Consistency of the posterior distribution and MLE for piecewise linear regression

Tristan Launay, Anne Philippe, Sophie Lamarche

▶ **To cite this version:**

Tristan Launay, Anne Philippe, Sophie Lamarche. Consistency of the posterior distribution and MLE for piecewise linear regression. 2012. ⟨hal-00681274v2⟩

HAL Id: hal-00681274

https://hal.science/hal-00681274v2

Preprint submitted on 20 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HAL Authorization

Consistency of the posterior distribution and MLE for piecewise linear regression

Tristan Launay[1,2] Anne Philippe[1] Sophie Lamarche[2]

[1] Laboratoire de Mathématiques Jean Leray,
2 Rue de la Houssinière – BP 92208, 44322 Nantes Cedex 3, France
[2] Electricité de France R&D, 1 Avenue du Général de Gaulle,
92141 Clamart Cedex, France

Abstract

We prove the weak consistency of the posterior distribution and that of the Bayes estimator for a two-phase piecewise linear regression mdoel where the break-point is unknown. The non-differentiability of the likelihood of the model with regard to the break-point parameter induces technical difficulties that we overcome by creating a regularised version of the problem at hand. We first recover the strong consistency of the quantities of interest for the regularised version, using results about the MLE, and we then prove that the regularised version and the original version of the problem share the same asymptotic properties.
keywords : consistency ; asymptotic distribution ; posterior distribution ; MLE ; piecewise regression.

1 Introduction

We consider a continuous segmented regression model with 2 phases, one of them (the rightmost) being zero. Let u be the unknown breakpoint and $\gamma \in \mathbb{R}$ be the unknown regression coefficient of the non zero phase. The observations $X_{1:n} = (X_1, \dots, X_n)$ depend on an exogenous variable that we denote $t_{1:n} = (t_1, \dots, t_n)$ via the model given for $i = 1, \dots, n$ by

$$X_i = \mu(\eta, t_i) + \xi_i := \gamma \cdot (t_i - u)\mathbb{1}_{[t_i, +\infty[}(u) + \xi_i, \tag{1.1}$$

where $(\xi_i)_{i\in\mathbb{N}}$ is a sequence of independent and identically distributed (i.i.d.) random variables with a common centered Gaussian distribution of unknown variance σ^2, $\mathcal{N}(0, \sigma^2)$, and where $\mathbb{1}_A$ denotes the indicator function of a set A.

Such a model is for instance used in practice to estimate and predict the heating part of the electricity demand in France. See Bruhns et al. (2005) for the definition of the complete model and Launay et al. (2012) for a Bayesian approach. In this particular case, u corresponds to the heating threshold above which the temperatures $t_{1:n}$ do not have any effect over the electricity load, and γ corresponds to the heating gradient i.e. the strength of the described heating effect.

The work presented in this paper is most notably inspired by the results developed in Ghosh et al. (2006) and Feder (1975).

Feder proved the weak consistency of the least squares estimator in segmented regression problems with a known finite number of phases under the hypotheses of his Theorem 3.10 and some additional assumptions disseminated throughout his paper, amongst which we find that the empirical cumulative distribution functions of the temperatures at the n-th step $t_{n1}, \dots, t_{nn}$ are required to converge to a cumulative distribution function, say F_n converges to F, which is of course to be compared to our own Assumption (A1). Feder also derived the asymptotic distribution of the least squares estimator under the same set of assumptions. Unfortunately there are a few typographical errors in his paper (most notably resulting in the disappearance of σ_0^2 from the

asymptotic variance matrix in his main theorems), and he also did not include $\widehat{\sigma}_n^2$ in his study of the asymptotic distribution.

The asymptotic behaviour of the posterior distribution is a central question that has already been raised in the past. For example, Ghosh et al. worked out the limit of the posterior distribution in a general and regular enough i.i.d. setup. In particular they manage to derive the asymptotic normality of the posterior distribution under third-order differentiability conditions. There are also a number of works dealing with some kind of non regularity, like these of Sareen (2003) which consider data the support of which depends on the parameters to be estimated, or those of Ibragimov and Has'minskii (1981) which offer the limiting behaviour of the likelihood ratio for a wide range of i.i.d. models whose likelihood may present different types of singularity. Unfortunately, the heating part model presented here does not fall into any of these already studied categories.

In this paper, we show that the results of Ghosh et al. can be extended to a non i.i.d. two-phase regression model. We do so by using the original idea found in Sylwester (1965)[1]: we introduce a new, regularised version of the problem called pseudo-problem, later reprised by Feder. The pseudo-problem consists in removing a fraction of the observations in the neighbourhood of the true parameter to obtain a differentiable likelihood function. We first recover the results of Ghosh et al. for this pseudo-problem and then extend these results to the (full) problem by showing that the estimates for the problem and the pseudo-problem have the same asymptotic behaviour.

From this point on, we shall denote the parameters $\theta = (\gamma, u, \sigma^2) = (\eta, \sigma^2)$ and θ_0 will denote the true value of θ. We may also occasionally refer to the intercept of the model as $\beta = -\gamma u$. The log-likelihood of the n first observations $X_{1:n}$ of the model will be denoted

$$l_{1:n}(X_{1:n}|\theta) = \sum_{i=1}^{n} l_i(X_i|\theta) \tag{1.2}$$

$$= -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \sum_{i=1}^{n}\frac{1}{2\sigma^2}\left(X_i - \gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right)^2, \tag{1.3}$$

where $l_i(X_i|\theta)$ designates the log-likelihood of the i-th observation X_i, i.e.

$$l_i(X_{1:n}|\theta) = -\frac{1}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(X_i - \gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right)^2. \tag{1.4}$$

Notice that we do not mention explicitly the link between the likelihood l and the sequence of temperatures $(t_n)_{n\in\mathbb{N}}$ in these notations, so as to keep them as minimal as possible. The least square estimator $\widehat{\theta}_n$ of θ being also the maximum likelihood estimator of the model, we refer to it as the MLE.

Throughout the rest of this paper we work under the following assumptions

Assumption (A1). The sequence of temperatures (exogenous variable) $(t_n)_{n\in\mathbb{N}}$ belongs to a compact set $[\underline{u}, \overline{u}]$ and the sequence of the empirical cumulative distribution functions $(F_n)_{n\in\mathbb{N}}$ of $(t_1, \ldots, t_n)$, defined by

$$F_n(u) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{[t_i,+\infty[}(u),$$

converges pointwise to a function F where F is a cumulative distribution function itself, which is continuously differentiable over $[\underline{u}, \overline{u}]$.

Remark 1. Due to a counterpart to Dini's Theorem (see Theorem 7.1 taken from Polya and Szegö, 2004, (p81)), F_n converges to F uniformly over $[\underline{u}, \overline{u}]$.

[1] Sylwester indeed considers the same model as we do here, however his asymptotic results are false due to an incorrect reparametrisation of the problem and an error in the proof of his Theorem 3.5.

Remark 2. Let h be a continuous, bounded function on $[\underline{u}, \overline{u}]$. As an immediate consequence of this assumption, for any interval $I \subset [\underline{u}, \overline{u}]$, we have, as $n \to +\infty$

$$\frac{1}{n}\sum_{i=1}^{n} h(t_i)\mathbb{1}_I(t_i) = \int_I h(t)\,\mathrm{d}F_n(t) \to \int_I h(t)\,\mathrm{d}F(t) = \int_I h(t)f(t)\,\mathrm{d}t,$$

the convergence holding true by definition of the convergence of probability measures (see Billingsley, 1999, pages 14–16). In particular, for $I = [\underline{u}, \overline{u}]$ and $I =]-\infty, u]$ we get, as $n \to +\infty$

$$\frac{1}{n}\sum_{i=1}^{n} h(t_i) \to \int_{\underline{u}}^{\overline{u}} h(t)f(t)\,\mathrm{d}t, \qquad \frac{1}{n}\sum_{i=1}^{n} h(t_i)\mathbb{1}_{[t_i,+\infty[}(u) \to \int_{\underline{u}}^{u} h(t)f(t)\,\mathrm{d}t.$$

Remark 3. It is a general enough assumption which encompasses both the common cases of i.i.d. continuous random variables and periodic (non random) variables under a continous (e.g. Gaussian) noise.

Assumption (A2). $\theta_0 \in \Theta$, where the parameter space Θ is defined (for identifiability) as

$$\Theta = \mathbb{R}^* \times]\underline{u}, \overline{u}[\times \mathbb{R}_+^*,$$

where $\mathbb{R}^* = \{x \in \mathbb{R}\,, x \neq 0\}$ and $\mathbb{R}_+^* = \{x \in \mathbb{R}\,, x > 0\}$.

Assumption (A3). $f = F'$ does not vanish (i.e. is positive) on $]\underline{u}, \overline{u}[$.

Assumption (A4). There exists $K \subset \Theta$ a compact subset of the parameter space Θ such that $\widehat{\theta}_n \in K$ for any n large enough.

The paper is organised as follows. In Section 2, we present the Bayesian consistency (the proofs involved there rely on the asymptotic distribution of the MLE) and introduce the concept of pseudo-problem. In Section 3, we prove that the MLE for the full problem is strongly consistent. In Section 4 we derive the asymptotic distribution of the MLE using the results of Section 3: to do so, we first derive the asymptotic distribution of the MLE for the pseudo-problem and then show that the MLEs for the pseudo-problem and the problem share the same asymptotic distribution. We discuss these results in Section 5. The extensive proofs of the main results are found in Section 6 while the most technical results are pushed back into Section 7 at the end of this paper.

Notations. Whenever mentioned, the O and o notations will be used to designate a.s. O and a.s. o respectively, unless there are indexed with $\mathbb{P}$ as in $O_{\mathbb{P}}$ and $o_{\mathbb{P}}$, in which case they will designate O and o in probability respectively.

Hereafter we will use the notation A^c for the complement of the set A and $B(x, r)$ for the open ball of radius r centred at x i.e. $B(x, r) = \{x',\ \|x' - x\| < r\}$.

2 Bayesian consistency

In this Section, we show that the posterior distribution of θ given $(X_1, \ldots, X_n)$ asymptotically favours any neighbourhood of θ_0 as long as the prior distribution itself charges a (possibly different) neighbourhood of θ_0 (see Theorem 2.1). We then present in Theorem 2.2 the main result of this paper i.e. the convergence of posterior distribution with suitable normalisation to a Gaussian distribution.

2.1 Consistency and asymptotic normality of the posterior distribution

Theorem 2.1. *Let $\pi(\cdot)$ be a prior distribution on θ, continuous and positive on a neighbourhood of θ_0 and let U be a neighbourhood of θ_0, then under Assumptions (A1)–(A4), as $n \to +\infty$,*

$$\int_U \pi(\theta|X_{1:n})\,\mathrm{d}\theta \xrightarrow{a.s.} 1. \tag{2.1}$$

Proof for Theorem 2.1. The proof is very similar to the one given in Ghosh and Ramamoorthi (2003) for a model with i.i.d. observations. Let $\delta > 0$ small enough so that $B(\theta_0, \delta) \subset U$. Since

$$\begin{aligned}\int_U \pi(\theta|X_{1:n})\,\mathrm{d}\theta &= \frac{1}{1 + \dfrac{\int_{U^c} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}{\int_U \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}} \\ &\leqslant \frac{1}{1 + \dfrac{\int_{B^c(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}{\int_{B(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}}\end{aligned}$$

it will suffice to show that

$$\frac{\int_{B^c(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}{\int_{B(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta} \xrightarrow{a.s.} 0. \tag{2.2}$$

To prove (2.2) we adequately majorate its numerator and minorate its denominator. The majoration mainly relies on Proposition 7.11 while the minoration is derived without any major difficulties. The comprehensive proof of (2.2) can be found in Section 6.1 on page 11. □

Let $\theta \in \Theta$, we now define $I(\theta)$, the asymptotic Fisher Information matrix $I(\theta)$ of the model, as the symmetric matrix given by

$$I(\theta) = \begin{bmatrix} \sigma^{-2}\displaystyle\int_{\underline{u}}^{u} (t-u)^2\,\mathrm{d}F(t) & -\sigma^{-2}\gamma\displaystyle\int_{\underline{u}}^{u} (t-u)\,\mathrm{d}F(t) & 0 \\ & \sigma^{-2}\gamma^2\displaystyle\int_{\underline{u}}^{u} 1\,\mathrm{d}F(t) & 0 \\ & & \dfrac{1}{2}\sigma^{-4} \end{bmatrix}. \tag{2.3}$$

It is obviously positive and definite since all its principal minor determinants are positive. The proof of the fact that it is indeed the limiting matrix of the Fisher Information matrix of the model is deferred to Lemma 7.10.

Theorem 2.2. *Let $\pi(\cdot)$ be a prior distribution on θ, continuous and positive at θ_0, and let $k_0 \in \mathbb{N}$ such that*

$$\int_\Theta \|\theta\|^{k_0}\pi(\theta)\,\mathrm{d}\theta < +\infty,$$

and denote

$$t = n^{\frac{1}{2}}(\theta - \widehat{\theta}_n), \tag{2.4}$$

and $\widetilde{\pi}_n(\cdot|X_{1:n})$ the posterior density of t given $X_{1:n}$, then under Assumptions (A1)–(A4), for any $0 \leqslant k \leqslant k_0$, as $n \to +\infty$,

$$\int_{\mathbb{R}^3} \|t\|^k \left|\widetilde{\pi}_n(t|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}t' I(\theta_0)t}\right| \mathrm{d}t \xrightarrow{\mathbb{P}} 0, \tag{2.5}$$

where $I(\theta)$ is defined in (2.3) and θ_0 the true value of the parameter.

The proof Theorem 2.2 relies on the consistency of the pseudo-problem, first introduced in Sylwester (1965), that we define in the next few paragraphs.

2.2 Pseudo-problem

The major challenge in proving Theorem 2.2 is that the typical arguments usually used to derive the asymptotic behaviour of the posterior distribution (see Ghosh et al., 2006, for example) do not directly apply here. The proof provided by Ghosh et al. requires a Taylor expansion of the likelihood of the model up to the third order at the MLE, and the likelihood of the model we consider here at the n-th step is very obviously not continuously differentiable w.r.t. u in each observed temperature t_i, $i = 1, \dots, n$. Note that the problem only grows worse as the number of observations increases.

To overcome this difficulty we follow the original idea first introduced in Sylwester (1965), and later used again in Feder (1975): we introduce a pseudo-problem for which we are able to recover the classical results and show that the differences between the estimates for the problem and the pseudo-problem are, in a sense, negligeable. The pseudo-problem is obtained by deleting all the observations within intervals D_n of respective sizes d_n centred around u_0. The intervals D_n are defined as

$$D_n = \left]u_0 - \frac{d_n}{2},\, u_0 + \frac{d_n}{2}\right[,$$

and their sizes d_n are chosen such that as $n \to +\infty$

$$d_n \to 0, \qquad\qquad n^{-\frac{1}{2}}(\log n) \cdot d_n^{-1} \to 0. \tag{2.6}$$

This new problem is called pseudo-problem because the value of u_0 is unknown and we therefore cannot in practice delete these observations. Note that the actual choice of the sequence $(d_n)_{n\in\mathbb{N}}$ does not influence the rest of the results in any way, as long as it satisfies to conditions (2.6). It thus does not matter at all whether one chooses (for instance) $d_n = n^{-\frac{1}{4}}$ or $d_n = \log^{-1} n$.

Let us denote n^{**} the number of observations deleted from the original problem, and $n^* = n - n^{**}$ the sample size of the pseudo-problem. Generally speaking, quantities annotated with a single asterisk * will refer to the pseudo-problem. $l^*_{1:n}(X_{1:n}|\theta)$ will thus designate the likelihood of the pseudo-problem i.e. (reindexing observations whenever necessary)

$$l^*_{1:n}(X_{1:n}|\theta) = -\frac{n^*}{2}\log\left(2\pi\sigma^2\right) - \sum_{i=1}^{n^*}\frac{1}{2\sigma^2}\left(X_i - \gamma\cdot(t_i - u)\mathbb{1}_{[t_i,\,+\infty[}(u)\right)^2. \tag{2.7}$$

On one hand, from an asymptotic point of view, the removal of those n^{**} observations should not have any kind of impact on the distribution theory. The intuitive idea is that deleting n^{**} observations takes away only a fraction n^{**}/n of the information which asymptotically approaches zero as will be shown below. The first condition (2.6) seems only a natural requirement if we ever hope to prove that the MLE for the problem and the pseudo-problem behave asymptotically in a similar manner (we will show they do in Theorem 4.2, see equation (4.1)).

On the other hand, assuming the MLE is consistent (we will show it is, in Theorem 3.3) and assuming that the sizes d_n are carefully chosen so that the sequence $(\widehat{u}_n)_{n\in\mathbb{N}}$ falls into the designed sequence of intervals $(D_n)_{n\in\mathbb{N}}$ (see Proposition 4.1, whose proof the second condition (2.6) is tailored for), these regions will provide open neighbourhoods of the MLE over which the likelihood of the pseudo-problem will be differentiable. The pseudo-problem can therefore be thought of as a locally regularised version of the problem (locally because we are only interested in the differentiability of the likelihood over a neighbourhood of the MLE). We should thus be able to retrieve the usual results for the pseudo-problem with a bit of work. It will be shown that this is indeed the case (see Theorem 2.3).

If the sequence $(d_n)_{n\in\mathbb{N}}$ satisfies to conditions (2.6), then as $n \to +\infty$,

$$\frac{n^{**}}{n} \to 0, \qquad\qquad \frac{n^*}{n} \to 1.$$

Using the uniform convergence of F_n to F over any compact subset (see Assumption (A1), and its Remark 1), we indeed find via a Taylor-Lagrange approximation

$$\begin{aligned}\frac{n^{**}}{n} &= F_n\left(u_0 + \frac{d_n}{2}\right) - F_n\left(u_0 - \frac{d_n}{2}\right) \\ &= F\left(u_0 + \frac{d_n}{2}\right) - F\left(u_0 - \frac{d_n}{2}\right) + \mathrm{o}(1) \\ &= d_n \cdot f(u_n) + \mathrm{o}(1),\end{aligned}$$

where $u_n \in D_n$, so that in the end, since $u_n \to u_0$ and f is continuous and positive at u_0, we have a.s.

$$\frac{n^{**}}{n} = d_n \cdot (f(u_0) + \mathrm{o}(1)) + \mathrm{o}(1) \to 0.$$

We now recover the asymptotic normality of the posterior distribution for the pseudo problem.

Theorem 2.3. *Let $\pi(\cdot)$ be a prior distribution on θ, continuous and positive at θ_0, and let $k_0 \in \mathbb{N}$ such that*

$$\int_\Theta \|\theta\|^{k_0} \pi(\theta)\,\mathrm{d}\theta < +\infty.$$

and denote

$$t^* = n^{\frac{1}{2}}(\theta - \widehat{\theta}_n^*), \tag{2.8}$$

and $\widetilde{\pi}_n^*(\cdot|X_{1:n})$ the posterior density of t^* given $X_{1:n}$, then under Assumptions (A1)–(A4) and conditions (2.6), for any $0 \leqslant k \leqslant k_0$, as $n \to +\infty$,

$$\int_{\mathbb{R}^3} \|t\|^k \left|\widetilde{\pi}_n^*(t|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}t' I(\theta_0)t}\right| \mathrm{d}t \xrightarrow{a.s.} 0, \tag{2.9}$$

where $I(\theta)$ is defined in (2.3).

