Scaling Limits for Interactive Hawkes Shot Noise Processes

BO LI* AND GUODONG PANG†

Abstract. We introduce an interactive Hawkes shot noise process, in which the shot noise process has a Hawkes arrival process whose intensity depends on the state of the shot noise process via the self-exciting function. Namely, the shot noise process and the Hawkes process are interactive. We prove a functional law of large numbers (FLLN) and a functional central limit theorem (FCLT) for the joint dynamics of shot noise process and the Hawkes process, and characterize the effect of the interaction between them. The FLLN limit is determined by a nonlinear function determined through an integral equation. The diffusion limit is a two-dimensional interactive stochastic differential equation driven by two independent time-changed Brownian motions. The limit of the CLT-scaled shot noise process itself can be also expressed equivalently in distribution as an Ornstein–Uhlenbeck process with time-dependent parameters, unlike being a Brownian motion in the standard case without interaction. The limit of the CLT-scaled Hawkes counting process can be expressed as a sum of two independent terms, one as a time-changed Brownian motion (just as the standard case), and the other as a (Volterra type) Gaussian process represented by an Itô integral with another time-changed Brownian motion, capturing the effect of the interaction in the self-exciting function with the state of the shot noise process. To prove the joint convergence of the co-dependent Hawkes and shot noise processes, the standard techniques for Hawkes processes using the immigration-birth representations and the associated renewal equations are no longer applicable. We develop novel techniques by constructing representations for the LLN and CLT scaled processes that resemble the limits together with the associated residual terms, and then use a localization technique together with some martingale properties to prove the residual terms converge to zero and hence the joint convergence of the scaled processes. We also consider an extension of our model, an interactive marked Hawkes shot noise process, where the intensity of the Hawkes arrivals also depends on an exogenous noise, and present the corresponding FLLN and FCLT limits.

1. Introduction

Shot noise processes have been extensively studied in the literature and used in various applications, for example, insurance and risk models [44, 42, 29, 35, 16], physics [36], telecommunications and queueing theory [12, 6], and so on. The arrival process of shots can be any point process, although it is usually assumed to be a Poisson or renewal process [31, 41, 22, 23]. Shot noise processes are often regarded as a generalization of compound (Poisson or renewal) processes, but capturing the enduring effects of the events over time. Hawkes process is a self-exciting point process, with a stochastic intensity process which depends on the past event times (or the history of the point process). There have been a lot of research activities recently on Hawkes process and its applications in finance, neural science, seismology, queueing and so on, see the recent surveys in [3, 18, 32]. In shot noise processes, the shot response function indicates the enduring effects from shots, and also captures the effect of noises. On the other hand, the self-exciting function indicates the enduring effects from the events, and can also depend on exogenous randomness (often called "marks").

*School of Mathematics, Nankai University, Tianjin, 300071 China

† Department of Computational Applied Mathematics and Operations Research, George R. Brown School of Engineering, Rice University, Houston, TX 77005

E-mail addresses: libo@nankai.edu.cn, gdpang@rice.edu.

Date: April 25, 2025.

Key words and phrases. Self-exciting shot noise process, Hawkes process, interaction, functional law of large numbers, functional central limit theorem, two-dimensional interactive stochastic differential equation.

Hawkes process can be used as the arrival process for a shot noise process, which is referred to as *Hawkes* shot noise process (similarly as *Poisson* shot noise process [31]). The analysis (particularly, scaling limits) for such a process is straightforward, following from the existing results for Hawkes processes and shot noise processes. In this modeling approach, Hawkes process is only used as an input point process, that is, the event times of the shots come from the Hawkes process, while the state of the shot noise process does not affect the dynamics of Hawkes process. Such a decoupling between point processes as arrivals and the state of shot noise processes is common in the existing models. However, in many applications, the state of the shot noise process may also affect the arrival process. For example, the arrival process for certain signal systems may not only be self-exciting, but also depend on the state of the signal transmission dynamics. For another example, shot noise processes can be used to model the cumulative claims with delayed settlements [1], and it is natural that there is certain dependence between the total claims and the claim settlement process.

In this paper, we introduce an *interactive* Hawkes shot noise process, in which Hawkes process, as the arrival process of the shot noise process, also depends on the state of the shot noise process through the self-exciting function. This has not been considered in the literature. The new interactive Hawkes shot noise process has many potential applications, as it is the case in shot noise processes and Hawkes processes mentioned above, particularly capturing the interactive nature of the arrival and state processes. As an example, in Section 3.3, we consider a risk process in which the cumulative claim amount up to each time is modeled as an inactive Hawkes shot noise process, see equation (3.24) and further discussions in that section.

We first show that the joint Hawkes counting process and shot noise process are well defined under mild conditions (Proposition 2.1). When both the self-exciting and shot-response functions are of exponential type, it can be shown that the joint dynamics of the Hawkes process, its intensity process and shot noise process is Markov, in particular, the intensity and shot noise process can be expressed via stochastic differential equations (SDEs) driven by the random walk of noises and the Hawkes jump process (see the discussions in Remark 2.5). However, for general self-exciting and shot-response functions, the exact analysis of the joint dynamics of Hawkes and shot noise processes becomes challenging.

Our focus will be to establish the functional law of large numbers (FLLN) and functional central limit theorem (FCLT) for the appropriately scaled sequence of the interactive Hawkes shot noise processes. We scale up time and space for the joint interactive processes - Hawkes counting process and shot noise process in the asymptotic regime. Functional limit theorems for shot noise processes with Poisson or renewal arrival processes in this asymptotic regime have been established in [29, 22, 23], and those for Hawkes processes have been established in [2]. For shot noise processes with Poisson or renewal arrival processes, because of the scaling in time in the shot response function, the scaled processes are then asymptotically equivalent to a compound process with a functional of noises only, so that the FCLT limit is simply a Brownian motion. For Hawkes processes, under a stability condition on the self-exciting function, it is shown in [2] that the limit is a Brownian motion with a constant variance coefficient.

For the new interactive Hawkes shot noise process, the FLLN limit for the joint Hawkes counting process and shot noise process is determined by a solution to a nonlinear integral equation involving the self-exciting function, which clearly shows the effect of the interaction; see equation (3.9) and the discussions in Remark 3.3. The FCLT limit is given by a two-dimensional interactive stochastic differential equation (SDE, see (3.15)) driven by two independent Brownian motions, which represents the randomness from the counting process and the noises. We also give alternative representations for the limit processes using simply stochastic integrals with respect to the two independent Brownian motions, and calculate the covariance functions of the two limit processes (see the discussions in Remarks 3.6, 3.7 and 3.8). One important observation is that the limit for the shot noise process is equivalent in distribution to an Ornstein-Uhlenbeck diffusion process with time-dependent parameters (see Remark 3.8). This is drastically different from the existing result for shot noise processes with an exogenously given arrival process (which can be shown to be a sum of two independent Brownian motions, resembling the limit for the compound Poisson/renewal process, see, e.g., [29, 22, 23]). Another important observation is that the limit for the Hawkes process, in comparison with that being a Brownian motion in the standard case proved in [2], has a decomposition of two independent terms, one being a time-changed Brownian motion as in the standard case, and the other being a (Volterra type) Gaussian process as an Itô integral with respect to another independent time-changed Brownian motion (see equations (3.18) and (3.19) and discussions in Remark 3.9). This additional independent term indicates the effect of the dependence of the Hawkes process on the shot noise process through the self-exciting function.

The proof for the standard Hawkes processes in [2] relies heavily on the renewal equation representation for the Hawkes counting process (as well as its cumulative intensity process) as a result of the cluster/immigration-birth representation. This has been the fundamental tool to establish functional limit theorems associated with Hawkes processes (see, e.g., [25, 26, 46, 21, 20]). For the new interactive Hawkes shot noise process, these existing renewal functions and methods cannot be directly applied or easily extended. Specifically, because of the dependence of the self-exciting function upon the state of the shot noise process, the birthing mechanism is no longer (conditionally) stationary and only a non-degenerate Volterra type equation for our counting process is expected. The explicit formula for the expectation of the counting process (also that of the cumulative intensity process) cannot be derived. See further detailed discussions in Remark 2.3. In [33], an extension of the renewal equation representation was developed in order to provide an expression for the expectation for the Hawkes counting process with non-stationary exogenous marks. However, that approach relies heavily on the exogeneity of the marks, and cannot be directly applied or possibly extended for our model, since the dependence upon the state of the shot noise process in the self-exciting function is highly endogenous. This forces us to develop new techniques to establish the FLLN and FCLT for the Hawkes process jointly with the shot noise processes.

We develop a novel approach to establish the FLLN and FCLT for the joint scaled processes. We start with an important observation for the LLN-scaled cumulative state-dependent intensity process, that is, given the scaling in time and space, the prelimit process resembles a truncated version of the desired limit because of the inside integral of the state-dependent self-exciting function in the domain that grows as the scaling parameter goes to infinity (see equation (5.3)). Moreover, the shot noise process itself has no Markov structure, and the noises have a memory effects through the shot response function over all time. Because of the scaling in time, the memory effects vanish asymptotically so that the scaled shot noise process itself can be approximated by a compound process with the compound variables of the infinite-time shot response function of noises, which therefore produces another residual term to take care of (see equation (5.4)). These residual terms have co-dependence of the shot noise process and Hawkes process, which causes substantial difficulties to analyze. Moreover, because of the co-dependence, the boundedness of the state processes cannot be pre-determined over all time, so we must employ a localization technique to control the increments of both the Hawkes process and the shot noise processes. Using this technique, together with martingale properties and various maximal inequalities over the localized times, we are able to prove that the LLN-scaled residual processes converge to zero, and hence the convergence of the LLN-scaled joint processes (see details in Section 5.2). For the CLT-scaled processes, the corresponding residual processes carry over in the diffusion scale; see equations (5.6) and (5.7). To prove these CLT-scaled residual terms also vanish turns out to be much harder, which requires more refined estimates the CLT-scaled Hawkes counting process and cumulative intensity process over the localized times. In addition, because of the co-dependence, it is also highly nontrivial to prove the joint convergence of the CLT-scaled Hawkes and shot noise processes; see the details in Section 5.3.

This approach does not involve any renewal equation representation of the Hawkes counting process and the cumulative intensity process as mentioned above. In fact, the technique can be also used to prove the FLLN and FCLT for the standard Hawkes process without invoking the renewal equation as in [2]. The proofs would be much simplified since no co-dependence upon another process with Hawkes process as input is involved. We believe that this new proof technique can be also used for other relevant studies in the future.

We have also considered an extension of the new process, *the interactive "marked" Hawkes shot noise process*, in which the self-exciting function, in addition to the state-dependency, also depends on an exogenous random variable as in the usual marked Hawkes processes [7, 8, 28, 19]. We present the FLLN and FCLT limits for this extended model in Theorem 4.1, specifically the FLLN limit takes exactly the same form with a modified function, and the FCLT limit is also a two-dimensional interactive SDE but driven by a two-dimensional Gaussian martingale process (resulting from the possibly corrected shot noises and random marks for Hawkes arrivals) and an independent Brownian motion term for the randomness of the counting process. The proof techniques we have developed can be applied to this extension, with some modifications to take into account the additional random term, and certain regularity conditions on the extended self-exciting function. The extended model, and the FLLN and FCLT results are stated in Section 4 and the sketch proofs are provided in Section 6. Many further extensions of the new model can be also studied in future work. For example, as studied in [39, 38, 33], both the noises $\{\xi_j\}$ and the marks $\{\eta_j\}$ can be non-stationary, in particular, dependent on the event times, that is, $\{\xi_j(\tau_j)\}$ and $\{\eta_j(\tau_j)\}$. In some sense, $Y(\tau_j-)$ can be also regarded as such a non-stationary noise, but not given exogenously. For another example, one can consider multivariate Hawkes shot noise processes with interactions among the components.

1.1. Organization of the paper.

The rest of the paper is organized as follows. We first describe the model in detail and discuss the basic properties in Section 2. We then present the FLLN and FCLT and discuss the properties of the limit processes in Section 3. The proofs are given in Section 5, with some preliminary results in Section 5.1 and the proofs for the FLLN and FCLT results in Sections 5.2 and 5.3, respectively. The extension to a model with marks is given in Section 4 with their sketch proofs in Section 6.

1.2. Notation.

All the processes and random variables are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ throughout the paper. $\mathbb{N}$ denotes the set of natural numbers, and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. $\mathbb{R}(\mathbb{R}_+)$ denotes the space of real (nonnegative) numbers, and specifically, $\mathbb{R}_+ = [0, \infty)$. Let $\mathbb{D} = \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ denote $\mathbb{R}$-valued function space of all càdlàg functions on $\mathbb{R}_+$, and also denote by $\mathbb{C}$ the subspace of continuous functions of $\mathbb{D}$. $(\mathbb{D}, J_1)$ denotes space $\mathbb{D}$ equipped with Skorohod J_1 topology, see [4], which is complete and separable. We use "u.o.c." as an abbreviation for "uniformly on compact sets". Notations $\to$ and $\Rightarrow$ mean convergence of real numbers and convergence in distribution, respectively. For an integrable variable ξ and an event A, we write $\mathbb{E}[\xi; A] = \mathbb{E}[\xi \times \mathbf{1}(A)]$ with $\mathbf{1}(\cdot)$ being the indicator function. Let ν be Lebesgue-Stieltjes measure induced by the right-continuous increasing function ν on $(0, \infty)$, that is, $\nu(a, b] = \nu(b) - \nu(a)$, we write $\int_a^b f(y)\, d\nu(y) = \int_{(a,b]} f(y)\, d\nu(y)$ and $\int_a^{b-} f(y)\, d\nu(y) = \int_{(a,b)} f(y)\, d\nu(y)$ for every measurable function f.

2. An interactive Hawkes shot noise process

We consider a shot noise process with an extended Hawkes arrival process, whose stochastic intensity depends on the state of the shot noise process through the self-exciting function. Specifically, let (N, Y) be a pair of Hawkes and shot noise processes taking values in $\mathbb{N}_0 \times \mathbb{R}$. The shot noise process $Y = \{Y(t) : t \ge 0\}$ is given by

$$Y(t) = \sum_{j=1}^{N(t)} \varphi\big(t - \tau_j, \xi_j\big) = \sum_{j\ge 1} \varphi\big(t - \tau_j, \xi_j\big)\mathbf{1}\big(\tau_j \le t\big), \tag{2.1}$$

where $\{\xi_j : j \in \mathbb{N}\}$ is a sequence of i.i.d. random variables taking values in a complete separable metric space $\mathcal{E}$ (for example, $\mathbb{R}^d$, $d \ge 1$), $\varphi : \mathbb{R}_+ \times \mathcal{E} \to \mathbb{R}$ is a deterministic and measurable function, referred to as the shot response function, and $\{\tau_j : j \in \mathbb{N}\}$ are the event/jump times associated with the Hawkes process N. The Hawkes process $N = \{N(t) : t \ge 0\}$ has the stochastic intensity $\lambda = \{\lambda(t) : t \ge 0\}$, that is,

$$\mathbb{P}\big(N \text{ has a jump in } [t, t + dt]\big|\mathscr{F}_{t-}\big) = \lambda(t)\, dt,$$

where $\{\mathscr{F}_t\}_{t\ge 0}$ is the filtration generated by N and $\{\xi_j, j \ge 1\}$, that is,

$$\mathscr{F}_t = \sigma\{N(s), 0 \le s \le t\} \vee \sigma\{\xi_j, j \le N(t)\} = \sigma\{(\tau_j, \xi_j), \tau_j \le t\}$$

for each $t \ge 0$, and the intensity $\lambda(t)$ is given by

$$\lambda(t) = \lambda_0 + \sum_{j\ge 1} \phi\big(t - \tau_j, Y(\tau_j -)\big)\mathbf{1}(\tau_j < t) = \lambda_0 + \int_0^{t-} \phi\big(t - s, Y(s-)\big)\, dN(s) \in \mathscr{F}_{t-}, \tag{2.2}$$

where $dN(s)$ is understood as the Lebesgue-Stieltjes integral induced by the right-continuous increasing process N, and where $\lambda_0 > 0$ is a constant, referred to as the baseline intensity, and $\phi : (0, \infty) \times \mathbb{R} \to \mathbb{R}_+$ is a deterministic and measurable function, referred to as the self-exciting function. We assume that the noises $\{\xi_j : j \in \mathbb{N}\}$ are exogenously given with a common distribution function $F(\cdot)$ on $\mathcal{E}$. More specifically, on the event $\{N(s) = k\}$,

$$\mathbb{P}\big(\xi_{k+1} \in B, \tau_{k+1} > t\big|\mathscr{F}_s\big) = \mathbb{P}(\xi \in B)\, \exp\Big(- \int_s^t \lambda_k(u)\, du\Big) \quad \forall t > s, \tag{2.3}$$

where ξ is a generic variable and $\mathscr{B}(\mathcal{E})$ denotes the collection of Borel sets on $\mathcal{E}$, and λ_k denotes the local hazard function for the $(k + 1)^{\text{th}}$-event time, given by

$$\lambda_k(t) = \lambda_0 + \sum_{1\le j\le k} \phi\big(t - \tau_j, Y(\tau_j -)\big) \in \mathscr{F}_{\tau_k} \quad \text{on the set } \{\tau_k < t\}, \tag{2.4}$$

with convention $\sum_{j=1}^0 = 0$. Note that $\lambda(u) = \lambda_k(u)$ on the set $t \in (\tau_k, \tau_{k+1}]$. One can see that $\{\xi_j : j \in \mathbb{N}\}$ is a stationary sequence of unpredictable variables. Let $\Lambda = \{\Lambda(t) : t \ge 0\}$ be the cumulative intensity process, that is, $\Lambda(t) = \int_0^t \lambda(s)\, ds$ for $t \ge 0$. It serves as the compensator for the ground counting process N.

Remark 2.1. It is clear by definition that λ in (2.2) is predictable, and the predictability is used to obtain the martingale property in Proposition 2.2 applying [11, Theorem 14.2.IV]. However, it is more convenient to work with its right-continuous version, that is,

$$\lambda_+(t) = \lambda_0 + \sum_{j\ge 1} \phi\big(t - \tau_j, Y(\tau_j -)\big)\mathbf{1}(\tau_j \le t) = \lambda_0 + \int_0^t \phi\big(t - s, Y(s-)\big)\, dN(s), \tag{2.5}$$

where the integral on the most right-hand side is often used in the literature. Moreover, since $\lambda_+(t) \ne \lambda(t)$ only at $t = \tau_k$, we have

$$\int_0^t f(s)\lambda(s)\, ds = \int_0^t f(s)\lambda_+(s)\, ds$$

holds for every bounded Borel function f. The identity above will be frequently used in the rest of the paper, from which the model stays unchanged with λ_+ in light of (2.3).

Assumption A1. *For all $y \in \mathbb{R}$ and $t > 0$, $\int_0^t \phi(u, y)\, du < \infty$.*

We say that the process (N, Y) is well-defined if the conditional intensity function determines the probability structure of N uniquely (see, e.g., [10, Proposition 7.3.IV]). We also say that the process N is non-explosive if $\tau_n < \infty$ and $\sup_n \tau_n = \infty$ with probability one, c.f. [27, page 280], or equivalently, $N(t) < \infty$ for every $t < \infty$. For the classical Hawkes model studied in [2], the integrability condition (i.e., $\int_0^t \phi(s)ds < \infty$ for each t, noting $\phi(\cdot)$ without the y component) implies the non-explosion property of N, c.f. [2, Lemma 1]. However, for our model (N, Y) described above, the integrability condition in Assumption A1 only ensures the well-posedness of (N, Y) up to the possibly explosive time $\tau_\infty = \lim_{n\to\infty} \tau_n$. In Remarks 2.2 and 2.4 below, we give examples in which $\mathbb{P}(N(t) = \infty) > 0$.

Proposition 2.1. *Under Assumption A1, (2.1), (2.2) and (2.3) uniquely define $\{(\tau_k, \xi_k)\}_{k\geq 1}$ as a regular point process with unpredictable marks on $\mathbb{R}_+ \times \mathcal{E}$ up to the explosive time $\tau_\infty = \lim_{n\to\infty} \tau_n \leq \infty$, and thus, they uniquely define the process (N, Y) on $\mathbb{N} \times \mathbb{R}$.*

Proof. For the well-definedness of the marked point process, we use a similar argument as in the proof of [10, Proposition 7.2.I], that is, a family of Janossy densities for $\{(\tau_k, \xi_k)\}_{k\geq 1}$ is defined and then [10, Proposition 5.3.II(ii)] can be applied.

More specifically, consider a product space $\mathbb{R}_+ \times \mathcal{E}$ with reference measure $F(\cdot)$ on $\mathcal{E}$. For every $T > 0$ and $n \in \mathbb{N}$, let $(t_k, z_k) \in [0, T] \times \mathcal{E}, 1 \leq k \leq n$ be one realization of the first n events of the point process $\{(\tau_k, \xi_k)\}_{k\leq n}$, where $0 = t_0 < t_1 < \cdots < t_n$ is the order statistics of the points of the ground process $N(t)$ on $[0, T]$. The conditional survival function is defined by

$$\begin{aligned} S_1(t) &:= \mathbb{P}(\tau_1 > t) = \exp\Big(-\int_0^t \lambda_0(u)\, du\Big)\,, \\ S_n(t) &:= \mathbb{P}\big(\tau_n > t \big| (t_k, z_k), k \leq n-1\big) \\ &= S_n\big(t\big|(t_k, z_k), k \leq n-1\big) = \exp\Big(-\int_{t_{n-1}}^t \lambda_{n-1}(u)\, du\Big), \quad \forall t > t_{n-1}, \end{aligned} \tag{2.6}$$

which characterizes the conditional tail distribution of the n^{th} jump time of N, and where λ_n is the local hazard function as defined in (2.4) with

$$Y(\tau_j-) = \sum_{i<j} \varphi\big((t_j - t_i)-, z_i\big) \in \mathscr{F}_{\tau_j-} \vee \sigma\{\tau_j\}$$

and we understand that $\lambda_0(t) = \lambda_0, Y(\tau_1-) = 0$. The integrability condition in Assumption A1 ensures the finiteness of the integrals in (2.6) above. Given the conditional survival functions above, $J_0(T)$ and the Janossy densities for $\{(\tau_k, \xi_k)\}_{k\geq 1}$ can be defined by

$$J_0(T) = S_1(T)\,,$$
$$j_n\big((t_1, z_1), \cdots, (t_n, z_n)\big|[0, T] \times \mathcal{E}\big) = \Big(\prod_{k=1}^n S_k(t_k)\lambda_k(t_k)\Big) \cdot S_{n+1}(T),$$

recalling that the reference measure on the mark space $\mathcal{E}$ is $F(\cdot)$. It is then straightforward to check that the symmetric version of Janossy densities above satisfies equation (5.3.9) in [10, Proposition 5.3.II(ii)], which gives the existence of the regular marked point process $\{(\tau_k, \xi_k)\}_{k\geq 1}$.

Moreover, the construction implies $\{(\tau_k, \xi_k)\}_{k\geq 1}$ having an $\{\mathscr{F}_t\}_{t\geq 0}$-intensity given by

$$\lambda(t, z) = \lambda(t) = \lambda_0 + \sum_{k\geq 1} \phi\big(t - \tau_k, y_k\big) \quad \text{with} \quad y_k = \sum_{i<k} \varphi\big((\tau_k - \tau_i)-, \xi_i\big),$$

and satisfies the condition of unpredictable marks in [10, Proposition 7.3.V]. For the last, applying [10, Proposition 7.3.IV] proves the uniqueness of the marked point process.

Given $\{y_k\}_{k\geq 1}$ above, the shot-noise process Y in (2.1) is the φ-interpolation of $\{y_k\}_{k\geq 1}$. Thus, (N,Y) is uniquely defined. □

Remark 2.2. We give an example in which $\mathbb{P}(N(t)=\infty)>0$ for every $t>0$. Consider a degenerate model (N,Y) with $\lambda_0=1$ and $\phi(t,y)=2y+3$ and $\varphi(t,y)=1$ for $t\geq 0$, that is, $Y(t)=N(t)$. It is clear that $\phi(t,y)$ is locally integrable in t and we have

$$\lambda(t)=1+\int_0^{t-}\big(2N(s-)+3\big)\,dN(s)=\big(1+N(t-)\big)^2.$$

One can find that N is itself a time-homogenous Markov process and

$$\begin{aligned}
&\mathbb{P}\big(N(t)=n-1\big|N(s)\geq n-1,\mathscr{F}_s\big)\\
=&\,\mathbf{1}(N(s)=n-1)\mathbb{P}\big(N(t)=n-1\big|N(s)=n-1,\mathscr{F}_s\big)\\
=&\,\mathbf{1}(N(s)=n-1)\mathbb{P}\big(\tau_n>t\big|N(s)=n-1,\mathscr{F}_s\big)=e^{-n^2(t-s)}\,\mathbf{1}(N(s)=n-1)
\end{aligned}$$

where (2.3) is applied and which implies for every $t>s>0$ and $n\in\mathbb{N}$,

$$\mathbb{P}\big(N(t)\geq n\big|N(s)\geq n-1,\mathscr{F}_s\big)\geq 1-e^{-n^2(t-s)}. \tag{2.7}$$

Consider a small $t_0>0$ satisfying

$$\sum_{k\geq n_0+1}k^{-2}\ln\frac{(k+1)(k+2)}{2}<t_0\leq\sum_{k\geq n_0}k^{-2}\ln\frac{(k+1)(k+2)}{2}<\infty$$

for some $n_0\in\mathbb{N}$, and define

$$s_j=t_0-\sum_{k\geq j+1}k^{-2}\ln\frac{(k+1)(k+2)}{2}>0,\quad\forall j\geq n_0.$$

Then $s_\infty=t_0$, we have from (2.7) with $(t,s)=(s_j,s_{j-1})$ in the following that for every $m\geq n_0+1$,

$$\begin{aligned}
&\mathbb{P}\big(N(t_0)\geq m\big|N(s_{n_0})\geq n_0\big)=\mathbb{P}\big(N(s_\infty)\geq m\big|N(s_{n_0})\geq n_0\big)\\
\geq&\,\mathbb{P}\big(N(s_k)\geq k,k=n_0+1,\cdots,m\big|N(s_{n_0})\geq n_0\big)\\
=&\prod_{k=n_0+1}^m\mathbb{P}\Big(N(s_k)\geq k\Big|N(s_j)\geq j,j=n_0,\cdots,k-1\Big)\\
\geq&\prod_{k=n_0+1}^m\big(1-e^{-k^2(s_k-s_{k-1})}\big)=\prod_{k=n_0+1}^m\bigg(1-\frac{2}{(k+1)(k+2)}\bigg)\\
=&\prod_{k=n_0+1}^m\frac{k(k+3)}{(k+1)(k+2)}=\frac{n_0+1}{n_0+3}\frac{m+3}{m+1}\to\frac{n_0+1}{n_0+3}\quad\text{as } m\to\infty,
\end{aligned}$$

which implies $\mathbb{P}(N(t_0)=\infty)>0$ since $\mathbb{P}\big(N(s_{n_0})\geq n_0\big)>0$. By the monotonicity of N, we further have $\mathbb{P}(N(t)=\infty)>0$ for every $t>0$.

