symmetry

MDPI

Article

Spatio-Functional Local Linear Asymmetric Least Square Regression Estimation: Application for Spatial Prediction of COVID-19 Propagation

Ali Laksaci [1], Salim Bouzebda [2,*], Fatimah Alshahrani [3], Ouahiba Litimein [4] and Boubaker Mechab [4]

1 Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; alikfa@kku.edu.sa
2 Laboratoire de Mathématiques Appliquées de Compiègne (L.M.A.C.), Université de Technologie de Compiègne, 60200 Compiègne, France
3 Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia; fmalshahrani@pnu.edu.sa
4 Laboratory of Statistics and Stochastic Processes, University of Djillali Liabes, BP 89, Sidi Bel Abbes 22000, Algeria; ouahiba.litimein@univ-sba.dz (O.L.); mechaboub@yahoo.fr (B.M.)
* Correspondence: salim.bouzebda@utc.fr

Abstract: The problem of estimating the spatio-functional expectile regression for a given spatial mixing structure $(X_{\mathbf{i}}, Y_{\mathbf{i}}) \in \mathcal{F} \times \mathbb{R}$, when $\mathbf{i} \in \mathbb{Z}^N, N \geq 1$ and $\mathcal{F}$ is a metric space, is investigated. We have proposed the M-estimation procedure to construct the Spatial Local Linear (SLL) estimator of the expectile regression function. The main contribution of this study is the establishment of the asymptotic properties of the SLL expectile regression estimator. Precisely, we establish the almost-complete convergence with rate. This result is proven under some mild conditions on the model in the mixing framework. The implementation of the SLL estimator is evaluated using an empirical investigation. A COVID-19 data application is performed, allowing this work to highlight the substantial superiority of the SLL-expectile over SLL-quantile in risk exploration.

Keywords: functional data; small ball probability; local linear estimation; Kolmogorov entropy; complete convergence; asymmetric least square regression

Citation: Laksaci, A.; Bouzebda, S.; Alshahrani, F.; Litimein, O.; Mechab, B. Spatio-Functional Local Linear Asymmetric Least Square Regression Estimation: Application for Spatial Prediction of COVID-19 Propagation. *Symmetry* **2023**, *15*, 2108. https://doi.org/10.3390/sym15122108

Academic Editor: Marek T. Malinowski

Received: 28 August 2023
Revised: 10 November 2023
Accepted: 20 November 2023
Published: 23 November 2023

1. Introduction

Spatial data is commonly generated in multiple fields of study such as econometrics, epidemiology, environmental science, image analysis, oceanography, meteorology, geostatistics, and others. Generally, the collection of this data occurs across various disciplines and is subsequently subjected to statistical analysis at designated measurement sites. Please refer to [1–5] in order to gain insights into various statistical applications. It is important to emphasize the significance of including a spatio-temporal framework in the modeling of some real problems. In this study, we employ the latest advancements in spatio-functional statistics to propose the Local Linear Free-Distribution (LLFD) modeling of Spatio-Functional Chronological Series Data (SFCSD).

In the context of nonparametric estimation for spatial data, the existing papers are mainly concerned with the estimation of probability density and regression functions; we cite a key reference: Tran [6]. He gave the asymptotic normality of the probability density function by the kernel estimation. Ref. [7] introduced a kernel method to estimate a spatial conditional regression under mixing spatial processes and investigated weak consistency and convergence rates. The general problem of the regression estimation for random fields is examined by [8]. The authors showed the uniform consistency on compact sets of the proposed spatial predictor as well as its asymptotic normality. Alternatively to the kernel Nadaraya–Watson, the LLFD was introduced by [9]. Under mild regularity assumptions, the authors established the asymptotic normality of the proposed estimator and its derivatives. The auto-regression function was investigated by [10]. The authors established the uniform convergence on compact sets under general conditions and the optimal rates of convergence in L_∞, while the spatial LLFD estimation was considered in [11]. In the same way, Li and Tran [12] combined the LLFD estimation with the nearest neighbor algorithm. For recent references on the topic, we refer to [13,14]. Concerning the SFCSD case, the initial exploration was conducted by [15]. In the last reference, the weak and strong consistencies of the estimate together with almost-sure rates of convergence are established. For further asymptotic results on this operator, one can refer to [16,17], while, for other functional models such as the modal regression and/or the quantile regression, we refer to [18–20]. Ref. [5] developed an asymptotic theory of conditional U-statistics for locally stationary random fields. The authors employed a stochastic sampling scheme that may create irregularly spaced sampling sites in a flexible manner and include both pure and mixed increasing domain frameworks.

In this paper, we investigate conditional expectile, which is based on the least asymmetrically weighted squares estimation, which was adopted from the econometrics literature and is one of the fundamental statistical application tools. This method frequently employs the [21] concept of expectiles, the least squares equivalent of the conventional quantiles. They were given this name because they resemble the quantiles of a random variable, but, unlike quantiles, they are based on a quadratic loss function, as in the case of the expectation; see [22–25] for more information. The expectile regression function has various uses in insurance, finance, and economics. In particular, it is used to assess the uncertain prospective positions of outcomes. The first investigations in this model were introduced in [26]. They utilized the parametric techniques to provide an estimator of the expectile model in unconditional and uni-dimensional cases. In the finite-dimensional case, the expectile operator was elaborated by [25] for the i.i.d. case and [27] for the strong mixing case. More alternative functional times series cases and or smoothing algorithms were developed in the literature for functional statistics. Such studies include the ergodic case in [28] and the k number of neighborhoods in [29]. In [30], a modification of ranked set sampling called moving extremes ranked set sampling is considered for the best linear unbiased estimator for the simple linear regression model. It is worth noting that the modeling of functional data has increasingly become an appealing avenue of research in mathematical statistics. This research direction has been popularized through numerous monographs or journal special issues (see, for instance, [31–33]). In this context, various regression models are introduced to appropriately fit this kind of data. We mention, for instance, the linear regression [34], the single index functional model (see [35], the classical regression [36], the functional partial linear regression (see [37], and the relative error regression [38]. For more recent references on the subject, refer to [39]. However, all the aforementioned models control the co-variability between the input and output variables through the central tendency. The expectile regression model fits this co-variability in a more comprehensive manner, allowing one to control the center as well as the tails of the data.

In this work, we investigate the spatio-functional estimation using the LLFD algorithm. We demonstrate the almost-complete convergence (a.c.c) rate of the constructed expectile regression estimators. We establish these results under general conditions, allowing the consideration of several particular situations. For instance, the strong mixing case is a special case of our spatial setting and the kernel method is a particular case of local linear strategy. This theoretical development has many applied derivatives, for example in financial risk assessment. It constitutes a good financial risk tool, such as for liquidity risk in banking or market risk in stock exchanges. The effectiveness of the proposed estimator is evaluated using a real data model and empirical data analysis.

The layout of the article is as follows. We present the spatio-functional model in Section 2. In Section 3, we specify the necessary conditions for the main results. The convergence rate of the proposed estimator is presented in Section 4. Section 5 is dedicated to discussing the computationability of the constructed expectile regression estimators.

Section 6 presents some concluding remarks. To prevent interrupting the flow of the presentation, all proofs are gathered in Appendix A.

2. Methodology

2.1. The Spatio-Functional Structure

Let $(\mathcal{X}_{\vec{\mathbf{i}}}, \mathcal{Y}_{\vec{\mathbf{i}}})$, $\vec{\mathbf{i}} \in \mathbb{Z}^N$ be a functional random field valued in $\mathcal{F} \times \mathbb{R}$. The functional space $(\mathcal{F}, Dis)$ is structured as a semi-metric space with distance Dis. Furthermore, let N be a nonnegative integer in $\mathbb{N}^*$ and suppose that $(\mathcal{X}_{\vec{\mathbf{i}}}, \mathcal{Y}_{\vec{\mathbf{i}}})$ is observed over a polyhedron area expressed by

$$I_{\vec{\mathbf{n}}} = \left\{ \vec{\mathbf{i}} = (i_1, \dots, i_N) \in \mathbb{Z}^N \ : \quad i_k = 1, 2, \dots n_k, \, k = 1, \dots, N \right\}.$$

The vector $\vec{\mathbf{i}} = (i_1, i_2, \dots, i_N)$ in $\mathbb{Z}^N$ is called a site and, for the N-uplet $\vec{\mathbf{n}} = (n_1, n_2, \dots, n_N)$ in $\mathbb{Z}^N$, we let $\hat{\vec{\mathbf{n}}} = \prod_{i=1}^{N} n_i$. The asymptotic design of this article is the increasing domain asymptotic. Formally, the latter is achieved when $\min\{n_i\} \to \infty$ and $|n_i/n_j| < C$ and/or C' for j, k such that $1 \leq j, k \leq N$, with C and/or C' being nonnegative constants. For this asymptotic design, we suppose that the functional $(\mathcal{X}_{\vec{\mathbf{i}}}, \mathcal{Y}_{\vec{\mathbf{i}}})$ for $\vec{\mathbf{i}} \in \mathbb{Z}^N$ has a strong mixing characteristic: there is a function $\psi(\cdot)$ such that $\psi(u) \downarrow 0$ as $u \to \infty$:

$$\begin{aligned} M_\alpha\Big(\mathfrak{F}(A), \mathfrak{F}(B)\Big) &= \sup_{E \in \mathfrak{F}(A),\, E' \in \mathfrak{F}(B)} \Big| \mathbb{P}(E \cap E') - \mathbb{P}(E)\mathbb{P}(E') \Big| \\ &\leq \phi\Big(\mathrm{Card}(A), \mathrm{Card}(B)\Big) \psi\Big(\mathrm{Dist}(A, B)\Big), \end{aligned} \tag{1}$$

where A, B are two subsets with finite cardinals and $\mathfrak{F}(\mathcal{C})$ is the sigma-algebra generated by the functional indexed by $\vec{\mathbf{i}} \in \mathcal{C}$. $\mathrm{Dist}\Big(A, B\Big)$ means the distance between A and B in the Euclidean sense and $Card(\mathcal{C})$ denotes the cardinal of $\mathcal{C}$. $\phi : \mathbb{N}^2 \to \mathbb{R}_+$ is a symmetric nondecreasing positive function in each variable. Finally, the functions $\phi(\cdot)$ and $\psi(\cdot)$ satisfy

$$\text{for all integers } n, m \qquad \phi(n, m) \leq C \min\,(n, m), \tag{2}$$

and

$$\text{for some } a > 0 \qquad \sum_{i=1}^{\infty} i^a \psi(i) < \infty. \tag{3}$$

Remark 1. *Notice that assumption (2) may be replaced by the following one:*

$$\psi(n, m) \leq C(n + m + 1)^\kappa, \textit{ for some } \kappa > 0.$$

Both the conditions (2) and (3) are used in [6,10]. It is important to observe that, when the value of N is equal to 1, the process $(\mathrm{X}_{\mathbf{i}}, \mathrm{Y}_{\mathbf{i}})$ is referred to as a strong mixing process. In his comprehensive analysis, [40] provided a detailed examination of mixing processes, illustrating his points with relevant examples. To facilitate the reader's comprehension of the spatio-functional data that meet the strong spatial mixing condition, as denoted by Equations (1)–(3), we provide an example of such data, namely, the spatial linear process. The definition and theoretical features of this process can be found in the works of [41,42]. Ref. [43] demonstrates that this particular process, given certain supplementary conditions, fulfills the assumption in (1).

