Functional Limit Theorems for A New Class of Non-Stationary Shot Noise Processes

GUODONG PANG AND YUHANG ZHOU

Abstract. We study a class of non-stationary shot noise processes which have a general arrival process of noises with non-stationary arrival rate and a general shot shape function. Given the arrival times, the shot noises are conditionally independent and each shot noise has a general (multivariate) cumulative distribution function (c.d.f.) depending on its arrival time. We prove a functional weak law of large numbers and a functional central limit theorem for this new class of non-stationary shot noise processes in an asymptotic regime with a high intensity of shot noises, under some mild regularity conditions on the shot shape function and the conditional (multivariate) c.d.f. We discuss the applications to a simple multiplicative model (which includes a class of non-stationary compound processes and applies to insurance risk theory and physics) and the queueing and work-input processes in an associated non-stationary infinite-server queueing system. To prove the weak convergence, we show new maximal inequalities and a new criterion of existence of a stochastic process in the space $\mathbb{D}$ given its consistent finite dimensional distributions, which involve a finite set function with the superadditive property.

1. Introduction

We consider a class of non-stationary shot noise processes $X := \{X(t) : t \geq 0\}$ described as follows. Let $A := \{A(t) : t \geq 0\}$ be a counting process with arrival times $\{\tau_i : i \in \mathbb{N}\}$. Let $\{Z_i : i \in \mathbb{N}\}$ be a sequence of conditionally independent $\mathbb{R}^k$-valued ($k \geq 1$) random vectors given the event times $\{\tau_i : i \in \mathbb{N}\}$. For each $i \in \mathbb{N}$, the distribution of Z_i depends on τ_i only. To indicate the dependence of Z_i on τ_i explicitly, we write $Z_i(\tau_i)$ for Z_i. The regular conditional probability for $Z_i(\tau_i)$ given that $\tau_i = t$, $t \geq 0$, is given by

$$P(Z_i(\tau_i) \leq x | \tau_i = t) = F_t(x), \quad t \geq 0, \quad x \in \mathbb{R}^k, \tag{1.1}$$

where for two vectors $x = (x_1, ..., x_k), y = (y_1, ..., y_k) \in \mathbb{R}^k$, $x \leq y$ means $x_i \leq y_i$ for each $i = 1, ..., k$ and $F_t(\cdot)$ is a joint/multivariate cumulative distribution function (c.d.f.) for each $t \geq 0$. Let $H : \mathbb{R}_+ \times \mathbb{R}^k \to \mathbb{R}$ be a deterministic measurable function representing the shot shape or the (impulse) response function. See the precise assumptions on $F_t(\cdot)$ and H in Assumption 2. Define the non-stationary shot noise process X by

$$X(t) := \sum_{i=1}^{A(t)} H(t - \tau_i, Z_i(\tau_i)), \quad t \geq 0. \tag{1.2}$$

Date: First version: August 30, 2016; This version: June 15, 2018.

Key words and phrases. Shot noise processes, functional weak law of large numbers, functional central limit theorem, Gaussian limit, non-stationarity, Skorohod J_1 topology, weak convergence, maximal inequalities, criterion of existence in the space $\mathbb{D}$.

The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA 16802 (`gup3@psu.edu` and `yxz197@psu.edu`).

In the literature the sequence of random variables $\{Z_i\}$ is often assumed to be i.i.d., independent of the arrival processes of shot noises (see, e.g., [9, 12, 13, 18, 27, 28, 29, 31, 40]). Limited work has studied for the sequence $\{Z_i\}$ with certain dependence structures. For example, in [33], $\{Z_i\}$ is modulated by a finite-state Markov chain and is conditionally independent with a distribution depending on the state of the chain at the arrival time of the shot noise (also modulated by the same chain). In [38], a cluster shot noise model is studied where $\{Z_i\}$ depends the same 'cluster mark' within each cluster. However, the 'non-stationarity' of shot noises has been neither explicitly modeled nor adequately studied, although it often occurs in stochastic systems (see, e.g., [2, 8, 15, 35, 36, 45] and references therein). In our model, the sequence $\{Z_i\}$ is assumed to be conditionally independent given the arrival times and the distribution depends upon the arrival times. We have explicitly modeled "non-stationarity" in the distribution of shot noises. In addition, the arrival process is also allowed to be a general non-stationary point process.

In this paper, we establish the functional weak law of large numbers (FWLLN) and functional central limit theorems (FCLTs) for this class of non-stationary shot noise processes in an asymptotic regime where the arrival rate is large while fixing the shot noise distributions $F_t(x)$ and shot shape function H (see Assumptions 1 and 2). (It is often referred to as the "high intensity/density regime" [4, 17, 19, 37].) Here we assume that the arrival process satisfies an FCLT with a continuous limiting process and a non-stationary arrival rate function. In the FCLT, we obtain a non-stationary stochastic process limit for the diffusion-scaled shot noise process (Theorem 2.2). The limit can be written as a sum of two independent processes, one as an integral functional of the limiting arrival process, and the other as a continuous Gaussian process. When the arrival limit is Gaussian, the limiting shot noise process becomes a Gaussian process. We also consider a finite collection of shot noise processes as defined in (1.2) with a family of shot shape functions but the same arrival process and noises, and prove their joint convergence in Theorem 2.3.

We discuss the applications of the FCLT to a simple multiplicative model and the queueing and work-input process for a non-stationary infinite-server queueing model ($G_t/G_t/\infty$) in Section 3. The simple multiplicative model requires that the shot shape function $H(t,x) = \tilde{H}(t)\varphi(x)$ for a nonnegative and monotone function $\tilde{H}(t)$ and a measurable function $\varphi : \mathbb{R}^k \to \mathbb{R}$. When $\tilde{H}(t) \equiv 1$ and $\varphi(x) = x$ for each $x \in \mathbb{R}_+$ ($k = 1$), the model becomes a non-stationary compound process, which is new in the literature. As a consequence of Theorem 2.2, we obtain an FCLT for such non-stationary compound processes; see Theorem 3.1. The multiplicative model has applications in insurance risk theory (setting $\varphi(x) = x$ for $x \in \mathbb{R}_+$), in particular, modeling the delay in insurance claim settlement [27]. It also has applications in physics [41], to study a damped harmonic oscillator subject to a random force, which has the new feature that the random forces depend on the arrival times.

For the non-stationary infinite-server queueing system, the queueing process requires that the shot shape function $H(t,x) = \mathbf{1}(t < x)$, while the work-input process has $H(t,x) = x\mathbf{1}(t < x)$, for $x \in \mathbb{R}_+$. The $G_t/G_t/\infty$ queueing model has been recently studied in [36]. In some sense, the class of non-stationary shot noise processes is a generalization of the non-stationary infinite-server queueing model studied in [36], where functional limit theorems are established for the associated two-parameter processes in addition to the total count process. Work-input processes are studied in [30] using Poisson shot noise processes, where fractional Brownian motion limits are obtained in the conventional scaling regime. We obtain an approximation for the joint queueing and work-input processes, which is a continuous two-dimensional non-stationary Gaussian process when the arrival limit is Gaussian.

To prove the FCLT, we employ a classical weak convergence criterion in the Skorohod J_1 topology, Theorem 13.3 in Billingsley [5]. It provides a sufficient condition involving a modulus of continuity (see (5.2)), which requires the maximal inequalities in Theorems 10.3 and 10.4 in [5] to establish. Maximal inequalities are usually very challenging to prove. In many applications, certain properties of the processes of interest (e.g., the martingale property) can often lead to useful maximal inequalities. The power of Theorems 10.3 and 10.4 in [5] lies in that sufficient conditions on the probability bounds for the increments of the process are provided in order to obtain the corresponding probability bounds in the maximal inequalities. Those conditions require that the probability bound involves a *finite measure* (see (5.4)), which can be often induced from the moment bounds for the increment of the processes. However, for the class of non-stationary shot noise processes, the probability bounds for the processes of interest do not provide such a convenient finite measure (see Lemmas 4.4–4.5 and discussions in Remark 4.2).

One main contribution of the paper is to prove new maximal inequalities involving a finite set function with the superadditive property, which generalize Theorems 10.3 and 10.4 in [5]; see Theorems 5.1 and 5.2. Their proofs in Section 7 are adaptations of the corresponding ones in [5]. We apply them to verify the sufficient condition with the modulus of continuity in Theorem 13.3 in [5] for the shot noise process in the proof of the FCLT. We also apply the maximal inequality in Theorem 5.1 to prove the FWLLN for the shot noise process. In addition, a criterion to prove that there exists a stochastic process in the space $\mathbb{D}$ given its consistent finite dimensional distributions is provided in Theorem 13.6 in [5]. That criterion also relies on the maximal inequalities in Theorems 10.3 and 10.4 in [5], and thus requires that the probability bound for the process increments involves a finite measure. We prove a new criterion of existence in Theorem 5.3, where the finite measure assumption in the probability bound for the increments of the process is relaxed to allow a finite set function with the superadditive property. It is then applied to prove the existence of the Gaussian limit process in the FCLT in the space $\mathbb{D}$ (in fact in $\mathbb{C}$, which requires additional continuity in mean square for the limit process; see the proof of Lemma 6.3). These generalizations of the maximal inequalities and the criterion of existence in the space $\mathbb{D}$ may be useful in future research on weak convergence of stochastic processes.

1.1. **Literature review.** Shot noise processes have been extensively studied, and have many applications in physics, insurance risk theory, telecommunications, and service systems. Functional limit theorems have been established in two asymptotic regimes.

In the asymptotic regime where the arrival rate of shot noises (the intensity/density) becomes large, only limited work has been done for some special classes of shot noise processes. One class includes the queueing, workload and work-input processes in infinite-server queueing systems with i.i.d. service times. For these models, the shot noises represent service times, and the shot shape function becomes an indicator function for the queueing process; see, e.g., Chapter 10 of [44] for a review. Networks of infinite-server queues with shot-noise-driven arrival intensities are recently studied in [26], in order to capture the strong fluctuations in the arrival process. Weak convergence of a certain class of compound stochastic processes was studied in [19] in this asymptotic regime. That paper includes a special class of shot noise processes with a renewal arrival process and i.i.d. shot noises in which the shot shape function $H(t, x)$ satisfies some regularity conditions (see the assumptions in Theorem 4.3 in [19]). See also the relevant discussions in Section 2.2. Our results are established for the most general setting with both non-stationary arrival processes and shot noises.

Although not proving functional limit theorems, some important asymptotic analysis has been also done for certain shot noise processes in this regime. Papoulis [37] first proved the normal approximation and its rate of convergence for the standard shot noise process with a Poisson arrival process and i.i.d. shot noises. Heinrich and Schmidt [17] studied multidimensional shot noise processes and proved the normal approximation and its rate of convergence in that regime. In both papers, the asymptotic behavior of the quantity $(X(t)-E[X(t)])/(Var(X(t)))^{1/2}$ is studied, which is different from the nature of our analysis. Recently, Biermé and Desolneux [4] studied the expected number of level crossings for the shot noise processes with a Poisson arrival process and i.i.d. shot noises, which has a shot shape function $H(t,x)=\tilde{H}(t)x$, for $t\geq 0$ and $x\in\mathbb{R}$, with a smooth function $\tilde{H}(\cdot)$ in this asymptotic regime.

The conventional scaling regime for shot noise processes is to scale up time and shot noises simultaneously. There is a vast literature of studies in that regime. Klüppelberg and Mikosch [27] proved an FCLT for explosive shot noise processes with a Poisson arrival process, which has a self-similar Gaussian limit. Klüppelberg et al. [28] proved an FCLT for Poisson shot noise processes which has an infinite-variance stable limit process. In [29], a fractional Brownian motion limit is proved for Poisson shot noise processes that capture long-range dependence. Iksanov [20] and Iksanov et al. [23] studied renewal shot noise processes where the shot shape function takes the form independent of shot noises $\{Z_i\}$, and proved FCLTs under various conditions on the shot shape function. Iksanov et al. [24, 25] recently studied renewal shot noise processes with immigration and proved scaling limits and convergence to stationarity. We refer to [21] for a thorough review on the subject. In [22], an infinite-server queueing model with correlated interarrival and service times is studied in the conventional scaling regime and a limiting Gaussian process is obtained in the FCLT assuming that the service time distributions are regularly varying. We also refer to the work in [6, 9, 12, 13, 18, 27, 28, 29, 31, 32, 34, 40] and references therein for relevant asymptotic properties of Poisson shot noise processes, and in [39, 43, 42, 33, 38] for more general shot noise processes, as well as their applications.

1.2. **Organization of the paper.** We summarize the notations used in the paper in the next subsection. The model and main results (FWLLN and FCLT) are presented in Section 2. We present the applications in Section 3. Preliminary results on the probability and moment bounds for some prelimit and limit processes are given in Section 4. We state the new maximal inequalities and criterion of existence in Section 5 and their proofs are given in Section 7. We prove the FCLT in Section 6. We collect additional proofs in Section 8.

1.3. **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 $\mathbb{R}^k$-valued function space of all 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} [5, 11, 44]. 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 $\mathbb{C}$ be the subset of $\mathbb{D}$ for 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 in 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. Functional Limit Theorems

In this section, we state the FWLLN and FCLT for the shot noise process X defined in (1.2). We consider a sequence of the non-stationary shot noise processes indexed by n and let $n \to \infty$. In particular, in the n^{th} system, we write A^n and X^n and the associated $\{\tau_i^n\}$ while the variables Z_i and the distributions F_t, $t \geq 0$, are fixed. We first make the following assumptions.

Assumption 1. *The sequence of arrival processes A^n satisfies an FCLT:*

$$\hat{A}^n := \sqrt{n}\big(\bar{A}^n - \Lambda\big) \Rightarrow \hat{A} \quad in \quad (\mathbb{D},\ J_1) \quad as \quad n \to \infty \tag{2.1}$$

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

Note that Assumption 1 implies an FWLLN for the fluid-scaled arrival process $\bar{A}^n$:

$$\bar{A}^n \Rightarrow \Lambda \quad \text{in} \quad (\mathbb{D},\ J_1) \quad \text{as} \quad n \to \infty. \tag{2.2}$$

A large class of models has a continuous Gaussian limit process $\hat{A}$. We provide several examples of arrival processes with a Gaussian limit that satisfy Assumption 1. A special case is $\hat{A}(t) = c_a B(\Lambda(t))$ for a standard Brownian motion B and a constant c_a capturing the variabilities in the arrival process. When the process A^n is a renewal process, c_a represents the coefficient of variation for the interarrival times. When the interarrival times are weakly dependent and satisfying the strong α-mixing condition, by Theorem 4.4.1 and Corollary 13.8.1 in [44], the arrival process $\hat{A}^n$ satisfies an FCLT with a Brownian motion limit, where the coefficient c_a captures the dependence among the interarrival times. When the arrival process is a Markov-modulated Poisson process, the limit $\hat{A}$ is a Brownian motion with c_a capturing the effect of the random environment (see Example 9.6.2 in [44] and also [1]). When the arrival process is a stationary Hawkes process (a class of simple point processes that are self-exciting and have clustering effect), whose intensity is the sum of a baseline intensity and a term depending upon the entire past history of the point process, the limit $\hat{A}$ is a non-Markov Gaussian process with dependent increments [14]. See also FCLTs with Brownian motion limits for nonlinear Hawkes processes in [3, 46].

To simplify notations, we define the following functions: for each $0 \leq u \leq s \leq t$,

$$G_k(t,u) := \int_{\mathbb{R}^k} H(t-u,x)^k dF_u(x), \quad k \in \mathbb{N},$$

$$\tilde{G}(t,u) := G_2(t,u) - G_1(t,u)^2 = \int_{\mathbb{R}^k} (H(t-u,x) - G_1(t,u))^2 dF_u(x),$$

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

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

$$\tilde{G}(t,s,u) := \check{G}_2(t,s,u) - \check{G}_1(t,s,u)^2,$$

$$\breve{G}(t,u) := \int_{\mathbb{R}^k} (H(t-u,x) - G_1(t,u))^4 dF_u(x),$$

and

$$\breve{G}(t,s,u) := \int_{\mathbb{R}^k} (H(t-u,x) - H(s-u,x) - G_1(t,u) + G_1(s,u))^4 dF_u(x).$$

We now state the following regularity conditions on the shot shape function H and the c.d.f. $F_t(\cdot)$. Note that conditions (i) and (ii) in Assumption 2 imply that all functions defined above are finite for $0 \le s \le t$. We use the convention that $H(t,x) \equiv 0$ for $t < 0$.

Assumption 2. *For each $t \ge 0$, the c.d.f. $F_t(\cdot)$ is continuous and has finite marginal mean. The shot shape function $H(\cdot,x) \in \mathbb{D}$ is monotone for each $x \in \mathbb{R}^k$. In addition, the following regularity conditions are satisfied:*

(i)
$$\sup_{0\le t\le T} V_0^T(G_1(t,\cdot)) < \infty, \tag{2.3}$$
where $V_0^T(G_1(t,\cdot))$ is the total variation of the function $G_1(t,\cdot)$ in the interval $[0,T]$, for each $0 \le t \le T$;

(ii) *for each $t \ge 0$,*
$$\sup_{0\le u\le t} \tilde{G}(t,u) < \infty \quad \textit{and} \quad \sup_{0\le u\le t} \breve{G}(t,u) < \infty; \tag{2.4}$$

(iii) *for each $T \ge t \ge 0$,*
$$\lim_{\delta\downarrow 0} \int_{[0,T]} \breve{G}_2(t,t-\delta,u) d\Lambda(u) = 0. \tag{2.5}$$

Remark 2.1. *We remark that the convergence in* (2.5) *always holds as $\delta \uparrow 0$ from the left given that $H(\cdot,x) \in \mathbb{D}$ for each $x \in \mathbb{R}^k$. Indeed, if $\delta \uparrow 0$, then $(t-\delta) \downarrow t$ for each $t \ge 0$. Since $H(\cdot,x) \in \mathbb{D}$ for each $x \in \mathbb{R}^k$, we have $H(t-u,x) - H(t-\delta-u,x) \to 0$ as $\delta \uparrow 0$ for each $0 \le u \le t$ and $x \in \mathbb{R}^k$. By the bounded convergence theorem we have $\breve{G}_2(t,t-\delta,u) \to 0$ as $\delta \uparrow 0$ from the left for each $u \ge 0$. Using the bounded convergence theorem again, we obtain the convergence in* (2.5) *as $\delta \uparrow 0$ from the left. Therefore, in condition (iii) we only require δ converges to 0 from the right.*

Note that the condition (2.5) implies that

$$\lim_{\delta\to 0} \int_{[0,T]} \breve{G}_1(t,t-\delta,u) d\Lambda(u) = 0. \tag{2.6}$$

Define the process $\bar{X}^n := \{\bar{X}^n(t) : t \ge 0\}$ by $\bar{X}^n(t) := n^{-1}X^n(t)$ for $t \ge 0$.