Proof of Theorem 2.3. The extensive proof, to be found in Section 6.1, was inspired by that of Theorem 4.2 in Ghosh et al. (2006) which deals with the case where the observations $X_1, \ldots, X_n$ are independent and identically distributed and where the (univariate) log-likelihood is differentiable in a fixed small neighbourhood of θ_0. We tweaked the original proof of Ghosh et al. so that we could deal with independent but not identically distributed observations and a (multivariate) log-likelihood that is guaranteed differentiable only on a decreasing small neighbourhood of θ_0. □

2.3 From the pseudo-problem to the original problem

We now give a short proof of Theorem 2.2. As we previously announced, it relies upon its counterpart for the pseudo-problem, i.e. Theorem 2.3.

Proof of Theorem 2.2. Recalling the definition of t and t^* given in (2.4) and (2.8) we observe that

$$t = t^* + n^{\frac{1}{2}}(\widehat{\theta}_n^* - \widehat{\theta}_n).$$

Thus the posterior distribution of t^* and that of t, given $X_{1:n}$ are linked together via

$$\widetilde{\pi}_n(t|X_{1:n}) = \widetilde{\pi}_n^*(t - \alpha_n|X_{1:n}) \tag{2.10}$$

where

$$\alpha_n = n^{\frac{1}{2}}(\widehat{\theta}_n^* - \widehat{\theta}_n).$$

Relationship (2.10) allows us to write

$$\begin{aligned}
&\int_{\mathbb{R}^3} \|t\|^k \left|\widetilde{\pi}_n(t|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}t'I(\theta_0)t}\right| \mathrm{d}t \\
&= \int_{\mathbb{R}^3} \|t\|^k \left|\widetilde{\pi}^*_n(t-\alpha_n|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}t'I(\theta_0)t}\right| \mathrm{d}t \\
&= \int_{\mathbb{R}^3} \|t+\alpha_n\|^k \left|\widetilde{\pi}^*_n(t|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}(t+\alpha_n)'I(\theta_0)(t+\alpha_n)}\right| \mathrm{d}t \\
&\leqslant \int_{\mathbb{R}^3} \|t+\alpha_n\|^k \left|\widetilde{\pi}^*_n(t|X_{1:n}) - (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2}t'I(\theta_0)t}\right| \mathrm{d}t \\
&\quad + (2\pi)^{-\frac{3}{2}}|I(\theta_0)|^{\frac{1}{2}} \int_{\mathbb{R}^3} \|t+\alpha_n\|^k \left|e^{-\frac{1}{2}(t+\alpha_n)'I(\theta_0)(t+\alpha_n)} - e^{-\frac{1}{2}t'I(\theta_0)t}\right| \mathrm{d}t
\end{aligned}$$

Theorem 2.3 ensures that the first integral on the right hand side of this last inequality goes to zero in probability. It therefore suffices to show that the second integral goes to zero in probability to end the proof, i.e. that as $n \to +\infty$

$$\int_{\mathbb{R}^3} \|t+\alpha_n\|^k \left|e^{-\frac{1}{2}(t+\alpha_n)'I(\theta_0)(t+\alpha_n)} - e^{-\frac{1}{2}t'I(\theta_0)t}\right| \mathrm{d}t \xrightarrow{\mathbb{P}} 0. \tag{2.11}$$

But the proof of (2.11) is straightforward knowing that $\alpha_n \xrightarrow{\mathbb{P}} 0$ (see (4.1)) and using dominated convergence. □

As an immediate consequence of Theorem 2.2 we want to mention the weak consistency of the Bayes estimator.

Corollary 2.4. *Let $\pi(\cdot)$ a prior distribution on θ, continuous and positive at θ_0, such that*

$$\int_\Theta \|\theta\|\pi(\theta)\,\mathrm{d}\theta < +\infty,$$

and denote

$$\widetilde{\theta}_n = \int_\Theta \theta\pi_n(\theta|X_{1:n})\,\mathrm{d}\theta,$$

the Bayes estimator of θ in the problem. Then under Assumptions (A1)–(A4), as $n \to +\infty$,

$$n^{\frac{1}{2}}(\widetilde{\theta}_n - \widehat{\theta}_n) \xrightarrow{\mathbb{P}} 0.$$

Proof of Corollary 2.4. By definition,

$$\widetilde{\theta}_n = \int_\Theta \theta\pi_n(\theta|X_{1:n})\,\mathrm{d}\theta$$

and this allows us to write

$$\begin{aligned}
n^{\frac{1}{2}}(\widetilde{\theta}_n - \widehat{\theta}_n) &= \int_\Theta n^{\frac{1}{2}}(\theta - \widehat{\theta}_n)\pi_n(\theta|X_{1:n})\,\mathrm{d}\theta \\
&= \int_{\mathbb{R}^3} t\widetilde{\pi}_n(t|X_{1:n})\,\mathrm{d}t \xrightarrow{\mathbb{P}} 0,
\end{aligned}$$

the last convergence being a direct consequence of Theorem 2.2 with $k_0 = 1$. □

Observe that, under conditions (2.6), the same arguments naturally apply to the pseudo-problem and lead to a strong consistency (a.s. convergence) of its associated Bayes estimator due to Theorem 2.3, thus recovering the results of Ghosh et al. (2006) for the regularised version of the problem.

3 Strong consistency of the MLE

In this Section we prove the strong consistency of the MLE over any compact set including the true parameter (see Theorem 3.1). It is a prerequisite for a more accurate version of the strong consistency (see Theorem 3.3) which lies at the heart of the proof of Theorem 2.3.

Theorem 3.1. *Under Assumptions (A1)–(A4), we have a.s., as* $n \to +\infty$,

$$\|\widehat{\theta}_n - \theta_0\| = \mathrm{o}(1).$$

Proof of Theorem 3.1. Recall that K is a compact subset of Θ, such that $\widehat{\theta}_n \in K$ for any n large enough. We denote

$$l_{1:n}(X_{1:n}|S) = \sup_{\theta \in S} l_{1:n}(X_{1:n}|\theta), \text{ for any } S \subset K,$$
$$K_n(a) = \{\theta \in \Theta,\ l_{1:n}(X_{1:n}|\theta) \geqslant \log a + l_{1:n}(X_{1:n}|K)\}, \text{ for any } a \in]0,\ 1[.$$

All we need to prove is that

$$\exists a \in]0,\ 1[,\ \mathbb{P}\left(\lim_{n \to +\infty} \sup_{\theta \in K_n(a)} \|\theta - \theta_0\| = 0\right) = 1. \tag{3.1}$$

since for any n large enough we have $\widehat{\theta}_n \in K_n(a)$ for any $a \in]0,\ 1[$. We control the likelihood upon the complement of a small ball in K and prove the contrapositive of (3.1) using compacity arguments. The extensive proof of (3.1) is to be found in Section 6.2 . □

We strengthen the result of Theorem 3.1 by giving a rate of convergence for the MLE (see Theorem 3.3). This requires a rate of convergence for the image of the MLE through the regression function of the model, that we give in the Proposition 3.2 below.

Proposition 3.2. *Under Assumptions (A1)–(A4), as* $n \to +\infty$, *a.s., for any open interval* $I \subset [\underline{u},\ \overline{u}]$,

$$\min_{t_i \in I,\ i \leqslant n} |\mu(\widehat{\eta}_n, t_i) - \mu(\eta_0, t_i)| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right).$$

Proof of Proposition 3.2. The proof is given in Section 6.2. □

Theorem 3.3. *Under Assumptions (A1)–(A4), we have a.s., as* $n \to +\infty$,

$$\|\widehat{\theta}_n - \theta_0\| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right). \tag{3.2}$$

Proof of Theorem 3.3. We show that a.s. (3.2) holds for each coordinate of $\widehat{\theta}_n - \theta_0$. The calculations for the variance σ^2 are pushed back into Section 6.2. We now prove the result for the parameters γ and u. It is more convenient to use a reparametrisation of the model in terms of slope γ and intercept β where $\beta = -\gamma u$.

Slope γ and intercept β. Let V_1 and V_2 be two non empty open intervals of $]\underline{u},\ u_0[$ such that their closures $\overline{V_1}$ and $\overline{V_2}$ do not overlap. For any $(t_1, t_2) \in V_1 \times V_2$, define $M(t_1, t_2)$ the obviously invertible matrix

$$M(t_1, t_2) = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \end{bmatrix},$$

and observe that for any $\tau = (\beta, \gamma)$,

$$M(t_1, t_2)\tau = \begin{bmatrix} \mu(\eta, t_1) \\ \mu(\eta, t_2) \end{bmatrix}.$$

Observe that by some basic linear algebra tricks we are able to write for any $(t_1, t_2) \in V_1 \times V_2$

$$\begin{aligned}\|\widehat{\tau}_n - \tau_0\|_\infty &= \|M(t_1,t_2)^{-1} M(t_1,t_2)(\widehat{\tau}_n - \tau_0)\|_\infty \\ &\leqslant \||M(t_1,t_2)^{-1}\|\|_\infty \cdot \|M(t_1,t_2)\widehat{\tau}_n - M(t_1,t_2)\tau_0\|_\infty \\ &\leqslant \frac{|t_2| + |t_1| + 2}{|t_2 - t_1|} \cdot \|M(t_1,t_2)\widehat{\tau}_n - M(t_1,t_2)\tau_0\|_\infty.\end{aligned}$$

Thus, using the equivalence of norms and a simple domination of the first term of the product in the inequality above, we find that there exists a constant $C \in \mathbb{R}_+^*$, such that for any $(t_1, t_2) \in V_1 \times V_2$

$$\|\widehat{\tau}_n - \tau_0\| \leqslant C \cdot \|M(t_1,t_2)\widehat{\tau}_n - M(t_1,t_2)\tau_0\|,$$

i.e.

$$\|\widehat{\tau}_n - \tau_0\| \leqslant C \cdot \left[\sum_{i=1}^{2} (\mu(\widehat{\eta}_n, t_i) - \mu(\eta_0, t_i))^2\right]^{\frac{1}{2}}. \tag{3.3}$$

Taking advantage of Proposition 3.2, we are able to exhibit two sequences of points $(t_{1,n})_{n\in\mathbb{N}}$ in V_1 and $(t_{2,n})_{n\in\mathbb{N}}$ in V_2 such that a.s., for $i = 1, 2$

$$|\mu(\widehat{\eta}_n, t_{i,n}) - \mu(\eta_0, t_{i,n})| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right). \tag{3.4}$$

Combining (3.3) and (3.4) together (using $t_i = t_{i,n}$ for every n), it is now trivial to see that a.s.

$$\|\widehat{\tau}_n - \tau_0\| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right),$$

which immediately implies the result for the γ and β components of θ.

Break-point u. Recalling that $u = -\beta\gamma^{-1}$ and thanks to the result we just proved, we find that a.s.

$$\begin{aligned}\widehat{u}_n = -\widehat{\beta}_n\widehat{\gamma}_n^{-1} &= -\left[\beta_0 + \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right)\right]\left[\gamma_0 + \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right)\right]^{-1} \\ &= -\beta_0\gamma_0^{-1} + \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right) = u_0 + \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right).\end{aligned}$$

□

4 Asymptotic distribution of the MLE

In this Section we derive the asymptotic distribution of the MLE for the pseudo-problem (see Proposition 4.1) and then show that the MLE of pseudo-problem and that of the problem share the same asymptotic distribution (see Theorem 4.2).

Proposition 4.1. *Under Assumptions (A1)–(A4) and conditions (2.6), as $n \to +\infty$*

$$n^{\frac{1}{2}}\left(\widehat{\theta}_n^* - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, I(\theta_0)^{-1}\right),$$

where the asymptotic Fisher Information Matrix $I(\cdot)$ is defined in (2.3).

Proof of Theorem 4.1. The proof is divided in two steps. We first show that the likelihood of the pseudo-problem is a.s. differentiable in a neighbourhood of the MLE $\widehat{\theta}_n^*$ for N large enough. We then recover the asymptotic distribution of the MLE following the usual scheme of proof, with a Taylor expansion of the likelihood of the pseudo-problem around the true parameter. The details of these two steps are given in Section 6.3. □

Theorem 4.2. *Under Assumptions (A1)–(A4) and conditions (2.6), as $n \to +\infty$,*

$$n^{\frac{1}{2}} \left(\widehat{\theta}_n - \theta_0 \right) \xrightarrow{d} \mathcal{N}\left(0, I(\theta_0)^{-1}\right),$$

where the asymptotic Fisher Information Matrix $I(\cdot)$ is defined in (2.3).

Proof of Theorem 4.2. It is a direct consequence of Proposition 4.1 as soon as we show that as $n \to +\infty$

$$\widehat{\theta}_n - \widehat{\theta}_n^* = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right). \tag{4.1}$$

To prove (4.1), we study each coordinate separately. For γ and u, we apply Lemmas 4.12 and 4.16 found in Feder (1975) with a slight modification: the rate of convergence d_n he uses may differ from ours but it suffices to formally replace $(\log\log n)^{\frac{1}{2}}$ by $(\log n)$ all throughout his paper and the proofs he provides go through without any other change. We thus get

$$\widehat{\gamma}_n - \widehat{\gamma}_n^* = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right), \qquad \widehat{u}_n - \widehat{u}_n^* = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right). \tag{4.2}$$

It now remains to show that

$$\widehat{\sigma}_n^2 - \widehat{\sigma}_n^{2*} = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right). \tag{4.3}$$

To do so, we use (4.2) and the decomposition (6.40)

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n) + \frac{2}{n}\sum_{i=1}^{n} \nu_i(\widehat{\eta}_n)\xi_i + \frac{1}{n}\sum_{i=1}^{n} \xi_i^2,$$

where $\nu_i(\widehat{\eta}_n) = \gamma_0 \cdot (t_i - u_0)\mathbb{1}_{[t_i, +\infty[}(u_0) - \widehat{\gamma}_n \cdot (t_i - \widehat{u}_n)\mathbb{1}_{[t_i, +\infty[}(\widehat{u}_n)$. The details of this are available in Section 6.3. □

5 Discussion

In this Section, we summarise the results presented in this paper. The consistency of the posterior distribution for a piecewise linear regression model is derived as well as its asymptotic normality with suitable normalisation. The proofs of these convergence results rely on the convergence of the MLE which is also proved here. In order to obtain all the asymptotic results, a regularised version of the problem at hand, called pseudo-problem, is first studied and the difference between this pseudo-problem and the (full) problem is then shown to be asymptotically negligeable.

The trick of deleting observations in a diminishing neighbourhood of the true parameter, originally found in Sylwester (1965) allows the likelihood of the pseudo-problem to be differentiated at the MLE, once the MLE is shown to asymptotically belong to that neighbourhood (this requires at least a small control of the rate of convergence of the MLE). This is the key argument needed to derive the asymptotic distribution of the MLE through the usual Taylor expansion of the likelihood at the MLE. Extending the results of Ghosh et al. (2006) to a non i.i.d. setup, the asymptotic normality of the posterior distribution for the pseudo-problem is then recovered from that of the MLE, and passes on almost naturally to the (full) problem.

The asymptotic normality of the MLE and the posterior distribution are proved in this paper in a non i.i.d. setup with a non continuously differentiable likelihood. In both cases we obtain the same asymptotic results as for an i.i.d. regular model: the rate of convergence is $\sqrt{n}$ and the limiting distribution is Gaussian (see Ghosh et al., 2006; Lehmann, 2004). For the piecewise linear regression model, the exogenous variable $t_{1:n}$ does not appear in the expression of the rate of convergence as opposed to what is known for the usual linear regression model (see Lehmann, 2004): this is due to our own Assumption (A1) which implies that $t'_{1:n}t_{1:n}$ is equivalent to n. Note that for a simple linear regression model, we also obtain the rate $\sqrt{n}$ under

Assumption (A1). In the litterature, several papers already highlighted the fact that the rate of convergence and the limiting distribution (when it exists) may be different for non regular models in the sense that the likelihood is either non continuous, or non continuously differentiable, or admits singularities (see Dacunha-Castelle, 1978; Ghosh et al., 1994; Ghosal and Samanta, 1995; Ibragimov and Has'minskii, 1981). For the piecewise regression model, the likelihood is continuous but non continuously differentiable on a countable set (but the left and right derivatives exist and are finite): the rate of convergence $\sqrt{n}$ is not so surprising in our case, because this rate was already obtained for a univariate i.i.d. model the likelihood of has the same non regularity at a single point. In that case, the rate of convergence of the MLE is shown to be n (see Dacunha-Castelle, 1978, for instance).

6 Extensive proofs

6.1 Proofs of Section 2

Proof of Theorem 2.1. To prove (2.2), we proceed as announced and deal with numerator and denominator in turn.

MAJORATION. From Proposition 7.11 with $\rho_n = 1$, for any given $\epsilon > 0$, we can choose $\delta > 0$ small enough so that a.s. for any n large enough

$$\sup_{\theta \in B^c(\theta_0,\delta)} \frac{1}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] \leqslant -\epsilon.$$

We thus obtain a.s. for any n large enough

$$\begin{aligned} 0 &\leqslant \int_{B^c(\theta_0,\delta)} \pi(\theta) \exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] \, \mathrm{d}\theta \\ &\leqslant e^{-n\epsilon} \int_{B^c(\theta_0,\delta)} \pi(\theta) \, \mathrm{d}\theta. \end{aligned} \tag{6.1}$$

MINORATION. Define $\theta_n \in \overline{B(\theta_0,\delta)}$ such that

$$\inf_{\theta \in B(\theta_0,\delta)} \frac{1}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] = \frac{1}{n}[l_{1:n}(X_{1:n}|\theta_n) - l_{1:n}(X_{1:n}|\theta_0)]$$

It is possible to define such a θ_n because $\overline{B(\theta_0,\delta)}$ is a compact subset of Θ for $\delta > 0$ small enough and $l_{1:n}(X_{1:n}|\cdot)$ is continuous as a function of θ. Let now

$$b_n(\theta) = \left(\frac{\sigma_0^2}{\sigma^2} - 1 - \log\frac{\sigma_0^2}{\sigma^2}\right) + \frac{1}{\sigma^2} \cdot \frac{1}{n} \sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2. \tag{6.2}$$

Recalling the definition of the log-likehood given in (1.2) and replacing X_i by its expression given

in (1.1) we find via straightforward algebra

$$\begin{aligned}
\frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] &= \log\frac{\sigma_0^2}{\sigma^2} + \left(\frac{1}{\sigma^2} - \frac{1}{\sigma_0^2}\right)\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2\right) \\
&\quad - \frac{1}{n\sigma^2}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]^2 - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]\xi_i \\
&= \log\frac{\sigma_0^2}{\sigma^2} + \left(\frac{1}{\sigma^2} - \frac{1}{\sigma_0^2}\right)\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2 + \sigma_0^2\right) \\
&\quad - \frac{1}{n\sigma^2}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]^2 - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]\xi_i \\
&= \left(\log\frac{\sigma_0^2}{\sigma^2} + 1 - \frac{\sigma_0^2}{\sigma^2}\right) + \frac{\sigma_0^2 - \sigma^2}{\sigma^2\sigma_0^2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right) \\
&\quad - \frac{1}{n\sigma^2}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]^2 - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]\xi_i \qquad (6.3) \\
&= -b_n(\theta) + \frac{\sigma_0^2 - \sigma^2}{\sigma^2\sigma_0^2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right) - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta,t_i)]\xi_i. \qquad (6.4)
\end{aligned}$$

