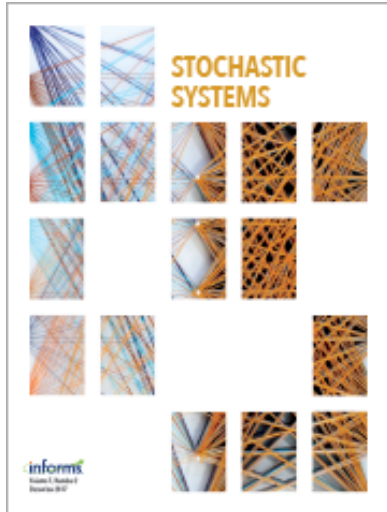




Stochastic Systems

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Functional Limit Theorems for Shot Noise Processes with Weakly Dependent Noises

Guodong Pang, Yuhang Zhou

To cite this article:

Guodong Pang, Yuhang Zhou (2020) Functional Limit Theorems for Shot Noise Processes with Weakly Dependent Noises. *Stochastic Systems* 10(2):99–123. <https://doi.org/10.1287/stsy.2019.0051>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Stochastic Systems*. Copyright © 2020 The Author(s). <https://doi.org/10.1287/stsy.2019.0051>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2020, The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Functional Limit Theorems for Shot Noise Processes with Weakly Dependent Noises

Guodong Pang,^a Yuhang Zhou^a

^aHarold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, Pennsylvania 16802

Contact: gup3@psu.edu,  <https://orcid.org/0000-0001-6468-0429> (GP); zhouyhnk@gmail.com (YZ)

Received: December 6, 2018

Revised: March 22, 2019; May 30, 2019

Accepted: July 28, 2019

Published Online in Articles in Advance:
June 2, 2020

<https://doi.org/10.1287/stsy.2019.0051>

Copyright: © 2020 The Author(s)

Abstract. We study shot noise processes when the shot noises are weakly dependent, satisfying the ρ -mixing condition. We prove a functional weak law of large numbers and a functional central limit theorem for this shot noise process in an asymptotic regime with a high intensity of shots. The deterministic fluid limit is unaffected by the presence of weak dependence. The limit in the diffusion scale is a continuous Gaussian process whose covariance function explicitly captures the dependence among the noises. The model and results can be applied in financial and insurance risks with dependent claims as well as queueing systems with dependent service times. To prove the existence of the limit process, we employ the existence criterion that uses a maximal inequality requiring a set function with a superadditivity property. We identify such a set function for the limit process by exploiting the ρ -mixing condition. To prove the weak convergence, we establish the tightness property and the convergence of finite dimensional distributions. To prove tightness, we construct two auxiliary processes and apply an Ottaviani-type inequality for weakly dependent sequences.



Open Access Statement: This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “Stochastic Systems. Copyright © 2020 The Author(s). <https://doi.org/10.1287/stsy.2019.0051>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Funding: This work is partly supported by the National Science Foundation [Grants CMMI-1538149, CMMI-1635410, and DMS/CMMI-1715875] and the Army Research Office [Grant W911NF-17-1-0019].

Keywords: shot noise processes • functional weak law of large numbers • functional central limit theorem • Gaussian process limit • weakly dependent shot noises • ρ -mixing • new existence criterion in the space $\mathbb{D}$ • set function with superadditivity property • Ottaviani-type inequality

1. Introduction

Shot noise processes have been extensively studied in applied probability and stochastic models. They have been used in risk and insurance theory, financial models, queueing theory, physics, and so on (see, e.g., Rice 1977, Lane 1984, Schmidt 1985, Samorodnitsky 1996, Budhiraja and Nyquist 2015, Iksanov 2016, Koops et al. 2017). A shot noise process $X := \{X(t) : t \geq 0\}$ is typically defined by

$$X(t) := \sum_{i=1}^{A(t)} H(t - \tau_i, Z_i), \quad t \geq 0, \quad (1)$$

where $A := \{A(t) : t \geq 0\}$ is a counting process of shots with arrival times $\{\tau_i : i \in \mathbb{N}\}$, $\{Z_i : i \in \mathbb{N}\}$ is a sequence of $\mathbb{R}^k$ -valued ($k \geq 1$) random vectors denoting the noises, and $H : \mathbb{R}_+ \times \mathbb{R}^k \rightarrow \mathbb{R}$ is a deterministic product-measurable function representing the shot shape or the impulse response function.

The sequence of random variables $\{Z_i : i \in \mathbb{N}\}$ in (1) is often assumed to be independent and identically distributed (i.i.d.) in the literature (see, e.g., Lane 1984, Huffer 1987, Doney and O’Brien 1991, Klüppelberg and Mikosch 1995, Samorodnitsky 1998, Klüppelberg et al. 2003, Klüppelberg and Kühn 2004, Ganesh et al. 2005, Ganesh and Torrisi 2006, Iksanov 2016). Because of the vast literature, we highlight the important results that are most relevant to our asymptotic analysis on functional limit theorems. With Poisson arrivals and i.i.d. noises, it is shown in Klüppelberg and Mikosch (1995) and Klüppelberg and Kühn (2004) that, under certain conditions on the shot shape function, the shot noise process presents long-range dependence phenomena, and the properly scaled process (under the conventional asymptotic regime scaling both time and space) results in a self-similar Gaussian process and fractional Brownian motion (FBM). Recently, in Pang and Taquq (2019),

with Poisson arrivals and power-law nonstationary noises, it is shown that the integrated shot noise process under the conventional scaling has a self-similar Gaussian process, which has nonstationary increments (a generalization of FBM). It is also shown in Klüppelberg et al. (2003), Iksanov (2013, 2016), and Iksanov et al. (2014, 2017a, b) that the scaling limit can be also an infinite-variance stable process with i.i.d. noises provided with certain power-law shot shape functions and/or renewal arrivals. As a special case of the compound stochastic processes studied by Iglehart (1973), it is shown that the scaled shot noise processes with a renewal arrival process and i.i.d. noises, in the large asymptotic regime (only scaling up the arrival rate and space without scaling time), result in a continuous Gaussian process. In addition, as a special case of the shot noise processes with conditionally independent and nonstationary noises in Pang and Zhou (2018a), when the noises are i.i.d., under more general conditions, a continuous Gaussian process limit is also established, extending the results in Iglehart (1973). It is worth noting that, unlike the conventional scaling regime, the Gaussian limiting processes in Iglehart (1973) and Pang and Zhou (2018a) do not exhibit long-range dependence phenomena.

However, the i.i.d. assumption on the noises $\{Z_i : i \in \mathbb{N}\}$ can be restrictive for certain applications. In insurance risk theory, ruin probabilities are studied with $\{Z_i : i \in \mathbb{N}\}$ being dependent claims in Valdez and Mo (2002). In Ramirez-Perez and Serfling (2001), a cluster shot noise model is studied in which $\{Z_i : i \in \mathbb{N}\}$ depends on the same “cluster mark” within each cluster so that the noises within each cluster are correlated. Because the queueing process (or the process counting the number of busy servers) of infinite-server queues is a special class of shot noise processes, the i.i.d. assumption on the service times does not hold for some applications, for example, the length of stay (treatment process) of patients at emergency rooms can be correlated (Pang and Whitt 2012, 2013). It is, therefore, important to understand the asymptotic behavior of the shot noise processes with dependent noises. In this paper, we focus on shot noises that are weakly dependent, satisfying the ρ -mixing condition. Many dependent variables satisfy this mixing condition, for example, strictly stationary, countable state, irreducible and aperiodic Markov chains, and the discrete autoregressive process of order one.

We prove a functional weak law of large numbers (FWLLN) and a functional central limit theorem (FCLT) for shot noise processes with weakly dependent noises in an asymptotic regime in which the arrival rate is large and the shot noise distributions F and shot shape function H are fixed (unscaled); see Assumptions 2 and 3. Such a regime is often referred to as the “high-intensity/density regime” (see, e.g., Papoulis 1971, Iglehart 1973, Heinrich and Schmidt 1985, Biermé and Desolneux 2012). In this regime, we assume that the arrival process, after being appropriately centered and scaled, converges weakly to a continuous limit process; see Assumption 3. This includes renewal processes and more general (nonstationary) counting processes as long as such an FCLT holds. It is even possible that the interarrival times are weakly dependent under mixing conditions, which results in a Brownian motion limit (see chapter 19 in Billingsley 1999 together with the inverse mapping in Whitt 2002). We establish a stochastic process limit for the similarly centered and scaled shot noise processes (Theorem 2). The limit process can be decomposed into a sum of two independent processes: one as an integral functional of the limit arrival process and the other as a continuous Gaussian process. When the arrival limit is Gaussian, the limit process becomes a Gaussian process. We give explicit characterizations for the limit processes in Theorem 2. It is worth mentioning that the weakly dependent assumption has no impact upon the fluid limit (Theorem 1). There is an extra term, compared with the i.i.d. case, in the covariance structure of the limit Gaussian process (the function $\Gamma(t, s, u)$ in Theorem 2), which itself captures all the dependence in the noises. It is also important to note that, like the i.i.d. and conditionally independent cases, the Gaussian limit in the case with weakly dependent noises does not exhibit long-range dependence phenomena. We provide two examples to explicitly calculate the extra term and demonstrate the impact of dependence among noises upon the covariance functions. They include an example of a first-order autoregressive process of shot noises and a shot noise process with batch arrivals in which noises within each batch are symmetrically correlated and noises across batches are mutually independent.

The general class of shot noise processes studied in this paper includes several well-known models as special cases. First, multiplicative shot noise models are an important class, in which the shot shape function H takes the form $H(t, x) = \tilde{H}(t)\varphi(x)$ for a monotone function $\tilde{H}(t)$ and a measurable function $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}$. Such a model has been widely studied in financial modeling (Chobanov 1999) and insurance risk (Ganesh and Torrisi 2006, Schmidt 2014), in which, for example, the function $\tilde{H}$ can have an exponential, power-law, or polynomial decay. Second, the compound process with weakly dependent variables is also a special case of the multiplicative models; see Theorem 3. It can be used to model batch arrival processes when the batch sizes are correlated or the work-input process when the work requirements are correlated provided that their correlation structure satisfies the ρ -mixing condition. Third, the queueing and work-input processes of infinite-server

queues are also special cases of shot noise processes although not belonging to the multiplicative models. In this setting, the shot shape function H takes the forms $H(t, x) = \mathbf{1}(t < x)$ and $H(t, x) = x\mathbf{1}(t < x)$ for the queue length and work-input processes, respectively. The (two-parameter) queueing process has been well studied in Pang and Whitt (2013) and Pang and Zhou (2018b). As a consequence of Theorem 2, we obtain an FCLT for the work-input process for the infinite-server queues with ρ -mixing service times (Theorem 4). In all these special models, we use the first-order autoregressive process of shot noises to demonstrate the impact of the noises upon the covariance functions.

The proof of the FCLT consists of three key components. The first is to establish the existence of the limiting Gaussian process in the space $\mathbb{C}$. For this, the standard existence criterion, theorem 13.6 in Billingsley (1999), cannot be used. Instead, we can apply the new existence criterion, theorem 5.3 in Pang and Zhou (2018a), which relies on a new maximal inequality using a set function with the superadditivity property instead of being a finite measure. By calculating the second moment of the process increment and exploiting the properties of the ρ -mixing condition, we are able to identify such a set function. It is worth noting that such a set function may not be always identified under any mixing condition; for example, the strong α -mixing condition fails. The second is to prove the convergence of finite dimensional distributions (Lemma 5), for which we adapt the central limit theorems (CLTs) for dependent variables under mixing conditions Withers (1981) and Dehling and Philipp (2002). The third is to show tightness, which is technically the most challenging component. We construct two auxiliary processes and use sophisticated moment bounds for dependent variables under strong mixing conditions (Andrews and Pollard 1994, Shao and Yu 1996) and apply an Ottaviani-type inequality (see, e.g., Bücher 2015) in a nontrivial manner. See also Remark 4.

Finally, it is worth highlighting that the high-intensity/density regime is different from the conventional regime in which the time and shot noises are being scaled up simultaneously. It remains open to study the shot noise processes with weakly dependent noises under the conventional scaling regime.

1.1. Organization of the Paper

The rest of the paper is organized as follows. In Section 1.2, we collect notation used throughout the paper. In Section 2, we describe the model in detail, state the main results, and present several examples. We also discuss the multiplicative models and the work-input processes in infinite-server queues in Section 2. The proofs are given in Section 3. In the appendix, we provide an overview of the new existence criterion in Pang and Zhou (2018a) and the Ottaviani-type inequality in Bücher (2015).

1.2. Notation

Throughout the paper, $\mathbb{N}$ denotes the set of natural numbers. $\mathbb{R}^k$ ($\mathbb{R}_+^k$) denotes the space of real-valued (non-negative) k -dimensional vectors, and we write $\mathbb{R}$ ($\mathbb{R}_+$) for $k = 1$. For $a, b \in \mathbb{R}$, we write $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Let $\mathbb{D}^k = \mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$ denote the space of $\mathbb{R}^k$ -valued càdlàg functions on $\mathbb{R}_+$. $(\mathbb{D}^k, J_1)$ denotes space $\mathbb{D}^k$ equipped with Skorohod J_1 topology with the metric d_{J_1} (see, e.g., (12.13) in Billingsley 1999 and (3.2) and its extension to domain $\mathbb{R}_+$ in Whitt 2002, chapter 3). Note that the space $(\mathbb{D}^k, J_1)$ is complete and separable. We write $\mathbb{D}$ for $\mathbb{D}^k$ when $k = 1$. Let $\mathbb{C} \subset \mathbb{D}$ be the subset of continuous functions. When considering functions defined on finite intervals, we write $\mathbb{D}([0, T], \mathbb{R})$ for $T > 0$. All random variables and processes are defined on a common complete probability space $(\Omega, \mathcal{F}, P)$. Notations $\rightarrow$ and $\Rightarrow$ mean convergence of real numbers and convergence in distribution, respectively. The abbreviation *a.s.* means *almost surely*. We use lower-case o notation for real-valued function f and nonzero g , and we write $f(x) = o(g(x))$ if $\lim_{x \rightarrow \infty} |f(x)/g(x)| = 0$.