Theorem 2.1. (*FWLLN*) *Under Assumptions 1–2,*
$$\bar{X}^n \Rightarrow \bar{X} \quad \textit{in} \quad (\mathbb{D},\ J_1) \quad \textit{as} \quad n \to \infty, \tag{2.7}$$
where $\bar{X} := \{\bar{X}(t) : t \ge 0\}$ is a continuous deterministic function, defined by
$$\bar{X}(t) := \int_{[0,t]} G_1(t,u) d\Lambda(u), \quad t \ge 0. \tag{2.8}$$

Define the process $\hat{X}^n := \{\hat{X}^n(t) : t \ge 0\}$ by
$$\hat{X}^n(t) := \sqrt{n}(\bar{X}^n(t) - \bar{X}(t)), \quad t \ge 0, \tag{2.9}$$
where $\bar{X}(t)$ is given in (2.8).

Theorem 2.2. (*FCLT*) *Under Assumptions 1–2,*

$$\hat{X}^n \Rightarrow \hat{X} \quad \text{in} \quad (\mathbb{D},\ J_1) \quad \text{as} \quad n \to \infty, \tag{2.10}$$

where $\hat{X} := \{\hat{X}(t) : t \geq 0\}$ *can be written as a sum of two independent stochastic processes* $\hat{X}_1 := \{\hat{X}_1(t) : t \geq 0\}$ *and* $\hat{X}_2 := \{\hat{X}_2(t) : t \geq 0\}$*, with*

$$\hat{X}_1(t) := \hat{A}(t)G_1(t,t) - \int_{(0,t]} \hat{A}(u) dG_1(t,u), \quad t \geq 0, \tag{2.11}$$

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

$$\hat{R}_2(t,s) := Cov\big(\hat{X}_2(t), \hat{X}_2(s)\big) = \int_{[0,t\wedge s]} \big(G_2(t,s,u) - G_1(t,u)G_1(s,u)\big) d\Lambda(u), \tag{2.12}$$

with

$$G_2(t,s,u) := \int_{\mathbb{R}^k} H(t-u,x)H(s-u,x) dF_u(x), \quad t \geq u \geq 0,\ s \geq u \geq 0. \tag{2.13}$$

Remark 2.2. *We remark that the limit process* $\hat{X}_1$ *has sample paths in* $\mathbb{D}$ *under Assumptions 1–2. See Lemma 6.1 and its proof. If, in addition,* $G_1(\cdot,u) \in \mathbb{C}$ *for each* $u \geq 0$*, then* $\hat{X}_1$ *is continuous.*

Remark 2.3. *If* $\hat{A}$ *is a continuous Gaussian process with mean 0 and covariance function* $\hat{R}_a(t,s)$*, then* $\hat{X}_1$ *is a Gaussian process with mean 0 and covariance function*

$$\hat{R}_1(t,s) := Cov(\hat{X}_1(t), \hat{X}_1(s)) = \int_{[0,t]} \int_{[0,s]} G_1(t,u)G_1(s,v) d\hat{R}_a(u,v), \quad t,s \geq 0.$$

In the special case that $\hat{A}(t) = c_a B(\Lambda(t))$ *is a time-changed Brownian motion, the covariance function* $\hat{R}_1(t,s)$ *becomes*

$$\hat{R}_1(t,s) = c_a^2 \int_{[0,t\wedge s]} G_1(t,u)G_1(s,u) d\Lambda(u), \quad t,s \geq 0.$$

It is worth noting that the limit $\hat{X}_2$ *in the FCLT only involves the fluid arrival limit* Λ.

2.1. **Joint convergence with a family of shot shape functions.** Consider a family of shot shape (impulse response) functions, $\{H^{(k)} : k \in \{1,\dots,K\}\}$ for some $K \in \mathbb{N}_+$. For each k, we denote the corresponding processes $\bar{X}^{n,(k)}$ and $\bar{X}^{(k)}$, $\hat{X}^{n,(k)}$ and $\hat{X}^{(k)}$, and the functions $G_i^{(k)}(t,u)$, $\check{G}_i^{(k)}(t,s,u)$, $i = 1,2$, $\tilde{G}^{(k)}(t,u)$ and $\tilde{G}^{(k)}(t,s,u)$. Suppose that the conditions in Assumption 2 hold for the associated functions with each k, which we refer to as Assumption $2'$.

Theorem 2.3. *Suppose that Assumptions 1 and $2'$ hold.*

(*i*) $\big(\bar{X}^{n,(1)},\dots,\bar{X}^{n,(K)}\big) \Rightarrow \big(\bar{X}^{(1)},\dots,\bar{X}^{(K)}\big)$ *in* $(\mathbb{D}^K, J_1)$ *as* $n \to \infty$*, where* $\bar{X}^{(k)} := \{\bar{X}^{(k)}(t) : t \geq 0\}$ *is a continuous deterministic function, defined by*

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

for $k = 1,\dots,K$.

(ii) $\big(\hat{X}^{n,(1)},\dots,\hat{X}^{n,(K)}\big) \Rightarrow \big(\hat{X}^{(1)},\dots,\hat{X}^{(K)}\big)$ *in* $(\mathbb{D}^K, J_1)$ *as* $n\to\infty$, *where* $\hat{X}^{(k)} := \{\hat{X}^{(k)}(t): t\ge 0\}$ *can be written as a sum of two independent processes* $\hat{X}_1^{(k)} := \{\hat{X}_1^{(k)}(t): t\ge 0\}$ *and* $\hat{X}_2^{(k)} := \{\hat{X}_2^{(k)}(t): t\ge 0\}$. *For* $k=1,\dots,K$, $\hat{X}_1^{(k)}$ *is defined by*

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

$\big(\hat{X}_2^{(1)},\dots,\hat{X}_2^{(K)}\big)$ *is a continuous multidimensional Gaussian processes with mean* 0 *and the covariance function*

$$Cov\big(\hat{X}_2^{(k)}(t),\hat{X}_2^{(j)}(s)\big) = \int_{[0,t\wedge s]} \big(G_2^{(k,j)}(t,s,u) - G_1^{(k)}(t,u)G_1^{(j)}(s,u)\big)d\Lambda(u)$$

where

$$G_2^{(k,j)}(t,s,u) := \int_{\mathbb{R}^k} H^{(k)}(t-u,x)H^{(j)}(s-u,x)dF_u(x)$$

for each $k,j=1,\dots,K$, *and* $t\ge u\ge 0,\ s\ge u\ge 0$.

Note that the statements in Remark 2.2 apply to each $\hat{X}^{(k)}$, $k=1,\dots,K$. If $\hat{A}$ is a continuous Gaussian process with mean 0 and covariance function $\hat{R}_a(t,s)$, then $\big(\hat{X}_1^{(1)},\dots,\hat{X}_1^{(K)}\big)$ is a multidimensional Gaussian process with mean 0 and covariance function

$$Cov\big(\hat{X}_1^{(k)}(t),\hat{X}_1^{(j)}(s)\big) = \int_{[0,t]}\int_{[0,s]} G_1^{(k)}(t,u)G_1^{(j)}(s,v)d\hat{R}_a(u,v)$$

for each $k,j=1,\dots,K$, and $t,s\ge 0$. In the special case that $\hat{A}(t) = c_aB(\Lambda(t))$ is a time-changed Brownian motion, the covariance function above becomes

$$c_a^2\int_{[0,t\wedge s]} G_1^{(k)}(t,u)G_1^{(j)}(s,u)d\Lambda(u),$$

for each $k,j=1,\dots,K$ and $t,s\ge 0$.

2.2. **Discussions on the model with i.i.d. shot noises.** When the shot noises are i.i.d. with the same distribution as a random vector Z having joint c.d.f F, the regularity conditions (i) and (iii) in Assumption 2 are not required. Indeed, the condition (i) is always satisfied given the monotonicity of $H(t,x)$ in t. As for condition (iii), first by Fubini's theorem, the left hand side of (2.5) becomes

$$\lim_{\delta\to 0}\int_{\mathbb{R}^k}\int_{[0,T]} \big(H(t-u,x) - H(t-\delta-u,x)\big)^2 d\Lambda(u)dF(x).$$

Next, since $|H(t-u,x) - H(t-\delta-u,x)| \le |H(T,x)|\vee|H(0,x)|$ for $t,u\in[0,T]$, $\delta>0$ and $x\in\mathbb{R}^k$, by the bounded convergence theorem, to show the limit above is equal to 0, it suffices to show that for each $x\in\mathbb{R}^k$,

$$\lim_{\delta\to 0}\int_{[0,T]} \big(H(t-u,x) - H(t-\delta-u,x)\big)^2 d\Lambda(u) = 0.$$

This equation holds since $\Lambda\in\mathbb{C}$ and $H(\cdot,x)\in\mathbb{D}$ has at most countably many discontinuous points for each $x\in\mathbb{R}^k$. However, in the i.i.d. case, we still require condition (ii), which holds if and only if

$$E[H(t,Z)^2] < \infty, \quad \text{for each} \quad t\ge 0.$$

We remark that in the i.i.d. case, with renewal arrivals, the model becomes a special case of the compound stochastic processes studied in [19] (by setting $\xi(t) = H(t,Z)$). Our assumptions on the function H satisfy the conditions (iii) and (iv) in Theorem 4.3 in [19], where we can change $(t-s)$ to $(\tilde{V}(t)-\tilde{V}(s))$ in the upper bounds, for $\tilde{V}(t)$ defined in (4.13) with a constant arrival rate λ, i.e., $\Lambda(t)=\lambda t$; see also the discussion in the last paragraph on page 24 of [19]. It is worth noting that the assumptions on the increments of the process ξ in [19] are made so that the maximal inequality involving a finite measure can be applied (see more discussions on the weak convergence and maximal inequalities in Section 5).

3. Applications

3.1. **A simple multiplicative model.** We consider a simple multiplicative model in which the shot shape function

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

where $\varphi:\mathbb{R}^k\to\mathbb{R}$ is a measurable function.

We discuss what the regularity conditions on the function H in Assumption 2 imply for this simple multiplicative model. First, we require that $\tilde{H}$ is in $\mathbb{D}$, nonnegative and monotone. Note that

$$G_1(t,u)=\tilde{H}(t-u)\tilde{G}_1(u), \quad \tilde{G}(t,u)=\tilde{H}(t-u)^2\tilde{G}(u),$$

and

$$\check{G}_1(t,s,u)=\big(\tilde{H}(t-u)-\tilde{H}(s-u)\big)\tilde{G}_1(u), \quad \check{G}_2(t,s,u)=\big(\tilde{H}(t-u)-\tilde{H}(s-u)\big)^2\tilde{G}_2(u),$$

where

$$\begin{aligned}\tilde{G}_1(u) &:= \int_{\mathbb{R}^k}\varphi(x)dF_u(x), \quad \tilde{G}_2(u):=\int_{\mathbb{R}^k}\varphi(x)^2dF_u(x),\\ \tilde{G}(u) &:= \tilde{G}_2(u)-\tilde{G}_1(u)^2, \quad u\geq 0.\end{aligned}$$

The condition in (2.3) requires that

$$\sup_{0\leq t\leq T} V_0^T\big(\tilde{H}(t-\cdot)\tilde{G}_1(\cdot)\big)<\infty, \tag{3.1}$$

while the condition in (2.4) requires that for each $t\geq 0$,

$$\sup_{0\leq u\leq t}\tilde{H}(t-u)\tilde{G}(u)<\infty. \tag{3.2}$$

The condition in (2.5) requires that for each $T\geq t\geq 0$,

$$\int_{[0,T]}\big(\tilde{H}(t-u)-\tilde{H}(t-\delta-u)\big)^2\tilde{G}_2(u)d\Lambda(u)\to 0 \quad \text{as} \quad \delta\to 0,$$

which is satisfied under the condition that $\tilde{H}$ is monotone and $\Lambda\in\mathbb{C}$. Thus, the regularity condition (iii) in Assumption 2 is not needed for the simple multiplicative model.

The limit process $\hat{X}_1$ in Theorem 2.2 becomes

$$\hat{X}_1(t) := \hat{A}(t)\tilde{H}(0)\tilde{G}_1(t)-\int_{(0,t]}\hat{A}(u)d\big(\tilde{H}(t-u)\tilde{G}_1(u)\big), \quad t\geq 0. \tag{3.3}$$

Note that $\hat{X}_1$ is continuous if $\tilde{H}(\cdot)\in\mathbb{C}$. (See Lemma 6.1 and its proof.) 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

$$\hat{R}_1(t,s)=\int_{[0,t]}\int_{[0,s]}\tilde{H}(t-u)\tilde{H}(s-v)\tilde{G}_1(u)\tilde{G}_1(v)d\hat{R}_a(u,v), \quad t,s\geq 0. \tag{3.4}$$

If $\hat{A}(t) = c_a B(\Lambda(t))$ is a time-changed Brownian motion, then the covariance function $\hat{R}_1(t,s)$ becomes

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

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

$$\hat{R}_2(t,s) = \int_{[0,t\wedge s]} \tilde{H}(t-u)\tilde{H}(s-u)\tilde{G}(u) d\Lambda(u), \quad t,s \geq 0. \tag{3.5}$$

There is one special case which is worth mentioning. Suppose that $x \in \mathbb{R}$, and $\varphi(x) = x$ for all $x \in \mathbb{R}$. In addition, if the conditional mean of shot noises is zero, that is, $\tilde{G}_1(u) = \int_{\mathbb{R}} x dF_u(x) = 0$ for each $u \geq 0$, then the fluid limit $\bar{X}(t)$ in (2.8) and the limit process $\hat{X}_1$ in Theorem 2.2 (see also equation (2.11)) both vanish. Thus, the limit for $\hat{X}^n$ only has one component $\hat{X}_2$, which is a continuous Gaussian process and has mean 0 and covariance function

$$\hat{R}_2(t,s) = \int_{[0,t\wedge s]} \tilde{H}(t-u)\tilde{H}(s-u)\tilde{G}_2(u) d\Lambda(u), \quad t,s \geq 0,$$

with $\tilde{G}_2(u) = \int_{\mathbb{R}} x^2 dF_u(x)$. It is somewhat surprising that in this special case, the stochastic variability in the arrival process vanishes, and the variability in the limit for the process $\hat{X}$ is only affected by the cumulative arrival rate function $\Lambda(t)$ from the arrival process.

We next discuss several special models.

3.1.1. *Non-stationary compound process.* Consider the following special case of the multiplicative model:

$$X(t) = \sum_{i=1}^{A(t)} Z_i(\tau_i), \quad t \geq 0, \tag{3.6}$$

where Z_i's are nonnegative as described above. Here $\tilde{H}(t) \equiv 1$, $t \in \mathbb{R}_+$ and $\varphi(x) = x$ for each $x \in \mathbb{R}_+$. This process can be regarded as a general *non-stationary compound process* with both non-stationarity in the arrival process and the sequence of random variables $\{Z_i(\tau_i)\}$. As a consequence of Theorem 2.2, we obtain the following theorem for the non-stationary compound processes.

Theorem 3.1. *Under Assumption 1 and assuming that for each $t \geq 0$, the c.d.f. $F_t(\cdot)$ is continuous and has finite variance,* (2.7) *in the FWLLN holds with the limit $\bar{X}$ given by*

$$\bar{X}(t) := \int_{[0,t]} \tilde{G}_1(u) d\Lambda(u), \quad t \geq 0, \tag{3.7}$$

and (2.10) *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, $\hat{X}_1$ is defined in* (3.3) *with $\tilde{H}(\cdot) \equiv 1$ and has continuous sample paths, and*

$$\hat{X}_2(t) = B_2\bigg(\int_{[0,t]} \tilde{G}(u) d\Lambda(u)\bigg), \quad t \geq 0, \tag{3.8}$$

for a standard Brownian motion B_2, with $\tilde{G}_1(u)$ and $\tilde{G}(u)$ being the conditional mean and variance of $\{Z_i(\tau_i)\}$:

$$\tilde{G}_1(u) = \int_{[0,\infty)} x dF_u(x), \quad \tilde{G}_2(u) = \int_{[0,\infty)} x^2 dF_u(x), \quad \tilde{G}(u) = \tilde{G}_2(u) - \tilde{G}_1(u)^2, \quad u \geq 0.$$

If $\hat{A}$ is a Gaussian process with mean 0 and covariance function $\hat{R}_a(s,t)$, then $\hat{X}_1$ is a continuous Gaussian process with mean 0 and covariance function

$$\hat{R}_1(t,s) = \int_{[0,t]} \int_{[0,s]} \tilde{G}_1(u)\tilde{G}_1(v) d\hat{R}_a(u,v), \quad t,s \geq 0. \tag{3.9}$$

In the special case that $\hat{A}(t) = c_a B(\Lambda(t))$ is a time-changed Brownian motion, the covariance function becomes

$$\hat{R}_1(t,s) = c_a^2 \int_{[0,t\wedge s]} \tilde{G}_1(u)^2 d\Lambda(u), \quad t,s \geq 0. \tag{3.10}$$

Observe that with i.i.d. random variables $\{Z_i\}$, under Assumption 1, we obtain the FCLT with $\hat{X}_2$ being a time-changed Brownian motion, that is, $\hat{X}_2(t) = \sigma_Z B_2(\Lambda(t))$ for each $t \geq 0$, where B_2 is a standard Brownian motion, independent of $\hat{X}_1$, and $\sigma_Z^2 = Var(Z_1)$. The covariance function for $\hat{X}_2$ in (3.8)

$$\hat{R}_2(t,s) = \int_{[0,t\wedge s]} \tilde{G}(u) d\Lambda(u), \tag{3.11}$$

is a natural generalization of that in the i.i.d. case.

3.1.2. *Application in insurance risk theory.* The simple multiplicative model has been used in insurance risk theory (see, e.g., [27]). Here $x \in \mathbb{R}_+$ since Z_i represents the pay-offs of insurance claims, and $\varphi(x) = x$ for all $x \in \mathbb{R}_+$. Specific examples for the function $\tilde{H}$ include:

(i) $\tilde{H}(t) = 1 - e^{-\gamma t}$ for some $\gamma > 0$. When $x \in \mathbb{R}_+$, this function is used to model delays in insurance claim settlement, referring to situations in which the pay-off of each claim decreases exponentially fast.

(ii) $\tilde{H}(t) = t^\gamma$ for some $\gamma > 0$. When $x \in \mathbb{R}_+$, the function $\tilde{H}$ is used to model the pay-off of each claim that decreases polynomially fast.

When the arrival process has a Gaussian limit, the FCLT provides a Gaussian approximation for the shot noise process, which can be then used to approximate the corresponding ruin probability. Recall that the ruin probability can be represented as the first-passage-time (hitting time) of a simple functional of the shot noise process; see Section 4.2 in [27]. Although the first-passage-time (hitting time) of Gaussian processes is difficult to explicitly characterize [7], numerical solutions can be easily obtained for the ruin probability by using the Gaussian approximations. It is worth noting that in the conventional regime, functional limit theorems are proved in [27, 28, 29] for Poisson shot noise processes, where the limit processes are Brownian motion, α-stable process and fractional Brownian motion under the appropriate assumptions upon the shot noises, respectively. See also Section 1.1 for relevant discussions.