Remark 2.3. If the self-exciting function is simply $\phi(t,\cdot)\equiv\phi(t)$, that is, there is no dependence of the Hawkes intensity process on the shot noise process, then N is a standard Hawkes process. By the immigration-birth representations, each individual produces children independently and identically following the same pattern of a ϕ-intensity inhomogeneous Poisson process. The local integrability of ϕ is referred to as the "non-explosion" criterion in [2, Lemma 1]. In this classical case, $\mathbb{E}[N(t)] < \infty$ satisfies a renewal equation

$$\mathbb{E}[N(t)] = \mathbb{E}[\Lambda(t)] = \lambda_0 t + \int_0^t \phi(t-s)\mathbb{E}\big[N(s)\big]\, ds.$$

Further denote the renewal kernel of ϕ by $\psi = \phi + \phi * \psi = \sum_{k\geq 1}\phi^{*k}$, where $f * g(t) := \int_0^t f(t-s)g(s)\, ds$ is the convolution for locally integrable functions f and g on $[0,\infty)$, and ϕ^{*k} is the k^{th} self-convolution of ϕ. It is easy to see that

$$\mathbb{E}\big[N(t)\big] = \lambda_0\Big(t + \int_0^t \psi(t-s)s\, ds\Big) = \lambda_0 \int_0^t \Big(1 + \int_0^s \psi(u)\, du\Big)\, ds, \tag{2.8}$$

c.f., [2, Lemma 4], where Fubini's theorem is applied in the second identity. From this expression, we observe that $\mathbb{E}[N(t)] \sim \frac{\lambda_0 t}{1-||\phi||_1}$ if and only if $||\phi||_1 := \int_0^\infty \phi(u)\, du \in (0,1)$. This condition $||\phi||_1 \in (0,1)$ is referred to as the stability condition for the Hawkes process N.

Moreover, for the study of standard Hawkes processes, the following martingale representation plays a crucial role: $X = N - \Lambda$ is a martingale with respect to the natural filtration with $\mathbb{E}[N(t)] < \infty$ for each $t > 0$, c.f. [10, Lemma 7.2.V] and [2, section 2]. Furthermore, with the renewal kernel above, we have

$$\Lambda(t) - \mathbb{E}\big[\Lambda(t)\big] = \int_0^t \psi(t-s)X(s)\, ds,$$

as shown in [2, Lemma 4].

However, because of the interaction between the Hawkes process and the shot noise process, these crucial expressions using renewal equations can no longer be derived, and hence, it is not easy to conclude the finiteness of $\mathbb{E}[N(t)]$ in general. We nevertheless have a local martingale property as stated below, which facilitates us to use localization arguments in the proofs in the paper.

Proposition 2.2. *Under Assumption A1, $X := N - \Lambda$ is an $\{\mathscr{F}_t\}_{t\geq 0}$-local martingale.*

Proof. By the construction of the process N in the proof of Proposition 2.1, we can apply [11, Theorem 14.2.IV] to show that Λ is the compensator for the process N, and the local martingale property is then implied by its definition [11, Proposition 14.2.I]. □

Remark 2.4. In this remark, we give a sufficient condition on ϕ, under which the process X is a $\{\mathscr{F}_t\}$-martingale. Namely, if $\sup_y \int_0^t e^{-\delta_0 u}\phi(u,y)\, du < 1$ for some $\delta_0 > 0$, then we have

$$\mathbb{E}\big[N(t)\big] < \infty \text{ and } X \text{ is a } \{\mathscr{F}_t\}\text{-martingale.}$$

Since X is a local martingale with quadratic variation N and $N(t\wedge\tau_k) \leq k$, one can have from the Burkholder-Davis-Gundy (BDG) inequality that $X(\cdot\wedge\tau_k)$ is a $\{\mathscr{F}_t\}$-martingale, c.f. the proof of Lemma 5.1, and so is $\int_0^t e^{-\delta_0 s}\, dX(s\wedge\tau_k)$. From this, one can obtain that for $t > 0$,

$$\mathbb{E}\Big[\int_0^{t\wedge\tau_k} e^{-\delta_0 s}\, dN(s)\Big] = \mathbb{E}\Big[\int_0^{t\wedge\tau_k} e^{-\delta_0 s}\lambda(s)\, ds\Big].$$

On the other hand, by Fubini's theorem, we have

$$\int_0^{t\wedge\tau_k} e^{-\delta_0 s}\lambda(s)\, ds = \lambda_0 \int_0^{t\wedge\tau_k} e^{-\delta_0 s}\, ds + \int_0^{t\wedge\tau_k} e^{-\delta_0 u}\int_0^u \phi\big(u-v, Y(v-)\big)\, dN(v)\, du.$$

Fixing $t > 0$, by Fubini's theorem, the second integral above equals

$$\int_0^{t\wedge\tau_k} dN(v) \int_v^{t\wedge\tau_k} e^{-\delta_0 u}\phi\big(u, Y(v-)\big)\, du$$
$$= \int_0^{t\wedge\tau_k} e^{-\delta_0 v}\, dN(v)\bigg(\int_0^{t\wedge\tau_k - v} e^{-\delta_0 u}\phi(u,y)\, du\bigg)\bigg|_{y=Y(v-)} \leq \alpha \int_0^{t\wedge\tau_k} e^{-\delta_0 v}\, dN(v),$$

for some $\alpha \in (0,1)$ by assumption, which gives

$$(1-\alpha)\cdot\mathbb{E}\bigg[\int_0^{t\wedge\tau_k} e^{-\delta_0 v}\, dN(v)\bigg] \leq \lambda_0 \int_0^\infty e^{-\delta_0 s}\, ds = \frac{\lambda_0}{\delta_0}$$

This implies that $\mathbb{E}\big[N(t)\big] \leq \frac{\lambda_0}{\delta_0(1-\alpha)} e^{\delta_0 t} < \infty$, and hence X is a $\{\mathscr{F}_t\}$-martingale.

Remark 2.5. When the shot response function is exponential and the arrival process is Poisson, the shot noise process is Markovian and can be expressed as a stochastic differential equation (SDE) [43, Example 2.1]. When the self-exciting function is exponential for the standard Hawkes process, the joint Hawkes counting and intensity process is also Markovian, and the intensity process can be expressed as an SDE [3, Proposition 2].

For the interactive Hawkes shot noise process, if the shot response and self-exciting functions are exponential, we can show that the joint process (Y, N, λ_+) is Markov and the processes Y and λ_+ can be also expressed using SDEs, where λ_+ is the right-continuous version in Remark 2.1. Specifically, let

$$\varphi(t,z) = 1 - \alpha e^{-\gamma t}\varphi(z) \quad \text{and} \quad \phi(t,y) = \kappa e^{-\kappa t} H(y)$$

for some $\kappa, \gamma > 0, \alpha \in \mathbb{R}$ and deterministic functions $H : \mathbb{R} \to \mathbb{R}_+$ and $\varphi : \mathcal{E} \to \mathbb{R}$. For the process (λ_+, Y) in (2.5) and (2.1), we have

$$\lambda_+(t) = \lambda_0 + \int_0^t \kappa e^{-\kappa(t-u)} H\big(Y(u-)\big)\, dN(u)\,,$$
$$Y(t) = N(t) - \alpha \int_0^t e^{-\gamma(t-u)}\, dW_\varphi\big(N(u)\big)\,,$$

where $W_\varphi(m) := \sum_{j=1}^m \varphi(\xi_j)$. By considering the dynamics for $\lambda_+ - \lambda_0$ and $Y - N$, one can obtain directly the following SDEs:

$$d\lambda_+(t) = \kappa\big(\lambda_0 - \lambda_+(t)\big)\, dt + \kappa H(Y(t-))\, dN(t)\,,$$
$$dY(t) = \gamma(Y(t) - N(t))\, dt + (1-\alpha)\, dW_\varphi(N(t))\,.$$

Hence, we obtain the Markovian property for (Y, N, λ_+).

Remark 2.6. Our model is very different from the (affine) self-exciting shot noise processes that have been considered in the literature [14, 43]: the pair (N, λ) represents a counting process with intensity λ, which is given by $d\lambda_t = \kappa(\theta - \lambda_t)\, dt + dN_t$. It can be shown that $\lambda_t = e^{-\kappa t}\lambda_0 + \theta(1 - e^{-\kappa t}) + \sum_{i=1}^{N_t} e^{-\kappa(t-\tau_i)}$, where τ_i's are the jump times of N_t. With $\lambda_0 = \theta = 0$, $\lambda(t)$ is often called as an (affine and Markovian) self-exciting shot noise process. Our model could be also called a self-exciting shot noise process in some sense. That is also one reason that we call our process as the interactive Hawkes shot noise process.

3. Functional limit theorems

In this paper, we aim to establish asymptotic approximations by establishing FLLN and FCLT for (N, Λ, Y). We will consider a sequence of interactive Hawkes shot noise processes $(N^{(n)}, Y^{(n)})$, indexed by n, that is,

$$Y^{(n)}(t) = \sum_{j\geq 1} \varphi^{(n)}\big(t - \tau_j^{(n)}, \xi_j^{(n)}\big)\,,$$

where we understand that $\varphi^{(n)}(t,z) = 0$ for $t < 0$, and $N^{(n)}$ has the stochastic intensity given by

$$\lambda^{(n)}(t) = \lambda_0^{(n)} + \int_0^{t-} \phi^{(n)}\big(t-u, Y^{(n)}(u-)\big)\, dN^{(n)}(u). \tag{3.1}$$

In addition, let $\Lambda^{(n)}(t)$ be the stochastic cumulative intensity process, that is,

$$\Lambda^{(n)}(t) = \int_0^t \lambda^{(n)}(u)\, du = \lambda_0^{(n)} t + \int_0^t \int_0^u \phi^{(n)}\big(u-v, Y^{(n)}(v-)\big)\, dN^{(n)}(v) du.$$

The FLLN and FCLT will be established under proper scalings of these processes.

3.1. FLLN.

Define the LLN-scaled processes

$$\big(\bar{N}^{(n)}(t), \bar{\Lambda}^{(n)}(t), \bar{Y}^{(n)}(t)\big) = \frac{1}{n}\big(N^{(n)}(nt), \Lambda^{(n)}(nt), Y^{(n)}(nt)\big), \quad t \geq 0. \tag{3.2}$$

We make the following assumptions.

Assumption A2. *Suppose that $\lambda_0^{(n)} = \lambda_0$, and the functions $\phi^{(n)}$ and $\varphi^{(n)}$ satisfy the following conditions.*

(1) For some deterministic and measurable function $\phi : (0,\infty) \times \mathbb{R} \to \mathbb{R}_+$,

$$\phi^{(n)}(t, ny) = \phi(t, y), \tag{3.3}$$

and for every $k > 0$, ϕ satisfies

$$I_k(t) := \sup_{|y|\leq k} \int_t^\infty \phi(u,y)\, du \to 0 \quad \text{as} \quad t \to \infty. \tag{3.4}$$

(2) Suppose $\varphi^{(n)}(t,z) = \varphi(t,z)$ for some measurable function φ satisfying

$$\big|\varphi(\infty, z) - \varphi(t,z)\big| \leq J(t) \cdot g(z) \tag{3.5}$$

where $J \in \mathbb{D}(\mathbb{R}_+)$ is a decreasing function with $J(0) = 1, J(\infty) = 0$ and $g : \mathcal{E} \to \mathbb{R}_+$.

(3) The noises $\{\xi_j^{(n)} : j \in \mathbb{N}\}$ are exogenously given with distribution $F(\cdot)$; with ξ denoting a generic variable for $\xi_j^{(n)}$, assume that the associated mean values satisfy

$$\mu_\varphi := \mathbb{E}\big[\varphi(\infty, \xi)\big] \in [0,\infty) \quad \text{and} \quad \mu_g := \mathbb{E}\big[g(\xi)\big] \in [0,\infty). \tag{3.6}$$

Remark 3.1. We remark that the condition on the scaling of $\phi^{(n)}$ in (3.3) is necessary since we expect some nontrivial convergence of $\bar{Y}^{(n)}$. As a result, for the scaled processes $\big(\bar{N}^{(n)}(t), \bar{\Lambda}^{(n)}(t), \bar{Y}^{(n)}(t)\big)$, the sequence $(N^{(n)}, \lambda^{(n)}, Y^{(n)})$ inherently contains a scaling through $\phi^{(n)}$ in the component of $Y^{(n)}$ specified in (3.3). Specifically, by Fubini's theorem and change of variables, we obtain from Remark

2.1 and Assumption A2-(1) that

$$\begin{aligned}\bar\Lambda^{(n)}(t) &= \lambda_0^{(n)}\, t + \frac{1}{n}\int_0^{nt}\Big(\int_0^u \phi^{(n)}\big(u-v, Y^{(n)}(v-)\big)\, dN^{(n)}(v)\Big)\, du \\ &= \lambda_0\, t + \int_0^t \Big(\int_0^u n\phi\big(n(u-v), \bar Y^{(n)}(v-)\big)\, d\bar N^{(n)}(v)\Big)\, du \\ &= \lambda_0\, t + \int_0^t \Big(\int_v^t n\phi\big(n(u-v), \bar Y^{(n)}(v-)\big)\, du\Big)\, d\bar N^{(n)}(v) \\ &= \lambda_0\, t + \int_0^t \Big(\int_0^{n(t-v)} \phi\big(u, \bar Y^{(n)}(v-)\big)\, du\Big)\, d\bar N^{(n)}(v)\,.\end{aligned} \tag{3.7}$$

This expression provides a hint on the proof; see also (5.3). The assumptions in (3.4) and (3.5) are only technicality conditions for the proof of convergence. This is because of the lack of the renewal equation representations as discussed in Remark 2.3. If ϕ takes the form $\phi(t,y) = \tilde h(t)\phi(y)$, as a multiplication of the two functions $\tilde h$ and ϕ, then the condition (3.4) requires that $\sup_{|y|\le k}\phi(y) < \infty$ and $\int_t^\infty \tilde h(u)du \to 0$ as $t\to\infty$, which are simply the local boundedness for ϕ and the integrability condition for $\tilde h$. It is clear that a multiplicative functions of φ, such as $\varphi(t,z) = c(t)\,\varphi(z)$ for some functions c and φ, will satisfy the condition (3.5) if $c(t)\to 1$ and $c(t)$ is increasing. Clearly, linear combinations of such multiplicative functions will also satisfy the condition. However, both conditions are more general than the usual multiplicative examples in the literature. For example, if $\varphi(t,z) = (1-\alpha e^{-\gamma t^{\ell(z)}})\varphi(z)$ for some $\ell(z) > 1$ and $\gamma > 0$ (a modified version of the exponential case in Remark 2.5), then $\varphi(\infty, z) = \varphi(z)$ and $|\varphi(\infty,z) - \varphi(t,z)| = \alpha e^{-\gamma t^{\ell(z)}}|\varphi(z)| \le \alpha e^{-\gamma t \mathbf{1}(t>1)}|\varphi(z)|$.

Assumption A3. *Assume that* $y \to H(y) := \int_0^\infty \phi(t,y)\, dt < 1$ *is continuous on* $\mathbb{R}$ *and*

$$\int_0^\infty \big(1 - H(\mu_\varphi\, y)\big)\, dy = \infty. \tag{3.8}$$

It is worth noting that Assumptions A3 and A2-(1) imply the condition in Assumption A1.

Remark 3.2. The condition $H(\cdot) < 1$ resembles the stability condition for standard Hawkes processes, the L_1 norm of the self-exciting function ϕ without dependence on y being strictly less than 1, that is, $||\phi||_1 < 1$ as discussed in Remark 2.3.

Theorem 3.1 (FLLN)**.** *Under Assumptions A2 and A3,*

$$\big(\bar N^{(n)}, \bar\Lambda^{(n)}, \bar Y^{(n)}\big) \to \big(\bar N, \bar\Lambda, \bar Y\big) \quad \textit{u.o.c. in probability as } n\to\infty,$$

where $\bar N = \bar\Lambda$ *and* $(\bar\Lambda, \bar Y)$ *is the unique solution to the equation*

$$\bar Y(t) = \mu_\varphi\, \bar\Lambda(t) \quad \textit{and} \quad \int_0^t \big(1 - H(\bar Y(s))\big)\, d\bar\Lambda(s) = \lambda_0\, t. \tag{3.9}$$

Remark 3.3. We remark that $H(\cdot) < 1$ in Assumption A3 ensures $\int_0^t (1 - H(\mu_\varphi y))\, dy < \infty$, which, together with the condition in (3.8), ensures the well-definedness of $\bar\Lambda$ in (3.9). The function $\bar\Lambda$ is clearly a finite, continuous and strictly increasing function with $\bar\Lambda(\infty) = \infty$. In particular, if $\mu_\varphi = 0$, (3.8) is always satisfied and

$$\bar\Lambda(t) = \frac{\lambda_0\, t}{1 - H(0)}. \tag{3.10}$$

If $\mu_\varphi > 0$, (3.8) is equivalent to $\int_0^\infty \big(1 - H(y)\big)\, dy = \infty$.

Nevertheless, if $\mu_\varphi > 0$ and Assumption A3 fails to hold, then the FLLN holds locally on $[0, t_0)$ where

$$t_0 = \frac{1}{\lambda_0 \mu_\varphi} \int_0^\infty \big(1 - H(y)\big)\, dy < \infty,$$

and $\bar\Lambda(t) \to \infty$ as $t \uparrow t_0$, that is, t_0 will be the explosive time for the limit. For example, if $1 - H(y) = \alpha e^{-\beta y}$ for some $\alpha \in (0,1), \beta > 0$, then (3.8) fails to hold.

Remark 3.4. It is also clear that if the self-exciting function ϕ in (2.2) is independent of the shot noise process (that is, $\phi(t,\cdot) \equiv \phi(t)$ for some $\phi \geq 0$) as in Remark 2.3, then the limit $\bar N(t)$ reduces to the well-known result for standard Hawkes process, that is, $\bar N(t) = \bar\Lambda(t) = \frac{\lambda_0 t}{1 - \|h\|_1}$ (see, e.g., [2]), and the limit $\bar Y(t)$ becomes the well-known result for shot noise processes, that is, $\bar Y(t) = \mu_\varphi \bar\Lambda(t) = \frac{\mu_\varphi \lambda_0 t}{1-\|h\|_1}$ (in the literature it is often assumed that $\mu_\varphi = 0$ so that $\bar Y(t) = 0$). In fact, for a general arrival process $N(t)$ satisfying $N(nt)/n \to \bar\Lambda(t)$ u.o.c. in probability, it can be shown that $\bar S^{(n)}(t) = \frac{1}{n}\sum_{i=1}^{N(nt)} \varphi(nt - \tau_i, \xi_i) \to \bar S(t) = \mathbb{E}[\varphi(\infty,\xi)] \cdot \bar\Lambda(t)$, u.o.c. in probability by adapting the arguments in [29] and [39]. We refer the readers to [46] for the functional limit theorems for the shot noise process with a Hawkes arrival process (no interaction).

3.2. **FCLT.** Define the CLT-scaled processes

$$\big(\hat N^{(n)}, \hat\Lambda^{(n)}, \hat Y^{(n)}\big) = \sqrt{n}\big(\bar N^{(n)} - \bar N, \bar\Lambda^{(n)} - \bar\Lambda, \bar Y^{(n)} - \bar Y\big), \tag{3.11}$$

where $\big(\bar N, \bar\Lambda, \bar Y\big) = (1, 1, \mu_\varphi)\,\bar\Lambda$ is given in (3.9). We impose the following conditions in addition to Assumptions A2 and A3.

Assumption A4. *Recall $I_k(t)$ and $J(t)$ in (3.4) and (3.5), respectively.*

(1) $H \in C^1(\mathbb{R})$ and for every $k > 0$,

$$\sqrt{t} \cdot I_k(t) \to 0 \quad \text{and} \quad \sqrt{t} \cdot J(t) \to 0 \quad \text{as} \quad t \to \infty. \tag{3.12}$$

(2) For the variable ξ, the associated variances satisfy

$$\sigma_\varphi^2 := Var\big(\varphi(\infty,\xi)\big) \in (0,\infty) \quad \text{and} \quad \sigma_g^2 := Var\big(g(\xi)\big) \in (0,\infty). \tag{3.13}$$

Remark 3.5. Assumption A4-(3.12) is a necessary condition to prove the convergence for the CLT-scaled processes, as in [34]. For the random partial sums associated with the noises $\{\xi_j\}$, that is, $\sum_{j=1}^n \varphi(\infty,\xi_j)$ and $\sum_{j=1}^n g(\xi_j)$, Assumption A2-(3.6) ensures that the corresponding FLLNs hold with a linear function limit (see equation (5.23)), while Assumption A4-(3.13) ensures the FCLTs with a Brownian motion limit (see Lemma 5.4). These Assumptions provide some "smoothness" for the noises terms, which is used to show the residual terms in the LLN and CLT scaled processes converges to 0.

Theorem 3.2 (FCLT)**.** *Under Assumptions A2–A4,*

$$\big(\hat N^{(n)}, \hat\Lambda^{(n)}, \hat Y^{(n)}\big) \Rightarrow \big(\hat N, \hat\Lambda, \hat Y\big) \quad \text{in} \quad (\mathbb{D}^3, J_1) \quad \text{as} \quad n \to \infty, \tag{3.14}$$

where $(\hat N, \hat\Lambda, \hat Y)$ is the unique solution to the following SDE:

$$\begin{aligned} d\hat N(t) &= \frac{1}{1 - H(\bar Y(t))}\, d\hat X(t) + \frac{\lambda_0 H'(\bar Y(t))}{(1 - H(\bar Y(t)))^2} \hat Y(t)\, dt\,, \\ d\hat\Lambda(t) &= d\hat N(t) - d\hat X(t)\,, \\ d\hat Y(t) &= d\hat Z_\varphi(t) + \mu_\varphi\, d\hat N(t)\,, \end{aligned} \tag{3.15}$$

with $\hat N(0) = \hat\Lambda(0) = \hat Y(0) = 0$, where $\bar Y$ and $\bar\Lambda$ are given by (3.9), and $\hat X$ and $\hat Z_\varphi$ are two independent time-changed Brownian motions, with zero drift and variances $\bar\Lambda(t)$ and $\sigma_\varphi^2 \cdot \bar\Lambda(t)$, respectively, that *is,* $\hat{X}(t) \stackrel{\mathrm{d}}{=} B_1(\bar{\Lambda}(t))$ *and* $\hat{Z}_\varphi(t) \stackrel{\mathrm{d}}{=} \sigma_\varphi B_2(\bar{\Lambda}(t))$ *for two independent standard Brownian motions* B_1 *and* B_2*.*

Remark 3.6. The limit process $(\hat{N}, \hat{Y})$ can be represented as a functional of $(\hat{X}, \hat{Z}_\varphi)$ in the theorem. By the fact that $\bar{Y}(t) = \mu_\varphi \bar{\Lambda}(t)$ and

$$\bar{\Lambda}'(t) = \frac{\lambda_0}{1 - H(\bar{Y}(t))} = \frac{\lambda_0}{1 - H(\mu_\varphi \bar{\Lambda}(t))} \tag{3.16}$$

which follows from (3.9). One can check from (3.15) that

$$\begin{aligned} d\big((1 - H(\bar{Y}(t)))\hat{N}(t)\big) &= (1 - H(\bar{Y}(t)))\, d\hat{N}(t) - \mu_\varphi \hat{N}(t) H'(\bar{Y}(t))\bar{\Lambda}'(t)\, dt \\ &= d\hat{X}(t) + H'(\bar{Y}(t))\bar{\Lambda}'(t)\cdot \hat{Y}(t)\, dt - H'(\bar{Y}(t))\bar{\Lambda}'(t)\cdot(\mu_\varphi \hat{N}(t))\, dt \\ &= d\hat{X}(t) + \bar{\Lambda}'(t)H'(\bar{Y}(t))\cdot \hat{Z}_\varphi(t)\, dt\,, \end{aligned}$$

and

$$\begin{aligned} d\big((1 - H(\bar{Y}(t)))\hat{Y}(t)\big) &= (1 - H(\bar{Y}(t)))\, d\hat{Y}(t) - \mu_\varphi \hat{Y}(t) H'(\bar{Y}(t))\bar{\Lambda}'(t)\, dt \\ &= (1 - H(\bar{Y}(t)))\, d\big(\hat{Z}_\varphi(t) + \mu_\varphi \hat{N}(t)\big) - \mu_\varphi \frac{\lambda_0 H'(\bar{Y}(t))}{1 - H(\bar{Y}(t))}\hat{Y}(t)\, dt \\ &= (1 - H(\bar{Y}(t)))\, d\hat{Z}_\varphi(t) + \mu_\varphi\, d\hat{X}(t)\,. \end{aligned}$$

Hence, we obtain a second expression for $\hat{N}, \hat{Y}$ as follows: for $t \geq 0$,

$$\begin{aligned} \hat{N}(t) &= \frac{1}{1 - H(\bar{Y}(t))}\Big(\hat{X}(t) + \int_0^t \frac{\lambda_0 H'(\bar{Y}(s))}{1 - H(\bar{Y}(s))}\hat{Z}_\varphi(s)\, ds\Big), \\ \hat{Y}(t) &= \frac{1}{1 - H(\bar{Y}(t))}\Big(\mu_\varphi \hat{X}(t) + \int_0^t \big(1 - H(\bar{Y}(s))\big)\, d\hat{Z}_\varphi(s)\Big). \end{aligned} \tag{3.17}$$

It is clear from the expression above that $(\hat{N}, \hat{Y})$ is a joint continuous mean zero Gaussian process, whose covariance functions can be explicitly calculated (see the next remark).

