Contribution to nonparametric regression estimation with general autocovariance error process and application to pharmacokinetics

Djihad Benelmadani

Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : Mathématiques Appliquées

Arrêté ministériel : 25 mai 2016

Présentée par

Djihad BENELMADANI

Thèse dirigée par **Sana LOUHICHI** et **Karim BENHENNI**

préparée au sein du **Laboratoire Jean Kuntzmann**
dans **l'École Doctorale Mathématiques, Sciences et technologies de l'information, Informatique**

Contribution à la régression non paramétrique avec un processus erreur d'autocovariance générale et application en pharmacocinétique

Thèse soutenue publiquement le **18 Septembre 2019,**
devant le jury composé de :

Mme Sana LOUHICHI
Professeur, Université Grenoble Alpes, Directrice de thèse
M. Karim BENHENNI
Maître de Conférences HDR, Université Grenoble Alpes, Directeur de thèse
M. Salim BOUZEBDA
Professeur, Université de Technologie de Compiègne, Rapporteur
Mme Sophie DABO-NIANG
Professeur, Université de Lille, Rapporteur
Mme Clémentine PRIEUR
Professeur, Université Grenoble Alpes, Présidente
Mme Marianne CLAUSEL
Professeur, Université de Loraine, Examinatrice
M. Didier Girard
Chargé de recherche CNRS, Invité

Acknowledgments

The present document, resulting of four years of research, would not have been possible without the guidance of God, and the support and help that I received from many people. I thank them all.

First of all, I would like to thank my supervisors, Sana Louhichi and Karim Benhenni, for their patient guidance, their help and support during my time as their student. I have been so lucky to have had supervisors who cared and believed in me.

I would also like to thank Sophie Dabo-Niang and Salim Bouzebda for accepting to review my thesis, and all the members of the committee, for their insightful remarks and their challenging questions. A special thanks to Didier Girard, for his helpful comments concerning the thesis simulations through these years.

I would like to thank all the members of Laboratoire Jean Kuntzmann, the stuff members for their helpful assistance, the teachers the students and my office colleagues for making my PhD life friendly on a daily basis. A special thanks to Alain Latour and Corinne Berzin for their help and their trust during my teaching life.

I gratefully acknowledge the funding received towards my PhD from the Algerian Government.

I would also like to thank my professors I had during my undergraduate and postgraduate studies, without whom I would never had been able to achieve this PhD. A special thanks to Fatiha Messaci and Zaher Mohdeb for, first entrusting me in Statistics, and all their help and their advices. Thank you for trusting me.

Whatever I say, I can never express my gratitude to my parents. Thank you so much for the unconditional love, the sacrifices, the support and encouragements, and for the prayers you showerd me with all my life. Thank you for the education you offered me leading to this PhD.

I would like to thank my husband, for always being there emotionally and scientifically. Thank you for the time you spent listening to my long complains and all your helpful scientific advices.

I would also like to thank my brother, my sisters and my little niece. For there is no joy they cannot magnify, no pain they cannot quell and no situation they cannot draw laughter from.

Finally, thank you to all my friends, who showed me that no distance can separate real caring hearts. Thank you for always being there for me.

Summary

In this thesis, we consider the fixed design regression model with repeated measurements, where the errors form a process with general autocovariance function, i.e. a second order process (stationary or nonstationary), with a non-differentiable covariance function along the diagonal. We are interested, among other problems, in the nonparametric estimation of the regression function of this model.

We first consider the well-known kernel regression estimator proposed by Gasser and Müller. We study its asymptotic performance when the number of experimental units and the number of observations tend to infinity. For a regular sequence of designs, we improve the higher rates of convergence of the variance and the bias. We also prove the asymptotic normality of this estimator in the case of correlated errors.

Second, we propose a new kernel estimator of the regression function based on a projection property. This estimator is constructed through the autocovariance function of the errors, and a specific function belonging to the Reproducing Kernel Hilbert Space (RKHS) associated to the autocovariance function. We study its asymptotic performance using the RKHS properties. These properties allow to obtain the optimal convergence rate of the variance. We also prove its asymptotic normality. We show that this new estimator has a smaller asymptotic variance then the one of Gasser and Müller. A simulation study is conducted to confirm this theoretical result.

Third, we propose a new kernel estimator for the regression function. This estimator is constructed through the trapezoidal numerical approximation of the kernel regression estimator based on continuous observations. We study its asymptotic performance, and we prove its asymptotic normality. Moreover, this estimator allow to obtain the asymptotic optimal sampling design for the estimation of the regression function. We run a simulation study to test the performance of the proposed estimator in a finite sample set, where we see its good performance, in terms of Integrated Mean Squared Error (IMSE). In addition, we show the reduction of the IMSE using the optimal sampling design instead of the uniform design in a finite sample set.

Finally, we consider an application of the regression function estimation in pharmacokinetics problems. We propose to use the nonparametric kernel methods, for the concentration-time curve estimation, instead of the classical parametric ones. We prove its good performance via simulation study and real data analysis. We also investigate the problem of estimating the Area Under the concentration Curve (AUC), where we introduce a new kernel estimator, obtained by the integration of the regression function estimator. We prove, using a simulation study, that the proposed estimators outperform the classical one in terms of Mean Squared Error. The crucial problem of finding the optimal sampling design for the AUC estimation is investigated using the Generalized Simulating Annealing algorithm.

Keywords. Nonparametric regression, correlated observations, autocovariance function, reproducing kernel Hilbert space, trapezoidal rule, asymptotic normality, pharmacokinetics.

Résumé

Dans cette thèse, nous considérons le modèle de régression avec plusieurs unités expérimentales, où les erreurs forment un processus d'autocovariance dans un cadre générale, c'est-à-dire, un processus du second ordre (stationnaire ou non stationnaire) avec une autocovariance non différentiable le long de la diagonale. Nous sommes intéressés, entre autres, à l'estimation non paramétrique de la fonction de régression de ce modèle.

Premièrement, nous considérons l'estimateur classique proposé par Gasser et Müller. Nous étudions ses performances asymptotiques quand le nombre d'unités expérimentales et le nombre d'observations tendent vers l'infini. Pour un échantillonnage régulier, nous améliorons les vitesses de convergence d'ordre supérieur de son biais et de sa variance. Nous montrons aussi sa normalité asymptotique dans le cas des erreurs corrélées.

Deuxièmement, nous proposons un nouvel estimateur à noyau pour la fonction de régression, basé sur une propriété de projection. Cet estimateur est construit à travers la fonction d'autocovariance des erreurs et une fonction particulière appartenant à l'Espace de Hilbert à Noyau Autoreproduisant (RKHS) associé à la fonction d'autocovariance. Nous étudions les performances asymptotiques de l'estimateur en utilisant les propriétés de RKHS. Ces propriétés nous permettent d'obtenir la vitesse optimale de convergence de la variance de cet estimateur. Nous prouvons sa normalité asymptotique, et montrons que sa variance est asymptotiquement plus petite que celle de l'estimateur de Gasser et Müller. Nous conduisons une étude de simulation pour confirmer nos résultats théoriques.

Troisièmement, nous proposons un nouvel estimateur à noyau pour la fonction de régression. Cet estimateur est construit en utilisant la règle numérique des trapèzes, pour approximer l'estimateur basé sur des données continues. Nous étudions aussi sa performance asymptotique et nous montrons sa normalité asymptotique. En outre, cet estimateur permet d'obtenir le plan d'échantillonnage optimal pour l'estimation de la fonction de régression. Une étude de simulation est conduite afin de tester le comportement de cet estimateur dans un plan d'échantillonnage de taille finie, en terme d'erreur en moyenne quadratique intégrée (IMSE). De plus, nous montrons la réduction dans l'IMSE en utilisant le plan d'échantillonnage optimal au lieu de l'échantillonnage uniforme.

Finalement, nous considérons une application de la régression non paramétrique dans le domaine pharmacocinétique. Nous proposons l'utilisation de l'estimateur non paramétrique à noyau pour l'estimation de la fonction de concentration. Nous vérifions son bon comportement par des simulations et une analyse de données réelles. Nous investiguons aussi le problème de l'estimation de l'Aire Sous la Courbe de concentration (AUC), pour lequel nous proposons un nouvel estimateur à noyau, obtenu par l'intégration de l'estimateur à noyau de la fonction de régression. Nous montrons, par une étude de simulation, que le nouvel estimateur est meilleur que l'estimateur classique en terme d'erreur en moyenne quadratique. Le problème crucial de l'obtention d'un plan d'échantillonnage optimale pour l'estimation de l'AUC est discuté en utilisant l'algorithme de recuit simulé généralisé.

Mots clés. Régression non paramétrique, observations corrélées, fonction d'autocovariance, espace de Hilbert à noyau autoreproduisant, règle des trapèzes, normalité asymptotique, pharmacocinétique.

Contents

List of Figures

List of Tables

Introduction (français)

Présentation générale

Les fonctions de régression non paramétrique ont été largement utilisées dans ces dernières décennies, pas seulement en Statistiques, mais dans plusieurs domaines tels que la médecine et le traitement du signal. La fonction de régression non paramétrique, est une fonction générale qui traduit la relation entre deux variables : une variable explicative X et une variable réponse Y, sans aucune restriction paramétrique sur cette fonction. L'un des problèmes rencontrés par les statisticiens dans leurs études est l'estimation de la fonction de régression, en se basant sur des observations partielles de cette fonction. En Statistiques, cela veut dire qu'on veut estimer la fonction $g(\cdot) = \mathbb{E}(Y|X = \cdot)$ à partir des observations $(X_i, Y_i)_{1 \leq i \leq n}$, qui représentent n copies de (X, Y). Ces observations sont souvent modélisées comme suit :

$$Y_i = g(X_i) + \varepsilon_i,$$

où les $(\varepsilon_i)_{1 \leq i \leq n}$ sont des variables aléatoires centrées appelées erreurs.

Dans cette thèse, nous considérons le cas où les $(X_i)_{1 \leq i \leq n}$ sont fixées dans un domaine et ne sont pas aléatoires (échantillonnage déterministe). C'est le cas, par exemple, des données longitudinales lorsque des unités expérimentales sont observées à des temps d'échantillonnage préalablement choisis. Par exemple, dans les problèmes de pharmacocinétique, les concentrations d'un certain médicament administré dans l'organisme sont observées toutes les demi-heures pendant 24 heures. De là, nous considérons le modèle de régression non paramétrique défini comme suit:

$$Y(t_{i,n}) = g(t_{i,n}) + \varepsilon(t_{i,n}) \quad \text{pour } i = 1, \cdots, n,$$

avec $0 \leq t_{1,n} < t_{2,n} < \cdots < t_{n,n} \leq 1$.

Le modèle le plus rencontré dans la littérature a été celui avec des erreurs indépendantes. Nous mentionnons, entre autres, les travaux de Priestly et Chao (1972), Benedetti (1977) et Gasser et Müller (1979). Ces auteurs ont proposé différents estimateurs à noyaux pour la fonction de régression. Gasser *et al.* (1984) ont utilisé l'estimateur à noyau dans le cas d'une courbe de croissance individuelle. Fan (1992) a considéré un estimateur linéaire local pondéré dans le cas d'observations indépendantes. Cependant, l'hypothèse que les observations sont indépendantes n'est pas toujours réaliste. Par exemple, les tailles observées sur le même enfant sont corrélées. Les observations de température mesurées au cours de la journée sont également corrélées. Pour cela, nous nous concentrons sur le problème de la régression non paramétrique où les observations sont corrélées. Ces modèles ont été utilisés par plusieurs chercheurs, parmi lesquels nous citons, entre autres, Hart et Wherly (1986) qui ont considéré des erreurs stationnaires (corrélées), Núñez-Antón *et al.* et Ferreira *et al.* (1997) ont examiné une classe spécifique d'erreurs non stationnaires avec des fonctions d'autovariance paramétriques, introduites par Núñez-Antón et Woodworth (1994) pour étudier l'efficacité des implants cochléaires. Pour plus de détails, voir Opsomer *et al.* (2001) pour une synthèse des résultats sur la régression non paramétrique en présence d'erreurs corrélées stationnaires.

En considérant des observations corrélées, Hart et Wherly (1986) ont utilisé un estimateur à noyau bien connu pour la fonction de régression, proposé par Gasser et Müller (1979). Ils ont montré que cet estimateur n'est pas consistant, dans le sens où sa variance ne tend pas vers zéro lorsque le nombre d'observations tend vers l'infini. Pour cela, il fallait considérer plusieurs unités expérimentales, disons m, sur lesquelles n observations sont effectuées. Dans ce cas, l'estimateur est consistant quand m tend vers l'infini. Pour des erreurs corrélées non stationnaires, Benhenni et Rachdi (2007) ont aussi considéré l'estimateur de Gasser et Müller.

Une autre conséquence de la présence de corrélation est l'échec de plusieurs méthodes de sélection de la fenêtre basées sur les données, telle que la méthode classique de validation croisée. Pour plus de détails sur cette question, voir par exemple, Altman (1990), Chiu (1989), Hart (1991, 1994) et Opsomer *et al.* (2001).

Dans cette thèse, nous considérons le modèle de régression non paramétrique avec des mesures répétées, définit par:

$$Y_j(t_{i,n}) = g(t_{i,n}) + \varepsilon_j(t_{i,n}), \quad \text{pour } i = 1, \cdots, n \text{ et } j = 1, \cdots, m, \tag{1}$$

où $\{\varepsilon_j, j = 1, \cdots, m\}$ est une suite de processus d'erreur centrées indépendantes et identiquement distribuées (i.i.d). avec la même distribution qu'un processus ε. La non corrélation des observations faites entre les unités expérimentales est une hypothèse naturelle, car en général, les dernières sont choisies indépendamment. Ce type de données a été étudié par plusieurs auteurs dans le cas paramétrique, nous mentionnons entre autres, Potthoff et Roy (1964) qui ont examiné certaines mesures dentaires de 11 filles et 16 garçons à 4 âges différents, Rao (1965, 1966), Grizzle et Allen (1969), Ghosh *et al.* (1973) ont considéré les méthodes non paramétriques dans les études longitudinales.

Nous notons ici que le modèle (1) peut être écrit comme suit:

$$\overline{Y}(t_{i,n}) = g(t_{i,n}) + \overline{\varepsilon}(t_{i,n}), \quad \text{pour} \quad i = 1, \cdots, n,$$

où $\overline{Y}(t_{i,n})$ et $\overline{\varepsilon}(t_{i,n})$ sont respectivement les moyennes empiriques de $\{Y_j(t_{i,n}), j = 1, \cdots, m\}$ et de $\{\varepsilon_j(t_{i,n}), j = 1, \cdots, m\}$. Cela montre que le problème de l'adaptation de la fonction de régression g à m courbes, pourrait être amené à ajuster cette dernière à la moyenne de l'échantillon $\overline{Y}(t_{i,n})$.

Un autre problème intéressant, concernant l'estimation de la fonction de régression g, est la dérivation d'un plan d'échantillonnage optimal $\{t_{i,n}, i = 1, \cdots, n\}$. Ce problème a été largement étudié dans le cas de l'estimation par régression paramétrique, nous mentionnons les travaux de Sacks et Ylvisaker (1966) et Belouni et Benhenni (2013). Ces auteurs ont considéré ce problème pour une classe particulière de fonctions de régression, et ont obtenu le plan d'échantillonnage optimal. Dette *et al.* (2017) ont construit une paire d'estimateurs linéaires sans biais avec l'échantillonnage optimale correspondant, pour comparer deux courbes de régression estimées à partir de deux échantillons finis des mesures dépendantes. Zhigljavsky *et al.* (2010) ont proposé une nouvelle approche d'échantillonnage expérimentale pour le modèle de localisation en présence de corrélation; ils ont montré la convergence faible de leur plan d'échantillonnage vers celui proposé par Bickel et Herzberg (1979).

Dans le cas non paramétrique, le problème de l'échantillonnage optimal a été moins développé dans la littérature, en particulier lorsque les observations sont corrélées. Müller (1984) a considéré un estimateur à noyau et il a introduit les points de l'échantillonnage optimal, lorsque les erreurs sont asymptotiquement non corrélées. Ce qui n'est pas toujours une hypothèse réaliste, pour les données longitudinales par exemple. Il a utilisé une suite d'échantillonnage régulier générée par une fonction de densité, et il a déduit la densité optimale qui minimise l'Erreur Moyenne Quadratique Intégrée (IMSE) asymptotique. Faraway (1990) a utilisé l'estimateur de Priestly et Chao, et Zhao et Yao (2012) ont utilisé un estimateur avec une fonction-poids générale. Ils ont considéré un échantillonnage séquentiel, pour l'estimation non paramétrique de la fonction de régression avec des observations corrélées. Efromovich (2008) a considéré le problème de l'échantillonnage optimal pour une hétéroscédasticité conditionnelle en utilisant l'approche par des séries de Fourier. Biedermann et Dette (2001) ont proposé une approche minimax pour obtenir l'échantillonnage optimal, par rapport à l'IMSE asymptotique, pour un bruit i.i.d.

Dans plusieurs problèmes de pharmacocinétique (PK), où nous étudions l'action des médicaments et leurs utilisations thérapeutiques, les scientifiques sont souvent amenés à calculer la surface sous la courbe de concentration (AUC). Cette mesure représente l'exposition totale de l'organisme au médicament administré. Elle est d'intérêt lors du calcul de la biodisponibilité du médicament, qui mesure le taux et l'étendue où le médicament atteint le site d'action, mais cela dépend aussi du mode d'administration du médicament. L'AUC est l'intégrale de la courbe de concentration sur la durée d'observation.

Pour estimer l'AUC, deux approches différentes sont possibles. L'approche paramétrique, où nous estimons les paramètres de la courbe de concentration, l'AUC est alors l'intégrale de l'estimation plug-in de la courbe de concentration. Cette approche a été utilisée par plusieurs auteurs, nous citons, le livre de Davidian et Giltinan (1995). La deuxième approche est l'approche non paramétrique, où l'intégrale AUC est estimée directement sans l'utilisation de la concentration. La méthode la plus utilisée pour l'estimation de l'AUC est l'approximation de l'intégrale par des méthodes quadratures. La règle des trapèzes est généralement utilisée, car elle donne une bonne estimation de l'AUC lorsque la concentration décroît de façon exponentielle. Pour plus de détails sur les différentes procédures de quadrature numérique et une comparaison entre elles, nous renvoyons le lecteur aux travaux de Bailer et Piegorsch (1990).

Les temps d'observations jouant un rôle crucial dans l'efficacité de l'estimation de l'AUC, il est intéressant de trouver le plan d'échantillonnage optimal pour cette estimation. Plusieurs approches pour obtenir les points d'échantillonnage optimaux ont été proposées par plusieurs auteurs. Nous mentionnons, parmi d'autres, Katz et D'argenio (1983) qui ont proposé un algorithme pour minimiser l'erreur moyenne quadratique (MSE) par rapport aux points d'échantillonnage. Duffull *et al.* (2002) ont introduit l'algorithme de recuit simulé (SA) pour obtenir l'échantillonnage optimal et l'ont comparé à d'autres algorithmes d'optimisation. Les travaux antérieurs supposent que les observations ne sont pas corrélées, ce qui n'est pas toujours une hypothèse réaliste. Pour une fonction de régression linéaire paramétrique g, Belouni et Benhenni (2015), ont introduit le plan d'échantillonnage optimal pour l'estimation de l'AUC lorsque les erreurs sont corrélées. Ils ont utilisé l'algorithme SA pour générer l'échantillonnage optimale et l'ont comparé à l'échantillonnage uniforme.

Organization de la thèse

Description courte: Dans cette thèse, nous considérons le modèle de régression non paramétrique donné par (1), avec un processus d'autocovariance générale stationnaire ou non stationnaire. Nous examinons dans un premier temps le problème de l'estimation de la fonction de régression g, où nous analysons l'estimateur à noyau proposé par Gasser et Müller (1979). Nous proposons également deux nouveaux estimateurs à noyau pour la fonction g, à savoir l'estimateur des trapèzes construit à partir de la règle numérique des trapèzes, et l'estimateur de projection construit à l'aide des espaces de Hilbert à noyau reproduisant. Nous étudions leurs comportements asymptotiques lorsque n et m tendent vers l'infini, en termes de vitesses de convergence et de distributions asymptotiques. Nous menons également une étude de simulation pour tester leurs performances pour un ensemble fini d'observations. Deuxièmement, nous examinons le problème de plan d'échantillonnage optimal pour estimer la fonction de régression g. Enfin, nous considérons une application de l'estimation de la fonction de régression et de son AUC ainsi que le plan d'échantillonnage optimal.

Dans la suite, nous décrivons le contenu de la thèse chapitre par chapitre.

Chapitre 1. Dans ce chapitre, nous considérons le problème de l'estimation de la fonction de régression g dans le modèle donné par (1). Nous nous concentrons sur l'estimateur à noyau proposé par Gasser et Müller (1979) donné, pour $x \in [0,1]$, par:

$$\hat{g}_{n,h}^{GM}(x) = \frac{1}{h}\sum_{i=1}^{n} \overline{Y}(t_{i,n}) \int_{m_{i-1,n}}^{m_{i,n}} K\Big(\frac{x-t}{h}\Big)\, dt, \tag{2}$$

où K est un noyau de support $[-1,1]$, $h = h(n,m)$ est une fenêtre et les points intermédiaires $m_{i,n}$ sont donnés par: $m_{0,n} = 0$, $m_{n,n} = 1$ et $m_{i,n} = (t_{i,n} + t_{i+1,n})/2$.

Dans cette thèse, contrairement au processus d'erreur considéré dans Gasser et Müller (1979) et de Hart et Wherly (1986), nous adaptons le modèle d'erreur considéré dans Benhenni et Rachdi (2007). Ce processus d'erreur ε est un processus du second ordre, avec une fonction d'autocovariance non différentiable, tels que les processus de Wiener et d'Ornstein-Uhlenbeck. Nous considérons la suite d'échantillonnage réguliers $\{(t_{i,n})_{1\le i\le n}, n \ge 1\}$, générée par une fonction de densité f, définie par Sacks et Ylvisaker (1970) comme suit:

$$t_{i,n} = F^{-1}\Big(\frac{i}{n}\Big) \quad \text{pour} \quad i = 1, \cdots, n, \tag{3}$$

où F est la fonction de distribution de la fonction de densité f. Nous montrons que nous pouvons améliorer les taux de convergence de la variance et du biais de $\hat{g}_{n,h}^{GM}$ en utilisant la suite (3) au lieu d'une suite d'échantillonnage engendré par une densité uniforme, comme ce fut le cas dans Benhenni et Rachdi (2007). En effet, nous obtenons le taux $\frac{1}{n^2h}$ au lieu de $\frac{1}{n}$ pour le biais (respectivement le taux $\frac{1}{mn^3h^2} + \frac{1}{n^2}$ au lieu de $\frac{1}{mn}$ pour la variance), voir Proposition 1.3.1 et Proposition 1.3.2.

Nous obtenons également la fenêtre optimale par rapport à l'IMSE asymptotique. De plus, sous des hypothèses classiques, nous démontrons la normalité asymptotique de l'estimateur $\hat{g}_{n,h}^{GM}$ lorsque n et m tendent vers l'infini.

L'amélioration des taux de convergence de la variance et du biais nous a été très utile pour une comparaison théorique de la performance l'estimateur $\hat{g}_{n,h}^{GM}$, à notre nouvel estimateur proposé au Chapitre 2.

Chapitre 2. Dans ce chapitre, nous considérons le même modèle de régression que dans le Chapitre 1. Nous construisons un nouvel estimateur de la fonction de régression g, que nous appelons estimateur de projection. Cet estimateur, qui est aussi un estimateur linéaire à noyau, est construit en utilisant l'inverse de la matrice d'autocovariance des observations, que nous supposons connue et inversible. Il est basé sur une propriété de projection et il est donné comme suit pour $x \in [0,1]$ (voir Définition 2.3.1):

$$\hat{g}_n^{pro}(x) = \sum_{i=1}^{n} m_{x,h}(t_i)\overline{Y}(t_{i,n}), \tag{4}$$

où, en posant $T_n = (t_{i,n})_{1 \leq i \leq n}$, les poids $(m_{x,h}(t_{i,n}))_{1 \leq i \leq n}$ sont déterminés par:

$$m'_{x,h_{|T_n}} = f_{x,h_{|T_n}}{}' R^{-1}_{|T_n}, \quad \text{où} \;\; f_{x,h}(t) = \frac{1}{h}\int_0^1 R(s,t) K\Big(\frac{x-s}{h}\Big)\, ds,$$

avec $f_{x,h_{|T_n}} := (f_{x,h}(t_{1,n}), \ldots, f_{x,h}(t_{n,n}))'$, $R_{|T_n} := (R(t_{i,n}, t_{j,n}))_{1 \leq i,j \leq n}$, $R^{-1}_{|T_n}$ l'inverse de $R_{|T_n}$ et $m_{x,h_{|T_n}} := (m_{x,h}(t_{1,n}), \ldots, m_{x,h}(t_{n,n}))'$.

Pour certains processus d'erreur classiques, lorsque l'inverse de la matrice d'autocovariance $R^{-1}_{|T_n}$ est analytiquement connue (e.g. processus de Wiener, processus généralisé de Wiener et processus d'Ornstein-Uhlenbeck), nous donnons une expression simplifiée de l'estimateur proposé (voir Proposition 2.3.1 et Proposition 2.3.2).

La construction de cet estimateur s'inspire des travaux de Sacks et Ylvisaker (1966, 1968, 1970), Cependant leur contexte est différent du nôtre. Ils ont considéré le modèle linéaire paramétrique $g(t) = \beta w(t)$ où β est un paramètre réel inconnu et w est une fonction connue appartenant à l'espace de Hilbert à noyau autoreproduisant associé à la fonction d'autocovariance du processus d'erreur ε, noté RKHS (R). Ils ont également supposé que la matrice d'autocovariance est connue et inversible. Dans notre cas, l'estimateur est construit à partir de la fonction $f_{x,h}$ que nous avons pu démontrer qu'elle appartient à l'espace RKHS (R).

Un rappel détaillé est dédié aux nombreuses techniques du RKHS (R) que nous avons utilisées pour obtenir nos résultats théoriques.

Nous étudions la performance asymptotique de l'estimateur proposé lorsque n et m tendent vers l'infini. Les propriétés de RKHS (R) permettent non seulement d'obtenir l'expression asymptotique de la variance, mais également de trouver le taux de convergence optimal de la variance résiduelle de cet estimateur, voir Proposition 2.4.5. Nous obtenons la fenêtre optimale h^* au sens de l'IMSE asymptotique, c'est à dire,

$$\overline{\lim_{n,m\to\infty}} \frac{\text{IMSE}(h^*)}{\text{IMSE}(h_{n,m})} \leq 1, \tag{5}$$

pour une fenêtre $h_{n,m}$ vérifiant : $\lim\limits_{n,m\to\infty} h_{n,m} = 0$ et $\overline{\lim\limits_{n,m\to\infty}}\, m h^3_{n,m} < +\infty$.

En outre, nous prouvons la normalité asymptotique de l'estimateur proposé, voir le Théorème 2.4.3. Nous effectuons également une comparaison théorique entre le nouvel estimateur $\hat{g}_n^{pro}$ et l'estimateur classique de Gasser et Müller $\hat{g}_{n,h}^{GM}$. Nous montrons que la variance de $\hat{g}_n^{pro}$ est plus petite que celle de $\hat{g}_{n,h}^{GM}$, alors qu'ils sont tous les deux asymptotiquement non biaisés. Le fait que l'estimateur de projection présente une variance asymptotique plus faible peut être argumenté par deux points. Premièrement, l'estimateur de Gasser et Müller $\hat{g}_{n,h}^{GM}$ ne tient pas compte de la corrélation des observations. Deuxièmement, l'estimateur $\hat{g}_{n,h}^{GM}$ est une approximation d'une intégrale et que la meilleure approximation linéaire d'une intégrale est basée sur une propriété de projection, voir par exemple Benhenni et Cambanis (1992).

Enfin, nous menons une étude de simulation afin d'étudier les performances de l'estimateur proposé $\hat{g}_n^{pro}$ dans un ensemble fini d'échantillonnage, où nous démontrons ses bonnes performances pour les échantillons de petite taille. Ensuite, nous le comparons avec l'estimateur de Gasser et Müller $\hat{g}_{n,h}^{GM}$, pour différentes valeurs du nombre d'unités expérimentales m et différentes valeurs de la taille de l'échantillonnage n. Cette simulation confirme nos résultats théoriques.

Chapitre 3. Dans ce chapitre, nous construisons un estimateur simple à noyau pour la fonction de régression dans le modèle donné par (1). Pour motiver cette construction, nous considérons l'estimateur à noyau de la fonction g basé sur des observations continues sur l'intervalle $[0,1]$ est donné, pour tout $x \in [0,1]$, par

$$\hat{g}_{[0,1]}(x) = \frac{1}{h}\int_0^1 K\Big(\frac{x-t}{h}\Big)\overline{Y}(t)\, dt \quad \text{avec} \quad \overline{Y}(t) = \frac{1}{m}\sum_{j=1}^{m} Y_j(t), \tag{6}$$

pour un noyau K de support $[-1,1]$ et une fenêtre $h = h(n,m)$.

Nous renvoyons le lecteur aux travaux de Blanke et Bosq (2008) ou de Didi et Louani (2013) pour plus de détails sur l'estimateur à noyau de la fonction de régression basé sur des observations continues. Dans les cas pratiques, où nous n'avons accès qu'à des observations discrètes, nous appliquons la règle numérique des trapèzes pour approcher l'estimateur continu, afin de construire un nouvel estimateur plus simple. L'estimateur des trapèzes basé sur les observations $(t_{i,n}, \overline{Y}(t_{i,n}))_{1\le i\le n}$, où $(t_{i,n})_{1\le i\le n}$ est une suite de plan d'échantillonnage régulier $\{T_n\}_{n\ge 1}$, générée par une fonction de densité f, est donné pour $x \in [0,1]$ par

$$\hat{g}_n^{\text{trap}}(x) = \frac{1}{2n}\sum_{k=1}^{N_{T_n}-1}\left\{\left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k}) + \left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k+1})\right\}, \tag{7}$$

où $t_{x,1} < \cdots < t_{x,N_{T_n}}$ sont les points de T_n dans $[x-h, x+h]$, $\varphi_{x,h}(t) = \frac{1}{h}K\big(\frac{x-t}{h}\big)$, K est un noyau de support $[-1,1]$ et $h = h(n,m)$ est une fenêtre avec $0 < h < 1$.

Nous étudions les propriétés asymptotiques de l'estimateur proposé $\hat{g}_n^{\text{trap}}$, lorsque le nombre d'unités expérimentales m et le nombre d'observations n tendent vers l'infini. De plus, nous démontrons la normalité asymptotique de cet estimateur et nous déduisons la fenêtre optimale au sens de l'IMSE asymptotique, donnée par (5).

Pour cet estimateur $\hat{g}_n^{\text{trap}}$, nous dérivons le plan d'échantillonnage asymptotiquement optimal, généré par la fonction de densité f^* qui minimise l'IMSE asymptotique. Pour obtenir cette fonction f^*, nous minimisons le terme, dans l'IMSE, qui dépend de la fonction de densité de l'échantillonnage donnée par:

$$\int_0^1 \frac{\alpha(x)}{f^2(x)} w(x)\, dx \triangleq \Psi_{(\alpha,w)}(f),$$

où α est la fonction de saut de la dérivée du premier ordre de la fonction d'autocovariance R sur la diagonale, et w est une fonction de densité. Nous devons alors résoudre le problème d'optimisation suivant:

$$f^* \in \underset{f>0,\ \int_0^1 f(x)dx=1}{\operatorname{argmin}} \Psi_{(\alpha,w)}(f).$$

Ce problème d'optimisation est résolu pour la densité optimale suivante:

$$f^*(t) = \frac{\{\alpha(t)w(t)\}^{1/3}}{\int_0^1 \{\alpha(s)w(s)\}^{1/3}\, ds} 1_{[0,1]}(t).$$

De plus, nous démontrons que cette densité optimale f^* satisfait le critère d'optimalité minimax donné par Biedermann et Dette (2001) pour des observations indépendantes, au sens qu'elle est robuste par rapport à la misspécification de la fonction d'autocovariance comme suit:

$$f^* \in \underset{f>0, \int_0^1 f(t)dt=1}{\operatorname{argmin}} \max_{(\alpha,w)\in\Lambda} \Psi_{(\alpha,w)}(f), \tag{8}$$

où,

$$\Lambda = \left\{(\alpha, w) \in (C[0,1])^2 \mid \int_0^1 \alpha(t)dt < \epsilon_1, \left(\int_0^1 w(s)^{1/2} ds\right)^2 < \epsilon_2\right\},$$

pour $\epsilon_1 > 0$ et $\epsilon_2 > 0$ fixés.

Pour tester la performance de l'estimateur proposé pour des ensembles finis d'échantillons (petits n et m), nous avons mené une vaste étude de simulation. Nous montrons que la performance de l'estimateur proposé s'améliore en augmentant m. Nous comparons également l'estimateur des trapèzes $\hat{g}_n^{\text{trap}}$ avec l'estimateur de Gasser et Müller $\hat{g}_n^{\text{GM}}$ pour différentes valeurs de n et m et différents "degrés" de corrélation. Nous montrons que les deux estimateurs ont à peu près la même performance, au sens de l'IMSE.

Enfin, nous menons une étude de simulation pour montrer la réduction de l'IMSE lors de l'utilisation du plan d'échantillonnage optimal, au lieu du plan uniforme dans un ensemble d'échantillons finis. Pour cela, nous avons choisi une grande classe de fonctions d'autocovariance paramétriques, pour lesquelles la densité de l'échantillonnage optimale dépend d'un paramètre inconnu. Nous utilisons ensuite l'algorithme de recuit simulé généralisé (GSG) pour estimer ce paramètre et nous obtenons ainsi l'échantillonnage optimal estimé par la méthode plug-in. Les simulations montrent que les plans d'échantillonnages optimaux théoriques et estimés réduisent d'une manière significative l'IMSE.

Chapitre 4. Dans ce chapitre, plusieurs problèmes de pharmacocinétique sont étudiés pour des données corrélées (simulées ou réelles). Nous examinons d'abord le problème de l'estimation de la fonction de concentration d'un certain médicament administré dans l'organisme, pour lequel nous proposons d'utiliser l'estimateur non paramétrique à noyau au lieu d'utiliser des méthodes paramétriques. Nous utilisons l'estimateur de Gasser et Müller et nous prouvons ses bonnes performances à travers une étude de simulation et une analyse de données réelles. Les données sont les concentrations plasmatiques de digoxine après l'administration orale du traitement considéré par Wagner et Yates (1973). Ensuite, nous étudions le problème de l'estimation de l'AUC:

$$AUC(g) = \int_0^T g(t)dt,$$

où T est le dernière temps d'observation. Nous introduisons un nouvel estimateur à noyau qui est l'intégrale de l'estimateur de la fonction de régression. Nous montrons, à l'aide d'une étude de simulation, que l'estimateur proposé est plus performant que l'estimateur classique en terme d'erreur moyenne quadratique. Enfin, le problème crucial de trouver le plan d'échantillonnage optimal pour l'estimation de l'AUC est étudié à l'aide de l'algorithme GSA.

Chapitre 5. Dans ce chapitre, nous terminons ce manuscrit par des conclusions et nous présentons quelques questions ouvertes et perspectives apparues au cours de la préparation de cette thèse.

Introduction

General presentation

Nonparametric regression functions have been extensively used in the past decades, not only in statistics but, in several domain such as, medicine and signal processing. The nonparametric regression function, is a general function that translates the relation between two variables: the explanatory variable X and the response variable Y, without any form or parametric restrictions on this function. One of the situations that statisticians encounter in their studies is the estimation of the regression function, based on partial observations of this function. In statistical terms, one wants to estimate the function $g(\cdot) = \mathbb{E}(Y|X = \cdot)$ based on the observations $(X_i, Y_i)_{1\leq i\leq n}$, which are n copies of (X, Y). These observations are often modeled as follows:

$$Y_i = g(X_i) + \varepsilon_i,$$

where $(\varepsilon_i)_{1\leq i\leq n}$ are centered random variables, called errors.

In this thesis, we consider the case where $(X_i)_{1\leq i\leq n}$ are fixed within some domain and not random (fixed design). This is the case, for instance, of the longitudinal data when experimental units are observed through sampling time points, chosen prior to the experiment. As a life time example, in pharmacokinetics problems, the concentrations of some drug administrated in the organism are observed every half an hour during 24 hours. Hence, we consider the so-called fixed design regression model defined as follows:

$$Y(t_{i,n}) = g(t_{i,n}) + \varepsilon(t_{i,n}) \quad \text{for } i = 1, \cdots, n,$$

with $0 \leq t_{1,n} < t_{2,n} < \cdots < t_{n,n} \leq 1$. The most thoroughly discussed model has been the one with independent errors. We mention, among others, the works of Priestly and Chao (1972) , Benedetti (1977) and Gasser and Müller (1979). The previous authors proposed different kernel regression estimators. Gasser *et al.* (1984) used the kernel estimator in the case of an individual growth curve. Fan (1992) considered a weighted local linear regression in the case of independent observations. However, considering that the observations are independent is not always a realistic assumption. For instance, the heights observed on the same child are correlated. The temperature observations measured along the day are also correlated. For this, we focus on the nonparametric regression estimation problem where the observations are correlated. These models were used by several authors, we mention among others, the work of Hart and Wherly (1986) who considered stationary (correlated) errors, Núñez-Antón *et al.* and Ferreira *et al.* (1997) considered a specific class of nonstationary errors with parametric autocovariance functions, introduced by Núñez-Antón and Woodworth (1994) to study the efficacy of cochlear implants. For a review on the nonparametric regression in the presence of stationary correlated errors, see Opsomer *et al.* (2001).

When considering correlated observations, Hart and Wherly (1986) used a well-known kernel regression estimator proposed by Gasser and Müller (1979). They showed that the kernel estimator is not consistent, in the sense that, its variance does not tend to zero when the number of observations n tends to infinity. For this, it was necessary to consider several experimental units, say m, on which n observations are taken; then the estimator is consistent when m tends to infinity. For a nonstationary correlated errors, Benhenni and Rachdi (2007) considered the Gasser and Müller estimator. Another consequence of the presence of correlation, is the breakdown of several data based bandwidth selection method, such as the classical cross-validation criterion. For details on this issue, see for instance, Altman (1990), Chiu (1989), Hart (1991, 1994) and Opsomer *et al.* (2001).

In this thesis, and for this type of data, we consider the so-called fixed design regression model with repeated measurements, given as follows:

$$Y_j(t_{i,n}) = g(t_{i,n}) + \varepsilon_j(t_{i,n}), \quad \text{for } i = 1, \cdots, n \text{ and } j = 1, \cdots, m, \tag{1}$$

where $\{\varepsilon_j, j = 1, \cdots, m\}$ is a sequence of i.i.d. centered error processes with the same distribution as a process ε. The non correlation of the observations made on different experimental units is a natural assumption, since in general, the later are chosen independently. This type of data was investigated by several authors in the parametric situation, we mention among others, Potthoff and Roy (1964) who considered some dental measurements of 11 girls and 16 boys at 4 different ages, Rao (1965, 1966), Grizzle and Allen (1969), Ghosh *et al.* (1973) considered the nonparametric methods in longitudinal studies.

It should be noted here, that Model (1) can be written as follows:

$$\overline{Y}(t_{i,n}) = g(t_{i,n}) + \overline{\varepsilon}(t_{i,n}), \quad \text{for} \quad i = 1, \cdots, n,$$

where $\overline{Y}(t_{i,n})$ and $\overline{\varepsilon}(t_{i,n})$ are respectively the sample mean of $\{Y_j(t_{i,n}), j = 1, \cdots, m\}$ and $\{\varepsilon_j(t_{i,n}), j = 1, \cdots, m\}$. This shows that the problem of fitting the regression function g to m curves, could be brought to fitting it to the sample means $\overline{Y}(t_{i,n})$.

Another interesting problem concerning the estimation of the regression function g, is the derivation the optimal sampling design $\{t_{i,n}, i = 1, \cdots, n\}$. This problem has been extensively studied in the case of the parametric regression estimation problems, we mention the works of Sacks and Ylvisaker (1966) and Belouni and Benhenni (2013). These authors considered this problem for a particular class of regression functions, and obtained the optimal sampling design. Dette *et al.* (2016,a) constructed a pair of linear unbiased estimators with corresponding optimal designs to compare two regression curves estimated from two finite samples of dependent measurements. Zhigljavsky *et al.* (2010) proposed a new approach of designing experiments for the location model in the presence of correlation; they showed the weakly convergence of their design to the one proposed by Bickel and Herzberg (1979).

In the nonparametric case, the problem of the optimal design was less developped in the literature, especially when the observations are correlated. Müller (1984) considered the kernel regression estimator, and introduced the optimal design points when the errors are asymptotically uncorrelated, which is not always a realistic assumption for the longitudinal data for instance. He used a regular design sequence generated by a density function, and derived the optimal density that minimizes the asymptotic Integrated Mean Squared Error (IMSE). Faraway (1990) used the Priestly and Chao estimator, and Zhao and Yao (2012) used a general weight function estimator. They considered a sequential design for the estimation of the regression function, when the observations are independent. Efromovich (2008) considered the problem of optimal design for a conditional heteroscedasticity using the Fourier series approach. Biedermann and Dette (2001) proposed a minimax approach to obtain the optimal design with respect to the asymptotic IMSE, for i.i.d. noise.

In several problems of pharmakokinetics (PK), where we study the action of drugs and their therapeutic use, scientists are often brought to calculate the Area Under the concentration Curve (AUC). This measure represents the total exposure of the body to the administrated drug. It is of interest when calculating the bioavailability of the drug, which measures the rate and extent to which a drug reaches the site of action, but it also depends on the way of administration of the drug. The AUC is the integral of the concentration curve over the observation time.

In order to estimate the AUC, two different approaches are possible. The parametric approach, where we estimate the parameters of the concentration curve, the AUC estimation is than the integral of the plug-in estimation of the concentration curve. This approach was used by several authors, we mention among others, the book of Davidian and Giltinan (1995). The second approach is the nonparametric approach, where the AUC is estimated directly without the use of the concentration. The most used method for the estimation of the AUC, is the approximation of the integral by quadrature methods. The trapezoidal rule is commonly used, since it gives good AUC estimation when the concentration decreases exponentially. For more details on the different Newton-Cotes numerical quadrature procedures and a comparison between them, we refer the reader to the work of Bailer and Piegorsch (1990).

Since the observation times play a crucual role in the efficiency of the AUC estimation, it is interesting to find the optimal sampling times for this estimation. Several approaches to obtain the optimal sampling points were proposed by several authors. We mention, among others, Katz and D'argenio (1983), who proposed an algorithm to minimize the MSE with respect to the sampling points. Duffull *et al.* (2002) introduced the Simulated Annealing algorithm (SA) to obtain the optimal design, and compared it to several other optimization algorithms. The previous works suppose that the observations are uncorrelated, which is not always a realistic assumption. For a parametric linear regression function g, Belouni and Benhenni (2015), introduced the optimal sampling design for the AUC estimation when the errors are correlated. They used the SA algorithm to generate the optimal design and they compared it to the uniform design.

Organization of the dissertation

Short description: In this thesis, we consider the nonparametric regression model given by (1), with general correlation error process which maybe stationary or nonstationary. We consider first, the problem of estimating the regression function g, where we analyse the well-known kernel regression estimator proposed by Gasser and Müller (1979). We also propose two new kernel estimators for the function g, namely the trapezoidal estimator constructed from a numerical rule, and the projection estimator constructed using the Reproducing Kernel Hilbert spaces. We study their asymptotic behaviors when n and m tend to infinity, in terms of rates of convergence and asymptotic distributions. We also conduct a simulation study to test their performances for a finite set of observations. Second, we consider the problem of finding the optimal sampling design for estimating the regression function g. Finally, we consider an application of the regression function estimation and its AUC along with the optimal sampling design.

In the sequel, we describe the content of the thesis chapter by chapter.

Chapter 1. In this chapter, we consider the problem of estimating the regression function g in the model given by (1). We consider the well-known kernel estimator proposed by Gasser and Müller (1979) given, for $x \in [0,1]$, by:

$$\hat{g}_{n,h}^{GM}(x) = \frac{1}{h}\sum_{i=1}^{n} \overline{Y}(t_{i,n}) \int_{m_{i-1,n}}^{m_{i,n}} K\Big(\frac{x-t}{h}\Big)\, dt, \tag{2}$$

where K is a kernel of support $[-1,1]$, $h = h(n,m)$ is a bandwidth and the midpoints $m_{i,n}$ are given by: $m_{0,n} = 0,\ m_{n,n} = 1$ and $m_{i,n} = (t_{i,n} + t_{i+1,n})/2$.

Here, the error process is different from that in Gasser and Müller (1979) and Hart and Wherly (1986) , but it is as in Benhenni and Rachdi (2007), i.e. ε is a second order process, with a non differentiable covariance function, such as the Wiener process and the Ornstein Uhlenbeck process. We consider the regular sequence of designs $\{(t_{i,n})_{1\le i\le n}, n \ge 1\}$, generated by a density function f, which was defined by Sacks and Ylvisaker (1970) as follows:

$$t_{i,n} = F^{-1}\Big(\frac{i}{n}\Big) \quad \text{for} \quad i = 1, \cdots, n, \tag{3}$$

where F is the distribution function of the density function f. We show that, we can improve the rates of convergence of the variance and the bias of $\hat{g}_{n,h}^{GM}$ when using (3), instead of other sequence of designs, as it was the case in Benhenni and Rachdi (2007). In fact, we obtain the rate $\frac{1}{n^2 h}$ instead of $\frac{1}{n}$ for the bias (respectively the rate $\frac{1}{mn^3h^2} + \frac{1}{n^2}$ instead of $\frac{1}{mn}$ for the variance), see Propositions (1.3.1) and (1.3.2).

We also derive the optimal bandwidth with respect to the asymptotic IMSE. In addition, we prove the asymptotic normality of the estimator $\hat{g}_{n,h}^{GM}$ when n and m tend to infinity, under classical assumptions.

The improvement of the rates of convergence of the variance and the bias, were very useful for deriving a theoretical comparison of the estimator $\hat{g}_{n,h}^{GM}$, to our new proposed estimator, given in Chapter 2.

Chapter 2. In this chapter, we consider the same regression model as in Chapter 1. We construct a new estimator of the regression function g. This estimator, which is also a linear kernel estimator, is constructed using the inverse of the autocovariance matrix of the observations, that we assume known and invertible. It is based on a projection property and is given, for $x \in [0,1]$, as follows (see Definition (2.3.1)):

$$\hat{g}_n^{pro}(x) = \sum_{i=1}^{n} m_{x,h}(t_i)\overline{Y}(t_{i,n}), \tag{4}$$

where the weights $(m_{x,h}(t_{i,n}))_{1\le i\le n}$ are being determined, letting $T_n = (t_{i,n})_{1\le i\le n}$, by:

$$m'_{x,h_{|T_n}} = f_{x,h_{|T_n}}{}' R_{|T_n}^{-1}, \quad \text{where} \quad f_{x,h}(t) = \frac{1}{h}\int_0^1 R(s,t) K\Big(\frac{x-s}{h}\Big)\, ds,$$

with $f_{x,h_{|T_n}} := (f_{x,h}(t_{1,n}), \ldots, f_{x,h}(t_{n,n}))'$, $R_{|T_n} := (R(t_{i,n}, t_{j,n}))_{1\le i,j\le n}$, $R_{|T_n}^{-1}$ the inverse of $R_{|T_n}$ and $m_{x,h_{|T_n}} := (m_{x,h}(t_{1,n}), \ldots, m_{x,h}(t_{n,n}))'$.

For some classical error processes, when the inverse of the autocovariance matrix is analytically known, we give a simplified expression of the proposed estimator, such as for the Wiener process, the generalized Wiener and the Ornstein-Uhlenbeck process (see Propositions (2.3.1) and (2.3.2)).

This estimator was inspired by the work of Sacks and Ylvisaker (1966, 1968, 1970) but there context is different than ours. They considered the parametric linear model with $g(t) = \beta w(t)$ where β is an unknown real parameter and w is a known function belonging to the Reproducing Kernel Hilbert Space associated to the autocovariance function of the error process ε, denoted by RKHS(R). They also assumed that the autocovariance matrix is known and invertible. In our case the function, through which the estimator is constructed, $f_{x,h}$ is proven to belong to RKHS(R).

A detailed recall is dedicated to the many techniques of the RKHS(R) that we used to obtain our theoretical results.

We investigate the asymptotic performance of the proposed estimator, when n and m tend to infinity. The properties of the RKHS(R) allow not only to obtain the asymptotic expression of the variance, but also to find the optimal rate of convergence of the residual variance of this estimator (see Proposition (2.4.5)). We derive the optimal bandwidth h^* with respect to the asymptotic IMSE, optimality in the sense that,

$$\overline{\lim_{n,m\to\infty}} \frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})} \leq 1, \tag{5}$$

for any sequence of bandwidths $h_{n,m}$ verifying: $\lim\limits_{n,m\to\infty} h_{n,m} = 0$ and $\overline{\lim\limits_{n,m\to\infty}} m h_{n,m}^3 < +\infty$.

In addition, we prove the asymptotic normality of the proposed estimator, see Theorem 2.4.3. We also give a theoretical comparison between the new estimator $\hat{g}_n^{pro}$ and Gasser and Müller's estimator $\hat{g}_{n,h}^{GM}$, proving that $\hat{g}_n^{pro}$ has an asymptotically smaller variance than $\hat{g}_{n,h}^{GM}$, whereas they are both asymptotically unbiased. The fact that the projection estimator has a smaller asymptotic variance can be argued by, in the one hand, the Gasser and Müller's estimator $\hat{g}_{n,h}^{GM}$ does not take into account the correlation requirement. On the other hand, the estimator $\hat{g}_{n,h}^{GM}$ is an approximation of an integral, and the best linear approximation of an integral is based on some projection property, see for instance Benhenni and Cambanis (1992).

Finally, we conduct a simulation study in order to investigate the performance of the proposed estimator $\hat{g}_n^{pro}$ in a finite sample set, where we prove its good performance for small sample sizes. Next we compare it with the Gasser and Müller's estimator $\hat{g}_{n,h}^{GM}$ for different values of the number of experimental units m and different values of the sample size n. This simulation confirms our theoretical results.

Chapter 3. In this chapter, we construct a simple kernel estimator for the regression function in the model given by (1). To motivate this construction, we consider the kernel estimator of g based on continuous observations on $[0, 1]$, given for any $x \in [0, 1]$ by,

$$\hat{g}_{[0,1]}(x) = \frac{1}{h} \int_0^1 K\Big(\frac{x-t}{h}\Big) \overline{Y}(t)\, dt \quad \text{with} \quad \overline{Y}(t) = \frac{1}{m} \sum_{j=1}^{m} Y_j(t), \tag{6}$$

for a kernel K of support $[-1, 1]$ and a bandwidth $h = h(n, m)$. We refer the reader to the works of Blanke and Bosq (2008) or Didi and Louani (2013) for more details on the kernel estimation of the regression function based on continuous observations. In practical cases, where we only have access to discrete observations, we apply the trapezoidal rule to approximate the continuous estimator to construct a new simpler estimator. The trapezoidal estimator based on the observations $(t_{i,n}, \overline{Y}(t_{i,n}))_{1\leq i\leq n}$ where $(t_{i,n})_{1\leq i\leq n}$ is a regular sequence of designs generated by some density function f is given, for $x \in [0,1]$, by

$$\hat{g}_n^{\text{trap}}(x) = \frac{1}{2n} \sum_{k=1}^{N_{T_n}-1} \left\{ \left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k}) + \left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k+1}) \right\}, \tag{7}$$

where $t_{x,1} < \cdots < t_{x,N_{T_n}}$ are the points of T_n in $[x-h, x+h]$, $\varphi_{x,h}(t) = \frac{1}{h}K\left(\frac{x-t}{h}\right)$, K is a kernel of support $[-1,1]$ and $h = h(n,m)$ is a bandwidth with $0 < h < 1$.

We investigate the asymptotic properties of the proposed estimator $\hat{g}_n^{\text{trap}}$, when both the number of experimental units m and the number of observations n tend to infinity. In addition, we prove the asymptotic normality of this estimator and we derive the optimal bandwidth with respect to the asymptotic IMSE, as given by (5).

For this estimator, $\hat{g}_n^{\text{trap}}$, we derive the asymptotic optimal sampling design, generated by the density function f^* that minimizes the asymptotic IMSE. To obtain this function f^*, we minimize the term of IMSE depending on the design density given by:

$$\int_0^1 \frac{\alpha(x)}{f^2(x)} w(x)\, dx \triangleq \Psi_{(\alpha,w)}(f),$$

where α is the jump function of the first order derivative of the autocovariance R at the diagonal, and w is any density function. We have then to solve the following optimization problem:

$$f^* \in \underset{f>0,\ \int_0^1 f(x)dx=1}{\text{argmin}} \Psi_{(\alpha,w)}(f).$$

This optimization problem is solved for the following optimal design density:

$$f^*(t) = \frac{\{\alpha(t)w(t)\}^{1/3}}{\int_0^1 \{\alpha(s)w(s)\}^{1/3}\, ds} 1_{[0,1]}(t).$$

Moreover, we prove that this optimal density f^* satisfies a minimax optimality criterion given by Biedermann and Dette for the independent observations, in the sense that it is robust with respect to the misspecification of the error's autocovariance function as follows:

$$f^* \in \underset{f>0, \int_0^1 f(t)dt=1}{\text{argmin}} \ \max_{(\alpha,w)\in\Lambda} \Psi_{(\alpha,w)}(f), \tag{8}$$

where, for fixed $\epsilon_1 > 0$ and $\epsilon_2 > 0$,

$$\Lambda = \left\{(\alpha, w) \in (C[0,1])^2 \ / \int_0^1 \alpha(t)dt < \epsilon_1, \left(\int_0^1 w(s)^{1/2} ds\right)^2 < \epsilon_2\right\}.$$

To test the performance of the proposed estimator in a finite sample sets (small n and m), we conduct an extensive simulation study. We show that the performance of the proposed estimator gets better as m increases. We also compare the trapezoidal estimator $\hat{g}_n^{\text{trap}}$ with the Gasser and Müller estimator $\hat{g}_n^{\text{GM}}$ for different values of n and m and different "degree" of correlation. We show that, both of the estimators have approximately the same performance, with respect to the IMSE.

Finally, we run a simulation study to show the reduction of the exact IMSE when using the optimal sampling design, instead of the uniform design in a finite sample set. For this, we chose a large class of parametric autocovariance functions, where the optimal design density depends on the autocovariance unknown parameter. We then use the Generalized Simulated Annealing algorithm (GSA) to estimate the parameter and so we obtain the plugin estimated optimal design. The simulations show that both the theoretical and the estimated optimal design reduce the IMSE significantly.

Chapter 4. In this chapter, several pharmacokinetics problems are investigated, for correlated data (simulated or real). We first consider the problem of estimating the concentration function of some administrated drug, where we propose to use the nonparametric kernel estimator instead of the parametric methods. We use the Gasser and Müller estimator, and we prove its good performance via simulation study and real data analysis. The data are digoxin plasma concentrations after an oral administration of a treatment, considered by Wagner and Yates (1973). Second, we investigate the problem of estimating the AUC:

$$AUC(g) = \int_0^T g(t)dt,$$

where T is the scientist's last sampling time. We introduce a new kernel estimator, which is the integration of the regression function estimator. We prove, using a simulation study, that the proposed estimators outperform the classical one in terms of Mean Squared Error. Finally, the crucial problem of finding the optimal sampling design for the AUC estimation is investigated using the GSA algorithm.

Chapter 5. In this chapter, we give a conclusion of the thesis and we present some opened questions and perspectives which appeared during the preparation of this thesis.

Author's contributions

Affiliation

- Currently a full time temporary research and teaching assistant (ATER) at "UFR Sciences de l'Homme et de la Société" in Grenoble Alpes University.

Publications

1. Benelmadani D, Benhenni K, Louhichi S. (2019) Trapezoidal rule and sampling designs for the nonparametric estimation of the regression function in models with correlated errors. Published in Annals of the Institute of Statistical Mathematics.

2. Benelmadani D, Benhenni K, Louhichi S. (2019) Optimal design for the nonparametric regression estimation applied to pharmacokinetic problems. Published in the proceedings of the 9th World Conference on Sampling and Blending. 329-337.

3. Benelmadani D, Benhenni K, Louhichi S. (2018) The reproducing kernel Hilbert space approach in nonparametric regression problems with correlated observations. In revision. Available at `https://arxiv.org/abs/1809.03754`.

Seminars and conferences

1. "Plans d'échantillonnage pour l'estimation non paramétrique de la fonction de régression pour des données dépendantes". Séminaire des doctorants du laboratoire Jean Kuntzmann. Grenoble, December-2015.

2. "Analyse de la fonction de régression dans des modèles non paramétriques et applications en pharmacologie". Septième Rencontre des jeunes statisticiens. Porquerolles, April-2017.

3. "Nonparametric estimation of the regression function in models with correlated errors". Séminaire des doctorants du laboratoire Jean Kuntzmann. Grenoble, November -2017.

4. "Estimation non paramétrique de la fonction de régression pour des modèles avec des erreurs corrélées". Cinquantième journées de statistiques. Paris Saclay, May-2018.

5. Poster: L'échantillonnage optimale pour l'estimation non paramétrique de la fonction de régression dans le cas des données corrélées. Journées MAS 2018, Processus et échantillonnage. Dijon, August-2018.

6. "Optimal design for the nonparametric regression estimation applied to pharmacokinetic problems". The 9th World on Sampling and Blending. Beijing, May-2019.

Projects

- Member of the project "Optimal trend estimation under errors with Matérn-type Autocovariance and applications to environmental data", funded by Grenoble Alpes Data Institute.

Chapter 1

Kernel estimation of the regression function in models with general correlated errors

Abstract: In this chapter, we consider a well-known kernel regression estimator proposed by Gasser and Müller (1979). We investigate its asymptotic behavior, when the number of experimental units and the number of observations tend to infinity. Our analysis of the asymptotic behavior concerns the case where the sampling points form a regular sequence of designs, as defined by Sacks and Ylvisaker (1970). Using this sequence of designs, we improve the rates of convergence of the bias and the variance, already studied by Benhenni and Rachdi (2007). The results of this chapter will be very useful to study the performance of a new kernel estimator, introduced in Chapter 3.

Key words: *Nonparametric regression, kernel estimator, correlated observations, repeated measurements, asymptotic normality.*

Résumé: Dans ce chapitre, nous considérons le célèbre estimateur à noyau de la fonction de régression proposé par Gasser et Müller en (1979). Nous examinons son comportement asymptotique, quand le nombre d'unités expérimentales et le nombre d'observations tendent vers l'infini. Notre analyse se porte sur le cas où le plan d'échantillonnage est régulier, défini par Sacks et Ylvisaker (1970). En utilisant ce plan, nous améliorons les vitesses de convergence du biais et de la variance, déjà étudiés par Benhenni et Rachdi (2016). Les résultats de ce chapitre seront très utiles pour étudier la performance d'un nouvel estimateur à noyau, introduit dans le chapitre 3.

Mots clés: *Régression non-paramétrique, estimateur à noyau, observations corrélées, mesures répétées, normalité asymptotique.*

1.1 Introduction, model and estimator

Suppose we have m experimental units, each of them is observed at n different design points (say $0 \leq t_1 < t_2 < \cdots < t_n \leq 1$). We consider the following so-called fixed design regression model, for $j = 1, \ldots, m$ and $i = 1, \ldots, n$,

$$Y_j(t_i) = g(t_i) + \varepsilon_j(t_i), \tag{1.1}$$

where g is the unknown regression function on $[0,1]$ and $\{\varepsilon_j(t), t \in [0,1]\}_j$ is a sequence of error processes. We assume that $g \in C^2([0,1])$ and that $(\varepsilon_j)_j$ are i.i.d. processes with the same distribution as a centered second order process ε, of autocovariance function R.

The kernel estimator, which will underlie the discussion in this chapter, is proposed by Gasser and Müller (1979) and is given, for any $x \in [0,1]$, by,

$$\hat{g}_n^{\mathrm{GM}}(x) = \sum_{i=1}^{n} \overline{Y}(t_i) \int_{m_{i-1}}^{m_i} \varphi_{x,h}(s)\, ds, \tag{1.2}$$

where $\overline{Y}(t_i) = \frac{1}{m}\sum_{j=1}^{m} Y_j(t_i)$, $\varphi_{x,h}(t) = \frac{1}{h}K(\frac{x-t}{h})$, K is a first order kernel of support $[-1,1]$, h is a bandwidth $(0 < h < 1)$ and $(m_i)_{1\leq i\leq n}$ is a sequence of midpoints defined as follows,

$$m_0 = 0, \quad m_i = \frac{t_i + t_{i+1}}{2} \quad \text{for} \quad i = 1, \ldots, n-1 \quad \text{and} \quad m_n = 1. \tag{1.3}$$

In this chapter, we investigate the asymptotic behavior of the estimator $\hat{g}_n^{\mathrm{GM}}$ when n and m tend to infinity. We derive higher order rates of convergence for the bias and the variance, in the case of the so-called regular sequence of designs $(t_{i,n})_{n\geq 1}$, defined by Sacks and Ylvisaker (1970) (see Definition 1.1.1); Than the ones obtained by Benhenni and Rachdi (2007), in the case of a Uniform design and correlated observations.

Definition 1.1.1 *Let F be a distribution function of some density f, with $\inf_{t\in[0,1]} f(t) > 0$ and $\sup_{t\in[0,1]} f(t) < \infty$. The so-called regular sequence of designs generated by a density f is defined by,*

$$T_n = \left\{ t_{i,n} = F^{-1}\left(\frac{i}{n}\right),\ i = 1, \ldots, n \right\} \quad \textit{for } n \geq 1.$$

1.2 Assumptions

In order to derive the asymptotic results, the following assumptions on the autocovariance function R and the kernel K are required.

(A) The autocovariance function R exists and is continuous on the square $[0,1]^2$.

(B) At the diagonal (i.e. when $t = s$ in the unit square), R has continuous left and right first-order derivatives, that is:

$$R^{(0,1)}(t,t^-) = \lim_{s\uparrow t} \frac{\partial R(t,s)}{\partial s} \quad \text{and} \quad R^{(0,1)}(t,t^+) = \lim_{s\downarrow t} \frac{\partial R(t,s)}{\partial s}.$$

The jump function along the diagonal $\alpha(t) \triangleq R^{(0,1)}(t,t^-) - R^{(0,1)}(t,t^+)$ is assumed to be continuous and not identically equal to zero.

(C) Off the diagonal (i.e. when $t \neq s$ in the unit square), R is assumed to have Lipshitz mixed partial derivatives up to order two which satisfies:

$$A^{(i,j)} \triangleq \sup_{0 \leq t \neq s \leq 1} |R^{(i,j)}(t,s)| < \infty \text{ for all integers } i,j \text{ such that } 0 \leq i+j \leq 2.$$

The previous assumptions are classical and were used in several works, see for instance, the works of Sacks and Ylvisaker (1966, 1968, 1970) and Benhenni and Rachdi (2007).

(D) The Kernel K is of support $[-1,1]$, at least in $C^2([-1,1])$, even and the second derivative K'' is Lipschitz.

