Functional Limit Theorems for Shot Noise Processes with Weakly Dependent Noises

GUODONG PANG AND YUHANG ZHOU

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 a new existence criterion established in [42] which 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.

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, earthquake models, physics, and so on (see, e.g., [47, 5, 52, 35, 8, 29, 49, 10, 22]). A shot noise process $X := \{X(t) : t \geq 0\}$ is typically defined by

$$X(t) := \sum_{i=1}^{A(t)} H(t - \tau_i, Z_i), \quad t \geq 0, \tag{1.1}$$

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 $H : \mathbb{R}_+ \times \mathbb{R}^k \to \mathbb{R}$ 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 (1.1) is often assumed to be i.i.d. in the literature; see, e.g., [12, 14, 15, 18, 30, 31, 32, 35, 50, 22]. Due to 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 [30, 32] 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 [36], with Poisson arrivals and power-law non-stationary noises, it is shown that the integrated shot noise process under the conventional scaling has a self-similar Gaussian process which has non-stationary increments (a generalization of FBM). It is also shown in [31, 22, 21, 24, 25, 26] 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 [20], 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 non-stationary noises in [42], when the noises are i.i.d., under more general conditions, a continuous Gaussian process limit is also established, extending the results in [20]. It is worth noting that unlike the conventional scaling regime, the Gaussian limiting processes in [20, 42] 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 [54]. In [46], a cluster shot noise model is studied where $\{Z_i : i \in \mathbb{N}\}$ depends the same 'cluster mark' within each cluster, so that the noises within each cluster are correlated. Since 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 [38, 39]. 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 where the arrival rate is large while 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., [4, 17, 20, 44]. 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 (non-stationary) 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 [6] together with the inverse mapping in [60]). We establish a stochastic process limit for the similarly centered and scaled shot noise processes (Theorem 2.2). 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 2.2. It is worth mentioning that the weakly dependence assumption has no impact upon the fluid limit (Theorem 2.1). 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 2.2), 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 provided 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 while 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, where 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 $\varphi : \mathbb{R}^k \to \mathbb{R}$. Such a model has been widely studied in financial modeling [9] and insurance risk [15, 51], where, 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 2.3. It can be used to model batch arrival processes where the batch sizes are correlated, or the work-input process where 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 [39, 43]. As a consequence of Theorem 2.2, we obtain an FCLT for the work-input process for the infinite-server queues with ρ-mixing service times (Theorem 2.4). 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 [6], cannot be used. Instead, we can apply the new existence criterion, Theorem 5.3 in [42], 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 3.5), for which we adapt the central limit theorems (CLTs) for dependent variables under mixing conditions [57] and [11]. 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 [1, 53], and apply an Ottaviani-type inequality (see, e.g., [2]) in a nontrivial manner. See also Remark 3.1.

Finally, it is worth highlighting that the high intensity/density regime is different from the conventional regime where 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 [42] and the Ottaviani-type inequality in [2].

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 (nonnegative) 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 cádlág 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 [6] and (3.2) and its extension to domain $\mathbb{R}_+$ in [60, 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 $\to$ 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 non-zero g, we write $f(x) = o(g(x))$ if $\lim_{x \to \infty} |f(x)/g(x)| = 0$.

2. Model and Results

We first state the assumptions to study the shot noise process X defined in (1.1). We start with the shot shape or the response function H. To simplify notations, we define the following functions: for each $0 \le u \le s \le t$,

$$G_1(t, u) := E[H(t - u, Z_1)] = \int_{\mathbb{R}^k} H(t - u, x) dF(x), \qquad (2.1)$$

$$G_2(t, u) := E\big[H(t - u, Z_1)^2\big] = \int_{\mathbb{R}^k} H(t - u, x)^2 dF(x),$$

$$\tilde{G}(t, u) := \mathrm{Var}\big(H(t - u, Z_1)\big) = G_2(t, u) - G_1(t, u)^2,$$

$$\check{G}_1(t, s, u) := E\big[H(t - u, Z_1) - H(s - u, Z_1)\big] = \int_{\mathbb{R}^k} \big(H(t - u, x) - H(s - u, x)\big) dF(x),$$

$$\check{G}_2(t, s, u) := E\big[\big(H(t - u, Z_1) - H(s - u, Z_1)\big)^2\big] = \int_{\mathbb{R}^k} \big(H(t - u, x) - H(s - u, x)\big)^2 dF(x),$$

$$\tilde{G}(t, s, u) := \mathrm{Var}\big(H(t - u, Z_1) - H(s - u, Z_1)\big) = \check{G}_2(t, s, u) - \check{G}_1(t, s, u)^2.$$

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 $t \ge 0$. For $0 \le t \le T$,*

$$\lim_{\delta \downarrow 0} \int_{[0,T]} \check{G}_2(t, t - \delta, u) d\Lambda(u) = 0. \qquad (2.2)$$

Remark 2.1. 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 3.6 and 3.8. The continuity of $G_1(t, \cdot)$ is needed for the continuity of the limit process $\hat{X}_1$ (see (2.8) in Theorem 2.1) and the convergence of finite dimensional distributions of $\hat{X}_2^n$ (Lemma 3.5). 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 while $H(\cdot, x)$ is not; see, e.g., 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 ρ-mixing condition was first introduced in [28] by Kolmogorov and Rozanov. We refer the reader to [7] for a thorough review on mixing conditions. Note that the ρ-mixing condition is weaker than uniform strong mixing (ϕ-mixing, [19]) but stronger than strong mixing (α-mixing, [48]).

Assumption 2. *The sequence of random vectors $\{Z_i : i \in \mathbb{N}\}$ is weakly dependent and stationary with a continuous 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, i.e., $F^c(x) = 1 - F(x)$ for $x \in \mathbb{R}^k$.

We provide an example where 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 $k \to \infty$; see page 201 in [6]. See also the example in Remark 2.3.

We consider a sequence of shot noise processes indexed by n and let $n \to \infty$. We write A^n and X^n and the associated $\{\tau_i^n\}$ for the nth process, while 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:*

$$\hat{A}^n := \sqrt{n}(\bar{A}^n - \Lambda) \Rightarrow \hat{A} \quad in \quad (\mathbb{D}, \ J_1) \quad as \quad n \to \infty, \tag{2.3}$$

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 $n \to \infty$. 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 (2.3) is satisfied by many models of arrival processes, including renewal process, stationary point processes with weakly dependent interarrival times, Markov-modulated Poisson process, Hawkes process, etc.; we refer the reader to the discussion below Assumption 1 in [42]. 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 [6]. 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 [60].

Define the process $\bar{X}^n := \{\bar{X}^n(t) : t \geq 0\}$ by $\bar{X}^n(t) := n^{-1}X^n(t)$ for $t \geq 0$, 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 2.1. (*FWLLN*) *Under Assumptions 1–3,*

$$\bar{X}^n(t) \Rightarrow \bar{X}(t) \quad in \quad (\mathbb{D}, J_1) \quad as \quad n \to \infty, \tag{2.4}$$

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

$$\bar{X}(t) := \int_{[0,t]} G_1(t,u)d\Lambda(u), \quad t \geq 0, \tag{2.5}$$

with $G_1(t,u)$ in (2.1).

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

$$\hat{X}^n(t) := \sqrt{n}(\bar{X}^n(t) - \bar{X}(t)), \quad t \geq 0, \tag{2.6}$$

where $\bar{X}(t)$ is given in (2.5).