Remark 3.7. If $\mu_\varphi = 0$ as in Remark 3.3, then the limit $(\hat{N}, \hat{Y})$ can be simplified as

$$\big(\hat{N}(t), \hat{Y}(t)\big) = \Big(\frac{\hat{X}(t)}{1 - H(0)} + \frac{\lambda_0 H'(0)}{(1 - H(0))^2}\int_0^t (t - s)\, d\hat{Z}_\varphi(s), \hat{Z}_\varphi(t)\Big), \quad t \geq 0. \tag{3.18}$$

Hence, we obtain that for $t, t' \geq 0$, recalling σ_φ in (3.13),

$$\begin{aligned} \mathbb{E}\big[\hat{N}(t)\hat{N}(t')\big] &= \frac{\lambda_0 (t \wedge t')}{(1 - H(0))^3} + \frac{\lambda_0^3 \sigma_\varphi^2 (H'(0))^2}{(1 - H(0))^5}\Big(\frac{(t \wedge t')^3}{3} + \frac{(t \wedge t')^2}{2}|t - t'|\Big), \\ \mathbb{E}\big[\hat{N}(t)\hat{Y}(t')\big] &= \frac{\lambda_0^2 \sigma_\varphi^2 H'(0)}{(1 - H(0))^3}\Big(\frac{(t \wedge t')^2}{2} + t'(t - t')^+\Big), \\ \mathbb{E}\big[\hat{Y}(t)\hat{Y}(t')\big] &= \frac{\lambda_0 \sigma_\varphi^2}{1 - H(0)}(t \wedge t'). \end{aligned}$$

If $\mu_\varphi \neq 0$, applying Ito formula to (3.17) we have

$$\hat{N}(t) = \frac{1}{1 - H(\bar{Y}(t))}\Big(\hat{X}(t) + \frac{1}{\mu_\varphi}\int_0^t \big(H(\bar{Y}(t)) - H(\bar{Y}(s))\big)\, d\hat{Z}_\varphi(s)\Big). \tag{3.19}$$

Hence, we obtain for $t, t' \geq 0$, with $y = \bar{Y}(t), y' = \bar{Y}(t')$,

$$\mathbb{E}\big[\hat{N}(t)\hat{N}(t')\big] = \frac{\mu_\varphi^{-1}\cdot(y\wedge y')}{(1-H(y))(1-H(y'))} + \frac{\sigma_\varphi^2}{\mu_\varphi^3}\int_0^{y\wedge y'}\frac{(H(y)-H(z))(H(y')-H(z))}{(1-H(y))(1-H(y'))}\,dz,$$

$$\mathbb{E}\big[\hat{N}(t)\hat{Y}(t')\big] = \frac{y\wedge y'}{(1-H(y))(1-H(y'))} + \frac{\sigma_\varphi^2}{\mu_\varphi^2}\int_0^{y\wedge y'}\frac{(H(y)-H(z))(1-H(z))}{(1-H(y))(1-H(y'))}\,dz,$$

$$\mathbb{E}\big[\hat{Y}(t)\hat{Y}(t')\big] = \frac{\mu_\varphi\cdot(y\wedge y')}{(1-H(y))(1-H(y'))} + \frac{\sigma_\varphi^2}{\mu_\varphi}\int_0^{y\wedge y'}\frac{(1-H(z))^2}{(1-H(y))(1-H(y'))}\,dz.$$

Remark 3.8. $(\hat{N}, \hat{\Lambda}, \hat{Y})$ in (3.15) can also be expressed in terms of Itô semi-martingales. Let

$$\hat{B}_1(t) := \int_0^t \frac{(1-H(\bar{Y}(s)))^{1/2}}{\sqrt{\lambda_0}}\,d\hat{X}(s) \quad\text{and}\quad \hat{B}_2(t) := \int_0^t \frac{(1-H(\bar{Y}(s)))^{1/2}}{\sigma_\varphi\sqrt{\lambda_0}}\,d\hat{Z}_\varphi(s). \tag{3.20}$$

One can find from (3.16) that $(\hat{B}_1, \hat{B}_2)$ are two independent standard Brownian motions and

$$\begin{aligned} d\hat{N}(t) &= \frac{\sqrt{\lambda_0}}{(1-H(\bar{Y}(t)))^{3/2}}\,d\hat{B}_1(t) + \frac{\lambda_0\,H'(\bar{Y}(t))}{(1-H(\bar{Y}(t)))^2}\,\hat{Y}(t)dt,\\ d\hat{Y}(t) &= \frac{\sigma_\varphi\sqrt{\lambda_0}}{(1-H(\bar{Y}(t)))^{1/2}}\,d\hat{B}_2(t) + \mu_\varphi\,d\hat{N}(t),\\ d\hat{\Lambda}(t) &= d\hat{N}(t) - \frac{\sqrt{\lambda_0}}{(1-H(\bar{Y}(t)))^{1/2}}\,d\hat{B}_1(t). \end{aligned} \tag{3.21}$$

Moreover, we also obtain that the limit $\hat{Y}$ is equal in distribution to an Ornstein-Uhlenbeck (OU) process with time-dependent parameters, which is the unique solution to the following SDE:

$$d\hat{Y}(t) = \; b_Y(t)\,\hat{Y}(t)\,dt + \sigma_Y(t)\,d\hat{B}(t)\,, \tag{3.22}$$

where $\hat{B}$ is a standard Brownian motion and

$$b_Y(t) = \frac{\lambda_0\mu_\varphi\,H'(\bar{Y}(t))}{(1-H(\bar{Y}(t)))^2} \quad\text{and}\quad \sigma_Y^2(t) = \frac{\lambda_0\mu_\varphi^2}{(1-H(\bar{Y}(t)))^3} + \frac{\lambda_0\sigma_\varphi^2}{(1-H(\bar{Y}(t)))},$$

The limit process $\hat{Y}$ can be expressed as

$$\begin{aligned} \hat{Y}(t) &= e^{\int_0^t b_Y(u)\,du}\int_0^t e^{-\int_0^u b_Y(v)\,dv}\sigma_Y(u)\,d\hat{B}(u)\\ &= \frac{1}{1-H(\bar{Y}(t))}\int_0^t \big(1-H(\bar{Y}(u))\big)\sigma_Y(u)\,d\hat{B}(u)\\ &= \frac{1}{1-H(\bar{Y}(t))}\int_0^t \sqrt{\bar{\Lambda}'(u)}\sqrt{\mu_\varphi^2+\sigma_\varphi^2(1-H(\bar{Y}(u)))^2}\,d\hat{B}(u)\,, \end{aligned}$$

Here we use the fact $\bar{Y} = \mu_\varphi\,\bar{\Lambda}$ and (3.16) to get

$$\begin{aligned} \int_0^t b_Y(u)\,du = \lambda_0\int_0^t \frac{\mu_\varphi\,H'(\bar{Y}(u))}{(1-H(\bar{Y}(u)))^2}\,du = \int_0^t \frac{H'(\bar{Y}(u))}{1-H(\bar{Y}(u))}\,d\bar{Y}(u)\\ = -\ln\big(1-H(\bar{Y}(u))\big)\Big|_0^t = \ln\Big(\frac{1-H(0)}{1-H(\bar{Y}(t))}\Big). \end{aligned}$$

Remark 3.9. If $\phi(t,y) \equiv \phi(t)$ for some $\phi \geq 0$, then $H(\cdot) \equiv ||\phi||_1 < 1$ and the Hawkes process does not depend on the shot noise process. We have in Theorem 3.1,

$$\bar{\Lambda}(t) = \frac{\lambda_0}{1-||\phi||_1}\cdot t.$$

In Theorem 3.2, by (3.17) and (3.20), we obtain

$$\begin{aligned}\big(\hat N(t), \hat Y(t) - \mu_\varphi \hat N(t)\big) &= \Big(\frac{\hat X(t)}{1-||\phi||_1}, \hat Z(t)\Big)\\ &= \Big(\frac{\lambda_0^{1/2}}{(1-||\phi||_1)^{3/2}}\hat B_1(t), \frac{\lambda_0^{1/2}\sigma_\varphi}{(1-||\phi||_1)^{1/2}}\hat B_2(t)\Big)\end{aligned}$$

where $\hat B_1$ and $\hat B_2$ are two independent standard Brownian motions. The limit $\hat N(t)$ coincides with the one-dimensional Hawkes limit in [2]. In comparison with the limits $\hat N$ in (3.18) and (3.19), we see that in our model with interactions between the Hawkes process and the shot noise process, in addition to the time-changed Brownian term $\hat X(t)$, there is an additional term as a functional of a time-changed Brownian motion $\hat Z(t)$, independent of $\hat X(t)$. However, the integrals with respect to $\hat Z(t)$, $\int_0^t (t-u)\, d\hat Z(u)$ and $\int_0^t \big(H(\bar Y(t)) - H(\bar Y(u))\big)\, d\hat Z(u)$ in (3.18) and (3.19), respectively, are Gaussian processes (of Volterra type). Hence, the effect of the shot noise process upon the Hawkes process is exhibited through the independent term involving an integral with respect to the time-changed Brownian motion $\hat Z(t)$, but as a (Volterra type) Gaussian process.

Moreover, it also gives the FCLT for the Hawkes shot noise process without interaction, that is,

$$\hat Y(t) = \frac{\mu_g \lambda_0^{1/2}}{(1-||\phi||_1)^{3/2}}\hat B_1(t) + \frac{\lambda_0^{1/2}\sigma_\varphi}{(1-||\phi||_1)^{1/2}}\hat B_2(t) \stackrel{\mathrm{d}}{=} \sigma_Y \hat B(t) \tag{3.23}$$

where $\hat B$ is a standard Brownian motion and

$$\sigma_Y^2 = \frac{\lambda_0}{1-||\phi||_1}\Big(\frac{\mu_\varphi^2}{(1-||\phi||_1)^2} + \sigma_\varphi^2\Big).$$

This result is also new to the literature, although the proof follows from an adaption of the arguments in [2, 29].

We also remark that the CLT-scaled process is centered by its FLLN limit in Theorem 3.1 instead of its expectation, which is different from the classical case in [2]. This is because the formula for $\mathbb{E}[\bar N^{(n)}(t)]$ is implicit for the new process.

3.3. **An application in risk processes.** Our newly introduced model can be potentially used in various applications where shot noise processes and/or Hawkes processes are used. For example, one can consider a risk process with the wealth $V(t)$ at time t given by

$$V(t) = v_0 + p(t) - Y(t) = v_0 + p(t) - \sum_{i=1}^{N(t)} \varphi(t-\tau_i, \xi_i) \tag{3.24}$$

where v_0 is the initial wealth, $p(t)$ is the premium collected, (Y, N) is the interactive Hawkes shot noise process as specified in (2.1)–(2.3). The standard risk process assumes $N(t)$ is Poisson and $\varphi(t,x) \equiv x$, so that the important measures such as ruin probability can be explicitly calculated, see, e.g., [1]. However, when the arrival process is more general, the exact analysis becomes challenging, so diffusion approximations of the wealth process become useful to approximate the ruin probabilities (see, e.g., the recent work in [9, 5]). In [9], the authors study a risk model with a Hawkes arrival process of claims and with $\varphi(t,x) \equiv x$ (that is, a compound Hawkes process with i.i.d. claims) where a diffusion approximation when the baseline rate of the Hawkes process is scaled up, is used to approximate the ruin probability. Our model in (3.24) is much more general, where the function $\varphi(t-s,x)$ indicates the delay effects in the claim settlement. Moreover, it is also natural that there is certain dependence between the total claims and the claim settlement process.

We consider the scaled wealth process

$$\hat V^{(n)}(t) = \sqrt{n}\big(v_0^{(n)} + p^{(n)}(t) - \bar Y^{(n)}(t)\big)\,,$$

where $\sqrt{n}v_0^{(n)} \to \hat{v}_0$, $p^{(n)}(t) = p^{(n)}\bar{Y}(t)$ with $p^{(n)} = 1 + n^{-1/2}\hat{p} + o(n^{-1/2})$ and $\bar{Y}(t)$ in (3.9), and $\bar{Y}^{(n)}(t)$ is as defined in (3.2). By Theorem 3.2, we approximate the scaled wealth process as

$$\hat{V}^{(n)}(t) \approx \hat{v}_0 + \hat{p}\,\bar{Y}(t) - \hat{Y}(t) =: \hat{V}(t)$$

where $\hat{Y}$ is given in (3.17). Here we assume $\mu_\varphi > 0$. Observe that $\hat{Y}$ can be equivalently expressed as follows,

$$\begin{aligned}\hat{Y}(t) &= \frac{1}{1-H(\bar{Y}(t))}\Big(\mu_\varphi\, B_1(\bar{\Lambda}(t)) + \sigma_\varphi \int_0^t \big(1 - H(\bar{Y}(u))\big)\, dB_2(\bar{\Lambda}(u))\Big)\\ &\stackrel{\mathrm{d}}{=} \frac{1}{1-H(\bar{Y}(t))} B\Big(\mu_\varphi^2\, \bar{\Lambda}(t) + \sigma_\varphi^2 \int_0^t \big(1 - H(\bar{Y}(s))\big)^2\, d\bar{\Lambda}(s)\Big)\end{aligned}$$

where B_1 and B_2 are given in Theorem 3.2, and B is a standard Brownian motion. One can see that

$$\begin{aligned}\Big\{\inf_{t>0} \hat{V}(t) < 0\Big\} \;&=\; \Big\{\inf_{t>0}\{\hat{v}_0 + \hat{p}\,\bar{Y}(t) - \hat{Y}(t)\} < 0\Big\}\\ =\Big\{\inf_{t>0}\Big\{(\hat{v}_0 + \hat{p}\,t)(1 - H(t)) &- B\Big(\mu_\varphi\, t + \frac{\sigma_\varphi^2}{\mu_\varphi}\int_0^t (1-H(y))^2\, dy\Big)\Big\} < 0\Big\},\end{aligned}$$

that is, the ruin probability for $\hat{V}$ is equivalent to the one-sided boundary passage problem for Brownian motion. One can then apply the existing methods to numerically approximate such boundary crossing probabilities for Brownian motions, c.f. [45], [17], [37].

4. An extension: Interactive marked Hawkes shot noise processes

The interactive Hawkes shot noise process introduced and studied above is of the simplest form, and can be extended in several ways. In this section, we introduce an interactive "marked" Hawkes shot noise process, that is, a joint process (N, Y) as defined in Section 2 except that the stochastic intensity process $\lambda(t)$ depends on some exogenous randomness, the so-called "marks", in addition to the dependence upon the state of the shot noise process, that is,

$$\lambda(t) = \lambda_0 + \sum_{j\ge 1} \phi\big(t - \tau_j, \eta_j, Y(\tau_j-)\big)\mathbf{1}(\tau_j < t), \tag{4.1}$$

where $\{\eta_j : j \ge 1\}$ is a sequence of stationary exogenous random variables on a measurable space $\mathcal{E}'$. In addition, we allow the unpredictable variables ξ_j and η_j to be correlated, that is, $\{(\xi_j, \eta_j)\}_{j\ge1}$ can be regarded as sequence of i.i.d. pairs of correlated variables. Let $F(dz, dz')$ be the joint distribution of the unpredictable variables (ξ, η) on $\mathcal{E}\times\mathcal{E}'$, and $F_\xi(dz), F_\eta(dz')$ be the associated marginal distributions.

The well-posedness of the process (N, Y) follows essentially from the same argument as in Proposition 2.1, under the following assumption analogous to Assumption A1. Indeed, one can consider a family of Janossy densities for $\{(\tau_k, \xi_k, \eta_k)\}_{k\ge1}$ as in the proof of Proposition 2.1.

Assumption A5. *For every $A > 0$, $\sup_{|y|\le A} \phi(\cdot, z', y)$ is locally integrable for F_η-a.s.*

We will consider a sequence of interactive Hawkes shot noise processes $(N^{(n)}, Y^{(n)})$, indexed by n, that is,

$$Y^{(n)}(t) = \sum_{j\ge1} \varphi^{(n)}\big(t - \tau_j^{(n)}, \xi_j^{(n)}\big)\,,$$

where we understand that $\varphi^{(n)}(t, z) = 0$ for $t < 0$, and $N^{(n)}$ has the stochastic intensity given by

$$\lambda^{(n)}(t) = \lambda_0^{(n)} + \sum_{j\ge1} \phi^{(n)}\big(t - \tau_j^{(n)}, \eta_j^{(n)}, Y^{(n)}(\tau_j^{(n)}-)\big)\mathbf{1}\big(\tau_j^{(n)} < t\big)\,,$$

where we understand that $\phi^{(n)}(t,z',y)=0$ for $t<0$. In addition, let $\Lambda^{(n)}(t)$ be the stochastic cumulative intensity process, that is,

$$\Lambda^{(n)}(t)=\int_0^t \lambda^{(n)}(u)\,du=\lambda_0^{(n)}t+\sum_{j\geq 1}\mathbf{1}\big(\tau_j^{(n)}<t\big)\int_0^{t-\tau_j^{(n)}}\phi^{(n)}\big(u,\eta_j^{(n)},Y^{(n)}(\tau_j^{(n)}-)\big)\,du\,.$$

The scalings for the processes remain the same for FLLN and FCLT. For the n^{th}-system, the conditions in Assumptions A2 and A3 are imposed with the changes with respect to the functions $\phi^{(n)}$ as stated in the following assumption.

Assumption A6. *The conditions in Assumptions A2 and A3 remain the same, with the following modifications concerning the functions $\phi^{(n)}$: for some measurable function $\phi(t,z',y)\geq 0$,*

$$\phi^{(n)}(t,z',ny)=\phi(t,z',y), \tag{4.2}$$

and for $z'\in\mathcal{E}'$ and $y\in\mathbb{R}$, denote by

$$h(z',y)=\int_0^\infty \phi(u,z',y)\,du \quad \text{and} \quad H(y)=\mathbb{E}\big[h(\eta,y)\big]. \tag{4.3}$$

Suppose for every $k>0$, there is a measurable function $f_k:\mathcal{E}'\times\mathbb{R}\to\mathbb{R}_+$ and a decreasing function $I_k\in\mathbb{D}(\mathbb{R}_+)$ with $I_k(0)=1$ and $I_k(\infty)=0$ such that

$$\int_t^\infty \phi(u,z',y)\,du\leq I_k(t)\cdot f_k(z',y),\quad \forall (t,z')\in\mathbb{R}_+\times\mathcal{E}', |y|\leq k. \tag{4.4}$$

For every $k>0$, $\{h(\eta,y),|y|\leq k\}$ and $\{f_k(\eta,y),|y|\leq k\}$ are uniformly integrable, that is,

$$\limsup_{r\to\infty}\sup_{|y|\leq k}\mathbb{E}\big[h(\eta,y);h(\eta,y)>r\big]=0, \tag{4.5}$$

$$\limsup_{r\to\infty}\sup_{|y|\leq k}\mathbb{E}\big[f_k(\eta,y);f_k(\eta,y)>r\big]=0. \tag{4.6}$$

Remark 4.1. With the presence of additional marks $\eta_j^{(n)}$, uniformly integrable conditions are imposed on h and f_k in (4.5) and (4.6), respectively, such that the FLLN holds for η-processes $\bar{Z}_h^{(n)}$ and $\bar{Z}_f^{(n)}$ in (6.1). Moreover, they also ensure the boundedness of $\mathbb{E}[h(\eta,y)]$ and $\mathbb{E}[f_k(\eta,y)]$ for all $|y|\leq k$. Comparing with Assumption A2-(3.4), I_k in (4.4) is normalized with $I_k(0)=1$ with the presence of f_k, where $\mathbb{E}[f_k(\eta,y)]$ can be greater than 1. Note that Assumption A3 holds with the modified function H in (4.3).

We impose the following conditions for the FCLT in addition to Assumption A6.

Assumption A7. *Recall (I_k,f_k) and (J,g) in Assumption A6, respectively. The conditions in Assumption A4 hold with the these modified functions. Suppose for every $k>0$, $\{h^2(\eta,y),|y|\leq k\}$ and $\{f_k^2(\eta,y),|y|\leq k\}$ are uniformly integrable, respectively, that is,*

$$\limsup_{r\to\infty}\sup_{|y|\leq k}\mathbb{E}\big[h^2(\eta,y);h(\eta,y)>r\big]=0, \tag{4.7}$$

$$\limsup_{r\to\infty}\sup_{|y|\leq k}\mathbb{E}\big[f_k^2(\eta,y);f(\eta,y)>r\big]=0. \tag{4.8}$$

In addition, for $h(z',y)$ in (4.3), assume that $\mathbb{E}\big[h^2(\eta,y)\big]$ and $\mathbb{E}\big[\varphi(\infty,\xi)h(\eta,y)\big]$ are continuous in $y\in\mathbb{R}$.

Remark 4.2. Comparing with Assumption A4, the uniform integrability conditions in (4.7) and (4.8) imply the uniform integrability conditions in (4.5) and (4.6), respectively, which ensure the $\mathbb{C}$-tightness of $\hat{Z}_h^{(n)}$ and $\hat{Z}_f^{(n)}$ in (6.2), respectively. The continuity condition of $\mathbb{E}\big[h^2(\eta,y)\big]$ and $\mathbb{E}\big[\varphi(\infty,\xi)h(\eta,y)\big]$ in y is a technical condition to apply [30, Theorem 2.2].

Before we state the theorem, we recall that a Gaussian martingale is a Gaussian process with martingale property, and thus has independent increment property, c.f. [24, Definition 4.28 & Proposition 4.30].

Theorem 4.1 (FCLT)**.** *Under Assumption A6, the FLLN in Theorem 3.1 holds with the limit* $\big(\bar N,\bar\Lambda,\bar Y\big)$ *having the same expression with the modified function* H *in* (4.3)*.*

Under Assumptions A6 and A7, the convergence in (3.14) *holds with the limit* $(\hat N,\hat\Lambda,\hat Y)$ *being the unique strong solution to the following SDE:*

$$\begin{aligned} d\hat N(t) &= \frac{\lambda_0 H'(\bar Y(t))}{(1-H(\bar Y(t)))^2}\hat Y(t)\,dt + \frac{1}{1-H(\bar Y(t))}\,d\big(\hat X+\hat Z_h\big)(t),\\ d\hat\Lambda(t) &= d\hat N(t) - d\hat X(t),\\ d\hat Y(t) &= \mu_\varphi\, d\hat N(t) + d\hat Z_\varphi(t), \end{aligned} \tag{4.9}$$

with $\hat N(0)=\hat\Lambda(0)=\hat Y(0)=0$*, where* $\big(\hat Z_\varphi,\hat Z_h\big)$ *is a two-dimensional mean-zero Gaussian martingale characterized by*

$$\mathbb{E}\big[(\alpha\,\hat Z^{(n)}_\varphi+\beta\,\hat Z_h)^2(t)\big] = \int_0^t Var\big(\alpha\,\varphi(\infty,\xi)+\beta\,h(\eta,\bar Y(s))\big)\,d\bar\Lambda(s), \tag{4.10}$$

for every $\alpha,\beta\in\mathbb{R}$ *and* $t\ge 0$*, and* $\hat X$ *is a mean-zero time-changed Brownian motion with variances* $\bar\Lambda(t)$ *and independent of* $\big(\hat Z_\varphi,\hat Z_h\big)$*, that is,* $\hat X(t)=B(\bar\Lambda(t))$ *for a standard Brownian motion* B*.*

Remark 4.3. Note that $\hat Z_\varphi$ and $\hat Z_\eta$ capture the randomness from $\{\xi_j\}$ and $\{\eta_j\}$, for which we allow dependence within each pair between ξ_j and η_j. The proof for this result follows from a direct extension of the model in (3.1).

