Investigating the integrate and fire model as the limit of a random discharge model: a stochastic analysis perspective

Jian-Guo Liu*   Ziheng Wang†   Yantong Xie‡   Yuan Zhang§
Zhennan Zhou¶

Abstract

In the mean field integrate-and-fire model, the dynamics of a typical neuron within a large network is modeled as a diffusion-jump stochastic process whose jump takes place once the voltage reaches a threshold. In this work the main goal is to establish the convergence relationship between a regularized process and the original one where in the regularized process, the jump mechanism is replaced by a Poisson dynamic, and jump intensity within the classically forbidden domain goes to infinity as the regularization parameter vanishes. On the macroscopic level, the Fokker-Planck equation for the process with random discharges (i.e. Poisson jumps) is defined on the whole space, while the equation for the limit process is defined on the half space. However, using an iteration scheme the difficulty due to the domain differences has been greatly mitigated and the convergence for the stochastic process and the firing rates can be established. Moreover, we find polynomial-order convergence for the distribution by a re-normalization argument in probability theory. Finally, using numerical experiments we quantitatively explore the rate and the asymptotic behavior of convergence for both linear and nonlinear models.

Keywords: Integrate and fire model; Fokker-Planck equation; Regularization; Weak convergence; Poisson jump; Neural network.
MSC2020 subject classifications: 60B10; 60H30; 92B20; 92C20.
Submitted to MNA on February 19, 2021, final version accepted on October 21, 2021.

1 Introduction

The classical description of the dynamics of a large set of neurons is based on deterministic/stochastic differential systems for the excitatory-inhibitory neuron network [26, 35, 40, 50]. One of the most famous models is the noisy leaky integrate and fire (LIF) model [35], where the collective behavior of a neural network can be averaged as a self-consistent environment and within the network the typical behavior of a neuron

---

*Department of Mathematics and Department of Physics, Duke University. E-mail: jliu@phy.duke.edu
†Mathematical Institute, University of Oxford. E-mail: wangz1@math.ox.ac.uk
‡School of Mathematical Sciences, Peking University. E-mail: 1901110045@pku.edu.cn
§School of Mathematical Sciences, Peking University. E-mail: zhangyuan@math.pku.edu.cn
¶Beijing International Center for Mathematical Research, Peking University. E-mail: zhennan@bicmr.pku.edu.cn

is approximated by a stochastic process [3, 4, 11, 15, 28, 33, 37, 38, 39, 46, 47, 49] and the influence from the network is given by an average synaptic input by a mean-field approximation [15, 33, 46, 49]. In this model, the membrane potential of a typical neuron within the network is denoted by the state variable X_t. When the synaptic input of the network (denoted by $I(t)$) vanishes, the membrane potential relaxes to the resting value V_L. In the single neuron approximation, the synaptic input $I(t)$, which itself is another stochastic process, is replaced by a continuous-in-time counterpart $I_c(t)$ (see e.g. [3, 4, 37, 42, 46, 47]), which takes the drift-diffusion form

$$I\,dt \approx I_c\,dt = \mu_c\,dt + \sigma_c\,dB_t. \tag{1.1}$$

Here, B_t is standard Brownian motion, and in principle the two processes $I_c(t)$ and $I(t)$ have the same mean and variance. Thus between the firing events, the evolution of the membrane potential is given by the following stochastic differential equation

$$dX_t = (-X_t + V_L + \mu_c)\,dt + \sigma_c\,dB_t. \tag{1.2}$$

Another important ingredient in the modeling is the firing-and-resetting mechanism: whenever the membrane voltage X_t reaches a threshold value V_F, it is immediately reset to a specific value $V_R < V_F$. Namely,

$$X_t = V_R \quad \text{if} \quad X_{t^-} = V_F. \tag{1.3}$$

The reader may refer to [46] for a thorough introduction to this subject.

From the perspective of probability theory, the jump-diffusion processes of type (1.2) and (1.3) were first introduced and studied by Feller [22, 23] (in terms of transition semigroups), which apparently was not motivated by applications in neuroscience. More specifically, [23] named such a process an "elementary return process" and presented its Fokker-Planck equation in a weak form, the proof of which was based on a Markov semigroup argument in [22]. In [1, 2, 43, 44], the authors were concerned with the spectral properties of the generator of the stochastic process or related models, and showed the exponential convergence in time towards the stationary distribution. In particular, [43] applied their results to a neuronal firing model driven by a Wiener process and computed the distribution of the first passage time. In the work [45, 48], the authors made assumptions on the stochastic process that were more relaxed than or modified from those in [27] and proved the existence of a pathwise solution of such process in a generalized sense.

For the jump-diffusion process (1.2) and (1.3), there has been a growing interest in studying the partial differential equation model for the dynamics of the probability density function with which the stochastic process X_t is associated. From the heuristic viewpoint using Itô's calculus, it is widely accepted ([5, 29]) that the evolution of the probability density $f(x, t) \geq 0$ of finding neurons at voltage x at time $t \geq 0$ satisfies the following Fokker-Planck equation on the half line with a singular source term

$$\frac{\partial f}{\partial t}(x,t) + \frac{\partial}{\partial x}[hf(x,t)] - a\frac{\partial^2 f}{\partial x^2}(x,t) = N(t)\delta(x - V_R), \quad x \in (-\infty, V_F), \quad t > 0, \tag{1.4}$$

where the drift velocity is $h = -x + V_L + \mu_c$, the diffusion coefficient is $a = \sigma_c^2/2$, $\delta(x)$ denotes the Dirac function and the precise definition of the mean firing rate $N(t)$ is given by (1.7). We complement (1.4) with the following Dirichlet and initial boundary conditions:

$$f(V_F, t) = 0, \quad f(-\infty, t) = 0, \quad f(x, 0) = f_{\text{in}}(x) \geq 0. \tag{1.5}$$

Equation (1.4) is supposed to be the evolution of a probability density, therefore

$$\int_{-\infty}^{V_F} f(x,t)\,dx = \int_{-\infty}^{V_F} f_{\text{in}}(x)\,dx = 1. \tag{1.6}$$

Due to the Dirichlet boundary condition at $x = V_F$, there is a time-dependent boundary flux escaping the domain, and a Dirac delta source term is added to the reset location $x = V_R$ to compensate for the loss. It is straightforward to check that the conservation equation (1.6) characterizes the mean firing rate $N(t)$ as the flux of neurons at the firing voltage, which is implicitly given by

$$N(t) := -a\frac{\partial f}{\partial x}(V_F, t) \geq 0. \tag{1.7}$$

We remark that this delta function source term on the right hand side of (1.4) is equivalent to setting the equation on $(-\infty, V_R) \cup (V_R, V_F)$ instead and imposing the following conditions

$$f(V_R^-, t) = f(V_R^+, t), \quad a\frac{\partial}{\partial x}f(V_R^-, t) - a\frac{\partial}{\partial x}f(V_R^+, t) = N(t), \quad \forall t \geq 0.$$

The firing events generate currents that propagate within the neuron network, which can be incorporated into this PDE model by expressing the drift velocity h and the diffusion coefficient a as functions of the mean-firing rate $N(t)$ (see e.g. [5, 6, 9, 31]). In the simplest form, the following choice has been widely considered

$$h(x, N(t)) = -x + bN(t), \quad a(N(t)) = a_0 + a_1(N(t)). \tag{1.8}$$

In particular, the term $-x$ describes the leaky behavior and b models the connectivity of the network: $b > 0$ describes excitatory networks and $b < 0$ describes inhibitory networks. In Sections 2 and 3, we only consider the simple case when $b = a_1 = 0$ and $a_0 = 1$, while in Section 4 we numerically investigate the nonlinear case when $b \neq 0$.

Many recent works are devoted to investigating the properties of solutions of such PDE models, including the finite-time blow-up of weak solutions, the multiplicity of steady solutions, the relative entropy estimate, the existence of classical solutions, structure-preserving numerical approximations, etc. (see e.g. [5, 6, 7, 8, 9, 31] and the references therein.) However to the best of our knowledge, due to the firing-and-resetting mechanism in (1.2), the rigorous derivation of the Fokker-Planck equation (1.4) in the classical sense from the microscopic stochastic process had not yet been achieved by conventional methods.

In [29], we established the rigorous connection between the linear Fokker-Planck equation (1.4) and the microscopic stochastic model (1.2) and (1.3). For simplicity, we assume for the rest of the work that

$$V_L = 0, \quad V_R = 0, \quad V_F = 1, \quad \mu_c = 0, \quad \sigma = \sqrt{2}, \tag{1.9}$$

which means the LIF model becomes a linear model with the interactions among the network neglected. With these assumptions, the process (1.2) and (1.3) becomes the standard O-U process with a "hard wall" at 1, i.e., whenever at time t, X_t hits 1, it immediately jumps to 0 and then we restart the O-U-like evolution independent of the past. Unlike the standard jump-diffusion process, the jumping time for X_t is determined by the hard wall boundary and thus the classical Itô calculus is not directly applicable. Inspired by the renewal nature of X_t that agrees with the pioneering work of Feller [23], a novel strategy based on an iterated scheme was proposed in [29] to show that the probability density function (abbreviated by p.d.f.) of X_t is the classical solution of its Fokker-Planck equation.

In fact, with the introduction of the auxiliary stochastic process counting the number of firing events, the density $f(x, t)$ of the potential X_t can be decomposed as a summation of sub-density functions $\{f_n(x, t)\}_{n=0}^{\infty}$. Each sub-density naturally links to a

less singular sub-PDE problem which is determined iteratively and of better regularity. Besides, with the exponential decay of decomposition all the regularities are preserved in the limit, and thus we concluded the desired properties for the PDE problem (1.4).

In this paper, we continue to study a family of related jump-diffusion processes (2.3) parameterized by δ, which are used to approximate the process X_t as $\delta \to 0^+$. Introduced in [7], such jump-diffusion processes are used to explore the reasonable modelling of the mean firing events on the macroscopic level, which are associated with the Fokker-Planck equations of neurons with random firing thresholds. As shown in [5], the solution of the Fokker-Planck equation with a deterministic firing potential may blow up in finite time, which is speculated to be related to the synchronization activity of neuronal networks. The random discharge mechanism in [7] provides an alternative scenario of incorporating the synchronous states besides introducing time delay, refractory states, etc (see [3, 4, 7, 13, 34]). However, all the regularization models are mostly based on scientific intuition or technical insights of the PDE theory, among which the random discharge model is tractable on the microscopic level and is reminiscent of the well-studied kinetic equations on the macroscopic level. Thus, we choose to focus on the random discharge model to rationalize the regularization effect and show the convergence relationship between such a model and its limit as the regularization parameter vanishes.

We denote such processes by X_t^δ, and the associated jumping rate $\lambda^\delta(X_t^\delta) = 0$ when $X_t^\delta < 1$, and $\lambda^\delta(X_t^\delta) = O(\delta^{-1})$ when $X_t^\delta > 1 + \delta$. Between the firing events, X_t^δ propagates along the O-U process

$$dX_t^\delta = -X_t^\delta \, dt + \sqrt{2} \, dB_t. \tag{1.10}$$

Recall that there exists a "hard wall" boundary for X_t, i.e. the firing event takes place whenever X_t reaches 1. However, the jumps of X_t^δ are determined by a state-dependent Poisson measure, for it evolves as the standard O-U process when $X_t^\delta < 1$ and can jump with a high rate once X_t^δ exceeds 1, and such a jump process can be interpreted as a "soft wall" boundary when $X_t^\delta \geq 1$. The precise definition of X_t^δ can be found in (2.2) and (2.3). The Poisson jump model frequently appears in kinetic models [10, 14, 17, 20] and with Itô's formula, we can derive its Dynkin's formula, forward and backward Kolmogorov equation and Feynman-Kac formula, etc. However, Itô calculus is not directly applicable when $\delta = 0$ and thus X_t is seen as the singular limit of a family of regularized processes. The primary goal of this paper is to justify whether and in what sense the regularized model converges to the original one.

Formally speaking, X_t can be seen as the limit of X_t^δ as $\delta \to 0^+$, for the distribution of the process X_t^δ is supposed to converge to the process X_t as $\delta \to 0^+$. The rigorous justification of such convergence is challenging due to the domain differences, and the main contribution of this paper is to rigorously establish the relationship between the two processes. We first present the strong Feller property (2.5) of the limit process X_t by comparing it with the regular O-U process. Then by a similar iteration strategy as that in [29], (see the decomposition (3.9) and the iteration relationship in Proposition 3.1 for details) we get convergence for the marginal distribution. Utilizing the strong Feller property for X_t, we easily generalize the convergence result to any finite dimensional distribution by induction, which together with the relatively compactness of $\{X_t^\delta\}_{\delta > 0}$ implies the weak convergence (2.6) for the processes. Also, by using the decomposition and iteration strategy, the weak convergence of the mean firing rate (2.7) can be established. Finally by a standard re-normalization method in probability theory, we rigorously prove a polynomial-order convergence rate (2.8) for the marginal distribution.