It is now easy to see that

$$\begin{aligned}
\inf_{\theta\in B(\theta_0,\delta)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] &= \frac{2}{n}[l_{1:n}(X_{1:n}|\theta_n) - l_{1:n}(X_{1:n}|\theta_0)] \\
&= -b_n(\theta_n) + \frac{\sigma_0^2 - \sigma_n^2}{\sigma_n^2\sigma_0^2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right) - \frac{2}{\sigma_n^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta_n,t_i)]\xi_i \\
&= -b_n(\theta_n) + \frac{1}{\sigma_n^2}\left[\frac{\sigma_0^2 - \sigma_n^2}{\sigma_0^2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right) - \frac{2}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta_n,t_i)]\xi_i\right] \\
&= -b_n(\theta_n) + \frac{1}{\sigma_n^2}R_n \\
&= \left(\log\frac{\sigma_0^2}{\sigma_n^2} + 1 - \frac{\sigma_0^2}{\sigma_n^2}\right) - \frac{1}{\sigma_n^2}\cdot\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0,t_i) - \mu(\eta_n,t_i)]^2 + \frac{1}{\sigma_n^2}R_n
\end{aligned}$$

where $R_n \xrightarrow{a.s.} 0$ because of the Law of Large Numbers and Lemma 7.6. Thanks to Lemma 7.3 we thus find that there exists $C \in \mathbb{R}_+^*$ such that

$$\begin{aligned}
\inf_{\theta\in B(\theta_0,\delta)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] &\geqslant \left(\log\frac{\sigma_0^2}{\sigma_n^2} + 1 - \frac{\sigma_0^2}{\sigma_n^2}\right) \\
&\quad - \frac{1}{\sigma_n^2}\left(C\|\theta_n - \theta_0\|^2 - R_n\right)
\end{aligned}$$

We now choose $\kappa > 0$ and $\delta > 0$ small enough so that

$$\sigma_n^2\left(\log\frac{\sigma_0^2}{\sigma_n^2} + 1 - \frac{\sigma_0^2}{\sigma_n^2}\right) \geqslant -\kappa, \qquad (6.5)$$

$$-\frac{3(\kappa + C\delta^2)}{2(\sigma_0^2 - \delta)} \geqslant -\frac{1}{2}\epsilon. \qquad (6.6)$$

Thanks to (6.5) and the definition of θ_n, we can now write that

$$\begin{aligned}\inf_{\theta\in B(\theta_0,\delta)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] &\geqslant -\frac{1}{\sigma_n^2}\left(\kappa + C\|\theta_n - \theta_0\|^2 - R_n\right)\\ &\geqslant -\frac{1}{\sigma_n^2}\left(\kappa + C\delta^2 - R_n\right).\end{aligned}$$

Since for any n large enough

$$|R_n| \leqslant \frac{1}{2}\left(\kappa + C\delta^2\right),$$

we find via (6.6) that for any n large enough

$$\begin{aligned}\inf_{\theta\in B(\theta_0,\delta)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] &\geqslant -\frac{3}{2\sigma_n^2}\left(\kappa + C\delta^2\right)\\ &\geqslant -\frac{3(\kappa + C\delta^2)}{2(\sigma_0^2 - \delta)} \geqslant -\frac{1}{2}\epsilon.\end{aligned}$$

We just proved that for any $\epsilon > 0$, we have a.s. for any n large enough

$$0 \geqslant \inf_{\theta\in B(\theta_0,\delta)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] \geqslant -\frac{1}{2}\epsilon,$$

which immediately implies

$$\int_{B(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta \geqslant e^{-\frac{1}{2}n\epsilon}\int_{B(\theta_0,\delta)} \pi(\theta)\,\mathrm{d}\theta. \tag{6.7}$$

Conclusion. Let now $\epsilon > 0$ and $\delta > 0$ small enough so that a.s. for any n large enough (6.1) and (6.7) both hold. We have a.s. for any n large enough

$$\frac{\int_{B^c(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta}{\int_{B(\theta_0,\delta)} \pi(\theta)\exp[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)]\,\mathrm{d}\theta} \leqslant \frac{\int_{B^c(\theta_0,\delta)} \pi(\theta)\,\mathrm{d}\theta}{\int_{B(\theta_0,\delta)} \pi(\theta)\,\mathrm{d}\theta} e^{-\frac{1}{2}n\epsilon} \to 0,$$

which ends the proof. □

Proof of Theorem 2.3. Because the posterior distribution of θ in the pseudo-problem, $\pi_n^*(\cdot|X_{1:n})$, can be written as

$$\pi_n^*(\theta|X_{1:n}) \propto \pi(\theta)\exp[l_{1:n}^*(X_{1:n}|\theta)],$$

the posterior density of $t^* = n^{\frac{1}{2}}(\theta - \widehat{\theta}_n^*) \in \mathbb{R}^3$ can be written as

$$\widetilde{\pi}_n^*(t|X_{1:n}) = C_n^{-1}\pi(\widehat{\theta}_n^* + n^{-\frac{1}{2}}t)\exp[l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^* + n^{-\frac{1}{2}}t) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)]$$

where

$$C_n = \int_{\mathbb{R}^3} \pi(\widehat{\theta}_n^* + n^{-\frac{1}{2}}t)\exp[l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^* + n^{-\frac{1}{2}}t) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)]\,\mathrm{d}t. \tag{6.8}$$

Denoting

$$\begin{aligned}g_n(t) = \pi(\widehat{\theta}_n^* + n^{-\frac{1}{2}}t)\exp[l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^* + n^{*-\frac{1}{2}}t) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)]\\ - \pi(\theta_0)e^{-\frac{1}{2}t'I(\theta_0)t},\end{aligned} \tag{6.9}$$

to prove (2.9) it suffices to show that for any $0 \leqslant k \leqslant k_0$,

$$\int_{\mathbb{R}^3} \|t\|^k |g_n(t)| \, \mathrm{d}t \xrightarrow{a.s.} 0. \tag{6.10}$$

Indeed, if (6.10) holds, $C_n \xrightarrow{a.s.} \pi(\theta_0)(2\pi)^{\frac{3}{2}}|I(\theta_0)|^{-\frac{1}{2}}$ $(k=0)$ and therefore, the integral in (2.9) which is dominated by

$$\begin{aligned} &C_n^{-1} \int_{\mathbb{R}^3} \|t\|^k |g_n(t)| \, \mathrm{d}t \\ &\quad + \int_{\mathbb{R}^3} \|t\|^k \left| C_n^{-1} \pi(\theta_0) e^{-\frac{1}{2} t' I(\theta_0) t} - (2\pi)^{-\frac{1}{2}} |I(\theta_0)|^{\frac{1}{2}} e^{-\frac{1}{2} t' I(\theta_0) t} \right| \mathrm{d}t \end{aligned}$$

also goes to zero a.s.

Let $0 < \delta$ to be chosen later, and let $0 \leqslant k \leqslant k_0$. To show (6.10), we break $\mathbb{R}^3$ into two regions

$$\begin{aligned} T_1(\delta) &= B^c(0, \delta n^{\frac{1}{2}} d_n) = \{t : \|t\| \geqslant \delta n^{\frac{1}{2}} d_n\} \\ T_2(\delta) &= B(0, \delta n^{\frac{1}{2}} d_n) = \{t : \|t\| < \delta n^{\frac{1}{2}} d_n\} \end{aligned}$$

and show that for $i = 1, 2$

$$\int_{T_i(\delta)} \|t\|^k |g_n(t)| \, \mathrm{d}t \xrightarrow{a.s.} 0. \tag{6.11}$$

Proof for $i = 1$. Note that $\int_{T_1(\delta)} \|t\|^k |g_n(t)|$ is dominated by

$$\begin{aligned} &\int_{T_1(\delta)} \|t\|^k \pi(\widehat{\theta}_n^* + n^{\frac{1}{2}} t) \exp[l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^* + n^{-\frac{1}{2}} t) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)] \, \mathrm{d}t \\ &\quad + \int_{T_1(\delta)} \|t\|^k \pi(\theta_0) e^{-\frac{1}{2} t' I(\theta_0) t} \, \mathrm{d}t. \end{aligned}$$

The second integral trivially goes to zero. For the first integral, we observe that it can be rewritten as

$$n^{\frac{1}{2}} \int_{B^c(\widehat{\theta}_n^*, \delta d_n)} n^{\frac{k}{2}} \|\theta - \widehat{\theta}_n^*\|^k \pi(\theta) \exp[l_{1:n}^*(X_{1:n}|\theta) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)] \, \mathrm{d}\theta.$$

The strong consistency of $\widehat{\theta}_n^*$ (see Theorem 3.3) implies that a.s., for any n large enough

$$\|\widehat{\theta}_n^* - \theta_0\| < \frac{1}{2} \delta d_n.$$

From this, we deduce that a.s., for any n large enough, $B^c(\widehat{\theta}_n^*, \delta d_n) \subset B^c(\theta_0, \frac{1}{2}\delta d_n)$ and thus that the first integral is dominated by

$$n^{\frac{k+1}{2}} \int_{B^c(\theta_0, \frac{1}{2}\delta d_n)} \|\theta - \widehat{\theta}_n^*\|^k \pi(\theta) \exp[l_{1:n}^*(X_{1:n}|\theta) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*)] \, \mathrm{d}\theta.$$

Recalling that $n^* \sim n$, Proposition 7.11 with $\rho_n = d_n$ implies that there a.s. exists $\epsilon > 0$ such that for any n large enough and any $\theta \in B^c(\theta_0, \frac{1}{2}\delta d_n)$ we have

$$l_{1:n}^*(X_{1:n}|\theta) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*) \leqslant -\epsilon n d_n^2.$$

It follows, using (2.6) that, a.s. for any n large enough the first integral is dominated by

$$\begin{aligned} n^{\frac{k+1}{2}} \exp(-\epsilon n d_n^2) \int_\Theta \|\theta - \widehat{\theta}_n^*\|^k \pi(\theta) \, \mathrm{d}t &= n^{\frac{k+1}{2}} \exp(-\epsilon n d_n^2) \cdot \mathrm{O}(1) \\ &\leqslant n^{\frac{k+1}{2}} n^{-\epsilon \log n} \cdot \mathrm{O}(1) \to 0, \end{aligned}$$

since by (2.6) we find that $nd_n^2 \geqslant (\log n)^2$ for any n large enough. Hence (6.11) holds for $i = 1$.

PROOF FOR $i = 2$. We first recall the multivariate Taylor expansion for a function g (k+1)-times continuously differentiable within a neighbourhood of $y \in \mathbb{R}^n$. With the usual differential calculus notations

$$\mathrm{D}^\alpha g(y) \cdot h^{(\alpha)} = \sum_{1 \leqslant i_1, \ldots, i_\alpha \leqslant n} \frac{\partial^\alpha g}{\partial_{i_1} \cdots \partial_{i_\alpha}}(y) \cdot h_{i_1} \cdots h_{i_\alpha}$$

we have

$$g(x) = \sum_{\alpha=0}^{k} \frac{1}{\alpha!} \mathrm{D}^\alpha g(y) \cdot (x-y)^{(\alpha)} + R_{k+1}(x) \tag{6.12}$$

where

$$R_{k+1}(x) = \frac{1}{(k+1)!} \int_0^1 (1-s)^k \mathrm{D}^{k+1} g(y + s(x-y)) \cdot (x-y)^{(k+1)} \, \mathrm{d}s. \tag{6.13}$$

Before expanding the log-likelihood over $T_2(\delta)$ in a such a way, we first have to make sure it is differentiable over the correct domain. Indeed, the strong consistency of $\widehat{\theta}_n^*$ (see Theorem 3.3) implies that a.s., whatever $\delta_0 > 0$, for n large enough,

$$\|\widehat{\theta}_n^* - \theta_0\| < \delta_0 d_n.$$

For δ chosen small enough, since $t \in T_2(\delta)$ implies

$$\|\theta - \widehat{\theta}_n^*\| < \delta d_n$$

it follows from the triangle inequality that a.s. for n large enough,

$$\|\theta - \theta_0\| < (\delta_0 + \delta) d_n < d_n.$$

A.s. for any n large enough, $t \in T_2(\delta)$ hence implies $\theta \in B(\theta_0, (\delta + \delta_0) d_n)$. We choose δ_0 and δ small enough so that $\delta + \delta_0 < 1$. This way, $\theta \mapsto l_{1:n}^*(X_{1:n}|\theta)$ is guaranteed to be infinitely continuously differentiable over $B(\theta_0, (\delta + \delta_0) d_n) \subset B(\theta_0, d_n)$.

Now expanding the log-likelihood in a Taylor series for any n large enough, and taking advantage of the fact that $l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*) = 0$, we define $B_{1:n}^*(\cdot)$ the symmetric matrix defined for $u \in D_n$ by

$$B_{1:n}^*(\theta) = -\begin{bmatrix} \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial\gamma\partial\gamma} & \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial\gamma\partial u} & \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial\gamma\partial\sigma^2} \\ & \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial u\partial u} & \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial u\partial\sigma^2} \\ & & \dfrac{\partial^2 l_{1:n}^*(X_{1:n}|\theta)}{\partial\sigma^2\partial\sigma^2} \end{bmatrix}. \tag{6.14}$$

and write that

$$\begin{aligned} l_{1:n}^*(X_{1:n}|\theta) - l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^*) = -\frac{1}{2}(\theta - \widehat{\theta}_n^*)' \left(B_{1:n}^*(\widehat{\theta}_n^*)\right)(\theta - \widehat{\theta}_n^*) \\ + R_{3,n}(\theta) \end{aligned} \tag{6.15}$$

where

$$R_{3,n}(\theta) = \frac{1}{3!} \int_0^1 (1-s)^2 \mathrm{D}^3 l_{1:n}^*(X_{1:n}|\widehat{\theta}_n^* + s(\theta - \widehat{\theta}_n^*)) \cdot (\theta - \widehat{\theta}_n^*)^{(3)} \, \mathrm{d}s. \tag{6.16}$$

Lemma 7.12 allows us to write that a.s. there exists a constant $C \in R_+^*$ such that for any n large enough, for any $t \in T_2(\delta)$

$$l^*_{1:n}(X_{1:n}|\widehat{\theta}^*_n + n^{-\frac{1}{2}}t) - l^*_{1:n}(X_{1:n}|\widehat{\theta}^*_n) = -\frac{1}{2}t'\left(n^{-1}B^*_{1:n}(\widehat{\theta}^*_n)\right)t + S_n(t) \tag{6.17}$$

where

$$|S_n(t)| \leqslant Cn^{-\frac{1}{2}} \cdot \|t\|^3. \tag{6.18}$$

From (6.18), we obtain that for any $t \in T_2(\delta)$, $S_n(t) \xrightarrow{a.s.} 0$. Because of Lemma 7.10, we have $n^{-1}B^*_{1:n}(\widehat{\theta}^*_n) \xrightarrow{a.s.} I(\theta_0)$, and it follows immediately that for any $t \in T_2(\delta)$,

$$g_n(t) \xrightarrow{a.s.} 0,$$

and thus that

$$\|t\|^k g_n(t) \xrightarrow{a.s.} 0.$$

From (6.18) we also obtain

$$|S_n(t)| \leqslant C\delta d_n \|t\|^2.$$

Lemma 7.10, combined with (2.6), (6.17) and the positivity of $I(\theta_0)$, ensures that a.s. for any n large enough

$$|S_n(t)| \leqslant \frac{1}{4}t'\left(n^{-1}B^*_{1:n}(\widehat{\theta}^*_n)\right)t,$$

so that from (6.17), a.s. for any n large enough

$$\exp[l^*_{1:n}(X_{1:n}|\widehat{\theta}^*_n + n^{-\frac{1}{2}}t) - l^*_{1:n}(X_{1:n}|\widehat{\theta}^*_n)] \leqslant e^{-\frac{1}{4}t'\left(n^{-1}B^*_{1:n}(\widehat{\theta}^*_n)\right)t} \leqslant e^{-\frac{1}{8}t'I(\theta_0)t}. \tag{6.19}$$

Therefore, for n large enough, $\|t\|^k|g_n(t)|$ is dominated by an integrable function on the set $T_2(\delta)$ and (6.11) holds for $i = 2$ which completes the proof. □

6.2 Proofs of Section 3

Proof of Theorem 3.1. From (6.4), it is easy to see that

$$\begin{aligned}\frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|K)] &\leqslant \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] \\ &\leqslant -b_n(\theta) + \frac{\sigma_0^2 - \sigma^2}{\sigma^2\sigma_0^2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right) \\ &\quad - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0, t_i) - \mu(\eta, t_i)]\xi_i.\end{aligned}$$

For any $\theta' \in \Theta$ and $r > 0$, let $B(\theta', r) = \{\theta, ;\ \|\theta' - \theta\|_1 < r\}$. It is now obvious that

$$\begin{aligned}&\frac{2}{n}[l_{1:n}(X_{1:n}|B(\theta', r)) - l_{1:n}(X_{1:n}|K)] \\ &\leqslant \sup_{\theta \in B(\theta', r)}\{-b_n(\theta)\} + \sup_{\theta \in B(\theta', r)}\left|\frac{\sigma_0^2 - \sigma^2}{\sigma^2\sigma_0^2}\right| \cdot \left|\frac{1}{n}\sum_{i=1}^n \xi_i^2 - \sigma_0^2\right| \\ &\quad + \sup_{\theta \in B(\theta', r)}\left\{\frac{2}{\sigma^2}\right\} \cdot \sup_{\theta \in B(\theta', r)}\left\{\left|\frac{1}{n}\sum_{i=1}^n [\mu(\eta_0, t_i) - \mu(\eta, t_i)]\xi_i\right|\right\}.\end{aligned} \tag{6.20}$$

Lemma 7.6 now ensures that

$$\sup_{\theta\in B(\theta',r)}\left|\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_0,t_i)-\mu(\eta,t_i)]\xi_i\right|\xrightarrow{a.s.}0,$$

and σ^2 being bounded away from 0 ensures the boundedness of $\sup_{\theta\in B(\theta',r)}\left\{\frac{2}{\sigma^2}\right\}$ which implies

$$\sup_{\theta\in B(\theta',r)}\left\{\frac{2}{\sigma^2}\right\}\cdot\sup_{\theta\in B(\theta',r)}\left\{\left|\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_0,t_i)-\mu(\eta,t_i)]\xi_i\right|\right\}\xrightarrow{a.s.}0.$$

Since σ^2 is bounded away from 0, taking advantage of the Strong Law of Large Numbers, we also obtain

$$\sup_{\theta\in B(\theta',r)}\left|\frac{\sigma^2-\sigma_0^2}{\sigma^2\sigma_0^2}\right|\cdot\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i^2-\sigma_0^2\right|\xrightarrow{a.s.}0.$$

We may thus rewrite (6.20) as

$$\frac{2}{n}[l_{1:n}(X_{1:n}|B(\theta',r))-l_{1:n}(X_{1:n}|K)]\leqslant\sup_{\theta\in B(\theta',r)}\{-b_n(\theta)\}+R_n,\tag{6.21}$$

where $R_n\xrightarrow{a.s.}0$.

