Supplementary Material for “Adaptive kernel estimation of the baseline function in the Cox model, with high-dimensional covariates”

Agathe Guilloux
Laboratoire de Statistique Théorique et Appliquée,
Université Pierre et Marie Curie - Paris 6
Centre de Mathématiques Appliquées, Ecole Polytechnique,
CNRS UMR 7641, 91128 Palaiseau, France
e-mail: agathe.guilloux@upmc.fr

Sarah Lemler
Laboratoire Mathématiques et Informatique pour la Complexité et les Systèmes,
École CentraleSupélec, France
e-mail: sarah.lemler@centralesupelec.fr

Marie-Luce Taupin
Laboratoire de Mathématiques et Modélisation d’Évry UMR CNRS 8071- USC INRA,
Université d’Évry Val d’Essonne, France
Unité MaIAGE, INRA Jouy-En-Josas, France
e-mail: marie-luce.taupin@genopole.cnrs.fr

1 Technical lemmas

1.1 Proof of Lemma ??
From Definition (??), we have $\alpha_h(t) = K_h * \alpha_0(t)$ for all $t \in [0, \tau]$ so that
\[
\mathbb{E}[||\tilde{\alpha}_h - \alpha_0||_2^2] = \mathbb{E}[||\tilde{\alpha}_h - \alpha_h||_2^2] + ||\alpha_h - \alpha_0||_2^2.
\]
The first term of the right part of this equality can be rewritten as
\[
\mathbb{E}[||\tilde{\alpha}_h - \alpha_h||_2^2] = \int_0^\tau \text{Var}[\tilde{\alpha}_h(t)]dt.
\]
It remains to bound $\text{Var}[\tilde{\alpha}_h(t)]$:
\[
\text{Var}[\tilde{\alpha}_h(t)] = \frac{1}{n} \text{Var} \left[ \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} dN_1(u) \right] \\
\leq \frac{1}{n} \mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} dN_1(u) \right)^2 \right].
\]

We apply the Doob-Meyer decomposition \( N_1 = M_1 + \Lambda_1 \):

\[
\text{Var}[\tilde{\alpha}_h(t)] \leq \frac{2}{n} \mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} dM_1(u) \right)^2 \right] + \frac{2}{n} \mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} \alpha_0(u) e^{\theta Z_1 Y_1(u)} du \right)^2 \right].
\]

Then, we have

\[
\mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} dM_1(u) \right)^2 \right] \leq \mathbb{E} \left[ \int_0^\tau K_h^2(t-u) \frac{1}{S^2(u, \beta_0)} \alpha_0(u) e^{\theta Z_1 Y_1(u)} du \right],
\]

and we finally get that

\[
\mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} dM_1(u) \right)^2 \right] \leq \frac{\mathbb{E}[e^{\theta T} Z_1] ||\alpha_0||_{\infty, \tau} ||K||_{L^2(\mathbb{R})}^2}{c_S^2}.
\]

Thus, the integrated variance term of the pseudo-estimator is bounded by

\[
\int_0^\tau \text{Var}[\tilde{\alpha}_h(t)] dt \leq \frac{2 ||\alpha_0||_{\infty, \tau} ||K||_{L^2(\mathbb{R})}^2}{c_S^2} \left( \mathbb{E}[e^{\theta T} Z_1] + ||\alpha_0||_{\infty, \tau} \mathbb{E}[e^{2\theta T} Z_1] \right) \frac{||K||_{L^2(\mathbb{R})}^2}{nh},
\]

which gives a bound of order \( 1/nh \).