As a complement and extension to the convergence justification, we numerically

explore the convergence rate and asymptotic behavior of the process X_t^δ and its density function for both the linear cases considered in previous analysis and nonlinear cases where the analysis is still intractable in the current framework. The numerical scheme is based on the Scharfetter-Gummel reformulation and preserves certain structures of the Fokker-Planck equation for X_t^δ [31]. We also find numerical evidence of the existence of an asymptotic profile. More specifically, the simulation results strongly suggest that one can numerically identify two parameters α and β such that

$$f^\delta(x,t) = \delta^\alpha \phi\left(\delta^\beta (x-1)\right) + o\left(\delta^\alpha\right), \quad \forall x \in [1, +\infty), \tag{1.11}$$

where ϕ is a profile function independent of δ. Equation (1.11) suggests how f^δ vanishes on the half-space $\{x \geq 1\}$ in a self-similar fashion.

It is worth mentioning that some recent works [12, 13] are devoted to exploring the long time behavior of some non-linear McKean-Vlasov type SDEs arising from neuron models, where the firing mechanism follows the same random discharge dynamic as the one in this work but the diffusion process is absent. In particular, it is proved in [12] that the SDE admits time periodic solutions through a Hopf bifurcation analysis. Although in this paper we only focus on the linear case, the convergence result may provide tools and insights for studying nonlinear models, especially for the long time behavior of the random discharge model. However, analysis of nonlinear models is beyond the scope of this paper.

The rest of the paper is outlined as follows. In Section 2, we first review the jump process X_t and give the precise definition of X_t^δ. After laying out their Fokker-Planck equations, we summarize the main convergence results between X_t^δ and X_t. In Section 3 we rigorously show by a probabilistic approach that the distribution and mean firing rate of the random discharge model weakly converge to those of the original model. Also, a polynomial order convergence rate for the marginal distribution is established. In Section 4, we investigate using numerical experiments the rates and the asymptotic behavior of related convergence for both linear and nonlinear models. In Section 5, we conclude this paper and give some future research directions. For the rest of this work, we use γ, C, C_0, C_k and C_T to denote generic constants.

2 Preliminaries and main results

We first briefly review the results for the jump process X_t in [29]. The stochastic process X_t has been formally defined in (1.2) and (1.3), and the interested readers may refer to [29] for the rigorous construction of such a process. In the integrate and fire model, the process X_t is used to describe the mean field behavior of the neuron network. Let the distribution of X_0 be denoted by μ. For technical reasons we suppose that μ is a probability measure on $(-\infty, 1-\beta]$ for some $\beta > 0$ and let $f_{\text{in}}(x)$ be its density. Then with the iteration idea, we have already shown in [29] that for any fixed $T > 0$, the Fokker-Planck equation for X_t is

$$\begin{cases} \dfrac{\partial f}{\partial t} - \dfrac{\partial}{\partial x}(xf) - \dfrac{\partial^2 f}{\partial x^2} = 0, \quad x \in (-\infty, 0) \cup (0, 1), t \in (0, T], \\[2mm] f(0^-, t) = f(0^+, t), \quad \dfrac{\partial}{\partial x}f(0^-, t) - \dfrac{\partial}{\partial x}f(0^+, t) = -\dfrac{\partial}{\partial x}f(1^-, t), \quad t \in (0, T], \\[2mm] f(-\infty, t) = 0, \quad f(1,t) = 0, \quad t \in [0, T], \\[2mm] f(x, 0) = f_{\text{in}}(x), \quad x \in (-\infty, 1). \end{cases} \tag{2.1}$$

In the rest of this paper, we define

$$N(t) := -\frac{\partial}{\partial x}f(1^-, t), \quad t > 0,$$

which serves as the definition of the mean firing rate.

In this paper, we consider a related family of jump-diffusion processes parameterized by δ as in (1.10), which is related to the Fokker-Planck equation for neurons with random firing thresholds, to approximate the original process X_t. Introduced in [7], we use the random discharge model to rationalize the regularization effect for X_t. The precise definition of X_t^δ is as follows.

For a fixed $\delta > 0$, we define the discharge rate function:

$$\lambda^\delta(x) = \begin{cases} 0, & x \leq 1 \\ \frac{x-1}{\delta^2}, & x \in [1, 1+\delta] \\ \frac{1}{\delta}, & x \geq 1+\delta \end{cases} \tag{2.2}$$

Then consider the following state-dependent jump-diffusion process as defined in Chapter VI of [30]:

$$dX_t^\delta = -X_t^\delta dt + \sqrt{2}dB_t + [-X_{t-}^\delta]dP^\delta(t, X_t^\delta) \tag{2.3}$$

where $P^\delta(t, X_t^\delta)$ is a Poisson point process with intensity $\lambda^\delta(X_t^\delta)$.

Remark 2.1. For the rest of this paper, we only consider the simplified initial condition for the processes, i.e. both X_t and X_t^δ start from a fixed point $x \in (-\infty, 1)$. In the following, $\mathbf{E}^x$ and $\mathbf{P}^x$ denote the expectation and probability of a stochastic process starting from x. The natural extension to general and proper initial conditions can be obtained by convolution.

First by Itô's formula, we can directly derive the Fokker-Planck equation for X_t^δ.

Theorem 2.2. *Let $f^\delta(x,t)$ denote the p.d.f. of the process X_t^δ starting from $y < 1$. It satisfies the following PDE problem in the sense of distributions*

$$\begin{cases} \dfrac{\partial f^\delta}{\partial t} - \dfrac{\partial}{\partial x}\left(xf^\delta\right) - \dfrac{\partial^2 f^\delta}{\partial x^2} = N^\delta(t)\delta(x) - \lambda^\delta(x)f^\delta(x,t), & x \in \mathbb{R}, \quad t > 0, \\ N^\delta(t) = \displaystyle\int_{\mathbb{R}} \lambda^\delta(z)f^\delta(z,t)dz, \quad t > 0, \\ f^\delta(-\infty, t) = f^\delta(+\infty, t) = 0, \quad t > 0, \\ f^\delta(x, 0) = \delta(x - y) \text{ in } \mathcal{D}'(\mathbb{R}), \end{cases} \tag{2.4}$$

where $N^\delta(t)$ is the modified mean firing rate and $\delta(x)$ denotes the Dirac function.

Remark 2.3. To prevent confusion, we clarify that $\delta(x)$ is Dirac function and δ denotes the parameter in the intensity (2.2).

Equation (2.4) is referred to as the Fokker-Planck equation for neuron networks with random discharges [7]. We observe that, in this variant model, the mean firing rate $N^\delta(t)$ is modified to an integral of the density function, which admits a global estimate as shown in [7]. The proof of Theorem 2.2 is elementary, and we choose to omit the details in this work.

However for the process X_t^δ, Itô's calculus is not directly applicable when $\delta = 0$ and the approximation error between the two processes is not quantifiable. The main contribution of this paper is to rigorously establish the relationship between the two processes. We first prove the strong Feller property of the limit process X_t, which plays a key role in the proof of weak convergence. With the iteration strategy, we can prove weak convergence between the stochastic processes and the convergence rate of the marginal distribution, which is summarized in the following theorem.

Theorem 2.4. *(i) (Strong Feller property.) For any bounded and Borel measurable function φ and any $t > 0$, $\mathbf{E}^x[\varphi(X_t)]$ is continuous against the starting point x. I.e., for*

any $\varphi \in \mathcal{B}_b(-\infty, 1)$, $\epsilon > 0$ and $x \in (-\infty, 1)$, there exists $\delta_0 \in (0, \varepsilon)$ such that for any $y \in (-\infty, 1)$ such that $|y - x| \leq \delta_0$, we have

$$\left|\mathbf{E}^x[\varphi(X_t)] - \mathbf{E}^y[\varphi(X_t)]\right| < \varepsilon. \tag{2.5}$$

(ii) (Weak convergence of stochastic process.) Let the stochastic processes X^δ and $X.$ start from any fixed $x \in (-\infty, 1)$. Without loss of generality, suppose that $X_0^\delta = X_0 = 0$. Then X^δ converges weakly to $X.$ as $\delta \to 0^+$, i.e., for any bounded and continuous function $\varphi \in C_b(D_{\mathbb{R}}[0, \infty))$,

$$\lim_{\delta \to 0^+} \mathbf{E}^0 \varphi(X_\cdot^\delta) = \mathbf{E}^0 \varphi(X_\cdot), \tag{2.6}$$

where $D_{\mathbb{R}}[0, \infty)$ denotes the space of right continuous functions $x : [0, \infty) \to \mathbb{R}$ with left limits.

(iii) (Convergence for the mean firing rate.) The modified mean firing rate $N^\delta(t)$ also converges weakly to the mean firing rate $N(t)$, i.e. for any fixed $T > 0$ and $\varphi \in C_b[0, T]$,

$$\lim_{\delta \to 0^+} \int_0^T \varphi(t) N^\delta(t) dt = \int_0^T \varphi(t) N(t) dt. \tag{2.7}$$

(iv) (Convergence rate for marginal distribution.) Let $F^\delta(x, t)$ and $F(x, t)$ denote the cumulative distribution functions of X_t^δ and X_t respectively. Without loss of generality, suppose that $X_0^\delta = X_0 = 0$. Then for any fixed $T < \infty$, there is a constant $\gamma \in (0, \infty)$ s.t. $\forall (x, t) \in \mathbb{R} \times [0, T]$, we have

$$\left|F^\delta(x, t) - F(x, t)\right| = O(\delta^\gamma) \quad \text{as} \quad \delta \to 0^+. \tag{2.8}$$

The detailed proof of the theorem is presented in Section 3, and is hereby outlined as follows.

- In Section 3.1, we show the strong Feller property (2.5) by using the connection between the jump diffusion process X_t and the standard O-U process.
- In Section 3.2.1, by the coupling method and the iteration approach in [29] we prove the convergence for the marginal distribution, i.e., for any fixed $T < \infty$ and $\forall \epsilon > 0$, there is a $\eta_0 > 0$, s.t. $\forall \delta < \eta_0$, we have

$$\left|F^\delta(x, t) - F(x, t)\right| \leq \epsilon \quad \text{for} \quad (x, t) \in \mathbb{R} \times [0, T]. \tag{2.9}$$

- Then in Section 3.2.2, by the strong Feller property for X_t, we get convergence for any finite marginal distribution by induction, which together with the relatively compactness of $\{X_t^\delta\}_{\delta > 0}$ gives the weak convergence for processes.
- In Section 3.3, we prove weak convergence of the mean firing rate by the decomposition and iteration approach in [29].
- Finally in Section 3.4, we rigorously prove a polynomial upper bound on the rate at which $F^\delta(x, t) \to F(x, t)$ as $\delta \to 0^+$ with a multiscale renormalization argument in probability theory.

3 The random discharge model and its convergence

In Section 2, we precisely defined a family of jump-diffusion processes X_t^δ that are associated with the Fokker-Planck equations with random discharges. With the jumping criterion slightly altered, we are able to derive the Fokker-Planck equation of X_t^δ by classical Itô's calculus, which is reckoned as a regularized model (see Theorem 6.1 in [7]). However, the rigorous justification of such convergence is challenging.

3.1 Strong Feller property

First, we prove the strong Feller property of process X_t, which is useful in getting convergence for the finite dimensional distributions. Dr. Lihu Xu at the University of Macau taught us the following nice and easy proof through direct communication.