2.2. Numerical Approximation of Expectile with Curve Regressor

In the rest of the paper, we assume that the functional random field $(\mathcal{X}_{\vec{\mathbf{i}}}, \mathcal{Y}_{\vec{\mathbf{i}}})$ satisfies the conditions (2) and (3). The LLFD of the expectile is obtained by assuming, for every $\mathcal{Z}$ in the vicinity of $\mathcal{X}$, for $p \in (0,1)$,

$$EXP_p(\mathcal{Z}) = EXP_p(\mathcal{X}) + EXP'_p(\mathcal{X})\alpha(\mathcal{Z},\mathcal{X}) + o(\alpha(\mathcal{Z},\mathcal{X})) \quad \text{with} \quad \alpha(\mathcal{Z},\mathcal{Z}) = 0. \tag{4}$$

where $\alpha(\cdot,\cdot)$ is a bilinear locating function such that

$$\text{For all} \quad \mathcal{X}' \in \mathcal{F}, \qquad C' Dis(\mathcal{X},\mathcal{X}') \leq |\alpha(\mathcal{X}',\mathcal{X})| \leq Dis(\mathcal{X},\mathcal{X}'). \tag{5}$$

Under this smoothing consideration, we define the LLFD of the expectile of $EXP_p(x)$ by finding the minimizers $(\widehat{\beta}_1, \widehat{\beta}_2)$ of

$$\min_{(\beta_1,\beta_2)\in\mathbb{R}^2} \sum_{\vec{\mathbf{i}}\in I_{\vec{\mathbf{n}}}} \left(\mathcal{Y}_{\vec{\mathbf{i}}} - \beta_1 - \beta_2\alpha(\mathcal{X}_{\vec{\mathbf{i}}}, x)\right)^2 \left| p - 1\!\!1_{\left\{\mathcal{Y}_{\vec{\mathbf{i}}} - \beta_1 - \beta_2\alpha(\mathcal{X}_{\vec{\mathbf{i}}}, x) \leq 0\right\}} \right| \mathcal{E}\left(\frac{Dis(x, \mathcal{X}_{\vec{\mathbf{i}}})}{\lambda}\right), \tag{6}$$

where λ is a positive real sequence and $\mathcal{E}$ is a kernel function. Recall that the definition (6) is motivated by the natural definition of the pth expectile of $\mathcal{Y}$, conditioned by $\mathcal{X} = x$, denoted by $EXP_p(x)$, that is, minimizer w.r.t. t, of the following minimization problem:

$$\min_t \mathbb{E}\left[\mathcal{L}_p(\mathcal{Y} - t) \mid \mathcal{X} = x\right], \tag{7}$$

where $\mathcal{L}_p(s) = |p - 1\!\!1_{\{s>0\}}|s^2$ and $1\!\!1_{\mathcal{A}}$ is the indicator of $\mathcal{A}$. Observe that, unlike the kernel estimator, the LLFD estimator is not explicitly defined. Thus, the establishment of the claimed asymptotic properties is a hard problem. In particular, this requires the representation of Bahadur associated with $\widehat{EXP}_p(y|x)$.

Remark 2.

- *The Nadaraya–Watson estimator employs local constant approximations. According to the numerical analyst [44], "Through all of scientific computing runs this common theme: Increase the accuracy at least to second order. What this means is: Get the linear term right". To clarify, a local constant approximation is deemed inadequate, whereas a local linear fit is considered preferable. Local linear fitting is an approach that is appealing from both a theoretical and practical perspective. The advantages of local linear fitting are discussed in the work of [45]. The proposed methodology demonstrates its adaptability to several design types, encompassing both random and fixed patterns, as well as highly clustered and virtually homogeneous designs. Moreover, it is worth noting that there is a lack of border effects observed in this context. The bias observed at the boundary remains consistent with that observed in the interior, without the need for the implementation of specific boundary kernels. No adjustments to the boundary are necessary when using local linear fitting, which is particularly advantageous in multidimensional scenarios where the boundary can involve a significant number of data points (see references [46,47]). Modifications to boundaries in higher dimensions pose significant challenges;*
- *It is clear that the regular regression can be viewed as particular case for our study. Indeed, if we put* $p = 0.5$, *the optimization problem (7) is equivalent to optimization with a scoring function associated to the least squared error. Thus, we can say that this also covers the local linear estimation of the regular regression studied, as constructed by [48].*

3. Hypotheses and Notation

Let (u_n), for $n \in \mathbb{N}$, be a sequence of real r.v.s. We say that (u_n) converges almost-completely (a.co.) toward zero if, and only if, for all

$$\epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|u_n| > \epsilon) < \infty.$$

Moreover, we say that the rate of the almost-complete convergence (a.c.c.) of (u_n) toward zero is on the order v_n (with $v_n \to 0$), and we write $u_n = O_{a.co.}(v_n)$ if, and only if, there exists $\epsilon > 0$ such that

$$\sum_{n=1}^{\infty} \mathbb{P}(|u_n| > \epsilon v_n) < \infty.$$

This kind of convergence implies both the almost-sure convergence and the convergence in probability. We aim to demonstrate the a.c.c. of the locally linear estimator $\widehat{EXP}_p(x)$ of $EXP_p(x)$. Firstly, we define

$$\begin{aligned}
G_p(y|x) &:= \mathbb{E}\left[\mathcal{L}_p(\mathcal{Y} - t) \,|\mathcal{X} = x\right], \\
\Gamma_1(y|x) &:= -\mathbb{E}\left[(\mathcal{Y} - y)\mathbb{1}_{\left\{\mathcal{Y}-y\leq 0\right\}} \mid \mathcal{X} = x\right], \\
\Gamma_2(y|x) &:= \mathbb{E}\left[(\mathcal{Y} - y)\mathbb{1}_{\left\{\mathcal{Y}-y> 0\right\}} \mid \mathcal{X} = x\right].
\end{aligned}$$

Next, assume the following:

(C1) The small function $\mathbb{P}(X \in B(x,\lambda)) = \varphi_x(\lambda)$ satisfies $\varphi_x(\lambda) > 0$. Moreover, there exists a function $\chi_x(\cdot)$ such that

$$\text{for all } \mathcal{S} \text{ in } [0,1], \quad \lim_{\lambda\to 0} \frac{\varphi_x(\mathcal{S}\lambda)}{\varphi_x(\lambda)} = \chi_x(\mathcal{S})$$

and the function $\alpha(\cdot,\cdot)$ exists such that

$$\sup_{\mathcal{U}\in B(x,\mathcal{R})} |\alpha(\mathcal{U},x) - Dist(x,\mathcal{U})| = o(\mathcal{R});$$

(C2) The operators $\Gamma_{i=1,2}(\cdot|x)$ are in class $C^1(\mathbb{R})$ and satisfy $\forall t_1, t_2 \in \mathbb{R}$, $\forall \mathbf{z}_1, \mathbf{z}_2 \in N_x$,

$$\left|\Gamma_i(t_2|\mathbf{z}_2) - \Gamma_i(t_1|\mathbf{z}_1)\right| \leq C(d^{k_i}(\mathbf{z}_1,\mathbf{z}_2) + |t_1 - t_2|^{k_i'}), \text{ for some, } k_i,\, k_i' > 0;$$

and $G_p(\cdot|x)$ verifies

$$\frac{\partial G_p(EXP_p(x)|x)}{\partial, y} < 0;$$

(C3) For all $\vec{\mathbf{j}} \neq \vec{\mathbf{i}}$,

$$0 < \sup_{\vec{\mathbf{i}}\neq\vec{\mathbf{j}}} \mathbb{P}\left[(\mathcal{X}_{\vec{\mathbf{j}}}, \mathcal{X}_{\vec{\mathbf{i}}}) \in B^2(x,\lambda)\right] \leq C(\varphi_x(\lambda))^{(a+1)/a},$$

for $C > 0$ and $1 < a < \frac{\delta}{N}$. Moreover, the random field $(X_{\mathbf{j}}, Y_{\mathbf{j}})_{\mathbf{j}\in\mathbb{N}}$ satisfies, for all $\vec{\mathbf{j}} \neq \vec{\mathbf{i}}$, almost surely,

$$\mathbb{E}\left[|Y_{\vec{\mathbf{i}}}Y_{\vec{\mathbf{j}}}|\Big|X_{\vec{\mathbf{i}}}, X_{\vec{\mathbf{j}}}\right] \leq C < \infty, \text{ and } \mathbb{E}\left[|Y_{\vec{\mathbf{i}}}|^q\Big|X_{\vec{\mathbf{i}}}\right] < C < \infty,$$

for some $q > 4$;

(C4) The kernel $\mathcal{E}(\cdot)$ is supported in $(-1,1)$, nonnegative, and differentiable in its support, satisfying that

$$D = \begin{pmatrix} \mathcal{E}(1) - \int_{-1}^{1} \mathcal{E}'(t)\chi_x(t)dt & \mathcal{E}(1) - \int_{-1}^{1} (t\mathcal{E}(t))'\chi_x(t)dt \\ \mathcal{E}(1) - \int_{-1}^{1} (t\mathcal{E}(t))'\chi_x(t)dt & \mathcal{E}(1) - \int_{-1}^{1} (t^2\mathcal{E}(t))'\chi_x(t)dt \end{pmatrix}$$

is a positive definite matrix;

(C5) There exists $\mathcal{V}_0 > 0$, such that

$$C\tilde{\mathbf{n}}^{\frac{5N}{\delta}-1+\mathcal{V}_0} \leq \varphi_x(\lambda); \quad \text{for} \quad C > 0.$$

Obviously, the five assumptions are not restrictive. They cover the functional aspect, the nonparametric feature, as well as the spatial dependency. Precisely, the functional path is explored by **(C1)**, and the nonparametric aspect is explored by **(C2)**, while the spatial dependency is evaluated by **(C3)**. The rest of the conditions can be considered as technical assumptions allowing the rate of the a.c.c. All the considered assumptions are compared to the previous works in nonparametric spatial functional time series data; for instance, see [29].

4. Main Results

The a.c.c. convergence rate of $\widehat{EXP}_p(x)$ to the expectile $EXP_p(x)$ is stated as follows.

Theorem 1. *If **(C1)–(C5)** hold, then, as $\vec{\boldsymbol{n}} \to \infty$,*

$$\left|\widehat{EXP}_p(x) - EXP_p(x)\right| = O(\lambda^\kappa) + O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\boldsymbol{n}}}{\tilde{\boldsymbol{n}}\varphi_x(\lambda)}}\right),$$

where κ is equal to $\min(k_1, k_2, k_1', k_2')$.

Proof. For the theorem's proof, we put $\alpha_{\vec{\mathbf{i}}} = \alpha(X, \mathcal{X}_{\vec{\mathbf{i}}})$ and $\vec{\mathbf{i}} \in \mathcal{I}_n$ $\mathcal{E}_{\vec{\mathbf{i}}} = \mathcal{E}(\lambda^{-1}Dis(x, \mathcal{X}_{\vec{\mathbf{i}}}))$. For this, we recall the following lemma. □

Lemma 1 (see [18])**.** *Consider $\mathcal{A}_{\vec{\boldsymbol{n}}}$ as a vectorial sequence of functions that satisfy the following:*

(i) *For every $\lambda \geq 1$ and multivariate ς:*

$$^\top\mathcal{A}_{\vec{\boldsymbol{n}}}(\lambda\varsigma) \leq {}^\top\varsigma\mathcal{A}_{\vec{\boldsymbol{n}}}(\varsigma);$$

(ii) *Let D a positive definite matrix and multivariate ς_0. Verify $\|\mathcal{A}_{\vec{\boldsymbol{n}}}(\varsigma_0)\| = o_{a.co.}(1)$ and*

$$\sup_{\|\varsigma\| \leq M} \|\mathcal{A}_{\vec{\boldsymbol{n}}}(\varsigma) + \lambda_0 D\varsigma - \mathcal{A}_{\vec{\boldsymbol{n}}}(\varsigma_0)\| = o_{a.co.}(1), \textit{ for } \lambda_0 > 0.$$

Then, for any multivariate sequence $\varsigma_{\vec{\boldsymbol{n}}}(\varsigma_0)$, in such a way that $\mathcal{A}_{\vec{\boldsymbol{n}}}(\varsigma_{\vec{\boldsymbol{n}}}) = o_{a.co.}(1)$, we have

$$\|\varsigma_{\vec{\boldsymbol{n}}}\| \leq M, \text{ a.co.} \tag{8}$$

For all $\varsigma = (\rho, v)$ in $\mathbb{R}^2$, we let

$$\Phi_{\vec{\mathbf{i}}}(\varsigma) = \mathcal{L}_p(\mathcal{Y}_{\vec{\mathbf{i}}} - (\rho + EXP_p(x)) - (\lambda^{-1}v + EXP_p'(x))\alpha(\mathcal{X}_{\vec{\mathbf{i}}}, x)),$$

where

$$\mathcal{L}_p(t) = t\Big(p1\!\!1_{[t\geq 0]} + (1-p)1\!\!1_{[t\leq 0]}\Big).$$

Observe that $\mathcal{L}_p(t)$ is the same as in (7). Thus, the main result is deduced from the use of Lemma 1 in [18] on

$$\mathcal{A}_{\mathbf{\vec{n}}}(\varsigma) = \frac{1}{\mathbf{\tilde{n}}\varphi_x(\lambda)} \begin{pmatrix} \sum_{\mathbf{\vec{i}} \in \mathcal{I}_n} \Phi_{\mathbf{\vec{i}}}(\varsigma)\mathcal{E}_{\mathbf{\vec{i}}} \\ \sum_{\mathbf{\vec{i}} \in \mathcal{I}_n} \Phi_{\mathbf{\vec{i}}}(\varsigma)\lambda^{-1}\alpha_{\mathbf{\vec{i}}}\mathcal{E}_{\mathbf{\vec{i}}} \end{pmatrix}.$$

Of course, we have to check the required conditions on

$$\mathcal{A}_{\mathbf{\vec{n}}}(\varsigma) \quad \text{and} \quad \varsigma_{\mathbf{\vec{n}}} = \begin{pmatrix} \widehat{EXP}_p(x) - EXP_p(x) \\ \lambda(\widehat{EXP'}_p(x) - EXP'_p(x)) \end{pmatrix}.$$

Subsequently, the theorem's proof is concluded from the following technical lemmas.

