=1}^{s}\sum_{j=1}^{n}\log I_{i,j}(\boldsymbol{Z}, \boldsymbol{\theta}).$$

Once again, the identifiability is not studied here for the same reasons as those explained in **Section 6.4.2**.

Chapter 7

Simulation study

7.1 Methods selection

The aim of this chapter is to investigate numerically the estimators quality for each method. For sake of simplicity, the estimation methods are abbreviated as follows:

- Moments estimation: ME;
- Maximum likelihood estimation: MLE;
- Maximum likelihood estimation using Monte Carlo approximation: MC;
- Maximum likelihood estimation using Quasi Monte Carlo approximation: QMC;
- Maximum likelihood estimation using half of the data which have an odd index: HDOI;
- Maximum likelihood estimation using half of the data which have an even index: HDEI;
- Maximum likelihood estimation using the composite likelihood based on the observations: CLO;
- Maximum likelihood estimation using the composite likelihood based on the increments: CLI;

The estimation methods are first tested on large samples in order to select the most efficient ones. To be able to compare results, we place ourselves within the same framework for each method. The model parameters and the observations characteristics are the following:

- Shape function parameter: $\boldsymbol{\xi} = (\alpha, \beta) = (1, 1)$;
- Scale parameter: $b = 1$;
- Maintenance actions efficiency: $\rho = 0.5$;
- Period of repairs: $T = 1$;
- Observations times: $\{jT^{-}; 1 \leq j \leq n\}$ with $n = 10$;
- Number of observed i.i.d. systems: $s = 1000$.

We generate 500 observations sets, and for each method except for MC and QMC an estimation of $\boldsymbol{\theta} = (\alpha, \beta, \rho, b)$ is computed for each set. These estimations are based on the minimization or maximization of the quantities provided in Equations (6.4), (6.15), (6.17), (6.18) and (6.19), based on a gradient method. The range for the parameters in the optimization procedure is the following:

- $\alpha \in [0.1, 5]$ for all methods except the ME method;
- $\beta \in [0.1, 5]$ for both ME and CLO methods, the range is reduced to $\beta \in [0.1, 2.5]$ for the HDEI, the HDOI and the CLI methods, due to numerical issues for large beta (see below);
- $\rho \in [0.01, 0.99]$ for all the methods;
- $b \in [0.1, 5]$ for the HDEI, the HDOI and the CLI methods.

The numerical issues for large beta are related to the computations of the integrals in the log-likelihood (see Equations (6.15), (6.17) and (6.19)). To be more precise, a large β induces high shape parameters for the involved gamma p.d.f., beyond 3150 when $(\alpha, \beta, \rho) = (1, 5, 0.5)$ and $n = 10$, leading to integrals which seem difficult to compute through the standard methods implemented in R. Reducing the maximum possible value of β to 2.5 allows the computations to be possible as the shape parameter does not exceed 28.
Concerning the initialization of the optimization procedure, we have experimented several different initial points. In each case, the final results were almost the same. Then, we have chosen to consider the true values of the parameters as initial points, which allows to slightly reduce the CPU times without affecting the numerical results.

The relative bias (in percentage) and the variances resulting of these estimations are shown in **Table 7.1** and **Table 7.2** respectively. Also, **Figures 7.1** to **7.5** represent the pairs plots of the estimations for each of the five methods.
The classical MLE is not treated here because it requires the use of MC and QMC approximations, and the large dimension ($n - 1 = 9$) induces too high computing times.

Table 7.1: Relative bias estimation

Methods / Relative bias (%)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
ME	0.14%	1.18%	2.43%	0.01%
HDEI	1.32%	0.11%	0.06%	0.63%
HDOI	0.11%	0.002%	0.17%	0.1%
CLO	0.08%	0.29%	0.02%	0.01%
CLI	2.79%	1.91%	10.24%	2.27%

Table 7.2: Variances estimation

Methods / Variances	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
ME	6.8×10^{-3}	1.0×10^{-2}	1.5×10^{-2}	1.7×10^{-3}
HDEI	6.0×10^{-3}	1.2×10^{-3}	2.3×10^{-4}	3.8×10^{-3}
HDOI	8.2×10^{-4}	2.6×10^{-4}	1.9×10^{-4}	7.7×10^{-4}
CLO	5.4×10^{-4}	1.6×10^{-3}	2.1×10^{-3}	1.9×10^{-5}
CLI	1.3×10^{-3}	2.3×10^{-4}	4.6×10^{-5}	5.2×10^{-4}

First, from **Tables 7.1** and **7.2** as well as **Figure 7.5**, it can be seen that the CLI method is biased, especially regarding the parameter ρ: the mean bias is beyond 10% associated with a variance less than

5.5×10^{-5}. This disqualifies the CLI method.
Among the remaining estimation methods, the ME method is globally less efficient than the other three: regarding the parameters β and ρ, the bias and the variances are higher than for any other method. This can also be seen by comparing **Figure 7.1** with **Figures 7.2**,**7.3**,**7.4**.
We observe the same characteristics for the HDEI regarding α and b, that is a higher bias and variance than for the HDOI and CLO methods. Moreover, extreme values can be seen in **Figure 7.2**. This is a lack of robustness of the HDEI method compared to alternative methods since for some data sets estimations can be very far from the true parameters values.

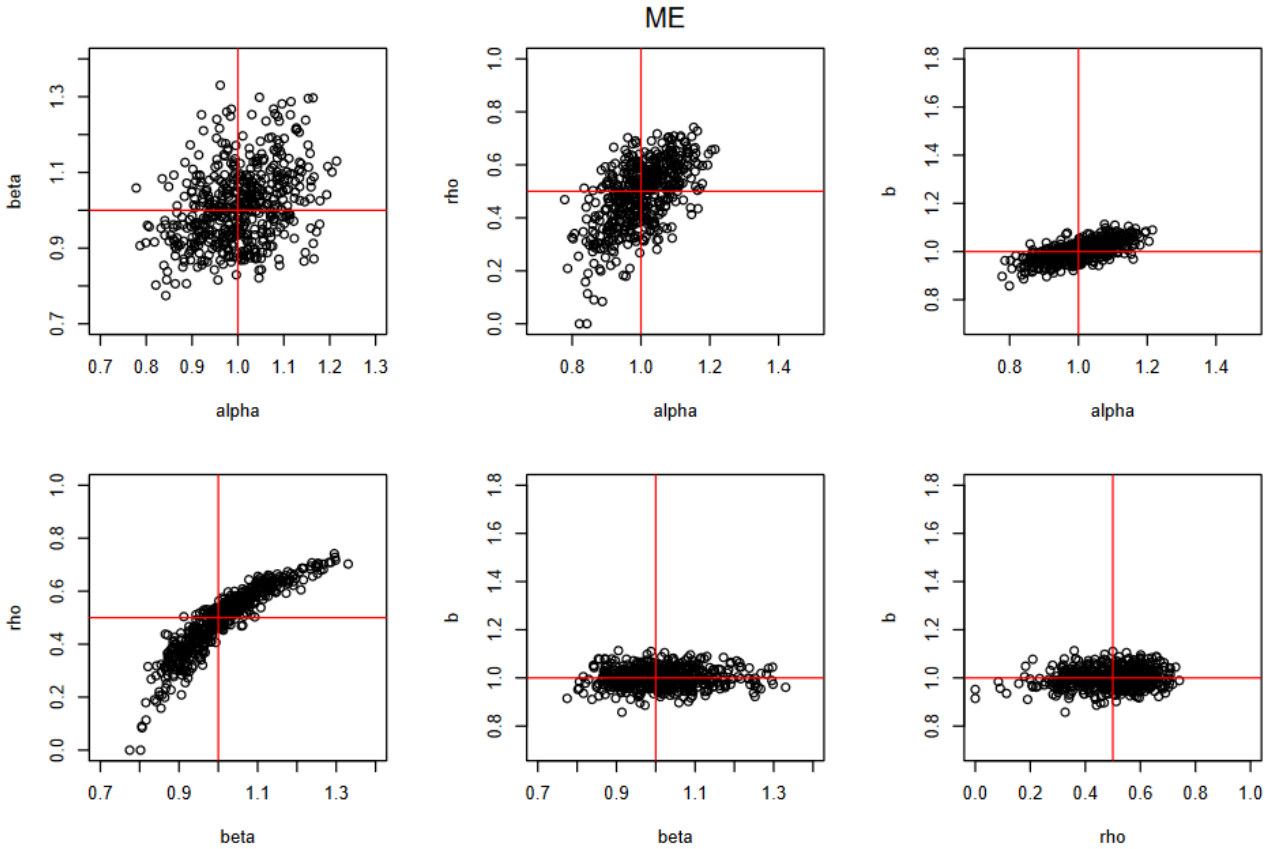

Figure 7.1: Pairs plots of ME method estimates. The red lines indicate the true value of each parameter.

In conclusion, the HDOI and the CLO methods are the most reliable with respect to the three other ones (ME, HDEI and CLI). Hence, in the following, these methods are investigated and compared in order to determine which one is the most efficient. Remember that the identifiability holds for $n \geq 5$ for the HDOI method while it holds as soon as $n \geq 3$ for the CLO method. Moreover, if $n = 2$, the identifiability holds for the MLE only. Hence, before going on further with the comparison between the HDOI and the CLO methods, the cases $n = 2, 3, 4$ are studied.

As already mentioned, the MC and QMC methods require large computing times. However, in order to increase the estimations accuracy, the number of points has to be large regarding n and this affects the computing time. Hence, the number of systems s and the number of points for the integral approximations are set to 100 and $2^{14} = 16,384$ respectively.

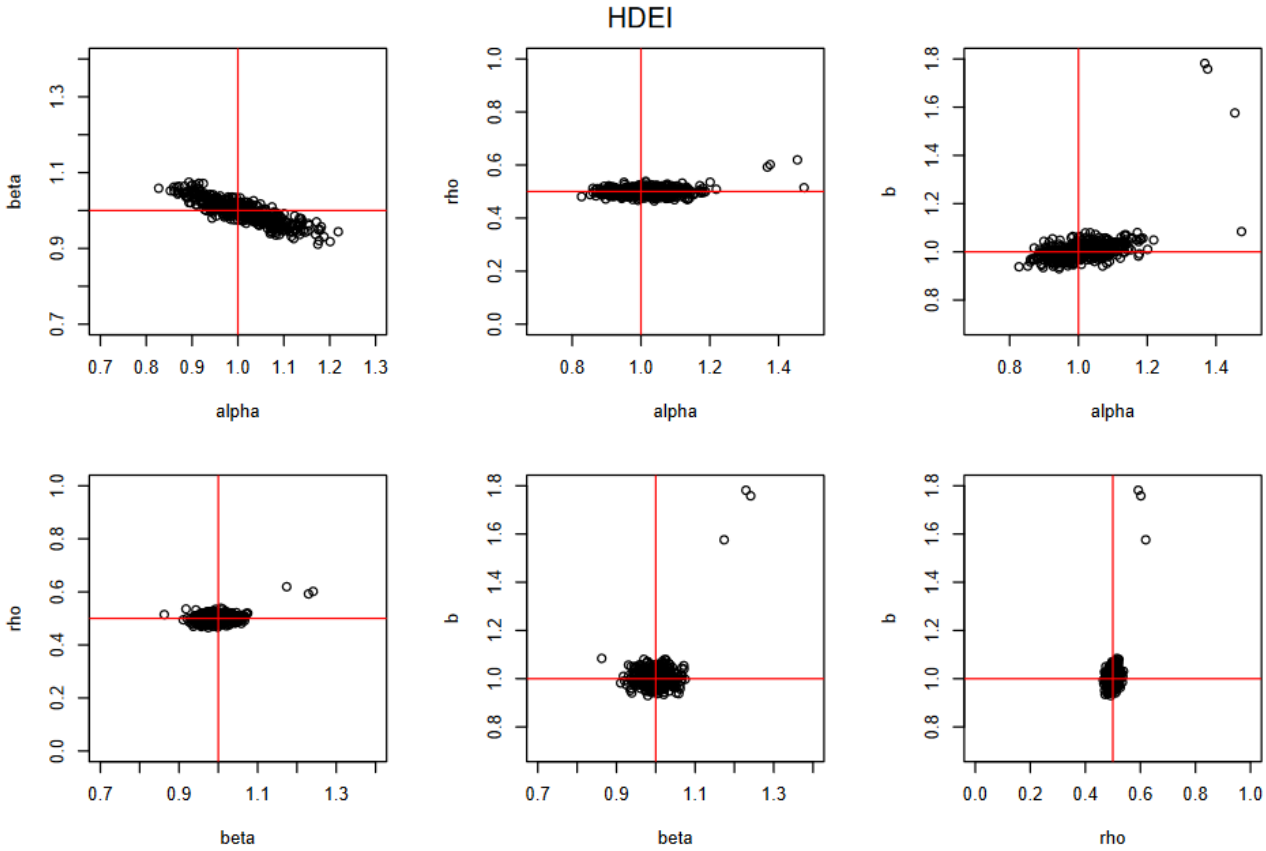

Figure 7.2: Pairs plots of HDEI estimates. The red lines indicate the true value of each parameter.

In the case where $n = 2$, only the MLE can be used because of identifiability issues. The computation of the integrals involved in the likelihood function is done through QMC and MC simulations, as well as through numerical integration based on adaptive quadrature. The results are shown in **Table 7.3**. Regarding the parameters α, β and b, the MC and the QMC methods are better than the MLE, while it is the opposite for the estimation of ρ.

Table 7.3: Relative bias and variances for the MLE, MC and QMC methods when $n = 2$

Methods / Relative bias (%)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
MLE	2.66%	1.12%	0.62%	3.05%
MC	0.08%	0.52%	7.01%	0.03%
QMC	0.45%	0.43%	3.52%	0.77%
Methods / Variances	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
MLE	9.7×10^{-3}	2.2×10^{-2}	4.7×10^{-3}	1.2×10^{-2}
MC	1.7×10^{-2}	3.7×10^{-2}	1.6×10^{-2}	2.0×10^{-2}
QMC	1.5×10^{-2}	3.9×10^{-2}	1.5×10^{-2}	2.2×10^{-2}

We now deal with the cases $n = 3, 4$, four estimation methods are tested: the MC and QMC based MLE, the CLO and the HDOI method. The other three ones are not considered due to their poor

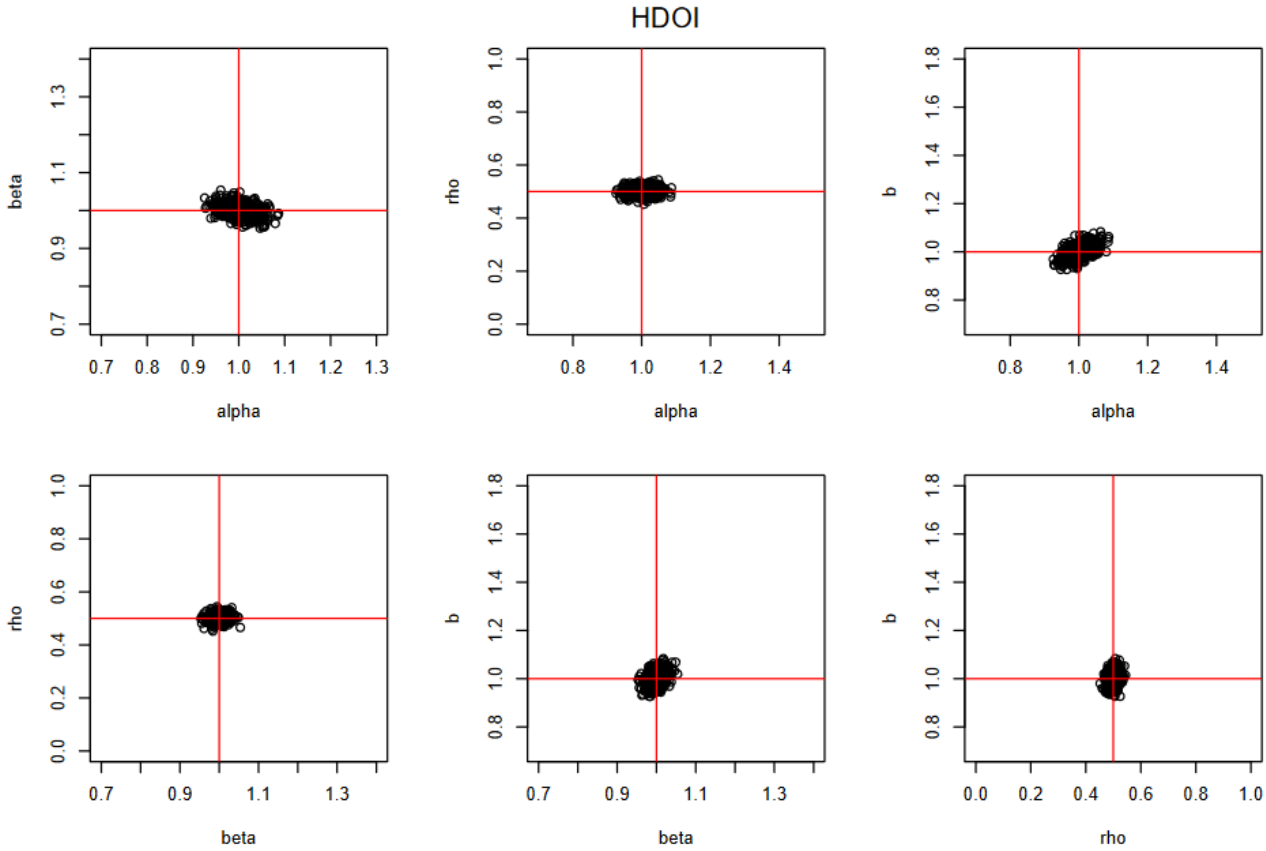

Figure 7.3: Pairs plots HDOI estimates. The red lines indicate the true value of each parameter.

quality. The estimation framework is the same as before except for the number of observed systems, which is reduced to $s = 100$. Note that even if $n < 5$ here, the identifiability for the HDOI method holds as $\beta = 1$ (**Corollary 2**). The results are given in **Table 7.4** to **Table 7.7**.

Table 7.4: Relative bias and variances for the CLO method when $n = 3, 4$

Relative bias (%)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	0.43%	42.56%	3.44%	0.39%
$n = 4$	0.29%	25.44%	6.36%	0.13%
Variances	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	5.7×10^{-3}	0.62	0.11	1.1×10^{-3}
$n = 4$	5.3×10^{-3}	0.39	0.08	7.0×10^{-4}

Table 7.5: Relative bias and variances for the MC method when $n = 3, 4$

Relative bias (%)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	2.98%	0.12%	1.46%	2.80%
$n = 4$	6.65%	0.28%	0.43%	4.34%
Variances	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	1.7×10^{-2}	2.6×10^{-2}	8.6×10^{-3}	1.8×10^{-2}
$n = 4$	8.8×10^{-3}	9.4×10^{-3}	3.3×10^{-3}	1.3×10^{-2}

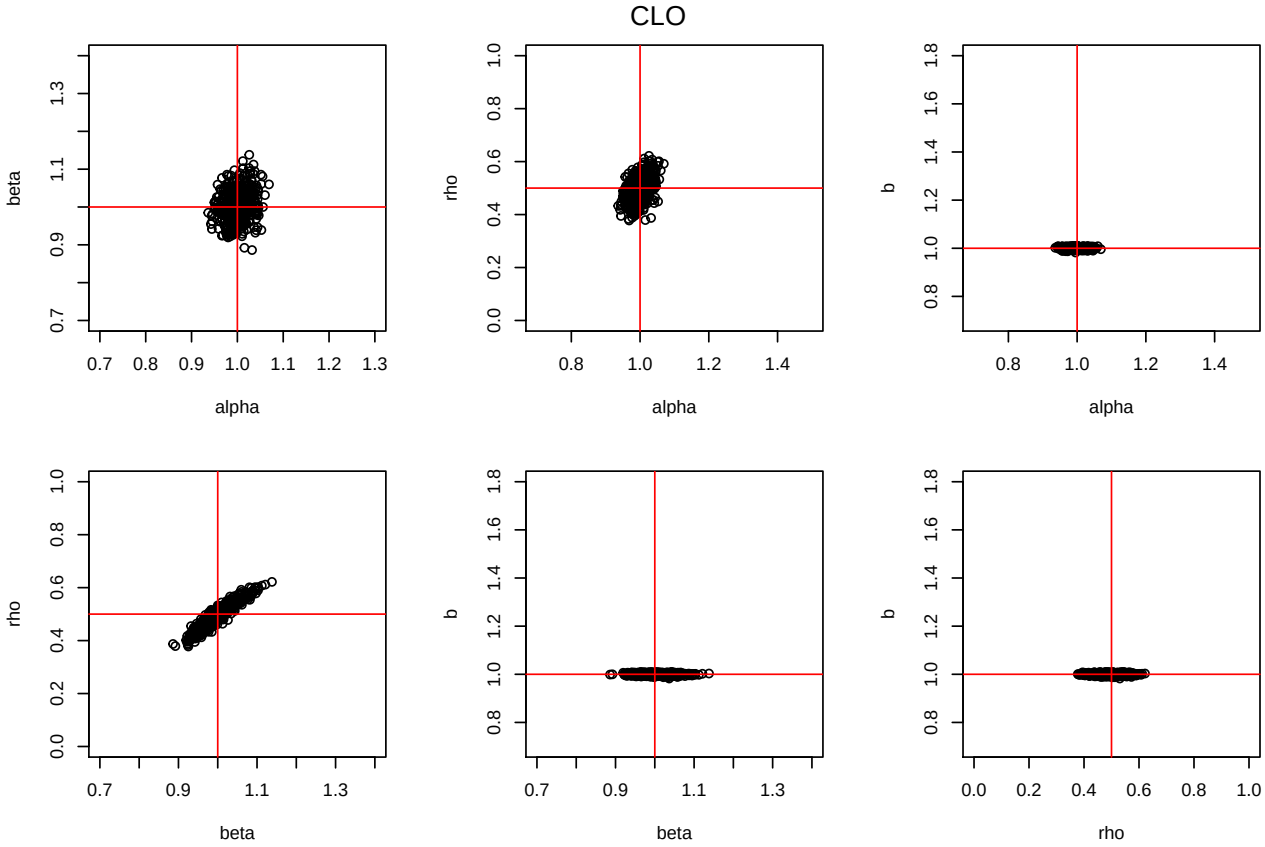

Figure 7.4: Pairs plots of CLO estimates. The red lines indicate the true value of each parameter.

Table 7.6: Relative bias and variances for the QMC method when $n = 3, 4$

Relative bias (%)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	6.14%	4.93%	3.19%	7.27%
$n = 4$	5.01%	1.92%	3.17%	4.57%
Variances	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
$n = 3$	1.5×10^{-2}	2.8×10^{-2}	7.1×10^{-3}	1.5×10^{-2}
$n = 4$	1.1×10^{-2}	1.1×10^{-2}	5.5×10^{-3}	1.0×10^{-2}

Table 7.7: Relative bias and variances for the HDOI method when $n = 3, 4$

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{b}$
Relative bias (%)	2.22%	2.19%	1.73%	3.54%
Variances	1.3×10^{-2}	1.7×10^{-2}	8.5×10^{-3}	1.8×10^{-2}

Regarding the small size of the data sets (100 trajectories observed at 4 times), the estimations of α and b by the CLO method are rather good. However, this method is not reliable for estimating parameters β and ρ.

In regards to the MC and QMC methods, the first one is clearly better in terms of bias while both have similar variances. Hence, the MC method is better than the QMC method. Also, unlike the CLO method, the MC method provides good results regarding β and ρ, and the estimation quality of α and b is acceptable.

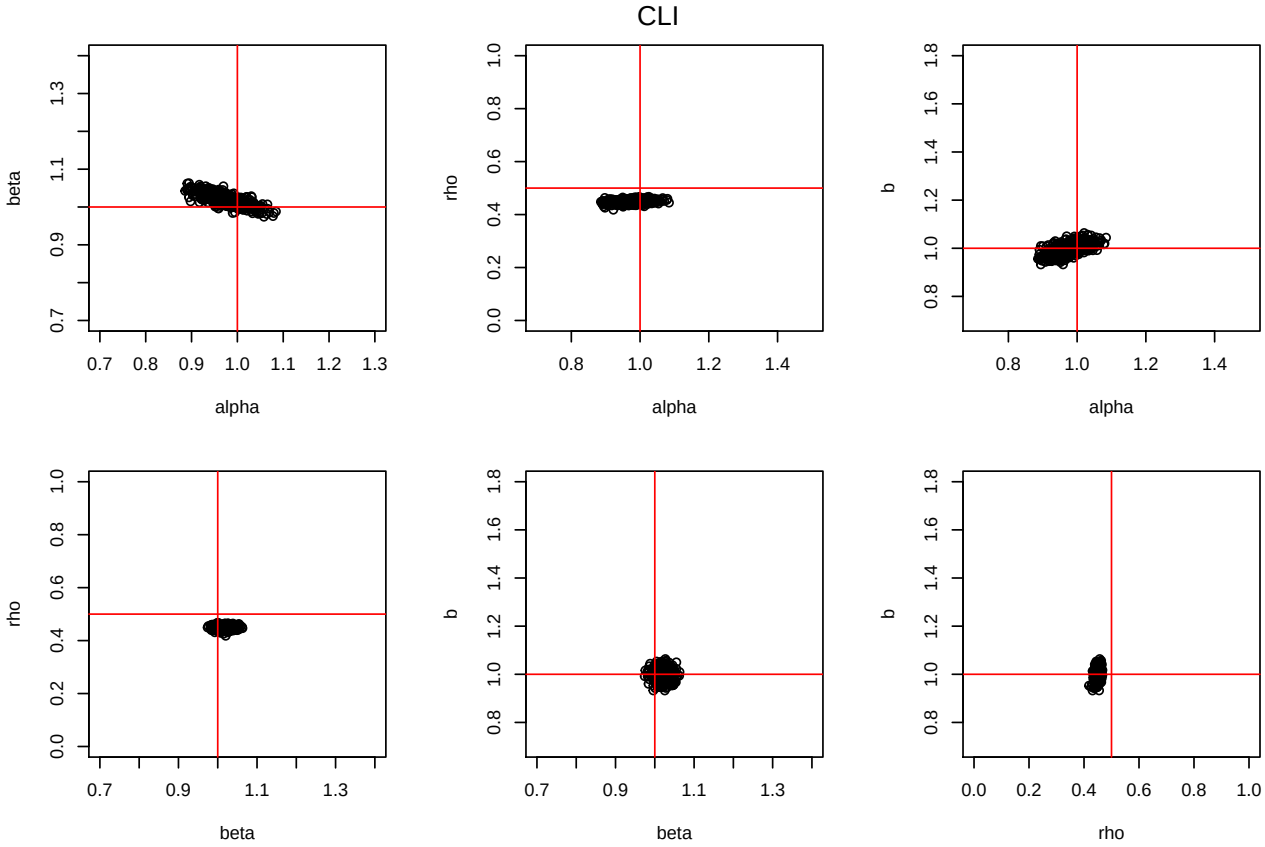

Figure 7.5: Pairs plots of CLI estimates. The red lines indicate the true value of each parameter.

Finally, the HDOI method does not differ for $n = 3$ and 4 since it considers the first and third observations only. This method is slightly more efficient for estimating b when $n = 3$ and α when $n = 3$ or 4 than the MC method, and both methods have similar variances.

If $n = 2, 3, 4$, no method stands out clearly from the others.

Let us recall that the MC and QMC approximations deal with the integrals of the function $G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{z_{i,n}})$ with respect to $\boldsymbol{x}$ over $[0,1]^{n-1}$, for $1 \leq i \leq s$, whose expression is given by

$$\begin{aligned} G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{z_{i,n}}) = & \prod_{j=1}^{n-1} f_{V_j}\left(\left(C_j^{(i)} - m_{i,j+1}\right) x_j + m_{i,j+1}\right) \\ & \times \prod_{j=1}^{n-1} f_{U_j}\left(z_j + \left(C_{j-1}^{(i)} - m_{i,j}\right) x_{j-1} - \left(C_j^{(i)} - m_{j+1}\right) x_j + m_{i,j} - m_{i,j+1}\right) \\ & \times f_{W_n}\left(z_{i,n} + x_{n-1}\left(C_{n-1}^{(i)} - m_{i,n}\right) + m_{i,n}\right) \end{aligned}$$

where $C_j = \sum_{i=1}^{j} z_j$ and $m_j = \max(0, -z_j)$. We now study the function G and the integral approximation of this function through the following example.

Example 9. *Let us set $n = 3$ and $s = 4$. An observations set $\boldsymbol{z_{4,3}}$ is generated within the framework of this chapter for the model parameters, leading to the following values:*

$$\boldsymbol{z_{4,3}} = \begin{pmatrix} 1.568 & -1.013 & -0.116 \\ 0.632 & 0.075 & 0.218 \\ 0.032 & 0.183 & 0.342 \\ 0.008 & 3.025 & 0.118 \end{pmatrix}$$

where the component $z_{i,j}$ is the degradation level of the i-th system at time jT^-. For each $1 \leq i \leq s$, the function $G_{\boldsymbol{\theta_0}}\big((x_1, x_2), \boldsymbol{z_{i,3}}\big)$ is evaluated for (x_1, x_2) in $(0,1)^2$. The level sets of this function are plotted in ***Figure 7.6*** *for $i = 1, 2, 3, 4$ from left to right, top to bottom. It can be seen that there is a region where the function G is zero, which highly depends on the observations and can reach more than half of the integration region. Regarding the remaining area where G is non-zero, the shape of the function is similar in each plot, that is:*

- *slowly increasing from the light grey area, where it is minimal, to the darker grey area;*
- *quickly increasing towards infinity at the edge (red area).*

Within this framework, the shape parameter of the U_j's and V_j's is 0.5. Hence, their p.d.f., which are considered in the expression of G, are equal to the function

$$y \mapsto \frac{\exp(-y)}{\Gamma(0.5)\sqrt{y}} \mathbb{1}_{(0,\infty)}(y).$$

Regarding the r.v. U_1, we have $y = z_1^{(i)} - \left(z_1^{(i)} - m_{i,2}\right) x_1 - m_{i,2}$ with $x_1 \in [0,1]$ and $m_{i,2} = \max(0, -z_{i,2})$. Hence, when x_1 tends towards 1 then y tends towards zero. In this case f_{U_1} tends to ∞, this explains why G tends towards ∞ at this edge, which also can be seen in ***Figure 7.6****. Similar results are obtained regarding the other distributions, which leads to the observed growth over almost all the edges. Note that this particularity is not an issue regarding the p.d.f. of W_n since its shape parameter is 1 here.*

Now, we are interested in the importance of the edge in the integral approximation. With this aim, the integrals are computed with roughly 5×10^5 points for the MC. The proportion of points belonging to the red area as well as the proportion of the integral arising from the red area are summarized in ***Table 7.8****.*

Table 7.8: Integral approximations characteristics on the edges

i	Integral of G over $(0,1)^2$	Proportion coming from the red area	Proportion of points falling in the red area
1	0.1	13%	0.7%
2	0.57	19%	0.3%
3	74.3	6%	0.2%
4	0.64	11%	0.2%

The weight of the edges is high because, for each computation, 6% to 19% of the integral arise from the red area while this area represents less than 1% of the integral domain.