Example 1.2.1 *Examples of processes with autocovariances satisfying Assumptions* $(A), (B)$ *and* (C) *are given as follows:*

1. *The Wiener process with autocovariance function* $R(s,t) = \sigma^2 \min(s,t)$*, has a constant jump function* $\alpha(t) = \sigma^2$ *and* $R^{(i,j)}(s,t) = 0$ *for all* i,j *such that* $i+j=2$ *and* $s \neq t$*.*

2. *The Ornstein-Uhlenbeck process with a stationary autcovariance* $R(s,t) = \sigma^2 e^{(-\lambda|s-t|)}$ *for* $\sigma > 0$ *and* $\lambda > 0$*. For this process* $\alpha(t) = 2\sigma^2\lambda$ *and* $R^{(0,2)}(s,t) = \sigma^2\lambda^2 e^{(-\lambda|s-t|)}$*.*

3. *A generalization of the Ornstein-Uhlenbeck process to a process with a nonstationary autocovariance function of the form:* $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ *for* $\sigma > 0$*,* $\lambda > 0$ *and* $0 < \rho < 1$*. For this process, the jump function is not constant and given by* $\alpha(t) = -2\sigma^2 ln(\rho) t^{\lambda-1}$*.*

4. *Sacks and Ylvisaker (1966) gave another general class of convex stationary autcovariance functions of the form,*

$$R(s,t) = \int_0^{1/|t-s|} (1 - \mu|t-s|) p(\mu) \, d\mu,$$

where p *is a probability density and* p'*, its derivative, are such that,*

$$\lim_{\mu \to \infty} \mu^3 p(\mu) < \infty, \quad \textit{and} \quad \int_a^\infty (\mu p'(\mu) + 3p(\mu))^2)\mu^6 d\mu < \infty,$$

for some finite constant a*. For this autocovariance function,* $\alpha(t) = 2\int_0^\infty \mu p(\mu) \, d\mu$ *for all* t*.*

Example 1.2.2 *Example of kernels satisfying Assumption* (D) *are given as follows:*

1. *The Quadratic kernel defined by* $K(u) = \frac{15}{16}(1-u^2)^2 \, \mathbb{1}_{\{|u| \leq 1\}}$*.*

2. *The Triweight kernel defined by* $K(u) = \frac{35}{32}(1-u^2)^3 \, \mathbb{1}_{\{|u| \leq 1\}}$*.*

1.3 Asymptotic results

The following propositions give the asymptotic expressions of the bias and the variance of the estimator $\hat{g}_n^{\mathrm{GM}}$.

Proposition 1.3.1 *Suppose that the kernel K satisfies Assumption (D). Moreover, assume that $f \in C^2([-1,1])$. If $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$ then for any $x \in]0,1[$,*

$$\mathbb{E}\left(\hat{g}_n^{GM}(x)\right) - g(x) = \frac{B}{2} h^2 g''(x) + o(h^2) + O\Big(\frac{1}{n^2 h}\Big),$$

where $B = \int_{-1}^{1} t^2 K(t) dt$.

Proposition 1.3.2 *Suppose that Assumptions (A)-(D) are satisfied and that $f \in C^2([-1,1])$. If $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$ then, for any $x \in]0,1[$,*

$$\mathrm{Var}\ \hat{g}_n^{GM}(x) = \frac{1}{m}\Big(R(x,x) - \frac{1}{2}\alpha(x) C_K h\Big) + o\Big(\frac{h}{m}\Big) + O\Big(\frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big),$$

where $C_K = \int_{-1}^{1}\int_{-1}^{1} |u-v| K(u) K(v) du dv$.

A comparison of the previous propositions to Theorem 2 of Benhenni and Rachdi (2007) yields that: the rate of convergence of the bias (respectively the variance) of the Gasser and Müller estimator to its limit, can be improved when using a regular sequence of designs. That is, we obtain the rate $\frac{1}{n^2h}$ instead of $\frac{1}{n}$ for the bias (respectively the rate $\frac{1}{mn^3h^2} + \frac{1}{n^2}$ instead of $\frac{1}{mn}$ for the variance).

Propositions 1.3.1 and 1.3.2 allow to derive the asymptotic expression of the mean squared error (MSE) of the estimator (1.2). The integrated mean squared error (IMSE) is then obtained by integrating the MSE with respect to a weight function w. The results are announced without proof in the following theorem, since it is a trivial consequence of the two propositions.

Theorem 1.3.1 *If all the assumptions of Propositions 1.3.1 and 1.3.2 are satisfied then for any $x \in]0,1[$,*

$$\begin{aligned}
\mathrm{MSE}(\hat{g}_n^{GM}(x)) &= \frac{1}{m}\Big(R(x,x) - \frac{1}{2}\alpha(x) C_K h\Big) + \frac{1}{4} h^4 (g''(x))^2 B^2 + o\Big(h^4 + \frac{h}{m}\Big) \\
&\quad + O\Big(\frac{h}{n^2} + \frac{1}{n^4h^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big). \\
\mathrm{IMSE}(\hat{g}_n^{GM}) &= \frac{1}{m}\int_0^1 R(x,x) w(x)\, dx - \frac{C_K h}{2m}\int_0^1 \alpha(x) w(x)\, dx \\
&\quad + \frac{B^2}{4} h^4 \int_0^1 [g''(x)]^2 w(x)\, dx + o\Big(h^4 + \frac{h}{m}\Big) \\
&\quad + O\Big(\frac{h}{n^2} + \frac{1}{n^4h^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big),
\end{aligned}$$

where w is a continuous density function, B and C_K are defined in the two propositions above.

The asymptotic optimal bandwidth is obtained by minimizing the asymptotic IMSE and is given by the following proposition.

Proposition 1.3.3 (Optimal bandwidth) *Suppose that the assumptions of Theorem 1.3.1 are satisfied. Moreover assume that $\frac{m}{n} = O(1)$ as $n, m \to \infty$. Denote by IMSE(h) the IMSE of the Gasser and Müller estimator when the bandwidth h is used. Then the bandwidth,*

$$h^* = \left(\frac{C_K \int_0^1 \alpha(x) w(x)\, dx}{2B^2 \int_0^1 [g''(x)]^2 w(x)\, dx}\right)^{1/3} m^{-1/3}, \tag{1.4}$$

is optimal in the sense that,

$$\overline{\lim_{n,m\to\infty}} \frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})} \leq 1,$$

for any sequence of bandwidths $h_{n,m}$ verifying:

$$\lim_{n,m\to\infty} h_{n,m} = 0 \quad \textit{and} \quad \overline{\lim_{n,m\to\infty}} \; m h_{n,m}^3 < +\infty,$$

where B and C_K are given in Propositions 1.3.1 and 1.3.2.

Finally, the following theorem gives the asymptotic normality of the Gasser and Müller estimator (1.2) for correlated errors.

Theorem 1.3.2 (Asymptotic normality) *Suppose that the assumptions of Theorem 1.3.1 are satisfied. If $\lim_{m\to\infty} \sqrt{m}h^2 = 0$ and $\frac{m}{n} = O(1)$ as $n, m \to \infty$ then for any $x \in]0,1[$,*

$$\sqrt{m}\Big(\hat{g}_n^{GM}(x) - g(x)\Big) \xrightarrow{\mathscr{D}} Z, \quad \textit{with } Z \sim \mathcal{N}(0, R(x,x)),$$

where $\mathscr{D}$ denotes the convergence in distribution and $\mathcal{N}$ is the normal distribution.

1.4 Proofs

Proof of Proposition 1.3.1.

Let $T_n = \{t_{i,n}, 1 \leq i \leq n\}$ (for the sake of clarity, we shall omit the n in $t_{i,n}$ when there is no ambiguity). Since $\mathbb{E}(\overline{Y}(t_i)) = g(t_i)$ for $i = 1, \ldots, n$ then for any $x \in]0,1[$ we have,

$$\mathbb{E}(\hat{g}_n^{GM}(x)) = \sum_{i=1}^{n} g(t_i) \int_{m_{i-1}}^{m_i} \varphi_{x,h}(s)\; ds.$$

Let $N_{T_n} = \text{Card } I_{x,h} = \text{Card}\{i = 1, \cdots, n / [m_{i-1}, m_i] \cap]x-h, x+h[\neq \emptyset\}$. The definition of $(m_i)_{1\leq i\leq n}$ yields that $N_{T_n} \geq 1$. Let $t_{x,i}$ be the points of T_n for which $i \in I_{x,h}$. For h small enough and without loss of generality, we take the following notation:

$$0 < m_{x,0} < x - h \leq m_{x,1} < \cdots < x + h \leq m_{x,N_{T_n}} < 1. \tag{1.5}$$

Thus, since $\varphi_{x,h}(s) = 0$ for $s \notin]x-h, x+h[$ then,

$$\mathbb{E}(\hat{g}_n^{GM}(x)) = \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\; ds.$$

Let,

$$l_h(x) \triangleq \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\; ds = \int_{m_{x,0}}^{m_{x,N_{T_n}}} \varphi_{x,h}(s) g(s)\; ds = \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s) g(s)\; ds,$$

and write,

$$\mathbb{E}(\hat{g}_n^{GM}(x)) = \mathbb{E}(\hat{g}_n^{GM}(x)) - l_h(x) + l_h(x) \triangleq \Delta_{x,h} + l_h(x). \tag{1.6}$$

We first control $\Delta_{x,h}$. We have,

$$\Delta_{x,h} = \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} (g(t_{x,i}) - g(s))\varphi_{x,h}(s)\, ds. \tag{1.7}$$

Recall that g and $\varphi_{x,h}$ are in C^2, then Taylor expansions of g and $\varphi_{x,h}$ for s in $]m_{x,i-1}, m_{x,i}[$ around $t_{x,i}$ yield,

$$g(s) = g(t_{x,i}) + (s - t_{x,i})g'(t_{x,i}) + \frac{1}{2}(s - t_{x,i})^2 g''(\theta_{x,i}),$$

and,

$$\varphi_{x,h}(s) = \varphi_{x,h}(t_{x,i}) + (s - t_{x,i})\varphi'_{x,h}(t_{x,i}) + \frac{1}{2}(s - t_{x,i})^2 \varphi''_{x,h}(s_{x,i}),$$

for some $\theta_{x,i}$ and $s_{x,i}$ between $t_{x,i}$ and s. Injecting these expansions in (1.7) gives,

$$\begin{aligned}
\Delta_{x,h} = & \sum_{i=1}^{N_{T_n}} g'(t_{x,i})\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (t_{x,i} - s)\, ds - \frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s - t_{x,i})^2 g''(\theta_{x,i})\, ds \\
& - \sum_{i=1}^{N_{T_n}} \varphi'_{x,h}(t_{x,i}) g'(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s - t_{x,i})^2\, ds - \frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi'_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s - t_{x,i})^3 g''(\theta_{x,i})\, ds. \\
& - \frac{1}{2}\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s - t_{x,i})^3 \varphi''_{x,h}(s_{x,i})\, ds - \frac{1}{4}\sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} (s - t_{x,i})^4 g''(\theta_{x,i})\varphi''_{x,h}(s_{x,i})\, ds.
\end{aligned} \tag{1.8}$$

To control these terms, we shall use the following lemma announced without proof (the proof is similar to that of Lemma 1 in Benelmadani *et al.* (2019a)).

Lemma 1.4.1 *Let $T_n = \{t_{i,n}, i = 1, \cdots, n\}$ for $n \geq 1$ be a regular sequence of designs (see Definition 1.1.1) and let $M_n = \{m_i, 0 \leq i \leq n\}$ (m_i are the midpoints defined by (1.3)). Suppose that $M_n \cap [x - h, x + h] \neq \emptyset$. If $nh \geq 1$ then,*

$$\sup_{0 \leq j \leq n} (t_{j+1,n} - t_{j,n}) = O\Big(\frac{1}{n}\Big) \quad \textit{and} \quad N_{T_n} = O(nh). \tag{1.9}$$

Recall that g' and g'' are both bounded and that for some appropriate constants c_j for $j = 0, 1, 2$,

$$\sup_{0 \leq t \leq 1} |\varphi^{(j)}_{x,h}(t)| \leq \frac{c_j}{h^{j+1}}. \tag{1.10}$$

On the one hand, we get using (1.9) and (1.10),

$$\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s-t_{x,i})^2 g''(\theta_{x,i})\ ds = O\Big(\frac{1}{n^2}\Big). \tag{1.11}$$

$$\sum_{i=1}^{N_{T_n}} \varphi'_{x,h}(t_{x,i}) g'(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s-t_{x,i})^2\ ds = O\Big(\frac{1}{n^2 h}\Big). \tag{1.12}$$

$$\sum_{i=1}^{N_{T_n}} \varphi'_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s-t_{x,i})^3 g''(\theta_{x,i})\ ds = O\Big(\frac{1}{n^3 h}\Big). \tag{1.13}$$

$$\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} (s-t_{x,i})^3 \varphi''_{x,h}(s_{x,i})\ ds = O\Big(\frac{1}{n^3 h^2}\Big). \tag{1.14}$$

$$\sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} (s-t_{x,i})^4 g''(\theta_{x,i}) \varphi''_{x,h}(s_{x,i})\ ds = O\Big(\frac{1}{n^4 h^2}\Big). \tag{1.15}$$

On the other hand, the definition of the midpoints $(m_i)_{0\le i\le n}$ yields,

$$(m_{x,i-1} - t_{x,i}) = \frac{1}{2}(t_{x,i-1} - t_{x,i}) \stackrel{\Delta}{=} -\frac{1}{2} d_{x,i-1} \quad \text{for} \quad i = 1, \cdots, N_{T_n}. \tag{1.16}$$

Using Equations (1.11)-(1.16) and the fact that $\lim\limits_{n\to\infty} nh = \infty$ we obtain,

$$\Delta_{x,h} = \frac{1}{8} \sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \varphi_{x,h}(t_{x,i}) (d^2_{x,i-1} - d^2_{x,i}) + O\Big(\frac{1}{n^2 h}\Big). \tag{1.17}$$

To control the first term of the right side of (1.17), we shall use the following lemma (its proof is given below).

Lemma 1.4.2 *Let $\{T_n, n \ge 1\}$ be a regular sequence of designs generated by a density function f (see Definition 1.1.1). If $f \in C^2([0,1])$ then for $i = 2, \cdots, n-1$,*

$$d_{i-1} - d_i = \frac{f'(t_i)}{2n^2 f(t_i)} \Big(\frac{1}{f^2(t_i^*)} + \frac{1}{f^2(t_{i-1}^*)} \Big) + O\Big(\frac{1}{n^3}\Big), \tag{1.18}$$

for some $t_i^* \in]t_i, t_{i+1}[$, recall that $d_i = t_{i+1} - t_i$.

Using the previous lemma we have,

$$\begin{aligned} \Delta_{x,h} = \frac{1}{16 n^2} \sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \varphi_{x,h}(t_{x,i}) \frac{f'(t_{x,i})}{f(t_{x,i})} \Big(\frac{1}{f^2(t^*_{x,i})} + \frac{1}{f^2(t^*_{x,i-1})} \Big) (d_{x,i-1} + d_{x,i}) \\ + O\Big(\frac{1}{n^3} + \frac{1}{n^2 h}\Big). \end{aligned}$$

Recall that g', f' and $\frac{1}{f}$ are all bounded. We obtain using Lemma 1.4.1, inequality (1.10),

$$\frac{1}{16 n^2} \sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \varphi_{x,h}(t_{x,i}) \frac{f'(t_{x,i})}{f(t_{x,i})} \Big(\frac{1}{f^2(t^*_{x,i})} + \frac{1}{f^2(t^*_{x,i-1})} \Big) (d_{x,i-1} + d_{x,i}) = O\Big(\frac{1}{n^2}\Big).$$

Since $\lim_{n\to\infty} nh = \infty$ then,

$$\Delta_{x,h} = O\Big(\frac{1}{n^2 h}\Big). \tag{1.19}$$

The control of $l_h(x)$ is classical and it can be seen from Gasser and Müller (1984) that,

$$l_h(x) = g(x) + \frac{1}{2}h^2 g''(x)\int_{-1}^{1} t^2 K(t)\ dt + o(h^2). \tag{1.20}$$

Finally, collecting (1.6), (1.19) and (1.20) we obtain,

$$\text{Bias}(\hat{g}_n^{GM}(x)) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{1}{n^2 h}\Big),$$

where $B = \int_{-1}^{1} t^2 K(t)\ dt$. This concludes the proof of Proposition 1.3.1. □

Proof of Lemma 1.4.2.

On the one hand, we have from the definition of the regular sequence of designs, for $k = 1, \cdots, n-1$,

$$F(t_{k+1}) - F(t_k) = \int_{t_k}^{t_{k+1}} f(t)\ dt = \frac{1}{n}.$$

The mean value theorem yields that for $k = 1, \cdots, n-1$,

$$d_k = t_{k+1} - t_k = \frac{1}{n f(t_k^*)}, \tag{1.21}$$

for some $t_k^* \in]t_k, t_{k+1}[$. On the other hand, note that for $k = 2, \dots, n-1$ we have,

$$\int_{t_k}^{t_{k+1}} f(t)\ dt - \int_{t_{k-1}}^{t_k} f(t)\ dt = \frac{1}{n} - \frac{1}{n} = 0.$$

Since $f \in C^2([0,1])$ then Taylor expansion of f around t_k yields,

$$\begin{aligned} f(t_k)d_k + f'(t_k)\int_{t_k}^{t_{k+1}} (t-t_k)\ dt + \frac{1}{2}\int_{t_k}^{t_{k+1}} (t-t_k)^2 f''(\epsilon_k)\ dt \\ - f(t_k)d_{k-1} - f'(t_k)\int_{t_{k-1}}^{t_k} (t-t_k)\ dt - \frac{1}{2}\int_{t_{k-1}}^{t_k} (t-t_k)^2 f''(\eta_k) = 0, \end{aligned}$$

for some $\epsilon_k \in]t_k, t_{k+1}[$ and some $\eta_k \in]t_{k-1}, t_k[$. Thus,

$$f(t_k)(d_k - d_{k-1}) + \frac{1}{2}f'(t_k)(d_k^2 + d_{k-1}^2) + \frac{1}{2}\int_{t_k}^{t_{k+1}} (t-t_k)^2 f''(\epsilon_k)\ dt - \frac{1}{2}\int_{t_{k-1}}^{t_k} (t-t_k)^2 f''(\eta_k) = 0.$$

Finally, using Equation (1.21), Lemma 1.4.1 and the fact that f'' is bounded we obtain,

$$d_k - d_{k-1} = -\frac{f'(t_k)}{2n^2 f(t_k)}\Big(\frac{1}{f^2(t_k^*)} + \frac{1}{f^2(t_{k-1}^*)}\Big) + O\Big(\frac{1}{n^3}\Big).$$

This concludes the proof of Lemma 1.4.2. □

Proof of Proposition 1.3.2.

For $i, j = 1, \cdots, n$ we have,

$$\text{Cov }(\overline{Y}(t_i), \overline{Y}(t_j)) = \frac{1}{m^2}\sum_{k=1}^{m}\sum_{l=1}^{m}\text{Cov }(\varepsilon_k(t_i), \varepsilon_l(t_j)) = \frac{1}{m^2}\sum_{k=1}^{m}\text{Cov }(\varepsilon_k(t_i), \varepsilon_k(t_j)) = \frac{1}{m}R(t_i, t_j).$$

We obtain using this last equality,

$$\text{Var }\hat{g}_n^{\text{GM}}(x) = \frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{n} R(t_i, t_j)\int_{m_{i-1}}^{m_i}\varphi_{x,h}(t)\ dt\int_{m_{j-1}}^{m_j}\varphi_{x,h}(s)\ ds.$$

Since $\varphi_{x,h}$ is of support $[x-h, x+h]$ we get, by taking the notation (1.5) as in the proof of Proposition 1.3.1,

$$\text{Var }\hat{g}_n^{\text{GM}}(x) = \frac{1}{m}\sum_{i=1}^{N_{T_n}}\sum_{j=1}^{N_{T_n}} R(t_{x,i}, t_{x,j})\int_{m_{x,i-1}}^{m_{x,i}}\varphi_{x,h}(t)\ dt\int_{m_{x,j-1}}^{m_{x,j}}\varphi_{x,h}(s)\ ds.$$

Let,

$$\begin{aligned}\sigma_{x,h}^2 &\triangleq \int_{x-h}^{x+h}\int_{x-h}^{x+h}\varphi_{x,h}(s)R(t,s)\varphi_{x,h}(t)ds\ dt = \int_{m_{x,0}}^{m_{x,N_{T_n}}}\int_{m_{x,0}}^{m_{x,N_{T_n}}}\varphi_{x,h}(s)R(t,s)\varphi_{x,h}(t)ds\ dt \\ &= \sum_{i=1}^{N_{T_n}}\sum_{j=1}^{N_{T_n}}\int_{m_{x,i-1}}^{m_{x,i}}\int_{m_{x,j-1}}^{m_{x,j}}\varphi_{x,h}(s)R(t,s)\varphi_{x,h}(t)ds\ dt,\end{aligned}$$

and write,

$$\text{Var }\hat{g}_n^{\text{GM}}(x) = \text{Var }\hat{g}_n^{\text{GM}}(x) - \frac{1}{m}\sigma_{x,h}^2 + \frac{1}{m}\sigma_{x,h}^2 \triangleq \Delta_{x,h} + \frac{1}{m}\sigma_{x,h}^2. \tag{1.22}$$

We first control $\Delta_{x,h}$. We have,

$$\Delta_{x,h} = \frac{1}{m}\sum_{i=1}^{N_{T_n}}\sum_{j=1}^{N_{T_n}}\int_{m_{x,i-1}}^{m_{x,i}}\int_{m_{x,j-1}}^{m_{x,j}}\varphi_{x,h}(t)\varphi_{x,h}(s)\big(R(t_{x,i}, t_{x,j}) - R(t,s)\big)\ ds\ dt.$$

For $i, j = 1, \ldots, N_{T_n}$ set,

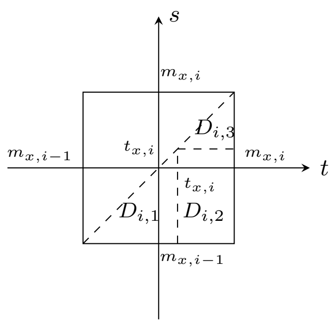
$$I_{i,j} = \int_{m_{x,i-1}}^{m_{x,i}}\int_{m_{x,j-1}}^{m_{x,j}}\varphi_{x,h}(t)\varphi_{x,h}(s)\big(R(t_{x,i}, t_{x,j}) - R(t,s)\big)\ ds\ dt. \tag{1.23}$$

Since R is a symmetric function then $\Delta_{x,h}$ can be written as follows.

$$\Delta_{x,h} = \frac{1}{m}\Big\{\sum_{i=1}^{N_{T_n}} I_{i,i} + 2\sum_{i=1}^{N_{T_n}}\sum_{j=i+1}^{N_{T_n}} I_{i,j}\Big\}. \tag{1.24}$$

We first control the diagonal terms $I_{i,i}$ for $i = 1, \cdots, N_{T_n}$. Since R is a symmetric function, it suffices to consider the integral over the lower half (triangular) of the square $[m_{x,i-1}, m_{x,i}]^2$. This triangular is further split into three triangulars as follows,

$$D_{i,1} = \{(t,s)\ :\ m_{x,i-1} \le s \le t < t_{x,i}\}.$$
$$D_{i,2} = \{(t,s)\ :\ m_{x,i-1} \le s < t_{x,i} < t < m_{x,i}\}.$$
$$D_{i,3} = \{(t,s)\ :\ t_{x,i} < s \le t < m_{x,i}\}.$$



The term $I_{i,i}$ is then written,

$$I_{i,i} = 2\sum_{k=1}^{3}\iint_{D_{i,k}} \varphi_{x,h}(t)\varphi_{x,h}(s)\big(R(t_{x,i},t_{x,i}) - R(t,s)\big)\ ds\ dt \triangleq 2\sum_{k=1}^{3} I_{i,i}^{(k)}.$$

We consider first the term $I_{i,i}^{(1)}$. We have,

$$I_{i,i}^{(1)} = \int_{m_{x,i-1}}^{t_{x,i}}\int_{m_{x,i-1}}^{t} \varphi_{x,h}(t)\varphi_{x,h}(s)\big(R(t_{x,i},t_{x,i}) - R(t,s)\big)\ ds\ dt. \tag{1.25}$$

Because of Assumptions (B) and (C), we can expand R in a Taylor series around $(t_{x,i},t_{x,i})$ for (t,s) in $D_{i,1}$ as follows,

$$\begin{aligned}
R(t,s) &= R(t_{x,i},s) + (t-t_{x,i})R^{(1,0)}(t_{x,i},s) + \frac{1}{2}(t-t_{x,i})^2R^{(2,0)}(\epsilon_{x,i}^{(1)},s) \\
&= R(t_{x,i},t_{x,i}) + (s-t_{x,i})R^{(0,1)}(t_{x,i},t_{x,i}^-) + \frac{1}{2}(s-t_{x,i})^2R^{(0,2)}(t_{x,i},\eta_{x,i}^{(1)}) \\
&\quad + (t-t_{x,i})R^{(1,0)}(t_{x,i},t_{x,i}^-) + (t-t_{x,i})(s-t_{x,i})R^{(1,1)}(t_{x,i},\eta_{x,i}^{(2)}) \\
&\quad + \frac{1}{2}(t-t_{x,i})^2R^{(2,0)}(\epsilon_{x,i}^{(1)},s),
\end{aligned}$$

for some $\epsilon_{x,i}^{(1)}$ in $]t,t_{x,i}[$ and some $\eta_{x,i}^{(1)},\eta_{x,i}^{(2)}$ in $]s,t_{x,i}[$. We obtain by inserting the above equation in (1.25),

$$\begin{aligned}
I_{i,i}^{(1)} &= R^{(0,1)}(t_{x,i},t_{x,i}^-)\int_{m_{x,i-1}}^{t_{x,i}}\int_{m_{x,i-1}}^{t} \varphi_{x,h}(t)\varphi_{x,h}(s)(t_{x,i}-s)\ ds\ dt \\
&\quad + R^{(1,0)}(t_{x,i},t_{x,i}^-)\int_{m_{x,i-1}}^{t_{x,i}}\int_{m_{x,i-1}}^{t} \varphi_{x,h}(t)\varphi_{x,h}(s)(t_{x,i}-t)\ ds\ dt \\
&\quad - \frac{1}{2}\int_{m_{x,i-1}}^{t_{x,i}}\int_{m_{x,i-1}}^{t} \varphi_{x,h}(t)\varphi_{x,h}(s)\Big((s-t_{x,i})^2R^{(0,2)}(t_{x,i},\eta_{x,i}^{(1)}) \\
&\quad + 2(t-t_{x,i})(s-t_{x,i})R^{(1,1)}(t_{x,i},\eta_{x,i}^{(2)}) + (t-t_{x,i})^2R^{(2,0)}(\epsilon_{x,i}^{(1)},s)\Big)\ ds\ dt \\
&\triangleq I_{i,i}^{(1,1)} + I_{i,i}^{(1,2)} + I_{i,i}^{(1,3)}.
\end{aligned} \tag{1.26}$$

Since $\varphi_{x,h}$ is in C^2 then Taylor expansion of $\varphi_{x,h}$ around $t_{x,i}$ for t and s in $D_{i,1}$ gives,

$$\varphi_{x,h}(t) = \varphi_{x,h}(t_{x,i}) + (t-t_{x,i})\varphi'_{x,h}(t_{x,i}) + \frac{1}{2}(t-t_{x,i})^2\varphi''_{x,h}(\epsilon_{x,i}^{(2)}),$$

and,

$$\varphi_{x,h}(s) = \varphi_{x,h}(t_{x,i}) + (s-t_{x,i})\varphi'_{x,h}(t_{x,i}) + \frac{1}{2}(s-t_{x,i})^2\varphi''_{x,h}(\eta_{x,i}^{(3)}),$$

for some $\epsilon_{x,i}^{(2)}$ in $]t, t_{x,i}[$ and some $\eta_{x,i}^{(3)}$ in $]s, t_{x,i}[$. Inserting these expansions in $I_{i,i}^{(1,1)}$ above yields,

$$\begin{aligned}
I_{i,i}^{(1,1)} &= R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi_{x,h}^2(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t_{x,i}-s)\, ds\, dt \\
&- R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi'_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t_{x,i}-s)^2\, ds\, dt \\
&+ R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi'_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t-t_{x,i})(t_{x,i}-s)\, ds\, dt \\
&+ \frac{1}{2}R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} \varphi''_{x,h}(\eta_{x,i}^{(3)})(t_{x,i}-s)^3\, ds\, dt \\
&- R^{(0,1)}(t_{x,i}, t_{x,i}^-){\varphi'_{x,h}}^2(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t-t_{x,i})(t_{x,i}-s)^2\, ds\, dt \\
&+ \frac{1}{2}R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi'_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} \varphi''_{x,h}(\eta_{x,i}^{(3)})(t-t_{x,i})(t_{x,i}-s)^3\, ds\, dt \\
&+ \frac{1}{2}R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} \varphi''_{x,h}(\epsilon_{x,i}^{(2)})(t-t_{x,i})^2(t_{x,i}-s)\, ds\, dt \\
&- \frac{1}{2}R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi'_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} \varphi''_{x,h}(\epsilon_{x,i}^{(2)})(t-t_{x,i})^2(t_{x,i}-s)^2\, ds\, dt \\
&+ \frac{1}{4}R^{(0,1)}(t_{x,i}, t_{x,i}^-) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} \varphi''_{x,h}(\eta_{x,i}^{(2)})\varphi''_{x,h}(\epsilon_{x,i}^{(2)})(t-t_{x,i})^2(t_{x,i}-s)^3\, ds\, dt.
\end{aligned}$$

We obtain using Assumption (C), Lemma 1.4.1, Inequality (1.10) and the fact that $\lim_{n\to\infty} nh = \infty$,

$$I_{i,i}^{(1,1)} = R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi_{x,h}^2(t_{x,i}) \int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t_{x,i}-s)\, ds\, dt + O\Big(\frac{1}{n^4h^3}\Big).$$

It is easy to verify that for $l, l' \in \{0,1,2\}$ we have,

$$\int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t-t_{x,i})^l (s-t_{x,i})^{l'}\, ds\, dt = (m_{x,i-1}-t_{x,i})^{l+l'+2}\Big(\frac{1}{(l+1)(l'+1)} - \frac{1}{(l+l'+2)(l'+1)}\Big).$$

From Equation (1.16) we get,

$$\int_{m_{x,i-1}}^{t_{x,i}} \int_{m_{x,i-1}}^{t} (t-t_{x,i})^l (s-t_{x,i})^{l'}\, ds\, dt \quad = \Big(\frac{-d_{x,i-1}}{2}\Big)^{l+l'+2}\Big(\frac{1}{(l+1)(l'+1)} - \frac{1}{(l+l'+2)(l'+1)}\Big). \tag{1.27}$$

We obtain using Equation (1.27),

$$I_{i,i}^{(1,1)} = \frac{1}{24}R^{(0,1)}(t_{x,i}, t_{x,i}^-)\varphi_{x,h}^2(t_{x,i})d_{x,i-1}^3 + O\Big(\frac{1}{n^4h^3}\Big)$$

Equation (1.21) then yields, for some $t^*_{x,i-1}$ in $]t_{x,i-1}, t_{x,i}[$,

$$I_{i,i}^{(1,1)} = \frac{1}{24n^2}\frac{\varphi_{x,h}^2(t_{x,i})}{f^2(t^*_{x,i-1})}R^{(0,1)}(t_{x,i}, t_{x,i}^-)d_{x,i-1} + O\Big(\frac{1}{n^4h^3}\Big),$$

We obtain using Lemma 1.4.1,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(1,1)} = \frac{1}{24n^2}\sum_{i=1}^{N_{T_n}} \frac{\varphi_{x,h}^2(t_{x,i})}{f^2(t_{x,i-1}^*)} R^{(0,1)}(t_{x,i}, t_{x,i}^-) d_{x,i-1} + O\Big(\frac{1}{n^3h^2}\Big).$$

Using a classical approximation of a sum by an integral (see for instance, Benelmadani *et al.* (2019a) c.f. Lemma 2 there) we obtain,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(1,1)} = \frac{1}{24n^2}\int_{x-h}^{x+h} R^{(0,1)}(t,t^-)\frac{\varphi_{x,h}^2(t)}{f^2(t)}dt \ + \ O\Big(\frac{1}{n^3h^2}\Big). \tag{1.28}$$

In a similar way, we verify that,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(1,2)} = \frac{1}{48n^2}\int_{x-h}^{x+h} R^{(1,0)}(t,t^-)\frac{\varphi_{x,h}^2(t)}{f^2(t)}dt \ + \ O\Big(\frac{1}{n^3h^2}\Big). \tag{1.29}$$

Assumption (C), Lemma 1.4.1 and Inequality (1.10) yield,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(1,3)} = O\Big(\frac{1}{n^3h}\Big). \tag{1.30}$$

We obtain collecting (1.28), (1.28), (1.29) and (1.30),

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(1)} = \frac{1}{24n^2}\int_{x-h}^{x+h} \Big(R^{(0,1)}(t,t^-) + \frac{1}{2}R^{(1,0)}(t,t^-)\Big)\frac{\varphi_{x,h}^2(t)}{f^2(t)}dt \ + \ O\Big(\frac{1}{n^3h^2}\Big). \tag{1.31}$$

Similarly we obtain,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(2)} = \frac{1}{16n^2}\int_{x-h}^{x+h} \Big(R^{(0,1)}(t,t^-) - R^{(1,0)}(t,t^-)\Big)\frac{\varphi_{x,h}^2(t)}{f^2(t)}dt \ + \ O\Big(\frac{1}{n^3h^2}\Big), \tag{1.32}$$

and,

$$\sum_{i=1}^{N_{T_n}} I_{i,i}^{(3)} = -\frac{1}{24n^2}\int_{x-h}^{x+h} \Big(\frac{1}{2}R^{(0,1)}(t^+,t) + R^{(1,0)}(t^+,t)\Big)\frac{\varphi_{x,h}^2(t)}{f^2(t)}dt \ + \ O\Big(\frac{1}{n^3h^2}\Big). \tag{1.33}$$

Finally, summing (1.31), (1.32) and (1.33) and using Assumption (B) we obtain,

$$\begin{aligned}
\sum_{i=1}^{N_{T_n}} I_{i,i} = 2\sum_{i=1}^{N_{T_n}}\sum_{k=1}^{3} I_{i,i}^{(k)} &= \frac{1}{24n^2}\int_{x-h}^{x+h} \frac{\varphi_{x,h}^2(t)}{f^2(t)}\Big(5R^{(0,1)}(t,t^-) - 2R^{(1,0)}(t,t^-) \\
&\quad - R^{(0,1)}(t^+,t) - 2R^{(1,0)}(t^+,t)\Big)\, dt + \ O\Big(\frac{1}{n^3h^2}\Big) \\
&= \frac{1}{6n^2}\int_{x-h}^{x+h} \frac{\varphi_{x,h}^2(t)}{f^2(t)}\Big(R^{(1,0)}(t^-,t) - R^{(1,0)}(t^+,t)\Big)\, dt + O\Big(\frac{1}{n^3h^2}\Big) \\
&= \frac{1}{6n^2}\int_{x-h}^{x+h} \frac{\varphi_{x,h}^2(t)}{f^2(t)}\alpha(t)\, dt + O\Big(\frac{1}{n^3h^2}\Big).
\end{aligned} \tag{1.34}$$

We now control the off diagonal terms $I_{i,j}$ given by (1.23) for $1 \leq i < j \leq N_{T_n}$. Using Assumption (C), Taylor-expansion of R at $(t_{x,i}, t_{x,j})$ for $(t,s) \in [m_{x,i-1}, m_{x,i}] \times [m_{x,j-1}, m_{x,j}]$ gives,

$$\begin{aligned}
R(t,s) = &R(t_{x,i}, t_{x,j}) + (t - t_{x,i})R^{(1,0)}(t_{x,i}, t_{x,j}) + (s - t_{x,j})R^{(0,1)}(t_{x,i}, t_{x,j}) \\
&+ \frac{1}{2}(t - t_{x,i})^2 R^{(2,0)}(\epsilon_{x,i}, t_{x,j}) + \frac{1}{2}(s - t_{x,j})^2 R^{(0,2)}(t_{x,i}, \eta_{x,j}) \\
&+ (t - t_{x,i})(s - t_{x,j})R^{(1,1)}(\epsilon_{x,i}, \eta_{x,j}),
\end{aligned}$$

for some $\epsilon_{x,i}$ between t and $t_{x,i}$ and some $\eta_{x,j}$ between s and $t_{x,j}$. Set,

$$I_{i,j} = I_{i,j}^{(1)} + I_{i,j}^{(2)} + I_{i,j}^{(3)} + I_{i,j}^{(4)} + I_{i,j}^{(5)}, \tag{1.35}$$

where,

$$\begin{aligned}
I_{i,j}^{(1)} &= R^{(1,0)}(t_{x,i}, t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i} - t)\varphi_{x,h}(t)\varphi_{x,h}(s)\, ds\, dt. \qquad (1.36)\\
I_{i,j}^{(2)} &= R^{(0,1)}(t_{x,i}, t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,j} - s)\varphi_{x,h}(t)\varphi_{x,h}(s)\, ds\, dt. \\
I_{i,j}^{(3)} &= -\frac{1}{2} \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} \varphi_{x,h}(t)\varphi_{x,h}(s)(t - t_{x,i})^2 R^{(2,0)}(\epsilon_{x,i}, t_{x,j})\, ds\, dt. \\
I_{i,j}^{(4)} &= -\frac{1}{2} \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} \varphi_{x,h}(t)\varphi_{x,h}(s)(s - t_{x,j})^2 R^{(0,2)}(t_{x,i}, \eta_{x,j})\, ds\, dt. \\
I_{i,j}^{(5)} &= - \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} \varphi_{x,h}(t)\varphi_{x,h}(s)(t - t_{x,i})(s - t_{x,j})R^{(1,1)}(\epsilon_{x,i}, \eta_{x,j})\, ds\, dt.
\end{aligned}$$

We first control the term $I_{i,j}^{(1)}$ for $i,j = 1, \cdots, N_{T_n}$. Since $\varphi_{x,h}$ is in C^2, Taylor expansions of $\varphi_{x,h}$ around $t_{x,i}$ for $t \in]m_{x,i-1}, m_{x,i}[$ yields,

$$\varphi_{x,h}(t) = \varphi_{x,h}(t_{x,i}) + (t - t_{x,i})\varphi'_{x,h}(t_{x,i}) + \frac{1}{2}(t - t_{x,i})^2 \varphi''_{x,h}(\epsilon_{x,i}^{(1)}), \tag{1.37}$$

for some $\epsilon_{x,i}^{(1)}$ between $t_{x,i}$ and t. Similarly, for $s \in]m_{x,j-1}, m_{x,j}[$ we obtain,

$$\varphi_{x,h}(s) = \varphi_{x,h}(t_{x,j}) + (s - t_{x,j})\varphi'_{x,h}(t_{x,j}) + \frac{1}{2}(s - t_{x,j})^2 \varphi''_{x,h}(\eta_{x,j}^{(1)}), \tag{1.38}$$

for some $\eta_{x,j}^{(1)}$ between $t_{x,j}$ and s. Using (1.36), (1.37) and (1.38) we obtain,

$$
\begin{aligned}
I_{i,j}^{(1)} &= R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i}-t)\, ds\, dt \\
&+ R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i})\varphi'_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i}-t)(s-t_{x,j})\, ds\, dt \\
&- R^{(1,0)}(t_{x,i},t_{x,j})\varphi'_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^2\, ds\, dt \\
&+ \frac{1}{2}R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i}-t)(s-t_{x,j})^2\varphi''_{x,h}(\eta_{x,j}^{(1)})\, ds\, dt \\
&- R^{(1,0)}(t_{x,i},t_{x,j})\varphi'_{x,h}(t_{x,i})\varphi'_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^2(s-t_{x,j})\, ds\, dt \\
&- \frac{1}{2}R^{(1,0)}(t_{x,i},t_{x,j})\varphi'_{x,h}(t_{x,i}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^2(s-t_{x,j})^2\varphi''_{x,h}(\eta_{x,j}^{(1)})\, ds\, dt \\
&- \frac{1}{2}R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^3\varphi''_{x,h}(\epsilon_{x,i}^{(1)})\, ds\, dt \\
&- \frac{1}{2}R^{(1,0)}(t_{x,i},t_{x,j})\varphi'_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^3(s-t_{x,j})\varphi''_{x,h}(\epsilon_{x,i}^{(1)})\, ds\, dt \\
&- \frac{1}{4}R^{(1,0)}(t_{x,i},t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^3(s-t_{x,j})^2\varphi''_{x,h}(\epsilon_{x,i}^{(1)})\varphi''_{x,h}(\eta_{x,j}^{(1)})\, ds\, dt.
\end{aligned}
$$

We obtain using Assumption (C), Lemma 1.4.1 and Inequality (1.10),

$$
\begin{aligned}
I_{i,j}^{(1)} &= R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i}-t)\, ds\, dt \\
&+ R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i})\varphi'_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t_{x,i}-t)(s-t_{x,j})\, ds\, dt \\
&- R^{(1,0)}(t_{x,i},t_{x,j})\varphi'_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,j}) \int_{m_{x,i-1}}^{m_{x,i}} \int_{m_{x,j-1}}^{m_{x,j}} (t-t_{x,i})^2\, ds\, dt + O\Big(\frac{1}{n^5h^4}\Big) \\
&\triangleq I_{i,j}^{(1,1)} + I_{i,j}^{(1,2)} + I_{i,j}^{(1,3)} + O\Big(\frac{1}{n^5h^4}\Big).
\end{aligned} \tag{1.39}
$$

We first control the term $I_{i,j}^{(1,1)}$. Basic integration together with Equation (1.16) yield,

$$
I_{i,j}^{(1,1)} = -\frac{1}{8}R^{(1,0)}(t_{x,i},t_{x,j})\varphi_{x,h}(t_{x,i})\varphi_{x,h}(t_{x,j})(m_{x,j}-m_{x,j-1})(d_{x,i}^2-d_{x,i-1}^2).
$$

Using the definition of the midpoints $(m_{x,j})_j$ and Lemma 1.4.2 we obtain, for some $t_{x,i}^*$ in $]t_{x,i}, t_{x,i+1}[$,

$$
\begin{aligned}
d_{x,i}^2 - d_{x,i-1}^2 &= (d_{x,i}+d_{x,i-1})(d_{x,i}-d_{x,i-1}) \\
&= 2(m_{x,i}-m_{x,i-1})(d_{x,i}-d_{x,i-1}) \\
&= -(m_{x,i}-m_{x,i-1})\frac{f'(t_{x,i})}{n^2 f(t_{x,i})}\Big(\frac{1}{f^2(t_{x,i}^*)} + \frac{1}{f^2(t_{x,i-1}^*)}\Big) + O\Big(\frac{1}{n^4}\Big). && (1.40) \\
&= O\Big(\frac{1}{n^3}\Big) && (1.41)
\end{aligned}
$$

Using Equation (1.41), Inequality (1.10), Assumption (C), Lemma 1.4.2 and the fact that f, f' and $\frac{1}{f}$ are all bounded,

$$I_{i,j}^{(1,1)} = O\Big(\frac{1}{n^4h^2}\Big).$$

Lemma 1.4.1 then yields,

$$\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1,1)} = O\Big(\frac{1}{n^2}\Big). \tag{1.42}$$

We control now the term $I_{i,j}^{(1,2)}$. Basic integration together with Equation (1.16) yield,

$$I_{i,j}^{(1,2)} = -\frac{1}{64} R^{(1,0)}(t_{x,i}, t_{x,j}) \varphi_{x,h}(t_{x,i}) \varphi'_{x,h}(t_{x,j}) (d_{x,i}^2 - d_{x,i-1}^2)(d_{x,j}^2 - d_{x,j-1}^2).$$

Using Equations (1.40) and (1.41), Inequality (1.10), Assumption (C), Lemma 1.4.2 and the boundedness of f, f' and $\frac{1}{f}$ we get,

$$\begin{aligned} I_{i,j}^{(1,2)} &= -\frac{1}{64n^4} R^{(1,0)}(t_{x,i}, t_{x,j}) \varphi_{x,h}(t_{x,i}) \varphi'_{x,h}(t_{x,j}) (m_{x,i} - m_{x,i-1})(m_{x,j} - m_{x,j-1}) \\ &\quad \times \frac{f'(t_{x,i})}{f(t_{x,i})} \Big(\frac{1}{f^2(t_{x,i}^*)} + \frac{1}{f^2(t_{x,i-1}^*)} \Big) \frac{f'(t_{x,j})}{f(t_{x,j})} \Big(\frac{1}{f^2(t_{x,j}^*)} + \frac{1}{f^2(t_{x,j-1}^*)} \Big) + O\Big(\frac{1}{n^7h^3}\Big) \\ &= O\Big(\frac{1}{n^6h^3}\Big). \end{aligned}$$

Thus,

$$\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1,2)} = O\Big(\frac{1}{n^4h}\Big). \tag{1.43}$$

We control now the term $I_{i,j}^{(1,3)}$. Using Equation (1.16) yields,

$$I_{i,j}^{(1,3)} = -\frac{1}{24} R^{(1,0)}(t_{x,i}, t_{x,j}) \varphi'_{x,h}(t_{x,i}) \varphi_{x,h}(t_{x,j}) (m_{x,j} - m_{x,j-1})(d_{x,i}^3 + d_{x,i-1}^3). \tag{1.44}$$

We obtain using the definition of the midpoints together with Equation (1.21),

$$\begin{aligned} d_{x,i}^3 + d_{x,i-1}^3 &= (d_{x,i} + d_{x,i-1})(d_{x,i}^2 - d_{x,i}d_{x,i-1} + d_{x,i-1}^2) \\ &= \frac{2}{n^2}(m_{x,i} - m_{x,i-1}) \Big(\frac{1}{f^2(t_{x,i}^*)} - \frac{1}{f(t_{x,i}^*)f(t_{x,i-1}^*)} + \frac{1}{f^2(t_{x,i-1}^*)} \Big), \end{aligned}$$

for some $t_{x,i}^*$ in $]t_{x,i}, t_{x,i+1}[$. Thus,

$$\begin{aligned} \sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1,3)} = -\frac{1}{12n^2} \sum_{\substack{i=1 \\ i<j}}^{N_{T_n}} \sum_{j=1}^{N_{T_n}} \Bigg(& R^{(1,0)}(t_{x,i}, t_{x,j}) \varphi'_{x,h}(t_{x,i}) \varphi_{x,h}(t_{x,j}) (m_{x,j} - m_{x,j-1}) \\ & \times (m_{x,i} - m_{x,i-1}) \Big(\frac{1}{f^2(t_{x,i}^*)} - \frac{1}{f(t_{x,i}^*)f(t_{x,i-1}^*)} + \frac{1}{f^2(t_{x,i-1}^*)} \Big) \Bigg). \end{aligned}$$

Using (twice) the classical approximation of a sum by an integral (see for instance Lemma 1 in Benelmadani *et al.* (2019a) we obtain,

$$\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1,3)} = -\frac{1}{12n^2}\int_{x-h}^{x+h}\int_{x-h}^{s} R^{(1,0)}(t,s)\varphi_{x,h}(s)\frac{\varphi'_{x,h}(t)}{f^2(t)}\,dt\,ds + O\Big(\frac{1}{n^3h^2}\Big). \tag{1.45}$$

We obtain collecting (1.39), (1.42), (1.43) and (1.45),

$$\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1)} = -\frac{1}{12n^2}\int_{x-h}^{x+h}\int_{x-h}^{s} R^{(1,0)}(t,s)\varphi_{x,h}(s)\frac{\varphi'_{x,h}(t)}{f^2(t)}\,dt\,ds + O\Big(\frac{1}{n^3h^2}+\frac{1}{n^2}\Big). \tag{1.46}$$

We shall prove that,

$$\frac{1}{12n^2}\int_{x-h}^{x+h}\int_{x-h}^{s} R^{(1,0)}(t,s)\varphi_{x,h}(s)\frac{\varphi'_{x,h}(t)}{f^2(t)}\,dt\,ds = -\frac{1}{12n^2}\int_{x-h}^{x+h} R^{(1,0)}(s^-,s)\varphi^2_{x,h}(s)\frac{1}{f^2(s)}\,ds. \tag{1.47}$$

For this, we use partial integral to get for $s \in]x-h, x+h[$,

$$\int_{x-h}^{s} R^{(1,0)}(t,s)\frac{\varphi'_{x,h}(t)}{f^2(t)}\,dt = R^{(1,0)}(t,s)\frac{\varphi_{x,h}(t)}{f^2(t)}\Big|_{x-h}^{s} - \int_{x-h}^{s}\varphi_{x,h}(t)\Big(\frac{1}{f^2(t)}R^{(2,0)}(t,s) - \frac{2f'(t)}{f^3(t)}R^{(1,0)}(t,s)\Big)\,dt.$$

Recall that $\varphi_{x,h}(x-h)=0$. We obtain using Inequality (1.10), Assumptions (C) and the fact that f' and $\frac{1}{f}$ are both bounded,

$$\int_{x-h}^{s}\varphi_{x,h}(t)\Big(\frac{1}{f^2(t)}R^{(2,0)}(t,s) - \frac{2f'(t)}{f^3(t)}R^{(1,0)}(t,s)\Big)\,dt = O(1).$$

Thus,

$$\int_{x-h}^{s} R^{(1,0)}(t,s)\frac{\varphi'_{x,h}(t)}{f^2(t)}\,dt = R^{(1,0)}(s^-,s)\frac{\varphi_{x,h}(s)}{f^2(s)} + O(1). \tag{1.48}$$

Finally, using (1.46) and (1.48) we obtain,

$$\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(1)} = -\frac{1}{12n^2}\int_{x-h}^{x+h} R^{(1,0)}(s^-,s)\frac{\varphi^2_{x,h}(s)}{f^2(s)}\,ds + O\Big(\frac{1}{n^3h^2}+\frac{1}{n^2}\Big). \tag{1.49}$$

Similarly we prove that,

$$\begin{aligned}\sum_{i=1}^{N_{T_n}} \sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(2)} &= -\frac{1}{12n^2}\int_{x-h}^{x+h}\int_{t}^{x+h} R^{(0,1)}(t,s)\varphi_{x,h}(t)\varphi'_{x,h}(s)\frac{1}{f^2(s)}\,ds\,dt + O\Big(\frac{1}{n^3h^2}+\frac{1}{n^2}\Big)\\ &= \frac{1}{12n^2}\int_{x-h}^{x+h} R^{(0,1)}(t,t^+)\frac{\varphi^2_{x,h}(t)}{f^2(t)}\,dt + O\Big(\frac{1}{n^3h^2}+\frac{1}{n^2}\Big).\end{aligned} \tag{1.50}$$

We verify using Assumption (C), Lemma 1.4.1 and Inequality (1.10) that,

$$\sum_{i=1}^{N_{T_n}}\sum_{j=i+1}^{N_{T_n}} I_{i,j}^{(3)} = O\Big(\frac{1}{n^2}\Big), \quad \sum_{\substack{i=1 \\ i<j}}^{N_{T_n}}\sum_{j=1}^{N_{T_n}} I_{i,j}^{(4)} = O\Big(\frac{1}{n^2}\Big) \text{ and } \sum_{\substack{i=1 \\ i<j}}^{N_{T_n}}\sum_{j=1}^{N_{T_n}} I_{i,j}^{(5)} = O\Big(\frac{1}{n^2}\Big). \tag{1.51}$$

Using (1.35), (1.49), (1.50) and (1.51) yields,

$$2\sum_{i=1}^{N_{T_n}}\sum_{j=i+1}^{N_{T_n}} I_{i,j} = -\frac{1}{6n^2}\int_{x-h}^{x+h}\big(R^{(1,0)}(t^-,t) - R^{(0,1)}(t,t^+)\big)\frac{\varphi_{x,h}^2(t)}{f^2(t)}\,dt + O\Big(\frac{1}{n^2}\Big). \tag{1.52}$$

Since $R^{(1,0)}(t^-,t) = \lim_{s\uparrow t}\frac{\partial R(s,t)}{\partial s} = \lim_{s\uparrow t}\frac{\partial R(t,s)}{\partial s} = R^{(0,1)}(t,t-)$ then,

$$2\sum_{i=1}^{N_{T_n}}\sum_{j=i+1}^{N_{T_n}} I_{i,j} = -\frac{1}{6n^2}\int_{x-h}^{x+h}\alpha(t)\frac{\varphi_{x,h}^2(t)}{f^2(t)}\,dt + O\Big(\frac{1}{n^2} + \frac{1}{n^3h^2}\Big). \tag{1.53}$$

Finally, we obtain by (1.24), (1.34) and (1.53),

$$\Delta_{x,h} = O\Big(\frac{1}{mn^2} + \frac{1}{mn^3h^2}\Big). \tag{1.54}$$

For the control of $\frac{1}{m}\sigma_{x,h}^2$, we use its asymptotic expression given by Benhenni and Rachdi (2007) by,

$$\sigma_{x,h}^2 = R(x,x) - \frac{1}{2}\alpha(x)C_K h + o(h). \tag{1.55}$$

Finally, the proof of Proposition 1.3.2 is concluded by collecting (1.22), (1.54) and (1.55). □

Proof of Proposition 1.3.3.

Let $I_1 = \int_0^1 R(x,x)w(x)\,dx$ and put,

$$\Psi(h,m) = -\frac{C_K h}{2m}\int_0^1 \alpha(x)w(x)\,dx + \frac{1}{4}h^4B^2\int_0^1 [g''(x)]^2 w(x)\,dx.$$

We have from Theorem 1.3.1,

$$\text{IMSE}(h) = \frac{I_1}{m} + \Psi(h,m) + o\Big(h^4 + \frac{h}{m}\Big) + O\Big(\frac{h}{n^2} + \frac{1}{n^4h^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big).$$

Let h^* be as defined in (1.4). It is clear that $h^* = \underset{0<h<1}{\text{argmin}}\ \Psi(h,m)$ so that $\Psi(h,m) \geq \Psi(h^*,m)$ for every $0 < h < 1$. Let $h_{n,m}$ be as defined in Corollary 1.3.3. We have,

$$\frac{\text{IMSE}(h^*)}{\text{IMSE}(h_{n,m})}$$

$$= \frac{\frac{I_1}{m} + \Psi(h^*,m) + o\Big(h^{*4} + \frac{h^*}{m}\Big) + O\Big(\frac{h^*}{n^2} + \frac{1}{n^4h^{*2}} + \frac{1}{mn^3h^{*2}} + \frac{1}{mn^2}\Big)}{\frac{I_1}{m} + \Psi(h_{n,m},m) + o\Big(h_{n,m}^4 + \frac{h_{n,m}}{m}\Big) + O\Big(\frac{h_{n,m}}{n^2} + \frac{1}{n^4h_{n,m}^2} + \frac{1}{mn^3h_{n,m}^2} + \frac{1}{mn^2}\Big)}$$

$$\leq \frac{I_1 + m\Psi(h^*,m) + o\Big(mh^{*4} + h^*\Big) + O\Big(\frac{mh^*}{n^2} + \frac{m}{n^4h^{*2}} + \frac{1}{n^3h^{*2}} + \frac{1}{n^2}\Big)}{I_1 + m\Psi(h_{n,m},m) + o\Big(mh_{n,m}^4 + h_{n,m}\Big) + O\Big(\frac{mh_{n,m}}{n^2} + \frac{m}{n^4h_{n,m}^2} + \frac{1}{n^3h_{n,m}^2} + \frac{1}{n^2}\Big)}.$$

Note that $m\Psi(h_{n,m}, m) = O(h_{n,m})$. Using the definition of h^*, the facts that: $mh^3_{n,m} = O(1)$, $\lim\limits_{n,m\to\infty} h_{n,m} = 0$ and the assumption $\frac{m}{n} = O(1)$ as $n, m \to \infty$, we obtain,

$$\overline{\lim_{n,m\to\infty}} \frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})} \leq 1.$$

This concludes the proof of Proposition 1.3.3. □

Proof of Theorem 1.3.2.

Let $x \in]0, 1[$ be fixed. We have,

$$\sqrt{m}\Big(\hat{g}^{GM}_{n,m}(x) - g(x)\Big) = \sqrt{m}\Big(\hat{g}^{GM}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{GM}_{n,m}(x)\big)\Big) + \sqrt{m}\ \mathrm{Bias}\Big(\hat{g}^{GM}_{n,m}(x)\Big). \tag{1.56}$$

Since $\lim\limits_{n,m\to\infty} \sqrt{m}h^2 = 0$ and $\frac{m}{n} = O(1)$ as $n, m \to \infty$ then Proposition 1.3.1 implies that,

$$\lim_{n,m\to\infty} \sqrt{m}\ \mathrm{Bias}\Big(\hat{g}^{GM}_{n,m}(x)\Big) = 0. \tag{1.57}$$

Consider now the first term of the right side of (1.56). We take the same notation as in the proof of Proposition 1.3.1 and recall that $\varphi_{x,h}$ is of support $[x-h, x+h]$. Since $\overline{Y}(t_{x,i}) - \mathbb{E}\big(\overline{Y}(t_{x,i})\big) = \overline{\varepsilon}(t_{x,i})$, we have, as done by Fraiman and Pérez Iribarren (1991),

$$\begin{aligned}
\sqrt{m}\Big(\hat{g}^{GM}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{GM}_{n,m}(x)\big)\Big) &= \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\ ds\ \varepsilon_j(t_{x,i}) \\
= \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\ ds\ \big(\varepsilon_j(t_{x,i}) - \varepsilon_j(x)\big) &+ \frac{1}{\sqrt{m}} \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\ ds \sum_{j=1}^{m} \varepsilon_j(x).
\end{aligned} \tag{1.58}$$

We start by controlling the second term of the right side of (3.67). On the one hand, we obtain using a classical approximation of a sum by an integral (see for instance, Benelmadani *et al.* (2019a) c.f. Lemma 2 there),

$$\sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\ ds \underset{n\to\infty}{\longrightarrow} \int_{-1}^{1} K(t)\ dt = 1.$$

On the other hand, the Central Limit Theorem for i.i.d. variables yields,

$$\frac{1}{\sqrt{m}} \sum_{j=1}^{m} \varepsilon_j(x) \underset{m\to\infty}{\overset{\mathscr{D}}{\longrightarrow}} Z \quad where \quad Z \sim \mathcal{N}(0, R(x,x)).$$

We shall prove now that the first term of the right side of (3.67) tends to 0 in probability as n, m tends to infinity. Let,

$$A_{m,n}(x) = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(s)\ ds\ \big(\varepsilon_j(t_{x,i}) - \varepsilon_j(x)\big) \triangleq \frac{1}{\sqrt{m}} \sum_{j=1}^{m} T_{n,j}(x).$$

From Chebyshev's Inequality, it suffices to prove that $\lim\limits_{n,m\to\infty} \mathbb{E}(A^2_{m,n}(x)) = 0$. We have for $j \neq l$, $\mathbb{E}(\varepsilon_j(x)\varepsilon_l(y)) = 0$ so $\mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = 0$. Hence,

$$\mathbb{E}(A^2_{m,n}(x)) = \frac{1}{m} \sum_{j=1}^{m} \sum_{l=1}^{m} \mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}(T^2_{n,j}(x)).$$

We have,

$$\begin{aligned}
\mathbb{E}(T_{n,j}^2(x)) &= \sum_{i=1}^{N_{T_n}} \sum_{k=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(t)\ dt \int_{m_{x,k-1}}^{m_{x,k}} \varphi_{x,h}(s)\ ds\ \mathbb{E}\Big(\big(\varepsilon_j(t_{x,i}) - \varepsilon_j(x)\big)\big(\varepsilon_j(t_{x,k}) - \varepsilon_j(x)\big)\Big) \\
&= \sum_{i=1}^{N_{T_n}} \sum_{k=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(t)\ dt \int_{m_{x,k-1}}^{m_{x,k}} \varphi_{x,h}(s)\ ds\Big(R(t_{x,i}, t_{x,k}) - R(t_{x,i}, x) - R(x, t_{x,k}) + R(x,x)\Big).
\end{aligned}$$

Since $\mathbb{E}((T_{n,j}^2(x))$ does not depend on j we get,

$$\begin{aligned}
\mathbb{E}(A_{m,n}^2(x)) &= \sum_{i=1}^{N_{T_n}} \sum_{k=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(t)\ dt \int_{m_{x,k-1}}^{m_{x,k}} \varphi_{x,h}(s)\ ds\Big(R(t_{x,i}, t_{x,k}) - R(t_{x,i}, x) - R(x, t_{x,k}) + R(x,x)\Big) \\
&\triangleq B_{n,1}(x) - B_{n,2}(x) - B_{n,3}(x) + B_{n,4}(x). \qquad (1.59)
\end{aligned}$$

We obtain using a classical approximation of a sum by an integral,

$$\begin{aligned}
B_{n,1}(x) &= \sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(t)\ dt\Big(\int_{x-h}^{x+h} \varphi_{x,h}(t) R(t_{x,i}, t)\ dt + O\Big(\frac{1}{nh}\Big)\Big) \\
&= \int_{x-h}^{x+h} \varphi_{x,h}(t)\Big(\sum_{i=1}^{N_{T_n}} \int_{m_{x,i-1}}^{m_{x,i}} \varphi_{x,h}(t)\ dt\ R(t_{x,i}, t)\Big) dt + O\Big(\frac{1}{nh}\Big) \\
&= \int_{x-h}^{x+h} \int_{x-h}^{x+h} \varphi_{x,h}(s)\varphi_{x,h}(t) R(s,t)\ ds\ dt + O(\frac{1}{nh}) = \sigma_{x,h}^2 + O\Big(\frac{1}{nh}\Big).
\end{aligned}$$

Using Equation (1.55) we obtain,

$$B_{n,1}(x) = R(x,x) - \frac{1}{2}\alpha(x) C_K h + o(h) + O\Big(\frac{1}{nh}\Big),$$

where $C_K = \int_{-1}^{1} \int_{-1}^{1} |u - v| K(u) K(v) du dv$. Since $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$. Thus,

$$\lim_{n\to\infty} B_{n,1}(x) = R(x,x). \qquad (1.60)$$

Consider now the term $B_{n,2}(x)$, we obtain using a classical approximation of a sum by an integral (see for instance, Benelmadani *et al.* (2019a) c.f. Lemma 2 there) twice,

$$\begin{aligned}
B_{n,2}(x) &= \int_{x-h}^{x+h} \int_{x-h}^{x+h} \varphi_{x,h}(s)\varphi_{x,h}(t) R(s,x)\ ds\ dt + O\Big(\frac{1}{nh}\Big) \\
&= \int_{x-h}^{x+h} \varphi_{x,h}(s) R(s,x)\ ds + O\Big(\frac{1}{nh}\Big) \\
&= \int_{-1}^{1} K(s) R(x - sh, x)\ ds + O\Big(\frac{1}{nh}\Big) \\
&= \int_{-1}^{0} K(s) R(x - sh, x)\ ds + \int_{0}^{1} K(s) R(x - sh, x)\ ds + O\Big(\frac{1}{nh}\Big).
\end{aligned}$$

For any $s \in]-1, 0[$, Taylor expansion of $R(\cdot, x)$ around x yields,

$$R(x - sh, x) = R(x,x) - shR^{(1,0)}(x^+, x) + o(h).$$

Similarly, for $s \in]0,1[$ we obtain,

$$R(x-sh,x) = R(x,x) - shR^{(1,0)}(x^-,x) + o(h).$$

Using Assumption (C) we get,

$$B_{n,2}(x) = R(x,x) - hR^{(1,0)}(x^+,x)\int_{-1}^{0} s\ K(s)\ ds - hR^{(1,0)}(x^-,x)\int_{0}^{1} s\ K(s)\ ds + o(h) + O\Big(\frac{1}{nh}\Big).$$

Hence,

$$\lim_{n\to\infty} B_{n,2}(x) = R(x,x). \tag{1.61}$$

Similarly,

$$\lim_{n\to\infty} B_{n,3}(x) = R(x,x). \tag{1.62}$$

It is easy to see that,

$$\begin{aligned}\lim_{n\to\infty} B_{n,4}(x) &= \lim_{n\to\infty} R(x,x)\sum_{i=1}^{N_{T_n}}\sum_{k=1}^{N_{T_n}}\int_{m_{x,i-1}}^{m_{x,i}}\varphi_{x,h}(t)\ dt\int_{m_{x,k-1}}^{m_{x,k}}\varphi_{x,h}(s)\ ds\\ &= R(x,x)\Big(\int_{-1}^{1} K(t)\ dt\Big)^2 = R(x,x). \qquad (1.63)\end{aligned}$$

Using (1.59), (1.60), (1.61), (1.62) and (1.63) we obtain,

$$\lim_{n,m\to\infty}\mathbb{E}(A^2_{m,n}(x)) = 0.$$

This concludes the proof of Theorem 1.3.2. □

Chapter 2

Reproducing kernel Hilbert Space approach in nonparametric regression problems with correlated observations

Abstract: In this chapter we investigate the problem of estimating the regression function in models with correlated observations. The data is obtained from several experimental units, each of them forms a time series. Using the properties of the Reproducing Kernel Hilbert spaces, we construct a new estimator based on the inverse of the autocovariance matrix of the observations. We give the asymptotic expressions of its bias and its variance. In addition, we give a theoretical comparison between this new estimator and the popular one proposed by Gasser and Müller, we show that the proposed estimator has an asymptotically smaller variance then the classical one. Finally, we conduct a simulation study to investigate the performance of the proposed estimator and to compare it to the Gasser and Müller's estimator in a finite sample set.