Assume now that $\theta'\neq\theta_0$, then we have

$$\sup_{\theta\in B(\theta',r)}|b_n(\theta)-b(\theta')|\leqslant\sup_{\theta\in B(\theta',r)}|b_n(\theta)-b_n(\theta')|+|b_n(\theta')-b(\theta')|.\tag{6.22}$$

Lemma 7.5 (see (7.9)) ensures the existence of a r small enough, say $r=r(\theta')$, such that

$$\sup_{\theta\in B(\theta',r(\theta'))}|b_n(\theta)-b_n(\theta')|\leqslant\frac{1}{4}b(\theta'),\tag{6.23}$$

uniformly in n. For n large enough, that same Lemma 7.5 (see (7.10)) also guarantees that

$$|b_n(\theta')-b(\theta')|\leqslant\frac{1}{4}b(\theta').\tag{6.24}$$

Adding inequalities (6.23) and (6.24) together and combining the result with (6.22), we deduce that for any n large enough

$$\sup_{\theta\in B(\theta',r(\theta'))}|b_n(\theta)-b(\theta')|\leqslant\frac{1}{2}b(\theta'),$$

i.e.

$$\sup_{\theta\in B(\theta',r(\theta'))}\{-b_n(\theta)\}\leqslant-\frac{1}{2}b(\theta'),$$

which finally gives together with (6.21)

$$\forall\theta'\neq\theta_0,\ \mathbb{P}\left(\limsup_{n\to+\infty}\frac{1}{n}[l_{1:n}(X_{1:n}|B(\theta',r(\theta')))-l_{1:n}(X_{1:n}|K)]\leqslant-\frac{1}{4}b(\theta')\right)=1.\tag{6.25}$$

Since Lemma 7.5 ensures that $b(\theta')>0$ for any $\theta'\neq\theta_0$, the previous statement implies

$$\forall\theta'\neq\theta_0,\ \mathbb{P}\left(\exists n(\theta')\in\mathbb{N},\ \forall n>n(\theta'),\ l_{1:n}(X_{1:n}|B(\theta',r(\theta')))-l_{1:n}(X_{1:n}|K)<-1\right)=1.\tag{6.26}$$

For a given $\delta > 0$, let us now define $K(\delta) = K \setminus B(\theta_0, \delta)$. $K(\delta)$ is obviously a compact set since K itself is a compact set. By compacity, from the covering

$$\bigcup_{\theta' \in K(\delta)} B(\theta', r(\theta')) \supset K(\delta),$$

there exists a finite subcovering, i.e.

$$\exists m(\delta) \in \mathbb{N},\ \bigcup_{j=1}^{m(\delta)} B(\theta'_j, r(\theta'_j)) \supset K(\delta).$$

In particular, (6.26) holds for $\theta' = \theta'_j, j = 1, \dots, m(\delta)$. Let us define

$$n_0(\delta) = \max_{j=1,\dots,m(\delta)} n(\theta'_j).$$

We may now write

$$\begin{aligned}&\forall \delta > 0,\ \exists n_0(\delta) \in \mathbb{N},\ \exists m(\delta) \in \mathbb{N},\ \forall j = 1, \dots, m(\delta),\\ &\quad \mathbb{P}\left(\forall n > n_0(\delta),\ l_{1:n}(X_{1:n}|B(\theta'_j, r(\theta'_j))) - l_{1:n}(X_{1:n}|K) < -1\right) = 1,\end{aligned}$$

which we turn into

$$\begin{aligned}&\forall \delta > 0,\ \exists n_0(\delta) \in \mathbb{N},\ \exists m(\delta) \in \mathbb{N},\\ &\quad \mathbb{P}\left(\forall n > n_0(\delta),\ \forall j = 1, \dots, m(\delta),\ l_{1:n}(X_{1:n}|B(\theta'_j, r(\theta'_j))) - l_{1:n}(X_{1:n}|K) < -1\right) = 1,\end{aligned}$$

thanks to the finiteness of $m(\delta)$, and finally into

$$\forall \delta > 0,\ \exists n_0(\delta) \in \mathbb{N},\ \mathbb{P}\left(\forall n > n_0(\delta),\ l_{1:n}(X_{1:n}|K(\delta)) - l_{1:n}(X_{1:n}|K) < -1\right) = 1,$$

because of the covering

$$\bigcup_{j=1}^{m(\delta)} B(\theta'_j, r(\theta'_j)) \supset K(\delta).$$

Let us now sum up what we have obtained so far. We proved that

$$\forall \delta > 0,\ \exists n_0(\delta) \in \mathbb{N},\ \mathbb{P}\left(\text{if } \forall n > n_0(\delta),\ l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|K) \geqslant \log e^{-1},\ \text{then } \theta \notin K(\delta)\right) = 1,$$

i.e.

$$\exists a = e^{-1} \in]0, 1[,\ \forall \delta > 0,\ \exists n_0(\delta) \in \mathbb{N},\ \mathbb{P}\left(\text{if } \forall n > n_0(\delta),\ \theta \in K_n(a),\ \text{then } \|\theta - \theta_0\|_1 < \delta\right) = 1,$$

that is to say

$$\exists a \in]0, 1[,\ \mathbb{P}\left(\lim_{n \to +\infty} \sup_{\theta \in K_n(a)} \|\theta - \theta_0\|_1 = 0\right) = 1.$$

□

Proof of Proposition 3.2. In this proof $\|\cdot\|$ will refer to the usual Euclidean norm. Reindexing whenever necessary, we also assume that the observations t_i are ordered, and we denote

$$t = (t_1, \dots, t_n), \qquad X = (X_1, \dots, X_n), \qquad \mu_0 = (\mu(\eta_0, t_1), \dots, \mu(\eta_0, t_n)),$$

$$N_{0,n} = \sup_{i \leqslant n}\{i,\ t_i < u_0\} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[t_i,\,+\infty[}(u_0), \qquad N_n = \sup_{i \leqslant n}\{i,\ t_i < \widehat{u}_n\} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[t_i,\,+\infty[}(\widehat{u}_n),$$

$$\zeta = \begin{cases} (0,\dots,0,\beta_0 + \gamma_0 t_{N_n+1},\dots,\beta_0 + \gamma_0 t_{N_{0,n}},0,\dots,0), & \text{if } N_n < N_{0,n} \\ (0,\dots,0), & \text{if } N_n = N_{0,n} \\ (0,\dots,0,\beta_0 + \gamma_0 t_{N_{0,n}+1},\dots,\beta_0 + \gamma_0 t_{N_n},0,\dots,0), & \text{if } N_n > N_{0,n} \end{cases},$$

Let $\mathcal{G}$ be the linear space spanned by the 2 linearly independent n-vectors

$$v_1 = (1,\dots,1,0,\dots,0) \qquad\qquad v_2 = (t_1,\dots,t_{N_n},0,\dots,0)$$

(both of which have their last $n - N_n$ coordinates valued to zero), and denote Q the orthogonal projection onto $\mathcal{G}$.

Let $\mathcal{G}^+$ denote the linear space spanned by v_1, v_2 and μ_0 and denote Q^+ the orthogonal projection onto $\mathcal{G}^+$. Observe that $\mathcal{G}^+$ is also spanned by v_1, v_2 and ζ.

Finally, denote μ^* the orthogonal projection of X onto $\mathcal{G}^+$ and $\widehat{\mu}$ the closest point to X in $\mathcal{G}^+$ satisfying the continuity assumption of the model, i.e.

$$\mu^* = Q^+ X, \qquad\qquad \widehat{\mu} = (\mu(\widehat{\eta}_n, t_1),\dots,\mu(\widehat{\eta}_n, t_n)).$$

We have

$$\begin{aligned} \|X - \mu^*\|^2 + \|\mu^* - \widehat{\mu}\|^2 = \|X - \widehat{\mu}\|^2 &\leqslant \|X - \mu_0\|^2, \\ \|X - \mu_0\|^2 - \|\mu^* - \mu_0\|^2 + \|\mu^* - \widehat{\mu}\|^2 &\leqslant \|X - \mu_0\|^2, \\ \|\mu^* - \mu_0\|^2 - 2\,\langle \mu^* - \mu_0, \widehat{\mu} - \mu_0\rangle + \|\widehat{\mu} - \mu_0\|^2 &\leqslant \|\mu^* - \mu_0\|^2. \end{aligned}$$

Thus

$$\|\widehat{\mu} - \mu_0\|^2 \leqslant 2\,\langle \mu^* - \mu_0, \widehat{\mu} - \mu_0\rangle \leqslant 2\|\mu^* - \mu_0\| \cdot \|\widehat{\mu} - \mu_0\|,$$

which leads to

$$\|\widehat{\mu} - \mu_0\| \leqslant 2\|\mu^* - \mu_0\| \leqslant 2\|Q^+\xi\|.$$

Our aim is to show that a.s.

$$\|Q^+\xi\| = \mathrm{O}\left(\log n\right). \tag{6.27}$$

If (6.27) held, then we would have a.s. $\|\widehat{\mu} - \mu_0\| = \mathrm{O}\left(\log n\right)$ i.e. a.s.

$$\sum_{i=1}^{n} \left(\mu(\widehat{\eta}_n, t_i - \mu(\eta_0, t_i)\right)^2 = \mathrm{O}\left(\log^2 n\right).$$

Hence, a.s. for any open interval $I \subset [\underline{u},\, \overline{u}]$ we would have

$$\sum_{i=1}^{n} \left(\mu(\widehat{\eta}_n, t_i - \mu(\eta_0, t_i)\right)^2 \mathbb{1}_I(t_i) = \mathrm{O}\left(\log^2 n\right).$$

This would immediately imply the desired result, i.e. that a.s.

$$\min_{t_i \in I,\ i \leqslant n} |\mu(\widehat{\eta}_n, t_i) - \mu(\eta_0, t_i)| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right),$$

since a.s.

$$\mathrm{O}\left(\log^2 n\right) = \sum_{i=1}^{n} \left(\mu(\widehat{\eta}_n, t_i - \mu(\eta_0, t_i)\right)^2 \mathbb{1}_I(t_i) \geqslant n \cdot \min_{t_i \in I,\ i \leqslant n} |\mu(\widehat{\eta}_n, t_i) - \mu(\eta_0, t_i)|^2 \cdot \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_I(t_i),$$

where (see Assumption (A1))

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_I(t_i) = \int_I \mathrm{d}F_n(t) \to \int_I \mathrm{d}F(t) = \int_I f(t)\,\mathrm{d}t > 0.$$

Let us now prove that (6.27) indeed holds. We consider the two following mutually exclusive situations.

Situation A: $\zeta = (0,\ldots,0)$. In this situation

$$\|Q^+\xi\| = \|Q\xi\|, \tag{6.28}$$

and Cochran's theorem guarantees that $\|Q\xi\|^2 \sim \chi^2(2)$ for $n \geqslant 2$. Hence, via Corollary 7.7, a.s.

$$\|Q\xi\| = \mathrm{O}\left(\log n\right), \tag{6.29}$$

and (6.27) follows from (6.28) and (6.29).

Situation B: $\zeta \neq (0,\ldots,0)$. Since

$$\frac{|\langle\zeta,\xi\rangle|}{\|\zeta\|} \sim \mathcal{N}(0,\sigma_0^2),$$

we also have, via Lemma 7.7, a.s.

$$\frac{|\langle\zeta,\xi\rangle|}{\|\zeta\|} = \mathrm{O}\left(\log n\right). \tag{6.30}$$

Notice that (6.27) follows from (6.29) and (6.30) if we manage to show that a.s.

$$\|Q^+\xi\| \leqslant \mathrm{O}(1)\cdot\left(\|Q\xi\| + \frac{|\langle\zeta,\xi\rangle|}{\|\zeta\|}\right). \tag{6.31}$$

It thus now suffices to prove that a.s., for any $g \in \mathcal{G}$

$$|\langle\zeta, g\rangle| = \|\zeta\|\,\|g\|\cdot\mathrm{o}(1), \tag{6.32}$$

where the o(1) mentioned in (6.32) is uniform in g over $\mathcal{G}$ (i.e. a.s. ζ is asymptotically uniformly orthogonal to $\mathcal{G}$), for (6.31) is a direct consequence of (6.32) and Lemma 6.1 whose proof is found in Feder (1975).

Lemma 6.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two linear subspaces of an inner product space $\mathcal{E}$. If there exists $\alpha < 1$ such that*

$$\forall (x,y) \in \mathcal{X}\times\mathcal{Y},\ |\langle x,y\rangle| \leqslant \alpha\|x\|\,\|y\|,$$

then

$$\|x+y\| \leqslant (1-\alpha)^{-1}(\|x^*\| + \|y^*\|),$$

where x^* (resp. y^*) is the orthogonal projection of $x+y$ onto $\mathcal{X}$ (resp. $\mathcal{Y}$).

Observe that, as a consequence of Assumption (A1) and Theorem 7.1, the three following convergences are uniform in u over $[\underline{u},\,\overline{u}]$ for $k = 0,1,2$,

$$\frac{1}{n}\sum_{i=1}^{n} t_i^k \mathbb{1}_{[t_i,\,+\infty[}(u) = \int_{\underline{u}}^{u} t^k\,\mathrm{d}F_n(t) \to \int_{\underline{u}}^{u} t^k\,\mathrm{d}F(t) = \int_{\underline{u}}^{u} t^k f(t)\,\mathrm{d}t. \tag{6.33}$$

We have a.s., for any $g(\phi) = (\cos\phi)v_1 + (\sin\phi)v_2 \in \mathcal{G}$, with $\phi \in [0,\, 2\pi]$

$$\begin{aligned}
|\langle \zeta, g(\phi)\rangle| &= \left| \sum_{i=1}^{N_n} (\beta_0 + \gamma_0 t_i)(\cos\phi + t_i \sin\phi) - \sum_{i=1}^{N_{0,n}} (\beta_0 + \gamma_0 t_i)(\cos\phi + t_i \sin\phi) \right| \\
&\leqslant (\max(|\underline{u}|, |\overline{u}|) + 1) \cdot \left| \sum_{i=1}^{N_n} |\beta_0 + \gamma_0 t_i| - \sum_{i=1}^{N_{0,n}} |\beta_0 + \gamma_0 t_i| \right| \\
&\leqslant (\max(|\underline{u}|, |\overline{u}|) + 1) \cdot \|\zeta\|_1 \\
&\leqslant (\max(|\underline{u}|, |\overline{u}|) + 1) \cdot \|\zeta\| \cdot n^{\frac{1}{2}} |N_n - N_{0,n}|^{\frac{1}{2}} \\
&\leqslant (\max(|\underline{u}|, |\overline{u}|) + 1) \cdot \|\zeta\| \cdot n^{\frac{1}{2}} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[t_i,\, +\infty[}(\widehat{u}_n) - \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[t_i,\, +\infty[}(u_0) \right|^{\frac{1}{2}},
\end{aligned}$$

i.e. we have a.s. for any $\phi \in [0,\, 2\pi]$

$$|\langle \zeta, g(\phi)\rangle| = n^{\frac{1}{2}} \|\zeta\| \cdot \mathrm{o}(1), \tag{6.34}$$

thanks to the strong consistency $\widehat{u}_n \xrightarrow{a.s.} u_0$ (see Theorem 3.1) and the uniform convergence mentioned in (6.33) with $(k = 0)$. Observe that the $\mathrm{o}(1)$ mentioned in (6.34) is uniform in ϕ over $[0,\, 2\pi]$. We also have a.s. for any $\phi \in [0,\, 2\pi]$

$$\begin{aligned}
\frac{1}{n}\|g(\phi)\|^2 &= \frac{1}{n} \sum_{i=1}^{n} (\cos\phi + t_i \sin\phi)^2 \mathbb{1}_{[t_i,\, +\infty[}(\widehat{u}_n) \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[t_i,\, +\infty[}(\widehat{u}_n) \cos^2\phi + 2\frac{1}{n} \sum_{i=1}^{n} t_i \mathbb{1}_{[t_i,\, +\infty[}(\widehat{u}_n) \cos\phi \sin\phi \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} t_i^2 \mathbb{1}_{[t_i,\, +\infty[}(\widehat{u}_n) \sin^2\phi \\
&\xrightarrow{a.s.} \cos^2\phi \int_{\underline{u}}^{u_0} f(t)\,\mathrm{d}t + \cos\phi \sin\phi \int_{\underline{u}}^{u_0} 2t f(t)\,\mathrm{d}t + \sin^2\phi \int_{\underline{u}}^{u_0} t^2 f(t)\,\mathrm{d}t,
\end{aligned}$$

once again making use of the strong consistency $\widehat{u}_n \xrightarrow{a.s.} u_0$ (see Theorem 3.1) and taking advantage of all three uniform convergences mentioned in (6.33). We thus obviously have a.s., uniformly in ϕ over $[0,\, 2\pi]$

$$\frac{1}{n}\|g(\phi)\|^2 \to \int_{\underline{u}}^{u_0} (\cos\phi + t\sin\phi)^2 f(t)\,\mathrm{d}t. \tag{6.35}$$

The limit in (6.35) is a positive and continuous function of ϕ, and is hence bounded, i.e. there exists $m > 0$ such that we have a.s.

$$\frac{1}{n}\|g(\phi)\|^2 \geqslant m + \mathrm{o}(1), \tag{6.36}$$

i.e.

$$\frac{1}{\|g(\phi)\|} = \mathrm{O}(n^{-\frac{1}{2}}), \tag{6.37}$$

where the $\mathrm{o}(1)$ mentioned in (6.36) and the $\mathrm{O}(n^{-\frac{1}{2}})$ mentioned in (6.37) are uniform in ϕ over $[0,\, 2\pi]$.

Combining (6.34) and (6.37) together, we have a.s. for any $\phi \in [0,\, 2\pi]$

$$|\langle \zeta, g(\phi)\rangle| = \|\zeta\|\, \|g(\phi)\| \cdot \mathrm{o}(1), \tag{6.38}$$

where the o(1) mentioned in (6.38) is uniform in ϕ over $[0,\,2\pi]$.

Hence, we have a.s, for any $r \in \mathbb{R}_+^*$, and any $\phi \in [0,\,2\pi]$, now denoting $g(\phi) = (r\cos\phi)v_1 + (r\sin\phi)v_2$ and applying (6.38) to $r^{-1}g(\phi)$

$$|\langle \zeta, g(\phi)\rangle| = r\left|\left\langle \zeta, r^{-1}g(\phi)\right\rangle\right| = r\cdot\|\zeta\|\,\|r^{-1}g(\phi)\|\cdot \mathrm{o}(1) = \|\zeta\|\,\|g(\phi)\|\cdot\mathrm{o}(1),$$

where the o(1) mentioned is uniform in ϕ over $[0,\,2\pi]$ and does not depend on r.

We immediately deduce that a.s. (6.32) holds i.e. a.s. ζ is asymptotically uniformly orthogonal to $\mathcal{G}$, which completes the proof. □

6.3 Proofs of Section 4

Proof of Proposition 4.1. We proceed as announced.

Step 1. We first prove that a.s.

$$\exists N \in \mathbb{N},\ \forall n > N,\ \widehat{u}_n^* \in D_n.$$

Let us notice that anything proven for the problem remains valid for the pseudo-problem. Because $n^* \sim n$, we have a.s., thanks to Theorem 3.3 and conditions (2.6), as $n \to +\infty$

$$n^{\frac{1}{2}}(\log^{-1} n)\cdot(\widehat{u}_n^* - u_0) = \mathrm{O}(1),$$
$$n^{\frac{1}{2}}(\log^{-1} n)\cdot d_n \to +\infty,$$

and thus deduce from the ratio of these two quantities that

$$\frac{\widehat{u}_n^* - u_0}{d_n} \xrightarrow{a.s.} 0,$$

and this directly implies the desired result.