The previous example leads us to the conclusion that the MC and QMC based methods could be improved by using importance sampling. Another alternative could be to consider a larger number of

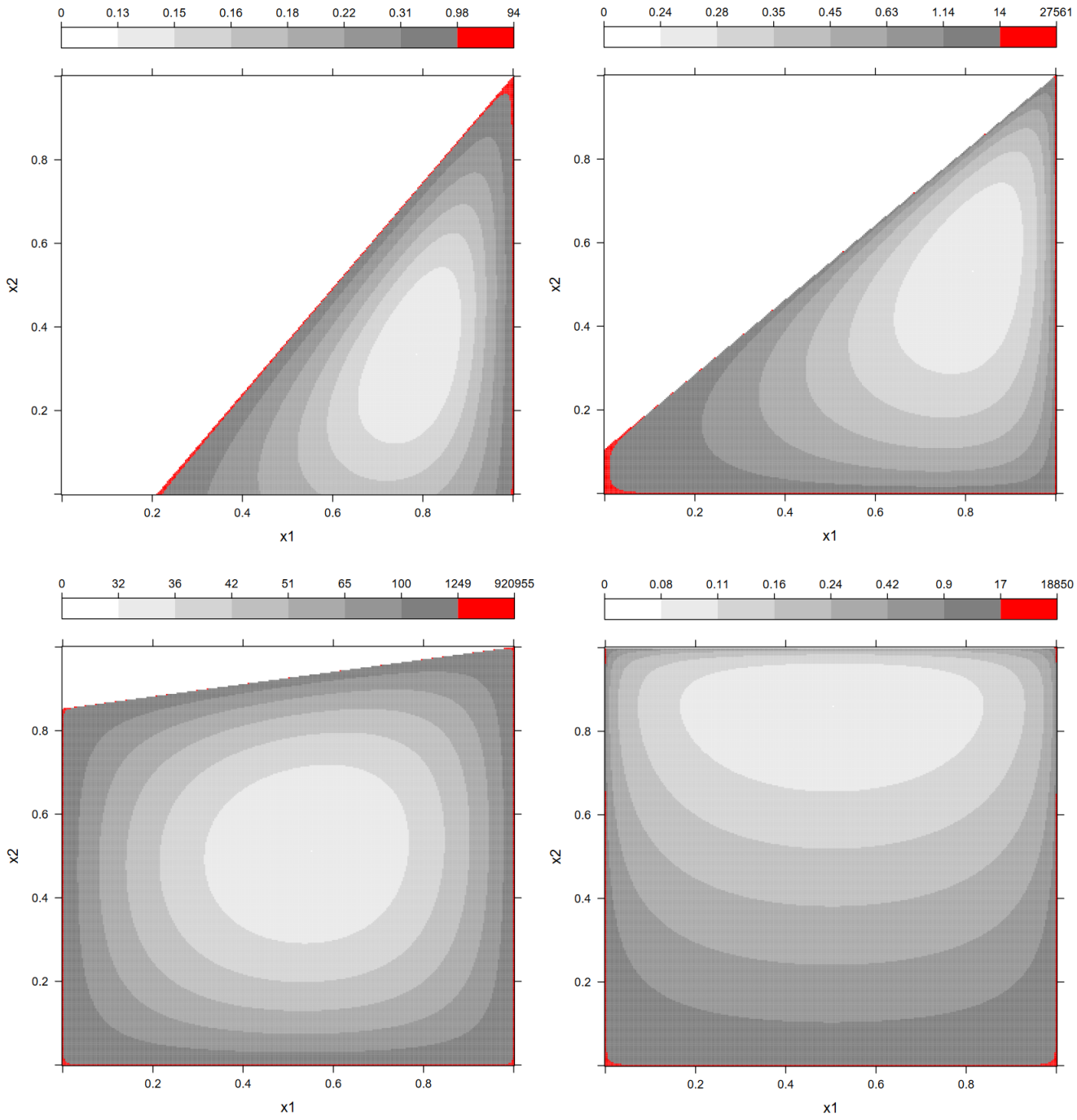

Figure 7.6: Levels of the function $G_{\boldsymbol{\theta_0}}(x_1, x_2, \boldsymbol{z_{i,3}})$ depending on x_1, x_2 and for $i = 1, 2, 3, 4$ from left to right, top to bottom.

points. However, the mean CPU time required for the retained estimation methods is another point that needs to be enlightened, our CPU times calculations are summarized in **Table 7.9**.

The mean CPU time for both MC and QMC methods is around 5 to 42 times higher than for the others methods, and this gap is proportional to the number of points for the integral approximations. Finally, the MC method is preferred to estimate the parameters when associated with a higher number of points, because of its efficiency. Also, the use of importance sampling could reduce the CPU time.

Table 7.9: Mean CPU times for one single estimation when $n = 2, 3, 4$, based on 500 repetitions

	MC and QMC	MLE	HDOI
$n = 2$	10min	2min	NA
$n = 3$	16min	NA	28s
$n = 4$	20min	NA	28s

However, due to their lower CPU times, the MLE method, without using neither MC nor QMC approximations, or the HDOI method can be more suitable to perform further investigation as the estimation of the variance estimators by boostrapping.

All the methods main characteristics are summarized in **Table 7.10**. Some of the proposed estimation methods can not be used in any case as they are not efficient enough when tested on large samples: the ME, the HDEI and the CLI methods. Also, depending on the number n of observations by trajectory, identifiability issues limit the number of possible methods.
As we just stated above, if $n < 5$, then either the MC method, or the MLE and the HDOI methods are preferred. When $n = 10$, the HDOI method as well as the CLO method provide good results. In order to determine which one of these two methods is the most efficient, they are tested on various samples of observations in the following section.

7.2 Large scale numerical tests

In this part, we deal with the CLO and the HDOI methods, which are tested on various observations sets. The considered model parameters as well as the observations characteristics are partly the same. To be more precise, the parameters α and b as well as the period of repairs T remain equal to 1, while the other parameters are varying as follows:

- Shape function parameter: $\beta = 0.5, 1, 1.5$;
- Maintenance actions efficiency: $\rho = 0.2, 0.5, 0.8$;
- Observations times: $\{jT^-; 1 \leq j \leq n\}$ with $n = 4, 16$;
- Number of observed i.i.d. systems: $s = 5, 25, 100, 400$

and the method is used 2000 times for each combination of those parameters. The estimations are computed as before, that is a maximization by the gradient method, searching (α, β, ρ) over $[0.1, 5]^2 \times [0.01, 0.99]$ for the CLO method, and (α, β, ρ, b) over $[0.1, 5] \times [0.1, 2.5] \times [0.01, 0.99] \times [0.1, 5]$ for the HDOI method. However, the initialization is different here: the initial value for the parameters is set at the middle of the search intervals, except in the case $\rho = 0.5$ where this parameter is initialized at 0.75.

Note that when $n = 4$ and $\beta = 1.5$, the identifiability for the HDOI method must be discussed depending on the value of ρ. When $\rho = 0.2, 0.5$, **Corollary 2** states that identifiability holds since $h_2(0.2, 1.5) < 1$ and $h_2(0.5, 1.5) < 1$, but this is not true if $\rho = 0.8$. However, the numerical study of the identifiability, whose results are illustrated in **Figure 6.4**, led us to conjecture that identifiability holds outside a given area, and the point $(\beta, r) = (\beta, 1 - \rho) = (1.5, 0.2)$ does not belong to this area.

Table 7.10: Characteristics summary for the methods selection

Methods	Identifiability	Numerical approximation of integrals	Estimations	Pros	Cons
ME	$n \geq 3$	None	Unbiased and high dispersion	No integral approximations and quick computation	Lack of accuracy
MLE	$n \geq 2$	Monte-Carlo and Quasi Monte-Carlo approximations (integrals of dimension $n-1$), or numerical approximations of one-dimensional integrals when $n = 2$	Unbiased only if the number of points for the integral approximation is sufficient, medium dispersion no matter the number of points. The MC based estimation is better	Identifiability holds as soon as $n \geq 2$	The number of points for the MC and QMC approximations does not affect the estimators' variances, but there a bias appears if it is too low. Unusable when n is too large because the number of points must be larger as well, which leads to non manageable computing time.
HDEI	Not studied	Numerical approximations of one-dimensional integrals	Unbiased and medium dispersion. Some estimates are very far from the true value	Accurate overall	It is possible for an estimation to be an outlier
HDOI	$n \geq 5$ (partially when $n = 3, 4$)	Numerical approximations of one-dimensional integrals	Unbiased and low dispersion	Accurate and reliable method	Identifiability holds over the entire parameters set only if $n \geq 5$
CLO	$n \geq 3$	None	Unbiased and medium to low dispersion	No integral approximations and quick computation, accurate and reliable method	The parameter β cannot be estimated properly (biased and high dispersion) when n is small, that is $n = 3, 4$.
CLI	Not studied	Numerical approximations of one-dimensional integrals	Highly biased (from 2% for $\hat{\alpha}$ to 10% for $\hat{\rho}$) and low dispersion	None	A high bias associated with a low dispersion make this method unusable

For each (β, ρ) in $\{0.5, 1, 1.5\} \times \{0.2, 0.5, 0.8\}$, the Relative Bias (RB) and the variances of the estimations are summarized in **Appendix C**. The figures deal either with the RB or with the variance, and are composed of four graphs: one for each parameter of the model. These graphs show four curves each, representing either the relative bias or the variance evolution with s for both the HDOI and the CLO methods when $n = 4$ and $n = 16$.
For the sake of clarity, as the RB can be higher than 200% as well as close to zero, instead of representing RB we chose to represent the quantity log(1+RB). Such a scale for the RB is unusual, hence, red dotted lines point out thresholds for which the RB is equal to 5%, 10%, 25% and 50%.
Finally, to be able to compare the results from one parameter to another, as well as the parameters sets, the y-axis range are the same for all plots. However, both the RB and the variance can be higher than the maximum value displayed in the y-axis. This signifies that if a curve does not appear on a graph, as the black one in the top right-hand graph of **Figure C.5**, each point of this curve is higher than this maximum.

As already stated in the previous part, the CLO method cannot be used if n is small because the estimations of β and ρ are biased. This can be seen once again when $n = 4$ through the figures exposed in **Appendix C**. In this case, the HDOI method is therefore the most reliable one whatever the value of β and ρ. Moreover, this method provides satisfying results even if only few observations are taken into account.

When $n = 16$, the value of the parameter ρ does not affect significantly the methods performance, unlike the parameter β. If $\beta = 0.5$, the CLO method appears to provide satisfying results while the HDOI method is biased with variances close to zero, and the value of s has little impact on both the bias and variances. This is clearly due to the number of observations the method deals with, because based on the same observations sets, this method is more efficient considering $n = 4$. We now try to understand why the performance of the HDOI method becomes poorer when n increases and look at an example.

Example 10. *We place ourselves within the following framework:* $n = 16$, $T = 1$ *and* $\boldsymbol{\theta} = (\alpha, \beta, \rho, b) = (1, 0.5, 0.5, 1)$. *The expectations of the components of* $\boldsymbol{Z}$ *are given by*

$$\mathbb{E}(\boldsymbol{Z}) = \left(\omega_j(\boldsymbol{\xi}) - \nu_{j-1}(\boldsymbol{\xi})\right)_{1 \leq j \leq 16}$$

and plotted in ***Figure 7.7****. This example shows that if the shape function for the underlying degradation process is concave, then the increments of the degradation get smaller and smaller, and vary more and more slowly as* j *increases.*

In **Example 10**, we can see that a small value for β quickly entails small deterioration increments between PM actions. Let us recall that the maintenance actions put back the system in the situation it was ρT units of times earlier. Hence, because of the slow degradation rate, the maintenance actions have a small impact on the degradation since the system was near the same situation before a repair, whatever its efficiency. As a result, this induces a numerical identifiability issue, leading to poor quality estimation results. Moreover, the HDOI method only takes into account one out of two increments, which reinforces this problem.

As a result, the first observations are the most important ones, and considering more observations which contain few additional information about the model parameters adversely affects the estimation quality. However, $\beta = 0.5$ leads to a concave shape function and thus the degradation intensity is

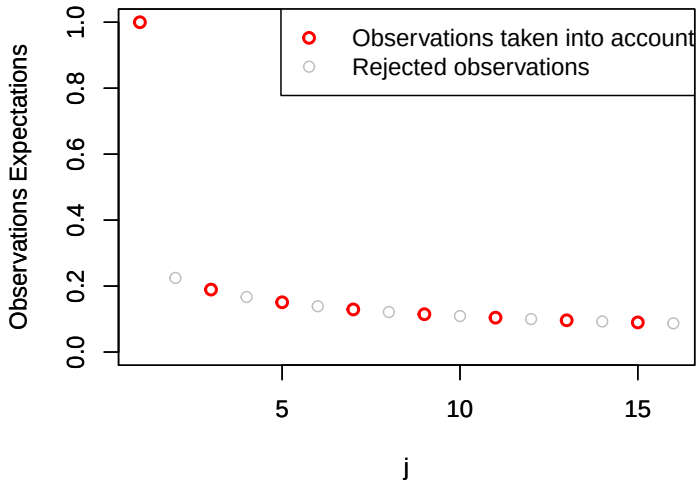

Figure 7.7: Expectations of the degradation increments with respect to j. The red dots matches the information taken into account in the HDOI method while the grey ones are not.

decreasing over time. Therefore, the maintenance actions reduce the deterioration level of the system, but they also induce an increase in the degradation intensity as the system is put back in a state it was earlier, which is inconsistent. As a conclusion, the ARA models are not the most suitable models in this case.

If $\beta = 1$, the CLO and the HDOI methods provide similar results for the estimations of α. Regarding b, the first one is better than the second one, while it is the opposite regarding β and ρ. Finally, when $\beta = 1.5$, both methods provide similar results for the estimations of β and ρ, and the CLO method is better in estimating α and b.

Given these characteristics, the choice of the method depends on the context, that is the degradation intensity, the maintenance actions efficiency and the number of observations by trajectory. This choice can be done through **Table 7.11**.

Table 7.11: Summary for the choice of the estimation method

	Decreasing degradation intensity ($\beta < 1$)	Constant degradation intensity ($\beta = 1$)	Increasing degradation intensity ($\beta > 1$)
$n = 4$	HDOI		
$n = 16$	CLO if ρ is small, otherwise HDOI reducing n to 4	HDOI	CLO

Overall, the HDOI method overcomes the CLO method. However, the CLO method always provides good estimations of α and b, whatever the values of β and n. Furthermore, the HDOI method better estimates β and ρ, reducing artificially the value of n in some cases. Hence, further work could be to improve the estimation quality by combining both the HDOI and the CLO methods. One approach is to keep the estimations of α and b provided by the CLO method and those of β and ρ from the other one.

In a similar way, plug-in estimations could be tested, that is substituting the values $\hat{\alpha}$ and $\hat{b}$ obtained by the CLO method in the log-likelihood of the HDOI method and then estimate β and ρ, as well as the opposite.

Part IV

Conclusion

This work has studied imperfect repair models in the context of gamma deteriorating systems. More specifically, estimation methods have been developed for the arithmetic reduction of degradation models of order one and infinity (ARD_1 and ARD_∞) in **Part II**, as well as for the arithmetic reduction of age model of order one (ARA_1) in **Part III**.

In **Part II**, we first dealt with two classical estimation methods in a fully parametric framework, namely the Moments Estimation and the Maximum Likelihood Estimation (ME and MLE) methods. The parameters identifiability was studied, estimators expressions were provided and finally the performance of the methods was illustrated. After that, arising from the study of the MLE method, an estimator for the maintenance actions efficiency (parameter ρ) was proposed in a semiparametric framework. In the case of one single trajectory and when the shape function of the underlying gamma process is concave, this estimator was proved to be strongly consistent as the number of repairs tends towards infinity, with a surprisingly high convergence rate, at least exponential for some particular cases. This work was next extended to the case where s independent and identical systems are observed. A similar semiparametric estimator was proposed for ρ and the strong consistency was proved to hold as s tends towards infinity, no matter the fixed number of repairs and out of any technical condition requirement. The convergence rate was also studied, which was shown to depend on the shape function of the gamma process and on the maintenance period, leading to a speed that can be either slower or faster than $\sqrt{s}$, according to the case.

In **Part III**, we proposed several estimations methods for an ARA_1 model, which are based on either the observations or on the increments, leading to six different estimation methods. Depending on the method, either expressions of the estimators or the log-likelihood are provided, and the identifiability of the model parameters was studied regarding four out of the six methods. In order to study and compare the performance of all these methods, numerical investigations based on simulated data have been conducted at a large scale. Two of them has appeared to be more efficient than the other ones: the Half Data based on the Odd Indexes (HDOI) and the Composite Likelihood based on the Observations (CLO) methods. Overall, however, none of these two methods stands out clearly from the other. Thus the choice of the method must be done with respect to the context. To be more specific, this choice depends on the number of repairs, the repairs efficiency and the degradation intensity evolution over time.

Several points of interest would be interesting to study in complement to this thesis. There exist many other models that extend the ARD models studied here, such as for instance the ARD_m model for which the basic idea is that a maintenance action removes a proportion of the degradation accumulated by the system from the last m maintenance actions. An idea could be to generalize the semiparametric estimate of the maintenance efficiency to such a model, which seems to be possible as this method relies on the non negativity of the gamma process. Moreover, the adaptation of this estimation procedure to another monotonous Lévy process than the gamma process would be interesting to study. Finally, based on a short numerical study, we have seen that the condition which ensures an exponential convergence rate was sufficient but not necessary. Further investigations could be done in order to see whether it could be possible to refine the mathematical conditions under which the different convergence rates are obtained.

Regarding the work done in **Part III**, several estimation procedures could be improved. As already stated, it could be interesting to see whether the estimation quality is better by combining the HDOI and the CLO methods. Moreover, we explained that the Monte Carlo and Quasi Monte Carlo integral

approximations required in the MLE method could be improved by using importance sampling. Further work should be done to address these possible improvements.

Beyond that, except in **Chapter 4**, asymptotic properties of the estimators were not investigated in this document and further work is required for a better understanding of their behaviour, beyond the numerical investigation performed in this thesis. One might study these properties as an extension of the present work.

Besides all of this, further work could be the development of estimation procedures in the context of an ARA_∞ model, for which a maintenance action removes a proportion of the age accumulated by the system since it was put into operation. More research might be done in this context, mimicking the work done regarding the ARA_1 model as a first step.

Finally, another point could be to consider an observation scheme which is decoupled from the maintenance schedule. It is not the case here as the degradation level is measured right before each repair. As an example, observations could still be conducted periodically while the system could be maintained at random times according to a Poisson process. However, preliminary investigations were done in this framework, and the degradation level of the maintained system seems difficult to write down with a simple expression, but that could remain exploitable for developing estimation procedures. Hence, this seems a challenging (but interesting) subject for further research.

Part V

References

Bibliography

[1] Abdel-Hameed, M. A Gamma Wear Process. *IEEE Transactions on Reliability*, R-24(2):152–153, 1975.

[2] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover, New York, ninth dover printing, tenth gpo printing edition, 1964.

[3] S. Alaswad and Y. Xiang. A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157:54–63, 2017.

[4] G. J. Babu. Subsample and half-sample methods. *Annals of the Institute of Statistical Mathematics*, 44(4):703–720, 1992.

[5] R. E. Barlow and F. Proschan. *Mathematical theory of reliability*. Wiley, New York, 1965.

[6] L. Bordes and S. Mercier. Extended geometric processes: semiparametric estimation and application to reliability. *Journal of the Iranian Statistical Society*, 12(1):1–34, 2013.

[7] L. Brenière, L. Doyen, and C. Bérenguer. Virtual age models with time-dependent covariates: A framework for simulation, parametric inference and quality of estimation. *Reliability Engineering & System Safety*, page 107054, 2020.

[8] I. T. Castro and S. Mercier. Performance measures for a deteriorating system subject to imperfect maintenance and delayed repairs. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 230(4):364–377, 2016.

[9] R. E. Chandler and S. Bate. Inference for clustered data using the independence loglikelihood. *Biometrika*, 94(1):167–183, 2007.

[10] E. Çinlar, Z. P. Bazant, and E. Osman. Stochastic process for extrapolating concrete creep. *Journal of the Engineering Mechanics Division*, 103(6):1069–1088, 1977.

[11] R. Cont and P. Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall / CRC Press, 2003.

[12] J.-Y. Dauxois, S. Gasmi, and O. Gaudoin. Semiparametric inference for an extended geometric failure rate reduction model. *Journal of Statistical Planning and Inference*, 199:14–28, 2019.

[13] M. L. G. de Toledo, M. A. Freitas, E. A. Colosimo, and G. L. Gilardoni. ARA and ARI imperfect repair models: Estimation, goodness-of-fit and reliability prediction. *Reliability Engineering & System Safety*, 140:107–115, 2015.

[14] *NIST Digital Library of Mathematical Functions.* http://dlmf.nist.gov/, Release 1.0.21 of 2018-12-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[15] P. Do, A. Voisin, E. Levrat, and B. Iung. A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. *Reliability Engineering & System Safety*, 133:22–32, 2015.

[16] L. Doyen and O. Gaudoin. Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering & System Safety*, 84(1):45–56, 2004.

[17] L. Doyen, O. Gaudoin, and A. Syamsundar. On geometric reduction of age or intensity models for imperfect maintenance. *Reliability Engineering & System Safety*, 168:40 – 52, 2017.

[18] M. Giorgio, F. Postiglione, and G. Pulcini. Bayesian estimation and prediction for the transformed Wiener degradation process. *Applied Stochastic Models in Business and Industry*, 2020.

[19] M. Giorgio and G. Pulcini. A new state-dependent degradation process and related model misidentification problems. *European Journal of Operational Research*, 267(3):1027–1038, 2018.

[20] M. Giorgio and G. Pulcini. A new age- and state-dependent degradation process with possibly negative increments. *Quality and Reliability Engineering International*, 35(5):1476–1501, 2019.

[21] J. A. Hartigan. Using subsample values as typical values. *Journal of the American Statistical Association*, 64(328):1303–1317, 1969.

[22] C.-H. Hu, M.-Y. Lee, and J. Tang. Optimum step-stress accelerated degradation test for wiener degradation process under constraints. *European Journal of Operational Research*, 241(2):412–4211, 2015.

[23] K. Huynh. Modeling past-dependent partial repairs for condition-based maintenance of continuously deteriorating systems. *European Journal of Operational Research*, 2019.

[24] W. Kahle. Imperfect repair in degradation processes: A kijima-type approach. *Applied Stochastic Models in Business and Industry*, 35(2):211–220, 2019.

[25] A. Khatab, C. Diallo, E.-H. Aghezzaf, and U. Venkatadri. Condition-based selective maintenance for stochastically degrading multi-component systems under periodic inspection and imperfect maintenance. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 232(4):447–463, 2018.

[26] M. Kijima. Some results for repairable systems with general repair. *Journal of Applied Probability*, 26(1):89–102, 1989.

[27] Y. Lam. *The geometric process and its applications.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007.

[28] C. Letot, P. Dehombreux, G. Fleurquin, and A. Lesage. An adaptive degradation-based maintenance model taking into account both imperfect adjustments and agan replacements. *Quality and Reliability Engineering International*, 33:2043–2058, 2017.

[29] B. Liu, S. Wu, M. Xie, and W. Kuo. A condition-based maintenance policy for degrading systems with age- and state-dependent operating cost. *European Journal of Operational Research*, 263(3):879–887, 2017.

[30] B. Liu, M. Xie, Z. Xu, and W. Kuo. An imperfect maintenance policy for mission-oriented systems subject to degradation and external shocks. *Computers & Industrial Engineering*, 102:21–32, 2016.

[31] R. Liu and T. Lux. Generalized method of moment estimation of multivariate multifractal models. *Economic Modelling*, 67:136–148, 2017.

[32] W. Q. Meeker and L. A. Escobar. *Statistical methods for reliability data*. Wiley series in probability and statistics. J. Wiley & Sons, New York, Chichester, 1998.

[33] S. Mercier and I. Castro. On the modelling of imperfect repairs for a continuously monitored gamma wear process through age reduction. *Journal of Applied Probability*, 50(4):1057–1076, 2013.

[34] S. Mercier and I. T. Castro. Stochastic comparisons of imperfect maintenance models for a gamma deteriorating system. *European Journal of Operational Research*, 273(1):237–248, 2019.

[35] R. Nicolai, J. Frenk, and R. Dekker. Modelling and optimizing imperfect maintenance of coatings on steel structures. *Structural Safety*, 31(3):234 – 244, 2009.

[36] H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

[37] A. Ponchet, M. Fouladirad, and A. Grall. Maintenance policy on a finite time span for a gradually deteriorating system with imperfect improvements. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 225(2):105–116, 2011.

[38] G. Salles, S. Mercier, and L. Bordes. Semiparametric estimate of the efficiency of imperfect maintenance actions for a gamma deteriorating system. *Journal of Statistical Planning and Inference*, 206:278 – 297, 2020.

[39] R. Storn and K. Price. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11:341–359, 1997.

[40] J. Van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1):2–21, 2009.

[41] C. Varin, N. Reid, and D. Firth. Composite likelihood methods. *Statistica Sinica*, 21(1):5–42, 2011.

[42] F. Wu, S. A. Niknam, and J. E. Kobza. A cost effective degradation-based maintenance strategy under imperfect repair. *Reliability Engineering & System Safety*, 144:234–243, 2015.

[43] S. Wu and I. T. Castro. Maintenance policy for a system with a weighted linear combination of degradation processes. *European Journal of Operational Research*, 280(1):124 – 133, 2020.

[44] Z.-S. Ye and N. Chen. The Inverse Gaussian Process as a Degradation Model. *Technometrics*, 56(3):302–311, 2014.

[45] M. Zhang, O. Gaudoin, and M. Xie. Degradation-based maintenance decision using stochastic filtering for systems under imperfect maintenance. *European Journal of Operational Research*, 245:531–541, 2015.

[46] M. Zhang and M. Xie. An ameliorated improvement factor model for imperfect maintenance and its goodness of fit. *Technometrics*, 59(2):237–246, 2017.

Part VI

Appendix

Appendix A

ARD_1 model, results of the simulation study

As already stated in **Section 3**, the results are exposed as follows.

For each possible combination of (β, ρ), the RB and the variances of the estimations are summarized in the following. The figures deal either with the RB or with the variance, and are composed of four graphs: one for each parameter of the model. These graphs show four curves each, representing either the relative bias or the variance evolution with s for both the ME and the MLE methods when $n = 5$ and $n = 20$.
To be able to compare the results from one parameter to another, as well as the parameters sets, the y-axis range is the same for all plots. However, both the RB and the variance can be higher than the maximum value displayed in the y-axis. This signifies that if a curve does not appear on a graph, as the black one in the top right-hand graph as well as the bottom left-hand one of **Figure A.1**, it means that all the points of the curve are out of range.

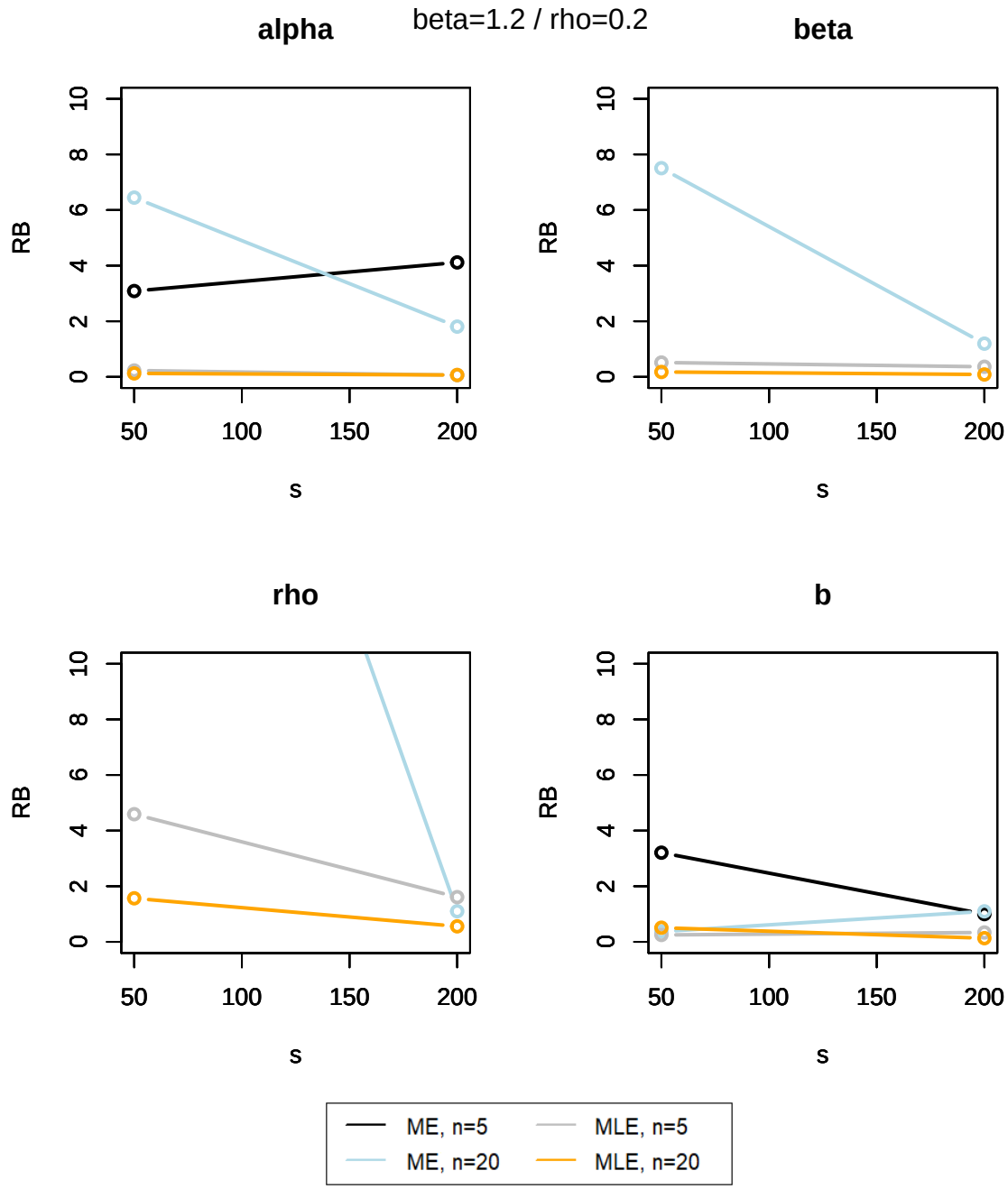

Figure A.1: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.2)$.

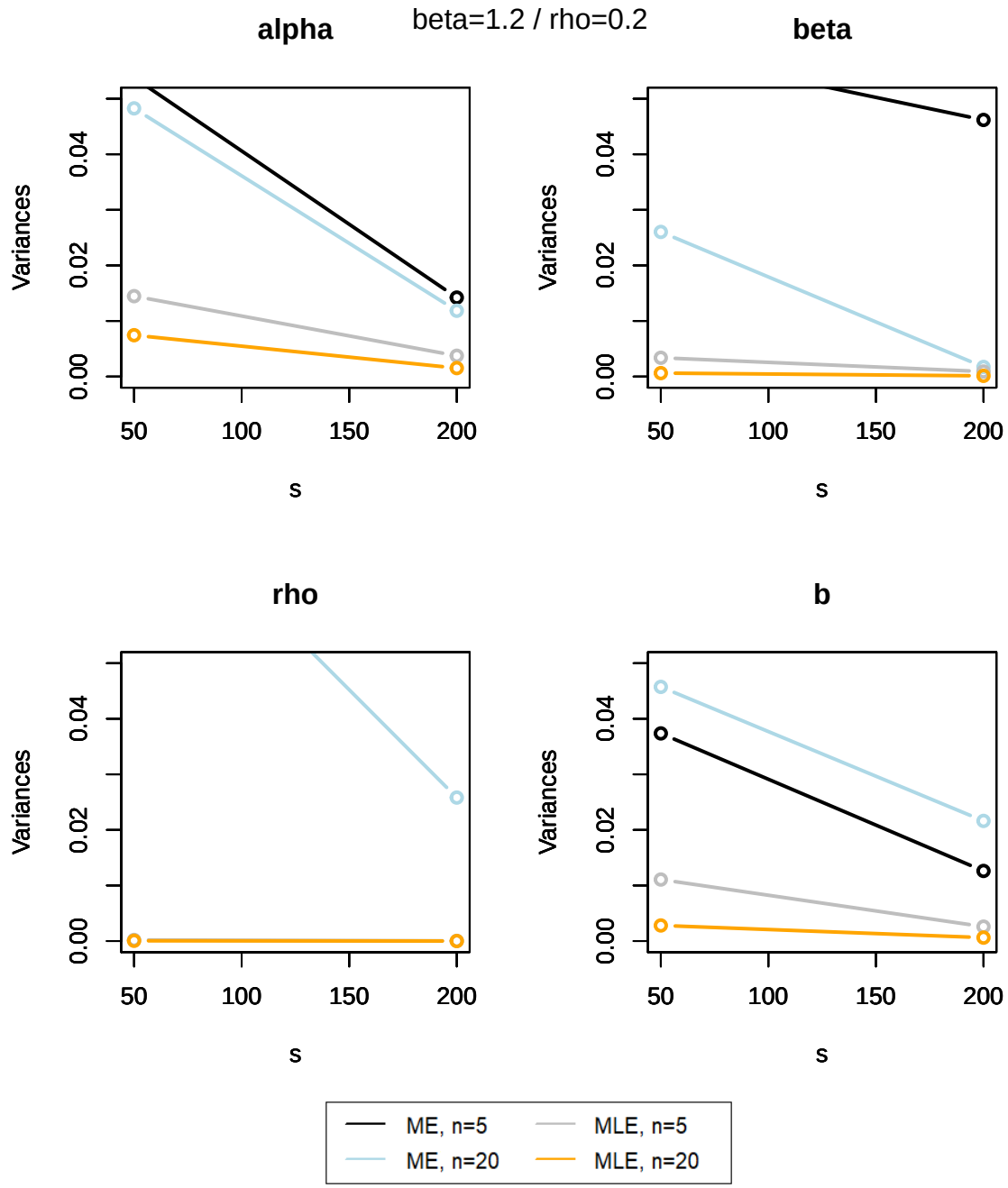

Figure A.2: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.2)$.