Proof of the strong Feller property in Theorem 2.4: Similarly to [29], we can strictly construct the jump process X_t starting from x and let T_1^x denote the first time it hits 1. By the proof of Theorem 2 in [29], we know that for any $\varepsilon_0 > 0$ the p.d.f $f_{T_1^y}(\cdot)$ of the first hitting time is uniformly continuous with respect to all $y < 1 - \varepsilon_0$. Thus for any fixed $t > 0$, $\varphi \in C_b(-\infty, 1)$ and $\forall \varepsilon > 0$, there exists $t_0 \in (0, t)$ such that for $\forall y \leq 1 - \varepsilon$,

$$\mathbf{P}(T_1^y \leq t_0) \leq \frac{\varepsilon}{8(\|\varphi\|_{L^\infty} + 1)}. \tag{3.1}$$

Recall that Skorohod [27] has proved that X_t is Markovian, thus

$$\phi_{\varphi, t - t_0}(x) := \mathbf{E}^x[\varphi(X_{t - t_0})]$$

is clearly bounded and measurable against $x \in (-\infty, 1)$. Now applying the strong Feller property of the regular OU process $\{OU_t\}_{t \geq 0}$ on $\phi_{\varphi, t - t_0}$, we have that for any $x \leq 1 - 2\varepsilon$, there exists $\delta_0 \in (0, \varepsilon)$ such that $\forall |y - x| \leq \delta_0$,

$$|\mathbf{E}^x[\phi_{\varphi, t - t_0}(OU_{t_0})] - \mathbf{E}^y[\phi_{\varphi, t - t_0}(OU_{t_0})]| \leq \frac{\varepsilon}{2}. \tag{3.2}$$

Now to compare $\mathbf{E}^x[\varphi(X_t)]$ and $\mathbf{E}^y[\varphi(X_t)]$, we have

$$\mathbf{E}^x[\varphi(X_t)] = \mathbf{E}^x[\varphi(X_t)\mathbb{1}_{T_1^x \leq t_0}] + \mathbf{E}^x[\varphi(X_t)\mathbb{1}_{T_1^x > t_0}] =: I_1^x + I_2^x.$$

Using (3.1), we immediately have that $|I_1^x| \leq \frac{\varepsilon}{8}$, and by the Markov property of X_t, we have

$$\begin{aligned}
I_2^x &= \mathbf{E}^x\left[\mathbf{E}^x\left[\varphi(X_t)\Big|X_{t_0}\right]\mathbb{1}_{T_1^x > t_0}\right] \\
&= \int_{-\infty}^1 \mathbf{E}^x[\varphi(X_t)\big|X_{t_0} = z]f_0^x(z, t_0)dz \\
&= \int_{-\infty}^1 \mathbf{E}^z[\varphi(X_{t - t_0})]f_0^x(z, t_0)dz \\
&= \int_{-\infty}^1 \phi_{\varphi, t - t_0}(z)f_{\text{ou}}^x(z, t_0)dz + \int_{-\infty}^1 \phi_{\varphi, t - t_0}(z)\left[f_0^x(z, t_0) - f_{\text{ou}}^x(z, t_0)\right]dz \\
&=: I_3^x + I_4^x,
\end{aligned} \tag{3.3}$$

where $f_0^x(y, t)dy = \mathbf{P}^x(X_t \in dy, T_1^x > t)$ and $f_{\text{ou}}^x(y, t)$ denotes the p.d.f. of standard OU process starting from x. Noting that $f_0^x(y, t)$ is the p.d.f. for the killed OU process, then $f_{\text{ou}}^x(y, t) \geq f_0^x(y, t)$ and thus for I_4^x, we have

$$\begin{aligned}
|I_4^x| &\leq \|\varphi\|_{L^\infty} \int_{-\infty}^1 f_{\text{ou}}^x(z, t_0) - f_0^x(z, t_0)dz \\
&\leq \|\varphi\|_{L^\infty}\left[1 - \int_{-\infty}^1 f_0^x(z, t_0)dz\right] \\
&= \|\varphi\|_{L^\infty}\mathbf{P}^x(T_1^x \leq t_0) \\
&\leq \frac{\varepsilon}{8}.
\end{aligned} \tag{3.4}$$

Similarly, for y we have

$$\mathbf{E}^y[\varphi(X_t)] = I_1^y + I_2^y$$

$$I_2^y = \int_{-\infty}^1 \phi_{\varphi, t-t_0}(z) f_{\text{ou}}^y(z, t_0) dz + \int_{-\infty}^1 \phi_{\varphi, t-t_0}(z) [f_0^y(z, t_0) - f_{\text{ou}}^y(z, t_0)] dz \tag{3.5}$$
$$=: I_3^y + I_4^y.$$

Then $|I_1^y| \leq \frac{\varepsilon}{8}$ by (3.1) and with the same argument in (3.4), we have $|I_4^y| \leq \frac{\varepsilon}{8}$ and thus

$$|\mathbf{E}^x[\varphi(X_t)] - \mathbf{E}^y[\varphi(X_t)]| \leq |I_1^x| + |I_1^y| + |I_4^x| + |I_4^y| + |I_3^x - I_3^y| \leq \varepsilon. \tag{3.6}$$

where the term $|I_3^x - I_3^y|$ is small because of (3.2) and since ε is arbitrary, the strong Feller property for X_t is valid. □

3.2 Weak Convergence

3.2.1 Convergence of the marginal distribution

Now we prove the marginal distribution convergence (2.9). For any $t > 0$, let $F^\delta(x, t) = P(X_t^\delta \leq x)$ denote the cumulative distribution function (abbreviated by c.d.f.) of X_t^δ, and $F(x, t)$ and $f(x, t)$ are the c.d.f. and p.d.f. of X_t respectively. In [29], we let n_t denote the counting process of jumping times of X_t and T_n be its n-th jumping time, with which we decompose $F(x, t)$ as the summation of sub-c.d.f. $F_n(x, t) = \mathbf{P}(X_t \leq x, n_t = n)$. Similarly for X_t^δ, define

$$n_t^\delta = \left| \{s : s \leq t, X_s^\delta \neq X_{s-}^\delta\} \right| \tag{3.7}$$

to be the counting process which denotes the number of jumping times before t. And for each $n \geq 1$, define the stopping times:

$$T_n^\delta = \inf\{t \geq 0 : n_t^\delta = n\}. \tag{3.8}$$

Let $F_{T_n^\delta}(t)$ and $f_{T_n^\delta}(t)$ be the c.d.f and p.d.f of T_n^δ respectively. Moreover, for each $n \geq 0$, we also define:

$$F_n^\delta(x, t) = \mathbf{P}\left(X_t^\delta \leq x, n_t^\delta = n\right). \tag{3.9}$$

Using similar arguments as in section 2 of [29], we have the following relationship and the exponential decay of $F_n^\delta(x, t)$ with respect to n.

Proposition 3.1. *For all $n \geq 1$,*

$$F_n^\delta(x, t) = \int_0^t F_{n-1}^\delta(x, t - s) dF_{T_1^\delta}(s) \tag{3.10}$$

$$F_{T_n^\delta}(t) = \int_0^t F_{T_{n-1}^\delta}(t - s) dF_{T_1^\delta}(s), \tag{3.11}$$

and there is a $\theta > 0$ such that for any $T \in (0, \infty)$,

$$F_n^\delta(x, t) \leq \exp(-\theta n + T) \tag{3.12}$$

for all $t \leq T$ and $x \in (-\infty, 1]$.

With the exponential decay of $F_n^\delta(x, t)$ with respect to n, we know that $F^\delta(x, t)$ is absolutely continuous with respect to the Lebesgue measure and use $f^\delta(x, t)$ to denote the p.d.f..

Before the discussion of technical details, we first outline the major steps as follows.

(i) We use the technique of coupling to compare the difference between $F_0^\delta(x, t)$ and $F_0(x, t)$ together with $F_{T_1^\delta}(t)$ and $F_{T_1}(t)$.

(ii) We prove the uniform continuity of $F_n(x,t)$ by the regularity of $F_0(x,t)$ and the iteration approach.

(iii) With the uniform continuity, parallel to step (i), we estimate the difference between $F_n(x,t)$ and $F_n^\delta(x,t)$.

(iv) With the exponential decay property of both $F_n(x,t)$ and $F_n^\delta(x,t)$, we complete the proof. In fact, we can decompose the difference between $F(x,t)$ and $F^\delta(x,t)$ into two terms. The first term is small because of the argument in step (iii), and the second term is small due to the exponential decay property.

We first state the following result for the c.d.f $F_{T_1}(t)$ for T_1; the proof can be found in section 2 of [29].

Proposition 3.2. *For any fixed $T > 0$, $F_{T_1}(\cdot)$ is uniformly continuous on $[0,T]$, i.e., $\forall \varepsilon > 0$, there exists $\eta_1 = \eta_1(T) > 0$, s.t. $\forall t, t' \in [0,T]$, $|t - t'| \leq \eta_1$, we have*

$$|F_{T_1}(t) - F_{T_1}(t')| < \varepsilon.$$

Now we compare the difference between $F_0^\delta(x,t)$ and $F_0(x,t)$ together with $F_{T_1^\delta}(t)$ and $F_{T_1}(t)$.

Lemma 3.3. *Fix any $T > 0$ and for any $\varepsilon > 0$, there is an $\eta_0 > 0$ such that for all $\delta \in (0, \eta_0]$ and all $(x,t) \in \mathbb{R} \times [0,T]$,*

$$\left| F_0^\delta(x,t) - F_0(x,t) \right| < \varepsilon. \tag{3.13}$$

At the same time, we have:

$$\left| F_{T_1}(t) - F_{T_1^\delta}(t) \right| < \varepsilon. \tag{3.14}$$

Proof. Noting that for any $y \in \mathbb{R}$,

$$Z_t^y = e^{-t}y + \sqrt{2}e^{-t}\int_0^t e^s dB_s \tag{3.15}$$

is an O-U process starting from y, we couple two stochastic processes $X_{t \wedge T_1}$ and $X_{t \wedge T_1^\delta}^\delta$ as follows:

(i) Let $\{Z_t^0\}_{t \geq 0}$ denote the standard O-U process starting from 0 and let Γ obey the exponential distribution $\exp(1)$ and be independent of the process Z_t^0.

(ii) Consider the following two stopping times:

$$T_1^0 = \inf\{t \geq 0 : Z_t^0 = 1\}, \quad T_1^{0,\delta} = \inf\{t \geq 0 : \int_0^t \lambda^\delta(Z_s^0)ds = \Gamma\}.$$

By definition, $Z_{t \wedge T_1^0}^0$ and $Z_{t \wedge T_1^{0,\delta}}^0$ are identically distributed as $X_{t \wedge T_1}$ and $X_{t \wedge T_1^\delta}^\delta$, and at the same time we have $T_1^0 \leq T_1^{0,\delta}$. Thus one has

$$F_0(x,t) = \mathbf{P}^0(Z_t^0 \leq x, T_1^0 > t) \leq F_0^\delta(x,t) = \mathbf{P}^0(Z_t^0 \leq x, T_1^{0,\delta} > t) \tag{3.16}$$

while

$$F_0^\delta(x,t) - F_0(x,t) \leq \mathbf{P}^0(T_1^0 \leq t, T_1^{0,\delta} > t). \tag{3.17}$$

By the strong Markov property of the O-U process, if we restart Z_t^0 at T_1^0, then $\{Z_{s+T_1^0}^0\}_{s \geq 0}$ forms a new O-U process starting at 1 which is independent of T_1^0. Denote this process by $\hat{Z}_s$. Moreover, defining a new stopping time $\hat{T}_1^{0,\delta} = T_1^{0,\delta} - T_1^0$ with respect to $\{\hat{Z}_s\}_{s \geq 0}$, one may have:

$$\mathbf{P}^0(T_1^0 \leq t, T_1^{0,\delta} > t) = \int_0^t \mathbf{P}^1(\hat{T}_1^{0,\delta} > t - s)dF_{T_1^0}(s). \tag{3.18}$$

As we have previously seen in Proposition 3.2, $F_{T_1}(s)$ is uniformly continuous on $[0, T]$ and note that T_1 and T_1^0 have the same distribution. Thus $\forall \, \varepsilon > 0$, $\exists \, \eta_1 = \eta_1(T) > 0$ s.t. $\forall s \geq 0$, we have $F_{T_1}(s + \eta_1) - F_{T_1}(s) < \varepsilon$. Then for $t < \eta_1$, we have

$$\mathbf{P}^0(T_1^0 \leq t, T_1^{0,\delta} > t) \leq \int_0^t dF_{T_1}(s) \leq F_{T_1}(\eta_1) < \epsilon.$$

For $t > \eta_1$, we have

$$\begin{aligned}
\mathbf{P}^0(T_1^0 \leq t, T_1^{0,\delta} > t) &= \int_0^t \mathbf{P}^1(\hat{T}_1^{0,\delta} > t - s) dF_{T_1}(s) \\
&= \int_0^{t-\eta_1} \mathbf{P}^1(\hat{T}_1^{0,\delta} > t - s) dF_{T_1}(s) + \int_{t-\eta_1}^t \mathbf{P}^1(\hat{T}_1^{0,\delta} > t - s) dF_{T_1}(s) \\
&\leq \mathbf{P}^1(\hat{T}_1^{0,\delta} > \eta_1) + (F_{T_1}(t) - F_{T_1}(t - \eta_1)) \\
&\leq \mathbf{P}^1(\hat{T}_1^{0,\delta} > \eta_1) + \varepsilon.
\end{aligned} \tag{3.19}$$

Hence, it suffices to prove that for any fixed $\eta_1 > 0$,

$$\lim_{\delta \to 0} \mathbf{P}^1(\hat{T}_1^{0,\delta} > \eta_1) = 0. \tag{3.20}$$

For any $\eta_1 > 0$ and $\delta > 0$, consider the following random subset generated by $\hat{Z}_s$ which denotes the time that $\hat{Z}_s$ is above the level $1 + \delta$ before η_1:

$$\hat{I}(\eta_1, \delta) := \{s < \eta_1 : \hat{Z}_s > 1 + \delta\}. \tag{3.21}$$

By definition, one may see that

$$\begin{aligned}
\left\{ \hat{T}_1^{0,\delta} \leq \eta_1 \right\} &\supset \left\{ \int_{T_1^0}^{T_1^0 + \eta_1} \lambda^\delta(Z_s^0) ds \geq \Gamma \right\} \\
&= \left\{ \int_{\{s < \eta_1 : \hat{X}_s \in [1, 1+\delta]\}} \lambda^\delta(\hat{Z}_s) ds + \int_{\{s < \eta_1 : \hat{Z}_s > 1+\delta\}} \lambda^\delta(\hat{Z}_s) ds \geq \Gamma \right\} \\
&\supset \left\{ \int_{\{s < \eta_1 : \hat{Z}_s > 1+\delta\}} \lambda^\delta(\hat{Z}_s) ds \geq \Gamma \right\} = \left\{ \frac{\mathfrak{L}\left(\hat{I}(\eta_1, \delta) \right)}{\delta} \geq \Gamma \right\}
\end{aligned} \tag{3.22}$$

Thus,

$$\mathbf{P}^1(\hat{T}_1^{0,\delta} \leq \eta_1) \geq \mathbf{P}^1\left(\frac{\mathfrak{L}\left(\hat{I}(\eta_1, \delta) \right)}{\delta} \geq \Gamma \right) \tag{3.23}$$

where $\mathfrak{L}$ denotes the Lebesgue measure in $\mathbb{R}$. Recall that $\Gamma \sim \exp(1)$. Thus it suffices to prove that

$$\lim_{\delta \to 0} \mathbf{P}^1\left(\frac{\mathfrak{L}\left(\hat{I}(\eta_1, \delta) \right)}{\delta} \geq \Gamma \right) = 1.$$

First we consider the case $\delta = 0$. With (3.15), we know that

$$\hat{Z}_t = e^{-t} + \sqrt{2} \int_0^t e^{-(t-s)} dB_s$$

Moreover, by the pathwise continuity of $\hat{X}_t$, one may see that $\hat{I}(\eta_1, 0)$ is a.s. either an empty set or an nonempty open set. We first show it is a.s. nonempty. Letting $\widetilde{T_1} = \inf\{t > 0 : \hat{Z}_t > 1\}$, it suffices to prove that

$$\mathbf{P}^1(\widetilde{T_1} = 0) = 1. \tag{3.24}$$

The proof of (3.24) relies on the $0 - 1$ Law for standard Brownian motion. See Theorem 7.2.3 on Page 362 of [19] for details.