2. Model and Results

We first state the assumptions to study the shot noise process X defined in (1). We start with the shot shape or the response function H . To simplify notations, we define the following functions: for each $0 \leq u \leq s \leq t$,

$$\begin{aligned} G_1(t, u) &:= E[H(t - u, Z_1)] = \int_{\mathbb{R}^k} H(t - u, x) dF(x), \\ G_2(t, u) &:= E[H(t - u, Z_1)^2] = \int_{\mathbb{R}^k} H(t - u, x)^2 dF(x), \\ \tilde{G}(t, u) &:= \text{Var}(H(t - u, Z_1)) = G_2(t, u) - G_1(t, u)^2, \end{aligned} \tag{2}$$

$$\begin{aligned}\check{G}_1(t, s, u) &:= E[H(t - u, Z_1) - H(s - u, Z_1)] = \int_{\mathbb{R}^k} (H(t - u, x) - H(s - u, x)) dF(x), \\ \check{G}_2(t, s, u) &:= E[(H(t - u, Z_1) - H(s - u, Z_1))^2] = \int_{\mathbb{R}^k} (H(t - u, x) - H(s - u, x))^2 dF(x), \\ \tilde{G}(t, s, u) &:= \text{Var}(H(t - u, Z_1) - H(s - u, Z_1)) = \check{G}_2(t, s, u) - \check{G}_1(t, s, u)^2.\end{aligned}$$

Assumption 1. For each $x \in \mathbb{R}^k$, the shot shape function $H(\cdot, x) \in \mathbb{D}$ is monotone. The function $G_1(t, u)$ is continuous in u for each $t \geq 0$. For $0 \leq t \leq T$,

$$\lim_{\delta \downarrow 0} \int_{[0, T]} \check{G}_2(t, t - \delta, u) d\Lambda(u) = 0. \quad (3)$$

Remark 1. The monotonicity condition of $H(t, x)$ in t is usually assumed in the literature of shot noise processes and is critical in this paper, for example, in the proofs of Lemmas 6 and 8. The continuity of $G_1(t, \cdot)$ is needed for the continuity of the limit process $\hat{X}_1$ (see (9) in Theorem 1) and the convergence of finite dimensional distributions of $\hat{X}_2^n$ (Lemma 5). When $H(\cdot, x)$ is continuous and monotone for each $x \in \mathbb{R}^k$, it is evident the assumption is satisfied. However, there are cases in which $G_1(t, \cdot)$ is continuous and $H(\cdot, x)$ is not; see, for example, the work-input process considered in Section 2.3.

For the sequence of noises $\{Z_i : i \in \mathbb{N}\}$, we impose the following ρ -mixing assumption. The notion of a ρ -mixing condition was first introduced in Kolmogorov and Rozanov (1960). We refer the reader to Bradley (2005) for a thorough review on mixing conditions. Note that the ρ -mixing condition is weaker than uniform strong mixing (ϕ -mixing, Ibragimov 1959) but stronger than strong mixing (α -mixing, Rosenblatt 1956).

Assumption 2. The sequence of random vectors $\{Z_i : i \in \mathbb{N}\}$ is weakly dependent and stationary with a continuous cumulative distribution function (c.d.f.) $F(x)$ for $x \in \mathbb{R}^k$. The sequence $\{Z_i : i \in \mathbb{N}\}$ satisfies that $C_\rho := \sum_{k=1}^{\infty} \rho_k < \infty$, where

$$\rho_k := \sup \left\{ \frac{|E[\xi\zeta] - E[\xi]E[\zeta]|}{\|\xi\|_2 \|\zeta\|_2} : \xi \in \mathcal{F}_m, \zeta \in \mathcal{G}_{m+k}, m \geq 1 \right\},$$

with $\mathcal{F}_k := \sigma\{Z_i : 1 \leq i \leq k\}$, $\mathcal{G}_k := \sigma\{Z_i : i \geq k\}$, and $\|\xi\|_2 := (E[\xi^2])^{1/2}$. Furthermore, for any $T \geq 0$, $\|H(T, Z_1)\|_{8+\delta} < \infty$ for some $\delta \in (0, \infty]$. Assume that $\sum_{k=1}^{\infty} k^2 \rho_k^{\gamma/(4+\gamma)} < \infty$ for some $0 < \gamma < 2$.

Denote F^c as the complement of c.d.f. F , that is, $F^c(x) = 1 - F(x)$ for $x \in \mathbb{R}^k$.

We provide an example in which the condition $C_\rho < \infty$ in Assumption 2 holds. Suppose that $\{Z_i : i \in \mathbb{N}\}$ is a strictly stationary Markov chain with a finite state space. If $\{Z_i : i \in \mathbb{N}\}$ is irreducible and aperiodic, then C_ρ is indeed finite. In particular, ρ_k decays exponentially fast as $k \rightarrow \infty$; see page 201 in Billingsley (1999). See also the example in Remark 3.

We consider a sequence of shot noise processes indexed by n and let $n \rightarrow \infty$. We write A^n and X^n and the associated $\{\tau_i^n\}$ for the n th process, and the sequence $\{Z_i : i \in \mathbb{N}\}$ and the c.d.f. F as well as the shot shape function H are fixed (independent of n). We assume that the arrival processes are independent of the noises. We make the following assumptions on arrival processes A^n .

Assumption 3. The sequence of arrival processes A^n with $A^n(0) = 0$ satisfies an FCLT:

$$\hat{A}^n := \sqrt{n}(\bar{A}^n - \Lambda) \Rightarrow \hat{A} \text{ in } (\mathbb{D}, J_1) \text{ as } n \rightarrow \infty, \quad (4)$$

where $\bar{A}^n := n^{-1}A^n$, $\Lambda := \{\Lambda(t) : t \geq 0\}$ is a deterministic, strictly increasing and continuous function with $\Lambda(0) = 0$, and $\hat{A}$ is a continuous process with mean zero and $\hat{A}(0) = 0$.

Note that Assumption 3 implies an FWLLN for the fluid-scaled arrival process $\bar{A}^n$: $\bar{A}^n \Rightarrow \Lambda$ in $(\mathbb{D}, J_1)$ as $n \rightarrow \infty$. The assumption of strictly increasing cumulative arrival rate function $\Lambda(t)$ in the limit means that the arrival rate $\lambda(t)$ (if it exists) is strictly positive. Common examples of the limit process $\hat{A}$ include a (time-changed) Brownian motion and a Gaussian process. In the former case, a typical limit process is $\hat{A}(t) = \mathcal{B}(c_a^2 \Lambda(t))$, where c_a^2 captures the variability and $\mathcal{B}$ is a standard Brownian motion. The condition (4) is satisfied by many models of arrival processes, including renewal process, stationary point processes with weakly dependent interarrival times, Markov-modulated Poisson processes, Hawkes processes, and so forth; we refer the reader to the discussion following assumption 1 in Pang and Zhou (2018a). In the case of renewal processes, c_a^2 represents the squared coefficient of variation (SCV) of the interarrival times. If the interarrival times form a stationary

ergodic sequence with a finite second moment satisfying certain mixing conditions, then the partial sum process satisfies an FCLT with a Brownian motion limit and its variance coefficient captures the correlation; see, for example, section 19 in Billingsley (1999). Then the FCLT for the arrival counting process follows directly by applying the continuous mapping theorem for the centered inverse mapping in chapter 13.7 in Whitt (2002).

Define the process $\bar{X}^n := \{\bar{X}^n(t) : t \geq 0\}$ by $\bar{X}^n(t) := n^{-1}X^n(t)$ for $t \geq 0$, and we have the following FWLLN for $\bar{X}^n$. We observe that the dependence in the noises does not affect the fluid limit, which is the same as in the case of i.i.d. noises.

Theorem 1 (FWLLN). *Under Assumptions 1–3,*

$$\bar{X}^n(t) \Rightarrow \bar{X}(t) \quad \text{in } (\mathbb{D}, J_1) \quad \text{as } n \rightarrow \infty, \quad (5)$$

where $\bar{X} := \{\bar{X}(t) : t \geq 0\}$ is a continuous deterministic function defined by

$$\bar{X}(t) := \int_{[0,t]} G_1(t, u) d\Lambda(u), \quad t \geq 0, \quad (6)$$

with $G_1(t, u)$ in (2).

Define the process $\hat{X}^n := \{\hat{X}^n(t) : t \geq 0\}$ by

$$\hat{X}^n(t) := \sqrt{n}(\bar{X}^n(t) - \bar{X}(t)), \quad t \geq 0, \quad (7)$$

where $\bar{X}(t)$ is given in (6).

Theorem 2 (FCLT). *Under Assumptions 1–3,*

$$\hat{X}^n \Rightarrow \hat{X} \quad \text{in } (\mathbb{D}, J_1) \quad \text{as } n \rightarrow \infty, \quad (8)$$

where $\hat{X} := \{\hat{X}(t) : t \geq 0\}$ can be written as a sum of two independent continuous stochastic processes $\hat{X}_1 := \{\hat{X}_1(t) : t \geq 0\}$ and $\hat{X}_2 := \{\hat{X}_2(t) : t \geq 0\}$. $\hat{X}_1$ is defined by

$$\hat{X}_1(t) := \hat{A}(t)G_1(t, t) - \int_{(0,t]} \hat{A}(u) dG_1(t, u), \quad t \geq 0. \quad (9)$$

$\hat{X}_2$ is a Gaussian process of mean zero and covariance function

$$\hat{R}_2(t, s) := \text{Cov}(\hat{X}_2(t), \hat{X}_2(s)) = \int_{[0, t \wedge s]} (G_2(t, s, u) - G_1(t, u)G_1(s, u) + \Gamma(t, s, u)) d\Lambda(u), \quad (10)$$

where

$$G_2(t, s, u) := \int_{\mathbb{R}^k} H(t - u, x)H(s - u, x) dF(x), \quad t \geq u \geq 0, \quad s \geq u \geq 0, \quad (11)$$

and

$$\begin{aligned} \Gamma(t, s, u) := & \sum_{l=2}^{\infty} (E[H(t - u, Z_1)H(s - u, Z_l)] - G_1(t, u)G_1(s, u) \\ & + E[H(s - u, Z_1)H(t - u, Z_l)] - G_1(s, u)G_1(t, u)). \end{aligned} \quad (12)$$

Remark 2. We make the following remarks.

- i. When the arrival limit $\hat{A}$ is Gaussian, $\hat{X}_1$ is also Gaussian; see remark 2.3 in Peng and Zhou (2018a).
- ii. The series in (12) is summable under Assumption 2. In particular, for each summand in (12), it is upper bounded by $2E[(|H(t - u, Z_1)| \vee |H(s - u, Z_1)|)^2] \rho_k$ so that the summation is bounded by $2E[(|H(t - u, Z_1)| \vee |H(s - u, Z_1)|)^2] C_\rho < \infty$.
- iii. Under Assumptions 1 and 2, it is easy to verify that $\Gamma(t, s, u)$ and $\hat{R}_2(s, t)$ are continuous in both s and t .
- iv. The limit process $\hat{X}_1$ is the same as that in the case of i.i.d. noises. The limit process $\hat{X}_2$ is different, and its covariance has the extra term $\Gamma(t, s, u)$ capturing the dependence in the noises. When the noises are i.i.d., the term $\Gamma(t, s, u)$ vanishes. It is well known in the limit theorems (CLT or FCLT) of partial sums of

weakly dependent variables that an extra term appears in the covariance of the limit capturing the correlations, separating from that in the i.i.d. case (see, e.g., chapter 4 in Billingsley 1999, Dehling and Philipp 2002, Bradley 2005). Our result indicates that the same separating property in the covariance functions of $\hat{X}_2$ also holds despite the effect of the impose response function H .

v. As in the case of i.i.d. noises (Peng and Zhou 2018a), the limit processes $\hat{X}_1$ and $\hat{X}_2$ are independent. The weak dependence among the noises does not affect the independence of the two processes. In addition, the variability in the arrival process, as captured in the limit $\hat{A}$, appears only in $\hat{X}_1$, and the fluid limit $\Lambda(t)$ of the arrival process appears in the covariance function of $\hat{X}_2$. The variability in the noises affects only $\hat{X}_2$. This can be regarded as a variability separation property of the arrival processes and the noises. This property has been observed in the study of infinite-server queues (Krichagina and Puhalskii 1997, Pang and Whitt 2010, Pang and Zhou 2017).

Example 1. We give an example in which $\hat{R}_2(t, s)$ can be further simplified. Let $\{Z_i : i \in \mathbb{N}\}$ be a first-order discrete autoregressive process, referred to as DAR(1) process (see Jacobs and Lewis 1978). Let Z_1 be distributed with c.d.f. F , and for $l \geq 2$,

$$Z_l = \delta_{l-1}Z_{l-1} + (1 - \delta_{l-1})S_l,$$

where $\{\delta_i : i \in \mathbb{N}\}$ is a sequence of i.i.d. Bernoulli random variables with $P(\delta_i = 1) = \alpha = 1 - P(\delta_i = 0)$ for $\alpha \in (0, 1)$, and $\{S_i : i \geq 2\}$, independent of $\{\delta_i : i \in \mathbb{N}\}$ and Z_1 , is a sequence of i.i.d. random variables with c.d.f. F . Then, it is easy to verify that $\{Z_i : i \in \mathbb{N}\}$ is stationary and satisfies the ρ -mixing condition (also ϕ and α -mixing conditions). In this case, the covariance function $\hat{R}_2(t, s)$ reduces to

$$\hat{R}_2(t, s) = \frac{1 + \alpha}{1 - \alpha} \int_{[0, t \wedge s]} (G_2(t, s, u) - G_1(t, u)G_1(s, u))d\Lambda(u), \quad (13)$$

for $t, s \geq 0$. In the case of i.i.d. noises, we have $\alpha = 0$, and thus, the effect of dependence among noises is evident with the covariance function $\hat{R}_2(t, s)$ increasing nonlinearly in the simple form of the multiplying coefficient $\frac{1+\alpha}{1-\alpha}$.