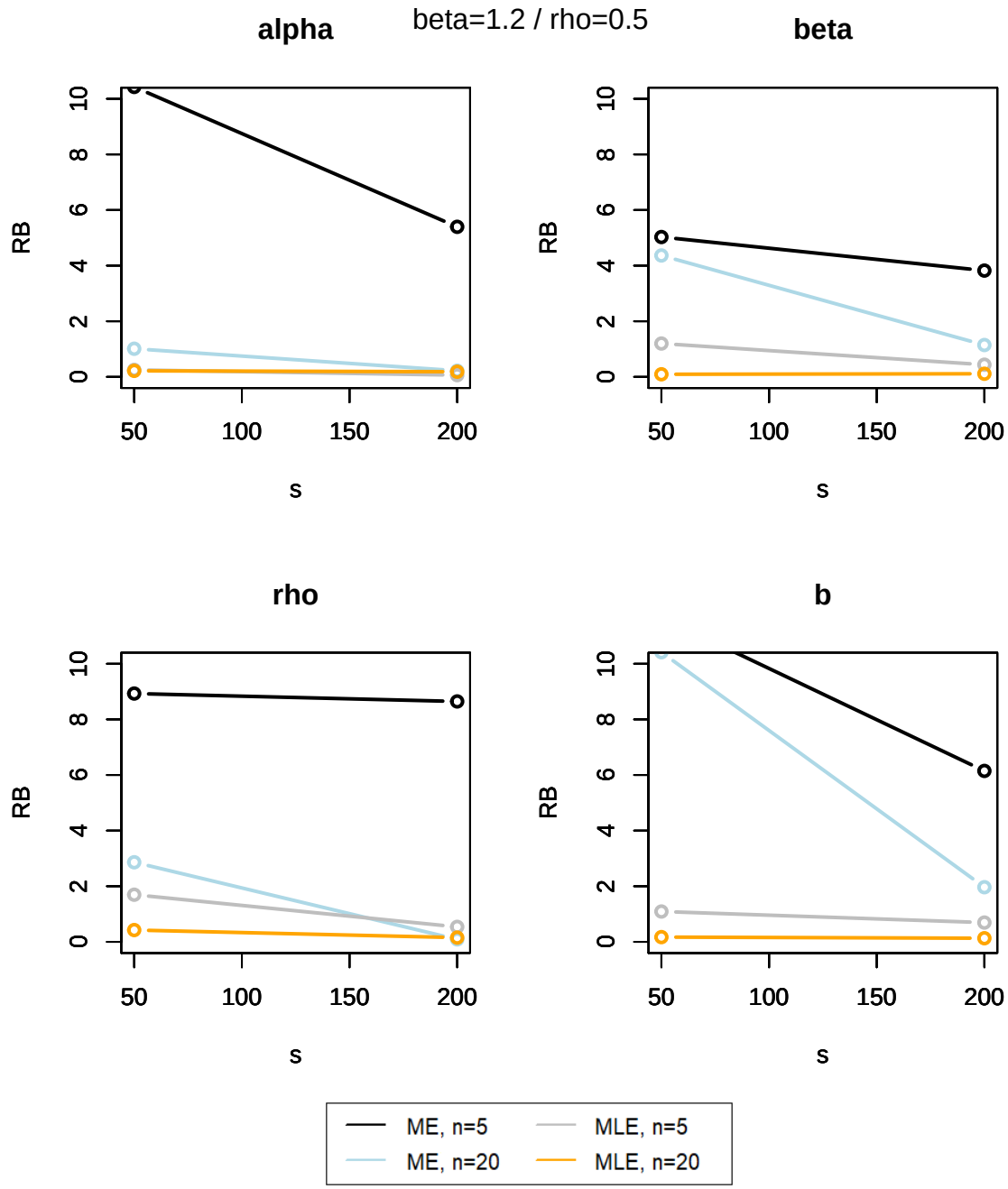

Figure A.3: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.5)$.

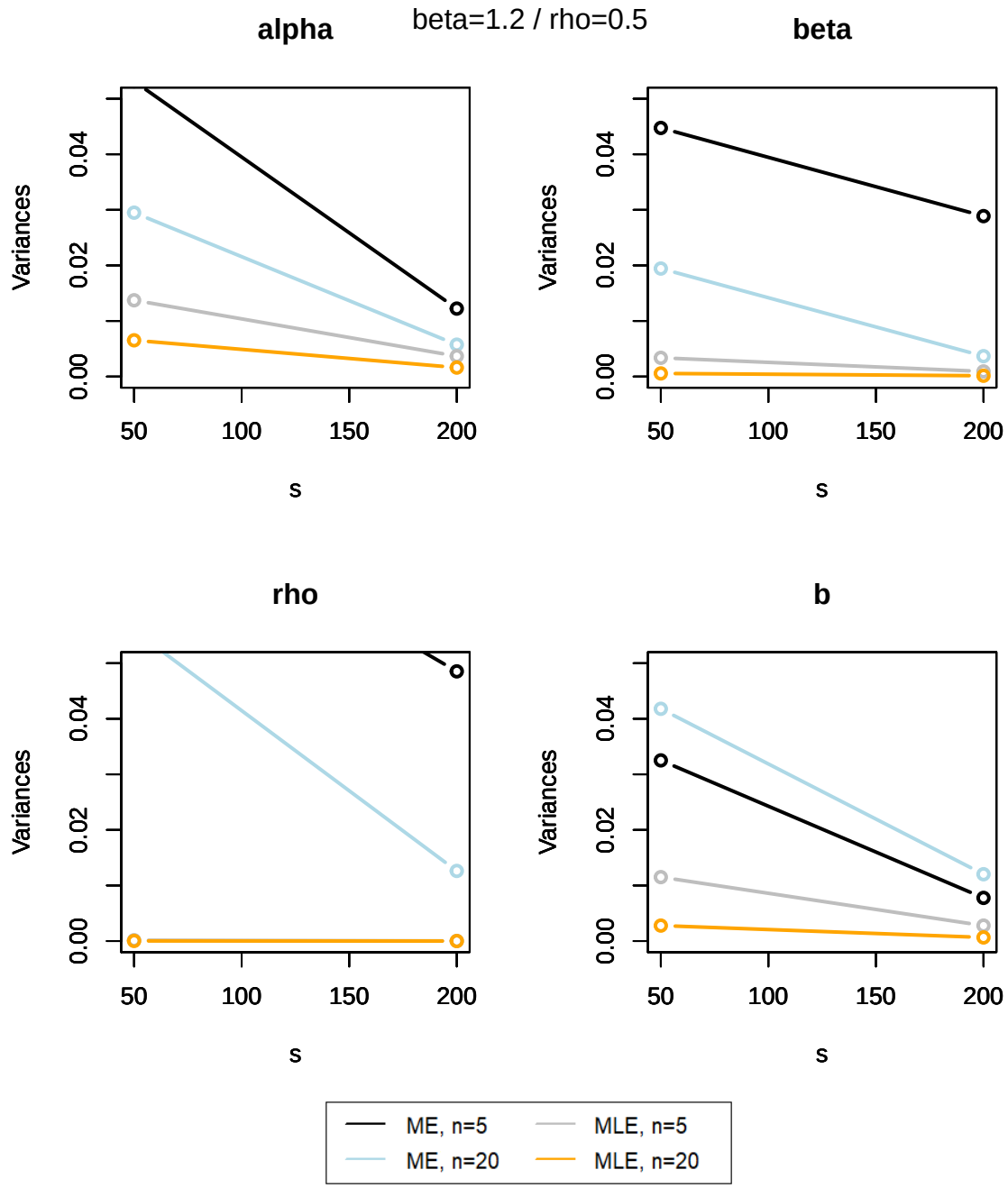

Figure A.4: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.5)$.

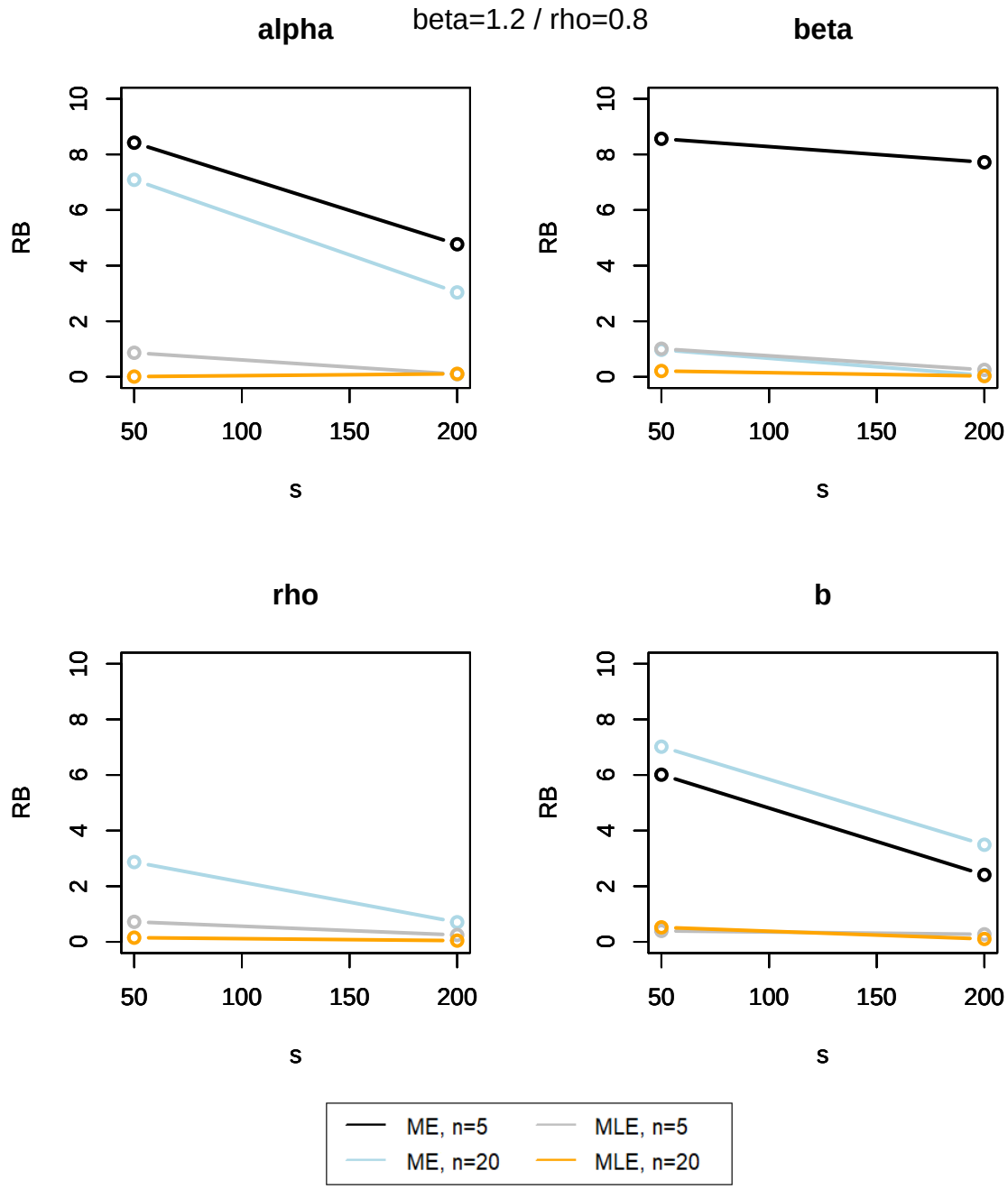

Figure A.5: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.8)$.

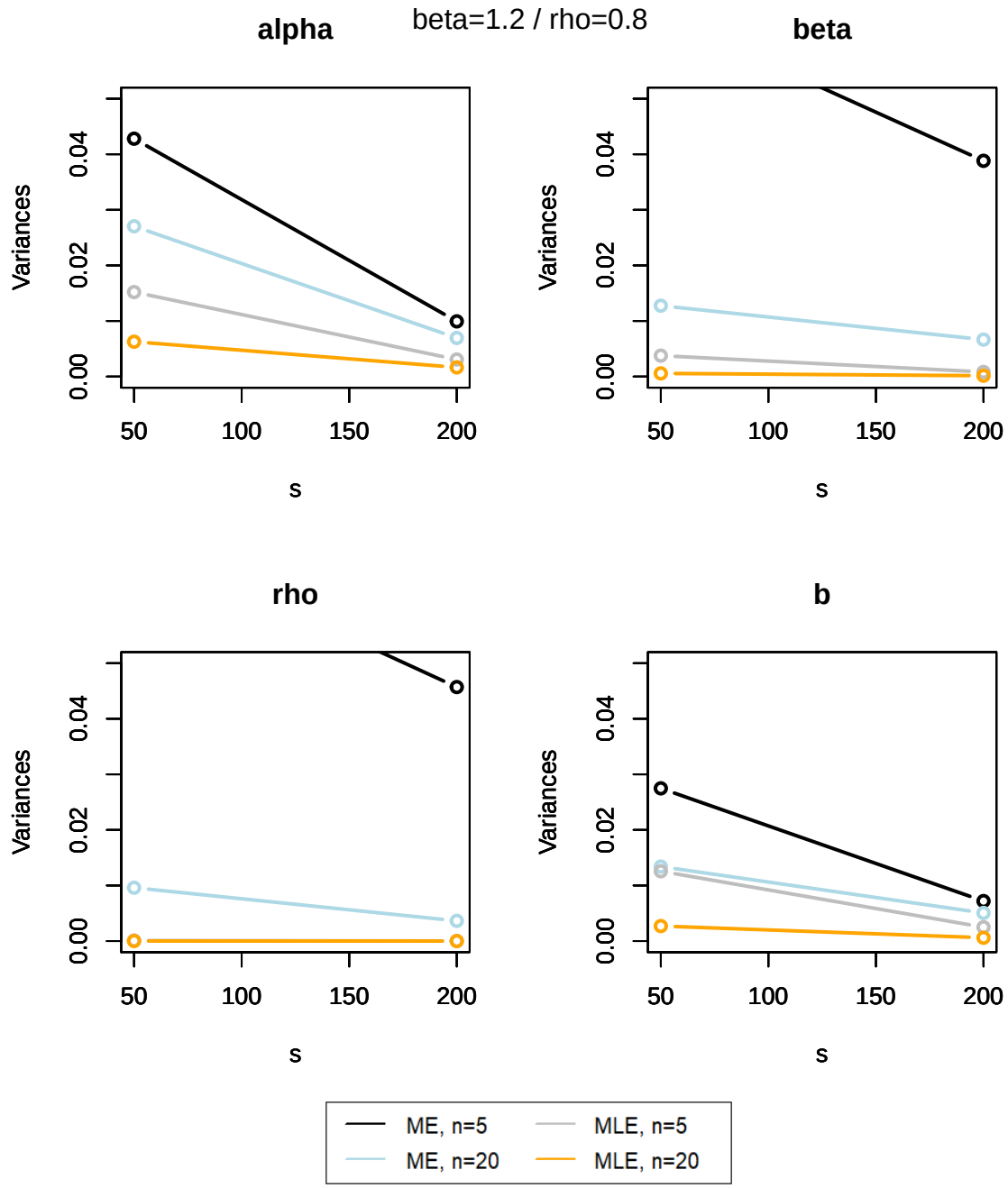

Figure A.6: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.8)$.

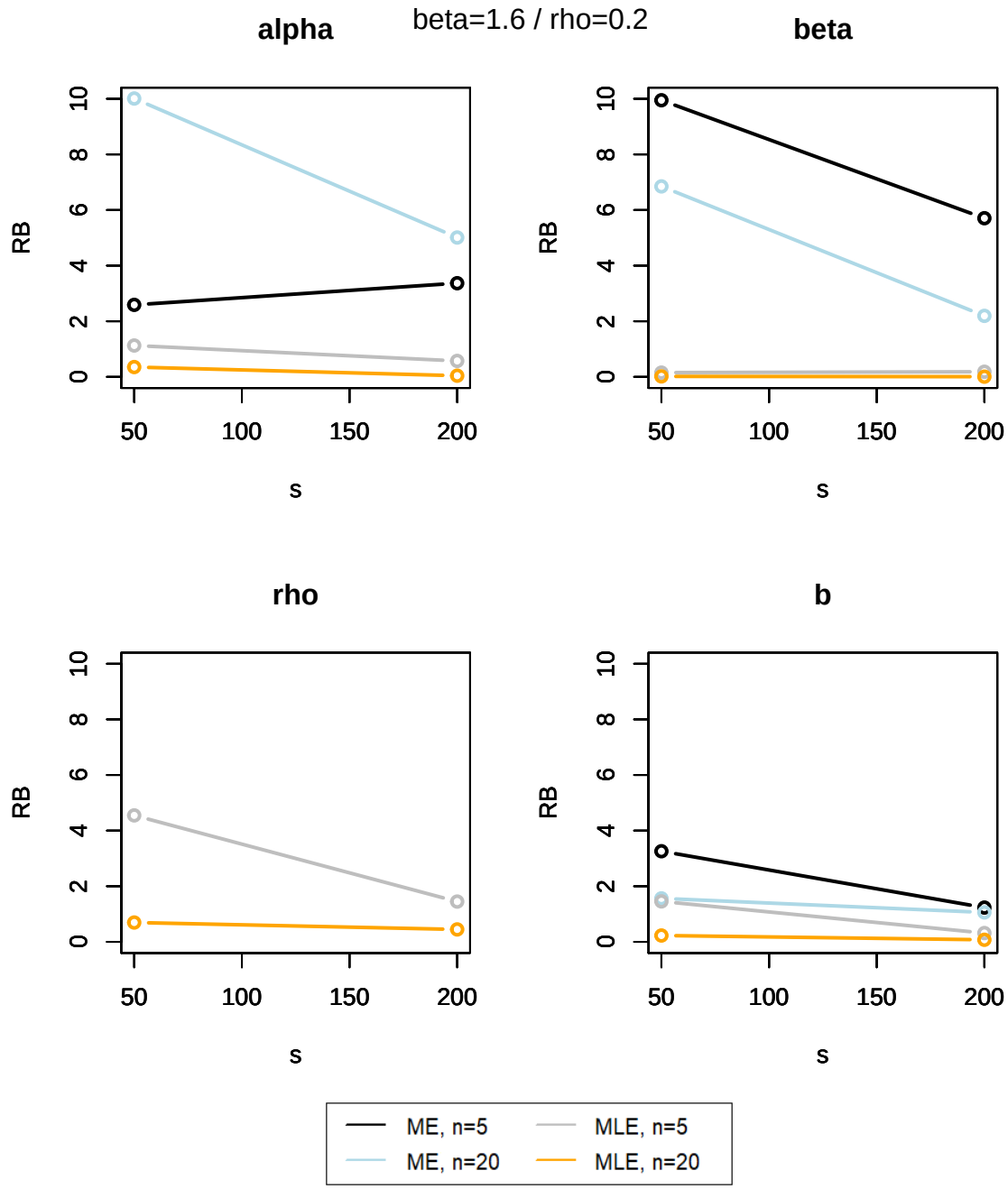

Figure A.7: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.2)$.

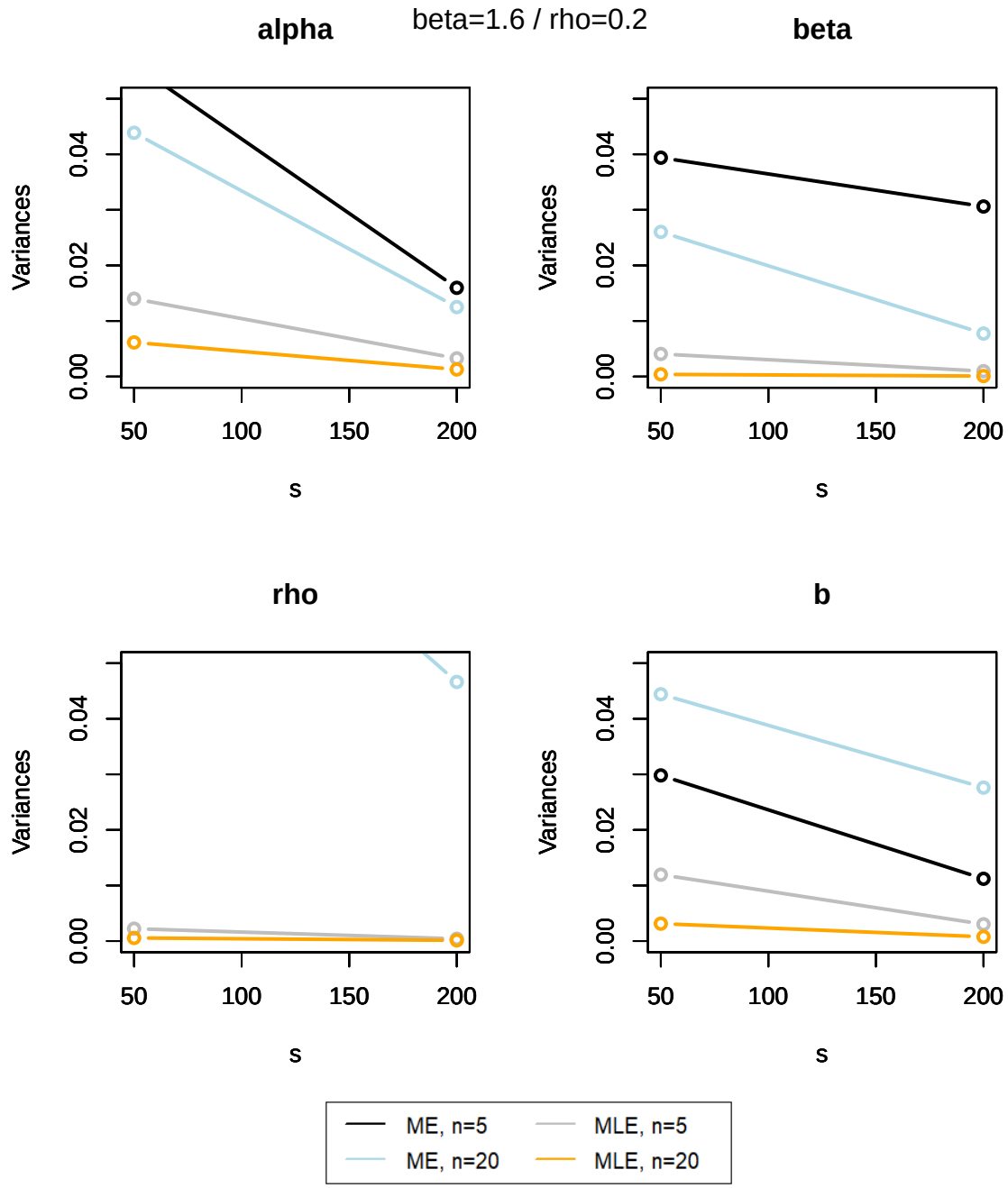

Figure A.8: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.2)$.

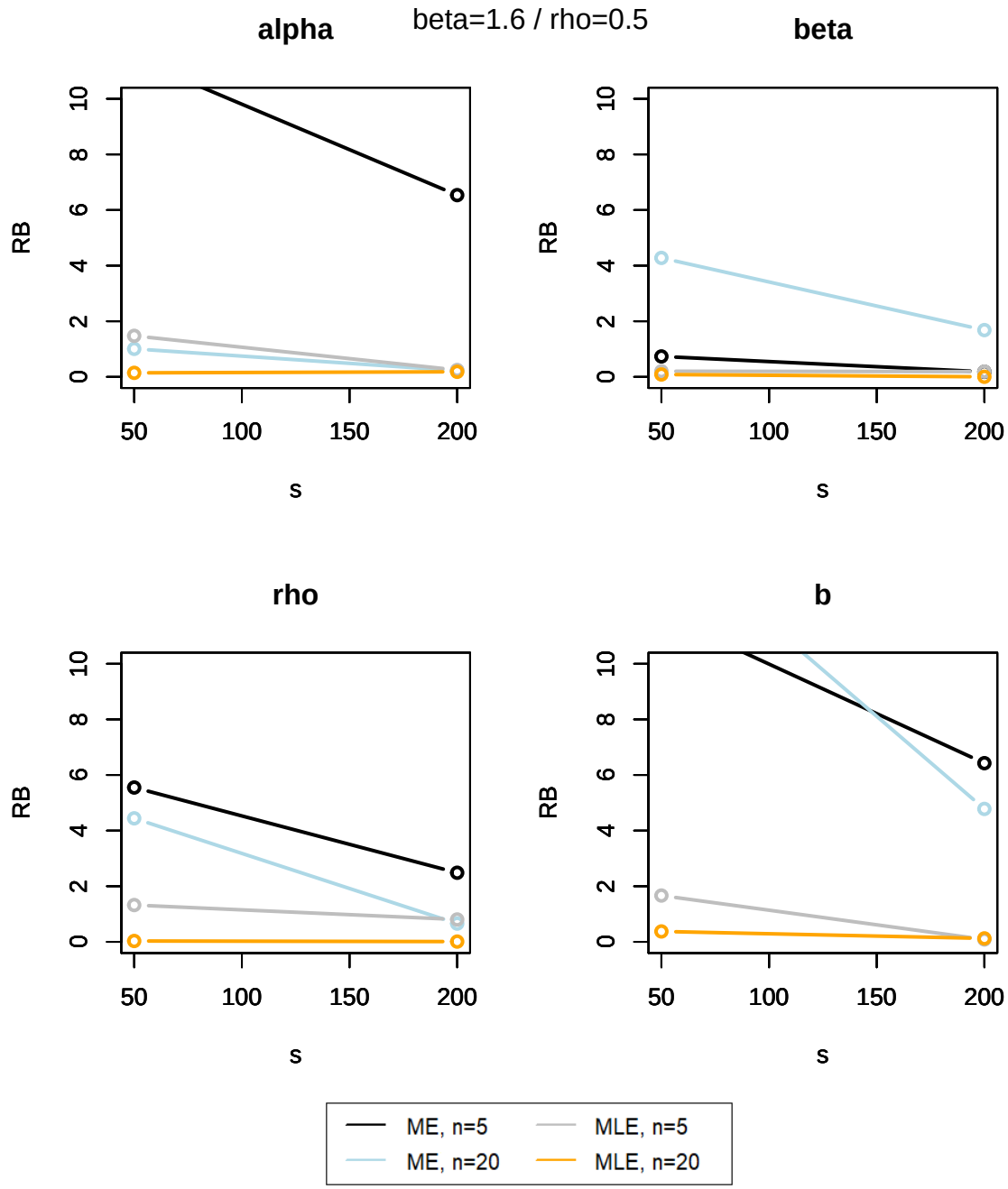

Figure A.9: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.5)$.

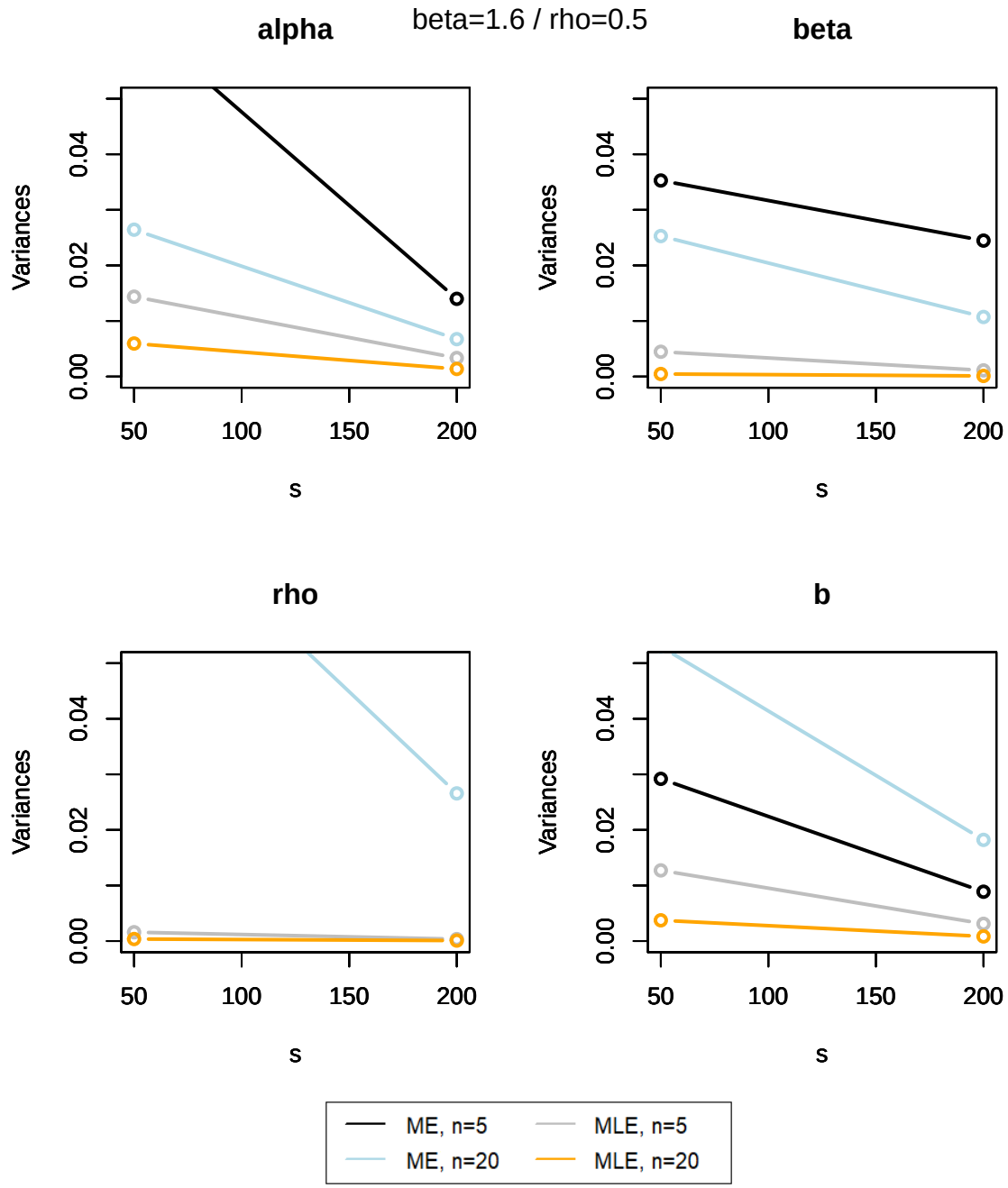

Figure A.10: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.5)$.