Lemma 2. *If* ***(C1)–(C5)*** *hold, then*

$$\|\mathcal{A}_{\vec{n}}(\varsigma_0)\| = O(\lambda^k) + O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\boldsymbol{n}}}{\tilde{\boldsymbol{n}}\varphi_x(\lambda)}}\right).$$

Lemma 3. *If* ***(C1)–(C5)*** *hold, then*

$$\sup_{\|\varsigma\| \le M} \| \mathbb{E}[\mathcal{A}_{\vec{n}}(\varsigma) - \mathcal{A}_{\vec{n}}(\varsigma_0)] + g_p(EXP_p(x)|x)D\varsigma \| = O(\lambda^\kappa),$$

where

$$g_p(y|x) = -\left(\frac{\partial}{\partial t}\Gamma_1(EXP_p(x)|x) + \frac{\partial}{\partial t}\Gamma_2(EXP_p(x)|x)\right).$$

Lemma 4. *If* ***(C1)–(C5)*** *hold, then*

$$\sup_{\|\varsigma\| \le M} \| \mathcal{A}_{\vec{n}}(\varsigma) - \mathcal{A}_{\vec{n}}(\varsigma_0) - \mathbb{E}[\mathcal{A}_{\vec{n}}(\varsigma) - \mathcal{A}_{\vec{n}}(\varsigma_0)] \| = O_{a.co.}\left(\sqrt{\frac{\ln \vec{\boldsymbol{n}}}{\tilde{\boldsymbol{n}}\varphi_x(\lambda)}}\right).$$

5. Real Data Application

Since COVID-19 has appeared, the health authorities in various countries have accelerated scientific research to control the propagation of the pandemic. At this stage, statistical modeling constitutes a principal tool to predict the future movement of the pandemic, allowing us to prevent the fast spread of the infection by this virus. The most appropriate models for these issues are those used to analyze the extreme values (see, [49,50]). The extreme values (EV) analysis is usually based on the estimation of the quantile function. Alternatively, we aim in this paper to implement the expectile model to fit the extreme values of the COVID-19 data. Recall that, as previously discussed, the expectile function has many advantages as risk models compared to the quantiles. In particular, the quantiles is an incoherent measure and it is defined by a backtesting measure based only on the frequencies of violations of fixed risk threshold, whereas the expectiles are coherent and elicitable with tail expectation. Therefore, as the expectation function relates the frequencies and values of data, the expectile model measures the risk's severity and frequency. On the other hand, the scoring measure of the expectile model is more regular and more smooth than the quantile. Thus, its implementation is very easy in practice. Next, the expectile is more sensitive to outliers, which is widely beneficial in risk investigation. In this sense, it detects the excessive propagation of risk better. For these reasons, the usefulness of the expectile regression in this kind of risk analysis is indisputable. To emphasize this great importance, we conduct a comparison study between both models (quantile and expectile). Note that the quantile estimator $\widehat{Qun}_p$ is obtained by taking in (6)

$$\mathcal{L}_p(s) = (2p-1)s + |s|.$$

Such comparative study is performed using COVID-19 data collected from 50 states in the USA during the period 3 April 2020 to 3 April 2021. The studied data are available on the website (https://covidtracking.com/data/, accessed on 1 August 2023). In this comparison study, we aimed to control the effect of the spatial interaction between the states on the propagation of the pandemic. Specifically, we predict the number of hospitalized cases one day ahead given the curves of the last 30 days of the positive tests in the neighboring states. Formally, we denote by $Y_{(i_1,i_2)}$ the number of hospitalized cases at day i_1 in the state i_2 and, in $X_{(i_1,i_2)}$, the curve of the last 30 days before i_1 at the state i_2. The spatio-temporal interaction of the data is shown in Figure 1.

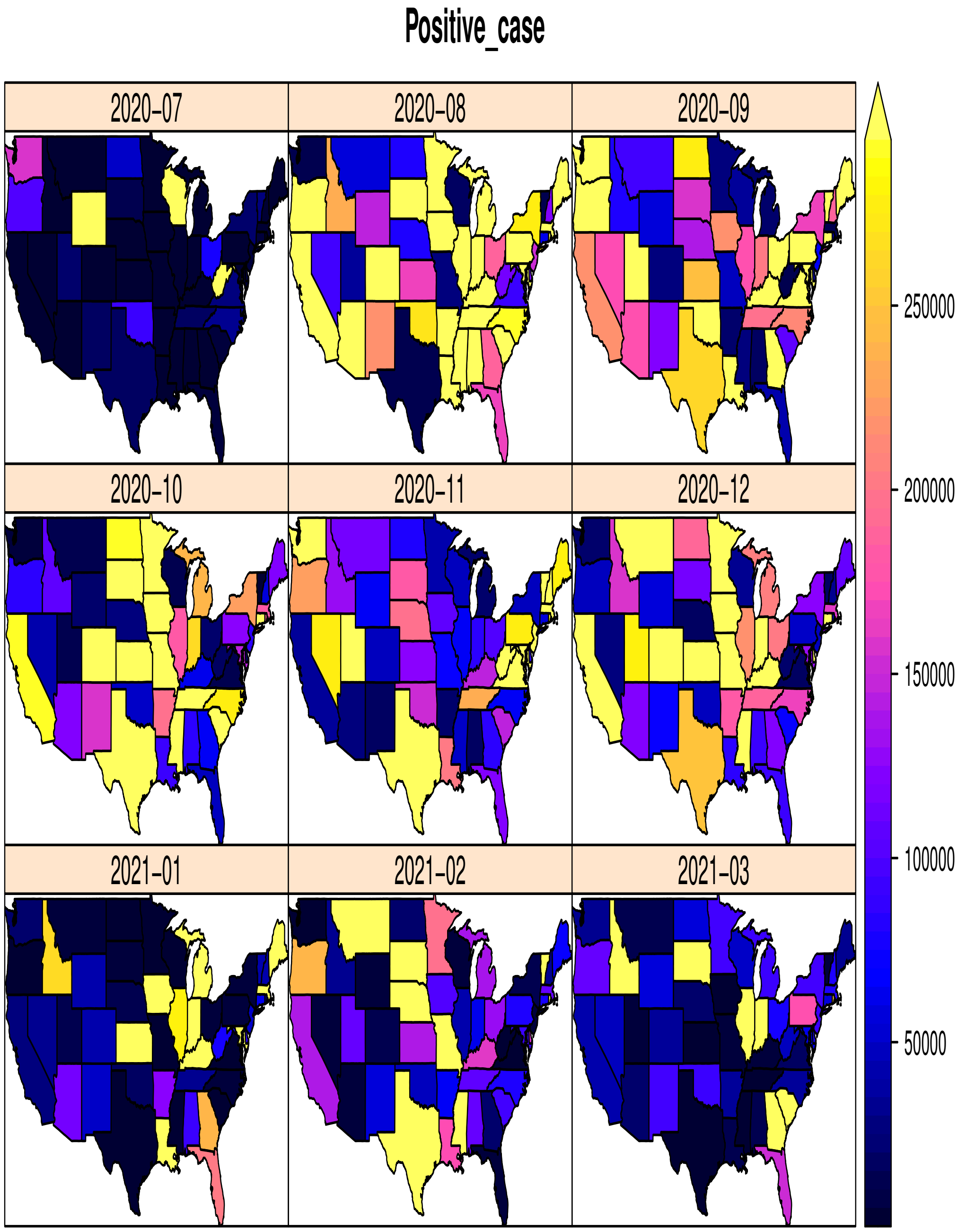

Figure 1. The spatio-temporal interaction of the hospitalized cases.

It is clear that the spatial vicinity of the states influences the propagation of the pandemic, in the sense that the propagation of the pandemic in the states affects others. Moreover, it clearly clearly that the data are affected by the presence of outliers in the hospitalized cases over the considered period, varying between 0 and 90,000 cases. Therefore, in order to accentuate the feasibility of the expectile and to detect the impact of the spatial interaction in the propagation of the pandemic, we compare the spatial prediction approach for the two models in both situations. In the first one, we neglect the spatial interaction within the data and we proceed without spatial trending, while the second one is based on the spatial detrending. Specifically, we control the spatial trending of two variables (response and explanatory) by the same approach as in [9], using the regression relationship defined by

$$\widetilde{X}_{\mathbf{i}} = m_1(\mathbf{i}) + X_{\mathbf{i}} \quad \text{and} \quad \widetilde{Y}_{\mathbf{i}} = m_2(\mathbf{i}) + Y_{\mathbf{i}}.$$

So, in the first situation, we compute the estimators $\widehat{EXP}_p$ and $\widehat{Qun}_p$ by $(\widehat{X}_{\mathbf{i}}, \widehat{Y}_{\mathbf{i}})_{\mathbf{i}}$, whereas, in the second situation, we construct the estimators from the initial data $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}}$. Obviously, the transformed data for the first situation are obtained by a pilot estimator for the functions m_1 and m_2. The latter is defined by

$$\widehat{m}_1(\mathbf{i}_0) = \frac{\sum_{\mathbf{i}\in\mathbf{I_n}} K(a_{\mathbf{n}}^{-1}\|\mathbf{i}_0 - \mathbf{i}\|)X_{\mathbf{i}}}{\sum_{\mathbf{i}\in\mathbf{I_n}} K(a_{\mathbf{n}}^{-1}\|\mathbf{i}_0 - \mathbf{i}\|)} \quad \left(\text{resp.} \quad \widehat{m}_2(\mathbf{j}_0) = \frac{\sum_{\mathbf{j}\in\mathbf{I_n}} K(b_{\mathbf{n}}^{-1}\|\mathbf{j}_0 - \mathbf{j}\|)Y_{\mathbf{j}}}{\sum_{\mathbf{j}\in\mathbf{I_n}} K(b_{\mathbf{n}}^{-1}\|\mathbf{j}_0 - \mathbf{j}\|)}\right),$$

where K is the kernel function and $a_{\mathbf{n}}$ and $b_{\mathbf{n}}$ constitute the bandwidth parameters within the real regression. Thus, the estimators $\widehat{EXP}_p$ and $\widehat{Qun}_p$ for the spatial detrending situation are obtained by

$$\widehat{Y}_{\mathbf{i}} = \widetilde{Y}_{\mathbf{i}} - \widehat{m}_2(\mathbf{i}) \quad \text{and} \quad \widehat{X}_{\mathbf{i}} = \widetilde{X}_{\mathbf{i}} - \widehat{m}_1(\mathbf{i}).$$

The real regressions m_1 and m_2 are obtained using the R-code `npreg` in the `np-R`-package. The bandwidth parameters $b_{\mathbf{n}}$ and $a_{\mathbf{n}}$ are selected by default, using the routine `npregbw` from the same `R`-package. The operator–estimators $\widehat{EXP}_p$ and $\widehat{Qun}_p$ are deduced from the $(0,1)$-quadratic kernel, and the smoothing sequence λ is selected locally by using a method of cross-validation over the k-nearest neighbors under the following `MSE` error:

$$\mathtt{MSE}(p) = \frac{1}{\mathbf{n}} \sum_{\mathbf{i}} \left(Y_{\mathbf{k}} - \widetilde{\xi}_{0.5}(X_{\mathbf{k}})\right)^2,$$

where $\widetilde{\xi}_p$ means both estimators $\widehat{EXP}_p$ and $\widehat{Qun}_p$. This rule is optimized from the subset

$$H_n = \left\{a \geq 0 : \sum_{i=1}^{n} 1\!\!1_{B(\mathfrak{z},a)}(X_i) = k\right\},$$

where $k \in \{5, 15, 25, \ldots, 0.5n\}$. Furthermore, the selection of the semi-metric is obtained by PCA-metric, which is more appropriate for this kind of discontinuous functional regressors. The EV comparison study is evaluated for the case $p = 0.01$, in the sense that we predict the 1% largest values of the parabolised hospitalized case for the 50 states at various periods. The prediction results are evaluated using the following backtesting measure:

$$Err = \frac{1}{50} \sum_{\mathbf{i}=1}^{50} \rho_{0.95}(Y_{\mathbf{i}} - \widetilde{\xi}_{0.95}(X_{\mathbf{i}})).$$

We evaluate this error for various periods. Specifically, we evaluate this error for 60 different days with both models and both situations. The box-plot of these errors is given in Figure 2.