Gathering the bias term \( ||\alpha_h - \alpha_0||_2^2 \) and the bound on variance term gives Inequality (??) in Lemma ??.

\[ \square \]

### 1.2 Proof of Lemma ??

The proof of Lemma ?? relies on an additional lemmas. First, write

\[
\hat{\alpha}_h(t) - \tilde{\alpha}_h(t) = \frac{1}{nh} \sum_{i=1}^n \int_0^\tau S(u, \beta_0) - S_n(u, \hat{\beta}) K \left( t - u \right) \frac{1}{S(u, \beta_0)S_n(u, \hat{\beta})} dN_i(u).
\]

We study the difference process \( \hat{\alpha}_h - \tilde{\alpha}_h \) on \( \Omega_k \), defined by (??) and on its complement. From Lemma ??, the process \( \hat{\alpha}_h - \tilde{\alpha}_h \) is controled on \( \Omega_k^c \). The following lemma allows to bound the difference process on \( \Omega_k \).
Lemma 1.1. Under Assumptions ?? ??, ?? ??, ?? ??, ?? and ???, for any $k \in \mathbb{N}$, we have

$$
\mathbb{E}[||\hat{\alpha}_{h} - \bar{\alpha}_{h}||_{2}^{2} \mathbb{1}(\Omega_{k})] \leq c_{4}(s)\frac{\log(nkp)}{n},
$$

where $c_{4}$ is a constant depending on $B, |\beta_{0}|_{1}, R, ||\alpha_{0}||_{\infty, \tau}, c_{S}, ||K||_{L^{2}(\mathbb{R})}, \tau$ and $s$ the sparsity index of $\beta_{0}$.

Gathering Lemmas ?? and 1.1, we finally get that, for a fixed $k$

$$
\mathbb{E}[||\hat{\alpha}_{h} - \bar{\alpha}_{h}||_{2}^{2}] \leq c(s)\frac{\log(nkp)}{n},
$$

with $c(s)$ a constant depending on $B, |\beta_{0}|_{1}, R, ||\alpha_{0}||_{\infty, \tau}, c_{S}, ||K||_{L^{2}(\mathbb{R})}, \tau$ and $s$ the sparsity index of $\beta_{0}$, and Lemma ?? is then proved. Let now prove the Lemmas ?? and 1.1.

1.3 Proof of Lemma ?? :

We have to bound $\mathbb{E}[||\hat{\alpha}_{h} - \bar{\alpha}_{h}||_{2}^{2} \mathbb{1}(\Omega_{k})]$, which is equal to

$$
\mathbb{E}[||\hat{\alpha}_{h} - \bar{\alpha}_{h}||_{2}^{2} \mathbb{1}(\Omega_{k})] = \int_{0}^{\tau} \mathbb{E}[(\hat{\alpha}_{h}^{2} - \bar{\alpha}_{h})^{2}] \mathbb{1}(\Omega_{k})]dt.
$$

First, let us focus on $\mathbb{E}[(\hat{\alpha}_{h}^{2} - \bar{\alpha}_{h})^{2}] \mathbb{1}(\Omega_{k})]$ defined by

$$
\mathbb{E}\left[\left(\frac{1}{n} \sum_{i=1}^{n} \int_{0}^{\tau} K_{h}(t-u) \mathbb{1}_{\{Y(u) > 0\}} S(u, \beta_{0}) - S_{n}(u, \hat{\beta}) \right) \mathbb{1}(\Omega_{k})\right].
$$

From Assumptions ??, $\hat{\beta}$ belongs to a ball $B(0, R)$ and $|\beta_{0}|_{1} < \infty$, so we have the following bound

$$
S_{n}(u, \hat{\beta}) - S(u, \beta_{0}) \leq 1 - \frac{n}{n} \sum_{i=1}^{n} e^{\beta_{0}Tz_{i}}e^{(\hat{\beta} - \beta_{0})^{T}z_{i}} \leq e^{2\beta_{0}^{1}}e^{BR}.
$$

For sake of simplicity, let us denote $C(B, R, |\beta_{0}|_{1})$ the bound in (1). From $\mathbb{1}_{\{Y(u) > 0\}}$ in the definition of $\hat{\alpha}_{h}$, there exists $i_{0} \in \{1, ..., n\}$ such that $Y_{i_{0}} \neq 0$, so that from Assumption ??