Similar to Remark 3.6, we can have an explicit formula for $(\hat N,\hat Y)$ in (4.9) as follows:

$$\begin{aligned} \hat N(t) &= \frac{1}{1-H(\bar Y(t))}\Big(\hat X(t)+\hat Z_\eta(t)+\int_0^t \bar\Lambda'(u)H'(\bar Y(u))\hat Z_\varphi(u)\,du\Big),\\ \hat Y(t) &= \frac{1}{1-H(\bar Y(t))}\Big(\mu_\varphi\,\hat X(t)+\mu_\varphi\,\hat Z_\eta(t)+\int_0^t\big(1-H(\bar Y(u))\big)\,d\hat Z_\varphi(u)\Big). \end{aligned} \tag{4.11}$$

Similar to Remark 3.7, if $\mu_\varphi\neq 0$

$$\hat N(t) = \frac{1}{1-H(\bar Y(t))}\Big(\hat X(t)+\hat Z_\eta(t)+\frac{1}{\mu_\varphi}\int_0^t\big(H(\bar Y(t))-H(\bar Y(u))\big)\,d\hat Z_\varphi(u)\Big),$$

and if $\mu_\varphi=0$, then

$$\big(\hat N(t),\hat Y(t)\big) = \Big(\frac{\hat X(t)+\hat Z_\eta(t)}{1-H(0)}+\frac{\lambda_0 H'(0)}{(1-H(0))^2}\int_0^t (t-u)\,d\hat Z_\varphi(u),\,\hat Z_\varphi(t)\Big).$$

Similar to Remark 3.8, we have following OU-type expression for $\hat Y$:

$$\begin{aligned} d\hat Y(t) &= \Big(d\hat Z_\varphi(t)+\frac{\mu_\varphi}{1-H(\bar Y(t))}\,(d\hat X+d\hat Z_\eta(t))\Big)+\frac{\lambda_0\mu_\varphi\,H'(\bar Y(t))}{(1-H(\bar Y(t)))^2}\hat Y(t)\,dt\\ &\overset{d}{=}\sigma_Y(t)\,dB(t)+b_Y(t)\hat Y(t)\,dt \end{aligned}$$

where B is a standard Brownian motion, $b_Y(t)=\dfrac{\lambda_0\mu_\varphi\,H'(\bar Y(t))}{(1-H(\bar Y(t)))^2}$ and

$$\sigma_Y^2(t) = \frac{\lambda_0\mu_\varphi^2}{(1-H(\bar Y(t)))^3}+\frac{\lambda_0}{1-H(\bar Y(t))}\mathrm{Var}\Big(g(\infty,\xi)+\frac{\mu_\varphi\,\tilde h(\eta,\bar Y(t))}{1-H(\bar Y(t))}\Big).$$

5. Proofs of Theorems 3.1 and 3.2

5.1. **Preliminaries.** First, we define the process $X^{(n)} = \{X^{(n)}(t) : t \geq 0\}$:

$$X^{(n)}(t) := N^{(n)}(t) - \Lambda^{(n)}(t).$$

It is evident that $X^{(n)}$ is a local martingale with respect to the filtration $\mathscr{F}_t^{(n)} := \sigma\{(\tau_j^{(n)}, \xi_j^{(n)}), j \leq N^{(n)}(t)\}$ by Proposition 2.1. We also define the associated LLN and CLT scaled processes:

$$\bar{X}^{(n)}(t) := \bar{N}^{(n)}(t) - \bar{\Lambda}^{(n)}(t) \quad \text{and} \quad \hat{X}^{(n)}(t) := \sqrt{n}\, \bar{X}^{(n)}(t) = \hat{N}^{(n)}(t) - \hat{\Lambda}^{(n)}(t), \tag{5.1}$$

recalling $\bar{N}^{(n)}$ and $\bar{\Lambda}^{(n)}$ in (3.2), and the LLN and CLT scaled auxiliary processes

$$\begin{aligned} \bar{Z}_\varphi^{(n)}(t) &:= \big(\bar{W}_\varphi^{(n)} - \mu_\varphi\big)\big(\bar{N}^{(n)}(t)\big) \qquad \text{and} \qquad \hat{Z}_\varphi^{(n)}(t) := \sqrt{n} \cdot \bar{Z}_\varphi^{(n)}(t), \\ \bar{Z}_g^{(n)}(t) &:= \big(\bar{W}_g^{(n)} - \mu_g\big)\big(\bar{N}^{(n)}(t)\big) \qquad \text{and} \qquad \hat{Z}_g^{(n)}(t) := \sqrt{n} \cdot \bar{Z}_g^{(n)}(t), \end{aligned} \tag{5.2}$$

where (φ, g) are functions given in Assumption A2-(2), and

$$\big(\bar{W}_\varphi^{(n)}(t), \bar{W}_g^{(n)}(t)\big) := \Big(n^{-1} \sum_{j \leq [nt]} \varphi\big(\infty, \xi_j^{(n)}\big), n^{-1} \sum_{j \leq [nt]} g\big(\xi_j^{(n)}\big)\Big).$$

By the assumption of unpredictable variables in (2.3), it is clear that $\bar{Z}_\varphi^{(n)}$ and $\bar{Z}_g^{(n)}$ are local martingales with respect to the filtration $\bar{\mathscr{F}}_t^{(n)} := \sigma\{(\bar{\tau}_j^{(n)}, \xi_j^{(n)}), \bar{\tau}_j^{(n)} \leq t\}$ where $\bar{\tau}_j^{(n)} := n^{-1}\tau_j^{(n)}$. More precisely, they are random time-changed discrete-time martingales with respect to $\{\xi_j^{(n)}\}_{j\geq 1}$.

We next provide the representations for the LLN-scaled and CLT-scaled processes, in which we use a direct construction of the terms that resembles those in the limit, and then show that the differences, referred to as the residual terms, will converge to zero.

For the LLN-scaled processes in (3.2), from (3.7) in Remark 3.1, we obtain

$$\begin{aligned} \bar{\Lambda}^{(n)}(t) &= \lambda_0\, t + \int_0^t \Big(\int_0^{n(t-v)} \phi\big(u, \bar{Y}^{(n)}(v-)\big)\, du \Big)\, d\bar{N}^{(n)}(v) \\ &= \lambda_0\, t + \int_0^t H\big(\bar{Y}^{(n)}(v-)\big)\, d\bar{N}^{(n)}(v) - \bar{\varepsilon}_1^{(n)}(t), \end{aligned} \tag{5.3}$$

recalling H in Assumption A3, and

$$\begin{aligned} \bar{Y}^{(n)}(t) &= \frac{1}{n} \sum_{j\geq 1} \varphi(\infty, \xi_j^{(n)})\, \mathbf{1}\big(\bar{\tau}_j^{(n)} \leq t\big) - \bar{\varepsilon}_2^{(n)}(t) \\ &= \mu_\varphi\, \bar{N}^{(n)}(t) + \bar{Z}_\varphi^{(n)}(t) - \bar{\varepsilon}_2^{(n)}(t), \end{aligned} \tag{5.4}$$

where

$$\begin{aligned} \bar{\varepsilon}_1^{(n)}(t) &:= \int_0^t \Big(\int_{n(t-v)}^\infty \phi\big(u, \bar{Y}^{(n)}(v-)\big)\, du \Big)\, d\bar{N}^{(n)}(v), \\ \bar{\varepsilon}_2^{(n)}(t) &:= \frac{1}{n} \sum_{j\geq 1} \Big(\varphi\big(\infty, \xi_j^{(n)}\big) - \varphi\big(n(t - \bar{\tau}_j^{(n)}), \xi_j^{(n)}\big) \Big)\, \mathbf{1}\big(\bar{\tau}_j^{(n)} \leq t\big). \end{aligned} \tag{5.5}$$

The terms $\bar{\varepsilon}_1^{(n)}$ and $\bar{\varepsilon}_2^{(n)}$ are the so-called LLN-scaled residual terms, and are expected to be negligible for large n.

Next, for the CLT-scaled processes in (3.11), from (5.1), (5.3), (3.9) and the fact that $\bar{N} = \bar{\Lambda}$, we have the following representation:

$$\begin{aligned}\hat{N}^{(n)}(t) &= \hat{X}^{(n)}(t) + \sqrt{n}\big(\bar{\Lambda}^{(n)}(t) - \bar{\Lambda}(t)\big) \\ &= \hat{X}^{(n)}(t) + \sqrt{n}\Big(\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\bar{N}^{(n)}(s) - \int_0^t H\big(\bar{Y}(s)\big)\, d\bar{\Lambda}(s)\Big) - \hat{\varepsilon}_1^{(n)}(t) \\ &= \hat{X}^{(n)}(t) + \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\hat{N}^{(n)}(s) \\ &\quad + \sqrt{n}\int_0^t \Big(H\big(\bar{Y}^{(n)}(s)\big) - H\big(\bar{Y}(s)\big)\Big)\, d\bar{\Lambda}(s) - \hat{\varepsilon}_1^{(n)}(t),\end{aligned} \tag{5.6}$$

and from (5.4) and the fact that $\bar{Y} = \mu_\varphi \bar{N}$, we also obtain

$$\hat{Y}^{(n)}(t) = \mu_\varphi \hat{N}^{(n)}(t) + \hat{Z}_\varphi^{(n)}(t) - \hat{\varepsilon}_2^{(n)}(t), \tag{5.7}$$

recalling $\hat{Z}_\varphi^{(n)}$ in (5.2), where

$$\big(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)}\big)(t) := \sqrt{n}\,\big(\bar{\varepsilon}_1^{(n)}, \bar{\varepsilon}_2^{(n)}\big)(t). \tag{5.8}$$

We also refer to the terms $\big(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)}\big)$ as the CLT-scaled residual terms. We observe the resemblance of these prelimit expressions with the limit $(\hat{X}, \hat{Z}_\varphi)$ as expressed in (3.17). We also expect the CLT-scaled residual terms $(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)})$ to be negligible for large n.

As we have observed, $\bar{N}^{(n)}$ may fail to have finite expectation and $\bar{X}^{(n)}$ is only a local-martingale by Proposition 2.1. To make use of its martingale property, we will use a localization technique in the proofs. Let $\bar{Y}$ be the solution to (3.9) in Theorem 3.1, which solves

$$\int_0^t \big(1 - H(\bar{Y}(s))\big)\, d\bar{Y}(s) = \lambda_0 \mu_\varphi\, t \quad \forall t > 0.$$

One can see that $\bar{Y}$ is the unique continuous and non-decreasing function satisfying the identity above for both $\mu_\varphi = 0$ and $\mu_\varphi > 0$. For fixed $T > 0$, let $k_0 > 0$ so that

$$\int_0^{k_0/2} \big(1 - H(y)\big)\, dy \geq \lambda_0 \mu_\varphi T. \tag{5.9}$$

Then, one can find that

$$0 \leq \sup_{t \leq T} \bar{Y}(t) = \bar{Y}(T) \leq k_0/2. \tag{5.10}$$

For the k_0 above, let the first passage time for $\bar{Y}^{(n)}$ be denoted by

$$\bar{\tau}_{k_0}^{(n)} := \inf\{t > 0, |\bar{Y}^{(n)}|(t) > k_0\}.$$

By the continuity of H and $H(y) < 1$ for all $y \in \mathbb{R}$ in Assumption A3, we have

$$\alpha_{k_0} := \sup_{|y| \leq k_0} H(y) = I_{k_0}(0) < 1, \tag{5.11}$$

recalling I_{k_0} in Assumption A2-(3.4). For notational brevity, we drop the subscript k_0 in $\bar{\tau}_{k_0}^{(n)}$, α_{k_0}, I_{k_0} and so on. c_0 denotes a constant that may vary from line to line, and c_1 and c_2 denote the constants from Burkholder-Davis-Gundy (BDG) inequality for the first and second moments of a local martingale, respectively. We have the following result, which will be used extensively in the proof.

Lemma 5.1. *Under Assumption A3, $\bar{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$ is a martingale with respect to $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\geq 0}$. There is a constant $c_0>0$ such that*

$$\mathbb{E}\Big[\sup_{s\leq t}\big(\bar{X}^{(n)}(s\wedge\bar{\tau}^{(n)})\big)^2\Big]\leq\frac{c_0}{n}\cdot t. \tag{5.12}$$

Hence, $\bar{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$ converges to 0 in L^2.

Proof. Since $\bar{\varepsilon}_1^{(n)}\geq 0$, we have from (5.1) and (5.3)

$$\begin{aligned}0&\leq\int_0^t\Big(1-H\big(\bar{Y}^{(n)}(s-)\big)\Big)\,d\bar{N}^{(n)}(s)\\&=\bar{X}^{(n)}(t)+\bar{\Lambda}^{(n)}(t)-\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s)\\&=\lambda_0\,t+\bar{X}^{(n)}(t)-\bar{\varepsilon}_1^{(n)}(t)\leq\lambda_0\,t+\bar{X}^{(n)}(t).\end{aligned} \tag{5.13}$$

Let $\{\bar{\vartheta}_k^{(n)}\}_{k\geq 1}$ be a localizing sequence of stopping times so that $\bar{X}^{(n)}(\cdot\wedge\bar{\vartheta}_k^{(n)})$ is a martingale and $\bar{\vartheta}_k^{(n)}\to\infty$ as $k\to\infty$. Then $\bar{X}^{(n)}(\,\cdot\wedge\bar{\tau}^{(n)}\wedge\bar{\vartheta}_k^{(n)})$ is also a martingale. Using the above inequality, we obtain

$$\mathbb{E}\Big[\int_0^{t\wedge\bar{\vartheta}_k^{(n)}}\Big(1-H\big(\bar{Y}^{(n)}(s-)\big)\Big)\,d\bar{N}^{(n)}\big(s\wedge\bar{\tau}^{(n)}\big)\Big]\leq\lambda_0\,t.$$

Letting $k\to\infty$, from the fact $H(y)<1$ for every $y\in\mathbb{R}$, we have

$$\mathbb{E}\Big[\int_0^{t}\Big(1-H\big(\bar{Y}^{(n)}(s-)\big)\Big)\,d\bar{N}^{(n)}\big(s\wedge\bar{\tau}^{(n)}\big)\Big]\leq\lambda_0\,t,\quad\forall t>0.$$

Applying the bound for $H(y)$ in (5.11) for every $y=\bar{Y}^{(n)}(s-)$ with $s\leq\bar{\tau}^{(n)}$, we obtain

$$\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big]\leq\frac{\lambda_0}{1-\alpha}\cdot t\,,\quad\forall t>0. \tag{5.14}$$

Applying the BGD inequality, c.f. [27, Theorem 20.12], to the local martingales $X^{(n)}$, which has quadratic variation $N^{(n)}$, gives

$$\mathbb{E}\Big[\sup_{s\leq t}\big(\bar{X}^{(n)}\big)^2(s\wedge\bar{\tau}^{(n)})\Big]\leq c_2\cdot\mathbb{E}\big[[\bar{X}^{(n)}](t\wedge\bar{\tau}^{(n)})\big]=\frac{c_2}{n}\cdot\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big],$$

where c_2 is the constant from the BDG inequality. Further applying (5.14) and [40, Theorem 51] proves the martingale property of $\bar{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$, which also shows the L^2 convergence of $\bar{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$. □

5.2. **Proof of the FLLN: Theorem 3.1.** For the fixed $T>0$ in (5.9), recall the following modulus of continuity of an arbitrary function x on $[0,T]$ from [4, equation (7.1)]:

$$w(x,\delta):=\sup_{0\leq u-v\leq\delta,\;0\leq u,v\leq T}\big|x(u)-x(v)\big|,\quad\forall\delta>0. \tag{5.15}$$

Lemma 5.2 (Continuity of $\bar{N}^{(n)}$)**.** *Under Assumptions A2 and A3, for some $c_0>0$,*

$$\limsup_{n\to\infty}\mathbb{E}\Big[w\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\Big]\leq c_0\,\delta\quad\textit{for every small}\;\;\delta>0.$$

Proof. We first show that $\bar{N}^{(n)}$ is equi-continuous in expectation, that is,

$$\lim_{s\to t}\limsup_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big]=0. \tag{5.16}$$

For every $t > s > 0$, one can check from (5.3) and the fact $\bar{\varepsilon}_1^{(n)} \geq 0$ that

$$\begin{aligned}\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{\Lambda}^{(n)}(s\wedge\bar{\tau}^{(n)}) &= \big(\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{\Lambda}^{(n)}(s)\big)\mathbf{1}(s\leq\bar{\tau}^{(n)})\\ &\leq \lambda_0\,(t-s) + \int_s^t H\big(\bar{Y}^{(n)}(u-)\big)\,d\bar{N}^{(n)}(u\wedge\bar{\tau}^{(n)}) + \bar{\varepsilon}_1^{(n)}(s)\mathbf{1}(s\leq\bar{\tau}^{(n)})\\ &\leq \lambda_0\,(t-s) + \alpha\left(\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\right) + \bar{\varepsilon}_1^{(n)}(s)\mathbf{1}(s\leq\bar{\tau}^{(n)}),\end{aligned}$$

where the bound (5.11) is applied. With the fact $\bar{N}^{(n)} = \bar{\Lambda}^{(n)} + \bar{X}^{(n)}$ in (5.1), it implies

$$\begin{aligned}&(1-\alpha)\left(\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\right)\\ &\leq \lambda_0\,(t-s) + \left(\bar{X}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{X}^{(n)}(s\wedge\bar{\tau}^{(n)})\right) + \bar{\varepsilon}_1^{(n)}(s)\mathbf{1}\big(s\leq\bar{\tau}^{(n)}\big).\end{aligned} \tag{5.17}$$

On the other hand, since I is a decreasing function on $\mathbb{R}_+$ by its definition in Assumption A2-(3.4), we also have from (5.5) for every $r \in (0, s)$

$$\begin{aligned}\bar{\varepsilon}_1^{(n)}(s)\mathbf{1}(s\leq\bar{\tau}^{(n)}) &\leq \int_0^s I\big(n(s-u)\big)\,d\bar{N}^{(n)}(u\wedge\bar{\tau}^{(n)})\\ &= \int_0^r I\big(n(s-u)\big)\,d\bar{N}^{(n)}(u\wedge\bar{\tau}^{(n)}) + \int_r^s I\big(n(s-u)\big)\,d\bar{N}^{(n)}(u\wedge\bar{\tau}^{(n)})\\ &\leq I\big(n(s-r)\big)\cdot\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) + \big(\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big).\end{aligned} \tag{5.18}$$

Substituting the inequality above into (5.17) gives for every $t > s > r > 0$,

$$\begin{aligned}&(1-\alpha)\left(\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\right)\\ &\leq \lambda_0\,(t-s) + \left(\bar{X}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{X}^{(n)}(s\wedge\bar{\tau}^{(n)})\right)\\ &\qquad + I\big(n(s-r)\big)\cdot\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) + \big(\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big).\end{aligned} \tag{5.19}$$

Taking expectations on both sides of (5.19), making use of (5.14) and the martingale property in Lemma 5.1, we will have for every $t > s > r > 0$,

$$\begin{aligned}&(1-\alpha)\,\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big]\\ &\leq \lambda_0\,(t-s) + I\big(n(s-r)\big)\cdot\mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big] + \mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big]\\ &\leq \lambda_0\,(t-s) + I\big(n(s-r)\big)\cdot\frac{\lambda_0}{1-\alpha}\cdot s + \mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big].\end{aligned}$$

This inequality implies that for some constant $c_0 > 0$ independent of (t, s, r, n),

$$\begin{aligned}&\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big]\\ &\leq c_0\cdot\mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big] + c_0\cdot\big((t-s) + s\cdot I\big(n(s-r)\big)\big).\end{aligned}$$

We claim that the inequality above can be used to show the continuity in (5.16). Letting $n \to \infty$, since $I(\infty) = 0$ under Assumption A2-(3.4), we have for every $t > s > r > 0$,

$$\begin{aligned}&\limsup_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big]\\ &\leq c_0\cdot\liminf_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big] + c_0\cdot(t-s),\end{aligned} \tag{5.20}$$

where $\limsup, \liminf$ above are indeed finite by the boundedness for $\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big]$ in (5.14). Moreover, by the additive property of $t \to \mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big]$, it is straightforward to have

$$\begin{aligned}&\liminf_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big]\\ &\geq \liminf_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\big] + \liminf_{n\to\infty}\mathbb{E}\big[\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(r\wedge\bar{\tau}^{(n)})\big].\end{aligned}$$

Therefore, we can always find $s_k \uparrow t$ such that

$$\lim_{k\to\infty}\liminf_{n\to\infty}\mathbb{E}\big[\bar N^{(n)}(s_k\wedge\bar\tau^{(n)})-\bar N^{(n)}(s_{k-1}\wedge\bar\tau^{(n)})\big]=0.$$

Letting $(s,r)=(s_k,s_{k-1})$ in (5.20), we have from the limit above that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}\big[\bar N^{(n)}(t\wedge\bar\tau^{(n)})-\bar N^{(n)}(s_k\wedge\bar\tau^{(n)})\big]=0.$$

The monotonicity of $s\to\mathbb{E}\big[\bar N^{(n)}(t\wedge\bar\tau^{(n)})-\bar N^{(n)}(s\wedge\bar\tau^{(n)})\big]$ can be used to prove the left-equicontinuity property in (5.16) at $t>0$. Next, fixing $s>0$ and letting $t\downarrow s, r\uparrow s$, the right-equicontinuity property in (5.16) at $s>0$ is proved by the left-equicontinuity.

Now for arbitrary small $\delta>0$, we take

$$t_0=0<t_1<\cdots<t_v=T\quad\text{with}\quad t_j-t_{j-1}=\delta\ \text{ for }\ 2\le j\le v\ \text{ and }\ t_1\le\delta.$$

By the monotonicity of $\bar N^{(n)}$, it is easy to check that

$$w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)}),\delta\big)\le 2\cdot\max_{1\le j\le v}\Big(\bar N^{(n)}(t_j\wedge\bar\tau^{(n)})-\bar N^{(n)}(t_{j-1}\wedge\bar\tau^{(n)})\Big),\tag{5.21}$$

similar to [4, equation (7.10)]. Further taking $(t,s,r)=(t_j,t_{j-1},t_{j-1}-\delta')$ in (5.19) for $2\le j\le v$ with $\delta'\in(0,\delta)$, we have

$$\begin{aligned}w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)}),\delta\big)&\le 2\cdot\bar N^{(n)}\big(\delta\wedge\bar\tau^{(n)}\big)\\&\quad+\frac{2}{1-\alpha}\cdot\Big(\lambda_0\,\delta+2\sup_{t\le T}\big|\bar X^{(n)}\big|(t\wedge\bar\tau^{(n)})+I(n\delta')\,\bar N^{(n)}(T\wedge\bar\tau^{(n)})\Big)\\&\quad+\frac{2}{1-\alpha}\cdot\sum_{j=2}^{v}\Big(\bar N^{(n)}(t_{j-1}\wedge\bar\tau^{(n)})-\bar N^{(n)}\big((t_{j-1}-\delta')\wedge\bar\tau^{(n)}\big)\Big).\end{aligned}\tag{5.22}$$

Taking expectation on both sides and making use of (5.14) gives

$$\begin{aligned}&\mathbb{E}\Big[w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)}),\delta\big)\Big]\\&\le\frac{2\,\lambda_0\cdot\delta}{1-\alpha}+\frac{2\,\lambda_0\cdot\delta}{1-\alpha}+\frac{4}{1-\alpha}\cdot\mathbb{E}\Big[\sup_{t\le T}\big|\bar X^{(n)}(t\wedge\bar\tau^{(n)})\big|\Big]+\frac{2\,\lambda_0\,T}{1-\alpha}\frac{I(n\delta')}{1-\alpha}\\&\quad+\frac{2}{1-\alpha}\cdot\sum_{j=2}^{v}\mathbb{E}\Big[\bar N^{(n)}\big(t_{j-1}\wedge\bar\tau^{(n)}\big)-\bar N^{(n)}\big((t_{j-1}-\delta')\wedge\bar\tau^{(n)}\big)\Big].\end{aligned}$$

Further applying Cauchy's inequality to (5.12) gives for some constant $c_0>0$,

$$\begin{aligned}\mathbb{E}\Big[w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)}),\delta\big)\Big]&\le c_0\cdot\big(\delta+\sqrt{T/n}+I(n\delta')\cdot T\big)\\&\quad+c_0\cdot\sum_{j=2}^{v}\mathbb{E}\Big[\bar N^{(n)}(t_{j-1}\wedge\bar\tau^{(n)})-\bar N^{(n)}\big((t_{j-1}-\delta')\wedge\bar\tau^{(n)}\big)\Big],\end{aligned}$$

where the last term of the partial sum is of $[T/\delta]$ many. Now, letting $n\to\infty$ with the fact $I(\infty)=0$, and then $\delta'\to 0+$ with the fact (5.16), we obtain the modulus of continuity in the lemma. □

We next show that $(\bar\varepsilon_1^{(n)},\bar\varepsilon_2^{(n)})$ in (5.5) is negligible for large n. Recalling $(\bar W_\varphi^{(n)},\bar W_g^{(n)},\bar Z_\varphi^{(n)},\bar Z_g^{(n)})$ defined in (5.2) and observing that $\{\xi_j^{(n)}\}_{j\ge1}$ in (2.3) are assumed to be unpredictable variables, identically distributed and have finite expectation under Assumption A2-(3), the FLLN for triangular arrays with i.i.d. variables holds (or more generally, a stationary and unpredictable sequence), c.f. [13, Theorem 2.2.11], that is, as $n\to\infty$,

$$\bar W_\varphi^{(n)}(t)\to\mu_\varphi\,t\quad\text{and}\quad\bar W_g^{(n)}(t)\to\mu_g\,t\quad\text{u.o.c. in probability.}\tag{5.23}$$

Moreover, since $\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big]$ is uniformly bounded in (5.14) for every $t>0$, one can have from the limit above that, despite of the limit behaviour for $\bar{N}^{(n)}$, as $n\to\infty$,

$$\bar{Z}_{\varphi}^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad\text{and}\quad \bar{Z}_{g}^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad \text{u.o.c. in probability.} \tag{5.24}$$

Lemma 5.3. *Under Assumptions A2 and A3, we have as $n\to\infty$,*

$$\bar{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad\textit{and}\quad \bar{\varepsilon}_2^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad \textit{u.o.c. in probability.}$$

Proof. Recalling $\bar{\varepsilon}_1^{(n)}$ in (5.5) and I in Assumption A2-(3.4), we have

$$0\le \bar{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\le \int_0^{t\wedge\bar{\tau}^{(n)}} I\big(n(t\wedge\bar{\tau}^{(n)}-u)\big)\,d\bar{N}^{(n)}(u). \tag{5.25}$$

Similar to (5.18), for every $t\in[0,T]$ and $\delta>0$, we have

$$\begin{aligned}\text{RHS of (5.25)} &\le \int_0^{t\wedge\bar{\tau}^{(n)}-\delta} I\big(n(t\wedge\bar{\tau}^{(n)}-u)\big)\,d\bar{N}^{(n)}(u)+I(0)\int_{t\wedge\bar{\tau}^{(n)}-\delta}^{t\wedge\bar{\tau}^{(n)}} d\bar{N}^{(n)}(u)\\ &\le I(n\delta)\cdot\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})+\big(\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}-\delta)\big),\end{aligned} \tag{5.26}$$

where we understand that $\bar{N}^{(n)}(u)=0$ for $u<0$, and the inequalities above hold on both $\{\delta\le t\wedge\bar{\tau}^{(n)}\}$ and $\{t\wedge\bar{\tau}^{(n)}<\delta\}$. Hence, we obtain for every $\delta>0$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\le T}\bar{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\Big] &\le I(n\delta)\cdot\mathbb{E}\big[\bar{N}^{(n)}(T\wedge\bar{\tau}^{(n)})\big]+\mathbb{E}\big[w\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\big]\\ &\le I(n\delta)\cdot\frac{\lambda_0 T}{1-\alpha}+\mathbb{E}\big[w\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\big],\end{aligned}$$

where (5.14) is applied. Letting $n\to\infty$ with the fact $I(\infty)=0$, then $\delta\to 0$ with Lemma 5.2 applied, the limit for $\bar{\varepsilon}_1^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$ is proved.