Step 2. Let $A_{1:n}^*(\cdot)$ be the column vector defined for $u \in D_n$ by

$$A_{1:n}^*(\theta) = \left(\left.\frac{\partial l_{1:n}^*(X_{1:n}|\theta)}{\partial\gamma}\right|_\theta,\ \left.\frac{\partial l_{1:n}^*(X_{1:n}|\theta)}{\partial u}\right|_\theta,\ \left.\frac{\partial l_{1:n}^*(X_{1:n}|\theta)}{\partial\sigma^2}\right|_\theta\right). \tag{6.39}$$

Step 1 allows us to expand a.s. $A_{1:n}^*(\widehat{\theta}_n^*)$ around θ_0 using a Taylor-Lagrange approximation

$$0 = A_{1:n}^*(\widehat{\theta}_n^*) = A_{1:n}^*(\theta_0) - B_{1:n}^*(\widetilde{\theta}_n)\left(\widehat{\theta}_n^* - \theta_0\right),$$

where $\widetilde{\theta}_n$ is a point between $\widehat{\theta}_n^*$ and θ_0 (see (6.14) for the definitions of $B_{1:n}^*$), and rewrite it as a.s.

$$\frac{1}{n^*}B_{1:n}^*(\widetilde{\theta}_n)\cdot n^{*\frac{1}{2}}\left(\widehat{\theta}_n^* - \theta_0\right) = n^{*-\frac{1}{2}}A_{1:n}^*(\theta_0).$$

Since $\widehat{\theta}_n^* \to \theta_0$, we also have $\widetilde{\theta}_n \to \theta_0$ and using both Lemmas 7.9 and 7.10 we immediately find that as $n \to +\infty$

$$I(\theta_0)\cdot n^{*\frac{1}{2}}\left(\widehat{\theta}_n^* - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, I(\theta_0)\right),$$

which means, remembering both that $n^* \sim n$ and that $I(\theta_0)$ is positive definite and thus invertible that as $n \to +\infty$

$$n^{\frac{1}{2}}\left(\widehat{\theta}_n^* - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, I(\theta_0)^{-1}\right).$$

□

Proof of Theorem 3.3. We now prove that

$$\|\widehat{\sigma}_n^2 - \sigma_0^2\| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right).$$

VARIANCE OF NOISE σ^2. Observe that

$$\begin{aligned}
\widehat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^{n} \left[X_i - \widehat{\gamma}_n \left(t_i - \widehat{u}_n\right)\mathbb{1}_{[t_i,+\infty[}(\widehat{u}_n)\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \left[\gamma_0 \cdot (t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \widehat{\gamma}_n \cdot (t_i - \widehat{u}_n)\mathbb{1}_{[t_i,+\infty[}(\widehat{u}_n) + \xi_i\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n) + \frac{2}{n}\sum_{i=1}^{n} \nu_i(\widehat{\eta}_n)\xi_i + \frac{1}{n}\sum_{i=1}^{n} \xi_i^2, \qquad (6.40)
\end{aligned}$$

where we denote for $i = 1, \ldots, n$,

$$\nu_i(\eta) = \gamma_0 \cdot (t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma \cdot (t_i - u)\mathbb{1}_{[t_i,+\infty[}(u). \qquad (6.41)$$

We have

$$\begin{aligned}
\sup_{i\in\mathbb{N}} |\nu_i(\widehat{\eta}_n)| &= \sup_{i\in\mathbb{N}} \left|\gamma_0 \cdot (t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \widehat{\gamma}_n \cdot (t_i - \widehat{u}_n)\mathbb{1}_{[t_i,+\infty[}(\widehat{u}_n)\right| \\
&\leqslant |\gamma_0 - \widehat{\gamma}_n| \cdot \sup_{i\in\mathbb{N}} \left|(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\right| \\
&\quad + |\widehat{\gamma}_n| \cdot \sup_{i\in\mathbb{N}} \left|(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - (t_i - \widehat{u}_n)\mathbb{1}_{[t_i,+\infty[}(\widehat{u}_n)\right| \\
&= \mathrm{O}\left(\gamma_0 - \widehat{\gamma}_n\right) + |\widehat{\gamma}_n|\,\mathrm{O}\left(u_0 - \widehat{u}_n\right), \qquad (6.42)
\end{aligned}$$

using straightforward dominations and Lemma 7.2, so that in the end, thanks to the previous results we have a.s.

$$\sup_{i\in\mathbb{N}} |\nu_i(\widehat{\eta}_n)| = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right). \qquad (6.43)$$

It is thus easy to see that a.s.

$$\frac{1}{n}\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n) = \mathrm{O}\left(n^{-1}\log^2 n\right) = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right), \qquad (6.44)$$

and also that, via Corollary 7.7, a.s.

$$\frac{2}{n}\sum_{i=1}^{n} \nu_i(\widehat{\eta}_n)\xi_i = \frac{2}{n}\left(\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n)\right)^{\frac{1}{2}} \cdot \mathrm{O}(\log n) = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right). \qquad (6.45)$$

From the Law of the Iterated Logarithm (see Breiman, 1992, Chapter 13, page 291) we have a.s.

$$\frac{1}{n}\sum_{i=1}^{n} \left(\xi_i^2 - \sigma_0^2\right) = \mathrm{O}\left(n^{-\frac{1}{2}} (\log\log n)^{\frac{1}{2}}\right) = \mathrm{O}\left(n^{-\frac{1}{2}} \log n\right) \qquad (6.46)$$

and the desired result follows from (6.44), (6.45) and (6.46) put together into (6.40). □

Proof of Theorem 4.2. To finish the proof, we need to show (4.3) i.e. that

$$\widehat{\sigma}_n^2 - \widehat{\sigma}_n^{2*} = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right).$$

We use the decomposition (6.40)

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n) + \frac{2}{n}\sum_{i=1}^{n} \nu_i(\widehat{\eta}_n)\xi_i + \frac{1}{n}\sum_{i=1}^{n} \xi_i^2,$$

where $\nu_i(\widehat{\eta}_n) = \gamma_0 \cdot (t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \widehat{\gamma}_n \cdot (t_i - \widehat{u}_n)\mathbb{1}_{[t_i,+\infty[}(\widehat{u}_n)$.

Having proved in Proposition 4.1 that

$$\widehat{\gamma}_n^* - \gamma_0 = \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right), \qquad \widehat{u}_n^* - u_0 = \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right)$$

we add these relationships to those from (4.2) and find that

$$\widehat{\gamma}_n - \gamma_0 = \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right), \qquad \widehat{u}_n - u_0 = \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right). \tag{6.47}$$

We now use (6.47) together with (6.42), we are able to write

$$\sup_{i\in\mathbb{N}} |\nu_i(\widehat{\eta}_n)| = \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right). \tag{6.48}$$

It is hence easy to see that

$$\frac{1}{n}\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n) = \mathrm{O}_{\mathbb{P}}\left(n^{-1}\right) = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right),$$

and also that

$$\frac{2}{n}\sum_{i=1}^{n} \nu_i(\widehat{\eta}_n)\xi_i = \frac{2}{n}\left(\sum_{i=1}^{n} \nu_i^2(\widehat{\eta}_n)\right)^{\frac{1}{2}} \cdot \mathrm{O}_{\mathbb{P}}(1) = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right),$$

which once both substituted into (6.40) yield

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} \xi_i^2 + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right).$$

What was done above with the problem and $\widehat{\sigma}_n^2$ can be done with the pseudo-problem and $\widehat{\sigma}_n^{2*}$ without any kind of modification so that

$$\widehat{\sigma}_n^{2*} = \frac{1}{n^*}\sum_{i=1}^{n^*} \xi_i^2 + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right).$$

We observe that

$$\begin{aligned}
\widehat{\sigma}_n^2 - \widehat{\sigma}_n^{2*} &= \frac{1}{n}\sum_{i=1}^{n} \xi_i^2 - \frac{1}{n^*}\sum_{i=1}^{n^*} \xi_i^2 + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) \\
&= \left[\frac{1}{n} - \frac{1}{n^*}\right] \cdot \sum_{i=1}^{n^*} \xi_i^2 + \frac{1}{n} \cdot \sum_{i=n^*+1}^{n} \xi_i^2 + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) \\
&= \frac{n^* - n}{n} \cdot \left(\frac{1}{n^*}\sum_{i=1}^{n^*} \xi_i^2\right) + \frac{n - n^*}{n} \cdot \left(\frac{1}{n - n^*}\sum_{i=n^*+1}^{n} \xi_i^2\right) + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) \\
&= \frac{n^* - n}{n} \cdot \left(\sigma_0^2 + \mathrm{O}_{\mathbb{P}}\left(n^{*-\frac{1}{2}}\right)\right) + \frac{n - n^*}{n} \cdot \left(\sigma_0^2 + \mathrm{O}_{\mathbb{P}}\left((n - n^*)^{-\frac{1}{2}}\right)\right) + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right),
\end{aligned}$$

using the Central Limit Theorem, and in the end we get

$$\begin{aligned}\widehat{\sigma}_n^2 - \widehat{\sigma}_n^{2*} &= \frac{n^* - n}{n} \cdot \mathrm{O}_{\mathbb{P}}\left(n^{*-\frac{1}{2}}\right) + \frac{n - n^*}{n} \cdot \mathrm{O}_{\mathbb{P}}\left((n - n^*)^{-\frac{1}{2}}\right) + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) \\ &= \mathrm{o}(1) \cdot \mathrm{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) + n^{-\frac{1}{2}} \cdot \mathrm{O}_{\mathbb{P}}\left(\left(\frac{n - n^*}{n}\right)^{\frac{1}{2}}\right) + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) \\ &= \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) + n^{-\frac{1}{2}} \cdot \mathrm{O}_{\mathbb{P}}\left(\mathrm{o}(1)\right) + \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right) = \mathrm{o}_{\mathbb{P}}\left(n^{-\frac{1}{2}}\right).\end{aligned}$$

□

7 Technical results

Theorem 7.1 (Polya's Theorem)**.** *Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of non decreasing (or non increasing) functions defined over $I = [a, b] \subset \mathbb{R}$. If g_n converges pointwise to g (i.e. $g_n(x) \to g(x)$ as $n \to +\infty$, for any $x \in I$) and g is continuous then*

$$\sup_{x\in I} |g_n(x) - g(x)| \xrightarrow[n\to+\infty]{} 0.$$

Proof of Lemma 7.1. Assume the functions g_n are non decreasing over I (if not, consider their opposites $-g_n$). g is continuous over I and thus bounded since I is compact. g is also non decreasing over I as the limit of a sequence of non decreasing functions. Let $\epsilon > 0$ and $k > \frac{g(b)-g(a)}{\epsilon}$ such that

$$\exists a = a_0 < \ldots < a_k = b \in I^{k+1},\ \forall i = 0, \ldots, k-1,\ g(a_{i+1}) - g(a_i) < \epsilon.$$

Now let $x \in I$ and let $i \in \mathbb{N}$ such that $a_i \leqslant x \leqslant a_{i+1}$. Since g_n and g are non decreasing, we find that

$$\begin{aligned} g_n(x) - g(x) &\leqslant g_n(a_{i+1}) - g(a_i) \leqslant g_n(a_{i+1}) - g(a_{i+1}) + \epsilon, \\ g_n(x) - g(x) &\geqslant g_n(a_i) - g(a_{i+1}) \geqslant g_n(a_i) - g(a_i) - \epsilon. \end{aligned}$$

The pointwise convergence of g_n to g and the finiteness of k together ensure that

$$\exists N_0 \in \mathbb{N},\ \forall n \geqslant N_0,\ \forall i = 0, \ldots, k,\ |g_n(a_i) - g(a_i)| < \epsilon,$$

which implies with both of the inequations mentioned above that

$$\exists N_0 \in \mathbb{N},\ \forall n \geqslant N_0,\ \forall x \in I,\ |g_n(x) - g(x)| < \epsilon.$$

□

Lemma 7.2. *Let $k \in \mathbb{N}^*$, there exists a constant $C \in \mathbb{R}_+^*$ such that for any $(u, u') \in [\underline{u}, \overline{u}]^2$*

$$\sup_{t\in[\underline{u},\,\overline{u}]} |(t-u')^k \mathbb{1}_{[t,\,+\infty[}(u') - (t-u)^k \mathbb{1}_{[t,\,+\infty[}(u)| = C|u - u'|. \tag{7.1}$$

Proof of Lemma 7.2. For any $(u, u') \in [\underline{u}, \overline{u}]^2$ we have

$$\begin{aligned}\sup_{t\in[\underline{u},\,\overline{u}]} |(t-u')^k \mathbb{1}_{[t,\,+\infty[}(u') - (t-u)^k \mathbb{1}_{[t,\,+\infty[}(u)| \leqslant &\sup_{t\in[\underline{u},\,\overline{u}]} \{|(t-u')^k - (t-u)^k| \mathbb{1}_{[t,\,+\infty[}(u')\} \\ &+ \sup_{t\in[\underline{u},\,\overline{u}]} \{|t-u|^k |\mathbb{1}_{[t,\,+\infty[}(u') - \mathbb{1}_{[t,\,+\infty[}(u)|\}.\end{aligned} \tag{7.2}$$

The mean value theorem guarantees that there exists v between u and u' such that

$$(t-u')^k - (t-u)^k = -k(t-v)^{k-1}(u'-u).$$

We thus have

$$\begin{aligned}\sup_{t\in[\underline{u},\,\overline{u}]}\{|(t-u')^k-(t-u)^k|\mathbb{1}_{[t,\,+\infty[}(u')\} &\leqslant \sup_{t\in[\underline{u},\,\overline{u}]}|(t-u')^k-(t-u)^k| \\ &\leqslant k|\overline{u}-\underline{u}|^{k-1}|u-u'|. \end{aligned} \tag{7.3}$$

Because $|t-u| \leqslant |u'-u|$ whenever $|\mathbb{1}_{[t,\,+\infty[}(u') - \mathbb{1}_{[t,\,+\infty[}(u)| \neq 0$, we also find that

$$\sup_{t\in[\underline{u},\,\overline{u}]}\{|t-u|^k|\mathbb{1}_{[t,\,+\infty[}(u')-\mathbb{1}_{[t,\,+\infty[}(u)|\} \leqslant |u-u'|^k \leqslant |u-u'||\overline{u}-\underline{u}|^{k-1}. \tag{7.4}$$

And now (7.1) is a simple consequence of (7.2), (7.3) and (7.4). □

Lemma 7.3. *For any $\eta' \in \mathbb{R}\times[\underline{u},\,\overline{u}]$, there exists $C \in \mathbb{R}_+^*$ such that for any $\eta \in \mathbb{R}\times[\underline{u},\,\overline{u}]$*

$$\sup_{t\in[\underline{u},\,\overline{u}]}|\mu(\eta,t)-\mu(\eta',t)| \leqslant C\|\eta-\eta'\|. \tag{7.5}$$

Proof of Lemma 7.3. We have indeed

$$\begin{aligned}\sup_{t\in[\underline{u},\,\overline{u}]}|\mu(\eta,t)-\mu(\eta',t)| &= \sup_{t\in[\underline{u},\,\overline{u}]}|\gamma\cdot(t-u)\mathbb{1}_{[t,\,+\infty[}(u)-\gamma'(t-u')\mathbb{1}_{[t,\,+\infty[}(u')| \\ &\leqslant \sup_{t\in[\underline{u},\,\overline{u}]}|[\gamma-\gamma'](t-u)\mathbb{1}_{[t,\,+\infty[}(u)| \\ &\quad + \sup_{t\in[\underline{u},\,\overline{u}]}|\gamma'[(t-u)\mathbb{1}_{[t,\,+\infty[}(u)-(t-u')\mathbb{1}_{[t,\,+\infty[}(u')]| \\ &\leqslant |\gamma-\gamma'|\cdot\sup_{t\in[\underline{u},\,\overline{u}]}|t-u| \\ &\quad + |\gamma'|\cdot\sup_{t\in[\underline{u},\,\overline{u}]}|(t-u)\mathbb{1}_{[t,\,+\infty[}(u)-(t-u')\mathbb{1}_{[t,\,+\infty[}(u')| \\ &\leqslant |\gamma-\gamma'|\cdot|\overline{u}-\underline{u}| + |\gamma'|\cdot\sup_{t\in[\underline{u},\,\overline{u}]}|(t-u)\mathbb{1}_{[t,\,+\infty[}(u)-(t-u')\mathbb{1}_{[t,\,+\infty[}(u')|. \end{aligned}$$

And now (7.5) is a simple consequence of Lemma 7.2. □

Lemma 7.4. *Let $A \subset \mathbb{R}\times[\underline{u},\,\overline{u}]$ be a bounded set. Then,*

$$\begin{aligned}&\forall\epsilon>0,\ \exists m(\epsilon)\in\mathbb{N},\ \exists\eta_1,\dots,\eta_{m(\epsilon)}\in A, \\ &\forall\eta,\eta'\in A,\ \exists j,j'\in\{1,\dots,m(\epsilon)\},\ \sup_{t\in[\underline{u},\,\overline{u}]}\big|\big[\mu(\eta,t)-\mu(\eta',t)\big]-\big[\mu(\eta_j,t)-\mu(\eta_{j'},t)\big]\big|<\epsilon,\end{aligned}$$

Proof of Lemma 7.4. It suffices to prove the following claim

$$\begin{aligned}&\forall\epsilon>0,\ \exists m(\epsilon)\in\mathbb{N},\ \exists\eta_1,\dots,\eta_{m(\epsilon)}\in A, \\ &\forall\eta\in A,\ \exists j\in\{1,\dots,m(\epsilon)\},\ \sup_{t\in[\underline{u},\,\overline{u}]}|\mu(\eta,t)-\mu(\eta_j,t)|<\epsilon.\end{aligned}$$

and then use the triangle inequality. To see that the claim holds, it suffices, thanks to Lemma 7.3, to exhibit a finite and tight enough grid of A such that any point in A lies close enough to a point of the grid. The existence of such a grid is obviously guaranteed since $A \subset \mathbb{R}^2$ is bounded. □

Lemma 7.5. *Recall the definition of b_n given in (6.2). Let*

$$b(\theta) = \left(\frac{\sigma_0^2}{\sigma^2} - 1 - \log\frac{\sigma_0^2}{\sigma^2}\right) + \frac{1}{\sigma^2}\int_{\underline{u}}^{\overline{u}}\left[\mu(\eta_0,t)-\mu(\eta,t)\right]^2 f(t)\,\mathrm{d}t. \tag{7.6}$$

Then, under Assumptions (A1)–(A4),

$$b_n(\theta) \geqslant 0. \tag{7.7}$$

$$b(\theta) \geqslant 0, \text{ with equality if and only if } \theta = \theta_0. \tag{7.8}$$

$$b_n(\theta') \to b_n(\theta), \text{ uniformly in } n, \text{ as } \theta' \to \theta. \tag{7.9}$$

$$b_n(\theta) \to b(\theta), \text{ as } n \to +\infty. \tag{7.10}$$

Proof of Lemma 7.5. We will prove each claim separately.

Proof of (7.7). That $b_n(\theta) \geqslant 0$ is trivial since the first term in (6.2) is non negative (having $x - 1 - \log x \geqslant 0$ with equality only if $x = 1$), and the second term in (6.2) is obviously non negative too.

Proof of (7.8). That $b(\theta) \geqslant 0$ is again easy enough to prove, both terms in (7.6) being trivially non negative. If $\theta \neq \theta_0$ then either $\sigma^2 \neq \sigma_0^2$ which implies the first term is positive, or $\mu(\eta_0, \cdot) \neq \mu(\eta, \cdot)$ which implies the second term is positive (since f is assumed positive on $[\underline{u}, \overline{u}]$). Hence if $\theta \neq \theta_0$ then $b(\theta) > 0$. That $\theta = \theta_0$ implies $b(\theta) = 0$ is of course straightforward.

Proof of (7.9). We first observe that

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{i=1}^{n} (\mu(\eta_0, t_i) - \mu(\eta', t_i))^2 - \frac{1}{n} \sum_{i=1}^{n} (\mu(\eta_0, t_i) - \mu(\eta, t_i))^2 \right| \\
&= \left| \frac{1}{n} \sum_{i=1}^{n} [2\mu(\eta_0, t_i) - \mu(\eta', t_i) - \mu(\eta, t_i)] \cdot [\mu(\eta, t_i) - \mu(\eta', t_i)] \right| \\
&\leqslant \frac{1}{n} \sum_{i=1}^{n} |2\mu(\eta_0, t_i) - \mu(\eta', t_i) - \mu(\eta, t_i)| \cdot |\mu(\eta, t_i) - \mu(\eta', t_i)| \\
&\leqslant \left(\sup_{t \in [\underline{u}, \overline{u}]} |\mu(\eta_0, t) - \mu(\eta, t)| + \sup_{t \in [\underline{u}, \overline{u}]} |\mu(\eta', t) - \mu(\eta, t)| \right) \cdot \sup_{t \in [\underline{u}, \overline{u}]} |\mu(\eta', t) - \mu(\eta, t)| \,.
\end{aligned} \tag{7.11}$$

As $\theta' \to \theta$, the convergence of the first term of b_n to the first term of b is obviously uniform in n since this part of b_n does not involve n at all. As $\theta' \to \theta$, via Lemma 7.3, we also obtain

$$\sup_{t \in [\underline{u}, \overline{u}]} |\mu(\eta', t) - \mu(\eta, t)| \to 0,$$

which ensures that the second part of (6.2) converges uniformly in n thanks to (7.11).