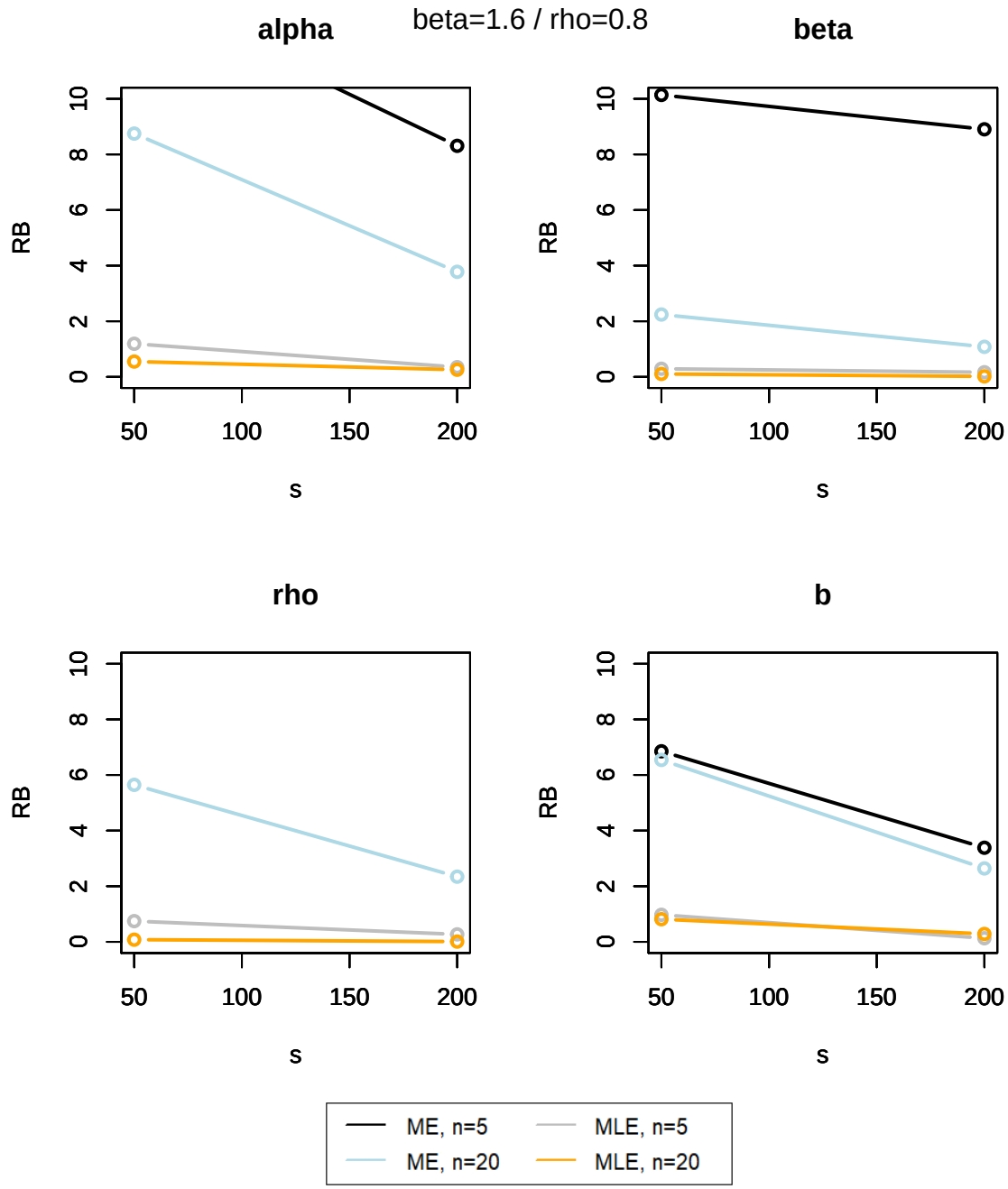

Figure A.11: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.8)$.

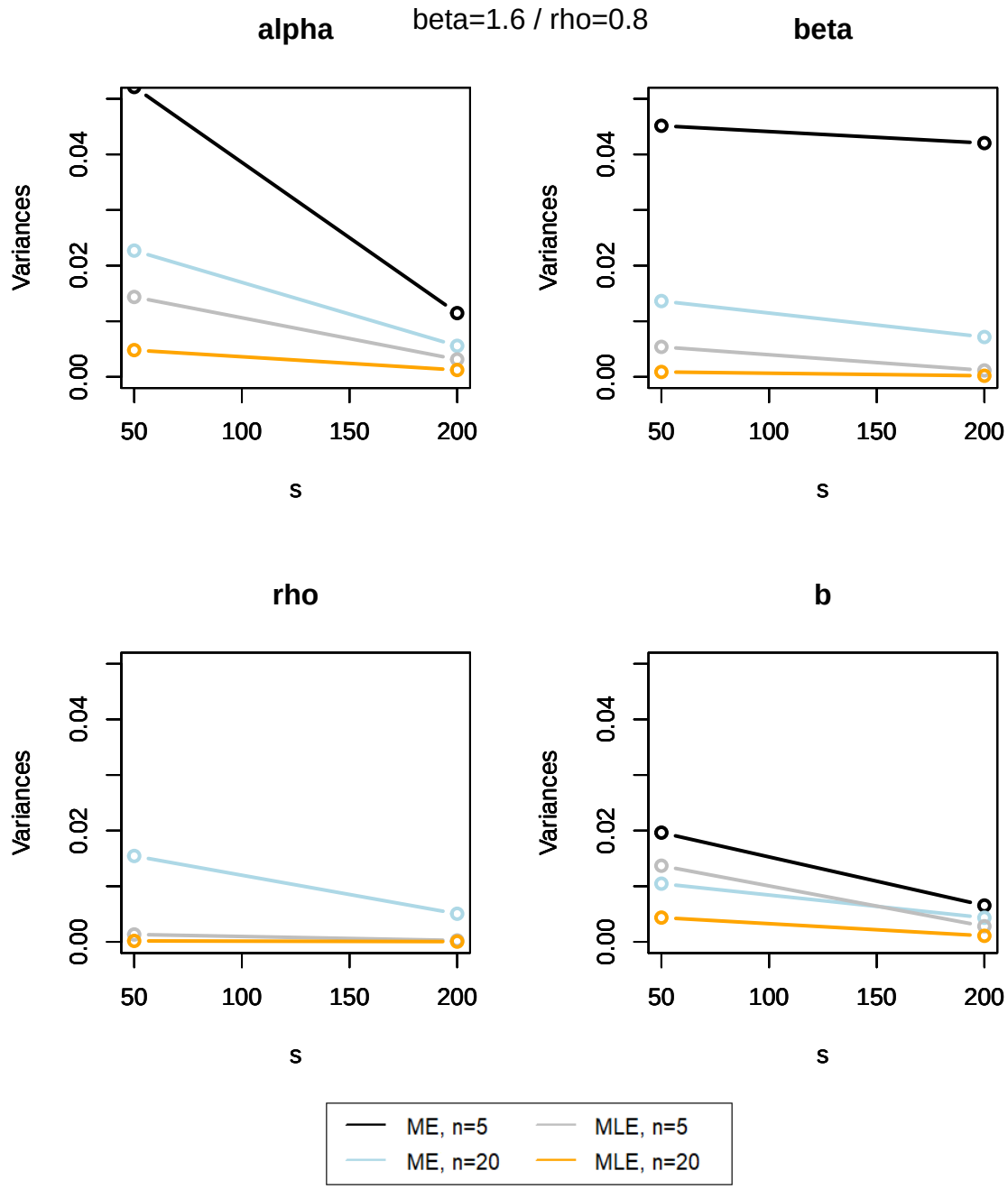

Figure A.12: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.8)$.

Appendix B

ARD_∞ model, results of the simulation study

As already stated in **Section 3**, the results are exposed as follows.

For each possible combination of (β, ρ), the RB and the variances of the estimations are summarized in the following. The figures deal either with the RB or with the variance, and are composed of four graphs: one for each parameter of the model. These graphs show four curves each, representing either the relative bias or the variance evolution with s for both the ME and the MLE methods when $n = 5$ and $n = 20$.
To be able to compare the results from one parameter to another, as well as the parameters sets, the y-axis range is the same for all plots. However, both the RB and the variance can be higher than the maximum value displayed in the y-axis. This signifies that if a curve does not appear on a graph, as the black one in the top right-hand graph as well as the bottom left-hand one of **Figure A.1**, it means that all the points of the curve are out of range.

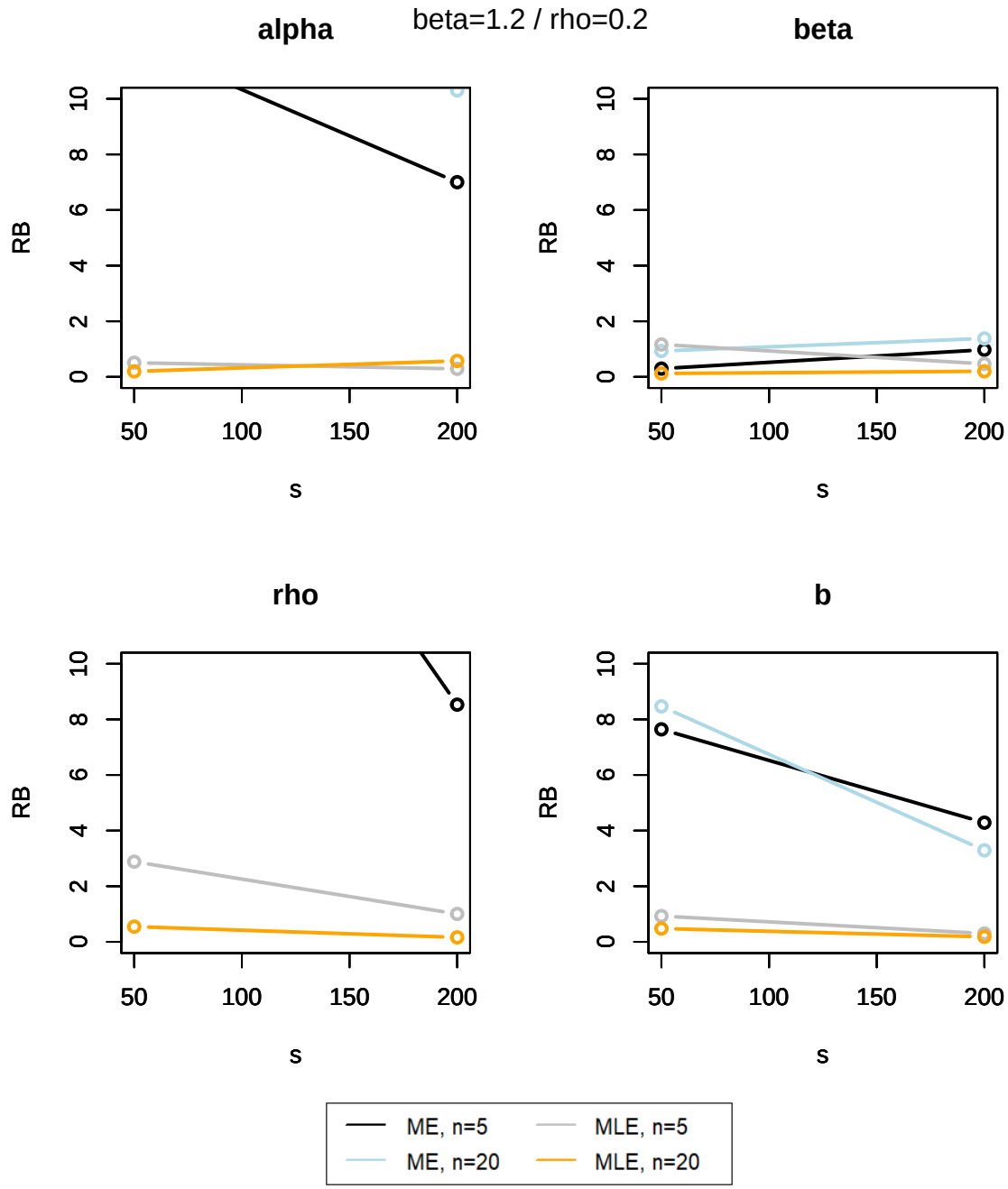

Figure B.1: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.2)$.

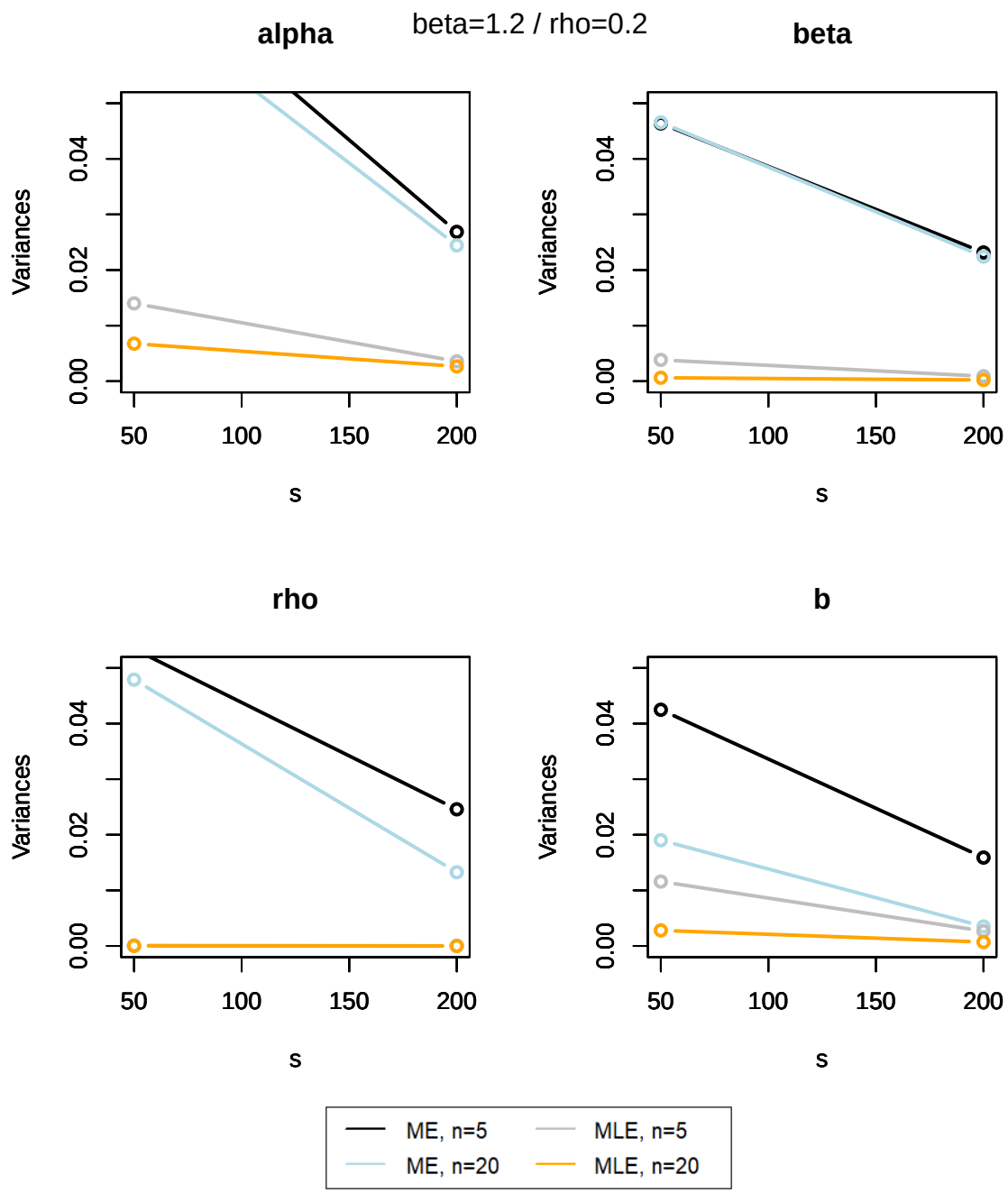

Figure B.2: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.2)$.

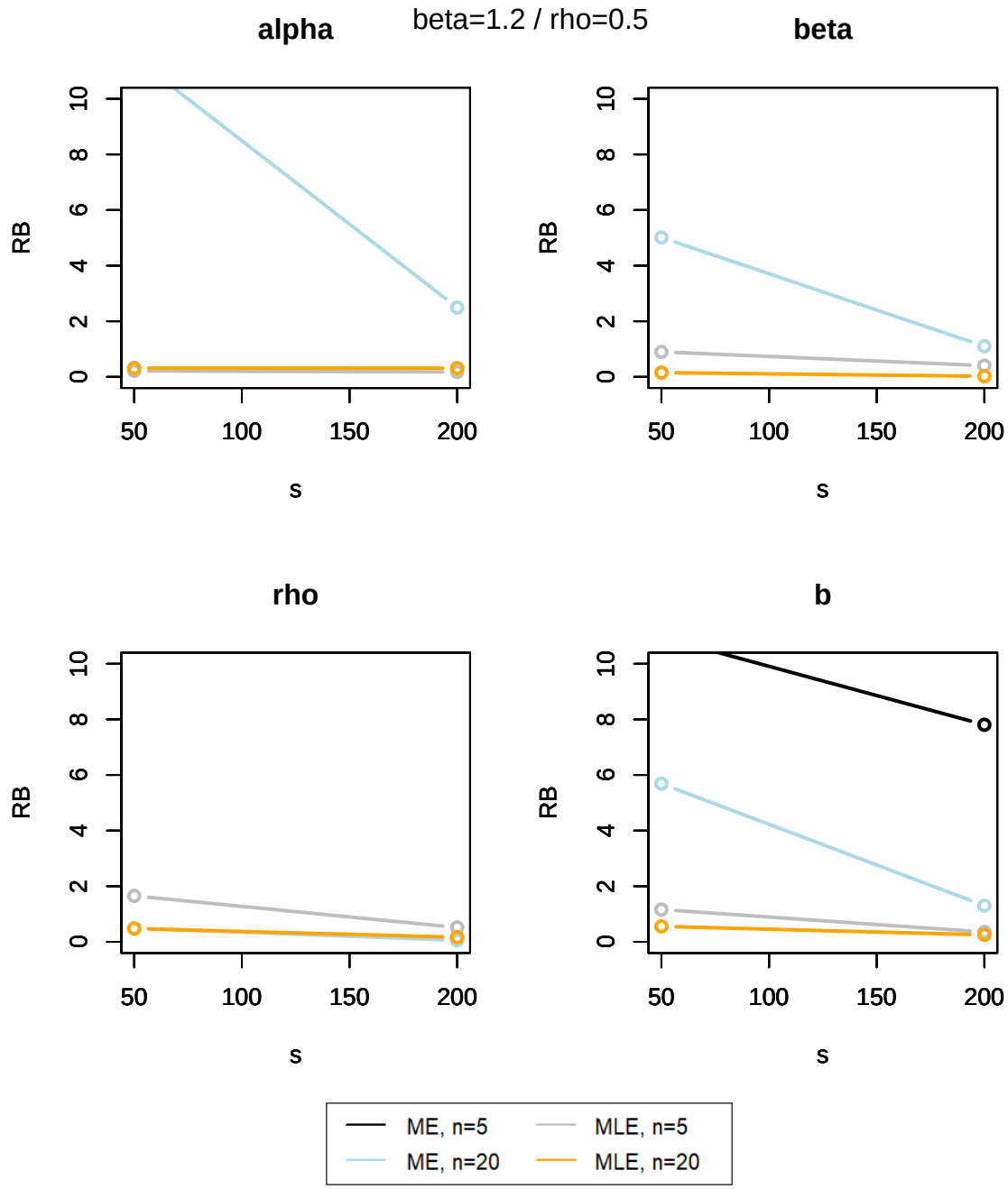

Figure B.3: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.5)$.

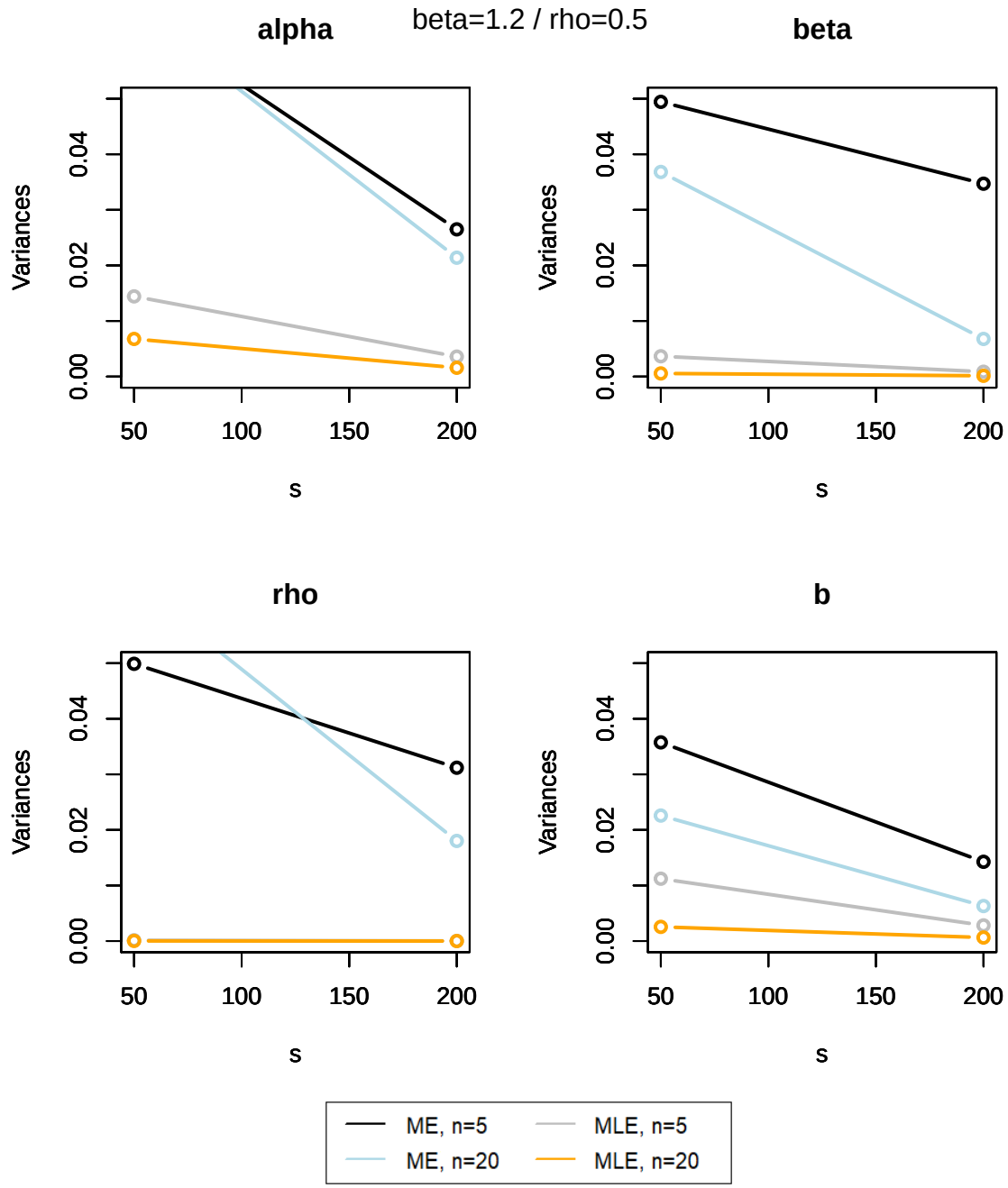

Figure B.4: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.5)$.

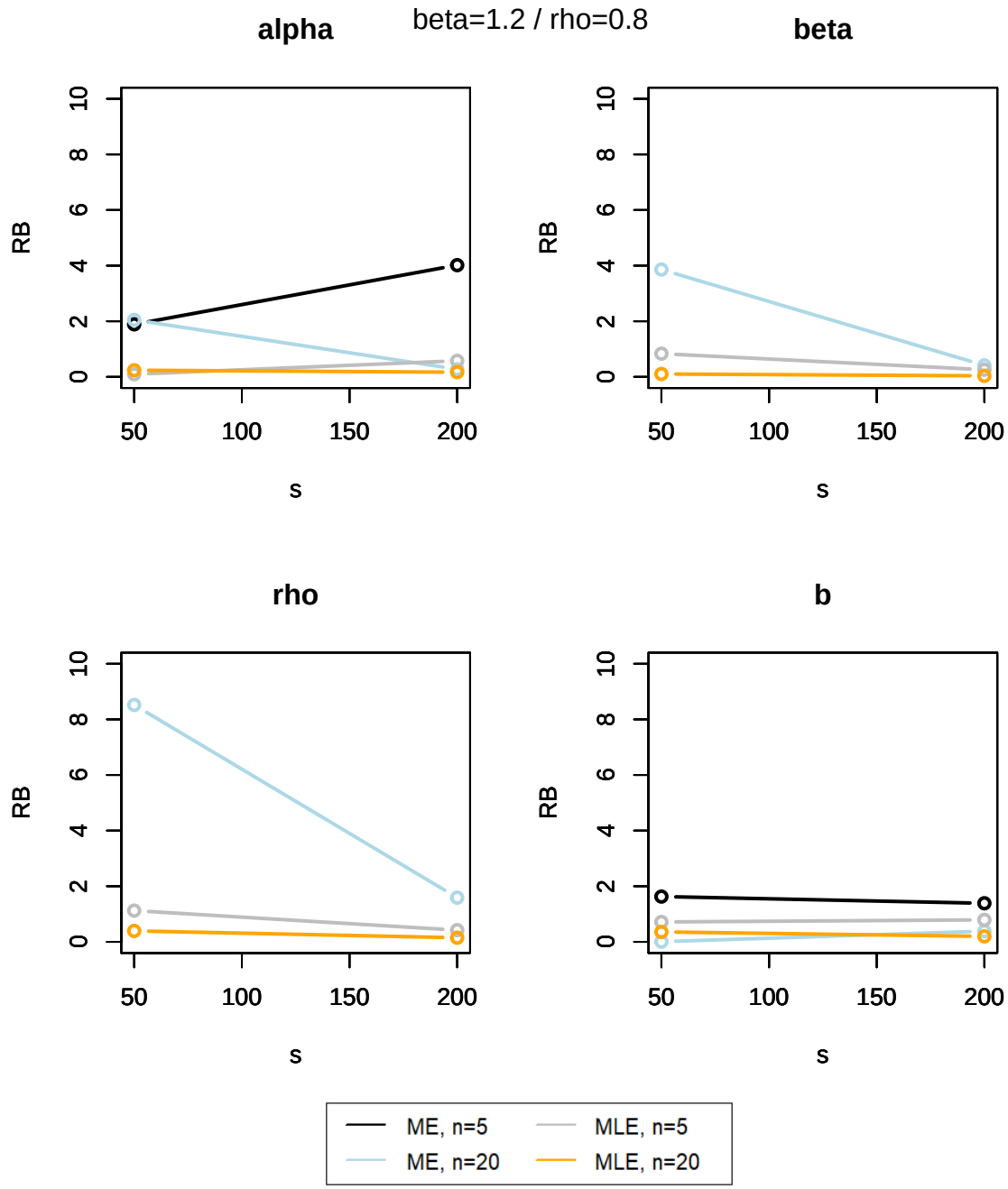

Figure B.5: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.8)$.

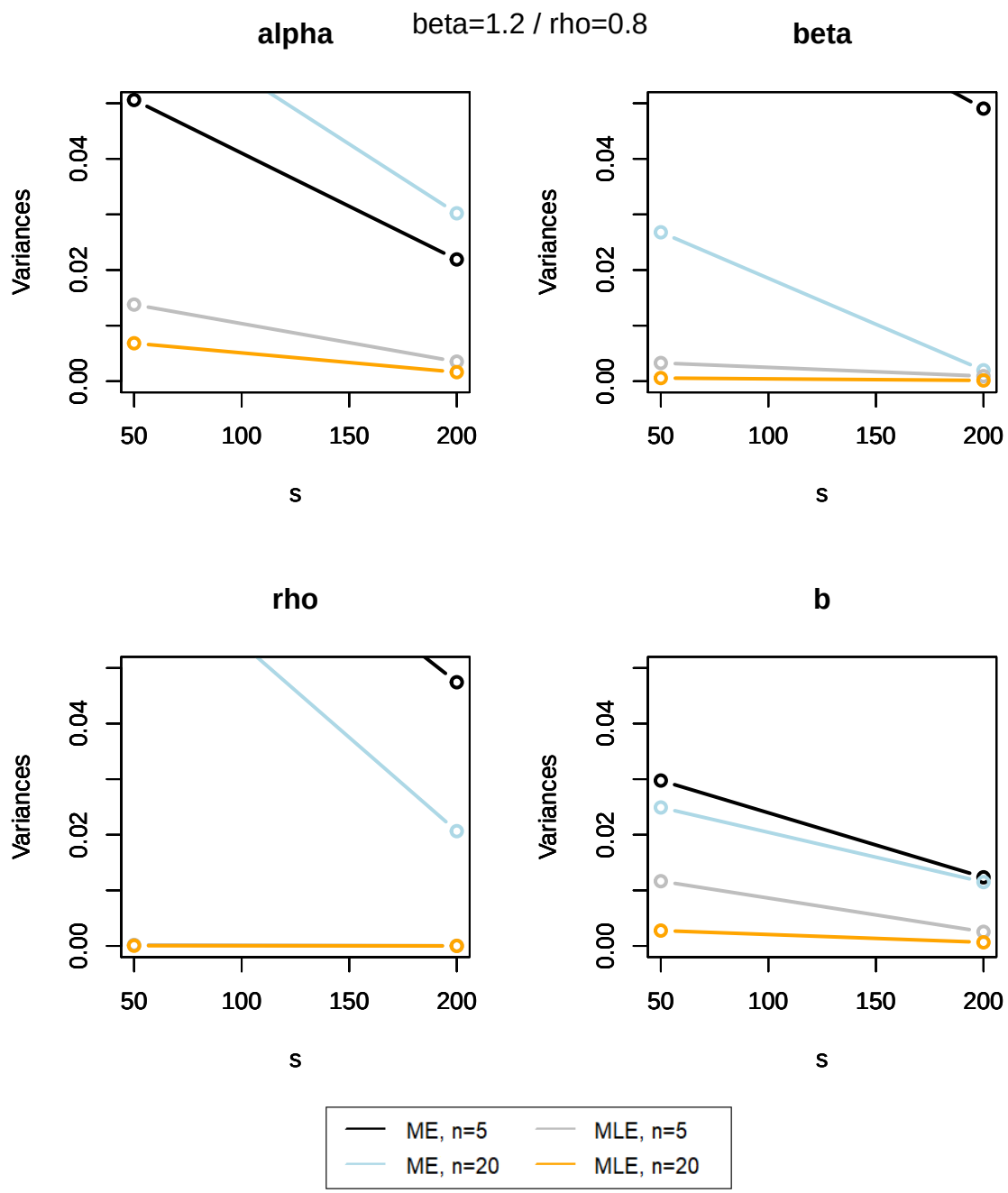

Figure B.6: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.2, 0.8)$.

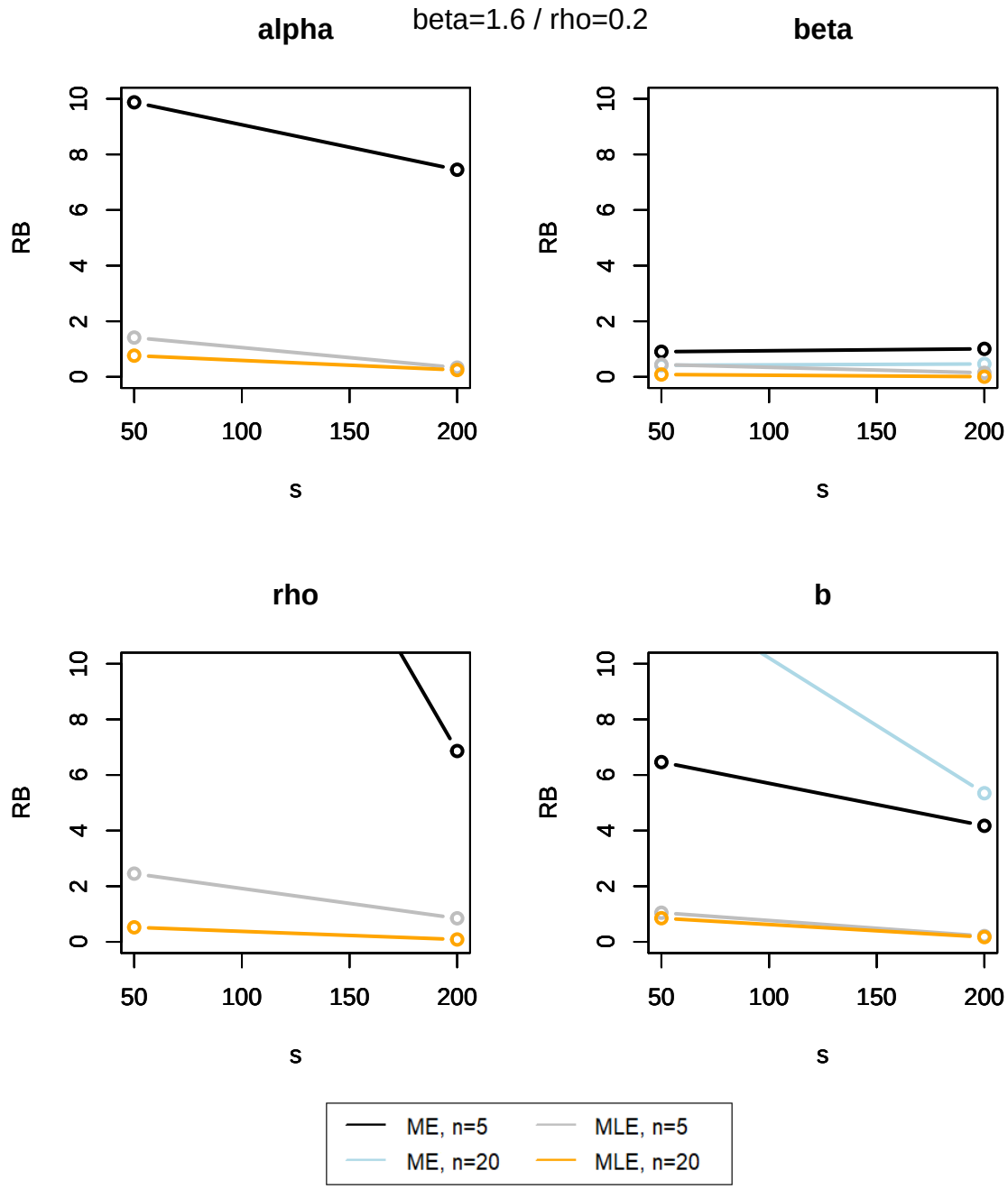

Figure B.7: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.2)$.