Keywords. *Nonparametric regression, correlated observations, growth curve, reproducing kernel Hilbert space, projection estimator, asymptotic normality* .

Résumé: Dans ce chapitre, nous considérons le problème d'estimation de la fonction de régression dans un modèle avec des erreurs corrélées. Les données sont obtenues à partir de plusieurs unités expérimentales, chacune représente une série temporelle. En utilisant les propriétés de l'espace de Hilbert à noyau autoreproduisant, nous construisons un nouvel estimateur basé sur l'inverse de la matrice d'autocovariance des observations. Nous donnons les expressions asymptotiques de son biais et de sa variance. En plus, nous faisons une comparaison théorique avec l'estimateur classique, proposé par Gasser et Müller. Nous montrons que la variance de l'estimateur proposé est asymptotiquement plus petite que celle de l'estimateur classique. Finalement, nous effectuons une étude de simulation, afin d'étudier la performance de l'estimateur proposé, et de le comparer avec l'estimateur de Gasser et Müller pour différentes tailles d'échantillons.

Mots clés: *Régression non paramétrique, observations corrélées, courbe de croissance, espace de Hilbert à noyau autoreproduisant, estimateur de projection, normalité asymptotique.*

2.1 Introduction

One of the situations that statisticians encounter in their studies is the estimation of a whole function based on partial observations of this function. For instance, in pharmacokinetics one wishes to estimate the concentration-time of some injected medicine in the organism, based on the observations of the concentration from blood tests over a period of time. In statistical terms, one wants to estimate a function, say g, relating two random variables: the explanatory variable X and the response variable Y, without any parametric restrictions on the function g. The statistical model often used is the following: $Y_i = g(X_i) + \varepsilon_i$ where $(X_i, Y_i)_{1\le i\le n}$ are n independent replicates of (X, Y) and $\{\varepsilon_i, i = 1, \cdots, n\}$ are centered random variables (called errors).

The most intensively treated model has been the one in which $(\varepsilon_i)_{1\le i\le n}$ are independent errors and $(X_i)_{1\le i\le n}$ are fixed within some domain. We mention the works of Priestly and Chao (1972), Benedetti (1977) and Gasser and Müller (1979) among others. However, the independence of the observations is not always a realistic assumption. For instance, the growth curve models are usually used in the case of longitudinal data, where the same experimental unit is being observed on multiple points of time. As a real life example, the heights observed on the same child are correlated. The temperature observations measured along the day are also correlated. For this, we focus, in this chapter, on the nonparametric kernel estimation problem where the observations are correlated.

In the current chapter, we consider a situation where the data is generated from m experimental units each of them having n measurements of the response. For this data, we consider the so-called fixed design regression model with repeated measurements given by,

$$Y_j(t_i) = g(t_i) + \varepsilon_j(t_i) \quad \text{for } i = 1, \cdots, n \text{ and } j = 1, \cdots, m, \tag{2.1}$$

where $\{\varepsilon_j, j = 1, \cdots, m\}$ is a sequence of i.i.d. centered error processes with the same distribution as a process ε. The non correlation of the errors $\{\varepsilon_j, j = 1, \cdots, m\}$ is a natural assumption since it is equivalent to assuming that the experimental units (in general individuals) are independent.

This model is usually used in the growth curve analysis and dose response problems, see for instance, the work of Azzalini (1984). It has also been considered by Müller (1984) with $m = 1$, where he supposed that the observations are asymptotically uncorrelated when the number of observations tends to infinity, i.e., $\mathrm{Cov}(\varepsilon(s), \varepsilon(t)) = O(1/n)$ for $s \neq t$, which is not a realistic assumption, for instance, in the growth curve analysis and temperature.

The correlated observations case was considered by Hart and Wherly (1986), who investigated the estimation of g in Model (2.1) where ε is a stationary error process. Using the kernel estimator proposed by Gasser and Müller, see Gasser and Müller (1979), they proved the consistency in $\mathbb{L}^2$ space of this estimator, when the number of experimental units m tends to infinity, but not when n tends to infinity as in the case of independent observations.

The assumption of stationarity made on the observations is however restrictive. In the previous pharmacokinetics example for instance, it is clear that the concentration of the medicine will be high at the beginning then decreases with time. For this, we shall investigate the estimation of g in Model (2.1) where ε is not necessarily a stationary error process. This case was partially investigated by Benhenni and Rachdi (2016) and Ferreira *et al.* (1997), where the Gasser and Müller's estimator was used.

In this chapter, we propose a new estimator for the regression function g in Model (2.1). This estimator, which is also a linear kernel estimator, is based on the inverse of the autocovariance matrix of the observations, that we assume known and invertible.

The proposed estimator was inspired by the work of Sacks and Ylvisaker (1966, 19868, 1970) but in a different context than ours. They considered the parametric model: $Y(t) = \beta f(t) + \varepsilon(t)$ where β is an unknown real parameter and f is a known function belonging to the Reproducing Kernel Hilbert Space associated to the autocovariance function of the error process ε, denoted by RKHS(R). They also assumed that the autocovariance matrix is known and invertible. It is worth noting that the Reproducing Kernel Hilbert Spaces have been used in several domains, for instance, in Statistics by Sacks and Ylvisaker (1966) and more recently by Dette *et al.* (2016), in Mathematical Analysis in Schwartz (1964) and in Signal Processing in Ramsay and Silverman (2005).

We also give the asymptotic statistical performance of the proposed estimator and we compare it to the classical Gasser and Müller's estimator (GM estimator), proving, in particular, that the proposed estimator outperforms the GM estimator, in the sense that it has an asymptotically smaller variance, wheras they both are asymptotically unbiased. This can be argued by the fact that, in statistics in general, the best linear estimator (or optimal predictor) is based on the inverse of the autocovariance matrix, see for instance, Benhenni and Cambanis (1992), whereas the GM estimator does not take into account this correlation requirement. In addition, the GM estimator is an approximation of an integral and, as known in statistics, the best linear approximation of an integral is based on some projection property.

This chapter is organized as follows. Section 2.2 is dedicated to a recall on the RKHS, which will be usefull for the construction of the new estimator, in addition to some technical details. In section 2.3, we construct our proposed estimator for the function g in Model (2.1) where ε is a centered, second order error process with a continuous autocovariance function R. It is constructed through the following function defined, for $x \in [0,1]$, by,

$$f_{x,h}(t) = \int_0^1 R(s,t)\varphi_{x,h}(t)\, ds \quad \text{where} \quad \varphi_{x,h}(t) = \frac{1}{h} K\Big(\frac{x-t}{h}\Big) \quad \text{for} \quad t \in [0,1], \tag{2.2}$$

where K is a Kernel and $h = h(n)$ is a bandwidth.

We shall see that this function belongs to the RKHS(R). This allows us to use the properties of this space to control the variance of the proposed estimator. These properties were introduced by Parzen (1959) to solve various problems in statistical inference on time series. We also give, in this section, the analytical expressions of this estimator for the generalised Wiener process and the Ornstein-Uhlenbeck process, since the analytical expression of the inverse of the autocovariance matrix can be derived for this class of processes.

In Section 2.4, we derive the asymptotic performances of this estimator. We give an asymptotic expression of the weights of this linear estimator, which is used to derive the asymptotic expression of its bias. The properties of the RKHS(R) not only allow us to obtain the asymptotic expression of the variance, but also to find the optimal rate of convergence of the residual variance. After obtaining the asymptotic expression of the Integrated Mean Squared Error (IMSE), we derive the asymptotic optimal bandwidth with respect to the IMSE criterion. Moreover, we prove the asymptotic normality of the proposed estimator.

In Section 2.5, we give a theoretical comparison between the new estimator and the Gasser and Müller's estimator. We prove that the proposed estimator has, asymptotically, a smaller variance than that of Gasser and Müller. Moreover, the proposed estimator has an asymptotically smaller IMSE, for instance, in the case of a Wiener process ε.

In Section 2.6, we conduct a simulation study in order to investigate the performance of the proposed estimator in a finite sample set, then we compare it with the Gasser and Müller's estimator for different values of the number of experimental units and different values of the sample size. Since the classical cross-validation criterion is shown to be inefficient in the presence of correlation (see for instance, Altman (1990), Chiu (1989) and Hart (1991, 1994), we use the optimal bandwidth that minimizes the exact IMSE, obtained using the Conjugated Gradient Algorithm. The results of this simulation study confirm our theoretical statements given in Section 3 and Section 4.

Finally, the supplementary materials section is dedicated to the proofs of the theoretical results.

In the following section, we introduce the Reproducing Kernel Hilbert Spaces, a quick recall about them which will be useful through out this chapter.

2.2 Reproducing Kernel Hilbert Spaces

Let $\varepsilon = (\varepsilon(t))_{t\in[0,1]}$ be a centered and a second order process of autocovariance R, such that R is invertible when restricted to any finite set on $[0,1]$. Let $L(\varepsilon(t), t \in [0,1])$ be the set of all random variables which maybe be written as a linear combinations of $\varepsilon(t)$ for $t \in [0,1]$, i.e., the set of random variables of the form $\sum_{i=1}^{l} \alpha_i \varepsilon(t_i)$ for some positive integer l and some constants α_i, $t_i \in [0,1]$ for $i = 1, \cdots, l$. Let also $L_2(\varepsilon)$ be the Hilbert space of all square integrable random variables in the linear manifold $L(\varepsilon(t), t \in [0,1])$, together with all random variables U that are limits in $\mathbb{L}^2$ of a sequence of random variables U_n in $L(\varepsilon(t), t \in [0,1])$, i.e, U is such that,

$$\exists\, (U_n)_{n\geq 0} \in L(\varepsilon(t), t \in [0,1]) : \lim_{n\to\infty} \mathbb{E}((U_n - U)^2) = 0.$$

Denote by $\mathcal{F}(\varepsilon)$ the family of functions g on $[0,1]$ defined by,

$$\mathcal{F}(\varepsilon) = \{g : [0,1] \to \mathbb{R} \text{ with } g(\cdot) = \mathbb{E}(U\varepsilon(\cdot)) \text{ where } U \in L_2(\varepsilon)\},$$

We note here that for every $g \in \mathcal{F}(\varepsilon)$, the associated U is unique. It is easy to verify that $\mathcal{F}(\varepsilon)$ is a Hilbert space equipped with the norm $||\ ||$ defined for $g \in \mathcal{F}(\varepsilon)$ by,

$$||g||^2 = \mathbb{E}(U^2).$$

In fact, let $g \in \mathcal{F}(\varepsilon)$, i.e, $g(\cdot) = \mathbb{E}(U\varepsilon(\cdot))$ for some $U \in L_2(\epsilon)$. We have,

- $||g|| = \sqrt{\mathbb{E}(U^2)} \geq 0$.
- $||g|| = \sqrt{\mathbb{E}(U^2)} = 0 \Rightarrow U = 0$ a.s. $\Rightarrow g = 0$.
- For $g \in \mathcal{F}(\varepsilon)$, i.e, $f(\cdot) = \mathbb{E}(V\varepsilon(\cdot))$ some $V \in L_2(\epsilon)$. We have,

$$\begin{aligned} ||g+f||^2 &= \mathbb{E}((U+V)^2) = \mathbb{E}(U^2) + \mathbb{E}(V^2) + 2\mathbb{E}(UV) \\ &\leq \mathbb{E}(U^2) + \mathbb{E}(V^2) + 2\sqrt{\mathbb{E}(U^2)}\sqrt{\mathbb{E}(V^2)} = \left(\sqrt{\mathbb{E}(U^2)} + \sqrt{\mathbb{E}(V^2)}\right)^2. \end{aligned}$$

Thus, $||g+f|| \leq \sqrt{\mathbb{E}(U^2)} + \sqrt{\mathbb{E}(V^2)} = ||g|| + ||f||$.

We now prove the completeness of $\mathcal{F}(\varepsilon)$. For this let $g_n(\cdot) = E(U_n\varepsilon(\cdot))$ be a Cauchy sequence in $\mathcal{F}(\varepsilon)$, i.e.,

$$\lim_{n,m\to\infty} ||g_n - g_m||^2 = 0.$$

From the definition of the norm $||\ ||$ we obtain,

$$\lim_{n,m\to\infty} \mathbb{E}((U_n - U_m)^2) = \lim_{n,m\to\infty} ||g_n - g_m||^2 = 0.$$

This yields that $(U_n)_{n\geq 1}$ is a Cauchy sequence in $L_2(\varepsilon)$, which is a Hilbert space as proven by Parzen (1959) (see page 8 there). Thus it exists $U \in L_2(\varepsilon)$ such that,

$$\lim_{n\to\infty} \mathbb{E}((U_n - U)^2) = 0.$$

Taking $g(\cdot) = \mathbb{E}(U\varepsilon(\cdot))$, which is clearly an element of $\mathcal{F}(\varepsilon)$ gives,

$$\lim_{n\to\infty} ||g_n - g||^2 = \lim_{n\to\infty} \mathbb{E}((U_n - U)^2) = 0.$$

This concludes the proof of completeness of $\mathcal{F}(\varepsilon)$.

The Hilbert space $\mathcal{F}(\varepsilon)$ can easily be identified as the Reproducing Kernel Hilbert Space associated to a reproducing kernel R (with $R(s,t) = \mathbb{E}(\varepsilon(s)\varepsilon(t))$), which is defined as follows.

Definition 2.2.1 *Parzen (1959) A Hilbert space H is said to be a Reproducing Kernel Hilbert Space associated to a reproducing kernel (or function) R (RKHS(R)), if its members are functions on some set T, and if there is a kernel R on $T \times T$ having the following two properties:*

$$\begin{cases} R(\cdot,t) \in H & \textit{for all } \ t \in T, \\ \langle g, R(\cdot,t)\rangle = g(t) & \textit{for all } \ t \in T \ \ \textit{and} \ \ g \in H, \end{cases} \tag{2.3}$$

where $\langle \cdot,\cdot \rangle$ is the inner (or scalar) product in H.

To prove this, we need to verify the properties given in (2.3). For $t \in [0,1]$ we have,

$$R(s,t) = \mathbb{E}(\varepsilon(s)\varepsilon(t)) \quad \text{for all} \ \ s \in [0,1].$$

Since $\varepsilon(s) \in L_2(\varepsilon)$ then $R(\cdot,t) \in \mathcal{F}(\varepsilon)$ for any fixed $t \in [0,1]$. Now let $g \in \mathcal{F}(\varepsilon)$, i.e.,

$$g(\cdot) = \mathbb{E}(U\varepsilon(\cdot)) \ \text{ for some } \ U \in L_2(\varepsilon).$$

Then,

$$\begin{aligned} \langle g, R(\cdot,t)\rangle &= \frac{1}{2}\big(||g||^2 + ||R(\cdot,t)||^2 - ||g - R(\cdot,t)||^2\big) = \frac{1}{2}\big(\mathbb{E}(U^2) + \mathbb{E}(\varepsilon(t)^2) - \mathbb{E}((U-\varepsilon(t))^2)\big) \\ &= \frac{1}{2}\mathbb{E}(2U\varepsilon(t)) = g(t). \end{aligned}$$

These properties together with the following theorem yield that $\mathcal{F}(\varepsilon)$ is the RKHS(R).

Theorem 2.2.1 (E. H. Moor) *Aronszajn (1944) A symmetric non-negative Kernel R generates a unique Hilbert space.*

In the sequel, we take R to be continuous on $[0,1]^2$ and we shall consider the function of interest given by (2.2). More generally, we consider the function f, defined for a continuous function φ and $t \in [0,1]$, by

$$f(t) = \int_0^1 R(s,t)\varphi(s)\, ds. \tag{2.4}$$

Lemma 2.2.1 *We have $f \in \mathcal{F}(\varepsilon)$, i.e., there exists $X \in L_2(\varepsilon)$ with,*

$$f(\cdot) = \mathbb{E}(X\varepsilon(\cdot)). \tag{2.5}$$

In addition,

$$\|f\|^2 = \mathbb{E}(X^2) = \int_0^1 \int_0^1 R(s,t)\varphi(s)\varphi(t)\, dt\, ds.$$

Now let $T_n = (t_1, t_2, \cdots, t_n)$ with $0 \leq t_1 < t_2 < \cdots < t_n \leq 1$ and let V_{T_n} be the subspace of $\mathcal{F}(\varepsilon)$ spanned by the functions $R(\cdot, t)$ for $t \in T_n$, i.e.,

$$V_{T_n} = \{g : [0,1] \to \mathbb{R} \ \text{ with } \ g(\cdot) = \mathbb{E}(U\varepsilon(\cdot)) \ \text{ where } \ U \in L(\varepsilon(t), t \in T_n)\}.$$

Our task is to prove that if $R_{|T_n} = (R(t_i, t_j)_{1 \leq i,j \leq n})$ is a non-singular matrix then V_{T_n} is a closed subspace of $\mathcal{F}(\varepsilon)$. For this let, $(g_m)_{m \geq 1}$ be a sequence in V_{T_n} converging to $g \in \mathcal{F}(\varepsilon)$. We shall prove that $g \in V_{T_n}$. Note that,

$$g_m(t) = \mathbb{E}(U_m \varepsilon(t)) \quad \text{with} \quad U_m = \sum_{i=1}^n a_{i,m}\varepsilon(t_i), \quad \text{where} \ \ (a_{i,m})_{m \geq 1} \in \mathbb{R}.$$

Since g_m converges in $\mathcal{F}(\varepsilon)$ then it is a Cauchy sequence, i.e.,

$$\lim_{m_1, m_2 \to \infty} ||g_{m_1} - g_{m_2}||^2 = 0.$$

By the definition of the norm on $\mathcal{F}(\varepsilon)$ we have,

$$\begin{aligned} ||g_{m_1} - g_{m_2}||^2 &= \mathbb{E}((U_{m_1} - U_{m_2})^2) = \mathbb{E}\Big(\big(\sum_{i=1}^n (a_{i,m_1} - a_{i,m_2})\varepsilon(t_i)\big)^2\Big) \\ &= \sum_{i=1}^n \sum_{j=1}^n (a_{i,m_1} - a_{i,m_2})(a_{j,m_1} - a_{j,m_2}) R(t_i, t_j) = A'_{m_1,m_2} R_{|T_n} A_{m_1,m_2}, \end{aligned}$$

where $A'_{m_1,m_2} = (a_{1,m_1} - a_{1,m_2}, \cdots, a_{n,m_1} - a_{n,m_2})'$. Thus,

$$\lim_{m_1, m_2 \to \infty} A'_{m_1,m_2} R_{|T_n} A_{m_1,m_2} = 0.$$

Since $R_{|T_n}$ is a symmetric positive matrix, we obtain,

$$\lim_{m_1, m_2 \to \infty} A'_{m_1,m_2} = \lim_{m_1, m_2 \to \infty} (a_{1,m_1} - a_{1,m_2}, \cdots, a_{n,m_1} - a_{n,m_2})' = (0, \ldots, 0)',$$

which yields that $(a_{i,m})_m$ is a Cauchy sequence on $\mathbb{R}$ for all $i = 1, \cdots, n$. Taking $a_i = \lim_{m \to \infty} a_{i,m}$ we obtain by the uniqueness of the limit,

$$g(\cdot) = \mathbb{E}(U\varepsilon(\cdot)) \ \text{ with } \ U = \sum_{i=1}^n a_i \varepsilon(t_i),$$

which yields that $g \in V_{T_n}$. Hence V_{T_n} is closed. $\square$

Since V_{T_n} is a closed subspace in the Hilbert space $\mathcal{F}(\varepsilon)$, one can define the orthogonal projection operator from $\mathcal{F}(\varepsilon)$ to V_{T_n} which we note by $P_{|T_n}$, i.e., for every $f \in \mathcal{F}(\varepsilon)$,

$$P_{|T_n} f = \underset{g \, \in V_{T_n}}{\text{argmin}} \, ||f - g||.$$

Par definition of $P_{|T_n}$, we have for any $g \in V_{T_n}$

$$\langle P_{|T_n} f - f, g \rangle = 0.$$

Now, for $t_i \in T_n$, $R(\cdot, t_i) \in V_{T_n}$. Hence, for every $i = 1, \ldots, n$.

$$\langle P_{|T_n} f - f, R(\cdot, t_i) \rangle = 0 \text{ or equivalently } \langle P_{|T_n} f, R(\cdot, t_i) \rangle = \langle f, R(\cdot, t_i) \rangle.$$

The last equality, together with (2.3), gives that,

$$P_{|T_n} f(\cdot) = f(\cdot) \text{ on } T_n. \ \square \tag{2.6}$$

2.3 Construction of the estimator using the RKHS approach

We consider Model (2.1) where g is the unknown regression function on $[0, 1]$ and $\{\varepsilon_j(t), t \in [0, 1]\}_j$ is a sequence of error processes. We assume that $g \in C^2([0, 1])$ and that $(\varepsilon_j)_j$ are i.i.d. processes with the same distribution as a centered second order process ε. We denote by R its autocovariance function, assumed to be known, continuous and forms a non singular matrix when restricted to $T \times T$ for any finite set $T \subset [0, 1]$.

2.3.1 Projection estimator

In this section, we shall give the definition of the new proposed estimator for the regression function g in Model (2.1). This estimator (see Definition 2.3.1 below) is constructed using the function $f_{x,h}$ given by (2.2) for $x \in [0, 1]$, $h \in]0, 1[$ and K is a first order kernel[1] of support $[-1, 1]$ belonging to C^1. This function is well known in time series analysis and has been used by several authors. We mention, among others, the work of Sacks and Ylvisaker (1966) and of Belouni and Benhenni (2015) for linear regression models with correlated errors. It is mainly used due to its belonging to the (RKHS(R)) (see Section 2.2 for more details). This space is spanned by the functions $\{R(\cdot, t_i)_{1 \le i \le n}\}$ forming a closed subspace on which an orthogonal projection of the function $f_{x,h}$ is feasible. We shall call the estimator obtained by this approach, the projection estimator.

The proposed estimator, which is a kernel estimator, is linear in the observations $\overline{Y}(t_i)$ and is given by the following definition.

Definition 2.3.1 *The projection estimator of the regression function g in Model (2.1) based on the observations $(t_i, Y_j(t_i))_{\substack{1 \le i \le n \\ 1 \le j \le m}}$ is given for any $x \in [0, 1]$ by,*

$$\hat{g}_n^{pro}(x) = \sum_{i=1}^{n} m_{x,h}(t_i) \overline{Y}(t_i), \tag{2.7}$$

where $\overline{Y}(t_i) = \frac{1}{m} \sum_{j=1}^{m} Y_j(t_i)$ and the weights $(m_{x,h}(t_i))_{1 \le i \le n}$ are being determined, letting $T_n = (t_i)_{1 \le i \le n}$, by,

$$m'_{x,h_{|T_n}} = {f_{x,h_{|T_n}}}' R_{|T_n}^{-1}, \tag{2.8}$$

with $f_{x,h_{|T_n}} := (f_{x,h}(t_1), \ldots, f_{x,h}(t_n))'$, $R_{|T_n} := (R(t_i, t_j))_{1 \le i,j \le n}$, $R_{|T_n}^{-1}$ the inverse of $R_{|T_n}$ and $m_{x,h_{|T_n}} := (m_{x,h}(t_1), \ldots, m_{x,h}(t_n))'$, where v' denotes the transpose of a vector v.

[1]The kernel K satisfies: $\int_{-1}^{1} K(t)dt = 1$, $\int_{-1}^{1} tK(t)dt = 0$ and $\int_{-1}^{1} t^2 K(t)dt < +\infty$.

Remark 2.3.1 *In order to motivate the proposed estimator, consider the regression model using m continuous experimental units, i.e.,*

$$Y_j(t) = g(t) + \varepsilon_j(t) \textit{ for } t \in [0,1] \quad \textit{and } j = 1, \cdots, m. \tag{2.9}$$

A continuous kernel estimator of g in Model (2.9) is given for any $x \in [0,1]$ by,

$$\hat{g}_{[0,1]}(x) = \int_0^1 \varphi_{x,h}(t)\overline{Y}(t)\ dt \quad \textit{with} \ \ \overline{Y}(t) = \frac{1}{m}\sum_{j=1}^{m} Y_j(t), \tag{2.10}$$

where $\varphi_{x,h}(t) = \frac{1}{h}K\left(\frac{x-t}{h}\right)$ for a kernel K and a bandwidth h. We refer the reader to the works of Blanke and Bosq (2008) and Didi and Louani (2013) for more details on the Kernel estimation of the regression function based on continuous observations.
Since in practice we only have access to discrete observations, then a linear approximation of the continuous kernel estimator should be of the form:

$$\hat{g}_n(x) = \sum_{i=1}^{n} W_{x,h}(t_i)\overline{Y}(t_i).$$

Now let,

$$f_{n,x}(t) = \sum_{i=1}^{n} W_{x,h}(t_i)R(t_i,t) \quad \textit{for} \ \ t \in [0,1].$$

Then the Mean Squared Error (MSE) of approximation can be written as:

$$\mathbb{E}\big(\hat{g}_{[0,1]}(x) - \hat{g}_n(x)\big)^2 = ||f_{x,h} - f_{n,x}||^2,$$

where $f_{x,h}$ is given by (2.2) and $||\cdot||$ is the norm of the RKHS(R)(see Section 2.2 for more details). Then the best linear predictor $\hat{g}_n^{pro}(x)$ of $\hat{g}_{[0,1]}(x)$ satisfies:

$$\inf_{W_{x,h_{|T_n}}} \mathbb{E}\big(\hat{g}_{[0,1]}(x) - \hat{g}_n(x)\big)^2 = ||f_{x,h} - P_{|T_n}f_{x,h}||^2,$$

where $P_{|T_n}f_{x,h}$ is the orthogonal projection of $f_{x,h}$ on the subspace of RKHS spanned by the function $\{R(\cdot,t_i), i = 1,\cdots,n\}$. The optimal coefficients $(W^*_{x,h}(t_i))_{1\le i\le n}$ can then be derived by using the fact that $P_{|T_n}f_{x,h}(t_i) = f_{x,h}(t_i)$ for $i = 1,\cdots,n$ (see Equation (2.6)) and this yields ${W^*_{x,h_{|T_n}}}' = {f_{x,h_{|T_n}}}'R^{-1}_{|T_n}$.

For some classical error processes, such as the Wiener and the Ornstein-Uhlenbeck processes, the estimator (2.7) has a simplified expression as shown in the following proposition.

Proposition 2.3.1 *Consider the regression model (2.1) where ε is of autocovariance function $R(s,t) = \int_0^{\min(s,t)} u^\beta\ du$ for a positive constant β. Let $t_0 = 0,\ t_{n+1} = 1$. Set $\overline{Y}(t_0) = 0$ and $\overline{Y}(t_{n+1}) = \overline{Y}(t_n)$. For any $x \in [0,1]$, the projection estimator (2.7) can be written as follows:*

$$\begin{aligned}\hat{g}_n^{pro}(x) = \frac{1}{\beta+1}\Bigg(&\sum_{i=1}^{n+1}\overline{Y}(t_i)\int_{t_{i-1}}^{t_i}\varphi_{x,h}(s)ds \\ &+ \sum_{i=0}^{n-1}\frac{\overline{Y}(t_{i+1}) - \overline{Y}(t_i)}{t_{i+1}^{\beta+1} - t_i^{\beta+1}}\int_{t_i}^{t_{i+1}}(s^{\beta+1} - t_{i+1}^{\beta+1})\varphi_{x,h}(s)ds\Bigg).\end{aligned} \tag{2.11}$$

Remark 2.3.2 *Taking $\beta = 0$ in the previous proposition gives the expression of the projection estimator* (2.7) *in the case where ε is the classical standard Wiener error process.*

Proposition 2.3.2 *If the error process ε in Model* (2.1) *is the Ornstein-Uhlenbeck process with $R(s,t) = e^{-|t-s|}$ then for any $x \in [0,1]$,*

$$\hat{g}_n^{pro}(x) = \sum_{i=2}^{n-1} \overline{Y}(t_i) \int_{t_{i-1}}^{t_{i+1}} e^{|s-t_i|} \varphi_{x,h}(s)\, ds + \overline{Y}(t_1) \int_0^{t_2} e^{s-t_1} \varphi_{x,h}(s)\, ds$$
$$+ \overline{Y}(t_n) \int_{t_{n-1}}^{1} e^{t_n - s} \varphi_{x,h}(s)\, ds - \sum_{i=1}^{n-1} \frac{e^{t_{i+1}} \overline{Y}(t_{i+1}) - e^{t_i} \overline{Y}(t_i)}{1 - e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{-s} \varphi_{x,h}(s)\, ds$$
$$+ \sum_{i=1}^{n-1} \frac{e^{-t_{i+1}} \overline{Y}(t_{i+1}) - e^{-t_i} \overline{Y}(t_i)}{1 - e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds,$$

where $\varphi_{x,h}$ is defined in the previous proposition.

Remark 2.3.3 *As the previous propositions show, the expression of $m_{x,h_{|T_n}}$ is known analytically for error processes of practical interest. For more complicated error processes, numerical methods can be used. For more general error processes, we will give an asymptotic simplified expression of the weights of the projection estimator (see Lemma 2.4.2 below).*

2.3.2 Assumptions and comments

In order to derive our asymptotic results, the following assumptions on the autocovariance function R and the Kernel K are required.

(A) R is continuous on the entire unit square and has left and right derivatives up to order two at the diagonal (i.e. when $s = t$), i.e.,

$$R^{(0,1)}(t, t^-) = \lim_{s \uparrow t} \frac{\partial R(t,s)}{\partial s} \quad \text{and} \quad R^{(0,1)}(t, t^+) = \lim_{s \downarrow t} \frac{\partial R(t,s)}{\partial s},$$

exist and are continuous. In a similar way we define $R^{(0,2)}(t,t^-)$ and $R^{(0,2)}(t,t^+)$.
Off the diagonal (i.e. when $s \neq t$ in the unit square), R has continuous derivatives up to order two.

For $t \in]0,1[$, let $\alpha(t) = R^{(0,1)}(t,t^-) - R^{(0,1)}(t,t^+)$. Assumption (A) gives the following lemma concerning the jump function α.

Lemma 2.3.1 *If Assumption (A) is satisfied then the jump function α is a positive function.*

To obtain our asymptotic results, we shall give next a stronger assumption on the jump function α.

(B) We assume that α is Lipschitz on $]0,1[$, $\inf_{0<t<1} \alpha(t) = \alpha_0 > 0$ and $\sup_{0<t<1} \alpha(t) = \alpha_1 < \infty$.

Assumptions (A) and (B) are classical regularity conditions and were used in several works, see for instance, Sacks and Ylvisaker (1966), Su and Cambanis (1993) and most recently Belouni and Benhenni (2015).

(C) For each $t \in [0,1]$, $R^{(0,2)}(.,t^+)$ is in the Reproducing Kernel Hilbert space associated to R, denoted by RKHS(R), equipped with the norm $||\cdot||$. In addition, $\sup_{0\le t\le 1} ||R^{(0,2)}(.,t^+)|| < \infty$ (see Section 2.2 for more details).

Assumption (C), which is more restrictive than (B) as indicated by Sacks and Ylvisaker (1966), is necessary to evaluate the weights of the projection estimator (see Lemma 2.4.2 below).

(D) K is an even function and K' is a Lipschitz function on $[-1,1]$.

Examples of autocovariance functions which satisfy Assumptions (A), (B) and (C) are given below.

Example 2.3.1

1. *The autocovariance function $R(s,t) = \sigma^2 min(s,t)$ of the Wiener process, has a constant jump function $\alpha(t) = \sigma^2$ and $R^{(i,j)}(s,t) = 0$ for all integers i,j such that $i+j=2$ and $s \neq t$.*

2. *The autocovariance function $R(s,t) = \sigma^2 e^{-\lambda|s-t|}$ of the stationary Ornstein-Uhnelbeck process with $\sigma > 0$ and $\lambda > 0$. For this process the jump function is $\alpha(t) = 2\sigma^2\lambda$ and $R^{(0,2)}(s,t) = \sigma^2\lambda^2 e^{-\lambda|s-t|}$.*

3. *Another general class of autocovariance functions was given by Sacks and Ylvisaker (1966) and has the form,*
$$R(s,t) = \int_0^{1/|t-s|} (1-\mu|t-s|)p(\mu)\, d\mu,$$
where p is a probability density and p' its derivative are such that,
$$\lim_{\mu\to\infty} \mu^3 p(\mu) < \infty, \quad \text{and} \quad \int_a^{\infty} (\mu p'(\mu) + 3p(\mu))^2)\mu^6 d\mu < \infty,$$
for some a. We have $\alpha(t) = 2\int_0^\infty up(u)\, du$.

2.4 Local asymptotic results

Let $T_n = (t_{i,n})_{1\le i\le n}$ for $n \ge 1$, be a fixed sequence of designs with $T_n \in D_n$, where,
$$D_n = \{(s_1, s_2, \ldots, s_n) \ : \ 0 \le s_1 < s_2 < \cdots < s_n \le 1\}.$$
Set $t_{0,n} = 0, t_{n+1,n} = 1$, $d_{j,n} = t_{j+1,n} - t_{j,n}$ and let for $x \in [0,1]$, $h = h(n)$,
$$I_{x,h} = \{i = 1, \cdots, n \ : \ [t_{i-1,n}, t_{i+1,n}] \cap]x-h, x+h[\neq \emptyset\}.$$
Denote by $N_{T_n} = \text{Card}(I_{x,h})$. Recall that $[x-h,x+h]$ is the support of the function $\varphi_{x,h}$. To obtain the asymptotic results, we require that the sequence $(T_n)_{n\ge1}$ satisfies the next assumption.

(E) $\lim_{n\to\infty} \sup_{0\le j\le n} d_{j,n} = 0$, $\lim_{n\to\infty}\left(\frac{1}{h}\sup_{0\le j\le n} d_{j,n}\right) = 0$, $\lim_{n\to\infty}\left(N_{T_n}\frac{1}{h^2}\sup_{0\le j\le n} d_{j,n}^2\right) = 0$ and $\limsup_{n\to\infty}\left(N_{T_n}^2\frac{1}{h^2}\sup_{0\le j\le n} d_{j,n}^2\right) < \infty$.

A simple sequence of designs that verifies Assumption (E) was presented by Sacks and Ylvisaker (1970) as follows.

Definition 2.4.1 *Let F be a distribution function of some density function f such that $\sup\limits_{0<t<1} f(t) < \infty$ and $\inf\limits_{0<t<1} f(t) > 0$. The so-called regular sequence of designs generated by f is defined by,*

$$T_n = \left\{ t_{i,n} = F^{-1}\Big(\frac{i}{n}\Big), i = 1, \dots, n \right\}.$$

In the sequel, the density f is assumed to be at least in $C^2([0,1])$. This sequence of designs verifies the following Lemma (see for instance Benelmadani *et al.* (2019a) for its proof).

Lemma 2.4.1 *Let $(T_n)_{n\geq 1}$ be a regular sequence of designs generated by some density function. For $x \in]0,1[$ and $h > 0$, suppose that $T_n \cap [x-h, x+h] \neq \emptyset$ and that $nh \geq 1$. Then,*

$$\sup_{0\leq j\leq n} d_{j,n} = O\Big(\frac{1}{n}\Big) \quad \text{and} \quad N_{T_n} = O(nh), \tag{2.12}$$

where N_{T_n} and $d_{j,n}$ are defined as above. In addition, if $\lim\limits_{n\to\infty} nh = \infty$ then the regular sequence of designs verifies Assumption (E).

2.4.1 Evaluation of the bias

In order to derive the asymptotic expression of the bias term of the projection estimator, we shall first give the asymptotic approximation of the weights $m_{x,h|T_n}$ (defined by (2.8)) in the following lemma.

Lemma 2.4.2 *Suppose that Assumptions $(A), (B)$ and (C) are satisfied. Then for any $x \in]0,1[$,*

$$m_{x,h}(t_{i,n}) = \begin{cases} \frac{1}{2}\varphi_{x,h}(t_{i,n})(t_{i+1,n} - t_{i-1,n}) + O\big(\alpha_{n,h} + \beta_{n,h}\big) & \text{if } i \notin \{1,n\} \text{ and} \\ & [t_{i-1,n}, t_{i+1,n}] \cap [x-h, x+h] \neq \emptyset, \\ O\big(N_{T_n}\alpha_{n,h} + n\beta_{n,h}\big) & \text{if } i \in \{1,n\}, \\ O\big(\beta_{n,h}\big) & \text{otherwise,} \end{cases}$$

where,

$$\alpha_{n,h} = \sup_{0\leq i\leq n} \sup_{t_{i,n}\leq s,t\leq t_{i+1,n}} d_{i,n}\ |\alpha(s)\varphi_{x,h}(s) - \alpha(t)\varphi_{x,h}(t)| = O\Big(\frac{1}{h^2}\sup_{0\leq j\leq n} d_j^2\Big),$$

$$\beta_{n,h} = \sup_{0\leq t\leq 1} \frac{1}{\alpha(t)}||R^{(0,2)}(.,t)||\frac{\sqrt{C}}{\sqrt{h}} \sup_{0\leq j\leq n} d_j^2 = O\Big(\frac{1}{\sqrt{h}} \sup_{0\leq j\leq n} d_j^2\Big),$$

and C is a positive constant defined in Proposition 2.4.3 below.

Remark 2.4.1 *This Lemma shows that the weights of the projection estimator are asymptoticly equivalent to those of some well known linear estimators of the regression function g. For instance,*

- *Priestly and Chao (1972) used the following weights:*
$$W_{x,h}(t_i) = (t_{i+1,n} - t_{i,n})\varphi_{x,h}(t_i) \quad \text{for } i = 1, \cdots, n.$$
- *Gasser and Müller (1979) used the following weights:*
$$W_{x,h}(t_i) = \int_{s_{i-1,n}}^{s_{i,n}} \varphi_{x,h}(s)\ ds \qquad \text{for } i = 1, \cdots, n,$$
where, $s_0 = 0$, $s_n = 1$ *and* $s_{i,n} = (t_{i+1,n} + t_{i,n})/2$ *for* $i = 1, \cdots, n-1$.
- *Cheng and Lin (1981) replaced* $s_{i,n}$ *by* $t_{i,n}$*, in the weights of the Gasser and Müller estimator.*

Using the asymptotic approximation of the weights given in Lemma 2.4.2, we can obtain the asymptotic expression of the bias of the projection estimator as shows the following proposition.

Proposition 2.4.1 *Suppose that Assumptions* $(A) - (D)$ *are satisfied. If* $T_n \cap]x-h, x+h[\neq \emptyset$ *and* $nh \geq 1$*, then for any* $x \in]0,1[$*,*
$$\mathbb{E}(\hat{g}_n^{pro}(x)) - g(x) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{N_{T_n}}{h^3} \sup_{0 \leq j \leq 1} d_{j,n}^3 + N_{T_n}\alpha_{n,h} + n\beta_{n,h}\Big),$$
where $\alpha_{n,h}$ *and* $\beta_{n,h}$ *are given in Lemma 2.4.2 and* $B = \int_{-1}^{1} t^2 K(t)\ dt$*.*

Remark 2.4.2 *Under the assumption of Lemma 2.4.1 we have,*
$$\mathbb{E}\ (\hat{g}_n^{pro}(x)) - g(x) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{1}{nh}\Big).$$

In the case of a Wiener error process, a direct computation of the bias term of the projection estimator (2.11), with $\beta = 0$, shows that the order term $O\Big(\frac{1}{nh}\Big)$ can be improved. The result is given by the following proposition.

Proposition 2.4.2 *Consider Model* (2.1) *with a Wiener error process of autocovariance function* $R(s,t) = \min(s,t)$*. Let* $(T_n)_{n \geq 1}$ *be a regular sequence of designs generated by a density function* f *(cf. Definition 2.4.1) and let* K *be a kernel satisfying Assumption* (D)*. If* $T_n \cap]x-h, x+h[\neq \emptyset$ *and* $nh \geq 1$ *then,*
$$\mathbb{E}\ (\hat{g}_n^{pro}(x)) - g(x) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{1}{n^2 h}\Big),$$
where B *is given in Proposition 2.4.1 above.*

2.4.2 Evaluation of the variance

It is shown in Lemma 2.2.1 of Section 2.2 that $f_{x,h}$ defined by (2.2) belongs to the RKHS(R) equipped with its norm $||\ ||$ and,
$$||f_{x,h}||^2 = \int_0^1 \int_0^1 \varphi_{x,h}(s)R(s,t)\varphi_{x,h}(t)ds\ dt \triangleq \sigma_{x,h}^2. \tag{2.13}$$

In addition if $P_{|T_n} f_{x,h}$ is the projection of $f_{x,h}$ on the subspace of $\mathcal{F}$ spanned by $\{R(.,t), t \in T_n\}$ then it is shown by (F2) in the supplementary facts of the Appendix that,
$$||P_{|T_n} f_{x,h}||^2 = m\text{Var }\hat{g}_n^{pro}(x). \tag{2.14}$$

The following proposition controls the residual variance $\frac{\sigma_{x,h}^2}{m} - \text{Var }\hat{g}_n^{pro}(x)$.

Proposition 2.4.3 *Suppose that Assumptions* (A) *and* (B) *are satisfied. Moreover, assume that* $\frac{1}{h}\sup_{1\le i\le n} d_i \le 1$ *and let,*

$$K_\infty = \sup_{t\in[-1,1]} |K(t)|, \quad R_1 = \sup_{t,s\in[0,1]} |R^{(1,1)}(s-,t+)| \quad \text{and} \quad R_2 = \sup_{t,s\in[0,1]} |R^{(0,2)}(s,t+)|.$$

Then we have, for any $x \in]0,1[$,

$$0 \le \frac{\sigma^2_{x,h}}{m} - \text{Var } \hat{g}^{pro}_n(x) \le \frac{C}{mh} \sup_{0\le j\le n} d^2_{j,n},$$

where $C = \begin{cases} K^2_\infty(\frac{4}{3}\alpha_1 + R_1 + \frac{4}{3}R_2) & \text{if } (x-h) \text{ and } (x+h) \in T_n, \\ K^2_\infty(\frac{8}{3}\alpha_1 + \frac{5}{3}R_1 + \frac{8}{3}R_2) & \text{otherwise.} \end{cases}$

If moreover $\{T_n, n\ge 1\}$ satisfies Assumption (E) then Proposition 2.4.3 gives,

$$\lim_{n,m\to\infty}\Big(\text{Var } \hat{g}^{pro}_n(x) - \frac{\sigma^2_{x,h}}{m}\Big) = 0.$$

The next proposition gives the rate of convergence of this residual variance.

Proposition 2.4.4 *Suppose that Assumptions* $(A),(B)$ *and* (C) *are satisfied. Moreover, assume that* $(T_n)_{n\ge1}$ *is a sequence of designs verifying Assumption* (E)*. Then for any* $x\in]0,1[$ *and for any positive integer* m,

$$\underline{\lim}_{n\to\infty} \frac{mN^2_{T_n}}{h}\Big(\frac{\sigma^2_{x,h}}{m} - \text{Var } \hat{g}^{pro}_n(x)\Big) \ge \frac{1}{12}\alpha(x)\Big\{\int_{-1}^{1} K^{2/3}(t)dt\Big\}^3, \tag{2.15}$$

where $\sigma^2_{x,h}$ *is given by* (2.13).

Using Propositions 2.4.3 and 2.4.4 we can obtain the optimal convergence rate $1/(mn^2h)$ of the residual variance. The result is given by the following proposition.

Proposition 2.4.5 *Suppose that all the assumptions of Lemma 2.4.1, Propositions 2.4.3 and 2.4.4 are satisfied. Then there exist some positive constants* C *and* C' *such that for any* $x\in]0,1[$ *and for any positive integer* m,

$$\overline{\lim}_{n\to\infty} mn^2h\Big(\frac{\sigma^2_{x,h}}{m} - \text{Var}(\hat{g}^{pro}_n(x))\Big) \le C, \tag{2.16}$$

and,

$$\underline{\lim}_{n\to\infty} mn^2h\Big(\frac{\sigma^2_{x,h}}{m} - \text{Var } \hat{g}^{pro}_n(x)\Big) \ge C'. \tag{2.17}$$

Under the stronger assumption (D) on the kernel K and using a regular sequence of designs (see Definition 2.4.1), we obtain the asymptotic expression of the variance as shown by the following proposition.

Proposition 2.4.6 *Suppose that Assumptions* $(A)-(D)$ *are satisfied. Moreover assume that* $(T_n)_{n\geq 1}$ *is a regular sequence of designs generated by a density function* f *(see Definition 2.4.1). If* $\lim_{n\to\infty} h=0$ *and* $\lim_{n\to\infty} nh=\infty$ *then for any* $x\in]0,1[$,

$$\mathrm{Var}(\hat{g}_n^{pro}(x))=\frac{\sigma^2_{x,h}}{m}-\frac{1}{12mn^2}\int_{x-h}^{x+h}\frac{\alpha(t)}{f^2(t)}\varphi^2_{x,h}(t)dt+O\Big(\frac{1}{mn^3h^2}\Big), \tag{2.18}$$

where $\sigma^2_{x,h}$ *is given by* (2.13).

The following lemma (proved in Benhenni and Rachdi (2007) gives the expression of the main term of the asymptotic variance $\sigma^2_{x,h}/m$ in terms of h.

Lemma 2.4.3 *Suppose that Assumptions* (A), (B) *and* (D) *are satisfied. If* $\lim_{n\to\infty} h=0$ *then, for any* $x\in]0,1[$, $\sigma^2_{x,h}$ *(as given by* (2.13)*) has the following asymptotic expression*

$$\sigma^2_{x,h}=\big(R(x,x)-\frac{1}{2}\alpha(x)C_Kh\big)+o(h), \tag{2.19}$$

where $C_K=\int_{-1}^{1}\int_{-1}^{1}|u-v|K(u)K(v)dudv$.

2.4.3 IMSE and optimal bandwidth

Proposition 2.4.6 and Remark 2.4.2 allow to derive the asymptotic expression of the Mean Squared Error (MSE) and the Integrated Mean Squared Error (IMSE) of the projection estimator (2.7) given, without proof, in the next theorem.

Theorem 2.4.1 *If all the assumptions of Propositions 2.4.1 and 2.4.6 are satisfied and if* $(T_n)_{n\geq 1}$ *is a regular sequence of designs generated by some density function (see Definition 2.4.1) then for any* $x\in]0,1[$,

$$\begin{aligned}\mathrm{MSE}(\hat{g}_n^{pro}(x))&=\frac{1}{m}\Big(R(x,x)-\frac{1}{2}\alpha(x)C_Kh\Big)+\frac{1}{4}h^4(g''(x))^2B^2+o\Big(h^4+\frac{h}{m}\Big)\\&\quad+O\Big(\frac{1}{mn^2h}+\frac{h}{n}+\frac{1}{n^2h^2}\Big),\end{aligned}$$

$$\begin{aligned}\mathrm{IMSE}(\hat{g}_n^{pro})&=\frac{1}{m}\int_0^1 R(x,x)w(x)\,dx-\frac{C_Kh}{2m}\int_0^1\alpha(x)w(x)\,dx\\&\quad+\frac{B^2}{4}h^4\int_0^1[g''(x)]^2w(x)\,dx+o\Big(h^4+\frac{h}{m}\Big)+O\Big(\frac{1}{mn^2h}+\frac{h}{n}+\frac{1}{n^2h^2}\Big),\end{aligned}$$

where w *is a positive density function,* B *is given in Proposition 2.4.1 and* C_K *is given in Lemma 2.4.3.*

Remark 2.4.3 *We note here that the term* $\frac{1}{12mn^2}\int_{x-h}^{x+h}\frac{\alpha(t)}{f^2(t)}\varphi^2_{x,h}(t)dt$ *appearing in the asymptotic variance, does not appear in the asymptotic* MSE *and* IMSE*, because it is negligible comparing to the squared bias, precisely due to the term* $O\Big(\frac{1}{nh}\Big)$.

However in the case of a Wiener error process, we have proven (see Proposition 2.4.2) that the previous term can be replaced by $O\Big(\frac{1}{n^2h}\Big)$ *when using exact weights of the projection estimator (and not their asymptotic expression). Therefor, when* ε *is a Wiener process, the asymptotic expressions of the* MSE *and* IMSE *of the projection estimator* (2.11) *(with* $\beta=0$*) are given by the following theorem.*

Theorem 2.4.2 *Consider Model* (2.1) *with a Wiener error process and suppose that the kernel* K *verifies Assumption* (D). *Moreover, assume that* $(T_n)_{n\geq 1}$ *is a regular sequence of designs generated by a function* f *(see Definition 2.4.1). If* $\lim_{n\to\infty} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$ *then for any* $x \in]0,1[$,

$$\begin{aligned}\mathrm{MSE}(\hat{g}_n^{pro}(x)) &= \frac{1}{m}\Big(R(x,x) - \frac{1}{2}\alpha(x)C_K h\Big) - \frac{1}{mn^2h}\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1} K^2(t)\,dt \\ &+ \frac{1}{4}h^4[g''(x)]^2B^2 + o\Big(\frac{h}{m} + h^4\Big) + O\Big(\frac{h}{n^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2} + \frac{1}{n^4h^2}\Big),\end{aligned}$$

and,

$$\begin{aligned}\mathrm{IMSE}(\hat{g}_n^{pro}) &= \frac{1}{m}\int_0^1 R(x,x)w(x)\,dx - \frac{C_K h}{2m}\int_0^1 \alpha(x)w(x)\,dx \\ &- \frac{A}{12mn^2h}\int_0^1 \frac{\alpha(x)}{f^2(x)}w(x)\,dx + \frac{B^2}{4}h^4\int_0^1 [g''(x)]^2 w(x)\,dx + o\Big(\frac{h}{m} + h^4\Big) \\ &+ O\Big(\frac{h}{n^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2} + \frac{1}{n^4h^2}\Big),\end{aligned}$$

where $A = \int_{-1}^{1} K^2(t)\,dt$, w, B *and* C_K *are given in Theorem 2.4.1.*

The asymptotic optimal bandwidth is obtained by minimizing the asymptotic IMSE and is given in the following corollary.

Corollary 2.4.1 (Optimal bandwidth) *Suppose that the assumptions of Theorem 2.4.1 are satisfied. Moreover assume that* $\frac{n}{m} = O(1)$ *as* $n, m \to \infty$. *Denote by IMSE(h) the IMSE of the projection estimator when the bandwidth h is used. Then the bandwidth,*

$$h^* = \left(\frac{C_K\int_0^1 \alpha(x)w(x)\,dx}{2B\int_0^1 [g''(x)]^2 w(x)\,dx}\right)^{1/3} m^{-1/3}, \tag{2.20}$$

is optimal in the sense that,

$$\overline{\lim_{n,m\to\infty}} \frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})} \leq 1,$$

for any sequence of bandwidths $h_{n,m}$ *verifying:*

$$\lim_{n,m\to\infty} h_{n,m} = 0 \quad \text{and} \quad \overline{\lim_{n,m\to\infty}} m h_{n,m}^3 < +\infty.$$

2.4.4 Asymptotic normality

The next theorem presents the asymptotic normality of the projection estimator (2.7) for any error process ε.

Theorem 2.4.3 *Suppose that the assumptions of Theorem 2.4.1 are satisfied. Moreover assume that* $\frac{n}{m} = O(1)$ *as* $n, m \to \infty$, *that* $\lim_{n,m\to\infty} nh^2 = \infty$ *and that* $\lim_{n,m\to\infty} \sqrt{m}h^2 = 0$. *Then for any* $x \in]0,1[$,

$$\sqrt{m}\Big(\hat{g}_n^{pro}(x) - g(x)\Big) \xrightarrow{\mathscr{D}} Z \quad \text{with } Z \sim \mathcal{N}(0, R(x,x)) \quad \text{as } n, m \to \infty,$$

where $\mathscr{D}$ *denotes the convergence in distribution and* $\mathcal{N}$ *is the normal distribution.*

2.5 Comparison with the Gasser and Müller's estimator

In this section, we shall perform a theoretical comparison between the projection estimator given in (2.7) and the classical estimator proposed by Gasser and Müller (1979) that we recall in the definition below.

Definition 2.5.1 *The Gasser and Müller's estimator of the regression function g based on the observations $(t_i, Y_j(t_i))_{\substack{1\le i\le n\\ 1\le j\le m}}$ is given for any $x \in [0,1]$ by,*

$$\hat{g}_n^{GM}(x) = \sum_{i=1}^{n} \overline{Y}(t_i) \int_{s_{i-1}}^{s_i} \varphi_{x,h}(s)\, ds\ , \tag{2.21}$$

where $\overline{Y}, \varphi_{x,h}$ and h are given in Definition 2.3.1. The midpoints $(s_i)_{1\le i\le n}$ are such that: $s_0 = 0, s_n = 1$ and for $i = 1, \ldots, n-1$, $s_i = (t_i + t_{i+1})/2$.

In order to compare this estimator to the projection estimator with respect to the IMSE, we recall in the next theorem the asymptotic expression of the IMSE of the Gasser and Müller's estimator (given in Chapter 1).

Theorem 2.5.1 *Suppose that Assumptions (A), (B) and (D) are satisfied. Moreover assume that $(T_n)_{n\ge1}$ is a regular sequence of designs generated by a density function f (see Definition 2.4.1). If $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$ then for any $x \in]0,1[$,*

$$\begin{aligned}\mathrm{MSE}(\hat{g}_n^{GM}(x)) &= \frac{1}{m}\Big(R(x,x) - \frac{1}{2}\alpha(x)C_K h\Big) + \frac{1}{4}h^4(g''(x))^2B^2 + o\Big(h^4 + \frac{h}{m}\Big)\\ &\quad + O\Big(\frac{h}{n^2} + \frac{1}{n^4h^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big),\end{aligned}$$

and,

$$\begin{aligned}\mathrm{IMSE}(\hat{g}_n^{GM}) &= \frac{1}{m}\int_0^1 R(x,x)w(x)\,dx - \frac{C_K h}{2m}\int_0^1 \alpha(x)w(x)\,dx + \frac{B^2}{4}h^4\int_0^1 [g''(x)]^2 w(x)\,dx\\ &\quad + o\Big(h^4 + \frac{h}{m}\Big) + O\Big(\frac{h}{n^2} + \frac{1}{n^4h^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big),\end{aligned}$$

where B and C_K are given in Propositions 2.4.1 and 2.4.6 and w is a continuous positive density.

The following theorem gives an asymptotic comparison in term of the variance of the projection estimator (2.7) and the Gasser and Müller's estimator (2.21).

Theorem 2.5.2 *Suppose that Assumptions $(A),(B)$ and (D) are satisfied. Moreover assume that $(T_n)_{n\ge1}$ is a regular sequence of designs generated by a density function f (see Definition 2.4.1). If $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$ then for any $x \in]0,1[$,*

$$\lim_{n,m\to\infty} mn^2h\,\Big(\mathrm{Var}\ \hat{g}_n^{GM}(x) - \mathrm{Var}\ \hat{g}_n^{pro}(x)\Big) = \frac{1}{12}\frac{\alpha(x)}{f^2(x)} > 0.$$

For a comparison of the bias of these estimators, we mention that the Gasser and Müller's estimator converges to zero slightly faster than the bias of the projection estimator, i.e., the term $O(\frac{1}{nh})$ in the bias of the projection estimator (see Remark 2.4.2) is replaced by $O(\frac{1}{n^2h})$ in the bias of the Gasser and Müller's estimator (see Proposition 1.3.1). However, for the Wiener error process both estimators have the same bias convergence rates, thus we can compare the asymptotic IMSE of both estimators in the following theorem.

Theorem 2.5.3 *Consider Model* (2.1) *where ε is a Wiener error process. Suppose that the assumptions of Theorem 2.4.1 are satisfied. Moreover, assume that $\lim_{n\to\infty} nh^2 = 0$ and that $\frac{m}{n} = O(1)$ then,*

$$\lim_{n,m\to\infty} mn^2h\,(\text{IMSE}\,(\hat{g}_n^{GM}) - \text{IMSE}\,(\hat{g}_n^{pro})) = \frac{\sigma^2}{12}\int_0^1 \frac{w(x)}{f^2(x)}\,dx > 0.$$

Remark 2.5.1 *Theorems 2.5.2 and 2.5.3 show that, the projection estimator has an asymptotically smaller variance than the Gasser and Müller's estimator for any error process, it also has an asymptotically smaller IMSE when ε is a Wiener error process. However the Gasser and Müller's estimator doesn't require the knowledge of the autocovariance function whereas the projection estimator does.*

2.6 Simulation study

In this section, we investigate the performance of the proposed estimator (2.7) using finite values of experimental units m and sampling points n. The following growth curves are considered:

$$\text{(M1)} \qquad g(x) = 10x^3 - 15x^4 + 6x^5 \quad \text{for} \;\; 0 < x < 1.$$
$$\text{(M2)} \qquad g(x) = x + 0.5\, e^{-80(x-0.5)^2} \quad \text{for} \;\; 0 < x < 1.$$

This growth curves were used by Hart and Wherly (1986) and Benhenni and Rachdi (2007) due to its similarity in shape to that of the logistic function, which is frequently found in growth curve analysis as noted by Hart and Wherly (1986). The sampling points are taken to be:

$$t_i = (i - 0.5)/n \;\; \text{for } i = 1, \cdots, n. \tag{2.22}$$

The error process ε is taken to be the Wiener error process with autocovariance function $R(s,t) = \sigma^2 \min(s,t)$. The Kernel used here is the quartic kernel given by $K(u) = \frac{15}{16}(1-u^2)^2 I_{[-1,1]}(u)$ and the bandwidth is the optimal one with respect to the exact IMSE, obtained using the Conjugated Gradient Algorithm (CGA). We consider the mean of all estimators obtained from 100 simulations. We take $\sigma^2 = 0.5$, simulations for other values of σ^2 gave similar results. The results are given in Figures 2.1 and 2.2 for Models (M1) and (M2) respectively, for a fixed number of observations $n = 100$ and three different values of experimental units $m = 5, 20, 50$.

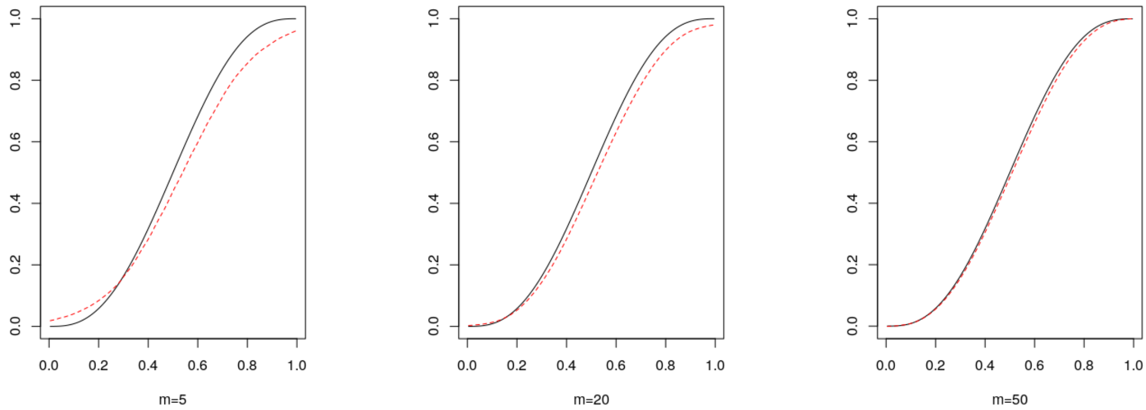


Figure 2.1: The regression function of model (M1) is in plain line and the projection estimator is in dashed line.

We can see for Model (M1), from Figure 2.1, that the projection estimator gets closer to the regression function when m gets bigger, which proves its good performance and consistency when m increases. These results are confirmed for the growth curve Model (M2) given in Figure 2.2.

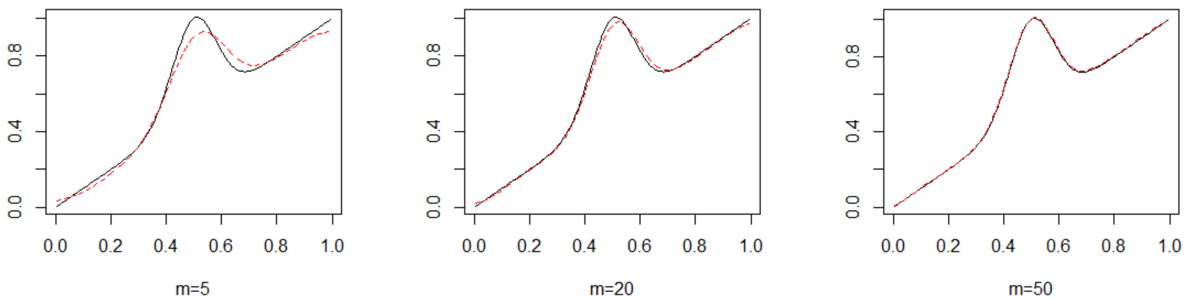


Figure 2.2: The regression function of model (M2) is in plain line and the projection estimator is in dashed line.

In this simulation study, we consider the comparison of the proposed estimator (2.7) to the Gasser and Müller (2.21) (referred by GM estimator) with respect to the exact IMSE in a finite sample set. For this, we consider the cubic growth curve of model (M1). We consider also the uniform design given by (2.22) and the quartic kernel $K(u) = \frac{15}{16}(1-u^2)^2 I_{[-1,1]}(u)$. For the error process, we shall consider both the Wiener of autocovariance function $R(s,t) = \min(s,t)$, and the Ornstein-Uhlenbeck process with autocovariance $R(s,t) = e^{-|s-t|}$.

The weight w, chosen here, is the uniform density on $[0,1]$, i.e., $w \equiv 1$ on $[0,1]$, we consider the optimal bandwidth with respect to the exact IMSE of the two estimators, i.e., $\inf_{0<h<1} \text{IMSE}(h)$. The bandwidth h is chosen through the algorithm CGA. The results are given in Tables 2.1 and 2.2 for $n = 10$ and for different values of m. These tables present the integrated bias squared denoted by $Ibias^2$, integrated variance denoted by $Ivar$ and the IMSE together with the optimal bandwidth associated to each estimator.

First, we can see from these two tables that, the optimal bandwidth decreases when m increases, as shown in Corollary 2.4.1. In addition, the optimal bandwidth of the projection estimator is slightly smaller than that of the GM estimator.