3.1.3. *Application in physics.* In [41], the multiplicative model is applied to study a damped harmonic oscillator subject to a random force, where the random forces are given by i.i.d. symmetric α-stable random variables ($\alpha \in (0,2]$). In the conventional scaling regime as reviewed in Section 1.1, the diffusion-scaled shot noise process is shown to converge to a stochastic integral with respect to a Lévy process. Here in the asymptotic regime with high intensity, our results provide a new limit under the assumptions that the random forces are conditionally independent with Gaussian distributions, having mean zero but the variance depending on the time when the forces occur. In addition, the random forces can be multidimensional.

Specifically, suppose that the random forces have a conditional Gaussian distribution with mean zero and covariance matrix Σ_t such that the density of $F_t(x)$ is

$$f_t(x) = \frac{1}{\sqrt{(2\pi)^k|\Sigma_t|}} \exp\bigg(- \frac{1}{2} x^{\mathrm{T}} \Sigma_t^{-1} x \bigg), \quad x \in \mathbb{R}^k \tag{3.12}$$

for each $t \geq 0$, where x^{T} is the transpose of $x \in \mathbb{R}^k$ and $|\Sigma_t|$ is the determinant of Σ_t. Then the limit process $\hat{X}$ in Theorem 2.2 has the covariance functions $\hat{R}_1(t,s)$ and $\hat{R}_2(t,s)$ in (3.4) and (3.5), respectively, where

$$\tilde{G}_1(u) = \int_{\mathbb{R}^k} \varphi(x) f_u(x) dx, \quad \tilde{G}_2(u) = \int_{\mathbb{R}^k} \varphi(x)^2 f_u(x) dx,$$

for each $u \geq 0$, with $f_t(u)$ given in (3.12).

3.2. **Queueing and work-input processes in non-stationary infinite-server queues.** Consider a non-stationary infinite-server queue with a time-varying arrival process A^n as in Assumption 1 and service times $\{Z_i\}$ as described above, denoted as "$G_t/G_t/\infty$" [36]. The total queue length process $Q^n := \{Q^n(t) : t \geq 0\}$ and the work-input process $W^n := \{W^n(t) : t \geq 0\}$ can be written as

$$Q^n(t) = \sum_{i=1}^{A^n(t)} \mathbf{1}(\tau_i^n + Z_i(\tau_i^n) > t), \quad t \geq 0, \tag{3.13}$$

and

$$W^n(t) = \sum_{i=1}^{A^n(t)} Z_i(\tau_i^n) \mathbf{1}(\tau_i^n + Z_i(\tau_i^n) > t), \quad t \geq 0. \tag{3.14}$$

The work-input processes are studied in [30], as a general class of Poisson shot noise processes. Functional limit theorems are proved under certain conditions, where the limit processes are fractional Brownian motions. Our results are distinct in two aspects: first, the model itself is new since the arrival process is more general, and non-stationarity lies in both arrival and service processes, and second, the scaling is different, since we let the arrival rate or intensity get large.

It is worth noting that the work-input process W^n is different from the remaining workload process in the system at each time. The latter can be obtained from the two-parameter process limits for the non-stationary infinite-server queueing model studied in [36]. The remaining workload process is the integral of the two-parameter queueing process tracking the residual service times with respect to the second time parameter.

For the queueing process Q^n, the shot shape function

$$H^{(1)}(t,x) = \mathbf{1}(t < x), \quad x \in \mathbb{R}_+,$$

and for the work-input process W^n,

$$H^{(2)}(t,x) = x\mathbf{1}(t < x), \quad x \in \mathbb{R}_+.$$

Since the queueing process Q^n has been studied in [36], we focus on the work-input process and its joint convergence with the queueing process. It is also worth noting that the queueing process is recently studied in [22] under the conventional scaling regime. Assuming that the interarrival and service times form a sequence of i.i.d. two-dimensional random vectors with a general bivariate distribution and that the service time distributions are regularly varying, they have proved the weak convergence of the queueing processes to a Gaussian process.

We discuss the regularity conditions in Assumption 2. It is easy to see that for each $x \in \mathbb{R}_+$, $H^{(1)}(t,x)$ and $H^{(2)}(t,x)$ are in $\mathbb{D}$ as a function of t, and they are nonnegative and nonincreasing in t. For $H^{(1)}(t,x)$, we have

$$G_1^{(1)}(t,u) = F_u^c(t-u), \quad G_2^{(1)}(t,u) = F_u^c(t-u), \quad \tilde{G}^{(1)}(t,u) = F_u(t-u)F_u^c(t-u),$$

and

$$\check{G}_1^{(1)}(t,s,u) = F_u(s-u) - F_u(t-u), \quad \check{G}_2^{(1)}(t,s,u) = F_u(t-u) - F_u(s-u).$$

Thus, conditions (i) and (ii) in Assumption 2 are always satisfied without additional assumptions on the distribution function $F_u(\cdot)$, and the condition in (2.5) becomes

$$\int_{[0,T]} (F_u(t-u) - F_u(t-\delta-u))d\Lambda(u) \to 0 \quad \text{as} \quad \delta \to 0,$$

for each $T \ge t \ge 0$. Since for each $u \ge 0$, $F_u(\cdot)$ is continuous (see Assumption 2), the integrand above converges to 0 as $\delta \to 0$. Thus, by the bounded convergence theorem, condition (2.5) holds. Therefore, for the queueing process, we impose the same assumptions on the arrival process and service times as those in [36]. In particular, we do not require additional moment conditions on the service times, which is distinct from the work-input process as shown below in (3.17).

For $H^{(2)}(t,x)$, we have

$$G_1^{(2)}(t,u) = \int_{(t-u,\infty)} x dF_u(x), \quad G_2^{(2)}(t,u) = \int_{(t-u,\infty)} x^2 dF_u(x), \tag{3.15}$$

$$\tilde{G}^{(2)}(t,u) = \int_{(t-u,\infty)} x^2 dF_u(x) - \Big(\int_{(t-u,\infty)} x dF_u(x) \Big)^2,$$

and

$$\check{G}_1^{(2)}(t,s,u) = -\int_{(s-u,t-u]} x dF_u(x), \quad \check{G}_2^{(2)}(t,s,u) = \int_{(s-u,t-u]} x^2 dF_u(x).$$

The condition in (2.3) requires that

$$\sup_{0 \le t \le T} V_0^T \bigg(\int_{(t-\cdot,\infty)} x dF_{\cdot}(x) \bigg) < \infty, \tag{3.16}$$

while the condition in (2.4) requires that for each $t \ge 0$,

$$\sup_{0 \le u \le t} \int_{(t-u,\infty)} x^2 dF_u(x) < \infty. \tag{3.17}$$

The condition in (2.5) requires that for each $T \ge t \ge 0$,

$$\int_{[0,T]} \int_{(t-\delta-u,t-u]} x^2 dF_u(x) d\Lambda(u) \to 0 \quad \text{as} \quad \delta \to 0, \tag{3.18}$$

which always holds since $F_u(\cdot)$ is continuous for each $u \ge 0$.

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

$$\bar{Q}(t) := \int_{[0,t]} F_u^c(t-u) d\Lambda(u), \tag{3.19}$$

$$\bar{W}(t) := \int_{[0,t]} G_1^{(2)}(t,u) d\Lambda(u) = \int_{[0,t]} \int_{(t-u,\infty)} x dF_u(x) d\Lambda(u), \tag{3.20}$$

for $t \geq 0$ and $G_1^{(2)}(t,u)$ is given in (3.15). It is easy to check that $\bar{Q}(t)$ and $\bar{W}(t)$ are continuous functions. Let the diffusion-scaled processes $\hat{Q}^n := \{\hat{Q}^n(t) : t \geq 0\}$ and $\hat{W}^n := \{\hat{W}^n(t) : t \geq 0\}$ be defined by $\hat{Q}^n(t) := \sqrt{n}(\bar{Q}^n(t) - \bar{Q}(t))$ and $\hat{W}^n(t) := \sqrt{n}(\bar{W}^n(t) - \bar{W}(t))$ for $t \geq 0$, respectively. By Theorem 2.3, we obtain the following theorem for the queueing and work-input processes.

Theorem 3.2. *Under Assumptions 1 and conditions in* (3.16)–(3.18),

$$(\bar{Q}^n, \bar{W}^n) \Rightarrow (\bar{Q}, \bar{W}) \quad in \quad (\mathbb{D}^2,\ J_1) \quad as \quad n \to \infty,$$

where $\bar{Q}$ and $\bar{W}$ are defined in (3.19)–(3.20), *and*

$$(\hat{Q}^n, \hat{W}^n) \Rightarrow (\hat{Q}, \hat{W}) \quad in \quad (\mathbb{D}^2,\ J_1) \quad as \quad n \to \infty,$$

where the limits $\hat{Q} = \{\hat{Q}(t) : t \geq 0\}$ and $\hat{W} = \{\hat{W}(t) : t \geq 0\}$ are continuous and can be written as $\hat{Q} = \hat{Q}_1 + \hat{Q}_2$ and $\hat{W} = \hat{W}_1 + \hat{W}_2$, $\hat{Q}_1$ and $\hat{Q}_2$ are independent, $\hat{W}_1$ and $\hat{W}_2$ are independent, $\hat{Q}_1$ and $\hat{W}_1$ are defined by

$$\hat{Q}_1(t) = \hat{A}(t) + \int_{(0,t]} \hat{A}(u) dF_u(t-u), \quad t \geq 0,$$

$$\hat{W}_1(t) = \hat{A}(t) G_1^{(2)}(t,t) - \int_{(0,t]} \hat{A}(u) dG_1^{(2)}(t,u), \quad t \geq 0,$$

and $\big(\hat{Q}_2, \hat{W}_2\big)$ is a two-dimensional Gaussian process with mean 0 *and covariance functions:*

$$Cov(\hat{Q}_2(t), \hat{Q}_2(s)) = \int_{[0,t\wedge s]} \big(F_u^c(t\wedge s - u) - F_u^c(t-u)F_u^c(s-u)\big) d\Lambda(u),$$

$$Cov(\hat{W}_2(t), \hat{W}_2(s)) = \int_{[0,t\wedge s]} \big(G_2^{(2)}(t\vee s, u) - G_1^{(2)}(t,u) G_1^{(2)}(s,u)\big) d\Lambda(u),$$

$$Cov(\hat{Q}_2(t), \hat{W}_2(s)) = \int_{[0,t\wedge s]} \big(G_1^{(2)}(t\vee s, u) - F_u^c(t-u) G_1^{(2)}(s,u)\big) d\Lambda(u),$$

for $t, s \geq 0$, where $G_1^{(2)}$ and $G_2^{(2)}$ are defined in (3.15).

It is evident that $G_1^{(1)}(\cdot, u) \in \mathbb{C}$ and $G_1^{(2)}(\cdot, u) \in \mathbb{C}$, and thus, by Lemma 6.1, $\hat{Q}_1$ and $\hat{W}_1$ are continuous. If $\hat{A}$ is a Gaussian process with mean 0 and covariance function $\hat{R}_a(s,t)$, then $\big(\hat{Q}_1, \hat{W}_1\big)$ is a two-dimensional continuous Gaussian process with mean 0 and covariance functions: for $t, s \geq 0$,

$$Cov(\hat{Q}_1(t), \hat{Q}_1(s)) = \int_{[0,t]} \int_{[0,s]} F_u^c(t-u) F_v^c(s-v) d\hat{R}_a(u,v)$$

$$Cov(\hat{W}_1(t), \hat{W}_1(s)) = \int_{[0,t]} \int_{[0,s]} \big(G_1^{(2)}(t,u) G_1^{(2)}(s,v)\big) d\hat{R}_a(u,v)$$

$$Cov(\hat{Q}_1(t), \hat{W}_1(s)) = \int_{[0,t]} \int_{[0,s]} \big(F_u^c(t-u) G_1^{(2)}(s,v)\big) d\hat{R}_a(u,v).$$

In the special case that $\hat{A}(t) = c_a B(\Lambda(t))$ is a time-changed Brownian motion, these covariance functions become

$$Cov(\hat{Q}_1(t), \hat{Q}_1(s)) = c_a^2 \int_{[0,t\wedge s]} F_u^c(t-u) F_u^c(s-u) d\Lambda(u),$$

$$Cov(\hat{W}_1(t), \hat{W}_1(s)) = c_a^2 \int_{[0,t\wedge s]} \big(G_1^{(2)}(t,u) G_1^{(2)}(s,u)\big) d\Lambda(u),$$

$$Cov(\hat{Q}_1(t), \hat{W}_1(s)) = c_a^2 \int_{[0,t\wedge s]} \big(F_u^c(t-u) G_1^{(2)}(s,u)\big) d\Lambda(u),$$

for each $t, s \geq 0$.

4. Preliminaries

In this section we provide some preliminaries for the proof of Theorem 2.2. We first give a representation for the process $\hat{X}^n$, which follows from simple calculations.

Lemma 4.1. *The process $\hat{X}^n$ defined in* (2.9) *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*

$$\begin{aligned}\hat{X}_1^n(t) &:= \int_{(0,t]} G_1(t,u) d\hat{A}^n(u) \\ &= \hat{A}^n(t) G_1(t,t) - \int_{(0,t]} \hat{A}^n(u-) dG_1(t,u), \quad t \geq 0,\end{aligned} \tag{4.1}$$

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

$$\hat{X}_2^n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(t)} \big(H(t-\tau_i^n, Z_i(\tau_i^n)) - G_1(t,\tau_i^n)\big), \quad t \geq 0. \tag{4.2}$$

By the covariance function $\hat{R}_2(t,s)$ of $\hat{X}_2$ in (2.12), we obtain the following moment properties for the increments of the limit process $\hat{X}_2$ (their proofs can be found in Appendix).

Lemma 4.2. *For each $0 \leq s \leq t$,*

$$E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big] = \int_{(s,t]} \tilde{G}(t,u) d\Lambda(u) + \int_{[0,s]} \tilde{G}(t,s,u) d\Lambda(u), \tag{4.3}$$

and

$$E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^4\big] = 3\bigg(\int_{(s,t]} \tilde{G}(t,u) d\Lambda(u) + \int_{[0,s]} \tilde{G}(t,s,u) d\Lambda(u)\bigg)^2. \tag{4.4}$$

Similarly, for the prelimit process $\hat{X}_2^n$, we have the following moment properties for its increments. The proof is given in the Appendix.

Lemma 4.3. *For each $0 \leq s \leq t$ and all n,*

$$E\big[\big|\hat{X}_2^n(s) - \hat{X}_2^n(t)\big|^2\big] = E\bigg[\int_{(s,t]} \tilde{G}(t,u) d\bar{A}^n(u) + \int_{(0,s]} \tilde{G}(t,s,u) d\bar{A}^n(u)\bigg], \tag{4.5}$$

and

$$\begin{aligned}E\big[\big|\hat{X}_2^n(s) - \hat{X}_2^n(t)\big|^4\big] &= 3E\bigg[\bigg(\int_{(s,t]} \tilde{G}(t,u) d\bar{A}^n(u) + \int_{(0,s]} \tilde{G}(t,s,u) d\bar{A}^n(u)\bigg)^2\bigg] \\ &\quad + \frac{1}{n^2} E\bigg[\sum_{i=A^n(r)+1}^{A^n(s)} \breve{G}(s,\tau_i^n)\bigg] - \frac{3}{n^2} E\bigg[\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n)^2\bigg]\end{aligned}$$

$$+\frac{1}{n^2}E\Big[\sum_{i=1}^{A^n(r)}\breve{G}(s,r,\tau_i^n)\Big]-\frac{3}{n^2}E\Big[\sum_{i=1}^{A^n(s)}\tilde{G}(s,r,\tau_i^n)^2\Big]. \quad (4.6)$$

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

$$\begin{aligned} V(s,t) &:= \tilde{C}\big(\Lambda(t)-\Lambda(s)\big)+\int_{[0,T]}\breve{G}_2(t,s,u)d\Lambda(u)\\ &= \tilde{C}\big(\Lambda(t)-\Lambda(s)\big)\\ &\quad +\int_{[0,T]}\int_{\mathbb{R}^k}\big(H(t-u,x)-H(s-u,x)\big)^2dF_u(x)d\Lambda(u), \end{aligned} \quad (4.7)$$

where $\tilde{C}=\sup_{0\le t,u\le T}\tilde{G}(t,u)<\infty$. To see that $\tilde{C}$ is indeed finite, recall that $\tilde{G}(t,u)=G_2(t,u)-G_1^2(t,u)$ and $H(t,x)$ is assumed monotone in t for each x. Then,

$$\sup_{0\le t,u\le T}\tilde{G}(t,u) \le \sup_{0\le t,u\le T}G_2(t,u)\le\sup_{0\le u\le T}\{G_2(T,u)\vee G_2(0,u)\}. \quad (4.8)$$

Consider $\sup_{0\le u\le T}G_2(T,u)$ first. We have

$$\sup_{0\le u\le T}G_2(T,u) = \sup_{0\le u\le T}\{\tilde{G}(T,u)+G_1^2(T,u)\}\le\sup_{0\le u\le T}\tilde{G}(T,u)+\sup_{0\le u\le T}G_1^2(T,u).$$

The first term is finite by Assumption 2 (ii). By Assumption 2 (i), we have $\sup_{0\le t,u\le T}|G_1(t,u)|<\infty$, which implies $\sup_{0\le t,u\le T}G_1^2(t,u)<\infty$, and thus, $\sup_{0\le u\le T}G_2(T,u)<\infty$. A similar argument applies to $\sup_{0\le u\le T}G_2(0,u)$. Thus, we have shown that $\tilde{C}$ is finite.

Remark 4.1. Note that the function $V(s,t)$ 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 $V(s,t)\le V(s,T)\le V(0,T)$ for each $s,t\ge0$;

(iii) $V(s,t)$ is continuous in both s and t. To see this, for any $\delta_1,\delta_2\in\mathbb{R}$,

$$\begin{aligned} &V(s,t)-V(s-\delta_1,t-\delta_2)\\ &=\tilde{C}\big(\Lambda(t)-\Lambda(t-\delta_2)-(\Lambda(s)-\Lambda(s-\delta_1))\big)\\ &\quad+\int_{[0,T]}\int_{\mathbb{R}^k}\big(H(t-u,x)-H(s-u,x)\big)^2dF_u(x)d\Lambda(u)\\ &\quad-\int_{[0,T]}\int_{\mathbb{R}^k}\big(H(t-\delta_2-u,x)-H(s-\delta_1-u,x)\big)^2dF_u(x)d\Lambda(u)\\ &=\tilde{C}\big(\Lambda(t)-\Lambda(t-\delta_2)-(\Lambda(s)-\Lambda(s-\delta_1))\big)\\ &\quad+\int_{[0,T]}\breve{G}_2(t,t-\delta_2,u)d\Lambda(u)+\int_{[0,T]}\breve{G}_2(s,s-\delta_1,u)d\Lambda(u)\\ &\quad+2\int_{[0,T]}\int_{\mathbb{R}^k}[H(t-u,x)-H(t-\delta_2-u,x)]\\ &\qquad\qquad\times[H(t-\delta_2-u,x)-H(s-u,x)]dF_u(x)d\Lambda(u)\\ &\quad-2\int_{[0,T]}\int_{\mathbb{R}^k}[H(s-u,x)-H(s-\delta_1-u,x)]\\ &\qquad\qquad\times[H(t-u,x)-H(s-\delta_1-u,x)]dF_u(x)d\Lambda(u). \end{aligned} \quad (4.9)$$

Note that the last two terms in (4.9) are bounded by

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

and

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

respectively, due to Cauchy-Schwarz inequality. By the continuity of Λ and Assumption 2 (iii) that for each $T \geq t \geq 0$, $\lim_{\delta\to 0}\int_{[0,T]} \check{G}_2(t,t-\delta,u)d\Lambda(u) = 0$, it is easy to verify that each term in (4.9) converges to 0 as $|\delta_1|+|\delta_2| \to 0$. Now the continuity of V has been proved.

