

# Self-Exciting Random Evolutions (SEREs) and Their Applications (Version 2)

Anatoliy Swishchuk<sup>1</sup>

Department Mathematics & Statistics

University of Calgary

Calgary, Canada T2N 1N4

aswish@ucalgary.ca

## Abstract

This paper is devoted to the study of a new class of random evolutions (RE), so-called self-exciting random evolutions (SEREs), and their applications. We also introduce a new random process  $x(t)$  such that it is based on a superposition of a Markov chain  $x_n$  and a Hawkes process  $N(t)$ , i.e.,  $x(t) := x_{N(t)}$ . We call this process self-walking imbedded semi-Hawkes process (Swish Process or SwishP). Then the self-exciting REs (SEREs) can be constructed in similar way as, e.g., semi-Markov REs, but instead of semi-Markov process  $x(t)$  we have SwishP. We give classifications and examples of self-exciting REs (SEREs). Then we consider two limit theorems for SEREs such as averaging (Theorem 1) and diffusion approximation (Theorem 2). Applications of SEREs are devoted to the so-called self-exciting traffic/transport process and self-exciting summation on a Markov chain, which are examples of continuous and discrete SERE. From these processes we can construct many other self-exciting processes, e.g., such as impulse traffic/transport process, self-exciting risk process, general compound Hawkes process for a stock price, etc. We present averaged and diffusion approximation of self-exciting processes. The novelty of the paper associated with new models, such as  $x(t)$  and SERE, and also new features of SEREs and their many applications, namely, self-exciting and clustering effects.

**Keywords:** self-exciting random evolutions (SEREs); self-walking imbedded semi-Hawkes process (SwishP); averaging; diffusion approximation; self-exciting traffic/transport process; self-exciting risk process; self-exciting summation on a Markov chain.

## 1 Introduction

The purpose of the present paper is to introduce and to study a new class of Random Evolutions (RE), such as self-exciting RE (SERE), which are

---

<sup>1</sup>The author thanks to two unanimous referees for their valuable comments and remarks with respect to the first version of this working paper (Available at SSRN: <https://ssrn.com/abstract=5055075>). The work on this working paper is still in progress.

operator-valued stochastic processes that possess two traits, self-exciting and clustering ones. These features make SERE completely different from other RE studied in the literature [8, 17, 14, 15, 23]. The results for SERE obtained in the paper (averaging (Theorem 1) and diffusion approximation (Theorem 2)) generalize many results in the literature devoted to the Hawkes-based models in finance and insurance, such as, e.g., risk model based on general compound Hawkes process [24] (self-exciting impulse traffic/transport process (8) in this paper) and general compound Hawkes processes in limit order books, see [25] (self-exciting summation on a Markov chain in this paper). We note that SERE is based on a new process  $x(t)$ , introduced in the paper, which is constructed as a superposition of a Markov chain  $x_n$  and a Hawkes process  $N(t)$ , i.e.,  $x(t) := x_{N(t)}$ . We call this process self-walking imbedded semi-Hawkes process (Swish Process or SwishP).

The Random Evolutions (REs) can be described by two objects: 1) operator dynamical system  $V(t)$  and 2) random process  $x(t)$ . Depending on structure of  $V(t)$  and properties of the stochastic process  $x_t$  we have different kind of REs: continuous, discrete, Markov, semi-Markov, etc.

From mathematical point of view, a RE is a solution of stochastic operator integral equation in a Banach space. The operator coefficients of such equations depend on random parameters. The RE, from physical point of view, is a model for a dynamical system whose state of evolution is subject to random variations. Such systems arise in many branches of science, e.g., random Hamiltonian and Shrödinger's equations with random potential in quantum mechanics, Maxwell's equation with a random reflective index in electrodynamics, transport equation, storage equation, etc. There is a lot of applications of REs in financial and insurance mathematics [23]. One of the recent applications of RE is associated with geometric Markov renewal processes which are a regime-switching models for a stock price in financial mathematics [23]. Another recent applications of RE is a semi-Markov risk process in insurance mathematics [23], and regime-switching financial models, including such models with jumps [23]. The REs are also examples of more general mathematical objects such as multiplicative operator functionals (MOFs) [17], which are random dynamical systems in Banach space. Different REs were introduced and studied by Hersh R. [8]. Semi-Markov REs were considered in [14, 15]. Inhomogeneous random evolutions, limit theorems for them and their applications were considered in [26]. Some results on random operators were introduced in many papers and books, for example, random linear operators were studied in [20], and integration theory in Hilbert spaces were investigated in [21].

This paper is devoted to the study of a new so-called self-exciting REs and their applications. We introduce a new random process  $x(t)$  such that it is based on a superposition of a Markov chain  $x_n$  and Hawkes process  $N(t)$ , i.e.,  $x(t) := x_{N(t)}$ . We call this process self-walking imbedded semi-Hawkes process (Swish Process or SwishP). Then the self-exciting REs can be constructed in similar way as, e.g., semi-Markov REs, but instead of semi-Markov process

$x(t)$  we have SwishP. We give first classifications and examples of self-exciting REs (SEREs). Then we consider limit theorems for SEREs such as averaging and diffusion approximation. Applications of SEREs are devoted to the so-called self-exciting traffic/transport process and self-exciting summation on a Markov chain, which are examples of continuous and discrete SERE. From these processes we can construct many other self-exciting processes, e.g., such as impulse traffic/transport process, self-exciting risk process, etc. We present averaged and diffusion approximation for self-exciting processes.

We note that Hawkes process was introduced in [9, 10] and was studied in many papers and books, and was applied to many problems in finance and insurance [12]. For example, risk process based on general compound Hawkes process was studied in [24], general compound Hawkes process applied to LOBs was intensively studied in [25]. Modelling financial contagion using mutually exciting jump processes were considered in [2]. Multivariate Hawkes processes with application to financial data were studied in [6]. Stability of nonlinear Hawkes processes were first considered in [4]. Some limit theorems for Hawkes processes and their application to financial statistics were investigated in [3]. Applications of limit theorems for marked Hawkes processes to a risk model was considered in [13].

The paper is organized as follows. Section 2 gives main definitions, conditions and classifications for SERE, including definition of a new Swish process. Section 3 presents some example of SERE, such as self-exciting impulse traffic/transport process, self-exciting geometric compound Swish process, self-exciting risk process, to name a few.. Section 4 describes martingale properties and weak convergence of SERE. Two main limit theorems for SERE, such as averaging (Theorem 1) and diffusion approximation (Theorem 2), are presented in section 5. Applications of limit theorems for SEREs are considered in section 6. These applications are devoted to the self-exciting traffic/transport process and self-exciting summation on a Markov chain, which are continuous and discrete SERE. Section 7, Appendix, contains proofs for Theorem 1 and Theorem 2. Discussions are considered in section 8 and section 9 concludes the paper.

## 2 Main Definitions and Conditions for Swish Process and SERE

### 2.1 Self-Walking Imbedded Semi-Hawkes Process (SwishP)

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathcal{P})$  be a probability space,  $t \in R_+ := [0, +\infty]$ , and  $(X, \mathcal{X})$  be a measurable phase space.