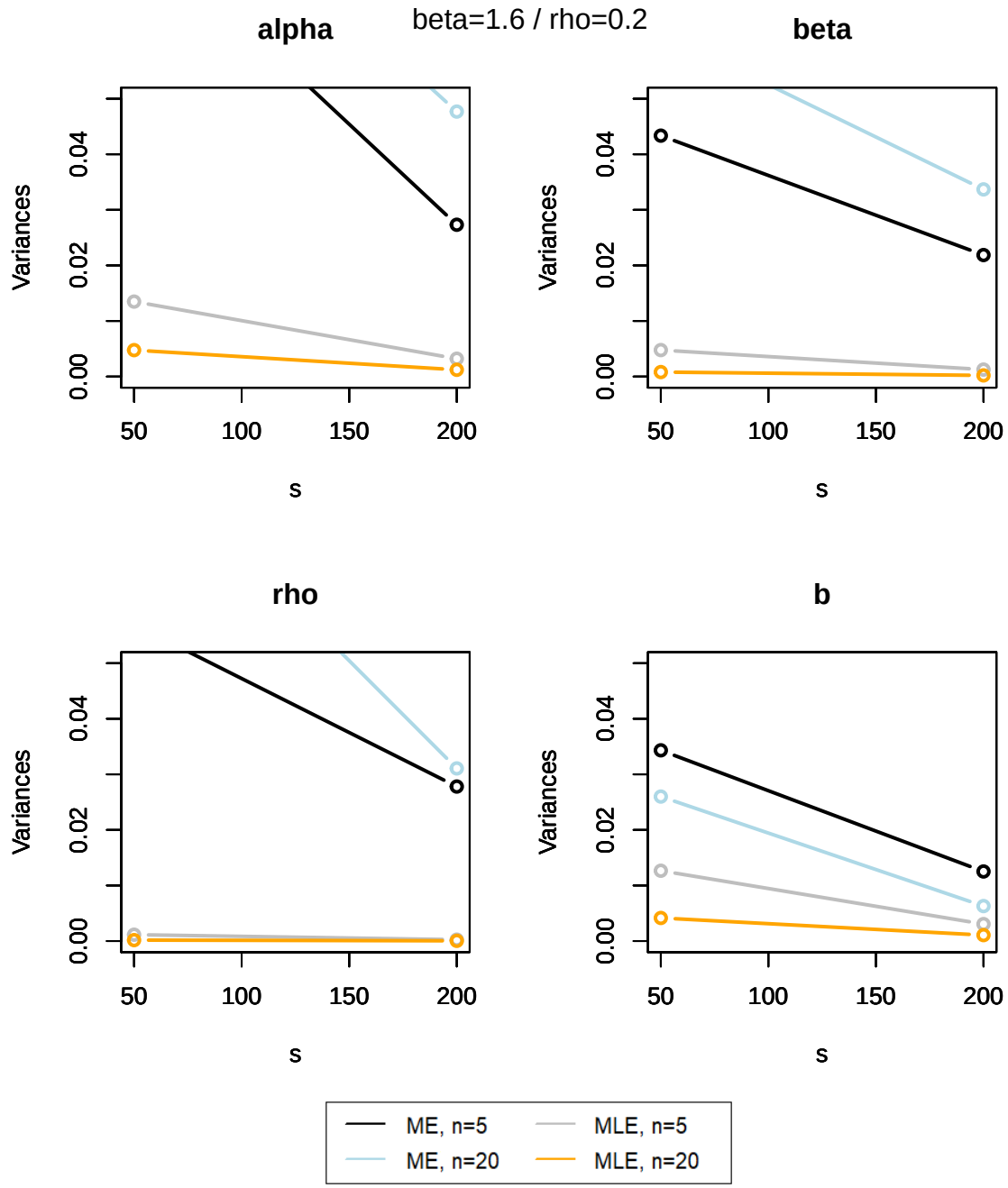

Figure B.8: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.2)$.

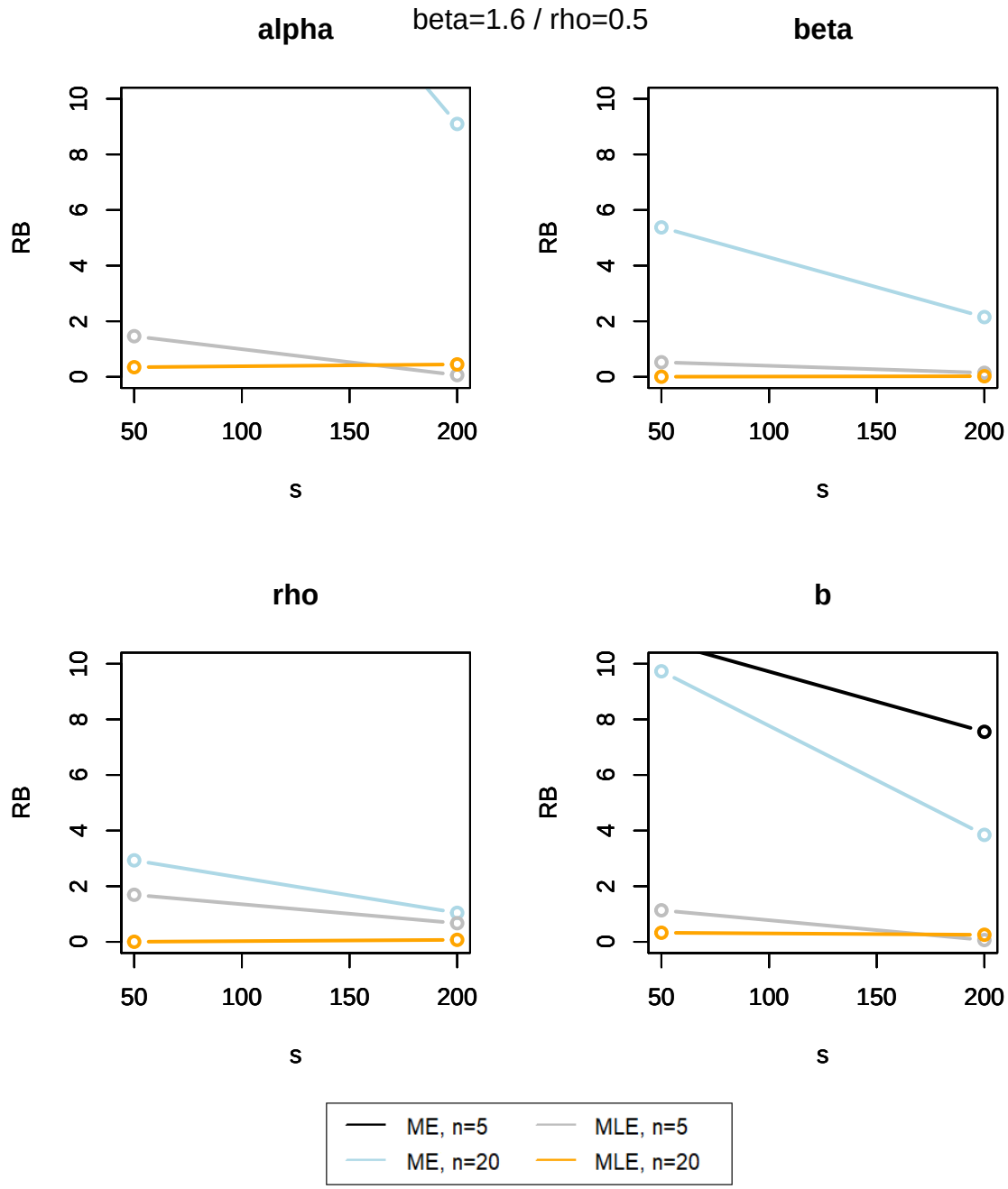

Figure B.9: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.5)$.

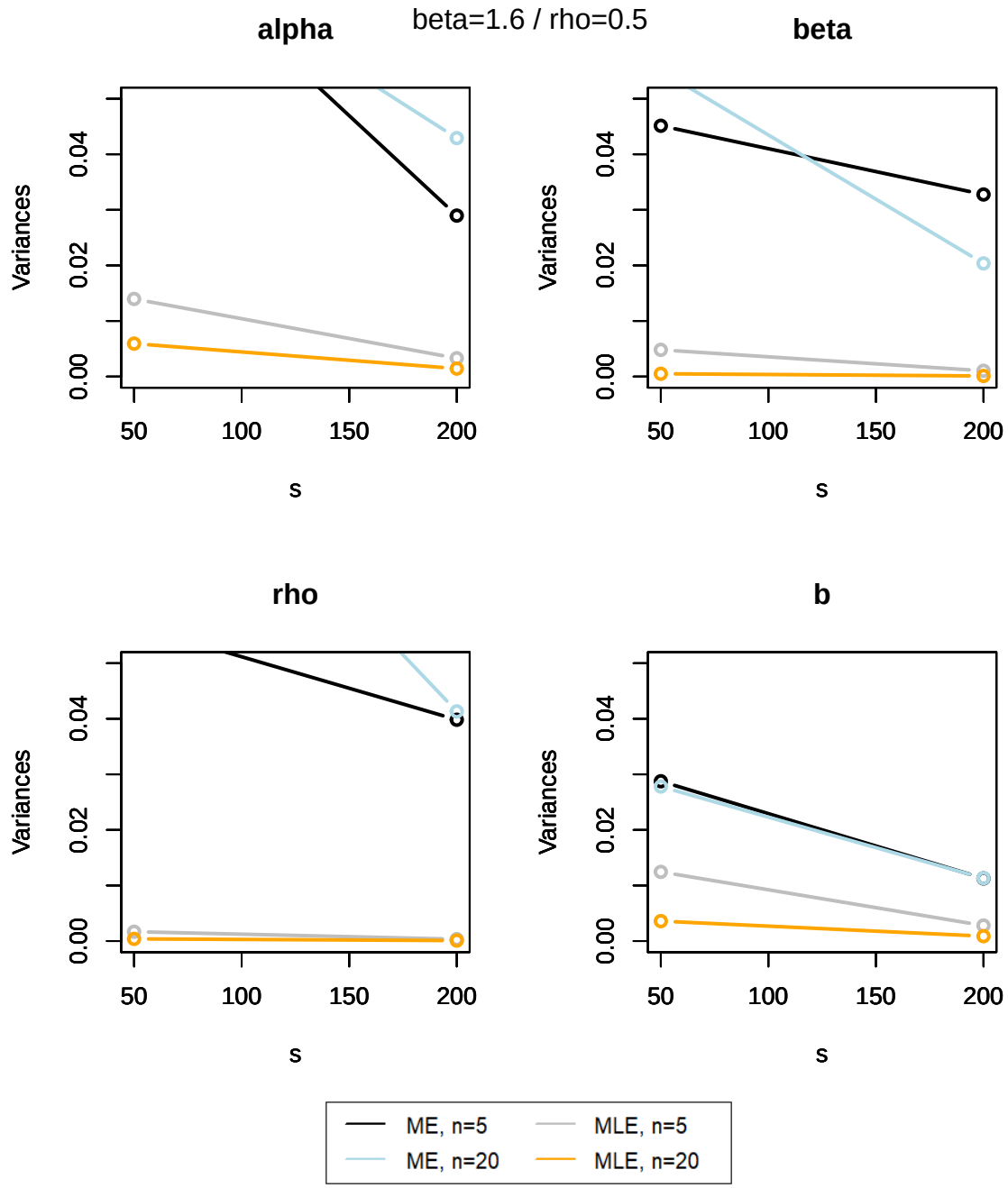

Figure B.10: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.5)$.

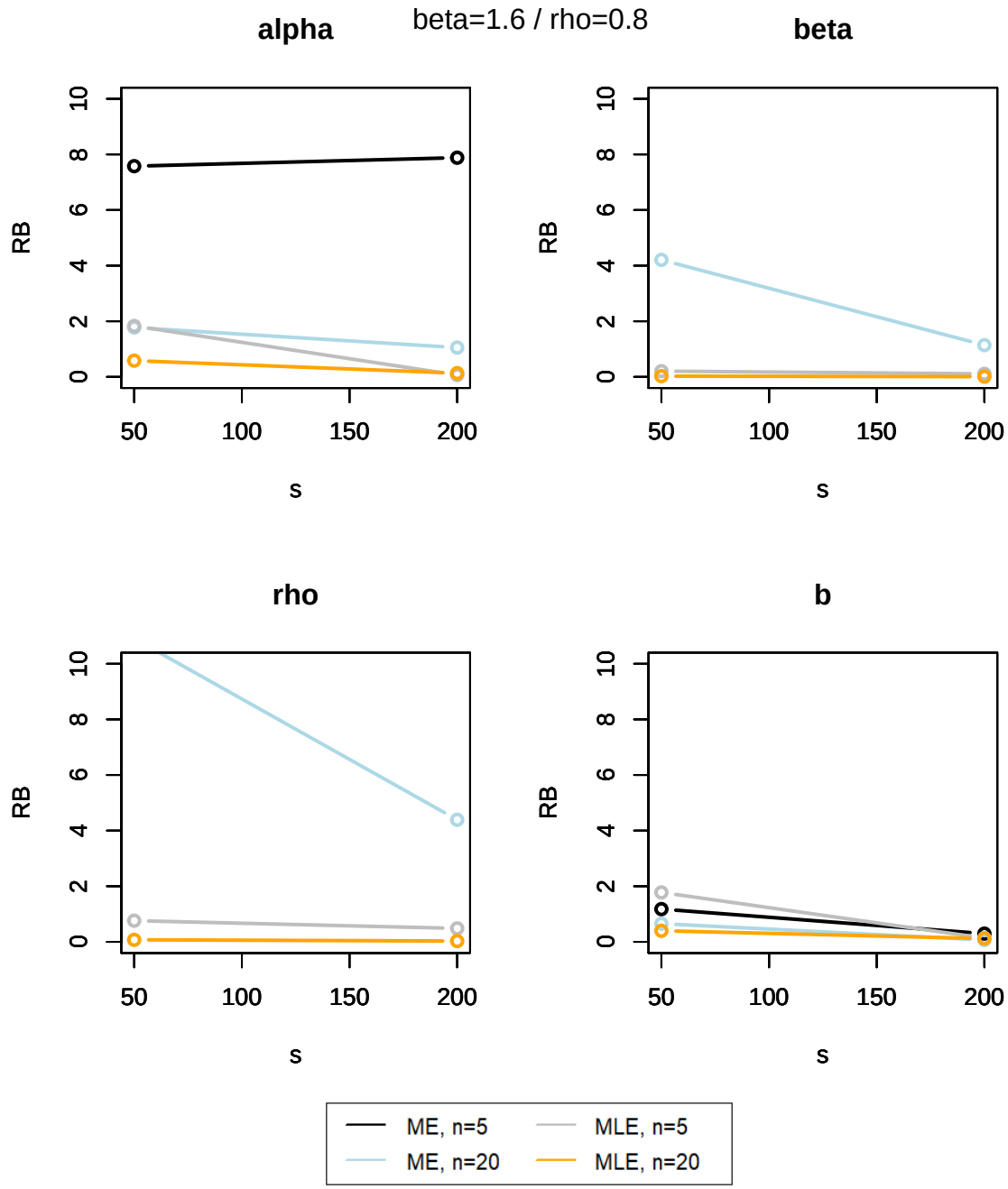

Figure B.11: Plots of the RB evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.8)$.

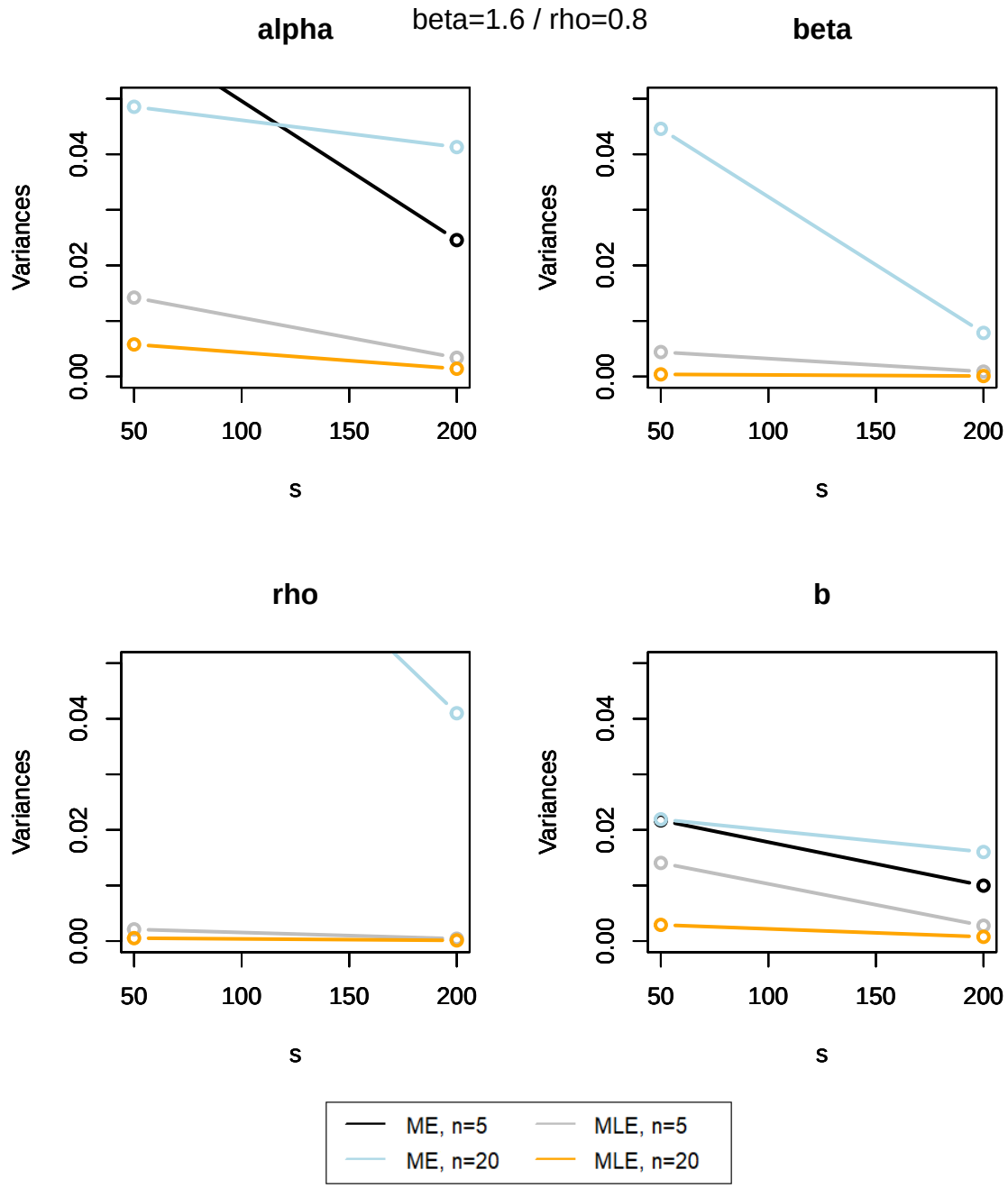

Figure B.12: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.6, 0.8)$.

Appendix C

ARA_1 model, results of the simulation study

As already stated in **Section 7.2**, the results are exposed as follows.

For each (β, ρ) in $\{0.5, 1, 1.5\} \times \{0.2, 0.5, 0.8\}$, the Relative Bias (RB) and the variances of the estimations are summarized in the following. The figures deal either with the RB or with the variance, and are composed of four graphs: one for each parameter of the model. These graphs show four curves each, representing either the relative bias or the variance evolution with s for both the HDOI and the CLO methods when $n = 4$ and $n = 16$.
For the sake of clarity, as the RB can be higher than 200% as well as close to zero, instead of representing RB we chose to represent the quantity log(1+RB). Such a scale for the RB is unusual, hence, red dotted lines point out thresholds for which the RB is equal to 5%, 10%, 25% and 50%.
Finally, to be able to compare the results from one parameter to another, as well as the parameters sets, the y-axis range are the same for all plots. However, both the RB and the variance can be higher than the maximum value displayed in the y-axis. This signifies that if a curve does not appear on a graph, as the black one in the top right-hand graph of **Figure C.5**, each point of this curve is higher than this maximum.

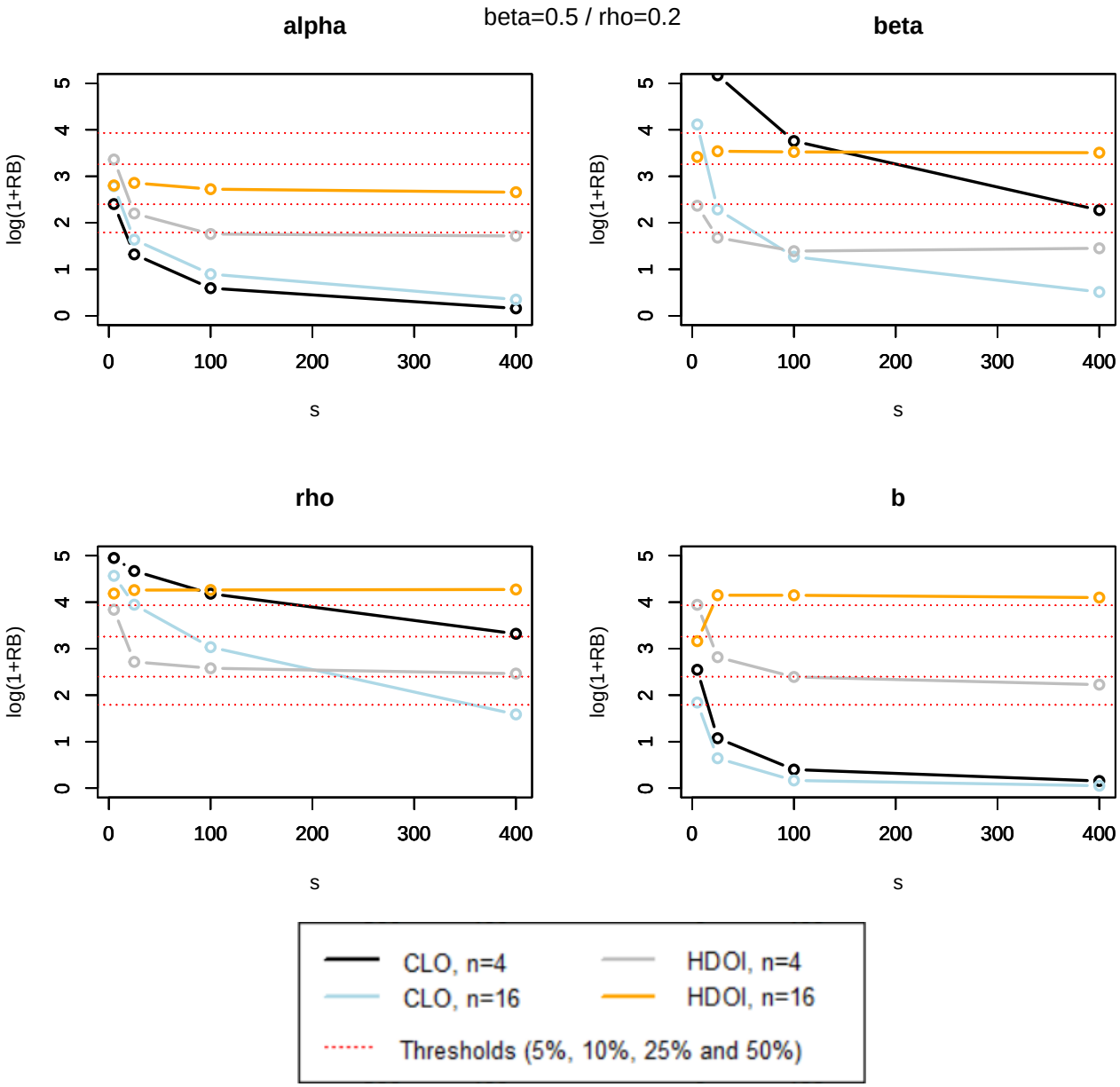

Figure C.1: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.2)$.

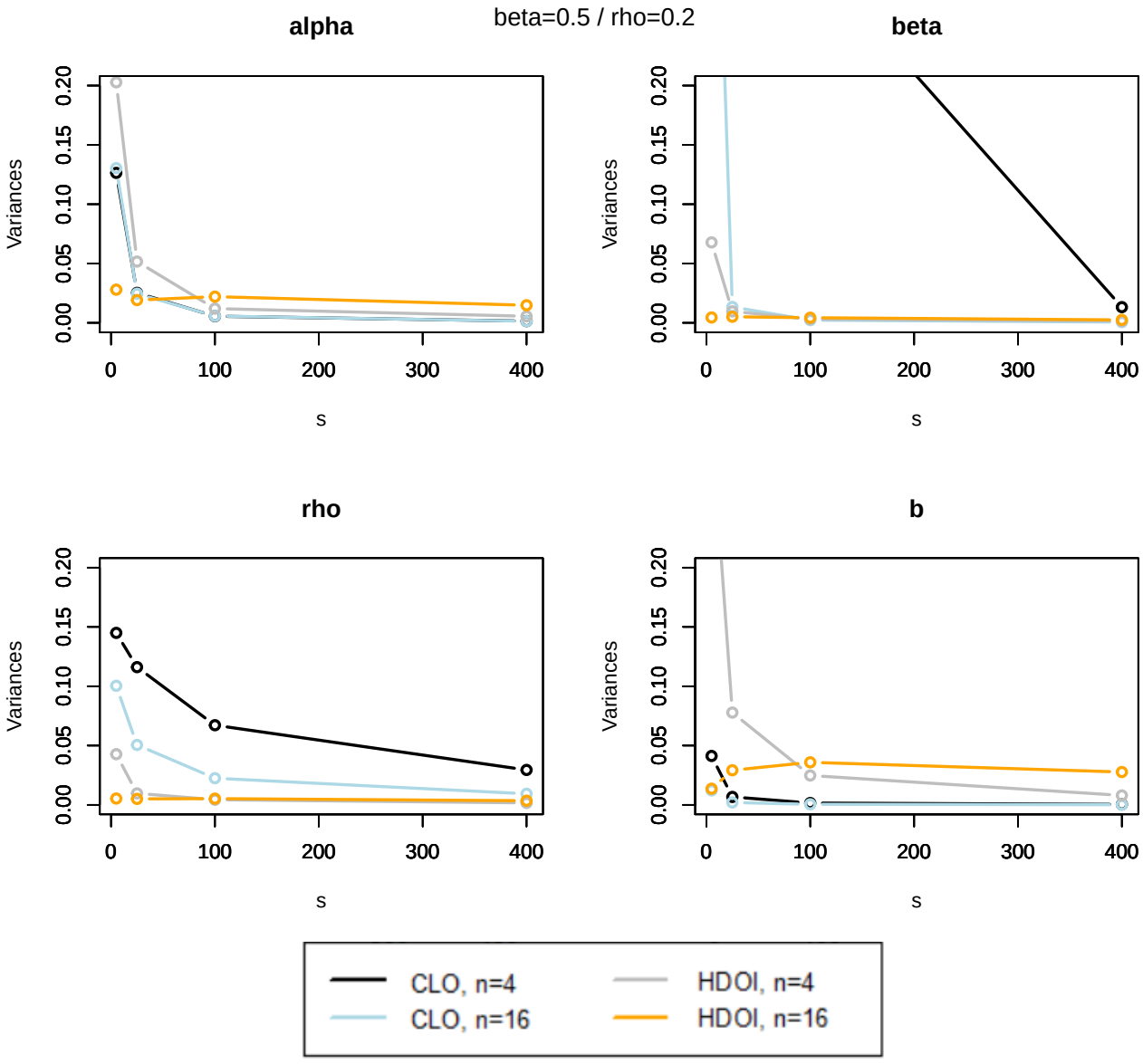

Figure C.2: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.2)$.

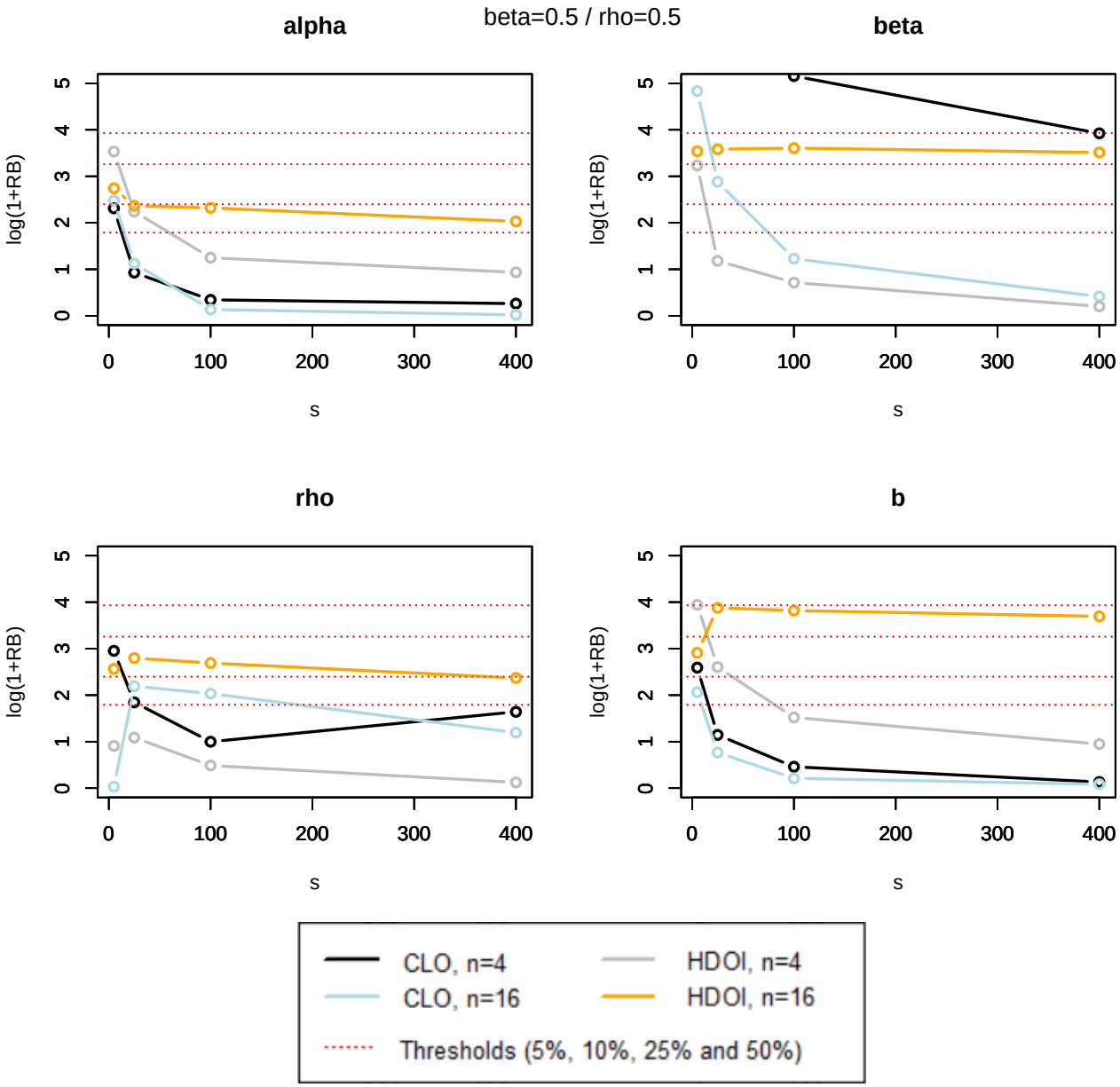

Figure C.3: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.5)$.