However, the function $V(s,t)$ cannot be regarded as a measure defined on $[0,T]$ due to the nonlinear integrand in (4.7), and cannot be written as the difference $\tilde{V}(t)-\tilde{V}(s)$ for some nondecreasing and continuous function $\tilde{V}$ on $\mathbb{R}_+$. Thus, the standard approach to prove weak convergence in $\mathbb{D}$ (e.g., Theorem 13.5 in [5]) and the existence of a stochastic process in the space $\mathbb{D}$ given its consistent finite-dimensional distributions (e.g., Theorem 13.6 in [5]) cannot be applied directly to the proof of the weak convergence $\hat{X}_2^n \Rightarrow \hat{X}_2$ in $(\mathbb{D}, J_1)$ as $n \to \infty$. This motivates us to prove new maximal inequalities that are necessary to prove the weak convergence in the next section. □

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

Lemma 4.4. *For $0 \leq r \leq s \leq t \leq T$ and any $\epsilon > 0$,*

$$P\big(\big|\hat{X}_2(r)-\hat{X}_2(s)\big| \wedge \big|\hat{X}_2(s)-\hat{X}_2(t)\big| \geq \epsilon\big) \leq \frac{3}{\epsilon^4}V(r,s)V(s,t). \tag{4.10}$$

Proof. We have

$$\begin{aligned} & P\big(\big|\hat{X}_2(r)-\hat{X}_2(s)\big| \wedge \big|\hat{X}_2(s)-\hat{X}_2(t)\big| \geq \epsilon\big) \\ \leq\ & \frac{1}{\epsilon^4}E\big[\big|\hat{X}_2(r)-\hat{X}_2(s)\big|^2\,\big|\hat{X}_2(s)-\hat{X}_2(t)\big|^2\big] \\ \leq\ & \frac{1}{\epsilon^4}\Big(E\big[\big|\hat{X}_2(r)-\hat{X}_2(s)\big|^4\big]\Big)^{1/2}\Big(E\big[\big|\hat{X}_2(s)-\hat{X}_2(t)\big|^4\big]\Big)^{1/2} \\ =\ & \frac{3}{\epsilon^4}\Big(\int_{(r,s]}\tilde{G}(s,u)d\Lambda(u)+\int_{[0,r]}\tilde{G}(s,r,u)d\Lambda(u)\Big) \\ & \quad\times\Big(\int_{(s,t]}\tilde{G}(t,u)d\Lambda(u)+\int_{[0,s]}\tilde{G}(t,s,u)d\Lambda(u)\Big) \\ \leq\ & \frac{3}{\epsilon^4}V(r,s)V(s,t), \end{aligned}$$

where the equality follows from Lemma 4.2. □

Similarly, we obtain the following probability bound for the increments of the prelimit process $\hat{X}_2^n$, whose proof is given in the Appendix.

Lemma 4.5. *For $0 \leq r \leq s \leq t \leq T$, all $n \geq 1$, and any $\epsilon > 0$,*

$$P\big(\big|\hat{X}_2^n(r)-\hat{X}_2^n(s)\big| \wedge \big|\hat{X}_2^n(s)-\hat{X}_2^n(t)\big| > \epsilon\big) \leq \frac{C}{\epsilon^4}V(r,s)V(s,t), \tag{4.11}$$

where $C > 0$ is a constant.

Remark 4.2. Note that even when the sequence $\{Z_i\}$ is i.i.d., the function $V(s,t)$ cannot be regarded as a measure defined on $[0,T]$, unless certain regularity conditions are imposed on the function H (see, e.g., Theorem 4.3 in [19]). For the exposition convenience in this remark, let us assume that $H(t,x)$ is nondecreasing in t for each $x \in \mathbb{R}^k$. It is possible to bound the function $V(s,t)$ by another function $\breve{V}$ which can be regarded as a measure on $[0,T]$, where

$$\begin{aligned}\breve{V}(s,t) \quad &:= \quad \tilde{C}\big(\Lambda(t)-\Lambda(s)\big) \\ &\qquad + \int_{[0,T]}\int_{\mathbb{R}^k}\big(H(T,x)-H(0,x)\big)\big(H(t-u,x)-H(s-u,x)\big)dF(x)d\Lambda(u).\end{aligned} \tag{4.12}$$

This function $\breve{V}(s,t)$ can be written as the difference $\tilde{V}(t)-\tilde{V}(s)$ for the nondecreasing and continuous function $\tilde{V}$ on $\mathbb{R}_+$, where

$$\begin{aligned}\tilde{V}(t) \quad &:= \quad \tilde{C}\Lambda(t) + \int_{[0,T]}\int_{\mathbb{R}^k}\big(H(T,x)-H(0,x)\big)H(t-u,x)dF(x)d\Lambda(u) \\ &= \quad \tilde{C}\Lambda(t) + \int_{\mathbb{R}^k}\big(H(T,x)-H(0,x)\big)\bigg(\int_{[0,T]}H(t-u,x)d\Lambda(u)\bigg)dF(x), \quad t\geq 0.\end{aligned} \tag{4.13}$$

It is easy to check that the function $\tilde{V}(t)$ is continuous without imposing any regularity conditions as remarked in Section 2.2. Thus, in this case, the weak convergence criterion in Theorem 13.5 of [5] can be applied with this bounding function $\breve{V}(s,t)$. See more discussions on this approach in Section 5.

Similarly, in the special case of the simple multiplicative model in Section 3.1, we also observe that the function $V(s,t)$ cannot be regarded as a measure defined on $[0,T]$, but we can bound the function $V(s,t)$ by a function $\breve{V}(s,t)$ which can be regarded as measure on $[0,T]$, where

$$\begin{aligned}\breve{V}(s,t) \quad &:= \quad \tilde{C}\big(\Lambda(t)-\Lambda(s)\big) \\ &\qquad + \big(\tilde{H}(T)-\tilde{H}(0)\big)\int_{[0,T]}\big(\tilde{H}(t-u)-\tilde{H}(s-u)\big)\tilde{G}_2(u)d\Lambda(u),\end{aligned}$$

where $\tilde{G}_2(u)$ is given in (3.1). Here we assume $\tilde{H}$ is nondecreasing for the convenience of exposition. As above, we can also apply Theorem 13.5 of [5] directly without imposing additional regularity conditions other than those in (3.1) and (3.2).

However, for our general non-stationary model, this bounding approach will give us the corresponding functions $\breve{V}(s,t)$ and $\tilde{V}(t)$ in (4.12) and (4.13), respectively, where F is replaced by F_u. The continuity of $\tilde{V}(t)$ is key in applying Theorem 13.5 of [5], which will require that

$$\lim_{\delta\downarrow 0}\int_{[0,T]}\int_{\mathbb{R}^k}\big(H(T,x)-H(0,x)\big)\big(H(t-u,x)-H(t-\delta-u,x)\big)dF_u(x)d\Lambda(u)=0.$$

Instead of making this assumption, we choose to work with the function $V(s,t)$ in (4.7) directly, while assuming the regularity condition (iii) in Assumption 2 (the continuity condition only requires (2.5)). Although under Assumption 2 we cannot apply Theorem 13.5 of [5], we are able to prove the weak convergence by providing new maximal inequalities in the next section. $\square$

5. A General Maximal Inequality and Criterion of Existence

A standard approach to prove weak convergence of stochastic processes $X^n \Rightarrow X$ in $(\mathbb{D}([0,T],\mathbb{R}), J_1)$ is stated as Theorem 13.5 in [5], which requires three conditions:

(i) convergence of finite-dimensional distributions, that is, $(X^n(t_1),\dots,X^n(t_k)) \Rightarrow (X(t_1),\dots,X(t_k))$ for continuity points $\{t_i : 1 \le i \le k\}$ of X;

(ii) $X(T) - X(T-\delta) \Rightarrow 0$ in $\mathbb{R}$ as $\delta \to 0$;

(iii) for $0 \le r \le s \le t \le T$, $n \ge 1$ and $\epsilon > 0$,

$$P\big(|X^n(s) - X^n(r)| \wedge |X^n(t) - X^n(s)| \ge \epsilon\big) \le \frac{1}{\epsilon^{4\beta}}(F(t) - F(r))^{2\alpha}, \tag{5.1}$$

where $\beta \ge 0$ and $\alpha > 1/2$, and F is a nondecreasing and continuous function on $[0,T]$.

The proof of this criterion relies on the maximal inequalities in Theorems 10.3 and 10.4 in [5]. Specifically, the proof requires verifying the third condition in Theorem 13.3 in [5], that is, for $\epsilon > 0$ and $\eta > 0$, there exists a $\delta \in (0,1)$ and n_0 such that

$$P(w''(X^n,\delta) \ge \epsilon) \le \eta, \quad n \ge n_0, \tag{5.2}$$

where $w''(x,\delta)$ is a modulus of continuity of a function $x \in \mathbb{D}$ defined by

$$w''(x,\delta) := \sup_{t_1 \le t \le t_2,\ t_2 - t_1 \le \delta} \{|x(t) - x(t_1)| \wedge |x(t_2) - x(t)|\}, \tag{5.3}$$

with the supremum over all triples t_1, t, t_2 in $[0,T]$ satisfying the constraints. To verify the condition in (5.2), the maximal inequalities in Theorems 10.3 and 10.4 of [5] play a key role. They provide conditions under which the probability bound on the increments as in (5.1) will imply the probability bound on the modulus of continuity as in (5.2).

We first review Theorem 10.3 of [5]. If $\mathcal{T}$ is a Borel subset of $[0,T]$ and μ is a *finite measure* on $[0,T]$ such that

$$P\big(|X(s) - X(r)| \wedge |X(t) - X(s)| \ge \epsilon\big) \le \frac{1}{\epsilon^{4\beta}}\big(\mu(\mathcal{T} \cap (r,t])\big)^{2\alpha}, \tag{5.4}$$

for $0 \le r \le s \le t \le T$, and for $\epsilon > 0$, $\alpha > 1/2$, and $\beta \ge 0$, then

$$P\left(\sup_{r \le s \le t,\ r,s,t \in \mathcal{T}} |X(s) - X(r)| \wedge |X(t) - X(s)| \ge \epsilon\right) \le \frac{K}{\epsilon^{4\beta}}(\mu(\mathcal{T}))^{2\alpha}, \tag{5.5}$$

for $\epsilon > 0$ and K being a constant depending only on α and β. Theorem 10.4 in [5] provides a further inequality which restricts the time intervals to be within δ distance. That is,

$$\begin{aligned} & P\left(\sup_{r \le s \le t, t - r \le \delta,\ r,s,t \in \mathcal{T}} |X(s) - X(r)| \wedge |X(t) - X(s)| \ge \epsilon\right) \\ \le\ & \frac{K}{\epsilon^{4\beta}}\mu(\mathcal{T}) \sup_{0 \le t \le T - 2\delta} \big(\mu(\mathcal{T} \cap [t, t+2\delta])\big)^{2\alpha - 1}, \end{aligned} \tag{5.6}$$

for $\epsilon > 0$ and K being a constant depending only on α and β.

A critical condition in these maximal inequalities requires that μ be a *finite measure* on $[0,T]$, and as a consequence, that results in the probability bound in (5.1) involving a nondecreasing and continuous function F (by simply taking $\mu(s,t] = F(t) - F(s)$). However, as we have observed in Lemmas 4.4–4.5, the probability bounds for $\hat{X}_2$ and $\hat{X}_2^n$ do not provide us a finite measure to work with. We will next prove new maximal inequalities by relaxing the finite measure condition. We give a definition of a set function below that will be used to replace the finite measure.

Definition 5.1. *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 $\mu(A) \le \mu(B)$;*
(iii) *μ is superadditive, that is, for any disjoint Borel sets A and B, $\mu(A) + \mu(B) \le \mu(A \cup B)$.*

By definition, the monotonicity implies that $\mu(s,t] \le \mu(0,T]$ for any $0 \le s \le t \le T$, and the superadditivity implies that $\mu(s,t] \le \mu(0,t] - \mu(0,s]$, for any $0 \le s \le t \le T$. It is evident that if μ is a measure, then the conditions (i)–(iii) are always satisfied.

Remark 5.1. *The function $V(s,t)$ defined in (4.7) naturally induces a set function ν satisfying the conditions in Definition 5.1. More precisely, for any Borel set $A \subset [0,T]$, define*

$$\nu(A) := \sup\left\{V(s,t) : (s,t] \subset A\right\}. \tag{5.7}$$

It is easy to check that ν satisfies all the conditions in Definition 5.1 and $\nu((s,t]) = V(s,t)$ for $0 \le s \le t \le T$. In particular, the condition (iii) is satisfied because of the inequality that $\sum_{i=1}^n x_i^2 \le \left(\sum_{i=1}^n x_i\right)^2$ for each $n \ge 1$ if all elements of $\{x_i : 1 \le i \le n\}$ have the same sign. Note that $V(s,t)$ is continuous in both s and t, while the continuity condition for the set function μ is not required in Theorems 5.1 and 5.2.

Let $\mathcal{T} \subset [0,T]$ be a Borel set and $\{X(t) : t \in \mathcal{T}\}$ is a stochastic process on $\mathcal{T}$. We assume that $X(t)$ is right-continuous in the sense that if for $k \ge 1$, $t_k \in \mathcal{T}$ such that $t_k \downarrow t \in \mathcal{T}$ as $k \to \infty$, we have $X(t_k) \to X(t)$ *a.s.* as $k \to \infty$. The following two theorems generalize Theorems 10.3 and 10.4 of [5], respectively. The proofs of Theorems 5.1–5.3 are given in Section 7, which are adapted from those in [5].

Theorem 5.1. *Suppose that $\alpha > 1/2$ and $\beta \ge 0$ and that μ is a finite set function in Definition 5.1 such that for any $r, s, t \in \mathcal{T}$ with $r \le s \le t$ and $\epsilon > 0$,*

$$P\big(|X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon\big) \le \frac{C_0}{\epsilon^{4\beta}}\big(\mu\big(\mathcal{T} \cap (r,t]\big)\big)^{2\alpha}, \tag{5.8}$$

where C_0 is a positive constant. Then

$$P\left(\sup_{r\le s\le t,\ r,s,t\in\mathcal{T}} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon\right) \le \frac{C_1}{\epsilon^{4\beta}}\big(\mu(\mathcal{T})\big)^{2\alpha}, \tag{5.9}$$

where C_1 is a positive constant that depends only on α, β and C_0.

Theorem 5.2. *Suppose that $\alpha > 1/2$ and $\beta \ge 0$, and for any $r \le s \le t$ with $t - r < 2\delta$, $\delta > 0$, $r, s, t \in \mathcal{T}$ and $\epsilon > 0$,*

$$P\big(|X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon\big) \le \frac{C_0}{\epsilon^{4\beta}}\big(\mu\big(\mathcal{T} \cap (r,t]\big)\big)^{2\alpha}, \tag{5.10}$$

where μ is a finite set function in Definition 5.1. Then

$$P\left(\sup_{\substack{r<s<t,\ r,s,t\in\mathcal{T}\\ t-r\le 2\delta}} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon\right)$$

$$\leq \frac{2C_1}{\epsilon^{4\beta}}\mu(\mathcal{T}) \sup_{0\leq t\leq T-2\delta} \big(\mu\big(\mathcal{T}\cap(t,t+2\delta]\big)\big)^{2\alpha-1}. \tag{5.11}$$

Similarly, a standard criterion to prove the existence of a stochastic process with sample paths in $\mathbb{D}$ given its finite dimensional distributions is given in Theorem 13.6 in [5]. That criterion also requires a probability bound in the same flavor as (5.1), and its proof relies on the same maximal inequalities in Theorems 10.3 and 10.4 of [5]. For our purpose, to prove the existence of the Gaussian process $\hat{X}_2$ in the space $\mathbb{D}$ (in fact in $\mathbb{C}$, see Lemma 6.3) given its finite dimensional distributions, the probability bound for the increments of $\hat{X}_2$ in Lemma 4.4 does not satisfy the condition in Theorem 13.6 in [5]. Therefore, we also generalize Theorem 13.6 in [5] by relaxing the condition on the probability bound for the increments of the process. We now state the new criterion of existence in the following theorem.

Theorem 5.3. *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 $0\leq r\leq s\leq t\leq T$, $\beta\geq 0$, $\alpha>1/2$ and $\epsilon>0$,*

$$P\big(|X(r)-X(s)|\wedge|X(s)-X(t)|\geq\epsilon\big)\leq\frac{C_2}{\epsilon^{4\beta}}\big(\mu(r,t]\big)^{2\alpha}, \tag{5.12}$$

where C_2 is a positive constant, μ is a finite set function in Definition 5.1 and $\mu(0,t]$ is continuous in t;

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

$$\lim_{\delta\downarrow 0} P\big(|X(t)-X(t+\delta)|>\epsilon\big)=0. \tag{5.13}$$

6. Proof of Theorems 2.1 and 2.2

In this section we prove Theorems 2.1 and 2.2.

Proof of Theorem 2.1. We first prove the continuity of $\bar{X}$. For each $t\geq 0$, let $t_k\geq 0$ be a sequence such that $\lim_{k\to\infty}t_k=t$. We have

$$\begin{aligned}
\bar{X}(t_k)-\bar{X}(t) &= \int_{[0,t_k]} G_1(t_k,u)d\Lambda(u)-\int_{[0,t]} G_1(t,u)d\Lambda(u)\\
&= \int_{[0,t]}[G_1(t_k,u)-G_1(t,u)]d\Lambda(u)+\int_{(t,t_k]} G_1(t_k,u)d\Lambda(u)\\
&= \int_{[0,t]}\int_{\mathbb{R}^k}[H(t_k-u,x)-H(t-u,x)]dF_u(x)d\Lambda(u)\\
&\quad+\int_{(t,t_k]}\int_{\mathbb{R}^k} H(t_k-u,x)dF_u(x)d\Lambda(u).
\end{aligned} \tag{6.1}$$

By applying Cauchy-Schwarz inequality twice, the first term on the right hand side of (6.1) is bounded by

$$\left(\Lambda(t)\int_{[0,t]}\check{G}_2(t,t_k,u)d\Lambda(u)\right)^{1/2}. \tag{6.2}$$

Under condition (2.5), it is easy to see that the quantity in (6.2) vanishes as $k \to \infty$ (see also Remark 2.1). For the second term in (6.1), Assumption 2 (i) implies that $\int_{\mathbb{R}^k} H(t_k - u, x) dF_u(x)$ has the same upper bound (and is thus integrable) for all k. By the continuity of Λ and the absolute continuity of Lebesgue-Stieltjes integration, the second term also converges to 0 as $k \to \infty$. Therefore, we conclude that $\bar{X}$ is continuous.