$$
S_{n}(u, \hat{\beta}) \geq \frac{1}{n} e^{-B|\beta_{0}|_{1}} e^{-B|\hat{\beta} - \beta_{0}|_{1}} \geq \frac{1}{n} e^{-2B|\beta_{0}|_{1}} e^{-BR}.
$$

Combining (1) and (2), for $C(B, R, |\beta_{0}|_{1}) = e^{8B|\beta_{0}|_{1}}e^{4BR}$, we obtain the following bound

$$
\mathbb{E}[(\hat{\alpha}_{h}^{2} - \bar{\alpha}_{h})^{2}] \mathbb{1}(\Omega_{k})] \leq \tilde{C}(B, R, |\beta_{0}|_{1}) \frac{n}{c_{S}^{2}} \mathbb{E}\left[\left(\int_{0}^{\tau} K_{h}(t-u) dN_{1}(u)\right)^{2} \mathbb{1}(\Omega_{k})\right].
$$

From Assumption ?? and the Cauchy-Schwarz Inequality, we get

$$
\mathbb{E}\left[\left(\int_{0}^{\tau} K_{h}(t-u) dN_{1}(u)\right)^{2} \mathbb{1}(\Omega_{k})\right] \leq \frac{||K||_{\infty}^{2}}{h^{2}} \mathbb{E}[N_{1}(\tau)^{2} \mathbb{1}(\Omega_{k})]
$$

$$
\leq \frac{||K||_{\infty}^{2}}{h^{2}} \mathbb{E}[N_{1}(\tau)^{4}]^{1/2} \mathbb{P}(\Omega_{k})^{1/2}.
$$
From the Doob-Meyer decomposition and the Bürkholder Inequality 2.4, we deduce that 
\( \mathbb{E}[N_1(\tau)] < \infty \).

Now we focus on 
\[ \int_0^\tau \mathbb{E}[(\hat{\alpha}_h^2 - \bar{\alpha}_h)^2(\Omega_k^c)] \, dt. \]
From the two bounds (3) and (4) obtained above, we have
\[
\int_0^\tau \mathbb{E}[(\hat{\alpha}_h^2 - \bar{\alpha}_h)^2(\Omega_k^c)] \, dt \leq \tilde{C}(B, R, ||\beta_0||_1) \frac{n^2}{c_S^2} \left( \frac{||K||_2^2}{h^2} \mathbb{E}^{1/2}[N_1(\tau)^4] \sqrt{\mathbb{P}(\Omega_k^c)} \right).
\]

Let introduce the following lemma that gives a bound for \( \mathbb{P}(\Omega_k^c) \).

**Lemma 1.2.** Under Assumptions ?? and ??, for all \( k \in \mathbb{N} \), there exists \( n_0 \in \mathbb{N} \), such that for \( n > n_0 \) we have
\[
\mathbb{P}[\Omega_k^c] \leq c_2 n^{-k}, \tag{5}
\]
where \( c_2 \) is a constant depending on \( B, ||\beta_0||_1 \) and \( s \).

From Lemma 1.2 and the fact that \( h^{-1} \leq n \) from Assumption ???, we get that
\[
\int_0^\tau \mathbb{E}[(\hat{\alpha}_h^2 - \bar{\alpha}_h)^2(\Omega_k^c)] \, dt \leq C(B, ||\beta_0||_1, R, ||\alpha_0||_{\infty, \tau}, c_S, \tau, ||K||_\infty) n^{4-k/2},
\]
where \( C(B, ||\beta_0||_1, R, ||\alpha_0||_{\infty, \tau}, c_S, \tau, ||K||_\infty) \) is a constant depending on elements in brackets. Finally, we obtain
\[
\mathbb{E}[||\hat{\alpha}_h^2 - \bar{\alpha}_h||_2^2(\Omega_k^c)] \leq C(B, ||\beta_0||_1, R, ||\alpha_0||_{\infty, \tau}, c_S, \tau, ||K||_\infty) n^{4-k/2},
\]
which ends the proof of Lemma ??.