It is also seen that both the $Ivar$ and the $Ibias^2$, of the two estimators decrease when m increases. In addition, the projection estimator has a smaller $Ibias^2$ and $Ivar$ than that of the GM estimator, which leads to a smaller IMSE.

Another way to look at these results is as follows: for a fixed number of experimental units $m = 10$ and when the error process is a Wiener process (similar results for the Ornstein-Uhlenbeck error process), the projection estimator would only need $n = 10$ observations on each experimental unit to obtain the performance IMSE $= 4.53 \times 10^{-2}$ (see Table 2.1), whereas the GM estimator would need to have $n = 18$ observations to obtain the same performance, and thus requires 80% more samples in order to achieve the same performance.

The results of this simulation study show that, even for small number of observations, the projection estimator outperforms the GM estimator with respect to IMSE.

It should be noted here that, in order to solve the problem at the edges $[0, h] \cap [1 - h, 1]$, it was necessary to adjust the kernel as suggested by Hart and Wherly (1986).

Table 2.1: The integrated squared bias, integrated variance, IMSE and the optimal bandwidth for $n = 10$ and different values of m under the Wiener error process, for the GM and the projection estimators.

$n = 10$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	10	1.508×10^{-3}	4.507×10^{-2}	4.658×10^{-2}	0.335
Pro		1.304×10^{-3}	4.399×10^{-2}	4.530×10^{-2}	0.321
GM	50	2.662×10^{-4}	9.494×10^{-3}	9.760×10^{-3}	0.198
Pro		1.981×10^{-4}	9.228×10^{-3}	9.426×10^{-3}	0.187
GM	100	1.505×10^{-4}	4.826×10^{-3}	4.977×10^{-3}	0.154
Pro		0.897×10^{-4}	4.689×10^{-3}	4.778×10^{-3}	0.142

Table 2.2: The integrated squared bias, integrated variance, IMSE and the optimal bandwidth for $n = 10$ and different values of m under the Ornstein-Uhlenbeck error process, for the GM and the projection estimators.

$n = 10$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	10	2.596×10^{-3}	8.821×10^{-2}	9.080×10^{-2}	0.387
Pro		2.494×10^{-3}	8.703×10^{-2}	8.952×10^{-2}	0.386
GM	50	4.481×10^{-4}	1.848×10^{-2}	1.893×10^{-2}	0.236
Pro		4.097×10^{-4}	1.822×10^{-2}	1.863×10^{-2}	0.237
GM	100	2.299×10^{-4}	9.390×10^{-3}	9.620×10^{-3}	0.186
Pro		1.885×10^{-4}	9.265×10^{-3}	9.453×10^{-3}	0.187

2.7 Proofs

In this section, we shall omit the index n in $t_{i,n}$ when there is no ambiguity.

Proof of Lemma 2.2.1

Define, for a suitable partition $(x_{i,n})_{i=1,\cdots,n}$ of $[0,1]$,

$$X_n = \sum_{i=1}^{n-1}(x_{i+1,n} - x_{i,n})\varphi(x_{i,n})\varepsilon(x_{i,n}) \in L_2(\varepsilon),$$

such that for any $t \in [0,1]$,

$$f(t) = \lim_{n\to\infty} \sum_{i=1}^{n-1}(x_{i+1,n} - x_{i,n})\varphi(x_{i,n})R(x_{i,n},t) = \lim_{n\to\infty} \mathbb{E}(X_n\varepsilon(t)).$$

We shall prove that $(X_n)_n$ converges to a certain element of $\mathbb{L}^2$, i.e.,

$$\exists\ X \in \mathbb{L}^2\ :\ \lim_{n\to\infty} \mathbb{E}\big((X_n - X)^2\big) = 0, \tag{2.23}$$

and by the definition of $L_2(\varepsilon)$ the limit in (2.23) proves that X is an element of $L_2(\varepsilon)$. Now the proof (2.23) is immediate, in fact it is easy to check that (X_n) id a Cauchy sequence in $\mathbb{L}^2$. By the completeness of $\mathbb{L}^2$, we deduce (2.23). In addition we have, $\lim_{n\to\infty} \mathbb{E}(X_n\varepsilon(t)) = \mathbb{E}(X\varepsilon(t))$, this is due to the following inequality,

$$\Big|\mathbb{E}(X_n\varepsilon(t)) - \mathbb{E}(X\varepsilon(t))\Big| \le \mathbb{E}\Big|(X_n - X)\varepsilon(t)\Big| \le \sqrt{\mathbb{E}((X_n - X)^2)}\sqrt{\mathbb{E}(\varepsilon(t)^2)},$$

and the fact that $\lim_{n\to\infty} \mathbb{E}((X_n - X)^2) = 0$ and $\mathbb{E}(\varepsilon(t)^2) < \infty$. The proof of (2.5) is concluded. Finally,

$$\mathbb{E}(X^2) = \lim_{n\to\infty} \mathbb{E}(X_n^2) = \lim_{n\to\infty} \sum_{i=1}^{n}\sum_{j=1}^{n}(x_{i+1,n} - x_{i,n})(x_{j+1,n} - x_{j,n})\varphi(x_{i,n})\varphi(x_{j,n})R(x_{i,n}, x_{j,n})$$
$$= \int_0^1 \int_0^1 \varphi(t)\varphi(t)R(s,t)\ ds\ dt.$$

This concludes the proof of Lemma 2.2.1. □

Proof of Proposition 2.3.1.

It is known that (see, for instance Su and Cambanis (1993) page 88) if $R(s,t) = \int_0^{\min(s,t)} u^\beta\ du$ then for any functions u and v and for any sampling design T_n we have,

$$u_{|T_n}{}' R_{|T_n}^{-1} v_{|T_n} = \frac{u(t_1)v(t_1)}{t_1^{\beta+1}} + \sum_{k=1}^{n-1} \frac{(u(t_{k+1}) - u(t_k))(v(t_{k+1}) - v(t_k))}{t_{k+1}^{\beta+1} - t_k^{\beta+1}}.$$

Replacing $u = f_{x,h}$ and $v = \overline{Y}$ we have,

$$\hat{g}_n^{pro}(x) = \frac{f_{x,h}(t_1)\overline{Y}(t_1)}{t_1^{\beta+1}} + \sum_{i=1}^{n-1} \frac{(f_{x,h}(t_{i+1}) - f_{x,h}(t_i))(\overline{Y}(t_{i+1}) - \overline{Y}(t_i))}{t_{i+1}^{\beta+1} - t_i^{\beta+1}}.$$

Recall that $R(s,t) = \frac{1}{\beta+1}\min(s,t)^{\beta+1}$ and,

$$f_{x,h}(t_i) = \int_0^1 R(s,t_i)\varphi_{x,h}(s)\ ds = \frac{1}{\beta+1}\Big(\int_0^{t_i} s^{\beta+1}\ \varphi_{x,h}(s)\ ds + t_i^{\beta+1} \int_{t_i}^1 \varphi_{x,h}(s)\ ds\Big).$$

Thus,

$$f_{x,h}(t_{i+1}) - f_{x,h}(t_i) = \frac{1}{\beta+1}\Big(\int_0^{t_{i+1}} s^{\beta+1}\varphi_{x,h}(s)\ ds + t_{i+1}^{\beta+1}\int_{t_{i+1}}^1 \varphi_{x,h}(s)\ ds$$
$$- \int_0^{t_i} s^{\beta+1}\varphi_{x,h}(s)\ ds - t_i^{\beta+1}\int_{t_i}^1 \varphi_{x,h}(s)\ ds + t_{i+1}^{\beta+1}\int_{t_i}^1 \varphi_{x,h}(s)\ ds - t_{i+1}^{\beta+1}\int_{t_i}^1 \varphi_{x,h}(s)\ ds\Big)$$
$$= \frac{1}{\beta+1}\Big(\int_{t_i}^{t_{i+1}} (s^{\beta+1} - t_{i+1}^{\beta+1})\varphi_{x,h}(s)\ ds + (t_{i+1}^{\beta+1} - t_i^{\beta+1})\int_{t_i}^1 \varphi_{x,h}(s)\ ds\Big).$$

Thus,

$$\hat{g}_n^{pro}(x) = \frac{f_{x,h}(t_1)\overline{Y}(t_1)}{t_1^{\beta+1}} + \frac{1}{\beta+1}\Big(\sum_{i=1}^{n-1}(\overline{Y}(t_{i+1}) - \overline{Y}(t_i))\int_{t_i}^1 \varphi_{x,h}(s)\ ds$$
$$+ \sum_{i=1}^{n-1} \frac{\overline{Y}(t_{i+1}) - \overline{Y}(t_i)}{t_{i+1}^{\beta+1} - t_i^{\beta+1}} \int_{t_i}^{t_{i+1}} (s^{\beta+1} - t_{i+1}^{\beta+1})\varphi_{x,h}(s)\ ds\Big)$$
$$= \frac{f_{x,h}(t_1)\overline{Y}(t_1)}{t_1^{\beta+1}} + \frac{1}{\beta+1}\Big(\sum_{i=2}^{n-1}\overline{Y}(t_i)\int_{t_{i-1}}^{t_i} \varphi_{x,h}(s)\ ds - \overline{Y}(t_1)\int_{t_1}^1 \varphi_{x,h}(s)\ ds$$
$$+ \overline{Y}(t_n)\int_{t_{n-1}}^1 \varphi_{x,h}(s)\ ds + \sum_{i=1}^{n-1} \frac{\overline{Y}(t_{i+1}) - \overline{Y}(t_i)}{t_{i+1}^{\beta+1} - t_i^{\beta+1}} \int_{t_i}^{t_{i+1}} (s^{\beta+1} - t_{i+1}^{\beta+1})\varphi_{x,h}(s)\ ds\Big).$$

Letting $t_0 = \overline{Y}(t_0) = 0$ we have,

$$\begin{aligned}\frac{f_{x,h}(t_1)\overline{Y}(t_1)}{t_1^{\beta+1}} &= \frac{1}{\beta+1}\Big(\frac{\overline{Y}(t_1)}{t_1^{\beta+1}}\int_0^{t_1} s^{\beta+1}\varphi_{x,h}(s)\ ds + \overline{Y}(t_1)\int_{t_1}^1 \varphi_{x,h}(s)\ ds\Big)\\ &= \frac{1}{\beta+1}\Big(\frac{\overline{Y}(t_1)-\overline{Y}(t_0)}{t_1^{\beta+1}-t_0^{\beta+1}}\int_0^{t_1}(s^{\beta+1}-t_1^{\beta+1})\varphi_{x,h}(s)\ ds\\ &\quad + \overline{Y}(t_1)\int_0^{t_1}\varphi_{x,h}(s)\ ds + \overline{Y}(t_1)\int_{t_1}^1\varphi_{x,h}(s)\ ds\Big).\end{aligned}$$

Finally,

$$\begin{aligned}\hat{g}_n^{pro}(x) &= \frac{1}{\beta+1}\Big(\sum_{i=2}^{n-1}\overline{Y}(t_i)\int_{t_{i-1}}^{t_i}\varphi_{x,h}(s)\ ds - \overline{Y}(t_1)\int_{t_1}^1\varphi_{x,h}(s)\ ds + \overline{Y}(t_n)\int_{t_{n-1}}^1\varphi_{x,h}(s)\ ds\\ &+ \overline{Y}(t_n)\int_{t_{n-1}}^{t_n}\varphi_{x,h}(s)\ ds - \overline{Y}(t_n)\int_{t_{n-1}}^{t_n}\varphi_{x,h}(s)\ ds\\ &+ \sum_{i=1}^{n-1}\frac{\overline{Y}(t_{i+1})-\overline{Y}(t_i)}{t_{i+1}^{\beta+1}-t_i^{\beta+1}}\int_{t_i}^{t_{i+1}}(s^{\beta+1}-t_{i+1}^{\beta+1})\varphi_{x,h}(s)\ ds\\ &+ \frac{\overline{Y}(t_1)-\overline{Y}(t_0)}{t_1^{\beta+1}-t_0^{\beta+1}}\int_0^{t_1}(s^{\beta+1}-t_1^{\beta+1})\varphi_{x,h}(s)\ ds + \overline{Y}(t_1)\int_0^{t_1}\varphi_{x,h}(s)\ ds + \overline{Y}(t_1)\int_{t_1}^1\varphi_{x,h}(s)\ ds\Big)\\ &= \frac{1}{\beta+1}\Big(\sum_{i=1}^{n+1}\overline{Y}(t_i)\int_{t_{i-1}}^{t_i}\varphi_{x,h}(s)\ ds + \sum_{i=0}^{n-1}\frac{\overline{Y}(t_{i+1})-\overline{Y}(t_i)}{t_{i+1}^{\beta+1}-t_i^{\beta+1}}\int_{t_i}^{t_{i+1}}(s^{\beta+1}-t_{i+1}^{\beta+1})\varphi_{x,h}(s)\ ds\Big),\end{aligned}$$

where $t_{n+1} = 1$ and $\overline{Y}(t_{n+1}) = \overline{Y}(t_n)$. This concludes the proof of Proposition 2.3.1. □

Proof of Proposition 2.3.2.

It is known (see Anderson (1960) page 210) that for every functions u and v and for every design T_n we have,

$$\begin{aligned}u'_{|T_n}R^{-1}_{|T_n}v_{|T_n} &= \frac{u(t_1)v(t_1)}{1-e^{-2(t_2-t_1)}} + \frac{u(t_n)v(t_n)}{1-e^{-2(t_n-t_{n-1})}} + \sum_{i=2}^{n-1}\frac{u(t_i)v(t_i)(1-e^{-2(t_{i+1}-t_{i-1})})}{(1-e^{-2(t_{i+1}-t_i)})(1-e^{-2(t_i-t_{i-1})})}\\ &\quad - \sum_{i=1}^{n-1}\frac{u(t_i)v(t_{i+1})+u(t_{i+1})v(t_i)}{1-e^{-2(t_{i+1}-t_i)}}e^{-(t_{i+1}-t_i)}.\end{aligned}$$

Taking $u = f_{x,h}$ and $v = \overline{Y}$ we get,

$$\begin{aligned}\hat{g}_n^{pro}(x) &= \frac{f_{x,h}(t_1)\overline{Y}(t_1)}{1-e^{-2(t_2-t_1)}} + \frac{f_{x,h}(t_n)\overline{Y}(t_n)}{1-e^{-2(t_n-t_{n-1})}} + \sum_{i=2}^{n-1}\frac{f_{x,h}(t_i)\overline{Y}(t_i)(1-e^{-2(t_{i+1}-t_{i-1})})}{(1-e^{-2(t_{i+1}-t_i)})(1-e^{-2(t_i-t_{i-1})})}\\ &\quad - \sum_{i=1}^{n-1}\frac{f_{x,h}(t_i)\overline{Y}(t_{i+1})+f_{x,h}(t_{i+1})\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}}e^{-(t_{i+1}-t_i)}\\ &\triangleq \frac{f_{x,h}(t_1)\overline{Y}(t_1)}{1-e^{-2(t_2-t_1)}} + \frac{f_{x,h}(t_n)\overline{Y}(t_n)}{1-e^{-2(t_n-t_{n-1})}} + A. \qquad (2.24)\end{aligned}$$

Note that,

$$1-e^{-2(t_{i+1}-t_{i-1})}=(1-e^{-2(t_{i+1}-t_i)})+(1-e^{-2(t_i-t_{i-1})})-(1-e^{-2(t_i-t_{i-1})})(1-e^{-2(t_{i+1}-t_i)}).$$

Thus,

$$\begin{aligned}
A &= \sum_{i=2}^{n-1}\frac{f_{x,h}(t_i)\overline{Y}(t_i)}{1-e^{-2(t_i-t_{i-1})}}+\sum_{i=2}^{n-1}\frac{f_{x,h}(t_i)\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}}-\sum_{i=2}^{n-1}f_{x,h}(t_i)\overline{Y}(t_i)\\
&\quad -\sum_{i=2}^{n}\frac{f_{x,h}(t_{i-1})\overline{Y}(t_i)}{1-e^{-2(t_i-t_{i-1})}}e^{-(t_i-t_{i-1})}-\sum_{i=1}^{n-1}\frac{f_{x,h}(t_{i+1})\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}}e^{-(t_{i+1}-t_i)}\\
&= \sum_{i=2}^{n-1}\frac{\overline{Y}(t_i)}{1-e^{-2(t_i-t_{i-1})}}\Big(f_{x,h}(t_i)-f_{x,h}(t_{i-1})e^{-(t_i-t_{i-1})}\Big)-\frac{f_{x,h}(t_{n-1})\overline{Y}(t_n)}{1-e^{-2(t_n-t_{n-1})}}e^{-(t_n-t_{n-1})}\\
&\quad +\sum_{i=2}^{n-1}\frac{\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}}\Big(f_{x,h}(t_i)-f_{x,h}(t_{i+1})e^{-(t_{i+1}-t_i)}\Big)-\frac{f_{x,h}(t_2)\overline{Y}(t_1)}{1-e^{-2(t_2-t_1)}}e^{-(t_2-t_1)}\\
&\quad -\sum_{i=2}^{n-1}f_{x,h}(t_i)\overline{Y}(t_i)
\end{aligned}\tag{2.25}$$

Simple calculations yield,

$$\begin{aligned}
&f_{x,h}(t_i)-f_{x,h}(t_{i-1})e^{-(t_i-t_{i-1})}=\\
&e^{-t_i}\int_{t_{i-1}}^{t_i}e^s\varphi_{x,h}(s)\,ds-e^{t_i}\int_{t_{i-1}}^{t_i}e^{-s}\varphi_{x,h}(s)\,ds+e^{t_i}(1-e^{-2(t_i-t_{i-1})})\int_{t_{i-1}}^{1}e^{-s}\varphi_{x,h}(s)\,ds.
\end{aligned}\tag{2.26}$$

In the same way we have,

$$\begin{aligned}
&f_{x,h}(t_i)-f_{x,h}(t_{i+1})e^{-(t_{i+1}-t_i)}=\\
&e^{t_i}\int_{t_i}^{t_{i+1}}e^{-s}\varphi_{x,h}(s)\,ds-e^{-t_i}\int_{t_i}^{t_{i+1}}e^{s}\varphi_{x,h}(s)\,ds+e^{-t_i}(1-e^{-2(t_{i+1}-t_i)})\int_{0}^{t_{i+1}}e^{s}\varphi_{x,h}(s)\,ds.
\end{aligned}\tag{2.27}$$

It is easy to verify that,

$$\sum_{i=2}^{n-1}f_{x,h}(t_i)\overline{Y}(t_i)=\sum_{i=2}^{n-1}\overline{Y}(t_i)e^{-t_i}\int_0^{t_i}e^s\varphi_{x,h}(s)\,ds+\sum_{i=2}^{n-1}\overline{Y}(t_i)e^{t_i}\int_{t_i}^{1}e^{-s}\varphi_{x,h}(s)\,ds.\tag{2.28}$$

We obtain using Equations (2.25), (2.26), (2.27) and (2.28),

$$
\begin{aligned}
A = & \sum_{i=2}^{n-1} \overline{Y}(t_i) e^{t_i} \int_{t_{i-1}}^{1} e^{-s} \varphi_{x,h}(s)\, ds + \sum_{i=2}^{n-1} \frac{\overline{Y}(t_i) e^{-t_i}}{1-e^{-2(t_i - t_{i-1})}} \int_{t_{i-1}}^{t_i} e^{s} \varphi_{x,h}(s)\, ds \\
& - \sum_{i=2}^{n-1} \frac{\overline{Y}(t_i) e^{t_i}}{1-e^{-2(t_i - t_{i-1})}} \int_{t_{i-1}}^{t_i} e^{-s} \varphi_{x,h}(s)\, ds + \sum_{i=2}^{n-1} \overline{Y}(t_i) e^{-t_i} \int_{0}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds \\
& + \sum_{i=2}^{n-1} \frac{\overline{Y}(t_i) e^{t_i}}{1-e^{-2(t_{i+1} - t_i)}} \int_{t_i}^{t_{i+1}} e^{-s} \varphi_{x,h}(s)\, ds - \sum_{i=2}^{n-1} \frac{\overline{Y}(t_i) e^{-t_i}}{1-e^{-2(t_{i+1} - t_i)}} \int_{t_i}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds \\
& - \frac{f_{x,h}(t_2) \overline{Y}(t_1)}{1-e^{-2(t_2 - t_1)}} e^{-(t_2 - t_1)} - \frac{f_{x,h}(t_{n-1}) \overline{Y}(t_n)}{1-e^{-2(t_n - t_{n-1})}} e^{-(t_n - t_{n-1})} \\
& - \sum_{i=2}^{n-1} \overline{Y}(t_i) e^{-t_i} \int_{0}^{t_i} e^{s} \varphi_{x,h}(s)\, ds - \sum_{i=2}^{n-1} \overline{Y}(t_i) e^{t_i} \int_{t_i}^{1} e^{-s} \varphi_{x,h}(s)\, ds.
\end{aligned}
$$

Replacing this expression of A in (2.24) gives,

$$
\begin{aligned}
\hat{g}_n^{pro}(x) = & \sum_{i=2}^{n-1} \overline{Y}(t_i) \int_{t_{i-1}}^{t_{i+1}} e^{|t_i - s|} \varphi_{x,h}(s)\, ds + \sum_{i=2}^{n-2} \frac{\overline{Y}(t_{i+1}) e^{-t_{i+1}} - \overline{Y}(t_i) e^{-t_i}}{1-e^{-2(t_{i+1} - t_i)}} \int_{t_i}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds \\
& - \sum_{i=2}^{n-2} \frac{\overline{Y}(t_{i+1}) e^{t_{i+1}} - \overline{Y}(t_i) e^{t_i}}{1-e^{-2(t_{i+1} - t_i)}} \int_{t_i}^{t_{i+1}} e^{-s} \varphi_{x,h}(s)\, ds + \frac{\overline{Y}(t_2) e^{-t_2}}{1-e^{-2(t_2 - t_1)}} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds \\
& - \frac{\overline{Y}(t_{n-1}) e^{-t_{n-1}}}{1-e^{-2(t_n - t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{s} \varphi_{x,h}(s)\, ds - \frac{\overline{Y}(t_2) e^{t_2}}{1-e^{-2(t_2 - t_1)}} \int_{t_1}^{t_2} e^{-s} \varphi_{x,h}(s)\, ds \\
& + \frac{\overline{Y}(t_{n-1}) e^{t_{n-1}}}{1-e^{-2(t_n - t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{-s} \varphi_{x,h}(s)\, ds + \frac{f_{x,h}(t_1) \overline{Y}(t_1)}{1-e^{-2(t_2 - t_1)}} - \frac{f_{x,h}(t_2) \overline{Y}(t_1)}{1-e^{-2(t_2 - t_1)}} e^{-(t_2 - t_1)} \\
& + \frac{f_{x,h}(t_n) \overline{Y}(t_n)}{1-e^{-2(t_n - t_{n-1})}} - \frac{f_{x,h}(t_{n-1}) \overline{Y}(t_n)}{1-e^{-2(t_n - t_{n-1})}} e^{-(t_n - t_{n-1})}.
\end{aligned}
\tag{2.29}
$$

Note that Equation (2.27) yields,

$$
\begin{aligned}
\frac{\overline{Y}(t_1)}{1-e^{-2(t_2 - t_1)}} \left(f_{x,h}(t_1) - f_{x,h}(t_2) e^{-(t_2 - t_1)} \right) = & \frac{\overline{Y}(t_1) e^{t_1}}{1-e^{-2(t_2 - t_1)}} \int_{t_1}^{t_2} e^{-s} \varphi_{x,h}(s)\, ds \\
& - \frac{\overline{Y}(t_1) e^{-t_1}}{1-e^{-2(t_2 - t_1)}} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds + \overline{Y}(t_1) e^{-t_1} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds.
\end{aligned}
\tag{2.30}
$$

Similarly, Equation (2.26) yields,

$$
\begin{aligned}
\frac{\overline{Y}(t_n)}{1-e^{-2(t_n - t_{n-1})}} \left(f_{x,h}(t_n) - f_{x,h}(t_{n-1}) e^{-(t_n - t_{n-1})} \right) = & \frac{\overline{Y}(t_n) e^{-t_n}}{1-e^{-2(t_n - t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{s} \varphi_{x,h}(s)\, ds \\
- \frac{\overline{Y}(t_n) e^{t_n}}{1-e^{-2(t_n - t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{-s} \varphi_{x,h}(s)\, ds + & \overline{Y}(t_n) e^{t_n} \int_{t_{n-1}}^{1} e^{-s} \varphi_{x,h}(s)\, ds.
\end{aligned}
\tag{2.31}
$$

We obtain using (2.30) and (2.31) in (2.29),

$$\begin{aligned}
\hat{g}_n^{pro}(x) &= \sum_{i=2}^{n-1} \overline{Y}(t_i) \int_{t_{i-1}}^{t_{i+1}} e^{|t_i-s|} \varphi_{x,h}(s)\, ds + \sum_{i=2}^{n-2} \frac{\overline{Y}(t_{i+1})e^{-t_{i+1}} - \overline{Y}(t_i)e^{-t_i}}{1-e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds \\
&- \sum_{i=2}^{n-2} \frac{\overline{Y}(t_{i+1})e^{t_{i+1}} - \overline{Y}(t_i)e^{t_i}}{1-e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{-s} \varphi_{x,h}(s)\, ds + \frac{\overline{Y}(t_2)e^{-t_2}}{1-e^{-2(t_2-t_1)}} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds \\
&- \frac{\overline{Y}(t_{n-1})e^{-t_{n-1}}}{1-e^{-2(t_n-t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{s} \varphi_{x,h}(s)\, ds - \frac{\overline{Y}(t_2)e^{t_2}}{1-e^{-2(t_2-t_1)}} \int_{t_1}^{t_2} e^{-s} \varphi_{x,h}(s)\, ds \\
&+ \frac{\overline{Y}(t_{n-1})e^{t_{n-1}}}{1-e^{-2(t_n-t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{-s} \varphi_{x,h}(s)\, ds + \frac{\overline{Y}(t_1)e^{t_1}}{1-e^{-2(t_2-t_1)}} \int_{t_1}^{t_2} e^{-s} \varphi_{x,h}(s)\, ds \\
&- \frac{\overline{Y}(t_1)e^{-t_1}}{1-e^{-2(t_2-t_1)}} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds + \overline{Y}(t_1)e^{-t_1} \int_{t_1}^{t_2} e^{s} \varphi_{x,h}(s)\, ds \\
&+ \frac{\overline{Y}(t_n)e^{-t_n}}{1-e^{-2(t_n-t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{s} \varphi_{x,h}(s)\, ds - \frac{\overline{Y}(t_n)e^{t_n}}{1-e^{-2(t_n-t_{n-1})}} \int_{t_{n-1}}^{t_n} e^{-s} \varphi_{x,h}(s)\, ds \\
&+ \overline{Y}(t_n)e^{t_n} \int_{t_{n-1}}^{1} e^{-s} \varphi_{x,h}(s)\, ds
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=2}^{n-1} \overline{Y}(t_i) \int_{t_{i-1}}^{t_{i+1}} e^{|s-t_i|} \varphi_{x,h}(s)\, ds + \overline{Y}(t_1) \int_0^{t_2} e^{s-t_1} \varphi_{x,h}(s)\, ds + \overline{Y}(t_n) \int_{t_{n-1}}^{1} e^{t_n-s} \varphi_{x,h}(s)\, ds \\
&- \sum_{i=1}^{n-1} \frac{e^{t_{i+1}}\overline{Y}(t_{i+1}) - e^{t_i}\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{-s} \varphi_{x,h}(s)\, ds \\
&+ \sum_{i=1}^{n-1} \frac{e^{-t_{i+1}}\overline{Y}(t_{i+1}) - e^{-t_i}\overline{Y}(t_i)}{1-e^{-2(t_{i+1}-t_i)}} \int_{t_i}^{t_{i+1}} e^{s} \varphi_{x,h}(s)\, ds.
\end{aligned}$$

This concludes the proof of Proposition 2.3.2. □

Proof of Lemma 2.3.1.

Let $(u,v) \in [-1,1]^2$. We first consider the triangle $\{-1 < u < v < 1\}$ which is further split into smaller triangles:

$$D_1 = \{0 < u < v < 1\},\ D_2 = \{-1 < u < 0 < v < 1\} \text{ and } D_3 = \{-1 < u < v < 0\}.$$

Let $b \in]0,1[$. For $(u,v) \in D_1$, using Assumption (A), Taylor expansion of R around (x,x) gives,

$$\begin{aligned}
R(x+bu, x+bv) &= R(x, x+bv) + buR^{(1,0)}(x, x+bv) + \frac{1}{2}b^2u^2R^{(2,0)}(\varepsilon_x, x+bv) \\
&= R(x,x) + bvR^{(0,1)}(x,\eta_x) + buR^{(1,0)}(x, x+bv) + \frac{1}{2}b^2u^2R^{(2,0)}(\varepsilon_x, x+bv),
\end{aligned}$$

where $x < \varepsilon_x < x+bu < x+bv$ and $x < \eta_x < x+bv$. Thus,

$$R(x+bu, x+bv) = R(x,x) + bvR^{(0,1)}(x,x^+) + buR^{(0,1)}(x,x^-) + o(b).$$

Now, for $(u,v) \in D_2$ we obtain in the same way,

$$
\begin{aligned}
R(x+bu, x+bv) &= R(x, x+bv) + buR^{(1,0)}(x, x+bv) + \frac{1}{2}b^2u^2R^{(2,0)}(\varepsilon_x, x+bv) \\
&= R(x,x) + bvR^{(0,1)}(x, \eta_x) + buR^{(1,0)}(x, x+bv) + \frac{1}{2}b^2u^2R^{(2,0)}(\varepsilon_x, x+bv),
\end{aligned}
$$

where $x+bu < \varepsilon_x < x < x+bv$ and $x < \eta_x < x+bv$. Thus,

$$R(x+bu, x+bv) = R(x,x) + bvR^{(0,1)}(x, x^+) + buR^{(0,1)}(x, x^-) + o(b).$$

Finally, for $(u,v) \in D_3$ we get,

$$
\begin{aligned}
R(x+bu, x+bv) &= R(x+bu, x) + bvR^{(0,1)}(x+bu, x) + \frac{1}{2}b^2v^2R^{(0,2)}(x+bu, \eta_x) \\
&= R(x,x) + ubR^{(1,0)}(\varepsilon_x, x) + bvR^{(0,1)}(x+bu, \eta_x) + \frac{1}{2}b^2v^2R^{(0,2)}(x+bu, \eta_x),
\end{aligned}
$$

where $x+hu < x+bv < \eta_x < x$ and $x+bu < \varepsilon_x < x$. Thus,

$$R(x+bu, x+bv) = R(x,x) + bvR^{(0,1)}(x, x^+) + buR^{(0,1)}(x, x^-) + o(b).$$

Hence for $v > u$ we have,

$$
\begin{aligned}
R(x+bu, x+bv) = R(x,x) &+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) + R^{(0,1)}(x,x^-)\big)(u+v) \\
&+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) - R^{(0,1)}(x,x^-)\big)(v-u) + o(b).
\end{aligned}
$$

Similarly, we obtain for the triangular $\{1 > u > v > -1\}$,

$$
\begin{aligned}
R(x+bu, x+bv) = R(x,x) &+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) + R^{(0,1)}(x,x^-)\big)(u+v) \\
&+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) - R^{(0,1)}(x,x^-)\big)(u-v).
\end{aligned}
$$

Thus, for $(u,v) \in [-1,1]^2$ we have,

$$
\begin{aligned}
R(x+bu, x+bv) = R(x,x) &+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) + R^{(0,1)}(x,x^-)\big)(u+v) \\
&+ \frac{B}{2}\big(R^{(0,1)}(x,x^+) - R^{(0,1)}(x,x^-)\big)|u-v|. \qquad (2.32)
\end{aligned}
$$

Consider now a function g, bounded and integrable on $[-1,1]$. The Dominated Convergence Theorem yields that $R(.,t) \times g$ is an integrable function for every $t \in [-1,1]$. Using (2.32) and putting,

$$\gamma(x) = \frac{1}{2}\big(R^{(0,1)}(x,x^+) + R^{(0,1)}(x,x^-)\big),$$

we obtain,

$$
\begin{aligned}
&\iint_{[-1,1]^2} R(x+bu, x+bv)g(u)g(v)dudv = R(x,x)\left(\int_{-1}^{1} g(u)du\right)^2 \\
&+ 2\gamma(x)b\int_{-1}^{1} g(u)du \int_{-1}^{1} vg(v)dv - \frac{B}{2}\alpha(x)\iint_{[-1,1]^2} g(u)g(v)|u-v|dudv + o(b).
\end{aligned}
\qquad (2.33)
$$

The left side of (2.33) is non-negative since the autocovariance function R is a non-negative definite function. Taking $g(u) = u1_{[-1,1]}(u)$ we obtain,

$$\int_{-1}^{1} g(u)du = 0 \quad \text{and} \iint_{[-1,1]^2} uv|u-v|dudv = -\frac{8}{15}.$$

Thus,

$$\frac{4}{15}\alpha(x) + o(b) \geq 0.$$

Taking b small enough concludes the proof of Lemma 2.3.1. □

Proof of Lemma 2.4.2.

The great lines of this proof are based on the work of Sacks and Ylvisaker (1966) (c.f. Lemma 3.2 there). Let $x, h \in]0,1[$ and put $g_n = P_{T_n} f_{x,h}$, it is shown by (2.102) in the Appendix that,

$$g_n(t_i) = \sum_{j=1}^{n} m_{x,h}(t_j) R(t_j, t_i) \quad \text{for all } i = 1, \cdots, n.$$

On the one hand, Assumption (A) yields that g_n is twice differentiable on $[0,1]$ except on T_n, but it has left and right derivatives. Thus, for every $i = 1, \ldots, n$ we have,

$$g_n'(t_i^-) = \sum_{j=1}^{n} m_{x,h}(t_j) R^{(0,1)}(t_j, t_i^-) \quad \text{and} \quad g_n'(t_i^+) = \sum_{j=1}^{n} m_{x,h}(t_j) R^{(0,1)}(t_j, t_i^+).$$

Since for $j \neq i$, $R^{(0,1)}(t_j, t_i^-) = R^{(0,1)}(t_j, t_i^+)$ then Assumption (B) yields,

$$g_n'(t_i^-) - g_n'(t_i^+) = \alpha(t_i) m_{x,h}(t_i). \tag{2.34}$$

On the other hand, Assumption (A) yields that $f_{x,h}$ (as defined by (2.2)) is twice differentiable on $]0,1[$, thus for $i = 1, \ldots, n-1$, Taylor expansion of $f_{x,h} - g_n$ around t_i gives,

$$f_{x,h}(t_{i+1}) - g_n(t_{i+1}) = (f_{x,h}(t_i) - g_n(t_i)) + d_i(f_{x,h}'(t_i) - g_n'(t_i^+)) + \frac{1}{2} d_i^2 (f_{x,h}''(\sigma_i) - g_n''(\sigma_i)),$$

where $d_i = t_{i+1} - t_i$ and $\sigma_i \in]t_i, t_{i+1}[$. Recall that, for all $i = 1, \ldots, n$, $f_{x,h}(t_i) = g_n(t_i)$ (see Equation (2.6)). Thus,

$$f_{x,h}'(t_i) - g_n'(t_i^+) = -\frac{1}{2} d_i (f_{x,h}''(\sigma_i) - g_n''(\sigma_i)), \tag{2.35}$$

Similarly, for $i = 2, \ldots, n$, we have,

$$f_{x,h}'(t_i) - g_n'(t_i^-) = \frac{1}{2} d_{i-1} (f_{x,h}''(\theta_i) - g_n''(\theta_i)), \tag{2.36}$$

for some $\theta_i \in]t_{i-1}, t_i[$. We obtain subtracting (2.36) from (2.36) and using (2.34) for $i = 2, \ldots, n-1$,

$$\alpha(t_i) m_{x,h}(t_i) = -\frac{1}{2} d_i (f_{x,h}''(\sigma_i) - g_n''(\sigma_i)) - \frac{1}{2} d_{i-1} (f_{x,h}''(\theta_i) - g_n''(\theta_i)). \tag{2.37}$$

We shall now control the last expression. On the one hand we have,

$$f'_{x,h}(t) = \int_0^t R^{(0,1)}(s,t^+)\varphi_{x,h}(s)\;ds + \int_t^1 R^{(0,1)}(s,t^-)\varphi_{x,h}(s)\;ds, \tag{2.38}$$

and,

$$\begin{aligned} f''_{x,h}(t) = {} & (R^{(0,1)}(t,t^+) - R^{(0,1)}(t,t^-))\varphi_{x,h}(t) + \int_0^1 R^{(0,2)}(s,t^+)\varphi_{x,h}(s)\;ds \\ & - \alpha(t)\varphi_{x,h}(t) + \int_0^1 R^{(0,2)}(s,t^+)\varphi_{x,h}(s)\;ds. \end{aligned} \tag{2.39}$$

On the other hand we know, using (F3) in the Appendix, that every function in the RKHS(R), noted by $\mathcal{F}(\varepsilon)$, is continuous, hence Assumption (C) implies that $R^{(0,2)}(\cdot,t^+)$ is a continuous function on $[0,1]$ for every fixed $t \in [0,1]$. Thus,

$$R^{(0,2)}(t,t^+) = \lim_{s\downarrow t} R^{(0,2)}(s,t^+) = \lim_{s\downarrow t} R^{(0,2)}(s,t^-) = R^{(0,2)}(t,t^-),$$

from which we get that $R^{(0,2)}(t,t)$ exists. Hence for $i = 1,\ldots,n$ we have,

$$g''_n(t_i^-) = g''_n(t_i^+) = \sum_{j=1}^n m_{x,h}(t_j) R^{(0,2)}(t_j,t_i). \tag{2.40}$$

In addition, it is shown by (F4) in the Appendix that for every $t \in [0,1]$,

$$f''_{x,h}(t) - g''_n(t) = -\alpha(t)\varphi_{x,h}(t) + \langle R^{(0,2)}(\cdot,t), f_{x,h} - g_n\rangle, \tag{2.41}$$

where $\langle\cdot,\cdot\rangle$ is the inner product on $\mathcal{F}(\varepsilon)$. Injecting (2.41) in (2.37) we obtain,

$$\begin{aligned} \alpha(t_i)m_{x,h}(t_i) = {} & \frac{1}{2}d_i\alpha(\sigma_i)\varphi_{x,h}(\sigma_i) + \frac{1}{2}d_{i-1}\alpha(\theta_i)\varphi_{x,h}(\theta_i) - \frac{1}{2}d_i\langle R^{(0,2)}(\cdot,\sigma_i), f_{x,h} - g_n\rangle \\ & - \frac{1}{2}d_{i-1}\langle R^{(0,2)}(\cdot,\theta_i), f_{x,h} - g_n\rangle. \end{aligned}$$

Using Assumption (B) we obtain for $i = 2,\ldots,n-1$,

$$\begin{aligned} m_{x,h}(t_i) = {} & \frac{1}{2}(d_i + d_{i-1})\varphi_{x,h}(t_i) \; + \frac{1}{2\alpha(t_i)}d_i\big(\alpha(\sigma_i)\varphi_{x,h}(\sigma_i) - \alpha(t_i)\varphi_{x,h}(t_i)\big) \\ & + \frac{1}{2\alpha(t_i)}d_{i-1}\big(\alpha(\theta_i)\varphi_{x,h}(\theta_i) - \alpha(t_i)\varphi_{x,h}(t_i)\big) - \frac{1}{2\alpha(t_i)}d_i\langle R^{(0,2)}(\cdot,\sigma_i), f_{x,h} - g_n\rangle \\ & - \frac{1}{2\alpha(t_i)}d_{i-1}\langle R^{(0,2)}(\cdot,\theta_i), f_{x,h} - g_n\rangle \\ \triangleq {} & \frac{1}{2}(d_i + d_{i-1})\varphi_{x,h}(t_i) \; + A_i^{(1)} + A_i^{(2)} - A_i^{(3)} - A_i^{(4)}, \end{aligned} \tag{2.42}$$

Using the Cauchy-Schwartz inequality, Assumption (C) and Equation (2.55) (in the proof of Proposition 2.4.3 below) we obtain,

$$|A_i^{(3)} + A_i^{(4)}| \le \sup_{0\le t\le 1} \frac{1}{2\alpha(t)}||R^{(0,2)}(.,t)||\frac{\sqrt{C}}{\sqrt{h}} \sup_{0\le j\le n} d_j^2 \triangleq \beta_{n,h}, \tag{2.43}$$

where C is a positive constant defined in Proposition 2.4.3 below.
Recall that $\varphi_{x,h}$ is of support $[x-h,x+h]$, thus for t_i such that $[t_{i-1},t_{i+1}]\cap]x-h,x+h[=\emptyset$, $\varphi_{x,h}(t)=0$ so that $A_i^{(1)}=0$ and $A_i^{(2)}=0$. For t_i such that $[t_{i-1},t_{i+1}]\cap]x-h,x+h[\neq\emptyset$, let,

$$\alpha_{n,h}=\sup_{0\le i\le n}\ \sup_{t_i\le s,t\le t_{i+1}}\frac{1}{2\alpha(t)}d_i|\alpha(s)\varphi_{x,h}(s)-\alpha(t)\varphi_{x,h}(t)|. \tag{2.44}$$

We obtain using (2.43) and (2.44) together with (2.42) for $i=2,\dots,n-1$,

$$m_{x,h}(t_i)=\begin{cases}\frac{1}{2}\varphi_{x,h}(t_i)(t_{i+1}-t_{i-1})\ +\ O\big(\alpha_{n,h}+\beta_{n,h}\big) & if\ [t_{i-1},t_{i+1}]\cap]x-h,x+h[\neq\emptyset\\ O\big(\beta_{n,h}\big) & otherwise.\end{cases}$$

After having obtained $m_{x,h}(t_i)$ for $i=2,\dots,n-1$, we are now able to obtain $m_{x,h}(t_1)$ and $m_{x,h}(t_n)$. We have for $i=1,\dots,n$,

$$R(t_1,t_i)m_{x,h}(t_1)+R(t_n,t_i)m_{x,h}(t_n)=f_{x,h}(t_i)-\sum_{j=2}^{n-1}m_{x,h}(t_j)R(t_j,t_i). \tag{2.45}$$

Recall that $N_{T_n}=\text{Card }I_{x,h}=\text{Card }\{i=1,\cdots,n\ :\ [t_{i-1},t_{i+1}]\cap]x-h,x+h[\neq\emptyset\}$ and that $t_{x,i}$ are the points of T_n for which $i\in I_{x,h}$. We have,

$$\sum_{j=2}^{n-1}m_{x,h}(t_j)R(t_j,t_i)=\sum_{j=1}^{N_{T_n}}m_{x,h}(t_{x,j})R(t_{x,j},t_i)+\sum_{j=2}^{n-1}1_{\{j\notin I_{x,h}\}}m_{x,h}(t_j)R(t_j,t_i).$$

On the one hand, we have using (2.42) (where $A_{x,j}$ stands for A_j with t_j replaced by $t_{x,j}$),

$$\begin{aligned}\sum_{j=2}^{n-1}m_{x,h}(t_j)R(t_j,t_i)&=\frac{1}{2}\sum_{j=1}^{N_{T_n}}(d_{x,j}+d_{x,j-1})\varphi_{x,h}(t_{x,j})R(t_{x,j},t_i)\\&\quad+\sum_{j=1}^{N_{T_n}}(A^1_{x,j}+A^2_{x,j}-A^3_{x,j}-A^4_{x,j})R(t_{x,j},t_i)-\sum_{j=2}^{n-1}1_{\{j\notin I_{x,h}\}}(A^3_j+A^4_j)R(t_j,t_i)\\&=\frac{1}{2}\sum_{j=1}^{N_{T_n}}(d_{x,j}+d_{x,j-1})\varphi_{x,h}(t_{x,j})R(t_{x,j},t_i)+\sum_{j=1}^{N_{T_n}}(A^1_{x,j}+A^2_{x,j})R(t_{x,j},t_i)-\sum_{j=1}^{n}(A^3_j+A^4_j)R(t_j,t_i).\end{aligned} \tag{2.46}$$

On the other hand,

$$\begin{aligned}f_{x,h}(t_i)&=\int_0^1R(s,t_i)\varphi_{x,h}(s)\,ds=\int_{x-h}^{x+h}R(s,t_i)\varphi_{x,h}(s)\,ds=\frac{1}{2}\sum_{j=1}^{N_{T_n}}\int_{t_{x,j-1}}^{t_{x,j+1}}R(s,t_i)\varphi_{x,h}(s)\,ds\\&=\frac{1}{2}\sum_{j=1}^{N_{T_n}}(d_{x,j}+d_{x,j-1})R(t_{x,j},t_i)\varphi_{x,h}(t_j)+\frac{1}{2}\sum_{j=1}^{N_{T_n}}\int_{t_{x,j-1}}^{t_{x,j+1}}(R(s,t_i)\varphi_{x,h}(s)-R(t_{x,j},t_i)\varphi_{x,h}(t_{x,j}))\,ds.\end{aligned} \tag{2.47}$$

Inserting (2.46) and (2.47) in (2.45) we obtain for $i=1,\dots,n$,

$$\begin{aligned}R(t_1,t_i)m_{x,h}(t_1)+R(t_n,t_i)m_{x,h}(t_n)&=\frac{1}{2}\sum_{j=1}^{N_{T_n}}\int_{t_{x,j-1}}^{t_{x,j+1}}(R(s,t_i)\varphi_{x,h}(s)-R(t_{x,j},t_i)\varphi_{x,h}(t_{x,j}))\,ds\\&-\sum_{j=1}^{N_{T_n}}(A^1_{x,j}+A^2_{x,j})R(t_{x,j},t_i)+\sum_{j=1}^{n}(A^3_j+A^4_j)R(t_j,t_i)\triangleq\Phi_{x,h}(t_i).\end{aligned}$$

We then obtain the following linear system,

$$\begin{cases} R(t_1,t_1)m_{x,h}(t_1) + R(t_n,t_1)m_{x,h}(t_1) = \Phi_{x,h}(t_1). \\ \\ R(t_1,t_n)m_{x,h}(t_1) + R(t_n,t_n)m_{x,h}(t_n) = \Phi_{x,h}(t_n). \end{cases} \tag{2.48}$$

Solving (2.48) for $m_{x,h}(t_1)$ and $m_{x,h}(t_n)$ we obtain,

$$m_{x,h}(t_1) = \frac{R(t_n,t_n)\Phi_{x,h}(t_1) - R(t_1,t_n)\Phi_{x,h}(t_n)}{R(t_1,t_1)R(t_n,t_n) - R(t_1,t_n)^2}.$$
$$m_{x,h}(t_n) = \frac{R(t_1,t_1)\Phi_{x,h}(t_n) - R(t_1,t_n)\Phi_{x,h}(t_1)}{R(t_1,t_1)R(t_n,t_n) - R(t_1,t_n)^2}.$$

Finally, simple calculations yield,

$$m_{x,h}(t_1) = O\big(N_{T_n}\alpha_{n,h} + n\beta_{n,h}\big) \quad \text{and} \quad m_{x,h}(t_n) = O\big(N_{T_n}\alpha_{n,h} + n\beta_{n,h}\big).$$

This completes the proof of Lemma 2.4.2. □

Proof of Proposition 2.4.1.

Recall that $N_{T_n} = \text{Card } I_{x,h} = \text{Card } \{i = 1, \cdots, n \; : \; [t_{i-1}, t_{i+1}] \cap]x-h, x+h[\neq \emptyset\}$ and denote by $t_{x,i}$ the points of T_n for which $i \in I_{x,h}$, that is $T_n \cap]x-h, x+h[= \{t_{x,2}, \cdots, t_{x,N_{T_n}-1}\}$. Since $\mathbb{E}(\overline{Y}(t_i)) = g(t_i)$ then,

$$\begin{aligned} \mathbb{E}(\hat{g}_n^{pro}(x)) &= \sum_{j=1}^{n} m_{x,h}(t_j)g(t_j) \\ &= \sum_{i=1}^{N_{T_n}} m_{x,h}(t_{x,i})g(t_{x,i}) + \sum_{j=2}^{n-1} 1_{\{i \notin I_{x,h}\}} m_{x,h}(t_j)g(t_j) + m_{x,h}(t_1)g(t_1) + m_{x,h}(t_n)g(t_n). \end{aligned}$$

Using the asymptotic approximation of $m_{x,h|T_n}$ given in Lemma 2.4.2 we obtain,

$$E(\hat{g}_n^{pro}(x)) = \frac{1}{2}\sum_{i=1}^{N_{T_n}} (t_{x,i+1} - t_{x,i-1})\varphi_{x,h}(t_{x,i})g(t_{x,i}) + O\big(N_{T_n}\alpha_{n,h} + n\beta_{n,h}\big), \tag{2.49}$$

For $x \in [0,1]$ let,

$$I_h(x) = \int_{x-h}^{x+h} \varphi_{x,h}(t)g(t)\ dt = \frac{1}{2}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i+1}} \varphi_{x,h}(t)g(t)\ dt,$$

and write,

$$\mathbb{E}(\hat{g}_n^{pro}(x)) = \mathbb{E}(\hat{g}_n^{pro}(x)) - I_h(x) + I_h(x) = \Delta_{x,h} + I_h(x) + O\big(N_{T_n}\alpha_{n,h} + n\beta_{n,h}\big), \tag{2.50}$$

where,

$$\Delta_{x,h} = \frac{1}{2}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i+1}} \Big(\varphi_{x,h}(t_{x,i})g(t_{x,i}) - \varphi_{x,h}(t)g(t)\Big)\ dt.$$

We first control $\Delta_{x,h}$. We have,

$$\Delta_{x,h} = \frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (g(t_{x,i}) - g(t))\ dt + \frac{1}{2}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i+1}} g(t)(\varphi_{x,h}(t_{x,i}) - \varphi_{x,h}(t))\ dt.$$

Since $\varphi_{x,h}$ is in C^1 and g is in C^2 then Taylor expansions of $\varphi_{x,h}$ and g give,

$$g(t) = g(t_{x,i}) + (t - t_{x,i})g'(t_{x,i}) + \frac{1}{2}(t - t_{x,i})^2 g''(\theta_{x,i}),$$

and,

$$\varphi_{x,h}(t) = \varphi_{x,h}(t_{x,i}) + (t - t_{x,i})\varphi'_{x,h}(\eta_{x,i}),$$

for some $\theta_{x,i}$ and $\eta_{x,i}$ between t and $t_{x,i}$. Thus,

$$\begin{aligned}
\Delta_{x,h} = &-\frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i})g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i})\ dt - \frac{1}{4}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} g''(\theta_{x,i})(t - t_{x,i})^2\ dt \\
&- \frac{1}{2}\sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} \varphi'_{x,h}(\eta_{x,i})(t - t_{x,i})\ dt - \frac{1}{2}\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} \varphi'_{x,h}(\eta_{x,i})(t - t_{x,i})^2\ dt \\
&- \frac{1}{4}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i+1}} g''(\theta_{x,i})\varphi'_{x,h}(\eta_{x,i})(t - t_{x,i})^3\ dt.
\end{aligned}$$

Recall that g' and g'' are both bounded and that,

$$\sup_{0\le t\le 1} |\varphi_{x,h}(t)| < \frac{c}{h} \quad \text{and} \quad \sup_{0\le t\le 1} |\varphi'_{x,h}(t)| < \frac{c'}{h^2}, \tag{2.51}$$

for appropriate positive constants c and c'. Using this we obtain,

$$\begin{aligned}
\frac{1}{4}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} g''(\theta_{x,i})(t - t_{x,i})^2\ dt &= O\Big(\frac{N_{T_n}}{h} \sup_{0\le j\le 1} d_{j,n}^3\Big) \\
\frac{1}{2}\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} \varphi'_{x,h}(\eta_{x,i})(t - t_{x,i})^2\ dt &= O\Big(\frac{N_{T_n}}{h^2} \sup_{0\le j\le 1} d_{j,n}^3\Big) \\
\frac{1}{4}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i+1}} g''(\theta_{x,i})\varphi'_{x,h}(\eta_{x,i})(t - t_{x,i})^2\ dt &= O\Big(\frac{N_{T_n}}{h^2} \sup_{0\le j\le 1} d_{j,n}^3\Big).
\end{aligned}$$

Thus,

$$\begin{aligned}
\Delta_{x,h} = &-\frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i})g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i})dt - \frac{1}{2}\sum_{i=1}^{N_{T_n}} g(t_{x,i})\varphi'_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i})\ dt \\
&- \frac{1}{2}\sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i})\Big(\varphi'_{x,h}(\eta_{x,i}) - \varphi'_{x,h}(t_{x,i})\Big)\ dt + O\Big(\frac{N_{T_n}}{h^2} \sup_{0\le j\le 1} d_{j,n}^3\Big).
\end{aligned}$$

Since $\varphi'_{x,h}$ is Lipschitz then,

$$\sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i}) \Big(\varphi'_{x,h}(\eta_{x,i}) - \varphi'_{x,h}(t_{x,i}) \Big) \ dt = O\bigg(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg).$$

Thus,

$$\Delta_{x,h} = -\frac{1}{2} \sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i}) dt - \frac{1}{2} \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \varphi'_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i+1}} (t - t_{x,i}) \ dt$$
$$+ O\bigg(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg).$$

Basic integration gives,

$$\Delta_{x,h} = -\frac{1}{4} \sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) g'(t_{x,i}) (d^2_{x,i} - d^2_{x,i-1}) - \frac{1}{4} \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \varphi'_{x,h}(t_{x,i}) (d^2_{x,i} - d^2_{x,i-1})$$
$$+ O\bigg(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg).$$

We shall show that,

$$A \triangleq \sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) g'(t_{x,i}) (d^2_{x,i} - d^2_{x,i-1}) = O\bigg(\frac{N_{T_n}}{h^2} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg),$$
$$B \triangleq \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \varphi'_{x,h}(t_{x,i}) (d^2_{x,i} - d^2_{x,i-1}) = O\bigg(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg).$$

Starting with the term A. Recall that, since φ is of support $[x-h, x+h]$ and $t_{x,1}, t_{x,N_{T_n}-1} \notin]x-h, x+h[$, then $\varphi_{x,h}(t_{x,N_{T_n}}) = \varphi_{x,h}(t_{x,1}) = 0$ thus,

$$\begin{aligned} A &= \sum_{i=2}^{N_{T_n}-1} \varphi_{x,h}(t_{x,i}) g'(t_{x,i}) d^2_{x,i} - \sum_{i=1}^{N_{T_n}-2} \varphi_{x,h}(t_{x,i+1}) g'(t_{x,i+1}) d^2_{x,i} \\ &= \sum_{i=2}^{N_{T_n}-2} \big(\varphi_{x,h}(t_{x,i}) g'(t_{x,i}) - \varphi_{x,h}(t_{x,i+1}) g'(t_{x,i+1}) \big) d^2_{x,i} + \bigg(\varphi_{x,h}(t_{x,N_{T_n}-1}) g'(t_{x,N_{T_n}-1}) d^2_{x,N_{T_n}-1} \\ &\qquad - \varphi_{x,h}(t_{x,2}) g'(t_{x,2}) d^2_{x,1} \bigg) \\ &\triangleq A_1 + A_2. \end{aligned}$$

On the one hand, Taylor expansions of $\varphi_{x,h}$ around $t_{x,N_{T_n}}$ and $t_{x,1}$ yield,

$$\varphi_{x,h}(t_{x,N_{T_n}-1}) = (t_{x,N_{T_n}-1} - t_{x,N_{T_n}}) \varphi'_{x,h}(\gamma_{x,N_{T_n}}),$$
$$\varphi_{x,h}(t_{x,2}) = (t_{x,2} - t_{x,1}) \varphi'_{x,h}(\gamma_{x,1}),$$

for some $\gamma_{x,N_{T_n}} \in]t_{x,N_{T_n}-1}, t_{x,N_{T_n}}[$ and some $\gamma_{x,1} \in]t_{x,1}, t_{x,2}[$. Using (2.51) and the fact that g' is bounded we obtain,

$$A_2 = O\bigg(\frac{1}{h^2} \sup_{0\le j\le 1} \ d^3_{j,n} \bigg).$$

On the other hand we have,

$$\begin{aligned}A_1 &= \sum_{i=2}^{N_{T_n}-2} \big(\varphi_{x,h}(t_{x,i})g'(t_{x,i}) - \varphi_{x,h}(t_{x,i+1})g'(t_{x,i+1})\big)d_{x,i}^2 \\ &= \sum_{i=2}^{N_{T_n}-2} \varphi_{x,h}(t_{x,i})\big(g'(t_{x,i}) - g'(t_{x,i+1})\big)d_{x,i}^2 + \sum_{i=2}^{N_{T_n}-2} g'(t_{x,i+1})\big(\varphi_{x,h}(t_{x,i}) - \varphi_{x,h}(t_{x,i+1})\big)d_{x,i}^2.\end{aligned}$$

Since $\varphi_{x,h}$ is in C^1 and g is in C^2 then using (2.51), we obtain,

$$A_1 = O\Big(\frac{N_{T_n}}{h^2} \sup_{0\le j\le 1} d_{j,n}^3\Big).$$

In a similar way and from Assumption (D), we obtain,

$$B = O\Big(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} d_{j,n}^3\Big).$$

Hence,

$$\Delta_{x,h} = O\Big(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} d_{j,n}^3\Big).$$

Thus using (2.50),

$$\mathbb{E}(\hat{g}_n^{pro}(x)) = I_h(x) + O(N_{T_n}\alpha_{n,h} + n\beta_{n,h}) + O\Big(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} d_{j,n}^3\Big).$$

The control of $I_h(x)$ is classical and it can bee seen from Gasser and Müller (1984) that,

$$I_h(x) = g(x) + \frac{1}{2}h^2 g''(x)\int_{-1}^{1} t^2 K(t)\ dt + o(h^2). \tag{2.52}$$

Finally,

$$\mathbb{E}(\hat{g}_n^{pro}(x)) - g(x) = \frac{1}{2}h^2 g''(x)\int_{-1}^{1} t^2 K(t)\ dt + o(h^2) + O\Big(\frac{N_{T_n}}{h^3} \sup_{0\le j\le 1} d_{j,n}^3 + N_{T_n}\alpha_{n,h} + n\beta_{n,h}\Big).$$

This concludes the proof of Proposition 2.4.1. □

Proof of Proposition 2.4.2.

Let $t_0 = 0,\ t_{n+1} = 1$ and set $\overline{Y}(t_0) = 0$ and $\overline{Y}(t_{n+1}) = \overline{Y}(t_n)$. Recall that,

$$\hat{g}_n^{pro}(x) = \sum_{i=1}^{n+1} \overline{Y}(t_i)\int_{t_{i-1}}^{t_i} \varphi_{x,h}(s)ds + \sum_{i=0}^{n} \frac{\overline{Y}(t_{i+1}) - \overline{Y}(t_i)}{t_{i+1} - t_i}\int_{t_i}^{t_{i+1}} (s - t_{i+1})\varphi_{x,h}(s)ds.$$

Since $\mathbb{E}\ (\overline{Y}(t_i)) = g(t_i)$ then,

$$\mathbb{E}\ (\hat{g}_n^{pro}(x)) = \sum_{i=1}^{n+1} g(t_i)\int_{t_{i-1}}^{t_i} \varphi_{x,h}(s)ds + \sum_{i=0}^{n} \frac{g(t_{i+1}) - g(t_i)}{t_{i+1} - t_i}\int_{t_i}^{t_{i+1}} (s - t_{i+1})\varphi_{x,h}(s)ds.$$

Recall that $N_{T_n} = \text{Card } I_{x,h} = \{i = 1, \cdots, n \; : \; [t_{i-1,n}, t_{i+1,n}] \cap]x-h, x+h[\neq \emptyset\}$ and denote by $t_{x,i}$ the points of T_n for which $i \in I_{x,h}$. Using the support of $\varphi_{x,h}$ we obtain,

$$\mathbb{E}\left(\hat{g}_n^{pro}(x)\right) = \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s)ds + \sum_{i=1}^{N_{T_n}} \frac{g(t_{x,i+1}) - g(t_{x,i})}{t_{x,i+1} - t_{x,i}} \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})\varphi_{x,h}(s)ds.$$

Let $d_{x,i} = t_{x,i+1} - t_{x,i}$. Since g is in C^2 and $\varphi_{x,h}$ is in C^1 then Taylor expansions of g around $t_{x,i}$ and of $\varphi_{x,h}$ around $t_{x,i+1}$ yield,

$$\begin{aligned} g(t_{x,i+1}) &= g(t_{x,i}) + d_{x,i}\; g'(t_{x,i}) + \frac{1}{2} d_{x,i}^2\; g''(\theta_{x,i}), \\ \varphi_{x,h}(s) &= \varphi_{x,h}(t_{x,i+1}) + (s - t_{x,i+1})\varphi'_{x,h}(s_i). \end{aligned}$$

for some $\theta_{x,i} \in]t_{x,i}, t_{x,i+1}[$ and some $s_i \in]s, t_{x,i+1}[$. Recall that, using the support of φ, $\varphi_{x,h}(t_{x,1}) = \varphi_{x,h}(t_{x,N_{T_n}}) = 0$ thus,

$$\begin{aligned} \mathbb{E}\left(\hat{g}_n^{pro}(x)\right) &= \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s)ds \;+ \sum_{i=1}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i+1}) \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})ds \\ &+ \sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})^2 \varphi'_{x,h}(s_i)ds + \frac{1}{2} \sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i+1}) g''(\theta_{x,i}) d_{x,i} \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})ds \\ &+ \frac{1}{2} \sum_{i=1}^{N_{T_n}} g''(\theta_{x,i}) d_{x,i} \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})^2 \varphi'_{x,h}(s_i)ds. \end{aligned}$$

Recall that g' and g'' are bounded, Lemma 2.4.1 yields $N_{T_n} = O(nh)$ and $d_{x,i} = O(\frac{1}{n})$ and using (2.51) we obtain,

$$\begin{aligned} &\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})^2 \varphi'_{x,h}(s_i)ds = O\left(\frac{1}{n^2 h}\right). \\ &\frac{1}{2} \sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i+1}) g''(\theta_{x,i}) d_{x,i} \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})ds = O\left(\frac{1}{n^2}\right). \\ &\frac{1}{2} \sum_{i=1}^{N_{T_n}} g''(\theta_{x,i}) d_{x,i} \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i+1})^2 \varphi'_{x,h}(s_i)ds = O\left(\frac{1}{n^3 h}\right). \end{aligned}$$

It follows that by simple integration,

$$\begin{aligned} \mathbb{E}\left(\hat{g}_n^{pro}(x)\right) &= \sum_{i=1}^{N_{T_n}} g(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s)ds - \frac{1}{2} \sum_{i=1}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i+1}) d_{x,i}^2 + O\left(\frac{1}{n^2 h}\right) \\ &= \sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s) g(s)\; ds + \sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s)(g(t_{x,i}) - g(s))\; ds \\ &\quad - \frac{1}{2} \sum_{i=1}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i+1}) d_{x,i}^2 + O\left(\frac{1}{n^2 h}\right). \end{aligned}$$

On the one hand, we have,

$$\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i}} \varphi_{x,h}(s) g(s)\ ds = \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\ ds.$$

On the other hand, Taylor expansion of g and $\varphi_{x,h}$ arround $t_{x,i}$ yield,

$$\begin{aligned} g(t_{x,i}) &= g(s) + (t_{x,i}-s) g'(t_{x,i}) - \frac{1}{2}(t_{x,i}-s)^2 g''(s'_i), \\ \varphi_{x,h}(s) &= \varphi_{x,h}(t_{x,i}) + (s - t_{x,i})\varphi'_{x,h}(s''_i). \end{aligned}$$

for some s'_i and s''_i in $]s, t_{x,i}[$. Thus,

$$\begin{aligned} \mathbb{E}\ (\hat{g}_n^{pro}(x)) &= \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\ ds + \sum_{i=2}^{N_{T_n}-1} g'(t_{x,i})\varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} (t_{x,i}-s)\ ds \\ &- \sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} (t_{x,i}-s)^2 \varphi'_{x,h}(s''_i)\ ds - \frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} g''(s'_i)(t_{x,i}-s)^2\ ds \\ &+ \frac{1}{2}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i}} g''(s'_i)\varphi'_{x,h}(s''_i)(t_{x,i}-s)^3\ ds - \frac{1}{2}\sum_{i=1}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i+1}) d_{x,i}^2 \\ &+ O\left(\frac{1}{n^2 h}\right). \end{aligned}$$

Using the boundedness of g' and g'' in addition to Lemma 2.4.1 and Equation (2.51), we obtain,

$$\begin{aligned} &\sum_{i=1}^{N_{T_n}} g'(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} (t_{x,i}-s)^2 \varphi'_{x,h}(s''_i)\ ds = O\left(\frac{1}{n^2 h}\right). \\ &\frac{1}{2}\sum_{i=1}^{N_{T_n}} \varphi_{x,h}(t_{x,i}) \int_{t_{x,i-1}}^{t_{x,i}} g''(s'_i)(t_{x,i}-s)^2\ ds = O\left(\frac{1}{n^2}\right). \\ &\frac{1}{2}\sum_{i=1}^{N_{T_n}} \int_{t_{x,i-1}}^{t_{x,i}} g''(s'_i)\varphi'_{x,h}(s''_i)(t_{x,i}-s)^3\ ds = O\left(\frac{1}{n^3 h}\right). \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}\ (\hat{g}_n^{pro}(x)) &= \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\ ds + \frac{1}{2}\sum_{i=2}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i}) d_{x,i-1}^2 \\ &\quad - \frac{1}{2}\sum_{i=1}^{N_{T_n}-2} g'(t_{x,i})\varphi_{x,h}(t_{x,i+1}) d_{x,i}^2 + O\left(\frac{1}{n^2 h}\right) \\ &= \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\ ds + \frac{1}{2}\sum_{i=1}^{N_{T_n}-2} \Big(g'(t_{x,i+1}) - g'(t_{x,i})\Big)\varphi_{x,h}(t_{x,i+1}) d_{x,i}^2 + O\left(\frac{1}{n^2 h}\right). \end{aligned}$$

Since g' is Lipschitz, then we have,

$$\mathbb{E}\ (\hat{g}_n^{pro}(x)) = \int_{x-h}^{x+h} \varphi_{x,h}(s)\ g(s)\ ds + O\left(\frac{1}{n^2 h}\right). \tag{2.53}$$

Finally, from (2.52) we obtain,

$$\mathbb{E}\ (\hat{g}_n^{pro}(x)) - g(x) = \frac{1}{2}h^2 g''(x) \int_{-1}^{1} t^2 K(t)dt + o(h^2) + O\Big(\frac{1}{n^2 h}\Big).$$

This concludes the proof of Proposition 2.4.2. □

Proof of Proposition 2.4.3.