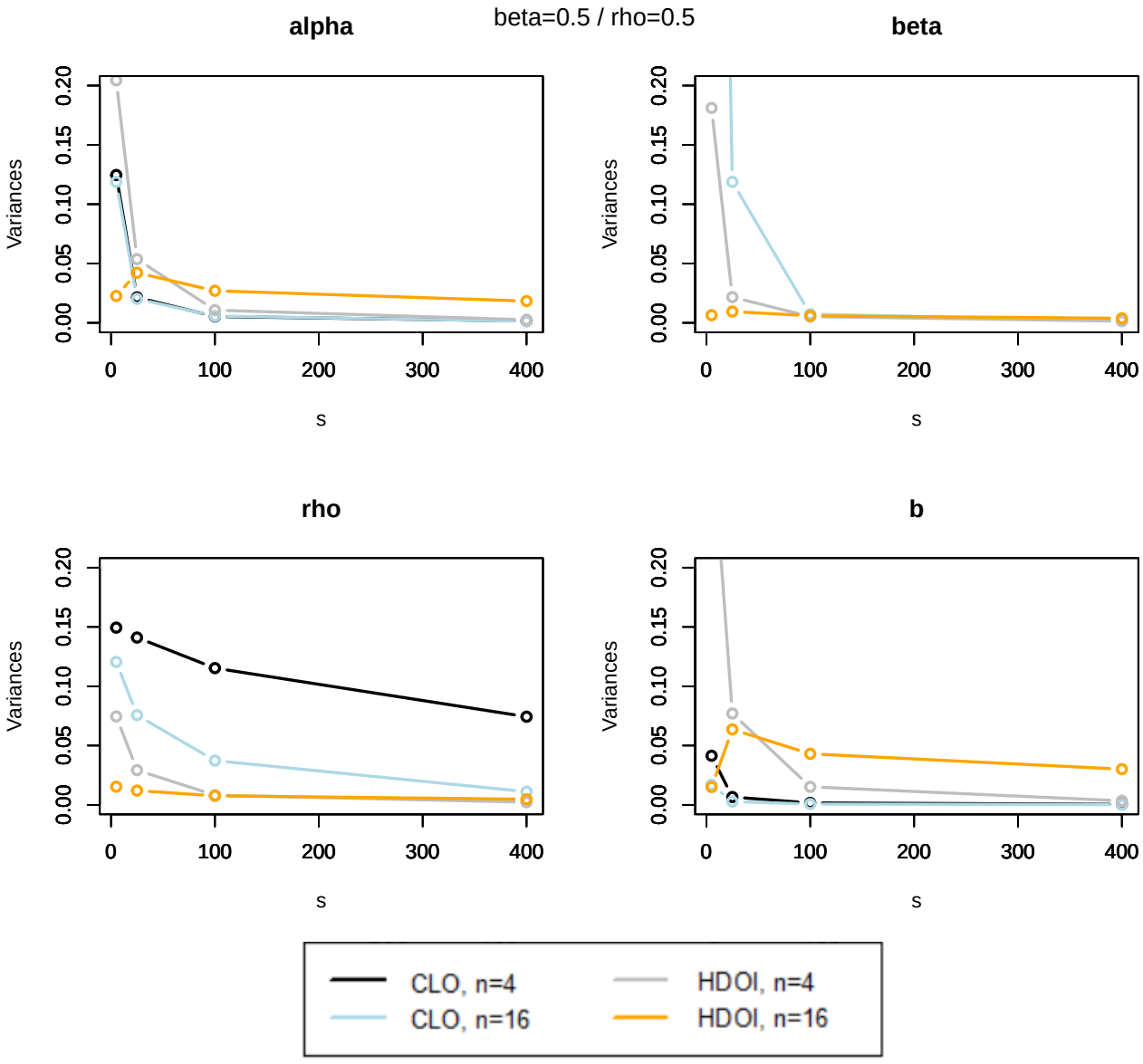

Figure C.4: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.5)$.

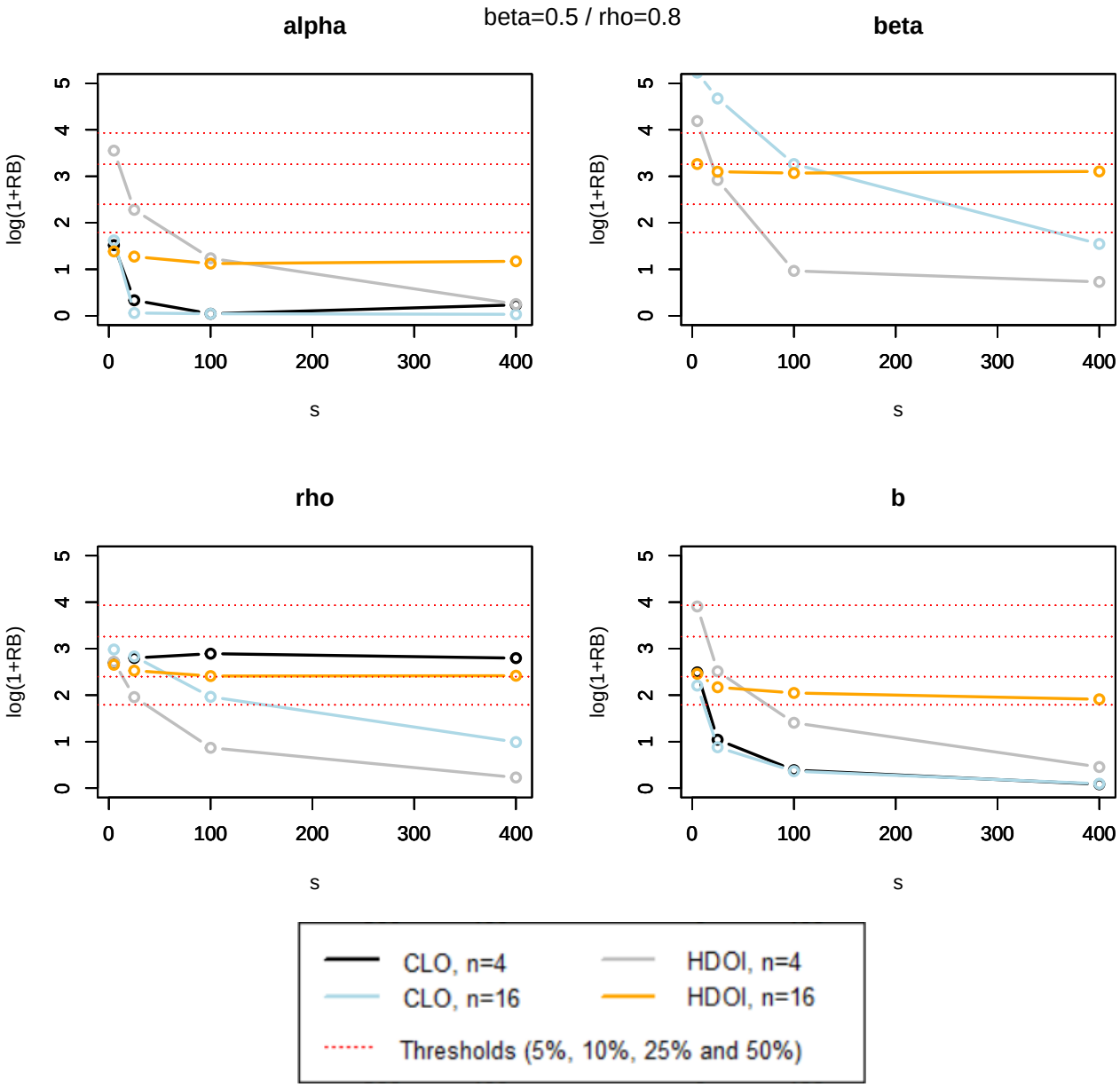

Figure C.5: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.5)$.

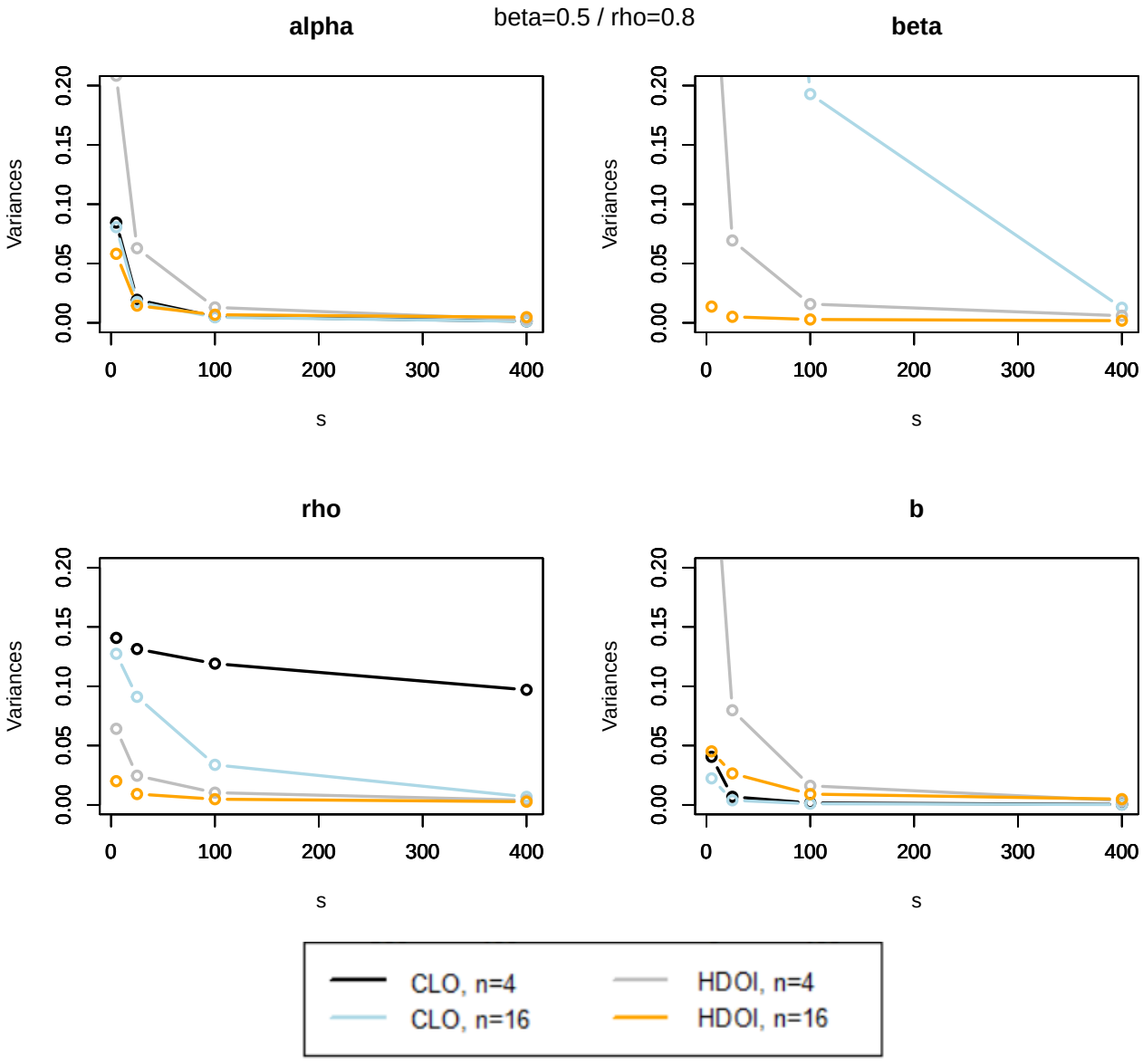

Figure C.6: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (0.5, 0.8)$.

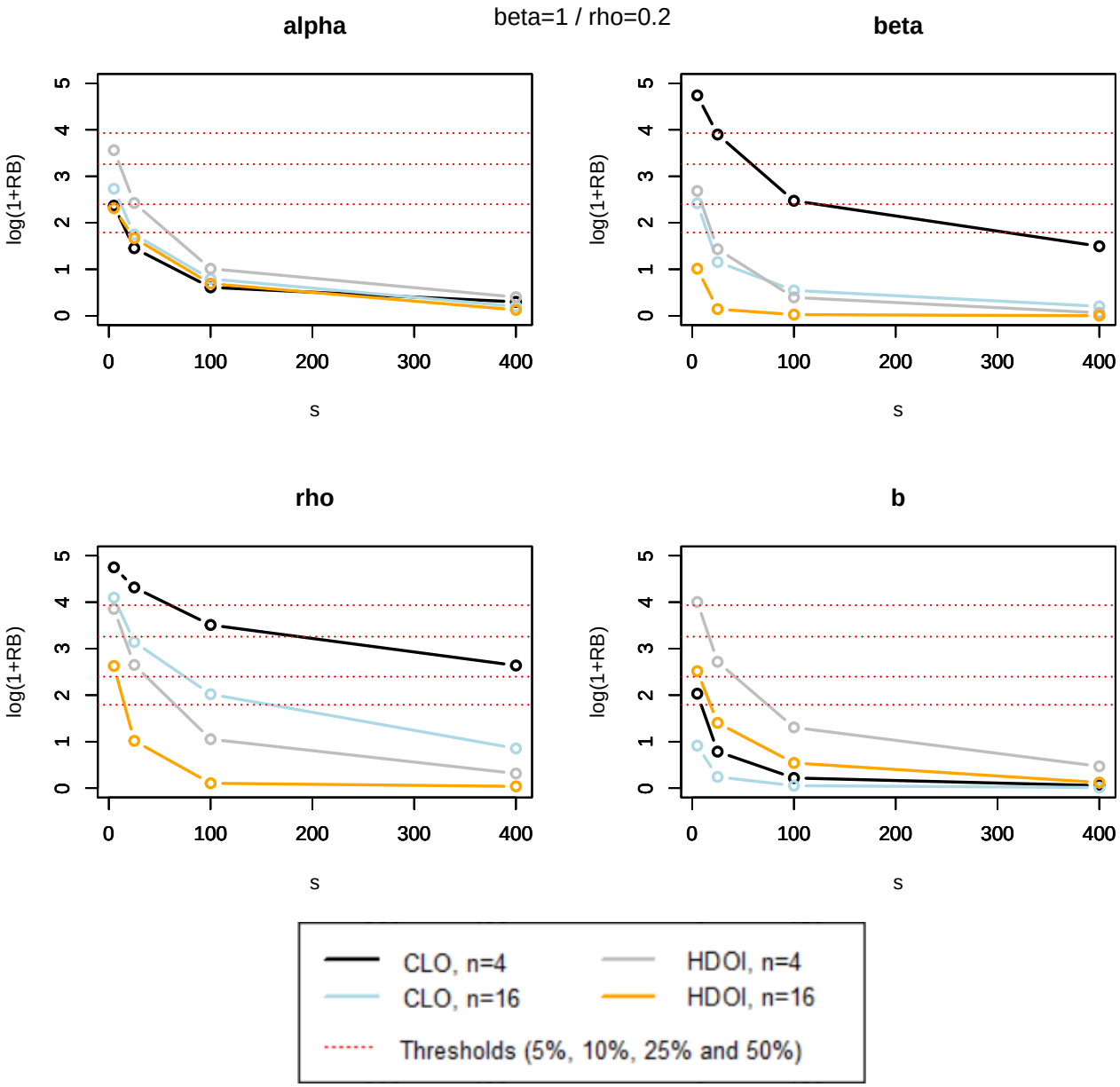

Figure C.7: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.2)$.

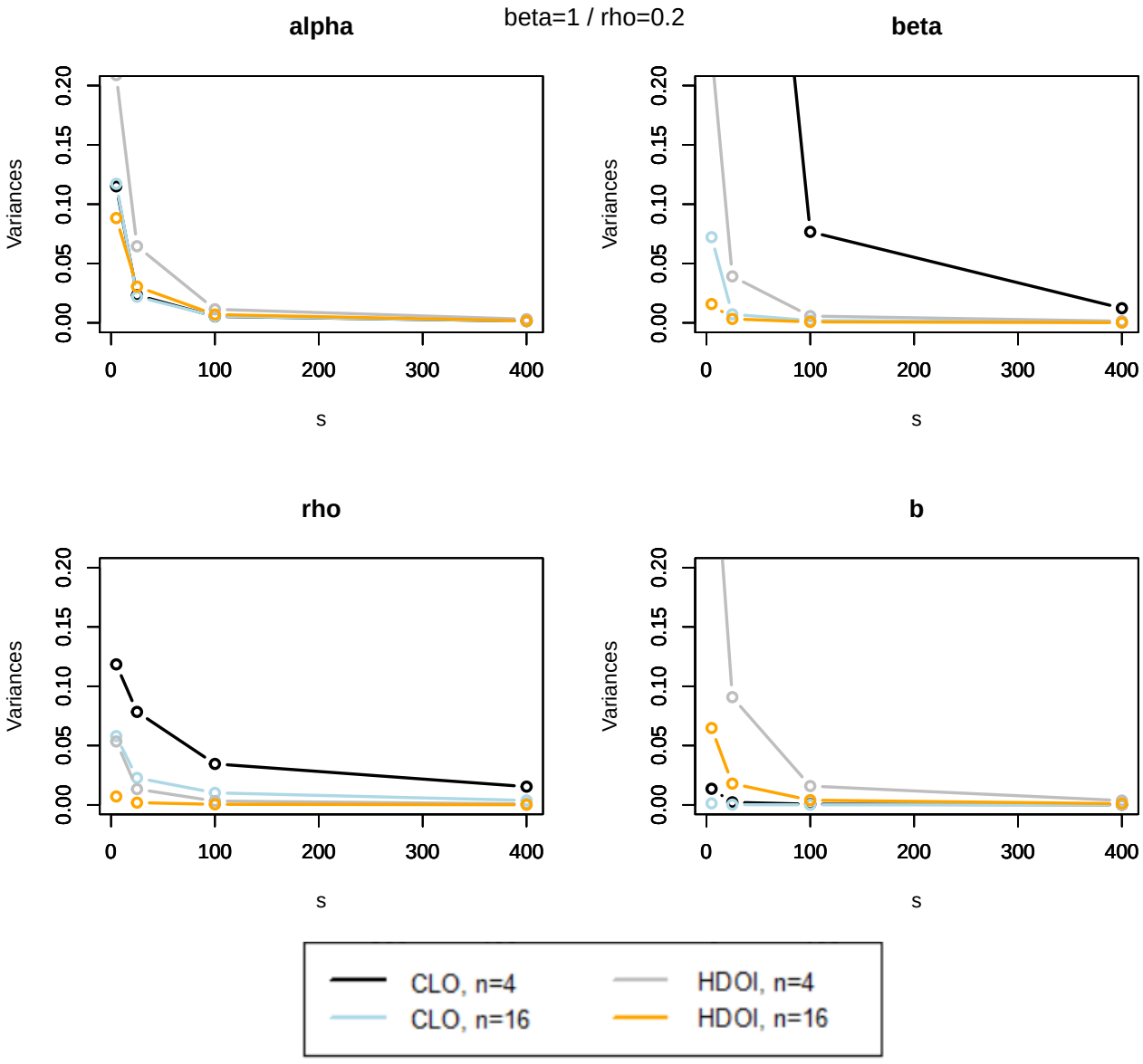

Figure C.8: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.2)$.

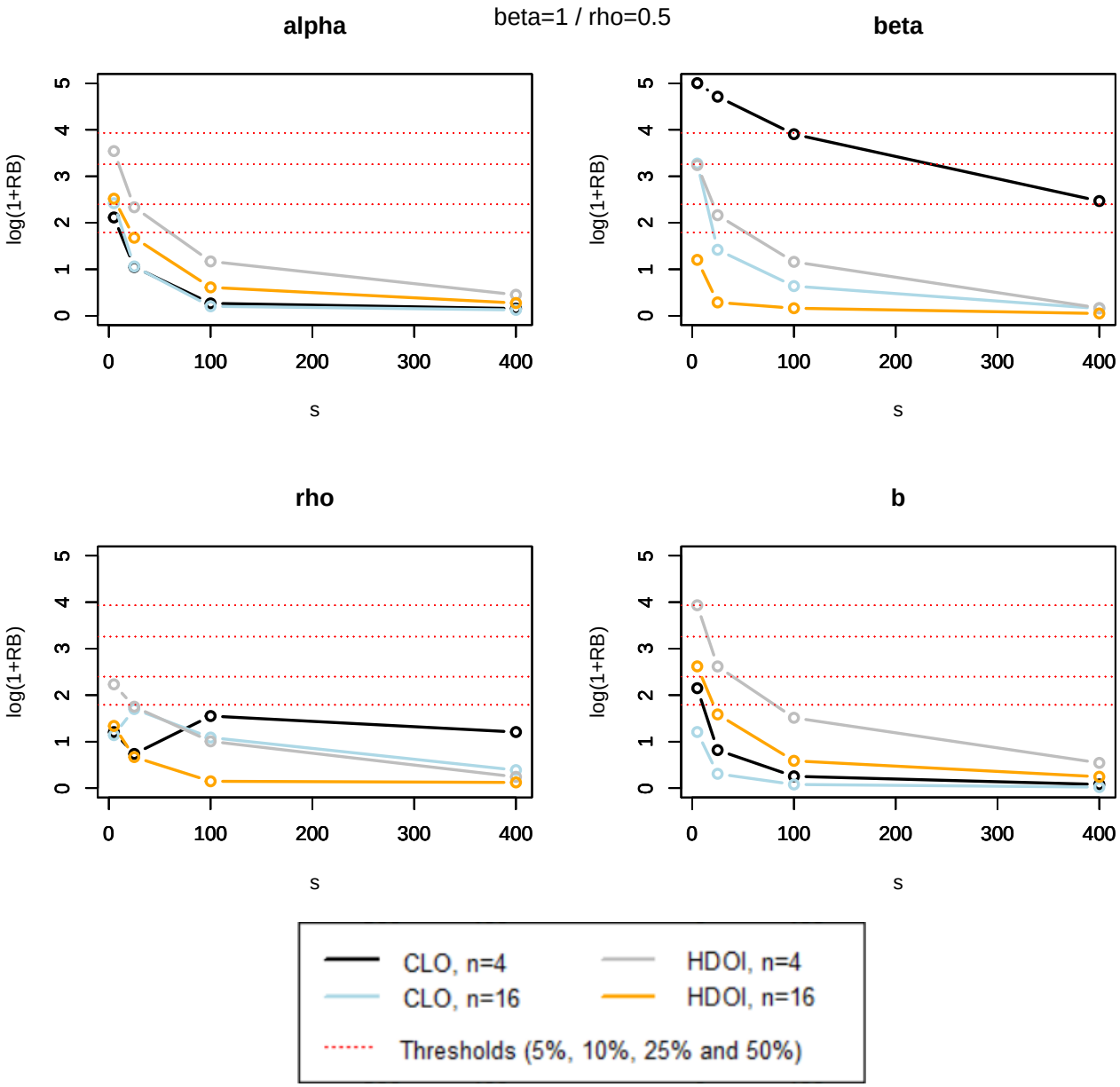

Figure C.9: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.5)$.

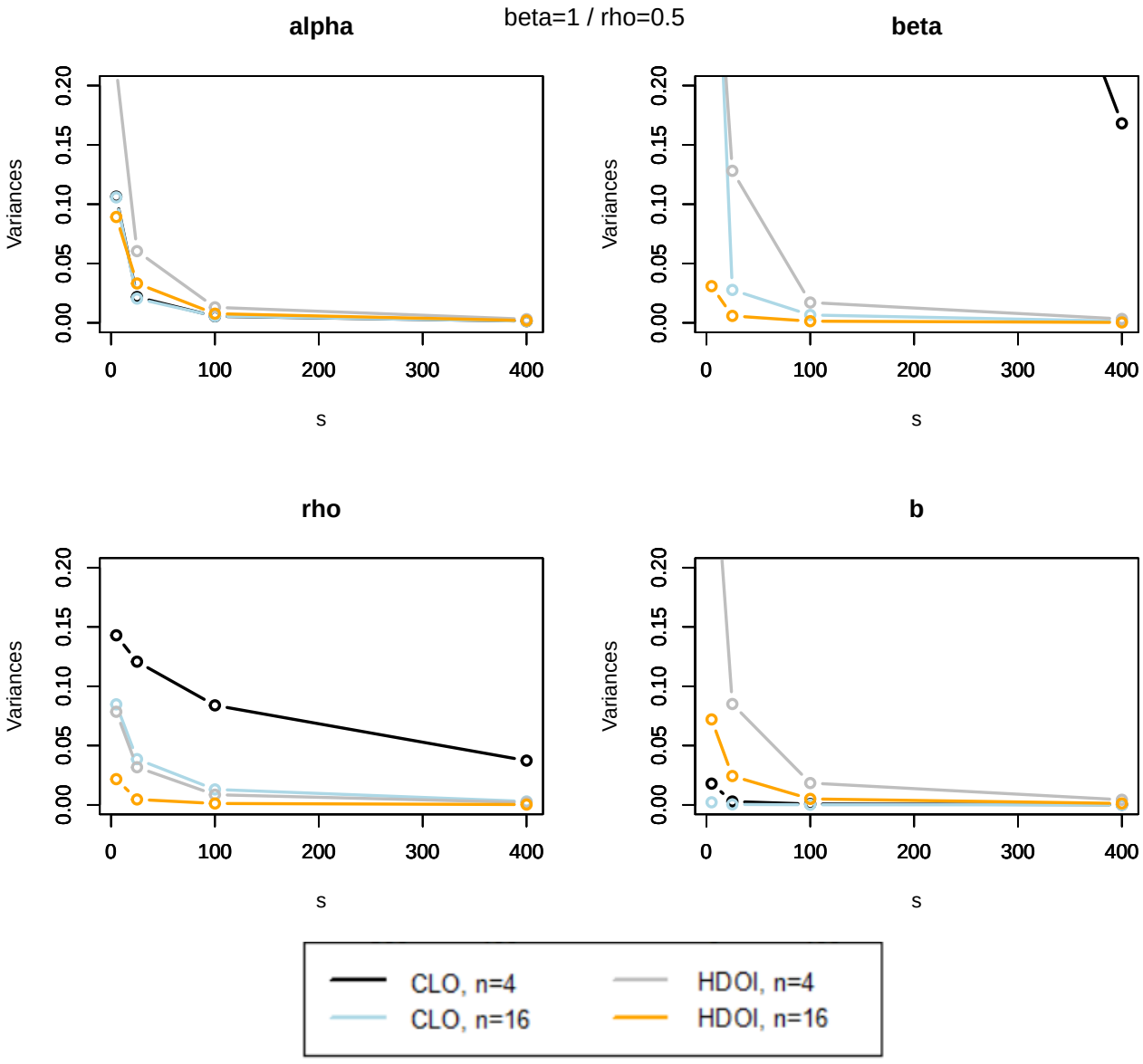

Figure C.10: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.5)$.

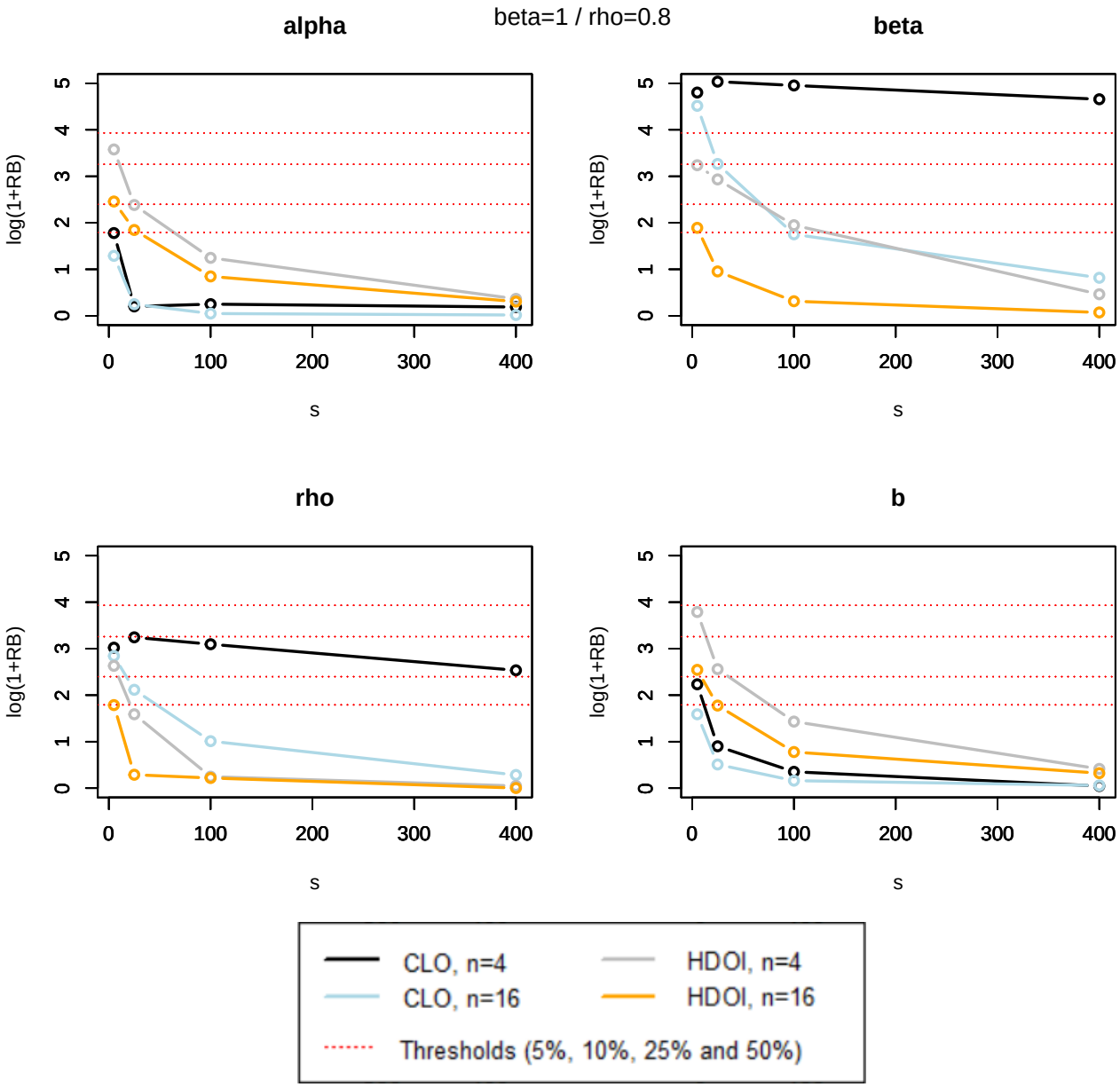

Figure C.11: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.8)$.