For $\bar{\varepsilon}_2^{(n)}$ in (5.5), recalling (J,g) in Assumption A2-(3.5), we obtain

$$\begin{aligned}\big|\bar{\varepsilon}_2^{(n)}(t)\big| &\le \frac{1}{n}\sum_{j\ge1} J\big(n(t-\bar{\tau}_j^{(n)})\big)\,g\big(\xi_j^{(n)}\big)=\int_0^t J\big(n(t-u)\big)\,d\bar{W}_g^{(n)}(u)\\ &=\int_0^t J\big(n(t-u)\big)\,d\bar{Z}_g^{(n)}(u)+\mu_g\int_0^t J\big(n(t-u)\big)\,d\bar{N}^{(n)}(u),\end{aligned} \tag{5.27}$$

recalling $\bar{Z}_g^{(n)}$ in (5.2), where we understand $J(u)=0$ for $u<0$. Observing that J is a decreasing function with $J(0)=1$ and $J(\infty)=0$, we regard ν_J as the Stieltjes measure induced by $(-J)$ on $(0,\infty)$. We have from Fubini's theorem that

$$\begin{aligned}\int_0^t J\big(n(t-u)\big)\,d\bar{Z}_g^{(n)}(u) &= \int_0^t\int_{t-u}^{\infty} d\nu_J(nv)\,d\bar{Z}_g^{(n)}(u)\\ =\int_0^{\infty} d\nu_J(nv)\int_{t-v}^{t} d\bar{Z}_g^{(n)}(u) &= \int_0^{\infty} d\nu_J(nv)\big(\bar{Z}_g^{(n)}(t)-\bar{Z}_g^{(n)}(t-v)\big),\end{aligned} \tag{5.28}$$

with the understanding that $\bar{Z}_g^{(n)}(u)=0$ for $u<0$. In particular, we have from (5.24) that

$$\begin{aligned}&\Big|\int_0^{t\wedge\bar{\tau}^{(n)}} J\big(n(t\wedge\bar{\tau}^{(n)}-u)\big)\,d\bar{Z}_g^{(n)}(u)\Big|\\ &\le \int_0^{\infty}\Big(\big|\bar{Z}_g^{(n)}\big|(t\wedge\bar{\tau}^{(n)})+\big|\bar{Z}_g^{(n)}\big|(t\wedge\bar{\tau}^{(n)}-v)\Big)\,d\nu_J(nv)\\ &\le 2\cdot\sup_{u\le t}\big|\bar{Z}_g^{(n)}\big|(u\wedge\bar{\tau}^{(n)})\to 0 \quad \text{u.o.c. in probability.}\end{aligned}$$

On the other hand, similar to the proof for $\bar{\varepsilon}_1^{(n)}$ in (5.25), one can also check that

$$\int_0^{t\wedge\bar{\tau}^{(n)}} J\big(n(t\wedge\bar{\tau}^{(n)}-u)\big)\,d\bar{N}^{(n)}(u)\to 0\quad \text{u.o.c. in probability.}$$

We thus obtain the uniform convergence in probability for $\bar{\varepsilon}_2^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$. □

Now, we are ready to show the FLLN for the LLN-scaled processes in (3.2). Recalling for every fixed $T>0$, $k_0>0$ is a constant defined in (5.9) so that (5.10) hold.

Proof of Theorem 3.1. Substituting (3.9) into (5.3) gives

$$\bar{\Lambda}^{(n)}(t)-\bar{\Lambda}(t)=\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s)-\int_0^t H\big(\bar{Y}(s)\big)\,d\bar{\Lambda}(s)-\bar{\varepsilon}_1^{(n)}(t).$$

Making use of $\bar{N}^{(n)}=\bar{X}^{(n)}+\bar{\Lambda}^{(n)}$ in (5.1) and the fact $\bar{Y}=\mu_\varphi\,\bar{\Lambda}$, we further obtain

$$\begin{aligned}\bar{\Lambda}^{(n)}(t)-\bar{\Lambda}(t)&=\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{X}^{(n)}(s)+\int_0^t\Big(H\big(\bar{Y}^{(n)}(s)\big)-H\big(\mu_\varphi\,\bar{\Lambda}^{(n)}(s)\big)\Big)\,d\bar{\Lambda}^{(n)}(s)\\&\quad+\Big(\int_0^t H\big(\mu_\varphi\,\bar{\Lambda}^{(n)}(s)\big)\,d\bar{\Lambda}^{(n)}(s)-\int_0^t H\big(\mu_\varphi\,\bar{\Lambda}(s)\big)\,d\bar{\Lambda}(s)\Big)-\bar{\varepsilon}_1^{(n)}(t)\\&=:\bar{\Lambda}_1^{(n)}(t)+\bar{\Lambda}_2^{(n)}(t)+\bar{\Lambda}_3^{(n)}(t)-\bar{\varepsilon}_1^{(n)}(t).\end{aligned}\tag{5.29}$$

Noticing that the integrators and the integrands in $\bar{\Lambda}_3^{(n)}$ above are strictly increasing and continuous, by change of variable we can rewrite the identity as

$$\bar{\Lambda}_3^{(n)}(t)=\int_0^{\bar{\Lambda}^{(n)}(t)}H(\mu_\varphi\,y)\,dy-\int_0^{\bar{\Lambda}(t)}H(\mu_\varphi\,y)\,dy=\int_{\bar{\Lambda}(t)}^{\bar{\Lambda}^{(n)}(t)}H(\mu_\varphi\,y)\,dy,$$

where we understand that $\int_a^b f(y)\,dy=-\int_b^a f(y)\,dy$ for every $a,b\in\mathbb{R}$. Applying the mean value theorem, for some $\tilde{\Lambda}^{(n)}$ taking value between $\bar{\Lambda}$ and $\bar{\Lambda}^{(n)}$, we obtain

$$\bar{\Lambda}_3^{(n)}(t)=\big(\bar{\Lambda}^{(n)}(t)-\bar{\Lambda}(t)\big)\cdot H\big(\mu_\varphi\,\tilde{\Lambda}^{(n)}(t)\big).$$

Plugging the identity above into (5.29) gives for every $t>0$,

$$\begin{aligned}&\big(1-H\big(\mu_\varphi\,\tilde{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})\big)\big)\cdot\big(\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{\Lambda}(t\wedge\bar{\tau}^{(n)})\big)\\&=\bar{\Lambda}_1^{(n)}(t\wedge\bar{\tau}^{(n)})+\bar{\Lambda}_2^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)}).\end{aligned}\tag{5.30}$$

Since $\mathbb{E}\big[\bar{\Lambda}^{(n)}(T\wedge\bar{\tau}^{(n)})\big]=\mathbb{E}\big[\bar{N}^{(n)}(T\wedge\bar{\tau}^{(n)})\big]$ has a bounded expectation in (5.14) in the proof of Lemma 5.1, for every $\varepsilon>0$, we have for some constant $c_\varepsilon>\bar{\Lambda}(T)>0$,

$$\mathbb{P}\big(\bar{\Lambda}^{(n)}(T\wedge\bar{\tau}^{(n)})>c_\varepsilon\big)\le\varepsilon,\quad\forall n\in\mathbb{N}.\tag{5.31}$$

Thus, $\sup_{t\le T}\tilde{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})\le c_\varepsilon$ on the set $\{\bar{\Lambda}^{(n)}(T\wedge\bar{\tau}^{(n)})\le c_\varepsilon\}$ by definition. Let $\gamma_\varepsilon=\sup_{|y|\le c_\varepsilon}H(\mu_\varphi y)$. Then, $\gamma_\varepsilon<1$ by the continuity of H on $\mathbb{R}_+$ and $H(y)<1$ in Assumption A3, and

$$\mathbb{P}\Big(\sup_{t\le T}H\big(\mu_\varphi\,\tilde{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})\big)\le\gamma_\varepsilon\Big)>1-\varepsilon,\quad\forall n\in\mathbb{N}.\tag{5.32}$$

By the local-martingale property of $\bar{X}^{(n)}$ from Proposition 2.2, $\bar{\Lambda}_1^{(n)}$ in (5.29) is also an $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\geq 0}$-adapted local martingale with quadratic variation given by $n^{-1}\int_0^t H^2\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s)$. Applying the BGD inequality, we obtain for some constant $c_2>0$ from the inequality that

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\leq T}(\bar{\Lambda}_1^{(n)})^2(t\wedge\bar{\tau}^{(n)})\Big] &\leq c_2\cdot\mathbb{E}\Big[\big[\bar{\Lambda}_1^{(n)}\big](T\wedge\bar{\tau}^{(n)})\Big]\\ &=\frac{c_2}{n}\cdot\mathbb{E}\Big[\int_0^T H^2\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\Big]\\ &\leq\frac{c_2}{n}\cdot\mathbb{E}\Big[\bar{N}^{(n)}\big(T\wedge\bar{\tau}^{(n)}\big)\Big]\leq\frac{c_2\,\lambda_0}{1-\alpha}\cdot\frac{T}{n},\end{aligned}$$

where the fact that $H(\cdot)<1$ and (5.14) are used. Thus as $n\to\infty$,

$$\bar{\Lambda}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0\quad\text{u.o.c. in probablity.}\tag{5.33}$$

Next, we obtain from (5.4) that

$$\bar{Y}^{(n)}(t)-\mu_\varphi\,\bar{\Lambda}^{(n)}(t)=\mu_\varphi\,\bar{X}^{(n)}(t)+\bar{Z}_\varphi^{(n)}(t)-\bar{\varepsilon}_2^{(n)}(t).\tag{5.34}$$

Applying (5.24), (5.12) and Lemma 5.3 gives

$$\bar{Y}^{(n)}(t\wedge\bar{\tau}^{(n)})-\mu_\varphi\,\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0\quad\text{u.o.c. in probability.}\tag{5.35}$$

Plugging into $\bar{\Lambda}_2^{(n)}$ in (5.29), applying (5.35) and using the continuity of H and the stochastically boundedness for $\bar{\Lambda}^{(n)}$ from (5.31), we have

$$\begin{aligned}\big|\bar{\Lambda}_2^{(n)}\big|(t\wedge\bar{\tau}^{(n)})&\leq\int_0^t\big|H(\bar{Y}^{(n)}(s))-H(\mu_\varphi\,\bar{\Lambda}^{(n)}(s))\big|\,d\bar{\Lambda}^{(n)}(s\wedge\bar{\tau}^{(n)})\\ &\to 0\quad\text{u.o.c. in probability.}\end{aligned}\tag{5.36}$$

Now, plugging the limits (5.33), (5.36) and Lemma 5.3 into (5.30) gives

$$\big(1-H\big(\mu_\varphi\,\tilde{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})\big)\big)\cdot\big(\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{\Lambda}(t\wedge\bar{\tau}^{(n)})\big)\to 0\quad\text{u.o.c. in probability.}$$

Further applying the fact in (5.32) shows

$$\bar{\Lambda}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{\Lambda}(t\wedge\bar{\tau}^{(n)})\to 0\quad\text{u.o.c. in probability.}$$

Moreover, by the fact that $\bar{Y}=\mu_\varphi\,\bar{\Lambda}$, we obtain from (5.35) that

$$\begin{aligned}&\big|\bar{Y}^{(n)}(t\wedge\bar{\tau}^{(n)})-\bar{Y}(t\wedge\bar{\tau}^{(n)})\big|\\ &\leq\big|\bar{Y}^{(n)}-\mu_\varphi\,\bar{\Lambda}^{(n)}\big|(t\wedge\bar{\tau}^{(n)})+\mu_\varphi\,\big|\bar{\Lambda}^{(n)}-\bar{\Lambda}\big|(t\wedge\bar{\tau}^{(n)})\to 0\quad\text{u.o.c. in probability.}\end{aligned}$$

For the last, on the set $\{\bar{\tau}_{k_0}^{(n)}\leq T\}$ for the k_0 in (5.9), we have from (5.10) that

$$\big|\bar{Y}^{(n)}(\bar{\tau}_{k_0}^{(n)})-\bar{Y}(\bar{\tau}_{k_0}^{(n)})\big|\geq\big|\bar{Y}^{(n)}(\bar{\tau}_{k_0}^{(n)})\big|-\bar{Y}(\bar{\tau}_{k_0}^{(n)})\geq k_0-k_0/2=k_0/2,$$

which implies

$$\mathbb{P}\big(T<\bar{\tau}_{k_0}^{(n)}\big)\to 1.\tag{5.37}$$

Therefore, the FLLN hold for $(\bar{N}^{(n)},\bar{Y}^{(n)},\bar{\Lambda}^{(n)})$ without stopping. This finishes the proof. □

5.3. **Proof of the FCLT: Theorem 3.2.** To prove FCLT for the process in (3.11), we will use the representation (5.6), $\hat{X}^{(n)}$ in (5.1), $\hat{Z}_\varphi^{(n)} = \hat{W}_\varphi^{(n)}(\bar{N}^{(n)})$ and $\hat{Z}_g^{(n)} = \hat{W}_g^{(n)}(\bar{N}^{(n)})$ in (5.2), where

$$\begin{aligned}\hat{W}_\varphi^{(n)}(t) &:= \sqrt{n}\left(\bar{W}_\varphi^{(n)}(t) - \mu_\varphi\, n^{-1}[n\,t]\right) = \frac{1}{\sqrt{n}}\sum_{j\le[nt]}\left(\varphi(\infty,\xi_j^{(n)}) - \mu_\varphi\right),\\ \hat{W}_g^{(n)}(t) &:= \sqrt{n}\left(\bar{W}_g^{(n)}(t) - \mu_g\, n^{-1}[n\,t]\right) = \frac{1}{\sqrt{n}}\sum_{j\le[nt]}\left(g(\xi_j^{(n)}) - \mu_g\right).\end{aligned} \tag{5.38}$$

We first prove the convergence of the local martingale terms.

Lemma 5.4. *Under Assumptions A2, A3 and A4, we have as $n\to\infty$*

$$\left(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}\right) \Rightarrow \left(\hat{X}, \hat{Z}_\varphi\right) \quad in \quad (\mathbb{D}^2, J_1)$$

and

$$\hat{Z}_g^{(n)} = \hat{W}_g^{(n)}(\bar{N}^{(n)}) \Rightarrow \hat{Z}_g \quad in \quad (\mathbb{D}, J_1),$$

where $\hat{X}$ and $\hat{Z}_\varphi$ are the two independent mean-zero time-changed Brownian motions with variances $\bar{\Lambda}(t)$ and $\sigma_\varphi^2\cdot\bar{\Lambda}(t)$, respectively, as given in Theorem 3.2, and $\hat{Z}_g$ is a mean-zero time-changed Brownian motion with variance $\sigma_g^2\cdot\bar{\Lambda}(t)$.

Proof. We apply [24, Theorem 3.22 in Chapter VIII]. By Proposition 2.2, the assumption of unpredictable variables in (2.3) and the finiteness of second moment for $\varphi(\infty,\xi)$ in Assumption A4-(3.13), $\left(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}\right)$ is an $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\ge 0}$-adapted locally-square-integrable martingale.

It is straightforward to check that $\hat{X}^{(n)}$ has bounded jump size of $1/\sqrt{n}$. By the definition of unpredictable variables, for every $\varepsilon>0$,

$$\begin{aligned}&\mathbb{E}\Big[\sum_{t\le T\wedge\bar{\tau}^{(n)}}\left(\Delta\,\hat{Z}_\varphi^{(n)}(t)\right)^2\mathbf{1}\left(\left|\Delta\,\hat{Z}_\varphi^{(n)}(t)\right|>\varepsilon\right)\Big]\\ &=\mathbb{E}\left[\bar{N}^{(n)}(T\wedge\bar{\tau}^{(n)})\right]\cdot\mathbb{E}\left[\left(\varphi(\infty,\xi)-\mu_\varphi\right)^2; |\varphi(\infty,\xi)-\mu_\varphi|>\sqrt{n}\,\varepsilon\right]\\ &\le\frac{2\,\lambda_0\,T}{1-\alpha}\cdot\mathbb{E}\left[\mu_\varphi^2+\varphi^2(\infty,\xi); |\varphi|(\infty,\xi)>\sqrt{n}\,\varepsilon-\mu_\varphi\right],\end{aligned} \tag{5.39}$$

where (5.14) is applied, and which converges to 0 as $n\to\infty$ by the finiteness of σ_φ^2. Thus, the condition 3.23 in [24, Theorem 3.22 in Chapter VIII] holds for $(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)})$.

Next, we check the condition $[\hat{\gamma}_5' - D]$. It is not hard to show that for every $a,b\in\mathbb{R}$

$$\begin{aligned}\left[a\,\hat{X}^{(n)} + b\,\hat{Z}_\varphi^{(n)}\right](t) &= \frac{1}{n}\sum_{j\ge 1}\Big(a + b\left(\varphi(\infty,\xi_j^{(n)})-\mu_\varphi\right)\Big)^2\mathbf{1}(\bar{\tau}_j^{(n)}\le t)\\ &\to\left(a^2+b^2\sigma_\varphi^2\right)\cdot\bar{N}(t)\quad\text{u.o.c. in probability,}\end{aligned}$$

where Theorem 3.1 and the FLLN for triangular arrays are applied, similar to (5.24).

This proves the limit for $(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)})$, and a similar argument can be applied to $\hat{Z}_g^{(n)}$ with $\varphi(\infty,z)$ above replaced by $g(z)$. $\square$

The main mathematic challenge for the proof of FCLT comes from showing that $(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)})$ is indeed negligible for n large enough. The idea used for $(\bar{\varepsilon}_1^{(n)}, \bar{\varepsilon}_2^{(n)})$ in (5.26) and (5.27) is not applicable here, which is essentially change of variable and relies on the "sufficient" continuity of $\bar{N}^{(n)}$. For the CLT-scaled representation in (5.6), taking $\hat{\varepsilon}_1^{(n)}$ in the following (5.40) for example, $I(t)$ is a decreasing function and expected to be of order $t^{-1/2}$ under Assumption A4-(1), which is the critical case for the classical Hawkes model as discussed in Remark 3.5. Thus, to apply change of variable for the proof, one may need that the increment of the integrator $\bar{N}^{(n)}$ in (5.40) has Hölder-continuity of order strictly larger than 1/2, which can not happen because of Brownian factor $\hat{X}^{(n)}$ in $\tilde{\varepsilon}_{1,1}^{(n)}$ defined later. Hence, we further investigate the underlying Volterra type equation for the error process $\hat{\varepsilon}_1^{(n)}$. More specifically, we rewrite the error process $\hat{\varepsilon}_1^{(n)}$ into three parts in (5.41), where the tightness for $\hat{X}^{(n)}$ and the idea used in [34] can be used to analyze the first two terms, and where the last part involves a defective renewal equation.

Lemma 5.5. *Under Assumptions A2, A3 and A4, we have as $n \to \infty$,*

$$\left(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)}\right) \to 0 \quad \textit{u.o.c. in probability.}$$

Proof. Recalling the definitions of $\hat{\varepsilon}_1^{(n)}$ in (5.8) and (5.5), and I_k in Assumption A4-(3.4) (the index $k = k_0$ is dropped for notational brevity), we have for $t \le \bar{\tau}^{(n)}$,

$$\begin{aligned} \hat{\varepsilon}_1^{(n)}(t) &= \sqrt{n} \int_0^t \Big(\int_{n(t-u)}^\infty \phi(v, \bar{Y}^{(n)}(u-))\, dv \Big) d\bar{N}^{(n)}(u) \\ &\le \sqrt{n} \int_0^t I\big(n(t-s)\big)\, d\bar{N}^{(n)}(s) =: \tilde{\varepsilon}_1^{(n)}(t). \end{aligned} \tag{5.40}$$

Making use of (5.1) and the definition for $\lambda^{(n)}$ in (3.1), we can rewrite

$$\begin{aligned} \tilde{\varepsilon}_1^{(n)}(t) &= \sqrt{n} \int_0^t I\big(n(t-s)\big) \left(d\bar{X}^{(n)}(s) + \lambda^{(n)}(ns)\, ds\right) \\ &= \int_0^t I\big(n(t-s)\big)\, d\hat{X}^{(n)}(s) + \lambda_0 \cdot \sqrt{n} \int_0^t I\big(n(t-s)\big)\, ds \\ &\qquad + \sqrt{n} \int_0^t I\big(n(t-u)\big) \int_0^u n\phi\big(n(u-v), \bar{Y}^{(n)}(v-)\big)\, d\bar{N}^{(n)}(v)\, du \\ &=: \tilde{\varepsilon}_{1,1}^{(n)}(t) + \tilde{\varepsilon}_{1,2}^{(n)}(t) + \tilde{\varepsilon}_{1,3}^{(n)}(t). \end{aligned} \tag{5.41}$$

Starting from $\tilde{\varepsilon}_{1,3}^{(n)}$ above, by Fubini's theorem and change of variable, we obtain

$$\tilde{\varepsilon}_{1,3}^{(n)}(t) = \sqrt{n} \int_0^t d\bar{N}^{(n)}(s) \int_0^{n(t-s)} I\big(n(t-s) - u\big) \phi\big(u, \bar{Y}^{(n)}(s-)\big)\, du. \tag{5.42}$$

Since I is a decreasing function on $[0, \infty)$ with $I(\infty) = 0$, we regard ν_I as the Stieltjes measure induced by $(-I)$ on $(0, \infty)$.For every $t > 0$ and $y \in \mathbb{R}$, similar to (5.28), one can find that

$$\begin{aligned} &\int_0^t I(t-u)\phi(u, y)\, du = \int_0^\infty d\nu_I(v) \int_{t-v}^t \phi(u, y)\, du \\ &= \int_0^t d\nu_I(v) \Big(\int_{t-v}^\infty \phi(u, y)\, du - \int_t^\infty \phi(u, y)\, du \Big) + \int_t^\infty d\nu_I(v) \int_0^t \phi(u, y)\, du \\ &= \int_0^t d\nu_I(v) \int_{t-v}^\infty \phi(u, y)\, du - \big(I(0) - I(t)\big) \int_t^\infty \phi(u, y)\, du + I(t) \int_0^t \phi(u, y)\, du \\ &= -\alpha \int_t^\infty \phi(u, y)\, du + \int_0^t d\nu_I(v) \int_{t-v}^\infty \phi(u, y)\, du + I(t)\, H(y), \end{aligned} \tag{5.43}$$

recalling the fact $I(0)=\alpha$ in (5.11) and H in Assumption A3, where we understand that $\phi(u,y)=0$ for $u<0$. Plugging into (5.42) and recalling $\hat{\varepsilon}_1^{(n)}$ and $\tilde{\varepsilon}_1^{(n)}$ in (5.40), we have for $t\le\bar{\tau}^{(n)}$,

$$\begin{aligned}\tilde{\varepsilon}_{1,3}^{(n)}(t) &= -\alpha\cdot\hat{\varepsilon}_1^{(n)}(t)+\int_0^t \hat{\varepsilon}_1^{(n)}(t-s)d\nu_I(ns)+\sqrt{n}\int_0^t I\big(n(t-s)\big)H\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s)\\ &\le -\alpha\cdot\hat{\varepsilon}_1^{(n)}(t)+\int_0^t \hat{\varepsilon}_1^{(n)}(t-s)d\nu_I(ns)+\alpha\cdot\tilde{\varepsilon}_1^{(n)}(t),\end{aligned}$$

where the bound for $H\big(\bar{Y}^{(n)}(s-)\big)$ in (5.11) is applied to every $s\le\bar{\tau}^{(n)}$ above. Substituting into (5.41) and eliminating $\alpha\cdot\tilde{\varepsilon}_1^{(n)}$ on both sides gives

$$(1-\alpha)\cdot\tilde{\varepsilon}_1^{(n)}(t)\le\tilde{\varepsilon}_{1,1}^{(n)}(t)+\tilde{\varepsilon}_{1,2}^{(n)}(t)-\alpha\cdot\hat{\varepsilon}_1^{(n)}(t)+\int_0^t\hat{\varepsilon}_1^{(n)}(t-s)d\nu_I(ns).$$

By the fact that $\hat{\varepsilon}_1^{(n)}\le\tilde{\varepsilon}_1^{(n)}$, the inequality above further implies for $t\le\bar{\tau}^{(n)}$,

$$0\le\hat{\varepsilon}_1^{(n)}(t)\le\tilde{\varepsilon}_{1,1}^{(n)}(t)+\tilde{\varepsilon}_{1,2}^{(n)}(t)+\int_0^t\hat{\varepsilon}_1^{(n)}(t-s)d\nu_I(ns).$$

In particular, taking the supremum in t over $[0,T\wedge\bar{\tau}^{(n)}]$, we have

$$\sup_{t\le T}\hat{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\le\sup_{t\le T}\big|\tilde{\varepsilon}_{1,1}^{(n)}\big|(t\wedge\bar{\tau}^{(n)})+\sup_{t\le T}\tilde{\varepsilon}_{1,2}^{(n)}(t)+I(0)\cdot\sup_{t\le T}\big|\hat{\varepsilon}_1^{(n)}\big|(t\wedge\bar{\tau}^{(n)}). \tag{5.44}$$

Applying Fubini's theorem, similar to (5.28), we obtain

$$\begin{aligned}\tilde{\varepsilon}_{1,1}^{(n)}(t) &= \int_0^t\int_{t-s}^\infty d\nu_I(nu)\,d\hat{X}^{(n)}(s)=\int_0^\infty d\nu_I(nu)\int_{t-u}^t d\hat{X}^{(n)}(s)\\ &=\int_0^\infty\big(\hat{X}^{(n)}(t)-\hat{X}^{(n)}(t-u)\big)d\nu_I(nu),\end{aligned} \tag{5.45}$$

where we understand that $\hat{X}^{(n)}(u)=0$ for $u<0$. Therefore, for every $t\le T$ and $\delta>0$,

$$\begin{aligned}\big|\tilde{\varepsilon}_{1,1}^{(n)}(t\wedge\bar{\tau}^{(n)})\big| &\le\int_0^\delta\Big|\hat{X}^{(n)}(t\wedge\bar{\tau}^{(n)})-\hat{X}^{(n)}(t\wedge\bar{\tau}^{(n)}-u)\Big|\,d\nu_I(nu)\\ &\qquad+\int_\delta^\infty\Big|\hat{X}^{(n)}(t\wedge\bar{\tau}^{(n)})-\hat{X}^{(n)}(t\wedge\bar{\tau}^{(n)}-u)\Big|\,d\nu_I(nu)\\ &\le w\big(\hat{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)+I(n\delta)\cdot 2\sup_{u\le T}\big|\hat{X}^{(n)}\big|(u\wedge\bar{\tau}^{(n)}),\end{aligned} \tag{5.46}$$

recalling the modulus of continuity in (5.15). Using a similar argument as in [34, Proposition 2], noticing that $\sup_{t>0}\sqrt{t}\,I(t)<\infty$ under Assumption A4-(1), we obtain for $\tilde{\varepsilon}_{1,2}^{(n)}$ in (5.41),

$$\begin{aligned}\tilde{\varepsilon}_{1,2}^{(n)}(t) &= \lambda_0\cdot\sqrt{n}\int_0^t I(nu)\,du\\ &\le\lambda_0\cdot\bigg(\int_0^\delta u^{-1/2}\big(\sqrt{nu}\,I(nu)\big)\,du+\mathbf{1}(t>\delta)\sqrt{n}\int_\delta^t I(nu)\,du\bigg)\\ &\le\lambda_0\cdot\sup_{s>0}\sqrt{s}I(s)\cdot\frac{\delta}{2}+\lambda_0\cdot\sqrt{n}\cdot I(n\delta)\cdot t\end{aligned} \tag{5.47}$$

where the monotonicity of I is applied and the inequality holds for every $\delta>0$ and $t>0$. Substituting (5.46) and (5.47) into (5.44) and by the fact $I(0)=\alpha<1$, we have in conclusion for some constant $c_0 > 0$ and every $\delta > 0$,

$$\begin{aligned}&\big(1-\alpha\big)\cdot\sup_{t\le T}\hat{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\\ &\le c_0\cdot\Big(w\big(\hat{X}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)+I(n\delta)\cdot\sup_{t\le T}\big|\hat{X}^{(n)}\big|(t\wedge\bar{\tau}^{(n)})+\delta+\sqrt{n}\,I(n\delta)\cdot T\Big).\end{aligned}$$

Making use of the weak convergence for $\hat{X}^{(n)}$ with Brownian limit in Lemma 5.4, letting $n\to\infty$ and then $\delta\to 0$, from the fact that $I(\infty)=0$, we obtain

$$\sup_{t\le T}\hat{\varepsilon}_1^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0\quad\text{in probability.}$$

Further applying (5.37) proves the convergence of $\hat{\varepsilon}_1^{(n)}$ without stopping.