Theorem 2.2. (*FCLT*) *Under Assumptions 1–3,*

$$\hat{X}^n \Rightarrow \hat{X} \quad in \quad (\mathbb{D},\ J_1) \quad as \quad n \to \infty, \tag{2.7}$$

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

$$\hat{X}_1(t) := \hat{A}(t)G_1(t,t) - \int_{(0,t]} \hat{A}(u)dG_1(t,u), \quad t \geq 0. \tag{2.8}$$

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

$$\hat{R}_2(t,s) := Cov\big(\hat{X}_2(t), \hat{X}_2(s)\big)$$
$$= \int_{[0,t \wedge s]} \big(G_2(t,s,u) - G_1(t,u)G_1(s,u) + \Gamma(t,s,u)\big)d\Lambda(u), \tag{2.9}$$

where

$$G_2(t,s,u) := \int_{\mathbb{R}^k} H(t-u,x)H(s-u,x)dF(x), \quad t \geq u \geq 0,\ s \geq u \geq 0, \tag{2.10}$$

and

$$\Gamma(t,s,u) \quad := \quad \sum_{l=2}^{\infty} \Big(E[H(t-u,Z_1)H(s-u,Z_l)] - G_1(t,u)G_1(s,u)$$

$$+ E[H(s-u,Z_1)H(t-u,Z_l)] - G_1(s,u)G_1(t,u)\Big). \tag{2.11}$$

Remark 2.2. 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 [42].

(ii) The series in (2.11) is summable under Assumption 2. In particular, for each summand in (2.11), 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 1–2, 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, whose covariance has the extra term $\Gamma(t, s, u)$ capturing the dependence in the noises. When the noises 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 [6], [7] and [11]. Our result indicates 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 [42], the limit processes $\hat{X}_1$ and $\hat{X}_2$ are independent. The weak dependence among the noises do 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 $\hat{X}_1$, while 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 [34, 37, 41].

Example 2.1. We give an example where $\hat{R}_2(t, s)$ 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 [27]. 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

$$\hat{R}_2(t, s) = \frac{1 + \alpha}{1 - \alpha} \int_{[0, t \wedge s]} \left(G_2(t, s, u) - G_1(t, u) G_1(s, u) d\Lambda(u), \right. \tag{2.12}$$

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 2.2. 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 ith 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 [38].)

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 $n \to \infty$, 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 is $\hat{A}_B(t) = \mathcal{B}(c_{a,B}^2 \Lambda_B(t))$ 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 below 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 $n \to \infty$, 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 2.2, we obtain that 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

$$\hat{R}_2(t,s) = m_B \int_{[0,t\wedge s]} \Big(\big(G_2(t,s,u) - G_1(t,u)G_1(s,u) \big) + \Gamma(t,s,u) \Big) d\Lambda_B(u), \qquad (2.13)$$

for $t,s \geq 0$, where $G_1(t,u)$ and $G_2(t,s,u)$ are given in (2.1) and (2.10), respectively, and

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

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., [59].

Observe that the impact of dependence among noises is captured in the double integral term in $\Gamma(t,s,u)$ above. Recall that any positively correlated bivariate c.d.f. can be approximated by a linear combination of the two extremals given the marginals [58], that is, $\Psi(x,y) \approx \widetilde{\Psi}_\rho(x,y)(x,y) = \rho F(x \wedge y) + (1-\rho)F(x)F(y)$ where $\rho \in (0,1)$ is the correlation parameter. Using this, we obtain an approximation of $\Gamma(t,s,u)$:

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

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

In [46], a shot noise process with cluster marks is studied, where cluster arrivals are Poisson and each entity of each cluster may arrive with certain random nonnegative "displacement" (delay). Our model with batch arrivals above 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 [46], then we can use the above 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, $Corr(Z_k, Z_l) = \frac{Var(\tilde{Z}_0)}{Var(\tilde{Z}_0)+Var(\tilde{Z}_1)}$ for each pair (k, l) in the cluster (batch). And we have $\Psi(x, y) = P(Z_k \le x, Z_l \le y) = P(\tilde{Z}_0 + \tilde{Z}_k \le x, \tilde{Z}_0 + \tilde{Z}_l \le y) = \int_{\mathbb{R}} \tilde{F}(x - z)\tilde{F}(y - z)d\tilde{F}_0(z)$, and $F(x) = P(\tilde{Z}_0 + \tilde{Z}_i \le x) = \int_{\mathbb{R}} \tilde{F}(x - z)d\tilde{F}_0(z)$. They can be easily used to calculate the covariance function $\hat{R}_2(t, s)$ in (2.13). In future work it would be interesting to establish the functional limit theorems for the shot noise process with general cluster marks where 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
$$H(t, x) := \tilde{H}(t)\varphi(x), \quad t \ge 0, \ x \in \mathbb{R}^k,$$
where $\varphi : \mathbb{R}^k \to \mathbb{R}$ is a measurable function. Assumption 1 requires that the function $\tilde{H}$ is monotone and continuous since
$$G_1(t, u) = E[H(t - u, Z_1)] = \tilde{H}(t - u)E[\varphi(Z_1)] \quad \text{for} \quad 0 \le u \le t.$$
In addition, it requires that $m_\varphi := E[\varphi(Z_1)]$ and $\sigma_\varphi^2 := \mathrm{Var}(\varphi(Z_1))$ are finite.

The limit process $\hat{X}_1$ in Theorem 2.2 becomes
$$\hat{X}_1(t) = m_\varphi \tilde{H}(0)\hat{A}(t) - m_\varphi \int_{(0,t]} \hat{A}(u)d\tilde{H}(t - u), \quad t \ge 0.$$

If $\hat{A}$ is a Gaussian process with mean 0 and covariance function $\hat{R}_a(t, s)$, then the covariance function of $\hat{X}_1$ is
$$\hat{R}_1(t, s) = m_\varphi^2 \int_{[0,t]} \int_{[0,s]} \tilde{H}(t - u)\tilde{H}(s - v)d\hat{R}_a(u, v), \quad t, s \ge 0.$$

If $\hat{A}(t) = c_a B(\Lambda(t))$ is a time-changed Brownian motion where 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
$$\hat{R}_1(t, s) = (c_a m_\varphi)^2 \int_{[0,t \wedge s]} \tilde{H}(t - u)\tilde{H}(s - u)d\Lambda(u), \quad t, s \ge 0.$$

The limit process $\hat{X}_2$ in Theorem 2.2 has the covariance function
$$\hat{R}_2(t, s) = \int_{[0,t \wedge s]} \left(\sigma_\varphi^2 \tilde{H}(t - u)\tilde{H}(s - u) + \Gamma(t, s, u) \right) d\Lambda(u), \quad t, s \ge 0,$$
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 2.3. We provide some examples of the shot shape functions. These are often used in the study of financial modeling and insurance risks [9, 15, 51].

  (i) *Exponential Decay.* The function $H(t,x)$ takes the form

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

  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 $0 \le u \le t$.

 (ii) *Power-law Decay.*

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

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

(iii) *Polynomial Decay.*

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

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

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

Example 2.4. With the DAR(1) noises $\{Z_i\}$ in Example 2.1, we obtain that

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

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:

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

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

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

As a consequence of Theorem 2.2, we obtain the following theorem for the compound processes under ρ-mixing condition.

Theorem 2.3. *Under Assumptions 2–3, (2.4) in the FWLLN holds with the limit $\bar{X}$ given by*

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

and (2.7) 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 $t, s \geq 0$,

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

Remark 2.3. With the DAR(1) noises $\{Z_i\}$ in Example 2.1, 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 increase, 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 [43]. 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 since 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 [43], 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 [33] for the FCLT with A^n being a Poisson process under the conventional scaling regime, where 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

$$G_1(t,u) = E[Z_1\mathbf{1}(Z_1 > t-u)] = \int_{(t-u,\infty)} x \, dF(x) \quad \text{for} \quad 0 \leq u \leq t. \qquad (2.14)$$

The conditions in Assumption 1 are satisfied since 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

$$\bar{W}(t) := \int_{[0,t]} \int_{(t-u,\infty)} x \, dF(x) d\Lambda(u), \quad t \geq 0. \qquad (2.15)$$

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 2.1–2.2, we obtain the following theorem for the work-input processes.