\qed

### 1.4 Proof of Lemma 1.1 :

On \( \Omega_k \), we have
\[
\mathbb{E}[(\hat{\alpha}_h^2 - \bar{\alpha}_h)^2(t)(\Omega_k^c)] \leq \frac{16B^2e^{4B ||\beta_0||_1^2}e^{2BR}C^2(s) \log(pn^k)}{c_S^2} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n \int_0^\tau |K_h(t-u)| S(u, \beta_0) \, dN_i(u) \right)^2 \right].
\]

Then, from the Doob-Meyer decomposition, we deduce that
\[
\mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n \int_0^\tau |K_h(t-u)| S(u, \beta_0) \, dN_i(u) \right)^2 \right] \leq \frac{2}{n} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n \int_0^\tau K_h(t-u) \, S(u, \beta_0) \, dM_1(u) \right)^2 \right] + \frac{2}{n} \mathbb{E} \left[ \left( \int_0^\tau K_h(t-u) \frac{1}{S(u, \beta_0)} \alpha_0(u) e^{\beta_0^T u} z_1 Y_1(u) \, du \right)^2 \right] \leq \frac{2||\alpha_0||_{\infty, \tau}}{c_S^2} (\mathbb{E}[e^{\beta_0^T z_1}] + ||\alpha_0||_{\infty, \tau} \mathbb{E}[e^{2\beta_0^T z_1}]) \frac{||K||_2^2}{c_S^2}.
\]

We finally obtain
\[
\mathbb{E}[||\hat{\alpha}_h^2 - \bar{\alpha}_h||_2^2(\Omega_k^c)] \leq C(c_S, ||\alpha_0||_{\infty, \tau}, ||\beta_0||_1, B, R, ||K||_2(\mathbb{R}), s) \frac{\log(pn^k)}{n}. \tag{6}
\]
\qed
1.5 Proof of Lemma 1.2

In order to calculate $\mathbb{P}(\Omega_k^c)$, let us begin by study the set $\Omega_{H,k}$ defined by (??). Let us introduce the two following sets

$$
\Omega_1 := \left\{ \omega : \forall u \in [0, \tau], |\hat{S}(u, \hat{\beta}) - S(u, \beta_0)| \leq B e^{BR}e^{2B|\beta_0|_1}C(s) \sqrt{\frac{\log(pmk)}{n}} \right\},
$$

$$
\Omega_2 := \left\{ \omega : \forall u \in [0, \tau], |S(u, \beta_0) - S(u, \beta_0)| \leq B e^{BR}e^{2B|\beta_0|_1}C(s) \sqrt{\frac{\log(pmk)}{n}} \right\}.
$$

We have $\Omega_{H,k} \supset \Omega_1 \cap \Omega_2$. We begin to calculate $\mathbb{P}(\Omega_1^c)$. By definition, we have

$$
|\hat{S}(u, \hat{\beta}) - S(u, \beta_0)| = \frac{1}{n} \sum_{i=1}^{n} \left( e^{\hat{\beta}^T Z_i} - e^{\beta_0^T Z_i} \right) Y_i(u)
$$

$$
\leq e^{B|\beta_0|_1} |e^{\hat{\beta} - \beta_0|_1} - 1|
$$

Under Assumptions ?? and ??, from Proposition ??, there exists a constant $c > 0$ such that, with probability larger than $1 - cn^{-k}$,

$$
|\hat{\beta} - \beta_0|_1 \leq C(s) \sqrt{\frac{\log(pmk)}{n}}.
$$

So, with probability larger than $1 - cn^{-k}$, using that $|e^x - e^y| \leq |x - y|e^{x+y}$ for all $x, y$, we have