Example 2. Consider the following model with a batch arrival process $A^n(t) = \sum_{i=1}^{A_B^n(t)} B_i$, $A_B^n(t)$ being the arrival process of batches and B_i being the size of the i th batch. Assume that the noises are independent of the batch arrival process. For simplicity, let the noises be real valued. Suppose that the noises within each batch are symmetrically correlated; that is, each pair of noises (Z_k, Z_l) within a batch has the same joint distribution function $\Psi(x, y) = P(Z_k \leq x, Z_l \leq y)$ for arbitrary pair (k, l) in the batch and $x, y \in \mathbb{R}$, and in addition, $\Psi(x, \infty) = \Psi(\infty, y) = F(x)$. (Note that an infinite-server queueing model with batch arrivals and with such symmetrically correlated service times within batches is studied in Pang and Whitt (2012).)

Assume that $A_B^n(t)$ satisfies an FCLT as in Assumption 3, that is, $\hat{A}_B^n := \sqrt{n}(\bar{A}_B^n - \Lambda_B) \Rightarrow \hat{A}_B$ in $(\mathbb{D}, J_1)$ as $n \rightarrow \infty$, where $\bar{A}_B^n := n^{-1}A_B^n$ and $\Lambda_B := \{\Lambda_B(t) : t \geq 0\}$ is the cumulative arrival rate of the batches (deterministic, strictly increasing and continuous with $\Lambda_B(0) = 0$), and the limit process $\hat{A}_B(t) = \mathcal{B}(c_{a,B}^2 \Lambda_B(t))$ is a time-changed Brownian motion with a standard Brownian motion $\mathcal{B}(t)$ and coefficient of variability $c_{a,B}^2$. (Note that, as discussed after Assumption 3, this process A_B^n itself can be a renewal process, stationary point process with weakly dependent interarrival times with $c_{a,B}^2$ capturing the variabilities and correlations, and so on.) Suppose that the batch size variables $\{B_i : i \geq 1\}$ are independent of the arrival process A_B^n and i.i.d. positive integer valued. Let each B_i have a probability mass function $\{p_k : k \geq 1\}$ with finite mean m_B and variance σ_B^2 and SCV $c_B^2 = \sigma_B^2/m_B^2$. Then it can be shown that $\hat{A}^n = \sqrt{n}(\bar{A}^n - \Lambda) \Rightarrow \hat{A}$ in $(\mathbb{D}, J_1)$ as $n \rightarrow \infty$, where $\Lambda(t) := m_B \Lambda_B(t)$ and $\hat{A}(t) = \mathcal{B}(c_a^2 \Lambda(t))$ with $c_a^2 := m_B(c_{a,B}^2 + c_B^2)$. It is easy to check that the sequence of noises satisfies the ρ -mixing condition (also, the ϕ and α mixing conditions).

Then, by Theorem 2, we obtain the weak convergence of the diffusion-scaled processes. The limit process $\hat{X}_1(t)$ is given by the Itô integral

$$\hat{X}_1(t) = \int_0^t G_1(t, u) d\mathcal{B}(c_a^2 \Lambda(u)), \quad t \geq 0,$$

and the limit process $\hat{X}_2(t)$ is a continuous Gaussian process with mean zero and covariance function

$$\hat{R}_2(t, s) = m_B \int_{[0, t \wedge s]} ((G_2(t, s, u) - G_1(t, u)G_1(s, u)) + \Gamma(t, s, u))d\Lambda_B(u), \quad (14)$$

for $t, s \geq 0$, where $G_1(t, u)$ and $G_2(t, s, u)$ are given in (2) and (11), respectively, and

$$\Gamma(t, s, u) = 2(E[B^*] - 1) \left(\int_{\mathbb{R}} \int_{\mathbb{R}} H(t - u, x) H(s - u, y) d\Psi(x, y) - G_1(t, u) G_1(s, u) \right),$$

with

$$E[B^*] = \frac{m_B(c_B^2 + 1) + 1}{2}.$$

Note that B^* has the stationary excess distribution of the bath size B_i , that is, $p_k^* = P(B^* = k) = \frac{1}{m_B} \sum_{j=k}^{\infty} p_j$ for $k \geq 1$ (see, e.g., Whitt 1983).

Observe that the impact of dependence among noises is captured in the double integral term in $\Gamma(t, s, u)$. Recall that any positively correlated bivariate c.d.f. can be approximated by a linear combination of the two extremals given the marginals (Whitt 1976), that is, $\Psi(x, y) \approx \tilde{\Psi}_\rho(x, y)(x, y) = \rho F(x \wedge y) + (1 - \rho)F(x)F(y)$, where $\rho \in (0, 1)$ is the correlation parameter. Using this, we obtain an approximation of $\Gamma(t, s, u)$:

$$\tilde{\Gamma}(t, s, u) = 2\rho(E[B^*] - 1) \left(\int_{\mathbb{R}} \int_{\mathbb{R}} H(t - u, x) H(s - u, y) dF(x \wedge y) - G_1(t, u) G_1(s, u) \right).$$

This implies that the covariance function is approximately increasing linearly in the correlation parameter ρ .

In Ramirez-Perez and Serfling (2001), a shot noise process with cluster marks is studied, in which cluster arrivals are Poisson and each entity of each cluster may arrive with certain random nonnegative “displacement” (delay). Our model with batch arrivals allows more general arrival processes of clusters, but all entities in a cluster arrive at the same time. The marks (noises) are mutually independent across clusters, but the noises within each cluster have a particular decomposable structure with two independent components, a common random variable for the cluster, and another i.i.d. random variable. If, in addition, we assume that the marks (noises) are an additive function of the two components as studied in section 4 of Ramirez-Perez and Serfling (2001), then we can use the preceding covariance formula for this particular shot noise process with cluster marks. In particular, suppose that in each batch, $Z_i = \tilde{Z}_0 + \tilde{Z}_i$ for a common random variable $\tilde{Z}_0$ with c.d.f., $\tilde{F}_0$, and i.i.d. random variables $\tilde{Z}_i$ with c.d.f. $\tilde{F}$, independent of $\tilde{Z}_0$. Then the marks (noises) within each cluster are symmetrically correlated; $\text{Corr}(Z_k, Z_l) = \frac{\text{Var}(\tilde{Z}_0)}{\text{Var}(\tilde{Z}_0) + \text{Var}(\tilde{Z}_1)}$ for each pair (k, l) in the cluster (batch). And we have

$$\Psi(x, y) = P(Z_k \leq x, Z_l \leq y) = P(\tilde{Z}_0 + \tilde{Z}_k \leq x, \tilde{Z}_0 + \tilde{Z}_l \leq y) = \int_{\mathbb{R}} \tilde{F}(x - z) \tilde{F}(y - z) d\tilde{F}_0(z),$$

and

$$F(x) = P(\tilde{Z}_0 + \tilde{Z}_i \leq x) = \int_{\mathbb{R}} \tilde{F}(x - z) d\tilde{F}_0(z).$$

They can be easily used to calculate the covariance function $\hat{R}_2(t, s)$ in (14). In future work, it would be interesting to establish the functional limit theorems for the shot noise process with general cluster marks in which entities of each cluster may arrive with random “displacement” and study the impact of such displacement.

2.1. Multiplicative Models

The multiplicative models have the shot shape function H of the form

$$H(t, x) := \tilde{H}(t) \varphi(x), \quad t \geq 0, \quad x \in \mathbb{R}^k,$$

where $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}$ is a measurable function. Assumption 1 requires that the function $\tilde{H}$ is monotone and continuous because

$$G_1(t, u) = E[H(t - u, Z_1)] = \tilde{H}(t - u) E[\varphi(Z_1)] \quad \text{for } 0 \leq u \leq t.$$

In addition, it requires that $m_\varphi := E[\varphi(Z_1)]$ and $\sigma_\varphi^2 := \text{Var}(\varphi(Z_1))$ are finite.

The limit process $\hat{X}_1$ in Theorem 2 becomes

$$\hat{X}_1(t) = m_\varphi \tilde{H}(0) \hat{A}(t) - m_\varphi \int_{(0, t]} \hat{A}(u) d\tilde{H}(t - u), \quad t \geq 0.$$

If $\hat{A}$ is a Gaussian process with mean zero and covariance function $\hat{R}_a(t, s)$, then the covariance function of $\hat{X}_1$ is

$$\hat{R}_1(t, s) = m_\varphi^2 \int_{[0, t]} \int_{[0, s]} \tilde{H}(t - u) \tilde{H}(s - v) d\hat{R}_a(u, v), \quad t, s \geq 0.$$

If $\hat{A}(t) = c_a B(\Lambda(t))$ is a time-changed Brownian motion in which the constant $c_a > 0$ can be regarded as a coefficient of variation and $\Lambda(t)$ is in Assumption 3, then the covariance function $\hat{R}_1(t, s)$ becomes

$$\hat{R}_1(t, s) = (c_a m_\varphi)^2 \int_{[0, t \wedge s]} \tilde{H}(t-u) \tilde{H}(s-u) d\Lambda(u), \quad t, s \geq 0.$$

The limit process $\hat{X}_2$ in Theorem 2 has the covariance function

$$\hat{R}_2(t, s) = \int_{[0, t \wedge s]} \left(\sigma_\varphi^2 \tilde{H}(t-u) \tilde{H}(s-u) + \Gamma(t, s, u) \right) d\Lambda(u), \quad t, s \geq 0,$$

where

$$\Gamma(t, s, u) = 2\tilde{H}(t-u)\tilde{H}(s-u) \sum_{l=2}^{\infty} \left[E[\varphi(Z_1)\varphi(Z_l)] - m_\varphi^2 \right].$$

Example 3. We provide some examples of the shot shape functions. These are often used in the study of financial modeling and insurance risks (Chobanov 1999, Ganesh and Torrisi 2006, Schmidt 2014).

i. Exponential decay: The function $H(t, x)$ takes the form

$$H(t, x) = xe^{-bt} \quad \text{for } b > 0.$$

Here $\tilde{H}(t) = e^{-bt}$, $\varphi(x) = x$, and $G_1(t, u) = E[e^{-b(t-u)}Z_1] = e^{-b(t-u)}E[Z_1]$ for $0 \leq u \leq t$.

ii. Power-law decay:

$$H(t, x) = \frac{x}{1+ct} \quad \text{for } c > 0.$$

Here $\tilde{H}(t) = 1/(1+ct)$, $\varphi(x) = x$, and $G_1(t, u) = E[Z_1]/(1+c(t-u))$ for $0 \leq u \leq t$. The effect decay is slower than for the exponential case, and the effect of the shot stays longer in the data.

iii. Polynomial decay:

$$H(t, x) = xt^{-\alpha}, \quad \text{for } \alpha > 0.$$

We have $\tilde{H}(t) = t^{-\alpha}$, $\varphi(x) = x$, and $G_1(t, u) = (t-u)^{-\alpha}E[Z_1]$ for $0 \leq u < t$.

It is evident that the conditions in Assumption 1 are satisfied for these functions H .

Example 4. With the DAR(1) noises $\{Z_i\}$ in Example 1, we obtain that

$$\hat{R}_2(t, s) = \frac{1+\alpha}{1-\alpha} \sigma_\varphi^2 \int_{[0, t \wedge s]} \tilde{H}(t-u) \tilde{H}(s-u) d\Lambda(u), \quad t, s \geq 0.$$

The impact of dependence among noises is captured in the coefficient $\frac{1+\alpha}{1-\alpha}$, increasing in the parameter α .

2.2. Compound Process Under ρ -Mixing Condition

Consider the following special case of the multiplicative model:

$$X(t) = \sum_{i=1}^{A(t)} Z_i, \quad t \geq 0,$$

where $\{Z_i : i \in \mathbb{N}\}$ are real-valued random variables. Here, we have $\tilde{H}(t) \equiv 1$, $t \in \mathbb{R}_+$, and $\varphi(x) = x$ for each $x \in \mathbb{R}_+$. Note that, in this case, the function $\Gamma(t, s, u)$ becomes

$$\Gamma(t, s, u) = 2 \sum_{l=2}^{\infty} \left(E[Z_1 Z_l] - (E[Z_1])^2 \right).$$

As a consequence of Theorem 2, we obtain the following theorem for the compound processes under the ρ -mixing condition.

Theorem 3. Under Assumptions 2 and 3, (5) in the FWLLN holds with the limit $\bar{X}$ given by

$$\bar{X}(t) := E[Z_1]\Lambda(t), \quad t \geq 0,$$

and (8) in the FCLT holds with the limit $\hat{X} = \hat{X}_1 + \hat{X}_2$, where $\hat{X}_1$ and $\hat{X}_2$ are independent, $\hat{X}_1 = E[Z_1]\hat{A}$, and $\hat{X}_2$ is a continuous Gaussian process with mean zero and covariance function: for $t, s \geq 0$,

$$\hat{R}_2(t, s) := \text{Cov}(\hat{X}_2(t), \hat{X}_2(s)) = \Lambda(t \wedge s) \left[\text{Var}(Z_1) + 2 \sum_{l=2}^{\infty} (E[Z_1 Z_l] - (E[Z_1])^2) \right].$$

Remark 3. With the DAR(1) noises $\{Z_i\}$ in Example 1, we obtain that the covariance function $\hat{R}_2(t, s)$ reduces to

$$\hat{R}_2(t, s) = \frac{1+\alpha}{1-\alpha} \Lambda(t \wedge s) \text{Var}[Z_1], \quad t, s \geq 0.$$

Observe that the process $\hat{X}_2$ becomes a Brownian motion with time change using $\Lambda(t)$, and the term $\frac{1+\alpha}{1-\alpha}$ captures the correlation among the noises. In the i.i.d. case ($\alpha = 0$), the covariance function $\hat{R}_2(t, s) = \Lambda(t \wedge s) \text{Var}[Z_1]$ for $t, s \geq 0$. It is evident that as α increases, the correlation increases the covariance function nonlinearly in the simple form $\frac{1+\alpha}{1-\alpha}$.