**Definition 1 (Hawkes Process [9, 10]).** The Hawkes process  $N(t)$  is a counting process with stochastic intensity function  $\lambda(t)$  :

$$\lambda(t) = \lambda + \int_0^t \mu(s) dN(s),$$

where  $\lambda > 0$  is a background intensity and  $\mu(t)$  is a self-exciting deterministic function. We assume that  $\mu(t) = \alpha e^{-\beta t}$ ,  $\alpha, \beta > 0$ , is an exponential self-exciting function. We suppose that  $\int_0^{+\infty} \mu(s) ds := \hat{\mu} < 1$ - stability condition to avoid accumulating of jumps. Here,  $\hat{\mu} = \alpha/\beta$ .

Moments of jumps (or arrival times of events) of the Hawkes process are  $\tau_n$ , and inter-arrival times are  $\theta_n \in R_+$ , therefore  $\tau_n := \sum_{k=0}^n \theta_k$ . Let we also have homogeneous Markov chain  $x_n \in X$ , where  $X$  is a space of states (or phase space) for  $x_n$ .

**Definition 2 (SwishP).** Swish Process (SwishP) is defined in the following way:

$$x(t) := x_{N(t)}, \quad (1)$$

where  $x_n$  is a Markov chain and  $N(t)$  is a Hawkes process, and they are independent.

As long as we suppose that  $x_n$  and  $\theta_n, n \geq 0$ , are independent, then

$$Q(x, A, t) := \mathcal{P}\{x_n \in A, \theta_n \leq t | x_0 = x\} = P(x, A)G(t), \quad (2)$$

where

$$\begin{aligned} P(x, A) &:= \mathcal{P}(x_{n+1} \in A | x_n = x), A \in \mathcal{X}, \\ G(t) &:= \mathcal{P}(\theta_{n+1} \leq t). \end{aligned} \quad (3)$$

## 2.2 Definition and Classification of SERE

Let  $(B, \mathcal{B}, \|\cdot\|)$  be a separable Banach space [1, 5, 26].

Let  $\{\Gamma(x); x \in X\}$  be a family of operators on the dense subspace  $B_0 \in B$ , which is common domain for  $\Gamma(x)$ , independent of  $x$ , non-commuting and unbounded in general, such that map  $\Gamma(x)f : X \rightarrow B$  is strongly  $\mathcal{X}/\mathcal{B}$ -measurable for all  $f \in B_0$ ,  $\forall t \in R_+$ . Also, let  $\{\mathcal{D}(x); x \in X\}$  is a family of bounded linear operators on  $B$ , such that map  $\mathcal{D}(x)f : X \rightarrow B$  is  $\mathcal{X}/\mathcal{B}$ -measurable,  $\forall f \in B$ .

**Definition 3 (Self-Exciting Random Evolution (SERE)).** SERE are defined by the solution of the following stochastic operator integral equation in separable Banach space  $B$ :

$$V(t)f = f + \int_0^t V(s)\Gamma(x(s))f ds + \sum_{k=1}^{N(t)} V(\tau_k)[\mathcal{D}(x_k) - I]f, \quad (4)$$

where  $N(t)$  is a Hawkes process,  $x(t)$  is a Swish process,  $x_k$  is a Markov chain,  $I$  is an identity operator on  $B$ ,  $f \in B_0$ . Because of the stability condition  $\int_0^{+\infty} \mu(s) ds := \hat{\mu} < 1$ , the sum in (4) is finite.

If  $x(t)$  in (1) is a Markov or semi-Markov process, then SERE in (4) is called a Markov or semi-Markov RE, respectively. In this case,  $N(t)$  is a renewal process, [23]. If  $\mathcal{D}(x) \equiv I$ ,  $\forall x \in X$ , then  $V(t)$  in (4) is called a continuous SERE. If  $\Gamma(x) \equiv 0$ ,  $\forall x \in X$ , is a zero operator on  $B$ , then  $V(t)$  in (4) is called a jump RE. SERE  $V_n := V(\tau_n)$  is called a discrete SERE.

Operators  $\Gamma(x)$ ,  $x \in X$ , describe a continuous component  $V^c(t)$  of SERE  $(V(t))_{t \geq 0}$  in (4), and operators  $\mathcal{D}(x)$  describe a jump component  $V^d(t)$  of SERE  $(V^d(t))_{t \geq 0}$  in (4).

In such a way, SERE is described by two objects: 1) operator dynamical system  $(V(t))_{t \geq 0}$ ; 2) random process  $x(t)$ .

We note, that it turned out to be (see [14] for more details)

$$V(t) = \Gamma_{x(t)}(t - \tau_{N(t)}) \prod_{k=1}^{N(t)} \mathcal{D}(x_k) \Gamma_{x_{k-1}}(\theta_k), \quad (5)$$

where  $\Gamma_x(t)$  are the semigroups of operators of  $t$  generated by the operators  $\Gamma(x)$ ,  $\forall x \in X$ . Because of the stability condition  $\int_0^{+\infty} \mu(s) ds := \hat{\mu} < 1$ , the product in (5) is finite.

We also note, that SERE in (5) is called a discontinuous SERE. Under conditions introduced in the next section, the solution  $V(t)$  of the equation (4) is unique and can be represented by product (5), that can be proved by constructive or iterative methods, similarly to the approach which was done in [14, 15] for semi-Markov REs.

## 2.3 Conditions for SERE

We suppose that the following conditions are satisfied (which will be used in Theorem 1 (Averaging of SERE) and Theorem 2 (DA of SERE)):

A) there exists Hilbert spaces  $H$  and  $H^*$  such that compactly imbedded in Banach spaces  $B$  and  $B^*$  respectively,  $H \subset B$ ,  $H^* \subset B^*$ , where  $B^*$  is a dual space to  $B$ , that divides points of  $B$ ;

B) operators  $\Gamma(x)$  and  $(\Gamma(x))^*$  are dissipative on any Hilbert space  $H$  and  $H^*$  respectively;

C) operators  $\mathcal{D}(x)$  and  $\mathcal{D}^*(x)$  are contractive on any Hilbert space  $H$  and  $H^*$  respectively;

D)  $(x_n; \quad n \geq 0)$  is a uniformly ergodic Markov chain with stationary distribution  $\rho(A)$ ,  $A \in \mathcal{X}$ ;

E)  $m_i := \int_0^\infty t^i G(dt)$  are uniformly integrable,  $\forall i = 1, 2, 3$ , where

$$G(t) := \mathcal{P}\{\omega : \theta_{n+1} \leq t\}; \quad (6)$$

F)

$$\begin{aligned} \int_X \rho(dx) \|\Gamma(x)f\|^k &< +\infty; \quad \int_X \rho(dx) \|P\mathcal{D}_j(x)f\|^k < +\infty; \quad j = 1, 2, \\ \int_X \rho(dx) \|\Gamma(x)f\|^{k-1} \cdot \|P\mathcal{D}_j(x)f\|^{k-1} &< +\infty; \quad \forall k = 1, 2, 3, 4, f \in B_0, \end{aligned} \quad (7)$$

where  $P$  is on operator generated by the transition probabilities  $P(x, A)$  of Markov chain  $(x_n; \quad n \geq 0)$ :

$$P(x, A) := \mathcal{P}\{\omega : x_{n+1} \in A/x_n = x\}, \quad (8)$$

and  $\mathcal{D}_j(x); \quad x \in X, \quad j = 1, 2\}$  is a family of some closed operators;

G)  $\int_0^{+\infty} \mu(s)ds < 1$ , where  $\mu(t)$  is a self-exciting exponential function for Hawkes process  $N(t)$ ;

H)  $\int_0^{+\infty} s\mu(s)ds < +\infty$ .