Since $\bar{X}$ is deterministic and continuous, to show that $\bar{X}^n \Rightarrow \bar{X}$ in $(\mathbb{D}, J_1)$ as $n \to \infty$, it suffices to show that for each $T > 0$ and $\epsilon > 0$,

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} |\bar{X}^n(t) - \bar{X}(t)| > \epsilon \bigg) = 0. \tag{6.3}$$

By the definitions of $\bar{X}^n$ and $\bar{X}$, we have

$$\begin{aligned}
& |\bar{X}^n(t) - \bar{X}(t)| \\
= & \left| \frac{1}{n} \sum_{i=1}^{A^n(t)} H(t - \tau_i^n, Z_i(\tau_i^n)) - \int_{[0,t]} G_1(t,u) d\Lambda(u) \right| \\
= & \left| \frac{1}{n} \sum_{i=1}^{A^n(t)} (H(t - \tau_i^n, Z_i(\tau_i^n)) - G_1(t, \tau_i^n)) + \int_{(0,t]} G_1(t,u) d(A^n(u) - \Lambda(u)) \right| \\
\leq & \left| \bar{X}_2^n(t) \right| + \left| \bar{A}^n(t) - \Lambda(t) \right| |G_1(t,t)| + \left| \int_{(0,t]} (\bar{A}^n(u) - \Lambda(u)) dG_1(t,u) \right|,
\end{aligned} \tag{6.4}$$

where $\bar{X}_2^n := \frac{1}{\sqrt{n}} \hat{X}_2^n$ for $\hat{X}_2^n$ defined in (4.2).

Recall that $\bar{A}^n \Rightarrow \Lambda$ in $(\mathbb{D}, J_1)$ in (2.2) and $\Lambda \in \mathbb{C}$, we have that for each $T > 0$ and $\epsilon > 0$,

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} \left| \bar{A}^n(t) - \Lambda(t) \right| > \epsilon \bigg) = 0. \tag{6.5}$$

For the second and third terms in (6.4), equation (6.5) together with the facts that $\sup_{t\in[0,T]} G_1(t,t)$ and $\sup_{t\in[0,T]} V_0^T(G_1(t,\cdot))$ are finite (recall (2.3)) implies that for $\epsilon > 0$

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} \left| \bar{A}^n(t) - \Lambda(t) \right| |G_1(t,t)| > \epsilon \bigg) = 0,$$

and

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} \left| \int_{(0,t]} (\bar{A}^n(u) - \Lambda(u)) dG_1(t,u) \right| > \epsilon \bigg) = 0.$$

The proof now reduces to show that for $\epsilon > 0$

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} \left| \bar{X}_2^n(t) \right| > \epsilon \bigg) = 0, \tag{6.6}$$

Since for any $0 \leq t \leq T$, $|\bar{X}_2^n(t)| \leq |\bar{X}_2^n(T) - \bar{X}_2^n(t)| + |\bar{X}_2^n(T)|$ and $|\bar{X}_2^n(t)| \leq |\bar{X}_2^n(t)| + |\bar{X}_2^n(T)|$, we have

$$|\bar{X}_2^n(t)| \leq |\bar{X}_2^n(t)| \wedge |\bar{X}_2^n(T) - \bar{X}_2^n(t)| + |\bar{X}_2^n(T)|.$$

Thus, to show (6.6), it suffices to prove that

$$\lim_{n\to\infty} P\bigg(\sup_{t\in[0,T]} \left| \bar{X}_2^n(t) \right| \wedge \left| \bar{X}_2^n(T) - \bar{X}_2^n(t) \right| > \epsilon \bigg) = 0, \tag{6.7}$$

and

$$\lim_{n\to\infty} P\big(\big|\bar{X}_2^n(T)\big| > \epsilon\big) = 0. \tag{6.8}$$

To show (6.7), recall Remark 5.1 that the function V in (4.7) induces a finite set function ν such that $\nu((s,t]) = V(s,t)$. By Lemma 4.5, condition (5.8) holds for $\hat{X}_2^n$ (with $\alpha=\beta=1$, $\mathcal{T}=[0,T]$ and ϵ being replaced by $\sqrt{n}\epsilon$), that is,

$$P\big(\big|\hat{X}_2^n(r)-\hat{X}_2^n(s)\big| \wedge \big|\hat{X}_2^n(s)-\hat{X}_2^n(t)\big| > \sqrt{n}\epsilon\big) \le \frac{\hat{C}_0}{n^2\epsilon^4} V(r,s)V(s,t),$$

for some constant $\hat{C}_0 > 0$. Thus, by Theorem 5.1, we obtain that as $n\to\infty$,

$$P\bigg(\sup_{0\le r<s<t\le T} \big|\hat{X}_2^n(r)-\hat{X}_2^n(s)\big| \wedge \big|\hat{X}_2^n(s)-\hat{X}_2^n(t)\big| \ge \sqrt{n}\epsilon\bigg) \le \frac{\hat{C}_0'}{n^2\epsilon^4} V^2(0,T) \to 0,$$

for some constant $\hat{C}_0' > 0$, which further implies (6.7) (by taking $r=0, s=t$ and $t=T$).

To prove (6.8), first note that by (6.5), there exists a large constant $K > 2\Lambda(T)$ such that $P\big(\bar{A}^n(T) > K\big) \to 0$ as $n\to\infty$. We then write

$$P\big(\big|\bar{X}_2^n(T)\big| > \epsilon\big) \le P\big(\bar{A}^n(T) > K\big) + P\big(\mathbf{1}(\bar{A}^n(T)\le K)\big|\bar{X}_2^n(T)\big| > \epsilon\big).$$

The second term on the right hand side is upper bounded by

$$\begin{aligned} & \frac{1}{n\epsilon^2} E\big[\mathbf{1}(\bar{A}^n(T)\le K)\big|\hat{X}_2^n(T)\big|^2\big] \\ = \;& \frac{1}{n\epsilon^2} E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\int_{(0,T]} \tilde{G}(T,u)d\bar{A}^n(u)\bigg] \\ \le \;& \frac{K}{n\epsilon^2} \sup_{0\le u\le T} \tilde{G}(T,u) \to 0 \quad \text{as} \quad n\to\infty, \end{aligned}$$

where the equality follows from (4.5). Note that $\sup_{0\le u\le T}\tilde{G}(T,u) < +\infty$ (see (2.4)). Thus, we have shown that (6.8) holds, which completes the proof of the theorem. $\square$

For the convergence of $\hat{X}_1^n$, we need the following lemma, whose proof is in the Appendix.

Lemma 6.1. *Define the mapping ψ on $\mathbb{D}$: for $z\in\mathbb{D}$,*

$$\psi(z)(t) := z(t)G_1(t,t) - \int_{(0,t]} z(u-)dG_1(t,u), \quad t\ge 0. \tag{6.9}$$

Then the following hold:

(i) *For any $z\in\mathbb{D}$, $\psi(z)\in\mathbb{D}$;*
(ii) *If $z\in\mathbb{C}$ and $G_1(\cdot,u)\in\mathbb{C}$ for each $u\ge 0$, then $\psi(z)\in\mathbb{C}$;*
(iii) *If $z_n\in\mathbb{D}$ for each $n\in\mathbb{N}$ and $z\in\mathbb{C}$ satisfy $z_n\to z$ in $(\mathbb{D},J_1)$ as $n\to\infty$, then $\psi(z_n)\to\psi(z)$ in $(\mathbb{D},J_1)$ as $n\to\infty$.*

Remark 6.1. *The integral in* (6.9) *is understood as a Lebesgue-Stieltjes integral, since for each $t\ge 0$, $G_1(t,\cdot)$ is of bounded variation under Assumption 2 (i) and $z\in\mathbb{D}$. Thus, the mapping ψ is well-defined.*

Lemma 6.2. *Under Assumptions 1–2, $\hat{X}_1^n \Rightarrow \hat{X}_1$ in $(\mathbb{D},J_1)$ as $n\to\infty$, where $\hat{X}_1$ is as given in Theorem 2.2.*

Proof. The claim follows from (4.1) and Lemma 6.1, and applying the continuous mapping theorem. $\square$

We next prove the convergence of the processes $\hat{X}_2^n$. This proceeds in the following steps:

Step 1: The existence of the limit Gaussian process $\hat{X}_2$ with sample paths in $\mathbb{C}$ (Lemma 6.3).

Step 2: The convergence of finite dimensional distributions of $\hat{X}_2^n$ to those of $\hat{X}_2$ (Lemma 6.6).

Step 3: Verifying the convergence criterion with the modulus of continuity as in Theorem 13.3 of [5] and completing the proof (Lemma 6.7).

Lemma 6.3. *The Gaussian process $\hat{X}_2$ with mean zero and covariance function in* (2.12) *has continuous sample paths.*

Proof. We first show that $\hat{X}_2 \in \mathbb{D}$ by verifying the conditions in Theorem 5.3. The finite-dimensional distributions of $\hat{X}_2$ are Gaussian with the covariance function $\hat{R}_2$ in (2.12). The consistency condition (i) is satisfied because of the Gaussian distributional property. Condition (ii) is satisfied by Lemma 4.4. To check condition (iii), it suffices to show that for all $t \in [0,T)$,

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

By (4.3) in Lemma 4.2, this holds since $H(\cdot, x) \in \mathbb{D}$ for all $x \in \mathbb{R}^k$. Thus we have shown that $\hat{X}_2 \in \mathbb{D}$.

Finally, to show that the Gaussian process $\hat{X}_2 \in \mathbb{C}$, given that $\hat{X}_2 \in \mathbb{D}$, it suffices to show that it is stochastically continuous (Theorem 1 in [16]). It is well known that a real-valued Gaussian process is continuous in quadratic mean if and only if it is stochastically continuous. Continuity in quadratic mean holds by (4.3) and (2.6) under Assumption 2 (iii). The proof is now complete. □

To prove the convergence of the finite-dimensional distributions of $\hat{X}_2^n$ to those of $\hat{X}_2$, we quote the following two lemmas in [10].

Lemma 6.4. *Let $z_1, ..., z_n$ and $w_1, ..., w_n$ be complex numbers of modulus less than 1. Then*

$$\left|\prod_{i=1}^n z_i - \prod_{i=1}^n w_i\right| \le \sum_{i=1}^n |z_i - w_i|.$$

Lemma 6.5. *For $x \in \mathbb{R}$,*

$$\left|e^{ix} - \sum_{m=0}^n \frac{(ix)^m}{m!}\right| \le \min\left\{\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right\}.$$

In particular, if b is a complex number with $|b| \le 1$, then $|e^b - (1+b)| \le |b|^2$.

Lemma 6.6. *The finite-dimensional distributions of the processes $\hat{X}_2^n$ converge to those of $\hat{X}_2$.*

Proof. We need to show that the l-dimensional random variables

$$\left(\hat{X}_2^n(t_j), 1 \le j \le l\right) \Rightarrow \left(\hat{X}_2(t_j), 1 \le j \le l\right) \quad \text{in} \quad \mathbb{R}^l \quad \text{as} \quad n \to \infty, \tag{6.10}$$

for any $0 \le t_1 \le ... \le t_l \le T$ and $l \ge 1$. We first consider the case when $l = 1$ (removing subscript 1 in t_1 for brevity below).

Before proceeding to the proof, for each $n \ge 1$, let the set Υ^n be the collection of the trajectories of $\{A^n(t) : t \ge 0\}$ such that for each $T \ge 0$, $\sup_{0 \le t \le T} |\hat{A}^n(t)| \le n^{1/4}$

and $\max_{1\le i\le A^n(T)}|\tau^n_{i+1}-\tau^n_i|\to 0$ as $n\to\infty$. It is evident that under Assumption 1, $P(\Upsilon^n)=P(A^n\in\Upsilon^n)\to 1$ as $n\to\infty$ and $A^n(t)$ increases without limit and is of order $O(n)$ on Υ^n.

Decompose $\hat{X}^n_2(t)$ as $\hat{X}^n_2(t)=\hat{X}^n_2(t)\mathbf{1}(A^n\in\Upsilon^n)+\hat{X}^n_2(t)\mathbf{1}(A^n\notin\Upsilon^n)$ and observe that for each $\epsilon>0$

$$P\big(|\hat{X}^n_2(t)|\mathbf{1}(A^n\notin\Upsilon^n)>\epsilon\big)\le P(A^n\notin\Upsilon^n)=1-P(\Upsilon^n)\to 0 \quad \text{as} \quad n\to\infty.$$

Thus, we obtain that $\hat{X}^n_2(t)\mathbf{1}(A^n\notin\Upsilon^n)\Rightarrow 0$ as $n\to\infty$. It then suffices to show that $\hat{X}^n_2(t)\mathbf{1}(A^n\in\Upsilon^n)\Rightarrow\hat{X}_2(t)$ in $\mathbb{R}$ as $n\to\infty$.

By the continuity theorem (see, e.g., [10]), it suffices to show that the characteristic function of $\hat{X}^n_2(t)\mathbf{1}(A^n\in\Upsilon^n)$, denoted by $\varphi^n_t(\theta)$, converges pointwise to that of $\hat{X}_2(t)$, denoted by $\varphi_t(\theta)$, and $\varphi_t(\theta)$ is continuous at $\theta=0$. Recall the covariance function of $\hat{X}_2$ in (2.12). For each $t\ge 0$, $\hat{X}_2(t)$ is a normal random variable with mean zero and variance $\int_{[0,t]}\tilde{G}(t,u)d\Lambda(u)$. Thus we have

$$\varphi_t(\theta)=E\big[\exp\big(\mathrm{i}\theta\hat{X}_2(t)\big)\big]=\exp\bigg(-\frac{\theta^2}{2}\int_{[0,t]}\tilde{G}(t,u)d\Lambda(u)\bigg), \tag{6.11}$$

and $\varphi_t(\theta)$ is continuous at $\theta=0$. For $\varphi^n_t(\theta)$, let $\mathcal{A}^n(t):=\sigma(A^n(s):0\le s\le t)\vee\mathcal{N}$ where $\mathcal{N}$ is the collection of P-null sets. Recall the definition of $\hat{X}^n_2$ in (4.2) and denote $\breve{H}^n_i(t):=H(t-\tau^n_i,Z_i(\tau^n_i))-\int_{\mathbb{R}^k}H(t-\tau^n_i,x)dF_{\tau^n_i}(x)$ and $\breve{A}^n(t):=A^n(t)\mathbf{1}(A^n\in\Upsilon^n)$ for brevity in the calculations below,

$$\begin{aligned}
\varphi^n_t(\theta) &= E\big[\exp\big(\mathrm{i}\theta\hat{X}^n_2(t)\mathbf{1}(A^n\in\Upsilon^n)\big)\big]=E\big[E\big[\exp\big(\mathrm{i}\theta\hat{X}^n_2(t)\mathbf{1}(A^n\in\Upsilon^n)\big)\ \big|\ \mathcal{A}^n(t)\big]\big]\\
&= E\bigg[E\bigg[\exp\bigg(\mathrm{i}\theta\frac{1}{\sqrt{n}}\sum_{i=1}^{A^n(t)}\breve{H}^n_i(t)\mathbf{1}(A^n\in\Upsilon^n)\bigg)\bigg|\mathcal{A}^n(t)\bigg]\bigg]\\
&= E\bigg[E\bigg[\exp\bigg(\mathrm{i}\theta\frac{1}{\sqrt{n}}\sum_{i=1}^{\breve{A}^n(t)}\breve{H}^n_i(t)\bigg)\bigg|\mathcal{A}^n(t)\bigg]\bigg]\\
&= E\bigg[E\bigg[\prod_{i=1}^{\breve{A}^n(t)}\exp\bigg(\mathrm{i}\theta\frac{1}{\sqrt{n}}\breve{H}^n_i(t)\bigg)\bigg|\mathcal{A}^n(t)\bigg]\bigg]\\
&= E\bigg[\prod_{i=1}^{\breve{A}^n(t)}E\bigg[\exp\bigg(\mathrm{i}\theta\frac{1}{\sqrt{n}}\breve{H}^n_i(t)\bigg)\bigg|\mathcal{A}^n(t)\bigg]\bigg]\\
&\le E\bigg[\prod_{i=1}^{\breve{A}^n(t)}\bigg[1-\frac{\theta^2}{2n}\tilde{G}(t,\tau^n_i)+\frac{\theta^2}{3!n}\min\bigg\{\frac{\theta E[(\breve{H}^n_i(t))^3|\mathcal{A}^n(t)]}{\sqrt{n}},6\tilde{G}(t,\tau^n_i)\bigg\}\bigg]\bigg],
\end{aligned}$$

where $\prod_{i=1}^0 x_i:=1$ for all $x_i\in\mathbb{R}$ whenever $\breve{A}^n(t)=0$ and the last inequality follows from the first part of Lemma 6.5. Notice that the minimum term above is smaller than $6\tilde{G}(t,\tau^n_i)$ and converges to 0 as $n\to\infty$, and thus it is of order $o(1/n)$ when being multiplied by $\frac{\theta^2}{3n}$. Therefore we can write

$$\varphi^n_t(\theta)=E\bigg[\prod_{i=1}^{\breve{A}^n(t)}\bigg[1-\frac{\theta^2}{2n}\tilde{G}(t,\tau^n_i)+o(n^{-1})\bigg]\bigg].$$

Thus, for large enough n (specified below), we have

$$\begin{aligned}
& |\varphi_t^n(\theta) - \varphi_t(\theta)| \\
= & \left| E\Big[\prod_{i=1}^{\breve{A}^n(t)} \Big[1 - \frac{\theta^2}{2n} \tilde{G}(t, \tau_i^n) + o(n^{-1}) \Big] \Big] - \exp\Big(- \frac{\theta^2}{2} \int_{[0,t]} \tilde{G}(t,u) d\Lambda(u) \Big) \right| \\
\leq & E\Big[\Big| \prod_{i=1}^{\breve{A}^n(t)} \Big[1 - \frac{\theta^2}{2n} \tilde{G}(t, \tau_i^n) + o(n^{-1}) \Big] - \prod_{i=1}^{\breve{A}^n(t)} \exp\Big(- \frac{\theta^2}{2n} \tilde{G}(t, \tau_i^n) \Big) \Big| \Big] \\
& + \left| E\left[\exp\left(-\frac{\theta^2}{2} \mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) \right) \right] - \exp\left(-\frac{\theta^2}{2} \int_{[0,t]} \tilde{G}(t,u) d\Lambda(u) \right) \right| \\
\leq & E\Big[\sum_{i=1}^{\breve{A}^n(t)} \left| \exp\left(-\frac{\theta^2}{2n} \tilde{G}(t, \tau_i^n) \right) - \left(1 - \frac{\theta^2}{2n} \tilde{G}(t, \tau_i^n) \right) \right| \Big] + E[o(1/n) \breve{A}^n(t)] \\
& + \left| E\left[\exp\left(-\frac{\theta^2}{2} \mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) \right) \right] - \exp\left(-\frac{\theta^2}{2} \int_{[0,t]} \tilde{G}(t,u) d\Lambda(u) \right) \right| \\
\leq & \frac{\theta^4}{4n} E\left[\mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}^2(t,u) d\bar{A}^n(u) \right] + o(1) \\
& + \left| E\left[\exp\left(-\frac{\theta^2}{2} \mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) \right) \right] - \exp\left(-\frac{\theta^2}{2} \int_{[0,t]} \tilde{G}(t,u) d\Lambda(u) \right) \right| \\
\to & \ 0, \quad \text{as} \quad n \to \infty.
\end{aligned} \tag{6.12}$$

Here the first inequality is by subtracting and adding the same term and the triangle inequality. The second inequality follows from Lemma 6.4. The third inequality follows from the definition of $\breve{A}^n$(actually Υ^n) and the second part of Lemma 6.5 for large n such that

$$\frac{\theta^2}{2n} \max_{1 \leq i \leq A^n(t)} \tilde{G}(t, \tau_i^n) \leq \frac{\theta^2}{2n} \sup_{0 \leq u \leq t} \tilde{G}(t,u) < 1.$$

Under Assumption 2 (ii), such a large n can always be found.