Proof of (7.10). Thanks to Assumption (A1), it is easy to see that

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} [\mu(\eta_0, t_i) - \mu(\eta, t_i)]^2 &= \int_{\underline{u}}^{\overline{u}} [\mu(\eta_0, t) - \mu(\eta, t)]^2 \, \mathrm{d}F_n(t) \\
&\to \int_{\underline{u}}^{\overline{u}} [\mu(\eta_0, t) - \mu(\eta, t)]^2 \, \mathrm{d}F(t) = \int_{\underline{u}}^{\overline{u}} [\mu(\eta_0, t) - \mu(\eta, t)]^2 f(t) \, \mathrm{d}t.
\end{aligned}$$

□

Lemma 7.6. *Let $A \subset \mathbb{R} \times [\underline{u}, \overline{u}]$ be a bounded set, and let $\eta_0 \in A$, then under Assumptions (A1)–(A4),*

$$\sup_{\eta \in A} \left| \frac{1}{n} \sum_{i=1}^{n} [\mu(\eta_0, t_i) - \mu(\eta, t_i)] \xi_i \right| \xrightarrow{a.s.} 0.$$

Proof of Lemma 7.6. Let $\epsilon > 0$, $\eta \in A$, and apply Lemma 7.4 to get the corresponding $m(\epsilon) \in \mathbb{N}$, $\{\eta_1, \dots, \eta_{m(\epsilon)}\} \subset A$, $j, j' \in \{1, \dots, m(\epsilon)\}$. We can write with the triangle inequality

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_0,t_i)-\mu(\eta,t_i)]\xi_i &= \frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{[\mu(\eta_0,t_i)-\mu(\eta,t_i)]-[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\right\}\xi_i \\ &\leqslant \sup_{(j,j')\in\{1,\dots,m(\epsilon)\}}\left\{\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i\right\} + \epsilon\cdot\frac{1}{n}\sum_{i=1}^{n}|\xi_i|\,.\end{aligned}$$

Hence

$$\begin{aligned}\sup_{\eta\in A}\left\{\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_0,t_i)-\mu(\eta,t_i)]\xi_i\right\} &\leqslant \sup_{(j,j')\in\{1,\dots,m(\epsilon)\}}\left\{\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i\right\} \\ &\quad + \epsilon\cdot\frac{1}{n}\sum_{i=1}^{n}|\xi_i|\,. \qquad (7.12)\end{aligned}$$

Let us now recall Kolmogorov's criterion, a proof of which is available in Section 17 of Loève (1991) on pages 250–251. This criterion guarantees that for any sequence $(Y_i)_{i\in\mathbb{N}}$ of independent random variables and any numerical sequence $(b_i)_{i\in\mathbb{N}}$ such that

$$\sum_{i=1}^{+\infty}\frac{\operatorname{Var} Y_i}{b_i^2} < +\infty,\ b_n \to +\infty,$$

we have

$$\frac{\sum_{i=1}^{n}(Y_i - \mathbb{E}Y_i)}{b_n} \xrightarrow{a.s.} 0.$$

For each couple $(j, j') \in \{1, \dots, m(\epsilon)\}$, Kolmogorov's criterion ensures that

$$\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i \xrightarrow{a.s.} 0,$$

for the coefficients $[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]$ are obviously bounded, and it suffices to pick $Y_i = [\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i$ and $b_i = i$. Having only a finite number of couples $(j, j') \in \{1, \dots, m(\epsilon)\}^2$ to consider allows us to write

$$\sup_{(j,j')\in\{1,\dots,m(\epsilon)\}}\frac{1}{n}\sum_{i=1}^{n}[\mu(\eta_j,t_i)-\mu(\eta_{j'},t_i)]\xi_i \xrightarrow{a.s.} 0. \qquad (7.13)$$

By (7.13), the first term on the right hand side of (7.12) converges almost surely to zero. The Strong Law of Large Numbers ensures that the second term on the right hand side of (7.12) converges almost surely to $\epsilon\cdot(2\pi^{-1}\sigma^2)^{\frac{1}{2}}$, and the result follows, since all the work done above for $(\xi_n)_{n\in\mathbb{N}}$ can be done again for $(-\xi_n)_{n\in\mathbb{N}}$. □

Lemma 7.7. *Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of independent identically distributed random variables such that for all $i \in \mathbb{N}$, either $Z_i \sim \mathcal{N}(0, \sigma^2)$ with $\sigma^2 > 0$, or $Z_i \sim \chi^2(k)$ with $k > 0$. Then a.s., as $n \to +\infty$*

$$Z_n = \mathrm{O}(\log n).$$

Proof of Lemma 7.7. Denote $Y_n = Z_n$ when the random variables are Gaussian, and $Y_n = Z_n/5$ when the random variables considered are chi-squared (so that $\mathbb{E}e^{2Y_1}$ and $\mathbb{E}e^{-2Y_1}$ are both finite). We will show that a.s. $Y_n = \mathrm{O}(\log n)$.

For any $\epsilon > 0$, from Markov's inequality we get:

$$\mathbb{P}\left(n^{-1}|e^{Y_n}| > \epsilon\right) = \mathbb{P}\left(n^{-2}e^{2Y_n} > \epsilon^2\right) \leqslant \epsilon^{-2}n^{-2}\mathbb{E}e^{2Y_1}.$$

From there it is easy to see that for any $\epsilon > 0$ we have

$$\sum_{n=1}^{+\infty} \mathbb{P}\left(n^{-1}|e^{Y_n}| > \epsilon\right) = \epsilon^{-2}\frac{\pi^2}{6}\mathbb{E}e^{2Y_1} < \infty,$$

which directly implies via Borel-Cantelli's Lemma (see for example Billingsley, 1995, Section 4, page 59) that a.s.

$$e^{Y_n} = \mathrm{o}(n).$$

In particular, a.s. for any n large enough,

$$Y_n \leqslant \log n.$$

What was done with $(Y_n)_{n\in\mathbb{N}}$ can be done again with $(-Y_n)_{n\in\mathbb{N}}$ so that in the end we have a.s for any n large enough,

$$-\log n \leqslant Y_n \leqslant \log n.$$

□

Lemma 7.8. *Under Assumptions (A1)–(A4), for any $\eta_0 \in \mathbb{R} \times [\underline{u}, \overline{u}]$, there exists $C \in \mathbb{R}_+^*$ such that for any n large enough, and for any η*

$$n^{-1}\sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \geqslant C\|\eta - \eta_0\|^2.$$

Proof of Lemma 7.8. We have already almost proved this result in (3.3) (see Theorem 3.3). There is however a small difficulty since the majoration was obtained for $\tau = (\beta, \gamma)$ and not $\eta = (\gamma, u)$.

Let V_1 and V_2 two non empty open intervals of $]\underline{u}, u_0[$ such that their closures $\overline{V_1}$ and $\overline{V_2}$ are do not overlap. We have

$$n^{-1}\sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \geqslant n^{-1}\left(\sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \mathbb{1}_{V_1}(t_i) + \sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \mathbb{1}_{V_2}(t_i)\right).$$

Using the same arguments we used to prove (3.3), we find that there exists $C \in \mathbb{R}_+^*$ such that (remembering the definition of the intercept β of the model)

$$n^{-1}\sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \geqslant \min\left(n^{-1}\sum_{i=1}^{n} \mathbb{1}_{V_1}(t_i), n^{-1}\sum_{i=1}^{n} \mathbb{1}_{V_2}(t_i)\right) \cdot C|\gamma - \gamma_0|^2,$$

$$n^{-1}\sum_{i=1}^{n} \left[\mu(\eta_0, t_i) - \mu(\eta, t_i)\right]^2 \geqslant \min\left(n^{-1}\sum_{i=1}^{n} \mathbb{1}_{V_1}(t_i), n^{-1}\sum_{i=1}^{n} \mathbb{1}_{V_2}(t_i)\right) \cdot C|\beta - \beta_0|^2,$$

and since for $j = 1, 2$ we have

$$n^{-1}\sum_{i=1}^{n} \mathbb{1}_{V_j}(t_i) \to \int_{V_j} f(t)\,\mathrm{d}t > 0,$$

there exists $C \in \mathbb{R}_+^*$ such that for any n large enough

$$n^{-1}\sum_{i=1}^{n}\left[\mu(\eta_0,t_i)-\mu(\eta,t_i)\right]^2 \geqslant C|\gamma-\gamma_0|^2,$$
$$n^{-1}\sum_{i=1}^{n}\left[\mu(\eta_0,t_i)-\mu(\eta,t_i)\right]^2 \geqslant C|\beta-\beta_0|^2.$$

Notice now that

$$\begin{aligned}
|u-u_0| &= |\gamma_0^{-1}\beta_0-\gamma^{-1}\beta| \\
&= |\gamma_0^{-1}||\beta_0-\gamma_0\gamma^{-1}\beta| \\
&\leqslant |\gamma_0^{-1}|\left\{|\beta_0-\beta|+|\beta-\gamma_0\gamma^{-1}\beta|\right\} \\
&\leqslant |\gamma_0^{-1}|\left\{|\beta_0-\beta|+|\gamma^{-1}\beta|\,|\gamma-\gamma_0|\right\} \\
&\leqslant |\gamma_0^{-1}|(|\beta_0-\beta|+|u|\,|\gamma-\gamma_0|) \\
&\leqslant |\gamma_0^{-1}|(1+\max(|\underline{u}|,|\overline{u}|))\cdot\max(|\beta_0-\beta|,|\gamma-\gamma_0|).
\end{aligned}$$

From here, since $u \in [\underline{u},\overline{u}]$ is bounded, it is straightforward that there exists $C \in \mathbb{R}_+^*$ such that for any n large enough

$$n^{-1}\sum_{i=1}^{n}\left[\mu(\eta_0,t_i)-\mu(\eta,t_i)\right]^2 \geqslant C|\gamma-\gamma_0|^2,$$
$$n^{-1}\sum_{i=1}^{n}\left[\mu(\eta_0,t_i)-\mu(\eta,t_i)\right]^2 \geqslant C|u-u_0|^2,$$

which ends the proof. □

Lemma 7.9. *Recall the definition of $A^*_{1:n}$ given in (6.39). Under Assumptions (A1)–(A4) and conditions (2.6), as $n \to +\infty$*

$$n^{-\frac{1}{2}}A^*_{1:n}(\theta_0) \xrightarrow{d} \mathcal{N}\left(0, I(\theta_0)\right). \tag{7.14}$$

Proof of Lemma 7.9. We will show that any linear combination of the coordinates of $A_{1:n}(\theta_0)$ is asymptotically normal using Lyapounov's Theorem. Let $\alpha \in \mathbb{R}^3$, $\|\alpha\| \neq 0$, so that differential calculus allows us to write

$$\begin{aligned}
\langle \alpha, A^*_{1:n}(\theta_0)\rangle &= \alpha_1 \cdot \left.\frac{\partial l^*_{1:n}(X_{1:n}|\theta)}{\partial\gamma}\right|_{\theta_0} + \alpha_2 \cdot \left.\frac{\partial l^*_{1:n}(X_{1:n}|\theta)}{\partial u}\right|_{\theta_0} + \alpha_3 \cdot \left.\frac{\partial l^*_{1:n}(X_{1:n}|\theta)}{\partial\sigma^2}\right|_{\theta_0} \\
&= \alpha_1 \cdot \frac{1}{\sigma_0^2}\sum_{i=1}^{n^*}\left[(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\xi_i\right] - \alpha_2\cdot\frac{\gamma_0}{\sigma_0^2}\sum_{i=1}^{n^*}\left[\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\xi_i\right] \\
&\quad + \alpha_3\cdot\frac{1}{2\sigma_0^2}\sum_{i=1}^{n^*}\left[\frac{1}{\sigma_0^2}\cdot\xi_i^2-1\right] \\
&= \sigma_0^{-2}\sum_{i=1}^{n^*} Z_i,
\end{aligned}$$

where we denote, for $i = 1,\ldots,n^*$

$$Z_i = \left[\left\{(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_1 - \gamma_0\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_2\right\}\cdot\xi_i + \frac{1}{2}\alpha_3\cdot\left\{\sigma_0^{-2}\cdot\xi_i^2-1\right\}\right]. \tag{7.15}$$

Since for $i = 1, \dots, n^*$ $\mathbb{E}[\xi_i] = 0$ and $\mathbb{E}[\xi_i^2] = \sigma^2$, we deduce that $\mathbb{E}[Z_i] = 0$, and hence that $\mathbb{E}[\langle \alpha, A^*_{1:n}(\theta_0)\rangle] = 0$.

Let us now find the expression of $\operatorname{Var}\langle \alpha, A^*_{1:n}(\theta_0)\rangle$. Because ξ_i and ξ_j are independent when $i \neq j$, so are Z_i and Z_j and we hence write

$$\begin{aligned}
\operatorname{Var}\langle \alpha, A^*_{1:n}(\theta_0)\rangle &= \sigma_0^{-4} \sum_{i=1}^{n^*} \operatorname{Var} Z_i \\
&= \sigma_0^{-4} \sum_{i=1}^{n^*} \left\{ \left[(t_i - u_0) \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_1 - \gamma_0 \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_2 \right]^2 \cdot \operatorname{Var} \xi_i \right. \\
&\qquad \left. + \frac{1}{4} \alpha_3^2 \cdot \operatorname{Var}\left[\sigma_0^{-2} \xi_i^2 - 1 \right] \right\},
\end{aligned}$$

because $\operatorname{Cov}\left[\xi_i, \left\{\sigma_0^{-2}\xi_i^2 - 1\right\}\right] = 0$, and we finally get

$$\operatorname{Var}\langle \alpha, A^*_{1:n}(\theta_0)\rangle = \sigma_0^{-4} \sum_{i=1}^{n^*} \left\{ \left[(t_i - u_0) \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_1 - \gamma_0 \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_2 \right]^2 \cdot \sigma_0^2 + \frac{1}{4} \alpha_3^2 \cdot 2 \right\}.$$

We can hence write

$$\begin{aligned}
n^{*-1} \operatorname{Var}\langle \alpha, A^*_{1:n}(\theta_0)\rangle &= \sigma_0^{-2} \frac{1}{n^*} \sum_{i=1}^{n^*} \left[(t_i - u_0) \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_1 - \gamma_0 \mathbb{1}_{[t_i, +\infty[}(u_0) \cdot \alpha_2 \right]^2 + \frac{1}{2} \sigma_0^{-4} \alpha_3^2 \\
&= \alpha_1^2 \cdot \sigma_0^{-2} \left\{ \frac{1}{n^*} \sum_{i=1}^{n^*} (t_i - u_0)^2 \mathbb{1}_{[t_i, +\infty[}(u_0) \right\} \\
&\quad - 2\alpha_1 \alpha_2 \cdot \sigma_0^{-2} \gamma_0 \left\{ \frac{1}{n^*} \sum_{i=1}^{n^*} (t_i - u_0) \mathbb{1}_{[t_i, +\infty[}(u_0) \right\} \\
&\quad + \alpha_2^2 \cdot \sigma_0^{-2} \gamma_0^2 \left\{ \frac{1}{n^*} \sum_{i=1}^{n^*} \mathbb{1}_{[t_i, +\infty[}(u_0) \right\} + \alpha_3^2 \cdot \frac{1}{2} \sigma_0^{-4} \\
&= \langle \alpha, I_{1:n}(\theta_0) \alpha \rangle,
\end{aligned}$$

where we denote

$$I^*_{1:n}(\theta) = \begin{bmatrix}
\sigma^{-2} \dfrac{1}{n^*} \displaystyle\sum_{i=1}^{n^*} (t_i - u)^2 \mathbb{1}_{[t_i, +\infty[}(u) & -\sigma^{-2}\gamma \dfrac{1}{n^*} \displaystyle\sum_{i=1}^{n^*} (t_i - u) \mathbb{1}_{[t_i, +\infty[}(u) & 0 \\
 & \sigma^{-2}\gamma^2 \dfrac{1}{n^*} \displaystyle\sum_{i=1}^{n^*} \mathbb{1}_{[t_i, +\infty[}(u) & 0 \\
 & & \dfrac{1}{2}\sigma^{-4}
\end{bmatrix}. \quad (7.16)$$

Remark that, by virtue of Assumption (A1), it is easy to check that for any $\theta \in \Theta$

$$I^*_{1:n}(\theta) \to I(\theta), \quad (7.17)$$

and observe that just like $I(\theta)$, $I^*_{1:n}(\theta)$ is positive definite, since all its principal minor determinants are positive.

Let us now check that the random variables Z_i meet Lyapounov's Theorem (see Billingsley, 1995, page 362) requirements before wrapping up this proof. The random variables Z_i are independent and trivially L^2. We denote $V_n^{*2} = \sum_{i=1}^{n^*} \operatorname{Var} Z_i$ and claim that Lyapounov's condition

holds, that is

$$\exists \delta > 0,\ \sum_{i=1}^{n^*} \mathbb{E}\left|\frac{Z_i - \mathbb{E}Z_i}{V_n^*}\right|^{2+\delta} = \mathrm{o}(1).$$

Indeed we have ($\delta = 1$)

$$\begin{aligned}
\sum_{i=1}^{n^*} \mathbb{E}\left|\frac{Z_i - \mathbb{E}Z_i}{V_n^*}\right|^3 &= \sum_{i=1}^{n^*} \mathbb{E}\left|\frac{Z_i}{V_n^*}\right|^3 \\
&= \frac{n^*}{\mathrm{Var}^{\frac{3}{2}}\langle \alpha, A^*_{1:n}(\theta_0)\rangle} \cdot \frac{1}{n^*} \sum_{i=1}^{n^*} \mathbb{E}\left|Z_i\right|^3 \\
&= \frac{1}{{n^*}^{\frac{1}{2}} \langle \alpha, I^*_{1:n}(\theta_0)\alpha\rangle^{\frac{3}{2}}} \cdot \frac{1}{n^*} \sum_{i=1}^{n^*} \mathbb{E}\left|Z_i\right|^3.
\end{aligned}$$

The first term of this last product is $\mathrm{O}\left(n^{*-\frac{1}{2}}\right)$ thanks to (7.17), and recalling the definition of Z_i from (7.15), there is no difficulty in showing that the last term of the product, namely $\frac{1}{n^*}\sum_{i=1}^{n^*} \mathbb{E}\left|Z_i\right|^3$ converges to a finite limit. Indeed we find, using trivial dominations and Assumption (A1) once again,

$$\begin{aligned}
|Z_i|^3 &= \left|\left\{(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_1 - \gamma_0 \mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_2\right\}\cdot\xi_i + \frac{1}{2}\alpha_3 \cdot \left\{\sigma_0^{-2}\cdot\xi_i^2 - 1\right\}\right|^3 \\
\mathbb{E}|Z_i|^3 &\leqslant \left(\left|(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_1 - \gamma_0 \mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_2\right| + \left|\frac{1}{2}\alpha_3\right|\right)^3 \\
&\quad \times \mathbb{E}\left(|\xi_i| + \left|\sigma_0^{-2}\cdot\xi_i^2 - 1\right|\right)^3 \\
\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|Z_i|^3 &\leqslant \frac{1}{n}\sum_{i=1}^{n}\left(\left|(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_1 - \gamma_0 \mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_2\right| + \left|\frac{1}{2}\alpha_3\right|\right)^3 \\
&\quad \times \mathbb{E}\left(|\xi_i| + \left|\sigma_0^{-2}\cdot\xi_i^2 - 1\right|\right)^3 \\
&\leqslant \mathrm{O}(1)\cdot\frac{1}{n}\sum_{i=1}^{n}\left(\left|(t_i - u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_1 - \gamma_0 \mathbb{1}_{[t_i,+\infty[}(u_0)\cdot\alpha_2\right| + \left|\frac{1}{2}\alpha_3\right|\right)^3 \\
&\leqslant \mathrm{O}(1).
\end{aligned}$$

Lyapounov's Theorem thus applies here and leads to

$$\sum_{i=1}^{n^*} \frac{Z_i - \mathbb{E}Z_i}{V_n^*} \xrightarrow{d} \mathcal{N}(0,1),$$

i.e. multiplying numerator and denominator by σ_0^{-2} we get

$$\frac{\langle \alpha, A^*_{1:n}(\theta_0)\rangle}{\mathrm{Var}^{\frac{1}{2}}\langle \alpha, A^*_{1:n}(\theta_0)\rangle} \xrightarrow{d} \mathcal{N}(0,1),$$

that is

$$\frac{\langle \alpha, A^*_{1:n}(\theta_0)\rangle}{{n^*}^{\frac{1}{2}}\langle \alpha, I^*_{1:n}(\theta_0)\alpha\rangle^{\frac{1}{2}}} \xrightarrow{d} \mathcal{N}(0,1),$$

and because of (7.17) we can also write,

$$\frac{\langle \alpha, A^*_{1:n}(\theta_0)\rangle}{n^{*\frac{1}{2}}\langle \alpha, I(\theta_0)\alpha\rangle^{\frac{1}{2}}} \xrightarrow{d} \mathcal{N}(0,1),$$

which, remembering that a.s. $n^* \sim n$, is equivalent to (7.14). □

Lemma 7.10. *Recall the definition of $B^*_{1:n}$ given in (6.14). Under Assumptions (A1)–(A4) and conditions (2.6), as $n \to +\infty$,*

$$\frac{1}{n}B^*_{1:n}(\theta_0) \xrightarrow{a.s.} I(\theta_0), \text{ as } n \to +\infty. \tag{7.18}$$

$$\frac{1}{n}B^*_{1:n}(\theta) \xrightarrow{a.s.} I(\theta_0), \text{ as } \theta \to \theta_0 \text{ and } n \to +\infty. \tag{7.19}$$

where the asymptotic Fisher Information Matrix $I(\cdot)$ is defined in (2.3).