If  $B := C_0(R)$ , then  $H := W^{l,2}(R)$  is a Sobolev space [1, 5, 22], and  $W^{l,2}(R) \subset C_0(R)$  and this imbedding is compact. For the spaces  $B := L_2(R)$  and  $H := W^{l,2}(R)$  we have the same situation.

Let us explain why we need all the conditions A)-H). It follows from the conditions A) - B) that operators  $\Gamma(x)$  and  $(\Gamma(x))^*$  generate a strongly continuous contractive semigroup of operators  $\Gamma_x(t)$  and  $\Gamma_x^*(t)$ ,  $\forall x \in X$ , in  $H$  and  $H^*$  respectively. From the conditions A)-C) it follows that SERE  $V(t)$  in (1) is a contractive operator in  $H$ ,  $\forall t \in R_+$ , and  $\|V(t)f\|_H$  is a semimartingale  $\forall f \in H$ . In such a way, the conditions A) - C) supply the following result:

**Proposition (Tightness of SERE).** SERE  $V(t)f$  is a tight process in  $B$ , namely,  $\forall \Delta > 0$  there exists a compact set  $K_\Delta$ :

$$\mathcal{P}\{V(t)f \in K_\Delta; 0 \leq t \leq T\} \geq 1 - \Delta. \quad (9)$$

**Proof.** This result follows from Kolmogorov - Doob inequality for semimartingale  $\|V(t)f\|_H$  [14].

Condition (9) is one of the main steps in the proving of limit theorems and rates of convergence for the sequence of SERE in series scheme. Conditions D)-E) need for asymptotical results for SERE associated with Markov chain  $x_n$  and r.v.  $\theta_n$ . We require technical condition D) for the averaged by  $\rho(dx)$  moments for the operators  $\Gamma(x)$  and  $\mathcal{D}_j(x), j = 1, 2$ , in the asymptotic expansions of operators  $\Gamma_x(t)$  (see (26)) and  $\mathcal{D}^\epsilon(x)$  (see (28) and (32)). Technical conditions G) and H), designed for the Hawkes process  $N(t)$ , needed for the asymptotic analysis of HP in the averaging (Theorem 1) and DA (Theorem 2) results for SERE.

### 3 Some Examples of SERE

Many examples for SERE can be obtained by assigning to every operator  $\Gamma(x)$  and  $\mathcal{D}(x)$  a specific map. For example, if we take

$$\Gamma(x) := v(x) \frac{d}{dz}, \quad \mathcal{D}(x) \equiv I, \quad B = C^1(R),$$

where  $C^1(R)$  is a space of differentiable functions on  $R$ , then we get a transport equation for (4) describing a motion of a particle with random velocity

$v(x(t))$ , i.e.,

$$z(t) = z + \int_0^t v(z(s), x(s)) ds.$$

Equation (4) in this case is

$$f(z(t)) = f(z) + \int_0^t v(z(s), x(s)) \frac{d}{dz} f(z(s)) ds.$$

Below we present many other examples of SERE.

**Example 1. Self-exciting Impulse Traffic/Transport process.** Let  $B = C(R)$  be a space of continuous functions on  $R$ , and operators  $\Gamma(x)$  and  $\mathcal{D}(x)$  are defined by the following way:

$$\Gamma(x)f(z) := v(z, x) \frac{d}{dz} f(z), \quad \mathcal{D}(x)f(z) := f(z + a(x)), \quad (10)$$

where functions  $v(z, x)$  and  $a(x)$  are continuous and bounded on  $R \times X$  and  $X$  respectively,  $\forall z \in R, \forall x \in X, f(z) \in C^1(R) := B_0$ . Then the equation (4) takes the form:

$$f(z(t)) = f(z) + \int_0^t v(z(s), x(s)) \frac{d}{dz} f(z(s)) ds + \sum_{k=1}^{N(t)} [f(z(\tau_k) + a(x_k)) - f(z(\tau_k))], \quad (11)$$

and SERE  $V(t)$  is defined by the relation:

$$V(t)f(z) = f(z(t)), \quad z(0) = z.$$

Equation (11) is a functional form for self-exciting impulse traffic process  $z_t$ , which satisfies the equation:

$$z(t) = z + \int_0^t v(z(s), x(s)) ds + \sum_{k=1}^{N(t)} a(x_k). \quad (12)$$

We note that self-exciting impulse traffic/transport process  $z(t)$  in (12) is a realization of discontinuous RE. We call  $z(t)$  a self-exciting traffic/transport process whenever  $a(x) = 0$ . We note, that if we take  $v(z, x) = v$  and  $a(x)$  takes negative values, then we get risk model based on general compound Hawkes process [24]. We note that sum in (12) cannot be exploded because of the stationary condition  $\hat{\mu} < 1$  (see condition G)).

**Example 2. Self-Exciting Summation on a Markov chain.** Put  $v(z, x) \equiv 0, \forall z \in R, \forall x \in X$ , in (12). Then the process

$$z(t) = z + \sum_{k=1}^{N(t)} a(x_k) \quad (13)$$

is a summation on a Markov chain  $(x_n; n \geq 0)$  and it is a realization of a jump SERE. We also called it general compound Hawkes process [25].

**Example 3. Self-Exciting Diffusion Process.** Let  $B = C(R)$ ,  $B_0 = C^2(R)$ -twice differentiable functions on  $R$ ,  $P_x(t, z, A)$  be a distribution function or transition probability kernel, which respects to the diffusion process  $\xi(t)$ , that is the solution of the stochastic differential equation in  $R$  with SwishP switchings:

$$d\xi(t) = a(\xi(t), x(t))dt + b(\xi(t), x(t))dw_t, \xi(0) = z, \quad (14)$$

where  $x(t)$  is a Swish process independent of a standard Wiener process  $w_t$ , coefficients  $a(z, x)$  and  $b(z, x)$  are bounded and continuous functions on  $R \times X$ . Let us define the following contraction semigroups of operators on  $B$ :

$$\Gamma_x(t)f(z) := \int_R P_x(t, z, dy)f(y), f(y) \in B, x \in X. \quad (15)$$

Their infinitesimal operators  $\Gamma(x)$  have the following form:

$$\Gamma(x)f(z) = a(z, x)\frac{d}{dz}f(z) + 2^{-1}b^2(z, x)\frac{d^2}{dz^2}f(z), f(z) \in B_0.$$

Here, the operators  $\mathcal{D}(x) \equiv I, \forall x \in X$ , are identity operators. Then the equation (4) takes the form:

$$f(\xi(t)) = f(z) + \int_0^t [a(\xi(s), x(s))\frac{d}{dz} + 2^{-1}b^2(\xi(s), x(s))\frac{d^2}{dz^2}]f(\xi(s))ds, \quad (16)$$

and SERE  $V(t)$  is defined by the relation

$$V(t)f(z) = E[f(\xi(t))/x(s); 0 \leq s \leq t; \xi(0) = z].$$

Equation (16) is a functional form for diffusion process  $\xi(t)$  in (14) in SwishP random media  $x(t)$ . We note that diffusion process  $\xi(t)$  in (14) is a realization of continuous SERE.