The great lines of this proof are based on Sacks and Ylvisaker (1966). From the definition of the orthogonal projection (see Section 2.2) and using the Pythagore theorem we obtain,

$$m\Big(\frac{\sigma_{x,h}^2}{m} - \mathrm{Var} g_n^{pro}(x)\Big) = ||f_{x,h}||^2 - ||P_{|T_n} f_{x,h}||^2 = ||f_{x,h} - P_{|T_n} f_{x,h}||^2, \tag{2.54}$$

where $P_{|T_n} f_{x,h}$ is the orthogonal projection of $f_{x,h}$ on the subspace of $\mathcal{F}(\varepsilon)$ spanned by $\{R(\cdot, t_i), t_i \in T_n\}$, denoted here by V_{T_n}. We shall then prove that,

$$||f_{x,h} - P_{|T_n} f_{x,h}||^2 \leq \frac{C}{h} \sup_{0 \leq j \leq n} d_{j,n}^2. \tag{2.55}$$

Recall that $N_{T_n} = \text{Card } I_{x,h} = \text{Card } I_{x,h} = \{i = 1, \cdots, n \ : \ [t_{i-1,n}, t_{i+1,n}] \cap]x-h, x+h[\neq \emptyset\}$ and denote by $t_{x,i}$ the points of T_n for which $i \in I_{x,h}$. Let $g_n := g_{n,x} = \sum_{i=1}^n \gamma_{x,i} R(\cdot, t_{x,i})$ with $\gamma_{x,i} = 0$ for every $i \notin I_{x,h}$. It is clear that $g_n \in V_{T_n}$ and thus from the definition of the orthogonal projection we have,

$$||f_{x,h} - P_{|T_n} f_{x,h}||^2 \leq ||f_{x,h} - g_n||^2.$$

Now using (F1) in the Appendix and the support of $\varphi_{x,h}$ we obtain,

$$\begin{aligned} ||f_{x,h} - g_n||^2 &= \int_0^1 (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)\ dt - \sum_{i=1}^{n} (f_{x,h}(t_i) - g_n(t_i))\gamma_{x,i} \\ &= \int_{x-h}^{x+h} (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)\ dt - \sum_{i=1}^{N_{T_n}} (f_{x,h}(t_{x,i}) - g_n(t_{x,i}))\gamma_{x,i} \end{aligned} \tag{2.56}$$

In what follows, we distinguish between three cases according to the location of $t_{x,1}$ and $t_{x,N_{T_n}}$ in the interval $[x-h, x+h]$.
First case. Suppose first that $t_{x,1} = x - h$ and $t_{x,N_{T_n}} = x + h$ and take,

$$\gamma_{x,i} = \int_{t_{x,i}}^{t_{x,i+1}} \varphi_{x,h}(t)\ dt \ \ \text{for} \ \ i = 1, \ldots, N_{T_n} - 1. \tag{2.57}$$

we have in this case,

$$||f_{x,h} - g_n||^2 = \sum_{i=1}^{N_{T_n}} \int_{t_{x,i}}^{t_{x,i+1}} \Big((f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,i}) - g_n(t_{x,i})) \Big) \varphi_{x,h}(t)\ dt. \tag{2.58}$$

Assumption (A) yields that $f_{x,h}$ is twice differentiable on $[0,1]$, while g_n is twice differentiable everywhere except on T_n, but it has left and right derivatives. Taylor expansion of $f_{x,h} - g_n$ around $t_{x,i}$ for $i = 1, \cdots, N_{T_n} - 1$ and $t \in]t_{x,i}, t_{x,i+1}[$ gives,

$$\begin{aligned} f_{x,h}(t) - g_n(t) &= (f_{x,h}(t_{x,i}) - g_n(t_{x,i})) + (t - t_{x,i})(f'_{x,h}(t_{x,i}) - g'_n(t_{x,i}^+)) \\ &\quad + \frac{1}{2}(t - t_{x,i})^2 (f''_{x,h}(\theta_{x,t}) - g''_n(\theta_{x,t}^+)), \end{aligned} \tag{2.59}$$

for some $\theta_{x,t} \in]t_{x,i}, t[$. On the one hand, we have,

$$g_n'(t_{x,i}^+) = \sum_{j=1}^{N_{T_n}-1} R^{(0,1)}(t_{x,j}, t_{x,i}^+)\gamma_{x,j}. \tag{2.60}$$

On the other hand, using (2.38) we obtain,

$$\begin{aligned} f_{x,h}'(t_{x,i}) &= \int_{x-h}^{x+h} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds = \sum_{j=1}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds \\ &= \sum_{\substack{j=1\\ j\neq i}}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds\ +\int_{t_{x,i}}^{t_{x,i+1}} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds. \end{aligned} \tag{2.61}$$

When $j \neq i$ we have,

$$\int_{t_{x,j}}^{t_{x,j+1}} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds = R^{(0,1)}(t_{x,j}, t_{x,i})\gamma_{x,j} + \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j})R^{(1,1)}(\delta_{s,j}, t_{x,i})\varphi_{x,h}(s)\ ds, \tag{2.62}$$

for some $\delta_{s,j} \in]t_{x,j}, s[$, while for $j = i$ we have,

$$\begin{aligned} \int_{t_{x,i}}^{t_{x,i+1}} R^{(0,1)}(s, t_{x,i}^+)\varphi_{x,h}(s)\ ds &= \int_{t_{x,i}}^{t_{x,i+1}} R^{(0,1)}(s, t_{x,i}^-)\varphi_{x,h}(s)\ ds \\ &= R^{(0,1)}(t_{x,i}, t_{x,i}^-)\gamma_{x,i} + \int_{t_{x,i}}^{t_{x,i+1}} (s - t_{x,i})R^{(1,1)}(\delta_{s,i}^+, t_{x,i}^-)\varphi_{x,h}(s)\ ds. \end{aligned} \tag{2.63}$$

Collecting (2.60), (2.61), (2.62) and (2.63) we obtain,

$$\begin{aligned} f_{x,h}'(t_{x,i}) - g_n'(t_{x,i}^+) &= \sum_{\substack{j=1\\ j\neq i}}^{N_{T_n}-1} R^{(0,1)}(t_{x,j}, t_{x,i})\gamma_{x,j} + \sum_{\substack{j=1\\ j\neq i}}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j})R^{(1,1)}(\delta_{s,j}, t_{x,i})\varphi_{x,h}(s)\ ds \\ &+ R^{(0,1)}(t_{x,i}, t_{x,i}^-)\gamma_{x,i} + \int_{t_{x,i}}^{t_{x,i+1}} R^{(1,1)}(\delta_{s,i}^+, t_{x,i}^-)\varphi_{x,h}(s)\ ds - \sum_{j=1}^{N_{T_n}-1} R^{(0,1)}(t_{x,j}, t_{x,i}^+)\gamma_{x,j} \\ &= \alpha(t_{x,i})\gamma_{x,i} + \sum_{j=1}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j})R^{(1,1)}(\delta_{s,j}^+, t_{x,i}^-)\varphi_{x,h}(s)\ ds. \end{aligned}$$

It is easy to see that,

$$\begin{aligned} |f_{x,h}'(t_{x,i}) - g_n'(t_{x,i}^+)| &\leq \alpha_1\gamma_{x,i} + \frac{K_\infty}{h}R_1 \sum_{j=1}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j})\ ds \\ &\leq \frac{K_\infty}{h}\alpha_1 d_{x,i} + \frac{K_\infty}{2h}R_1 \sum_{j=1}^{N_{T_n}-1} d_{x,j}^2. \end{aligned} \tag{2.64}$$

We deduce from (2.39) that for all $\theta_{x,t} \in]t_{x,i}, t_{x,i+1}[$ we have,

$$|f''_{x,h}(\theta_{x,t})| \leq \frac{K_\infty}{h}\alpha_1 + \frac{K_\infty}{h}R_2 \times 2h = \frac{K_\infty}{h}\alpha_1 + 2K_\infty R_2.$$

In addition, for $\theta_{x,t} \in]t_{x,i}, t_{x,i+1}[$ we have,

$$|g''_n(\theta^+_{x,t})| = \left| \sum_{j=1}^{N_{T_n}-1} R^{(0,2)}(t_{x,j}, \theta^+_{x,t})\gamma_{x,j} \right| \leq \frac{K_\infty}{h}R_2 \sum_{j=1}^{N_{T_n}-1} d_{x,j} = \frac{K_\infty}{h}R_2 \times 2h = 2K_\infty R_2,$$

Thus,

$$|f''_{x,h}(\theta_{x,t}) - g''_n(\theta^+_{x,t})| \leq \frac{K_\infty}{h}\alpha_1 + 4K_\infty R_2. \tag{2.65}$$

Equations (2.59), (2.64) and (2.65) yield that for $i = 1, \cdots, N_{T_n} - 1$,

$$\begin{aligned}
&\left| \int_{t_{x,i}}^{t_{x,i+1}} \big[(f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,i}) - g_n(t_{x,i}))\big]\varphi_{x,h}(t)\, dt \right| \\
&\leq \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})|f'_{x,h}(t_{x,i}) - g'_n(t^+_{x,i})||\varphi_{x,h}(t)|\, dt \\
&\qquad + \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2|f''_{x,h}(\theta_{x,t}) - g''_n(\theta^+_{x,t})||\varphi_{x,h}(t)|\, dt \\
&\leq \left(\frac{K_\infty}{h}\alpha_1 d_{x,i} + \frac{K_\infty}{2h}R_1 \sum_{j=1}^{N_{T_n}-1} d^2_{x,j}\right) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})|\varphi_{x,h}(t)|\, dt \\
&\qquad + \frac{1}{2}\left(\frac{K_\infty}{h}\alpha_1 + 4K_\infty R_2\right) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2|\varphi_{x,h}(t)|\, dt \\
&\leq \left(\frac{K_\infty}{h}\alpha_1 d_{x,i} + \frac{K_\infty}{2h}R_1 \sum_{j=1}^{N_{T_n}-1} d^2_{x,j}\right)\frac{K_\infty}{2h}d^2_{x,i} + \frac{1}{2}\left(\frac{K_\infty}{h}\alpha_1 + 4K_\infty R_2\right)\frac{K_\infty}{3h}d^3_{x,i} \\
&\leq \frac{K^2_\infty}{4h^2}R_1 d^2_{x,i} \sum_{j=1}^{N_{T_n}-1} d^2_{x,j} + \frac{2K^2_\infty}{3h}\Big(\frac{\alpha_1}{h} + R_2\Big)d^3_{x,i}.
\end{aligned} \tag{2.66}$$

Injecting this inequality in (2.58) yields,

$$\begin{aligned}
||f_{x,h} - P_{|T_n}f_{x,h}||^2 &\leq \frac{K^2_\infty}{4h^2}R_1\left(\sum_{i=1}^{N_{T_n}-1} d^2_{x,i}\right)^2 + \frac{2K^2_\infty}{3h}\Big(\frac{\alpha_1}{h} + R_2\Big)\sum_{i=1}^{N_{T_n}-1} d^3_{x,i} \\
&\leq \frac{K^2_\infty}{4h^2}R_1 \sup_{1\leq i\leq n} d^2_{i,n}\left(\sum_{i=1}^{N_{T_n}-1} d_{x,i}\right)^2 + \frac{2K^2_\infty}{3h}\big(\frac{\alpha_1}{h} + R_2\big) \sup_{1\leq i\leq n} d^2_{i,n} \sum_{i=1}^{N_{T_n}-1} d_{x,i}.
\end{aligned}$$

Since $\sum_{i=1}^{N_{T_n}-1} d_{x,i} = 2h$ then,

$$||f_{x,h} - P_{|T_n}f_{x,h}||^2 \leq \left(\frac{4}{3h}\alpha_1 + R_1 + \frac{4}{3}R_2\right)K^2_\infty \sup_{1\leq i\leq n} d^2_{i,n}$$

Finally, since $h < 1$ then,

$$||f_{x,h} - P_{|T_n}f_{x,h}||^2 \leq \left(\frac{4}{3}\alpha_1 + R_1 + \frac{4}{3}R_2\right)K^2_\infty \frac{1}{h}\sup_{1\leq i\leq n} d^2_{i,n}.$$

Proposition 2.4.3 is then proved for the first case.

Second case. Consider now the case where $t_{x,1} < x-h$ and $t_{x,N_{T_n}} > x+h$. For $i = 2, \ldots, N_{T_n}-2$ set,

$$\gamma_{x,i} = \int_{t_{x,i}}^{t_{x,i+1}} \varphi_{x,h}(t)\, dt,\ \gamma_{x,1} = \int_{x-h}^{t_{x,2}} \varphi_{x,h}(t)\, dt,\ \gamma_{x,N_{T_n}-1} = \int_{t_{x,N_{T_n}-1}}^{x+h} \varphi_{x,h}(t)\, dt \text{ and } \gamma_{x,N_{T_n}} = 0.$$

Using this we obtain,

$$\begin{aligned}
||f_{x,h} - g_n||^2 &= \int_{x-h}^{t_{x,2}} \Big((f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,1}) - g_n(t_{x,1}))\Big)\varphi_{x,h}(t)\, dt \\
&+ \sum_{i=2}^{N_{T_n}} \int_{t_{x,i}}^{t_{x,i+1}} \Big((f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,i}) - g_n(t_{x,i}))\Big)\varphi_{x,h}(t)\, dt \\
&+ \int_{t_{x,N_{T_n}}}^{x+h} \Big((f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,N_{T_n}}) - g_n(t_{x,N_{T_n}}))\Big)\varphi_{x,h}(t)\, dt. \qquad (2.67)
\end{aligned}$$

We first control the first term of (2.67). Let,

$$A_{x,h}^{(1)} = \int_{x-h}^{t_{x,2}} \Big((f_{x,h}(t) - g_n(t)) - (f_{x,h}(t_{x,1}) - g_n(t_{x,1}))\Big)\varphi_{x,h}(t)\, dt.$$

For $t \in]x-h, t_{x,2}[$ we have,

$$\begin{aligned}
f_{x,h}(t) - g_n(t) &= (f_{x,h}(t_{x,1}) - g_n(t_{x,1})) + (t - t_{x,1})(f'_{x,h}(t_{x,1}) - g'_n(t_{x,1}^+)) \\
&+ \frac{1}{2}(t - t_{x,1})^2 (f''_{x,h}(\theta_{x,1}) - g''_n(\theta_{x,1}^+)), \qquad (2.68)
\end{aligned}$$

for some $\theta_{x,1} \in]x-h, t[$. Equation (2.38) yields,

$$\begin{aligned}
f'_{x,h}(t_{x,1}) &= \int_{x-h}^{x+h} R^{(0,1)}(s, t_{x,1}^+)\varphi_{x,h}(s)\, ds = \sum_{j=1}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} R^{(0,1)}(s, t_{x,1}^+)\varphi_{x,h}(s)\, ds \\
&= \int_{x-h}^{t_{x,2}} R^{(0,1)}(s, t_{x,1}^-)\varphi_{x,h}(s)\, ds + \sum_{j=2}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} R^{(0,1)}(s, t_{x,1}^+)\varphi_{x,h}(s)\, ds \\
&= R^{(0,1)}(t_{x,1}, t_{x,1}^-)\gamma_{x,1} + \int_{x-h}^{t_{x,2}} (s - t_{x,1}) R^{(1,1)}(\delta_{s,1}^+, t_{x,1}^-)\varphi_{x,h}(s)\, ds \\
&+ \sum_{j=2}^{N_{T_n}-1} R^{(0,1)}(t_{x,j}, t_{x,1})\gamma_{x,j} + \sum_{j=2}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j}) R^{(1,1)}(\delta_{s,j}, t_{x,1}^+)\varphi_{x,h}(s)\, ds. \qquad (2.69)
\end{aligned}$$

Recall that,

$$g'_n(t_{x,1}^+) = R^{(0,1)}(t_{x,1}, t_{x,1}^+)\gamma_{x,1} + \sum_{j=2}^{N_{T_n}-1} R^{(0,1)}(t_{x,j}, t_{x,1})\gamma_{x,j}. \qquad (2.70)$$

Equations (2.69) and (2.70) give,

$$\begin{aligned}
f'_{x,h}(t_{x,1}) - g'_n(t_{x,1}^+) &= \alpha(t_{x,1})\gamma_{x,1} + \sum_{j=2}^{N_{T_n}-1} \int_{t_{x,j}}^{t_{x,j+1}} (s - t_{x,j}) R^{(1,1)}(\delta_{s,j}, t_{x,1}^+)\varphi_{x,h}(s)\, ds \\
&+ \int_{x-h}^{t_{x,2}} (s - t_{x,1}) R^{(1,1)}(\delta_{s,1}^+, t_{x,1}^-)\varphi_{x,h}(s)\, ds.
\end{aligned}$$

Note that $t_{x,2}-(x-h)\leq \sup_{1\leq i\leq n} d_{i,n}$. We obtain,

$$\begin{aligned}
|f'_{x,h}(t_{x,1})-g'_n(t^-_{x,1})| &\leq \frac{K_\infty}{h}\alpha_1 \sup_{1\leq i\leq n} d_{i,n} + \frac{K_\infty}{2h}R_1 \sum_{j=2}^{N_{T_n}-1} d^2_{x,j} + \frac{K_\infty}{2h}R_1 \sup_{1\leq i\leq n} d^2_{i,n} \\
&\leq \frac{K_\infty}{h}\alpha_1 \sup_{1\leq i\leq n} d_{i,n} + K_\infty R_1 \sup_{1\leq i\leq n} d_{i,n} + \frac{K_\infty}{2h}R_1 \sup_{1\leq i\leq n} d^2_{i,n} \\
&\leq K_\infty\Big(\frac{\alpha_1}{h}+\frac{3}{2}R_1\Big) \sup_{1\leq i\leq n} d_{i,n} \qquad (2.71)
\end{aligned}$$

By (2.65) we have,

$$|f''_{x,h}(\theta_{x,t})-g''_n(\theta^-_{x,t})| \leq \frac{K_\infty}{h}\alpha_1 + 4K_\infty R_2. \qquad (2.72)$$

Equations (2.68), (2.71) and (2.72) yield,

$$\begin{aligned}
|A^{(1)}_{x,h}| \leq \; & |f'_{x,h}(t_{x,1})-g'_n(t^+_{x,1})| \int_{x-h}^{t_{x,2}} (t-t_{x,1})|\varphi_{x,h}(t)|\, dt \\
& + \frac{1}{2}\int_{x-h}^{t_{x,2}} (t-t_{x,1})^2 |f''_{x,h}(\theta_{x,1})-g''_n(\theta^+_{x,1})||\varphi_{x,h}(t)|\, dt \\
\leq & \left(K_\infty\Big(\frac{\alpha_1}{h}+\frac{3}{2}R_1\Big)\sup_{1\leq i\leq n} d_{i,n}\right)\frac{K_\infty}{2h}\sup_{1\leq i\leq n} d^2_{i,n} + \Big(\frac{K_\infty}{h}\alpha_1+4K_\infty R_2\Big)\frac{K_\infty}{6h}\sup_{1\leq i\leq n} d^3_{i,n} \\
\leq & \left(\frac{2}{3}\alpha_1+\frac{3}{4}R_1+\frac{2}{3}R_2\right)\frac{K^2_\infty}{h^2}\sup_{1\leq i\leq n} d^3_{i,n}. \qquad (2.73)
\end{aligned}$$

Similarly we obtain,

$$\begin{aligned}
A^{(2)}_{x,h} &\triangleq \int_{t_{x,N_{T_n}}}^{x+h} \Big((f_{x,h}(t)-g_n(t)) - (f_{x,h}(t_{x,N_{T_n}})-g_n(t_{x,N_{T_n}}))\Big)\varphi_{x,h}(t)\, dt \\
|A^{(2)}_{x,h}| &\leq \left(\frac{2}{3}\alpha_1+\frac{3}{4}R_1+\frac{2}{3}R_2\right)\frac{K^2_\infty}{h^2}\sup_{1\leq i\leq n} d^3_{i,n}. \qquad (2.74)
\end{aligned}$$

Thus,

$$|A^{(1)}_{x,h}+A^{(2)}_{x,h}| \leq \left(\frac{4}{3}\alpha_1+\frac{3}{2}R_1+\frac{4}{3}R_2\right)\frac{K^2_\infty}{h^2}\sup_{1\leq i\leq n} d^3_{i,n}.$$

For $i=2,\ldots,N_{T_n}-2$, similar calculations as those leading to (2.66) give,

$$\begin{aligned}
&\Big|\int_{t_{x,i}}^{t_{x,i+1}} \big((f_{x,h}(t)-g_n(t))-(f_{x,h}(t_{x,i})-g_n(t_{x,i}))\big)\varphi_{x,h}(t)\, dt\Big| \\
&\leq \frac{K^2_\infty}{4h^2}R_1 d^2_{x,i}\sum_{j=1}^{N_{T_n}} d^2_{x,j} + \frac{2K^2_\infty}{3h}(\frac{\alpha_1}{h}+R_2)d^3_{x,i}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\Big|\sum_{i=2}^{N_{T_n}-2}\int_{t_{x,i}}^{t_{x,i+1}} \Big((f_{x,h}(t)-g_n(t))-(f_{x,h}(t_{x,i})-g_n(t_{x,i}))\Big)\varphi_{x,h}(t)\, dt\Big| \\
&\leq \left(\frac{4}{3}\alpha_1+R_1+\frac{4}{3}R_2\right)\frac{K^2_\infty}{h}\sup_{1\leq i\leq n} d^2_{i,n}. \qquad (2.75)
\end{aligned}$$

Then, Equations (2.73), (2.74) and (2.75) yield,

$$\begin{aligned}||f_{x,h} - P_{T_n} f_{x,h}||^2 &\leq \left(\frac{4}{3}\alpha_1 + \frac{3}{2}R_1 + \frac{4}{3}R_2\right)\frac{K_\infty^2}{h}\sup_{1\leq i\leq n} d_{i,n}^2 + \left(\frac{4}{3}\alpha_1 + R_1 + \frac{4}{3}R_2\right)\frac{K_\infty^2}{h^2}\sup_{1\leq i\leq n} d_{i,n}^3 \\ &= \left(\frac{8}{3}\alpha_1 + \frac{5}{2}R_1 + \frac{8}{3}R_2\right)\frac{K_\infty^2}{h}\sup_{1\leq i\leq n} d_{i,n}^2\end{aligned}$$

Third case. Suppose now that $t_{x,1} = x - h$ and $t_{x,N_{T_n}} > x + h$ (respectively $t_{x,1} < x - h$ and $t_{x,N_{T_n}} = x + h$). Let $T_{n-1} = T_n - \{x - h\}$ (respectively $T_{n-1} = T_n - \{x + h\}$). Since $P_{T_{n-1}} f_{x,h} \in V_{T_n}$ we obtain,

$$||f_{x,h} - P_{T_n} f_{x,h}||^2 \leq ||f_{x,h} - P_{T_{n-1}} f_{x,h}||^2,$$

we can then apply the result of the second case to the right side of the previous inequality. The proof of Proposition 2.4.3 is complete. □

Proof of Proposition 2.4.4.

The great lines of this proof are based on the work of Sacks and Ylvisaker (1966). Keeping Equation (2.54) in mind we deduce that Equation (2.15) is equivalent to,

$$\varliminf_{n\to\infty} \frac{N_{T_n}^2}{h}||f_{x,h} - P_{|T_n} f_{x,h}||^2 \geq \frac{1}{12}\alpha(x)\left\{\int_{-1}^{1} K^{2/3}(t)dt\right\}^3. \tag{2.76}$$

We shall take the same notation as in the previous proof. Let $g_n = P_{|T_n} f_{x,h}$, it is shown by Equation (2.102) in the Appendix that:

$$g_n(t_i) = f_{x,h}(t_i) = \sum_{j=1}^{n} R(t_j, t_i) m_{x,h}(t_j), \quad \text{for } i = 1, \cdots, n.$$

We have from (F1) in the Appendix that,

$$\begin{aligned}||f_{x,h} - g_n||^2 &= \int_0^1 (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)\, dt - \sum_{i=1}^{n} m_{x,h}(t_i)(f_{x,h}(t_i) - g_n(t_i)) \\ &= \int_{x-h}^{x+h} (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)dt.\end{aligned}$$

Suppose first that $t_{x,1} = x - h$ and $t_{x,N_{T_n}} = x + h$, then the last equalities give,

$$||f_{x,h} - g_n||^2 = \sum_{i=1}^{N_{T_n}-1} \int_{t_{x,i}}^{t_{x,i+1}} (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)dt. \tag{2.77}$$

Under Assumptions (A) and (B), the function $f_{x,h}$ is twice differentiable at every $t \in [0, 1]$ and g_n is twice differentiable at every $t \in [0, 1]$ except on T_n, however, it has left and right derivatives. We expand $(f_{x,h} - g_n)$ in a Taylor series around $t_{x,i}$ for $t \in]t_{x,i}, t_{x,i+1}[$ up to order 2 we obtain,

$$\begin{aligned}f_{x,h}(t) - g_n(t) = (f_{x,h}(t_{x,i}) - g_n(t_{x,i})) &+ (t - t_{x,i})(f'_{x,h}(t_{x,i}) - g'_n(t_{x,i}^+)) \\ &+ \frac{1}{2}(t - t_{x,i})^2(f''_{x,h}(\sigma_{x,t}) - g''_n(\sigma_{x,t}^+)),\end{aligned}$$

for some $\sigma_{x,t} \in]t_{x,i}, t[$. Since $g_n(t_{x,i}) = f_{x,h}(t_{x,i})$ then,

$$f_{x,h}(t) - g_n(t) = (t - t_{x,i})(f'_{x,h}(t_{x,i}) - g'_n(t^+_{x,i})) + \frac{1}{2}(t - t_{x,i})^2(f''_{x,h}(\sigma_{x,t}) - g''_n(\sigma^+_{x,t})), \tag{2.78}$$

On the one hand, we have for $i \in 1, \ldots, N_{T_n} - 1$,

$$f_{x,h}(t_{x,i+1}) - g_n(t_{x,i+1}) = d_{x,i}(f'_{x,h}(t_{x,i}) - g'_n(t^+_{x,i})) + \frac{1}{2}d^2_{x,i}(f''_{x,h}(\sigma_{x,i}) - g''_n(\sigma^+_{x,i})).$$

for some $\sigma_{x,i} \in]t_{x,i}, t_{x,i+1}[$. Thus,

$$f'_{x,h}(t_{x,i}) - g'_n(t^+_{x,i}) = -\frac{1}{2}d_{x,i}(f''_{x,h}(\sigma_{x,i}) - g''_n(\sigma^+_{x,i})). \tag{2.79}$$

On the other hand, it is shown by (F4) in the Appendix that,

$$f''_{x,h}(t) - g''_n(t^+) = -\alpha(t)\varphi_{x,h}(t) + \langle R^{(0,2)}(\cdot, t^+), f_{x,h} - g_n\rangle. \tag{2.80}$$

Injecting (2.79) and (2.80) in (2.78) gives,

$$\begin{aligned}
f_{x,h}(t) - g_n(t) &= -\frac{1}{2}(t - t_{x,i})d_{x,i}(f''_{x,h}(\sigma_{x,i}) - g''_n(\sigma^+_{x,i})) + \frac{1}{2}(t - t_{x_i})^2(f''_{x,h}(\sigma_{x,i}) - g''_n(\sigma^+_{x,i})) \\
&= \frac{1}{2}d_{x,i}(t - t_{x,i})\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i}) - \frac{1}{2}(t - t_{x,i})^2\alpha(\sigma_{x,t})\varphi_{x,h}(\sigma_{x,t}) \\
&\quad - \frac{1}{2}d_{x,i}(t - t_{x,i})\langle R^{(0,2)}(\cdot, \sigma^+_{x,i}), f_{x,h} - g_n\rangle + \frac{1}{2}(t - t_{x,i})^2\langle R^{(0,2)}(\cdot, \sigma^+_{x,t}), f_{x,h} - g_n\rangle.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\int_{t_{x,i}}^{t_{x,i+1}} (f_{x,h}(t) - g_n(t))\varphi_{x,h}(t)\, dt = \\
&\frac{1}{2}d_{x,i}\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})\varphi_{x,h}(t)\, dt - \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2\alpha(\sigma_{x,t})\varphi_{x,h}(\sigma_{x,t})\varphi_{x,h}(t)\, dt \\
&\quad - \frac{1}{2}d_{x,i}\langle R^{(0,2)}(\cdot, \sigma^+_{x,i}), f_{x,h} - g_n\rangle \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})\varphi_{x,h}(t)\, dt \\
&\quad + \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2\langle R^{(0,2)}(\cdot, \sigma^+_{x,t}), f_{x,h} - g_n\rangle\varphi_{x,h}(t)\, dt \\
&= \frac{1}{4}d^3_{x,i}\alpha(\sigma_{x,i})\varphi^2_{x,h}(\sigma_{x,i}) - \frac{1}{6}d^3_{x,i}\alpha(\sigma_{x,i})\varphi^2_{x,h}(\sigma_{x,i}) \\
&\quad + \frac{1}{2}d_{x,i}\alpha(\sigma_{x_i})\varphi_{x,h}(\sigma_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})[\varphi_{x,h}(t) - \varphi_{x,h}(\sigma_{x,i})]\, dt \\
&\quad - \frac{1}{2}\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2[\varphi_{x,h}(t) - \varphi_{x,h}(\sigma_{x,i})]\, dt \\
&\quad - \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2[\alpha(\sigma_{x,t})\varphi_{x,h}(\sigma_{x,t}) - \alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i})]\varphi_{x,h}(t)\, dt \\
&\quad - \frac{1}{2}d_{x,i}\langle R^{(0,2)}(\cdot, \sigma^+_{x,i}), f_{x,h} - g_n\rangle \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})\varphi_{x,h}(t)\, dt \\
&\quad + \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2\langle R^{(0,2)}(\cdot, \sigma^+_{x,t}), f_{x,h} - g_n\rangle\varphi_{x,h}(t)\, dt \\
&= \frac{1}{12}d^3_{x,i}\alpha(\sigma_{x,i})\varphi^2_{x,h}(\sigma_{x,i}) + A^{(1)}_{x,i} - A^{(2)}_{x,i} - A^{(3)}_{x,i} - A^{(4)}_{x,i} + A^{(5)}_{x,i},
\end{aligned} \tag{2.81}$$

$$\text{where,}\quad A_{x,i}^{(1)} = \frac{1}{2}d_{x,i}\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i})\int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})[\varphi_{x,h}(t)-\varphi_{x,h}(\sigma_{x,i})]\ dt.$$
$$A_{x,i}^{(2)} = \frac{1}{2}\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i})\int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})^2[\varphi_{x,h}(t)-\varphi_{x,h}(\sigma_{x,i})]\ dt.$$
$$A_{x,i}^{(3)} = \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})^2[\alpha(\sigma_{x,t})\varphi_{x,h}(\sigma_{x,t})-\alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i})]\varphi_{x,h}(t)\ dt.$$
$$A_{x,i}^{(4)} = \frac{1}{2}d_{x,i}\langle R^{(0,2)}(\cdot,\sigma_{x,i}^{+}), f_{x,h}-g_n\rangle\int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})\varphi_{x,h}(t)\ dt.$$
$$A_{x,i}^{(5)} = \frac{1}{2}\int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})^2\langle R^{(0,2)}(\cdot,\sigma_{x,t}^{+}), f_{x,h}-g_n\rangle\varphi_{x,h}(t)\ dt.$$

We shall now control these quantities. Let,

$$B_{x,i}^{(1)} = \sup_{t_{x,i}<s,t<t_{x,i+1}}|\varphi_{x,h}(t)-\varphi_{x,h}(s)| \text{ and } B_{x,i}^{(2)} = \sup_{t_{x,i}<s,t<t_{x,i+1}}|\alpha(t)\varphi_{x,h}(t)-\alpha(s)\varphi_{x,h}(s)|.$$

Since α and $\varphi_{x,h}$ are Lipschitz then,

$$\sup_{0\le i\le n} B_{x,i}^{(1)} = O\Big(\frac{1}{h^2}\sup_{0\le j\le n} d_{j,n}\Big) \quad\text{and}\quad \sup_{0\le i\le n} B_{x,i}^{(2)} = O\Big(\frac{1}{h^2}\sup_{0\le j\le n} d_{j,n}\Big). \tag{2.82}$$

Elementary calculations show that,

$$|A_{x,i}^{(1)}| \le \frac{a_1}{h}B_{x,i}^{(1)}d_{x,i}^3, \qquad |A_{x,i}^{(2)}| \le \frac{a_2}{h}B_{x,i}^{(1)}d_{x,i}^3 \quad\text{and } |A_{x,i}^{(3)}| \le \frac{a_3}{h}B_{x,i}^{(2)}d_{x,i}^3, \tag{2.83}$$

for appropriate constants a_1, a_2 and a_3. We obtain from the Cauchy-Schwartz inequality, Assumption (C) and Proposition 2.4.3 that,

$$|A_{x,i}^{(4)}| + |A_{x,i}^{(5)}| \le \frac{a_4}{h}d_{x,i}^3||f_{x,h}-g_n|| \le \frac{1}{h}d_{x,i}^3\underbrace{a_4\sqrt{\frac{C}{h}}}_{a_h}\sup_{0\le j\le n} d_{j,n}, \tag{2.84}$$

for an appropriate constant a_4 (C is defined in Proposition 2.4.3). Thus,

$$\begin{aligned}
&\int_{t_{x,i}}^{t_{x,i+1}}(f_{x,h}(t)-g_n(t))\varphi_{x,h}(t)\ dt \\
&= \frac{1}{12}d_{x,i}^3\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}) + A_{x,i}^{(1)} - A_{x,i}^{(2)} - A_{x,i}^{(3)} - A_{x,i}^{(4)} + A_{x,i}^{(5)} \\
&\ge \frac{1}{12}d_{x,i}^3\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}) - d_{x,i}^3\Big(\frac{a_1}{h}B_{x,i}^{(1)} + \frac{a_2}{h}B_{x,i}^{(2)} + \frac{a_h}{h}\sup_{0\le j\le n} d_{j,n}\Big).
\end{aligned} \tag{2.85}$$

Let,

$$\rho_{h,N_{T_n}} = \sup_{0\le i\le N_{T_n}}\Big(\frac{a_1}{h}B_{x,i}^{(1)} + \frac{a_2}{h}B_{x,i}^{(2)} + \frac{a_h}{h}\sup_{0\le j\le n} d_{j,n}\Big).$$

Equation (2.82) implies that for an appropriate constant c and c' we have,

$$|\rho_{h,N_{T_n}}| \le \Big(\frac{c}{h^3}\sup_{0\le j\le n} d_{j,n} + \frac{c'}{h^{3/2}}\sup_{0\le j\le n} d_{j,n}\Big).$$

Using (2.85) and (2.77) together with Equation (2.85) in (2.77) we obtain,

$$\begin{aligned}||f_{x,h}-g_n||^2 &\geq \sum_{i=1}^{N_{T_n}-1}\Big(\frac{1}{12}\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i})-\rho_{h,N_{T_n}}\Big)d_{x,i}^3\\ &\geq \frac{1}{12}\sum_{i=1}^{N_{T_n}-1}\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i})d_{x,i}^3-\frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n}d_{j,n}^4-\frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n}d_{j,n}^4. \qquad (2.86)\end{aligned}$$

Then the Hölder's inequality gives,

$$||f_{x,h}-g_n||^2 \geq \frac{1}{12(N_{T_n}-1)^2}\Big\{\sum_{j=1}^{N_{T_n}-1}[\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i})]^{\frac{1}{3}}d_{x,i}\Big\}^3-\frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n}d_{j,n}^4-\frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n}d_{j,n}^4.$$

We shall now control the first term of the right side of this inequality. We have,

$$\begin{aligned}&\Big\{\sum_{j=1}^{N_{T_n}-1}\Big(\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i})\Big)^{\frac{1}{3}}d_{x,i}\Big\}^3\\ &=\Big\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}dt-\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\Big((\alpha(t)\varphi_{x,h}^2(t))^{\frac{1}{3}}-(\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}))^{\frac{1}{3}}\Big)\Big\}^3\\ &=\Big\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}dt\Big\}^3-\Big\{\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\Big((\alpha(t)\varphi_{x,h}^2(t))^{\frac{1}{3}}-(\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}))^{\frac{1}{3}}\Big)\Big\}^3\\ &\quad-3\Big\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}dt\Big\}^2\Big\{\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\Big((\alpha(t)\varphi_{x,h}^2(t))^{\frac{1}{3}}-(\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}))^{\frac{1}{3}}\Big)\Big\}\\ &\quad+3\Big\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}dt\Big\}\Big\{\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\Big((\alpha(t)\varphi_{x,h}^2(t))^{\frac{1}{3}}-(\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}))^{\frac{1}{3}}\Big)\Big\}^2\\ &\triangleq\Big\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}dt\Big\}^3+B,\end{aligned}$$

We obtain using (2.51) and the fact that α is Lipschitz,

$$B=O\bigg(\Big(\frac{N_{T_n}}{h^{5/3}}\sup_{0\leq j\leq n}d_{j,n}^2\Big)^3\bigg)+O\bigg(\Big(\frac{N_{T_n}}{h^{5/3}}\sup_{0\leq j\leq n}d_{j,n}^2\Big)\,h^{2/3}\bigg)+O\bigg(\Big(\frac{N_{T_n}}{h^{5/3}}\sup_{0\leq j\leq n}d_{j,n}^2\Big)^2h^{1/3}\bigg).$$

Assumption (E) implies that for an appropriate constant c'' we have,

$$|B|\leq\frac{c''N_{T_n}}{h}\sup_{0\leq j\leq n}d_{j,n}^2.$$

Using the Riemann integrability of α and $\varphi_{x,h}$ we get,

$$
\begin{aligned}
||f_{x,h} - g_n||^2 &\geq \frac{1}{12(N_{T_n}-1)^2}\left\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}\,dt\right\}^3 - \frac{c''}{N_{T_n}h}\sup_{0\leq j\leq n} d_{j,n}^2 \\
&- \frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n} d_{j,n}^4 - \frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n} d_{j,n}^4 \\
&\geq \frac{1}{12N_{T_n}^2}\left\{\int_{x-h}^{x+h}\Big(\alpha(t)\varphi_{x,h}^2(t)\Big)^{\frac{1}{3}}\,dt\right\}^3 - \frac{c''}{N_{T_n}h}\sup_{0\leq j\leq n} d_{j,n}^2 - \frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n} d_{j,n}^4 - \frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n} d_{j,n}^4 \\
&= \frac{1}{12h^2N_{T_n}^2}\left\{\int_{x-h}^{x+h}\Big(\alpha(t)K^2(\frac{x-t}{h})\Big)^{\frac{1}{3}}\,dt\right\}^3 - \frac{c''}{N_{T_n}h}\sup_{0\leq j\leq n} d_{j,n}^2 \\
&- \frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n} d_{j,n}^4 - \frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n} d_{j,n}^4 \\
&= \frac{h}{12N_{T_n}^2}\left\{\int_{-1}^{1}\Big(\alpha(x-th)K^2(t)\Big)^{\frac{1}{3}}\,dt\right\}^3 - \frac{c''}{N_{T_n}h}\sup_{0\leq j\leq n} d_{j,n}^2 \\
&- \frac{cN_{T_n}}{h^3}\sup_{0\leq j\leq n} d_{j,n}^4 - \frac{c'N_{T_n}}{h^{3/2}}\sup_{0\leq j\leq n} d_{j,n}^4.
\end{aligned}
$$

Assumption (E) implies that,

$$\lim_{n\to\infty}\frac{1}{h^2}N_{T_n}\sup_{0\leq j\leq n} d_{j,n}^2 = 0, \quad \lim_{n\to\infty}\frac{1}{h^4}\sup_{0\leq j\leq n} d_{j,n}^4 N_{T_n}^3 = 0 \quad \text{and} \quad \lim_{n\to\infty}\frac{c'}{h^{3/2}}\sup_{0\leq j\leq n} d_{j,n}^4 N_{T_n} = 0.$$

Finally the continuity of α yields,

$$\underline{\lim}_{n\to\infty}\frac{N_{T_n}^2}{h}||f_{x,h} - g_n||^2 \geq \frac{1}{12}\alpha(x)\left\{\int_{-1}^{1}K(t)^{\frac{2}{3}}\,dt\right\}^3.$$

Inequality (2.76) is then proved for a sequence of designs containing $x-h$ and $x+h$. Consider now any sequence of designs $\{T_n, n\geq 1\}$ satisfying Assumption (E) we can adjoin the points $\{x-h, x+h\}$ to T_n (if they aren't present). Hence we form a sequence $\{S_n, n\geq 1\}$ with $S_n \in D_{n+2}$ and satisfying (2.76). We have,

$$||f_{x,h} - P_{|S_n}f_{x,h}||^2 \leq ||f_{x,h} - P_{|T_n}f_{x,h}||^2.$$

Then,

$$N_{S_n}^2||f_{x,h} - P_{|S_n}f_{x,h}||^2 \leq N_{S_n}^2||f_{x,h} - P_{|T_n}f_{x,h}||^2. \tag{2.87}$$

We know that $N_{S_n}\in\{N_{T_n}+1, N_{T_n}+2\}$, replacing N_{S_n} in the right term of (2.87) by $(N_{T_n}+2)$ (or $(N_{T_n}+1)$) gives,

$$\frac{N_{S_n}^2}{h}||f_{x,h} - P_{|S_n}f_{x,h}||^2 - \frac{(4+2N_{T_n})}{h}||f_{x,h} - P_{|T_n}f_{x,h}||^2 \leq \frac{N_{T_n}^2}{h}||f_{x,h} - P_{|T_n}f_{x,h}||^2.$$

Assumption (E) and Equation (2.55) yield,

$$\underline{\lim}_{n\to\infty}\frac{(4+2N_{T_n})}{h}||f_{x,h} - P_{|T_n}f_{x,h}||^2 = 0.$$

Hence, for any sequence $\{T_n, n\geq 1\}$ we have,

$$\underline{\lim}_{n\to\infty}\frac{N_{T_n}^2}{h}||f_{x,h} - P_{|T_n}f_{x,h}||^2 \geq \frac{1}{12}\alpha(x)\left\{\int_{-1}^{1}K^{2/3}(t)dt\right\}^3.$$

This completes the proof of Proposition 2.4.4. $\square$

Proof of Proposition 2.4.5.

On the one hand, Proposition 2.4.3 yields that there exists a constant $c > 0$ such that,

$$0 \leq \frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x) \leq \frac{c}{mh} \sup_{0 \leq j \leq n} d_{j,n}^2.$$

Lemma 2.4.1 implies that there exists a constant $c' > 0$ such that,

$$\sup_{0 \leq j \leq n} d_{j,n}^2 \leq \frac{c'}{n^2}.$$

Thus, for $n \geq 1$ we have,

$$0 \leq \frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x) \leq \frac{c'c}{mn^2h}.$$

Finally, taking $C = cc'$ we obtain,

$$\overline{\lim_{n \to \infty}} \; mn^2h\Big(\frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x)\Big) \leq C.$$

Inequality (2.16) is then proved. On the other hand, Proposition 2.4.4 yields,

$$\frac{mN_{T_n}^2}{h}\Big(\frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x)\Big) \geq \frac{1}{12}\alpha(x)\Big\{\int_{-1}^{1} K^{2/3}(t)dt\Big\}^3.$$

Lemma 2.4.1 implies that there exists a constant $c'' > 0$ such that,

$$N_{T_n} < c''nh,$$

which implies that,

$$c''mn^2h\,\Big(\frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x)\Big) \geq \frac{1}{12}\alpha(x)\Big\{\int_{-1}^{1} K^{2/3}(t)dt\Big\}^3.$$

Finally, taking $C' = \frac{1}{12c''}\alpha(x)\Big\{\int_{-1}^{1} K^{2/3}(t)dt\Big\}^3$ we obtain,

$$\underline{\lim}_{n \to \infty} \; mn^2h\Big(\frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x)\Big) \geq C'.$$

This concludes the proof of Proposition 2.4.5. □

Proof of Proposition 2.4.6

The first part of this proof is the same as that of Proposition (2.4.4). Recall that,

$$m\Big(\frac{\sigma_{x,h}^2}{m} - \text{Var } \hat{g}_n^{pro}(x)\Big) = ||f_{x,h}||^2 - ||P_{|T_n}f_{x,h}||^2 = ||f_{x,h} - P_{|T_n}f_{x,h}||^2.$$

Using (2.77) and (2.81) we obtain,

$$\begin{aligned}\text{Var } \hat{g}_n^{pro}(x) - \frac{\sigma_{x,h}^2}{m} &= -\frac{1}{m}||f_{x,h} - P_{|T_n}f_{x,h}||^2 \\ &= -\frac{1}{m}\sum_{i=1}^{N_{T_n}}\Big(\frac{1}{12}d_{x,i}^3\alpha(\sigma_{x,i})\varphi_{x,h}^2(\sigma_{x,i}) + A_{x,i}^{(1)} - A_{x,i}^{(2)} - A_{x,i}^{(3)} - A_{x,i}^{(4)} + A_{x,i}^{(5)}\Big),\end{aligned} \tag{2.88}$$

for some $\sigma_{x,i} \in]t_{x,i}, t_{x,i+1}[$ and some $\sigma_{x,t} \in]t_{x,i}, t[$, where,

$$\begin{aligned}
A_{x,i}^{(1)} &= \frac{1}{2} d_{x,i} \alpha(\sigma_{x,i}) \varphi_{x,h}(\sigma_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})[\varphi_{x,h}(t) - \varphi_{x,h}(\sigma_{x,i})]\, dt. \\
A_{x,i}^{(2)} &= \frac{1}{2} \alpha(\sigma_{x,i}) \varphi_{x,h}(\sigma_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2[\varphi_{x,h}(t) - \varphi_{x,h}(\sigma_{x,i})]\, dt. \\
A_{x,i}^{(3)} &= \frac{1}{2} \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2[\alpha(\sigma_{x,t})\varphi_{x,h}(\sigma_{x,t}) - \alpha(\sigma_{x,i})\varphi_{x,h}(\sigma_{x,i})]\varphi_{x,h}(t)\, dt. \\
A_{x,i}^{(4)} &= \frac{1}{2} d_{x,i} \langle R^{(0,2)}(\cdot, \sigma_{x,i}^{+}), f_{x,h} - g_n \rangle \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})\varphi_{x,h}(t)\, dt. \\
A_{x,i}^{(5)} &= \frac{1}{2} \int_{t_{x,i}}^{t_{x,i+1}} (t - t_{x,i})^2 \langle R^{(0,2)}(\cdot, \sigma_{x,t}^{+}), f_{x,h} - g_n \rangle \varphi_{x,h}(t)\, dt.
\end{aligned}$$

From the definition of the regular sequence of designs (see Definition 2.4.1) and the mean value theorem we have for $i = 1, \cdots, N_{T_n}$,

$$d_{x,i} = t_{x,i+1} - t_{x,i} = F^{-1}\Big(\frac{i+1}{n}\Big) - F^{-1}\Big(\frac{i}{n}\Big) = \frac{1}{n f(t_{x,i}^{*})},$$

where $t_{x,i}^{*} \in]t_{x,i}, t_{x,i+1}[$. Using this together with (2.88) we obtain,

$$\begin{aligned}
\text{Var } \hat{g}_n^{pro}(x) - \frac{\sigma_{x,h}^2}{m} &= -\frac{1}{12mn^2} \sum_{i=1}^{N_{T_n}} d_{x,i} \frac{1}{f^2(t_{x,i}^{*})} \alpha(\sigma_{x,i}) \varphi_{x,h}^2(\sigma_{x,i}) \\
&\quad - \frac{1}{m} \sum_{i=1}^{N_{T_n}} \Big(A_{x,i}^{(1)} - A_{x,i}^{(2)} - A_{x,i}^{(3)} - A_{x,i}^{(4)} + A_{x,i}^{(5)} \Big).
\end{aligned}$$

Lemma 2.4.1 yields that $N_{T_n} = O(nh)$. Using (2.83), (2.84) and (2.82) we obtain,

$$A_{x,i}^{(1)} = O\Big(\frac{1}{n^4 h^3}\Big), \quad A_{x,i}^{(2)} = O\Big(\frac{1}{n^4 h^3}\Big), \quad A_{x,i}^{(3)} = O\Big(\frac{1}{n^4 h^3}\Big) \quad \text{and} \quad A_{x,i}^{(4)} + A_{x,i}^{(5)} = O\Big(\frac{1}{n^4 h^{3/2}}\Big).$$

Finally,

$$\text{Var } \hat{g}_n^{pro}(x) - \frac{\sigma_{x,h}^2}{m} = -\frac{1}{12mn^2} \sum_{i=1}^{N_{T_n}} d_{x,i} \frac{1}{f^2(t_{x,i}^{*})} \alpha(\sigma_{x,i}) \varphi_{x,h}^2(\sigma_{x,i}) + O\Big(\frac{1}{mn^3h^2} + \frac{1}{mn^3\sqrt{h}}\Big).$$

Using a classical approximation of a sum by an integral (see for instance, Lemma 2 in Benelmadani *et al.* (2019a) and the fact that $0 < h < 1$ we obtain,

$$\text{Var } \hat{g}_n^{pro}(x) - \frac{\sigma_{x,h}^2}{m} = -\frac{1}{12mn^2} \int_{x-h}^{x+h} \frac{\alpha(t)}{f^2(t)} \varphi_{x,h}^2(t)\, dt + O\Big(\frac{1}{mn^3h^2}\Big).$$

This concludes the proof of Proposition 2.4.6. □

Proof of Theorem 2.4.1.

First, note that since α and f are Lipschitz functions then the asymptotic expression of the integral in (2.18) is:

$$\begin{aligned}\frac{1}{mn^2}\int_{x-h}^{x+h}\frac{\alpha(t)}{f^2(t)}\varphi_{x,h}^2(t)dt &= \frac{1}{mn^2h}\int_{-1}^{1}\frac{\alpha(x-th)}{f^2(x-th)}K^2(t)\,dt\\ &= \frac{1}{mn^2h}\Big(\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1}K^2(t)\,dt+\int_{-1}^{1}\Big(\frac{\alpha(x-th)}{f^2(x-th)}-\frac{\alpha(x)}{f^2(x)}\Big)K^2(t)\,dt\Big)\\ &= \frac{1}{mn^2h}\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1}K^2(t)\,dt+O\Big(\frac{1}{mn^2}\Big).\end{aligned}$$

This last equality together with Proposition 2.4.6 and Proposition 2.4.2 concludes the proof of Theorem 2.4.1. □

Proof of Corollary 2.4.1.

Let $I_1=\int_0^1 R(x,x)w(x)\,dx$ and put,

$$\Psi(h,m)=-\frac{C_K h}{2m}\int_0^1\alpha(x)w(x)\,dx+\frac{1}{4}h^4B^2\int_0^1[g''(x)]^2w(x)\,dx.$$

We have from Theorem 2.4.1,

$$\text{IMSE}(h)=\frac{I}{m}+\Psi(h,m)+o\Big(h^4+\frac{h}{m}\Big)+O\Big(\frac{1}{mn^2h}+\frac{h}{n}+\frac{1}{n^2h^2}\Big),$$

Let h^* be as defined by (2.20). It is clear that $h^*=\underset{0<h<1}{\text{argmin}}\ \Psi(h,m)$ so that $\Psi(h,m)\geq\Psi(h^*,m)$ for every $0<h<1$. Let $h_{n,m}$ be as defined in Corollary 2.4.1. We have,

$$\frac{\text{IMSE}(h^*)}{\text{IMSE}(h_{n,m})}=\frac{\frac{I_1}{m}+\Psi(h^*,m)+o\Big(h^{*4}+\frac{h^*}{m}\Big)+O\Big(\frac{1}{mn^2h^*}+\frac{h^*}{n}+\frac{1}{n^2h^{*2}}\Big)}{\frac{I_1}{m}+\Psi(h_{n,m},m)+o\Big(h_{n,m}^4+\frac{h_{n,m}}{m}\Big)+O\Big(\frac{1}{mn^2h_{n,m}}+\frac{h_{n,m}}{n}+\frac{1}{n^2h_{n,m}^2}\Big)}$$

$$\leq\frac{I_1+m\Psi(h_{n,m},m)+o\Big(mh^{*4}+h^*\Big)+O\Big(\frac{1}{n^2h^*}+\frac{mh^*}{n}+\frac{m}{n^2h^2}\Big)}{I_1+m\Psi(h_{n,m},m)+o\Big(mh_{n,m}^4+h_{n,m}\Big)+O\Big(\frac{1}{n^2h_{n,m}}+\frac{mh_{n,m}}{n}+\frac{m}{n^2h_{n,m}^2}\Big)}.$$

We have, using the definition of h^*, $mh_{n,m}^3=O(1)$, $\lim\limits_{n,m\to\infty}h_{n,m}=0$ and using the assumption $\frac{m}{n}=O(1)$ as $n,m\to\infty$ we know that $m\Psi(h_{n,m},m)=O(h_{n,m})$. Thus,

$$\overline{\lim_{n,m\to\infty}}\frac{\text{IMSE}(h^*)}{\text{IMSE}(h_{n,m})}\leq 1.$$

This concludes the proof of Corollary 2.4.1. □

Proof of Theorem 2.4.3.

Let $x \in]0,1[$ be fixed. We have the following decomposition,

$$\sqrt{m}\Big(\hat{g}^{pro}_{n,m}(x) - g(x)\Big) = \sqrt{m}\Big(\hat{g}^{pro}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{pro}_{n,m}(x)\big)\Big) + \sqrt{m}\,\Big(\mathbb{E}(\hat{g}^{pro}_{n,m}(x)) - g(x)\Big). \tag{2.89}$$

Since $\lim_{n,m\to\infty} \sqrt{m}h = 0$, $\frac{n}{m} = O(1)$ as $n,m \to \infty$ and $\lim_{n,m\to\infty} nh^2 = \infty$ then Remark 2.4.2 implies that,

$$\lim_{n,m\to\infty} \sqrt{m}\,\Big(\mathbb{E}(\hat{g}^{pro}_{n,m}(x)) - g(x)\Big) = 0.$$

Consider now the first term of the right side of (2.89). Since $\overline{Y}(t_{x,i}) - \mathbb{E}(\overline{Y}(t_{x,i})) = \overline{\varepsilon}(t_{x,i})$, we have, as done by Fraiman and Iribarren (1991),

$$\begin{aligned}\sqrt{m}\Big(\hat{g}^{pro}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{pro}_{n,m}(x)\big)\Big) &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\sum_{i=1}^{n} m_{x,h}(t_i)\varepsilon_j(t_i) \\ = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\sum_{i=1}^{n} m_{x,h}(t_i)\big(\varepsilon_j(t_i) - \varepsilon_j(x)\big) &+ \Big(\sum_{i=1}^{n} m_{x,h}(t_i)\Big)\Big(\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\varepsilon_j(x)\Big).\end{aligned} \tag{2.90}$$

We start by controlling the second term of this last equation. Using Lemma 2.4.2 together with Lemma 2.4.1 we obtain,

$$m_{x,h}(t_{i,n}) = \begin{cases} \frac{1}{2}\varphi_{x,h}(t_{i,n})(t_{i+1,n} - t_{i-1,n}) + O\Big(\frac{1}{n^2h^2} + \frac{1}{n^2\sqrt{h}}\Big) & if \;\; i \notin \{1,n\} \text{ and} \\ & [t_{i-1,n}, t_{i+1,n}] \cap [x-h, x+h] \neq \emptyset, \\ O\Big(\frac{1}{n^2h^2} + \frac{1}{n^2\sqrt{h}}\Big) & \text{if } \; i \in \{1,n\}, \\ O\Big(\frac{1}{n^2\sqrt{h}}\Big) & otherwise. \end{cases}$$

Recall that $N_{T_n} = \text{Card } I_{x,h} = \text{Card }\{i = 1, \cdots, n \;:\; [t_{i-1}, t_{i+1}] \cap]x-h, x+h[\neq \emptyset\}$ and denote by $t_{x,i}$ the points of T_n for which $i \in I_{x,h}$, Lemma 2.4.1 yields that $N_{T_n} = O(nh)$. Thus,

$$\sum_{i=1}^{n} m_{x,h}(t_i) = \frac{1}{2}\sum_{i=2}^{N_{T_n}-1} \varphi_{x,h}(t_{x,i})(t_{x,i+1} - t_{x,i-1}) + O\Big(\frac{1}{nh}\Big).$$

Since $\lim_{n\to\infty} nh = +\infty$, then using the Riemann integrability of K, we obtain,

$$\lim_{n,m\to\infty}\sum_{i=1}^{n} m_{x,h}(t_i) = \frac{1}{2}\lim_{n,m\to\infty}\sum_{i=2}^{N_{T_n}-1} \varphi_{x,h}(t_{x,i})(t_{x,i+1} - t_{x,i-1}) = \int_{-1}^{1} K(t)\; dt = 1.$$

The Central Limit Theorem for i.i.d. variables yields,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\varepsilon_j(x) \xrightarrow[m\to\infty]{\mathscr{D}} Z \quad where \;\; Z \sim \backslash(0, R(x,x)).$$

We shall prove now that the first term of Equation (2.90) tends to 0 in probability as n,m tends to infinity. Let,

$$A_{m,n}(x) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\sum_{i=1}^{n} m_{x,h}(t_i)\big(\varepsilon_j(t_i) - \varepsilon_j(x)\big) \triangleq \frac{1}{\sqrt{m}}\sum_{j=1}^{m} T_{n,j}(x).$$

From the Chebyshev inequality, it suffices to prove that $\lim_{n,m\to\infty} \mathbb{E}(A^2_{m,n}(x)) = 0$. We have for $j \neq l$, $\mathbb{E}(\varepsilon_j(x)\varepsilon_l(y)) = 0$ then $\mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = 0$. Hence,

$$\mathbb{E}(A^2_{m,n}(x)) = \frac{1}{m}\sum_{j=1}^{m}\sum_{l=1}^{m}\mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = \frac{1}{m}\sum_{j=1}^{m}\mathbb{E}(T^2_{n,j}(x)).$$

We have,

$$\begin{aligned}
\mathbb{E}(T^2_{n,j}(x)) &= \sum_{i=1}^{n}\sum_{k=1}^{n} m_{x,h}(t_i)m_{x,h}(t_k)\mathbb{E}\Big(\big(\varepsilon_j(t_i)-\varepsilon_j(x)\big)\big(\varepsilon_j(t_k)-\varepsilon_j(x)\big)\Big)\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} m_{x,h}(t_i)m_{x,h}(t_k)\Big(R(t_i,t_k)-R(t_i,x)-R(x,t_k)+R(x,x)\Big).
\end{aligned}$$

Note that $\mathbb{E}(T^2_{n,j}(x))$ does not depend on j hence,

$$\begin{aligned}
\mathbb{E}(A^2_{m,n}(x)) &= \sum_{i=1}^{n}\sum_{k=1}^{n} m_{x,h}(t_i)m_{x,h}(t_k)\Big(R(t_i,t_k)-R(t_i,x)-R(x,t_k)+R(x,x)\Big)\\
&\triangleq B_{n,1}(x)-B_{n,2}(x)-B_{n,3}(x)+B_{n,4}(x). \qquad (2.91)
\end{aligned}$$

Using Lemma 2.4.2 and the approximation of a sum by an integral (see, for instance, Lemma 2 in Benelmadani *et al.* (2019a) we obtain,

$$B_{n,1}(x) = \int_{x-h}^{x+h}\int_{x-h}^{x+h}\varphi_{x,h}(s)\varphi_{x,h}(t)R(s,t)\ ds\ dt + O\Big(\frac{1}{nh}\Big) = \sigma^2_{x,h} + O\Big(\frac{1}{nh}\Big).$$

Using Equation (2.19) we obtain,

$$B_{n,1}(x) = R(x,x) - \frac{1}{2}\alpha(x)C_K h + o(h) + O\Big(\frac{1}{nh}\Big).$$

where $C_K = \int_{-1}^{1}\int_{-1}^{1}|u-v|K(u)K(v)dudv$. Since $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$ then,

$$\lim_{n\to\infty} B_{n,1}(x) = R(x,x). \qquad (2.92)$$

Consider now the term $B_{n,2}(x)$. We obtain using Lemma 2.4.2 and the approximation of a sum by an integral,

$$\begin{aligned}
B_{n,2}(x) &= \int_{x-h}^{x+h}\int_{x-h}^{x+h}\varphi_{x,h}(s)\varphi_{x,h}(t)R(s,x)\ ds\ dt + O\Big(\frac{1}{nh}\Big)\\
&= \int_{x-h}^{x+h}\varphi_{x,h}(s)R(s,x)\ ds + O\Big(\frac{1}{nh}\Big)\\
&= \int_{-1}^{1}K(s)R(x-hs,x)\ ds + O\Big(\frac{1}{nh}\Big)\\
&= \int_{-1}^{0}K(s)R(x-hs,x)\ ds + \int_{0}^{1}K(s)R(x-hs,x)\ ds + O\Big(\frac{1}{nh}\Big).
\end{aligned}$$

For $s \in]-1,0[$, Taylor expansion of $R(\cdot,x)$ around x yields,

$$R(s,x) = R(x-sh,x) - shR^{(1,0)}(x^+,x) + o(h).$$

Similarly for $s \in]0,1[$ we obtain,

$$R(x-sh,x) = R(x,x) - shR^{(1,0)}(x^-,x) + o(h).$$

Thus,

$$B_{n,2}(x) = R(x,x) - hR^{(1,0)}(x^+,x)\int_{-1}^{0} s\ k(s)\ ds - hR^{(1,0)}(x^-,x)\int_{0}^{1} s\ k(s)\ ds + o(h) + O\Big(\frac{1}{nh}\Big).$$

Hence,

$$\lim_{n\to\infty} B_{n,2}(x) = R(x,x). \tag{2.93}$$

Similarly,

$$\lim_{n\to\infty} B_{n,3}(x) = R(x,x). \tag{2.94}$$

It is easy to verify that,

$$\lim_{n\to\infty} B_{n,4}(x) = R(x,x). \tag{2.95}$$

Inserting (2.92), (2.93), (2.94) and (2.95) in (2.91) yields,

$$\lim_{n,m\to\infty} \mathbb{E}(A^2_{m,n}(x)) = 0.$$

This concludes the proof of Theorem 2.4.3. □

Proof of Theorem 2.5.2.