(i) For any $\Delta t > 0$, $\{\widetilde{T_1} < \Delta t\} \supset \{\hat{Z}_{\Delta t} > 1\} = \{e^{-\Delta t} - \sqrt{2} \int_0^{\Delta t} e^{-(\Delta t - s)} dB_s > 1\} \in \mathcal{F}_{\Delta t}$, where $\mathcal{F}$ is the natural filtration generated by $\hat{Z}_s$.

(ii) Thus for $\Delta t_n \to 0$, $\{\widetilde{T_1} = 0\} \supset \{\hat{X}_{\Delta t_n} > 1, \text{i.o.}\} \in \mathcal{F}_{0+}^B$, where $\mathcal{F}_{0+}^B$ is the infinitesimal increment σ-field of Brownian motion $\{B_t\}_{t \geq 0}$ and i.o. stands for infinitely often.

(iii) By the $0 - 1$ Law, we now only need to prove that $\mathbf{P}^1(\hat{Z}_{\Delta t_n} > 1, \text{i.o.}) > 0$. At the same time, with (3.15) we have

$$\mathbf{P}^1(\hat{X}_{\Delta t_n} > 1) = \mathbf{P}^1\left(N(0,1) > \frac{1 - e^{\Delta t_n}}{\sqrt{1 - e^{-2\Delta t_n}}}\right). \tag{3.25}$$

Now noting that $\frac{1 - e^{\Delta t_n}}{\sqrt{1 - e^{-2\Delta t_n}}} = O(\Delta t_n) \to 0$ as $n \to \infty$, we have $\lim_{n \to \infty} \mathbf{P}^1(\hat{Z}_{\Delta t_n} > 1) = \frac{1}{2} > 0$. Thus we have proved (3.24) and hence

$$\mathbf{P}^1(\mathfrak{L}(\hat{I}(\eta_1, 0)) > 0) = 1. \tag{3.26}$$

Note that events $\{\mathfrak{L}(\hat{I}(\eta_1, \frac{1}{n})) > \frac{1}{n}\Gamma\} \to \{\mathfrak{L}(\hat{I}(\eta_1, 0)) > 0\}$ as $n \to \infty$. Thus for any $\varepsilon > 0$, $\exists N$ s.t. for all $n \geq N$

$$\mathbf{P}^1\left(\mathfrak{L}(\hat{I}(\delta_1, \frac{1}{n})) > \frac{1}{n}\Gamma\right) \geq 1 - \varepsilon. \tag{3.27}$$

Fixing any $\delta \leq \frac{1}{N}$ and recalling that ε is arbitrary, together with (3.23), we get (3.20). Combining (3.20), (3.17) and (3.19), when δ is small we have

$$F_0^\delta(x, t) - F_0(x, t) \leq \mathbf{P}^0(T_1^0 \leq t, T_1^{0,\delta} > t) \leq \mathbf{P}^1(\hat{T}_1^{0,\delta} > \eta_1) + \varepsilon \leq 2\varepsilon.$$

and thus the proof of Lemma 3.3 is complete.

□

Before proceeding with the iterative argument, we need to iteratively derive the uniform continuity of $F_n(x, t)$. Recalling Proposition 3.1 of [29], we immediately get the following essential uniform continuity of $F_0(x, t)$.

Proposition 3.4. *Fix any $T > 0$ and for the $\eta_1 = \eta_1(T) > 0$ in Proposition 3.2, there exists $\eta_2 = \eta_2(T) \in (0, \eta_1)$ such that for all $x \in \mathbb{R}$ and $t, t' \in [\eta_1, T]$, $|t' - t| < \eta_2$, we have*

$$|F_0(x, t) - F_0(x, t')| < \varepsilon. \tag{3.28}$$

Next for $F_1(x, t)$, by Lemma 2.2 of [29], we have

$$F_1(x, t) = \int_0^t F_0(x, t - s) dF_{T_1}(s),$$

$$F_1(x, t') = \int_0^{t'} F_0(x, t' - s) dF_{T_1}(s).$$

Now we prove:

Corollary 3.5. *Fix any $T > 0$ and recall the definition of η_1, η_2 in Proposition 3.2-3.4. For all $0 \leq t < t' < T$ such that $t' - t < \eta_2$, and any $x \in \mathbb{R}$, we always have*

$$|F_1(x,t) - F_1(x,t')| \leq 3\epsilon. \tag{3.29}$$

Remark 3.6. Here we no longer need t, t' to be away from 0.

Proof. First, supposing $t \in [\eta_1, T]$, we may write:

$$F_1(x,t) = \int_0^{t-\eta_1} F_0(x, t-s) dF_{T_1}(s) + \int_{t-\eta_1}^t F_0(x, t-s) dF_{T_1}(s)$$
$$=: A_1 + A_2,$$

while

$$F_1(x,t') = \int_0^{t-\eta_1} F_0(x, t'-s) dF_{T_1}(s) + \int_{t-\eta_1}^{t'} F_0(x, t'-s) dF_{T_1}(s)$$
$$=: B_1 + B_2.$$

Using the uniform continuity of $F_{T_1}(t)$ and since $F_0(x,t) \leq 1$, we have

$$\begin{cases} A_2 \leq \int_{t-\eta_1}^t dF_{T_1}(s) \leq \varepsilon \\ B_2 \leq \int_{t-\eta_1}^{t'-\eta_1} dF_{T_1}(s) + \int_{t'-\eta_1}^{t'} dF_{T_1}(s) \leq 2\varepsilon, \end{cases}$$

which together with Proposition 3.4 imply that

$$|F_1(x,t) - F_1(x,t')| \leq |A_1 - B_1| + |A_2 - B_2|$$
$$\leq \int_0^{t-\eta_1} |F_0(x, t'-s) - F_0(x, t-s)| \, dF_{T_1}(s) + 2\varepsilon$$
$$\leq \int_0^{t-\eta_1} \varepsilon dF_{T_1}(s) + 2\varepsilon \leq 3\varepsilon.$$

When $t \leq \eta_1$, we have

$$F_1(x,t) = \int_0^t F_0(x, t-s) dF_{T_1}(s) \leq \int_0^{\eta_1} dF_{T_1}(s) = F_{T_1}(\eta_1) \leq \varepsilon.$$

And note that $\eta_2 \leq \eta_1$, while $t' - t \in (0, \eta_2)$

$$F_1(x,t') \leq \int_0^t dF_{T_1}(s) + \int_t^{t'} dF_{T_1}(s) \leq 2\varepsilon.$$

Thus $|F_1(x,t) - F_1(x,t')| \leq 2\varepsilon$.

$\square$

Similarly, one may inductively prove:

Corollary 3.7. *Fix any $T > 0$ and for all $n \geq 1$, any $x \in \mathbb{R}$ and all $0 \leq t \leq t' < T$ such that $t' - t < \eta_2$, we have*

$$|F_n(x,t) - F_n(x,t')| \leq (2n+1)\varepsilon.$$

Now with the uniform continuity of $F_n(x,t)$, we can continue the proof of Theorem 2.4. First, parallel to the proof of Lemma 3.3, we can consider the difference between $F_1(x,t)$ and $F_1^\delta(x,t)$. Specifically, we have the following lemma.

Lemma 3.8. *Fix any $T > 0$ and recall the definition of δ in Lemma 3.3. We have for any $(x, t) \in \mathbb{R} \times [0, T]$, $|F_1(x, t) - F_1^\delta(x, t)| \leq 5\varepsilon$.*

Proof. Note that

$$\begin{cases} F_1(x, t) = \int_0^t F_0(x, t - s) dF_{T_1}(s), \\ F_1^\delta(x, t) = \int_0^t F_0^\delta(x, t - s) dF_{T_1^\delta}(s). \end{cases}$$

For any $t \in [0, T]$, we introduce the intermediate term:

$$\int_0^t F_0(x, t - s) dF_{T_1^\delta}(s).$$

Recalling (3.13) and (3.14) together with the fact that $F_0^\delta, F_0 \in [0, 1]$, we have

$$\left| F_1(x, t) - F_1^\delta(x, t) \right|$$

$$\leq \int_0^t \left| F_0^\delta(x, t - s) - F_0(x, t - s) \right| dF_{T_1^\delta}(s) + \left| \int_0^t F_0(x, t - s) d\left(F_{T_1}(s) - F_{T_1^\delta}(s) \right) \right| \quad (3.30)$$

$$\leq \varepsilon F_{T_1^\delta}(t) + \left| \mathbf{E}^0 \left[F_0(x, t - T_1^0) \mathbb{1}_{\{T_1^0 \leq t\}} \right] - \mathbf{E}^0 \left[F_0(x, t - T_1^{0,\delta}) \mathbb{1}_{\{T_1^{0,\delta} \leq t\}} \right] \right|.$$

Now recall the definition of η_2 in Proposition 3.4 and consider the following "good event"

$$G := \{ T_1^0 \leq t - \eta_1 - \eta_2, T_1^{0,\delta} - T_1^0 \leq \eta_2 \}. \quad (3.31)$$

Recalling Proposition 3.2-3.4 and (3.20), we have

$$\left| \mathbf{E}^0 \left[F_0(x, t - T_1^0) \mathbb{1}_{\{T_1^0 \leq t\}} \right] - \mathbf{E}^0 \left[F_0(x, t - T_1^{0,\delta}) \mathbb{1}_{\{T_1^{0,\delta} \leq t\}} \right] \right|$$

$$\leq \mathbf{E}^0 \left[\left| F_0(x, t - T_1^0) - F_0(x, t - T_1^{0,\delta}) \right| \cdot \mathbb{1}_G \right] + \mathbf{P}^0(G^C \cap \{T_1^0 \leq t\}) \quad (3.32)$$

$$\leq \varepsilon + \mathbf{P}^0(T_1^0 \in (t - \eta_1 - \eta_2, t]) + \mathbf{P}^0(T_1^{0,\delta} - T_1^0 > \eta_2) \leq 4\varepsilon.$$

Combining with (3.30), we complete the proof.