**Example 4. Self-Exciting Geometric Compound Process (SEGCP).**

Let  $x(t)$  be a Swish process on the phase space  $X \times \mathbf{R}^+$ , i.e.,  $x(t) := x_{Nt}$ . Let  $c(x)$  be a bounded continuous function on  $X$  such that  $c(x) > -1$ . We define a SEGCP  $S_t$  with SwishP  $x(t)$  as follows:

$$S_t := S_0 \prod_{k=0}^{N(t)} (1 + c(x_k)), \quad (17)$$

where  $S_0 > 0$  is the initial value of  $S_t$ . We call the process Self-Exciting Geometric Compound Process (SEGCP) which is similar to the geometric compound Poisson process

$$S_t = S_0 \prod_{k=1}^{N(t)} (1 + Y_k),$$



where  $S_0 > 0$ ,  $N(t)$  is a standard Poisson process,  $(Y_k)_{k \in \mathbb{Z}_+}$  are i. i. d. random variable, which is a trading model in many financial applications as a pure jump model (see [23]). We note that process  $S_t$  cannot be exploded because of the stationary condition  $\hat{\mu} < 1$ .

Let  $B := C_0(\mathbf{R}_+)$  be a space of continuous functions on  $\mathbf{R}_+$ , vanishing at infinity, and let us define a family of bounded contracting operators  $D(x)$  on  $C_0(\mathbf{R}_+)$  :

$$D(x)f(s) := f(s(1 + c(x))), \quad x \in X, s \in \mathbf{R}_+. \quad (18)$$

With these contraction operators  $D(x)$  we define the following jump SERE  $V(t)$  of SEGCP in (17)

$$V(t) = \prod_{k=0}^{N(t)} D(x_k) := D(x_{N(t)}) \circ D(x_{N(t)-1}) \circ \dots \circ D(x_1) \circ D(x_0). \quad (19)$$

Using (18) we obtain from (19)

$$V(t)f(s) = \prod_{k=0}^{N(t)} D(x_k)f(s) = f(s \prod_{k=0}^{N(t)} (1 + c(x_k))) = f(S_t), \quad (20)$$

where  $S_t$  is defined in (17) and  $S_0 = s$ .

**Example 5. Self-Exciting Risk Process (SERP).** SERSP is defined as

$$R(t) = u + ct - \sum_{k=1}^{N(t)} a(x_k), \quad (21)$$

where  $u$  is an initial capital of an insurance company,  $c$  is a premium,  $a(x)$  is a measurable bounded function on  $X$ , and  $N(t)$  is a hakes process. Here,

$$\Gamma_x(t)f(u) = \Gamma(t)f(u) := f(u + ct), \quad \mathcal{D}(x)f(u) := f(u - a(x)).$$

## 4 Martingale Property and Weak Convergence of SERE

We use martingale methods to study SERE and limit theorems for them. The main idea is that process

$$M_n f := V_n f - f - \sum_{k=0}^{n-1} E[V_{k+1} f - V_k f / \mathcal{F}_k], \quad V_0 = I, f \in B, \quad (22)$$

is an  $\mathcal{F}_n$ -martingale in  $B$ , where  $\sigma$ -algebra

$$\mathcal{F}_n := \sigma\{x_k, \tau_k; 0 \leq k \leq n\}, \quad V_n := V(\tau_n),$$

$E$  is an expectation by probability  $\mathcal{P}$ . It is defined as a Bochner integral in Banach space  $B$  (see [26]). For definitions of expectations of operator-valued processes and martingale meaning in this context, see [26], Chapter 1, sec.1.1 Representation of the martingale  $M_n$  (see (22)) in the form of martingale-difference

$$M_n f = \sum_{k=0}^{n-1} [V_{k+1} f - E(V_{k+1} f / \mathcal{F}_k)] \quad (23)$$

gives us the possibility to calculate the weak quadratic variation:

$$< l(M_n f) > := \sum_{k=0}^{n-1} E[l^2((V_{k+1} - V_k)f) / \mathcal{F}_k], \quad (24)$$

where  $l \in B^*$  is a linear functional from  $B^*$ , and  $B^*$  is a dual space to  $B$ , dividing points of  $B$ .

In general, the weak convergence of SERE in series scheme is obtained from the criterion of weak compactness of the processes with values in separable Banach space. The limiting SERE we obtain from the solution of a martingale problem in the form of some integral operator equations in Banach space  $B$ . In particular, the martingale method of obtaining of the limit theorems for the sequence of SERE is based on the solution of the following problems (see [26], Chapter 1, sec. 1.2): 1) weak compactness of the family of measures generated by the sequences of SERE (see [26], Chapter 1, Theorem 1.9, and **Proposition (Tightness of SERE)**, sec. 2.3); 2) any limiting point of this family of measures is the solution of a martingale problem; 3) the solution of martingale problem is unique. The conditions 1)-2) guarantee the existence of weakly convergent subsequence, and condition 3) gives the uniqueness of the weak limit. It follows from 1)-3) that consequence of SERE converges weakly to the unique solution of a martingale problem.

We also use the representation

$$V_{k+1} f - V_k f = [\Gamma_{x_k}(\theta_{k+1}) \mathcal{D}(x_k) - I] V_k f, \quad V_k := V(\tau_k), \quad (25)$$

and the following expression for semigroups of operators  $\Gamma_x(t)$  (see [5], [26]):

$$\Gamma_x(t) f = f + \sum_{k=1}^{n-1} \frac{t^k}{k!} \Gamma^k(x) f + (n-1)^{-1} \int_0^t (t-s)^n \Gamma_x(s) \Gamma^n(x) f ds, \quad \forall x \in X, \quad (26)$$

$\forall f \in \cap_{x \in X} \text{Dom}(\Gamma^n(x))$ . Taking into account (22)-(26) and conditions A)-H) from sec. 2, we obtain the limit theorems for SERE in the next sections.

## 5 Main Results: Averaging and Diffusion Approximation of SERE

### 5.1 Averaging of Self-Exciting Random Evolutions

Let's consider a SERE in series scheme:

$$V_\varepsilon(t) = f + \int_0^t \Gamma(x(s/\varepsilon))V_\varepsilon(s)f ds + \sum_{k=1}^{N(t/\varepsilon)} [\mathcal{D}^\varepsilon(x_k) - I]V_\varepsilon(\varepsilon\tau_k)f, \quad (27)$$

where

$$\mathcal{D}^\varepsilon(x) = I + \varepsilon D_1(x) + o(\varepsilon), \quad (28)$$

$\{D_1(x); x, y \in X\}$  is a family of closed linear operators,  $\|0(\varepsilon)f\|/\varepsilon \rightarrow 0$  as  $\varepsilon \rightarrow 0$ ,  $\varepsilon$  is a small parameter,

$$f \in B_0 := \bigcap_{x,y \in X} \text{Dom}(\Gamma^2(x)) \cap \text{Dom}(D_1^2(x, y)). \quad (29)$$

Another form for  $V_\varepsilon(t)$  in (29) is (see also (5)):

$$V_\varepsilon(t) = \Gamma_{x(t/\varepsilon)}(t - \varepsilon\tau_{N(t/\varepsilon)}) \prod_{k=1}^{N(t/\varepsilon)} \mathcal{D}^\varepsilon(x_k)\Gamma_{k-1}(\varepsilon\theta_k). \quad (30)$$

**Theorem 1 (Averaging of SERE).** Let the conditions A)-D), E) with  $i = 2$ , F) with  $k = 2, j = 1$ , and G) are satisfied (see sec. 2.3). Then SERE  $V_\varepsilon(t)$  in (27) or in (30) converges weakly as  $\varepsilon \rightarrow 0$  to the following limiting process  $V_0(t)$  satisfying the equation

$$V_0(t)f - f - \hat{\lambda} \int_0^t (\hat{\Gamma} + \hat{D})V_0(s)f ds = 0, \quad (31)$$

where

$$\begin{aligned} \hat{\Gamma} &:= m \int_x \rho(dx) \Gamma(x), \\ \hat{D} &:= \int_x \rho(dx) D_1(x), \end{aligned}$$

$m := m_1$ ,  $\hat{\lambda}$  is defined by

$$\hat{\lambda} := \frac{\lambda}{1 - \hat{\mu}},$$

where  $\lambda$  and  $\hat{\mu}$  are defined in sec. 2.1. Proof of Theorem 1 is given in the Appendix.