2.3. The Work-Input Process in Infinite-Server Queues

Consider an infinite-server queue with an arrival process A^n as in Assumption 3 and service times $\{Z_i : i \in \mathbb{N}\}$ as in Assumption 2. This model has been studied in Pang and Zhou (2018b). The total queue length process $Q^n := \{Q^n(t) : t \geq 0\}$ can be written as

$$Q^n(t) = \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + Z_i > t), \quad t \geq 0.$$

It is a shot noise process with $H(t, x) = \mathbf{1}(x > t)$ and $G_1(t, u) = E[\mathbf{1}(Z_1 > t - u)] = F^c(t - u)$ for $0 \leq u \leq t$. The conditions in Assumption 1 are satisfied because F is continuous. The FWLLN and FCLT for the queueing process Q^n are obtained as a consequence of the two-parameter process limits in Pang and Zhou (2018b), and also as a special case of the results in this paper.

Here we focus on the work-input process $W^n := \{W^n(t) : t \geq 0\}$ defined by

$$W^n(t) = \sum_{i=1}^{A^n(t)} Z_i \mathbf{1}(\tau_i^n + Z_i > t), \quad t \geq 0.$$

We refer the reader to Kurtz (1996) for the FCLT with A^n being a Poisson process under the conventional scaling regime, in which the limit process is a fractional Brownian motion. The process W^n is a shot noise process with $H(t, x) = x \mathbf{1}(x > t)$. For each $x \in \mathbb{R}$, $H(t, x) = x \mathbf{1}(x > t)$ is monotone. The function G_1 becomes

$$G_1(t, u) = E[Z_1 \mathbf{1}(Z_1 > t - u)] = \int_{(t-u, \infty)} x dF(x) \quad \text{for } 0 \leq u \leq t. \quad (15)$$

The conditions in Assumption 1 are satisfied because F is continuous.

Let the fluid-scaled processes $\bar{W}^n := \{\bar{W}^n(t) : t \geq 0\}$ be defined by $\bar{W}^n(t) := n^{-1} W^n(t)$ for $t \geq 0$. Define the deterministic functions

$$\bar{W}(t) := \int_{[0, t]} \int_{(t-u, \infty)} x dF(x) d\Lambda(u), \quad t \geq 0. \quad (16)$$

It is easy to check that $\bar{W}(t)$ is a continuous function. Let the diffusion-scaled processes $\hat{W}^n := \{\hat{W}^n(t) : t \geq 0\}$ be defined by $\hat{W}^n(t) := \sqrt{n}(\bar{W}^n(t) - \bar{W}(t))$ for $t \geq 0$. By Theorems 1 and 2, we obtain the following theorem for the work-input processes.

Theorem 4. Under Assumptions 2 and 3,

$$\bar{W}^n \Rightarrow \bar{W} \quad \text{in } (\mathbb{D}, J_1) \quad \text{as } n \rightarrow \infty,$$

where $\bar{W}$ is defined in (16), and

$$\hat{W}^n \Rightarrow \hat{W} \quad \text{in } (\mathbb{D}, J_1) \quad \text{as } n \rightarrow \infty,$$

where the limit $\hat{W} = \{\hat{W}(t) : t \geq 0\}$ is continuous and can be written as $\hat{W} = \hat{W}_1 + \hat{W}_2$ with $\hat{W}_1$ and $\hat{W}_2$ being independent. $\hat{W}_1$ is defined by

$$\hat{W}_1(t) = \hat{A}(t)G_1(t, t) - \int_{(0, t]} \hat{A}(u)dG_1(t, u), \quad t \geq 0,$$

with $G_1(t, u)$ in (15), and $\hat{W}_2$ is a continuous Gaussian process with mean zero and covariance function

$$\text{Cov}(\hat{W}_2(t), \hat{W}_2(s)) = \int_{[0, t \wedge s]} (G_2(t, s, u) - G_1(t, u)G_1(s, u) + \Gamma_W(t, s, u))d\Lambda(u),$$

where, by denoting $\tilde{Z}_l(x) = Z_l \mathbf{1}(Z_l > x)$ for any $l \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$\begin{aligned} \Gamma_W(t, s, u) = & \sum_{l=2}^{\infty} \left([E[\tilde{Z}_1(t-u)\tilde{Z}_l(s-u)] - E[\tilde{Z}_1(t-u)]E[\tilde{Z}_l(s-u)]] \right. \\ & \left. + [E[\tilde{Z}_1(s-u)\tilde{Z}_l(t-u)] - E[\tilde{Z}_1(s-u)]E[\tilde{Z}_l(t-u)]] \right), \end{aligned}$$

for $t, s \geq 0$.

Example 5. Suppose that the service times $\{Z_i\}$ are DAR(1) as in Example 1. We have the covariance function

$$\text{Cov}(\hat{W}_2(t), \hat{W}_2(s)) = \frac{1+\alpha}{1-\alpha} \int_{[0, t \wedge s]} (G_2(t, s, u) - G_1(t, u)G_1(s, u))d\Lambda(u),$$

for $t, s \geq 0$, where

$$G_2(t, s, u) = E[Z_1^2 \mathbf{1}(Z_1 > t \vee s - u)] = \int_{(t \vee s - u, \infty)} x^2 dF(x),$$

and $G_1(t, u)$ is given in (15). The impact of dependence among service times is captured by the multiplicative coefficient $\frac{1+\alpha}{1-\alpha}$.

3. Proof of Theorem 2

Because Theorem 1 is directly implied by Theorem 2, we focus on the proof of Theorem 2 in this section. Note that, in the proofs, we assume that, for each $x \in \mathbb{R}^k$, $H(t, x)$ is nondecreasing in t and nonnegative for brevity, and it can be easily verified that the proofs extend to the general monotone assumption.

Recall $G_1(t, u)$ in (2). We first give a representation for the process $\hat{X}^n$, which follows from simple calculations.

Lemma 1. The process $\hat{X}^n$ defined in (7) can be written as $\hat{X}^n = \hat{X}_1^n + \hat{X}_2^n$, where the processes $\hat{X}_1^n$ and $\hat{X}_2^n$ are given by

$$\begin{aligned} \hat{X}_1^n(t) &:= \int_{(0, t]} G_1(t, u)d\hat{A}^n(u) \\ &= \hat{A}^n(t)G_1(t, t) - \int_{(0, t]} \hat{A}^n(u-)dG_1(t, u), \quad t \geq 0, \end{aligned}$$

where $\hat{A}^n(u-)$ denotes the left limit of $\hat{A}^n$ at time u , and

$$\hat{X}_2^n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} \left(H(t - \tau_i^n, Z_i) - G_1(t, \tau_i^n) \right), \quad t \geq 0.$$

For the convergence of $\hat{X}_1^n$, by lemma 6.1 in Pang and Zhou (2018a) (with the distribution function $F_t(\cdot)$ replaced with $F(\cdot)$), the proof becomes a direct application of continuous mapping theorem. We omit the proof for brevity.

Theorem 5. Under Assumptions 2 and 3, $\hat{X}_1^n \Rightarrow \hat{X}_1$ in $(\mathbb{D}, J_1)$ as $n \rightarrow \infty$, where $\hat{X}_1$ is as given in Theorem 2.

We prove the convergence of the processes $\hat{X}_2^n$ in three steps:

Step 1: The existence of the limit Gaussian process $\hat{X}_2$ in $\mathbb{C}$ (Section 3.1, Theorem 7).

Step 2: The convergence of finite dimensional distributions of $\hat{X}_2^n$ to those of $\hat{X}_2$ (Section 3.2, Lemma 5).

Step 3: The tightness of $\hat{X}_2^n$ in the space $\mathbb{D}$ (Section 3.3, Lemma 11).

We summarize these results in the following theorem.

Theorem 6. Under Assumptions 2 and 3, $\hat{X}_2^n \Rightarrow \hat{X}_2$ in $(\mathbb{D}, J_1)$ as $n \rightarrow \infty$, where $\hat{X}_2$ is as given in Theorem 2.

3.1. Existence of the Limit Process $\hat{X}_2$ in $\mathbb{C}$

To prove the existence theorem, we first provide some properties on the increments of the limiting process $\hat{X}_2$.

Lemma 2. For each $0 \leq s \leq t$,

$$\begin{aligned} E[|\hat{X}_2(s) - \hat{X}_2(t)|^2] &= \int_{(s,t]} (\tilde{G}(t, u) + \Gamma(t, t, u)) d\Lambda(u) \\ &\quad + \int_{[0,s]} (\tilde{G}(t, s, u) + \Gamma(t, t, u) + \Gamma(s, s, u) - 2\Gamma(t, s, u)) d\Lambda(u). \end{aligned} \quad (17)$$

The proof of (17) follows by (10) and direct calculations.

Definition 1. Fix $T > 0$. For any $0 \leq s \leq t \leq T$, define a nonnegative function $V : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ by

$$\begin{aligned} V(s, t) &:= \tilde{C}(\Lambda(t) - \Lambda(s)) \\ &\quad + (1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} (H(t - u, x) - H(s - u, x))^2 dF(x) d\Lambda(u). \end{aligned} \quad (18)$$

where

$$\tilde{C} := \sup_{0 \leq t, u \leq T} [\tilde{G}(t, u) + \Gamma(t, t, u)] < \infty.$$

We need the following definition of a set function with a superadditivity property.

Definition 2. Let μ be a set function from the Borel subset of $\mathbb{R}_+$ into $\mathbb{R}_+ \cup \{\infty\}$ such that

- μ is nonnegative and $\mu(\emptyset) = 0$.
- μ is monotone; that is, if $A \subseteq B \subset \mathbb{R}_+$, then $\mu(A) \leq \mu(B)$.
- μ is superadditive; that is, for any disjoint Borel sets A and B , $\mu(A) + \mu(B) \leq \mu(A \cup B)$.

Proposition 1. The function V introduced in Definition 1 has the following properties:

- $V(t, t) = 0$ for each $t \in [0, T]$.
- $V(s, t)$ is nondecreasing in t for each s and nonincreasing in s for each t , and thus, it is evident that $V(s, t) \leq V(s, T) \leq V(0, T)$ for each $s, t \geq 0$.
- $V(s, t)$ is continuous in both s and t .
- For any Borel set $A \subset [0, T]$, let

$$v(A) := \sup\{V(s, t) : (s, t] \subset A\}.$$

Here v satisfies all the conditions in Definition 2 and $v((s, t]) = V(s, t)$ for $0 \leq s \leq t \leq T$.

Proof of Proposition 1. The properties (i) and (ii) are straightforward to check. For (iii), for any $\delta_1, \delta_2 \in \mathbb{R}$,

$$\begin{aligned}
 & V(s, t) - V(s - \delta_1, t - \delta_2) \\
 &= \tilde{C}(\Lambda(t) - \Lambda(t - \delta_2) - (\Lambda(s) - \Lambda(s - \delta_1))) \\
 &\quad + (1 + 2C_\rho) \int_{[0, T]} \int_{\mathbb{R}^k} (H(t - u, x) - H(s - u, x))^2 dF(x) d\Lambda(u) \\
 &\quad - (1 + 2C_\rho) \int_{[0, T]} \int_{\mathbb{R}^k} (H(t - \delta_2 - u, x) - H(s - \delta_1 - u, x))^2 dF(x) d\Lambda(u) \\
 &= \tilde{C}(\Lambda(t) - \Lambda(t - \delta_2) - (\Lambda(s) - \Lambda(s - \delta_1))) \\
 &\quad + (1 + 2C_\rho) \int_{[0, T]} \check{G}_2(t, t - \delta_2, u) d\Lambda(u) + (1 + 2C_\rho) \int_{[0, T]} \check{G}_2(s, s - \delta_1, u) d\Lambda(u) \\
 &\quad + 2(1 + 2C_\rho) \int_{[0, T]} \int_{\mathbb{R}^k} [H(t - u, x) - H(t - \delta_2 - u, x)][H(t - \delta_2 - u, x) - H(s - u, x)] dF(x) d\Lambda(u) \\
 &\quad - 2(1 + 2C_\rho) \int_{[0, T]} \int_{\mathbb{R}^k} [H(s - u, x) - H(s - \delta_1 - u, x)][H(t - u, x) - H(s - \delta_1 - u, x)] dF(x) d\Lambda(u).
 \end{aligned} \tag{19}$$

Note that the last two terms in (19) are bounded by

$$2(1 + 2C_\rho) \left(\int_{[0, T]} \check{G}_2(t, t - \delta_2, u) d\Lambda(u) \right)^{1/2} \left(\int_{[0, T]} \check{G}_2(t - \delta_2, s, u) d\Lambda(u) \right)^{1/2},$$

and

$$2(1 + 2C_\rho) \left(\int_{[0, T]} \check{G}_2(s, s - \delta_1, u) d\Lambda(u) \right)^{1/2} \left(\int_{[0, T]} \check{G}_2(t, s - \delta_1, u) d\Lambda(u) \right)^{1/2},$$

respectively, because of the Cauchy–Schwarz inequality. By the continuity of Λ and (3) in Assumption 1 that, for each $T \geq t \geq 0$, $\lim_{\delta \rightarrow 0} \int_{[0, T]} \check{G}_2(t, t - \delta, u) d\Lambda(u) = 0$, we show that each term in (19) converges to zero as $\delta_1 \rightarrow 0$ and $\delta_2 \rightarrow 0$. Now the continuity of V has been proved.

For (iv), the superadditivity property follows from the inequality $\sum_{i=1}^n x_i^2 \leq (\sum_{i=1}^n x_i)^2$ for each $n \geq 1$ if all elements of $\{x_i : 1 \leq i \leq n\}$ have the same sign.