Let $x \in]0,1[$. On the one hand, we have from Proposition 2.4.6 and Remark 2.4.3,

$$\text{Var}\ \hat{g}_n^{pro}(x) = \frac{\sigma^2_{x,h}}{m} - \frac{A}{12mn^2h}\frac{\alpha(x)}{f^2(x)} + O\Big(\frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big), \tag{2.96}$$

where $A = \int_{-1}^{1} K^2(t)\ dt$. On the other hand, it can be seen from the proof of Proposition 1.3.2 that,

$$\text{Var}\ \hat{g}_n^{GM}(x) = \frac{\sigma^2_{x,h}}{m} + O\Big(\frac{1}{mn^2} + \frac{1}{mn^3h^2}\Big). \tag{2.97}$$

Equations (2.96) and (2.97) then yield,

$$mn^2h\ \Big(\text{Var}\ \hat{g}_n^{GM} - \text{Var}\ \hat{g}_n^{pro}\Big) = \frac{A}{12}\frac{\alpha(x)}{f^2(x)} + O\Big(h + \frac{1}{nh}\Big).$$

Recall that $\alpha(x) > 0$ and that $\frac{1}{f(x)} > 0$. Since $h \to 0$ and $nh \to \infty$ as $n,m \to \infty$ we obtain,

$$\lim_{n,m\to\infty} mn^2h\ \Big(\text{Var}\ \hat{g}_n^{GM}(x) - \text{Var}\ \hat{g}_n^{pro}(x)\Big) = \frac{A}{12}\frac{\alpha(x)}{f^2(x)} > 0.$$

This concludes the proof of Theorem 2.5.2. □

Proof of Theorem 2.5.3.

We have from the proof of Proposition 2.4.2 (Equation (2.53)) for any $x \in]0,1[$,

$$\mathbb{E}(\hat{g}^{pro}_{n,m}(x)) - g(x) = I_h(x) - g(x) + O\Big(\frac{1}{n^2h}\Big), \tag{2.98}$$

where,

$$I_h(x) = \int_{x-h}^{x+h} \varphi_{x,h}(s) g(s)\ ds.$$

Hence, using (2.96) and (2.98) we get for a positive density measure w,

$$\begin{aligned}\mathrm{IMSE}(\hat{g}^{pro}_n) = \frac{1}{m}\int_0^1 \sigma^2_{x,h}\ w(x)\ dx - \frac{A}{12mn^2h}\int_0^1 \frac{\alpha(x)}{f^2(x)}\ w(x)\ dx + \int_0^1 \big(I_h(x) - g(x)\big)^2\ w(x)\ dx \\ + O\Big(\frac{1}{n^4h^2} + \frac{h}{n^2} + \frac{1}{mn^3h^2} + \frac{1}{mn^2}\Big).\end{aligned} \tag{2.99}$$

It can be seen from Proposition 1.3.1 that,

$$\mathbb{E}(\hat{g}^{GM}_{n,m}(x)) - g(x) = I_h(x) - g(x) + O\Big(\frac{1}{n^2h}\Big). \tag{2.100}$$

Using (2.97) and (2.100) yield,

$$\begin{aligned}\mathrm{IMSE}(\hat{g}^{GM}_n) = \frac{1}{m}\int_0^1 \sigma^2_{x,h}\ w(x)\ dx + \int_0^1 \big(I_h(x) - g(x)\big)^2\ w(x)\ dx \\ + O\Big(\frac{1}{n^4h^2} + \frac{h}{n^2} + \frac{1}{mn^2} + \frac{1}{mn^3h^2}\Big).\end{aligned} \tag{2.101}$$

Then, Equations (2.99) and (2.101) yield,

$$mn^2h\ \Big(\mathrm{IMSE}\ (\hat{g}^{GM}_n) - \mathrm{IMSE}\ (\hat{g}^{pro}_n)\Big) = \frac{A}{12}\int_0^1 \frac{\alpha(x)}{f^2(x)}\ w(x)\ dx + O\Big(\frac{m}{n^2h} + mh^2 + h + \frac{1}{nh}\Big).$$

Since $\frac{m}{n} = O(1)$ and $mh^2 \to 0$ as $n, m \to \infty$ we obtain,

$$\lim_{n,m\to\infty} mn^2h\ \Big(\mathrm{IMSE}\ (\hat{g}^{GM}_n) - \mathrm{IMSE}\ (\hat{g}^{pro}_n)\Big) = \frac{A}{12}\int_0^1 \frac{\alpha(x)}{f^2(x)}\ w(x)\ dx > 0.$$

This concludes the proof of Theorem 2.5.3. $\square$

2.8 Appendix

Supplementary facts

(F1) Let f be defined by (2.4). We shall prove that if $g \in V_{T_n}$, i.e., if $g(\cdot) = \sum_{j=1}^n a_j R(t_j, \cdot)$ for some $a_i \in \mathbb{R}$, then

$$||f - g||^2 = \int_0^1 \varphi(s)(f(s) - g(s))\ ds - \sum_{i=1}^n a_i(f(t_i) - g(t_i)).$$

In fact,

$$||f - g||^2 = \langle f - g, f - g\rangle = \langle f, f - g\rangle - \langle g, f - g\rangle$$

On the one hand, note that $f - g \in \mathcal{F}(\varepsilon)$ and by using (2.3) we obtain,

$$\langle g, f - g\rangle = \sum_{i=1}^{n} a_i \langle R(t_i, \cdot), f - g\rangle = \sum_{i=1}^{n} a_i (f(t_i) - g(t_i)).$$

On the another hand, Lemma 2.2.1 and its proof yield that $f(\cdot) = \mathbb{E}(X\varepsilon(\cdot))$ where $X \in L_2(\varepsilon)$ and that,

$$\lim_{l\to\infty} \mathbb{E}(X_l - X)^2 = 0 \quad \text{where} \quad X_l = \sum_{j=1}^{l-1} (x_{j+1,l} - x_{j,l})\varphi_{x,h}(x_{j,l})\varepsilon(x_{j,l}),$$

where $(x_{j,l})_{j=1,\cdots,l}$ is a suitable partition of $[0, 1]$. Let $F_l(\cdot) = \mathbb{E}(X_l\varepsilon(\cdot))$ which is an element of $\mathcal{F}(\varepsilon)$. Clearly,

$$\langle f, f - g\rangle = \langle f - F_l, f - g\rangle + \langle F_l, f - g\rangle.$$

We have,

$$|\langle f - F_l, f - g\rangle| \le ||f - F_l||\ ||f - g|| \le \sqrt{\mathbb{E}((X_l - X)^2)}||f - g||.$$

Thus $\lim\limits_{l\to\infty} \langle f - F_l, f - g\rangle = 0$. In addition,

$$\begin{aligned}\langle F_l, f - g\rangle &= \Big\langle \sum_{j=1}^{l-1} (x_{j+1,l} - x_{j,l})\varphi(x_{j,l})R(x_{j,l}, \cdot), f - g \Big\rangle \\ &= \sum_{j=1}^{l-1} (x_{j+1,l} - x_{j,l})\varphi(x_{j,l})\langle R(x_{j,l}, \cdot), f - g\rangle = \sum_{j=1}^{l-1} (x_{j+1,l} - x_{j,l})\varphi(x_{j,l})(f(x_{j,l}) - g(x_{j,l})).\end{aligned}$$

Hence,

$$\lim_{l\to\infty} \langle F_l, f - g\rangle = \int_0^1 \varphi(t)(f(t) - g(t))\, dt.$$

Finally,

$$\langle f, f - g\rangle = \int_0^1 \varphi(t)(f(t) - g(t))\, dt. \ \square$$

(F2) For $x \in [0, 1]$, let $f_{x,h}$ be defined by (2.2). We shall prove that,

$$m\text{Var}(\hat{g}_n^{pro}(x)) = ||P_{|T_n} f_{x,h}||^2.$$

In fact, by the definition of the projection operator $P_{|T_n}$, we have $P_{|T_n} f_{x,h} \in V_{T_n}$ and for $t \in [0, 1]$,

$$P_{|T_n} f_{x,h}(t) = \sum_{i=1}^{n} a_i R(t_i, t) = \mathbb{E}(\sum_{i=1}^{n} a_i \epsilon(t_i)\epsilon(t)) \quad \text{for some} \quad a_i \in \mathbb{R} \text{ for } i = 1, \cdots, n,$$

and then,

$$||P_{|T_n} f_{x,h}||^2 = \mathbb{E}\left(\sum_{i=1}^{n} a_i \epsilon(t_i)\right)^2 = \sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j R(t_i, t_j) = \sum_{i=1}^{n} a_i P_{|T_n} f_{x,h}(t_i).$$

Recall that $m_{x,h}{'}_{|T_n} = f_{x,h}{'}_{|T_n} R_{|T_n}$ and using (2.6) we obtain,

$$P_{|T_n} f_{x,h}(t_i) = f_{x,h}(t_i) = \sum_{j=1}^{n} m_{x,h}(t_j) R(t_i, t_j). \tag{2.102}$$

We have then, using (2.102),

$$\begin{aligned}||P_{|T_n} f_{x,h}||^2 &= \sum_{i=1}^{n} a_i \sum_{j=1}^{n} m_{x,h}(t_j) R(t_i, t_j) = \sum_{j=1}^{n} m_{x,h}(t_j) \sum_{i=1}^{n} a_i R(t_i, t_j) \\ &= \sum_{j=1}^{n} m_{x,h}(t_j) \sum_{i=1}^{n} m_{x,h}(t_i) R(t_i, t_j) = m \mathrm{Var}(\hat{g}_n^{pro}(x)). \ \square\end{aligned}$$

(F3) We shall now prove prove that every function in $\mathcal{F}(\varepsilon)$ is continuous on $[0,1]$. In fact let $g \in \mathcal{F}(\varepsilon)$, i.e.,

$$g(\cdot) = \mathbb{E}(U\varepsilon(\cdot)) \quad \text{for some} \quad U \in L_2(\varepsilon).$$

For $s, t \in [0,1]$, (2.3) and Cauchy-Swartz inequality yields,

$$\begin{aligned}|g(t) - g(s)| &= |\langle R(\cdot,t), g\rangle - \langle R(\cdot,s), g\rangle| = |\langle R(\cdot,t) - R(\cdot,s), g\rangle| \\ &\leq ||R(\cdot,t) - R(\cdot,s)|| \ ||g|| = ||R(\cdot,t) - R(\cdot,s)|| \ \sqrt{\mathbb{E}(U^2)}.\end{aligned}$$

Since ε is of second order process then $\mathbb{E}(U^2) < \infty$ and since R is continuous on $[0,1]^2$ we obtain,

$$\lim_{s \to t} ||R(\cdot,t) - R(\cdot,s)||^2 = \lim_{s \to t} (R(t,t) + R(s,s) - 2R(s,t)) = 0,$$

which yields that $\lim_{s \to t} |g(t) - g(s)| = 0$. Hence g is continuous. $\square$

(F4) Suppose that R verifies Assumptions $(A), (B)$ and (C). Let f be defined by (2.4). We shall prove that if $g \in V_{T_n}$, i.e., $g(\cdot) = \sum_{j=1}^{n} a_j R(t_j, \cdot)$ with $(a_i)_i \in \mathbb{R}$ then,

$$f''(t) - g''(t^+) = -\alpha(t)\varphi(t) + \langle R^{(0,2)}(\cdot, t^+), f - g\rangle.$$

In fact, we have, as in Equation (2.39),

$$f''(t) = -\alpha(t)\varphi(t) + \int_0^1 R^{(0,2)}(s, t^+)\varphi(s)\, ds.$$

In addition, we have clearly

$$g''(t^+) = \sum_{j=1}^{n} a_j R^{(0,2)}(t_j, t^+).$$

Thus,

$$f''(t) - g''(t^+) = -\alpha(t)\varphi(t) + \int_0^1 R^{(0,2)}(s,t^+)\varphi(s)\, ds - \sum_{j=1}^n a_j R^{(0,2)}(t_j,t^+).$$

We have,

$$\langle R^{(0,2)}(\cdot,t^+), f-g\rangle = \langle R^{(0,2)}(\cdot,t^+), f\rangle - \langle R^{(0,2)}(\cdot,t^+), g\rangle$$

On the one hand, since by Assumption (C), $R^{(0,2)}(\cdot,t^+)$ is in $\mathcal{F}(\varepsilon)$ then (2.3) yields,

$$\langle R^{(0,2)}(\cdot,t^+), g\rangle = \sum_{j=1}^n a_j \langle R^{(0,2)}(\cdot,t^+), R(\cdot,t_j)\rangle = \sum_{j=1}^n a_j R^{(0,2)}(t_j,t^+). \tag{2.103}$$

On the other hand, from Lemma 2.2.1 we have $f(\cdot) = \mathbb{E}(X\varepsilon(\cdot))$ where $X \in L_2(\varepsilon)$ and,

$$\lim_{l\to\infty} \mathbb{E}(X_l - X)^2 = 0 \quad \text{with} \quad X_l = \sum_{j=1}^{l-1}(x_{j+1,l} - x_{j,l})\varphi(x_{j,l})\varepsilon(x_{j,l}),$$

where $(x_{j,l})_{j=1,\cdots,l}$ is a suitable partition of $[0,1]$. Let $F_l(\cdot) = \mathbb{E}(X_l\varepsilon(\cdot)) \in \mathcal{F}(\varepsilon)$, we have,

$$\langle R^{(0,2)}(\cdot,t^+), f\rangle = \langle R^{(0,2)}(\cdot,t^+), f - F_l\rangle + \langle R^{(0,2)}(\cdot,t^+), F_l\rangle, \tag{2.104}$$

and,

$$|\langle R^{(0,2)}(\cdot,t^+), f - F_l\rangle| \le ||R^{(0,2)}(\cdot,t^+)||\ ||f - F_l|| = ||R^{(0,2)}(\cdot,t^+)||\sqrt{\mathbb{E}((X_l - X)^2)}.$$

The last bound together with Assumption (C) gives $\lim_{l\to\infty} |\langle R^{(0,2)}(\cdot,t^+), f - F_l\rangle| = 0$, in addition,

$$\begin{aligned}\langle R^{(0,2)}(\cdot,t^+), F_l\rangle &= \sum_{j=1}^{l-1}(x_{j+1,l} - x_{j,l})\varphi(x_{j,l})\langle R^{(0,2)}(\cdot,t^+), \varepsilon(x_{j,l})\rangle \\ &= \sum_{j=1}^{l-1}(x_{j+1,l} - x_{j,l})\varphi(x_{j,l})R^{(0,2)}(x_{j,l},t^+).\end{aligned}$$

Thus,

$$\lim_{l\to\infty}\langle R^{(0,2)}(\cdot,t^+), F_l\rangle = \int_0^1 \varphi(s)R^{(0,2)}(s,t^+)\, ds. \tag{2.105}$$

Finally, using (2.103), (2.104) and (2.105) yield,

$$\langle R^{(0,2)}(\cdot,t^+), f-g\rangle = \int_0^1 \varphi(s)R^{(0,2)}(s,t^+)\, ds - \sum_{j=1}^n a_j R^{(0,2)}(t_j,t^+).\ \square$$

Chapter 3

Trapezoidal rule and sampling designs for the nonparametric estimation of the regression function in models with correlated errors

Abstract: The problem of estimating the regression function in a fixed design models with correlated observations is considered. Such observations are obtained from several experimental units, each of them forms a time series. Based on the trapezoidal rule, we propose a simple kernel estimator and we derive the asymptotic expression of its integrated mean squared error IMSE and its asymptotic normality. The problems of the optimal bandwidth and the optimal design with respect to the asymptotic IMSE are also investigated. Finally, a simulation study is conducted to study the performance of the new estimator and to compare it with the classical estimator of Gasser and Müller in a finite sample set. In addition, we study the robustness of the optimal design with respect to the misspecification of the autocovariance function.

Key words: *Nonparametric regression, optimal design, autocovariance function, trapezoidal rule, asymptotic normality.*

Résumé: Le problème d'estimation de la fonction de régression est considéré, dans un modèle avec des erreurs corrélées. Les observations sont obtenues à partir de plusieurs unités expérimentales, chacune forme une série temporelle. Nous proposons un nouvel estimateur à noyau, en se basant sur la règle des trapèzes. Nous étudions son comportement asymptotique et nous montrons sa normalité asymptotique. Le problème de la fenêtre optimal et l'échantillonnage optimal sont investigués dans un contexte asymptotique. Finalement, nous affections une étude de simulation afin de tester la performance de l'estimateur proposé, pour des petites tailles d'échantillonnage. En outre, nous étudions la robustesse de l'échantillonnage optimale par rapport à la mauvaise spécification de la fonction d'autocovariance.

Mot clés: *Régression non parametrique, plans d'échantillonnage optimale, fonction d'autocovariance, règle des trapèzes, normalité asymptotique.*

3.1 Introduction

A classical problem in Statistics is the nonparametric estimation of the regression function of a response variable Y given an explanatory variable X, i.e, estimating the function g defined by $g(t) = \mathbb{E}(Y|X = t)$, based on the observations of $(X_i, Y_i)_{1 \leq i \leq n}$ which are copies of (X, Y). These observations are often modeled as follows: $Y_i = g(t_i) + \varepsilon_i$ where g is the unknown regression function to be estimated, the $\{t_i, i = 1, \cdots, n\}$ is the sampling design and $\{\varepsilon_i, i = 1, \cdots, n\}$ are centered errors. Typically when $(\varepsilon_i)_i$ are i.i.d. the estimation of g has been extensively investigated by several authors. We mention, among others, the work of Priestly and Chao (1972), Benedetti (1977) and Gasser and Müller (1979, 1984). However, considering that the observations are independent is not always a realistic assumption. In pharmacokinetics for instance, one wishes to estimate the concentration-time of some injected medicine in the organism, based on the observation of blood tests over a period of time. It is clear that the observations provided from the same individual are correlated. For this reason, we shall investigate in this chapter the nonparametric regression estimation problem where the observations are correlated.
We consider the so-called fixed design regression model with repeated measurements, i.e.,

$$Y_j(t_i) = g(t_i) + \varepsilon_j(t_i) \quad \text{for } i = 1, \cdots, n \text{ and } j = 1, \cdots, m, \tag{3.1}$$

where $\{\varepsilon_j, j = 1, \cdots, m\}$ is a sequence of i.i.d. centered error processes with the same distribution as a process ε. Such models are well known in growth curve analysis and in dose response curves. They can be obtained, as noted by Azzalini (1984), from m individual being observed on a period of time. Generally, observations between different individuals will be uncorrelated. Hence, it is of interest to relax the assumption of correlation between the experimental units.

Müller (1984) considered Model (3.1) for $m = 1$ (observations on one experimental unit) and he supposed that, for $s \neq t$, the covariance $\mathrm{Cov}(\varepsilon_j(t), \varepsilon_j(s))$ tends to 0 as n tends to infinity, which is not a realistic assumption, as indicated by Hart and Wherly (1986), in the growth curve problems. They investigated the estimation of g in Model (3.1) with a stationary error process. They used the estimator proposed by Gasser and Müller (1979), and they showed that, in order to obtain the consistency of the kernel estimator in the presence of correlations, it is necessary to take m experimental units and to let m tends to infinity.

The stationarity assumption is however restrictive, for instance, in the previous pharmacokinetics example, it is clear that the concentration of the medicine will be high at the beginning then decreases with time. For this, we shall investigate the estimation of g in Model (3.1) where ε is a nonstationary error process. This case was partially investigated by Ferreira *et al.* (1997) and Benhenni and Rachdi (2007), where the Gasser and Müller estimator was used.

In this chapter, we propose a new estimator for the regression function g as an approximation of the kernel estimator based on continuous observations in the whole interval $[0, 1]$ constructed through a stochastic integral. See, for instance, Blanke and Bosq (2008), Didi and Louani (2013). When only discrete observations are available, we use the "best" approximation of the stochastic integral, which is obtained by using the trapezoidal rule based on discrete observations at appropriate n sampling points generated by a sampling density in the interval $[0, 1]$.

This estimator has a relatively simpler expression than the kernel estimator proposed by Gasser and Müller (1979). Moreover, since this last one depends on n integrals of a kernel at middle samples; and may be subject to numerical (computational) instability, for instance when a Gaussian kernel is used, whereas the proposed estimator depends only on the observations and the values of the kernel at the sampling points.

In addition to its simple expression, the proposed estimator allows to bring an answer to another important and open statistical problem under correlated errors, which is the optimal design problem. For instance, in the previous pharmacokinetic example, one wishes to find the best moments for the blood testing to be made in order to have a better estimate of the concentration curve.

The optimal design problem has been extensively studied in parametric regression. We mention the work of Sacks and Ylvisaker (1966), Belouni and Benhenni (2015) and more recently Dette *et al.* (2016) among others. In the nonparametric case, Müller (1984) introduced the optimal design points when the errors are asymptotically independent. He used a regular design sequence generated by a density function f, i.e, $t_i = F^{-1}(\frac{i}{n})$, where F is the distribution function associated to f. He derived the optimal design generated by a density that minimizes the asymptotic Integrated Mean Squared Error (IMSE). To the best of our knowledge, there exists no result concerning the problem of optimal design for nonparametric regression estimation in models under more general class of error processes.

We also investigate the problem of the asymptotic optimal bandwidth. We mention, for the nonparametric case, the work of Hart and Wherly (1986) and Benhenni and Rachdi (2007). For results on the break down of some data based methods for bandwidth selection in the presence of correlation, for instance the cross validation, and other alternative methods, the reader is referred to Chiu (1989), Altman (1990), Hart (1991, 1994) among others.

This chapter is organized as follows. In Section 2, we present the new estimator of the regression function g in Model (3.1) where ε is a centered error process. In Section 3, we give the asymptotic expressions of the bias, the variance and the IMSE. We then derive the asymptotic optimal bandwidth with respect to the asymptotic IMSE. In addition, we obtain the optimal design density with respect to the asymptotic IMSE, and we prove that it is minimax optimal. We also prove the asymptotic normality of the proposed estimator. In Section 4, we conduct a simulation study to investigate the performance of the new estimator and then to compare it with that of Gasser and Müller. We also conducted a study to compare the uniform and the optimal sampling designs, and to study the robustness of the optimal design, with respect to the misspecification of the autocovariance function. Since the classical cross validation criteria turned out to be inefficient in the presence of correlation, we use the bandwidth that minimizes the exact IMSE, the comparison is performed for different numbers of experimental units and design points. Finally, Section 5 is dedicated to the proofs of our theoretical results.

3.2 Model and estimator

We consider m experimental units, each of them having n different measurements of the response (say $0 \le t_1 < t_2 < \cdots < t_n \le 1$). The so-called fixed design regression model is defined as follows:

$$Y_j(t_i) = g(t_i) + \varepsilon_j(t_i) \text{ where } j = 1, \ldots, m \text{ and } i = 1, \ldots, n, \tag{3.2}$$

where g is the unknown regression function on $[0, 1]$ and $\{\varepsilon_j(t), t \in [0, 1]\}_j$ is a sequence of error processes.

We assume that $g \in C^2([0, 1])$ and that $(\varepsilon_j)_j$ are i.i.d. processes with the same distribution as a centered second order process ε. We denote by R its autocovariance function.

3.2.1 Simple estimator and sampling design

In order to motivate the construction of our new estimator, we consider the regression model using m continuous experimental units, i.e,

$$Y_j(t) = g(t) + \varepsilon_j(t) \text{ for } t \in [0, 1] \text{ and } j = 1, \cdots, m. \tag{3.3}$$

A continuous kernel estimator of g in Model (3.3) is given for any $x \in [0,1]$ by,

$$\hat{g}_{[0,1]}(x) = \int_0^1 \varphi_{x,h}(t)\overline{Y}(t)\ dt \quad \text{with} \quad \overline{Y}(t) = \frac{1}{m}\sum_{j=1}^{m} Y_j(t), \tag{3.4}$$

where $\varphi_{x,h}(t) = \frac{1}{h}K\left(\frac{x-t}{h}\right)$ for a kernel K and a bandwidth h. For details on the Kernel estimation of the regression function based on continuous observations see, for instance, Blanke and Bosq (2008) or Didi and Louani (2013).

In the practical case where we only have access to discrete observations, we apply the trapezoidal rule to approximate the continuous Kernel estimator given by (3.4). We construct then a new simple estimator of the regression function that we shall call the trapezoidal estimator.

Before introducing the proposed estimator, we begin with defining a sequence of designs which will be used in its construction. This class of designs was considered by Sacks and Ylvisaker (1970).

Definition 3.2.1 *Let F be a distribution function of some density f satisfying $\inf_{t\in[0,1]} f(t) > 0$ and $\sup_{t\in[0,1]} f(t) < \infty$. The so-called regular sequence of designs generated by a density f is defined by,*

$$T_n = \left\{ t_{i,n} = F^{-1}\left(\frac{i}{n}\right),\ i = 1,\ldots,n. \right\} \text{ for } n \geq 1.$$

Such a sequence of designs verifies the next useful lemma.

Lemma 3.2.1 *For $n \geq 1$ let $T_n = (t_{i,n})_{i=1,\cdots,n}$ be a regular sequence of designs generated by some density function. Let $x \in]0,1[$, $h > 0$ and note by $N_{T_n} \triangleq \text{Card } (T_n \cap [x-h, x+h])$. Suppose that $N_{T_n} \neq 0$ and that $nh \geq 1$. Then,*

$$\sup_{0\leq j\leq n} (t_{j+1,n} - t_{j,n}) = O\left(\frac{1}{n}\right) \quad \text{and} \quad N_{T_n} = O(nh). \tag{3.5}$$

We shall now give the definition of the trapezoidal estimator, obtained from a discrete approximation of the continuous estimator $\hat{g}_{[0,1]}$ given by (3.4).

Definition 3.2.2 *The trapezoidal estimator of the regression function g based on the observations $(t_{i,n}, Y_j(t_{i,n}))_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$, where $T_n = (t_{i,n})_{1\leq i\leq n}$ is a regular sequence of designs generated by a density function f of support intersecting $[x-h, x+h]$ is given, for any $x \in [0,1]$, by,*

$$\hat{g}_n^{trap}(x) = \frac{1}{2n}\sum_{k=1}^{N_{T_n}-1} \left\{ \left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k}) + \left(\frac{\varphi_{x,h}}{f}\overline{Y}\right)(t_{x,k+1}) \right\}, \tag{3.6}$$

where $t_{x,1} < \cdots < t_{x,N_{T_n}}$ are the points of T_n in $[x-h, x+h]$, $\varphi_{x,h}(t) = \frac{1}{h}K\left(\frac{x-t}{h}\right)$, $\overline{Y}$ is given in (3.4), K is a kernel of support $[-1,1]$ and $h = h(n,m)$ is a bandwidth with $0 < h < 1$.

In order to derive our asymptotic results, the following assumptions on the autocovariance function R and the kernel K are required.

3.2.2 Assumptions

(A) The autocovariance function R exists and is continuous on the square $[0,1]^2$.

(B) At the diagonal (when $t = s$ in the unit square), R has continuous left and right first-order derivatives, that is:

$$R^{(0,1)}(t,t^-) = \lim_{s \uparrow t} \frac{\partial R(t,s)}{\partial s} \quad \text{and} \quad R^{(0,1)}(t,t^+) = \lim_{s \downarrow t} \frac{\partial R(t,s)}{\partial s}.$$

The jump function along the diagonal $\alpha(t) \triangleq R^{(0,1)}(t,t^-) - R^{(0,1)}(t,t^+)$ is assumed to be continuous and not identically equal to zero.

(C) Off the diagonal (when $t \neq s$ in the unit square), R is assumed to have continuous mixed partial derivatives up to order two and,

$$A^{(i,j)} \triangleq \sup_{0 \le t \neq s \le 1} |R^{(i,j)}(t,s)| < \infty \text{ for } i,j \text{ such that } 0 \le i+j \le 2.$$

(D) The Kernel K is even at least in $C^2([-1,1])$ and K'' is Lipschitz on [-1,1].

Examples of processes with autocovariances satisfying Assumptions $(A), (B)$ and (C) are given as follows.

Example 3.2.1

1. *The Wiener process with autocovariance function $R(s,t) = \sigma^2 \min(s,t)$, has a constant jump function $\alpha(t) = \sigma^2$ and $R^{(i,j)}(s,t) = 0$ for all i,j such that $i+j=2$ and $s \neq t$.*

2. *The Ornstein-Uhlenbeck process with a stationary autcovariance $R(s,t) = \sigma^2 \exp(-\lambda|s-t|)$ for $\sigma > 0$ and $\lambda > 0$. For this process $\alpha(t) = 2\sigma^2\lambda$ and $R^{(0,2)}(s,t) = \sigma^2\lambda^2 \exp(-\lambda|s-t|)$.*

3. *A generalization of the Ornstein-Uhlenbeck process to a process with a nonstationary autocovariance function of the form: $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ for $\sigma > 0$, $\lambda > 0$ and $0 < \rho < 1$. For this process the jump function, which is not constant when $\lambda \neq 1$, is given by $\alpha(t) = -2\sigma^2 ln(\rho) t^{\lambda-1}$.*

4. *Sacks and Ylvisaker (1966) gave another general class of convex stationary autcovariance functions of the form,*

$$R(s,t) = \int_0^{1/|t-s|} (1 - \mu|t-s|) p(\mu)\, d\mu,$$

where p is a probability density and p' its derivative are such that,

$$\lim_{\mu \to \infty} \mu^3 p(\mu) < \infty, \quad \text{and} \quad \int_a^\infty (\mu p'(\mu) + 3p(\mu))^2)\mu^6 d\mu < \infty,$$

for some finite constant a. For this autocovariance function, $\alpha(t) = 2\int_0^\infty \mu p(\mu)\, d\mu$ for all t.

The following kernels satisfy Assumption (D).

Example 3.2.2

1. *The Quadratic kernel defined by $K(u) = \frac{15}{16}(1-u^2)^2\, \mathbb{1}_{\{|u| \le 1\}}$.*

2. *The Triweight kernel defined by $K(u) = \frac{35}{32}(1-u^2)^3\, \mathbb{1}_{\{|u| \le 1\}}$.*

3.3 Asymptotic results

The following propositions give the asymptotic expressions of the bias and the variance of the trapezoidal estimator as defined by (3.6).

Proposition 3.3.1 *Suppose that Assumption* (D) *is satisfied. Moreover assume that* $f \in C^2([0,1])$ *and* f'', g'' *are Lipschitz functions on* $[0,1]$. *If* $\lim_{n\to\infty} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$ *then for any* $x \in]0,1[$,

$$\text{Bias}(\hat{g}_n^{trap}(x)) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{1}{n^3h^3}\Big),$$

where $B = \int_{-1}^{1} t^2 K(t)\ dt$.

Proposition 3.3.2 *Suppose that Assumptions* $(A), (B), (C)$ *and* (D) *are satisfied. Moreover assume that* $f \in C^2([0,1])$ *and for any* $t \in [0,1]$, f'' *and* $R^{(0,2)}(t,.)$ *are all Lipschitz on* $[0,1]$. *If* $\lim_{n\to\infty} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$ *then for any* $x \in]0,1[$,

$$\begin{aligned}\text{Var}(\hat{g}_n^{trap}(x)) = &\frac{1}{m}\Big(R(x,x) - \frac{h}{2}C_K\alpha(x)\Big) + \frac{V}{12mn^2h}\frac{\alpha(x)}{f^2(x)} \\ &+ o\Big(\frac{h}{m}\Big) + O\Big(\frac{1}{mn^2} + \frac{1}{mn^3h^3}\Big),\end{aligned}$$

where $V = \int_{-1}^{1} K^2(t)\ dt$ *and* $C_K = \int_{-1}^{1}\int_{-1}^{1} |u-v|K(u)K(v)dudv$.

Propositions 3.3.1 and 3.3.2 allow to derive the asymptotic expression of the mean squared error (MSE) of the Trapezoidal estimator (3.6). The integrated mean squared error (IMSE) is then obtained by integrating the MSE with respect to some weight function w. The results are announced, without proof, in the following theorem.

Theorem 3.3.1 *If all the assumptions of Propositions 3.3.1 and 3.3.2 are satisfied then for any* $x \in]0,1[$,

$$\begin{aligned}\text{MSE}(\hat{g}_n^{trap}(x)) = &\frac{1}{m}\Big(R(x,x) - \frac{h}{2}\alpha(x)C_K\Big) + \frac{V}{12mn^2h}\frac{\alpha(x)}{f^2(x)} + \frac{1}{4}h^4[g''(x)]^2B^2 \\ &+ o\big(h^4 + \frac{h}{m}\big) + O\Big(\frac{1}{n^3h} + \frac{1}{mn^3h^3} + \frac{1}{mn^2} + \frac{1}{n^6h^6}\Big),\end{aligned}$$

$$\begin{aligned}\text{IMSE}(\hat{g}_n^{trap}) = &\frac{1}{m}\int_0^1 \Big(R(x,x) - \frac{h}{2}\alpha(x)C_K\Big)w(x)\ dx + \frac{V}{12mn^2h}\int_0^1 \frac{\alpha(x)}{f^2(x)}w(x)\ dx \\ &+ \frac{1}{4}h^4B^2\int_0^1 [g''(x)]^2 w(x)\ dx + o\big(h^4 + \frac{h}{m}\big) \\ &+ O\Big(\frac{1}{n^3h} + \frac{1}{mn^3h^3} + \frac{1}{mn^2} + \frac{1}{n^6h^6}\Big), \end{aligned} \tag{3.7}$$

where w *is a continuous density function,* V, B *and* C_K *are given in Propositions 3.3.1, 3.3.2.*

The previous Theorem shows, the efficiency of the Trapezoidal estimator, since the IMSE tends to 0 when $m \to \infty$, $h \to 0$ and $nh \to \infty$ as $n \to \infty$.

The asymptotic optimal bandwidth is obtained by minimizing the asymptotic IMSE as given by the following proposition.

Proposition 3.3.3 (Optimal bandwidth) *Suppose that the assumptions of Theorem 3.3.1 are satisfied. Moreover assume that $\frac{m}{n} = O(1)$ as $n, m \to \infty$. Denote by IMSE(h) the IMSE of the trapezoidal estimator when the bandwidth h is used. Then the bandwidth,*

$$h^* = \left(\frac{C_K \int_0^1 \alpha(x) w(x)\, dx}{2B^2 \int_0^1 [g''(x)]^2 w(x)\, dx} \right)^{1/3} m^{-1/3}, \tag{3.8}$$

is optimal in the sense that,

$$\overline{\lim_{n,m\to\infty}} \frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})} \le 1,$$

for any sequence of bandwidths $h_{n,m}$ verifying:

$$\lim_{n,m\to\infty} h_{n,m} = 0 \quad \textit{and} \quad \overline{\lim_{n,m\to\infty}} \; m h_{n,m}^3 < +\infty,$$

where B and C_K are given in Propositions 3.3.1 and 3.3.2.

We are interested now in finding the optimal design density, i.e, a function f^* according to the criteria $f^* \in \operatorname*{argmin}_f$ IMSE, where the minimum is taken with respect to the class of positive densities defined on $[0,1]$. In view of Theorem 3.3.1, the asymptotic optimal design density verifies,

$$f^* \in \operatorname*{argmin}_{f>0,\ \int_0^1 f(x)dx=1} \int_0^1 \frac{\alpha(x)}{f^2(x)} w(x)\, dx.$$

This optimization problem is solved in the following corollary.

Corollary 3.3.1 (Optimal design) *Suppose that the assumptions of Theorem 3.3.1 are satisfied. If $\lim_{n\to\infty} nh^2 = \infty$ and $\lim_{n,m\to\infty} \frac{n}{m} = \infty$, then the optimal sampling density with respect to the asymptotic* IMSE *is given by,*

$$f^*(t) = \frac{\{\alpha(t) w(t)\}^{1/3}}{\int_0^1 \{\alpha(s) w(s)\}^{1/3}\, ds} 1_{[0,1]}(t). \tag{3.9}$$

Let $\hat{g}_{n,f^*}^{trap}$ be the Trapezoidal estimator (3.6) with $f = f^*$ defined by (3.9). We have,

$$\begin{aligned}
\mathrm{IMSE}(\hat{g}_{n,f^*}^{trap}) = {} & \frac{1}{m} \int_0^1 \Big(R(x,x) - \frac{1}{2}\alpha(x) C_K h \Big) w(x)\, dx \\
& + \frac{V}{12mn^2h} \Big(\int_0^1 (\alpha(x) w(x))^{1/3}\, dx \Big)^3 + \frac{1}{4} h^4 B^2 \int_0^1 [g''(x)]^2 w(x)\, dx \\
& + o\big(h^4 + \frac{h}{m}\big) + O\Big(\frac{1}{n^3 h} + \frac{1}{mn^3h^3} + \frac{1}{mn^2} + \frac{1}{n^6h^6} \Big).
\end{aligned}$$

Remark 3.3.1 *Let $\hat{g}_{n,unif}^{trap}$ be the Trapezoidal estimator* (3.6) *with a uniform density, i.e, $f = f_{unif}$ the identity in $[0,1]$. The asymptotic IMSE of $\hat{g}_{n,unif}^{trap}$ is given by,*

$$\begin{aligned}
\mathrm{IMSE}(\hat{g}_{n,unif}^{trap}) = {} & \frac{1}{m} \int_0^1 \Big(R(x,x) - \frac{1}{2}\alpha(x) C_K h \Big) w(x)\, dx + \frac{V}{12mn^2h} \int_0^1 \alpha(x) w(x)\, dx \\
& + \frac{1}{4} h^4 B^2 \int_0^1 [g''(x)]^2 w(x)\, dx + o\big(h^4 + \frac{h}{m}\big) \\
& + O\Big(\frac{1}{n^3 h} + \frac{1}{mn^3h^3} + \frac{1}{mn^2} + \frac{1}{n^6h^6} \Big).
\end{aligned}$$

The reduction of the residual IMSE, $\overline{\text{IMSE}} \triangleq \text{IMSE} - \sigma^2_{x,h}/m$, *by using the asymptotic optimal design over the uniform design is then,*

$$rIMSE = \frac{\overline{\text{IMSE}}(\hat{g}^{trap}_{n,unif}) - \overline{\text{IMSE}}(\hat{g}^{trap}_{n,f^*})}{\overline{\text{IMSE}}(\hat{g}^{trap}_{n,unif})} \sim 1 - \frac{\left(\int_0^1 (\alpha(x)w(x))^{1/3}\, dx\right)^3}{\int_0^1 \alpha(x)w(x)\, dx}.$$

For instance, if $R(s,t) = st\min(s,t)$ *then* $\alpha(t) = t^2$. *Taking* $w \equiv 1$ *gives* $rIMSE \sim 35\%$.

Finally, the next theorem gives the asymptotic normality of the Trapezoidal estimator (3.6).

The optimal design, generated by the density function (3.9), may not be robust with respect to the missspecification of the autocovariance jump function α, and the weight function w. For this, we shall use a minimax criterion to obtain the optimal sampling design. Biedermann and Dette (2001) gave the following criterion, a density function f^* is said to be minimax optimal if,

$$f^* \in \underset{f>0, \int_0^1 f(t)dt=1}{\operatorname{argmin}} \; \max_{(\alpha,w)\in\Lambda} \Psi_{(\alpha,w)}(f), \tag{3.10}$$

where,

$$\Psi_{(\alpha,w)}(f) = \int_0^1 \frac{\alpha(t)}{f^2(t)} w(t)dt,$$

and,

$$\Lambda = \left\{(\alpha,w) \in (C[0,1])^2 \;/\; \int_0^1 \alpha(t)dt < \epsilon_1, \left(\int_0^1 w(s)^{1/2} ds\right)^2 < \epsilon_2\right\}.$$

The following theorem assures that the asymptotic optimal design density, defined in Corollary 3.3.1, is optimal in the sense of minimax.

Theorem 3.3.2 (Minimax optimality) *Suppose that the assumptions of Theorem 3.3.1 are satisfied. The function* f^* *given by* (3.9) *is optimal with respect to the minimax criterion* (3.10).

Finally, we conclude our theoretical results by the asymptotic normality of the trapezoidal estimator, presented in the following theorem.

Theorem 3.3.3 (Asymptotic normality) *Suppose that the assumptions of Theorem 3.3.1 are satisfied. If* $\lim_{m\to\infty} \sqrt{m}h^2 = 0$ *and* $\lim_{n\to\infty} nh^2 = \infty$ *then for any* $x \in]0,1[$,

$$\sqrt{m}\Big(\hat{g}^{trap}_n(x) - g(x)\Big) \xrightarrow{\mathscr{D}} Z, \quad \textit{with } Z \sim \mathcal{N}(0, R(x,x)),$$

where $\mathscr{D}$ *denotes the convergence in distribution and* $\mathcal{N}$ *is the normal distribution.*

3.4 Simulation study

3.4.1 Performance of the estimator

In this section, we investigate the performance of our estimator (3.6) in a finite sample set. We shall use the cubic growth curve, used by Hart and Wherly (1986) and Benhenni and Rachdi (2007),

$$g(x) = 10x^3 - 15x^4 + 6x^5 \quad \text{for} \quad 0 < x < 1. \tag{3.11}$$

This function was mainly used due to its similarity to the logistic function which is frequently found in growth curve analysis. The sampling points are taken to be:

$$t_i = (i - 0.5)/n \quad \text{for } i = 1, \cdots, n. \tag{3.12}$$

The error process ε is taken to be the Wiener error process with autocovariance function $R(s,t) = \sigma^2 \min(s,t)$. The Kernel used here is the quadratic kernel given by $K(u) = (15/16)(1-u^2)^2 I_{[-1,1]}(u)$. The bandwidth used in this study is the optimal bandwidth with respect to the exact IMSE.

We consider the mean of all estimators obtained from 100 simulations. We take $\sigma^2 = 0.5$ and simulations for other values of σ^2 gave similar results, they are given in Figure 3.1 for a fixed number of observations $n = 100$ and three different values of experimental units $m = 5, 20, 100$.

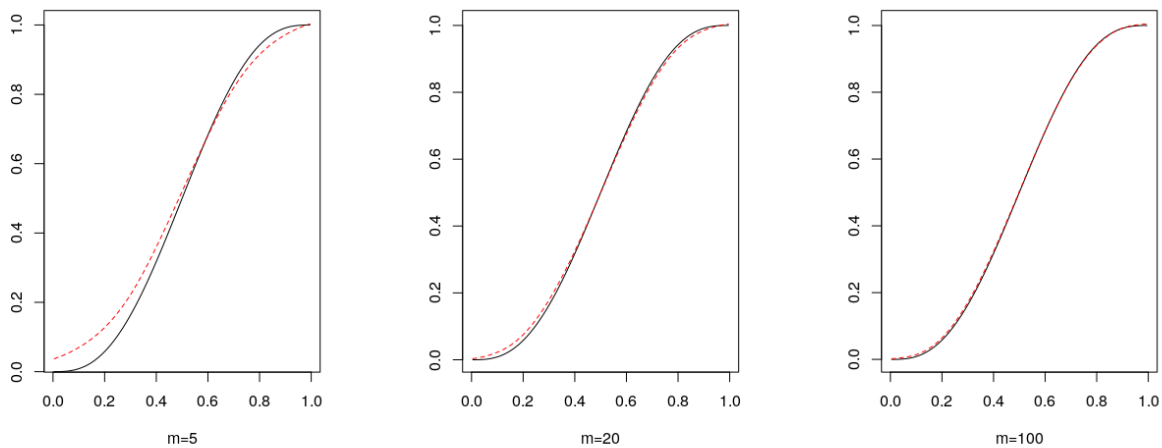


Figure 3.1: Cubic regression function is in plain line and the trapezoidal estimator is in dashed one.

It is clear that, the performance of the trapezoidal estimator gets better as m increases.

Our aim now is to compare the trapezoidal estimator to that of Gasser and Müller (1979) (referred by GM estimator), given for any $x \in]0,1[$ by,

$$\hat{g}_n^{GM}(x) = \sum_{i=1}^{n} \int_{m_{i-1}}^{m_i} \varphi_{x,h}(t)dt \; \overline{Y}(t_i), \tag{3.13}$$

where $m_0 = 0$, $m_n = 1$ and $m_i = (t_i + t_{i+1})/2$ for $i = 2, \cdots, n-1$, $\varphi_{x,h}(t) = (1/h)K((x-t)/h)$ and $\overline{Y}(t_i) = (1/m)\sum_{j=1}^{m} Y_j(t_i)$.

This comparison is conducted with respect to the non-asymptotic IMSE and under different types of correlation errors. We consider again the cubic regression function, the design given by (3.12) and the quadratic kernel. The two error processes considered here are the stationary Ornstein-Uhlenbeck process with $R(s,t) = \exp(-\lambda|s-t|)$, and the nonstationary Wiener process with $R(s,t) = \sigma^2 \min(s,t)$. We investigate various "amount" of correlation by taking different values of both σ^2 and λ.

We take the weight density w to be uniform on $[0,1]$, and we compare the optimal non-asymptotic IMSE of the two estimators, i.e., $\inf\limits_{0<h<1} \text{IMSE}(h)$. The bandwidth h is chosen over a grid from 0.09 to 0.5. The results are given in Tables 3.1-3.6 for $n = 30$ and for different values of m. The tables present the integrated bias squared denoted by $Ibias^2$, integrated variance denoted by $Ivar$ and the IMSE together with the optimal bandwidth associated to the smallest non-asymptotic IMSE for each estimator. The tables are organized according to the "degree" of correlation of the errors.

It can be seen that the optimal bandwidth is the same for both estimators, in addition, as expected, it decreases as m increases.

Consider first the case of strong correlated errors, i.e, for a large σ^2 and a small λ. In Table 3.1, for the Wiener process with $\sigma^2 = 1$, it appears that the G-M estimator has a slightly smaller $Ibias^2$ while the trapezoidal estimator has a slightly smaller $Ivar$ and since the $Ibias^2$ is too small compared to the $Ivar$ then the trapezoidal estimator has a slightly smaller IMSE. For the Ornstein-Uhlenbeck with $\lambda = 1$ (c.f. Table 3.2) it can be seen that the trapezoidal estimator has a slightly better performance because of a smaller IMSE, due to a smaller $Ibias^2$ and a smaller $Ivar$.

Consider now the case of moderate correlated errors. In Table 3.3 (for the Wiener process with $\sigma^2 = 0.5$) it seems that the G-M estimator has a slightly smaller $Ibias^2$ while the trapezoidal estimator has a slightly smaller $Ivar$ and smaller IMSE. While for the Ornstein-Uhlenbeck process with $\lambda = 25$, presented in table 3.4, the G-M estimator has slightly smaller IMSE due to a smaller $Ibias^2$ and a smaller $Ivar$.

Finally, consider the weakly correlated errors, i.e, for a small value of σ^2 and a large value of λ. In table 3.5, for the Wiener process with $\sigma^2 = 0.06$. it appears that the G-M estimator has a slightly smaller $Ibias^2$ while the trapezoidal estimator has a smaller $Ivar$ and smaller IMSE. However, for the Ornstein-Uhlenbeck process with $\lambda = 50$ (c.f. Table 3.6) the trapezoidal estimator has a slightly smaller $Ibias^2$ while the G-M estimator has a slightly smaller $Ivar$ and IMSE.

Overall, the two estimators, i.e, the trapezoidal estimator and the Gasser and Müller estimator, have "approximately" the same performance. Hence, the proposed estimator, which has a simpler expression, is as efficient as the classical Gasser and Müller estimator.

In all the previous cases, it appears that $Ibias^2$ is always smaller than $Ivar$. It should be noted here that, both of the estimators have boundary problems. A modified kernel at the edges, as suggested by Hart and Wherly (1986), was used in this simulation.

3.4.2 Optimal design

Another important aspect we looked at in this simulation study was the use of the asymptotic optimal design in a finite sample set. We consider the class of autocovariance functions introduced in Example 3.2.1 as follows:

$$R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}, \quad \sigma^2 > 0, \lambda > 0 \text{ and } \; 0 < \rho < 1,$$

for which the jump function $\alpha(t) = -2\sigma^2 ln(\rho) t^{\lambda-1}$. In particular when $\lambda = 1$ we obtain an Ornstein-Uhlenbeck stationary error process. In our numerical studies we will consider the nonstationary case, $\lambda \neq 1$. This class of nonstationary parametric autocovariance functions was introduced by Núñez-Antón and Woodworth (1994) to study the efficacy of cochlear implants. It was also used by several other authors, by Ferreira *et al.* (1997) who were interested in obtaining the optimal bandwidth for the Gasser and Müller estimator, by Ziemmerman *et al.* (1998), and then by Núñez-Antón (1997) to study the speech recognition data.

We compare, for $m \in \{5, 10, 20, 30\}$ and for instance $h = 0.123$, the non-asymptotic IMSE (taking $w \equiv 1$) of the trapezoidal estimator (3.6), using both the uniform design (3.12), i.e., $f \equiv 1$ and the optimal design generated by f^* given in (3.9), i.e.,

$$f^*_\lambda(t) = \frac{\lambda+2}{3}\, t^{(\lambda-1)/3}\, 1_{[0,1]}(t) \quad \text{and} \quad t^*_{\lambda,i} = \left(\frac{i}{n}\right)^{3/(\lambda+2)}.$$

Robustness of the optimal design

The optimal design depends on the autocovariance parameter λ, which is not known in practice, therefore we cannot use this design to compute the estimate of the regression function g. As an alternative, we can estimate first the autocovariance parameter λ, from the observations obtained following a uniform design, then we obtain the estimated optimal design defined as follows:

$$f^*_{\widehat{\lambda}}(t) = \frac{\widehat{\lambda}+2}{3}\, t^{(\widehat{\lambda}-1)/3}\, 1_{[0,1]}(t) \quad \text{and} \quad t^*_{\widehat{\lambda},i} = \left(\frac{i}{n}\right)^{3/(\widehat{\lambda}+2)}.$$

The estimation of the autocovariance parameters is obtained by minimizing the following criterion, as done for instance in Ferreira *et al.* (1997):

$$Q_{n,m}(\sigma^2, \lambda, \rho) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\widehat{R}(t_i, t_j) - R(t_i, t_j)\Big)^2, \tag{3.14}$$

where the empirical correlation estimator is given as follows:

$$\widehat{R}(t_i, t_j) = \frac{1}{m-1}\sum_{k=1}^{m}\big(Y_k(t_i) - \overline{Y}(t_i)\big)\big(Y_k(t_j) - \overline{Y}(t_j)\big) \quad \text{for } i, j = 1, \cdots, n.$$

From Amemiya (1985), as noted by Ferreira *et al.* (1997), it is known that the non linear least square estimator $(\widehat{\sigma^2}, \widehat{\lambda}, \widehat{\rho})$ is consistent.

In our simulation study, we fixed $\lambda = 4, \sigma^2 = 0.5$ and $\rho = 0.5$. To estimate $(\lambda, \sigma^2, \rho)$, we generated 100 matrices $(Y_j(t_i))_{\substack{1\le i\le n\\ 1\le j\le m}}$ of observations using the uniform design. For every matrix, we used the Generalized Simulated Annealing (GSA) algorithm to minimize the function (3.14), the estimation $(\widehat{\lambda}, \widehat{\sigma^2}, \widehat{\rho})$ is then the median of the 100 estimated values. For more details on the use of the software R algorithm function, see Xiang *et al.* (2013). This algorithm is essentially known for its ability to handle very complex non-linear objective functions with a very large number of optima.

The results are given in Tables 3.7-3.10, where the reduction in the IMSE by taking the optimal design instead of the uniform design is given by,

$$rIMSE_\lambda = \frac{\mathrm{IMSE}(\hat{g}^{trap}_{n,unif}) - \mathrm{IMSE}(\hat{g}^{trap}_{n,f^*_\lambda})}{\mathrm{IMSE}(\hat{g}^{trap}_{n,unif})},$$

and the reduction in the IMSE by taking the plug-in estimated optimal design instead of the uniform design is given by,

$$rIMSE_{\widehat{\lambda}} = \frac{\mathrm{IMSE}(\hat{g}^{trap}_{n,unif}) - \mathrm{IMSE}(\hat{g}^{trap}_{n,f^*_{\widehat{\lambda}}})}{\mathrm{IMSE}(\hat{g}^{trap}_{n,unif})}.$$

It can be seen in Tables 3.7-3.10 that there exists a reduction of the IMSE of the Trapezoidal estimator when using the optimal design, even for small values of the sampling size n and the number of experimental units m. Likewise, the estimated optimal design obtained by estimating the covariance parameter, still provides a reduction of the IMSE over the uniform design. This reduction is close to the one using the theoretical optimal design, this shows that the optimal design is robust when the covariance parameter has to be estimated.

Table 3.1: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Wiener error process with $\sigma^2 = 1$, for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	2.8832×10^{-3}	8.4967×10^{-2}	8.7850×10^{-2}	0.411
Trap		2.8833×10^{-3}	8.4959×10^{-2}	8.7843×10^{-2}	0.411
GM	15	1.04816×10^{-3}	2.9293×10^{-2}	3.0341×10^{-2}	0.322
Trap		1.04856×10^{-3}	2.9276×10^{-2}	3.0325×10^{-2}	0.322
GM	30	2.7691×10^{-4}	1.5169×10^{-2}	1.5446×10^{-2}	0.233
Trap		2.8535×10^{-4}	1.5124×10^{-2}	1.5409×10^{-2}	0.233

Table 3.2: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Ornstein-Uhlenbeck error process with $\lambda = 1$ for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	4.57002×10^{-3}	1.70570×10^{-1}	1.75140×10^{-1}	0.46
Trap		4.57001×10^{-3}	1.70565×10^{-1}	1.75135×10^{-1}	0.46
GM	15	1.31050×10^{-3}	5.8884×10^{-2}	6.0194×10^{-2}	0.34
Trap		1.30997×10^{-3}	5.8857×10^{-2}	6.0167×10^{-2}	0.34
GM	30	7.7889×10^{-4}	2.9818×10^{-2}	3.0597×10^{-2}	0.30
Trap		7.7828×10^{-4}	2.9791×10^{-2}	3.0569×10^{-2}	0.30

Table 3.3: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Wiener error process with $\sigma^2 = 0.5$ for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	1.0481×10^{-3}	4.3939×10^{-2}	4.4988×10^{-2}	0.322
$Trap$		1.0485×10^{-3}	4.3915×10^{-2}	4.4963×10^{-2}	0.322
GM	15	2.7691×10^{-4}	1.5169×10^{-2}	1.5446×10^{-2}	0.233
$Trap$		2.8535×10^{-4}	1.5124×10^{-2}	1.5409×10^{-2}	0.233
GM	30	1.1792×10^{-4}	7.7228×10^{-3}	7.8407×10^{-3}	0.188
$Trap$		1.4175×10^{-4}	7.6733×10^{-3}	7.8150×10^{-3}	0.188

Table 3.4: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Ornstein-Uhlenbeck error process with $\lambda = 25$ for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	4.3931×10^{-3}	2.7163×10^{-2}	3.1556×10^{-2}	0.455
$Trap$		4.3930×10^{-3}	2.7165×10^{-2}	3.1558×10^{-2}	0.455
GM	15	1.7942×10^{-3}	1.2819×10^{-2}	1.4613×10^{-2}	0.366
$Trap$		1.7935×10^{-3}	1.2824×10^{-2}	1.4618×10^{-2}	0.366
GM	30	1.0481×10^{-3}	7.0808×10^{-3}	8.1290×10^{-3}	0.322
$Trap$		1.0485×10^{-3}	7.0855×10^{-3}	8.1341×10^{-3}	0.322

Table 3.5: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Wiener error process with $\sigma^2 = 0.06$ for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	9.9714×10^{-5}	5.5781×10^{-3}	5.6778×10^{-3}	0.181
$Trap$		1.2841×10^{-4}	5.5373×10^{-3}	5.6657×10^{-3}	0.181
GM	15	9.9714×10^{-5}	4.6484×10^{-3}	4.7481×10^{-3}	0.181
$Trap$		1.2841×10^{-4}	4.6145×10^{-3}	4.7429×10^{-3}	0.181
GM	30	9.9714×10^{-4}	3.9844×10^{-3}	4.0841×10^{-3}	0.181
$Trap$		1.2841×10^{-4}	3.9552×10^{-3}	4.0836×10^{-3}	0.181

Table 3.6: The integrated squared bias, Integrated variance, IMSE and the optimal bandwidth in terms of m under the Ornstein-Uhlenbeck error process with $\lambda = 50$ for the GM and the trapezoidal estimator.

$n = 20$	m	$Ibias^2$	$Ivar$	IMSE	h_{opt}
GM	5	4.3496×10^{-3}	1.9905×10^{-2}	2.4255×10^{-2}	0.454
$Trap$		4.3494×10^{-3}	1.9907×10^{-2}	2.4257×10^{-2}	0.454
GM	15	2.8194×10^{-3}	1.8049×10^{-2}	2.0868×10^{-2}	0.408
$Trap$		2.8192×10^{-3}	1.8053×10^{-2}	2.0872×10^{-2}	0.408
GM	30	2.8194×10^{-3}	1.5470×10^{-2}	1.8290×10^{-2}	0.408
$Trap$		2.8192×10^{-3}	1.5474×10^{-2}	1.8293×10^{-2}	0.408

Table 3.7: The IMSE and the reductions in the IMSE of $\hat{g}_n^{trap}$ using the uniform design, theoretical optimal design and estimated optimal design when $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ and $n = 5$.

m	$Trap_{unif}$	$Trap_{opt}$	$rIMSE_\lambda$	$Trap_{\widehat{opt}}$	$rIMSE_{\widehat{\lambda}}$	$\widehat{\lambda}$
5	0.3661	0.3138	14.28%	0.3167	13.50%	5.15
10	0.3537	0.2988	15.54%	0.2992	15.41%	4.09
20	0.3475	0.2912	16.20%	0.2928	15.74%	4.40
30	0.3454	0.2887	16.42%	0.2844	17.67%	3.45

Table 3.8: The IMSE and the reductions in the IMSE of $\hat{g}_n^{trap}$ using the uniform design, theoretical optimal design and estimated optimal design when $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ and $n = 10$.

m	$Trap_{unif}$	$Trap_{opt}$	$rIMSE_\lambda$	$Trap_{\widehat{opt}}$	$rIMSE_{\widehat{\lambda}}$	$\widehat{\lambda}$
5	0.1969	0.1771	10.06%	0.1822	7.50%	5.06
10	0.1674	0.1494	10.79%	0.1487	11.19%	3.91
20	0.1527	0.1355	11.26%	0.1305	14.54 %	3.21
30	0.1477	0.1309	11.43%	0.1346	8.87%	4.50

Table 3.9: The IMSE and the reductions in the IMSE of $\hat{g}_n^{trap}$ using the uniform design, theoretical optimal design and estimated optimal design when $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ and $n = 20$.

m	$Trap_{unif}$	$Trap_{opt}$	$rIMSE_\lambda$	$Trap_{\widehat{opt}}$	$rIMSE_{\widehat{\lambda}}$	$\widehat{\lambda}$
5	0.1699	0.1487	12.52%	0.1457	14.26%	4.35
10	0.1274	0.1096	12.14%	0.1106	11.34%	3.82
20	0.1022	0.0901	11.86%	0.0885	13.39%	4.34
30	0.0947	0.0836	11.73%	0.0839	11.31%	3.90

Table 3.10: The IMSE and the reductions in the IMSE of $\hat{g}_n^{trap}$ using the uniform design, theoretical optimal design and estimated optimal design when $R(s,t) = \sigma^2 \rho^{|s^\lambda - t^\lambda|/\lambda}$ and $n = 30$.

m	$Trap_{unif}$	$Trap_{opt}$	$rIMSE_\lambda$	$Trap_{\widehat{opt}}$	$rIMSE_{\widehat{\lambda}}$	$\widehat{\lambda}$
5	0.1682	0.1488	11.56%	0.1434	14.78%	4.46
10	0.1201	0.1056	12.09%	0.0973	19.03%	4.86
20	0.0961	0.0840	12.57%	0.0861	10.4%	3.69
30	0.0881	0.0768	12.78%	0.7586	13.88%	4.14

3.5 Proofs

Proof of Lemma 3.2.1.

For the sake of clarity, we omit the n in $t_{i,n}$. For $i = 1, \cdots, n-1$ the Mean Value Theorem (m.v.t) yields that there exists $\eta_i \in]t_i, t_{i+1}[$ such that,

$$t_{i+1} - t_i = F^{-1}(\frac{i+1}{n}) - F^{-1}(\frac{i}{n}) = \frac{1}{nf(\eta_i)}.$$

Since $\inf_{0 \le t \le 1} f(t) > 0$ then $t_{i+1} - t_i = O(\frac{1}{n})$. We shall now prove the second part of the Lemma. Since $T_n \cap [x-h, x+h] \neq \emptyset$, there exist i_1, i_N indexes in $\{1, \ldots, n\}$ such that,

$$N_{T_n} \le i_N - i_1 + 1.$$

From the definition of the regular sequence we have for all $i = 1, ..., n$,

$$t_i = F^{-1}\Big(\frac{i}{n}\Big) \quad \text{thus} \quad i = nF(t_i).$$

Using this and the m.v.t we obtain for some $\epsilon_x \in]t_{i_1}, t_{i_N}[$,

$$N_{T_n} \le n\big(F(t_{i_N}) - F(t_{i_1})\big) + 1 = n(t_{i_N} - t_{i_1}) f(\epsilon_x) + 1,$$

The boundedness of f and the fact that $t_{i_N} - t_{i_1} \le 2h$ yield,

$$N_{T_n} \le (2 \sup_{0 \le t \le 1} f(t))\; nh + 1.$$

This concludes the proof of the second part of Lemma 3.2.1 since $1 \le nh$. $\square$

Proof of Proposition 3.3.1.