## 5.2 Diffusion Approximation (Functional Central Limit Theorem) of Self-Exciting Random Evolutions

Let us consider SERE  $V^\varepsilon(t)$ , which is defined as as

$$V^\varepsilon(t) = f + \int_0^t \Gamma(x(s/\varepsilon^2))V^\varepsilon(s)f ds + \sum_{k=1}^{N(t/\varepsilon^2)} [\mathcal{D}^\varepsilon(x_k) - I]V^\varepsilon(\varepsilon\tau_k)f,$$

$$\mathcal{D}^\varepsilon(x) := I + \varepsilon D_1(x) + \varepsilon^2 D_2(x) + o(\varepsilon^2), \quad (32)$$

$\{D_i(x); \quad x \in X, \quad i = 1, 2\}$  are closed linear operators and  $\|0(\varepsilon^2)f\|/\varepsilon^2 \rightarrow 0, \varepsilon \rightarrow 0$

$$\begin{aligned} \forall f \in B_0 &:= \bigcap_{x \in X} \text{Dom}(\Gamma^4(x)) \bigcap \text{Dom}(D_2(x)), \\ \text{Dom}(D_2(x)) &\subseteq \text{Dom}(D_1(x)); \quad D_1(x) \subseteq \text{Dom}(D_1(x)), \\ \forall x \in X, \quad \Gamma^i(x) &\subset \text{Dom}(D_2(x)), \quad i = \overline{1, 3}. \end{aligned} \quad (33)$$

In such a way,  $V^\varepsilon(t)$  can be also presented in the following way (see also (5)):

$$V^\varepsilon(t) = \Gamma_{x(t/\varepsilon^2)}(t/\varepsilon - \varepsilon \tau_{N(t/\varepsilon^2)}) \prod_{k=1}^{N(t/\varepsilon^2)} \mathcal{D}^\varepsilon(x_k) \Gamma_{x_{k-1}}(\varepsilon \theta_k), \quad (34)$$

where  $\mathcal{D}^\varepsilon(x)$  are defined in (32).

**Theorem 2 (Diffusion Approximation of SERE).** Let the conditions A)-D), E) with  $i = 3$ , F) with  $k = 3, j = 2$ , and G)-H) are satisfied (see sec. 2.3). Also, let the **balance condition** be satisfied:

$$\int_X \rho(dx) [m\Gamma(x) + D_1(x)]f = 0, \quad \forall f \in B_0 \quad (35)$$

Then the SERE  $V^\varepsilon(t)$  converges weakly as  $\varepsilon \rightarrow 0$  to the limiting process  $V^0(t)$  which satisfies the following equation

$$V^0(t)f = f + \hat{\lambda} \cdot \int_0^t \hat{L} \cdot V^0(s)f ds + \int_0^t \int_X \sigma \Gamma(x) V^0(s)f W(dx, ds), \quad (36)$$

where

$$\begin{aligned} \hat{L} : &= \int_X \rho(dx) L(x), \\ L(x) &:= (m\Gamma(x) + D_1(x))(R_0 - I)(m\Gamma(x) + D_1(x)) + \\ &\quad + m_2 \Gamma^2(x)/2 + mD_1(x)\Gamma(x) + D_2(x), \end{aligned}$$

$R_0$  is a potential operator of  $(x_n; \quad n \geq 0)$  (see [14, 15, 23]),  $W(dx, ds)$  is the Wiener orthogonal martingale measure with quadratic variation  $\rho(dx)ds$  (see [15, 26]) and  $\hat{\lambda}$  is defined in (31). Proof of Theorem 2 is given in the Appendix.

## 6 Applications of SERE in Limit Theorems

Applications of SERE will be given to self-exciting traffic/transport process and self-exciting summation on a Markov chain, which are continuous and discrete SERE. From these processes we can construct many other self-exciting processes, e.g., such as impulse traffic/transport process, self-exciting risk process, etc. We will present averaged and diffusion approximation of self-exciting processes. All convergences are in Skorokhod topology [19] when  $\epsilon \rightarrow 0$ . The results here follow from sec. 5.1-5.2 (Theorem 1 and Theorem 2).

### 6.1 Averaged Self-exciting Traffic/Transport and Summation on a Markov Chain Processes (see sec. 4.2)

If we have self-exciting traffic/transport process in series scheme  $t/\epsilon$

$$z^\epsilon(t) = z + \int_0^t v(z^\epsilon(s), x(s/\epsilon)) ds, \quad (37)$$

then the averaged self-exciting traffic/transport process satisfies the ODE (weak convergence when  $\epsilon \rightarrow 0$  in Skorokhod topology [19])

$$\frac{d\hat{z}(t)}{dt} = \hat{v}(\hat{z}(t)), \quad \hat{z}(0) = z, \quad (38)$$

where  $\hat{v}(z) := \hat{\lambda} m \int_X \rho(dx) v(z, x)$ ,  $\hat{\lambda} := \frac{\lambda}{1-\hat{\mu}}$ .

If we have self-exciting summation on a Markov chain process in series scheme  $t/\epsilon$ ,

$$z_t^\epsilon = z + \epsilon \sum_{k=1}^{N(t/\epsilon)} a(x_k), \quad (39)$$

then the averaged self-exciting summation on a Markov chain process has the following form: (weak convergence when  $\epsilon \rightarrow 0$  in Skorokhod topology [19])

$$\hat{z}_t = z + \hat{a}t, \quad (40)$$

where  $\hat{a} := \hat{\lambda} \int_X \rho(dx) a(x)$ .

### 6.2 Diffusion Self-exciting Traffic/Transport and Summation on a Markov Chain Processes

If we have self-exciting traffic/transport process in series scheme  $t/\epsilon^2$

$$z^\epsilon(t) = z + \int_0^t v(z^\epsilon(s), x(s/\epsilon^2)) ds, \quad (41)$$

with balance condition  $\hat{v}(z) = \int_X \rho(dx) v(z, x) = 0$ , then the diffusion self-exciting traffic/transport process has the following SDE form (weak convergence when  $\epsilon \rightarrow 0$  in Skorokhod topology [19])

$$d\hat{z}(t) = b(\hat{z}(t))dt + \sigma(\hat{z}(t))dw(t), \quad (42)$$

where  $b(z) := \hat{\lambda} \int_X \rho(dx) [m^2 v(z, x)(R_0 - I)v'(z, x) + m_2 v(z, x)v'(z, x)/2]$  and  $\sigma^2(z) := 2\hat{\lambda} \int_X \rho(dx) [m^2 v(z, x)(R_0 - I)v(z, x) + m_2 v^2(z, x)/2]$ ,  $R_0$  is a potential of the Markov chain  $(x_n, n \geq 0)$ ,  $w(t)$  is a standard Wiener process.