A similar argument can be applied to $\hat{\varepsilon}_2^{(n)}$, where we mainly highlight the difference in its proof below.

Recalling $\hat{Z}_g^{(n)}$ in Lemma 5.4, we have from (5.27) that, for $t\le\bar{\tau}^{(n)}$,

$$\begin{aligned}\hat{\varepsilon}_2^{(n)}(t)&\le\int_0^t J\big(n(t-s)\big)d\hat{Z}_g^{(n)}(s)+\mu_g\sqrt{n}\int_0^t J\big(n(t-s)\big)d\bar{N}^{(n)}(s)\\ &=:\tilde{\varepsilon}_{2,1}^{(n)}(t)+\mu_g\,\tilde{\varepsilon}_{2,2}^{(n)}(t).\end{aligned}\tag{5.48}$$

Applying the argument used for $\tilde{\varepsilon}_1^{(n)}$ in (5.41) to $\tilde{\varepsilon}_{2,2}^{(n)}$, we have

$$\begin{aligned}\tilde{\varepsilon}_{2,2}^{(n)}(t)&=\int_0^t J(n(t-s))d\hat{X}^{(n)}(s)+\lambda_0\cdot\sqrt{n}\int_0^t J\big(n(t-u)\big)\,du\\ &\qquad+\sqrt{n}\int_0^t J(n(t-u)\big)\int_0^u n\phi\big(n(u-v),\bar{Y}^{(n)}(v-)\big)\,d\bar{N}^{(n)}(v)du\\ &=:\tilde{\varepsilon}_{2,3}^{(n)}(t)+\tilde{\varepsilon}_{2,4}^{(n)}(t)+\tilde{\varepsilon}_{2,5}^{(n)}(t).\end{aligned}\tag{5.49}$$

Recalling that J in Assumption A2-(2) is decreasing with $J(0)=1$ and $J(\infty)=0$, one can have following identity similar to (5.43), for every $y\in\mathbb{R}$,

$$\int_0^t J(t-u)\phi(u,y)\,du=-\int_t^\infty\phi(u,y)\,du+\int_0^t d\nu_J(u)\int_{t-u}^\infty\phi(v,y)\,dv+J(t)\cdot H(y).$$

Therefore, recalling $\hat{\varepsilon}_1^{(n)}$ in (5.40) and $\tilde{\varepsilon}_{2,2}^{(n)}$ in (5.48), for $t\le\bar{\tau}^{(n)}$,

$$\tilde{\varepsilon}_{2,5}^{(n)}(t)\le\int_0^t\hat{\varepsilon}_1^{(n)}(t-s)d\nu_J(ns)-\hat{\varepsilon}_1^{(n)}(t)+\alpha\cdot\tilde{\varepsilon}_{2,2}^{(n)}(t).$$

Plugging into (5.49) and eliminating $\alpha\cdot\tilde{\varepsilon}_{22}^{(n)}$ gives for $t\le\bar{\tau}^{(n)}$,

$$(1-\alpha)\cdot\tilde{\varepsilon}_{2,2}^{(n)}(t)\le\tilde{\varepsilon}_{2,3}^{(n)}(t)+\tilde{\varepsilon}_{2,4}^{(n)}(t)-\hat{\varepsilon}_1^{(n)}(t)+\int_0^t\hat{\varepsilon}_1^{(n)}(t-s)d\nu_J(ns).$$

Plugging into (5.48) leads to the conclusion that for $t\le\bar{\tau}^{(n)}$,

$$\hat{\varepsilon}_2^{(n)}(t)\le\tilde{\varepsilon}_{2,1}^{(n)}(t)+\frac{\mu_g}{1-\alpha}\Big(\tilde{\varepsilon}_{2,3}^{(n)}(t)+\tilde{\varepsilon}_{2,4}^{(n)}(t)+\int_0^t\hat{\varepsilon}_1^{(n)}(t-s)d\nu_J(ns)-\hat{\varepsilon}_1^{(n)}(t)\Big).$$

Following the argument used for $\tilde{\varepsilon}_{1,1}^{(n)}$ in (5.45) and (5.46) to both $\tilde{\varepsilon}_{2,1}^{(n)}$ and $\tilde{\varepsilon}_{2,3}^{(n)}$, we obtain for $t \le T \wedge \bar{\tau}^{(n)}$,

$$\tilde{\varepsilon}_{2,1}^{(n)}(t) \le w\big(\hat{Z}_g^{(n)}(\cdot \wedge \bar{\tau}^{(n)}), \delta\big) + 2 \cdot J(n\delta) \cdot \sup_{u \le T} \big|\hat{Z}_g^{(n)}(u \wedge \bar{\tau}^{(n)})\big|,$$

$$\tilde{\varepsilon}_{2,3}^{(n)}(t) \le w\big(\hat{X}^{(n)}(\cdot \wedge \bar{\tau}^{(n)}), \delta\big) + 2 \cdot J(n\delta) \cdot \sup_{u \le T} \big|\hat{X}^{(n)}(u \wedge \bar{\tau}^{(n)})\big|.$$

The upper bounds on the right hand sides above converge to 0 in probability by the weak convergence of $\hat{Z}_g^{(n)}$ and $\hat{X}^{(n)}$ in Lemma 5.4, as well as the fact that $J(\infty) = 0$. Applying the argument used in (5.47) for $\tilde{\varepsilon}_{1,2}^{(n)}$ to $\tilde{\varepsilon}_{2,4}^{(n)}$, we also obtain

$$\tilde{\varepsilon}_{2,4}^{(n)}(t) \le \lambda_0 \cdot \sup_{s>0} \sqrt{s} J(s) \cdot \frac{\delta}{2} + \lambda_0 \cdot \sqrt{n} \cdot J(n\delta) \cdot t \to 0 \quad \text{u.o.c.},$$

under Assumption A4-(3.12) by letting $n \to \infty$ and $\delta \to 0$. For the last, applying the convergence for $\hat{\varepsilon}_1^{(n)}$ and (5.37), we can show the convergence of $\hat{\varepsilon}_2^{(n)}$. This finishes the proof. □

Now, we are ready to prove Theorem 3.2. To apply [30, Theorem 2.2] for convergence of stochastic integrals, the joint convergence of the component processes is necessary, for which Lemma 5.4 for the martingales and Theorem 3.1 are ready to be applied.

Proof of Theorem 3.2. Our proof starts with the joint convergence of $\big(\hat{N}^{(n)}, \hat{\Lambda}^{(n)}, \hat{Y}^{(n)}\big)$ defined in (3.11), where [30, Theorem 2.2] concerning weak convergence of stochastic integrals is used. Recalling that as $n \to \infty$, $\big(\bar{N}^{(n)}, \bar{\Lambda}^{(n)}, \bar{Y}^{(n)}\big)$ converges to a deterministic limit in Theorem 3.1, we already have from Lemma 5.4 that

$$\big(\bar{N}^{(n)}, \bar{\Lambda}^{(n)}, \bar{Y}^{(n)}, \hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}\big) \Rightarrow \big(\bar{N}, \bar{\Lambda}, \bar{Y}, \hat{X}, \hat{Z}_\varphi\big) \quad \text{in} \quad (\mathbb{D}^5, J_1),$$

and from (5.30) that

$$\begin{aligned} &\big(1 - H\big(\mu_\varphi \tilde{\Lambda}^{(n)}(t)\big)\big) \cdot \hat{\Lambda}^{(n)}(t) \\ &= \sqrt{n}\,\big(\bar{\Lambda}_1^{(n)}(t) + \bar{\Lambda}_2^{(n)}(t) - \bar{\varepsilon}_1^{(n)}(t)\big) = \hat{\Lambda}_1^{(n)}(t) + \hat{\Lambda}_2^{(n)}(t) - \hat{\varepsilon}_1^{(n)}(t). \end{aligned} \tag{5.50}$$

Making use of $\hat{X}^{(n)}$ in (5.1), we can write

$$\hat{\Lambda}_1^{(n)}(t) = \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\hat{X}^{(n)}(s).$$

Applying [30, Theorem 2.2], we can have

$$\big(\bar{N}^{(n)}, \bar{\Lambda}^{(n)}, \bar{Y}^{(n)}, \hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}, \hat{\Lambda}_1^{(n)}\big) \Rightarrow \big(\bar{N}, \bar{\Lambda}, \bar{Y}, \hat{X}, \hat{Z}_\varphi, \hat{\Lambda}_1\big) \quad \text{in} \quad (\mathbb{D}^6, J_1), \tag{5.51}$$

where

$$\hat{\Lambda}_1(t) = \int_0^t H(\bar{Y}(s))\, d\hat{X}(s).$$

Next, applying the mean value theorem, we have

$$\begin{aligned} \hat{\Lambda}_2^{(n)}(t) &= \sqrt{n} \int_0^t \big(H(\bar{Y}^{(n)}(u)) - H(\mu_\varphi \bar{\Lambda}^{(n)}(u))\big)\, d\bar{\Lambda}^{(n)}(u) \\ &= \int_0^t H'\big(\tilde{Y}_1^{(n)}(u)\big) \sqrt{n}\big(\bar{Y}^{(n)}(u) - \mu_\varphi \bar{\Lambda}^{(n)}(u)\big)\, d\bar{\Lambda}^{(n)}(u), \end{aligned} \tag{5.52}$$

for some $\tilde{Y}_1^{(n)}$ taking value between $\bar{Y}^{(n)}$ and $\mu_\varphi \bar{\Lambda}^{(n)}$. Moreover, applying (5.34) we have

$$\sqrt{n}\big(\bar{Y}^{(n)}(t) - \mu_\varphi \bar{\Lambda}^{(n)}(t)\big) = \hat{Z}_\varphi^{(n)}(t) + \mu_\varphi \hat{X}^{(n)}(t) - \hat{\varepsilon}_2^{(n)}(t). \tag{5.53}$$

Now, applying (5.51) to $\tilde{\Lambda}^{(n)}$ in (5.50), $\tilde{Y}_1^{(n)}$ in (5.52) and (5.53), we obtain

$$\begin{aligned}&\big(\hat{X}^{(n)},\hat{Z}_{\varphi}^{(n)},\hat{\Lambda}_1^{(n)},\tilde{\Lambda}^{(n)},\tilde{Y}_1^{(n)},\sqrt{n}\big(\bar{Y}^{(n)}-\mu_{\varphi}\,\bar{\Lambda}^{(n)}\big)\big)\\ \Rightarrow&\big(\hat{X},\hat{Z}_{\varphi},\hat{\Lambda}_1,\bar{\Lambda},\bar{Y},\hat{Z}_{\varphi}+\mu_{\varphi}\,\hat{X}\big)\quad\text{in}\quad(\mathbb{D}^6,J_1),\end{aligned}$$

where the limit for $\hat{\varepsilon}_2^{(n)}$ in Lemma 5.5 is also applied. Plugging the limit above into (5.52) for $\hat{\Lambda}_2^{(n)}$ and applying [30, Theorem 2.2], we obtain

$$\begin{aligned}&\big(\hat{X}^{(n)},\hat{Z}_{\varphi}^{(n)},\hat{\Lambda}_1^{(n)},\tilde{\Lambda}^{(n)},\sqrt{n}\big(\bar{Y}^{(n)}-\mu_{\varphi}\,\bar{\Lambda}^{(n)}\big),\hat{\Lambda}_2^{(n)}\big)\\ \Rightarrow&\big(\hat{X},\hat{Z}_{\varphi},\hat{\Lambda}_1,\bar{\Lambda},\hat{Z}_{\varphi}+\mu_{\varphi}\,\hat{X},\hat{\Lambda}_2\big)\quad\text{in}\quad(\mathbb{D}^6,J_1),\end{aligned}$$

where

$$\hat{\Lambda}_2(t)=\int_0^t H'(\bar{Y}(s))\big(\hat{Z}_{\varphi}(s)+\mu_{\varphi}\,\hat{X}(s)\big)\,d\bar{\Lambda}(s).$$

Plugging into (5.50) and applying Lemma 5.5 gives

$$\big(\hat{X}^{(n)},\sqrt{n}\big(\bar{Y}^{(n)}-\mu_{\varphi}\,\bar{\Lambda}^{(n)}\big),\hat{\Lambda}^{(n)}\big)\Rightarrow\big(\hat{X},\hat{Z}_{\varphi}+\mu_{\varphi}\,\hat{X},\hat{\Lambda}\big)\quad\text{in}\quad(\mathbb{D}^3,J_1),\tag{5.54}$$

where we also need the property in (5.32), and

$$\hat{\Lambda}(t)=\frac{\hat{\Lambda}_1(t)+\hat{\Lambda}_2(t)}{1-H(\bar{Y}(t))}.$$

By the fact $\bar{Y}=\mu_{\varphi}\bar{\Lambda}$, we can rewrite (5.53) as

$$\sqrt{n}\big(\bar{Y}^{(n)}-\mu_{\varphi}\,\bar{\Lambda}^{(n)}\big)=\hat{Y}^{(n)}-\mu_{\varphi}\,\hat{\Lambda}^{(n)}.$$

Thus, with the FCLT for $\hat{\Lambda}^{(n)}$ proved above, we can further rewrite (5.54) as

$$\big(\hat{X}^{(n)},\hat{\Lambda}^{(n)},\hat{Y}^{(n)}\big)\Rightarrow\big(\hat{X},\hat{\Lambda},\hat{Y}\big)\quad\text{in}\quad(\mathbb{D}^3,J_1).$$

Noticing that $\hat{N}^{(n)}=\hat{\Lambda}^{(n)}+\hat{X}^{(n)}$, we further have the joint convergence

$$\big(\hat{\Lambda}^{(n)},\hat{Y}^{(n)},\hat{N}^{(n)}\big)\Rightarrow\big(\hat{\Lambda},\hat{Y},\hat{N}\big)\quad\text{in}\quad(\mathbb{D}^3,J_1).$$

Finally, we derive the SDE for $(\hat{N},\hat{Y})$ in (3.15). Applying the mean value theorem to (5.6), for some process $\tilde{Y}_2^{(n)}$ between $\bar{Y}^{(n)}$ and $\bar{Y}$, we have

$$\hat{N}^{(n)}(t)=\hat{X}^{(n)}(t)+\int_0^t H\big(\bar{Y}^{(n)}(u-)\big)\,d\hat{N}^{(n)}(u)+\int_0^t H'\big(\tilde{Y}_2^{(n)}(u)\big)\hat{Y}^{(n)}(u)\,d\bar{\Lambda}(u)-\hat{\varepsilon}_1^{(n)}(t).$$

Letting $n\to\infty$, together with (5.7) and the convergence above, we obtain

$$\begin{aligned}d\hat{N}(t)&=d\hat{X}(t)+H(\bar{Y}(t))\,d\hat{N}(t)+H'(\bar{Y}(t))\hat{Y}(t)\,d\bar{\Lambda}(t),\\ d\hat{Y}(t)&=d\hat{Z}_{\varphi}(t)+\mu_{\varphi}\,d\hat{N}(t).\end{aligned}$$

This gives the desired SDE for $(\hat{N},\hat{Y})$ in (3.15) with the expression in (3.16). □

6. Proof of Theorem 4.1

For the extension in Section 4, the same proof ideas can be applied. Abuse of notion, we let $\bar{X}^{(n)}$ and $\hat{X}^{(n)}$ be the associated LLN and CLT scaled processes defined in (5.1), and $(\bar{Z}_\varphi^{(n)}, \bar{Z}_g^{(n)})$ be the auxiliary processes defined in (5.2). With the presence of the additional marks in ϕ in (4.1), we further define the following LLN and CLT scaled auxiliary processes

$$\begin{aligned} \bar{Z}_h^{(n)}(t) &= \frac{1}{n}\sum_{j\geq 1}\Big(h(\eta_j^{(n)}, y_j) - H(y_j)\Big)\Big|_{y_j=\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)}\mathbf{1}\big(\bar{\tau}_j^{(n)}\leq t\big),\\ \bar{Z}_{f_k}^{(n)}(t) &= \frac{1}{n}\sum_{j\geq 1}\Big(f_k\big(\eta_j^{(n)}, y_j\big) - \mathbb{E}\big[f_k(\eta, y_j)\big]\Big)\Big|_{y_j=\bar{Y}^{(n)}(\tau_j^{(n)}-)}\mathbf{1}\big(\bar{\tau}_j^{(n)}\leq t\big), \end{aligned} \tag{6.1}$$

recalling the new functions h and f_k in Assumption A6-(4.3) and (4.4), and let

$$\big(\hat{Z}_h^{(n)}, \hat{Z}_{f_k}^{(n)}\big) = \sqrt{n}\,\big(\bar{Z}_h^{(n)}, \bar{Z}_{f_k}^{(n)}\big). \tag{6.2}$$

For every fixed $k>0$, by the same reasoning of $\{\eta_j^{(n)}\}_{j\geq 1}$ being unpredictable marks, $\bar{Z}_h^{(n)}$ and $\bar{Z}_{f_k}^{(n)}$ above are also $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\geq 0}$-adapted local martingales. With the new notion above, slightly different from (5.3), we have following representations for the LLN-scaled process:

$$\begin{aligned} \bar{\Lambda}^{(n)}(t) &= \lambda_0 t + \frac{1}{n}\sum_{j\geq 1} h\big(\eta_j^{(n)}, \bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\mathbf{1}\big(\bar{\tau}_j^{(n)}\leq t\big) - \bar{\varepsilon}_1^{(n)}(t)\\ &= \lambda_0 t + \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\,d\bar{N}^{(n)}(s) + \bar{Z}_h^{(n)}(t) - \bar{\varepsilon}_1^{(n)}(t), \end{aligned} \tag{6.3}$$

where LLN-scaled residual term takes the form

$$\bar{\varepsilon}_1^{(n)}(t) = \frac{1}{n}\sum_{j\geq 1}\Big(\int_{n(t-\bar{\tau}_j^{(n)})}^{\infty}\phi\big(u, \eta_j^{(n)}, \bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\,du\Big)\mathbf{1}\big(\bar{\tau}_j^{(n)}\leq t\big). \tag{6.4}$$

Since $\bar{Y}^{(n)}$ in the extended model takes the same form as the original one (2.1), we still have the expression of $\bar{Y}^{(n)}(t)$ in (5.4) with the same formula of $\bar{\varepsilon}_2^{(n)}(t)$ in (5.5). Therefore, we only focus on the proofs of $\bar{\Lambda}^{(n)}$, $\bar{N}^{(n)}$, $\bar{\varepsilon}_1^{(n)}$ related variables and highlight the differences.

For the localization technique, let $k_0>0$ be defined in (5.9), with the modified H in (4.3), and for such k_0, let the first passage time for $\bar{Y}^{(n)}$ be defined by

$$\bar{\tau}_{k_0}^{(n)} := \inf\{t>0, |\bar{Y}^{(n)}|(t) > k_0\}.$$

By the assumption of $H\in\mathbb{C}(\mathbb{R})$ and the uniformly integrability on $\{f_{k_0}(\eta, y), |y|\leq k_0\}$ in Assumption A6, we have

$$\alpha := \sup_{|y|\leq k_0} H(y) = \sup_{|y|\leq k_0}\mathbb{E}\big[h(\eta, y)\big] < 1 \quad\text{and}\quad \sup_{|y|\leq k_0}\mathbb{E}\big[f_{k_0}(\eta, y)\big]\leq \beta < \infty \tag{6.5}$$

for some $\beta>0$, where we drop the subscript k_0 in the associated notations below.

6.1. Sketch proof of the convergence of LLN-scaled processes $\big(\bar{N}^{(n)}, \bar{\Lambda}^{(n)}, \bar{Y}^{(n)}\big)$.

First, the results in Lemma 5.1 hold under Assumption A6, and in addition, $\bar{Z}_h^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$, $\bar{Z}_f^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$ are also martingales with respect to $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\geq 0}$ and

$$\bar{Z}_h^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad\text{and}\quad \bar{Z}_f^{(n)}(t\wedge\bar{\tau}^{(n)})\to 0 \quad\text{u.o.c. in probability.} \tag{6.6}$$

To prove the martingale property, from (6.3), we obtain

$$\begin{aligned}0 &\le \int_0^t \Big(1 - H\big(\bar{Y}^{(n)}(s-)\big)\Big)\, d\bar{N}^{(n)}(s)\\ &= \bar{X}^{(n)}(t) + \bar{\Lambda}^{(n)}(t) - \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\bar{N}^{(n)}(s)\\ &\le \lambda_0\, t + \bar{X}^{(n)}(t) + \bar{Z}_h^{(n)}(t) - \bar{\varepsilon}_1^{(n)}\,.\end{aligned}$$

By a similar argument to get (5.14), using the bound in (6.5) also gives

$$\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big] \le \frac{\lambda_0}{1-\alpha}\cdot t \quad \forall t>0. \tag{6.7}$$

Applying the BDG-inequality, we obtain for some constant $c_2>0$ from the inequality that

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\le t}\big(\bar{X}^{(n)}(s\wedge\bar{\tau}^{(n)})\big)^2\Big] &\le c_2\,\mathbb{E}\Big[[\bar{X}^{(n)}](t\wedge\bar{\tau}^{(n)})\Big]\\ &= \frac{c_2}{n}\,\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big] \le \frac{c_2}{n}\frac{\lambda_0}{1-\alpha}\cdot t\,, \quad \forall t>0.\end{aligned} \tag{6.8}$$

Observe the additional term $\bar{Z}_h^{(n)}$ comparing with (5.13). By the assumption on the marks $\{\eta_j^{(n)}\}_{j\ge1}$ being unpredictable, $\bar{Z}_h^{(n)}$ in (6.1) is a local martingale process with quadratic variation given by

$$\big[\bar{Z}_h^{(n)}\big](t) = \frac{1}{n^2}\sum_{j\ge1}\Big(h(\eta_j^{(n)},y_j) - H(y_j)\Big)^2\Big|_{y_j=\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)}\mathbf{1}\big(\bar{\tau}_j^{(n)}\le t\big).$$

Making use of the fact that $(x^2+y^2)\le(|x|+|y|)^2$, one have

$$\big[\bar{Z}_h^{(n)}\big]^{1/2}(t) \le \frac{1}{n}\sum_{j\ge1}\Big(h(\eta_j^{(n)},y_j) + H(y_j)\Big)\Big|_{y_j=\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)}\mathbf{1}\big(\bar{\tau}_j^{(n)}\le t\big).$$

Applying the BDG inequality and recalling the fact $H(y)=\mathbb{E}[h(\eta,y)]<1$, we have for some constant c_1,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\le t}|\bar{Z}_h^{(n)}|(s\wedge\bar{\tau}^{(n)})\Big] &\le c_1\,\mathbb{E}\Big[\big[\bar{Z}_h^{(n)}\big]^{1/2}(t\wedge\bar{\tau}^{(n)})\Big]\\ &\le 2\,c_1\,\mathbb{E}\Big[\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\bar{N}^{(n)}(ds\wedge\bar{\tau}^{(n)})\Big]\\ &\le 2\,c_1\,\mathbb{E}\big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)})\big] \le \frac{2\,\lambda_0\,c_1}{1-\alpha}\cdot t,\end{aligned}$$

which implies the martingale property of $\bar{Z}_h^{(n)}(\cdot\wedge\bar{\tau}^{(n)})$.