Without surprise, the efficiency of $\widehat{EXP}_p$ and $\widehat{Qun}_p$ are strongly affected by the spatial correlation as well as the choice of the model, in the sense that the Spatial Expectile With Detrending (SEWD) performs better than the other models. It is clear that the SEWD shows preferment over the Spatial Expectile Without Detrending (SEWOD), the Spatial Quantile With Detrending (SQWD), and the Spatial Quantile Without Detrending (SQWOD). Such a statement confirms the spatial interaction within the data, which is that the error in the spatial detrending is smaller than the case when the spatial dependency is not taken into account.

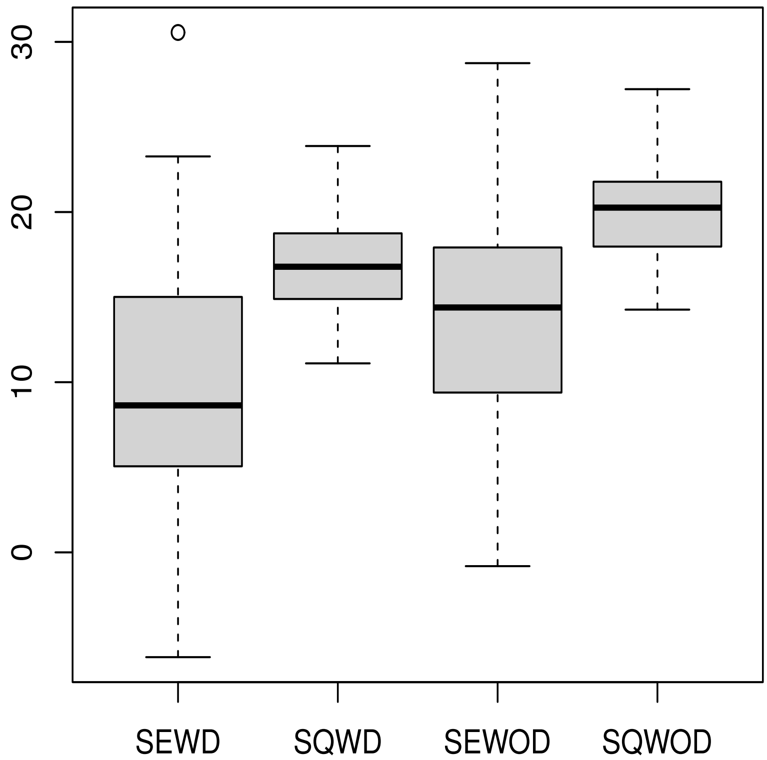

Figure 2. The spatio-temporal interaction among the hospitalized cases.

6. Concluding Remarks

As a risk model within regression settings, the M-estimation technique is employed to construct a Spatial Local Linear (SLL) estimate for the expectile function. As asymptotic behavior, we explicitly define the convergence rate for the obtained estimator. Two principal features characterize this contribution. The first one is the strong mixing property of the spatial correlation, while the second one concerns the dimension of the input random variable, which is is not necessarily finite. Such consideration allows one to improve the asymptotic property of the constructed estimator in spatio-functional time series analysis. Moreover, the expression of the convergence rate explores various factors of this study, including correlation, data functionality, and the functional class of the distribution. The implementation of the SLL estimator is assessed through empirical investigation. A real data application is conducted to showcase the superiority of the SLL-expectile over the SLL-quantile in risk assessment. The outcomes of the computational part confirm the advantages of the expectile over the quantiles as a risk analyzer. This is mainly due to the high sensitivity of the outliers exhibited by the expectile model. The extreme events have great consideration in risk analysis because they generate an important cost in practice. In addition to these important outcomes, the present paper introduces significant avenues for future exploration. Specifically, forthcoming research could delve into adapting our framework to handle censored data scenarios, which hold promise for intriguing findings. Another pivotal question involves delving into the limiting distributions of the estimators under investigation. This endeavor involves intricate mathematical complexities that transcend the scope of the present paper.

Furthermore, the path of investigation leads to the consideration of a functional kNN local linear approach for expectile regression estimators. This potential avenue presents the prospect of an alternative estimator that combines the merits of both methodologies—the local linear technique and the kNN approach.

The literature on nonparametric regression analysis, specifically where both the outcome and regressor variables are of functional nature, is still limited in the literature. Moreover, the application of our findings to this particular scenario is an inherent possibility within the scope of our current study. It should be noted that the concept of expectile, as employed in this paper, is not applicable when the variable Y is of a functional nature. This is due to the inherent inability to establish an order among functional variables. However, it is possible to utilize [51]'s concept for situations where the answer variable is multi-dimensional.

Another potential direction for future research involves the exploration of more intricate dependence structures, such as the ergodic spatial dependence or the quasi-association functional random fields.

Author Contributions: The authors contributed approximately equally to this work. Formal analysis, F.A.; Validation, O.L. and B.M.; Writing—review and editing, A.L. and S.B. All authors have read and agreed to the final version of the manuscript.

Funding: The authors thank and extend their appreciation to the funders of this project (1) Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R358), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; (2) The Deanship of Scientific Research at King Khalid University through the Research Groups Program under grant number R.G.P. 1/177/44.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available through the link https://covidtracking.com/data/ (accessed on 1 August 2023).

Acknowledgments: The authors would like to thank the Editor-in-Chief, an Associate-Editor, and three referees for their extremely helpful remarks, which resulted in a substantial improvement to the original form of the work and a presentation that was more sharply focused.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

This section is devoted to the proof of our results. The aforementioned notation is also used in what follows.

Proof of Lemma 2. Let us define, for $j = 0, 1$,

$$A^j_{\tilde{\mathbf{n}}} = \frac{1}{\tilde{\mathbf{n}}\varphi_x(\lambda)} \sum_{\vec{\mathbf{i}} \in \mathcal{I}_n} \Phi_{\vec{\mathbf{i}}}(\varsigma_0)\lambda^{-j}\alpha^j_{\vec{\mathbf{i}}}\mathcal{E}_{\vec{\mathbf{i}}}, \qquad j = 0, 1.$$

Thus, it suffices to prove

$$A^j_{\tilde{\mathbf{n}}} = O(\lambda^k) + O_{a.co.}\left(\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}\right)^{1/2}, \quad j = 0, 1.$$

Therefore, we split the proof into two assertions:

$$\left|A^j_{\tilde{\mathbf{n}}} - \mathbb{E}\left[A^j_{\tilde{\mathbf{n}}}\right]\right| = O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right) \qquad \text{for } j = 0, 1. \tag{A1}$$

and

$$\mathbb{E}\left[A^j_{\tilde{\mathbf{n}}}\right] = O(\lambda^\kappa), \qquad \text{for } j = 0, 1, \tag{A2}$$

starting with the deterministic part (A2). Using the fact that

$$(1-p)\Gamma_1(EXP_p(x)|x) + p\Gamma_2(EXP_p(x)|x) = 0,$$

we readily infer that

$$\begin{aligned}
&\mathbb{E}\left[A^0_{\vec{\mathbf{n}}}\right] \\
&= \frac{1-p}{\varphi_x(\lambda)}\mathbb{E}\left[\mathcal{E}_1(\mathcal{Y}_1 - EXP_p(x) + EXP'_p(x)\alpha_{\vec{\mathbf{i}}}\mathbb{1}_{[\mathcal{Y}_1 \le EXP_p(x)+EXP'_p(x)\alpha_{\vec{\mathbf{i}}})]}\right] \\
&\quad + \frac{p}{\varphi_x(\lambda)}\mathbb{E}\left[\mathcal{E}_1(\mathcal{Y}_1 - EXP_p(x) + EXP'_p(x)\alpha_{\vec{\mathbf{i}}}\mathbb{1}_{[\mathcal{Y}_1 \ge EXP_p(x)+EXP'_p(x)\alpha_{\vec{\mathbf{i}}}]}\right] \\
&\le \frac{1-p}{\varphi_x(\lambda)}\mathbb{E}\left[\left|\Gamma_1(EXP_p(x) + EXP'_p(x)\alpha_{\vec{\mathbf{i}}}|(\mathcal{X}_{\vec{\mathbf{i}}}) - \Gamma_1(EXP_p(x)|x)\right|\mathcal{E}_1\right] \\
&\quad + \frac{p}{\varphi_x(\lambda)}\mathbb{E}\left[\left|\Gamma_2(EXP_p(x) + EXP'_p(x)\alpha_{\vec{\mathbf{i}}}|(\mathcal{X}_{\vec{\mathbf{i}}}) - \Gamma_2(EXP_p(x)|x)\right|\mathcal{E}_1\right].
\end{aligned}$$

Making use of the condition **(C2)**, we obtain

$$\mathbb{E}\left[A^0_{\vec{\mathbf{n}}}\right] \le C\lambda^{\kappa}.$$

Conversely, for $\mathbb{E}\left[A^1_{\vec{\mathbf{n}}}\right]$, we use

$$\lambda^{-1}\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1\mathbb{1}_{B(x,h)}(\mathcal{X}_1)| \le \mathcal{E}_1,$$

implying that

$$\mathbb{E}\left[A^1_{\vec{\mathbf{n}}}\right] \le \mathbb{E}\left[A^0_{\vec{\mathbf{n}}}\right] \le C\lambda^{\kappa}.$$

Now, to investigate (A1), let us define

$$A^j_{\vec{\mathbf{n}}} = \frac{1}{\tilde{\mathbf{n}}\varphi_x(\lambda)}\sum_{\vec{\mathbf{i}}\in\mathcal{I}_n}\varsigma^j_{\vec{\mathbf{i}}}$$

where

$$\varsigma^j_{\vec{\mathbf{i}}} = (\Phi_{\vec{\mathbf{i}}}(\varsigma_0)\lambda^{-j}\alpha^j_{\vec{\mathbf{i}}}\mathcal{E}_{\vec{\mathbf{i}}} - \mathbb{E}[\Phi_{\vec{\mathbf{i}}}(\varsigma_0)\lambda^{-j}\alpha^j_{\vec{\mathbf{i}}}\mathcal{E}_{\vec{\mathbf{i}}}]) \quad \text{for} \quad j = 0,1.$$

Next, consider a spatial block composition as in [6]. This decomposition splits the sum on 2^N sums. Indeed, for a given $p_{\vec{\mathbf{n}}}$, we define

$$\begin{aligned}
T(1,x,\vec{\mathbf{n}},\vec{\mathbf{j}}) &= \sum_{i_k=2j_k+1 k=1,2,\ldots,N}^{2j_k p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}} \varsigma^j_{\vec{\mathbf{i}}}, \\
T(2,x,\vec{\mathbf{n}},\vec{\mathbf{j}}) &= \sum_{i_k=2j_k+1 k=1,2,\ldots,N-1}^{2j_k p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}} \sum_{i_{\vec{\mathbf{n}}}=2j_{\vec{\mathbf{n}}}p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}+1}^{(j_{\vec{\mathbf{n}}}+1)p_{\vec{\mathbf{n}}}} \varsigma^j_{\vec{\mathbf{i}}}, \\
T(3,x,\vec{\mathbf{n}},\vec{\mathbf{j}}) &= \sum_{i_k=2j_k+1 k=1,2,\ldots,N-2}^{2j_k p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}} \sum_{i_{N-1}=2j_{N-1}p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}+1}^{2(j_{N-1}+1)p_{\vec{\mathbf{n}}}} \sum_{i_{\vec{\mathbf{n}}}=2j_n p_{\vec{\mathbf{n}}}+1}^{2j_n p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}} \varsigma^j_{\vec{\mathbf{i}}}, \\
T(4,x,\vec{\mathbf{n}},\vec{\mathbf{j}}) &= \sum_{i_k=2j_k p_{\vec{\mathbf{n}}}+1 k=1,2,\ldots,N-2}^{2j_k p_{\vec{\mathbf{n}}}} \sum_{i_{N-1}=2j_{N-1}p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}+1}^{2(j_{N-1}+1)p_{\vec{\mathbf{n}}}} \sum_{i_{\vec{\mathbf{n}}}=2j_{\vec{\mathbf{n}}}p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}+1}^{2(j_{\vec{\mathbf{n}}}+1)p_{\vec{\mathbf{n}}}} \varsigma^j_{\vec{\mathbf{i}}},
\end{aligned}$$

and so on. Next, let

$$T(2^{N-1},x,\vec{\mathbf{n}},\vec{\mathbf{j}}) = \sum_{i_k=2j_k p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}+1 k=1,2,\ldots,N-1}^{2(j_k+1)p_{\vec{\mathbf{n}}}} \sum_{i_{\vec{\mathbf{n}}}=2j_{\vec{\mathbf{n}}}p_{\vec{\mathbf{n}}}+p_{\vec{\mathbf{n}}}}^{2j_{\vec{\mathbf{n}}}p_{\vec{\mathbf{n}}}+1} \varsigma^j_{\vec{\mathbf{i}}},$$

$$T(2^N, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}) \quad = \sum_{i_k = 2j_k p_{\vec{\mathbf{n}}}+1 k=1,2,\ldots,N}^{2(j_k+1)p_{\vec{\mathbf{n}}}} \varsigma^j_{\vec{\mathbf{i}}}.$$