Theorem 2.4. *Under Assumptions 2–3,*

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

where $\bar{W}$ is defined in (2.15), and

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

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 (2.14), and $\hat{W}_2$ is a continuous Gaussian process with mean 0 and covariance function:

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

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

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

$$\left. + \Big[E[\tilde{Z}_1(s-u)\tilde{Z}_l(t-u)] - E[\tilde{Z}_1(s-u)]E[\tilde{Z}_l(t-u)]\Big] \right)$$

for $t,s \geq 0$.

Example 2.5. Suppose that the service times $\{Z_i\}$ are DAR(1) as in Example 2.1. We have the covariance function

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

for $t,s \geq 0$, where

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

and $G_1(t,u)$ is given in (2.14). The impact of dependence among service times is captured by the multiplicative coefficient $\frac{1+\alpha}{1-\alpha}$.

3. Proof of Theorem 2.2

Since Theorem 2.1 is directly implied by Theorem 2.2, we focus on the proof of Theorem 2.2 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 (2.1). We first give a representation for the process $\hat{X}^n$, which follows from simple calculations.

Lemma 3.1. *The process $\hat{X}^n$ defined in (2.6) 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*

$$\hat{X}_1^n(t) \;\; := \;\; \int_{(0,t]} G_1(t,u) d\hat{A}^n(u)$$

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

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

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

For the convergence of $\hat{X}_1^n$, by Lemma 6.1 in [42] (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 3.1. *Under Assumptions 2 and 3, $\hat{X}_1^n \Rightarrow \hat{X}_1$ in $(\mathbb{D}, J_1)$ as $n \to \infty$, where $\hat{X}_1$ is as given in Theorem 2.2.*

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 4.1, Theorem 3.3).

Step 2: The convergence of finite dimensional distributions of $\hat{X}_2^n$ to those of $\hat{X}_2$ (Section 3.2, Lemma 3.5).

Step 3: The tightness of $\hat{X}_2^n$ in the space $\mathbb{D}$ (Section 3.3, Lemma 3.11).

We summarize these results in the following theorem.

Theorem 3.2. *Under Assumptions 2–3, $\hat{X}_2^n \Rightarrow \hat{X}_2$ in $(\mathbb{D}, J_1)$ as $n \to \infty$, where $\hat{X}_2$ is as given in Theorem 2.2.*

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 3.2. *For each $0 \leq s \leq t$,*

$$E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big]$$

$$= \int_{(s,t]} \big(\tilde{G}(t,u) + \Gamma(t,t,u) \big) d\Lambda(u)$$

$$+ \int_{[0,s]} \big(\tilde{G}(t,s,u) + \Gamma(t,t,u) + \Gamma(s,s,u) - 2\Gamma(t,s,u) \big) d\Lambda(u). \tag{3.1}$$

Proof. The proof of (3.1) follows by (2.9) and direct calculations. $\qquad\square$

Definition 3.1. *Fix $T > 0$. For any $0 \leq s \leq t \leq T$, define a nonnegative function $V : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ by*

$$V(s,t) \;\; := \;\; \tilde{C}\big(\Lambda(t) - \Lambda(s)\big)$$

$$+ (1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} \big(H(t-u,x) - H(s-u,x) \big)^2 dF(x) d\Lambda(u). \tag{3.2}$$

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

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

Proposition 3.1. *The function V introduced in Definition 3.1 has the following properties:*
  (i) *$V(t,t) = 0$ for each $t \in [0,T]$;*
 (ii) *$V(s,t)$ 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 $s,t \geq 0$;*
(iii) *$V(s,t)$ is continuous in both s and t;*
(iv) *For any Borel set $A \subset [0,T]$, let*

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

ν satisfies all the conditions in Definition 3.2 and $\nu((s,t]) = V(s,t)$ for $0 \leq s \leq t \leq T$.

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

$$
\begin{aligned}
&V(s,t) - V(s - \delta_1, t - \delta_2) \\
&= \tilde{C}\big(\Lambda(t) - \Lambda(t - \delta_2) - (\Lambda(s) - \Lambda(s - \delta_1))\big) \\
&\quad + (1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} \big(H(t - u, x) - H(s - u, x)\big)^2 dF(x) d\Lambda(u) \\
&\quad - (1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} \big(H(t - \delta_2 - u, x) - H(s - \delta_1 - u, x)\big)^2 dF(x) d\Lambda(u) \\
&= \tilde{C}\big(\Lambda(t) - \Lambda(t - \delta_2) - (\Lambda(s) - \Lambda(s - \delta_1))\big) \\
&\quad + (1 + 2C_\rho) \int_{[0,T]} \check{G}_2(t, t - \delta_2, u) d\Lambda(u) + (1 + 2C_\rho) \int_{[0,T]} \check{G}_2(s, s - \delta_1, u) d\Lambda(u) \\
&\quad + 2(1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} [H(t - u, x) - H(t - \delta_2 - u, x)] \\
&\qquad\qquad\qquad \times [H(t - \delta_2 - u, x) - H(s - u, x)] dF(x) d\Lambda(u) \\
&\quad - 2(1 + 2C_\rho) \int_{[0,T]} \int_{\mathbb{R}^k} [H(s - u, x) - H(s - \delta_1 - u, x)] \\
&\qquad\qquad\qquad \times [H(t - u, x) - H(s - \delta_1 - u, x)] dF(x) d\Lambda(u). \tag{3.3}
\end{aligned}
$$

Note that the last two terms in (3.3) 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, due to Cauchy-Schwarz inequality. By the continuity of Λ and (2.2) in Assumption 1 that for each $T \geq t \geq 0$, $\lim_{\delta \to 0} \int_{[0,T]} \check{G}_2(t, t - \delta, u) d\Lambda(u) = 0$, we show that each term in (3.3) converges to 0 as $\delta_1 \to 0$ and $\delta_2 \to 0$. Now the continuity of V has been proved.

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

Lemma 3.3. *For $0 \leq s \leq t \leq T$,*

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

Proof. First, by direct calculations we obtain

$$\Gamma(t, t, u) + \Gamma(s, s, u) - 2\Gamma(t, s, u)$$

$$= \; 2 \sum_{l=2}^{\infty} E\Big[\big[H(t - u, Z_1) - H(s - u, Z_1)\big] \times \big[H(t - u, Z_l) - H(s - u, Z_l)\big]$$

$$- \big(G_1(t, u) - G_1(s, u)\big)^2 \Big]. \qquad (3.4)$$

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 (3.4) is upper bounded by

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

Therefore, we have

$$\Gamma(t, t, u) + \Gamma(s, s, u) - 2\Gamma(t, s, u) \quad \leq \quad 2\|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 \, C_\rho. \qquad (3.5)$$

By Lemma 3.2 and (3.5), we have

$$E\big[\big|\hat{X}_2(r) - \hat{X}_2(s)\big|^2\big]$$

$$\leq \quad \tilde{C}\big(\Lambda(t) - \Lambda(s)\big)$$

$$+ \int_{[0,s]} \Big(\|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 + 2C_\rho \|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 \Big) d\Lambda(u)$$

$$\leq \quad \tilde{C}\big(\Lambda(t) - \Lambda(s)\big) + (1 + 2C_\rho) \int_{[0,T]} \|H(t - u, Z_1) - H(s - u, Z_1)\|_2^2 d\Lambda(u),$$

which is equal to $V(s, t)$ in (3.2). $\qquad\square$

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

Lemma 3.4. *For $0 \leq r \leq s \leq t \leq T$ and any $\epsilon > 0$,*

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

Proof. We have

$$P\big(\big|\hat{X}_2(r) - \hat{X}_2(s)\big| \wedge \big|\hat{X}_2(s) - \hat{X}_2(t)\big| \geq \epsilon\big)$$

$$\leq \quad \frac{1}{\epsilon^4} E\big[\big|\hat{X}_2(r) - \hat{X}_2(s)\big|^2 \, \big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big]$$

$$\leq \quad \frac{1}{\epsilon^4} \Big(E\big[\big|\hat{X}_2(r) - \hat{X}_2(s)\big|^4\big]\Big)^{1/2} \Big(E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^4\big]\Big)^{1/2}$$

$$= \quad \frac{3}{\epsilon^4} E\big[\big|\hat{X}_2(r) - \hat{X}_2(s)\big|^2\big] E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big]$$

where the equality follows from the fact that the kurtosis of a normal random variable is 3. The claim then follows from Lemma 3.3. □

We are now ready to prove the existence theorem. Note that the ρ-mixing condition is important here since 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 Lemma 3.3 and 3.4 do not hold.