$$
|\hat{S}(u, \hat{\beta}) - S(u, \beta_0)| \leq e^{B|\beta_0|_1}B|\hat{\beta} - \beta_0|_1 e^{B|\hat{\beta} - \beta_0|_1}
$$

$$
\leq Be^{BR}e^{2B|\beta_0|_1}C(s) \sqrt{\frac{\log(pmk)}{n}}.
$$

We deduce that

$$
\mathbb{P}(\Omega_1^c) \leq cn^{-k}. \quad (7)
$$

To calculate $\mathbb{P}(\Omega_2^c)$, we remark that

$$
n(S(u, \beta_0) - S(u, \beta_0)) = \sum_{i=1}^{n} \left( e^{\beta_0^T Z_i} Y_i(u) - \mathbb{E} \left[ e^{\beta_0^T Z_i} Y_i(u) \right] \right).
$$

As $0 \leq e^{\beta_0^T Z_i} Y_i(u) \leq e^{B|\beta_0|_1}$, we apply a Hoeffding inequality:

$$
\mathbb{P} \left( \left| S(u, \beta_0) - S(u, \beta_0) \right| \geq \frac{y}{n} \right) \leq 2 \exp \left( - \frac{2y^2}{ne^{2B|\beta_0|_1}} \right),
$$
and with \( y = Be^{BR}e^{2B|\beta_0|}C(s)\sqrt{n \log(pn^k)/2} \), we finally get

\[
\mathbb{P}\left( |S_n(u, \beta_0) - S(u, \beta_0)| \leq C(s)Be^{BR}e^{2B|\beta_0|}\sqrt{\log(pn^k)/n} \right)
\leq 2\exp\left( -\frac{2B^2e^{2BR}e^{4B|\beta_0|}C^2(s)\log(pn^k)}{e^{2B|\beta_0|}} \right)
\leq \frac{2}{pn^k}.
\]

We conclude that there exists a constant \( c_7 > 0 \) such that

\[
\mathbb{P}(\Omega_2^c) \leq c_7n^{-k}. \tag{8}
\]

Gathering (7) and (8), we obtain

\[
\mathbb{P}(\Omega_{H,k}^c) \leq \mathbb{P}(\Omega_1^c) + \mathbb{P}(\Omega_2^c) \leq \tilde{c}n^{-k}, \tag{9}
\]

where \( \tilde{c} > 0 \) is a constant. It remains to calculate \( \mathbb{P}(\Omega_{S_n}^c) \), with \( \Omega_{S_n}^c \) defined by (??), to obtain \( \mathbb{P}(\Omega_k^c) \). We decompose

\[
S_n(u, \hat{\beta}) - S_n(u, \beta_0) = S_n(u, \hat{\beta}) - S_n(u, \beta_0) + S_n(u, \beta_0) - S(u, \beta_0).
\]

On \( \Omega_1 \cap \Omega_2 \),

\[
S_n(u, \hat{\beta}) - S_n(u, \beta_0) \geq -2Be^{BR}e^{2B|\beta_0|}\sqrt{\log(pn^k)/n} \in (-\infty, 0)
\]

So for \( n \) large enough, we have that \( S_n(u, \hat{\beta}) - S_n(u, \beta_0) \geq -cs/2 \). For \( n \) large enough, \( \Omega_1 \cap \Omega_2 \subset \Omega_{S_n} \), and

\[
\mathbb{P}(\Omega_{S_n}^c) \leq \mathbb{P}(\Omega_1^c) + \mathbb{P}(\Omega_2^c) \leq \tilde{c}n^{-k}. \tag{10}
\]

Gathering (9) and (10), we finally obtain for \( n \) large enough that \( \mathbb{P}(\Omega_k^c) \leq c_2n^{-k} \), where \( c_2 \) is a constant depending on \( B, |\beta_0|_1 \) and \( s \).

\[\square\]

2 Classical results

In this appendix, some classical technical lemmas and a theorem needed for the proofs of the two main theorems of the chapter, are listed. We do not give the proofs of these well-known results but we give the references where to find their proofs.