Lemma 3. For $0 \leq s \leq t \leq T$,

$$E\left[\left|\hat{X}_2(s) - \hat{X}_2(t)\right|^2\right] \leq V(s, t).$$

Proof of Lemma 3. First, by direct calculations, we obtain

$$\begin{aligned}
 & \Gamma(t, t, u) + \Gamma(s, s, u) - 2\Gamma(t, s, u) \\
 &= 2 \sum_{l=2}^{\infty} E\left[[H(t - u, Z_1) - H(s - u, Z_1)] \times [H(t - u, Z_l) - H(s - u, Z_l)] - (G_1(t, u) - G_1(s, u))^2\right].
 \end{aligned} \tag{20}$$

Notice that, for each $l \geq 2$,

$$(G_1(t, u) - G_1(s, u))^2 = E[H(t - u, Z_1) - H(s - u, Z_1)]E[H(t - u, Z_l) - H(s - u, Z_l)].$$

By the definition of the ρ -mixing condition, each summand in (20) is upper bounded by

$$2\|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 \rho_k.$$

Therefore, we have

$$\Gamma(t, t, u) + \Gamma(s, s, u) - 2\Gamma(t, s, u) \leq 2\|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 C_\rho. \tag{21}$$

By Lemma 2 and (21), we have

$$\begin{aligned} & E\left[|\hat{X}_2(r) - \hat{X}_2(s)|^2\right] \\ & \leq \tilde{C}(\Lambda(t) - \Lambda(s)) \\ & \quad + \int_{[0,s]} (\|H(t-u, Z_1) - H(s-u, Z_1)\|_2^2 + 2C_\rho \|H(t-u, Z_1) - H(s-u, Z_1)\|_2^2) d\Lambda(u) \\ & \leq \tilde{C}(\Lambda(t) - \Lambda(s)) + (1 + 2C_\rho) \int_{[0,T]} \|H(t-u, Z_1) - H(s-u, Z_1)\|_2^2 d\Lambda(u), \end{aligned}$$

which is equal to $V(s, t)$ in (18).

We now state the probability bound for the increments of the limit process $\hat{X}_2$.

Lemma 4. For $0 \leq r \leq s \leq t \leq T$ and any $\epsilon > 0$,

$$P\left(|\hat{X}_2(r) - \hat{X}_2(s)| \wedge |\hat{X}_2(s) - \hat{X}_2(t)| \geq \epsilon\right) \leq \frac{3}{\epsilon^4} V(r, s) V(s, t).$$

Proof of Lemma 4. We have

$$\begin{aligned} & P(|\hat{X}_2(r) - \hat{X}_2(s)| \wedge |\hat{X}_2(s) - \hat{X}_2(t)| \geq \epsilon) \\ & \leq \frac{1}{\epsilon^4} E\left[|\hat{X}_2(r) - \hat{X}_2(s)|^2 |\hat{X}_2(s) - \hat{X}_2(t)|^2\right] \\ & \leq \frac{1}{\epsilon^4} \left(E\left[|\hat{X}_2(r) - \hat{X}_2(s)|^4\right]\right)^{1/2} \left(E\left[|\hat{X}_2(s) - \hat{X}_2(t)|^4\right]\right)^{1/2} \\ & = \frac{3}{\epsilon^4} E\left[|\hat{X}_2(r) - \hat{X}_2(s)|^2\right] E\left[|\hat{X}_2(s) - \hat{X}_2(t)|^2\right], \end{aligned}$$

where the equality follows from the fact that the kurtosis of a normal random variable is three. The claim then follows from Lemma 3.

We are now ready to prove the existence theorem. Note that the ρ -mixing condition is important here because we are unable to obtain a set function satisfying the superadditivity property if the ρ -mixing condition is relaxed; for example, under the strong α -mixing condition, the bounds in Lemmas 3 and 4 do not hold.

Theorem 7. The centered Gaussian process $\hat{X}_2$ with covariance function given in (10) has continuous sample paths a.s.

Proof of Theorem 7. The proof is done in two steps. First, we show that $\hat{X}_2 \in \mathbb{D}$ by applying Theorem A.1 in the appendix (a generalization of the classical existence criterion in theorem 13.6 in Billingsley 1999). In the new existence criterion, we need to verify three conditions: (i) Consistency of the finite dimensional distributions satisfying the conditions of Kolmogorov's existence theorem, which is easily satisfied by the Gaussian distributional property of the process $\hat{X}_2$. (ii) Condition (A.1): a probability inequality for the increment of the process, which requires the set function satisfying the superadditivity property. This is implied by the probability bound in Lemma 4. (iii) The continuity property: for any $\epsilon > 0$ and $t \in [0, T)$, $\lim_{\delta \downarrow 0} P(|\hat{X}_2(t) - \hat{X}_2(t + \delta)| > \epsilon) = 0$. This is implied by the following: for all $t \in [0, T)$,

$$\lim_{\delta \downarrow 0} E\left[|\hat{X}_2(t) - \hat{X}_2(t + \delta)|^2\right] = 0.$$

This follows by the formula in Lemma 2 and the continuity properties of $\Gamma(t, s, u)$ in both s and t and Assumption 1.

In the second step, we show the existence in $\mathbb{C}$ given the existence in $\mathbb{D}$, for which it suffices to show that the process $\hat{X}_2$ is stochastically continuous (theorem 1 in Hahn 1978) or, equivalently, that it is continuous in quadratic mean because the process is Gaussian. This follows from Lemma 2 and the continuity properties of Γ and Assumption 1. This completes the proof.

3.2. Convergence of Finite Dimensional Distributions

Before proceeding to the proof, for each $n \geq 1$, we define the set Υ^n to be the collection of the trajectories of $\{A^n(t) : t \geq 0\}$ as

$$\Upsilon^n = \left\{ A^n : \sup_{0 \leq t \leq T} \left| \frac{A^n(t)}{n} - \Lambda(t) \right| \leq \epsilon(n) \quad \text{and} \quad \max_{1 \leq i \leq A^n(T)} |\tau_{i+1}^n - \tau_i^n| \leq \epsilon(n) \right\}, \quad (22)$$

where $\epsilon(n) \rightarrow 0$ as $n \rightarrow \infty$ is chosen such that $P(\Upsilon^n) \rightarrow 1$ as $n \rightarrow \infty$. It is evident that, under Assumption 3, such a function $\epsilon(n)$ exists. Thus, we can fix a trajectory $A^n \in \Upsilon$ and regard it as deterministic without loss of generality.

Lemma 5. *The finite dimensional distributions of $\hat{X}_2^n$ converge weakly to those of $\hat{X}_2$ as $n \rightarrow \infty$.*

Proof of Lemma 5. Fix any $0 \leq t \leq T$ and trajectory $A^n \in \Upsilon^n$ (defined in (22)), and we first prove that $\hat{X}_2^n(t) \Rightarrow \hat{X}_2(t)$ in $\mathbb{R}$ as $n \rightarrow \infty$. To simplify the notation, we denote

$$\mathcal{H}_i^n(t) := H(t - \tau_i^n, Z_i) - G_1(t, \tau_i^n). \quad (23)$$

We apply theorem 2.1 and proposition 2.1(a) in Withers (1981) with $S_n = \sum_{j=1}^{A^n(t)} X_{jn}$, $X_{jn} = \mathcal{H}_j^n(t)$, and $S_n(a, b) = \sum_{j=a+1}^{a+b} X_{jn}$, $0 \leq a, 1 \leq b \leq n - a$ in their notation. To that end, we need to verify

- i. $\sup_{n \in \mathbb{N}, 1 \leq j \leq n} \|\mathcal{H}_j^n(t)\|_{2+\delta} < +\infty$, where $0 < \delta \leq \infty$.
- ii. $\sup_{a, b, n} E[S_n(a, b)^2]/b < \infty$.
- iii. $\rho_n = o(1)$.
- iv. $\text{Var}(S_n) \rightarrow \infty$ as $n \rightarrow \infty$ and $\sum_{\ell=0}^{\infty} \tilde{c}(\ell) < \infty$, where

$$\tilde{c}(\ell) = \max_{n: \ell < n} \tilde{c}_n(\ell),$$

and

$$\tilde{c}_n(\ell) = \sup_{p, m: |p-m| \geq \ell, 1 \leq p \leq n, 1 \leq m \leq n} |E[\mathcal{H}_p^n \mathcal{H}_m^n]|.$$

We remark that conditions (i) and (ii) are required by proposition 2.1 in Withers (1981) with $\epsilon = \gamma = 0$. The l -mixing condition in theorem 2.1 in Withers (1981) is implied by condition (iii) because ρ -mixing is stronger than l -mixing; see page 513 in Withers (1981). Conditions (iv) is additionally required by theorem 2.1 in the last-mentioned reference.

Conditions (i) and (iii) are directly implied by Assumption 2.

For condition (ii), we have

$$\begin{aligned} \frac{1}{b} E[S_n(a, b)^2] &= \frac{1}{b} E \left[\left(\sum_{i=a+1}^{a+b} \mathcal{H}_i^n(t) \right)^2 \right] \\ &= \frac{1}{b} \sum_{i=a+1}^{a+b} E[(\mathcal{H}_i^n(t))^2] + \frac{2}{b} \sum_{i, j=a+1, i < j}^{a+b} E[\mathcal{H}_i^n(t) \mathcal{H}_j^n(t)] \\ &\leq \frac{1}{b} \sum_{i=a+1}^{a+b} \|H(T, Z_1)\|_2^2 + 2 \|H(T, Z_1)\|_2^2 / b \sum_{i, j=a+1, i < j}^{a+b} \rho_{|j-i|} \\ &\leq (1 + 2C_\rho) \|H(T, Z_1)\|_2^2 < \infty. \end{aligned}$$

For condition (iv), we have

$$\begin{aligned} \frac{\text{Var}(S_n)}{n} &= \frac{1}{n} \text{Var} \left(\sum_{j=1}^{A^n(t)} \mathcal{H}_j^n(t) \right) \\ &= \frac{1}{n} \sum_{i=1}^{A^n(t)} E[(\mathcal{H}_i^n(t))^2] + \frac{2}{n} \sum_{i, j=1, i < j}^{A^n(t)} E[\mathcal{H}_i^n(t) \mathcal{H}_j^n(t)] \\ &= \frac{1}{n} \sum_{i=1}^{A^n(t)} E[(\mathcal{H}_i^n(t))^2] + \frac{2}{n} \sum_{i, j=1, i < j}^{A^n(t)} r_{ij}^n \end{aligned} \quad (24)$$

where

$$\begin{aligned} r_{ij}^n &:= E[\mathcal{H}_i^n(t)\mathcal{H}_j^n(t)] \\ &= E\left[H(t - \tau_i^n, Z_i)H(t - \tau_j^n, Z_j)\right] - G_1(t, \tau_i^n)G_1(t, \tau_j^n). \end{aligned}$$

By stationarity of $\{Z_i : i \in \mathbb{N}\}$, for fixed ℓ , we consider

$$\begin{aligned} R_\ell^n &:= \sum_{i=1}^{A^n(t)-\ell} r_{i,i+\ell}^n = \sum_{i=1}^{A^n(t)-\ell} \left(E[H(t - \tau_i^n, Z_i)H(t - \tau_{i+\ell}^n, Z_{i+\ell})] - G_1(t, \tau_i^n)^2 \right) \\ &\quad + \sum_{i=1}^{A^n(t)-\ell} \Delta_{i,\ell}^{n,(1)} + \sum_{i=1}^{A^n(t)-\ell} \Delta_{i,\ell}^{n,(2)}, \end{aligned}$$

where

$$\begin{aligned} \Delta_{i,\ell}^{n,(1)} &:= E[H(t - \tau_i^n, Z_i)[H(t - \tau_{i+\ell}^n, Z_{i+\ell}) - H(t - \tau_i^n, Z_{i+\ell})]], \\ \Delta_{i,\ell}^{n,(2)} &:= G_1(t, \tau_i^n)[G_1(t, \tau_i^n) - G_1(t, \tau_{i+\ell}^n)]. \end{aligned}$$

Observe that by the Cauchy–Schwarz inequality, we have

$$\begin{aligned} \Delta_{i,\ell}^{n,(1)} &\leq \left(E[H(t - \tau_i^n, Z_i)^2] E[(H(t - \tau_{i+\ell}^n, Z_{i+\ell}) - H(t - \tau_i^n, Z_{i+\ell}))^2] \right)^{1/2} \\ &= \left(G_2(t, \tau_i^n) \int_{\mathbb{R}^k} (H(t - \tau_{i+\ell}^n) - H(t - \tau_i^n))^2 dF(x) \right)^{1/2}. \end{aligned}$$

Now, given the trajectories of A^n in Υ^n and continuity of $G_1(t, \cdot)$ and the condition on $\check{G}_2(t, s, \cdot)$ in (3) (recall Assumption 1), for each fixed ℓ and $t \geq 0$, we obtain that

$$\max_{1 \leq i \leq a+b-\ell} \Delta_{i,\ell}^{n,(1)} \rightarrow 0 \quad \text{and} \quad \max_{1 \leq i \leq a+b-\ell} \Delta_{i,\ell}^{n,(2)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus, we have

$$R_\ell^n = \int_0^t E[H(t - u, Z_1)H(t - u, Z_{1+\ell}) - G_1(t, u)^2] dA^n(u) + o(n).$$

Therefore, as $n \rightarrow \infty$,

$$\begin{aligned} \frac{2}{n} \sum_{i < j}^{A^n(t)} r_{ij}^n &= \frac{2}{n} \sum_{\ell=1}^{A^n(t)-1} R_\ell^n \rightarrow 2 \sum_{\ell=1}^{\infty} \int_0^t E[\mathcal{H}_1(t)\mathcal{H}_{1+\ell}(t)] d\Lambda(u) \\ &= \int_0^t \Gamma(t, t, u) d\Lambda(u). \end{aligned}$$

Thus, by (24), we obtain

$$\frac{\text{Var}(S_n)}{n} \rightarrow \sigma^2 := \int_0^t [\check{G}(t, u) + \Gamma(t, t, u)] d\Lambda(u) \quad \text{as } n \rightarrow \infty. \quad (25)$$

We have verified conditions (i)–(iv), and by theorem 2.1 and proposition 2.1(a) in Withers (1981), we have

$$S_n / \sqrt{\text{Var}(S_n)} \Rightarrow N(0, 1) \quad \text{as } n \rightarrow \infty. \quad (26)$$

By definition of S_n , we have $S_n = \sqrt{n} \hat{X}_2^n(t)$. Thus, by (25) and (26), we have

$$\hat{X}_2^n(t) \Rightarrow N(0, \sigma^2) \stackrel{\text{d}}{=} \hat{X}_2(t) \quad \text{as } n \rightarrow \infty,$$

where “ $\stackrel{\text{d}}{=}$ ” denotes “equal in distribution.”