Proof of Lemma 7.10. We will prove each claim separately.

Proof of (7.18). Differential calculus provides the following expressions for the coefficients of $\frac{1}{n^*}B^*_{1:n}(\theta)$.

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{11} = \sigma^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u)^2\mathbb{1}_{[t_i,+\infty[}(u),$$

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{12} = \sigma^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)-2\gamma\cdot(t_i-u)\right]\mathbb{1}_{[t_i,+\infty[}(u),$$

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{13} = \sigma^{-4}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\right](t_i-u)\mathbb{1}_{[t_i,+\infty[}(u),$$

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{22} = \sigma^{-2}\gamma^2\frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u),$$

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{23} = -\sigma^{-4}\gamma\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\right]\mathbb{1}_{[t_i,+\infty[}(u),$$

$$\left(\frac{1}{n^*}B^*_{1:n}(\theta)\right)_{33} = -\frac{1}{2}\sigma^{-4}+\sigma^{-6}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right]^2.$$

The convergence we claim is then a direct consequence of Assumption (A1) and the fact that $n^* \sim n$ and, depending on the coefficients, either the Strong Law of Large Numbers or Kolmogorov's criterion. Notice that

$$\frac{1}{n^*}B^*_{1:n}(\theta_0) - I^*_{1:n}(\theta_0) \xrightarrow{a.s.} 0,$$

where $I^*_{1:n}$ is defined in (7.16).

Proof of (7.19). We will show that in fact, as $n \to +\infty$ and $\theta \to \theta_0$,

$$C^*_{1:n}(\theta) = \frac{1}{n^*}B^*_{1:n}(\theta_0) - \frac{1}{n^*}B^*_{1:n}(\theta) \xrightarrow{a.s.} 0,$$

which will end the proof since $n^* \sim n$. We will consider each coefficient of $C^*_{1:n}(\theta)$ in turn, making use of Assumption (A1) once again and apply repeatedly the Strong Law of Large Numbers and Kolmogorov's criterion as well as Lemma 7.2, whenever needed.

$$
\begin{aligned}
C^*_{1:n}(\theta)_{11} &= \sigma_0^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u_0)^2\mathbb{1}_{[t_i,+\infty[}(u_0) - \sigma^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u)^2\mathbb{1}_{[t_i,+\infty[}(u)\\
&= \left(\sigma_0^{-2}-\sigma^{-2}\right)\cdot\frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u_0)^2\mathbb{1}_{[t_i,+\infty[}(u_0)\\
&\quad + \sigma^{-2}\cdot\left(\frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u_0)^2\mathbb{1}_{[t_i,+\infty[}(u_0) - \frac{1}{n^*}\sum_{i=1}^{n^*}(t_i-u)^2\mathbb{1}_{[t_i,+\infty[}(u)\right)\\
&= \mathrm{o}(1)\cdot\mathrm{O}(1) + \mathrm{O}(1)\cdot\mathrm{O}\left(u-u_0\right) \to 0.
\end{aligned}
$$

then last equality holding true because of Lemma 7.2.

$$
\begin{aligned}
C^*_{1:n}(\theta)_{22} &= \sigma_0^{-2}\gamma_0^2\frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u_0) - \sigma^{-2}\gamma^2\frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u)\\
&= \left(\sigma_0^{-2}\gamma_0^2-\sigma^{-2}\gamma^2\right)\cdot\frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u_0)\\
&\quad + \sigma^{-2}\gamma^2\cdot\left[\frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u_0) - \frac{1}{n^*}\sum_{i=1}^{n^*}\mathbb{1}_{[t_i,+\infty[}(u)\right]\\
&= \mathrm{o}(1)\cdot\mathrm{O}(1) + \mathrm{O}(1)\cdot\left[\{F_{n^*}(u_0)-F(u_0)\}+\{F(u_0)-F(u)\}+\{F(u)-F_{n^*}(u)\}\right]\\
&= \mathrm{o}(1) + \mathrm{O}(1)\cdot\left[\mathrm{o}(1)+\mathrm{o}(1)+\mathrm{o}(1)\right] \to 0,
\end{aligned}
$$

the last equality holding true because of the uniform convergence of F_{n^*} to F over any compact subset such as $[\underline{u}, \overline{u}]$ (see Assumption (A1), and its Remark 1).

$$
\begin{aligned}
C^*_{1:n}(\theta)_{33} &= \frac{1}{2}\sigma^{-4} - \sigma^{-6}\frac{1}{n^*}\sum_{i=1}^{n^*}[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)]^2\\
&\quad - \left(\frac{1}{2}\sigma_0^{-4} - \sigma_0^{-6}\frac{1}{n^*}\sum_{i=1}^{n^*}\xi_i^2\right)\\
&= \frac{1}{2}(\sigma^{-4}-\sigma_0^{-4}) - (\sigma^{-6}-\sigma_0^{-6})\cdot\frac{1}{n^*}\sum_{i=1}^{n^*}\xi_i^2\\
&\quad - \sigma^{-6}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right]\xi_i\\
&\quad - \sigma^{-6}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right]^2\\
&= \mathrm{o}(1) + \mathrm{o}(1)\cdot\frac{1}{n^*}\sum_{i=1}^{n^*}\xi_i^2 + \mathrm{o}(1) + \mathrm{o}(1) \xrightarrow{a.s.} 0,
\end{aligned}
$$

where the two last o(1) are direct consequences of Lemmas 7.3 and 7.6.

Those same Lemmas used together with Lemma 7.2, the Strong Law of Large Numbers as well as the well-known Cauchy-Schwarz inequality imply that a.s.

$$\begin{aligned}
C^*_{1:n}(\theta)_{23} &= \sigma^{-4}\gamma\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i + \gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\right]\mathbb{1}_{[t_i,+\infty[}(u) \\
&\quad - \sigma_0^{-4}\gamma_0\frac{1}{n^*}\sum_{i=1}^{n^*}\xi_i\mathbb{1}_{[t_i,+\infty[}(u_0) \\
&= \frac{1}{n^*}\sum_{i=1}^{n^*}\left[\sigma^{-4}\gamma\mathbb{1}_{[t_i,+\infty[}(u) - \sigma_0^{-4}\gamma_0\mathbb{1}_{[t_i,+\infty[}(u_0)\right]\xi_i \\
&\quad + \frac{1}{n^*}\sum_{i=1}^{n^*}\left[\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\right]\sigma^{-4}\gamma\mathbb{1}_{[t_i,+\infty[}(u) \\
&= \mathrm{o}(1) + \mathrm{o}(1) \xrightarrow{a.s.} 0,
\end{aligned}$$

and also that a.s.

$$\begin{aligned}
C^*_{1:n}(\theta)_{13} &= \sigma_0^{-4}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\right]\xi_i \\
&\quad - \sigma^{-4}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i + \gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\right](t_i-u)\mathbb{1}_{[t_i,+\infty[}(u) \\
&= \frac{1}{n^*}\sum_{i=1}^{n^*}\left[\sigma_0^{-4}(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \sigma^{-4}(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u)\right]\xi_i \\
&\quad - \sigma^{-4}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \gamma\cdot(t_i-u)\right](t_i-u)\mathbb{1}_{[t_i,+\infty[}(u) \\
&= \mathrm{o}(1) + \mathrm{o}(1) \xrightarrow{a.s.} 0.
\end{aligned}$$

and finally that a.s.

$$\begin{aligned}
C^*_{1:n}(\theta)_{12} &= \sigma_0^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i - \gamma_0\cdot(t_i-u_0)\right]\mathbb{1}_{[t_i,+\infty[}(u_0) \\
&\quad - \sigma^{-2}\frac{1}{n^*}\sum_{i=1}^{n^*}\left[\xi_i + \gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - 2\gamma\cdot(t_i-u)\right]\mathbb{1}_{[t_i,+\infty[}(u) \\
&= \frac{1}{n^*}\sum_{i=1}^{n^*}\xi_i\cdot\left[\sigma_0^{-2}\mathbb{1}_{[t_i,+\infty[}(u_0) - \sigma^{-2}\mathbb{1}_{[t_i,+\infty[}(u)\right] \\
&\quad + \frac{1}{n^*}\sum_{i=1}^{n^*}\left[-\sigma_0^{-2}\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0) - \sigma^{-2}(\gamma_0\cdot(t_i-u_0)\mathbb{1}_{[t_i,+\infty[}(u_0)\right. \\
&\qquad\qquad \left. -2\gamma\cdot(t_i-u)\mathbb{1}_{[t_i,+\infty[}(u))\right] \\
&= \mathrm{o}(1) + \mathrm{o}(1) \xrightarrow{a.s.} 0.
\end{aligned}$$

□

Proposition 7.11. *Let $0 < \delta$, and let $(\rho_n)_{n\in\mathbb{N}}$ be a positive sequence such that, as $n \to +\infty$*

$$\rho_n = \mathrm{O}(1) \tag{7.20}$$

$$n^{-\frac{1}{2}}(\log n)\cdot\rho_n^{-1} \to 0 \tag{7.21}$$

and denote

$$B^c(\theta_0, \delta\rho_n) = \{\theta \in \Theta,\ \|\theta - \theta_0\| \geqslant \delta\rho_n\},$$

Then, under Assumptions (A1)–(A4), a.s., there exists $\epsilon > 0$ such that, for any n large enough

$$\sup_{\theta \in B^c(\theta_0, \delta\rho_n)} \frac{1}{n\rho_n^2}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\widehat{\theta}_n)] \leqslant -\epsilon. \tag{7.22}$$

$$\sup_{\theta \in B^c(\theta_0, \delta\rho_n)} \frac{1}{n\rho_n^2}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] \leqslant -\epsilon. \tag{7.23}$$

Proof of Proposition 7.11. This proposition is to be compared to the regularity condition imposed in Ghosh et al. (2006) (see their condition (A4) in Chapter 4). The aim of this proposition is to show that our model satisfies to a somewhat stronger version of that condition.

Let $0 < \delta$. Notice first that, similarly to what was done in (6.20), we are able to deduce that a.s.

$$\frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\widehat{\theta}_n)] \leqslant \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] =: i_n(\theta). \tag{7.24}$$

where i_n is defined over $\mathbb{R} \times [\underline{u}, \overline{u}] \times \mathbb{R}_+^* \supset \Theta \supset B^c(\theta_0, \delta\rho_n)$ by

$$i_n(\theta) = \log\frac{\sigma_0^2}{\sigma^2} + 1 + \frac{1}{n\sigma_0^2}\sum_{i=1}^n(\xi_i^2 - \sigma_0^2) - \frac{1}{n\sigma^2}\sum_{i=1}^n[\xi_i + \mu(\eta_0, t_i) - \mu(\eta, t_i)]^2. \tag{7.25}$$

$$= \log\frac{\sigma_0^2}{\sigma^2} + 1 - \frac{\sigma_0^2}{\sigma^2} + \frac{1}{n\sigma_0^2}\sum_{i=1}^n(\xi_i^2 - \sigma_0^2) - \frac{1}{n\sigma^2}\sum_{i=1}^n\left\{[\xi_i + \mu(\eta_0, t_i) - \mu(\eta, t_i)]^2 - \sigma_0^2\right\}. \tag{7.26}$$

From (7.24) it is clear that we need only prove (7.23) to end the proof.

The rest of this proof is divided into 6 major steps. Step 1 shows that for a given n the supremum considered is reached on a point θ_n. Step 2 and 3 focus on obtaining useful majorations of the supremum. Step 4 is dedicated to proving that the sequence θ_n admits an accumulation point (the coordinates of which satisfy to some conditions), while step 5 makes use of this last fact to effectively dominate the supremum. Step 6 wraps up the proof.

Step 1. We first show that a.s. for any n there exists $\theta_n \in \mathbb{R}\times[\underline{u}, \overline{u}]\times\mathbb{R}_+^*$ such that $\|\theta_n - \theta_0\| \geqslant \delta\rho_n$ and

$$i_n(\theta_n) = \sup_{\Theta \in B^c(\theta_0, \delta\rho_n)} i_n(\theta). \tag{7.27}$$

Let $n \in \mathbb{N}$ and let $(\theta_{n,k})_{k\in\mathbb{N}}$ be a sequence of points in $B^c(\theta_0, \delta\rho_n)$ such that

$$\lim_{k\longrightarrow+\infty} i_n(\theta_{n,k}) = \sup_{\Theta \in B^c(\theta_0, \delta\rho_n)} i_n(\theta).$$

From (7.25) it is obvious that $\sigma_{n,k}^2$ is bounded: if it was not, we would be able to extract a subsequence such that σ_{n,k_j}^2 would go to $+\infty$ and thus $i_n(\theta_{n,k_j})$ would go to $-\infty$. For the very same reason, $\gamma_{n,k}$ too is bounded. Recalling that $u_{n,k}$ is bounded too by definition, we now see that there exists a subsequence $(\theta_{n,k_j})_{j\in\mathbb{N}}$ in $B^c(\theta_0, \delta\rho_n)$ and a point θ_n in $\overline{B^c(\theta_0, \delta\rho_n)}$ (i.e. in $\mathbb{R} \times [\underline{u}, \overline{u}] \times \mathbb{R}_+$, and such that $\|\theta_n - \theta_0\| \geqslant \delta\rho_n$) such that $(\theta_{n,k_j})_{j\in\mathbb{N}} \xrightarrow[j\longrightarrow+\infty]{} \theta_n$.

Finally from (7.25) again it is easy to see that $\sigma_n^2 > 0$ for if it was not $i_n(\theta_{n,k_j})$ would go to $-\infty$ once again, unless (by continuity of μ with regard to η) $\xi_i + \mu(\eta_0, t_i) - \mu(\eta_n, t_i) = 0$ for all $i \leqslant n$ which a.s. does not happen.

Step 2. From the previous step and the continuity of i_n with regard to θ we are able to write

$$\sup_{\Theta \in B^c(\theta_0, \delta\rho_n)} \frac{2}{n}[l_{1:n}(X_{1:n}|\theta) - l_{1:n}(X_{1:n}|\theta_0)] = i_n(\theta_n). \tag{7.28}$$

where $(\theta_n)_{n\in\mathbb{N}}$ is the sequence defined in Step 1. We now derive a convenient majoration of $i_n(\theta_n)$. Expanding from (7.26) we get

$$\begin{aligned} i_n(\theta_n) &= \left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)+\frac{1}{n\sigma_0^2}\sum_{i=1}^n(\xi_i^2-\sigma_0^2)-\frac{1}{n\sigma^2}\sum_{i=1}^n\left\{[\xi_i+\mu(\eta_0,t_i)-\mu(\eta,t_i)]^2-\sigma_0^2\right\} \\ &= \left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)+\frac{\sigma_n^2-\sigma_0^2}{n\sigma_0^2\sigma_n^2}\sum_{i=1}^n(\xi_i^2-\sigma_0^2)-\frac{1}{n\sigma_n^2}\sum_{i=1}^n[\mu(\eta_0,t_i)-\mu(\eta_n,t_i)]^2 \\ &\quad -\frac{2}{n\sigma_n^2}\sum_{i=1}^n[\mu(\eta_0,t_i)-\mu(\eta_n,t_i)]\xi_i \end{aligned}$$

Thanks to Lemma 7.8, we know that there exists $C_1\in\mathbb{R}_+^*$ such that

$$\begin{aligned} i_n(\theta_n) &\leqslant \left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)+\frac{\sigma_n^2-\sigma_0^2}{n\sigma_0^2\sigma_n^2}\sum_{i=1}^n(\xi_i^2-\sigma_0^2) \\ &\quad -\frac{1}{\sigma_n^2}C_1\|\eta_n-\eta_0\|^2-\frac{2}{n\sigma_n^2}\sum_{i=1}^n[\mu(\eta_0,t_i)-\mu(\eta_n,t_i)]\xi_i \end{aligned}$$

From there, the Law of the Iterated Logarithm and a factorisation of the last term together with Corollary 7.7 lead to:

$$\begin{aligned} i_n(\theta_n) &\leqslant \left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)+\frac{1}{\sigma_n^2}|\sigma_n^2-\sigma_0^2|R_{1,n}-\frac{1}{\sigma_n^2}C_1\|\eta_n-\eta_0\|^2 \\ &\quad +\frac{1}{n\sigma_n^2}\left(\sum_{i=1}^n[\mu(\eta_0,t_i)-\mu(\eta_n,t_i)]^2\right)^{\frac{1}{2}}R_{2,n} \end{aligned}$$

where a.s. $R_{1,n}=\mathrm{O}\left(n^{-\frac{1}{2}}(\log\log n)^{\frac{1}{2}}\right)$ and $R_{2,n}=\mathrm{O}\left(\log n\right)$. Lemma 7.3 ensures there exists $C_2\in\mathbb{R}_+^*$ such that

$$i_n(\theta_n)\leqslant\left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)+\frac{1}{\sigma_n^2}|\sigma_n^2-\sigma_0^2|R_{1,n}-\frac{1}{\sigma_n^2}C_1\|\eta_n-\eta_0\|^2+\frac{1}{n\sigma_n^2}C_2n^{\frac{1}{2}}\|\eta_n-\eta_0\|R_{2,n}$$

We thus deduce that there exists $C\in\mathbb{R}_+^*$ such that:

$$i_n(\theta_n)\leqslant\left(\log\frac{\sigma_0^2}{\sigma_n^2}+1-\frac{\sigma_0^2}{\sigma_n^2}\right)-\frac{1}{\sigma_n^2}C\|\eta_n-\eta_0\|^2+\frac{1}{\sigma_n^2}\|\theta_n-\theta_0\|R_n \tag{7.29}$$

where a.s. $R_n=\mathrm{O}\left(n^{-\frac{1}{2}}\log n\right)$. Notice in particular that, due to (7.21), $R_n=\mathrm{o}(\rho_n)$.