To show the convergence for $\bar{Z}_h^{(n)}$, the method of localization used for FLLN of triangular array is applied, c.f. [13, Theorem 2.2.11]. For every $\varepsilon>0$, let $\tilde{h}(z,y) := h(z,y)\mathbf{1}(h(z,y)\le n\varepsilon)$ and

$$\bar{Z}_{\tilde{h}}^{(n)}(t) := \frac{1}{n}\sum_{j\ge1}\Big(\tilde{h}\big(\eta_j^{(n)},y_j\big) - \mathbb{E}\big[\tilde{h}(\eta,y_j)\big]\Big)\Big|_{y_j=\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)}\mathbf{1}\big(\bar{\tau}_j^{(n)}\le t\big)$$

comparing with $\bar{Z}_h^{(n)}$ in (6.1). Since $\tilde{h}\le h\wedge(n\varepsilon)$, we have for every $y\in\mathbb{R}$

$$\mathbb{E}\big[\tilde{h}^2(\eta,y)\big] \le \mathbb{E}\big[n\varepsilon\cdot\tilde{h}(\eta,y)\big] \le n\varepsilon\,\mathbb{E}\big[h(\eta,y)\big] = n\varepsilon\, H(y) \le n\varepsilon.$$

Applying the BDG inequality, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{s\le t}\big(\bar Z^{(n)}_{\tilde h}(s\wedge\bar\tau^{(n)})\big)^2\Big]\le c_2\,\mathbb{E}\Big[\big[\bar Z^{(n)}_{\tilde h}\big](t\wedge\bar\tau^{(n)})\Big]\\
\le&\frac{c_2}{n}\,\mathbb{E}\Big[\int_0^t\mathbb{E}\big[\tilde h^2(\eta,\bar Y^{(n)}(s-))\big]\,d\bar N^{(n)}(ds\wedge\bar\tau^{(n)})\Big] \qquad (6.9)\\
\le& c_2\,\varepsilon\cdot\mathbb{E}\big[\bar N^{(n)}(t\wedge\bar\tau^{(n)})\big]\le\frac{\lambda_0c_2\,t}{1-\alpha}\cdot\varepsilon.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{s\le t}\big|\bar Z^{(n)}_h-\bar Z^{(n)}_{\tilde h}\big|(s\wedge\bar\tau^{(n)})\Big]\\
\le& 2\cdot\mathbb{E}\Big[\int_0^t\mathbb{E}\big[h(\eta,y);h(\eta,y)>n\varepsilon\big]\Big|_{y=\bar Y^{(n)}(s-)}\bar N^{(n)}(ds\wedge\bar\tau^{(n)})\Big] \qquad (6.10)\\
\le& 2\sup_{|y|\le k_0}\mathbb{E}\big[h(\eta,y);h(\eta,y)>n\varepsilon\big]\cdot\mathbb{E}\big[\bar N^{(n)}(t\wedge\bar\tau^{(n)})\big],
\end{aligned}$$

where the last line above converges to 0 as $n\to\infty$ by the assumption of uniform integrability in Assumption A6-(4.5) and (6.7). The inequalities above can be used to show the limit for $\bar Z^{(n)}_h$ in (6.6). The martingale property and the convergence for $\bar Z^{(n)}_f$ can be proved similarly.

We next show that $\bar N^{(n)}$ is a tight family in $\mathbb{D}[0,T]$ in the sense of the uniformity, similar to Lemma 5.2, that is, under Assumption A6, we claim for every $\varepsilon>0$,

$$\limsup_{\delta\to0}\limsup_{n\to\infty}\mathbb{P}\Big(w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)},\delta)>\varepsilon\Big)=0.$$

To prove this, we first show that $\bar N^{(n)}$ is equi-continuous in expectation, that is,

$$\lim_{s\to t}\limsup_{n\to\infty}\mathbb{E}\big[\bar N^{(n)}(t\wedge\bar\tau^{(n)})-\bar N^{(n)}(s\wedge\bar\tau^{(n)})\big]=0. \qquad (6.11)$$

For $t>s>0$, we have from (6.3) and the fact $\bar\varepsilon^{(n)}_1\ge0$ that

$$\begin{aligned}
\bar\Lambda^{(n)}(t\wedge\bar\tau^{(n)})-\bar\Lambda^{(n)}(s\wedge\bar\tau^{(n)})&=\big(\bar\Lambda^{(n)}(t\wedge\bar\tau^{(n)})-\bar\Lambda^{(n)}(s)\big)\mathbf{1}(s\le\bar\tau^{(n)})\\
&\le\lambda_0(t-s)+\big(\bar Z^{(n)}_h(t\wedge\bar\tau^{(n)})-\bar Z^{(n)}_h(s\wedge\bar\tau^{(n)})\big)\\
&\quad+\int_s^tH\big(\bar Y^{(n)}(u-)\big)\,d\bar N^{(n)}(u\wedge\bar\tau^{(n)})+\bar\varepsilon^{(n)}_1(s)\mathbf{1}(s\le\bar\tau^{(n)}).
\end{aligned}$$

For $\bar\varepsilon^{(n)}_1$ in (6.4), recalling (I,f) in Assumption A6-(4.4), we have

$$\bar\varepsilon^{(n)}_1(s)\mathbf{1}(s\le\bar\tau^{(n)})\le\frac1n\sum_{j\ge1}I\big(n(s-\bar\tau^{(n)}_j)\big)\,f\big(\eta^{(n)}_j,\bar Y^{(n)}(\bar\tau^{(n)}_j-)\big)\mathbf{1}\big(\bar\tau^{(n)}_j\le s\wedge\bar\tau^{(n)}\big). \qquad (6.12)$$

By a similar argument as in the proof of Lemma 5.2, we obtain that for every $t>s>r>0$,

$$\begin{aligned}
&(1-\alpha)\,\big(\bar N^{(n)}(t\wedge\bar\tau^{(n)})-\bar N^{(n)}(s\wedge\bar\tau^{(n)})\big)\\
\le&\lambda_0\,(t-s)+\beta\,I\big(n(s-r)\big)\,\bar N^{(n)}(r\wedge\bar\tau^{(n)})+\beta\,\big(\bar N^{(n)}(s\wedge\bar\tau^{(n)})-\bar N^{(n)}(r\wedge\bar\tau^{(n)})\big)\\
&+2\,\Big(\sup_{u\le t}\big|\bar X^{(n)}\big|(u\wedge\bar\tau^{(n)})+\sup_{u\le t}\big|\bar Z^{(n)}_h\big|(u\wedge\bar\tau^{(n)})+\sup_{u\le t}\big|\bar Z^{(n)}_f\big|(u\wedge\bar\tau^{(n)})\Big).
\end{aligned}$$

Thus applying the inequality (5.21), we have for every $\delta > \delta' > 0$,

$$\begin{aligned} & w\big(\bar N^{(n)}(\cdot\wedge\bar\tau^{(n)}),\delta\big) \\ & \le 2\,\bar N^{(n)}(\delta\wedge\bar\tau^{(n)}) + \frac{2}{1-\alpha}\Big(\lambda_0\delta + I(n\delta')\cdot\beta\,\bar N^{(n)}(T\wedge\bar\tau^{(n)})\Big) \\ & \quad + \frac{2\beta}{1-\alpha}\sum_{j=1}^{v}\Big(\bar N^{(n)}(t_j\wedge\bar\tau^{(n)}) - \bar N^{(n)}\big((t_j-\delta')\wedge\bar\tau^{(n)}\big)\Big) \\ & \quad + \frac{4}{1-\alpha}\Big(\sup_{u\le T}|\bar X^{(n)}|(u\wedge\bar\tau^{(n)}) + \sup_{u\le T}|\bar Z_h^{(n)}|(u\wedge\bar\tau^{(n)}) + \sup_{u\le T}|\bar Z_f^{(n)}|(u\wedge\bar\tau^{(n)})\Big). \end{aligned} \tag{6.13}$$

Observe the last two additional η-terms in the inequality above comparing with (5.22). Then the claim will follow from the assumptions and the convergence result in the first step.

We will then show that the convergence of $\bar\varepsilon_1^{(n)}$ and $\bar\varepsilon_2^{(n)}$ as in Lemma 5.3. Here we have

$$\begin{aligned} \bar\varepsilon_1^{(n)}(t\wedge\bar\tau^{(n)}) & \le \frac{1}{n}\sum_{j\ge1} I\big(n(t\wedge\bar\tau^{(n)} - \bar\tau_j^{(n)})\big)\, f\big(\eta_j^{(n)},\bar Y^{(n)}(\bar\tau_j^{(n)}-)\big)\mathbf{1}\big(\bar\tau_j^{(n)}\le s\wedge\bar\tau^{(n)}\big) \\ & = \int_0^{t\wedge\bar\tau^{(n)}} I\big(n(t\wedge\bar\tau^{(n)}-u)\big)\Big(d\bar Z_f^{(n)}(u) + \mathbb{E}\big[f(\eta,\bar Y^{(n)}(u-))\big]d\bar N^{(n)}(du)\Big) \\ & \le \int_0^{t\wedge\bar\tau^{(n)}} I\big(n(t\wedge\bar\tau^{(n)}-u)\big)\Big(d\bar Z_f^{(n)}(u) + \beta\,d\bar N^{(n)}(u)\Big)\,. \end{aligned} \tag{6.14}$$

Comparing with (5.25), we use the bound in (6.5) above. Then the convergence of $\bar\varepsilon_2^{(n)}(\cdot\wedge\bar\tau^{(n)})$ will follow from a similar argument as in the proof of Lemma 5.3. Moreover, the convergence of $\bar\varepsilon_2^{(n)}(\cdot\wedge\bar\tau^{(n)})$ can be proved following the same procedure as in the proof of Lemma 5.3.

Finally, given the above results, the proof for the FLLN will follow the same steps as the proof of Theorem 3.1.

6.2. **Sketch proof of the convergence of CLT-scaled processes $\big(\hat N^{(n)},\hat\Lambda^{(n)},\hat Y^{(n)}\big)$.** We give first following representation for the CLT-scaled $\hat N^{(n)}$ from (6.3):

$$\begin{aligned} \hat N^{(n)}(t) & = \hat X^{(n)}(t) + \sqrt{n}\big(\bar\Lambda^{(n)}(t) - \bar\Lambda(t)\big) \\ & = \hat X^{(n)}(t) + \int_0^t H\big(\bar Y^{(n)}(s-)\big)\,d\hat N^{(n)}(s) \\ & \quad + \sqrt{n}\int_0^t \Big(H\big(\bar Y^{(n)}(s)\big) - H\big(\bar Y(s)\big)\Big)\,d\bar\Lambda(s) + \hat Z_h^{(n)}(t) - \hat\varepsilon_1^{(n)}(t)\,. \end{aligned} \tag{6.15}$$

Observe the additional term $\hat Z_h^{(n)}$ in (6.2) of $\{\eta_j^{(n)}\}_{j\ge1}$ comparing with (5.6). We still have the same representation for the CLT-scaled $\hat Y^{(n)}$ as in (5.7). We will mainly focus on the proofs related to $\hat\Lambda^{(n)}$, $\hat N^{(n)}$, $\hat\varepsilon_1^{(n)}$ and highlight the differences. Recalling $\hat X^{(n)}$ in (5.1), $\hat Z_\varphi^{(n)}$, $\hat Z_g^{(n)}$ in (5.2), and $\hat Z_h^{(n)}$, $\hat Z_f^{(n)}$ in (6.2), we first prove the convergence of the local martingale terms above analogy to Lemma 5.4. Under Assumption A7-(4.7) and (4.8), we have

$$\sup_{|y|\le k_0}\mathbb{E}\big[h^2(\eta,y)\big]\le\gamma \quad\text{and}\quad \sup_{|y|\le k_0}\mathbb{E}\big[f^2(\eta,y)\big]\le\gamma, \tag{6.16}$$

for some $\gamma>0$, in addition to (6.5).

Lemma 6.1. *Under Assumptions A6 and A7, we have as $n\to\infty$,*

$$\big(\hat X^{(n)},\hat Z_\varphi^{(n)},\hat Z_h^{(n)}\big)\Rightarrow\big(\hat X,\hat Z_\varphi,\hat Z_h\big)\quad in\quad(\mathbb{D}^3,J_1),$$

where $\big(\hat{Z}_\varphi, \hat{Z}_h\big)$ *is a two dimensional mean-zero Gaussian martingales characterized by* (4.10)*, and* $\hat{X}$ *is a mean-zero time-changed Brownian motion with variances* $\bar{\Lambda}(t)$ *and independent of* $\big(\hat{Z}_\varphi, \hat{Z}_h\big)$*, as given in Theorem 4.1.* $\{\hat{Z}_f^{(n)}\}_{n\geq 1}$ *is a* $\mathbb{C}$*-tight families in* $\mathbb{D}$*.*

Proof. We apply [24, Theorem 3.22 in Chapter VIII] as the proof of Lemma 5.4, where by the assumption of unpredictable marks in (2.3), $\big(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}, \hat{Z}_h^{(n)}\big)$ is an $\{\bar{\mathscr{F}}_t^{(n)}\}_{t\geq 0}$-adapted locally-square-integrable martingale, so is $\hat{Z}_f^{(n)}$.

We first check the condition 3.23 in [24, Theorem 3.22 in Chapter VIII] for $\big(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}, \hat{Z}_h^{(n)}\big)$. It is true that $\hat{X}^{(n)}$ has bounded jump size of $1/\sqrt{n}$, and (5.39) still hold for $\hat{Z}_\varphi^{(n)}$. We also have for $\hat{Z}_h^{(n)}$ in (6.2), for every $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{s\leq T\wedge\bar{\tau}^{(n)}} \Big(\Delta\,\hat{Z}_h^{(n)}(s)\Big)^2 \mathbf{1}\big(\big|\Delta\,\hat{Z}_h^{(n)}(s)\big| > \varepsilon\big)\Big] \\
&= \mathbb{E}\Big[\int_0^{T\wedge\bar{\tau}^{(n)}} \mathbb{E}\Big[\big(h(\eta,y)-H(y)\big)^2; |h(\eta,y)-H(y)| > \sqrt{n}\varepsilon\Big]\Big|_{y=\bar{Y}^{(n)}(s-)} d\bar{N}^{(n)}(s)\Big] \\
&\leq \frac{2\,\lambda_0\,T}{1-\alpha}\cdot \sup_{|y|\leq k_0} \mathbb{E}\Big[1+h^2(\eta,y); h(\eta,y) > \sqrt{n}\varepsilon\Big]
\end{aligned} \tag{6.17}$$

where we make use of the fact that $H(y) < 1$. By the uniform integrability of $\{h^2(\eta,y), |y|\leq k_0\}$ in Assumption A7-(4.7), the last line above converges to 0 as $n\to\infty$ and thus the condition 3.23 mentioned above for $\big(\hat{X}^{(n)}, \hat{Z}_\varphi^{(n)}, \hat{Z}_h^{(n)}\big)$ hold.

Next, we check the condition $[\hat{\gamma}_5' - D]$. For arbitrary $a, b, c \in \mathbb{R}$, it is not hard to find that $\big(a\,\hat{X}^{(n)} + b\,\hat{Z}_\varphi^{(n)} + c\,\hat{Z}_h^{(n)}\big)$ is a local martingale with quadratic variation given by

$$\big[a\,\hat{X}^{(n)} + b\,\hat{Z}_\varphi^{(n)} + c\,\hat{Z}_h^{(n)}\big](t) = \frac{1}{n}\sum_{j\geq 1} \kappa_j^{(n)}\big(\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\mathbf{1}(\bar{\tau}_j^{(n)}\leq t),$$

where for simplicity we denote by for $y\in\mathbb{R}$,

$$\kappa_j^{(n)}(y) = \Big(a + b\,\big(\varphi(\infty,\xi_j^{(n)}) - \mu_\varphi\big) + c\,\big(h(\eta_j^{(n)},y) - H(y)\big)\Big)^2.$$

It is straightforward to check that $\big\{\kappa_j^{(n)}(y), |y|\leq k_0\big\}$ is uniformly integrable under Assumption A7 and (4.7). Therefore, one can check

$$\begin{aligned}
&\big[a\,\hat{X}^{(n)} + b\,\hat{Z}_\varphi^{(n)} + c\,\hat{Z}_h^{(n)}\big](t) - \int_0^t \mathbb{E}\big[\kappa(y)\big]\Big|_{y=\bar{Y}^{(n)}(s-)} d\bar{N}^{(n)}(s) \\
&= \frac{1}{n}\sum_{j\geq 1}\Big(\kappa_j^{(n)}(y_j) - \mathbb{E}\big[\kappa_j^{(n)}(y_j)\big]\Big)\Big|_{y_j=\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)} \mathbf{1}(\bar{\tau}_j^{(n)}\leq t) \to 0 \quad \text{u.o.c. in probability.}
\end{aligned} \tag{6.18}$$

Moreover, under Assumption A7, we have

$$\mathbb{E}\big[\kappa_j^{(n)}(y)\big] = \mathbb{E}\big[\kappa(y)\big] = a^2 + \mathrm{Var}\big(b\,\varphi(\infty,\xi) + c\,h(\eta,y)\big) \in \mathbb{C}(\mathbb{R}).$$

By the fact $\big(\bar{Y}^{(n)}, \bar{N}^{(n)}\big) \to (\bar{Y}, \bar{N})$ u.o.c. in probability, we can apply [30, Theorem 2.2] to the integral term in (6.18) and find that

$$\int_0^t \mathbb{E}\big[\kappa(y)\big]\Big|_{y=\bar{Y}^{(n)}(s-)} d\bar{N}^{(n)}(s) \to \int_0^t \mathbb{E}\big[\kappa(y)\big]\Big|_{y=\bar{Y}(s)} d\bar{N}(s) \quad \text{u.o.c. in probability.}$$

Or equivalently, for every $a, b, c \in \mathbb{R}$ and $t > 0$, we have shown that

$$\begin{aligned}
&\big[a\,\hat{X}^{(n)} + b\,\hat{Z}^{(n)}_{\varphi} + c\,\hat{Z}^{(n)}_{h}\big](t) \\
&\to \int_0^t \Big(a^2 + \mathrm{Var}\big(b\,\varphi(\infty,\xi) + c\,h(\eta,y)\big)\Big|_{y=\bar{Y}^{(n)}(s)}\Big)\, d\bar{\Lambda}(s) \\
&= \mathbb{E}\big[(a\hat{X}(t) + b\,\hat{Z}_{\varphi}(t) + c\,\hat{Z}_h(t))^2\big] \quad \text{u.o.c. in probability.}
\end{aligned}$$

This proves the convergence of $\big(\hat{X}^{(n)}, \hat{Z}^{(n)}_{\varphi}, \hat{Z}^{(n)}_{h}\big)$.

To prove the $\mathbb{C}$-tightness of $\{\hat{Z}^{(n)}_f\}_{n\geq 1}$ in (6.2), we apply [15, Theorem 8.6 & Remark 8.7] to show that $\{\hat{Z}^{(n)}_f\}_{n\geq 1}$ is a tight family in $\mathbb{D}$, and then claim that its maximum jump sizes converges to 0 in probability. By the unpredictability of marks $\{\eta^{(n)}_j\}_{j\geq 1}$, for every $t - s \leq \delta$ and $0 < s < t \leq T$,

$$\begin{aligned}
&\mathbb{E}\Big[\big(\hat{Z}^{(n)}_f(t\wedge\bar{\tau}^{(n)}) - \hat{Z}^{(n)}_f(s\wedge\bar{\tau}^{(n)})\big)^2\Big|\bar{\mathscr{F}}^{(n)}(s)\Big] \\
&= \mathbb{E}\Big[\int_s^t \mathrm{Var}\big(f(\eta,y)\big)\Big|_{y=\bar{Y}^{(n)}(u-)}\, d\bar{N}^{(n)}(u\wedge\bar{\tau}^{(n)})\Big|\bar{\mathscr{F}}^{(n)}(s)\Big] \\
&\leq \sup_{|y|\leq k_0} \mathbb{E}\big[f^2(\eta,y)\big]\cdot\mathbb{E}\Big[\bar{N}^{(n)}(t\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\Big|\bar{\mathscr{F}}^{(n)}(s)\Big] \\
&\leq \gamma\cdot\mathbb{E}\Big[w\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\Big|\bar{\mathscr{F}}^{(n)}(s)\Big],
\end{aligned} \tag{6.19}$$

recalling the bound for $\mathbb{E}[f^2(\eta, y)]$ in (6.16) and the modulus of continuity in (5.15), which gives the inequality [15, eqn (8.33)]. Then we verify [15, eqn (8.29)], that is,

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\mathbb{E}\Big[w\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\Big] = 0. \tag{6.20}$$

Applying (6.13) and (6.7), we have for every $\delta' \in (0, \delta)$,

$$\begin{aligned}
\mathbb{E}\Big[w\,\big(\bar{N}^{(n)}(\cdot\wedge\bar{\tau}^{(n)}),\delta\big)\Big] &\leq \frac{2\lambda_0\,\delta}{1-\alpha} + \frac{2\lambda_0\,\delta}{1-\alpha} + \frac{2\beta\,\lambda_0\,T}{1-\alpha}\frac{I(n\delta')}{1-\alpha} \\
&+ \frac{2\beta}{1-\alpha}\sum_{j=1}^{v}\mathbb{E}\Big[\bar{N}^{(n)}(t_j\wedge\bar{\tau}^{(n)}) - \bar{N}^{(n)}((t_j-\delta')\wedge\bar{\tau}^{(n)})\Big] \\
&+ \frac{4}{1-\alpha}\,\mathbb{E}\Big[\sup_{t\leq T}|\bar{X}^{(n)}|(t\wedge\bar{\tau}^{(n)}) + \sup_{t\leq T}|\bar{Z}^{(n)}_h|(t\wedge\bar{\tau}^{(n)}) + \sup_{t\leq T}|\bar{Z}^{(n)}_f|(t\wedge\bar{\tau}^{(n)})\Big].
\end{aligned} \tag{6.21}$$

Applying the BDG's inequality and the bound for $\mathbb{E}[h^2(\eta, y)]$ in (6.16), similar to (6.9) but the finiteness of L^2-norm, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\leq T}\big(\bar{Z}^{(n)}_h\big)^2(t\wedge\bar{\tau}^{(n)})\Big] \leq c_2\cdot\mathbb{E}\Big[\big[\bar{Z}^{(n)}_h\big](T\wedge\bar{\tau}^{(n)})\Big] \\
&\leq \frac{c_2}{n}\cdot\mathbb{E}\Big[\int_0^T \mathbb{E}\big[h^2(\eta,y)\big]\Big|_{y=\bar{Y}^{(n)}(s-)}\, d\bar{N}^{(n)}(s\wedge\bar{\tau}^{(n)})\Big] \\
&\leq \frac{c_2}{n}\cdot\sup_{|y|\leq k_0}\mathbb{E}\big[h^2(\eta,y)\big]\cdot\mathbb{E}\big[\bar{N}^{(n)}(T\wedge\bar{\tau}^{(n)})\big] \leq \frac{c_2\,\lambda_0\,\gamma}{1-\alpha}\cdot\frac{T}{n},
\end{aligned}$$

where (6.7) is also applied in the last line above. Similarly,

$$\mathbb{E}\big[\sup_{t\leq T}\big(\bar{Z}^{(n)}_f\big)^2(t\wedge\bar{\tau}^{(n)})\big] \leq \frac{c_2}{n}\cdot\sup_{|y|\leq k_0}\mathbb{E}\big[f^2(\eta,y)\big]\cdot\mathbb{E}\big[\bar{N}^{(n)}(T\wedge\bar{\tau}^{(n)})\big] \leq \frac{c_2\,\lambda_0\,\gamma}{1-\alpha}\cdot\frac{T}{n}.$$

Substituting into (6.21), applying Cauchy's inequality and (6.8) gives for every $\delta > \delta' > 0$

$$\mathbb{E}\Big[w\big(\bar{N}^{(n)}(\cdot \wedge \bar{\tau}^{(n)}), \delta\big)\Big] \leq c_0 \cdot \big(\delta + I(n\delta') + n^{-1/2}\big) \\ + c_0 \cdot \sum_{j=1}^{v} \mathbb{E}\Big[\bar{N}^{(n)}(t_j \wedge \bar{\tau}^{(n)}) - \bar{N}^{(n)}((t_j - \delta') \wedge \bar{\tau}^{(n)})\Big].$$

Letting $n \to \infty$ with the fact $I(\infty) = 0$, and then letting $\delta' \to 0$ with (6.11) applied, we have

$$\limsup_{n\to\infty} \mathbb{E}\Big[w\big(\bar{N}^{(n)}(\cdot \wedge \bar{\tau}^{(n)}), \delta\big)\Big] \leq c_0 \cdot \delta,$$

which gives (6.20) and proves the tightness of $\{\hat{Z}_f^{(n)}\}_{n\geq 1}$ in $\mathbb{D}$ by applying [15, Theorem 8.6 & Remark 8.7]. Next, similar to (6.17), we can obtain that for every $\varepsilon > 0$,

$$\mathbb{E}\Big[\sum_{t\leq T\wedge\bar{\tau}^{(n)}} \Big(\Delta\,\hat{Z}_f^{(n)}(t)\Big)^2 \mathbf{1}\big(|\Delta\,\hat{Z}_f^{(n)}|(t) > \varepsilon\big)\Big] \\ = \mathbb{E}\Big[\int_0^{T\wedge\bar{\tau}^{(n)}} \mathbb{E}\Big[\big(f(\eta, y) - \mathbb{E}[f(\eta, y)]\big)^2; \big|f(\eta, y) - \mathbb{E}[f(\eta, y)]\big| > \sqrt{n}\varepsilon\Big]\Big|_{y=\bar{Y}^{(n)}(s-)} d\bar{N}^{(n)}(s)\Big] \\ \leq \frac{2\,\lambda_0\,T}{1-\alpha} \cdot \sup_{|y|\leq k_0} \mathbb{E}\Big[\gamma + f^2(\eta, y); f(\eta, y) > \sqrt{n}\varepsilon\Big],$$

recalling the bound for $\mathbb{E}[f^2(\eta, y)]$ in (6.16), and which converges to 0 as $n \to \infty$ by the uniform integrability of $f^2(\eta, y)$ in Assumption A7-(4.8). This implies the convergence of the maximum jump sizes of $\hat{Z}_f^{(n)}$ to zero and hence proves the $\mathbb{C}$-tightness of $\{\hat{Z}_f^{(n)}\}_{n\geq 1}$. This finishes the proof. □

Next, we show that the CLT-scaled residual term $\hat{\varepsilon}_1^{(n)}$ and $\hat{\varepsilon}_2^{(n)}$ are negligible for large n, similar to Lemma 5.5. Notice that the presence of marks $\{\eta_j^{(n)}\}_{j\geq 1}$ makes the inequality from (4.4) much different from that from (3.4).