To finish the proof, by the Cramér–Wold theorem, it suffices to show that, for any $m \in \mathbb{N}^+$ and $0 \leq t_1 < t_2 < \dots < t_m \leq T$ and $\{a_i \in \mathbb{R} : i = 1, \dots, m\}$,

$$\sum_{i=1}^m a_i \hat{X}_2^n(t_i) \Rightarrow \sum_{i=1}^m a_i \hat{X}_2(t_i) \quad \text{as } n \rightarrow \infty.$$

We have proved the case when $m = 1$. Now we consider the case when $m = 2$. By simple algebra, we can write

$$\begin{aligned} \sqrt{n} \sum_{i=1}^2 a_i \hat{X}_2^n(t_i) &= a_1 \sum_{i=1}^{A^n(t_1)} \mathcal{H}_i^n(t_1) + a_2 \sum_{i=1}^{A^n(t_2)} \mathcal{H}_i^n(t_2) \\ &= \sum_{i=1}^{A^n(t_2)} \mathcal{H}_i^{n,*}, \end{aligned}$$

where

$$\mathcal{H}_i^{n,*} := \begin{cases} a_1 \mathcal{H}_i^n(t_1) + a_2 \mathcal{H}_i^n(t_2) & \text{for } 1 \leq i \leq A^n(t_1), \\ a_2 \mathcal{H}_i^n(t_2) & \text{for } A^n(t_1) + 1 \leq i \leq A^n(t_2). \end{cases}$$

Because the randomness of $\mathcal{H}_i^{n,*}$ comes only from Z_i , the dependence between $\mathcal{H}_i^{n,*}$ is the same between $\mathcal{H}_i^n$. Therefore, the similar arguments for the first case apply. It is clear that this argument can be extended for any general $m > 2$. This completes the proof.

3.3. Tightness

For any real-valued function f on $[0, T]$, denote its modulus of continuity by

$$\omega_\delta(f) := \sup_{|x-y| \leq \delta, x, y \in [0, T]} |f(x) - f(y)|, \quad \delta > 0. \quad (27)$$

Define the processes $\hat{D}^n := \{\hat{D}^n(x) : x \in [0, T]\}$ by

$$\hat{D}^n(x) := \frac{1}{\sqrt{n}} \sum_{i=1}^n (H(x, Z_i) - E[H(x, Z_1)]), \quad x \in [0, T]. \quad (28)$$

We next show the tightness of $\{\hat{D}^n : n \geq 1\}$ in the space $\mathbb{D}$.

Lemma 6. Under Assumptions 1 and 2, for each $\eta > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P(\omega_\delta(\hat{D}^n) > \eta) = 0. \quad (29)$$

Proof of Lemma 6. We prove that

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \left\| \sup_{\varrho(x, y) \leq \delta} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 = 0, \quad (30)$$

where

$$\varrho(x, y) := \|H(x, Z_1) - H(y, Z_1)\|_2, \quad \text{for } x, y \in [0, T], \quad (31)$$

This implies that

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} E[\omega_\delta(\hat{D}^n)^2] = 0. \quad (32)$$

To see that, suppose (30) holds; then, for any $\epsilon > 0$, there exists $\delta_0 > 0, n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$,

$$\left\| \sup_{\varrho(x, y) \leq \delta_0} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 \leq \epsilon. \quad (33)$$

By Assumption 1, there exists $\delta_1 > 0$ such that $\varrho(x, y) \leq \delta_0$ when $|x - y| \leq \delta_1$. That is, the set $\{(x, y) : |x - y| \leq \delta_1\}$ is a subset of $\{(x, y) : \varrho(x, y) \leq \delta\}$. Thus,

$$\left\| \sup_{|x-y| \leq \delta_1} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 \leq \left\| \sup_{\varrho(x,y) \leq \delta_0} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 \leq \epsilon. \quad (34)$$

We proceed to prove (30). Define the bracket number $N(\delta, T)$ as the minimal number of N that there exist points $t_1, \dots, t_N \in [0, T]$ such that, for each $t \in [0, T]$, there exists a t_i for which $\varrho(t, t_i) \leq \delta$. To show (30), we apply theorem 2.2 in Andrews and Pollard (1994) (under strong α -mixing condition) with $Q = 4$, $0 < \gamma < 2$ in their notation. The first summable condition there obviously holds by our Assumption 2. It then suffices to show that $N(\delta, T) = O(\delta^{-2})$ so that the integral condition $\int_{[0,1]} \delta^{-\gamma/(2+\gamma)} N(\delta, T)^{1/4} d\delta < \infty$ is fulfilled.

Define $g(t) = E[H(t, Z_1)H(T, Z_1)]$ for $t \in [0, T]$. Because $H \in \mathbb{D}$ is nondecreasing, then the function g is also nondecreasing. Thus, for a given $\delta > 0$, we can choose a partition of $[0, T]$ by $0 = t_0 < t_1 < \dots < t_N = T$ such that $g(t_{i+1}) - g(t_i) \leq \delta^2$ for every i . Then, for each i and $s \in [t_i, t_{i+1})$,

$$\varrho(s, t_i)^2 \leq E[(H(s, Z_1) - H(t_i, Z_1))H(T, Z_1)] \leq g(t_{i+1}) - g(t_i) \leq \delta^2. \quad (35)$$

Because $g \in \mathbb{D}$, the number of jump points with jump size greater than δ^2 is finite; see theorem 12.2.1 in Whitt (2002). Therefore, after making sure that the aforementioned jump points are among $t_0, \dots, t_N$, the number N , an upper bound for the bracket number $N(\delta, T)$, is bounded by a constant times $1/\delta^2$. The proof is now complete.

Lemma 7. Under Assumptions 1 and 2,

$$\sup_{n \in \mathbb{N}} E \left[\sup_{x \in [0, T]} \hat{D}^n(x)^2 \right] < \infty. \quad (36)$$

Proof of Lemma 7. Fix $\delta > 0$, choose $\{x_i : 0 \leq i \leq [T/\delta] + 1\}$ such that $0 = x_0 < x_1 < \dots < x_{[T/\delta]+1} = T$ and $x_{i+1} - x_i = T/\delta$ for each i . Then, for any $n_0 \in \mathbb{N}$,

$$\sup_{n \in \mathbb{N}} E \left[\sup_{x \in [0, T]} \hat{D}^n(x)^2 \right] \leq \max_{1 \leq n \leq n_0} E \left[\sup_{x \in [0, T]} \hat{D}^n(x)^2 \right] + \sup_{n \in \mathbb{N}} \sum_{i=1}^{[T/\delta]+1} E \left[\hat{D}^n(x_i)^2 \right] + \sup_{n \geq n_0} E \left[\omega_\delta(\hat{D}^n)^2 \right]. \quad (37)$$

The first summand is the maximum over finite many finite values. For the second summand, by similar calculation and arguments in (25), we conclude that it is finite. The finiteness of the last summand (when n_0 large enough) is an immediate consequence of Lemma 6.

Lemma 8. Under Assumptions 1 and 2,

$$\hat{D}^n \Rightarrow \hat{D} \quad \text{in } (\mathbb{D}, J_1) \quad \text{as } n \rightarrow \infty, \quad (38)$$

where $\hat{D} = \{\hat{D}(x) : x \geq 0\}$ is a continuous centered Gaussian process with covariance function

$$\begin{aligned} \text{Cov}(\hat{D}(x), \hat{D}(y)) &= \left(E[H(x, Z_1)H(y, Z_1)] - \int_{\mathbb{R}^k} H(x, z)dF(z) \int_{\mathbb{R}^k} H(y, z)dF(z) \right) \\ &\quad + \sum_{j=2}^{\infty} \left[\left(E[H(x, Z_1)H(y, Z_j)] - \int_{\mathbb{R}^k} H(x, z)dF(z) \int_{\mathbb{R}^k} H(y, z)dF(z) \right) \right. \\ &\quad \left. + \left(E[H(x, Z_j)H(y, Z_1)] - \int_{\mathbb{R}^k} H(x, z)dF(z) \int_{\mathbb{R}^k} H(y, z)dF(z) \right) \right]. \end{aligned}$$

Proof of Lemma 8. The tightness property is proved in Lemmas 6 and 7. The convergence of finite dimensional distributions can be done by a similar but simpler argument as in the proof of Lemma 5, and thus, its proof is omitted. We, thus, obtain the weak convergence of $\hat{D}^n$ to $\hat{D}$.

For the existence of the process $\hat{D}$ in $\mathbb{C}$, we calculate the second moment of the increment:

$$\begin{aligned} E[|\hat{D}(x) - \hat{D}(y)|^2] &= \text{Var}(H(x, Z_1) - H(y, Z_1)) \\ &\quad + 2 \sum_{j=2}^{\infty} \left[E[[H(x, Z_1) - H(y, Z_1)][H(x, Z_j) - H(y, Z_j)]] - \left(\int_{\mathbb{R}^k} (H(x, z) - H(y, z)) dF(z) \right)^2 \right]. \end{aligned} \quad (39)$$

By Assumption 2, we obtain that

$$E[|\hat{D}(x) - \hat{D}(y)|^2] \leq (1 + 2C_\rho) E[[H(x, Z_1) - H(y, Z_1)]^2]. \quad (40)$$

Because H is nondecreasing, then we can apply theorem 13.6 in Billingsley (1999) to prove the existence in $\mathbb{D}$, and thus, in $\mathbb{C}$ by the continuity in quadratic mean. This completes the proof.

Define the two-parameter processes $\hat{V}^n(t, x)$ by

$$\hat{V}^n(t, x) := \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} (H(x, Z_i) - E[H(x, Z_1)]), \quad t, x \geq 0. \quad (41)$$

Lemma 9. Under Assumptions 1–3, for any $\epsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\sup_{t \in [0, T]} \sup_{|x-y| \leq \delta, x, y \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(t, y)| > \epsilon \right) = 0. \quad (42)$$

Proof of Lemma 9. Recall the definition of $\hat{D}^n$ in (28). Observe that

$$\begin{aligned} &\sup_{t \in [0, T]} \sup_{|x-y| \leq \delta, x, y \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(t, y)| \\ &\leq \max_{\ell \leq A^n(T)} \sup_{|x-y| \leq \delta, x, y \in [0, T]} |\hat{V}^n(\tau_\ell^n, x) - \hat{V}^n(\tau_\ell^n, y)| \\ &= \max_{\ell \leq A^n(T)} \sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell). \end{aligned} \quad (43)$$

For $x, y \in [0, T]$, define

$$L_i(x, y) := H(x, Z_i) - E[H(x, Z_1)] - H(y, Z_i) + E[H(y, Z_1)]. \quad (44)$$

Let $l_n = \lfloor n^{(5-\kappa)/8} \rfloor = o(\sqrt{n})$ for $\kappa > 1$. Then, applying the Ottaviani-type inequality under strong α -mixing in Lemma A.1 (noting $\alpha_n \leq \rho_n/4$, see Bradley 2005; and also Bücher 2015, lemma 3), we obtain

$$P \left(\max_{\ell \leq A^n(T)} \sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell) > 3\epsilon \right) \leq \frac{B_{n,1}(\delta) + B_{n,2}(\delta) + \frac{1}{4} [A^n(T)/l_n] \rho_{l_n}}{1 - \max_{\ell=1, \dots, A^n(T)} P(\sqrt{\ell/n} \omega_\delta(\hat{D}^\ell) > \epsilon)}, \quad (45)$$

where

$$B_{n,1}(\delta) := P \left(\sqrt{\frac{A^n(T)}{n}} \omega_\delta(\hat{D}^n) > \epsilon \right),$$

and

$$B_{n,2}(\delta) := P \left(\max_{\substack{p, q \in \{1, \dots, A^n(T)\}, \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x, y \in [0, T], \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_i(x, y) \right| > \epsilon \right). \quad (46)$$

First, by the trajectories of A^n defined in (22),

$$\sqrt{\frac{A^n(T)}{n}} \rightarrow \sqrt{\Lambda(T)} < \infty \quad \text{as } n \rightarrow \infty.$$

Thus, by Lemma 6,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} B_{n,1}(\delta) = 0.$$

Next, for large enough n , we have (when $1 < \kappa < 3$), $[A^n(T)/l_n]\rho_{l_n} = O(n^{(\kappa^2-4\kappa+3)/8}) = o(1)$, and thus,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} [A^n(T)/l_n]\rho_{l_n} = 0.$$

As for $B_{n,2}(\delta)$, we truncate $L_i(x, y)$ as follows. For any $M > 0$ and $i \in \mathbb{N}$,

$$L_i(x, y) = L_{i,M}(x, y) + L_{i,M^c}(x, y), \quad (47)$$

where

$$\begin{aligned} L_{i,M}(x, y) &:= L_i(x, y)\mathbf{1}(|L_i(x, y)| \leq M), \\ L_{i,M^c}(x, y) &:= L_i(x, y)\mathbf{1}(|L_i(x, y)| > M). \end{aligned}$$

(Such a convention, denoting the truncated function by subscripting M or M^c , applies to other functions or processes as well and will not cause confusion in the context.) Then we have

$$\begin{aligned} & \max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_i(x, y) \right| \\ & \leq \max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_{i,M}(x, y) \right| \\ & \quad + \max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_{i,M^c}(x, y) \right| \\ & \leq 2Ml_n/\sqrt{n} + \max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_{i,M^c}(x, y) \right|. \end{aligned}$$

Evidently, for any fixed $M > 0$, $\lim_{n \rightarrow \infty} 2Ml_n/\sqrt{n} = 0$. On the other hand, by stationarity of $\{Z_n : n \geq 1\}$,

$$\begin{aligned} & P \left(\max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_{i,M^c}(x, y) \right| > \epsilon/2 \right) \\ & \leq P \left(\max_{\substack{p,q \in \{1, \dots, A^n(T)\} \\ 0 < q-p \leq 2l_n}} \sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \sum_{i=p+1}^q |L_{i,M^c}(x, y)| > \epsilon/2 \right) \quad (\text{triangle inequality}) \\ & \leq 2l_n(A^n(T) - 2l_n)P \left(\sup_{\substack{x,y \in [0,T] \\ |x-y| \leq \delta}} \frac{1}{\sqrt{n}} \sum_{i=1}^{2l_n} |L_{i,M^c}(x, y)| > \epsilon/2 \right) \quad (\text{by stationarity of } \{Z_i\}) \\ & \leq 2l_n(A^n(T) - 2l_n)P \left(\frac{2}{\sqrt{n}} \sum_{i=1}^{2l_n} |H(T, Z_i) + E[H(T, Z_i)]| \mathbf{1}(|H(T, Z_i)| > M/2) > \epsilon/2 \right). \\ & \hspace{15em} (\text{by definition of } L_i) \end{aligned} \quad (48)$$

We then apply theorem 4.1 in Shao and Yu (1996). Choose $p = 4, r = 8$ in their notation, then by (4.4) in Shao and Yu (1996),

$$E \left[\left(\frac{2}{\sqrt{n}} \sum_{i=1}^{2l_n} |H(T, Z_i) + E[H(T, Z_i)]| \mathbf{1}(|H(T, Z_i)| > M/2) \right)^4 \right] \leq K \frac{l_n^2}{n^2},$$

where K is a constant because it is assumed that $\|H(T, Z_1)\|_{8+\delta} < \infty$ for some $\delta > 0$ in Assumption 2.