The final convergence to zero is implied by the facts that $\bar{A}^n \Rightarrow \Lambda$ in $\mathbb{D}$, $P(\Upsilon^n) \to 1$, the continuous mapping theorem and the uniform integrability of the two sequences for each $t \geq 0$:

$$\left\{ \mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) : n \geq 1 \right\},$$

and

$$\left\{ \exp\left(-\frac{\theta^2}{2} \mathbf{1}(A^n \in \Upsilon^n) \int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) \right) : n \geq 1 \right\},$$

since

$$\sup_n E\left[\mathbf{1}(A^n \in \Upsilon^n) \Big(\int_{[0,t]} \tilde{G}(t,u) d\bar{A}^n(u) \Big)^2 \right] < \infty$$

by the definition of Υ^n and Assumption 2 (ii).

Therefore, we have shown that for each fixed $t \geq 0$,

$$\hat{X}_2^n(t) \Rightarrow \hat{X}_2(t) \quad \text{in} \quad \mathbb{R} \quad \text{as} \quad n \to \infty.$$

To generalize to the case $l > 1$, we prove that for any $(\theta_1, ..., \theta_l) \in \mathbb{R}^l$ and $0 \leq t_1 < \cdots < t_l \leq T'$,

$$E\left[\exp\left(\mathrm{i}\sum_{i=1}^{l}\theta_i \hat{X}_2^n(t_i)\right)\right] \to E\left[\exp\left(\mathrm{i}\sum_{i=1}^{l}\theta_i \hat{X}_2(t_i)\right)\right] \quad \text{as} \quad n \to \infty,$$

and the limit is continuous at $(0, ..., 0) \in \mathbb{R}^l$. By definition, $\sum_{i=1}^{l}\theta_i \hat{X}_2(t_i)$ is a normal random variable with mean zero and variance

$$\sum_{i=1}^{l}\sum_{j=1}^{l}\theta_i\theta_j \hat{R}_2(t_i, t_j),$$

for the covariance function $\hat{R}_2$ defined in (2.12). Thus we have

$$E\left[\exp\left(\mathrm{i}\sum_{i=1}^{l}\theta_i \hat{X}_2(t_i)\right)\right] = \exp\left(-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\theta_i\theta_j \hat{R}_2(t_i, t_j)\right), \tag{6.13}$$

and it is continuous at $(0, ..., 0) \in \mathbb{R}^l$.

By definition,

$$\sum_{i=1}^{l}\theta_i \hat{X}_2^n(t_i) = \sum_{i=1}^{l}\theta_i\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{A^n(t_i)}\left(H(t_i - \tau_k^n, Z_k(\tau_k^n)) - \int_{\mathbb{R}^k} H(t_i - \tau_k^n, x)dF_{\tau_k^n}(x)\right)\right).$$

Thus, a direct calculation as in (6.12) shows that

$$\begin{aligned} &E\left[\exp\left(\mathrm{i}\sum_{i=1}^{l}\theta_i \hat{X}_2^n(t_i)\mathbf{1}(\hat{A}^n \in \Upsilon^n)\right)\right] \\ &= E\left[1 - \frac{1}{2}\mathbf{1}(\hat{A}^n \in \Upsilon^n)\sum_{i=1}^{l}\sum_{j=1}^{l}\theta_i\theta_j\int_0^{t_i \wedge t_j}(G_2(t_i, t_j, u) - G_1(t_i, u)G_1(t_j, u))d\bar{A}^n(u)\right] + o(n^{-1}). \end{aligned} \tag{6.14}$$

The convergence of (6.14) to (6.13) can be shown in a similar way as in (6.12) by Lemmas 6.4–6.5. This completes the proof of the convergence of the finite-dimensional distributions. □

Lemma 6.7. *Under Assumptions 1–2, $\hat{X}_2^n \Rightarrow \hat{X}_2$ in $(\mathbb{D}, J_1)$ as $n \to \infty$, where $\hat{X}_2$ is as given in Theorem 2.2.*

Proof. Given the convergence of finite-dimensional distributions of $\hat{X}_2^n$ in Lemma 6.6, by Theorem 13.3 in [5], it suffices to show that for each $\epsilon > 0$

$$\lim_{\delta \to 0} P\left(\left|\hat{X}_2(T) - \hat{X}_2(T - \delta)\right| \geq \epsilon\right) = 0, \tag{6.15}$$

and

$$\lim_{\delta \to 0}\limsup_{n} P\left(\sup_{\substack{0 \leq r < s < t \leq T \\ t - r \leq \delta}}\left|\hat{X}_2^n(r) - \hat{X}_2^n(s)\right| \wedge \left|\hat{X}_2^n(s) - \hat{X}_2^n(t)\right| \geq \epsilon\right) = 0. \tag{6.16}$$

(6.15) is implied by

$$\lim_{\delta\to 0} E\Big[\big|\hat{X}_2(T)-\hat{X}_2(T-\delta)\big|^2\Big]=0. \tag{6.17}$$

By Lemma 4.2,

$$\begin{aligned} & E\Big[\big|\hat{X}_2(T)-\hat{X}_2(T-\delta)\big|^2\Big] \\ = & \int_{(T-\delta,T]} \tilde{G}(T,u)d\Lambda(u)+\int_{[0,T-\delta]} \tilde{G}(T,T-\delta,u)d\Lambda(u). \end{aligned} \tag{6.18}$$

Recall that $\Lambda(\cdot)\in\mathbb{C}$. The first term in the summation above vanishes as δ goes to zero due to absolute continuity of the Lebesgue–Stieltjes integral. The second term vanishes as $\delta\to 0$ by Assumption 2 (iii).

Finally, (6.16) is easily implied by Lemma 4.5 and Theorem 5.2. Specifically, recall Remark 5.1 that the function V induces a set function ν such that $\nu((s,t])=V(s,t)$. By Lemma 4.5, condition (5.10) holds for all $\hat{X}_2^n$ (with $\alpha=\beta=1$ and $\mathcal{T}=[0,T]$). Then by Theorem 5.2, we obtain that

$$\begin{aligned} & P\Bigg(\sup_{\substack{0\le r<s<t\le T\\ t-r\le\delta}} \big|\hat{X}_2^n(r)-\hat{X}_2^n(s)\big|\wedge\big|\hat{X}_2^n(s)-\hat{X}_2^n(t)\big|\ge\epsilon\Bigg) \\ \le & \frac{C_2'}{\epsilon^4}V(0,T)\sup_{0\le t\le T-2\delta} V(t,t+2\delta). \end{aligned} \tag{6.19}$$

By the uniform continuity of V, we obtain (6.16) holds. The proof is now complete. □

Proof of Theorem 2.2. We begin by defining an auxiliary process $\tilde{X}_2^n=\{\tilde{X}_2^n(t):t\ge 0\}$ by

$$\tilde{X}_2^n(t):=\frac{1}{\sqrt{n}}\sum_{i=1}^{[n\Lambda(t)]}\left(H(t-u_i^n,Z_i(u_i^n))-\int_{\mathbb{R}^k} H(t-u_i^n,x)dF_{u_i^n}(x)\right),\quad t\ge 0, \tag{6.20}$$

where $[x]$ denotes the largest integer less than or equal to x, $u_i^n=\Lambda^{-1}(\frac{i}{n})$ for $i=1,...,[n\Lambda(t)]$, and Λ^{-1} is the inverse function of Λ defined by $\Lambda^{-1}(t):=\inf\{u\ge 0:\Lambda(u)\ge t\}$ for $t\ge 0$. Note that, comparing with the definition of $\hat{X}_2^n(t)$ in (4.2), we replace $A^n(t)$ by $[n\Lambda(t)]$ and τ_i^n by u_i^n in the definition of $\tilde{X}_2^n$. Thus, the only source of randomness in $\tilde{X}_2^n$ comes from the sequence $\{Z_i(u_i^n),i\ge 1\}$. All the arguments in Lemmas 6.6 and 6.7 hold true with A^n replaced by $[n\Lambda]$ and associated τ_i^n replaced by u_i^n, since the only requirement on A^n in those lemmas is Assumption 1, which is obviously satisfied by taking $A^n=[n\Lambda]$. Thus, we have

$$\tilde{X}_2^n\Rightarrow\hat{X}_2\quad\text{in}\quad\mathbb{D}\quad\text{as}\quad n\to\infty. \tag{6.21}$$

Moreover, since $\{Z_i(u_i^n),i\ge 1\}$ and A^n are assumed to be mutually independent for each n, $\tilde{X}_2^n$ and $\hat{X}_1^n$ are independent. Thus, we obtain the joint convergence

$$\big(\hat{X}_1^n,\tilde{X}_2^n\big)\Rightarrow\big(\hat{X}_1,\hat{X}_2\big)\quad\text{in}\quad\mathbb{D}^2\quad\text{as}\quad n\to\infty, \tag{6.22}$$

and the limits $\hat{X}_1$ and $\hat{X}_2$ are also independent.

Then, by Lemma 6.7 and (6.21), we obtain that for any $\zeta>0$ and $T>0$,

$$\lim_{n\to\infty} P\Bigg(\sup_{t\in[0,T]}\big|\hat{X}_2^n(t)-\tilde{X}_2^n(t)\big|>\zeta\Bigg)=0.$$

Thus, we have proved the joint convergence

$$\big(\hat{X}_1^n, \hat{X}_2^n\big) \Rightarrow \big(\hat{X}_1, \hat{X}_2\big) \quad \text{in} \quad \mathbb{D}^2 \quad \text{as} \quad n \to \infty.$$

By continuity of addition in $\mathbb{D}$ (e.g., Corollary 12.7.1 in [44]) and the continuous mapping theorem, we obtain that $\hat{X}^n = \hat{X}_1^n + \hat{X}_2^n \Rightarrow \hat{X}_1 + \hat{X}_2$ in $\mathbb{D}$ as $n \to \infty$. This completes the proof. □

6.1. **Proof of Theorem 2.3.** In this subsection, we sketch the proof of Theorem 2.3. For each $j = 1, ..., K$, $\hat{X}^{n,(j)}$ is first decomposed into a summation of $\hat{X}_1^{n,(j)}$ and $\hat{X}_2^{n,(j)}$, defined similarly as in Lemma 4.1. The weak convergence of $\hat{X}_1^{n,(j)}$ to $\hat{X}_1^{(j)}$ in $\mathbb{D}$ follows from similar arguments as that of $\hat{X}_1^n$ to $\hat{X}_1$ in Lemma 6.1–6.2, with the mapping ψ modified accordingly to $\psi^{(j)}$. Thus, the joint weak convergence of $\big(\hat{X}_1^{n,(1)}, ..., \hat{X}_1^{n,(K)}\big)$ follows from the continuity of the joint mapping $\big(\psi^{(1)}, ..., \psi^{(K)}\big)$ from $\mathbb{D}^K$ to $\mathbb{D}^K$ since each component is continuous.

For each $j = 1, ..., K$, by similar arguments in Lemmas 6.6–6.7, $\hat{X}_2^{n,(j)}$ converges weakly to $\hat{X}_2^{(j)}$ in $\mathbb{D}$. Thus, the tightness of the joint process $\big(\hat{X}_2^{n,(1)}, ..., \hat{X}_2^{n,(K)}\big)$ follows from the tightness of each component. It remains to show the convergence of finite dimensional distributions of $\big(\hat{X}_2^{n,(1)}, ..., \hat{X}_2^{n,(K)}\big)$, that is, for any $l \in \mathbb{N}$, we need to show that as $n \to \infty$,

$$\big(\hat{X}_2^{n,(j)}(t_m), 1 \leq j \leq K, 1 \leq m \leq l\big) \Rightarrow \big(\hat{X}_2^{(j)}(t_m), 1 \leq j \leq K, 1 \leq m \leq l\big) \quad \text{in} \quad \mathbb{R}^{lK}.$$

We first consider $l = 1$ and write $t_1 = t$ for brevity. By Cramér-Wold theorem, it is equivalent to prove that for any $(\theta_j, j = 1, ..., K) \in \mathbb{R}^K$,

$$\sum_{j=1}^{K} \theta_j \hat{X}_2^{n,(j)}(t) \Rightarrow \sum_{j=1}^{K} \theta_j \hat{X}_2^{(j)}(t) \quad \text{in} \quad \mathbb{R} \quad \text{as} \quad n \to \infty. \tag{6.23}$$

A simple algebra shows that (6.23) follows from the arguments in Lemma 6.6 with slight changes. It is also evident that the case when $l > 1$ can be proved in the same way as its counterpart in Lemma 6.6 with slight changes. That completes the proof of the finite dimensional distributions.

By constructing auxiliary processes $\big(\tilde{X}_2^{n,(1)}, ..., \tilde{X}_2^{n,(K)}\big)$ as in (6.20), we first obtain that $\big(\tilde{X}_2^{n,(1)}, ..., \tilde{X}_2^{n,(K)}\big)$ converges to the same limit process as $\big(\hat{X}_2^{n,(1)}, ..., \hat{X}_2^{n,(K)}\big)$. We can then conclude the independence and continuity statements in the theorem similarly as in the proof of Theorem 2.2, which completes the proof.

7. Proofs of Theorems 5.1–5.3

Proof of Theorem 5.1. We adapt the proof of Theorem 10.3 in [5] and modify some arguments for the relaxed condition on μ being a finite set function with the superadditive property as given in Definition 5.1. We consider three cases.

Case 1: Suppose that $\mathcal{T} = [0, T]$ and $\mu(0, t]$ is continuous in t.

Suppose further that $\mu(0, t]$ is strictly increasing in t. For each $k \in \mathbb{N}$, let $D_k \subset [0, T]$ be the set of $\{z_i^{(k)} : i \geq 0\}$ such that $z_0^{(k)} = 0$ and

$$\mu\big(0, z_i^{(k)}\big] = \frac{i}{2^k}\mu(0, T]. \tag{7.1}$$

It is easy to see that for each k, the set $D_k \subset D_{k+1}$. Let B_k be the maximum of $|X(s) - X(r)| \wedge |X(t) - X(s)|$ over triples in D_k satisfying $0 \leq r \leq s \leq t \leq T$. Let A_k be the same maximum but with further constraint that r,s,t are adjacent: $\exists i \in \mathbb{N}$ such that $(r,s,t) = (z_{i-1}^{(k)}, z_i^{(k)}, z_{i+1}^{(k)})$. For any $t = z_j^{(k)} \in D_k$, define a point $t' \in D_{k-1}$ by

$$t' = \begin{cases} t & \text{if } t \in D_{k-1}, \\ z^{(k-1)}_{\frac{j-1}{2}} & \text{if } t \notin D_{k-1} \text{ and } |X(t) - X(z_{j-1}^{(k)})| \le |X(t) - X(z_{j+1}^{(k)})|, \\ z^{(k-1)}_{\frac{j+1}{2}} & \text{if } t \notin D_{k-1} \text{ and } |X(t) - X(z_{j-1}^{(k)})| > |X(t) - X(z_{j+1}^{(k)})|. \end{cases} \tag{7.2}$$

Then $|X(t) - X(t')| \le A_k$ for $t \in D_k$. As in the proof of Case 1 for Theorem 10.3 in [5], we obtain that

$$\sup_{0\le r\le s\le t\le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \le 2\sum_{k=1}^{\infty} A_k.$$

Now we need to control $\sum_{k=1}^{\infty} A_k$. Suppose that $0 < \theta < 1$ and choose c such that $c \cdot \sum_{k=1}^{\infty} \theta^k = 1/2$. Then

$$\begin{aligned} & P\left(\sup_{0\le r\le s\le t\le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon \right) \\ \le\ & P\left(2\sum_{k=1}^{\infty} A_k \ge \epsilon \right) \le \sum_{k=1}^{\infty} P\big(A_k \ge c\epsilon\theta^k\big) \\ \le\ & \sum_{k=1}^{\infty}\sum_{i=1}^{2^k} P\left(\big|X(z_{i-1}^{(k)}) - X(z_i^{(k)})\big| \wedge \big|X(z_i^{(k)}) - X(z_{i+1}^{(k)})\big| \ge c\epsilon\theta^k \right). \end{aligned} \tag{7.3}$$

By assumptions and (7.1),

$$\begin{aligned} & P\left(\big|X(z_{i-1}^{(k)}) - X(z_i^{(k)})\big| \wedge \big|X(z_i^{(k)}) - X(z_{i+1}^{(k)})\big| \ge c\epsilon\theta^k \right) \\ \le\ & \frac{C_0}{(c\epsilon\theta^k)^{4\beta}} \big(\mu(z_{i-1}^{(k)}, z_{i+1}^{(k)}]\big)^{2\alpha} \\ \le\ & \frac{C_0}{(c\epsilon\theta^k)^{4\beta}} \big(\mu(0, z_{i+1}^{(k)}] - \mu(0, z_{i-1}^{(k)}]\big)^{2\alpha} \\ =\ & \frac{C_0\big(2\mu(0,T]\big)^{2\alpha}}{(c\epsilon)^4} \left(\frac{1}{\theta^{4\beta} 2^{2\alpha}} \right)^k, \end{aligned} \tag{7.4}$$

where in the second inequality the superadditive property of the set function μ in Definition 5.1 (iii) is used. Therefore, (7.3) becomes

$$\begin{aligned} & P\left(\sup_{0\le r\le s\le t\le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon \right) \\ \le\ & \sum_{k=1}^{\infty} 2^k \frac{C_0\big(2\mu(0,T]\big)^{2\alpha}}{(c\epsilon)^4} \left(\frac{1}{\theta^{4\beta} 2^{2\alpha}} \right)^k \\ =\ & \frac{C_0\big(2\mu(0,T]\big)^{2\alpha}}{(c\epsilon)^4} \sum_{k=1}^{\infty} \left(\frac{1}{\theta^{4\beta} 2^{2\alpha-1}} \right)^k. \end{aligned}$$

Since $4\beta \ge 0$ and $2\alpha - 1 > 0$, there exists a $\theta \in (0,1)$ for which the series above converges, and this shows how to define the constant C_1.