$\square$

Now one may inductively prove

Lemma 3.9. *Fix any $T > 0$ and recall the definition of δ in Lemma 3.3. We have for all $n \geq 1$ and any $x \in \mathbb{R} \times [0, T]$, $|F_n(x, t) - F_n^\delta(x, t)| \leq (n + 2)^2 \varepsilon$.*

Proof. By Lemma 3.3 and Lemma 3.8, the result has been shown to be true for $n = 0$ and 1. Now suppose Proposition 1 holds for all $k \leq n - 1$. Now for $k = n$, we have:

$$\begin{cases} F_n(x, t) = \int_0^t F_{n-1}(x, t - s) dF_{T_1}(s), \\ F_n^\delta(x, t) = \int_0^t F_{n-1}^\delta(x, t - s) dF_{T_1^\delta}(s). \end{cases} \quad (3.33)$$

Again there is

$$\left| F_n(x, t) - F_n^\delta(x, t) \right| \leq \int_0^t \left| F_{n-1}^\delta(x, t - s) - F_{n-1}(x, t - s) \right| dF_{T_1^\delta}(s)$$

$$+ \left| \mathbf{E}^0 \left[F_{n-1}(x, t - T_1^0) \mathbb{1}_{\{T_1^0 \leq t\}} \right] - \mathbf{E}^0 \left[F_{n-1}(x, t - T_1^{0,\delta}) \mathbb{1}_{\{T_1^{0,\delta} \leq t\}} \right] \right|.$$

$$(3.34)$$

Recall the "good event" $G = \{T_1^0 \leq t - \eta_2, T_1^{0,\delta} - T_1^0 \leq \eta_2\}$. We have

$$
\begin{aligned}
&\left| \mathbf{E}^0 \left[F_{n-1}(x, t - T_1^0) \mathbb{1}_{\{T_1^0 \leq t\}} \right] - \mathbf{E}^0 \left[F_{n-1}(x, t - T_1^{0,\delta}) \mathbb{1}_{\{T_1^{0,\delta} \leq t\}} \right] \right| \\
\leq & \mathbf{E}^0 \left[\left| F_{n-1}(x, t - T_1^0) - F_{n-1}(x, t - T_1^{0,\delta}) \right| \cdot \mathbb{1}_G \right] + \mathbf{P}^0(G^C \cap \{T_1^0 \leq t\}) \\
\leq & (2n-1)\varepsilon + \mathbf{P}^0(T_1^0 \in (t - \eta_2, t]) + \mathbf{P}^0(T_1^{0,\delta} - T_1^0 > \eta_2) \\
\leq & (2n+1)\varepsilon.
\end{aligned}
\tag{3.35}
$$

Thus we have

$$
\left| F_n(x, t) - F_n^\delta(x, t) \right| \leq [(n+1)^2 + (2n+1)]\varepsilon \leq (n+2)^2 \varepsilon.
$$

$\square$

Finally, for all $T < \infty$, and any $t \in [0, T]$, $x \leq 1$,

$$
F_n(x, t) \leq \mathbf{P}^0(T_n \leq t), \quad F_n^\delta(x, t) \leq \mathbf{P}^0(T_n^\delta \leq t).
$$

By the argument in Lemma 2.3 of [29], we have already implied that $\exists$ a constant C depending only on T such that

$$
\mathbf{P}^0(T_n^\delta \leq t) \leq \mathbf{P}^0(T_n \leq t) \leq \exp(-Cn).
\tag{3.36}
$$

What's more, we can decompose the difference between $F(x, t)$ and $F^\delta(x, t)$ into two terms. That is,

$$
\left| F(x, t) - F^\delta(x, t) \right| \leq \sum_{i=0}^{n} \left| F_i(x, t) - F_i^\delta(x, t) \right| + \sum_{i=n+1}^{+\infty} \left| F_i(x, t) - F_i^\delta(x, t) \right|.
\tag{3.37}
$$

Now using Lemma 3.9 we know that the first term of (3.37) is small, while the second term is small due to the exponential decay property (3.36). Thus the proof of (2.9) is complete.

3.2.2 Weak convergence in the sense of process

Now we can prove weak convergence in the sense of stochastic process $X^\delta_{\cdot} \to X_{\cdot}$ as $\delta \to 0^+$. We first prove the convergence of the finite dimensional marginal distribution and then use the relative compactness to conclude weak convergence for process.

Proposition 3.10. *For all integer $n > 0$, $0 < t_1 < t_2 < \cdots < t_n < \infty$ and $-\infty < a_i < b_i < 1$, $i = 1, 2, \cdots, n$, we have*

$$
\mathbf{P}^0(X^\delta_{t_i} \in (a_i, b_i], i = 1, 2, \cdots, n) \to \mathbf{P}^0(X_{t_i} \in (a_i, b_i], i = 1, 2, \cdots, n) \quad as \quad \delta \to 0^+,
\tag{3.38}
$$

i.e.,

$$
(X^\delta_{t_1}, X^\delta_{t_2}, \cdots, X^\delta_{t_n}) \to (X_{t_1}, X_{t_2} \cdots, X_{t_n}).
\tag{3.39}
$$

where $\mathbf{P}^0(\cdot)$ denotes the process starts from 0.

When $n = 1$, the result for X_t, X_t^δ starting from 0 has been proved in Section 3.2.1. By the same proof we have convergence with respect to any $X_0^\delta = X_0 = x_0$, where x_0 belongs to some compact subset of $(-\infty, 1)$. Moreover, the δ can be chosen to be uniform, i.e.,

Corollary 3.11. *For any $-\infty < a < b < 1$, and any $x_0 \in [a, b]$, we let $F^{\delta, x_0}(x, t)$ and $F^{x_0}(x, t)$ be the c.d.f. of X_t^{δ} and X_t starting from x_0. Then for any $T < \infty$ and $\epsilon > 0$, there exists $\eta_0 = \eta(T, a, b, \epsilon) > 0$ s.t. for all $0 < \delta < \eta_0$, $x_0 \in [a, b]$, $(x, t) \in \mathbb{R} \times [0, T]$, we have*

$$|F^{\delta, x_0}(x, t) - F^{x_0}(x, t)| \le \epsilon. \tag{3.40}$$

Thus we conclude the case for $n = 1$ and then by the strong Feller property for X_t and induction, we can get Proposition 3.10; the detailed proof can be found in Appendix A.

With the finite dimensional weak convergence shown as above, the rest of the proof of weak convergence for processes follows from a standard relative compactness argument. Note that $(\mathbb{R}, \mathcal{B})$ is clearly complete and separable. By Theorem 3.7.2 and Theorem 3.7.8 of [21], in order to show that $X^{\delta}_{\cdot} \to X_{\cdot}$, it suffices to prove the following:
(i) For any $T < \infty$ and $\epsilon > 0$, $\exists$ a compact set $\Gamma \subset (-\infty, +\infty)$ such that for any $\delta > 0$, we have

$$\mathbf{P}^0(X_t^{\delta} \in \Gamma, \forall t \in [0, T]) \ge 1 - \epsilon. \tag{3.41}$$

(ii) For any $T < \infty$ and $\epsilon > 0$, $\exists \sigma > 0$ such that for any $\delta > 0$, we have

$$\mathbf{P}^0(w'(X_{\cdot}^{\delta}, \sigma, T) \ge \epsilon) \le \epsilon. \tag{3.42}$$

where

$$w'(X_{\cdot}^{\delta}, \sigma, T) = \inf_{\{t_i\}} \max_i \sup_{s, t \in [t_{i-1}, t_i)} |X_s^{\delta} - X_t^{\delta}|$$

and $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$ with $t_i - t_{i-1} > \sigma$ for all i. See (6.2) in Chapter 3 of [21] for details.

In order to verify the conditions above, we first need to recall the constructions in Section 2.2 of [29]. For claim (i), note that $T_n \le T_n^{\delta}$ for all $\delta > 0$ and by exponential decay, we have for any fixed $T < \infty$ and $\epsilon > 0$, $\exists n_0$, s.t.

$$\mathbf{P}^0(T_{n_0}^{\delta} \le T) \le \mathbf{P}^0(T_{n_0} \le T) < \frac{\epsilon}{2}.$$

Note that by Doob's inequality, we have that there is an $M_0 < \infty$ s.t.

$$\mathbf{P}^0(\max_{t \le T} |Y_t^{(i)}| > M_0) \le \mathbf{P}^0(\max_{t \le T} |\int_0^t e^s dB_s| > M_0) < \frac{\epsilon}{2n_0}.$$

Now consider $\Gamma = [-M_0, M_0]$ and event $A = \{T_{n_0}^{\delta} > T\} \cap \cup_{i=1}^{n_0} \{\max_{t \le T} |Y_t^{(i)}| < M_0\}$, where $\mathbf{P}^0(A) > 1 - \frac{\epsilon}{2} - n_0 \cdot \frac{\epsilon}{2n_0} = 1 - \epsilon$. Then recalling that the trajectory of X_t^{δ} can be decomposed by parts of $Y_t^{(i)}, t \in [0, T]$, thus in event A for any $\delta > 0, t \in [0, T]$, X_t^{δ} belongs to Γ, which gives (3.41).

Finally, in order to check claim (ii), we first note that by Proposition 3.1 of [29], $f_{T_1}(\cdot) \in C[0, T]$ with $f_{T_1}(0) = 0$. Thus for all $\epsilon > 0$, $\exists n_0 < \infty$, s.t. $f_{T_1}(t) < \frac{\epsilon}{2T}$ for all $t \le \frac{T}{n_0}$. Then consider the event $B_1 = \cap_{i=1}^{n_0} \{\tau_i > \frac{T}{n_0}\}$, where

$$\mathbf{P}^0(B_1^c) \le n_0 \mathbf{P}^0(\tau_1 \le \frac{T}{n_0}) \le n_0 \cdot \int_0^{n_0} \frac{\epsilon}{2T} dt < \frac{\epsilon}{2}. \tag{3.43}$$

Moreover, note that the OU-process $Y_t^{(i)}$ is a.s. uniformly continuous. Thus $\exists \sigma_1 > 0$ s.t. for each i,

$$\mathbf{P}^0(\exists s < t \in [0, T], t - s < \sigma_1, |Y_t^{(i)} - Y_s^{(i)}| \ge \epsilon) < \frac{\epsilon}{2n_0}. \tag{3.44}$$

Then consider event

$$B_2 = \cap_{i=1}^{n_0} (\forall s < t \in [0, T], t - s < \sigma_1, |Y_t^{(i)} - Y_s^{(i)}| < \epsilon),$$

and $\mathbf{P}^0(B_2^c) \leq n_0 \cdot \frac{\epsilon}{2n_0} \leq \frac{\epsilon}{2}$. Thus let $\sigma = \min\{\frac{\sigma_1}{3}, \frac{T}{2n_0}\}$ and $B = B_1 \cap B_2$. Then in the event B, the jump-diffusion process $\{X_t^\delta\}_{t \in [0,T]}$ is composed of at most n_0 uniformly continuous O-U process each with length at least $\frac{T}{n_0} \geq \sigma$. Thus one may always construct a partition with

$$\max_i \sup_{s,t \in [t_{i-1}, t_i)} |X^\delta(s) - X^\delta(t)| < \epsilon.$$

So we have

$$\mathbf{P}^0(w'(X_\cdot^\delta, \sigma, T) \leq \epsilon) \geq \mathbf{P}^0(B) > 1 - \epsilon,$$

which gives (3.42) and thus the proof of the weak convergence for process is complete.

3.3 Weak Convergence for the Mean Firing Rate

In this section we consider the convergence of the mean firing rate by the iteration approach. Clearly, the density function $f^\delta(x, t)$ for the jump-diffusion process X_t^δ in (2.3) and the mean firing rate $N^\delta(t) = \int_{\mathbb{R}} f^\delta(y, t)\lambda^\delta(y)dy$ admit the following expansions

$$f^\delta(x, t) = \sum_{n=0}^{\infty} f_n^\delta(x, t), \quad N^\delta(t) = \sum_{n=0}^{\infty} N_n^\delta(t). \tag{3.45}$$

where $f_n^\delta(x, t)$ is the density function of the measure induced by $F_n^\delta(\cdot, t)$ in (3.9) and for $n \geq 0$,

$$N_n^\delta(t) := \int_{\mathbb{R}} f_n^\delta(y, t)\lambda^\delta(y)dy.$$

To prove (2.7), we need to build a connection between $N_n^\delta(t)$ and the p.d.f. $f_{T_n^\delta}(t)$ of the jumping time in (3.8). We first derive the Dynkin's formula for the killed process $\widetilde{X_t^\delta}$ that is obtained by stopping the process X_t^δ at the first jumping time T_1^δ. To be specific,

$$\widetilde{X_t^\delta} = \begin{cases} X_t^\delta, & t < T_1^\delta, \\ X_{T_1^\delta}^\delta, & t \geq T_1^\delta, \end{cases} \tag{3.46}$$

First, we derive the Fokker-Planck equation for f_0^δ and its decay property for further iteration calculations.

Lemma 3.12. *Let $f_0^\delta(x, t)$ be the density of the measure induced by $F_0^\delta(\cdot, t)$ defined in (3.9). Then it is the classical solution of the following equation*

$$\begin{cases} \dfrac{\partial f_0^\delta}{\partial t} - \dfrac{\partial}{\partial x}\left(x f_0^\delta\right) - \dfrac{\partial^2 f_0^\delta}{\partial x^2} = -\lambda^\delta(x) f_0^\delta(x, t), & x \in \mathbb{R}, \quad t > 0, \\ f_0^\delta(x, 0) = \delta(x) \text{ in } \mathcal{D}'(-\infty, +\infty), \end{cases} \tag{3.47}$$

where $\delta(x)$ denotes the Dirac function. Moreover, for any $t > 0$, and $|x|$ sufficiently large, one should have $\exists\, C > 0$ s.t.

$$\max\left\{|f_0^\delta(x, t)|, \quad \left|\dfrac{\partial}{\partial x} f_0^\delta(x, t)\right|, \quad \left|\dfrac{\partial^2}{\partial x^2} f_0^\delta(x, t)\right|\right\} \leq \exp(-Cx^2), \tag{3.48}$$

Remark 3.13. The proof follows the standard argument as in [16].