For h small enough and since $T_n \cap [x-h, x+h] \neq \emptyset$ we take $t_{x,1} < t_{x,2} < \cdots < t_{x,N_{T_n}}$ the points of T_n in $[x-h, x+h]$. Since $\mathbb{E}(\overline{Y}(t_i)) = g(t_i)$ for all $i = 1, \cdots, n$ we have,

$$\mathbb{E}(\hat{g}_n^{\text{trap}}(x)) = \frac{1}{2n}\Big\{ \sum_{k=1}^{N_{T_n}-1} \big(\frac{\varphi_{x,h}}{f} g\big)(t_{x,k}) + \big(\frac{\varphi_{x,h}}{f} g\big)(t_{x,k+1}) \Big\}.$$

From the definition of the regular sequence of designs we have for $k = 1, \dots, N_{T_n} - 1$,

$$F(t_{x,k+1}) - F(t_{x,k}) = \frac{1}{n} \iff \int_{t_{x,k}}^{t_{x,k+1}} f(t)\, dt = \frac{1}{n}. \tag{3.15}$$

Thus,

$$\mathbb{E}(\hat{g}_n^{\text{trap}}(x)) = \frac{1}{2} \sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \left\{ \big(\frac{\varphi_{x,h}}{f} g\big)(t_{x,k}) + \big(\frac{\varphi_{x,h}}{f} g\big)(t_{x,k+1}) \right\} f(t)\, dt.$$

Let,

$$\begin{aligned} I_h(x) &= \int_{x-h}^{x+h} \varphi_{x,h}(t) g(t)\, dt \\ &= \sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \varphi_{x,h}(t) g(t)\, dt + \int_{x-h}^{t_{x,1}} \varphi_{x,h}(t) g(t)\, dt + \int_{t_{x,N_{T_n}}}^{x+h} \varphi_{x,h}(t) g(t)\, dt, \end{aligned}$$

and write,

$$\mathbb{E}(\hat{g}_n^{\text{trap}}(x)) = \mathbb{E}(\hat{g}_n^{\text{trap}}(x)) - I_h(x) + I_h(x) \triangleq \Delta_{x,h} + I_h(x). \tag{3.16}$$

We first control $\Delta_{x,h}$. Let,

$$\Delta_{x,h} = \Delta_{x,h}^1 + \Delta_{x,h}^2, \tag{3.17}$$

where,

$$\begin{aligned} \Delta_{x,h}^1 &= \frac{1}{2} \sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \left(\Big(\frac{\varphi_{x,h}}{f} g\Big)(t_{x,k}) f(t) - \varphi_{x,h}(t) g(t) \right) dt \\ &\quad - \frac{1}{2} \int_{x-h}^{t_{x,1}} \varphi_{x,h}(t) g(t)\, dt - \frac{1}{2} \int_{t_{x,N_{T_n}}}^{x+h} \varphi_{x,h}(t) g(t)\, dt. \\ \Delta_{x,h}^2 &= \frac{1}{2} \sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \left(\Big(\frac{\varphi_{x,h}}{f} g\Big)(t_{x,k+1}) f(t) - \varphi_{x,h}(t) g(t) \right) dt \\ &\quad - \frac{1}{2} \int_{x-h}^{t_{x,1}} \varphi_{x,h}(t) g(t)\, dt - \frac{1}{2} \int_{t_{x,N_{T_n}}}^{x+h} \varphi_{x,h}(t) g(t)\, dt. \end{aligned}$$

For $t \in [x-h, t_{x,1}]$, Taylor expansion of $\varphi_{x,h}$ around $(x-h)$ yields,

$$\varphi_{x,h}(t) = \varphi_{x,h}(x-h) + (t-(x-h))\varphi'_{x,h}(x-h) + \frac{1}{2}(t-(x-h))^2 \varphi''_{x,h}(\theta_{x,h}), \tag{3.18}$$

for some $\theta_{x,h} \in]x-h, t_{x,1}[$. Recall that by definition of $\varphi_{x,h}$ we have,

$$\sup_{0 \le t \le 1} |\varphi_{x,h}^{(j)}(t)| \le \frac{c_j}{h^{j+1}} \quad \text{for } j = 0, 1, 2, \tag{3.19}$$

for some appropriate constants c_j where $j = 0, 1, 2$. In addition, since $\varphi_{x,h}$ is in C^2 and of support $[x-h, x+h]$ then,

$$\varphi_{x,h}(x-h) = \varphi_{x,h}(x+h) = \varphi'_{x,h}(x-h) = \varphi'_{x,h}(x+h) = 0. \tag{3.20}$$

Using (3.20) and (3.19) in (3.18) and using Lemma (3.2.1) we obtain for $t \in [x-h, t_{x,1}]$,

$$\varphi_{x,h}(t) = \frac{1}{2}(t-(x-h))^2 \varphi''_{x,h}(\theta_{x,h}) = O\Big(\frac{1}{n^2h^3}\Big), \tag{3.21}$$

Likewise, for $t \in [t_{x,N_{T_n}}, x+h]$ we have,

$$\varphi_{x,h}(t) = \frac{1}{2}(t-(x+h))^2 \varphi''_{x,h}(\theta'_{x,h}) = O\Big(\frac{1}{n^2h^3}\Big), \tag{3.22}$$

where $\theta'_{x,h} \in]t_{x,N_{T_n}}, x+h[$. Hence,

$$\int_{x-h}^{t_{x,1}} \varphi_{x,h}(t)g(t)\, dt = O\Big(\frac{1}{n^3h^3}\Big) \text{ and } \int_{t_{x,N_{T_n}}}^{x+h} \varphi_{x,h}(t)g(t)\, dt = O\Big(\frac{1}{n^3h^3}\Big).$$

Thus,

$$\Delta^1_{x,h} = \frac{1}{2}\sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \Big(\Big(\frac{\varphi_{x,h}}{f}g\Big)(t_{x,k}) - \Big(\frac{\varphi_{x,h}}{f}g\Big)(t)\Big) f(t)\, dt + O\Big(\frac{1}{n^3h^3}\Big),$$

and,

$$\Delta^2_{x,h} = \frac{1}{2}\sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} \Big(\Big(\frac{\varphi_{x,h}}{f}g\Big)(t_{x,k+1}) - \Big(\frac{\varphi_{x,h}}{f}g\Big)(t)\Big) f(t)\, dt + O\Big(\frac{1}{n^3h^3}\Big).$$

Recall that $\varphi_{x,h}$ is in C^2 and $f, g \in C^2([0,1])$, then for any $t \in]t_{x,k}, t_{x,k+1}[$ Taylor expansions of $\frac{\varphi_{x,h}}{f}g$ and f around $t_{x,k}$ give,

$$\begin{aligned}
\Delta^1_{x,h} = &\frac{1}{2}\sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k}) f(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t_{x,k}-t)\, dt \\
&- \frac{1}{2}\sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k}) f'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t-t_{x,k})^2\, dt \\
&- \frac{1}{4}\sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t-t_{x,k})^3 f''(\eta_{x,k})\, dt \\
&- \frac{1}{4}\sum_{k=1}^{N_{T_n}-1} f(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t-t_{x,k})^2 \Big(\frac{\varphi_{x,h}}{f}g\Big)''(\theta_{x,k})\, dt \\
&- \frac{1}{4}\sum_{k=1}^{N_{T_n}-1} f'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t-t_{x,k})^3 \Big(\frac{\varphi_{x,h}}{f}g\Big)''(\theta_{x,k})\, dt \\
&- \frac{1}{8}\sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} (t-t_{x,k})^4 \Big(\frac{\varphi_{x,h}}{f}g\Big)''(\theta_{x,k}) f''(\eta_{x,k})\, dt + O\Big(\frac{1}{n^3h^3}\Big),
\end{aligned}$$

where $\theta_{x,k}$ and $\eta_{x,k}$ are in $]t_{x,k}, t[$. Recall that the functions $g^{(j)}, f^{(j)}$ for $j = 0, 1, 2$ are all bounded, then using(3.19) and Lemma 3.2.1 we get,

$$\sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t - t_{x,k})^3 f''(\eta_{x,k})\, dt = O\Big(\frac{1}{n^3 h}\Big). \tag{3.23}$$

$$\sum_{k=1}^{N_{T_n}-1} f'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t - t_{x,k})^3 \Big(\frac{\varphi_{x,h}}{f} g\Big)''(\theta_{x,k})\, dt = O\Big(\frac{1}{n^3 h^2}\Big). \tag{3.24}$$

$$\sum_{k=1}^{N_{T_n}-1} \int_{t_{x,k}}^{t_{x,k+1}} (t - t_{x,k})^4 \Big(\frac{\varphi_{x,h}}{f} g\Big)''(\theta_{x,k}) f''(\eta_{x,k})\, dt = O\Big(\frac{1}{n^4 h^2}\Big). \tag{3.25}$$

Note that, since $\varphi''_{x,h}$, g'' and f'' are all lipschitz then,

$$\begin{aligned}\Big(\frac{\varphi_{x,h}}{f} g\Big)''(\theta_{x,k}) &= \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k}) + \Big[\Big(\frac{\varphi_{x,h}}{f} g\Big)''(\theta_{x,k}) - \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k})\Big] \\ &= \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k}) + O\Big(\frac{1}{n h^4}\Big). \end{aligned} \tag{3.26}$$

Injecting (3.23), (3.24), (3.25) and (3.26) in $\Delta^1_{x,h}$ we have,

$$\begin{aligned}\Delta^1_{x,h} = \frac{1}{2} &\sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k}) f(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t_{x,k} - t)\, dt \\ &- \frac{1}{2} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k}) f'(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t - t_{x,k})^2\, dt \\ &- \frac{1}{4} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k}) f(t_{x,k}) \int_{t_{x,k}}^{t_{x,k+1}} (t - t_{x,k})^2\, dt + O\Big(\frac{1}{n^3 h^3}\Big).\end{aligned}$$

Let $d_{x,k} = t_{x,k+1} - t_{x,k}$. We obtain by basic integration,

$$\begin{aligned}\Delta^1_{x,h} = &-\frac{1}{4} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k}) f(t_{x,k}) d^2_{x,k} - \frac{1}{6} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k}) f'(t_{x,k}) d^3_{x,k} \\ &- \frac{1}{12} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k}) f(t_{x,k}) d^3_{x,k} + O\Big(\frac{1}{n^3 h^3}\Big).\end{aligned} \tag{3.27}$$

Similarly we verify that,

$$\begin{aligned}\Delta^2_{x,h} = &\frac{1}{4} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k+1}) f(t_{x,k+1}) d^2_{x,k} - \frac{1}{6} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k+1}) f'(t_{x,k+1}) d^3_{x,k} \\ &- \frac{1}{12} \sum_{k=1}^{N_{T_n}-1} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k+1}) f(t_{x,k+1}) d^3_{x,k} + O\Big(\frac{1}{n^3 h^3}\Big),\end{aligned} \tag{3.28}$$

Summing (3.27) and (3.28) gives,

$$\begin{aligned}
&\Delta_{x,h} = \Delta^1_{x,h} + \Delta^2_{x,h} \\
&= \frac{1}{4}\sum_{k=1}^{N_{T_n}-1} d^2_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k+1})f(t_{x,k+1}) - \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k})f(t_{x,k})\Big] \\
&\quad - \frac{1}{6}\sum_{k=1}^{N_{T_n}-1} d^3_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k+1})f'(t_{x,k+1}) + \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k})f'(t_{x,k})\Big] \\
&\quad - \frac{1}{12}\sum_{k=1}^{N_{T_n}-1} d^3_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k+1})f(t_{x,k+1}) + \Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k})f(t_{x,k})\Big] + O\Big(\frac{1}{n^3h^3}\Big).
\end{aligned}$$

Since $\varphi'_{x,h}$ is in C^1 and $g', f' \in C^1([0,1])$, Taylor expansion of $\Big(\frac{\varphi_{x,h}}{f}g\Big)'f$ around $t_{x,k}$ yields,

$$\Big(\big(\frac{\varphi_{x,h}}{f}g\big)'f\Big)(t_{x,k+1}) = \Big(\big(\frac{\varphi_{x,h}}{f}g\big)'f\Big)(t_{x,k}) + d_{x,k}\Big(\big(\frac{\varphi_{x,h}}{f}g\big)'f\Big)'(\nu_{x,k}),$$

where $\nu_{x,k} \in]t_{x,k}, t_{x,k+1}[$. We then have,

$$\begin{aligned}
\Delta_{x,h} &= \frac{1}{4}\sum_{k=1}^{N_{T_n}-1} d^3_{x,k}\Big(\big(\frac{\varphi_{x,h}}{f}g\big)'f\Big)'(\nu_{x,k}) \\
&- \frac{1}{6}\sum_{k=1}^{N_{T_n}-1} d^3_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k+1})f'(t_{x,k+1}) + \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k})f'(t_{x,k})\Big] \\
&- \frac{1}{12}\sum_{k=1}^{N_{T_n}-1} d^3_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k+1})f(t_{x,k+1}) + \Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k})f(t_{x,k})\Big] + O\Big(\frac{1}{n^3h^3}\Big).
\end{aligned}$$

From the definition of the regular sequence of designs and using the m.v.t. we obtain for $k = 1, \cdots, N_{T_n} - 1$,

$$\int_{t_{x,k}}^{t_{x,k+1}} f(t)\,dt = \frac{1}{n} \iff d_{x,k} = \frac{1}{nf(t^*_{x,k})} \text{ for some } t^*_{x,k} \in]t_{x,k}, t_{x,k+1}[. \tag{3.29}$$

This equation yields,

$$\begin{aligned}
\Delta_{x,h} = &\frac{1}{4n^2}\sum_{k=1}^{N_{T_n}-1} d_{x,k}\frac{1}{f^2(t^*_{x,k})}\Big(\big(\frac{\varphi_{x,h}}{f}g\big)'f\Big)'(\nu_{x,k}) \\
&- \frac{1}{6n^2}\sum_{k=1}^{N_{T_n}-1} d_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k+1})\frac{f'(t_{x,k+1})}{f^2(t^*_{x,k})} + \Big(\frac{\varphi_{x,h}}{f}g\Big)'(t_{x,k})\frac{f'(t_{x,k})}{f^2(t^*_{x,k})}\Big] \\
&- \frac{1}{12n^2}\sum_{k=1}^{N_{T_n}-1} d_{x,k}\Big[\Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k+1})\frac{f(t_{x,k+1})}{f^2(t^*_{x,k})} + \Big(\frac{\varphi_{x,h}}{f}g\Big)''(t_{x,k})\frac{f(t_{x,k})}{f^2(t^*_{x,k})}\Big] \\
&+ O\Big(\frac{1}{n^3h^3}\Big).
\end{aligned}$$

Using the Riemann integrability of $\varphi_{x,h}^{(j)}, f^{(j)}$ and $g^{(j)}$ for $j = 0, 1, 2$ and applying Lemma 3.5.1 in the Appendix with $u(t) = \frac{1}{f^2(t)}$ and $v(t) = \Big((\frac{\varphi_{x,h}}{f} g)' f \Big)'(t)$ we obtain,

$$\sum_{k=1}^{N_{T_n}-1} d_{x,k} \frac{1}{f^2(t^*_{x,k})} \Big((\frac{\varphi_{x,h}}{f} g)' f \Big)'(\nu_{x,k}) = \int_{x-h}^{x+h} \frac{1}{f^2(t)} \Big((\frac{\varphi_{x,h}}{f} g)' f \Big)'(t)\ dt + O\Big(\frac{1}{nh^3}\Big).$$

Similarly, taking $u(t) = \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t)$ and $v(t) = \frac{f'(t)}{f^2(t)}$ in Lemma 3.5.1 we obtain,

$$\sum_{k=1}^{N_{T_n}-1} d_{x,k} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t_{x,k+1}) \frac{f'(t_{x,k+1})}{f^2(t^*_{x,k})} = \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t) \frac{f'(t)}{f^2(t)}\ dt + O\Big(\frac{1}{nh^3}\Big).$$

Again taking $u(t) = \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t)$ and $v(t) = \frac{f(t)}{f^2(t)}$ we obtain,

$$\sum_{k=1}^{N_{T_n}-1} d_{x,k} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t_{x,k+1}) \frac{f(t_{x,k+1})}{f^2(t^*_{x,k})} = \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t) \frac{1}{f(t)}\ dt + O\Big(\frac{1}{nh^3}\Big).$$

Hence,

$$\begin{aligned}
\Delta_{x,h} = {} & \frac{1}{4n^2} \int_{x-h}^{x+h} \frac{1}{f^2(t)} \Big((\frac{\varphi_{x,h}}{f} g)' f \Big)'(t)\ dt - \frac{1}{3n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t) \frac{f'(t)}{f^2(t)}\ dt \\
& - \frac{1}{6n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t) \frac{1}{f(t)}\ dt + O\Big(\frac{1}{n^3h^3}\Big).
\end{aligned}$$

Simple derivations yield,

$$\begin{aligned}
\Delta_{x,h} = {} & \frac{1}{4n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t) \frac{1}{f(t)}\ dt + \frac{1}{4n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t) \frac{f'(t)}{f^2(t)}\ dt \\
& - \frac{1}{3n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t) \frac{f'(t)}{f^2(t)}\ dt - \frac{1}{6n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t) \frac{1}{f(t)}\ dt + O\Big(\frac{1}{n^3h^3}\Big) \\
= {} & \frac{1}{12n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)''(t) \frac{1}{f(t)}\ dt - \frac{1}{12n^2} \int_{x-h}^{x+h} \Big(\frac{\varphi_{x,h}}{f} g\Big)'(t) \frac{f'(t)}{f^2(t)}\ dt + O\Big(\frac{1}{n^3h^3}\Big) \\
= {} & \frac{1}{12n^2} \int_{x-h}^{x+h} \Big(\Big(\frac{\varphi_{x,h}}{f} g\Big)' \frac{1}{f} \Big)'(t)\ dt + O\Big(\frac{1}{n^3h^3}\Big).
\end{aligned}$$

Finally,

$$\Delta_{x,h} = \frac{1}{12n^2} \Big(\Big(\frac{\varphi_{x,h}}{f} g\Big)' \frac{1}{f} \Big)(x+h) - \Big(\Big(\frac{\varphi_{x,h}}{f} g\Big)' \frac{1}{f} \Big)(x-h) + O\Big(\frac{1}{n^3h^3}\Big).$$

The last equation together with (3.20) yield,

$$\Delta_{x,h} = O\Big(\frac{1}{n^3h^3}\Big). \tag{3.30}$$

The control of $I_h(x)$ is classical and it can be seen from Gasser and Müller that,

$$I_h(x) = g(x) + \frac{1}{2} h^2 g''(x) \int_{-1}^{1} t^2 K(t)\ dt + o(h^2). \tag{3.31}$$

Finally, collecting (3.16), (3.30) and (3.31) gives,

$$\mathrm{Bias}(\hat{g}_n^{trap}(x)) = \frac{1}{2}h^2 g''(x)B + o(h^2) + O\Big(\frac{1}{n^3h^3}\Big),$$

where $B = \int_{-1}^{1} t^2K(t)\ dt$. This concludes the proof of Proposition 3.3.1. □

Proof of Proposition 3.3.2.

The greatest lines of this proof are based on the work of Belouni and Benhenni (2015). For h small enough and since $T_n \cap [x-h, x+h] \neq \emptyset$ we have,

$$0 \leq t_1 < \cdots < x-h \leq t_{x,1} < \cdots < t_{x,N_{T_n}} \leq x+h < \cdots < t_n \leq 1.$$

Let,

$$\Phi(t,s) = \big(\frac{\varphi_{x,h}}{f}\big)(t)R(t,s)\big(\frac{\varphi_{x,h}}{f}\big)(s),$$

and,

$$\sigma_{x,h}^2 = \int_{x-h}^{x+h}\int_{x-h}^{x+h} \varphi_{x,h}(t)R(t,s)\varphi_{x,h}(s)\ ds\ dt. \tag{3.32}$$

On the one hand,

$$\begin{aligned}\mathrm{Var}(\hat{g}_n^{trap}(x)) = \frac{1}{4mn^2}\sum_{i=1}^{N_{T_n}-1}\sum_{j=1}^{N_{T_n}-1}\Big\{&\Phi(t_{x,i},t_{x,j}) + \Phi(t_{x,i},t_{x,j+1}) + \Phi(t_{x,i+1},t_{x,j})\\ &+ \Phi(t_{x,i+1},t_{x,j+1})\Big\}\end{aligned}$$

Using (3.15) one can write,

$$\begin{aligned}\mathrm{Var}(\hat{g}_n^{trap}(x)) = \frac{1}{4m}\sum_{i=1}^{N_{T_n}-1}\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\Big\{&\Phi(t_{x,i},t_{x,j}) + \Phi(t_{x,i},t_{x,j+1})\\ &+ \Phi(t_{x,i+1},t_{x,j}) + \Phi(t_{x,i+1},t_{x,j+1})\Big\}f(s)\ f(t)\ ds\ dt.\end{aligned}$$

On the other hand we have,

$$\begin{aligned}\sigma_{x,h}^2 &= \sum_{i=1}^{N_{T_n}-1}\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\ f(s)\ ds\ dt\\ &+ 2\int_{x-h}^{t_{x,1}}\int_{t_{x,N_{T_n}}}^{x+h} \Phi(t,s)f(t)\ f(s)\ ds\ dt + \int_{t_{x,N_{T_n}}}^{x+h}\int_{t_{x,N_{T_n}}}^{x+h} \Phi(t,s)f(t)\ f(s)\ ds\ dt\\ &+ \int_{x-h}^{t_{x,1}}\int_{x-h}^{t_{x,1}} \Phi(t,s)f(t)\ f(s)\ ds\ dt + 2\sum_{j=1}^{N_{T_n}-1}\int_{x-h}^{t_{x,1}}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\ f(s)\ ds\ dt\\ &+ 2\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,N_{T_n}}}^{x+h}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\ f(s)\ ds\ dt.\end{aligned}$$

Recall that Lemma 3.2.1 yields $N_{T_n} = O(nh)$ and $\sup_{1\le i\le n} d_{x,i} = O(\frac{1}{n})$. Using (3.21) and (3.22) we have,

$$\sup_{(x-h)\le t\le t_{x,1}} |\varphi_{x,h}(t)| = O\Big(\frac{1}{n^2h^3}\Big) \quad \text{and} \quad \sup_{t_{x,N_{T_n}}\le t\le (x+h)} |\varphi_{x,h}(t)| = O\Big(\frac{1}{n^2h^3}\Big). \tag{3.33}$$

Since f and R are bounded, using (3.19) and (3.33) we obtain,

$$\int_{x-h}^{t_{x,1}}\int_{t_{x,N_{T_n}}}^{x+h} \Phi(t,s)f(t)\, f(s)\, ds\, dt = O\Big(\frac{1}{n^6h^6}\Big),$$
$$\int_{t_{x,N_{T_n}}}^{x+h}\int_{t_{x,N_{T_n}}}^{x+h} \Phi(t,s)f(t)\, f(s)\, ds\, dt = O\Big(\frac{1}{n^6h^6}\Big),$$
$$\int_{x-h}^{t_{x,1}}\int_{x-h}^{t_{x,1}} \Phi(t,s)f(t)\, f(s)\, ds\, dt = O\Big(\frac{1}{n^6h^6}\Big),$$
$$\sum_{j=1}^{N_{T_n}-1}\int_{x-h}^{t_{x,1}}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\, f(s)\, ds\, dt = O\Big(\frac{1}{n^3h^3}\Big),$$
$$\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,N_{T_n}}}^{x+h}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\, f(s)\, ds\, dt = O\Big(\frac{1}{n^3h^3}\Big).$$

Thus,

$$\sigma^2_{x,h} = \sum_{i=1}^{N_{T_n}-1}\sum_{j=1}^{N_{T_n}-1}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} \Phi(t,s)f(t)\, f(s)\, ds\, dt + O\Big(\frac{1}{n^3h^3}\Big).$$

We shall control the residual variance $\mathrm{Var}(\hat{g}_n^{trap}(x)) - \frac{\sigma^2_{x,h}}{m}$. For this, let,

$$N_{i,j}(t,s) = \Phi(t_{x,i},t_{x,j}) + \Phi(t_{x,i+1},t_{x,j}) + \Phi(t_{x,i},t_{x,j+1}) + \Phi(t_{x,i+1},t_{x,j+1}) - 4\Phi(t,s), \tag{3.34}$$

and put,

$$I_{i,j} = \frac{1}{4m}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} N_{i,j}(t,s)f(t)\, f(s)\, ds\, dt. \tag{3.35}$$

The residual variance can then be written as follows,

$$\mathrm{Var}(\hat{g}_n^{trap}(x)) - \frac{\sigma^2_{x,h}}{m} = \sum_{i=1}^{N_{T_n}-1} I_{i,i} + \sum\sum_{i\neq j=1}^{N_{T_n}-1} I_{i,j} + O\Big(\frac{1}{mn^3h^3}\Big), \tag{3.36}$$

Starting with the diagonal terms $I_{i,i}$. Since for any $s,t\in[0,1]$, we have $N_{i,i}(s,t) = N_{i,i}(t,s)$, then we can write,

$$I_{i,i} = \frac{1}{2m}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,i}}^{t} N_{i,i}(t,s)f(t)\, f(s)\, ds\, dt. \tag{3.37}$$

Because of Assumption (B), $N_{i,i}$ has left and right first order derivatives on the diagonal on $[0,1]^2$. For any s,t such that $(t_{x,i} < s \le t < t_{x,i+1})$, Taylor expansion of Φ around $(t_{x,i},t_{x,i})$

gives,

$$\begin{aligned}\Phi(t,s) &= \Phi(t,t_{x,i}) + (s-t_{x,i})\Phi^{(0,1)}(t,t_{x,i}) + \frac{1}{2}(s-t_{x,i})^2\Phi^{(0,2)}(t,\eta_{s,i}^{(1)})\\ &= \Phi(t_{x,i},t_{x,i}) + (t-t_{x,i})\Phi^{(1,0)}(\epsilon_{t,i}^{(1)},t_{x,i}) + (s-t_{x,i})\Phi^{(0,1)}(\epsilon_i,t_{x,i})\\ &\quad + (s-t_{x,i})(t-\epsilon_i)\Phi^{(1,1)}(\epsilon_{t,i}^{(2)},t_{x,i}) + \frac{1}{2}(s-t_{x,i})^2\Phi^{(0,2)}(t,\eta_{s,i}^{(1)}),\end{aligned}$$

for some $\epsilon_i \in]t_{x,i},t_{x,i+1}[$, some $\epsilon_{t,i}^{(1)}$ in $]t_{x,i},t[$, some $\epsilon_{t,i}^{(2)}$ between t and ϵ_i and some $\eta_{t,i}^{(1)}$ in $]t_{x,i},s[$. We have,

$$\begin{aligned}\Phi(t,s) &= \Phi(t_{x,i},t_{x,i}) + (t-t_{x,i})\Phi^{(1,0)}(\epsilon_i,t_{x,i}) + (s-t_{x,i})\Phi^{(0,1)}(\epsilon_i,t_{x,i})\\ &\quad + (t-t_{x,i})\Big(\Phi^{(1,0)}(\epsilon_{t,i}^{(1)},t_{x,i}) - \Phi^{(1,0)}(\epsilon_i,t_{x,i})\Big)\\ &\quad + (s-t_{x,i})(t-\epsilon_i)\Phi^{(1,1)}(\epsilon_{t,i}^{(2)},t_{x,i}) + \frac{1}{2}(s-t_{x,i})^2\Phi^{(0,2)}(t,\eta_{s,i}^{(1)}).\end{aligned}$$

For l and l' integers such that $l+l' \le 2$, Assumption (C) yields,

$$\sup_{s\neq t} |\Phi^{(l,l')}(t,s)| = O\Big(\frac{1}{h^{l+l'+2}}\Big). \tag{3.38}$$

In addition, since $\varphi_{x,h}, \varphi'_{x,h}, \frac{1}{f}, R$ and $R(\cdot,t_{x,i})$ are all continuous on $]t_{x,i},t_{x,i+1}[$, then fo $s\neq t$ in $]t_{x,i},t_{x,i+1}[$ we have,

$$\begin{aligned}&\Big|\Phi^{(1,0)}(s,t_{x,i}) - \Phi^{(1,0)}(t,t_{x,i})\Big| = \Big|\frac{\varphi_{x,h}}{f}(t_{x,i})\Big|\Big|R(s,t_{x,i})\Big(\frac{\varphi'_{x,h}}{f}(s) - \frac{\varphi'_{x,h}}{f}(t)\Big)\\ &+ R^{(1,0)}(s,t_{x,i})\Big(\frac{\varphi_{x,h}}{f}(s) - \frac{\varphi_{x,h}}{f}(t)\Big) + \frac{\varphi_{x,h}}{f}(t)\Big(R^{(1,0)}(s,t_{x,i}) - R^{(1,0)}(t,t_{x,i})\Big)\\ &+ \frac{\varphi'_{x,h}}{f}(t)\Big(R(s,t_{x,i}) - R(t,t_{x,i})\Big)\Big| = O\Big(\frac{1}{nh^4}\Big).\end{aligned}$$

Finally, using this equation together with Lemma 3.2.1 we obtain,

$$\Phi(t,s) = \Phi(t_{x,i},t_{x,i}) + (t-t_{x,i})\Phi^{(1,0)}(\epsilon_i,t_{x,i}) + (s-t_{x,i})\Phi^{(0,1)}(\epsilon_i,t_{x,i}) + O\Big(\frac{1}{n^2h^4}\Big). \tag{3.39}$$

Similarly we verify that,

$$\Phi(t_{x,i+1},t_{x,i+1}) = \Phi(t_{x,i},t_{x,i}) + d_{x,i}\Phi^{(1,0)}(\epsilon_i,t_{x,i}) + d_{x,i}\Phi^{(0,1)}(\epsilon_i,t_{x,i}) + O\Big(\frac{1}{n^2h^4}\Big), \tag{3.40}$$

and that,

$$\Phi(t_{x,i+1},t_{x,i}) = \Phi(t_{x,i},t_{x,i}) + d_{x,i}\Phi^{(1,0)}(\epsilon_i,t_{x,i}) + O\Big(\frac{1}{n^2h^4}\Big). \tag{3.41}$$

Inserting (3.39), (3.40) and (3.41) in (3.34) for $i=j$ and using (3.38) and Lemma 3.2.1, we obtain,

$$\begin{aligned}N_{i,i}(t,s) &= 3d_{x,i}\Phi^{(1,0)}(\epsilon_i,t_{x,i}) - 4(t-t_{x,i})\Phi^{(1,0)}(\epsilon_i,t_{x,i})\\ &\quad + d_{x,i}\Phi^{(0,1)}(\epsilon_i,t_{x,i})) - 4(s-t_{x,i})\Phi^{(0,1)}(\epsilon_i,t_{x,i}) + O\Big(\frac{1}{n^2h^4}\Big).\end{aligned}$$

Replacing this expression in (3.37), and using the boundedness of f and Lemma 3.2.1, we obtain,

$$I_{i,i} = \frac{1}{2m}\Bigg(d_{x,i}\Big(3\Phi^{(1,0)}(\epsilon_i, t_{x,i}) + \Phi^{(0,1)}(\epsilon_i, t_{x,i})\Big) \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,i}}^{t} f(t)\, f(s)\, ds\, dt \\ - 4\Phi^{(1,0)}(\epsilon_i, t_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,i}}^{t} (t - t_{x,i}) f(t)\, f(s)\, ds\, dt \\ - 4\Phi^{(0,1)}(\epsilon_i, t_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,i}}^{t} (s - t_{x,i}) f(t)\, f(s)\, ds\, dt \Bigg) + O\Big(\frac{1}{mn^4h^4}\Big). \tag{3.42}$$

Recall that f is in $C^2([0,1])$ and that $d_{x,i} = O(\frac{1}{n})$ from Lemma 3.2.1. It can easily be verified that for any integers l and l':

$$\int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,i}}^{t} (s - t_{x,i})^{l'} (s - t_{x,i})^{l} f(t)\, f(s)\, ds\, dt = \frac{f^2(t_{x,i})\, d_{x,i}^{(l+l'+2)}}{(l'+1)(l+l'+2)} + O\Big(\frac{1}{n^{l+l'+3}}\Big).$$

Using this last Equation together with (3.38) in (3.42) above, and (3.29) we obtain,

$$\begin{aligned} I_{i,i} &= \frac{1}{12m}\Big(\Phi^{(1,0)}(\epsilon_i, t_{x,i}) - \Phi^{(0,1)}(\epsilon_i, t_{x,i})\Big) f^2(t_{x,i}) d_{x,i}^3 + O\Big(\frac{1}{mn^4h^4}\Big) \\ &= \frac{1}{12mn^2}\Big(\Phi^{(1,0)}(\epsilon_i, t_{x,i}) - \Phi^{(0,1)}(\epsilon_i, t_{x,i})\Big) \frac{f^2(t_{x,i})}{f^2(t^*_{x,i})} d_{x,i} + O\Big(\frac{1}{mn^4h^4}\Big). \end{aligned}$$

Finally using Lemma 3.2.1, the integrability of $\varphi_{x,h}, \varphi'_{x,h}, f, f'$ and $R^{(0,1)}(.,t)$ and applying Lemma 3.5.1 in the Appendix, we obtain,

$$\begin{aligned} \sum_{i=1}^{N_{T_n}-1} I_{i,i} &= \frac{1}{12mn^2} \sum_{i=1}^{N_{T_n}-1} \Big(\Phi^{(1,0)}(\epsilon_i, t_{x,i}) - \Phi^{(0,1)}(\epsilon_i, t_{x,i})\Big) \frac{f^2(t_{x,i})}{f^2(t^*_{x,i})} d_{x,i} + O\Big(\frac{1}{mn^3h^3}\Big) \\ &= \frac{1}{12mn^2} \int_{x-h}^{x+h} \Big(\Phi^{(1,0)}(t^+, t) - \Phi^{(0,1)}(t^+, t)\Big)\, dt + O\Big(\frac{1}{mn^3h^3}\Big). \end{aligned} \tag{3.43}$$

Since $\Phi^{(0,1)}(t^+, t) = \Phi^{(0,1)}(t, t^-) = \Phi^{(1,0)}(t^-, t)$, then,

$$\sum_{i=1}^{N_{T_n}-1} I_{i,i} = -\frac{1}{12mn^2} \int_{x-h}^{x+h} \Big(\Phi^{(1,0)}(t^-, t) - \Phi^{(1,0)}(t^+, t)\Big)\, dt + O\Big(\frac{1}{mn^3h^3}\Big). \tag{3.44}$$

Now, it remains to handle the off diagonal term. Assumption (B) yields that $N_{i,j}$ for $i \neq j$ is twice differentiable off the diagonal on $[0,1]^2$. Taylor expansion of $N_{i,j}$ around $(t_{x,i}, t_{x,j})$ for $i \neq j$ up to order 2 gives,

$$\begin{aligned} \Phi(t,s) = \Phi(t_{x,i}, t_{x,j}) &+ (t - t_{x,i})\Phi^{(1,0)}(t_{x,i}, t_{x,j}) + (s - t_{x,j})\Phi^{(0,1)}(t_{x,i}, t_{x,j}) \\ &+ \frac{1}{2}(t - t_{x,i})^2 \Phi^{(2,0)}(\epsilon^{(1)}_{x,i}, t_{x,j}) + \frac{1}{2}(s - t_{x,j})^2 \Phi^{(0,2)}(t_{x,i}, \eta^{(1)}_{x,j}) \\ &+ (t - t_{x,i})(s - t_{x,j})\Phi^{(1,1)}(\epsilon^{(1)}_{x,i}, \eta^{(1)}_{x,j}), \end{aligned} \tag{3.45}$$

for some $\epsilon^{(1)}_{x,i}$ between $t_{x,i}$ and t and some $\eta^{(1)}_{x,j}$ between $t_{x,j}$ and s. Taking $t = t_{x,i+1}$ and $s = t_{x,j}$ in (3.45), we obtain,

$$\Phi(t_{x,i+1}, t_{x,j}) = \Phi(t_{x,i}, t_{x,j}) + d_{x,i}\Phi^{(1,0)}(t_{x,i}, t_{x,j}) + \frac{1}{2} d_{x,i}^2 \Phi^{(2,0)}(\epsilon^{(2)}_{x,i}, t_{x,j}), \tag{3.46}$$

for some $\epsilon_{x,i}^{(2)}$ in $]t_{x,i}, t_{x,i+1}[$. Taking $t = t_{x,i}$ and $s = t_{x,j+1}$ in (3.45), we obtain,

$$\Phi(t_{x,i}, t_{x,j+1}) = \Phi(t_{x,i}, t_{x,j}) + d_{x,j}\Phi^{(0,1)}(t_{x,i}, t_{x,j}) + \frac{1}{2}d_{x,j}^2\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(2)}), \tag{3.47}$$

for some $\eta_{x,j}^{(2)}$ in $]t_{x,j}, t_{x,j+1}[$. Taking $t = t_{x,i+1}$ and $s = t_{x,j+1}$ in (3.45), we obtain,

$$\begin{aligned}\Phi(t_{x,i+1}, t_{x,j+1}) = {} & \Phi(t_{x,i}, t_{x,j}) + d_{x,i}\Phi^{(1,0)}(t_{x,i}, t_{x,j}) + d_{x,j}\Phi^{(0,1)}(t_{x,i}, t_{x,j}) \\ & + \frac{1}{2}d_{x,i}^2\Phi^{(2,0)}(\epsilon_{x,i}^{(3)}, t_{x,j}) + \frac{1}{2}d_{x,j}^2\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(3)}) \\ & + d_{x,i}d_{x,j}\Phi^{(1,1)}(\epsilon_{x,i}^{(3)}, \eta_{x,j}^{(3)}),\end{aligned} \tag{3.48}$$

We obtain by inserting (3.45), (3.46), (3.47) and (3.48) in (3.34),

$$\begin{aligned}N_{i,j}(t,s) = {} & \Phi^{(1,0)}(t_{x,i}, t_{x,j})\big(2d_{x,i} - 4(t - t_{x,i})\big) + \Phi^{(0,1)}(t_{x,i}, t_{x,j})\big(2d_{x,j} - 4(s - t_{x,j})\big) \\ & + \frac{1}{2}d_{x,i}^2\big(\Phi^{(2,0)}(\epsilon_{x,i}^{(2)}, t_{x,j}) + \Phi^{(2,0)}(\epsilon_{x,i}^{(3)}, t_{x,j})\big) - 2(t - t_{x,i})^2\Phi^{(2,0)}(\epsilon_{x,i}^{(1)}, t_{x,j}) \\ & + \frac{1}{2}d_{x,j}^2\big(\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(2)}) + \Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(3)})\big) - 2(s - t_{x,j})^2\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(1)}) \\ & + d_{x,i}d_{x,j}\ \Phi^{(1,1)}(\epsilon_{x,i}^{(3)}, \eta_{x,j}^{(3)}) - 4(t - t_{x,i})(s - t_{x,j})\Phi^{(1,1)}(\epsilon_{x,i}^{(1)}, \eta_{x,j}^{(1)}).\end{aligned}$$

We obtain inserting the last equation in (3.35),

$$I_{i,j} = \frac{1}{4m}\sum_{l=1}^{5} I_{i,j}^{(l)}, \tag{3.49}$$

where,

$$
\begin{aligned}
I_{i,j}^{(1)} &= \Phi^{(1,0)}(t_{x,i}, t_{x,j})\Big(2d_{x,i}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} f(t)f(s)dtds \\
&\quad - 4\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})f(t)f(s)dtds\Big). \\
I_{i,j}^{(2)} &= \Phi^{(0,1)}(t_{x,i}, t_{x,j})\Big(2d_{x,j}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} f(t)f(s)dtds \\
&\quad - 4\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(s-t_{x,j})f(t)f(s)dtds\Big). \\
I_{i,j}^{(3)} &= \frac{1}{2}d_{x,i}^2\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\big(\Phi^{(2,0)}(\epsilon_{x,i}^{(2)}, t_{x,j}) + \Phi^{(2,0)}(\epsilon_{x,i}^{(3)}, t_{x,j})\big)f(t)f(s)dtds \\
&\quad - 2\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})^2\Phi^{(2,0)}(\epsilon_{x,i}^{(1)}, t_{x,j})f(t)f(s)dtds. \\
I_{i,j}^{(4)} &= \frac{1}{2}d_{x,j}^2\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\big(\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(2)}) + \Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(3)})\big)f(t)f(s)dtds \\
&\quad - 2\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(s-t_{x,j})^2\Phi^{(0,2)}(t_{x,i}, \eta_{x,j}^{(1)})f(t)f(s)dtds. \\
I_{i,j}^{(5)} &= d_{x,i}d_{x,j}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\Phi^{(1,1)}(\epsilon_{x,i}^{(3)}, \eta_{x,j}^{(3)})f(t)f(s)dtds \\
&\quad - 4\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})(s-t_{x,j})\Phi^{(1,1)}(\epsilon_{x,i}^{(1)}, \eta_{x,j}^{(1)})f(t)f(s)dtds.
\end{aligned}
$$

We first consider the term $I_{i,j}^{(1)}$. For $l = 0, 1, 2$, let,

$$\omega_{i,l} = \int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})^l f(t)dt \tag{3.50}$$

The term $I_{i,j}^{(1)}$ can then be written as,

$$I_{i,j}^{(1)} = \Phi^{(1,0)}(t_{x,i}, t_{x,j})\Big(2d_{x,i}\omega_{i,0}\omega_{j,0} - 4\omega_{i,1}\omega_{j,0}\Big). \tag{3.51}$$

Expanding f around $t_{x,i}$ yields,

$$
\begin{aligned}
\omega_{i,l} &= \int_{t_{x,i}}^{t_{x,i+1}}(t-t_{x,i})^l\big(f(t_{x,i}) + (t-t_{x,i})f'(t_{x,i}) + \frac{1}{2}(t-t_{x,i})^2 f''(\epsilon_{x,i}^{(4)})\big)\ dt \\
&= \frac{d_{x,i}^{(l+1)}}{(l+1)}f(t_{x,i}) + \frac{d_{x,i}^{(l+2)}}{(l+2)}f'(t_{x,i}) + O\Big(\frac{1}{n^{(l+3)}}\Big),
\end{aligned} \tag{3.52}
$$

for some $\epsilon_{x,i}^{(4)}$ in $]t_{x,i}, t_{x,i+1}[$. Thus for $l = 0, 1, 2$,

$$\begin{aligned}
I_{i,j}^{(1)} &= \Phi^{(1,0)}(t_{x,i}, t_{x,j}) \Bigg(2d_{x,i} \Bigg(d_{x,i} f(t_{x,i}) + \frac{d_{x,i}^2}{2} f'(t_{x,i}) + O\Big(\frac{1}{n^3}\Big) \Bigg) \\
&\times \Bigg(d_{x,j} f(t_{x,j}) + \frac{d_{x,j}^2}{2} f'(t_{x,i}) + O\Big(\frac{1}{n^3}\Big) \Bigg) \\
&- 4\Big(\frac{d_{x,i}^2}{2} f(t_{x,i}) + \frac{d_{x,i}^3}{3} f'(t_{x,i}) + O\Big(\frac{1}{n^4}\Big) \Big) \Big(d_{x,j} f(t_{x,j}) + \frac{d_{x,j}^2}{2} f'(t_{x,i}) + O\Big(\frac{1}{n^3}\Big) \Big) \Bigg) \\
&= \Phi^{(1,0)}(t_{x,i}, t_{x,j}) \Big(-\frac{1}{3} f'(t_{x,i}) f(t_{x,j}) d_{x,i}^3 d_{x,j} + O\Big(\frac{1}{n^5}\Big) \Big).
\end{aligned}$$

We obtain using Equations (3.38) and (3.29),

$$\begin{aligned}
I_{i,j}^{(1)} &= -\frac{1}{3} \Phi^{(1,0)}(t_{x,i}, t_{x,j}) f'(t_{x,i}) f(t_{x,j}) d_{x,i}^3 d_{x,j} + O\Big(\frac{1}{n^5 h^3}\Big) \\
&= -\frac{1}{3n^2} \Phi^{(1,0)}(t_{x,i}, t_{x,j}) \frac{f'(t_{x,i})}{f^2(t_{x,i}^*)} f(t_{x,j}) d_{x,i} d_{x,j} + O\Big(\frac{1}{n^5 h^3}\Big),
\end{aligned}$$

for some $t_{x,i}^*$ in $]t_{x,i}, t_{x,i+1}[$. Using Lemma 3.2.1 and the integrability of $\varphi_{x,h}, \varphi'_{x,h}, f$, and of $R^{(0,1)}(., t)$ and applying Lemma 3.5.1 twice, we obtain,

$$\begin{aligned}
\sum_{i \neq j = 1}^{N_{T_n}-1}\sum I_{i,j}^{(1)} &= -\frac{1}{3n^2} \sum_{i \neq j = 1}^{N_{T_n}-1}\sum \Phi^{(1,0)}(t_{x,i}, t_{x,j}) \frac{f'(t_{x,i})}{f^2(t_{x,i}^*)} f(t_{x,j}) d_{x,i} d_{x,j} + O\Big(\frac{1}{n^3 h}\Big) \\
&= -\frac{1}{3n^2} \int_{x-h}^{x+h} \int_{x-h}^{x+h} \Phi^{(1,0)}(t, s) \frac{f'(t)}{f^2(t)} f(s) 1_{\{s \neq t\}} \, dt \, ds + O\Big(\frac{1}{n^3 h^2}\Big). \qquad (3.53)
\end{aligned}$$

Similarly we verify that,

$$\begin{aligned}
\sum_{i \neq j = 1}^{N_{T_n}-1}\sum I_{i,j}^{(2)} &= -\frac{1}{3n^2} \int_{x-h}^{x+h} \int_{x-h}^{x+h} \Phi^{(0,1)}(t, s) \frac{f'(s)}{f^2(s)} f(t) 1_{\{s \neq t\}} \, dt \, ds + O\Big(\frac{1}{n^3 h^2}\Big) \\
&= -\frac{1}{3n^2} \int_{x-h}^{x+h} \int_{x-h}^{x+h} \Phi^{(1,0)}(t, s) \frac{f'(t)}{f^2(t)} f(s) 1_{\{s \neq t\}} \, dt \, ds + O\Big(\frac{1}{n^3 h^2}\Big). \qquad (3.54)
\end{aligned}$$

We now control the term $I_{i,j}^3$. We have,

$$\begin{aligned}
I_{i,j}^{(3)} &= d_{x,i}^2 \Phi^{(2,0)}(t_{x,i}, t_{x,j}) \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,j}}^{t_{x,j+1}} f(t) f(s) dt ds \\
&- 2\Phi^{(2,0)}(t_{x,i}, t_{x,j}) \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,j}}^{t_{x,j+1}} (t - t_{x,i})^2 f(t) f(s) dt ds \\
&+ \frac{1}{2} d_{x,i}^2 \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,j}}^{t_{x,j+1}} \Phi^{(2,0)}(\epsilon_{x,i}^{(2)}, t_{x,j}) + \Phi^{(2,0)}(\epsilon_{x,i}^{(3)}, t_{x,j}) - 2\Phi^{(2,0)}(t_{x,i}, t_{x,j}) f(t) f(s) dt ds \\
&- 2 \int_{t_{x,i}}^{t_{x,i+1}} \int_{t_{x,j}}^{t_{x,j+1}} (t - t_{x,i})^2 \big(\Phi^{(2,0)}(\epsilon_{x,i}^{(1)}, t_{x,j}) - \Phi^{(2,0)}(t_{x,i}, t_{x,j}) \big) f(t) f(s) dt ds.
\end{aligned}$$

Using (3.38), Lemma 3.2.1 and Equation (3.50) we get,

$$I_{i,j}^{(3)} = d_{x,i}^2\Phi^{(2,0)}(t_{x,i},t_{x,j})\omega_{i,0}\omega_{j,0} - 2\Phi^{(2,0)}(t_{x,i},t_{x,j})\omega_{i,2}\omega_{i,0} + O\Big(\frac{1}{n^5h^5}\Big).$$

Note first that, using (3.52) for $l=0$ along with $l=2$ and Lemma 3.2.1, we obtain,

$$\begin{aligned} I_{i,j}^{(3)} &= \frac{1}{3}\Phi^{(2,0)}(t_{x,i},t_{x,j})d_{x,i}^3 d_{x,j} f(t_{x,i})f(t_{x,j}) + O\Big(\frac{1}{n^5h^5}\Big) \\ &= \frac{1}{3n^2}\Phi^{(2,0)}(t_{x,i},t_{x,j})\frac{f(t_{x,i})}{f^2(t_{x,i}^*)}f(t_{x,j})d_{x,i}d_{x,j} + O\Big(\frac{1}{n^5h^5}\Big), \end{aligned}$$

Likewise, using Lemma 3.2.1 and the integrability of $\varphi_{x,h}^{(k)}$, $f^{(k)}$ for $k=0,1,2$ we have,

$$\sum_{i\neq j=1}^{N_{T_n}-1}\sum I_{i,j}^{(3)} = \frac{1}{3n^2}\int_{x-h}^{x+h}\int_{x-h}^{x+h}\Phi^{(2,0)}(t,s)\frac{f(s)}{f(t)}1_{\{s\neq t\}}\,dt\,ds + O\Big(\frac{1}{n^3h^3}\Big). \tag{3.55}$$

Similarly, we obtain,

$$\begin{aligned} \sum_{i\neq j=1}^{N_{T_n}-1}\sum I_{i,j}^{(4)} &= \frac{1}{3n^2}\int_{x-h}^{x+h}\int_{x-h}^{x+h}\Phi^{(0,2)}(t,s)\frac{f(t)}{f(s)}1_{\{s\neq t\}}\,dt\,ds + O\Big(\frac{1}{n^3h^3}\Big) \\ &= \frac{1}{3n^2}\int_{x-h}^{x+h}\int_{x-h}^{x+h}\Phi^{(2,0)}(t,s)\frac{f(s)}{f(t)}1_{\{s\neq t\}}\,dt\,ds + O\Big(\frac{1}{n^3h^3}\Big). \end{aligned} \tag{3.56}$$

Finally, for the term $I_{i,j}^{(5)}$, we have,

$$\begin{aligned} I_{i,j}^{(5)} &= d_{x,i}d_{x,j}\Phi^{(1,1)}(t_{x,i},t_{x,j})\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}} f(t)f(s)dtds \\ &- 4\Phi^{(1,1)}(t_{x,i},t_{x,j})\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})(s-t_{x,j})f(t)f(s)dtds \\ &+ d_{x,i}d_{x,j}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\Big(\Phi^{(1,1)}(\epsilon_{x,i}^{(3)},\eta_{x,j}^{(3)}) - \Phi^{(1,1)}(t_{x,i},t_{x,j})\Big)f(t)f(s)dtds \\ &- 4\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})(s-t_{x,j})\Big(\Phi^{(1,1)}(\epsilon_{x,i}^{(1)},\eta_{x,j}^{(1)}) - \Phi^{(1,1)}(t_{x,i},t_{x,j})\Big)f(t)f(s)dtds \\ &= d_{x,i}d_{x,j}\Phi^{(1,1)}(t_{x,i},t_{x,j})\omega_{i,0}\omega_{j,0} - 4\Phi^{(1,1)}(t_{x,i},t_{x,j})\omega_{i,1}\omega_{j,1} \\ &+ d_{x,i}d_{x,j}\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}\Big(\Phi^{(1,1)}(\epsilon_{x,i}^{(3)},\eta_{x,j}^{(3)}) - \Phi^{(1,1)}(t_{x,i},t_{x,j})\Big)f(t)f(s)dtds \\ &- 4\int_{t_{x,i}}^{t_{x,i+1}}\int_{t_{x,j}}^{t_{x,j+1}}(t-t_{x,i})(s-t_{x,j})\Big(\Phi^{(1,1)}(\epsilon_{x,i}^{(1)},\eta_{x,j}^{(1)}) - \Phi^{(1,1)}(t_{x,i},t_{x,j})\Big)f(t)f(s)dtds. \end{aligned}$$

Recall that $f, f', \frac{1}{f}$ are all bounded and using (3.38) and (3.52) with $l=l'=1$ we obtain,

$$I_{i,j}^{(5)} = O\Big(\frac{1}{n^5h^5}\Big).$$

Finally, since $N_{T_n} = O(nh)$ from Lemma 3.2.1, we obtain,

$$\sum\sum_{i\neq j=1}^{N_{T_n}-1} I_{i,j}^{(5)} = O\Big(\frac{1}{n^3h^3}\Big). \tag{3.57}$$

Replacing (3.53), (3.54), (3.55), (3.56) and (3.57) in (3.49) we obtain,

$$\begin{aligned}
\sum\sum_{i\neq j=1}^{N_{T_n}-1} I_{i,j} &= \frac{1}{6mn^2}\int_{x-h}^{x+h}\int_{x-h}^{x+h}\Big(\frac{\Phi^{(2,0)}(t,s)f(t)-\Phi^{(1,0)}(t,s)f'(t)}{f^2(t)}\Big)1_{\{s\neq t\}}f(s)dtds \\
&\quad + O\Big(\frac{1}{mn^3h^3}\Big) \\
&= \frac{1}{6mn^2}\int_{x-h}^{x+h}\Big(\int_{x-h}^{s}\frac{\partial}{\partial s}\Big(\frac{\Phi^{(1,0)}(t,s)}{f(t)}\Big)\Big)dt\ f(s)ds + O\Big(\frac{1}{mn^3h^3}\Big) \\
&\quad + \frac{1}{6mn^2}\int_{x-h}^{x+h}\Big(\int_{s}^{x+h}\frac{\partial}{\partial s}\Big(\frac{\Phi^{(1,0)}(t,s)}{f(t)}\Big)\Big)dt\ f(s)ds + O\Big(\frac{1}{mn^3h^3}\Big) \\
&= \frac{1}{6mn^2}\int_{x-h}^{x+h}\big(\Phi^{(1,0)}(s^-,s)-\Phi^{(1,0)}(s^+,s)\big)ds + O\Big(\frac{1}{mn^3h^3}\Big) \\
&\quad + \frac{1}{6mn^2}\int_{x-h}^{x+h}\Big(\frac{\Phi^{(1,0)}(x+h,s)}{f(x+h)} - \frac{\Phi^{(1,0)}(x-h,s)}{f(x-h)}\Big)f(s)\ ds + O\Big(\frac{1}{mn^3h^3}\Big).
\end{aligned}$$

Note that for $t\neq s$,

$$\Phi^{(1,0)}(t,s) = \Big(\frac{\varphi'_{x,h}(t)f(t)-\varphi_{x,h}(t)f'(t)}{f^2(t)}R(t,s) + \frac{\varphi_{x,h}(t)}{f(t)}R^{(1,0)}(t,s)\Big)\frac{\varphi_{x,h}(s)}{f(s)}. \tag{3.58}$$

It follows from (3.20) that,

$$\frac{\Phi^{(1,0)}(x+h,s)}{f(x+h)} = \frac{\Phi^{(1,0)}(x-h,s)}{f(x-h)} = 0 \quad \text{for all } s\in]x-h,x+h[.$$

Thus,

$$\sum\sum_{i\neq j=1}^{N_{T_n}-1} I_{i,j} = \frac{1}{6mn^2}\int_{x-h}^{x+h}\big(\Phi^{(1,0)}(t^-,t)-\Phi^{(1,0)}(t^+,t)\big)dt + O\Big(\frac{1}{mn^3h^3}\Big). \tag{3.59}$$

Inserting (3.44) and (3.59) in (3.36), we obtain,

$$\begin{aligned}
\mathrm{Var}(\hat g_n^{trap}(x)) = \frac{1}{m}\sigma_{x,h}^2 + \frac{1}{12mn^2}\int_{x-h}^{x+h}\big(\Phi^{(1,0)}(t^-,t)-\Phi^{(1,0)}(t^+,t)\big)\ dt \\
+ O\Big(\frac{1}{mn^3h^3}\Big).
\end{aligned} \tag{3.60}$$

Applying (3.58) it follows that,

$$\Phi^{(1,0)}(t^-,t)-\Phi^{(1,0)}(t^+,t) = \frac{\varphi_{x,h}^2(t)}{f^2(t)}\Big(R^{(1,0)}(t^-,t)-R^{(1,0)}(t^-,t)\Big) = \frac{\varphi_{x,h}^2(t)}{f^2(t)}\alpha(t). \tag{3.61}$$

Replacing (3.61) in (3.60) we obtain,

$$\mathrm{Var}(\hat{g}_n^{trap}(x)) = \frac{1}{m}\sigma_{x,h}^2 + \frac{1}{12mn^2}\int_{x-h}^{x+h}\frac{\varphi_{x,h}^2(t)}{f^2(t)}\alpha(t)\,dt + O\Big(\frac{1}{mn^3h^3}\Big). \tag{3.62}$$

Since α and f are continuous on $[0,1]$, then one can write,

$$\begin{aligned}\int_{x-h}^{x+h}\frac{\alpha(t)}{f^2(t)}\varphi_{x,h}^2(t)dt &= \frac{1}{h}\int_{-1}^{1}\frac{\alpha(x-th)}{f^2(x-th)}K^2(t)\,dt\\ &= \frac{1}{h}\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1}K^2(t)\,dt + \frac{1}{h}\int_{-1}^{1}\Big(\frac{\alpha(x-th)}{f^2(x-th)} - \frac{\alpha(x)}{f^2(x)}\Big)K^2(t)\,dt\\ &= \frac{1}{h}\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1}K^2(t)\,dt + O(1).\end{aligned} \tag{3.63}$$

Recall that for an even kernel, we have a simplified expression of $\sigma_{x,h}^2$ given by Benhenni and Rachdi (2007) as follows,

$$\sigma_{x,h}^2 = R(x,x) - \frac{1}{2}\alpha(x)C_K h + o(h), \tag{3.64}$$

where $C_K = \int_{-1}^{1}\int_{-1}^{1}|u-v|K(u)K(v)dudv$.
Finally, using (3.63) and (3.64) in (3.62) yields,

$$\begin{aligned}\mathrm{Var}(\hat{g}_n^{trap}(x)) = &\frac{1}{m}\Big(R(x,x) - \frac{1}{2}\alpha(x)C_K h\Big) + \frac{1}{12mn^2h}\frac{\alpha(x)}{f^2(x)}\int_{-1}^{1}K^2(t)\,dt\\ &+ o\Big(\frac{h}{m}\Big) + O\Big(\frac{1}{mn^2} + \frac{1}{mn^3h^3}\Big).\end{aligned}$$

This concludes the proof of Proposition 3.3.2. □

Proof of Proposition 3.3.3.

Let $I_1 = \int_0^1 R(x,x)w(x)\,dx$, $I_2 = \int_0^1 \frac{\alpha(x)}{f^2(x)}w(x)\,dx$ and put,

$$\Psi(h,m) = -\frac{C_K h}{2m}\int_0^1 \alpha(x)w(x)\,dx + \frac{1}{4}h^4B^2\int_0^1 [g''(x)]^2 w(x)\,dx.$$

We have from Equation (3.7) in Theorem 3.3.1,

$$\mathrm{IMSE}(h) = \frac{I_1}{m} + \Psi(h,m) + \frac{VI_2}{12mn^2h} + o\Big(h^4 + \frac{h}{m}\Big) + O\Big(\frac{1}{n^3h} + \frac{1}{mn^3h^3} + \frac{1}{mn^2} + \frac{1}{n^6h^6}\Big).$$

Let h^* be as defined in (3.8). It is clear that $h^* = \underset{0<h<1}{\mathrm{argmin}}\ \Psi(h,m)$ so that $\Psi(h,m) \geq \Psi(h^*,m)$ for every $0<h<1$. Let $h_{n,m}$ be as defined in Corollary 3.3.3. We have,

$$\begin{aligned}&\frac{\mathrm{IMSE}(h^*)}{\mathrm{IMSE}(h_{n,m})}\\ &= \frac{\frac{I_1}{m} + \Psi(h^*,m) + \frac{VI_2}{12mn^2h^*} + o\Big(h^{*4} + \frac{h^*}{m}\Big) + O\Big(\frac{1}{n^3h^*} + \frac{1}{mn^3h^{*3}} + \frac{1}{mn^2} + \frac{1}{n^6h^{*6}}\Big)}{\frac{I_1}{m} + \Psi(h_{n,m},m) + \frac{VI_2}{12mn^2h_{n,m}} + o\Big(h_{n,m}^4 + \frac{h_{n,m}}{m}\Big) + O\Big(\frac{1}{n^3h_{n,m}} + \frac{1}{mn^3h_{n,m}^3} + \frac{1}{mn^2} + \frac{1}{n^6h_{n,m}^6}\Big)}\\ &\leq \frac{I_1 + m\Psi(h_{n,m},m) + \frac{VI_2}{12n^2h^*} + o\Big(mh^{*4} + h^*\Big) + O\Big(\frac{m}{n^3h^*} + \frac{1}{n^3h^{*3}} + \frac{1}{n^2} + \frac{m}{n^6h^{*6}}\Big)}{I_1 + m\Psi(h_{n,m},m) + \frac{VI_2}{12n^2h_{n,m}} + o\Big(mh_{n,m}^4 + h_{n,m}\Big) + O\Big(\frac{m}{n^3h_{n,m}} + \frac{1}{n^3h_{n,m}^3} + \frac{1}{n^2} + \frac{m}{n^6h_{n,m}^6}\Big)}.\end{aligned}$$

Using the definition of h^*, $mh_{n,m}^3 = O(1)$, $\lim_{n,m\to\infty} h_{n,m} = 0$ and the assumption $\frac{m}{n} = O(1)$ as $n, m \to \infty$ we know that $m\Psi(h_{n,m}, m) = O(h_{n,m})$. Then,

$$\overline{\lim_{n,m\to\infty}} \frac{\text{IMSE}(h^*)}{\text{IMSE}(h_{n,m})} \leq 1.$$

This concludes the proof of Proposition 3.3.3. □

Proof of Corollary 3.3.1.

Let f^* be as defined in (3.9). Let $D(f) = \int_0^1 \frac{\alpha(x)}{f^2(x)} w(x)\, dx$, then it is sufficient to prove that:

$$D(f^*) \leq D(f) \quad \text{for every positive density } f \quad \text{on } [0,1].$$

Applying Hölder's inequality, we get,

$$\begin{aligned} D(f^*) &= \Big(\int_0^1 \{\alpha(x)w(x)\}^{1/3}\, dx \Big)^3 = \Big(\int_0^1 \Big(\frac{\alpha(x)w(x)}{f^2(x)} \Big)^{1/3} f^{2/3}(x)\, dx \Big)^3 \\ &\leq \Big(\int_0^1 \frac{\alpha(x)w(x)}{f^2(x)}\, dx \Big) \Big(\int_0^1 f(x)\, dx \Big)^2 = D(f). \end{aligned}$$

Hence,

$$\underset{\{f>0 \text{ density on } [0,1]\}}{\text{argmin}} D(f) = f^*.$$

This completes the proof of Corollary 3.3.1. □

Proof of Theorem 3.3.2.

Let f^* be as defined in (3.9). The proof of this theorem will be done in two steps:

1. $\sup\{\Psi_{(\alpha,w)}(f^*)/(\alpha, w) \in \Lambda\} \leq \epsilon_1\epsilon_2$.
2. $\forall f, \exists(\alpha, w) \in \Lambda : \Psi_{(\alpha,w)}(f) \geq \epsilon_1\epsilon_2$.

First step: By direct application of the Hölder's inequality we have:

$$\begin{aligned} \Psi_{(\alpha,w)}(f^*) &= \Big(\int_0^1 \{\alpha(s)w(s)\}^{1/3}\, ds \Big)^3 = \Big(\int_0^1 \alpha(s)^{1/3} \sqrt{w(s)}^{2/3}\, ds \Big)^3 \\ &\leq \Big(\int_0^1 \alpha(s) ds \Big) \Big(\int_0^1 \sqrt{w(s)} ds \Big)^2 \leq \epsilon_1\epsilon_2. \end{aligned}$$

Second step: Let f be an arbitrary positive density. Take $\alpha^* \equiv \epsilon_1$ and $w^* \equiv \epsilon_2$, then $(\alpha^*, w^*) \in \Lambda$ and:

$$\Psi_{(\alpha^*,w^*)}(f) = \int_0^1 \frac{\alpha^*(s)w^*(s)}{f^2(s)}\, ds = \epsilon_1\epsilon_2 \int_0^1 \frac{1}{f^2(s)}\, ds \geq \epsilon_1\epsilon_2\ ,$$

since, using the Hölder's inequality we have:

$$1 = \int_0^1 f^{2/3}(s) \Big(\frac{1}{f^2(s)} \Big)^{1/3} ds \leq \Big(\int_0^1 f(s)\, ds \Big)^{2/3} \Big(\int_0^1 \frac{1}{f^2(s)}\, ds \Big)^{1/3} = \Big(\int_0^1 \frac{1}{f^2(s)}\, ds \Big)^{1/3}.$$

This completes the proof of Theorem 3.3.2. □

Proof of Theorem 3.3.3.