An application of Markov's inequality to (48) yields an upper bound:

$$2l_n(A^n(T) - 2l_n)Kl_n^2/n^2 \leq O(l_n^3/n) = o(1),$$

when $\kappa \in (7/3, 3)$.

To finish the treatment for the second summand in (A.1), we need to show that the denominator in (45) is bounded above from zero for large enough n and small enough δ . By Lemma 6, there exist $\delta_0 > 0$ and $n_0 = n_0(\delta_0)$ such that $P(\omega_\delta(\hat{D}_n) > \epsilon) < 1/2$ for all $\delta < \delta_0$ and $n > n_0$. Thus, for $\delta < \delta_0$,

$$\max_{\ell=n_0, \dots, A^n(T)} P \left(\sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell) > \epsilon \right) \leq \max_{\ell=n_0, \dots, A^n(T)} P(\omega_\delta(\hat{D}^\ell) > \epsilon) < 1/2.$$

For $\ell < n_0$ and $\delta > 0$, we have

$$\omega_\delta(\hat{D}^\ell) \leq 2 \sup_{x \in [0, T]} |\hat{D}^\ell(x)| \leq 2 \sup_{x \in [0, T]} |\hat{D}_M^\ell(x)| + 2 \sup_{x \in [0, T]} |\hat{D}^\ell(x) - \hat{D}_M^\ell(x)|,$$

where $\hat{D}_M^\ell(x)$ is defined as

$$\hat{D}_M^\ell(x) := \frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} (H(x, Z_i) - E[H(x, Z_1)]) \mathbf{1}(|H(x, Z_i)| \leq M/2), \quad x \in [0, T].$$

Thus,

$$\begin{aligned} & \max_{\ell=1, \dots, n_0-1} P \left(\sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell) > 2\epsilon \right) \\ & \leq \max_{\ell=1, \dots, n_0-1} P \left(\sqrt{\frac{\ell}{n}} \sup_{x \in [0, T]} |\hat{D}_M^\ell(x)| > \epsilon/2 \right) \\ & \quad + \max_{\ell=1, \dots, n_0-1} P \left(\sqrt{\frac{\ell}{n}} \sup_{x \in [0, T]} |\hat{D}^\ell(x) - \hat{D}_M^\ell(x)| > \epsilon/2 \right). \end{aligned}$$

Note that $\sup_{x \in [0, T]} |\hat{D}_M^\ell(x)| \leq \sqrt{\ell} M n_0$. Thus, the first term on the right-hand side converges to zero as $n \rightarrow \infty$. The second term is upper bounded by

$$\frac{16n_0}{n\epsilon^2} \sum_{i=1}^{n_0} E[|H(T, Z_i)|^2 \mathbf{1}(|H(T, Z_i)| > M/2)] = o(1).$$

Therefore, the denominator in (45) is bounded above from zero for large enough n and small enough δ . This completes the proof of the lemma.

Lemma 10. Under Assumptions 1–3, for any $\epsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\sup_{\substack{|s-t| \leq \delta \\ s, t \in [0, T]}} \sup_{x \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(s, x)| > \epsilon \right) = 0. \quad (49)$$

Proof of Lemma 10. It suffices to show that

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\max_{\substack{0 \leq j \leq T \\ j \in \mathbb{N}}} \sup_{j\delta \leq t \leq (j+1)\delta} \sup_{x \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(j\delta, x)| > 3\epsilon \right) = 0.$$

The probability in this equation can be bounded by

$$\begin{aligned} & \sum_{j=1}^{\lceil T/\delta \rceil + 1} P \left(\sup_{j\delta \leq t \leq (j+1)\delta} \sup_{x \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(j\delta, x)| > 3\epsilon \right) \\ &= \sum_{j=1}^{\lceil T/\delta \rceil + 1} P \left(\max_{\ell=A^n(j\delta)+1, \dots, A^n((j+1)\delta)} \sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=A^n(j\delta)+1}^{\ell} (H(x, Z_i) - E[H(x, Z_i)]) \right| > 3\epsilon \right). \end{aligned}$$

Applying the Ottaviani-type inequality in Lemma A.1 (lemma 3 in Bücher 2015), we obtain the preceding quantity is further bounded by

$$\sum_{j=1}^{\lceil T/\delta \rceil + 1} \left(C_{j,n,1}(\delta) + C_{j,n,2}(\delta) + \frac{1}{4} [(A^n((j+1)\delta) - A^n(j\delta))/l_n] \rho_{l_n} \right) / (1 - C_{j,n,3}) \quad (50)$$

where

$$C_{j,n,1}(\delta) := P \left(\sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=A^n(j\delta)+1}^{A^n((j+1)\delta)} (H(x, Z_i) - E[H(x, Z_i)]) \right| > \epsilon \right), \quad (51)$$

$$C_{j,n,2}(\delta) := P \left(\max_{\substack{p, q \in \{A^n(j\delta)+1, \dots, A^n((j+1)\delta)\} \\ 0 < q - p < 2l_n}} \sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q (H(x, Z_i) - E[H(x, Z_i)]) \right| > \epsilon \right), \quad (52)$$

and

$$C_{j,n,3}(\delta) := \max_{\ell=A^n(j\delta)+1, \dots, A^n((j+1)\delta)} P \left(\sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=\ell}^{A^n((j+1)\delta)} (H(x, Z_i) - E[H(x, Z_i)]) \right| > \epsilon \right). \quad (53)$$

Recall the trajectories of $\{A^n(t) : t \geq 0\}$ defined in (22).

We first treat $C_{j,n,1}(\delta)$ in (51). Recall $\hat{D}^n$ defined in (28). We have

$$C_{j,n,1}(\delta) = P \left(\sup_{x \in [0, T]} \left| \sqrt{A^n((j+1)\delta)/n} \hat{D}^{A^n((j+1)\delta)}(x) - \sqrt{A^n(j\delta)/n} \hat{D}^{A^n(j\delta)}(x) \right| > \epsilon \right).$$

By (22), Lemma 8 and the Portmanteau theorem, we obtain

$$\limsup_{n \rightarrow \infty} C_{j,n,1}(\delta) \leq P \left(\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \sup_{x \in [0, T]} |\hat{D}(x)| > \epsilon \right) = P \left(\sup_{x \in [0, T]} |\hat{D}(x)| > \epsilon / \left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \right). \quad (54)$$

Because $\sup_{x \in [0, T]} |\hat{D}(x)|$ possesses moments of any order, the preceding probability converges to zero faster than any order of δ by the continuity of Λ and proposition A.2.3 in van der Vaart and Wellner (1996). Therefore, we have

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \sum_{j=1}^{\lceil T/\delta \rceil + 1} C_{j,n,1} \leq \lim_{\delta \downarrow 0} \sum_{j=1}^{\lceil T/\delta \rceil + 1} P \left(\sup_{x \in [0, T]} |\hat{D}(x)| > \epsilon / \left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \right) = 0. \quad (55)$$

Next, we turn to $C_{j,n,2}(\delta)$ in (52). Recall the definition of $L_i(x, y)$ following (43). We have

$$C_{j,n,2}(\delta) = P \left(\max_{\substack{p, q \in \{A^n(j\delta)+1, \dots, A^n((j+1)\delta)\} \\ 0 < q - p < 2l_n}} \sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^q L_i(x, 0) \right| > \epsilon \right). \quad (56)$$

Then $\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \sum_{j=1}^{\lceil T/\delta \rceil + 1} C_{j,n,2} = 0$ follows from similar arguments for B_{n2} from (47) and (48).

Next, for $\delta \leq 1$, we have

$$\begin{aligned} & \left| \sum_{j=1}^{\lfloor T/\delta \rfloor + 1} [(A^n((j+1)\delta) - A^n(j\delta))/l_n] \rho_{l_n} \right| \\ & \leq \left(\sum_{j=1}^{\lfloor T/\delta \rfloor + 1} (\Lambda((j+1)\delta) - \Lambda(j\delta)) \right) n \rho_{l_n} / l_n + 2(\lfloor T/\delta \rfloor + 1) \epsilon(n) n \rho_n / l_n \\ & = (\Lambda(\lfloor T/\delta \rfloor + 2) - \Lambda(\delta) + 2(\lfloor T/\delta \rfloor + 1) \epsilon(n)) n \rho_n / l_n \\ & = o(1), \end{aligned} \quad (57)$$

where the last equality follows from $\epsilon(n) = o(1)$ and $n \rho_n / l_n = o(1)$.

Given (55)–(57), we have

$$\sum_{j=1}^{\lfloor T/\delta \rfloor + 1} \left(C_{j,n,1}(\delta) + C_{j,n,2}(\delta) + \frac{1}{4} [(A^n((j+1)\delta) - A^n(j\delta))/l_n] \rho_{l_n} \right) = 0. \quad (58)$$

To prove (50), it now suffices to show that there exists $\epsilon_0 > 0$ such that for sufficiently large n , small δ , and each $1 \leq j \leq \lfloor T/\delta \rfloor + 1$, we have $C_{j,n,3}(\delta) \leq 1 - \epsilon_0$.

Recall the definition of $C_{j,n,3}$ in (53). For each $\delta > 0$ and $j \in \mathbb{N}$, choose a constant $k_{\delta,j}^n$ such that $A^n(j\delta) + 1 \leq k_{\delta,j}^n \leq A^n((j+1)\delta)$. Then, by a similar argument for $C_{j,n,2}$, we obtain that

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \max_{\ell = A^n(j\delta) + 1, \dots, k_{\delta,j}^n} P \left(\sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=\ell}^{A^n((j+1)\delta)} (H(x, Z_i) - E[H(x, Z_i)]) \right| > \epsilon \right) = 0.$$

On the other hand, by a similar argument as (54), we have from the Portmanteau theorem that, for $\ell \geq A^n(j\delta) + 1$,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} P \left(\sup_{x \in [0, T]} \frac{1}{\sqrt{n}} \left| \sum_{i=\ell}^{A^n((j+1)\delta)} (H(x, Z_i) - E[H(x, Z_i)]) \right| > \epsilon \right) \\ & \leq P \left(\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \sup_{x \in [0, T]} |\hat{D}(x)| > \epsilon \right) \\ & = P \left(\sup_{x \in [0, T]} |\hat{D}(x)| > \epsilon / \left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \right) \\ & < 1. \end{aligned}$$

Because $\Lambda \in \mathbb{C}$, we conclude that there exists ϵ_0 such that, for sufficiently large n , small δ , and each $1 \leq j \leq \lfloor T/\delta \rfloor + 1$, we have $C_{j,n,3}(\delta) \leq 1 - \epsilon_0$. The proof is now complete.

Lemma 11. The sequence $\{\hat{X}_2^n : n \geq 1\}$ is tight in space $\mathbb{D}$.

Proof of Lemma 11. By the triangle inequality, we have

$$\sup_{\substack{|s-t| \leq \delta \\ s, t \in [0, T]}} |\hat{X}_2^n(t) - \hat{X}_2^n(s)| \leq \sup_{\substack{|s-t| \leq \delta \\ s, t \in [0, T]}} \sup_{x \in [0, T]} |\hat{V}^n(t, x) - \hat{V}^n(s, x)| + \sup_{t \in [0, T]} \sup_{\substack{|x-y| \leq \delta \\ x, y \in [0, T]}} |\hat{V}^n(t, x) - \hat{V}^n(t, y)|.$$

Then the proof is complete by Lemmas 9 and 10.

Remark 4. As can be seen from the proof, tightness also holds under the strong α -mixing condition because the moment bounds and the Ottaviano-type inequality we use are established under the α -mixing condition. However, as mentioned, the proof of existence of the limit process $\hat{X}_2$ in $\mathbb{C}$ in Theorem 7 can be carried out only under the ρ -mixing condition. In the i.i.d. case, the tightness proof is done in Pang and Zhou (2018a) by applying the new sufficient convergence criterion in $(\mathbb{D}, J_1)$ (extending the classical result in theorem 13.5 of Billingsley 1999) by verifying the probability bound for the process increments. That approach relies on the new maximality inequality, extending theorems 10.3 and 10.4 in Billingsley (1999), that uses the set function with a superadditivity property

instead of a finite measure. However, that approach does not work for weakly dependent variables under mixing conditions; specifically, the probability bound for the increments in Lemma 4 cannot be established for the scaled processes $\hat{X}_2^n$.