If we have self-exciting summation on a Markov Chain Processes in series scheme  $t/\epsilon^2$

$$z_t^\epsilon = z + \epsilon \sum_{k=1}^{N(t/\epsilon^2)} a(x_k), \quad (43)$$

with a balance condition  $\hat{a} = \int_X \rho(dx) a(x) = 0$ , then the diffusion summation on a Markov chain (weak convergence when  $\epsilon \rightarrow 0$  in Skorokhod topology [19]) is a Wiener process with variance

$$\sigma^2 := 2\hat{\lambda} \int_X \rho(dx) [m^2 a(x)(R_0 - I)a(x) + m_2 a^2(x)/2], \quad (44)$$

i.e.,  $z_t^\epsilon \rightarrow \hat{z}_t$  when  $\epsilon \rightarrow 0$ , and

$$\hat{z}_t = z + \sigma w(t).$$

## 7 Appendix: Proofs of Theorem 1 (Averaging of SERE) and Theorem 2 (DA of SERE)

### 7.1 Proof of Averaging of SERE (Theorem 1)

Under conditions A) - C) the sequence of SERE  $V_\epsilon(t)f$  is tight (see **Proposition (Tightness of SERE)**, sec. 4).

Under conditions D), E),  $i = 2, F)$ ,  $k = 2, j = 1$ , the sequence of SERE  $V_\epsilon(t)f$  is weakly compact  $\rho - a.s.$  in  $D_B[0, +\infty)$  with limit points in  $C_B[0, +\infty)$ ,  $f \in B_0$ . Here:  $D_B[0, +\infty)$  is a space of  $B$ -valued càdlàg function on  $[0, +\infty)$ , and  $C_B[0, +\infty)$  is a space of  $B$ -valued continuous function on  $[0, +\infty)$ .

Let's consider the following process in  $D_B[0, +\infty)$ :

$$M_{\nu(t/\epsilon)}^\epsilon f^\epsilon := V_{\nu(t/\epsilon)}^\epsilon f^\epsilon - f^\epsilon - \sum_{k=0}^{N(t/\epsilon)-1} E_\rho[V_{k+1}^\epsilon f_{k+1}^\epsilon - V_k^\epsilon f_k^\epsilon / \mathcal{F}_k], \quad (45)$$

where  $V_n^\epsilon := V_\epsilon(\epsilon \tau_n)$ ,

$$f^\epsilon := f + \epsilon f_1(x(t/\epsilon)),$$

$$f_k^\varepsilon := f^\varepsilon(x_k),$$

function  $f_1(x)$  is defined from the equation

$$\begin{aligned} (P - I)f_1(x) &= \left[ (\hat{\Gamma} + \hat{D}) - (m\Gamma(x) + PD_1(x)) \right] f, \\ \hat{\Gamma} &:= m \int_x \rho(dx) \Gamma(x), \quad \hat{D} := \int_x \rho(dx) D_1(x), \\ m &:= m_1 \end{aligned} \tag{46}$$

(see E)),  $f \in B_0$ .

The process  $M_{N(t/\varepsilon)}^\varepsilon f^\varepsilon$  is an  $\mathcal{F}_t^\varepsilon$ -martingale with respect to the  $\sigma$ -algebra  $\mathcal{F}_t^\varepsilon := \sigma\{x(s/\varepsilon); 0 \leq s \leq t\}$ .

The martingale  $M_{N(t/\varepsilon)}^\varepsilon f^\varepsilon$  in (45) has the asymptotic representation:

$$M_{N(t/\varepsilon)}^\varepsilon f^\varepsilon = V_{N(t/\varepsilon)}^\varepsilon f - f - \varepsilon \sum_{k=0}^{N(t/\varepsilon)} (\hat{\Gamma} + \hat{D}) V_k^\varepsilon f + 0_f(\varepsilon), \tag{47}$$

where  $\hat{\Gamma}, \hat{D}, f, f^\varepsilon$  are defined in (46) and

$$\|0_f(\varepsilon)\|/\varepsilon \rightarrow \text{const} \quad \text{as } \varepsilon \rightarrow 0, \quad \forall f \in B_0.$$

The families  $l(M_{N(t/\varepsilon)}^\varepsilon f^\varepsilon)$  and

$$l \left( \sum_{k=0}^{N(t/\varepsilon)} E_\rho[(V_{k+1}^\varepsilon f_{k+1}^\varepsilon - V_k^\varepsilon f_k^\varepsilon)/\mathcal{F}_k] \right)$$

are weakly compact for all  $l \in B_0^*$  is a some dense subset from  $B^*$ . Let  $V_0(t)$  be a limit process for  $V_\varepsilon(t)$  as  $\varepsilon \rightarrow 0$ .

Since (see (25))

$$[V_\varepsilon(t) - V_{N(t/\varepsilon)}^\varepsilon] = [\Gamma_{x(t/\varepsilon)}(t - \varepsilon \tau_{N(t/\varepsilon)}) - I] \cdot V_{N(t/\varepsilon)}^\varepsilon \tag{48}$$

and the right-hand side in (48) tends to zero as  $\varepsilon \rightarrow 0$ , then it is clearly that the limits for  $V_\varepsilon(t)$  and  $V_{N(t/\varepsilon)}^\varepsilon$  are the same, namely,  $V_0(t) = \rho - a.s.$

Taking into account the LLN for Hawkes process,  $N(t)/t \rightarrow_{t \rightarrow +\infty} \lambda/(1 - \hat{\mu})$ , see, e.g., [3, 4, 7, 27], the sum  $\varepsilon \cdot \sum_{k=0}^{N(t/\varepsilon)} (\hat{\Gamma} + \hat{D}) V_k^\varepsilon f$  converges strongly as  $\varepsilon \rightarrow 0$  to the integral

$$\hat{\lambda} \cdot \int_0^t (\hat{\Gamma} + \hat{D}) V_0(s) f ds,$$

where

$$\hat{\lambda} := \frac{\lambda}{1 - \hat{\mu}}. \tag{49}$$

The quadratic variation of the martingale  $l(M_{\nu(t/\varepsilon)}^\varepsilon f^\varepsilon)$  tends to zero, and, hence,

$$M_{\nu(t/\varepsilon)}^\varepsilon f^\varepsilon \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad \forall f \in B_0, \quad \forall e \in B_0^*.$$

Passing to the limit in (47) as  $\varepsilon \rightarrow 0$  and taking into account the all previous reasonings we obtain that the limit process  $V_0(t)$  satisfies the equation:

$$V_0(t)f - f - \hat{\lambda} \int_0^t (\hat{\Gamma} + \hat{D})V_0(s)f ds = 0, \quad (50)$$

where  $\hat{\lambda}$  is defined in (49).