Proof. Recall in (3.15) that we use Z_t^0 to denote an O-U process starting from 0. The proof of (3.48) follows the same argument as in Lemma 3.1 of [29]. For any fixed $T > 0$, according to Theorem 3.5 in Chapter V of [25] by Garroni and Menaldi, there exists a unique Green's function $G : \mathbb{R} \times [0,T] \times \mathbb{R} \times [0,T] \ni (y,s,x,t) \mapsto G(y,s,x,t)$ for the parabolic operator

$$\mathcal{L}_y = -y\partial_y \cdot + \partial_{yy}^2 \cdot - \lambda^\delta \cdot,$$

That is for a given $(x,t) \in \mathbb{R} \times [0,T]$, the function $\mathbb{R} \times [0,t) \ni (y,s) \mapsto G(y,s,x,t)$ is a solution of the PDE

$$\begin{cases} \partial_s G(y,s,x,t) + \mathcal{L}_y G(y,s,x,t) = 0, & (y,s) \in \mathbb{R} \times [0,t), \\ G(y,t,x,t) = \delta(y-x) \text{ in } \mathcal{D}'(\mathbb{R}) \end{cases} \tag{3.49}$$

Following Theorem 5 in Chap. 9 of [24], for any given $(y,s) \in \mathbb{R} \times [0,T)$, the function $\mathbb{R} \times (s,T] \ni (x,t) \mapsto G(y,s,x,t)$ is also known to be Green's function of the adjoint operator

$$\mathcal{L}_x^* = \partial_x[x\cdot] + \partial_{xx}^2 \cdot - \lambda^\delta.$$

i.e.

$$\begin{cases} \partial_t G(y,s,x,t) = \mathcal{L}_x^* G(y,s,x,t), & (x,t) \in \mathbb{R} \times (s,T], \\ G(y,s,x,s) = \delta(x-y) \text{ in } \mathcal{D}'(\mathbb{R}), \end{cases} \tag{3.50}$$

Morever, it belongs to $C^{2,1}$ in x,t and satisfies the following estimate:

$$\left| \partial^\ell G(y,s,x,t) \right| \leq C(t-s)^{-\frac{1+\ell}{2}} \exp\left(-C_0 \frac{(x-y)^2}{t-s} \right), \quad 0 \leq s < t \leq T. \tag{3.51}$$

where $\ell = 0,1,2$, $\partial^\ell = \partial_{xt}^\ell = \partial_x^m \partial_t^n$, $\ell = 2m + n$, for $m, n \in \mathbb{N}_0$.

Thus given any smooth test function $\phi : \mathbb{R} \times [0,T]$ with compact support, the PDE problem

$$\begin{cases} \partial_s u(y,s) = y\partial_y u(y,s) - \partial_{yy} u(y,s) + \lambda^\delta(y)u(y,s) - \phi(y,s), & (y,s) \in \mathbb{R} \times [0,T), \\ u(y,T) = 0 \quad y \in \mathbb{R} \end{cases} \tag{3.52}$$

admits a unique classical solution

$$u(y,s) = \int_s^T \int_{-\infty}^{+\infty} G(y,s,x,t)\phi(x,t)dxdt. \tag{3.53}$$

Set

$$M_t := u(Z_t^0, t) \quad \text{and} \quad N_t := \exp\{-\int_0^t \lambda^\delta(Z_s^0)ds\}$$

— they are both semimartingales. Then by Itô's formula (see Exercise 5.32 on Page 209 of [36] for details), we have

$$d(M_t N_t) = M_t dN_t + N_t dM_t + d\langle M, N\rangle_t. \tag{3.54}$$

Note that $dN_t = -N_t \lambda^\delta(Z_t^0)dt$, thus N_t is of bounded variation and then the quadratic variation $\langle N\rangle_t = 0$. By $\langle M, N\rangle_t \leq \langle M\rangle_t \langle N\rangle_t$, we know the covariance process $\langle M, N\rangle_t$ is equal to 0. Hence

$$d(M_t N_t) = -u(Z_t^0,t)\exp\{-\int_0^t \lambda^\delta(Z_s^0)ds\}\lambda^\delta(Z_t^0)dt + \exp\{-\int_0^t \lambda^\delta(Z_s^0)ds\}$$
$$\cdot \left(\left[u_t(Z_t^0,t) - u_x(Z_t^0,t)Z_t^0 + u_{xx}(Z_t^0,t) \right] dt + \sqrt{2}u_x(Z_t^0,t)dB_t \right). \tag{3.55}$$

With (3.55) and the boundary condition of u in (3.52), we have

$$
\begin{aligned}
0 &= u(Z_T^0, T) \exp(-\int_0^T \lambda^\delta(Z_t^0)dt) \\
&= u(0,0) - \int_0^T u(Z_t^0, t)\lambda^\delta(Z_t^0) \exp(-\int_0^t \lambda^\delta(Z_s^0)ds)dt \\
&\quad + \int_0^T \left[u_t(Z_t^0, t) - xu_x(Z_t^0, t) + u_{xx}(Z_t^0, t) \right] \exp(-\int_0^t \lambda^\delta(Z_s^0)ds)dt \\
&\quad + \sqrt{2}\int_0^T u_x(Z_t^0, t) \exp(-\int_0^t \lambda^\delta(Z_s^0)ds)dB_t.
\end{aligned}
\tag{3.56}
$$

Taking the expectation of (3.56) and recalling (3.52), we have

$$
\begin{aligned}
u(0,0) &= \mathbf{E}^0 \left[\int_0^T \phi(Z_t^0, t) \exp(-\int_0^t \lambda^\delta(Z_s^0)ds)dt \right] \\
&= \int_0^T \mathbf{E}^0 \left[\phi(Z_t^0, t) \exp(-\int_0^t \lambda^\delta(Z_s^0)ds) \right] dt
\end{aligned}
\tag{3.57}
$$

Now applying formula (8.2.10) on Page 139 of [41] with $f(\cdot) = \phi(\cdot, t)$,

$$
\mathbf{E}^0 \left[\phi(Z_t^0, t) \exp(-\int_0^t \lambda^\delta(Z_s^0)ds) \right] = \mathbf{E}^0 \left[\phi(X_t^\delta, t)\mathbb{1}_{\{T_1^\delta > t\}} \right] = \int_{-\infty}^{+\infty} \phi(x, t)f_0^\delta(x, t)dx.
\tag{3.58}
$$

By (3.53) we have

$$
\int_0^T \int_{-\infty}^{+\infty} \phi(x,t)G(0,0,x,t)dxdt = u(0,0) = \int_0^T \int_{-\infty}^{+\infty} \phi(x,t)f_0^\delta(x,t)dxdt,
$$

which implies that $G(0,0,x,t)$ coincides with $f_0^\delta(x,t)$. Thus we conclude that $f_0^\delta(x,t)$ satisfies (3.47) and the decay property (3.48) is valid because of (3.51). $\qquad \square$

Next we can prove

Lemma 3.14. *For any $n \geq 1$ and $t > 0$,*

$$
f_{T_n^\delta}(t) = N_{n-1}^\delta(t).
\tag{3.59}
$$

Proof. We prove (3.59) inductively. First for the case when $n = 1$, with (3.47) and (3.48), one has for any $t > 0$,

$$
\begin{aligned}
f_{T_1^\delta}(t) &= \frac{d}{dt}\mathbf{P}^0(T_1^\delta \leq t) = -\frac{d}{dt}\int_{-\infty}^{+\infty} f_0^\delta(x,t)dx \\
&= -\int_{-\infty}^{+\infty} \frac{d}{dt}f_0^\delta(x,t)dx \\
&= \int_{-\infty}^{+\infty} \left[\lambda^\delta(x)f_0^\delta(x,t) - \frac{\partial^2 f_0^\delta}{\partial x^2} - \frac{\partial}{\partial x}\left(xf_0^\delta\right) \right] dx \\
&= \int_{-\infty}^{+\infty} \lambda^\delta(x)f_0^\delta(x,t)dx.
\end{aligned}
\tag{3.60}
$$

Now we assume that (3.59) holds for all $k \leq n$ and note that

$$
f_{T_{n+1}^\delta}(t) = \int_0^t f_{T_n^\delta}(t-s)f_{T_1^\delta}(s)ds = \int_0^t \int_{-\infty}^{+\infty} f_{n-1}^\delta(x, t-s)\lambda^\delta(x)dx f_{T_1^\delta}(s)ds.
$$

By Fubini's formula,

$$f_{T_{n+1}^\delta}(t) = \int_{-\infty}^{+\infty} \int_0^t f_{n-1}^\delta(x, t-s) f_{T_1^\delta}(s) ds \lambda^\delta(x) dx = \int_{-\infty}^{+\infty} f_n^\delta(x, t) \lambda^\delta(x) dx = N_n(t).$$

(3.61)

$\square$

Now we can show the weak convergence of N^δ as $\delta \to 0^+$. More precisely, given $T > 0$, for any smooth test function $\phi(t) \in C_b[0, T]$, we have

$$\left| \int_0^T \phi(t) N^\delta(t) dt - \int_0^T \phi(t) N(t) dt \right| \to 0, \quad \text{as} \quad \delta \to 0^+.$$

(3.62)

Notice that both $N(t)$ and $N^\delta(t)$ have series representations. In light of the following decomposition

$$\left| \int_0^T \phi(t) N^\delta(t) dt - \int_0^T \phi(t) N(t) dt \right|$$

$$\leq \sum_{i=1}^N \left| \int_0^T \phi(t) F'_{T_n^\delta}(t) dt - \int_0^T \phi(t) F'_{T_n}(t) dt \right| + \sum_{i=N+1}^{+\infty} \left| \int_0^T \phi(t) F'_{T_n^\delta}(t) dt - \int_0^T \phi(t) F'_{T_n}(t) dt \right|$$

$$=: I_1 + I_2.$$

Due to the exponential decay property (3.36), $I_2 \to 0$ as $\delta \to 0^+$ and thus it suffices to estimate I_1, noting that ϕ is bounded,

$$I_1 = \sum_{i=1}^N \left| \int_0^T \phi(t) dF_{T_n^\delta}(t) - \int_0^T \phi(t) dF_{T_n}(t) \right|$$

Noting that $T_1^\delta \to T_1$ in distribution and T_n, T_n^δ are the i.i.d. summation of T_1, T_1^δ respectively, we have $T_n^\delta \to T_n$ in distribution as $\delta \to 0^+$ and thus for any n,

$$\lim_{\delta \to 0^+} \left| \int_0^T \phi(t) dF_{T_n^\delta}(t) - \int_0^T \phi(t) dF_{T_n}(t) \right| = 0.$$

Noting that the summation in I_1 is finite, we conclude that $I_1 \to 0$ as $\delta \to 0^+$. Hence, the proof for the weak convergence of N^δ is complete.

3.4 Convergence rate

Finally, we aim to rigorously prove a polynomial-order convergence rate of $F^\delta(x, t) \to F(x, t)$ as $\delta \to 0^+$ as in (2.8). For any sufficiently small $\epsilon > 0$, let $\epsilon_0 = \epsilon^2$ and $n_0 = \frac{1}{C_T} \log(\epsilon_0^{-1})$. By lemma 3.3 and Lemma 3.9, for all $\epsilon_0 > 0$, $\exists \eta_0 = \eta_0(\epsilon_0) > 0$ s.t. $\forall \delta \leq \eta_0$, $n \geq 0$ and $(x, t) \in \mathbb{R} \times [0, T]$,

$$\left| F_n^\delta(x, t) - F_n(x, t) \right| \leq (n+2)^2 \epsilon_0.$$

Moreover, recalling (3.36), for any fixed $T < \infty$, there $\exists C_T > 0$ s.t.

$$\mathbf{P}^0(T_n^\delta \leq t) \leq \mathbf{P}^0(T_n \leq t) \leq \exp(-C_T n).$$

Then for all $\delta < \eta_0(\epsilon_0)$, $t \in [0, T]$ and $x \in \mathbb{R}$,

$$
|F^\delta(x, t) - F(x, t)|
$$
$$
\leq \sum_{n=0}^{n_0-1} |F_n^\delta(x, t) - F_n(x, t)| + 2\mathbf{P}^0(T_{n_0} \leq t) \tag{3.63}
$$
$$
\leq (n_0 + 2)^3 \epsilon_0 + 2\epsilon_0 < \epsilon.
$$

Thus to get the convergence rate of a polynomial-order, we only need to find a lower bound for η_0, which is polynomial with respect to ϵ. In the order to prove the result of interest, it suffices to show the following polynomial-order relationship:

(i) The η_1 and η_2 defined in Proposition 3.2 and Proposition 3.4 are both of a polynomial-order of ϵ.

(ii) The η_0 in Lemma 3.3 is of a polynomial order of η_1 and η_2, and thus also a polynomial-order of ϵ.

Proof of (i): For η_1, we can always set $\eta_1 = \frac{\epsilon}{\max_{t \leq T} f_{T_1}(t) + 1}$ by Proposition 3.2. As for η_2, when $\epsilon < 1$ is sufficiently small, one may let $\eta_2 = \eta_1^3 = \frac{\epsilon^3}{[\max_{t \leq T} f_{T_1}(t) + 1]^3}$ and prove that it satisfies the condition in Proposition 3.4. For any $\eta_1 < t' < t \leq T$, $t - t' \leq \eta_2$, we have that for any $x \in \mathbb{R}$,

$$
|F_0(x, t) - F_0(x, t')| = \left| \int_{-\infty}^x [f_0(y, t) - f_0(y, t')] \, dy \right| \leq \int_{-\infty}^x \int_{t'}^t \left| \frac{df_0}{ds}(y, s) \right| ds dy. \tag{3.64}
$$

By the estimation of the Green's function (38) of [29], $\exists\, C, C_0 < \infty$ that depend only on T s.t. for any $s \in [t', t]$,

$$
\left| \frac{df_0}{dt}(y, s) \right| \leq C s^{-1} \exp\left(-C_0 \frac{y^2}{s} \right) \leq C \eta_1^{-1} \exp\left(-C_0 \frac{y^2}{T} \right).
$$

Thus

$$
|F_0(x, t) - F_0(x, t')| \leq \eta_1^2 \cdot C \int_{-\infty}^1 \exp\left(-C_0 \frac{y^2}{T} \right) dy < \epsilon. \tag{3.65}
$$

$\square$

For (ii), recalling the proof of Lemma 3.3, the choice of η_0 is decided by (3.20), thus with the following lemma we can find a polynomial order of η_1 as a lower bound for η_0.