2.1 A Cauchy-Schwarz Inequality

The following lemma gives a useful inequality concerning integrals with respect to the counting process \( N \).
Lemma 2.1 (Cauchy-Schwarz). For all function $g$ bounded on $[0, \tau]$,  
\[
N(\tau) \int_{\tau_1}^{\tau_2} g^2(s) dN(s) \geq \left( \int_{\tau_1}^{\tau_2} g(s) dN(s) \right)^2, \quad 0 \leq \tau_1 \leq \tau_2 \leq \tau
\]
We refer to ? for the proof of this lemma.

2.2 Young Inequality

The following lemma provides an inequality that bounds a norm of the convolution product of two functions by a product of norms of each function.

Lemma 2.2 (Young Inequality). Let $p, q \in [1, +\infty)$ such that $1/p + 1/q \geq 1$. If $s \in L^p(\mathbb{R})$ and $t \in L^q(\mathbb{R})$, then $s$ and $t$ are convolable. Moreover, if $1/r = 1/p + 1/q - 1$, then $f * g \in L^r(\mathbb{R})$ and  
\[
||s * t||_r \leq ||s||_p ||t||_q
\]
This convolution inequality is proved in ? (Theorem 3.4 p.149).

2.3 Talagrand Inequality

The following Talagrand Inequality is a concentration inequality that allows to control the supremum of an empirical process.

Theorem 2.3 (Talagrand Inequality). Let $\xi_1, \ldots, \xi_n$ be independent random values, and let
\[
\nu_{n, \xi}(f) = \frac{1}{n} \sum_{i=1}^{n} \{ f(\xi_i) - \mathbb{E}[f(\xi_i)] \}.
\]
Then, for a countable class of functions $\mathcal{F}$ uniformly bounded and $\varepsilon > 0$, we have
\[
\mathbb{E} \left[ \sup_{f \in \mathcal{F}} \nu_{n, \xi}^2(f) - 2(1 + 2\varepsilon^2)H^2 \right] \leq \frac{4}{d} \left( \frac{W}{n} e^{-d\varepsilon^2 n H^2 W} + \frac{98M^2}{dn^2 \varphi^2(\varepsilon)} e^{-2d\varphi(\varepsilon) n H / M} \right),
\]
with $\varphi(\varepsilon) = \sqrt{1 + \varepsilon^2} - 1$, $d = 1/6$ and
\[
\sup_{f \in \mathcal{F}} ||f||_\infty \leq M, \quad \mathbb{E} \left[ \sup_{f \in \mathcal{F}} \nu_{n, \xi}(f) \right] \leq H, \quad \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} \text{Var}[f(\xi)] \leq W.
\]
This theorem is a useful corollary from the classical Talagrand established by ?. The proof of Theorem 2.3 can be found in ? (Lemma 6.1). The proof of the theorem follows from a concentration Inequality in ? and arguments that can be found in ?.
2.4 A classical inequality: the Bürkholder Inequality

The last technical result is a Bürkholder Inequality that gives a norm relation between a martingale and its optional process. We refer to p.75, for the proof of this result.

**Theorem 2.4** (Bürkholder Inequality). If \( M = (M_t, \mathcal{F}_t)_{t \geq 0} \) is a martingale, then there are universal constants \( \gamma_b \) and \( \kappa_b \) (independent of \( M \)) such that for every \( t \geq 0 \)

\[
\gamma_b \| \sqrt{[M]_t} \|_2 \leq \| M_t \|_2 \leq \kappa_b \| \sqrt{[M]_t} \|_2,
\]

where \([M]_t\) is the quadratic variation of \( M_t \).

This theorem is used to prove Lemma ?? and in the oracle inequalities of Theorem ??, the constants depend on \( \kappa_b \).