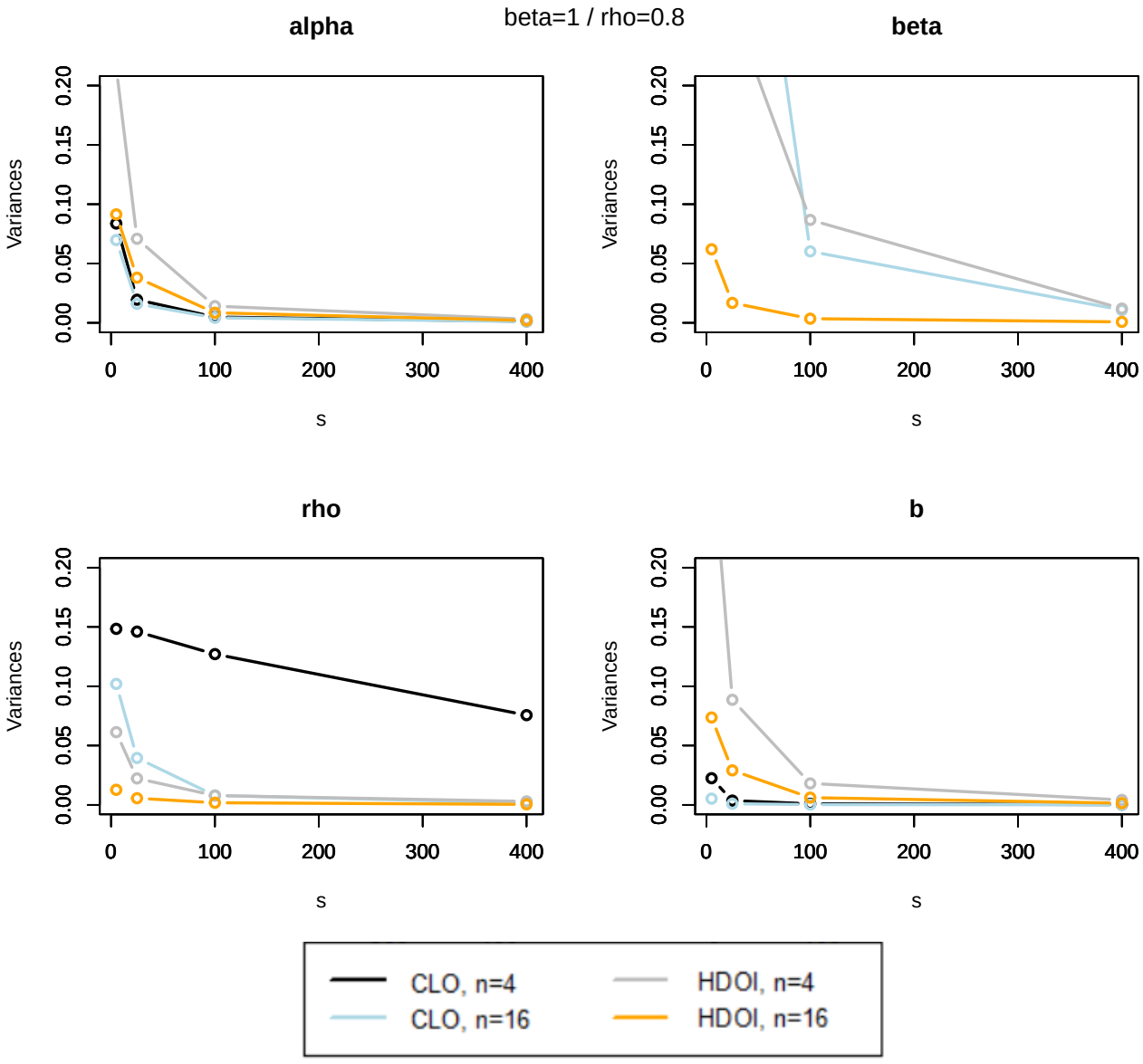

Figure C.12: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1, 0.8)$.

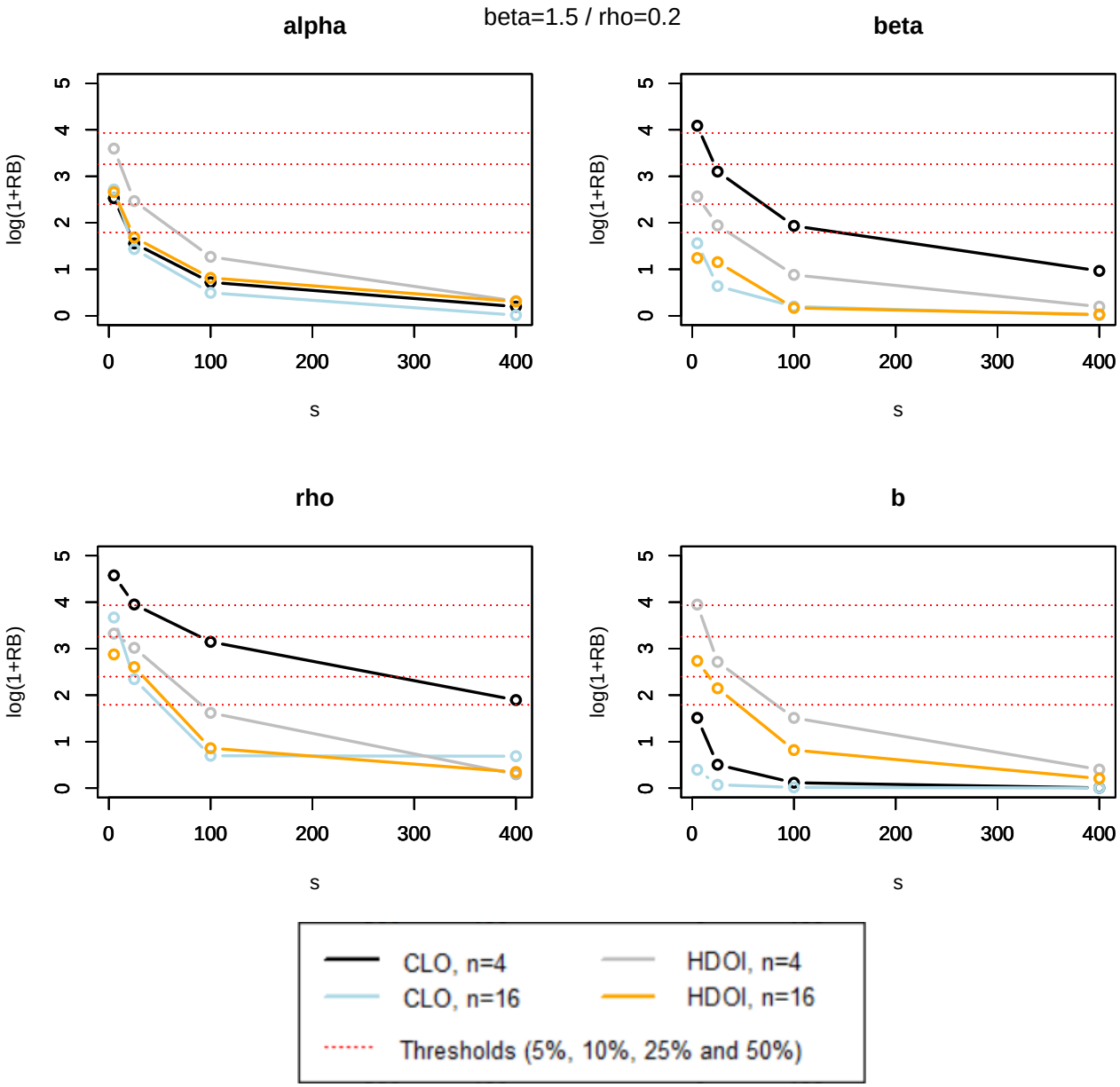

Figure C.13: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.2)$.

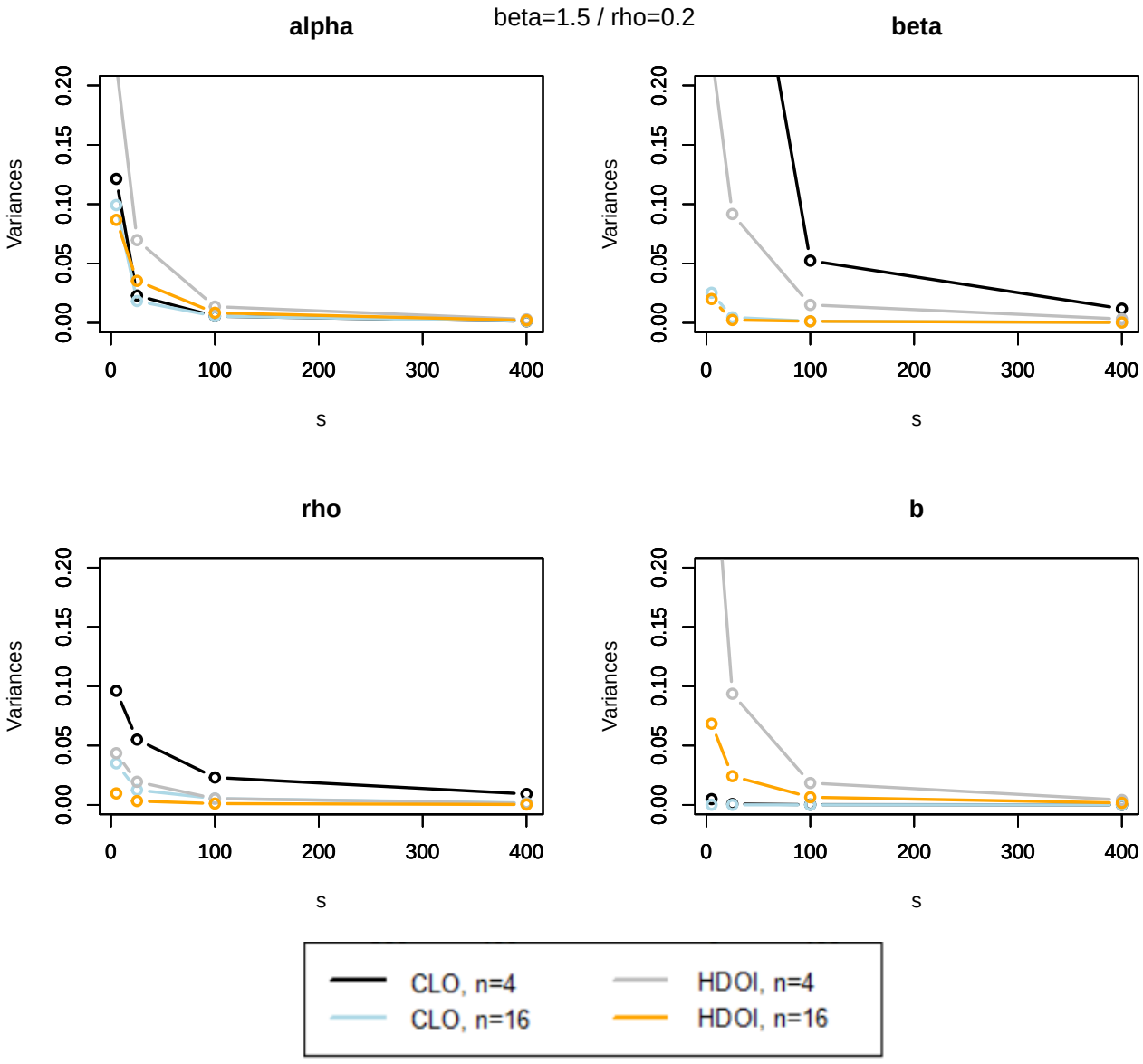

Figure C.14: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.2)$.

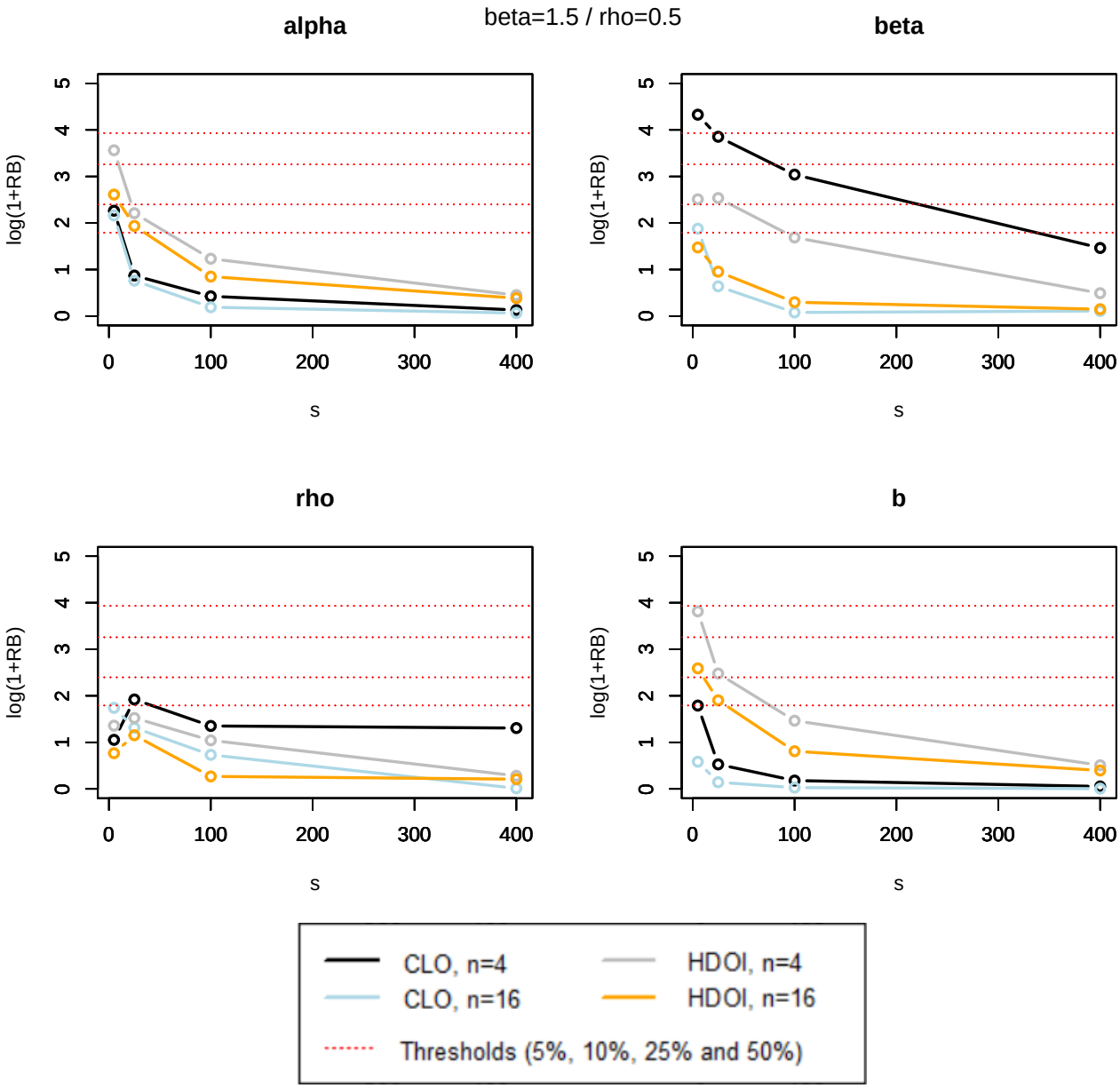

Figure C.15: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.5)$.

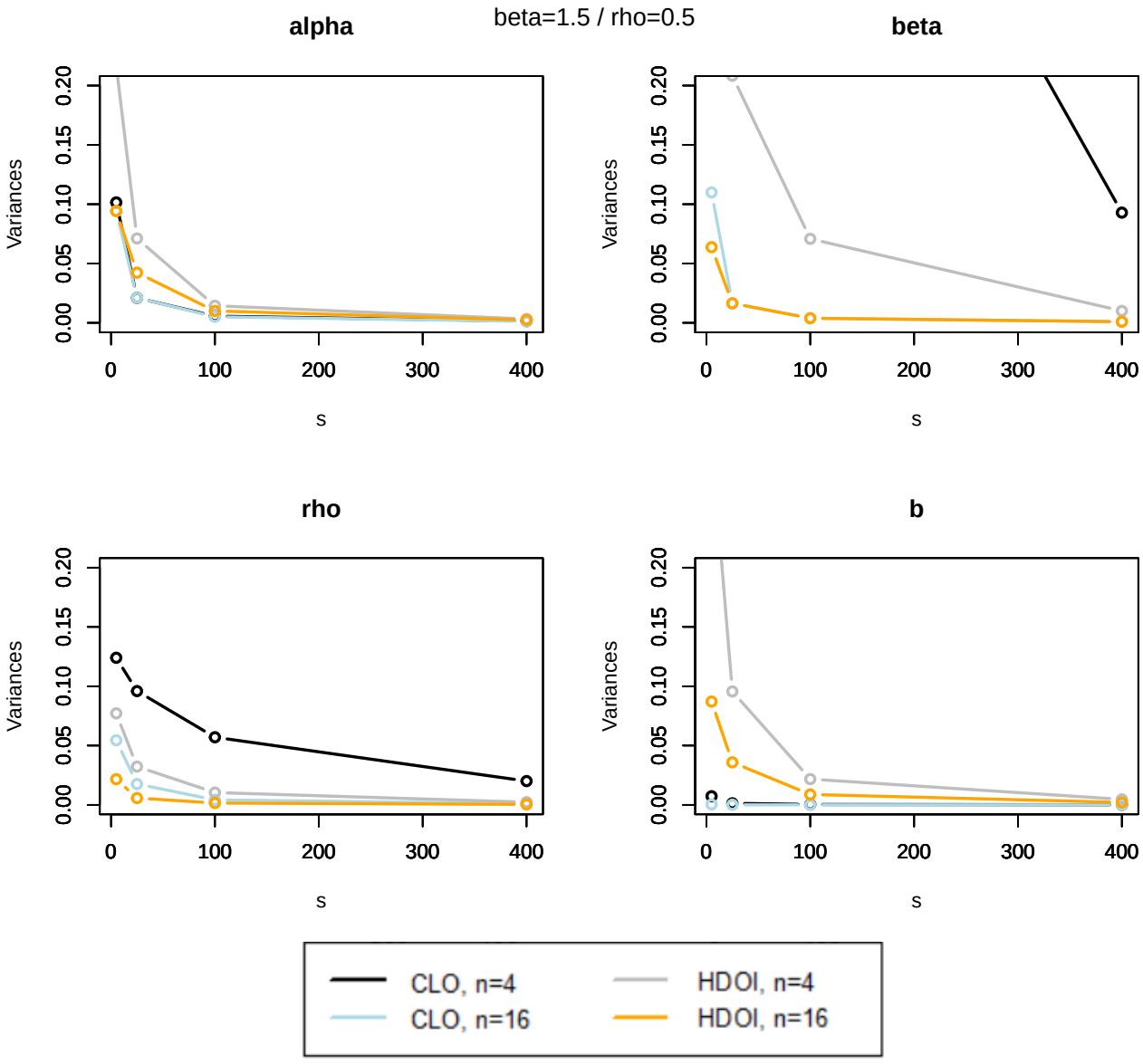

Figure C.16: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.5)$.

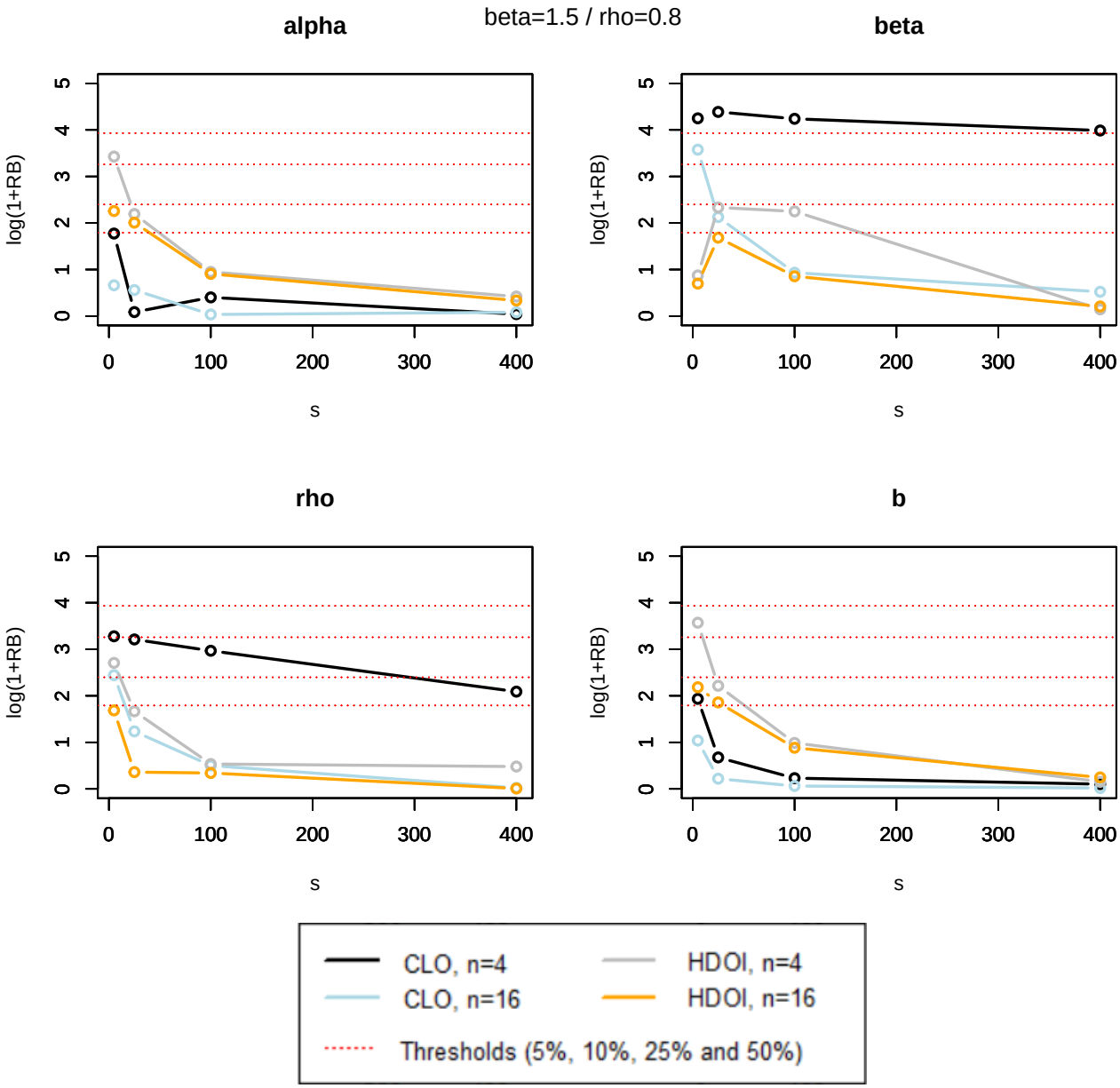

Figure C.17: Plots of the quantity log(1+RB) related to the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.8)$.

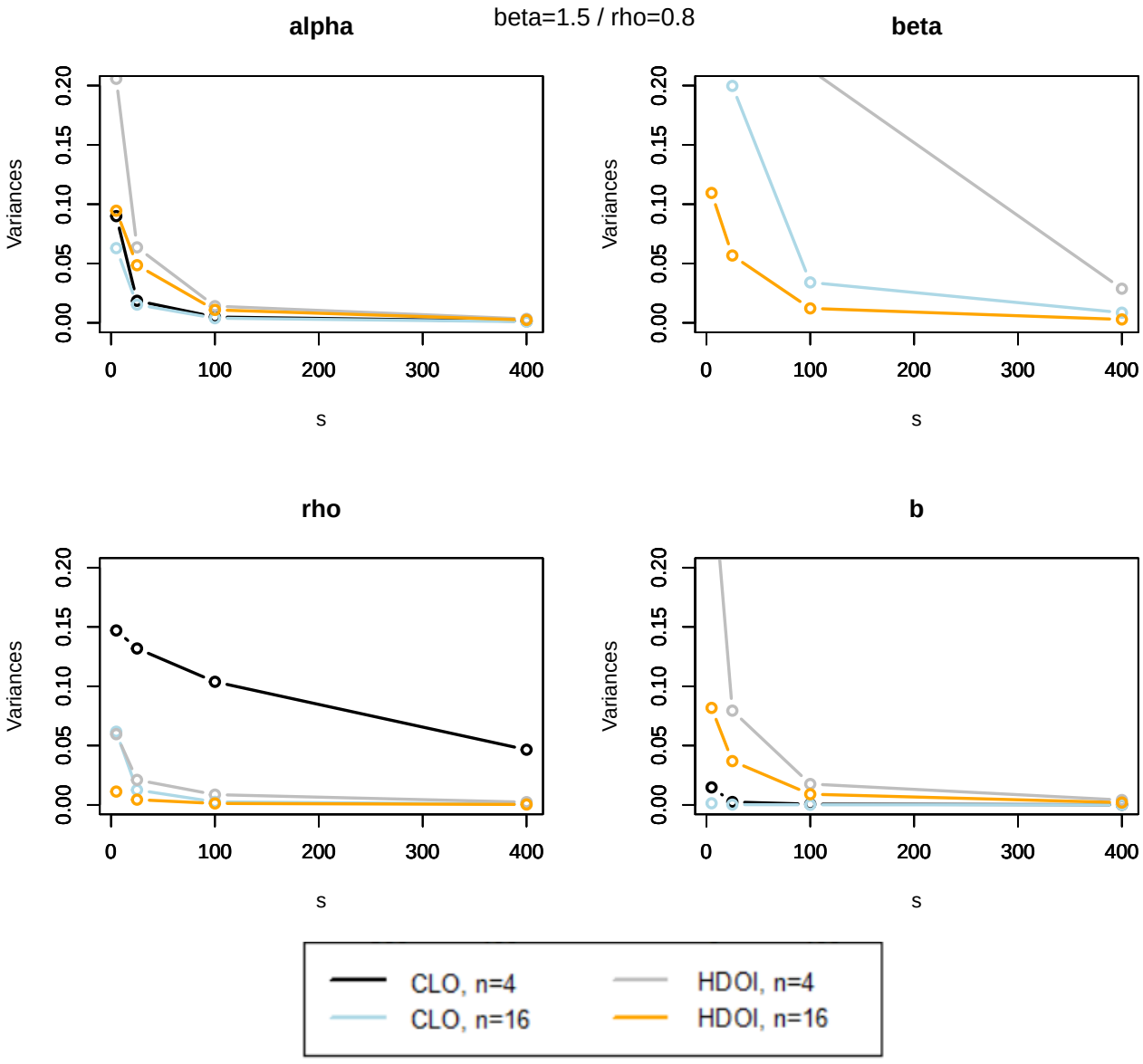

Figure C.18: Plots of the variance evolution of the parameters estimations depending on s, when $(\beta, \rho) = (1.5, 0.8)$.

Appendix D

Introduction (french version)

La sécurité et la fiabilité sont deux indicateurs cruciaux pour un grand nombre de structures (telles que, par exemple, les chemins de fer, les moteurs d'aéronefs ou les centrales nucléaires), ce qui a conduit au développement de la théorie de la fiabilité. Durant des années, seules les données de durée de vie étaient disponibles et les premières études fiabilistes se concentraient sur l'analyse de ces données (voir, par exemple, [32]), ce qui présente encore aujourd'hui un intérêt dans un large variété de cas. Dans ce contexte et dans le cas de système réparables, maintenus de manière instantanée, les instants de défaillances (ou de maintenances) successifs sont vus comme des points d'arrivée d'un processus de comptage, et les défaillances correspondent donc à des événements récurrents. Il existe plusieurs types de maintenance. Les modèles classiques considèrent des maintenances parfaites et minimales, dont les processus de comptage sous-jacents sont respectivement décrits par des processus de renouvellement et des processus de Poisson non homogènes (voir [5]). En s'intéressant à l'intensité de défaillance, par exemple, l'effet de ces deux types de maintenance est illustré sur la **Figure D.1** : une maintenance minimale n'affecte pas l'intensité de défaillance, tandis qu'une maintenance parfaite la réduit à sa valeur initiale. Cependant, la réalité se situe souvent entre ces deux extrêmes, d'où l'introduction des maintenances imparfaites (aussi représentée sur la **Figure D.1**). Dans la littérature, un grand nombre de modèles ont été envisagés pour les modéliser, tels que, par exemple, les modèles d'âge virtuel introduits par Kijima [26], traités de façon plus approfondie dans [13, 16], et davantage dans [7] où les auteurs ajoutent des covariables au modèle d'âge virtuel. D'autres possibilités sont les processus géométriques [27] (plus largement étudié dans [6], et plus récemment dans [12]) ou, comme évoqués plus haut, les modèles basés sur la réduction de l'intensité de défaillance [13, 16]. Voir, par exemple, [17] pour un compte rendu récent et des extensions pour de tels modèles. Voir aussi [36] pour plus de références et d'autres modèles.

Aujourd'hui, le développement de la surveillance en ligne et l'utilisation croissante de capteurs permettent d'obtenir des informations spécifiques sur l'état d'un système et sur son évolution au cours du temps, sans avoir à attendre une panne du système. L'information obtenu est souvent traduite par un nombre réel, pouvant représenter par exemple la longueur d'une fissure, l'épaisseur d'un câble, le niveau de corrosion, ... Cet indicateur peut être considéré comme une mesure du niveau de dégradation du système. De nos jours, l'évolution de ce nombre réel au cours du temps est communément modélisée par un processus stochastique, souvent considéré comme ayant une tendance positive. Les modèles classiques incluent le processus inverse Gaussien [44] ou les processus de Wiener (avec tendance) [22, 29, 45]. Dernièrement, le processus de Wiener transformé a également été introduit par [20], et davantage étudié dans [18], où les incréments de la dégradation peuvent être négatifs. Tous ces processus stochastiques sont relativement commun dans des domaines variés autre que celui de la fiabilité, tel que la finance, les assurances

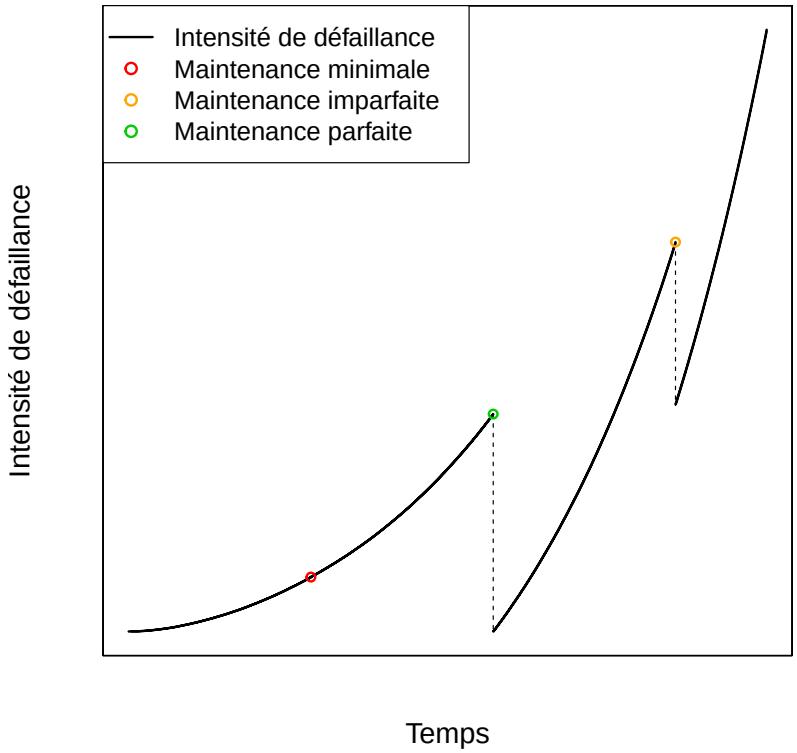

Figure D.1

ou l'épidémiologie. Cette thèse est dédiée au processus gamma, largement répandus depuis qu'ils ont été introduits dans le domaine de la fiabilité par Abdel-Hameed [1] et Çinlar [10]. Ce processus étant monotone, il est parfaitement adapté pour la modélisation de l'évolution de dégradation croissante.

Avant de définir le processus gamma, la loi gamma est définie, permettant d'introduire les notations et la paramétrisation utilisées dans la suite du document.

Une variable aléatoire X suit une loi gamma de paramètre de forme $a > 0$ et de paramètre d'échelle $b > 0$ (noté $X \sim \Gamma(a, b)$), si sa fonction de densité est telle que :

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \, \mathbf{1}_{\mathbb{R}^+}(x)$$

par rapport à la mesure de Lebesgue. L'espérance et la variance d'une telle variable sont données par $\mathbb{E}(X) = a/b$ et $\mathbb{V}(X) = a/b^2$.

On rappelle maintenant quelques propriétés très connues de la loi gamma :
Soient X_1 et X_2 deux variables aléatoires indépendantes suivant des lois gamma $\Gamma(a_1\,, b)$ et $\Gamma(a_2\,, b)$ respectivement, avec $a_1, a_2, b > 0$. Ainsi, pour tout c strictement positif, les variables aléatoires cX_1 et $X_1 + X_2$ suivent également des lois gamma $\Gamma\,(a_1\,, b/c)$ et $\Gamma(a_1 + a_2\,, b)$ respectivement.