The uniqueness of the limiting SERE  $V_0(t)f$  in averaging scheme follows from the equation (50) and the fact that if the operator  $\hat{\Gamma} + \hat{D}$  generates a semigroup, then  $V_0(t)f = \exp\{\hat{\lambda}(m\hat{\Gamma} + \hat{D}) \cdot t\}f$  and this representation is unique. *Q.E.D.*

**Remark.** In a similar way we can consider averaging of self-exciting random evolutions in reducible phase space, or merged self-exciting random evolutions, as it was done in [14, 15, 23] for semi-Markov RE.

## 7.2 Proof of DA of SERE (Theorem 2)

Under conditions A) - C) the sequence of SERE  $V^\varepsilon(tf)$  is tight (see **Proposition (Tightness of SERE)**, sec. 4).

Under conditions  $D), E), i = 3, F), k = 4, j = 2$ , the sequence of SERE  $V_\varepsilon(t/\varepsilon)f$  is weakly compact  $\rho - a.s.$  in  $D_B[0, +\infty)$  with limit points in  $C_B[0, +\infty)$ ,  $f \in B_0$ .

Let us consider the following process in  $D_B[0, +\infty)$  :

$$M_{N(t/\varepsilon^2)}^\varepsilon f^\varepsilon := V_{N(t/\varepsilon^2)}^\varepsilon f^\varepsilon - f^\varepsilon - \sum_{k=0}^{N(t/\varepsilon^2)-1} E_\rho[V_{k+1}^\varepsilon f_{k+1}^\varepsilon - V_k^\varepsilon f_k^\varepsilon / \mathcal{F}_k], \quad (51)$$

where  $f^\varepsilon := f + \varepsilon f_1(x(t/\varepsilon^2)) + \varepsilon^2 f_2(x(t/\varepsilon^2))$ , and functions  $f_1$  and  $f_2$  are defined from the following equations:

$$\begin{aligned} (P - I)f_1(x) &= -[m\Gamma(x) + PD_1(x, \cdot)]f, \\ (P - I)f_2(x) &= [\hat{L} - L(x)]f, \\ \hat{L} &= \int_X \rho(dx)L(x), \end{aligned} \quad (52)$$

$$\begin{aligned} L(x) &:= (m\Gamma(x) + D_1(x))(R_0 - I)(m\Gamma(x) + D_1(x)) + \\ &\quad + m_2\Gamma^2(x)/2 + mD_1(x)\Gamma(x) + D_2(x), \end{aligned}$$



$R_0$  is a potential operator of  $(x_n; \quad n \geq 0)$  (see [14, 15, 23]).

The balance condition (35) and condition  $\int_X \rho(dx)(\hat{L} - L(x)) = 0$  give the solvability of the equations in (52).

The process  $M_{\nu(t/\varepsilon^2)}^\varepsilon f^\varepsilon$  is an  $\mathcal{F}_t^\varepsilon$ -martingale with respect to the  $\sigma$ -algebra  $\mathcal{F}_t^\varepsilon := \sigma\{x(s/\varepsilon^2); 0 \leq s \leq t\}$ .

This martingale has the asymptotic representation:

$$M_{N(t/\varepsilon^2)}^\varepsilon f^\varepsilon = V_{N(t/\varepsilon^2)}^\varepsilon f - f - \varepsilon^2 \sum_{k=0}^{N(t/\varepsilon^2)-1} \hat{L} V_k^\varepsilon f - 0_f(\varepsilon), \quad (53)$$

where  $\hat{L}$  is defined in (52) and

$$\|0_f(\varepsilon)\|/\varepsilon \rightarrow \text{const} \quad \varepsilon \rightarrow 0, \quad \forall f \in B_0.$$

The families  $l(M_{N(t/\varepsilon^2)}^\varepsilon f^\varepsilon)$  and  $l(\sum_{k=0}^{N(t/\varepsilon^2)} E_\rho[(V_{k+1}^\varepsilon f_{k+1}^\varepsilon - V_k^\varepsilon f_k^\varepsilon)/\mathcal{F}_k])$  are weakly compact for all  $l \in B_0^*$ ,  $f \in B_0$ .

Set  $V^0(t)$  for the limit process for  $V_\varepsilon(t/\varepsilon)$  as  $\varepsilon \rightarrow 0$ .

Similar to (48) we obtain that the limits for  $V^\varepsilon(t)$  and  $V_{N(t/\varepsilon^2)}^\varepsilon$  are the same, namely,  $V^0(t)$ , because

$$[V^\varepsilon(t) - V_{N(t/\varepsilon^2)}^\varepsilon] = [\Gamma_{x(t/\varepsilon^2)}(t - \varepsilon \tau_{N(t/\varepsilon^2)}) - I] \cdot V_{N(t/\varepsilon^2)}^\varepsilon.$$

Taking into account the LLN for Hawkes process,  $N(t)/t \rightarrow_{t \rightarrow +\infty} \lambda/(1 - \hat{\mu})$ , see e.g., [3, 4, 7, 27], the sum  $\varepsilon^2 \sum_{k=0}^{N(t/\varepsilon^2)} \hat{L} V_k^\varepsilon f$  converges strongly as  $\varepsilon \rightarrow 0$  to the integral  $\hat{\lambda} \int_0^t \hat{L} V^0(s) f ds$ , where  $\hat{\lambda}$  is defined in (49).

Set  $M^0(t)f$  be a limit martingale for  $M_{\nu(t/\varepsilon^2)}^\varepsilon f^\varepsilon$  as  $\varepsilon \rightarrow 0$ .

Then, from (32)–(35) and previous reasonings (51)–(53), we have as  $\varepsilon \rightarrow 0$ :

$$M^0(t)f = V^0(t)f - f - \hat{\lambda} \cdot \int_0^t \hat{L} V^0(s) f ds. \quad (54)$$

The quadratic variation of the martingale  $M^0(t)f$  has the form:

$$< l(M^0(t)f) > = \int_0^t \int_X l^2(\sigma \Gamma(x) V^0(s) f) \rho(dx) ds, \quad (55)$$

where

$$\sigma^2 := \hat{\lambda}[m_2 - m^2],$$

where  $m_2, m$  are defined in E), (6), and  $\hat{\lambda}$  in (49).

The solution of martingale problem for  $M^0(t)$  (namely, to find the representation of  $M^0(t)$  with quadratic variation (55)) is expressed by the integral

over Wiener orthogonal martingale measure  $W(dx, ds)$  with quadratic variation  $\rho(dx) \cdot ds$ :

$$M^0(t)f = \sigma \int_0^t \int_x \Gamma(x) V^0(s) f W(dx, ds). \quad (56)$$

In such a way, the limit process  $V^0(t)$  satisfies the following equation (see (54) and (56)):

$$V^0(t)f = f + \hat{\lambda} \cdot \int_0^t \hat{L} \cdot V^0(s) f ds + \sigma \int_0^t \int_x \Gamma(x) V^0(s) f W(dx, ds). \quad (57)$$

If the operator  $\hat{L}$  generates the semigroup  $U(t)$  then the process  $V^0(t)f$  in (57) satisfied equation:

$$V^0(t)f = U(t)f + \sigma \int_0^t \int_x U(t-s) \Gamma(x) V^0(s) f W(dx, ds). \quad (58)$$

The uniqueness of the limit evolution  $V^0(t)f$  in diffusion approximation scheme follows from the uniqueness of the solution of martingale problem for  $V^0(t)f$  (see ((54)-(56)). The latter is proved by dual SERE in series scheme by constructing the limit equation in diffusion approximation and by using a dual identify. For more details, see [14]. *Q.E.D.*

**Remark.** In a similar way, we can consider diffusion approximation of self-exciting random evolutions in reducible phase space, as we did it in [14, 15, 23] for semi-Markov RE.