Theorem 3.3. *The centered Gaussian process $\hat{X}_2$ with covariance function given in (2.9) has continuous sample paths a.s.*

Proof. The proof is done in two steps. First, we show that $\hat{X}_2 \in \mathbb{D}$ by applying Theorem 4.1 in the Appendix (a generalization of the classical existence criterion in Theorem 13.6 in [6]). 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 (4.1): 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 3.4. (iii) The continuity property: for any $\epsilon > 0$ and $t \in [0, T)$, $\lim_{\delta \downarrow 0} P\big(\big|\hat{X}_2(t) - \hat{X}_2(t + \delta)\big| > \epsilon\big) = 0$. This is implied by the following: for all $t \in [0, T)$,

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

This follows by the formula in Lemma 3.2 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 [16]) or equivalently, it is continuous in quadratic mean since the process is Gaussian. This follows from Lemma 3.2 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

$$\Upsilon^n = \left\{ A^n : \sup_{0 \leq t \leq T} \left| \frac{A^n(t)}{n} - \Lambda(t) \right| \leq \epsilon(n) \quad \text{and} \quad \max_{1 \leq i \leq A^n(T)} |\tau_{i+1}^n - \tau_i^n| \leq \epsilon(n) \right\}, \quad (3.6)$$

where $\epsilon(n) \to 0$ as $n \to \infty$ is chosen such that $P(\Upsilon^n) \to 1$ as $n \to \infty$. 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 3.5. *The finite dimensional distributions of $\hat{X}_2^n$ converge weakly to those of $\hat{X}_2$ as $n \to \infty$.*

Proof. Fix any $0 \leq t \leq T$ and trajectory $A^n \in \Upsilon^n$ (defined in (3.6)), and we first prove that $\hat{X}_2^n(t) \Rightarrow \hat{X}_2(t)$ in $\mathbb{R}$ as $n \to \infty$. To simplify the notation, we denote

$$\mathcal{H}_i^n(t) := H(t - \tau_i^n, Z_i) - G_1(t, \tau_i^n). \quad (3.7)$$

We apply Theorem 2.1 and Proposition 2.1(a) in [57], 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): $\mathrm{Var}(S_n) \to \infty$ as $n \to \infty$ and $\sum_{l=0}^{\infty} \tilde{c}(\ell) < \infty$, where
$$\tilde{c}(\ell) = \max_{n:\ell<n} \tilde{c}_n(\ell),$$

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

We remark that conditions (i) and (ii) are required by Proposition 2.1 in [57] with $\epsilon = \gamma = 0$. The l-mixing condition in Theorem 2.1 in [57] is implied by condition (iii) since ρ-mixing is stronger than l-mixing; see page 513 in [57]. Conditions (iv) is additionally required by Theorem 2.1 in the last-mentioned reference.

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

For condition (ii), we have

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

For condition (iv), we have

$$
\begin{aligned}
\frac{\mathrm{Var}(S_n)}{n} &= \frac{1}{n}\mathrm{Var}\bigg(\sum_{j=1}^{A^n(t)} \mathcal{H}_i^n(t)\bigg) \\
&= \frac{1}{n}\sum_{i=1}^{A^n(t)} E\big[(\mathcal{H}_i^n(t))^2\big] + \frac{2}{n}\sum_{i,j=1,i<j}^{A^n(t)} E\big[\mathcal{H}_i^n(t)\mathcal{H}_j^n(t)\big] \\
&= \frac{1}{n}\sum_{i=1}^{A^n(t)} E\big[(\mathcal{H}_i^n(t))^2\big] + \frac{2}{n}\sum_{i,j=1,i<j}^{A^n(t)} r_{i,j}^n, \qquad (3.8)
\end{aligned}
$$

where

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

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

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

where

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

Observe that by the Cauchy–Schwarz inequality, we have

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

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 (2.2) (recall Assumption 1), for each fixed ℓ and $t \ge 0$, we obtain that

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

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 $n \to \infty$,

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

Thus, by (3.8), we obtain

$$\frac{\text{Var}(S_n)}{n} \to \sigma^2 := \int_0^t [\tilde{G}(t, u) + \Gamma(t, t, u)]d\Lambda(u) \quad \text{as} \quad n \to \infty. \tag{3.9}$$

We have verified conditions (i)–(iv), and by Theorem 2.1 and Proposition 2.1(a) in [57], we have

$$S_n/\sqrt{\text{Var}(S_n)} \Rightarrow N(0,1) \quad \text{as} \quad n \to \infty. \tag{3.10}$$

By definition of S_n, we have $S_n = \sqrt{n}\hat{X}_2^n(t)$. Thus, by (3.9) and (3.10), we have

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

where "$\stackrel{d}{=}$" denotes "equal in distribution".

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

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

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

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

$$= \quad \sum_{i=1}^{A^n(t_2)} \mathcal{H}_i^{n,*},$$

where

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

Since 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. $\qquad \square$

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

$$\omega_\delta(f) := \sup_{|x-y| \le \delta, \ x,y \in [0,T]} |f(x) - f(y)|, \quad \delta > 0. \tag{3.11}$$

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

$$\hat{D}^n(x) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big(H(x, Z_i) - E[H(x, Z_1)] \Big), \quad x \in [0, T]. \tag{3.12}$$

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

Lemma 3.6. *Under Assumptions 1–2, for each $\eta > 0$,*

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P\Big(\omega_\delta(\hat{D}^n) > \eta\Big) = 0. \tag{3.13}$$

Proof. We prove that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \Big\| \sup_{\varrho(x,y) \le \delta} |\hat{D}^n(x) - \hat{D}^n(y)| \Big\|_2 = 0, \tag{3.14}$$

where

$$\varrho(x, y) := \|H(x, Z_1) - H(y, Z_1)\|_2, \quad \text{for} \quad x, y \in [0, T]. \tag{3.15}$$

This implies that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E\big[\omega_\delta(\hat{D}^n)^2\big] = 0. \tag{3.16}$$

To see that, suppose (3.14) holds, then for any $\epsilon > 0$, there exits $\delta_0 > 0, n_0 \in \mathbb{N}$ such that for all $n \ge n_0$,

$$\Big\| \sup_{\varrho(x,y) \le \delta_0} |\hat{D}^n(x) - \hat{D}^n(y)| \Big\|_2 \le \epsilon. \tag{3.17}$$

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,

$$\left\| \sup_{|x-y| \leq \delta_1} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 \leq \left\| \sup_{\varrho(x,y) \leq \delta_0} |\hat{D}^n(x) - \hat{D}^n(y)| \right\|_2 \leq \epsilon. \tag{3.18}$$