Proposition 3.15. *For all sufficiently small $\epsilon > 0$, $\exists\, \gamma \in (0, \infty)$ such that $\forall \delta < \epsilon^\gamma$, we have*

$$
\mathbf{P}^1(\hat{T}_1^{0,\delta} > \epsilon) < \epsilon. \tag{3.66}
$$

In order to prove (3.66), recalling (3.22) we have

$$
\left\{ \hat{T}_1^{0,\delta} \leq \epsilon \right\} \supset \left\{ \frac{\mathfrak{L}\left(\hat{I}(\epsilon, \delta) \right)}{\delta} \geq \Gamma \right\}
$$

Thus it suffices to show that for any $\delta < \epsilon^\gamma$

$$
\mathbf{P}^1\left(\frac{\mathfrak{L}\left(\hat{I}(\epsilon, \delta) \right)}{\delta} \geq \Gamma \right) = \mathbf{P}^1\left(\frac{\int_0^\epsilon \mathbb{1}_{\{Z_t^1 > 1 + \delta\}} dt}{\delta} \geq \Gamma \right) \geq 1 - \epsilon, \tag{3.67}
$$

where Z_t^1 denotes a standard O-U process starting from 1 as in (3.15) where Γ is an independent exponential distribution obeying $\exp(1)$.

Now we use a renormalization argument which is standard for Brownian Motion (abbreviated by B.M.) to prove (3.67). We first introduce a sequence of scales as follows: define $\epsilon_n = \epsilon \cdot 2^{-n}$ for all $n \geq 0$ and a decreasing sequence of stopping times,

$$\tau_n = \inf\{t \geq 0, \ |Z_t^1 - 1| = \epsilon_n\}. \tag{3.68}$$

We hereby outline our argument as follows:

1. We first introduce a sequence of geometrically shrinking "boxes", all centered at 1, where the size of box n equals ϵ_n, which is half that of its predecessor.

2. Note that by $t = \epsilon$, an O-U process starting from 1 on average will wander a distance at least $O(\epsilon^{0.5}) \gg \epsilon$ away from 1. So with high probability, the O-U process has already escaped the largest box by time ϵ.

3. By scaling invariance of B.M., i.e. for any $a > 0$, $\frac{1}{\sqrt{a}} B_{at} \overset{d}{=} B_t$, we can prove that for all $n \geq 0$ with at least a uniformly positive probability $p > 0$, an O-U process stays at the right of 1 for some positive fraction of time between τ_n and τ_{n-1} to trigger the Poisson jump under the intensity (2.2), and we say the O-U process "succeeds" in the nth step when such an event happens.

4. We can choose an appropriate constant γ independent of ϵ whose exact value can be found in (3.76), and $n_1 = O(\log \epsilon^{-1})$ s.t. $(1-p)^{n_1} < \frac{\epsilon}{3}$. Then when $\delta < \epsilon^{4\gamma}$ with high probability, any success in step $n \leq n_0$ can trigger our Poisson jump.

5. In order not to trigger the Poisson jump, the O-U process must not jump in all of the first n_1 steps. Thus by the strong Markov property, we know that the probability of not jumping is no larger than the product of these uniform upper bounds $(1-p)^{n_1} < \epsilon/3$.

Now returning to the detailed proof, we firstly show that with high probability $\tau_0 \leq \epsilon$.

Lemma 3.16. *For τ_0 in (3.68), $\exists \alpha > 0$ s.t. $\mathbf{P}^1(\tau_0 > \epsilon) \leq \exp(-\epsilon^\alpha)$ for all sufficiently small $\epsilon > 0$.*

Proof. Recall that

$$Z_t^1 = 1 - \int_0^t Z_s^1 ds + B_t.$$

Let $\bar{\tau}_\epsilon = \inf\{t > 0, \ |B_t| = \epsilon^{\frac{2}{3}}\}$. Noting that $\{\tau_0 > \epsilon\} \subset \{\bar{\tau}_\epsilon > \epsilon\}$ and by the scaling invariance and the Markov property of B.M. B_t,

$$\begin{aligned}
\mathbf{P}^1(\tau_0 > \epsilon) &\leq \mathbf{P}^0(\bar{\tau}_\epsilon > \epsilon) \\
&= \mathbf{P}^0(\bar{\tau}_1 > \epsilon^{-\frac{1}{3}}) \\
&\leq \left(1 - \min_{z \in [-1,1]} \mathbf{P}^z(|B_1| > 1)\right)^{\lfloor \epsilon^{-\frac{1}{3}} \rfloor} \\
&\leq (1 - \mathbf{P}^0(|B_1| > 2))^{\lfloor \epsilon^{-\frac{1}{3}} \rfloor} \\
&\leq \exp\left(-\epsilon^{-\frac{1}{4}}\right)
\end{aligned} \tag{3.69}$$

for all sufficiently small ϵ, where $\lfloor \cdot \rfloor$ denotes the integer part. Thus let $\alpha = \frac{1}{4}$ and the proof is complete. $\square$

Now for any integer $n \geq 1$, we say step n is a "success" if the event

$$A_n := \left\{\int_{\tau_n}^{\tau_{n-1}} \mathbb{1}_{\{Z_t^1 > 1 + \epsilon_n\}} dt > \epsilon_n^2\right\} \tag{3.70}$$

happens. To find a lower bound for the probabilities of success in each step, we first consider the following technical lemma for B.M. at scale of order 1:

Lemma 3.17. *For B_t the standard B.M., let*

$$\Gamma_x = \inf\{t > 0 : B_t = x\}. \tag{3.71}$$

Then

$$\mathbf{P}\left(\Gamma_4 < \Gamma_{-\frac{1}{2}}, \Gamma_4 < 4, \int_0^{\Gamma_{\frac{2}{3}}} \mathbb{1}_{\{B_t \in [\frac{1}{4}, \frac{1}{2}]\}} dt > 1, \int_0^{\Gamma_{2.8}} \mathbb{1}_{\{B_t \in [2.2, 2.5]\}} dt > 1\right) =: p > 0. \tag{3.72}$$

Proof. Note that a standard B.M. can approximate any continuous function starting from 0 with a positive probability. (See Theorem 5.4 on Page 206 of [18] for details.) Thus by easily choosing a continuous function satisfying the condition in the event of (3.72), we conclude the result of interest by noting that Brownian motion can approximate such a continuous function with positive probability. $\square$

Remark 3.18. The seemingly mysterious constants in (3.72) are purposely chosen to meet the later needs in the proof of Lemma 3.20. Particularly, we need these constants to create a certain level of "redundancy" so that after introducing the drift term, our O-U process can still stay within the intervals of interest.

Then by the scaling invariance of B_t, we immediately have for all $\theta > 0$,

$$\mathbf{P}\left(\Gamma_{4\theta} < \Gamma_{-\frac{1}{2}\theta}, \Gamma_{4\theta} < 4\theta^2, \int_0^{\Gamma_{\frac{2}{3}\theta}} \mathbb{1}_{\{B_t \in [\frac{1}{4}\theta, \frac{1}{2}\theta]\}} dt > \theta^2, \int_0^{\Gamma_{2.8\theta}} \mathbb{1}_{\{B_t \in [2.2\theta, 2.5\theta]\}} dt > \theta^2\right) = p. \tag{3.73}$$

Now we return to the probability of A_n. Let $Z_t^{1+\epsilon_n}$ be the O-U process starting from $1 + \epsilon_n$, i.e., $Z_t^{1+\epsilon_n} = 1 + \epsilon_n - \int_0^t Z_s^{1+\epsilon_n} ds + B_t$. Then by the strong Markov property for the O-U process,

$$\mathbf{P}(A_n) \geq \min\left\{\mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1+\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right), \mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1-\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right)\right\}$$

Now letting $\theta = \epsilon_n$, we look at the event

$$E^{\epsilon_n} := \left\{\Gamma_{4\epsilon_n} < \Gamma_{-\frac{1}{2}\epsilon_n}, \Gamma_{4\epsilon_n} < 4\epsilon_n^2, \int_0^{\Gamma_{\frac{2}{3}\epsilon_n}} \mathbb{1}_{\{B_t \in [\frac{1}{4}\epsilon_n, \frac{1}{2}\epsilon_n]\}} dt > \epsilon_n^2, \int_0^{\Gamma_{2.8\epsilon_n}} \mathbb{1}_{\{B_t \in [2.2\epsilon_n, 2.5\epsilon_n]\}} dt > \epsilon_n^2\right\}. \tag{3.74}$$

Lemma 3.19. *Given the event E^{ϵ_n}, we have $Z_t^{1+\epsilon_n} \in [0, 2]$ a.s. $\forall t \in [0, \Gamma_{4\epsilon_n}]$.*

Proof. Otherwise let $\bar{\Gamma}_0$ and $\bar{\Gamma}_2$ be the first time $Z_t^{1+\epsilon_n}$ hits 0 or 2 respectively. Without loss of generality, suppose $Z_t^{1+\epsilon_n}$ hits 0 before 2 at $[0, \Gamma_{4\epsilon_n}]$ in the event E^{ϵ_n} and thus we look at the event $\{\bar{\Gamma}_0 < \Gamma_{4\epsilon_n}, \bar{\Gamma}_0 < \bar{\Gamma}_2\} \cap E^{\epsilon_n}$. Then within this event there is a.s. $t < 4\epsilon_n^2$ s.t. $Z_t^{1+\epsilon_n} \leq 0$, $B_t \geq -\frac{1}{2}\epsilon_n$ and $Z_s^{1+\epsilon_n} < 2, \forall s < t$. Note that

$$Z_t^{1+\epsilon_n} = 1 + \epsilon_n - \int_0^t Z_s^{1+\epsilon_n} ds + B_t.$$

However,

$$\text{RHS} \geq (1 + \epsilon_n) - 2 \cdot 4\epsilon_n^2 - \frac{1}{2}\epsilon_n > 0$$

which implies that

$$\mathbf{P}\left(\{\bar{\Gamma}_0 < \Gamma_{4\epsilon_n}, \bar{\Gamma}_0 < \bar{\Gamma}_2\} \cap E^{\epsilon_n}\right) = 0.$$

Similarly, one also has

$$\mathbf{P}\left(\{\bar{\Gamma}_2 < \Gamma_{4\epsilon_n}, \bar{\Gamma}_2 < \bar{\Gamma}_0\} \cap E^{\epsilon_n}\right) = 0.$$

$\square$

Now we are able to show that with at least a uniformly positive probability $p > 0$, an O-U process stays to the right of 1 for some positive fraction of time between τ_n and τ_{n-1} to trigger the Poisson jump; the detailed proof can be found in Appendix B.

Lemma 3.20. *For all sufficiently small $\epsilon > 0$ and $n \geq 1$,*

$$
\mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1+\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right) \geq \mathbf{P}(E^{\epsilon_n}) = p,
$$
$$
\mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1-\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right) \geq \mathbf{P}(E^{\epsilon_n}) = p. \tag{3.75}
$$

Thus we have

$$
\mathbf{P}(A_n) \geq \mathbf{P}(E^{\epsilon_n}) = p > 0.
$$

With the above preparation, we can finish the proof of (3.67).