Additionally, we set

$$U(x, \vec{\mathbf{n}}, i) = \sum_{\vec{\mathbf{j}} \in \mathcal{J}} T(i, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}),$$

with

$$\mathcal{J} = \left\{ \vec{\mathbf{j}} = (j_k)_{1 \le k \le N} \text{ with } 0 \le j_k \le r_k - 1 \right\},$$

and $r_{\vec{\mathbf{i}}} = 2n_{\vec{\mathbf{i}}} p_{\vec{\mathbf{n}}}^{-1}$, $i = 1, \ldots, N$. Remark that

$$\mid \mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0) - \mathbb{E}[\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0)] \mid = \frac{1}{\tilde{\mathbf{n}} \varphi_x(\lambda)} \sum_{\vec{\mathbf{i}}=1}^{2^N} U(x, \vec{\mathbf{n}}, i). \tag{A3}$$

So, it suffices to compute

$$\mathbb{P}(U(x, \vec{\mathbf{n}}, i) \ge \eta \tilde{\mathbf{n}} \varphi_x(\lambda)) \text{ for all } i = 1, \ldots, 2^N.$$

We prove only the case $i = 1$; the proof of the other cases is obtained using the same ideas. It is obtained by enumerating the $M = \prod_{k=1}^N r_k = 2^{-N} \tilde{\mathbf{n}} p_{\vec{\mathbf{n}}}^{-N} \le \tilde{\mathbf{n}} p_{\vec{\mathbf{n}}}^{-N}$ random variables $T(1, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}); \vec{\mathbf{j}} \in \mathcal{J}$ in the line order $Z_1, \ldots, Z_M$. Thus, each Z_j is

$$Z_j = \sum_{i \in I(1, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}_j)} \varsigma^j_{\vec{\mathbf{i}}},$$

with

$$I(1, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}_j) = \left\{ \vec{\mathbf{i}} : 2j_{k_j} p_{\vec{\mathbf{n}}} + 1 \le i_k \le 2j_{k_j} p_{\vec{\mathbf{n}}} + p_{\vec{\mathbf{n}}}; k = 1, 2, \ldots N \right\}.$$

Clearly, each set $I(1, x, \vec{\mathbf{n}}, \vec{\mathbf{j}}_j)$ contains $p_{\vec{\mathbf{n}}}^N$ sites and are distant by at least $p_{\vec{\mathbf{n}}}^N$. Therefore, the variables $Z_1, Z_2, \ldots, Z_M$ can be approximated by independent variables $Z_1^*, Z_2^*, \ldots, Z_M^*$ identically distributed as $Z_{j=1,\ldots,M}$, such that

$$\sum_{i=1}^r \mathbb{E} \mid Z_i - Z_i^* \mid \le 2CM p_{\vec{\mathbf{n}}}^N \phi((M-1) p_{\vec{\mathbf{n}}}^N, p_{\vec{\mathbf{n}}}^N) \psi(p_{\vec{\mathbf{n}}}^N).$$

Furthermore,

$$\mathbb{P}(U(x, \vec{\mathbf{n}}, 1) \ge \eta \tilde{\mathbf{n}} \varphi_x(\lambda)) \le \mathcal{V}_1 + \mathcal{V}_2 n,$$

where

$$\mathcal{V}_1 = \mathbb{P}\left(\mid \sum_{j=1}^n Z^*_{\vec{\mathbf{j}}} \mid \ge \frac{\eta \tilde{\mathbf{n}} \varphi_x(\lambda)}{2} \right)$$

and

$$\mathcal{V}_2 = \mathbb{P}\left(\sum \mid Z_{\vec{\mathbf{j}}} - Z^*_{\vec{\mathbf{j}}} \mid \ge \frac{\eta \tilde{\mathbf{n}} \varphi_x(\lambda)}{2} \right).$$

For $\mathcal{V}_2$, we use the Markov inequality to obtain

$$\begin{aligned} \mathcal{V}_2 \quad &= \quad \mathbb{P}\left(\sum \mid Z_{\vec{\mathbf{j}}} - Z^*_{\vec{\mathbf{j}}} \mid \ge \frac{\eta \tilde{\mathbf{n}} \varphi_x(\lambda)}{2} \right) \\ &\le \quad \frac{1}{\eta \tilde{\mathbf{n}} \varphi_x(\lambda)} \sum \mathbb{E}\Big[\mid Z_j - Z_j^* \mid \Big] \\ &\le \quad 2M p_{\vec{\mathbf{n}}}^N (\eta \tilde{\mathbf{n}} \varphi_x(\lambda))^{-1} \phi((M-1) p_{\vec{\mathbf{n}}}^N, p_{\vec{\mathbf{n}}}^N) \psi(p_{\vec{\mathbf{n}}}). \end{aligned}$$

As $\tilde{\mathbf{n}} = 2^N M p_{\vec{\mathbf{n}}}^N$ and $\phi((M-1)p_{\vec{\mathbf{n}}}^N, p_{\vec{\mathbf{n}}}^N) \leq p_{\vec{\mathbf{n}}}^N$, then, for $\eta = \eta_0 \sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}$, we have

$$\mathcal{V}_2 \leq \tilde{\mathbf{n}} p_{\vec{\mathbf{n}}}^N (\ln \tilde{\mathbf{n}})^{-1/2} (\tilde{\mathbf{n}}\varphi_x(\lambda))^{-1/2} \psi(p_{\vec{\mathbf{n}}}).$$

By choosing $p_{\vec{\mathbf{n}}} = C\left(\frac{\tilde{\mathbf{n}}\phi_x(\lambda)}{\ln \tilde{\mathbf{n}}}\right)^{1/2N}$, we have

$$\mathcal{V}_2 \leq \tilde{\mathbf{n}} \psi(p_{\vec{\mathbf{n}}}). \tag{A4}$$

Consequently, combining (3) and (**C4**), we conclude that

$$\sum_{\vec{\mathbf{n}}} \tilde{\mathbf{n}} \psi(p_{\vec{\mathbf{n}}}) < \infty.$$

Concerning $\mathcal{V}_1$, we write

$$\begin{aligned}
\mathcal{V}_1 &= \mathbb{P}\left(\mid \sum_{j=1}^{n} Z_{\vec{\mathbf{j}}}^* \mid \geq \frac{M\eta\tilde{\mathbf{n}}\varphi_x(\lambda)}{2M} \right) \\
&\leq 2\exp\left(\frac{-(\eta\tilde{\mathbf{n}}\varphi_x(\lambda))^2}{MVar[Z_1^*] + Cp_{\vec{\mathbf{n}}}^N \eta \tilde{\mathbf{n}}\varphi_x(\lambda)} \right).
\end{aligned}$$

Next, we asymptotically evaluate $Var[Z_1^*]$. Indeed,

$$Var[Z_1^*] = \sum_{\vec{\mathbf{i}} \in I(1,x,\vec{\mathbf{n}},1)} Var[\varsigma_{\vec{\mathbf{i}}}] + R_{\vec{\mathbf{n}}},$$

where

$$R_{\vec{\mathbf{n}}} = \sum_{\vec{\mathbf{i}} \neq j \in I(1,x,\vec{\mathbf{n}},1)} \mid cov(\varsigma_{\vec{\mathbf{i}}}, \varsigma_{\vec{\mathbf{j}}}) \mid .$$

By Assumptions (**C1**) and the first part of (**C3**), we have

$$Var[\varsigma_{\vec{\mathbf{i}}}] \leq C(\varphi_x(\lambda) + (\varphi_x(\lambda))^2).$$

Hence,

$$\sum_{\vec{\mathbf{i}} \in I(1,x,\vec{\mathbf{n}},1)} Var[\varsigma_{\vec{\mathbf{i}}}] = O(p_{\vec{\mathbf{n}}}^N \varphi_x(\lambda)).$$

Now, for $R_{\vec{\mathbf{n}}}$, we split the sum over

$$\begin{aligned}
S_1 &= \{\vec{\mathbf{i}}, \vec{\mathbf{j}} \in I(1,x,\vec{\mathbf{n}},1) : 0 < \| \vec{\mathbf{i}} - \vec{\mathbf{j}} \| \leq C_{\vec{\mathbf{n}}}\}, \\
S_2 &= \{\vec{\mathbf{i}}, \vec{\mathbf{j}} \in I(1,x,\vec{\mathbf{n}},1) : \| \vec{\mathbf{i}} - \vec{\mathbf{j}} \| > C_{\vec{\mathbf{n}}}\},
\end{aligned}$$

where $C_{\vec{\mathbf{n}}}$ goes to $+\infty$ when $\vec{\mathbf{n}} \to \infty$. Therefore,

$$\begin{aligned}
R_{\vec{\mathbf{n}}} &= \sum_{(\vec{\mathbf{i}},\vec{\mathbf{j}}) \in S_1} \mid cov(\varsigma_{\vec{\mathbf{i}}}, \varsigma_{\vec{\mathbf{j}}}) \mid + \sum_{(\vec{\mathbf{i}},\vec{\mathbf{j}}) \in S_2} \mid cov(\varsigma_{\vec{\mathbf{i}}}, \varsigma_{\vec{\mathbf{j}}}) \mid \\
&= R_{\vec{\mathbf{n}}}^1 + R_{\vec{\mathbf{n}}}^2.
\end{aligned}$$

On the one hand, we have

$$\begin{aligned}
R_{\vec{\mathbf{n}}}^1 &\leq C \sum_{(\vec{\mathbf{i}},\vec{\mathbf{j}}) \in S_1} \mid \mathbb{E}[\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{E}_{\vec{\mathbf{j}}}] \mid + \mid \mathbb{E}[\mathcal{E}_{\vec{\mathbf{i}}}]\mathbb{E}[\mathcal{E}_{\vec{\mathbf{j}}}] \mid) \\
&\leq Cp_{\vec{\mathbf{n}}}^N \varphi_x(\lambda) \Big((\varphi_x(\lambda))^{1/a} + \varphi_x(\lambda) \Big)
\end{aligned}$$

$$\leq \quad C p_{\vec{\mathbf{n}}}^{N} (\varphi_x(\lambda))^{(a+1)/a}.$$

On the other hand, we have

$$R_{\vec{\mathbf{n}}}^{2} = \sum_{(\vec{\mathbf{i}},\vec{\mathbf{j}})\in S_2} \mid cov(\varsigma_{\vec{\mathbf{i}}}, \varsigma_{\vec{\mathbf{j}}}) \mid .$$

As the random variables $\mathcal{E}_j$ are bounded, we deduce from covariance inequality in [52] that

$$\mid cov(\varsigma_{\vec{\mathbf{i}}}, \varsigma_{\vec{\mathbf{j}}}) \mid \leq C\psi(\parallel \vec{\mathbf{i}} - \vec{\mathbf{j}} \parallel).$$