We proceed to prove (3.14). Define the bracket number $N(\delta, T)$ as the minimal number of N that there exist points $t_1, ..., t_N \in [0, T]$ such that for each $t \in [0, T]$, there exists a t_i for which $\varrho(t, t_i) \leq \delta$. To show (3.14), we apply Theorem 2.2 in [1] (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 $t \in [0, T]$. Since $H \in \mathbb{D}$ is nondecreasing, then the function g is also nondecreasing. Thus, for a given $\delta > 0$, we can choose a partition of $[0, T]$ by $0 = t_0 < t_1 < ... < t_N = T$ such that $g(t_{i+1}-) - g(t_i) \leq \delta^2$ for every i. Then, for each i and $s \in [t_i, t_{i+1})$,

$$\varrho(s, t_i)^2 \leq E[(H(s, Z_1) - H(t_i, Z_1))H(T, Z_1)] \leq g(t_{i+1}-) - g(t_i) \leq \delta^2. \tag{3.19}$$

Since $g \in \mathbb{D}$, the number of jump points with jump size greater than δ^2 is finite; see Theorem 12.2.1 in [60]. Therefore, after making sure that above mentioned jump points are among $t_0, ..., t_N$, 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 3.7. *Under Assumptions 1–2,*

$$\sup_{n \in \mathbb{N}} E\left[\sup_{x \in [0,T]} \hat{D}^n(x)^2 \right] < \infty. \tag{3.20}$$

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

$$\sup_{n \in \mathbb{N}} E\left[\sup_{x \in [0,T]} \hat{D}^n(x)^2 \right] \leq \max_{1 \leq n \leq n_0} E\left[\sup_{x \in [0,T]} \hat{D}^n(x)^2 \right] + \sup_{n \in \mathbb{N}} \sum_{i=1}^{[T/\delta]+1} E[\hat{D}^n(x_i)^2]$$
$$+ \sup_{n \geq n_0} E[\omega_\delta(\hat{D}^n)^2]. \tag{3.21}$$

The first summand is the maximum over finite many finite values. For the second summand, by similar calculation and arguments in (3.9), we conclude that it is finite. The finiteness of the last summand (when n_0 large enough) is an immediate consequence of Lemma 3.6. □

Lemma 3.8. *Under Assumptions 1–2,*

$$\hat{D}^n \Rightarrow \hat{D} \quad in \quad (\mathbb{D}, J_1) \quad as \quad n \to \infty, \tag{3.22}$$

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

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

$$+ \left(E\big[H(x, Z_j)H(y, Z_1)\big] - \int_{\mathbb{R}^k} H(x, z)dF(z) \int_{\mathbb{R}^k} H(y, z)dF(z) \right) \Bigg].$$

Proof. The tightness property is proved in Lemmas 3.6 and 3.7. The convergence of finite dimensional distributions can be done by a similar but simpler argument as in the proof of Lemma 3.5, 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:

$$E\big[|\hat{D}(x) - \hat{D}(y)|^2\big] = Var\big(H(x, Z_1) - H(y, Z_1)\big)$$

$$+ 2\sum_{j=2}^{\infty} \Bigg[E\Big[\big[H(x, Z_1) - H(y, Z_1)\big]\big[H(x, Z_j) - H(y, Z_j)\big]\Big]$$

$$- \Big(\int_{\mathbb{R}^k} \big(H(x, z) - H(y, z)\big)dF(z) \Big)^2 \Bigg]. \tag{3.23}$$

By Assumption 2, we obtain that

$$E\big[|\hat{D}(x) - \hat{D}(y)|^2\big] \leq (1 + 2C_\rho)E\Big[\big[H(x, Z_1) - H(y, Z_1)\big]^2\Big]. \tag{3.24}$$

Since H is nondecreasing, then we can apply Theorem 13.6 in [6] to prove the existence in $\mathbb{D}$, and thus in $\mathbb{C}$ by the continuity in quadratic mean. This completes the proof. $\square$

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

$$\hat{V}^n(t, x) := \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} \big(H(x, Z_i) - E[H(x, Z_1)]\big), \quad t, x \geq 0. \tag{3.25}$$

Lemma 3.9. *Under Assumptions 1–3, for any $\epsilon > 0$,*

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P\left(\sup_{t \in [0,T]} \sup_{|x-y| \leq \delta, x,y \in [0,T]} \big|\hat{V}^n(t, x) - \hat{V}^n(t, y)\big| > \epsilon \right) = 0. \tag{3.26}$$

Proof. Recall the definition of $\hat{D}^n$ in (3.12). Observe that

$$\sup_{t \in [0,T]} \sup_{|x-y| \leq \delta, x,y \in [0,T]} \big|\hat{V}^n(t, x) - \hat{V}^n(t, y)\big|$$

$$\leq \max_{\ell \leq A^n(T)} \sup_{|x-y| \leq \delta \ x,y \in [0,T]} \big|\hat{V}^n(\tau_\ell^n, x) - \hat{V}^n(\tau_\ell^n, y)\big|$$

$$= \max_{\ell \leq A^n(T)} \sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell). \tag{3.27}$$

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

$$L_i(x, y) := H(x, Z_i) - E[H(x, Z_1)] - H(y, Z_i) + E[H(y, Z_1)]. \tag{3.28}$$

Let $l_n = [n^{(5-\kappa)/8}] = o(\sqrt{n})$ for $\kappa > 1$. Then, applying the Ottaviani-type inequality under strong α-mixing in Lemma 4.1 (noting $\alpha_n \leq \rho_n/4$, see [7], and also [2, Lemma 3]), we obtain

$$P\left(\max_{\ell \leq A^n(T)} \sqrt{\frac{\ell}{n}} \omega_\delta(\hat{D}^\ell) > 3\epsilon \right) \leq \frac{B_{n,1}(\delta) + B_{n,2}(\delta) + \frac{1}{4}[A^n(T)/l_n]\rho_{l_n}}{1 - \max_{\ell=1,\ldots,A^n(T)} P\left(\sqrt{\ell/n}\, \omega_\delta(\hat{D}^\ell) > \epsilon \right)}, \tag{3.29}$$

where

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

and

$$B_{n,2}(\delta) := P\left(\max_{\substack{p,q\in\{1,\ldots,A^n(T)\},\\0<q-p\leq 2l_n}}\sup_{\substack{x,y\in[0,T],\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}L_i(x,y)\right|>\epsilon\right). \tag{3.30}$$

First, by the trajectories of A^n defined in (3.6),

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

Thus, by Lemma 3.6,

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

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

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

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

$$L_i(x,y)=L_{i,M}(x,y)+L_{i,M^c}(x,y), \tag{3.31}$$

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

$$\max_{\substack{p,q\in\{1,\ldots,A^n(T)\}\\0<q-p\leq 2l_n}}\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}L_i(x,y)\right|$$

$$\leq\quad\max_{\substack{p,q\in\{1,\ldots,A^n(T)\}\\0<q-p\leq 2l_n}}\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}L_{i,M}(x,y)\right|$$

$$+\max_{\substack{p,q\in\{1,\ldots,A^n(T)\}\\0<p-q\leq 2l_n}}\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}L_{i,M^c}(x,y)\right|$$

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

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

$$P\left(\max_{\substack{p,q\in 1,\ldots,A^n(T)\\0<q-p\leq 2l_n}}\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}L_{i,M^c}(x,y)\right|>\epsilon/2\right) \tag{3.32}$$

$$\leq \ P\Bigg(\max_{\substack{p,q\in 1,\ldots,A^n(T)\\0<q-p\leq 2l_n}}\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\sum_{i=p+1}^{q}|L_{i,M^c}(x,y)|>\epsilon/2\Bigg)\quad\text{(triangle inequality)}$$

$$\leq \ 2l_n(A^n(T)-2l_n)P\Bigg(\sup_{\substack{x,y\in[0,T]\\|x-y|\leq\delta}}\frac{1}{\sqrt{n}}\sum_{i=1}^{2l_n}|L_{i,M^c}(x,y)|>\epsilon/2\Bigg)\quad\text{(by stationarity of $\{Z_i\}$)}$$

$$\leq \ 2l_n(A^n(T)-2l_n)P\Bigg(\frac{2}{\sqrt{n}}\sum_{i=1}^{2l_n}|H(T,Z_i)+E[H(T,Z_i)]|\,\mathbf{1}(|H(T,Z_i)|>M/2)>\epsilon/2\Bigg).$$

$$\text{(by definition of L_i)}$$

We then apply Theorem 4.1 in [53]. Choose $p=4, r=8$ in their notation, then by (4.4) in [53],