If $\mu(0,t]$ is not strictly increasing in t. Consider first the set function $\mu_\kappa(s,t] := \mu(s,t] + \kappa(t-s)$ for all $0 \le s \le t \le T$ where κ is a positive constant and then let κ go to 0.

Case 2: Suppose that $\mathcal{T}$ is finite. Without loss of generality, we may assume that $\mathcal{T} = \{t_i : 0 \le i \le v\}$ such that

$$0 = t_0 < t_1 < ... < t_v = T.$$

Define the processes $X' := \{X'(t) : t \in \mathcal{T}\}$ by

$$X'(t) := \begin{cases} X(t_i) & \text{if } t_i \le t < t_{i+1}, \quad 0 \le i < v, \\ X(T) & \text{if } t = T. \end{cases} \tag{7.5}$$

It is easy to see that $|X'(r) - X'(s)| \wedge |X'(s) - X'(t)| > 0$ only if r, s and t fall into different subintervals of $[t_i, t_{i+1})$. Suppose that

$$r \in [t_i, t_{i+1}), \quad s \in [t_j, t_{j+1}), \quad t \in [t_k, t_{k+1}), \quad i < j < k.$$

Then by the definition of X' in (7.5) and the assumption in (5.8),

$$\begin{aligned} & P\big(|X'(r) - X'(s)| \wedge |X'(s) - X'(t)| > \epsilon\big) \\ = \; & P\big(|X(t_i) - X(t_j)| \wedge |X(t_j) - X(t_k)| > \epsilon\big) \\ \le \; & \frac{C_0}{\epsilon^{4\beta}} \big(\mu\big(\mathcal{T} \cap (t_i, t_k]\big)\big)^{2\alpha}. \end{aligned} \tag{7.6}$$

Now define an measure ν on $[0,T]$ such that for each $1 \le i \le v-1$, over the interval $[t_i, t_{i+1}]$, ν has a uniform distribution of mass $\mu\big(\mathcal{T} \cap (0, t_{i+1}]\big) - \mu\big(\mathcal{T} \cap (0, t_{i-1}]\big)$, and over the interval $[0, t_1]$, ν has a uniform distribution of mass $\mu\big(\mathcal{T} \cap [0, t_1]\big)$. Note that by definition,

$$\nu(0,T] \le \mu\big(\mathcal{T} \cap (0, t_v]\big) + \mu\big(\mathcal{T} \cap (0, t_{v-1}]\big) \le 2\mu(\mathcal{T}). \tag{7.7}$$

Then by the property (iii) of the set function μ in Definition 5.1,

$$\begin{aligned} \mu\big(\mathcal{T} \cap (t_i, t_k]\big) & \le \mu\big(\mathcal{T} \cap (0, t_k]\big) - \mu\big(\mathcal{T} \cap (0, t_i]\big) \\ & \le \nu(t_{i+1}, t_k] \le \nu(r, t]. \end{aligned}$$

Then (7.6) implies that

$$P\big(|X'(r) - X'(s)| \wedge |X'(s) - X'(t)| > \epsilon\big) \le \frac{C_0}{\epsilon^{4\beta}} \big(\nu(r,t]\big)^{2\alpha}. \tag{7.8}$$

Even though the assumption $t \in [t_k, t_{k+1})$ requires that $t < T$, the inequality above also holds for $t = T$ by similar arguments.

By Theorem 10.3 in [5], (7.7) and the definition of X', we obtain that

$$\begin{aligned} & P\Big(\sup_{0 \le r < s < t \le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon\Big) \\ = \; & P\Big(\sup_{0 \le r < s < t \le T} |X'(r) - X'(s)| \wedge |X'(s) - X'(t)| \ge \epsilon\Big) \\ \le \; & \frac{C_3}{\epsilon^4} \big(\nu(0,T)\big)^{2\alpha} \le \frac{2^{2\alpha} C_3}{\epsilon^{4\beta}} \big(\mu(\mathcal{T})\big)^{2\alpha}, \end{aligned} \tag{7.9}$$

where $C_3 > 0$ is a constant depending only on C_0.

Case 3: Consider the general $\mathcal{T}$ and set function μ. Let $\{T_n\}$ be finite sets

$$T_n := \{0 = t_{n,0} < t_{n,1} < ... < t_{n,v_n} = T\}, \quad n \in \mathbb{N},$$

such that $T_n \subset T_{n+1}$ and $\cup_{n=1}^{\infty} T_n$ is dense in $\mathcal{T}$. Let μ_n be a measure having a mass $\mu\big(\mathcal{T}\cap(0,t_{n,i}]\big)-\mu\big(\mathcal{T}\cap(0,t_{n,i-1}]\big)$ at points $t_{n,i}$. Define the processes $X^{n,\prime} := \{X^{n,\prime}(t) : t \in \mathcal{T}\}$ similarly as X' above, that is,

$$X^{n,\prime}(t) := \begin{cases} X(t_{n,i}) & \text{if } t_{n,i} \le t < t_{n,i+1}, \quad 0 \le i < v_n, \\ X(T) & \text{if } t = T. \end{cases} \tag{7.10}$$

Denote the event

$$E^n_\epsilon := \left\{ \sup_{0\le r<s<t\le T} |X^{n,\prime}(r) - X^{n,\prime}(s)| \wedge |X^{n,\prime}(s) - X^{n,\prime}(t)| \ge \epsilon \right\}$$

for each n and $\epsilon > 0$. Then by *Case 2*,

$$P(E^n_\epsilon) \le \frac{2^{2\alpha} C_3}{\epsilon^{4\beta}} \big(\mu_n(T_n)\big)^{2\alpha} = \frac{2^{2\alpha} C_3}{\epsilon^{4\beta}} \big(\mu(\mathcal{T})\big)^{2\alpha}. \tag{7.11}$$

Since $X^{n,\prime}$ and X are right continuous, by the construction of $X^{n,\prime}$ from X above, we have that

$$\begin{aligned} &\sup_{0\le r<s<t\le T} |X^{n,\prime}(r) - X^{n,\prime}(s)| \wedge |X^{n,\prime}(s) - X^{n,\prime}(t)| \\ &\quad \longrightarrow \sup_{0\le r<s<t\le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \quad a.s. \quad \text{as} \quad n \to \infty. \end{aligned} \tag{7.12}$$

Thus, by (7.12) and Fatou's lemma, for each $\epsilon > \kappa > 0$,

$$\begin{aligned} & P\left(\sup_{0\le r<s<t\le T} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon \right) \\ \le\ & P\big(\liminf_n E^n_{\epsilon-\kappa}\big) \le \liminf_n P(E^n_{\epsilon-\kappa}) \le \frac{2^{2\alpha} C_3}{(\epsilon-\kappa)^{4\beta}} \big(\mu(\mathcal{T})\big)^{2\alpha}. \end{aligned}$$

where the last inequality follows (7.11) . The proof is complete by letting κ go to 0. This completes the proof. □

Proof of Theorem 5.2. Take $v = [T/\delta]$, $t_i = i\delta$ for $0 \le i < v$ and $t_v = T$. Then, by Theorem 5.1, we obtain that for $\epsilon \ge 0$, and for each $1 \le i \le v-1$,

$$P\left(\sup_{t_{i-1}\le r\le s\le t\le t_i} |X(r) - X(s)| \wedge |X(s) - X(t)| \ge \epsilon \right) \le \frac{C_1}{\epsilon^{4\beta}} \big(\mu\big(\mathcal{T}\cap(t_{i-1}, t_{i+1}]\big)\big)^{2\alpha}. \tag{7.13}$$

For the upper bound, we have

$$\begin{aligned} & \sum_{i=1}^{v-1} \big(\mu\big(\mathcal{T}\cap(t_{i-1}, t_{i+1}]\big)\big)^{2\alpha} \\ \le\ & \sum_{i=1}^{v-1} \mu\big(\mathcal{T}\cap(t_{i-1}, t_{i+1}]\big) \times \max_{0\le i\le v-1} \big(\mu\big(\mathcal{T}\cap(t_{i-1}, t_{i+1}]\big)\big)^{2\alpha-1} \\ \le\ & 2\mu(\mathcal{T}) \times \max_{0\le i\le v-1} \big(\mu\big(\mathcal{T}\cap(t_{i-1}, t_{i+1}]\big)\big)^{2\alpha-1}, \end{aligned} \tag{7.14}$$

where the second inequality follows from the superadditive property of the set function μ in Definition 5.1 (iii). We can then conclude the probability bound in (5.11). □

Proof of Theorem 5.3. We modify the proof of Theorem 13.6 in [5] by relaxing the condition (13.15) with our condition (ii).

For each $n \in \mathbb{N}$, consider the points $t_i^n = iT/2^n$ for $i = 0, ..., 2^n$, and let $\tilde{X}^n$ be a random function that is constant over each $[t_{i-1}^n, t_i^n)$ and for which $(\tilde{X}^n(t_0^n), ..., \tilde{X}^n(t_{2^n}^n))$ has the same distribution of $(X(t_0^n), ..., X(t_{2^n}^n))$. Thus, $\tilde{X}^n$ are elements of $\mathbb{D}$. We need to show that the distributions of $\{\tilde{X}^n\}$ are tight and X is the limit in distribution of some subsequence of $\{\tilde{X}^n\}$. To prove tightness of $\tilde{X}^n$, we apply Theorem 13.2 of [5] with the condition (13.5) replaced by (13.8) using the modulus of continuity w'' in (5.3).

We first provide a proof for the first condition in (13.8) of [5], that is, for each $\epsilon > 0$ and $\eta > 0$, there exists a $\delta \in (0, 1)$ and an integer n_0 such that

$$P\big(w''(\tilde{X}^n, \delta) \geq \epsilon\big) \leq \eta, \quad \text{for} \quad n \geq n_0. \tag{7.15}$$

Now let $\tilde{Y}^n$ be the process $\tilde{X}^n$ with the time-set cut back to $T_n = \{t_i^n\}$. Let $\tilde{\mu}_n$ be a finite measure having mass $\mu(0, t_i^n] - \mu(0, t_{i-1}^n]$ at t_i^n for $i = 0, ..., 2^n$. By (5.10), for $0 \leq r \leq s \leq t \leq T$,

$$P\left(\big|\tilde{Y}^n(s) - \tilde{Y}^n(r)\big| \wedge \big|\tilde{Y}^n(t) - \tilde{Y}^n(s)\big| \geq \epsilon\right) \leq \frac{C_0}{\epsilon^{4\beta}}\big(\tilde{\mu}_n(r, t]\big)^{2\alpha}.$$

It follows by Theorem 10.4 in [5] that

$$\begin{aligned} & P\Bigg(\sup_{\substack{0\leq r<s<t\leq T\\ t-r\leq\delta}} \big|\tilde{Y}^n(r) - \tilde{Y}^n(s)\big| \wedge \big|\tilde{Y}^n(s) - \tilde{Y}^n(t)\big| \geq \epsilon\Bigg) \\ \leq\ & \frac{C_4}{\epsilon^{4\beta}}\tilde{\mu}_n(0, T] \sup_{0\leq t\leq T-2\delta}\big(\tilde{\mu}_n(t, t + 2\delta]\big)^{2\alpha-1}, \end{aligned} \tag{7.16}$$

for some constant $C_4 > 0$.

By the definition of $\tilde{\mu}_n$, we have that

$$\tilde{\mu}_n(0, T] = \tilde{\mu}_n(\{t_i^n\}) = \sum_{i=0}^{2^n}\big(\mu(0, t_i^n] - \mu(0, t_{i-1}^n]\big) \leq \mu(0, T], \tag{7.17}$$

which follows from the superadditive property of the set function μ in Definition 5.1 (iii).

By the construction of $\tilde{Y}^n$ from $\tilde{X}^n$, we obtain that when $T/2^n \leq \delta$,

$$w''(\tilde{X}^n, \delta) \leq \sup_{\substack{0\leq r<s<t\leq T\\ t-r\leq 2\delta}} \big\{\big|\tilde{Y}^n(r) - \tilde{Y}^n(s)\big| \wedge \big|\tilde{Y}^n(s) - \tilde{Y}^n(t)\big|\big\}. \tag{7.18}$$

Now by the definition of μ^n, we have that when $T/2^n \leq \delta$,

$$\begin{aligned} \mu^n(t, t + 4\delta] &= \mu^n\big(\{t_i^n : t < t_i^n \leq t + 4\delta\}\big) = \sum_{t_i^n : t < t_i^n \leq t+4\delta}\big(\mu(0, t_i^n] - \mu(0, t_{i-1}^n]\big) \\ &\leq \mu(0, t + 4\delta] - \mu(0, t - T/2^n] \\ &\leq \sup_{0<t-s\leq 5\delta}\big(\mu(0, t] - \mu(0, s]\big), \end{aligned} \tag{7.19}$$

where the two inequalities follows from the superadditive property of the set function μ in Definition 5.1 (iii).

Inequalities (7.16)–(7.19) together with the uniform continuity of $\mu(0, t]$ imply (7.15) holds.

Then, the verifications of the condition (13.4) in Theorem 13.2 of [5], and the second and third conditions in (13.8) of [5] follow exactly the same arguments as in the proof of Theorem 13.6 of [5]. The proof is complete. □

8. Appendix: Proofs of Lemmas 4.2, 4.3, 4.5 and 6.1

Proof of Lemma 4.2. Recall the covariance function $\hat{R}_2$ of the Gaussian process $\hat{X}_2$ defined in (2.12). We obtain that for $0 \le s \le t$,

$$\begin{aligned}
& E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big] \\
= \;& \hat{R}_2(s,s) + \hat{R}_2(t,t) - 2\hat{R}_2(s,t) \\
= \;& \int_{[0,s]} \big(G_2(s,s,u) - G_1^2(s,u)\big) d\Lambda(u) + \int_{[0,t]} \big(G_2(t,t,u) - G_1^2(t,u)\big) d\Lambda(u) \\
& \qquad\qquad -2\int_{[0,s]} \big(G_2(t,s,u) - G_1(t,u)G_1(s,u)\big) d\Lambda(u) \\
= \;& \int_{(s,t]} \tilde{G}(t,u) d\Lambda(u) + \int_{[0,s]} \tilde{G}(t,s,u) d\Lambda(u).
\end{aligned}$$

Since $\hat{X}_2(s) - \hat{X}_2(t)$ is normal and the kurtosis of a normal random variable is 3, we obtain

$$\begin{aligned}
E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^4\big] &= 3\big(E\big[\big|\hat{X}_2(s) - \hat{X}_2(t)\big|^2\big]\big)^2 \\
&= 3\bigg(\int_{(s,t]} \tilde{G}(t,u) d\Lambda(u) + \int_{[0,s]} \tilde{G}(t,s,u) d\Lambda(u)\bigg)^2.
\end{aligned}$$

□

Proof of Lemma 4.3. Let

$$\check{H}_i(s) := H(s - \tau_i^n, Z_i(\tau_i^n)) - \int_{\mathbb{R}^k} H(s - \tau_i^n, x) dF_{\tau_i^n}(x), \quad i \in \mathbb{N}, \quad s \ge 0.$$

(Note that we omit the dependence of $\check{H}_i$ on n for brevity.) By definition, for each $0 \le r \le s$, we have

$$\hat{X}_2^n(s) - \hat{X}_2^n(r) = \frac{1}{\sqrt{n}} \sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s) + \frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(r)} \big(\check{H}_i(s) - \check{H}_i(r)\big). \tag{8.1}$$

Recall $\mathcal{A}^n$ defined in the proof of Lemma 6.6, the two summation terms in (8.1) are conditional independent given $\mathcal{A}^n$, and the conditional expectations of both terms equal to zero a.s. By conditioning and direct calculations, we obtain that

$$E\Bigg[\bigg(\frac{1}{\sqrt{n}} \sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)\bigg)^2\Bigg] = E\left[\int_{(r,s]} \tilde{G}(s,u) d\bar{A}^n(u)\right],$$

and

$$E\Bigg[\bigg(\frac{1}{\sqrt{n}} \sum_{i=1}^{A^n(r)} \big(\check{H}_i(s) - \check{H}_i(r)\big)\bigg)^2\Bigg] = E\left[\int_{[0,r]} \tilde{G}(s,r,u) d\bar{A}^n(u)\right].$$

For the fourth moment, by conditioning and direct calculations, we have

$$E\big[\big|\hat{X}_2^n(s) - \hat{X}_2^n(r)\big|^4\big]$$

$$
\begin{aligned}
&= \frac{1}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)\Big)^4\Big] + \frac{1}{n^2}E\Big[\Big(\sum_{i=1}^{A^n(r)} \big[\check{H}_i(s)-\check{H}_i(r)\big]\Big)^4\Big] \\
&\quad + \frac{6}{n^2}E\Big[\Big(\sum_{i=1}^{A^n(s)} \check{H}_i(s)\Big)^2\Big(\sum_{i=1}^{A^n(r)} \big[\check{H}_i(s)-\check{H}_i(r)\big]\Big)^2\Big] \\
&= \frac{1}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)^4\Big] + \frac{6}{n^2}E\Big[\sum_{i,j=A^n(r)+1,\ i\neq j}^{A^n(s)} \check{H}_i(s)^2\check{H}_j(s)^2\Big] \\
&\quad + \frac{1}{n^2}E\Big[\sum_{i=1}^{A^n(r)} \big(\check{H}_i(s)-\check{H}_i(r)\big)^4\Big] + \frac{6}{n^2}E\Big[\sum_{i,j=1,\ i\neq j}^{A^n(s)} \big(\check{H}_i(s)-\check{H}(r)\big)^2\big(\check{H}_j(s)-\check{H}_j(r)\big)^2\Big] \\
&\quad + \frac{6}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)\Big)^2\Big(\sum_{i=1}^{A^n(r)} \big[\check{H}_i(s)-\check{H}_i(r)\big]\Big)^2\Big] \\
&= \frac{1}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)^4\Big] + \frac{6}{n^2}E\Big[\sum_{i,j=A^n(r)+1,\ i\neq j}^{A^n(s)} \tilde{G}(s,\tau_i^n)\tilde{G}(s,\tau_j^n)\Big] \\
&\quad + \frac{1}{n^2}E\Big[\sum_{i=1}^{A^n(r)} \big(\check{H}_i(s)-\check{H}_i(r)\big)^4\Big] + \frac{6}{n^2}E\Big[\sum_{i,j=1,\ i\neq j}^{A^n(s)} \tilde{G}(s,r,\tau_i^n)\tilde{G}(s,r,\tau_j^n)\Big] \\
&\quad + \frac{6}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)\Big)^2\Big(\sum_{i=1}^{A^n(r)} \big[\check{H}_i(s)-\check{H}_i(r)\big]\Big)^2\Big] \\
&= \frac{1}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \check{H}_i(s)^4\Big] + \frac{3}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n)\Big)^2\Big] - \frac{3}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n)^2\Big] \\
&\quad + \frac{1}{n^2}E\Big[\sum_{i=1}^{A^n(r)} \big(\check{H}_i(s)-\check{H}_i(r)\big)^4\Big] + \frac{3}{n^2}E\Big[\Big(\sum_{i=1}^{A^n(s)} \tilde{G}(s,r,\tau_i^n)\Big)^2\Big] - \frac{3}{n^2}E\Big[\sum_{i=1}^{A^n(s)} \tilde{G}(s,r,\tau_i^n)^2\Big] \\
&\quad + \frac{6}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n)\Big)\Big(\sum_{i=1}^{A^n(r)} \tilde{G}(s,r,\tau_i^n)\Big)\Big] \\
&= \frac{3}{n^2}E\Big[\Big(\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n) + \sum_{i=1}^{A^n(r)} \tilde{G}(s,r,\tau_i^n)\Big)^2\Big] \\
&\quad + \frac{1}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \breve{G}(s,\tau_i^n)\Big] - \frac{3}{n^2}E\Big[\sum_{i=A^n(r)+1}^{A^n(s)} \tilde{G}(s,\tau_i^n)^2\Big] \\
&\quad + \frac{1}{n^2}E\Big[\sum_{i=1}^{A^n(r)} \breve{G}(t,s,\tau_i^n)\Big] - \frac{3}{n^2}E\Big[\sum_{i=1}^{A^n(s)} \tilde{G}(s,r,\tau_i^n)^2\Big]. \qquad (8.2)
\end{aligned}
$$

We further obtain that the first term in (8.2) is equal to

$$3E\bigg[\bigg(\int_{(r,s]}\tilde{G}(s,u)d\bar{A}^n(u)+\int_{[0,r]}\tilde{G}(s,r,u)d\bar{A}^n(u)\bigg)^2\bigg]. \tag{8.3}$$

This completes the proof. □

Proof of Lemma 4.5. First we observe that for any $K\in\mathbb{N}$, $nK=A^n(\tau^n_{nK})$. On $\{A^n(T)\le nK=A^n(\tau^n_{nK})\}$, we have $t=t\wedge\tau^n_{nK}$ for $t\le T$. Thus, $A^n(t)=A^n(t\wedge\tau^n_{nK})$ and $\hat{X}^n_2(t)=\hat{X}^n_2(t\wedge\tau^n_{nK})$ on $\{A^n(T)\le nK\}$.