Therefore, we obtain

$$\begin{aligned} R_{\vec{\mathbf{n}}}^{2} &\leq C \sum_{(\vec{\mathbf{i}},\vec{\mathbf{j}})\in S_2} \psi(\parallel \vec{\mathbf{i}} - \vec{\mathbf{j}} \parallel) \leq C p_{\vec{\mathbf{n}}}^{N} \sum_{\vec{\mathbf{i}}:\|\vec{\mathbf{i}}\|\geq C_{\vec{\mathbf{n}}}} \psi(\parallel \vec{\mathbf{i}} \parallel) \\ &\leq C p_{\vec{\mathbf{n}}}^{N} C_{\vec{\mathbf{n}}}^{-Na} \sum_{\vec{\mathbf{i}}:\|\vec{\mathbf{i}}\|\geq C_{\vec{\mathbf{n}}}} \parallel \vec{\mathbf{i}} \parallel^{Na} \psi(\parallel \vec{\mathbf{i}} \parallel). \end{aligned}$$

Let us denote

$$C_{\vec{\mathbf{n}}} = (\varphi_x(\lambda))^{-1/Na}.$$

Then, we have

$$\begin{aligned} R_{\vec{\mathbf{n}}}^{2} &\leq C p_{\vec{\mathbf{n}}}^{N} C_{\vec{\mathbf{n}}}^{-Na} \sum_{\vec{\mathbf{i}}:\|\vec{\mathbf{i}}\|\geq C_{\vec{\mathbf{n}}}} \parallel \vec{\mathbf{i}} \parallel^{Na} \psi(\parallel \vec{\mathbf{i}} \parallel) \\ &\leq C p_{\vec{\mathbf{n}}}^{N} (\varphi_x(\lambda)) \sum_{\vec{\mathbf{i}}:\|\vec{\mathbf{i}}\|\geq C_{\vec{\mathbf{n}}}} \parallel \vec{\mathbf{i}} \parallel^{Na} \psi(\parallel \vec{\mathbf{i}} \parallel). \end{aligned}$$

Once again, by (3) and (**C4**), we obtain

$$R_{\vec{\mathbf{n}}}^{2} \leq C p_{\vec{\mathbf{n}}}^{N} (\varphi_x(\lambda)) \quad \text{and} \quad R_{\vec{\mathbf{n}}}^{1} \leq C p_{\vec{\mathbf{n}}}^{N} (\varphi_x(\lambda)).$$

Consequently, we have

$$Var[Z_1^*] = O(p_{\vec{\mathbf{n}}}^{N} (\phi_x(\lambda))),$$

implying

$$\mathcal{V}_1 \leq \exp(-C(\eta_0 \ln \tilde{\mathbf{n}})).$$

Finally, we conclude that

$$\left| A_{\vec{\mathbf{n}}}^{j} - \mathbb{E}\left[A_{\vec{\mathbf{n}}}^{j} \right] \right| = O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}} \varphi_x(\lambda)}} \right).$$

Since $\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0) = \begin{pmatrix} A_{\vec{\mathbf{n}}}^{0} \\ A_{\vec{\mathbf{n}}}^{1} \end{pmatrix}$, then

$$\|\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0)\| = O(\lambda^k) + O_{a.co.}\left(\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}} \varphi_x(\lambda)} \right)^{1/2}.$$

Hence, the proof is complete. □

Proof of Lemma 3. We write

$$\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma) = \begin{pmatrix} \mathcal{A}_{\vec{\mathbf{n}}}^{0}(\varsigma) \\ \mathcal{A}_{\vec{\mathbf{n}}}^{1}(\varsigma) \end{pmatrix} \quad \text{with} \quad \mathcal{A}_{\vec{\mathbf{n}}}^{j}(\varsigma) = \frac{1}{\tilde{\mathbf{n}} \varphi_x(\lambda)} \sum_{\vec{\mathbf{i}}\in\mathcal{I}_n} \Phi_{\vec{\mathbf{i}}}(\varsigma) \lambda^{-j} \alpha_{\vec{\mathbf{i}}}^{j} \mathcal{E}_{\vec{\mathbf{i}}}.$$

The bias term of $\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma)$ is

$$\begin{aligned}\mathbb{E}\Big[\mathcal{A}^0_{\vec{\mathbf{n}}}(\varsigma)\Big] &= \frac{1-p}{\varphi_x(\lambda)}\mathbb{E}\Big[\Gamma_1(EXP_p+\rho+(\lambda^{-1}v+EXP'_p)\alpha_{\vec{\mathbf{i}}}|\mathcal{X}_1)\mathcal{E}_1\Big] \\ &\quad+\frac{p}{\varphi_x(\lambda)}\mathbb{E}\Big[\Gamma_2(EXP_p+\rho+(\lambda^{-1}v+EXP'_p)\alpha_{\vec{\mathbf{i}}}|\mathcal{X}_1)\mathcal{E}_1\Big],\end{aligned}$$

while, for $\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma)$,

$$\begin{aligned}\mathbb{E}\Big[\mathcal{A}^1_{\vec{\mathbf{n}}}(\varsigma)\Big] &= \frac{1-p}{\varphi_x(\lambda)}\mathbb{E}\Big[\Gamma_1(EXP_p+\rho+(\lambda^{-1}v+EXP'_p)\alpha_{\vec{\mathbf{i}}}|\mathcal{X}_1)\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1\Big] \\ &\quad+\frac{p}{\varphi_x(\lambda)}\mathbb{E}\Big[\Gamma_2(EXP_p+\rho+(\lambda^{-1}v+EXP'_p)\alpha_{\vec{\mathbf{i}}}|\mathcal{X}_1)\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1\Big].\end{aligned}$$

By simple analytical arguments, we obtain

$$\begin{aligned}&\mathbb{E}\Big[\mathcal{A}^0_{\vec{\mathbf{n}}}(\varsigma)\Big]-\mathbb{E}\Big[\mathcal{A}^0_{\vec{\mathbf{n}}}(\varsigma_0)\Big] \\ &= +\frac{1-p}{\varphi_x(\lambda)}\left(\frac{\partial}{\partial t}\Gamma_1(EXP_p,x)(\mathbb{E}[\mathcal{E}_1],\lambda^{-1}\mathbb{E}[\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1])\varsigma\right) \\ &\quad+\frac{p}{\varphi_x(\lambda)}\left(\frac{\partial}{\partial t}\Gamma_2(EXP_p|x)(\mathbb{E}[\mathcal{E}_1],\lambda^{-1}\mathbb{E}[\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1])\varsigma\right) \\ &\quad+O(\lambda^\kappa)+o(\|\varsigma\|),\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\Big[\mathcal{A}^1_{\vec{\mathbf{n}}}(\varsigma)\Big]-\mathbb{E}\Big[\mathcal{A}^1_{\vec{\mathbf{n}}}(\varsigma_0)\Big] \\ &= +\frac{1-p}{\varphi_x(\lambda)}\left(\frac{\partial}{\partial t}\Gamma_1(EXP_p|x)(\lambda^{-1}\mathbb{E}[\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1],\lambda^{-2}\mathbb{E}[\alpha^2_{\vec{\mathbf{i}}}\mathcal{E}_1])\varsigma\right) \\ &\quad+\frac{p}{\varphi_x(\lambda)}\left(\frac{\partial}{\partial t}\Gamma_2(EXP_p|x)(\lambda^{-1}\mathbb{E}[\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1],\lambda^{-2}\mathbb{E}[\alpha^2_{\vec{\mathbf{i}}}\mathcal{E}_1])\varsigma\right) \\ &\quad+O(\lambda^\kappa)+o(\|\varsigma\|).\end{aligned}$$

Finally, we obtain

$$\begin{aligned}&\mathbb{E}[\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma)-\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0)] \\ &= -\frac{g_p(EXP_p|x)}{\varphi_x(\lambda)}\begin{pmatrix}\mathbb{E}[\mathcal{E}_{\vec{\mathbf{i}}}] & \mathbb{E}[\lambda^{-1}\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1] \\ \mathbb{E}[\lambda^{-1}\alpha_{\vec{\mathbf{i}}}\mathcal{E}_1] & \mathbb{E}\Big[\lambda^{-2}\alpha^2_{\vec{\mathbf{i}}}\mathcal{E}_1\Big]\end{pmatrix}\varsigma \\ &\quad+O(\lambda^\kappa)+o(\|\varsigma\|).\end{aligned}$$

It is shown in [53] that

$$\begin{aligned}\lambda^{-a}\mathbb{E}[\alpha^a_{\vec{\mathbf{i}}}\mathcal{E}^b_i] &= \varphi_x(\lambda)\left(\mathcal{E}^b(1)-\int_{-1}^1(u^a\mathcal{E}^b(u))'\chi_x(u)du\right) \\ &\quad+o(\varphi_x(\lambda)).\end{aligned}$$

Hence, we conclude that

$$\sup_{\|\varsigma\|\le M}\|\,\mathbb{E}[\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma)-\mathcal{A}_{\vec{\mathbf{n}}}(\varsigma_0)]+g(EXP_p|x)D\varsigma+o(\|\,\varsigma\,\|)\,\|=O(\lambda^\kappa).$$

Thus, the proof is completed. □

Proof of Lemma 4. We use the compactness of the ball $B(0,M)$ in $\mathbb{R}^2$, that is,

$$B(0,M) \subset \bigcup_{j=1}^{d_{\tilde{\mathbf{n}}}} B(\vartheta_j, l_{\tilde{\mathbf{n}}}),\ \ \vartheta_j = \begin{pmatrix} \rho_j \\ \upsilon_j \end{pmatrix} \text{ and } l_{\tilde{\mathbf{n}}} = d_{\tilde{\mathbf{n}}}^{-1} = 1/\sqrt{n}.$$

Then, $\forall \varsigma \in B(0,M)$; we put $j(\varsigma) = \arg\min_j \mid \varsigma - \vartheta_j \mid$ and we write

$$\begin{aligned}
&\sup_{\|\varsigma\|\leq M} \| \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma) - \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma_0) - \mathbb{E}[\mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma) - \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma_0)] \| \\
&\quad \leq \sup_{\|\varsigma\|\leq M} \| \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma) - \mathcal{A}_{\tilde{\mathbf{n}}}(\vartheta_j) \| \\
&\qquad + \sup_{\|\varsigma\|\leq M} \| \mathcal{A}_{\tilde{\mathbf{n}}}(\vartheta_j) - \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma_0) - \mathbb{E}[\mathcal{A}_{\tilde{\mathbf{n}}}(\vartheta_j) - \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma_0)] \| \\
&\qquad + \sup_{\|\varsigma\|\leq M} \| \mathbb{E}[\mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma) - \mathcal{A}_{\tilde{\mathbf{n}}}(\vartheta_j)] \| .
\end{aligned}$$

We combine the inequalities

$$|\mathcal{L}_p(t) - \mathcal{L}_p(t_0) \,|\leq C|t-t_0| + t_0|1\!\!1_{[t<0]} - 1\!\!1_{[t_0<0]}|,$$

and

$$|1\!\!1_{[\mathcal{Y}<a]} - 1\!\!1_{[\mathcal{Y}<b]}| \leq 1\!\!1_{[|\mathcal{Y}-b|\leq|a-b|]},$$

to prove

$$\sup_{\|\varsigma\|\leq M} \| \mathcal{A}_{\tilde{\mathbf{n}}}(\varsigma) - \mathcal{A}_{\tilde{\mathbf{n}}}(\vartheta_j) \| \leq 2 \sum_{\vec{\mathbf{i}}\in\mathcal{I}_n} W_{\vec{\mathbf{i}}},$$

where

$$W_{\vec{\mathbf{i}}} = W_{1\vec{\mathbf{i}}} + W_{2\vec{\mathbf{i}}} + W_{3\vec{\mathbf{i}}},$$

and

$$\begin{aligned}
W_{1\vec{\mathbf{i}}} &= \frac{1}{\tilde{\mathbf{n}}\varphi_x(\lambda)} \sup_{\|\varsigma\|\leq M} 1\!\!1_{\{[|\mathcal{Y}_{\vec{\mathbf{i}}}-(\rho_j+EXP_p(x))-(\lambda^{-1}\upsilon_j+EXP'_p(x))\alpha_{\vec{\mathbf{i}}}|\leq Cl_{\tilde{\mathbf{n}}}]\}} \left\| \begin{pmatrix} 1 \\ \lambda^{-1}\alpha_{\vec{\mathbf{i}}} \end{pmatrix} \right\| \mathcal{E}_{\vec{\mathbf{i}}}, \\
W_{2\vec{\mathbf{i}}} &= \frac{1}{\tilde{\mathbf{n}}\varphi_x(\lambda)} \sup_{\|\varsigma\|\leq M} 1\!\!1_{\{[|\mathcal{Y}_{\vec{\mathbf{i}}}-(\rho_j+EXP_p(x))-(\lambda^{-1}\upsilon_j+EXP'_p(x))\alpha_{\vec{\mathbf{i}}}|\leq Cl_{\tilde{\mathbf{n}}}]\}} \left\| \begin{pmatrix} 1 \\ \lambda^{-1}\alpha_{\vec{\mathbf{i}}} \end{pmatrix} \right\| \mathcal{E}_{\vec{\mathbf{i}}}\mathcal{Y}_{\vec{\mathbf{i}}}, \\
W_{3\vec{\mathbf{i}}} &= \frac{l_{\tilde{\mathbf{n}}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)} \left\| \begin{pmatrix} 1 \\ \lambda^{-1}\alpha_{\vec{\mathbf{i}}} \end{pmatrix} \right\| \mathcal{E}_{\vec{\mathbf{i}}}.
\end{aligned}$$