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

where K is a constant since 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 (3.32) 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 (3.43), we need to show that the denominator in (3.29) is bounded above from zero for large enough n and small enough δ. By Lemma 3.6, there exist $\delta_0>0$ and $n_0=n_0(\delta_0)$ such that $P\big(\omega_\delta(\hat{D}_n)>\epsilon\big)<1/2$ for all $\delta<\delta_0$ and $n>n_0$. Thus, for $\delta<\delta_0$,

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

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

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

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

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

Thus,

$$\max_{\ell=1,\ldots,n_0-1}P\Big(\sqrt{\frac{\ell}{n}}\,\omega_\delta(\hat{D}^\ell)>2\epsilon\Big)$$

$$\leq \ \max_{\ell=1,\ldots,n_0-1}P\Big(\sqrt{\frac{\ell}{n}}\,\sup_{x\in[0,T]}\big|\hat{D}_M^\ell(x)\big|>\epsilon/2\Big)$$

$$+\max_{\ell=1,\ldots,n_0-1}P\Big(\sqrt{\frac{\ell}{n}}\,\sup_{x\in[0,T]}\big|\hat{D}^\ell(x)-\hat{D}_M^\ell(x)\big|>\epsilon/2\Big).$$

Note that $\sup_{x\in[0,T]}\left|\hat{D}^{\ell}_M(x)\right| \leq \sqrt{\ell}Mn_0$. Thus the first term on the right hand side converges to 0 as $n \to \infty$. The second term is upper bounded by

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

Therefore, the denominator in (3.29) is bounded above from zero for large enough n and small enough δ. This completes the proof of the lemma. $\qquad\square$

Lemma 3.10. *Under Assumptions 1–3, for any $\epsilon > 0$,*

$$\lim_{\delta\downarrow0}\limsup_{n\to\infty} P\left(\sup_{\substack{|s-t|\leq\delta\\s,t\in[0,T]}}\sup_{x\in[0,T]}\left|\hat{V}^n(t,x)-\hat{V}^n(s,x)\right| > \epsilon\right) = 0. \tag{3.33}$$

Proof. It suffices to show that

$$\lim_{\delta\downarrow0}\limsup_{n\to\infty} P\left(\max_{\substack{0\leq j\delta\leq T\\j\in\mathbb{N}}}\sup_{j\delta\leq t\leq(j+1)\delta}\sup_{x\in[0,T]}\left|\hat{V}^n(t,x)-\hat{V}^n(j\delta,x)\right| > 3\epsilon\right) = 0.$$

The probability in this equation can be bounded by

$$\sum_{j=1}^{[T/\delta]+1} P\left(\sup_{j\delta\leq t\leq(j+1)\delta}\sup_{x\in[0,T]}\left|\hat{V}^n(t,x)-\hat{V}^n(j\delta,x)\right| > 3\epsilon\right)$$

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

Applying the Ottaviani-type inequality in Lemma 4.1 (Lemma 3 in [2]), we obtain the quantity above is further bounded by

$$\sum_{j=1}^{[T/\delta]+1} \left(C_{j,n,1}(\delta) + C_{j,n,2}(\delta) + \frac{1}{4}[(A^n((j+1)\delta)-A^n(j\delta))/l_n]\rho_{l_n}\right)/\left(1-C_{j,n,3}\right) \tag{3.34}$$

where

$$C_{j,n,1}(\delta) := P\left(\sup_{x\in[0,T]}\frac{1}{\sqrt{n}}\left|\sum_{i=A^n(j\delta)+1}^{A^n((j+1)\delta)}\left(H(x,Z_i)-E[H(x,Z_i)]\right)\right| > \epsilon\right), \tag{3.35}$$

$$C_{j,n,2}(\delta) := P\left(\max_{\substack{p,q\in\{A^n(j\delta)+1,\dots,A^n((j+1)\delta)\}\\0<q-p<2l_n}}\sup_{x\in[0,T]}\frac{1}{\sqrt{n}}\left|\sum_{i=p+1}^{q}\left(H(x,Z_i)-E[H(x,Z_i)]\right)\right| > \epsilon\right), \tag{3.36}$$

and

$$C_{j,n,3}(\delta) := \max_{\ell=A^n(j\delta)+1,\dots,A^n((j+1)\delta)} P\left(\sup_{x\in[0,T]}\frac{1}{\sqrt{n}}\left|\sum_{i=\ell}^{A^n((j+1)\delta)}\left(H(x,Z_i)-E[H(x,Z_i)]\right)\right| > \epsilon\right). \tag{3.37}$$

Recall the trajectories of $\{A^n(t): t \geq 0\}$ defined in (3.6).

We first treat $C_{j,n,1}(\delta)$ in (3.35). Recall $\hat{D}^n$ defined in (3.12). 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 (3.6), Lemma 3.8 and the Portmanteau Theorem, we obtain

$$
\begin{aligned}
\limsup_{n\to\infty} C_{j,n,1}(\delta) &\leq P\left(\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)}\right) \sup_{x\in[0,T]} |\hat{D}(x)| > \epsilon \right) \\
&= P\left(\sup_{x\in[0,T]} |\hat{D}(x)| > \epsilon/\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)}\right) \right). \quad (3.38)
\end{aligned}
$$

Since $\sup_{x\in[0,T]} |\hat{D}(x)|$ possesses moments of any order, the probability above converges to 0 faster than any order of δ by the continuity of Λ and Proposition A.2.3 in [56]. Therefore we have

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \sum_{j=1}^{[T/\delta]+1} C_{j,n,1} \leq \lim_{\delta\downarrow 0} \sum_{j=1}^{[T/\delta]+1} P\left(\sup_{x\in[0,T]} |\hat{D}(x)| > \epsilon/\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)}\right) \right) = 0.$$
$$(3.39)$$

Next we turn to $C_{j,n2}(\delta)$ in (3.36). Recall the definition of $L_i(x,y)$ below (3.27). We have

$$C_{j,n,2}(\delta) = P\left(\max_{\substack{p,q\in\{A^n(j\delta)+1,...,A^n((j+1)\delta)\}\\ 0<q-p<2l_n}} \sup_{x\in[0,T]} \frac{1}{\sqrt{n}} \left| \sum_{i=p+1}^{q} L_i(x,0)] \right| > \epsilon \right). \quad (3.40)$$

Then $\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \sum_{j=1}^{[T/\delta]+1} C_{j,n,2} = 0$ follows from similar arguments for B_{n2} from (3.31) to (3.32).

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

$$
\begin{aligned}
&\left| \sum_{j=1}^{[T/\delta]+1} [(A^n((j+1)\delta) - A^n(j\delta))/l_n]\rho_{l_n} \right| \\
&\leq \left(\sum_{j=1}^{[T/\delta]+1} (\Lambda((j+1)\delta) - \Lambda(j\delta)) \right) n\rho_{l_n}/l_n + 2([T/\delta]+1)\epsilon(n)n\rho_n/l_n \\
&= \left(\Lambda([T/\delta]+2) - \Lambda(\delta) + 2([T/\delta]+1)\epsilon(n) \right) n\rho_n/l_n \\
&= o(1), \quad (3.41)
\end{aligned}
$$

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

Given (3.39), (3.40) and (3.41), we have

$$\sum_{j=1}^{[T/\delta]+1} \left(C_{j,n,1}(\delta) + C_{j,n,2}(\delta) + \frac{1}{4}[(A^n((j+1)\delta) - A^n(j\delta))/l_n]\rho_{l_n} \right) = 0. \quad (3.42)$$

To prove (3.34), it now suffices to show that there exists $\epsilon_0 > 0$ such that for sufficiently large n, small δ and each $1 \leq j \leq [T/\delta] + 1$, we have $C_{j,n,3}(\delta) \leq 1 - \epsilon_0$.