## 8 Discussion

The main difference between SERE and, e.g., semi-Markov REs (SMRE) is that it involves new features for the REs: self-exciting and clustering ones, which is not the case for the SMRE. This was incorporated into the Swish process  $x(t) := x_{N(t)}$ , where  $N(t)$  is a Hawkes process, and into the parameter  $\hat{\lambda} := \frac{\lambda}{1-\hat{\mu}}$ ,  $\hat{\mu} := \int_0^{+\infty} \mu(t) dt = \alpha/\beta$ , see, e.g., sec. 6.1, for the averaged self-exciting summation on a Markov chain. This parameter is presented in all results for the SERE, including averaging and diffusion approximation, and their applications. These features show originality and novelty of the paper. The results for SERE obtained in the paper generalize many results in the literature devoted to the Hawkes-based models in finance and insurance, such as, e.g., risk model based on general compound Hawkes process [24] (self-exciting impulse traffic/transport process (8) in this paper) and general compound Hawkes processes in limit order books [25] (self-exciting

summation on a Markov chain in this paper). In applications, such as high-frequency and algorithmic trading (HFT) and in limit order books modelling (when transactions happen every milliseconds), one usually considers the long scale  $tn$  (minutes, hours, etc.) instead of  $t$  (milliseconds) to have time for solutions many problems like liquidation, acquisitions, and market making, to name a few. In this case, the limit theorems such as law of large numbers (averaging) and functional central limit theorem (diffusion approximation) for the case when  $n \rightarrow +\infty$  and  $t$  is fixed, are very useful. In this paper, we used the scaling  $t/\epsilon$  to get our results when  $\epsilon \rightarrow 0$ . Of course, it is similar to the case  $tn$ , when  $n \rightarrow +\infty$  (see, e.g., [25]). Therefore, many applications of SERE already exist in finance [25] and risk theory [24]. However, many more applications of SERE to the examples of SERE in sec. 3 could be considered space permitting for the current paper. Finally, we note that some new results such as normal deviations and rates of convergence for SERE and their applications could also be formulated space permitting for the paper. Therefore, the author leaves preparing those results and applications to the future research work.

## 9 Conclusion

In this paper we have developed a new class of so-called self-exciting random evolutions (SEREs) and their applications. We have first introduced a new random process  $x(t)$  such that it is based on a superposition of a Markov chain  $x_n$  and a Hawkes process  $N(t)$ , i.e.,  $x(t) := x_{N(t)}$ . We call this process self-walking imbedded semi-Hawkes process (Swish Process or SwishP). Then the self-exciting REs (SEREs) have been constructed in similar way as, e.g., semi-Markov REs, but instead of semi-Markov process  $x(t)$  we have SwishP. We gave classifications and examples of self-exciting REs (SEREs). Then we considered two limit theorems for SEREs such as averaging (Theorem 1) and diffusion approximation (Theorem 2). Applications of SEREs have been devoted to the so-called self-exciting traffic/transport process and self-exciting summation on a Markov chain, which are examples of continuous and discrete SERE. From these processes we can constructed many other self-exciting processes, e.g., such as impulse traffic/transport process, self-exciting risk process, etc. We also presented averaged and diffusion approximation for some self-exciting processes.

## Acknowledgment

The author thanks NSERC for continuing support.

## References

- [1] Adams, R (1979): *Sobolev Spaces*. Academic Press.
- [2] Y. Ait-Sahalia, J. Cacho-Diaz, R. Laeven, (2015): Modelling financial contagion using mutually exciting jump processes, *J. Financial Econ.* 117, 58-606.
- [3] E. Bacry, S. Delattre, M. Hoffmann, J. F. Muzy, (2013): Some limit theorems for Hawkes processes and application to financial statistics., *Stoch. Proc. Appl.* 123, 2475-2499.
- [4] Bremaud, P., and L. Massoulié. (1996): Stability of nonlinear Hawkes processes. *Annals of Probability*. 24, 1563-1588.
- [5] Dunford, N. and Schwartz, J. (1958): *Linear Operators. General Theory*. Wiley.
- [6] Embrechts, P., Liniger, T. and Lin, L. (2011): Multivariate Hawkes Processes: an Application to Financial Data. *Applied Probability Trust*, 7 March.
- [7] Fierro, R., Leiva, V. and Moeller, J. (2015): The Hawkes process with different exciting functions and its asymptotic behaviour. *J. Appl. Prob.* 52, 37-54.
- [8] Hersh, R. (1974): Random evolutions: A survey of results and problems. *Rocky Mount. Math. J.*, 4(3), 443-478.
- [9] Hawkes, A.(1971): Point spectra of some mutually-exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)* 33: 438-43.
- [10] Hawkes, A. (1971): Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58, 83-90.
- [11] Hawkes, A. G. and Oakes, D. (1974): A cluster representation of a self-exciting process. *J. Appl. Prob.* 11, 493-503.
- [12] A.G. Hawkes, (2018): Hawkes processes and their applications to finance: a review, *Quantitative Finance*, 18:2, 193-198
- [13] D. Karabash, L. Zhu, (2015): Limit theorems for marked Hawkes processes with application to a risk model. *Stoch. Models*, 31, 433-451.
- [14] Korolyuk, V. S. and Swishchuk A. V. (1995): *Evolution of Systems in Random Media*. Chapman & Hall CRC, 1995.
- [15] Korolyuk, V. S. and Swishchuk A. V. (1995): *Semi-Markov Random Evolutions*. Kluwer AP, 1995.

- [16] Li, B. and Pang, G. (2022): Functional limit theorems for non-stationary marked Hawkes processes in the high intensity regime. *Stoch. Proc. Appl.*, v 143, Jan 2022, 285-339.
- [17] Pinsky, M. (1991): *Lectures on Random Evolutions*. World Scientific Publishers.
- [18] Y. Seol, (2020): Limit theorems for a jump-diffusion model with Hawkes jumps, *Mathematical Communications*, Volume 26, 53-67.
- [19] Skorokhod, A. (1965): *Studies in the Theory of Random Processes*. Addison-Wesley.
- [20] Skorokhod, A. (1984): *Random Linear Operators*. Dordrecht, Boston, Lancaster; D. Reidel Publishing Co.
- [21] Skorokhod, A. (1974): *Integration in Hilbert Spaces*, New York, Heidelberg; Springer-Verlag.
- [22] Sobolev, S. L. (1991): *Some Applications of Functional Analysis in Mathematical Physics*. American Mathematical Society.
- [23] Swishchuk, A. V. (2000): *Random Evolutions and Their Applications. New Trends*. Kluwer AP.
- [24] Swishchuk, A. (2018): Risk Model Based on General Compound Hawkes Process. *Wilmott*, 2018(94), 50-57.
- [25] Swishchuk, A. and Huffman, A. (2020): General Compound Hawkes Processes in Limit Order Books. *Risks* 2020, 8(1), 28.
- [26] Swishchuk, A. (2019): *Inhomogeneous Random Evolutions and Their Applications*. CRC Press.
- [27] Zhu, L. (2013): Central limit theorem for nonlinear Hawkes processes. *J. Applied Probability*. 50, 760-771.