3.4. Completing the Proof of Theorem 2

Proof of Theorem 2. Recall the set Υ^n of trajectories $\{A^n(t) : t \geq 0\}$ defined in (22), which is used in the proofs of the convergence of the process $\hat{X}_2^n$, both in convergence of finite-dimensional distributions in Lemma 5 and tightness in Lemmas 9 and 10.

We directly obtain the tightness of $\{\hat{X}_1^n + \hat{X}_2^n : n \in \mathbb{N}\}$ on Υ^n given the tightness property shown for $\{\hat{X}_1^n\}$ and $\{\hat{X}_2^n\}$. For the convergence of their finite dimensional distributions, we first consider the convergence of $\hat{X}_1^n(t) + \hat{X}_2^n(t)$ for some $t \geq 0$ on Υ^n . Let $\mathcal{A}^n(t) = \sigma\{A^n(s) : 0 \leq s \leq t\}$ be the filtration generated by the arrival process. We have, for $\vartheta \in \mathbb{R}$,

$$\begin{aligned} E\left[e^{i\vartheta(\hat{X}_1^n(t) + \hat{X}_2^n(t))}\right] &= E\left[E\left[e^{i\vartheta(\hat{X}_1^n(t) + \hat{X}_2^n(t))} \middle| \mathcal{A}^n(t)\right]\right] \\ &= E\left[E\left[e^{i\vartheta\hat{X}_1^n(t)} e^{i\vartheta\hat{X}_2^n(t)} \middle| \mathcal{A}^n(t)\right]\right] = E\left[e^{i\vartheta\hat{X}_1^n(t)} E\left[e^{i\vartheta\hat{X}_2^n(t)} \middle| \mathcal{A}^n(t)\right]\right]. \end{aligned}$$

On Υ^n , by the convergence of $\hat{X}_1^n$ and by the proof of Lemma 5 for the convergence of finite-dimensional distributions of $\hat{X}_2^n$, we obtain that the preceding equation converges to $E[e^{i\vartheta(\hat{X}_1(t) + \hat{X}_2(t))}]$. The convergence of finite-dimensional distributions follows similarly. Because $P(\Upsilon^n) \rightarrow 1$ as $n \rightarrow \infty$, the proof is complete.

Acknowledgments

The authors thank Harold Dehling for helpful discussions. They also appreciate the helpful comments by the associate editor and two anonymous referees.

Appendix

A.1. Existence Criterion in the Space $\mathbb{D}$

A classical existence criterion for a stochastic process with sample paths in $\mathbb{D}$ given its finite dimensional distributions is given in theorem 13.6 in Billingsley (1999). That criterion requires a probability bound for the increments of the processes in which the upper bound involves a finite measure. Its proof relies on the classical maximal inequalities in theorems 10.3 and 10.4 in Billingsley (1999). Using the new maximal inequalities requiring a set function with the superadditivity property, a new existence criterion in the space $\mathbb{D}$ is established in Pang and Zhou (2018a). We now state this criterion of existence in the following theorem. It is used in the proof of Theorem 7.

Recall the definition of a set function with the superadditivity property in Definition 2.

Theorem A.1 (Pang and Zhou 2018a, theorem 5.3). *There exists a random element X in $\mathbb{D}([0, T], \mathbb{R})$ with finite-dimensional distributions $\pi_{t_1, \dots, t_k}$ for any $0 \leq t_1 < \dots < t_k \leq T$, that is, $\pi_{t_1, \dots, t_k}(x_1, \dots, x_k) = P(X(t_1) \leq x_1, \dots, X(t_k) \leq x_k)$ for $x_i \in \mathbb{R}$, $i = 1, \dots, k$, if the following conditions are satisfied:*

- i. *the finite dimensional distributions $\pi_{t_1, \dots, t_k}$ are consistent, satisfying the conditions of Kolmogorov's existence theorem.*
- ii. *For any $0 \leq r \leq s \leq t \leq T$ and $\epsilon > 0$,*

$$P(|X(r) - X(s)| \wedge |X(s) - X(t)| \geq \epsilon) \leq \frac{C}{\epsilon^{4\beta}} (\mu(r, t))^{2\alpha}, \quad (\text{A.1})$$

for some $\beta \geq 0$ and $\alpha > 1/2$, where C is a positive constant, μ is a finite set function in Definition 2 and $\mu(0, t]$ is continuous in t .

- iii. *For any $\epsilon > 0$ and $t \in [0, T)$,*

$$\lim_{\delta \downarrow 0} P(|X(t) - X(t + \delta)| > \epsilon) = 0. \quad (\text{A.2})$$

4.2. An Ottaviani-Type Inequality

Let $\{X_n : n \in \mathbb{Z}\}$ be a sequence of random elements in some Banach space $\mathbb{E}$ satisfying the strong α -mixing condition (Bradley 2005) with the mixing coefficients $\{\alpha_n\}$. Let $\mathbb{T}$ be some arbitrary index set. For each $i \in \mathbb{Z}$, let $G_i \in \ell^\infty(\mathbb{E} \times \mathbb{T})$. For each t , let $Y_i(t) = G_i(X_i, t)$ and $S_n(t) = \sum_{i=1}^n Y_i(t)$ for $n \geq 1$ with $S_0 \equiv 0$. Let $\|S_n\| := \sup_{t \in \mathbb{T}} |S_n(t)|$.

Lemma A.1 (Bücher 2015, lemma 3). Suppose that $\|S_m - S_n\|$ is measurable for each $0 \leq n < m$. Then, for each $\epsilon > 0$ and $1 \leq k < n$,

$$\begin{aligned} & P\left(\max_{j=1,\dots,n} \|S_j\| > 3\epsilon\right) \times \left(1 - \max_{j=1,\dots,n} P\left(\|S_n - S_j\| > \epsilon\right)\right) \\ & \leq P(\|S_n\| > \epsilon) + P\left(\max_{j,l \in \{1,\dots,n\}, 0 < l-j \leq 2k} \|S_l - S_j\| > \epsilon\right) + \lfloor n/k \rfloor \times \alpha_k. \end{aligned} \quad (\text{A.3})$$

References

- Andrews DW, Pollard D (1994) An introduction to functional central limit theorems for dependent stochastic processes. *Internat. Statist. Rev.* 62(1):119–132.
- Biermé H, Desolneux A (2012) Crossings of smooth shot noise processes. *Ann. Appl. Probab.* 22(6):2240–2281.
- Billingsley P (1999) *Convergence of Probability Measures* (John Wiley & Sons, New York).
- Bradley RC (2005) Basic properties of strong mixing conditions: A survey and some open questions. *Probab. Surveys* 2(2):107–144.
- Bücher A (2015) A note on weak convergence of the sequential multivariate empirical process under strong mixing. *J. Theoretical Probab.* 28(3):1028–1037.
- Budhiraja A, Nyquist P (2015) Large deviations for multidimensional state-dependent shot noise processes. *J. Appl. Probab.* 52(4):1097–1114.
- Chobanov G (1999) Modeling financial asset returns with shot noise processes. *Math. Comput. Model.* 29(10–12):17–21.
- Dehling H, Philipp W (2002) *Empirical Process Techniques for Dependent Data* (Birkhäuser, Boston).
- Doney RA, O'Brien GL (1991) Loud shot noise. *Ann. Appl. Probab.* 1(1):88–103.
- Ganesh A, Torrisi GL (2006) A class of risk processes with delayed claims: Ruin probability estimates under heavy tail conditions. *J. Appl. Probab.* 43(4):916–926.
- Ganesh A, Macci C, Torrisi GL (2005) Sample path large deviations principles for Poisson shot noise processes and applications. *Electronic J. Probab.* 10:1026–1043.
- Hahn MG (1978) Central limit theorem in $D[0, 1]$. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 44(2):89–101.
- Heinrich L, Schmidt V (1985) Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. Appl. Probab.* 17(4):709–730.
- Huffer FW (1987) Inequalities for the $M/G/\infty$ queue and related shot noise processes. *J. Appl. Probab.* 24(4):978–989.
- Ibragimov IA (1959) Some limit theorems for stochastic processes stationary in the strict sense. *Dokl. Akad. Nauk SSSR.* 125(26):711–714.
- Iglehart DL (1973) Weak convergence of compound stochastic process, I. *Stochastic Process. Appl.* 1(1):11–31.
- Iksanov A (2013) Functional limit theorems for renewal shot noise processes with increasing response functions. *Stochastic Process. Appl.* 123(6):1987–2010.
- Iksanov A (2016) *Renewal Theory for Perturbed Random Walks and Similar Processes* (Birkhäuser, Cham).
- Iksanov A, Marynych A, Meiners M (2014) Limit theorems for renewal shot noise processes with eventually decreasing response functions. *Stochastic Processes Their Appl.* 124(6):2132–2170.
- Iksanov A, Marynych A, Meiners M (2017a) Asymptotics of random processes with immigration I: Scaling limits. *Bernoulli* 23(2):1233–1278.
- Iksanov A, Marynych A, Meiners M (2017b) Asymptotics of random processes with immigration II: Convergence to stationarity. *Bernoulli* 23(2):1279–1298.
- Jacobs PA, Lewis PA (1978) Discrete time series generated by mixtures. I: Correlational and runs properties. *J. Royal Statist. Soc. B.* 40(1):94–105.
- Kluppelberg C, Kühn C (2004) Fractional Brownian motion as a weak limit of Poisson shot noise processes—With applications to finance. *Stochastic Process. Appl.* 113(2):333–351.
- Kluppelberg C, Mikosch T (1995) Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* 1(1–2):125–147.
- Kluppelberg C, Mikosch T, Schärff A (2003) Regular variation in the mean and stable limits for Poisson shot noise. *Bernoulli* 9(3):467–496.
- Kolmogorov AN, Rozanov YA (1960) On strong mixing conditions for stationary Gaussian processes. *Theory Probab. Appl.* 5(2):204–208.
- Koops D, Mandjes M, Boxma O (2017) Networks of $/G/\infty$ server queues with shot-noise-driven arrival intensities. *Queueing Systems* 86(3–4):301–325.
- Krichagina EV, Puhalskii AA (1997) A heavy-traffic analysis of a closed queueing system with a GI/∞ service center. *Queueing Systems* 25(1–4):235–280.
- Kurtz TG (1996) Limit theorems for workload input models. Kelly FP, Zachary S, Ziedins I, eds. *Stochastic Networks: Theory and Applications*, Royal Statistical Society Lecture Note Series, vol. 4 (Oxford Science Publications, Oxford, UK), 119–140.
- Lane JA (1984) The central limit theorem for the Poisson shot-noise process. *J. Appl. Probab.* 21(2):287–301.
- Pang G, Taquu MS (2019) Non-stationary self-similar Gaussian processes as scaling limits of power-law shot noise processes and generalizations of fractional Brownian motion. *High Frequency* 2(2):95–112.
- Pang G, Whitt W (2010) Two-parameter heavy-traffic limits for infinite-server queues. *Queueing Systems* 65(4):325–364.
- Pang G, Whitt W (2012) Infinite-server queues with batch arrivals and dependent service times. *Probab. Engrg. Inform. Sci.* 26(2):197–220.
- Pang G, Whitt W (2013) Two-parameter heavy-traffic limits for infinite-server queues with dependent service times. *Queueing Systems* 73(2):119–146.
- Pang G, Zhou Y (2017) Two-parameter process limits for an infinite-server queue with arrival dependent service times. *Stochastic Processes Their Appl.* 127(5):1375–1416.
- Pang G, Zhou Y (2018a) Functional limit theorems for a new class of non-stationary shot noise processes. *Stochastic Processes Their Appl.* 128(2):505–544.
- Pang G, Zhou Y (2018b) Two-parameter process limits for infinite-server queues with dependent service times via chaining bounds. *Queueing Systems* 88(1–2):1–25.
- Papoulis A (1971) High density shot noise and Gaussianity. *J. Appl. Probab.* 18(1):118–127.
- Ramirez-Perez F, Serfling R (2001) Shot noise on cluster processes with cluster marks, and studies of long range dependence. *Adv. Appl. Probab.* 33(3):631–651.
- Rice J (1977) On generalized shot noises. *Adv. Appl. Probab.* 9(3):553–565.

- Rosenblatt M (1956) A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* 42(1):43–47.
- Samorodnitsky G (1996) A class of shot noise models for financial applications. Heyde CC, Prohorov YV, Pyke R, Rachev ST, eds. *Athens Conf. Appl. Probab. Time Series Anal.* (Springer, New York), 332–353.
- Samorodnitsky G (1998) Tail behavior of some shot noise processes. Adler RJ, Feldman RE, Taqqu MS, eds. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* (Birkhäuser, Basel), 473–486.
- Schmidt T (2014) Catastrophe insurance modeled by shot-noise processes. *Risks* 2(1):3–24.
- Schmidt V (1985) On finiteness and continuity of shot noise processes. *Optim.* 16(6):921–933.
- Shao QM, Yu H (1996) Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* 24(4):2098–2127.
- Valdez EA, Mo K (2002). Ruin probabilities with dependent claims. Working paper, University of New South Wales, Sydney, Australia.
- van der Vaart AW, Wellner JA (1996) *Weak Convergence and Empirical Process with Applications to Statistics* (Springer, New York).
- Withers CS (1981) Central limit theorems for dependent variables. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 57(4):509–534.
- Whitt W (1976) Bivariate distributions with given marginals. *Ann. Statist.* 4(6):1280–1289.
- Whitt W (1983) Comparing batch delays and customer delays. *Bell Systems Tech. J.* 62(7):2001–2009.
- Whitt W (2002) *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and Their Applications to Queues* (Springer-Verlag, New York).