For $W_{1\vec{\mathbf{i}}}$ and $W_{3\vec{\mathbf{i}}}$, we adopt the same techniques as in Lemma (2), where $\varsigma_{\vec{\mathbf{i}}}$ is replaced by $W_{1\vec{\mathbf{i}}}$ and $W_{3\vec{\mathbf{i}}}$. Thus, since $\mathbb{E}[W_{1\vec{\mathbf{i}}}] = O(l_n\varphi_x(\lambda))$ and $\mathbb{E}[W_{3\vec{\mathbf{i}}}] = O(l_{\tilde{\mathbf{n}}}^2\varphi_x(\lambda))$, we obtain

$$|W_{1\vec{\mathbf{i}}}| = O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right) \quad \text{and} \quad |W_{3\vec{\mathbf{i}}}| = O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right).$$

However, as $W_{2\vec{\mathbf{i}}}$ is unbounded, we analyze via the truncation method. Indeed, we define

$$W_{2\vec{\mathbf{i}}} = \frac{1}{\tilde{\mathbf{n}}\varphi_x(\lambda)} \sup_{\|\varsigma\|\leq M} 1\!\!1_{\{[|\mathcal{Y}_{\vec{\mathbf{i}}}-(\rho_j+EXP_p(x))-(\lambda^{-1}\upsilon_j+EXP'_p(x))\alpha_{\vec{\mathbf{i}}}|\leq Cl_{\tilde{\mathbf{n}}}]\}} \left\| \begin{pmatrix} 1 \\ \lambda^{-1}\alpha_{\vec{\mathbf{i}}} \end{pmatrix} \right\| \mathcal{E}_{\vec{\mathbf{i}}}\mathcal{Y}^*_{\vec{\mathbf{i}}},$$

such that

$$\mathcal{Y}^*_{\vec{\mathbf{i}}} = \mathcal{Y}_{\vec{\mathbf{i}}} 1\!\!1_{\{\mathcal{Y}_{\vec{\mathbf{i}}}|<\gamma_{\mathbf{n}}\}}.$$

So, the convergence of $W_{2\vec{\mathbf{i}}}$ is a consequence of

$$\mathbb{P}\left(\left|W^*_{2\vec{\mathbf{i}}} - W_{2\vec{\mathbf{i}}}\right| > \eta\sqrt{\frac{\log \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right), \tag{A5}$$

$$\left|\mathbb{E}[W^*_{2\vec{\mathbf{i}}}] - \mathbb{E}[W_{2\vec{\mathbf{i}}}]\right| = o\left(\sqrt{\frac{\log \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right), \tag{A6}$$

and

$$\mathbb{P}\left(\left|W^*_{2\vec{\mathbf{i}}} - \mathbb{E}[W^*_{2\vec{\mathbf{i}}}\right| > \eta\sqrt{\frac{\log \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right). \tag{A7}$$

Concerning (A5). By Markov's inequality,

$$\begin{aligned}\mathbb{P}\left(\left|W^*_{2\vec{\mathbf{i}}} - W_{2\vec{\mathbf{i}}}\right| > \epsilon_0\left(\sqrt{\frac{\log \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\,\varphi_x(\lambda)}}\right)\right) &\leq \sum_{\vec{\mathbf{i}}\in\mathcal{I}_{\tilde{\mathbf{n}}}} \mathbb{P}\left(\mathcal{Y}_{\vec{\mathbf{i}}} > \gamma_{\mathbf{n}}\right) \\ &\leq \tilde{\mathbf{n}}\gamma_{\mathbf{n}}^{-q}\mathbb{E}\left[\mathcal{Y}^q\right].\end{aligned}$$

It follows that

$$\mathbb{P}\left(\left|W^*_{2\vec{\mathbf{i}}} - W_{2\vec{\mathbf{i}}}\right| > \eta\sqrt{\frac{\log \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right) \leq C\tilde{\mathbf{n}}\gamma_{\mathbf{n}}^{-q}. \tag{A8}$$

For (A6). By Holder's inequality with $\iota_1 = \frac{q}{2}$, for ι_2, such that

$$\frac{1}{\iota_1} + \frac{1}{\iota_2} = 1,$$

we have

$$\begin{aligned}&\left|\mathbb{E}[W^*_{2\vec{\mathbf{i}}}] - \mathbb{E}[W_{2\vec{\mathbf{i}}}]\right| \\ &\leq \frac{C}{\mathbb{E}\left[\mathcal{E}_{\mathbf{1}}\right]}\mathbb{E}\left[|\mathcal{Y}|\mathbb{1}_{\left\{\mathcal{Y}\geq\gamma_{\mathbf{n}}\right\}}\mathcal{E}_{\vec{\mathbf{i}}}\right] \\ &\leq \frac{\gamma_{\mathbf{n}}^{-1}}{\mathbb{E}\left[\mathcal{E}_{\mathbf{1}}\right]}\mathbb{E}^{1/\iota_1}\left[|\mathcal{Y}^{2\iota_1}|\right]\mathbb{E}^{1/\iota_2}\left[\mathcal{E}_{\mathbf{1}}^{\iota_2}\right] \\ &\leq \frac{C\gamma_{\mathbf{n}}^{-1}}{\mathbb{E}\left[\mathcal{E}_{\mathbf{1}}\right]}\mathbb{E}^{1/\iota_2}\left[\mathcal{E}_{\mathbf{1}}^{\iota_2}\right],\end{aligned}$$

allowing

$$\left|\mathbb{E}[W^*_{2\vec{\mathbf{i}}}] - \mathbb{E}[W_{2\vec{\mathbf{i}}}]\right| \leq C\gamma_{\mathbf{n}}^{-1}\varphi_x^{(1-\iota_2)/\iota_2}(\lambda). \tag{A9}$$

Concerning (A7). We adopt spatial block techniques as in Lemma 2, where $\varsigma_{\vec{\mathbf{i}}}$ is replaced by $W^*_{2\vec{\mathbf{i}}}$. The main difference is in the variance term. For $W^*_{2\vec{\mathbf{i}}}$, we have

$$\begin{aligned}Var\left[W^*_{2\vec{\mathbf{i}}}\right] &\leq Cl_{\vec{\mathbf{n}}};\mathbb{E}\left[\mathcal{E}^2_{\vec{\mathbf{i}}}\mathcal{Y}^{*2}_{\vec{\mathbf{i}}}\right] \leq C\,Cl_{\vec{\mathbf{n}}}\mathbb{E}\left[\mathcal{E}^2_{\vec{\mathbf{i}}}\mathcal{Y}^2_{\vec{\mathbf{i}}}\right] \\ &\leq Cl_{\vec{\mathbf{n}}}\mathbb{E}\left[\mathcal{E}^2_{\vec{\mathbf{i}}}\mathbb{E}\left[\mathcal{Y}^2_{\vec{\mathbf{i}}}|X_{\vec{\mathbf{i}}}\right]\right] \leq Cl_{\vec{\mathbf{n}}}\varphi_x(\lambda).\end{aligned}$$

It follows that

$$\sum_{\vec{\mathbf{i}}\in I(1,\vec{\mathbf{n}},\mathbf{1})} Var\left[W^*_{2\vec{\mathbf{i}}}\right] = O\left(l_{\vec{\mathbf{n}}}p^N_{\vec{\mathbf{n}}}\varphi_x(\lambda)\right).$$

Next, for $\vec{\mathbf{i}} \neq \vec{\mathbf{j}}$, we have

$$\begin{aligned} cov(\Lambda_{\vec{\mathbf{i}}}, \Lambda_{\vec{\mathbf{j}}}) &\leq C\, \mathbb{E}\Big[\mathcal{E}_{\vec{\mathbf{i}}}|\mathcal{Y}^*_{\vec{\mathbf{i}}}|\mathcal{E}_{\vec{\mathbf{j}}}|\mathcal{Y}^*_{\vec{\mathbf{j}}}|\Big] \\ &\leq C\, \mathbb{E}\Big[\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{E}_{\mathbf{j}}|\mathcal{Y}_{\vec{\mathbf{i}}}\mathcal{Y}_{\vec{\mathbf{j}}}|\Big] \\ &\leq C\, \mathbb{E}\Big[\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{E}_{\mathbf{j}}\mathbb{E}\Big[|\mathcal{Y}_{\vec{\mathbf{i}}}\mathcal{Y}_{\vec{\mathbf{j}}}||X_{\vec{\mathbf{i}}}X_{\vec{\mathbf{j}}}\Big]\Big] \\ &\leq C\, \mathbb{E}\Big[\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{E}_{\vec{\mathbf{j}}}\Big] \leq C\, \varphi_x x^{(a+1)/a}(\lambda). \end{aligned}$$

But, as $\mathbb{E}\Big[\mathcal{Y}^p_{\mathbf{i}}|X_{\vec{\mathbf{i}}}\Big] < \infty$, we give $\forall \vec{\mathbf{i}} \neq \vec{\mathbf{j}}$

$$\begin{aligned} cov(\Lambda_{\vec{\mathbf{i}}}, \Lambda_{\vec{\mathbf{j}}}) &\leq \|\Lambda_{\vec{\mathbf{i}}}\|_p^2 \varphi^{1-2/p}(\|\vec{\mathbf{i}} - \vec{\mathbf{j}}\|) \\ &\leq C\, \|\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{Y}^*_{\vec{\mathbf{i}}}\|_p^2 \varphi^{1-2/p}(\|\vec{\mathbf{i}} - \vec{\mathbf{j}}\|) \\ &\leq C\, \|\mathcal{E}_{\vec{\mathbf{i}}}\mathcal{Y}_{\vec{\mathbf{i}}}\|_p^2 \varphi^{1-2/p}(\|\vec{\mathbf{i}} - \vec{\mathbf{j}}\|) \\ &\leq C\, \|\mathcal{E}_{\vec{\mathbf{i}}}\|_p^2 \varphi^{1-2/p}(\|\vec{\mathbf{i}} - \vec{\mathbf{j}}\|) \\ &\leq C\, \varphi_x x^{2/p}(\lambda) \varphi^{1-2/p}(\|\vec{\mathbf{i}} - \vec{\mathbf{j}}\|)). \end{aligned}$$

Therefore, for $c_{\mathbf{n}} = \varphi_x x(\lambda)^{2/Np(a+1)-1/Na}$, we infer

$$\begin{aligned} &\sum_{\vec{\mathbf{i}} \neq \vec{\mathbf{j}} \in I(1,\vec{\mathbf{n}},\mathbf{1})} \Big| cov(\Lambda_{\vec{\mathbf{i}}}, \Lambda_{\vec{\mathbf{j}}})\Big| \\ &\leq \sum_{\Big\{\vec{\mathbf{i}},\vec{\mathbf{j}} \in I(1,\vec{\mathbf{n}},\mathbf{1}) \,\Big\|\vec{\mathbf{i}}-\vec{\mathbf{j}}\Big\| \leq c_{\vec{\mathbf{n}}}\Big\}} \Big| cov(\Lambda_{\vec{\mathbf{i}}}, \Lambda_{\vec{\mathbf{j}}})\Big| \\ &\quad + \sum_{\Big\{\vec{\mathbf{i}},\vec{\mathbf{j}} \in I(1,\vec{\mathbf{n}},\mathbf{1}) \,\Big\|\vec{\mathbf{i}}-\vec{\mathbf{j}}\Big\| > c_{\vec{\mathbf{n}}}\Big\}} \Big| cov(\Lambda_{\vec{\mathbf{i}}}, \Lambda_{\vec{\mathbf{j}}})\Big| \\ &\leq C\, p^N_{\vec{\mathbf{n}}} \varphi_x x(\lambda) \Big(c^N_{\vec{\mathbf{n}}} \varphi_x x(\lambda)^{1/a} \\ &\quad + c_{\vec{\mathbf{n}}}^{-Na} \varphi_x x^{2/p-1}(\lambda) \sum_{\vec{\mathbf{i}}:\|\vec{\mathbf{i}}\| \geq c_{\vec{\mathbf{n}}}} \Big\|\vec{\mathbf{i}}\Big\|^{Na} \varphi^{1-2/p}\Big(\Big\|\vec{\mathbf{i}}\Big\|\Big)\Big) \\ &\leq C p^N_{\vec{\mathbf{n}}} \varphi_x x(\lambda). \end{aligned}$$

Finally,

$$Var\Big[\sum_{\vec{\mathbf{i}} \in I(1,\vec{\mathbf{n}},\mathbf{1})} \Lambda_{\vec{\mathbf{i}}}\Big] = O\Big(p^N_{\vec{\mathbf{n}}} \varphi_x x(\lambda)\Big),$$

allowing

$$\mathcal{V}_1 \leq \exp\Big(- C(\eta_0) \log \tilde{\mathbf{n}}\Big). \tag{A10}$$

This implies

$$\big|W_{1\vec{\mathbf{i}}}\big| = O_{a.co.}\left(\sqrt{\frac{\ln \tilde{\mathbf{n}}}{\tilde{\mathbf{n}}\varphi_x(\lambda)}}\right).$$

which deduces the result of this lemma.