Recall the definition of $C_{j,n,3}$ in (3.37). 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

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

On the other hand, by similar argument as (3.38), we have from the Portmanteau Theorem that for $\ell \geq A^n(j\delta) + 1$,

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

$$\leq \quad P\left(\left(\sqrt{\Lambda((j+1)\delta)} - \sqrt{\Lambda(j\delta)} \right) \sup_{x \in [0,T]} |\hat{D}(x)| > \epsilon \right)$$

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

$$< \quad 1.$$

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

Lemma 3.11. *The sequence $\{\hat{X}_2^n : n \geq 1\}$ is tight in space $\mathbb{D}$.*

Proof. By triangle inequality, we have

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

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

Then the proof is complete by Lemmas 3.9–3.10. □

Remark 3.1. As can be seen from the proof, tightness also holds under the strong α-mixing condition since the moment bounds and the Ottavian-type inequality we use are established under the α-mixing condition. However, as mentioned above, the proof of existence of the limit process $\hat{X}_2$ in $\mathbb{C}$ in Theorem 3.3 can be carried out only under the ρ-mixing condition. In the i.i.d. case, the tightness proof is done in [42] by applying the new sufficient convergence criterion in $(\mathbb{D}, J_1)$ (extending the classical result in Theorem 13.5 of [6]) 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 [6], 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 3.4 cannot be established for the scaled processes $\hat{X}_2^n$.

3.4. **Completing the proof of Theorem 2.2.**

Proof of Theorem 2.2. Recall the set Υ^n of trajectories $\{A^n(t) : t \geq 0\}$ defined in (3.6), 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 3.5 and tightness in Lemmas 3.9 and 3.10.

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 $\vartheta \in \mathbb{R}$,

$$E\left[e^{i\vartheta(\hat{X}_1^n(t)+\hat{X}_2^n(t))}\right] = E\left[E\left[e^{i\vartheta(\hat{X}_1^n(t)+\hat{X}_2^n(t))}\Big|\mathcal{A}^n(t)\right]\right]$$

$$= E\left[E\left[e^{i\vartheta\hat{X}_1^n(t)}e^{i\vartheta\hat{X}_2^n(t)}\Big|\mathcal{A}^n(t)\right]\right] = E\left[e^{i\vartheta\hat{X}_1^n(t)}E\left[e^{i\vartheta\hat{X}_2^n(t)}\Big|\mathcal{A}^n(t)\right]\right].$$

On Υ^n, by the convergence of $\hat{X}_1^n$, and by the proof of Lemma 3.5 for the convergence of finite-dimensional distributions of $\hat{X}_2^n$, we obtain that the above equation converges to $E\left[e^{i\vartheta(\hat{X}_1(t)+\hat{X}_2(t))}\right]$. The convergence of finite-dimensional distributions follows similarly. Since $P(\Upsilon^n) \to 1$ as $n \to \infty$, the proof is complete. $\qquad\square$

4. Appendix

4.1. **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 [6]. That criterion requires a probability bound for the increments of the processes where the upper bound involves a finite measure. Its proof relies on the classical maximal inequalities in Theorems 10.3 and 10.4 in [6]. 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 [42]. We now state this criterion of existence in the following theorem. It is used in the proof of Theorem 3.3.

Recall the definition of a set function with the superadditivity property in Definition 3.2.

Theorem 4.1 (Theorem 5.3 in [42]). *There exists a random element X in $\mathbb{D}([0,T],\mathbb{R})$ with finite-dimensional distributions $\pi_{t_1,\cdots,t_k}$ for any $0 \leq t_1 < \cdots < t_k \leq T$, that is, $\pi_{t_1,\cdots,t_k}(x_1,\ldots,x_k) = P(X(t_1) \leq x_1,\ldots,X(t_k) \leq x_k)$ for $x_i \in \mathbb{R}$, $i = 1,...,k$, if the following conditions are satisfied:*

(i) *the finite dimensional distributions $\pi_{t_1,\cdots,t_k}$ are consistent, satisfying the conditions of Kolmogorov's existence theorem;*

(ii) *for any $0 \leq r \leq s \leq t \leq T$ and $\epsilon > 0$,*

$$P\big(|X(r) - X(s)| \wedge |X(s) - X(t)| \geq \epsilon\big) \leq \frac{C}{\epsilon^{4\beta}}\big(\mu(r,t]\big)^{2\alpha}, \qquad (4.1)$$

for some $\beta \geq 0$ and $\alpha > 1/2$, where C is a positive constant, μ is a finite set function in Definition 3.2 and $\mu(0,t]$ is continuous in t;

(iii) *for any $\epsilon > 0$ and $t \in [0,T)$,*

$$\lim_{\delta\downarrow 0} P\big(|X(t) - X(t+\delta)| > \epsilon\big) = 0. \qquad (4.2)$$

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 [7] 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 4.1 (Lemma 3 in [2])**.** *Suppose that* $\|S_m - S_n\|$ *is measurable for each* $0 \le n < m$. *Then, for each* $\epsilon > 0$ *and* $1 \le k < n$,

$$P\Big(\max_{j=1,\dots,n} \|S_j\| > 3\epsilon \Big) \times \Big(1 - \max_{j=1,\dots,n} P(\|S_n - S_j\|) > \epsilon \Big)$$

$$\le P(\|S_n\| > \epsilon) + P\Big(\max_{j,l \in \{1,\dots,n\}, \, 0 < l-j \le 2k} \|S_l - S_j\| > \epsilon \Big) + \lfloor n/k \rfloor \times \alpha_k. \qquad (4.3)$$

Acknowledgement

References

[1] D. W. Andrews and D. Pollard. (1994) An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review/Revue Internationale de Statistique.* 62(1), 119–132.

[2] A. Bücher. (2015) A note on weak convergence of the sequential multivariate empirical process under strong mixing. *Journal of Theoretical Probability.* 28(3), 1028-1037.

[3] R. J. Adler and J. E. Taylor. (2009) *Random Fields and Geometry.* Springer Science & Business Media.

[4] H. Biermé and A. Desolneux. (2012) Crossings of smooth shot noise processes. *Annals of Applied Probability.* 22(6), 2240–2281.

[5] L. Bondesson. (1988). Shot noise processes and distributions. *Encyclopedia of Statistical Sciences.*

[6] P. Billingsley. (1999) *Convergence of Probability Measures.* John Wiley & Sons.

[7] R. C. Bradley (2005). Basic properties of strong mixing conditions: A survey and some open questions. *Probability Surveys.* 2(2), 107–144.

[8] A. Budhiraja and P. Nyquist. (2015) Large deviations for multidimensional state-dependent shot noise processes. *Journal of Applied Probability.* 52(4), 1097–1114.

[9] G. Chobanov (1999) Modeling financial asset returns with shot noise processes. *Mathematical and Computer Modelling.* 29(10-12), 17–21.

[10] J. P. Conte and B. F. Peng (1996) Nonstationary earthquake ground motion model. *In Proceedings of the 11th World Conference on Earthquake Engineering,* Acapulco, Mexico. Paper (No. 301).

[11] H. Dehling and W. Philipp. (2002) *Empirical Process Techniques for Dependent Data.* Birkhäuser, Boston.

[12] R. A. Doney and G. L. O'Brien. (1991) Loud shot noise. *Annals of Applied Probability.* 1(1), 88–103.

[13] S. N. Ethier and T. G. Kurtz. (1986) *Markov Processes: Characterization and Convergence.* John Wiley & Sons.