Lemma 6.2. *Under Assumptions A6 and A7, we have as $n \to \infty$,*

$$\big(\hat{\varepsilon}_1^{(n)}, \hat{\varepsilon}_2^{(n)}\big) \to 0 \quad \textit{u.o.c. in probability.}$$

Proof. Recalling (I, f) in Assumption A6-(4.4), we have for $t \leq \bar{\tau}^{(n)}$,

$$\hat{\varepsilon}_1^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j\geq 1} \Big(\int_{n(t-\bar{\tau}_j^{(n)})}^{\infty} \phi\big(u, \eta_j^{(n)}, \bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\,du\Big)\mathbf{1}\big(\bar{\tau}_j^{(n)} \leq t\big) \\ \leq \frac{1}{\sqrt{n}} \sum_{j\geq 1} I\big(n(t - \bar{\tau}_j^{(n)})\big)\,f\big(\eta_j^{(n)}, \bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big) =: \tilde{\varepsilon}_1^{(n)}(t) \tag{6.22}$$

where we understand that $I(u)=0$ for $u<0$. Recalling $\hat{X}^{(n)}$ in (5.1), $\hat{Z}_f^{(n)}$ in (6.2) and the bound for $\mathbb{E}\big[f(\eta,y)\big]$ in (6.5), we further have for $t\le\bar{\tau}^{(n)}$,

$$\begin{aligned}\tilde{\varepsilon}_1^{(n)}(t) &= \sqrt{n}\int_0^t I\big(n(t-s)\big)\left(d\bar{Z}_f^{(n)}(s)+\mathbb{E}\big[f(\eta,\bar{Y}^{(n)}(s-))\big]\,d\bar{N}^{(n)}(s)\right)\\ &\le \sqrt{n}\int_0^t I\big(n(t-s)\big)\left(d\bar{Z}_f^{(n)}(s)+\beta\,d\bar{N}^{(n)}(s)\right)\\ &= \int_0^t I\big(n(t-s)\big)\,d\hat{Z}_f^{(n)}(s)+\beta\int_0^t I\big(n(t-s)\big)\,d\hat{X}^{(n)}(s)\\ &\qquad+\beta\cdot\sqrt{n}\int_0^t I\big(n(t-s)\big)\big(\lambda_0\,ds+(\lambda^{(n)}(ns)-\lambda_0)\,ds\big)\\ &=: \tilde{\varepsilon}_2^{(n)}(t)+\beta\cdot\tilde{\varepsilon}_3^{(n)}(t)+\lambda_0\beta\cdot\tilde{\varepsilon}_4^{(n)}(t)+\beta\cdot\tilde{\varepsilon}_5^{(n)}(t).\end{aligned}\tag{6.23}$$

Notice the additional term $\tilde{\varepsilon}_2^{(n)}$ with respect to $\{\eta_j^{(n)}\}_{j\ge1}$ and the bound β instead of α comparing with (5.41). By Fubini's theorem, we have from the definition of $\lambda^{(n)}$ in (4.1) that

$$\begin{aligned}\tilde{\varepsilon}_5^{(n)}(t) &= \sqrt{n}\int_0^t I\big(n(t-s)\big)\Big(\sum_{j\ge1}\phi\big(n(s-\bar{\tau}_j^{(n)}),\eta_j^{(n)},\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\Big)\,ds\\ &= \frac{1}{\sqrt{n}}\sum_{j\ge1}\mathbf{1}(\bar{\tau}_j^{(n)}\le t)\int_0^{n(t-\bar{\tau}_j^{(n)})} I\big(n(t-\bar{\tau}_j^{(n)})-s\big)\phi\big(s,\eta_j^{(n)},\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big)\,ds\end{aligned}\tag{6.24}$$

where we understand that $\phi(u,z',y)=0$ for $u<0$. Moreover, similar to (5.43), we obtain for every $z'\in\mathcal{E}'$, $y\in\mathbb{R}$ and $t\ge0$,

$$\begin{aligned}\int_0^t I(t-s)\phi(s,z',y)\,ds &= \int_0^\infty d\nu_I(u)\int_{t-u}^t \phi\big(s,z',y\big)\,ds\\ &= \int_0^t d\nu_I(u)\int_{t-u}^\infty \phi\big(s,z',y\big)-\int_t^\infty \phi\big(s,z',y\big)\,ds+I(t)\cdot h\big(z',y\big)\end{aligned}$$

recalling that $I(0)=1$ and h in Assumption A6-(4.3). Substituting into (6.24) gives

$$\tilde{\varepsilon}_5^{(n)}(t)=\int_0^t \hat{\varepsilon}_1^{(n)}(t-u)d\nu_I(nu)-\hat{\varepsilon}_1^{(n)}(t)+\tilde{\varepsilon}_6^{(n)}(t),\tag{6.25}$$

recalling $\hat{\varepsilon}_1^{(n)}$ in (6.22), where

$$\tilde{\varepsilon}_6^{(n)}(t):=\frac{1}{\sqrt{n}}\sum_{j\ge1} I\big(n(t-\bar{\tau}_j^{(n)})\big)\,h\big(\eta_j^{(n)},\bar{Y}^{(n)}(\bar{\tau}_j^{(n)}-)\big).$$

Now, applying the same procedure of calculations in (6.23) to $\tilde{\varepsilon}_6^{(n)}$ above, with f in (6.22) replaced by h, we obtain

$$\begin{aligned}\tilde{\varepsilon}_6^{(n)}(t) &\le \int_0^t I\big(n(t-s)\big)\,d\hat{Z}_h^{(n)}(s)+\alpha\int_0^t I\big(n(t-s)\big)\,d\hat{X}^{(n)}(s)\\ &\qquad\alpha\cdot\sqrt{n}\int_0^t I\big(n(t-s)\big)\big(\lambda_0\,ds+(\lambda^{(n)}(ns)-\lambda_0)\,ds\big)\\ &=: \tilde{\varepsilon}_7^{(n)}(t)+\alpha\cdot\tilde{\varepsilon}_3^{(n)}(t)+\lambda_0\alpha\cdot\tilde{\varepsilon}_4^{(n)}(t)+\alpha\cdot\tilde{\varepsilon}_5^{(n)}(t).\end{aligned}\tag{6.26}$$

with the bound β replaced by α for $\mathbb{E}[h(\eta,y)]=H(y)$ in (6.5) and $\tilde{\varepsilon}_2^{(n)}$ replaced by $\tilde{\varepsilon}_7^{(n)}$.

In conclusion, we have showed in (6.23), (6.25) and (6.26) that

$$\tilde{\varepsilon}_1^{(n)}(t) \le \tilde{\varepsilon}_2^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_3^{(n)}(t) + \lambda_0\beta \cdot \tilde{\varepsilon}_4^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_5^{(n)}(t), \tag{6.27}$$

$$\tilde{\varepsilon}_5^{(n)}(t) = \int_0^t \hat{\varepsilon}_1^{(n)}(t-u)d\nu_I(nu) - \hat{\varepsilon}_1^{(n)}(t) + \tilde{\varepsilon}_6^{(n)}(t), \tag{6.28}$$

$$\tilde{\varepsilon}_6^{(n)}(t) \le \tilde{\varepsilon}_7^{(n)}(t) + \alpha \cdot \tilde{\varepsilon}_3^{(n)}(t) + \lambda_0\alpha \cdot \tilde{\varepsilon}_4^{(n)}(t) + \alpha \cdot \tilde{\varepsilon}_5^{(n)}(t), \tag{6.29}$$

where recalling in (6.23) and (6.26) we have defined

$$\begin{aligned} \tilde{\varepsilon}_2^{(n)}(t) &= \int_0^t I\big(n(t-s)\big)\, d\hat{Z}_f^{(n)}(s) \qquad \text{and} \qquad \tilde{\varepsilon}_3^{(n)}(t) = \int_0^t I\big(n(t-s)\big)\, d\hat{X}^{(n)}(s), \\ \tilde{\varepsilon}_4^{(n)}(t) &= \int_0^t I(ns)\, ds \qquad\qquad\qquad \text{and} \qquad \tilde{\varepsilon}_7^{(n)}(t) = \int_0^t I\big(n(t-s)\big)\, d\hat{Z}_h^{(n)}(s). \end{aligned} \tag{6.30}$$

Substituting (6.29) into (6.28) and canceling $\alpha \cdot \tilde{\varepsilon}_5^{(n)}$ on both sides gives

$$(1-\alpha)\cdot \tilde{\varepsilon}_5^{(n)}(t) \le \int_0^t \hat{\varepsilon}_1^{(n)}(t-u)d\nu_I(nu) - \hat{\varepsilon}_1^{(n)}(t) + \tilde{\varepsilon}_7^{(n)}(t) + \alpha \cdot \tilde{\varepsilon}_3^{(n)}(t) + \lambda_0\alpha \cdot \tilde{\varepsilon}_4^{(n)}(t)$$

Further substituting into (6.27) shows

$$\begin{aligned} (1-\alpha)\cdot \tilde{\varepsilon}_1^{(n)}(t) \le (1-\alpha)\cdot \tilde{\varepsilon}_2^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_3^{(n)}(t) + \lambda_0\beta \cdot \tilde{\varepsilon}_4^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_7^{(n)}(t) \\ + \beta \cdot \Big(\int_0^t \hat{\varepsilon}_1^{(n)}(t-u)d\nu_I(nu) - \hat{\varepsilon}_1^{(n)}(t) \Big). \end{aligned}$$

By the fact that $\hat{\varepsilon}_1^{(n)} \le \tilde{\varepsilon}_1^{(n)}$, we eventually obtain

$$\begin{aligned} \big(1+\beta-\alpha\big)\cdot \hat{\varepsilon}_1^{(n)}(t) \le (1-\alpha)\cdot \tilde{\varepsilon}_2^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_3^{(n)}(t) + \lambda_0\beta \cdot \tilde{\varepsilon}_4^{(n)}(t) + \beta \cdot \tilde{\varepsilon}_7^{(n)}(t) \\ + \beta \cdot \int_0^t \hat{\varepsilon}_1^{(n)}(t-u)d\nu_I(nu). \end{aligned}$$

Taking the supremum in t gives

$$\begin{aligned} \sup_{t\le T} \hat{\varepsilon}_1^{(n)}(t\wedge \bar{\tau}^{(n)}) \le \sup_{t\le T} \tilde{\varepsilon}_2^{(n)}(t\wedge \bar{\tau}^{(n)}) + \frac{\beta}{1-\alpha} \sup_{t\le T} \tilde{\varepsilon}_3^{(n)}(t\wedge \bar{\tau}^{(n)}) \\ + \frac{\lambda_0\beta}{1-\alpha} \sup_{t\le T} \tilde{\varepsilon}_3^{(n)}(t) + \frac{\beta}{1-\alpha} \sup_{t\le T} \tilde{\varepsilon}_7^{(n)}(t\wedge \bar{\tau}^{(n)}). \end{aligned} \tag{6.31}$$

Recalling the new $\tilde{\varepsilon}^{(n)}$'s in (6.30), it can be checked that

$$\sup_{t\le T} \tilde{\varepsilon}_3^{(n)}(t) = \sqrt{n}\int_0^T I(ns)\, ds \to 0$$

similar to the proof of $\tilde{\varepsilon}_{1,2}^{(n)}$ in (5.47). Applying Fubini's theorem, similar to (5.45), we obtain

$$\begin{aligned} \tilde{\varepsilon}_2^{(n)}(t) &= \int_0^\infty d\nu_I(nu)\big(\hat{Z}_f^{(n)}(t) - \hat{Z}_f^{(n)}(t-u)\big), \\ \tilde{\varepsilon}_3^{(n)}(t) &= \int_0^\infty d\nu_I(nu)\big(\hat{X}^{(n)}(t) - \hat{X}^{(n)}(t-u)\big), \\ \tilde{\varepsilon}_7^{(n)}(t) &= \int_0^\infty d\nu_I(nu)\big(\hat{Z}_h^{(n)}(t) - \hat{Z}_h^{(n)}(t-u)\big). \end{aligned}$$

By the $\mathbb{C}$-tightnesses of $\hat{X}^{(n)}$, $\hat{Z}_f^{(n)}$ and $\hat{Z}_h^{(n)}$ proved in Lemma 6.1, applying the same idea of proof for $\tilde{\varepsilon}_{1,1}^{(n)}$ in (5.46), we conclude

$$\tilde{\varepsilon}_2^{(n)} \to 0, \quad \tilde{\varepsilon}_3^{(n)} \to 0 \quad \text{and} \quad \tilde{\varepsilon}_7^{(n)} \to 0 \quad \text{u.o.c. in probability.}$$

Plugging into (6.31) proves the convergence of $\hat{\varepsilon}_1^{(n)}$ to zero.

The convergence of $\hat{\varepsilon}_2^{(n)}$ to zero is proved following the same procedure as in the proof of Lemma 5.5 and hence omitted. □

Completing the proof of Theorem 4.1. Given the results in Lemmas 6.1 and 6.2, the convergence of the CLT-scaled processes in Theorem 4.1 is proved following the same steps as the proof of Theorem 3.2. Here we give a stretch proof.

Recalling the representation (6.3), we have

$$\begin{aligned}
\hat{\Lambda}^{(n)}(t) &= \sqrt{n}\Big(\int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\bar{N}^{(n)}(s) - \int_0^t H\big(\bar{Y}(s)\big)\, d\bar{\Lambda}(s)\Big) + \hat{Z}_h^{(n)}(t) - \hat{\varepsilon}_1^{(n)}(t)\\
&= \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\hat{X}^{(n)}(s) + \sqrt{n}\int_0^t \Big(H\big(\bar{Y}^{(n)}(s-)\big) - H\big(\mu_\varphi\, \bar{\Lambda}^{(n)}(s)\big)\Big)\, d\bar{\Lambda}^{(n)}(s)\\
&\quad + \sqrt{n}\Big(\int_0^t H\big(\mu_\varphi\, \bar{\Lambda}^{(n)}(s)\big)\, d\bar{\Lambda}^{(n)}(s) - \int_0^t H\big(\mu_\varphi\, \bar{\Lambda}(s)\big)\, d\bar{\Lambda}(s)\Big) + \hat{Z}_h^{(n)}(t) - \hat{\varepsilon}_1^{(n)}(t)\\
&= \hat{\Lambda}_1^{(n)}(t) + \hat{\Lambda}_2^{(n)}(t) + \hat{\Lambda}_3^{(n)}(t) + \hat{Z}_h^{(n)}(t) - \hat{\varepsilon}_1^{(n)}(t).
\end{aligned}$$

By change of variable and the mean value theorem, we write

$$\hat{\Lambda}_3^{(n)}(t) = \sqrt{n}\int_{\bar{\Lambda}(t)}^{\bar{\Lambda}^{(n)}(t)} H(\mu_\varphi\, y)\, dy = \hat{\Lambda}^{(n)}(t)\cdot H\big(\mu_\varphi\, \tilde{\Lambda}^{(n)}(t)\big),$$

for some $\tilde{\Lambda}^{(n)}$ between $\bar{\Lambda}^{(n)}$ and $\bar{\Lambda}$, which gives, similar to (5.50),

$$\big(1 - H\big(\mu_\varphi\, \tilde{\Lambda}^{(n)}(t)\big)\big)\cdot \hat{\Lambda}^{(n)}(t) = \hat{\Lambda}_1^{(n)}(t) + \hat{\Lambda}_2^{(n)}(t) + \hat{Z}_h^{(n)}(t) - \hat{\varepsilon}_1^{(n)}(t). \tag{6.32}$$

For each component processes, applying [30, Theorem 2.2], by the convergence of $(\bar{Y}^{(n)}, \hat{X}^{(n)})$ from FLLN and Lemma 6.1, we have

$$\hat{\Lambda}_1^{(n)}(t) = \int_0^t H\big(\bar{Y}^{(n)}(s-)\big)\, d\hat{X}^{(n)}(s) \Rightarrow \int_0^t H\big(\bar{Y}(s)\big)\, d\hat{X}(s) = \hat{\Lambda}_1(t), \quad \text{in} \quad (\mathbb{D}, J_1).$$

For $\hat{\Lambda}_2^{(n)}$ in (6.32), we have from Lemma 6.1 first that

$$\sqrt{n}\big(\bar{Y}^{(n)}(t) - \mu_\varphi\, \bar{\Lambda}^{(n)}(t)\big) = \mu_\varphi\, \hat{X}^{(n)}(t) + \hat{Z}_\varphi^{(n)}(t) - \hat{\varepsilon}_2^{(n)}(t) \Rightarrow \mu_\varphi \hat{X}(t) + \hat{Z}_\varphi(t) \quad \text{in} \quad (\mathbb{D}, J_1).$$

Therefore, applying the mean value theorem gives

$$\hat{\Lambda}_2^{(n)}(t) = \int_0^t H'\big(\tilde{Y}^{(n)}(u)\big)\sqrt{n}\big(\bar{Y}^{(n)}(u) - \mu_\varphi\, \bar{\Lambda}^{(n)}(u)\big)\, d\bar{\Lambda}^{(n)}(u),$$

for some $\tilde{Y}^{(n)}$ between $\bar{Y}^{(n)}$ and $\mu_\varphi\, \bar{\Lambda}^{(n)}$. Further applying [30, Theorem 2.2], we obtain

$$\hat{\Lambda}_2^{(n)}(t) \Rightarrow \hat{\Lambda}_2(t) = \int_0^t H'\big(\bar{Y}(s)\big)\big(\mu_\varphi \hat{X}(s) + \hat{Z}(s)\big)\, d\bar{\Lambda}(s), \quad \text{in} \quad (\mathbb{D}, J_1).$$

We then obtain the joint convergence

$$\begin{aligned}
&\big(\tilde{\Lambda}^{(n)}, \hat{X}^{(n)}, \hat{Z}_h^{(n)}, \hat{Z}_\varphi^{(n)}, \sqrt{n}\big(\bar{Y}^{(n)}(t) - \mu_\varphi\, \bar{\Lambda}^{(n)}(t)\big), \hat{\Lambda}_1^{(n)}, \hat{\Lambda}_2^{(n)}\big)\\
&\Rightarrow \big(\bar{\Lambda}, \hat{X}, \hat{Z}_h, \hat{Z}_\varphi, \mu_\varphi \hat{X}(t) + \hat{Z}_\varphi(t), \hat{\Lambda}_1, \hat{\Lambda}_2\big) \quad \text{in} \quad (\mathbb{D}^7, J_1).
\end{aligned}$$

Notice that $\hat{Z}_h$ and $\hat{Z}_\varphi$ can be correlated. Substituting into (6.32) gives

$$\big(\hat{X}^{(n)}, \hat{\Lambda}^{(n)}, \hat{Y}^{(n)}\big) \Rightarrow \big(\hat{X}, \hat{\Lambda}, \hat{Y}\big) \quad \text{in} \quad (\mathbb{D}^3, J_1).$$

Finally, given the convergence results above, the SDE for $(\hat{N}, \hat{Y})$ in (4.9) can be derived as the derivation of (3.15). □

Acknowledgement

We thank the referee for the helpful comments which have led to the improvement of the presentation of the paper. G. Pang is partly funded by the NSF Grant DMS 2216765.

References

[1] S. Asmussen and H. Albrecher. *Ruin Probabilities*. World Scientific, 2010.

[2] E. Bacry, S. Delattre, M. Hoffmann, and J.-F. Muzy. Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*, 123(7):2475–2499, 2013.

[3] E. Bacry, I. Mastromatteo, and J.-F. Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.

[4] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[5] J. Blanchet and H. Lam. A heavy traffic approach to modeling large life insurance portfolios. *Insurance: Mathematics and Economics*, 53(1):237–251, 2013.

[6] O. Boxma and M. Mandjes. Shot-noise queueing models. *Queueing Systems*, 99(1-2):121–159, 2021.

[7] P. Brémaud and L. Massoulié. Power spectra of general shot noises and Hawkes point processes with a random excitation. *Advances in Applied Probability*, 34(1):205–222, 2002.

[8] P. Brémaud, G. Nappo, and G. Torrisi. Rate of convergence to equilibrium of marked Hawkes processes. *Journal of Applied Probability*, 39(1):123–136, 2002.

[9] Z. Cheng and Y. Seol. Diffusion approximation of a risk model with non-stationary Hawkes arrivals of claims. *Methodology and Computing in Applied Probability*, 22:555–571, 2020.

[10] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods*. Springer, New York, 2nd edition, Aug 2003.

[11] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes Volume II: General Theory and Structure*. Springer New York, 2008.

[12] A. Daw and J. Pender. Queues driven by Hawkes processes. *Stochastic Systems*, 8(3):192–229, 2018.

[13] R. Durrett. *Probability Theory and Examples*. Cambridge University Press, fifth edition, 2019.

[14] E. Errais, K. Giesecke, and L. R. Goldberg. Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665, 2010.

[15] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, Inc., mar 1986.

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

[17] E. Hashorva. Boundary Crossings of Brownian Motion. *Electronic Communications in Probability*, 10(none), Jan. 2005.

[18] A. G. Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198, 2018.

[19] U. Horst and W. Xu. Functional limit theorems for marked Hawkes point measures. *Stochastic Processes and their Applications*, 134:94–131, 2021.

[20] U. Horst and W. Xu. Functional limit theorems for Hawkes processes. *arXiv:2401.11495*, 2024.

[21] U. Horst, W. Xu, and R. Zhang. Convergence of heavy-tailed Hawkes processes and the microstructure of rough volatility. *arXiv:2312.08784*, 2023.

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

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

[24] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*. Springer Science & Business Media, 2013.

[25] T. Jaisson and M. Rosenbaum. Limit theorems for nearly unstable Hawkes processes. *The Annals of Applied Probability*, 25(2):600–631, 2015.

[26] T. Jaisson and M. Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *The Annals of Applied Probability*, 26(5):2860–2882, 2016.

[27] O. Kallenberg. *Foundations of Modern Probability*. Springer International Publishing, 2021.
[28] D. Karabash and L. Zhu. Limit theorems for marked Hawkes processes with application to a risk model. *Stochastic Models*, 31(3):433–451, 2015.
[29] C. Klüppelberg and T. Mikosch. Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli*, pages 125–147, 1995.
[30] T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*, 19(3):1035–1070, 1991.
[31] J. A. Lane. The central limit theorem for the Poisson shot-noise process. *Journal of Applied Probability*, 21(2):287–301, 1984.
[32] P. J. Laub, Y. Lee, and T. Taimre. *The elements of Hawkes processes*. Springer, 2021.
[33] B. Li and G. Pang. Functional limit theorems for nonstationary marked Hawkes processes in the high intensity regime. *Stochastic Processes and their Applications*, 143:285–339, 2022.
[34] B. Li and G. Pang. Heavy-traffic limits for parallel single-server queues with randomly split Hawkes arrival processes. *Journal of Applied Probability*, 61(2):490–514, 2024.
[35] C. Macci, G. Stabile, and G. Luca Torrisi. Lundberg parameters for non standard risk processes. *Scandinavian Actuarial Journal*, 2005(6):417–432, 2005.
[36] T. M. Niebauer, R. Schilling, K. Danzmann, A. Rüdiger, and W. Winkler. Nonstationary shot noise and its effect on the sensitivity of interferometers. *Phys. Rev. A*, 43:5022–5029, May 1991.
[37] A. Novikov, V. Frishling, and N. Kordzakhia. Approximations of boundary crossing probabilities for a Brownian motion. *Journal of Applied Probability*, 36(4):1019–1030, Dec. 1999.
[38] G. Pang and M. S. Taqqu. Nonstationary self-similar Gaussian processes as scaling limits of power-law shot noise processes and generalizations of fractional brownian motion. *High Frequency*, 2(2):95–112, 2019.
[39] G. Pang and Y. Zhou. Functional limit theorems for a new class of non-stationary shot noise processes. *Stochastic Processes and their Applications*, 128(2):505–544, 2018.
[40] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer Berlin Heidelberg, 2005.
[41] J. Rice. On generalized shot noise. *Advances in Applied Probability*, 9(3):553–565, 1977.
[42] G. Samorodnitsky. A class of shot noise models for financial applications. In *Athens Conference on Applied Probability and Time Series Analysis: Volume I: Applied Probability In Honor of JM Gani*, pages 332–353. Springer, 1996.
[43] T. Schmidt. Shot-noise processes in finance. *From Statistics to Mathematical Finance: Festschrift in Honour of Winfried Stute*, pages 367–385, 2017.
[44] T. Schmidt and W. Stute. Shot-noise processes and the minimal martingale measure. *Statistics & Probability Letters*, 77(12):1332–1338, 2007.
[45] L. Wang and K. Pötzelberger. Boundary crossing probability for Brownian motion and general boundaries. *Journal of Applied Probability*, 34(1):54–65, Mar. 1997.
[46] W. Xu. Diffusion approximations for self-excited systems with applications to general branching processes. *The Annals of Applied Probability*, 34(3):2650–2713, 2024.