Now, for $K\in\mathbb{N}$ such that $K>\Lambda(T)$ and $\epsilon>0$,

$$\begin{aligned}
& P\big(\big|\hat{X}^n_2(r)-\hat{X}^n_2(s)\big|\wedge\big|\hat{X}^n_2(s)-\hat{X}^n_2(t)\big|\ge\epsilon\big)\\
\le\ & P(A^n(T)\ge nK)\\
&\quad +P\left(A^n(T)<nK,\big|\hat{X}^n_2(r)-\hat{X}^n_2(s)\big|\wedge\big|\hat{X}^n_2(s)-\hat{X}^n_2(t)\big|\ge\epsilon\right)\\
\le\ & P(\bar{A}^n(T)\ge K)\\
&\quad +\frac{1}{\epsilon^4}E\left[\mathbf{1}(\bar{A}^n(T)\le K)\cdot\big|\hat{X}^n_2(r)-\hat{X}^n_2(s)\big|^2\cdot\big|\hat{X}^n_2(s)-\hat{X}^n_2(t)\big|^2\right]\\
\le\ & P(\bar{A}^n(T)\ge K)\\
&\quad +\frac{1}{\epsilon^4}\left(E\left[\big|\hat{X}^n_2(r\wedge\tau^n_{nK})-\hat{X}^n_2(s\wedge\tau^n_{nK})\big|^4\right]\right)^{1/2}\\
&\qquad\qquad\times\left(E\left[\big|\hat{X}^n_2(s\wedge\tau^n_{nK})-\hat{X}^n_2(t\wedge\tau^n_{nK})\big|^4\right]\right)^{1/2},
\end{aligned} \tag{8.4}$$

where the last inequality is from Cauchy-Schwarz inequality and from the observation that $\hat{X}^n_2(t)=\hat{X}^n_2(t\wedge\tau^n_{nK})$ for $t\le T$ on $\{\bar{A}^n(T)\le K\}$. Since $\bar{A}^n(T)\Rightarrow\Lambda(T)$ as $n\to\infty$ by Assumption 1, we have

$$P(\bar{A}^n(T)\ge K)\to P(\Lambda(T)\ge K)=0\quad\text{as}\quad n\to\infty$$

for the chosen $K>\Lambda(T)$. Therefore, due to (8.4), Lemma 4.5 is implied by

$$E\big[\big|\hat{X}^n_2(r\wedge\tau^n_{nK})-\hat{X}^n_2(s\wedge\tau^n_{nK})\big|^4\big]\le C_5V(r,s)^2, \tag{8.5}$$

for $n\in\mathbb{N}$, $0\le r\le s\le T$ and some positive constant C_5. By similar calculations in Lemma 4.3, we obtain

$$\begin{aligned}
& E\left[\big|\hat{X}^n_2(r\wedge\tau^n_{nK})-\hat{X}^n_2(s\wedge\tau^n_{nK})\big|^4\right]\\
=\ & 3E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\bigg(\int_{(r,s]}\tilde{G}(s,u)d\bar{A}^n(u)+\int_{[0,r]}\tilde{G}(s,r,u)d\bar{A}^n(u)\bigg)^2\bigg]\\
&+\frac{1}{n^2}E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\sum_{i=A^n(r)+1}^{A^n(s)}\breve{G}(s,\tau^n_i)\bigg]-\frac{3}{n^2}E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\sum_{i=A^n(r)+1}^{A^n(s)}\tilde{G}(s,\tau^n_i)^2\bigg]\\
&+\frac{1}{n^2}E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\sum_{i=1}^{A^n(r)}\breve{G}(t,s,\tau^n_i)\bigg]-\frac{3}{n^2}E\bigg[\mathbf{1}(\bar{A}^n(T)\le K)\sum_{i=1}^{A^n(s)}\tilde{G}(s,r,\tau^n_i)^2\bigg]
\end{aligned}$$

On $\{\bar{A}^n(T) \le K\}$,

$$\begin{aligned}
& E\Big[\Big(\int_{(r,s]} \tilde{G}(s,u)d\bar{A}^n(u) + \int_{[0,r]} \tilde{G}(s,r,u)d\bar{A}^n(u)\Big)^2\Big] \\
\xrightarrow{n\to\infty} \; & \left(\int_{(r,s]} \tilde{G}(s,u)d\Lambda(u) + \int_{[0,r]} \tilde{G}(s,r,u)d\Lambda(u)\right)^2 \\
\le \; & \left(\int_{(r,s]} \tilde{G}(s,u)d\Lambda(u) + \int_{[0,T]} \tilde{G}(s,r,u)d\Lambda(u)\right)^2 \\
\le \; & \left(\int_{(r,s]} G_2(s,u)d\Lambda(u) + \int_{[0,T]} \check{G}_2(s,r,u)d\Lambda(u)\right)^2 = V^2(r,s). \qquad (8.6)
\end{aligned}$$

Here the convergence is implied by the uniform integrability (see, e.g., Theorem 3.5 in [5]) of the sequence for each $r \le s$:

$$\left\{\left(\int_{(r,s]} \tilde{G}(s,u)d\bar{A}^n(u) + \int_{[0,r]} \tilde{G}(s,r,u)d\bar{A}^n(u)\right)^2 : n \ge 1\right\}$$

under Assumptions 1 and 2 (ii). (8.6) implies (8.5), which further implies Lemma 4.5. □

Proof of Lemma 6.1. We first prove (i). Fix $0 \le t \le T$. To show that $\psi(z)(t)$ is right continuous at t, let $\{t_k : k \ge 1\}$ converge to t from the right (i.e., $t_k \ge t$ for each k) as $k \to \infty$ and we prove that $\psi(z)(t_k) \to \psi(z)(t)$ as $k \to \infty$.

By the definition of ψ,

$$\begin{aligned}
& \psi(z)(t_k) - \psi(z)(t) \\
= \; & z(t_k)G_1(t_k,t_k) - \int_{(0,t_k]} z(u-)dG_1(t_k,u) \\
& \quad -z(t)G_1(t,t) + \int_{(0,t]} z(u-)dG_1(t,u) \\
= \; & z(t)[G_1(t_k,t) - G_1(t,t)] - \int_{(0,t]} z(u-)d(G_1(t_k,u) - G_1(t,u)) \\
& \quad +z(t_k)G_1(t_k,t_k) - z(t)G_1(t_k,t) - \int_{(t,t_k]} z(u-)dG_1(t_k,u). \qquad (8.7)
\end{aligned}$$

Recall the definition of $G_1(t,u) = \int_{\mathbb{R}^k} H(t-u,x)dF_u(x)$ and the assumption that $H(\cdot,x) \in \mathbb{D}$ for each $x \in \mathbb{R}^k$. It is easy (by the bounded convergence theorem) to see that $G_1(\cdot,u) \in \mathbb{D}$ for each $u \ge 0$. Thus, the first term $z(t)[G_1(t_k,t) - G_1(t,t)]$ converges to 0 as $k \to \infty$.

By Theorem 12.2.2 in [44], any function in $\mathbb{D}$ can be approximated by piecewise-constant functions. In particular, for any $\epsilon > 0$, there exists finitely many points $\bar{t}_i$ such that $0 \equiv \bar{t}_0 < \bar{t}_1 < \cdots < \bar{t}_{m-1} \le \bar{t}_m \equiv t < \bar{t}_{m+1} < \cdots < \bar{t}_M \equiv T$ and z_c is constant on the intervals $[\bar{t}_{i-1}, \bar{t}_i)$, $1 \le i \le M-1$, and $[\bar{t}_{M-1}, T]$ such that $\|z - z_c\|_T \le \epsilon$.

For the second term in (8.7), we can write

$$\int_{(0,t])} z(u-)d(G_1(t_k,u) - G_1(t,u))$$

$$= \int_{(0,t]} z_c(u-)d(G_1(t_k,u) - G_1(t,u)) + \int_{(0,t]} (z(u-) - z_c(u-))d(G_1(t_k,u) - G_1(t,u)).$$

By the definition of z_c, the first integral on the right hand side is equal to

$$\sum_{i=0}^{m} z_c(\bar{t}_{i+1})\big[G_1(t_k,\bar{t}_{i+1}) - G_1(t,\bar{t}_{i+1}) - (G_1(t_k,\bar{t}_i) - G_1(t,\bar{t}_i))\big]. \tag{8.8}$$

Each summand above converges to 0 as $k \to \infty$ by the fact that $G_1(\cdot,u) \in \mathbb{D}$ for each $u \geq 0$. Thus the summation also vanishes when $k \to \infty$. For the second integral, it is bounded by

$$\epsilon \cdot 2 \sup_{0\leq t\leq T} V_0^T(G_1(t,\cdot)),$$

where the coefficient of ϵ is finite under Assumption 2 (i). Since ϵ is arbitrary, we have shown that the second term in (8.7) converges to 0 as $k \to \infty$.

When k is large enough, we have $(t,t_k] \subset [\bar{t}_m(\equiv t), \bar{t}_{m+1})$. In that case, z_c is constant on $[t,t_k]$, yielding 0 if we replace z by z_c in the last line in (8.7). Observe that

$$\begin{aligned}&(z(t_k) - z_c(t_k))G_1(t_k,t_k) - (z(t) - z_c(t))G_1(t_k,t) - \int_{(t,t_k]} (z(u-) - z_c(u-))dG_1(t_k,u)\\ &\leq \Big(2 \sup_{0\leq t,u\leq T} G_1(t,u) + \sup_{0\leq t\leq T} V_0^T(G_1(t,\cdot)\Big) \cdot \epsilon.\end{aligned}$$

The coefficient of ϵ is finite under Assumption 2. Since ϵ is arbitrary, this completes the proof of right continuity.

The existence of a left limit for $\psi(t)$ at each $0 < t \leq T$ follows the similar argument above. In particular, if t_k converges to t from left, then the first term $z(t)[G_1(t_k,t) - G_1(t,t)]$ in (8.7) has a limit since $G_1(\cdot,u) \in \mathbb{D}$ for each $u \geq 0$. Similarly, each summand in (8.8) also has a limit, so does the summation. The last line in (8.7) in this case still converges to 0 (and also has a limit). Thus, $\psi(z)$ has left limit at $0 < t \leq T$. The proof for $\psi(z) \in \mathbb{D}$ is complete.

The claim in (ii) follows directly from the above argument while imposing the conditions $z \in \mathbb{C}$ and $G_1(\cdot,u) \in \mathbb{C}$ for each $u \geq 0$.

For (iii), we need to show that for $z_n \in \mathbb{D}$ and $z \in \mathbb{C}$ and for $T > 0$, if $\|z_n - z\|_T \to 0$ as $n \to \infty$, then $d_{J_1}(\psi(z_n), \psi(z)) \to 0$ as $n \to \infty$. Since the J_1 metric is bounded by the uniform norm (see, e.g., Section 3.3 in [44]), it suffices to prove that $\|\psi(z_n) - \psi(z)\|_T \to 0$ as $n \to \infty$. Recalling that $H(\cdot,x)$ is monotone for each $x \in \mathbb{R}^k$, we obtain

$$\begin{aligned}&\|\psi(z_n) - \psi(z)\|_T\\ \leq\ &\|z_n - z\|_T \sup_{0\leq t\leq T} |G_1(t,t)| + \|z_n - z\|_T \sup_{0\leq t\leq T} V_0^T(G_1(t,\cdot)).\end{aligned} \tag{8.9}$$

Therefore, by the assumptions, the two terms in (8.9) converge to zero as $n \to \infty$. □

Acknowledgement

We thank two anonymous reviewers for their comments and suggestions which have greatly improved our paper. This work is partly supported by an NSF grant (CMMI-1538149), an Army Research Office grant W911NF-17-1-0019, and Marcus endowment grant at Penn State University.

References

[1] D. Anderson, J. Blom, M. Mandjes, H. Thorsdottir, and K. de Turck. (2016) A functional central limit theorem for a Markov-modulated infinite-server queue. *Methodology and Computing in Applied Probability.* 18(1), 153–168.

[2] M. Armony, S. Israelit, A. Mandelbaum, Y. N. Marmor, Y. Tseytlin and G. B. Yom-Tov. (2015) Patient flow in hospitals: a data-based queueing-science perspective. *Stochastic Systems.* 5(1), 146–194.

[3] E. Bacry, S. Delattre, M. Hoffmann and J. F. Muzy. (2013) Scaling limits for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications.* 123(7), 2475–2499.

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

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

[6] A. Budhiraja and P. Nyquist. (2015) Large deviations for multidimensional state-dependent shot noise processes. *J. Appl. Probab.* 52(4), 1097–1114.

[7] L. Decreusefond and D. Nualart. (2008) Hitting times for Gaussian processes. *Annals of Probability.* 36(1), 319–330.

[8] J.-D. Deschênes. *Non-stationary photodetection shot noise in frequency combs: a signal processing perspective.* Ph.D. Disseration. Université Laval.

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

[10] R. Durrett. (2005) *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4th edition.

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

[12] 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.

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

[14] X. Gao and L. Zhu. (2016) A functional central limit theorem for stationary Hawkes process and its application to infinite-server queues. *Working paper.* https://arxiv.org/abs/1607.06624

[15] M. Gyllenberg and D. Silvestrov. (2008) *Quasi-Stationary Phenomena in Nonlinearly Perturbed Stochastic Systems.* Berlin, Boston: De Gruyter.

[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. *Adv. Appl. Prob.* 17(4), 709–730.

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

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

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

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

[22] A. Iksanov, W. Jedidi and F. Bouzeffour. (2016) Functional limit theorems for the number of busy servers in a $G/G/\infty$ queue. *Preprint.* https://arxiv.org/abs/1610.08662

[23] 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.

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

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

[26] D. Koops, M. Mandjes and O. Boxma. (2016) Networks of $\cdot/G/\infty$ server queues with shot-noise-driven arrival intensities. *Working paper.* http://arxiv.org/abs/1608.04924

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

[28] 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.

[29] 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.
[30] 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.
[31] J. A. Lane. (1984) The central limit theorem for the Poisson shot-noise process. *J. Appl. Probab.* 21(2), 287–301.
[32] R. B. Lund, W. P. McCormick and Y. Xiao. (2004) Limiting properties of Poisson shot noise processes. *J. Appl. Probab.* 41(3), 911-918.
[33] C. Macci, G. Stabile and G. L. Torrisi. (2005) Lundberg parameters for nonstandard risk processes. *Scandinavian Actuarial Journal.* 6, 417–432.
[34] W. P. McCormick. (1997) Extremes for shot noise processes with heavy tailed amplitude. *J. Appl. Probab.* 34(3), 643–656.
[35] T. M. Niebauer, R. Schilling, K. Danzmann, A. Rüdiger and W. Winkler. (1991) Nonstationary shot noise and its effect on the sensitivity of interferometers. *Phys. Rev. A.* 43(9), 5022–5029.
[36] G. Pang and Y. Zhou. (2017) Two-parameter process limits for an infinite-server queue with arrival dependent service times. *Stochastic Processes and their Applications.* 127(5), 1375–1416.
[37] A. Papoulis. (1971) High density shot noise and Gaussianity. *J. Appl. Probab.* 18(1), 118–127.
[38] F. Ramirez-Perez and R. Serfling. (2001) Shot noise on cluster processes with cluster marks, and studies of long range dependence. *Adv. Appl. Prob.* 33(3), 631–651.
[39] J. Rice. (1977) On generalized shot noises. *Adv. Appl. Prob.* 9(3), 553–565.
[40] 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.
[41] E. Scalas and N. Viles. (2013) A functional limit theorem for stochastic integrals driven by a time-changed symmetric α-stable Lévy process. *Stochastic Processes and their Applications.* 124(1), 385–410.
[42] V. Schmidt. (1985) On finiteness and continuity of shot noise processes. *Optimization.* 16, 921–933.
[43] M. Westcott. (1976) On the existence of a generalized shot-noise process. *Studies in Probability and Statistics (papers in honor of Edwin J. G. Pitman).* 73–88. North-Holland.
[44] W. Whitt. (2002) *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and Their Applications to Queues.* Springer.
[45] X. Zhang and W. Whitt. (2017) A data-driven model of an emergency department. *Operations Research for Health Care.* 12(1), 1–15.
[46] L. Zhu. (2013) Central limit theorem for nonlinear Hawkes processes. *J. Appl. Prob.* 50, 760–771.