□

References

1. Ripley, B.D. *Spatial Statistics*; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons, Inc.: New York, NY, USA, 1981; p. x+252.
2. Rosenblatt, M. *Stationary Sequences and Random Fields*; Birkhäuser Boston, Inc.: Boston, MA, USA, 1985; p. 258. [CrossRef]
3. Guyon, X. *Random Fields on a Network: Modeling, Statistics, and Applications*; Probability and Its Applications (New York); Translated from the 1992 French original by Carenne Ludeña; Springer: New York, NY, USA, 1995; p. xii+255.
4. Cressie, N.A.C. *Statistics for Spatial Data*, revised ed.; Paperback edition of the 1993 edition [MR1239641]; Wiley Classics Library, John Wiley & Sons, Inc.: New York, NY, USA, 2015; p. xx+900.
5. Bouzebda, S.; Soukarieh, I. Non-Parametric Conditional U-Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design. *Mathematics* **2023**, *11*, 16. [CrossRef]
6. Tran, L.T. Kernel density estimation on random fields. *J. Multivar. Anal.* **1990**, *34*, 37–53. [CrossRef]
7. Lu, Z.; Chen, X. Spatial kernel regression estimation: Weak consistency. *Statist. Probab. Lett.* **2004**, *68*, 125–136. [CrossRef]
8. Biau, G.; Cadre, B. Nonparametric spatial prediction. *Stat. Inference Stoch. Process.* **2004**, *7*, 327–349. [CrossRef]
9. Hallin, M.; Lu, Z.; Tran, L.T. Local linear spatial regression. *Ann. Statist.* **2004**, *32*, 2469–2500. [CrossRef]
10. Carbon, M.; Francq, C.; Tran, L.T. Kernel regression estimation for random fields. *J. Statist. Plann. Inference* **2007**, *137*, 778–798. [CrossRef]
11. Xu, R.; Wang, J. L_1-estimation for spatial nonparametric regression. *J. Nonparametr. Stat.* **2008**, *20*, 523–537. [CrossRef]
12. Li, J.; Tran, L.T. Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference* **2009**, *139*, 164–175. [CrossRef]
13. Bouzebda, S.; Slaoui, Y. Nonparametric recursive method for kernel-type function estimators for spatial data. *Statist. Probab. Lett.* **2018**, *139*, 103–114. [CrossRef]
14. Bouzebda, S.; Slaoui, Y. Large and moderate deviation principles for recursive kernel estimators of a regression function for spatial data defined by stochastic approximation method. *Statist. Probab. Lett.* **2019**, *151*, 17–28. [CrossRef]
15. Dabo-Niang, S.; Rachdi, M.; Yao, A.F. Kernel regression estimation for spatial functional random variables. *Far East J. Theor. Stat.* **2011**, *37*, 77–113.
16. Dabo-Niang, S.; Kaid, Z.; Laksaci, A. On spatial conditional mode estimation for a functional regressor. *Statist. Probab. Lett.* **2012**, *82*, 1413–1421. [CrossRef]
17. Dabo-Niang, S.; Kaid, Z.; Laksaci, A. Asymptotic properties of the kernel estimate of spatial conditional mode when the regressor is functional. *AStA Adv. Stat. Anal.* **2015**, *99*, 131–160. [CrossRef]
18. Laksaci, A.; Maref, F. Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *C. R. Math. Acad. Sci.* **2009**, *347*, 1075–1080. [CrossRef]
19. Al-Awadhi, F.A.; Kaid, Z.; Laksaci, A.; Ouassou, I.; Rachdi, M. Functional data analysis: Local linear estimation of the L_1-conditional quantiles. *Stat. Methods Appl.* **2019**, *28*, 217–240. [CrossRef]
20. Rachdi, M.; Laksaci, A.; Al-Kandari, N.M. Expectile regression for spatial functional data analysis (sFDA). *Metrika* **2022**, *85*, 627–655. [CrossRef]
21. Newey, W.K.; Powell, J.L. Asymmetric least squares estimation and testing. *Econometrica* **1987**, *55*, 819–847. [CrossRef]
22. Aigner, D.J.; Amemiya, T.; Poirier, D.J. On the estimation of production frontiers: Maximum likelihood estimation of the parameters of a discontinuous density function. *Internat. Econ. Rev.* **1976**, *17*, 377–396. [CrossRef]
23. Efron, B. Regression percentiles using asymmetric squared error loss. *Statist. Sin.* **1991**, *1*, 93–125.
24. Abdous, B.; Rémillard, B. Relating quantiles and expectiles under weighted-symmetry. *Ann. Inst. Statist. Math.* **1995**, *47*, 371–384. [CrossRef]
25. Mohammedi, M.; Bouzebda, S.; Laksaci, A. The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data. *J. Multivar. Anal.* **2021**, *181*, 104673. [CrossRef]
26. Bellini, F.; Bignozzi, V.; Puccetti, G. Conditional expectiles, time consistency and mixture convexity properties. *Insur. Math. Econ.* **2018**, *82*, 117–123. [CrossRef]
27. Almanjahie, I.M.; Bouzebda, S.; Kaid, Z.; Laksaci, A. Nonparametric estimation of expectile regression in functional dependent data. *J. Nonparametr. Stat.* **2022**, *34*, 250–281. [CrossRef]
28. Alshahrani, F.; Almanjahie, I.M.; Elmezouar, Z.C.; Kaid, Z.; Laksaci, A.; Rachdi, M. Functional Ergodic Time Series Analysis Using Expectile Regression. *Mathematics* **2022**, *10*, 3919. [CrossRef]
29. Almanjahie, I.M.; Bouzebda, S.; Chikr Elmezouar, Z.; Laksaci, A. The functional *k*NN estimator of the conditional expectile: uniform consistency in number of neighbors. *Stat. Risk Model.* **2022**, *38*, 47–63. [CrossRef]
30. Yao, D.S.; Chen, W.X.; Long, C.X. Parametric estimation for the simple linear regression model under moving extremes ranked set sampling design. *Appl. Math. J. Chin. Univ. Ser. B* **2021**, *36*, 269–277. [CrossRef]
31. Chowdhury, J.; Chaudhuri, P. Convergence rates for kernel regression in infinite-dimensional spaces. *Ann. Inst. Statist. Math.* **2020**, *72*, 471–509. [CrossRef]
32. Aneiros, G.; Horová, I.; Hušková, M.; Vieu, P. On functional data analysis and related topics. *J. Multivar. Anal.* **2022**, *189*, 104861. [CrossRef]
33. Rachdi, M. Special Issue on Functional Data Analysis: Theory and Applications to Different Scenarios. *Mathematics* **2022–2023**, *10*. Available online: https://www.mdpi.com/journal/mathematics/special_issues/45POZ9BG9S (accessed on 21 October 2022).

34. Cardot, H.; Crambes, C.; Sarda, P. Quantile regression when the covariates are functions. *J. Nonparametr. Stat.* **2005**, *17*, 841–856. [CrossRef]
35. Novo, S.; Aneiros, G.; Vieu, P. Automatic and location-adaptive estimation in functional single-index regression. *J. Nonparametr. Stat.* **2019**, *31*, 364–392. [CrossRef]
36. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis*; Theory and Practice; Springer Series in Statistics; Springer: New York, NY, USA, 2006; p. xx+258.
37. Aneiros-Pérez, G.; Vieu, P. Automatic estimation procedure in partial linear model with functional data. *Statist. Pap.* **2011**, *52*, 751–771. [CrossRef]
38. Demongeot, J.; Hamie, A.; Laksaci, A.; Rachdi, M. Relative-error prediction in nonparametric functional statistics: theory and practice. *J. Multivar. Anal.* **2016**, *146*, 261–268. [CrossRef]
39. Soukarieh, I.; Bouzebda, S. Weak Convergence of the Conditional *U*-statistics for Locally Stationary Functional Time Series. In *Statistical Inference for Stochastic Processes*; Springer Nature: Berlin/Heidelberg, Germany, 2024; pp. 1–85.
40. Doukhan, P. *Mixing: Properties and Examples*; Lecture Notes in Statistics; Springer: New York, NY, USA, 1994; Volume 85, p. xii+142. [CrossRef]
41. Tjøstheim, D. Statistical spatial series modelling. *Adv. Appl. Probab.* **1978**, *10*, 130–154. [CrossRef]
42. Moore, M. Spatial linear processes. *Comm. Statist. Stoch. Model.* **1988**, *4*, 45–75. [CrossRef]
43. Guyon, X. Estimation d'un champ par pseudo-vraisemblance conditionnelle: Étude asymptotique et application au cas markovien. In *Spatial Processes and Spatial Time Series Analysis (Brussels, 1985)*; Travaux Rech.; Publication des Facultés Universitaires Saint-Louis: Saint-Louis, Brussels, 1987; Volume 11, pp. 15–62.
44. Strang, G. Wavelet transforms versus Fourier transforms. *Bull. Am. Math. Soc.* **1993**, *28*, 288–305. [CrossRef]
45. Fan, J.; Gijbels, I. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability*; Chapman & Hall: London, UK, 1996; Volume 66, p. xvi+341.
46. Silverman, B.W. *Density Estimat Ion for Statistics and Data Analysis*; Monographs on Statistics and Applied Probability; Chapman & Hall: London, UK, 1986; p. x+175. [CrossRef]
47. Cheng, M.Y.; Fan, J.; Marron, J.S. On automatic boundary corrections. *Ann. Statist.* **1997**, *25*, 1691–1708. [CrossRef]
48. Barrientos-Marin, J.; Ferraty, F.; Vieu, P. Locally modelled regression and functional data. *J. Nonparametr. Stat.* **2010**, *22*, 617–632. [CrossRef]
49. Canton Enriquez, D.; Niembro-Ceceña, J.A.; Muñoz Mandujano, M.; Alarcon, D.; Arcadia Guerrero, J.; Gonzalez Garcia, I.; Montes Gutierrez, A.A.; Gutierrez-Lopez, A. Application of probabilistic models for extreme values to the COVID-2019 epidemic daily dataset. *Data Brief* **2022**, *40*, 107783. [CrossRef]
50. Daouia, A.; Gijbels, I.; Stupfler, G. Extremile regression. *J. Am. Statist. Assoc.* **2022**, *117*, 1579–1586. [CrossRef]
51. Maume-Deschamps, V.; Rullière, D.; Said, K. Multivariate extensions of expectiles risk measures. *Depend. Model.* **2017**, *5*, 20–44. [CrossRef]
52. Ibragimov, I.A.; Linnik, Y.V. *Independent and Stationary Sequences of Random Variables*; Wolters-Noordhoff Publishing: Groningen, The Netherlands, 1971; p. 443; With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.
53. Rachdi, M.; Laksaci, A.; Demongeot, J.; Abdali, A.; Madani, F. Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Comput. Statist. Data Anal.* **2014**, *73*, 53–68. [CrossRef]