Step 3. We obtain two majorations, (7.31) and (7.32), that we will make use of in the coming steps. Using a conversion of $\theta=(\gamma,u,\sigma^2)$ into the spherical coordinate system we write θ_n as

$$\theta_n=(r_n\cos\psi_n\cos\phi_n,r_n\sin\psi_n\cos\phi_n,r_n\sin\phi_n),$$

where

$$(r_n,\psi_n,\phi_n)\in\mathbb{R}_+^*\times[0,\,2\pi]\times]0,\,\pi[,$$

and deduce from (7.29) that

$$\begin{aligned} i_n(\theta_n) &\leqslant \left(\log\frac{\sigma_0^2}{r_n\sin\phi_n}+1-\frac{\sigma_0^2}{r_n\sin\phi_n}\right)-Cr_n\frac{\cos^2\phi_n}{\sin\phi_n}+\frac{1}{\sin\phi_n}R_n \qquad (7.30)\\ &\leqslant \left(\log\frac{\sigma_0^2}{r_n\sin\phi_n}+1-\frac{\sigma_0^2}{r_n\sin\phi_n}\right)+\frac{1}{\sin\phi_n}\left[R_n-Cr_n\cos^2\phi_n\right]. \qquad (7.31) \end{aligned}$$

From (7.30) we also get the following majoration

$$i_n(\theta_n) \leqslant \left(\log \frac{\sigma_0^2}{r_n \sin \phi_n} + 1 - \frac{\sigma_0^2}{r_n \sin \phi_n} \right) + \frac{1}{\sin \phi_n} R_n. \tag{7.32}$$

Step 4. We show that the sequence $(\theta_n)_{n \in \mathbb{N}}$ we built, converges to a finite limit θ_∞ (extracting a subsequence if necessary). Extracting a subsequence if necessary, we can assume that $(\psi_n, \phi_n) \to (\psi_\infty, \phi_\infty) \in [0,\, 2\pi] \times [0,\, \pi]$. We consider the two following mutually exclusive situations.

Situation A: $\phi_\infty = 0 \mod \pi$. In this situation, there exists $\epsilon > 0$ such that for any n large enough,

$$\begin{aligned} \left[R_n - C r_n \cos^2 \phi_n \right] &= \left(\frac{R_n}{r_n} - C \cos^2 \phi_n \right) r_n \\ &\leqslant -\epsilon r_n, \end{aligned}$$

because a.s. $R_n = \mathrm{o}(r_n)$ (since $R_n = \mathrm{o}(\rho_n)$ and $r_n \leqslant \rho_n$). Used together with (7.31), this leads to

$$i_n(\theta_n) \leqslant \left(\frac{\sigma_0^2}{r_n \sin \phi_n} - 1 - \log \frac{\sigma_0^2}{r_n \sin \phi_n} \right) - \epsilon \frac{r_n}{\sin \phi_n},$$

for any n large enough and hence $i_n(\theta_n) \to -\infty$ whether r_n goes to zero or not.

Situation B: $\phi_\infty \neq 0 \mod \pi$. In this situation, from (7.32), we see that $r_n \to 0$ and $r_n \to +\infty$ both lead to $i_n(\theta_n) \to -\infty$.

Observing that $i_n(\theta)$ converges a.s. to a finite value for any $\theta \in \Theta$ as $n \to +\infty$, we see that $\lim_{n \to \infty} i_n(\theta_n) = -\infty$ is not possible by construction of the sequence θ_n, and deduce that, extracting a subsequence if necessary, there exists

$$(r_\infty, \psi_\infty, \phi_\infty) \in \mathbb{R}_+^* \times [0,\, 2\pi] \times]0,\, \pi[,$$

such that $\lim_{n \to +\infty} \theta_n = \theta_\infty$. Notice that in particular, $\sigma_\infty^2 > 0$.

Step 5. We will now end the proof by showing that there exists $\epsilon > 0$ such that for any n large enough

$$i_n(\theta_n) \leqslant -\epsilon \rho_n^2. \tag{7.33}$$

We consider the two following mutually exclusive situations.

Situation A: $\sigma_\infty^2 \neq \sigma_0^2$. In this situation, from (7.29) we get

$$i_n(\theta_n) \leqslant \left(\log \frac{\sigma_0^2}{\sigma_n^2} + 1 - \frac{\sigma_0^2}{\sigma_n^2} \right) + \frac{1}{\sigma_n^2} \left\| \theta_n - \theta_0 \right\| R_n$$

and the right-hand side converges to

$$\left(\log \frac{\sigma_0^2}{\sigma_\infty^2} + 1 - \frac{\sigma_0^2}{\sigma_\infty^2} \right) < 0.$$

There hence exists $\epsilon > 0$ such that for any n large enough

$$i_n(\theta_n) \leqslant -\epsilon.$$

Since $\rho_n = \mathrm{O}(1)$ by (7.20), (7.33) is a direct consequence of this.

Situation B: $\sigma_\infty^2 = \sigma_0^2$. In this situation, recalling that for any $x > 0$

$$\log x + 1 - x \leqslant -\frac{(x-1)^2}{2} + \frac{(x-1)^3}{3},$$

we deduce from (7.29) that for any n large enough

$$\begin{aligned} i_n(\theta_n) &\leqslant -\frac{1}{2}\left(\frac{\sigma_0^2}{\sigma_n^2}-1\right)^2+\frac{1}{3}\left(\frac{\sigma_0^2}{\sigma_n^2}-1\right)^3-\frac{1}{\sigma_n^2}C\|\eta_n-\eta_0\|^2+\frac{1}{\sigma_n^2}\|\theta_n-\theta_0\|R_n \\ &\leqslant \left(\frac{\sigma_0^2}{\sigma_n^2}-1\right)^2\left[\frac{1}{3}\left(\frac{\sigma_0^2}{\sigma_n^2}-1\right)-\frac{1}{2}\right]-\frac{1}{\sigma_n^2}C\|\eta_n-\eta_0\|^2+\frac{1}{\sigma_n^2}\|\theta_n-\theta_0\|R_n \\ &\leqslant -\frac{1}{4}\left(\frac{\sigma_0^2}{\sigma_n^2}-1\right)^2-\frac{1}{\sigma_n^2}C\|\eta_n-\eta_0\|^2+\frac{1}{\sigma_n^2}\|\theta_n-\theta_0\|R_n \\ &\leqslant \frac{1}{\sigma_n^2}\left\{-c\left[(\sigma_0^2-\sigma_n^2)^2-\|\eta_n-\eta_0\|^2\right]+\|\theta_n-\theta_0\|R_n\right\} \end{aligned}$$

where $c=\min(1/4,C)>0$. It follows that for any n large enough

$$\begin{aligned} i_n(\theta_n) &\leqslant \frac{1}{\sigma_n^2}\left(-c\|\theta_n-\theta_0\|^2+\|\theta_n-\theta_0\|R_n\right) \\ &\leqslant \frac{1}{\sigma_n^2}\|\theta_n-\theta_0\|\left(R_n-c\|\theta_n-\theta_0\|\right). \end{aligned}$$

Thus, for any n large enough

$$i_n(\theta_n)\leqslant \frac{1}{\sigma_n^2}\frac{\|\theta_n-\theta_0\|}{\rho_n}\left(\frac{R_n}{\rho_n}-c\frac{\|\theta_n-\theta_0\|}{\rho_n}\right)\rho_n^2.$$

Recalling that

$$\begin{aligned} \|\theta_n-\theta_0\| &\geqslant \delta\rho_n, \\ R_n &= \mathrm{o}(\rho_n), \\ \sigma_n^2 &\to \sigma_\infty^2>0 \end{aligned}$$

we obtain for any n large enough,

$$i_n(\theta_n)\leqslant \frac{1}{\sigma_n^2}\frac{\|\theta_n-\theta_0\|}{\rho_n}\left(-c\frac{\delta}{2}\right)\rho_n^2\leqslant -\frac{c\delta^2}{2\sigma_n^2}\rho_n^2\leqslant -\frac{c\delta^2}{3\sigma_\infty^2}\rho_n^2.$$

Hence (7.33) holds in this situation too: it suffices to take $\epsilon=\dfrac{c\delta^2}{3\sigma_\infty^2}$.

We just proved that (7.33) holds in both cases considered.

Step 6. (7.23) is a consequence of (7.28) and (7.33).

□

Lemma 7.12. *Let $0<\delta<1$ then under Assumptions (A1)–(A4) and conditions (2.6), a.s. there exists a constant $C\in R_+^*$ such that for any n large enough and for any $1\leqslant i_1,i_2,i_3\leqslant 3$*

$$\left|\frac{1}{n}\frac{\partial^3 l_{1:n}^*(X_{1:n}|\theta)}{\partial_{i_1}\partial_{i_2}\partial_{i_3}}\right|\leqslant C \tag{7.34}$$

for any $\theta\in B(\theta_0,\delta d_n)$.

Proof of Lemma 7.12. Let $0<\delta<1$. We will prove (7.34) stands true for any $1\leqslant i_1,i_2,i_3\leqslant 3$. First notice that for n large enough, $\theta\mapsto l_{1:n}^*(X_{1:n}|\theta)$ is indeed infinitely continuously differentiable over $B(\theta_0,\delta d_n)$ by definition of the pseudo-problem. Any θ subsequently considered within this proof is assumed to belong to $B(\theta_0,\delta d_n)$. Any convergence subsequently mentioned within this proof is uniform in θ for $\theta\in B(\theta_0,\delta d_n)$ for any n large enough thanks to Theorem 7.1 and Lemma 7.6.

Proof of (7.34) for $\beta = (3,0,0)$.

$$\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial\gamma)^3} = 0.$$

Proof of (7.34) for $\beta = (2,1,0)$.

$$\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial\gamma)^2\partial u}\right| = \frac{2}{\sigma^2}\left|\frac{1}{n}\sum_{i=1}^{n^*}(t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right| \xrightarrow[n\to+\infty]{} \frac{2}{\sigma^4}\left|\int_{\underline{u}}^{u}(t-u)f(t)\,\mathrm{d}t\right| \leqslant \frac{2}{\sigma^2}|\overline{u}-\underline{u}|.$$

Proof of (7.34) for $\beta = (2,0,1)$.

$$\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial\gamma)^2\partial\sigma^2}\right| = \frac{1}{\sigma^4}\left|\frac{1}{n}\sum_{i=1}^{n^*}(t_i-u)^2\mathbb{1}_{]t_i,+\infty[}(u)\right| \xrightarrow[n\to+\infty]{} \frac{1}{\sigma^4}\left|\int_{\underline{u}}^{u}(t-u)^2f(t)\,\mathrm{d}t\right| \leqslant \frac{1}{\sigma^4}|\overline{u}-\underline{u}|^2.$$

Proof of (7.34) for $\beta = (1,2,0)$.

$$\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{\partial\gamma\cdot(\partial u)^2}\right| = \frac{2}{\sigma^2}\left|\gamma\frac{1}{n}\sum_{i=1}^{n^*}\mathbb{1}_{]t_i,+\infty[}(u)\right| \xrightarrow[n\to+\infty]{} \frac{2}{\sigma^2}\left|\gamma\int_{\underline{u}}^{u}f(t)\,\mathrm{d}t\right| \leqslant \frac{2}{\sigma^2}|\gamma|.$$

Proof of (7.34) for $\beta = (1,1,1)$.

$$\begin{aligned}\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{\partial\gamma\partial u\partial\sigma^2}\right| &= \frac{1}{\sigma^4}\left|\frac{1}{n}\sum_{i=1}^{n^*}(X_i-2\gamma\cdot(t_i-u))\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &= \frac{1}{\sigma^4}\left|\frac{1}{n}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{]t_i,+\infty[}(u_0)-2\gamma\cdot(t_i-u)\right]\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &\xrightarrow[n\to+\infty]{a.s.} \frac{1}{\sigma^4}\left|\int_{\underline{u}}^{\min(u,u_0)}\gamma_0\cdot(t-u_0)f(t)\,\mathrm{d}t-2\int_{\underline{u}}^{u}\gamma\cdot(t-u)f(t)\,\mathrm{d}t\right|\end{aligned}$$

And this limit is bounded by $\frac{3}{\sigma^4}|\overline{u}-\underline{u}|(|\gamma|+|\gamma_0|)$.

Proof of (7.34) for $\beta = (1,0,2)$.

$$\begin{aligned}\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{\partial\gamma\cdot(\partial\sigma^2)^2}\right| &= \frac{2}{\sigma^6}\left|\frac{1}{n}\sum_{i=1}^{n^*}\left[x_i-\gamma\cdot(t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right](t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &= \frac{2}{\sigma^6}\left|\frac{1}{n}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{]t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\right](t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &\xrightarrow[n\to+\infty]{a.s.} \frac{2}{\sigma^6}\left|\int_{\underline{u}}^{\min(u,u_0)}\gamma_0\cdot(t-u_0)(t-u)f(t)\,\mathrm{d}t-\int_{\underline{u}}^{u}\gamma\cdot(t-u)^2f(t)\,\mathrm{d}t\right|\end{aligned}$$

And this limit is bounded by $\frac{4}{\sigma^6}|\overline{u}-\underline{u}|^2(|\gamma|+|\gamma_0|)$.

Proof of (7.34) for $\beta = (0,3,0)$.

$$\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial u)^3}\right| = 0.$$

Proof of (7.34) for $\beta = (0,2,1)$.

$$\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial u)^2\partial\sigma^2}\right| = \frac{1}{\sigma^4}\gamma^2\left|\frac{1}{n}\sum_{i=1}^{n^*}\mathbb{1}_{]t_i,+\infty[}(u)\right| \xrightarrow[n\to+\infty]{} \frac{1}{\sigma^4}\gamma^2\left|\int_{\underline{u}}^{u} f(t)\,\mathrm{d}t\right| \leqslant \frac{1}{\sigma^4}\gamma^2.$$

Proof of (7.34) for $\beta = (0,1,2)$.

$$\begin{aligned}\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{\partial u(\partial\sigma^2)^2}\right| &= \frac{2}{\sigma^6}\left|\frac{1}{n}\sum_{i=1}^{n^*}\left[x_i-\gamma\cdot(t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right]\gamma\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &= \frac{2}{\sigma^6}\left|\frac{1}{n}\sum_{i=1}^{n^*}\left[\xi_i+\gamma_0\cdot(t_i-u_0)\mathbb{1}_{]t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\right]\gamma\mathbb{1}_{]t_i,+\infty[}(u)\right| \\ &\xrightarrow[n\to+\infty]{a.s.} \frac{2}{\sigma^6}\left|\int_{\underline{u}}^{u_0}\gamma\gamma_0\cdot(t-u_0)f(t)\,\mathrm{d}t-\int_{\underline{u}}^{u}\gamma^2(t-u)f(t)\,\mathrm{d}t\right|\end{aligned}$$

And this limit is bounded by $\frac{2}{\sigma^6}|\overline{u}-\underline{u}|(|\gamma^2|+|\gamma_0\gamma|)$.

Proof of (7.34) for $\beta = (0,0,3)$.

$$\begin{aligned}\left|\frac{1}{n}\frac{\partial^3 l^*_{1:n}(X_{1:n}|\theta)}{(\partial\sigma^2)^3}\right| &= \frac{1}{\sigma^6}\left|-1+\frac{3}{\sigma^2}\frac{1}{n}\sum_{i=1}^{n^*}\left[x_i-\gamma\cdot(t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right]^2\right| \\ &= \frac{1}{\sigma^6}\left|-1+\frac{3}{\sigma^2}\frac{1}{n}\sum_{i=1}^{n^*}\left[\xi_i+\gamma\cdot(t_i-u_0)\mathbb{1}_{]t_i,+\infty[}(u_0)-\gamma\cdot(t_i-u)\mathbb{1}_{]t_i,+\infty[}(u)\right]^2\right| \\ &\xrightarrow[n\to+\infty]{a.s.} \frac{1}{\sigma^8}\left|-\sigma^2+3\left(\sigma_0^2+\int_{\underline{u}}^{u_0}\gamma_0^2(t-u_0)^2f(t)\,\mathrm{d}t\right.\right. \\ &\qquad\left.\left.-2\int_{\underline{u}}^{\min(u,u_0)}\gamma\gamma_0\cdot(t-u_0)(t-u)f(t)\,\mathrm{d}t+\int_{\underline{u}}^{u}\gamma^2(t-u)^2f(t)\,\mathrm{d}t\right)\right|\end{aligned}$$

And this limit is bounded by $\frac{1}{\sigma^8}\left[3\sigma_0^2+\sigma^2+(|\gamma|+|\gamma_0|)^2(\overline{u}-\underline{u})^2\right]$.

(7.34) is thus a direct consequence of both the uniform convergences mentioned above and the trivial majoration of all the limits involved by a fixed constant C for any n large enough. □

Acknowledgements

The authors would like to thank the referees and associate editor for their constructive comments.

References

Billingsley, P. (1995). *Probability and Measure*. Wiley, 3rd edition.

Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley, 2nd edition.

Breiman, L. (1992). *Probability*. SIAM.

Bruhns, A., Deurveilher, G., and Roy, J. (2005). A non-linear regression model for mid-term load forecasting and improvements in seasonnality. *Proceedings of the 15th Power Systems Computation Conference 2005, Liege Belgium*.

Dacunha-Castelle, D. (1978). Vitesse de convergence pour certains problèmes statistiques. In *École d'Été de Probabilités de Saint-Flour, VII (Saint-Flour, 1977)*, volume 678 of *Lecture Notes in Math.*, pages 1–172. Springer, Berlin.

Feder, P. I. (1975). On asymptotic distribution theory in segmented regression problems – identified case. *The Annals of Statistics*, 3(1):49–83.

Ghosal, S. and Samanta, T. (1995). Asymptotic behaviour of Bayes estimates and posterior distributions in multiparameter nonregular cases. *Math. Methods Statist.*, 4(4):361–388.

Ghosh, J. K., Delampady, M., and Samanta, T. (2006). *An Introduction to Bayesian Analysis, Theory and Methods*. Springer.

Ghosh, J. K., Ghosal, S., and Samanta, T. (1994). Stability and convergence of the posterior in non-regular problems. In *Statistical decision theory and related topics, V (West Lafayette, IN, 1992)*, pages 183–199. Springer, New York.

Ghosh, J. K. and Ramamoorthi, R. V. (2003). *Bayesian Nonparametrics*. Springer.

Ibragimov, I. and Has'minskii, R. (1981). *Statistical Estimation Asymptotic Theory*. Springer.

Launay, T., Philippe, A., and Lamarche, S. (2012). Construction of an informative hierarchical prior distribution. application to electricity load forecasting. *Preprint*. arXiv:1109.4533.

Lehmann, E. L. (2004). *Elements of large-sample theory*. Springer Texts in Statistics. Springer-Verlag, New York.

Loève, M. (1991). *Proability Theory I*. Springer, 4th edition.

Polya, G. and Szegö, G. (2004). *Problems and Theorems in Analysis I*. Springer.

Sareen, S. (2003). Reference bayesian inference in non-regular models. *Journal of Econometrics*, 113:265–288.

Sylwester, D. L. (1965). On maximum likelihood estimation for two-phase linear regression. Technical Report 11, Department of Statistics, Stanford University.