On pose maintenant la fonction $a(\cdot) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ croissante et continue, telle que $a(0) = 0$, et on pose $b > 0$. Soit $(X_t)_{t\geq 0}$ un processus stochastique continue à droite avec limites à gauche. Le processus $(X_t)_{t\geq 0}$ est un processus gamma non homogène ayant pour fonction de forme $a(\cdot)$ et pour paramètre

d'échelle b, si

- $X_0 = 0$ presque sûrement;
- les incréments du processus sont indépendants;
- les incréments suivent des lois gamma, autrement dit pour tout $0 \leq s < t$, on a $X_t - X_s \sim \Gamma(a(t) - a(s), b)$,

(voir, par exemple, [1]). Voir [40] et ses références pour un large aperçu des processus gamma. Un exemple de trajectoires simulées selon un processus gamma est donné sur la **Figure D.2**.

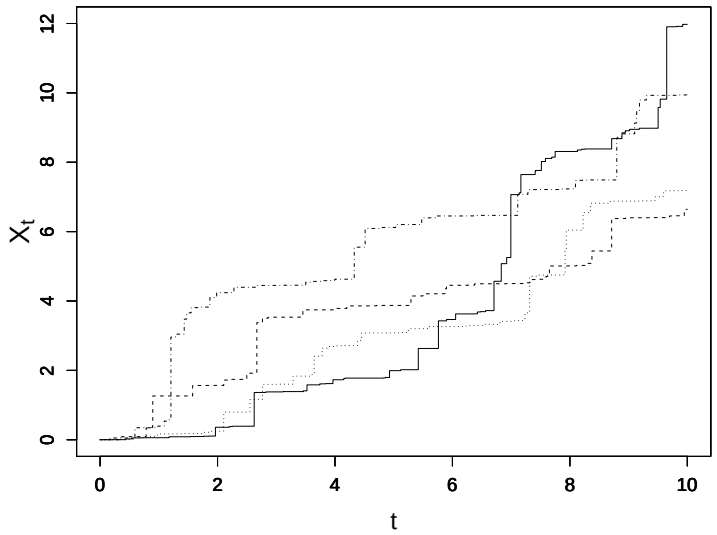

Figure D.2: Trajectoires simulées selon un processus gamma ayant pour fonction de forme $a(t) = t$ et pour paramètre d'échelle $b = 1$.

Dans le but d'atténuer la dégradation du système au cours du temps et ainsi rallonger sa durée de vie, des actions de maintenance préventives peuvent être effectuées, en plus des maintenances correctives qui sont réalisées lors d'une défaillance. Du point de vue de la dégradation, un grand nombre de politiques de maintenance présentes dans la littérature considèrent des maintenances conditionnelles, où une maintenance préventive est effectuée dès lors que la dégradation atteint un seuil donné. Dans ce cas, *la plupart des modèles de maintenance conditionnelle existants se limitent à des maintenances parfaites* ("most of the existing CBM models have been limited to perfect maintenance actions"), comme relevé dans [3] (ou encore [45]). Cependant, plusieurs modèles de maintenance imparfaite apparaissent dans la littérature récente, dans le contexte de la dégradation de systèmes, voir [3] pour une récente revue de la littérature. Certains modèles sont basés sur la notion d'âge virtuel, précédemment introduite dans le cadre des évènements récurrents (voir, par exemple, [19, 33]), pour lesquels le système est rajeuni par une maintenance. D'autres modèles supposent qu'un maintenance imparfaite réduit le niveau de dégradation du système, tels que [25, 28, 35, 37, 42], et pouvant également être accompagnée par une augmentation de l'intensité de dégradation, comme dans [15]. Dans d'autres articles, l'efficacité des maintenances est supposée décroissante avec le nombre de maintenances (voir, par exemple, [30, 46]), et d'autres études plus poussées, comme dans [23], considèrent des modèles de maintenance imparfaite tels que (i) les maintenances ont une efficacité aléatoire (ii) l'intensité de dégradation augmente avec le nombre de maintenance.

Dans tous ces articles, cependant, le point essentiel est l'optimisation de la politique de maintenance, en considérant à la fois des maintenances parfaites (remplacements) et imparfaites. À notre connaissance, peu d'articles traitent des modèles de maintenance imparfaite d'un point de vue statistique, excepté [45], où les auteurs proposent un méthode du maximum de vraisemblance pour estimer les paramètres du processus de Wiener (décrivant la dégradation du système non maintenu) associée à une procédure itérative basée sur un filtre de Kalman pour les différents facteurs impliqués dans les maintenances imparfaites successives.

Dans le cadre de systèmes se détériorant et sujets à des maintenances imparfaites, l'estimation des paramètres du processus de dégradation sous-jacent et de l'efficacité des maintenances est d'une grande utilité pour l'optimisation des politiques de maintenance. En effet, une fois les paramètres estimés, des prédictions sur le devenir du système maintenu peuvent être faites, permettant alors par exemple d'adapter (optimiser) la périodicité des actions de maintenance et de planifier une révision générale. D'un point de vue sécurité, la principale préoccupation est de s'assurer que les maintenances sont suffisamment efficaces, afin de maintenir avec une forte probabilité le niveau de dégradation sous un seuil donné. Tant que ce seuil n'est pas atteint, les actions de maintenance peuvent être ajustées, soit en adaptant leur périodicité, soit en augmentant leur efficacité (lorsque c'est possible). Mis à part la sécurité, les coûts de maintenance est évidemment un autre axe important. C'est par exemple ce qui est fait dans [42], où la minimisation des coûts est basée sur la durée de surveillance du système et de l'efficacité des actions de maintenance. Dans [23], l'auteur considère un seuil de dégradation au delà duquel un maintenance imparfaite est réalisée. L'optimisation est faite en fonction de ce seuil ainsi que la périodicité des inspections. Finalement, dans [43], une politique de maintenance est proposée, où un remplacement est effectué soit lorsque la dégradation dépasse un certain seuil, soit lorsqu'un nombre fixé de maintenances imparfaites ont été réalisées. Voir, par exemple, ces trois articles et leurs références pour un aperçu en rapport avec l'optimisation de politiques de maintenance.

Cette thèse est dédiée au développement et à l'application de procédures d'estimation pour trois modèles spécifiques de maintenance imparfaite dans le cadre de systèmes se dégradant suivant un processus gamma. Le document est divisé en quatre partie, incluant cette introduction.

Dans la **Partie II**, deux modèles de maintenance imparfaite sont étudiés. Il s'agit des modèles de réduction arithmétique de la dégradation d'ordre un et infini (ARD_1 et ARD_∞, ARD signifiant *Arithmetic Reduction of Degradation*), pour lesquels chaque maintenance réduit le niveau de dégradation du système. Le modèle ARD_1 a tout d'abord été introduit dans [8] et davantage étudié dans [34]. En se basant sur le modèle de réduction arithmétique de l'intensité (de défaillance) d'ordre un dans le cadre des évènements récurrents, l'idée de ce modèle est qu'une maintenance retire une proportion ρ de la dégradation accumulée par le système depuis la dernière maintenance (où $\rho \in [0, 1[$). Partant de cette même idée, [16] a également défini le modèle de réduction arithmétique de l'intensité d'ordre infini pour les évènements récurrents, que l'on adapte ici dans le contexte de la dégradation, conduisant à la définition du modèle ARD_∞. Concernant ce modèle, chaque maintenance réduit de $\rho\%$ le niveau de dégradation actuel, c'est-à-dire la dégradation accumulée par le système depuis sa mise en service. Une fois ces modèles définis, on se place dans un cadre paramétrique et le schéma d'observation est lui aussi défini. La méthode des moments et la Méthode du Maximum de Vraisemblance (MMV) sont développées pour les deux modèles dans le **Chapitre 2**. Plus précisément, l'identifiabilité est étudiée, et on donne ensuite une expression de l'estimateur des paramètres. Des tests numériques basés sur des jeux de données simulés sont réalisés au sein du **Chapitre 3**. Dans le **Chapitre 4**, un estimateur original de ρ est proposé, qui

ne dépend pas des paramètres du processus gamma sous-jacent, mais uniquement des observations. Le cadre est alors semi paramétrique. L'idée de cet estimateur provient d'une étude préliminaire dans le cadre du maximum de vraisemblance. Il a été observé que lorsqu'un unique système est considéré, la valeur minimale possible pour ρ se comporte de façon intéressante dans le cas où la fonction de forme a du processus gamma sous-jacent est concave : cette valeur se rapproche très rapidement de la vraie valeur (inconnue) de l'efficacité des actions de maintenance lorsque le nombre de maintenance augmente. Cet estimateur semi paramétrique a tout d'abord été étudié dans le cadre du modèle ARD_1, travail publié par la suite dans un journal international (voir [38]) et qui est reproduit ici à l'identique dans les **Sections 4.1** à **4.6**. Dans la **Section 4.7**, cette procédure d'estimation est adapté pour le modèle ARD_∞, et ce travail est spécifique au présent document.

Dans la **Partie III**, on considère un modèle de maintenance imparfaite basé sur l'âge virtuel et introduit par [34] : le modèle de réduction arithmétique de l'âge (virtuel) d'ordre un (ARA_1, ARA signifiant *Arithmetic Reduction of Age*). Suivant l'idée de [16] dans le cadre des évènements récurrents, ce modèle est tel que chaque action de maintenance retire une proportion ρ de l'âge du système accumulé depuis la dernière maintenance. Contrairement au modèle ARD_1, le système est ici rajeuni, c'est-à-dire qu'il est remis dans l'état exact où il se trouvait a un instant donnée précédent la maintenance. Les premiers axes des travaux de cette partie sont similaires à ceux de la partie précédente : le modèle et le schéma d'observation sont définis, la méthode des moments est étudiée d'un point de vue identifiabilité, puis on donne une expression de l'estimateur des paramètres du modèle. Ensuite, la MMV est développée. Cependant, à cause d'un problème de dépendance, la fonction de vraisemblance se révèle être un produit d'intégrales en grande dimension, et cela pose problème pour estimer les paramètres numériquement de façon classique. Ainsi, la MMV nécessite d'approcher numériquement les intégrales par des méthodes de Monte Carlo et Quasi Monte Carlo randomisé. Mis à part cela, afin de contourner l'approximation numérique d'intégrales en grande dimension, des méthodes alternatives basées sur le maximum de vraisemblance sont étudiées : le maximum de vraisemblance composite et la méthode du demi échantillon. La première méthode consiste à supposer que les observations sont indépendantes, tandis que la seconde ne prend en compte qu'une observation sur deux. Une fois de plus, le développement de ces méthodes se traduit par l'étude de l'identifiabilité des paramètres, qui a été complexe pour une des méthodes, ainsi que l'expression des estimateurs (ou de la log-vraisemblance). Cependant, l'identifiabilité n'a pas été vérifiée concernant deux des six méthodes proposées car les tests numériques ont révélé des problèmes d'identifiabilité pour ces deux méthodes. Enfin, dans le **Chapitre 7**, les performances numériques des méthodes sont illustrées en deux temps. Une première étude permet d'éliminer les méthodes peu performantes, et dans un second temps, une étude plus poussée nous permet de sélectionner, parmi les méthodes restantes, la plus appropriée en fonction du cadre d'application.

Pour finir, la **Partie IV** met en évidence les conclusions de la thèse ainsi que les perspectives.

Appendix E

Conclusion (french version)

Dans ce document, des modèles de maintenance imparfaite ont été étudiés dans le cadre de systèmes se dégradant suivant un processus gamma. Plus précisément, des méthodes d'estimation ont été développées pour les modèles de réduction arithmétique de la dégradation (ARD pour *Arithmetic Reduction of Degradation*) d'ordre un et infini (notés ARD_1 et ARD_∞) dans la **Partie II**, ainsi que pour le modèle de réduction arithmétique de l'âge (ARA pour *Arithmetic Reduction of Age*) d'ordre un (noté ARA_1) dans la **Partie III**.

Dans la **Partie II**, on traite tout d'abord deux méthodes d'estimation classiques dans un cadre paramétrique : la méthode des moments et celle du maximum de vraisemblance. L'identifiabilité des paramètres a été étudiée, les expressions des estimateurs ont été donnée et la performance des méthodes a été illustrée. Par la suite, découlant de l'étude de la méthode du maximum de vraisemblance, un estimateur de l'efficacité des maintenances (le paramètre ρ) est proposé dans un cadre semi paramétrique. Dans le cas où un unique système est observé et lorsque la fonction de forme du processus gamma sous-jacent est concave, la convergence presque sûre de cet estimateur lorsque le nombre de maintenance tend vers l'infini est démontrée, associée à une vitesse de convergence étonnamment élevée, au moins exponentielle dans certains cas particuliers. Ces travaux sont ensuite généralisés au cas où s systèmes identiques et indépendants sont observés. Un estimateur semi paramétrique similaire est proposé pour ρ, et la convergence presque sûre de cet estimateur lorsque s tend vers l'infini est vérifiée, et ce peu importe le nombre de réparations n (fixé) et sous aucune hypothèse. La vitesse de convergence a aussi été étudiée, il est démontré qu'elle dépend de la fonction de forme du processus gamma et de la périodicité des maintenances, conduisant à une vitesse de convergence pouvant être plus lente ou plus rapide que $\sqrt{s}$ selon les cas.

Dans la **Partie III**, plusieurs méthodes d'estimation sont proposées dans le cadre d'une modèle ARA_1, qui sont basées soit sur les observations, soit sur les incréments, conduisant à six méthodes différentes. Selon la méthode traitée, soit une expression des estimateurs, ou soit la log-vraisemblance, est donnée, et l'identifiabilité des paramètres du modèle a été étudiée pour quatre des six méthodes. Dans le but d'étudier et comparer les performances de ces méthodes, des tests numériques basés sur des données simulées ont été menés à grande échelle. Deux méthodes se sont révélées être plus efficaces que les autres : la méthode du demi échantillon basée sur les indices impairs, ainsi que la méthode du maximum de vraisemblance composite basée sur les observations. Cependant, d'un point de vue global, aucune des deux méthodes ne se détache. Ainsi, le choix de la méthode doit être fait selon le contexte. Plus précisément, ce choix dépend du nombre de maintenances, de leur efficacité et de l'évolution de l'intensité de

dégradation au cours du temps.

Il y a plusieurs points non abordés qu'il serait intéressant d'étudier en complément de cette thèse. Il existe beaucoup d'autres modèles dans la continuité des modèles ARD étudiés ici, par exemple les modèles ARD_m qui reposent sur l'idée qu'une action de maintenance réduit d'une certaine proportion la dégradation accumulée par le système depuis les m dernières maintenances. Une des axes de travail envisageable est de généraliser l'estimateur semi paramétrique de l'efficacité des maintenances à un tel modèle, ce qui semble possible car cette méthode repose sur la positivité des incréments du processus gamma. De plus, il serait intéressant d'adapter cette procédure d'estimation en considérant un processus de Lévy monotone autre que le processus gamma. Finalement, en se basant sur une rapide analyse numérique, il a été illustré que la condition qui assure une vitesse de convergence exponentielle est suffisante, mais non nécessaire. Des études plus approfondies pourraient être menées dans le but de savoir s'il est possible d'affiner les conditions sous lesquelles les différentes vitesses de convergence sont obtenues.

Concernant les travaux menés dans la **Partie III**, des méthodes d'estimations pourraient être améliorées parmi celles qui sont proposées. Comme il a été précisé, il serait intéressant de voir si la qualité des estimations est améliorée en combinant la méthode du demi échantillon basée sur les indices impairs avec la méthode du maximum de vraisemblance composite basée sur les observations. De plus, il a été expliqué que les approximations des intégrales par les méthodes de Monte Carlo et Quasi Monte Carlo pouvaient être améliorées en ayant recours à de l'échantillonnage préférentiel. Des travaux supplémentaires pourraient être faits pour aborder ces possibles améliorations.

En plus de cela, mis à part dans le **Chapitre 4**, les propriétés asymptotiques des estimateurs n'ont pas été étudiées dans ce document et une étude complémentaire est requise pour une meilleure compréhension de leur comportement asymptotique, au delà des expériences numériques effectuées dans cette thèse. Ces propriétés pourraient alors être étudiées dans la continuité des travaux actuels.

Au delà de ces points spécifiques, un étude future pourrait concerner le développement de procédures d'estimation dans le cadre d'un modèle ARA_∞, pour lequel une maintenance réduit d'une proportion donnée l'âge du système accumulé depuis sa mise en service. D'avantage de recherches peuvent être faites dans ce contexte, en imitant les travaux effectués pour le modèle ARA_1 dans un premier temps.

Finalement, un dernier point à étudier peut être de considérer un schéma d'observation indépendant des opérations de maintenance, ce qui n'est pas le cas ici étant donnée que le niveau de dégradation est mesuré juste avant chaque maintenance. Par exemple, les observations pourraient toujours être conduites de manière périodique, tandis que le système pourrait être maintenu à des instants aléatoires suivant un processus de Poisson. Cependant, des études préliminaires ont été faites dans ce cadre, et le niveau de dégradation du système maintenu semble difficile à exprimer, mais reste exploitable pour le développement de procédures d'estimation. Ainsi, cela semble être un ambitieux, mais intéressant, projet pour de futures recherches.

Appendix F

Abstract (french version)

La thèse s'intéresse à l'étude de modèles de maintenance imparfaite dans le cadre d'une système se dégradant suivant un processus gamma. Plus précisément, trois modèles particuliers sont considérés, et le but est de développer des méthodes d'estimation pour les paramètres de chacun de ces modèles. Pour commencer, une définition brève du processus gamma est donnée.

Soit la fonction $a(\cdot) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ croissante et continue, et telle que $a(0) = 0$, et $b > 0$. Soit $(X_t)_{t\geq 0}$ un processus stochastique continue à droite avec limites à gauche. Le processus $(X_t)_{t\geq 0}$ est un processus gamma non homogène ayant pour fonction de forme $a(\cdot)$ et pour paramètre d'échelle b, si

- $X_0 = 0$ presque sûrement;
- le processus est à incréments indépendants;
- les incréments suivent des lois gamma, c'est-à-dire que pour tout $0 \leq s < t$, on a $X_t - X_s \sim \Gamma(a(t) - a(s), b)$,

(voir, par exemple, [1]). Dans l'ensemble des travaux, on considère un système dont le niveau de dégradation évolue selon un processus gamma $(X_t)_{t\geq 0}$ comme défini ci-dessus, et ayant pour fonction de forme et paramètre de d'échelle $a(\cdot)$ et b respectivement. On suppose également que le système est maintenu de manière imparfaite, et les actions de maintenance sont supposées instantanées et périodiques de période T connue.

La première partie étudie un type spécifique de modèle de maintenance imparfaite : les modèles de réduction arithmétique de la dégradation, dits ARD pour *Arithmetic Reduction of Degradation*. Plus exactement, deux modèles ARD sont considérés, qui sont les modèles ARD d'ordre un et d'ordre infini (respectivement ARD_1 et ARD_∞). On donne maintenant la définition du modèle ARD_1, la définition de l'autre modèle sera ensuite donnée.

Pour un modèle ARD_1, une action de maintenance a pour effet de réduire d'une proportion donnée $\rho \in [0, 1)$ la dégradation du système accumulée depuis la dernière maintenance. L'efficacité des actions de maintenance est alors mesurée par le paramètre ρ, ainsi, l'action de maintenance à l'instant jT réduit de $\rho\%$ la dégradation accumulée par le système durant l'intervalle de temps $[\,(j-1)T\,, jT\,[$.

On note $(Y_t)_{t\geq 0}$ le processus stochastique décrivant l'évolution du niveau de dégradation du système maintenu suivant le modèle ARD_1. On note également $(X^{(j)})_{j\in\mathbb{N}^*}$ une suite de copies indépendantes de

$(X_t)_{t\geq 0}$, où $X^{(j)}$ correspond à la dégradation intrinsèque (sans maintenance) du système entre les instants $(j-1)T$ et jT. Le système est supposé en parfait état à l'instant initial $t=0$, soit $Y_0 = X_0^{(1)} = 0$. Entre l'instant initial et la première maintenance à l'instant T, le système se dégrade suivant le processus $X_t^{(1)}$. Ainsi

$$Y_t = X_t^{(1)} \text{ et } Y_T = (1-\rho)\left(X_T^{(1)} - X_0^{(1)}\right) = (1-\rho)X_T^{(1)}.$$

Pour $t \in [T\,, 2T[$, la dégradation du système correspond à la somme de la dégradation accumulée sur $[t, 2T[$, soit $X_t^{(2)} - X_T^{(2)}$, et du niveau de dégradation juste après la première maintenance, donc

$$Y_t = Y_T + \left(X_t^{(2)} - X_T^{(2)}\right)$$

et

$$Y_{2T} = Y_{2T^-} + \rho\left(X_{2T^-}^{(2)} - X_T^{(2)}\right) = (1-\rho)\left(\left(X_{2T}^{(2)} - X_T^{(2)}\right) + \left(X_T^{(1)} - X_0^{(1)}\right)\right).$$

Ainsi, pour tout t dans $[\,nT\,, (n+1)T\,[$, avec n dans $\mathbb{N}^*$, on a :

$$Y_{nT} = (1-\rho)\sum_{j=1}^{n}\left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right)$$

et

$$Y_t = Y_{nT} + \left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right).$$

où les variables aléatoires Y_{nT} et $\left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right)$ suivent des lois gamma $\Gamma\left(a(nT)\,, b/(1-\rho)\right)$ et $\Gamma\left(a(t) - a(nT)\,, b\right)$ respectivement.

Concernant le modèle ARD_∞, une action de maintenance à l'instant jT réduit le niveau de dégradation à cet instant de $\rho\%$, soit la dégradation du système accumulée depuis l'instant initial $t=0$. Ce modèle et le précédent se comportent de manière identique sur l'intervalle de temps $[0\,, 2T\,[$. Ainsi, en conservant les mêmes notations que pour le modèle précédent, on peut directement écrire que pour tout t dans $[0\,, T\,[$ on a

$$Y_t = X_t^{(1)} \text{ et } Y_T = (1-\rho)\left(X_T^{(1)} - X_0^{(1)}\right) = (1-\rho)X_T^{(1)},$$

et pour tout t dans $[\,T\,, 2T\,[$

$$Y_t = Y_T + \left(X_t^{(2)} - X_T^{(2)}\right).$$

Lorsque la seconde maintenance survient, une proportion ρ de la dégradation accumulée depuis l'instant $t=0$ est retirée, ainsi $Y_{2T} = (1-\rho)\left(Y_T + \left(X_{2T}^{(2)} - X_T^{(2)}\right)\right)$, qui peut être écrit comme

$$Y_{2T} = (1-\rho)^2\left(X_T^{(1)} - X_0^{(1)}\right) + (1-\rho)\left(X_{2T}^{(2)} - X_T^{(2)}\right).$$

Finalement, pour tout t dans $[nT\,,(n+1)T\,[$ et n dans $\mathbb{N}$, on déduit une expression pour Y_t, qui est :

$$Y_t = \sum_{j=1}^{n}(1-\rho)^{n-j+1}\left(X_{jT}^{(j)} - X_{(j-1)T}^{(j)}\right) \;+\; \left(X_t^{(n+1)} - X_{nT}^{(n+1)}\right),$$

et où les incréments sont indépendants et suivent des lois gamma.

Une fois établie la définition formelle de ces modèles, la suite de cette première partie est elle même divisée en deux axes, tous deux basés sur le développement de méthodes d'estimation pour chacun des modèles. Pour cela, un schéma d'observation est tout d'abord défini, où l'on considère que s systèmes indépendants et identiquement distribués (i.i.d.) sont observés n fois chacun, juste avant les n premières maintenances, soit aux instants jT^- pour $1 \leq j \leq n$. On se place ensuite dans le contexte particulier où la fonction de forme a du processus gamma sous-jacent est de type power law, soit $a : t \mapsto \alpha t^\beta$ avec α, $\beta > 0$. Ainsi, les paramètres d'intérêt sont les paramètres de la fonction de forme α et β, le paramètre d'échelle b, ainsi que l'efficacité des actions de maintenance ρ.

Dans un premier temps, deux méthodes d'estimation classiques sont développées. Il s'agit de la méthode des moments et de celle du maximum de vraisemblance. Une brève introduction de chaque méthode est faite, l'identifiabilité des paramètres est ensuite étudiée et enfin une expression des estimateurs est donnée. Ensuite, des tests numériques sont menés sur des jeux de données simulés, et pour finir les résultats sont analysés.

En se basant sur une particularité révélée lors du l'étude du maximum de vraisemblance, un un estimateur original pour l'efficacité des actions de maintenance est développé. Cet estimateur dépend uniquement des observations, et non des paramètres du processus gamma sous-jacent, il s'agit donc d'un estimateur semi paramétrique.

Une fois cette estimateur défini et plusieurs résultats techniques intermédiaires donnés, cette méthode d'estimation est en premier lieu étudiée en considérant un unique système ($s = 1$). Lorsque la fonction de forme du processus gamma sous-jacent est concave, la convergence presque sûre de cet estimateur est vérifiée lorsque n tend vers l'infini. De plus, sous certaines hypothèses techniques plus fortes, la vitesse de convergence de cet estimateur est particulièrement élevée, jusqu'à atteindre une vitesse exponentielle. Cette méthode est ensuite généralisée au cas où plusieurs systèmes i.i.d. sont observés. La convergence presque sûre est encore une fois vérifiée lorsque n est fixé et s tend vers l'infini, et ce sous aucune hypothèse sur la fonction de forme. La vitesse de convergence est ici plus faible que précédemment, mais peut tout de même est plus haute que la vitesse de convergence classique $\sqrt{s}$.

Dans la seconde partie, un autre modèle de maintenance imparfaite est étudié, qui est un modèle d'âge virtuel. Il s'agit du modèle de réduction arithmétique de l'âge (virtuel) d'ordre un (ARA_1 pour *Arithmetic Reduction of Age of order one*). Pour un tel modèle, une action de maintenance réduit l'âge virtuel du système de ρT unités de temps. En d'autre termes, une maintenance retire $\rho\%$ de l'âge accumulé par le système depuis la dernière maintenance. Ainsi, on peut écrire que

$$V(t) = t - \rho k T \text{ pour } kT \leq t < (k+1)\,T$$

où $V(t)$ est l'âge virtuel du système à l'instant t. On remarque que $V(kT) = (1-\rho)\,kT$ et $V(kT^-) = [(1-\rho)\,k + \rho]\,T$. De plus, la k-ème action de maintenance remet le système dans l'état où il se trouvait à l'instant $kT - \rho T = (k-\rho)\,T$. L'âge virtuel correspondant est

$$V((k-\rho)\,T) = (k-\rho)\,T - \rho\,(k-1)\,T = (1-\rho)\,kT = V(kT)\,.$$

ce qui est cohérent car à l'instant kT, le système est remis dans l'état où il se trouvait à l'instant $(k-\rho)\,T$. Cela implique que

$$Z_{(k-\rho)T} = Z_{kT}$$

et finalement

$$V(t) = \sum_{j\geq 0} (t - \rho jT)\,\mathbf{1}_{[jT,(j+1)T)}(t).$$

Ainsi, le niveau de dégradation Z_t pour $nT < t \leq (n+1)T$, avec n dans $\mathbb{N}^*$, peut s'écrire comme

$$Z_t = Z_{nT} + \left(X^{(n+1)}_{V(t)} - X^{(n+1)}_{V(nT)}\right) = Z_{nT} + \left(X^{(n+1)}_{t-\rho nT} - X^{(n+1)}_{(1-\rho)nT}\right)$$

avec

$$Z_{nT} = Z_{(n-1)T} + \left(X^{(n)}_{V(nT)} - X^{(n)}_{V((n-1)T)}\right) = Z_{(n-1)T} + \left(X^{(n)}_{(1-\rho)nT} - X^{(n)}_{(1-\rho)(n-1)T}\right),$$

et où $\left(X^{(n+1)}_{t-\rho nT} - X^{(n+1)}_{(1-\rho)nT}\right)$ suit une loi gamma $\Gamma\big(a(t-\rho nT) - a((1-\rho)nT), b\big)$. De plus, on peut écrire Z_{nT} comme

$$Z_{nT} = \sum_{j=1}^{n} \left(X^{(j)}_{(1-\rho)jT} - X^{(j)}_{(1-\rho)(j-1)T}\right),$$

qui suit également une loi gamma dont les paramètres de forme et d'échelle sont $a((1-\rho)nT)$ et b respectivement. Ainsi, Z_t est la somme de deux variables aléatoires indépendantes suivant des lois gamma de même paramètre d'échelle, ce qui entraîne que Z_t suit aussi une loi gamma $\Gamma\big(a(t-\rho nT), b\big)$.

Des méthodes d'estimations sont ensuite développées dans le même contexte que celui de la première partie, c'est-à-dire que l'on conserve ici le même schéma d'observation ainsi que la même fonction de forme tels qu'ils sont définis plus haut. Ainsi, les paramètres à estimer restent inchangés. Les méthodes développées sont basées soit sur les observations, soit sur les incréments des observations, i.e. les variables aléatoires $Z_{jT^-} - Z_{(j-1)T^-}$ pour $1 \leq j \leq n$. Une fois encore la méthode des moments est traitée, basée sur les observations. La méthode du maximum de vraisemblance est ensuite abordée, cette fois-ci basée sur les incréments. Cependant, ces incréments n'étant pas indépendants, la fonction de vraisemblance se révèle être un produit d'intégrales en dimension $n-1$, et cela pose problème pour estimer les paramètres numériquement de façon classique. Ainsi, un travail préliminaire est fait dans le but de ramener le domaine d'intégration à $[0,1]^{n-1}$, puis le maximum de vraisemblance est traité en utilisant les méthodes de Monte Carlo et Quasi Monte Carlo randomisé pour les approximations d'intégrales.

En plus de cela, afin de contourner le problème des intégrales en grande dimension, quatre méthodes alternatives basées sur le maximum de vraisemblance sont étudiées : la méthode du demi échantillon et le maximum de vraisemblance composite. La première méthode, basée sur les incréments, consiste à ne considérer qu'un incrément sur deux. Étant donné que le j-ème incrément est dépendant des $j-1$-ème et $j+1$-ème, n'en choisir qu'un sur deux permet d'obtenir l'indépendance. Ainsi, suivant le principe du maximum de vraisemblance, une expression de la vraisemblance basée sur la moitié des incréments est donnée, nécessitant dans ce cas l'approximation d'intégrales de dimensions 1 seulement. Deux méthodes

sont développées suivant ce procédé, la première considérant les incréments dont les indices sont impairs et la seconde ceux dont les indices sont pairs. Le maximum de vraisemblance composite considérée par la suite repose sur une hypothèse simple, celle de l'indépendance. Dans ce cas deux méthodes sont également développées, l'une basée sur les observations (supposées indépendantes) et l'autre sur les incréments (supposés indépendants). Concernant cette méthode, lorsque les incréments sont considérés, la vraisemblance nécessite une fois de plus l'approximation d'intégrales de dimensions 1, tandis qu'il n'y a aucune intégrale dans l'expression de la vraisemblance composite basée sur les observations. En parallèle, l'identifiabilité des paramètres du modèle est étudiée pour chacune des méthodes, exceptées la méthode du demi échantillon basée sur les indices pairs et la méthode du maximum de vraisemblance composite basée sur les incréments. En effet, lors des tests numériques effectués, ces deux méthodes ont révélé des problèmes d'identifiabilité.

Pour finir, l'ensemble de ces méthodes ont été testées numériquement sur des données simulées. Cette étude s'est déroulée en deux temps. Tout d'abord, les tests ont été réalisés sur de larges jeux de données, dans le but de voir quelles méthodes sont les plus fiables et lesquelles ne le sont pas. Deux méthodes se sont révélées être plus efficaces, et ont ensuite été testées à plus grande échelle. Finalement, les résultats sont analysés dans le but de conclure quelle méthode est la plus fiable. Cependant, aucune des méthodes ne se détache nettement. Ainsi, le choix de la méthode se fait en fonction du contexte d'application. Plus précisément, le choix se fait en fonction du nombre de maintenance n, de leur efficacité ρ, ainsi qu'en fonction de l'évolution de l'intensité de dégradation, plus précisément de la convexité de la fonction de forme (qui dépend de la valeur du paramètre β).