Let $x \in]0,1[$ be fixed. We have,

$$\sqrt{m}\Big(\hat{g}^{trap}_{n,m}(x) - g(x)\Big) = \sqrt{m}\Big(\hat{g}^{trap}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{trap}_{n,m}(x)\big)\Big) + \sqrt{m}\ \mathrm{Bias}\Big(\hat{g}^{trap}_{n,m}(x)\Big). \tag{3.65}$$

Since $\lim\limits_{n,m\to\infty} \sqrt{m}h^2 = 0$ and $\lim\limits_{n,m\to\infty} nh^2 = \infty$ then Proposition 3.3.1 implies that,

$$\lim_{n,m\to\infty} \sqrt{m}\ \mathrm{Bias}\Big(\hat{g}^{trap}_{n,m}(x)\Big) = 0. \tag{3.66}$$

Consider now the first term of the right side of (3.65). Since $\overline{Y}(t_{x,i}) - \mathbb{E}(\overline{Y}(t_{x,i})) = \overline{\varepsilon}(t_{x,i})$, we have, as done by Fraiman and Pérez Iribarren (1991),

$$\begin{aligned}
\sqrt{m}\Big(\hat{g}^{trap}_{n,m}(x) - \mathbb{E}\big(\hat{g}^{trap}_{n,m}(x)\big)\Big) &= \frac{1}{\sqrt{m}}\Big\{\sum_{j=1}^{m}\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\Big(\big(\frac{\varphi_{x,h}}{f}\varepsilon_j\big)(t_{x,i}) + \big(\frac{\varphi_{x,h}}{f}\varepsilon_j\big)(t_{x,i+1})\Big)\Big\}\\
= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i})\big(\varepsilon_j(t_{x,i}) - \varepsilon_j(x)\big)&\\
+ \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i+1})\big(\varepsilon_j(t_{x,i+1}) - \varepsilon_j(x)\big)&\\
+ \Big(\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\Big(\frac{\varphi_{x,h}}{f}(t_{x,i}) + \frac{\varphi_{x,h}}{f}(t_{x,i+1})\Big)\Big)\Big(\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\varepsilon_j(x)\Big).&
\end{aligned} \tag{3.67}$$

We start by controlling the last term of this last equation. Recall that Equation (3.29) yields for some $t^*_{x,i} \in]t_{x,i}, t_{x,i+1}[$ that $\frac{1}{n} = (t_{x,i+1} - t_{x,i})f(t^*_{x,i})$. From the Riemann integrability of $\varphi_{x,h}$ and f and Lemma 3.5.1 we obtain,

$$\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\Big(\frac{\varphi_{x,h}}{f}(t_{x,i}) + \frac{\varphi_{x,h}}{f}(t_{x,i+1})\Big) =$$

$$\frac{1}{2}\sum_{i=1}^{N_{T_n}-1}\Big(\frac{\varphi_{x,h}}{f}(t_{x,i}) + \frac{\varphi_{x,h}}{f}(t_{x,i+1})\Big)f(t^*_{x,i})(t_{x,i+1} - t_{x,i}) \underset{m,n\to\infty}{\longrightarrow} \int_{-1}^{1} K(t)\ dt = 1.$$

where $d_{x,i} = t_{x,i+1} - t_{x,i}$ and $t^*_{x,i} \in]t_{x,i}, t_{x,i+1}[$. The Central Limit Theorem for i.i.d. variables yields,

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\varepsilon_j(x) \underset{m\to\infty}{\overset{\mathscr{D}}{\longrightarrow}} Z \quad \textit{where} \quad Z \sim \mathcal{N}(0, R(x,x)).$$

We shall prove now that the two first terms of Equation (3.67) tend to 0 in probability as n, m tends to infinity. We will only study the first term, the second one is treated analogously. Let,

$$A_{m,n}(x) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\frac{1}{2n}\sum_{i=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i})\big(\varepsilon_j(t_{x,i}) - \varepsilon_j(x)\big) \triangleq \frac{1}{\sqrt{m}}\sum_{j=1}^{m}T_{n,j}(x).$$

From the Chebyshev inequality, it suffices to prove that $\lim\limits_{n,m\to\infty}\mathbb{E}(A^2_{m,n}(x)) = 0$. We have for $j \neq l$, $\mathbb{E}(\varepsilon_j(x)\varepsilon_l(y)) = 0$ so $\mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = 0$. Hence,

$$\mathbb{E}(A^2_{m,n}(x)) = \frac{1}{m}\sum_{j=1}^{m}\sum_{l=1}^{m}\mathbb{E}(T_{n,j}(x)T_{n,l}(x)) = \frac{1}{m}\sum_{j=1}^{m}\mathbb{E}(T^2_{n,j}(x)).$$

We have,

$$
\begin{aligned}
&\mathbb{E}(T_{n,j}^2(x)) = \\
&\frac{1}{4n^2}\sum_{i=1}^{N_{T_n}-1}\sum_{k=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i})\frac{\varphi_{x,h}}{f}(t_{x,k})\mathbb{E}\Big(\big(\varepsilon_j(t_{x,i})-\varepsilon_j(x)\big)\big(\varepsilon_j(t_{x,k})-\varepsilon_j(x)\big)\Big) \\
&= \frac{1}{4n^2}\sum_{i=1}^{N_{T_n}-1}\sum_{k=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i})\frac{\varphi_{x,h}}{f}(t_{x,k})\Big(R(t_{x,i},t_{x,k})-R(t_{x,i},x)-R(x,t_{x,k})+R(x,x)\Big).
\end{aligned}
$$

Since $\mathbb{E}((T_{n,j}^2(x))$ does not depend on j we get,

$$
\begin{aligned}
&\mathbb{E}(A_{m,n}^2(x)) = \\
&\frac{1}{4n^2}\sum_{i=1}^{N_{T_n}-1}\sum_{k=1}^{N_{T_n}-1}\frac{\varphi_{x,h}}{f}(t_{x,i})\frac{\varphi_{x,h}}{f}(t_{x,k})\Big(R(t_{x,i},t_{x,k})-R(t_{x,i},x)-R(x,t_{x,k})+R(x,x)\Big) \\
&\triangleq \frac{1}{4}\Big(B_{n,1}(x)-B_{n,2}(x)-B_{n,3}(x)+B_{n,4}(x)\Big). \qquad (3.68)
\end{aligned}
$$

We obtain using Equation (3.29) for $t_{x,i}^* \in]t_{x,i}, t_{x,i+1}[$,

$$
B_{n,1}(x) = \sum_{i=1}^{N_{T_n}-1}\sum_{k=1}^{N_{T_n}-1} f(t_{x,i}^*)f(t_{x,k}^*)\ \frac{\varphi_{x,h}}{f}(t_{x,i})\frac{\varphi_{x,h}}{f}(t_{x,k})R(t_{x,i},t_{x,k})d_{x,i}d_{x,k}.
$$

The use of Lemma 3.5.1 twice yields,

$$
\begin{aligned}
B_{n,1}(x) &= \sum_{i=1}^{N_{T_n}-1} f(t_{x,i}^*)\frac{\varphi_{x,h}}{f}(t_{x,i})d_{x,i}\Big\{\int_{x-h}^{x+h}\varphi_{x,h}(t)R(t_{x,i},t)\ dt + O(\frac{1}{nh})\Big\} \\
&= \int_{x-h}^{x+h}\varphi_{x,h}(t)\Big\{\sum_{i=1}^{N_{T_n}-1} f(t_{x,i}^*)\frac{\varphi_{x,h}}{f}(t_{x,i})R(t_{x,i},t)d_{x,i}\Big\}dt + O(\frac{1}{nh}) \\
&= \int_{x-h}^{x+h}\int_{x-h}^{x+h}\varphi_{x,h}(s)\varphi_{x,h}(t)R(s,t)\ ds\ dt + O(\frac{1}{nh}) = \sigma_{x,h}^2 + O(\frac{1}{nh}).
\end{aligned}
$$

Using (3.64) we obtain,

$$
B_{n,1}(x) = R(x,x) - \frac{1}{2}\alpha(x)C_K h + o(h) + O(\frac{1}{nh}).
$$

where $C_K = \int_{-1}^1\int_{-1}^1 |u-v|K(u)K(v)dudv$. Since $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$. Thus,

$$
\lim_{n\to\infty} B_{n,1}(x) = R(x,x). \qquad (3.69)
$$

Consider now the term $B_{n,2}(x)$. We obtain using Lemma 3.5.1 twice,

$$\begin{aligned}
B_{n,2}(x) &= \int_{x-h}^{x+h}\int_{x-h}^{x+h} \varphi_{x,h}(s)\varphi_{x,h}(t)R(s,x)\ ds\ dt + O(\frac{1}{nh}) \\
&= \int_{x-h}^{x+h} \varphi_{x,h}(s)R(s,x)\ ds + O(\frac{1}{nh}) \\
&= \int_{-1}^{1} K(s)R(x-hs,x)\ ds + O(\frac{1}{nh}) \\
&= \int_{-1}^{0} K(s)R(x-hs,x)\ ds + \int_{0}^{1} K(s)R(x-hs,x)\ ds + O(\frac{1}{nh}).
\end{aligned}$$

For $s \in]-1,0[$, Taylor expansion of $R(\cdot,x)$ around x yields,

$$R(s,x) = R(x-sh,x) - shR^{(1,0)}(x+,x) + o(h).$$

Similarly for $s \in]0,1[$ we obtain,

$$R(x-sh,x) = R(x,x) - shR^{(1,0)}(x-,x) + o(h).$$

Thus,

$$\begin{aligned}
B_{n,2}(x) = R(x,x) &- hR^{(1,0)}(x+,x)\int_{-1}^{0} s\ K(s)\ ds \\
&- hR^{(1,0)}(x-,x)\int_{0}^{1} s\ K(s)\ ds + o(h) + O(\frac{1}{nh}).
\end{aligned}$$

Hence,

$$\lim_{n\to\infty} B_{n,2}(x) = R(x,x). \tag{3.70}$$

Similarly,

$$\lim_{n\to\infty} B_{n,3}(x) = R(x,x). \tag{3.71}$$

It is easy to see that,

$$\begin{aligned}
\lim_{n\to\infty} B_{n,4}(x) &= \lim_{n\to\infty} R(x,x) \sum_{i=1}^{N_{T_n}-1}\sum_{k=1}^{N_{T_n}-1} \frac{\varphi_{x,h}}{f}(t_{x,i})\frac{\varphi_{x,h}}{f}(t_{x,k})) \\
&= R(x,x)\Big(\int_{-1}^{1} K(t)\ dt\Big)^2 = R(x,x).
\end{aligned} \tag{3.72}$$

Inserting (3.69), (3.70), (3.71) and (3.72) in (3.68) yields,

$$\lim_{n,m\to\infty} \mathbb{E}(A^2_{m,n}(x)) = 0.$$

This concludes the proof of Theorem 3.3.3. □

Appendix

Lemma 3.5.1 (Integral approximation of a sum) *Let u and v be two Lipschitz functions on $[x-h, x+h]$, i.e, there exists two positive numbers l_1 and l_2 such that,*

$$|u(s) - u(t)| \le l_1|s-t|, \qquad |v(s) - v(t)| \le l_2|s-t|.$$

Let $t_{x,1} < \cdots < t_{x,N_{T_n}}$ be points in $[x-h, x+h]$ and put $d_{x,i} = t_{x,i+1} - t_{x,i}$. Then,

$$\sum_{i=1}^{N_{T_n}-1} u(t_{x,i})v(t'_{x,i})d_{x,i} = \int_{x-h}^{x+h} u(t)v(t)\ dt + \Delta_{n,h},$$

for any $t'_{x,i} \in [t_{x,i}, t_{x,i+1}]$ for all $i = 1, \cdots, n$ and for some appropriate positive constants c_1, c_2 and c_3,

$$|\Delta_{n,h}| \le c_1\ l_1 \frac{h}{n} \sup_{t\in[0,1]} |v(t)|\ + c_2\ l_2 \frac{h}{n} \sup_{t\in[0,1]} |u(t)| + 2\frac{c_3}{n} \sup_{\substack{t\in[x-h,t_{x,1}] \\ \cup[t_{x,N_{T_n}},x+h]}} |v(t)u(t)|.$$

Proof of Lemma 3.5.1. In fact, let $\Delta_{x,h} = A - B$ where,

$$A = \sum_{i=1}^{N_{T_n}-1} u(t_{x,i})v(t'_{x,i})d_{x,i} \quad \text{and} \quad B = \int_{x-h}^{x+h} u(t)v(t)\ dt.$$

We have,

$$B = \sum_{i=1}^{N_{T_n}-1} \int_{t_{x,i}}^{t_{x,i+1}} u(t)v(t)\ dt + \int_{x-h}^{t_{x,1}} u(t)v(t)\ dt + \int_{t_{x,N_{T_n}}}^{x+h} u(t)v(t)\ dt \triangleq B_1 + B_2,$$

where $B_2 = \int_{x-h}^{t_{x,1}} u(t)v(t)\ dt + \int_{t_{x,N_{T_n}}}^{x+h} u(t)v(t)\ dt$. On the one hand, since $(t_{x,1} - (x-h)) \le \sup_{1\le i\le n} d_{x,i}$ and $(x+h-t_{x,N_{T_n}}) \le \sup_{1\le i\le n} d_{x,i}$ we have,

$$|B_2| \le 2c_3 \sup_{\substack{t\in[x-h,t_{x,1}] \\ \cup[t_{x,N_{T_n}},x+h]}} |v(t)u(t)| \sup_{1\le i\le n} d_{x,i}.$$

On the other hand, we have,

$$\begin{aligned} A - B_1 &= \sum_{i=1}^{N_{T_n}-1} \int_{t_{x,i}}^{t_{x,i+1}} \big(u(t_{x,i})v(t'_{x,i}) - u(t)v(t)\big)\ dt \\ &= \sum_{i=1}^{N_{T_n}-1} v(t'_{x,i}) \int_{t_{x,i}}^{t_{x,i+1}} \big(u(t_{x,i}) - u(t)\big)\ dt + \sum_{i=1}^{N_{T_n}-1} \int_{t_{x,i}}^{t_{x,i+1}} u(t)\big(v(t'_{x,i}) - v(t)\big)\ dt. \end{aligned}$$

Since u and v are Lipschitz continuous we obtain,

$$|A - B_1| \le N_{T_n} \sup_{t\in[0,1]} |v(t)| l_1 \sup_{1\le i\le n} d^2_{x,i} + N_{T_n} \sup_{t\in[0,1]} |u(t)| l_2 \sup_{1\le i\le n} d^2_{x,i}.$$

Since $nh \geq 1$, Lemma 3.2.1 yields that $\sup\limits_{1 \leq i \leq n} d_{x,i} = O(\frac{1}{n})$ and $N_{T_n} = O(nh)$. Hence,

$$\begin{aligned}
|\Delta_{n,h}| &= |A - B| \leq |A - B_1| + |B_2| \\
&\leq c_1 \; l_1 \frac{h}{n} \sup_{t \in [0,1]} |v(t)| \; + c_2 \; l_2 \frac{h}{n} \sup_{t \in [0,1]} |u(t)| + 2 \frac{c_3}{n} \sup_{\substack{t \in [x-h, t_{x,1}] \\ \cup [t_{x,N_{T_n}}, x+h]}} |v(t) u(t)|.
\end{aligned}$$

This concludes the proof of Lemma 3.5.1. □

Chapter 4

Optimal design for the nonparametric regression estimation applied to pharmacokinetics problems

Abstract: Several problems of pharmacokinetics are investigated. The concentration-time curve estimation is considered using a nonparametric kernel estimator through a simulation study and real data analysis. For the Area Under the concentration Curve (AUC), we introduce a new kernel estimator and we show, through simulation study, that it outperforms the classical trapezoidal estimator in term of the estimation error. The problem of estimating the bioavailability is also considered. The crucial problem of finding the optimal sampling design for the AUC estimation is investigated using the General Simulated Annealing algorithm. The digoxin plasma concentration is used in the simulation studies, for both correlated and uncorrelated observations.

Key words: *Pharmacokinetics, concentration curve, nonparametric estimation, AUC, bioavailability, optimal sampling design, General Simulated Annealing Algorithm.*

Résumé: Plusieurs problèmes de pharmacocinétique sont traités. L'estimation de la fonction de concentration-temps est considérée en utilisant un estimateur à noyau, à travers une étude de simulation ainsi qu'une analyse des données réelles. Pour l'estimation de l'Aire Sous la Courbe de concentration (AUC), nous introduisons un nouvel estimateur à noyau et nous montrons, à travers une étude de simulation, que le nouvel estimateur est meilleur que l'estimateur trapezoidal classique en terme de l'erreur d'estimation. Le problème de l'estimation de la biodisponibilité est aussi considéré. Le problème crucial de l'obtention d'un plan d'échantillonnage optimal pour l'estimation de l'AUC est aussi traité, en utilisant l'algorithme de gradient conjugué ou l'algorithme de recuit simulé généralisé. La concentration plasmatique de digoxine est utilisée dans l'étude de simulation, pour donnés corrélées et non corrélées.

Mots clés: *Pharmacocinétique, courbe de concentration, estimation non paramétrique, AUC, biodisponibilité, plan d'échantillonnage optimal, algorithme de recuit simulé généralisé.*

4.1 Introduction

Some of the most important problems that pharmacokinetics researchers are brought to investigate, are the estimation of the concentration-time curve of some administrated drug (or any pharmacological agent) and the Area Under this Curve (AUC). This area represents the drug exposure of the organism over time and is critical in estimating the efficiency of absorption (or bioavailability), since the later involves ratios of this area.

In this paper, the digoxin plasma concentration is considered in different administration ways. The concentration function is presented by a sum of exponential terms. In particular, following an oral administration, the concentration is presented by a three exponential terms function, whereas after intravenous injection of the drug, only a two exponential concentration curve is taken. The number of terms describes the number of kinetically homogeneous compartments that the drug invades in the body. The compartmental model we considered is a classical model that mimic the dynamical processes of absorption, distribution and elimination of a drug in the body. In pharmacokinetics, the compartments are usually different tissue organs within which the concentration of a drug is assumed to be kinetically homogeneous. For more details on the compartmental models in pharmacokinetics, we refer the reader to the work of Shargel *et al.* (1970) and Gibaldi and Perrier (1975). We also focus, in the simulation study, on both correlated and uncorrelated observations, since in practice when measurements are taking from the same experimental units they are most likely to be correlated.

In pharmacokinetics studies, after the administration of a drug in the organism, blood samples are taken from subjects according to a specific sampling plan. For instance, blood samples are taken every half an hour in a period of 6 hours, but other sampling plans (or designs) maybe more efficient as it will be shown in this paper (section 4.5). When measurements of the drug concentrations from the blood samples are taken, they are not the exact values of the real concentration function, that we are aiming to estimate, but rather some approached values. For instance, if we observe the concentrations obtained from the same blood sample in Table 4.1, they are different from an assay to another due to some measurements errors. For this reason, statisticians add a residual component (called error) to the observed concentration function. Hence, we shall consider the so-called fixed design regression model given by,

$$Y(t_{i,n}) = g(t_{i,n}) + \varepsilon(t_{i,n}) \quad \text{for } i = 1, \cdots, n \tag{4.1}$$

where $Y(t_{i,n})$ is the observed drug concentration at time $t_{i,n}$, g is the unknown concentration function, the sampling design times $(t_{i,n})_i$ are such that $t_{1,n} < \cdots < t_{n,n}$ and ε is an error process.

In this chapter, several problems of pharmacokinetics are investigated. First, the estimation of the concentration curve g in Model (4.1) is considered. Scientists often use parametric methods to estimate the concentration curve, see for instance Gibaldi and Perrier (1975) for more details. However, these methods suppose that the concentration curve has a specific shape and that it depends on some parameters to be estimated. For the cases where no prior knowledge concerning g is possible so that only the estimation of g is of interest. Then, we propose a nonparametric approach to estimate the concentration curve g based on the kernel estimator given by Gasser and Müller (1979). Other nonparametric kernel estimators could be used, see for instance Cheng and lin (1981) and more recently Benelmadani *et al.* (2019a, 2019b) . We prove, through a simulation study and by considering a real data analysis, the good behaviour of the proposed estimator. The estimation of the AUC is also investigated, for this, we propose a new nonparametric estimator which is obtained by an integration of the nonparametric kernel estimator of g. This AUC estimator is shown, through a simulation study, to outperform the classical estimator, which is based on the trapezoidal rule for approximating an integral, in terms of the Mean Squared Error (MSE). The bioavailability estimation problem is also considered in this simulation study.

Finally, we investigate the crucial problem of finding the optimal sampling design times for the AUC estimation. In other words, the goal is to choose the best sampling times with respect to some criterion in order to estimate efficiently the AUC. This problem was first investigated by Chernoff (1953) and Box and Lucas (1959). Several approaches were proposed to obtain the optimal sampling times, to estimate the AUC by the trapezoidal approximation, using different objective functions and different algorithms. We refer to the work of Choi *et al.* (2007) for a review and a comparison of these methods.

In this work, we considered the MSE as an objective function, since in statistics it is an important tool to study the performance of an estimator. It measures, in average, the squared difference between the estimated values and what we seek to estimate. Then minimizing it with respect to sampling designs, would lead to a better estimate. For uncorrelated observations, the Conjugated Gradient Algorithm is used and for correlated observations, the General Simulating Annealing Algorithm is considered. The choice of the latest was essentially due to its ability to handle very complex non-linear objective functions with a very large number of local optima

This chapter is organized as follows. In Section 4.2, we propose the nonparametric kernel estimator of the concentration curve and we study, through a simulation experiment and a real data analysis, the performance of this estimator. In Section 4.3, the estimation of the AUC is considered by proposing a new kernel estimator. We conduct a simulation study to illustrate its good behaviour and to compare it to the classical trapezoidal estimator, for both correlated and uncorrelated observations. In Section 4.4, we consider the estimation of the bioavailability through a simulation study. In section 4.5, we investigate the problem of finding the optimal sampling points using MSE as the objective function and under appropriate algorithms. Finally, Section 4.6 presents some comments and a conclusion of this work.

4.2 Estimation of the concentration curve

In this section, we consider the estimation of the digoxin plasma concentration over time. We use both simulated data and numerical data set generated in a pharmaceutical study conducted by Wagner and Yates (1973). To estimate the concentration, we use the nonparametric kernel estimator given by Gasser and Müller (1979)and Hart and Wherly (1986) for $t \in [0, 1]$ as follows,

$$g_{n,h}^{\mathrm{GM}}(t) = \sum_{i=1}^{n} Y(t_{i,n}) \int_{m_{i-1,n}}^{m_{i,n}} \varphi_{t,h}(s)\ ds, \tag{4.2}$$

where $\varphi_{t,h}(s) = \frac{1}{h}K\big(\frac{t-s}{h}\big)$, K is a kernel of support $[-1, 1]$, $h = h(n, m)$ is a bandwidth with $0 < h < 1$ and the midpoints m_i are given by: $m_{0,n} = 0$, $m_{n,n} = 1$ and $m_{i,n} = (t_i + t_{i+1})/2$. In our analysis we consider the quartic kernel given by $K(u) = \frac{15}{16}(1 - u^2)^2 1_{|u|\leq 1}$.

4.2.1 Simulation study

We consider first the hypothetical model of digoxin plasma concentration after an oral dosage given in Wagner and Ayres (1977), by the following three exponential function:

$$g_{\mathrm{oral}}(t) = -2.4e^{-10t} - 2e^{-0.65t} - 0.4e^{-0.0146t} \quad \text{for} \quad t \in [0, 96]. \tag{4.3}$$

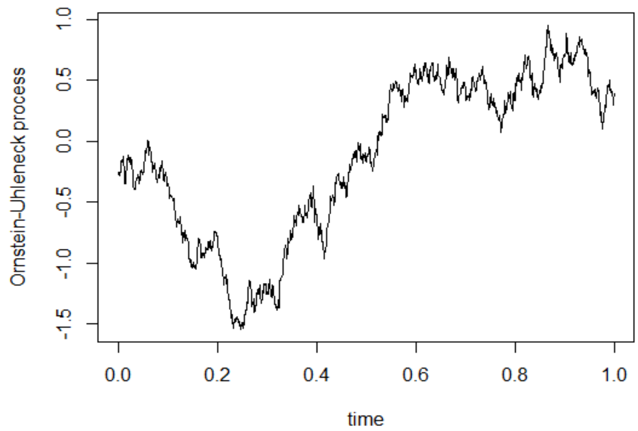


Figure 4.1: A trajectory of the Ornstein-Uhlenbeck process.

This model was used in several other works, see for instance, Katz and D'Argenio (1983), Piegorsh and Bailer (1989) and, more recently, by Belouni and Benhenni (2013), in order to estimate the AUC and to obtain an optimal sampling design using a classical estimator based on the trapezoidal rule. In our work, we still consider the same model for the same purpose, but in another context. In fact, in this chapter we use a nonparametric regression approach for estimating the concentration curve which does not require a specific form of this curve, whereas in the previous works, specific parametric models were imposed to the concentration function.
In our simulation study, the error process has an autocovariance function:

$$R(s,t) = \mathrm{Cov}(\varepsilon(s), \varepsilon(t)) = \sigma(s)\sigma(t)\rho(s,t), \tag{4.4}$$

where ρ is the autocorrelation function (specified later) and σ is the heteroscedastic standard deviation given, for a regression function g, by:

$$\sigma(t) = 0.05 + 0.1g(t).$$

We consider both uncorrelated errors generated by an error process with the correlation function:

$$\rho(s,t) = \begin{cases} 1 & \text{if } s = t, \\ 0 & \text{if } s \neq t, \end{cases} \tag{4.5}$$

and correlated errors, generated by the Ornstein-Uhlenbeck error process, with the correlation function:

$$\rho(s,t) = \mathrm{e}^{\,-\lambda|s-t|}. \tag{4.6}$$

In the sequel we take $\lambda = 1$. A visualization of this Ornstein-Uhlenbeck error process is displayed in Figure 4.1. We fixed the sampling points number $n = 13$ and considered the conventional sampling design $(t_i)_{1\leq i\leq n}$ given by Wagner and Ayres (1977) as follows:

$$0,\ 0.25,\ 0.5,\ 0.75,\ 1,\ 1.5,\ 3,\ 5,\ 12,\ 24,\ 48,\ 72,\ 96. \tag{4.7}$$

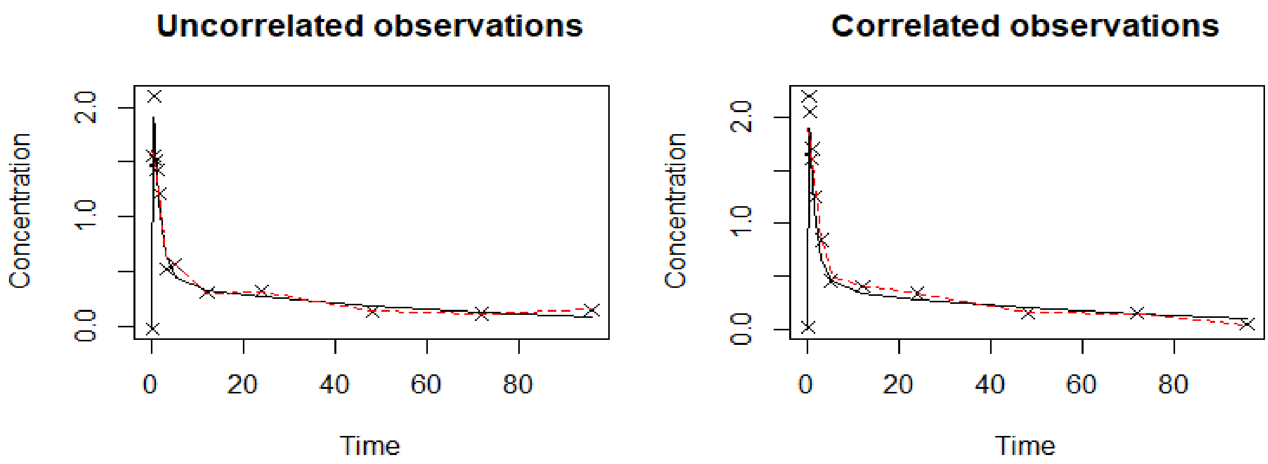


Figure 4.2: The observations are in black crosses, the estimator is in dashed red line and the concentration curve (4.3) is in plain line.

In this simulation study, we considered the mean of 10 vectors of generated observations following Model (4.1) where g is given by (4.3) and ρ is given by (4.5) or (4.6).
Based on the Gasser and Müller estimator $g_{n,h}^{\mathrm{GM}}$ of g_{oral}, the optimal bandwidth is obtained using the conjugate-gradient algorithms (CGA) (see Fletcher and Reeves (1946)), where the objective function is the Integrated Mean Squared Error (IMSE) given by,

$$\begin{aligned}\mathrm{IMSE}_n^{\mathrm{GM}}(h) &= \int_0^1 \mathbb{E}\left(g_{\mathrm{oral}}(t) - g_{n,h}^{\mathrm{GM}}(t)\right)^2 dt\\ &= \int_0^1 \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma(t_{i,n})\sigma(t_{j,n})\rho(t_{i,n},t_{j,n}) \int_{m_{i-1,n}}^{m_{i,n}} \int_{m_{j-1,n}}^{m_{j,n}} \varphi_{x,h}(s)\varphi_{x,h}(u)\, ds\, du\right.\\ &\quad \left. + \left(\sum_{i=1}^{n} (g(t_{i,n}) - g(t)) \int_{m_{i-1,n}}^{m_{i,n}} \varphi_{x,h}(s)\, ds\right)^2\right) dt,\end{aligned}$$

here the $\mathbb{E}(X)$ stands for the expectation of a random variable X. Figure 4.2 shows that the dashed curve which represents the estimator, approaches well the plain curve that represents the real concentration curve to be estimated. This visualisation shows that our nonparametric approach gives a good estimator for the concentration curve.

4.2.2 Real data analysis

In this section, we present our analysis of a data set considered by Wagner and Yates (1973). The data are digoxin plasma concentrations after an oral administration of a treatment which consists of a $0.25mg$ tablets of digoxin. The measurements were taken following a chosen time line as presented in Table 4.1. For one subject, duplicate assays were run on each plasma sample and we took the average of the two assays to present the concentrations.

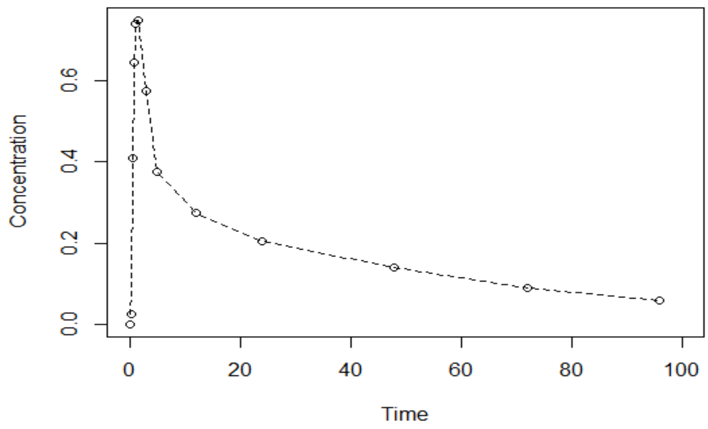


Figure 4.3: The observations in Table 4.1 are in circles and the estimator $g_{n,h}^{\mathrm{GM}}$ is in dashed line.

Table 4.1: Digoxin plasma concentration (mg/ml) in a subject administered orally.

Hours	0	0.25	0.5	0.75	1	1.5	3	5	12	24	48	72	96
assay 1	0	0.03	0.39	0.65	0.74	0.79	0.6	0.38	0.28	0.21	0.13	0.09	0.08
assay 2	0	0.02	0.43	0.64	$\cdots$	0.71	0.55	0.37	0.27	0.20	0.15	0.09	0.04

We use the estimator $g_{n,h}^{\mathrm{GM}}$ given by (4.2) with the quartic kernel defined above. The bandwidth we choose in our analysis is obtained using the cross validation criterion (see for instance Härdle (1990)), since this method is well-known in nonparametric regression problems. The results given in Figure 4.3 confirm those presented previously in our simulation study, in other words, the estimator $g_{n,h}^{\mathrm{GM}}$ tracks very faithfully the data.

Remark 4.2.1 *We note here that, in our simulation study, the function g_{oral} has been adjusted to use samples in $[0,1]$ as follows:*

$$\tilde{g}_{oral}(t) = -2.4e^{-10t\times 96} - 2e^{-0.65t\times 96} - 0.4e^{-0.0146t\times 96}.$$

In addition, in order to consider the boundary effect, that is to reduce the bias $\mathbb{E}(g_(n,h)^{GM}(t)) - g_{oral}(t)$ near the boundaries $[0,h]\cap[1-h,1]$, Hart and Wherly (1986) modified the estimator, so that it becomes a weighted average of the observations in the whole interval $[0,1]$, instead of the reduced interval $[h,1-h]$. This modification is done by taking the following kernel at the edges:

$$\tilde{K}(t) = K(t)\Big/\int_{t-h}^{t+h} K(u)du, \quad \textit{for} \quad t\in[0,h]\cap[1-h,1].$$

4.3 Estimation of the AUC

In this section we consider the problem of estimating the "partial" area under the concentration curve, i.e.,

$$\mathrm{AUC}(g) = \mathrm{AUC}^T(g) = \int_0^T g(t)dt, \tag{4.8}$$

where T is the investigator's last sampling time. For more details on the use of partial or total (from 0 to ∞) areas under the concentration curve, one can refer to the work of Lovering *et al.* (1975) and Wagner and Ayres (1977).

To estimate this area, scientists often use quadrature methods, see for instance the work of Katz and D'Argenio (1983), Belouni and Benhenni (2013). Their estimator is based on the trapezoidal numerical rule and is given by:

$$\widehat{\mathrm{AUC}}_n(g) = \frac{1}{2}\sum_{i=1}^{n-1}\Big\{Y(t_{i,n}) + Y(t_{i+1,n})\Big\}(t_{i+1,n} - t_{i,n}), \tag{4.9}$$

An intuitive estimator of AUC is the integral of the estimator of the concentration curve. Hence, we propose a new kernel estimator for $\mathrm{AUC}(g)$, which is obtained by integrating the estimator $g_{n,h}^{\mathrm{GM}}$, given by (4.2), as follows:

$$\widehat{\mathrm{AUC}}_{n,h}^{\mathrm{GM}}(g) = \sum_{i=1}^{n} Y(t_{i,n}) \int_0^1 \int_{m_{i-1,n}}^{m_{i,n}} \varphi_{t,h}(s)\ dsdt, \tag{4.10}$$

where $\varphi_{t,h}$ and the midpoints $m_{i,n}$ are defined in Section 2.

In order to compare the proposed estimator (4.10) to the trapezoidal estimator (4.9), we consider simulated observations from two different concentration functions. The first one represents the digoxin plasma concentration after an oral injection given by g_{oral} in (4.3), and the second one represents its concentration after an intravenous injection, given by Katz and D'Argenio (1983), by:

$$g_{\mathrm{int}}(t) = 3.117e^{-0.65t} + 0.6657e^{-0.0146t}, \tag{4.11}$$

where *int* stands for intravenous. We generate data from these two concentration functions (the mean of 100 vectors) with the autocovariance defined by (4.4), for both uncorrelated (4.5) and correlated (4.6) errors, and the conventional design (4.7). Likewise, we use the quartic kernel and the bandwidth obtained by the CGA, to estimate $\mathrm{AUC}(g_{\mathrm{oral}})$ and $\mathrm{AUC}(g_{\mathrm{int}})$ over the time interval $[0, 96]$. The results are displayed in Tables 4.2 and 4.3.

It is shown from these numerical results, in terms of MSE, that we can obtain smaller errors (up to 3% less in our simulation) and better estimations when using the proposed nonparametric kernel estimator, than the classical trapezoidal estimator, for both correlated and uncorrelated observations.

Table 4.2: The MSE, the expectation of AUC(g_{oral}) estimator $\pm$ the standard deviation SD, the optimal bandwidth and the estimation value under the correlations (4.5) and (4.6); AUC(g_{oral}) = 23.49.

$n = 13$	Estimator	MSE	$\mathbb{E}(\hat{\text{AUC}}) \pm SD$	Optimal h	Estimation
Uncorrelated	$\widehat{\text{AUC}}_n$	8.65	23.78 $\pm$ 2.93	-	23.79
observations	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	8.60	23.57 $\pm$ 2.93	0.092	23.57
Correlated	$\widehat{\text{AUC}}_n$	8.77	23.78 $\pm$ 2.95	-	23.71
observations	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	8.70	23.54 $\pm$ 2.95	0.098	23.47

Table 4.3: The MSE, the expectation of AUC(g_{int}) estimator $\pm$ the standard deviation SD, the optimal bandwidth and the estimation value under the correlations (4.5) and (4.6); AUC(g_{int}) = 39.17.

$n = 13$	Estimator	MSE	$\mathbb{E}(\hat{\text{AUC}}) \pm SD$	Optimal h	Estimation
Uncorrelated	$\widehat{\text{AUC}}_n$	12.65	39.81 $\pm$ 3.50	-	39.85
observations	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	12.26	39.20 $\pm$ 3.50	0.064	39.23
Correlated	$\widehat{\text{AUC}}_n$	12.92	39.81 $\pm$ 3.54	-	39.93
observations	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	12.48	39.15 $\pm$ 3.53	0.071	39.26

4.4 Estimation of the bioavailability

The absolute bioavailability is defined by the following formula:

$$F_{abs} = \frac{\text{AUC}_c}{\text{AUC}_i}\frac{D_i}{D_c},$$

where AUC_c and D_c are respectively the exposure and the dose of the medicament administrated in some chosen way (such as: oral, ocular, rectal $\cdots$), AUC_i and D_i are the exposure and the dose of the medicament administrated intravenously.
We conducted a simulation study to estimate the absolute bioavailability of the digoxin when the same dose is administrated in both ways. We considered the same simulated data presented in the previous section, i.e., the function g_{oral} defined by (4.3) to represent the concentration curve after oral administration and g_{int} defined by (4.11) for intravenous administration. The natural

Table 4.4: The estimated bioavailability (the real value of F_{abs} is 0.600).

$n = 13$	Estimator	Estimation
Uncorrelated	$\hat{F}_n$	0.597
observations	$\hat{F}_{n,h}^{\mathrm{GM}}$	0.601
Correlated	$\hat{F}_n$	0.594
observations	$\hat{F}_{n,h}^{\mathrm{GM}}$	0.598

estimations we used are:

$$\hat{F}_n = \frac{\widehat{\mathrm{AUC}}_n(g_{\mathrm{oral}})}{\widehat{\mathrm{AUC}}_n(g_{\mathrm{int}})}, \qquad \hat{F}_{n,h}^{\mathrm{GM}} = \frac{\widehat{\mathrm{AUC}}_{n,h}^{\mathrm{GM}}(g_{\mathrm{oral}})}{\widehat{\mathrm{AUC}}_{n,h}^{\mathrm{GM}}(g_{\mathrm{int}})}.$$

The results, given in Table 4.4, are a consequence of the AUC estimations presented in the previous section, i.e., better estimations of AUC give "slightly" better estimation of bioavailability.

4.5 Optimal design for AUC estimation

In this section, we are interested in finding the optimal sampling design with respect to the MSE for estimating the AUC(g_{oral}). The first and last sampling times were fixed to 0 and 96, the CGA was used to obtain the optimal design when the errors are uncorrelated. However, when the errors are correlated according to the autocovariance function (4.6), the CGA could not be applied since it requires the differentiablity of the objective function, we used instead the Generalized Simulated Annealing Algorithm (GSAA), see Xiang *et al.* (2013). This algorithm is essentially known for its ability to handle very complex non-linear objective functions with a very large number of optima. We used the optimal bandwidth for the GM estimator obtained in Section 4.3. The results are presented in Tables 4.5 and 4.6 and in Figure 4.4.

Tables 4.2 and 4.5 show that, we can significantly decrease the MSE up to 50% when using the optimal design instead of the conventional design, for instance when using the Gasser and Müller estimator we can reduce the MSE = 8.70 for the conventional design to MSE = 4.29 using the optimal design, which corresponds to decrease of 50%. Figure 4.4 and Table 4.6 show that most of the optimal sampling design points are located in the elimination phase and fewer are in the absorption phase and near the peak, in contrary to the conventional sampling design presented by (4.7).

4.6 Conclusion

In this work, a nonparametric regression method was applied to pharmacokinetics problems. In particular, we were interested in three major problems: the nonparametric estimation of the concentration curve, of its area and the derivation of the optimal sampling design. First, we used

Table 4.5: The MSE, the expectation of AUC(g_{oral}) estimator $\pm$ the standard deviation SD, using the optimal design and the estimation value under the correlations (4.5) and (4.6); AUC(g_{oral}) = 23.49.

$n = 13$	Estimator	MSE	$\mathbb{E}(\hat{\text{AUC}}) \pm SD$	Estimation
Uncorrelated observations	$\widehat{\text{AUC}}_n$	4.23	23.42 $\pm$ 2.06	23.42
	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	4.27	23.43 $\pm$ 2.07	23.51
Correlated observations	$\widehat{\text{AUC}}_n$	4.26	23.42 $\pm$ 2.06	23.42
	$\widehat{\text{AUC}}_{n,h}^{\text{GM}}$	4.29	23.43 $\pm$ 2.07	23.45

Table 4.6: Optimal sampling design obtained by CGA and GSAA for estimating $\widehat{\text{AUC}}(g_{\text{oral}})$.

	Uncorrelated errors (CGA)		Correlated errors (GSAA)	
	$\widehat{\text{AUC}}_n$	$\widehat{\text{AUC}}_n^{\text{GM}}$	$\widehat{\text{AUC}}_n$	$\widehat{\text{AUC}}_n^{\text{GM}}$
Absorption phase	0	0	0	0
	1.92	1.92	1.92	1.92
	8.64	6.72	7.68	6.72
Elimination phase	14.40	13.44	15.36	14.40
	24.00	22.08	22.08	22.08
	28.80	27.84	29.76	28.80
	40.32	38.40	38.40	37.44
	44.16	43.20	46.08	45.12
	57.60	56.64	55.68	54.72
	60.48	60.48	63.36	61.44
	76.80	76.80	74.88	73.92
	77.76	77.76	79.68	79.68
	96	96	96	96

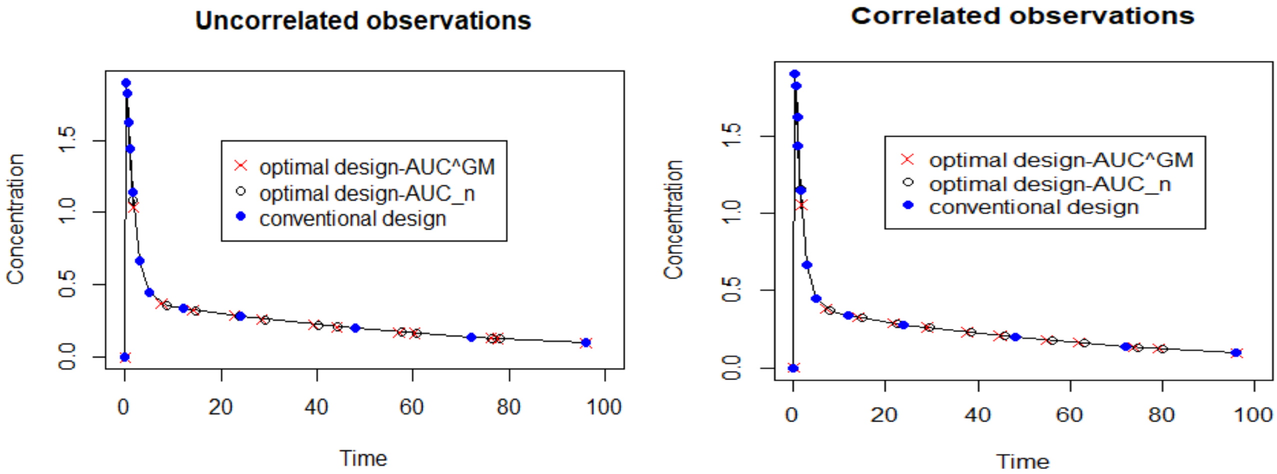


Figure 4.4: Comparison of the optimal designs and the conventional design for the estimation of AUC(g_{oral}).

the nonparametric kernel estimator to estimate the concentration curve, without the knowledge of its shape, in contrary to the parametric classical methods. We showed, through simulation studies and real data analysis, the good behaviour of the nonparametric kernel estimator.
We proposed, then, a new kernel estimator for the area under the concentration curve, which is constructed by integrating the concentration estimator. We showed by simulation study, that the proposed estimator outperforms the classical AUC estimator based on the trapezoidal rule, in the sense that it has a smaller MSE.
Finally, we investigated the problem of finding the optimal sampling design in the sense of minimizing the MSE with respect to all sampling design of size n. The generated optimal design enabled us to decrease significantly up to 50% the error of estimation (MSE), then the classical conventional design. Moreover, it is shown in our simulation study that there is no significant effect of correlated observations (through the auto covariance function) on the performance of our estimations in the pharmacokinetic problems that we investigated.
As noted by Katz and D'Argenio (1983), the proposed sampling points, in contrary to the conventional one, requires relatively large number of sampling times in the elimination phase and does not take into consideration the absorption phase and concentration peak. However, as observed by Belouni and Benhenni (2013), the selection of an optimal design method could be adjusted, by fixing certain sampling times in the absorption phase and around the peak and then obtain the remaining optimal sampling times.

ACKNOWLEDGEMENT: This work was developed in the framework of Grenoble Alpes Data Institute, supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).

Chapter 5

Conclusion and perspectives

In this thesis, we were interested in the nonparametric estimation of the regression function, in the case where the error process is of general autocovariance function, this includes the stationary and nonstationary processes. The model we considered is the fixed-design regression model with repeated measurements, where m experimental units are being observed each on n sampling points.

We began by considering a well known kernel regression estimator, proposed by Gasser and Müller. We studied its asymptotic behavior when the number of experimental units and the number of observations tend to infinity. We focused our studies in the case of the regular design sequence where the sampling points are generated using a specific design density.

We then proposed two new kernel estimators for the regression function. The first one is called the projection estimator, it was constructed through a projection property and using the Reproducing Kernel Hilbert Spaces propoerties. The second one is named the trapezoidal estimator, constructed using a numerical rule. We studied their asymptotic behaviors when n and m tend to infinity, proved their asymptotic normality and derived their optimal asymptotic bandwidths. We also obtained the optimal regular sampling design for the trapezoidal estimator. We then conducted a simulation studies to test their performances in a finite sample sets, and to compare them to the classical Gasser and Müller estimator. Our simulations confirmed our theoretical results.

Finally, we considered an application to pharmacokinetics, where we proposed the use of the nonparametric regression estimators to estimate the concentration function of a given drug. We proposed a new kernel estimators to estimate the Area Under the concentration Curve (AUC) and we showed its good performances via a simulation study and a real data analysis. We also consider the problem of obtaining the optimal sampling design for the AUC estimation using the General Simulating Annealing algorithm.

During the preparation of this thesis, some issues and questions appeared, leading to some future research, that we develop hereafter.

Models with dependent experimental units

In our thesis work, we considered the kernel regression model, where the errors $\{\varepsilon_j, j = 1, \cdots, m\}$ are i.i.d. processes with the same distribution as a centered, second order process ε. The independence of the experimental units is a realistic assumption, for instance, in the longitudinal data when observing the heights of children over the years. However, if we consider a situation where we measure some variable on a group of students, the students who were in particular classes in particular schools tend to be more similar and hence dependent. For this, it is also interesting to consider the kernel regression model with repeated, dependent observations and generalize our studies.

Different sampling times for each experimental units

During this thesis, we considered the nonparametric regression model with repeated measurements, where the observations are made on the same sampling times $(t_{i,n})_{i=1,\cdots,n}$ for each experimental unit. This is the case when choosing the sampling times is possible, for instance, to estimate the concentration-time curve, the scientists fix the sampling time prior to the experiment. However, imagine a set of panel count data, where the experimental units are observed continuously and for each subject, only the number of occurrences of the event are known at a finite distinct observation time points. Hence the times may vary from an experimental unit to another. Sun *et al.* (2007) considered this type of data and studied the problem of estimating the parameters of regression function. Núñez-Antón *et al.* (1991) developed a three-stage approach to estimate the regression function for such type of data. In the nonparametric case, one possible extension of the thesis work is to consider the following nonparametric regression model:

$$Y_j(t_{i,j}) = g(t_{i,j}) + \varepsilon_j(t_{i,j}) \quad \text{for} \quad j = 1, \cdots, m \quad \text{and } i = 1, \cdots, n_j,$$

where $\{\varepsilon_j, j = 1, \cdots, m\}$ are i.i.d. with the same distribution as a centered, second ordered error processe. An intuitive kernel estimator one can use to estimate the regression function g is given, for $x \in [0, 1]$ by:

$$\hat{g}_{n,m}(x) = \sum_{j=1}^{m} \sum_{i=1}^{n_j} W_{i,j}(x) Y_j(t_{i,j}),$$

where $W_{i,j}$ are some precised weights. One can use the Gasser and Müller, the projection or the trapezoidal weights. It is interesting to study the asymptotic behavior of this estimator, when the number of experimental units m and the numbers of observations n_j for $j = 1, \cdots, m$ tend to infinity.

Data based bandwidth selection methods for correlated errors

The data based bandwidth selection methods are widely developed for uncorrelated observations. It has been shown that many of them, including the well know cross-validation, breakdown when dealing with correlated observation, see for instance Altman (1990), Chiu (1989) and Hart (1991, 1994). Opsomer *et al.* (2001) explained the practical consequences of the sensitivity in the presence of correlation.

A very interesting problem to improve the performance of the kernel estimator, is than to propose a selection method to obtain the optimal bandwidth in the presence of correlation, which is the topic of our current project "Optimal trend estimation under errors with Matérn-type Autocovariance and applications to environmental data", funded by Grenoble Alpes Data Institute.

Estimating the change points for a non-differentiable regression function

A classical hypothesis that we made on the nonparametric regression model, is that the regression function is twice differentiable. This was useful, especially when studying the asymptotic performances of the kernel estimators. This assumption, as we have seen by many examples, is realistic for many type of data, such as: the concentration-time curve and the heights of children. Sometimes, even a smooth function may contain a a discontinuity, or what is called a change point. For instance, Cobb (1978) considered the annual volume of the Nile river from 1871-1970, where there was an abrupt change in the rainfall activity near the end. He suggested that the change was occurred in the year of 1898. Müller (1992) proposed a new method to estimate the change point while considering that the observations are i.i.d. Many other authors have focused on estimating the change point for uncorrelated error, and hence the problem for correlated observations is still an open question.

Nonparametric regression estimation for more regular autocovariance function

In our work, we assumed that the error process is a second order process, with a non-differentiable covariance function. This is the case, for instance, for the Wiener process and the Ornstein-Uhlenbeck process. The work can be brought to other processes with a more regular autocovariance functions, such as the iterated Brownian motion, which was considered by Benhenni *et al.* (2013) to study the effect of the regularity on the Gasser and Müller estimator.

Bibliography

[1] Altman, N. S. (1990) Kernel Smoothing of Data with Correlated Errors. American Statistical Association, 85. 749-759.

[2] Amemiya, T. (1985) Advanced Econometrics. Harvard University Press, Cambridge, MA.

[3] Anderson, T.W. (1960). Some Stochastic Process Models For Intelligence test scores. Mathematical Methods in Social Sciences. Stanford University. Press, Stanford, CA, 205-220.

[4] Aronszajn, N. (1944). La Theorie des Noyaux Reproduisants et ses Applications. Mathematical Proceedings of the Cambridge Philosophical Society, 39. 133-153.

[5] Azzalini, A. (1984). Estimation and Hypothesis Testing For Collections of Autoregressive Time Series. Biometrika, 71. 85-90.

[6] Bailer, A.J., Piegorsch, W.W. (1990). Estimating integrals using quadrature methods with and application in pharmacokinetics. Biometrics, 46. 1201-1211.

[7] Belouni, M., Benhenni K. (2015). Optimal and robust designs for estimating the concentration curve and the AUC. Scandinavian Journal of Statistics: Theory and Application. 42 453-470.

[8] Benedetti, J. (1977). On the Nonparametric Estimation of the Regression Function. Journal of the Royal Statistical Society, 39. 248-253.

[9] Benelmadani, D., Benhenni, K., Louhichi, S. (2019a). Trapezoidal rule and sampling desings for the nonparametric estimation of the regression function in models with correlated errors. https://arxiv.org/abs/1806.04896

[10] Benelmadani, D., Benhenni, K., Louhichi, S. (2019b). The reproducing kernel Hilbert space approach in nonparametric regression problems with correlated observations. Annals of the Institute of Statistical Mathematics.

[11] Benhenni, K., Cambanis, S. (1992). Sampling Designs for Estimating Integrals of Stochastic Processes. The Annals of Statistics, 20. 161-194 .

[12] Benhenni, K., Rachdi, M. (2007). Nonparametric Estimation of Average Growth Curve with General Nonstationary Error Process. Communications in Statistics: Theory and Methods, 36. 1137-1186.

[13] Benhenni, K., Rachdi, M., Su, Y. (2013). The effect of the regularity of the error process on the performance of the kernel regression estimators. Metrika, 76. 765-781.

[14] Bickel, P.J., Herzberg A.M. (1979) Robustness of design against autocorrelation in time, I. Asymptotic theory, optimality for location and linear regression. Annals od statistics, 7. 77-95.

[15] Biedermann, S., Dette, H. (2001) Minimax optimal design for nonparametric regression: a further optimality property of the uniform distribution. MODA 6: Advances in Model-Oriented Design and Analysis. (Atkinson, A.C., Hackl, P., Müller, W.G.,eds). Physica-Verlag, Heidelberg. 13-20.

[16] Blanke, D., Bosq, D. (2008). Regression Estimation and Prediction in Continuous Time. Journal of the Japan Statistical Society (Nihon Tôkei Gakkai Kaihô), 38. 15-26.

[17] Box, G.E.P., Lucas, H.L. (1959). Design of Experiments in non linear situations. Biometrika, 46. 77-90.

[18] Cheng, K.F., Lin, P.E. (1981). Nonparametric Estimation of the Regression Function. Z. Wahrscheinlichkeitstheorie verw. Gebiete, 57. 223-233.

[19] Chernoff, H. (1953). Locally optimum designs for estimating parameters. Annals of Mathematical Statistics, 24. 586-602.

[20] Cobb, G.W. (1978). The problem of the Nile: Conditional solution to a change point problem. Biometrika, 65. 243-251.

[21] S. T. Chiu (1989). Bandwidth Selection for Kernel Estimation with Correlated Noise. Statistics and Probability Letters, 8. 347-354.

[22] Choi, L., Caffo, B., Rohde, C. (2007). Optimal sampling times in bioequivalence studies using a simulated annealing algorithm. Statistics and Computing, 17. 337-347

[23] Davidian, M, Giltinan, D.M. (1995). Nonlinear models for repeated measurement data. Chapman and Hall.

[24] Dette, H., Pepelyshev, A., Zhigljavsky A. (2016). Optimal designs in regression with correlated errors. The Annals of Statistics, 44. 113-152.

[25] Dette, H., Schorning, K., Konstantinou M., (2017). Optimal designs for comparing regression models with correlated observations. Computational Statistics and Data Analysis. `http://dx.doi.org/10.1016/j.csda.2016.06.017`.

[26] Didi, S., and Louani, D. (2013). Asymptotic Results for the Regression Function Estimate on Continuous Time Stationary and Ergodic Data. Journal of Statistics and Risk Modelling, 31. 129-150.

[27] Duffull, S.B., Retout, S., Mentré, F. (2002). The use of simulated annealing for fitting optimal population designs. Computer Methods and Programs in Biomedicine, 69. 25-35.

[28] Efromovich, S. (2008). Optimal sequential design in a controlled nonparametric regression. Scandinavian Journal of Statistics, 35. 226-285.

[29] Fan, J. (1992). Design-adaptive nonparametric regression. Journal of the American Statistical Association, 87. 998-1004.

[30] Faraway J.J. (1990). Sequantial Design for the nonparametric regression of curves and surfaces. Computer Science and Statistics: Proceedings of the 22nd Annual Symposium on the Interface, Michigan State University, MI. 104-110.

[31] Ferreira, E., Núñez-Antón, V., Rodríguez-Póo, J. (1997). Kernel Regression Estimates of Growth Curves Using Nonstationary Correlated Errors. Statistics & Probability Letters, 34. 413-423.

[32] Fletcher, R. Reeves, C.M. (1946). Function minimization by conjugate gradients. The Computer Journal, 7. 149-154.

[33] Fraiman, R., Pérez Iribarren, G. (1991). Nonparametric Regression in Models with Weak Error's Structure. Journal of Multivariate Analysis, 37. 180-196.

[34] Gasser, T., Müller, H.G. (1979). Kernel Estimation of Regression Functions. Lecture Notes in Mathematics, 757. 23-68.

[35] Gasser, T., Müller, H-G. (1984). Estimating Regression Function And Their Derivatives By The Kernel Method. Scandinavian Journal of Statistics, 11. 171-185.

[36] Gasser, T., Müller, H.G. Köhler, W., Molinari, L. and Prader, A. (1984). Noparametric regression analysis of growth curves. The annals of statistics 12, 210-229.

[37] Gibaldi, M., Perrier, D. (1975). Pharmacokinetics. Drugs and Pharmaceutical sciences. Second edition.

[38] Ghosh, M., Grizzle. J.E., Sen, P.K (1973). Nonparametric methods in longitudinal studies. American Statistical Association, 68. 29-36.

[39] Grizzle, J.E., Allen. D.M. (1969). Analysis of growth and dose response curves. Biometrics 25, 357-381.

[40] Härdle W. (1990). Applied Nonparametric Regression. Econometric Society Monographs 19.

[41] Hart, J.D., Wherly. T.E. (1986). Kernel Regression Estimation Using Repeated Measurements Data. Journal of the American Statistical Association 81, 1080-1088.

[42] Hart, J.D. (1991). Kernel Regression Estimation with Time Series Errors. Royal Statistical Society B. 53 173-187.

[43] Hart J,D.(1994). Automated Kernel Smoothing of Dependent Data by Using Time Series Cross Validation. Royal Statistical Society B. 56 529-542.

[44] Katz D., D'Argenio. (1983). Experimental design for estimating integrals by numerical quadrature with applications to pharmacokinetc studies. Biometrics, 391, 621-628.

[45] Lovering E.G., McGilveray I.J., McMillan I., Tostowaryk W. (1975). Comparative bioavailabilities from truncated blood level curves. Journal of Pharmaceutical Sciences, 64, 1521-1524.

[46] Müller, G.H. (1984). Optimal Designs For Nonparametric Kernel Regression. Statistics And Probability Letters 2, 285-290.

[47] Müller, G.H. (1991). Smooth Optimum Kernel Estimators Near Endpoints . Biometrika. 78,3, 521-530.

[48] Müller, G.H. (1992). Change-points in nonparametric regression analysis. The Annals of Statistics. 20. 737-761.

[49] Núñez-Antón V. (1997). Longitudinal data analysis: Non-stationary error structures and antedependent models. Applied Stochastic Models and Data Analysis, 13. 279-287.

[50] Núñez-Antón, V. Rodriquz-Póo, J.M., Vieu, P. (1999). Longitudinal Data with nonstationary errors: a nonparametric three-stage approach. Sociedad de Estadística e Investigación Operativa. Test 8. 201-231.

[51] Núñez-Antón, V. Woodworth, G. (1994). Analysis of Longitudinal Data with Unequally Spaced Observations and Time-Dependent Correlated Errors. BIOMETRICS, 50. 445-456.

[52] Opsomer, J., Wang, Y., Yang, Y. (2001). Nonparametric Regression with correlated errors. Statistical Science, 16. 134-153.

[53] Parzen, E. (1959). Statistical Inference on Time Series by Hilbert Space Methods. Department of Statistics. Stanford University. Technical Report. 23.

[54] Piegorsh W.W., Bailer A.J. (1989). Optimal design allocations for estimating area under curves for studies employing destructive sampling. Journal of Pharmacokinetics and Biopharmaceutics, 17, 423-507.

[55] Potthoff, R.F., Roy, S.N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika, 51. 313-326.

[56] Priestly, M.B., Chao, M.T. (1972). Nonparametric function fitting. Journal of Royal Statistical Society, 34. 384-392.

[57] Ramsay, J.O., Silverman, B.W. (2005). Functional Data Analysis. Springer Series in Statistics.

[58] Rao, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, 52. 447-458.

[59] Rao, C.R. (1966). Covariance adjustment and related problems in multivariate analysis. Multivariate analysis. Academic press, New York,87-103.

[60] Sacks, J., Ylvisaker, D. (1966). Designs For Regression Problems With Correlated Errors. The Annals of Mathematical Statistics. 37, 66-89.

[61] Sacks, J., Ylvisaker, D. (1968). Designs For Regression Problems With Correlated Errors: many parameters. The Annals of Mathematical Statistics. 39, 69-79.

[62] Sacks, J., Ylvisaker, D. (1970). Designs For Regression Problems With Correlated Errors III. The Annals of Mathematical Statistics. 41, 2057-2074.

[63] Schwartz, L. (1964). Sous Espace Hilbertiens d'Espace Vectoriels Topologiques et Noyaux Associés (Noyaux Reproduisant) . Journal d'Analyse Mathématiques, 13. 115-256.

[64] Shargel, L. Wu-Pong, S. Andrew, B.C. (1970). Applied Biopharmaceutics and Pharmacokinetics, sixth edition.

[65] Su, Y., Cambanis, S. (1993). Sampling Designs for Estimation of a Random Process. Stochastic Processes and their Applications, 46. 47-89.

[66] Sun, J., Tong, X., He, X. (2007). Regression analysis of a panel count data with dependent observation times. Stochastic Processes and their Applications, 63. 1053-1059.

[67] Wagner, J.G., Yates, J.D. (1973). Equivalence Lack in Digoxin Plasma Levels. The Journal of the American Medical Association 2, 199-204.

[68] Wagner, J.G., Ayres, J.W. (1977). Bioavailability assessment: Methods to estimate total area and total amount excreted and importance of blood and urine sampling scheme with application to digoxin. Journal of pharmacokinetics and biopharmaceutics, 5. 533-557.

[69] Xiang, Y. Gubian S, Suomela B, Hoeng, J. (2013). Generalized Simulated Annealing for efficient global optimization: the GenSA Package for R. The R Journal, 5. 13-28.

[70] Zimmerman D L, Núñez-Antón V, El Barmi H. (1998). Computational aspects of likelihood-based estimation of first-order antedependencc models, J. Statist. Computut. Simul, 60. 67-84.

[71] Zhao, Z., Yao W. (2012). Sequential design for nonparametric inference. The Canadian Journal of Statistics, 40. 362-377.

[72] Zhigljavsky, A. Dette, H. Pepelyshev, A. (2010). A new approach to optimal design for linear models with correlated observations. Journal of the American Statistical Association, 105. 1093-1103.