[14] A. Ganesh, C. Macci and G. L. Torrisi. (2005) Sample path large deviations principles for Poisson shot noise processes and applications. *Electronic Journal of Probability.* Vol. 10, 1026–1043.

[15] A. Ganesh and G. L. Torrisi. (2006) A class of risk processes with delayed claims: ruin probability estimates under heavy tail conditions. *Journal of Applied Probability.* 43(4), 916–926.

[16] M. G. Hahn. (1978) Central limit theorem in $D[0,1]$. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete.* 44(2), 89–101.

[17] L. Heinrich and V. Schmidt. (1985) Normal convergence of multidimensional shot noise and rates of this convergence. *Advances in Applied Probability.* 17(4), 709–730.

[18] F. W. Huffer. (1987) Inequalities for the $M/G/\infty$ queue and related shot noise processes. *Journal of Applied Probability.* 24(4), 978–989.

[19] I. A. Ibragimov (1959). Some limit theorems for stochastic processes stationary in the strict sense. *Dokl. Akad. Nauk SSSR* (125(711–714), p. 26).

[20] D. L. Iglehart. (1973) Weak convergence of compound stochastic process, I. *Stochastic Processes and their Applications.* 1(1), 11–31.

[21] A. Iksanov. (2013) Functional limit theorems for renewal shot noise processes with increasing response functions. *Stochastic Processes and their Applications.* 123(6), 1987–2010.

[22] A. Iksanov. (2016) *Renewal Theory for Perturbed Random Walks and Similar Processes*. Birkhäuser.

[23] A. Iksanov, W. Jedidi and F. Bouzeffour. (2018) Functional limit theorems for the number of busy servers in a $G/G/\infty$ queue. *Journal of Applied Probability*. 55(1), 15–29.

[24] A. Iksanov, A. Marynych and M. Meiners. (2014) Limit theorems for renewal shot noise processes with eventually decreasing response functions. *Stochastic Processes and their Applications*. 124(6), 2132–2170.

[25] A. Iksanov, A. Marynych and M. Meiners. (2017a) Asymptotics of random processes with immigration I: Scaling limits. *Bernoulli*. 23(2), 1233–1278.

[26] A. Iksanov, A. Marynych and M. Meiners. (2017b) Asymptotics of random processes with immigration II: Convergence to stationarity. *Bernoulli*. 23(2), 1279–1298.

[27] P. A. Jacobs and P. A. Lewis (1978). Discrete time series generated by mixtures. I: Correlational and runs properties. *Journal of the Royal Statistical Society. Series B (Methodological)*. 40(1), 94–105.

[28] A.N. Kolmogorov and Y.A. Rozanov. (1960). On strong mixing conditions for stationary Gaussian processes. *Theory of Probability & Its Applications*. 5(2), 204–208.

[29] D. Koops, M. Mandjes and O. Boxma. (2017) Networks of $\cdot/G/\infty$ server queues with shot-noise-driven arrival intensities. *Queueing Systems*. 86(3-4), 301–325.

[30] C. Klüppelberg and T. Mikosch. (1995) Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli*. 1(1-2), 125–147.

[31] C. Klüppelberg, T. Mikosch and A. Schärf. (2003) Regular variation in the mean and stable limits for Poisson shot noise. *Bernoulli*. 9(3), 467–496.

[32] C. Klüppelberg and C. Kühn. (2004) Fractional Brownian motion as a weak limit of Poisson shot noise processes - with applications to finance. *Stochastic Processes and their Applications*. 113(2), 333–351.

[33] T. G. Kurtz. (1996) Limit theorems for workload input models. *Stochastic Networks: Theory and Applications*, F. P Kelly, S. Zachary, I. Ziedins, eds. Royal Statistical Society Lecture Note Series 4. Oxford Science Publications, Oxford, UK.

[34] E. V. Krichagina and A. A. Puhalskii. (1997) A heavy-traffic analysis of a closed queueing system with a GI/∞ service center. *Queueing Systems*. Vol. 25, 235–280.

[35] J. A. Lane. (1984) The central limit theorem for the Poisson shot-noise process. *Journal of Applied Probability*. 21(2), 287–301.

[36] G. Pang and M. S. Taqqu. (2019) Non-stationary self-similar Gaussian processes as scaling limits of power-law shot noise processes and generalizations of fractional Brownian motion. *High Frequency*. Vol. 2, No. 2, 95–112.

[37] G. Pang and W. Whitt. (2010) Two-parameter heavy-traffic limits for infinite-server queues. *Queueing Systems*. Vol. 65, 325–364.

[38] G. Pang and W. Whitt. (2012) Infinite-server queues with batch arrivals and dependent service times. *Probability in Engineering and Informational Sciences*. 26(2), 197–220.

[39] G. Pang and W. Whitt. (2013). Two-parameter heavy-traffic limits for infinite-server queues with dependent service times. *Queueing Systems*. 73(2), 119–146.

[40] G. Pang and Y. Zhou. (2015) $G/G/\infty$ queues with renewal alternating interruptions. *Advances in Applied Probability*. 48(3), 812–831.

[41] G. Pang and Y. Zhou. (2017a) Two-parameter process limits for an infinite-server queue with arrival dependent service times. *Stochastic Processes and their Applications*. 127(5), 1375–1416.

[42] G. Pang and Y. Zhou. (2018) Functional limit theorems for a new class of non-stationary shot noise processes. *Stochastic Processes and their Applications*. 128(2), 505–544.

[43] G. Pang and Y. Zhou. (2018) Two-parameter process limits for infinite-server queues with dependent service times via chaining bounds. *Queueing Systems*. 88(1-2), 1–25.

[44] A. Papoulis. (1971) High density shot noise and Gaussianity. *Journal of Applied Probability*. 18(1), 118–127.

[45] E. Rio (2017). *Asymptotic Theory of Weakly Dependent Random Processes*. Springer Berlin.

[46] F. Ramirez-Perez and R. Serfling. (2001) Shot noise on cluster processes with cluster marks, and studies of long range dependence. *Advances in Applied Probability*. 33(3), 631–651.

[47] J. Rice. (1977) On generalized shot noises. *Advances in Applied Probability*. 9(3), 553–565.

[48] M. Rosenblatt. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences*. 42(1), 43–47.

[49] G. Samorodnitsky (1996). A class of shot noise models for financial applications. In Athens Conference on Applied Probability and Time Series Analysis (pp. 332–353). Springer, New York.

[50] G. Samorodnitsky. (1998) Tail behavior of some shot noise processes. In *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* Editors: R. J. Adler, R. E. Feldman and M. S. Taqqu. Birkhäuser, Basel.

[51] T. Schmidt (2014). Catastrophe insurance modeled by shot-noise processes. *Risks*, 2(1), 3–24.

[52] V. Schmidt. (1985) On finiteness and continuity of shot noise processes. *Optimization.* 16, 921–933.

[53] Q. M. Shao and H. Yu (1996). Weak convergence for weighted empirical processes of dependent sequences. *Annals of Probability.* 24(4), 2098–2127.

[54] E. A. Valdez and K. Mo (2002). Ruin probabilities with dependent claims. Working paper, UNSW, Sydney, Australia.

[55] W. B. Wu (2007). Strong invariance principles for dependent random variables. *Annals of Probability.* 35(6), 2294–2320.

[56] A. W. van der Vaart, J. A. Wellner. (1996) *Weak Convergence and Empirical Process with Applications to Statistics.* Springer, New York.

[57] C. S. Withers. (1981). Central limit theorems for dependent variables. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete.* 57(4), 509–534.

[58] W. Whitt. (1976) Bivariate distributions with given marginals. *Annals of Statistics.* 4(6), 1280–1289.

[59] W. Whitt. (1983) Comparing batch delays and customer delays. *Bell System Tech J.* 62(7), 2001–2009.

[60] W. Whitt. (2002) *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and Their Applications to Queues.* Springer.