Proof of Proposition 3.15: For any $\epsilon > 0$ (without loss of generality $\epsilon < \frac{1}{3}$), define $n_1 = \left\lfloor \frac{\log(\epsilon^{-1}) + \log 3}{\log\left(\frac{1}{1-p}\right)} \right\rfloor + 1$ and a constant $\gamma \in (1, +\infty)$ that depends only on p as

$$
\gamma := \frac{4 \log 2}{\log\left(\frac{1}{1-p}\right)} + 1 \tag{3.76}
$$

so we have $\epsilon_{n_1} = \epsilon \cdot 2^{-n_1} \geq \epsilon^\gamma$. Recall that Γ obeys the exponential distribution $\exp(1)$ and we define a "globally failed event",

$$
B = \left\{\Gamma \geq \epsilon^{-\gamma}\right\} \cup \{\tau_0 > \epsilon\} \cup \cap_{n=1}^{n_1} A_n^c
$$

and call B^c the globally successful event. By Lemmas 3.19 and 3.20 we have

$$
\mathbf{P}(B) \leq \exp(-\epsilon^{-\gamma}) + \exp(-\epsilon^{-\frac{1}{4}}) + (1-p)^{n_1}
$$
$$
< \frac{1}{3}\epsilon + \frac{1}{3}\epsilon + \frac{1}{3}\epsilon < \epsilon. \tag{3.77}
$$

Recalling the random set $\hat{I}(\eta_1, \delta)$ in (3.22), then for any $\delta < \epsilon^{4\gamma} < \epsilon_{n_1}$, within B^c we have $\{\tau_0 \leq \epsilon\}$ and $\exists k \in [1, n_1]$ s.t.

$$
\frac{\mathfrak{L}\left(\hat{I}(\epsilon, \delta)\right)}{\delta} = \frac{\int_0^\epsilon \mathbb{1}_{\{Z_t^1 > 1+\delta\}} dt}{\delta} \geq \frac{\int_{\tau_k}^{\tau_{k-1}} \mathbb{1}_{\{Z_t^1 > 1+\epsilon_k\}} dt}{\delta} \geq \frac{\epsilon_{n_1}^2}{\delta} \geq \epsilon^{-2\gamma} > \Gamma, \tag{3.78}
$$

which gives (3.67). Thus we have found a polynomial order of η_1 as the lower bound for η_0 and the same for η_2 and the proof of (2.8) is complete. $\qquad\square$

4 Numerical tests

In previous sections we have shown that the state-dependent jump-diffusion process X_t^δ converges to X_t in distribution with a polynomial-order convergence rate. However, quantifying the correct convergence rate of X_t^δ remains an open question. Recently, a structure-preserving numerical scheme for the Fokker-Planck equation (2.1) based on Scharfetter-Gummel reformulation was proposed in [31]. With this numerical scheme, we aim to explore the convergence structure and study X_t^δ in terms of density functions through the Fokker-Planck equation (2.4) together with its nonlinear cases. Numerical study of the density function f^δ not only provides numerical evidence of the convergence rate of the process, but also indicates signs of self-similar structure when $\delta \to 0$.

This section is outlined as follows. First, in Section 4.1, we introduce the Scharfetter-Gummel reformulation on Fokker-Planck equation and the detailed construction of the

numerical schemes. Then, in Section 4.2, we numerically examine the convergence rate of the approximation error. Last, in Section 4.3, we study the self-similar structure of f^δ when it vanishes on $(1, +\infty)$ as $\delta \to 0$. We also note that the Fokker-Planck equation (3.47) for the killing process $\widetilde{X_t^\delta}$ (defined in Equation (3.46)) is also considered in this section.

4.1 Introduction to the scheme

First we introduce the nonlinear extensions of the Fokker-Planck equations for the jump-diffusion process X_t^δ and the killing process $\widetilde{X_t^\delta}$, which are similar to the Fokker-Planck equation (1.4) associated with X_t. We only show the following nonlinear equation for the density function f^δ of the jump-diffusion process X_t^δ; the nonlinear equation for the density function f_0^δ of the killing process $\widetilde{X_t^\delta}$ can be derived in a similar way.

$$
\begin{cases}
\dfrac{\partial f^\delta}{\partial t}(x,t) + \dfrac{\partial}{\partial x}[(-v + bN^\delta(t))f^\delta(x,t)] - a(N^\delta(t))\dfrac{\partial^2 f^\delta}{\partial x^2}(x,t) \\
= N^\delta(t)\delta(x) - \lambda^\delta(x)f^\delta(x,t), \quad (x,t) \in (-\infty, 1) \times [0, +\infty), \\
N^\delta(t) = \displaystyle\int_{\mathbb{R}} \lambda^\delta(y)f^\delta(y,t)dy,\, f^\delta(-\infty,t) = f^\delta(+\infty,t) = 0, \quad t > 0,
\end{cases}
\tag{4.1}
$$

where the terms $a(N^\delta(t))$ and $-v + bN^\delta(t)$ incorporate the effect of the mean firing rate on the dynamics of the density function at the macroscopic level. In particular, b models the connectivity of the neuron networks: $b > 0$ describes excitatory networks and $b < 0$ describes inhibitory networks. In this section, we assume $a \equiv 1$ and we are concerned with the convergence behavior with different connectivity parameters b.

The Scharfetter-Gummel reformulation on Equation (4.1) is given as follows:

$$
\dfrac{\partial f^\delta}{\partial t}(x,t) - a\dfrac{\partial}{\partial x}\left(M^\delta(x,t)\partial_x\left(\dfrac{f^\delta(x,t)}{M^\delta(x,t)}\right)\right) = N^\delta(t)\delta(x) - \lambda^\delta(x)f^\delta(x,t),
$$
$$
(x,t) \in (-\infty, 1) \times [0, +\infty), \tag{4.2}
$$

where

$$
M^\delta(x,t) = \exp\left(-\dfrac{(x - bN^\delta(t))^2}{2a}\right).
$$

The numerical scheme for Equation (4.1) is based on this reformulated equation.

Even though the jump-diffusion process X_t^δ is of better regularity than X_t, numerical approximation of $f^\delta(x,t)$ near $x = 1$ is still at risk of being inaccurate especially when δ is close to 0. Therefore, we apply the logistic scaling of the density function to partition a denser grid around $x = 1$. We take the computation domain as $[-4, 4]$ and assume homogeneous Dirichlet boundary condition for the density functions. We make the substitution

$$
y = h_L(x) = \dfrac{1}{1 + e^{-(x-1)}}, \quad x \in [-4, 4], \tag{4.3}
$$

and denote $q^\delta(y,t) = f^\delta(g_L(y),t)$, where g_L stands for inverse function of the logistic function h_L. Figure 1 shows an illustration of the scaling.

Then we put Equation (4.3) into Equation (4.2) to derive an equation for $q^\delta(y,t)$ on the computational domain

$$
\dfrac{\partial q^\delta}{\partial t}(y,t) - \dfrac{a}{g_L'(y)}\partial_y\left(\dfrac{M^\delta(g_L(y),t)}{g_L'(y)}\partial_y\left(\dfrac{q^\delta(y,t)}{M^\delta(g_L(y),t)}\right)\right) = N^\delta(t)\delta(y - y_r)
$$
$$
-\lambda^\delta(g_L(y))q^\delta(y,t), \quad (y,t) \in \left[\dfrac{1}{1 + e^5}, \dfrac{1}{1 + e^{-3}}\right] \times [0, +\infty), \tag{4.4}
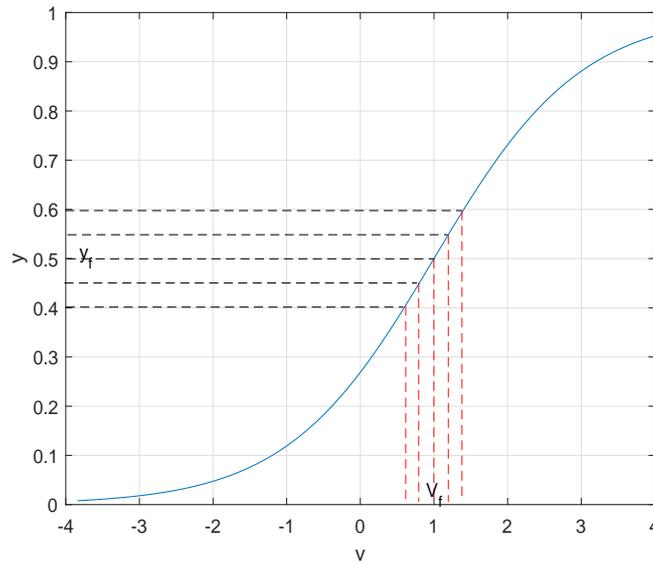$$

Figure 1: Illustration of the logistic scaling. In the figure, we partition a denser grid near $V_f = 1$ through the logistic scaling.

where $g'_L(y) = \frac{1}{y-y^2}$ and the reset point $y_r = h_L(0) = \frac{1}{1+e}$. In addition, the mean firing rate function N^δ is given as follows:

$$N^\delta(t) = \int_{\frac{1}{1+e^5}}^{\frac{1}{1+e^{-3}}} g'_L(z) \lambda^\delta\left(g_L(z)\right) q^\delta(z,t) dz. \tag{4.5}$$

The numerical scheme applied in this section is based on discretization of Equations (4.4) and (4.5). Let $q^\delta_{j,m}$ stand for the numerical value of q^δ at $y_j = jh + \frac{1}{1+e^5} \in \left[\frac{1}{1+e^5}, \frac{1}{1+e^{-3}}\right]$ and $t_m = m\tau \geq 0$, where h and τ denote spatial and temporal step lengths. Let N^δ_m denote the numerical approximation of the firing rate function N^δ at t_m. Also note that the reset point y_r is a grid point denoted as $y_r = y_D$ where $D \in \mathbb{N}^*$.

We apply the semi-implicit scheme to discretize the equations [31]. In other words, we treat q^δ_j implicitly but treat N^δ explicitly (including the N^δ in the term $M^\delta(x,t)$). The scheme is as follows:

$$\frac{q^\delta_{j,m+1} - q^\delta_{j,m}}{\tau} - \frac{a}{h^2 g'_L(y_j)} \left(\frac{M^\delta\left(g_L\left(y_{j+\frac{1}{2}}\right), t_m\right)}{g'_L\left(y_{j+\frac{1}{2}}\right)} \left(\frac{q^\delta_{j+1,m+1}}{M^\delta\left(g_L\left(y_{j+1}\right), t_m\right)} - \frac{q^\delta_{j,m+1}}{M^\delta\left(g_L\left(y_j\right), t_m\right)} \right) \right.$$

$$\left. - \frac{M^\delta\left(g_L\left(y_{j-\frac{1}{2}}\right), t_m\right)}{g'_L\left(y_{j-\frac{1}{2}}\right)} \left(\frac{q^\delta_{j,m+1}}{M^\delta\left(g_L\left(y_j\right), t_m\right)} - \frac{q^\delta_{j-1,m+1}}{M^\delta\left(g_L\left(y_{j-1}\right), t_m\right)} \right) \right)$$

$$= \frac{1}{h} N^\delta_n I(y_j = y_D) - \lambda^\delta\left(g_L\left(y_j\right)\right) q^\delta_{j,m+1}, \tag{4.6}$$

and

$$N^\delta_m = h \sum_j g'_L\left(y_j\right) \lambda^\delta\left(g_L\left(y_j\right)\right) q^\delta_{j,m}, \tag{4.7}$$

where $I(y)$ is the indicator function.

In our numerical simulations, we simulate excitatory and inhibitory networks with different connectivity parameters b. We use the same Gaussian distribution as the initial condition for all of the numerical tests:

$$f_G(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(x-x_0)^2}{2\sigma_0^2}}, \tag{4.8}$$

where $x_0 = -1$ and $\sigma_0^2 = 0.01$ are two given parameters. Since we assume that X_t^δ and $\widetilde{X_t^\delta}$ start at a given point in previous sections, we choose σ_0 to be very small in order to approximate the one-point initial distribution of the processes. The computing time is fixed to $t_{\max} = 1$. In addition, we consider the rate function $\lambda^\delta(x)$ as follows:

$$\lambda^\delta(x) = \begin{cases} 0, & x \leq 1, \\ \frac{1}{\delta}, & x \geq 1, \end{cases} \tag{4.9}$$

which is slightly different from the continuous rate of Equation (2.2).

Remark 4.1. In fact, with the same initial data, numerical solutions of the Fokker-Planck equations with two different rate functions are almost the same. Convergence and asymptotic behavior for the two cases (Equations (2.2) and (4.9)) are similar, though the convergence exponents are slightly different. However, (4.9) is of a simpler form, which facilitates the convergence study. Hence, we only consider the rate function defined in Equation (4.9) in numerical tests.

4.2 Convergence rates

In this subsection we aim to investigate the convergence of the jump-diffusion process X_t^δ (defined in Equation (1.10)) and the killing process $\widetilde{X_t^\delta}$ (defined in Equation (3.46)) as $\delta \to 0$ through numerical examination of the Fokker-Planck equations of the two processes (see Equations (2.4) and (3.47)). We compute the discrepancies between the density functions and firing rate functions (defined in Equation (4.1)) of the two processes, i.e. we consider density discrepancy

$$D^f(\delta) = \left\| f^\delta(x,1) - f(x,1) \right\|_\infty, \quad D^{f_0}(\delta) = \left\| f_0^\delta(x,1) - f_0(x,1) \right\|_\infty, \tag{4.10}$$

and firing rate discrepancy

$$D^N(\delta) = \left\| N^\delta(t) - N(t) \right\|_\infty, \quad D^{N_0}(\delta) = \left\| N_0^\delta(t) - N_0(t) \right\|_\infty, \tag{4.11}$$

where $\|\cdot\|_\infty$ denotes the L^∞ norm in space or time. Here the density functions $f(x,t), f_0(x,t)$ and firing rate functions $N(t), N_0(t)$ of X_t and its killing process $\widetilde{X_t}$ (defined in a similar way using Equation (3.46)) are obtained by numerically solving the nonlinear Fokker-Planck equation (1.4) with nonlinear drift and diffusion term (1.8) using the scheme in Section 4.1.

In Figure 2, we show the results of simulating the cases $\delta = \frac{1}{2^k}$ ($0 \leq k \leq 7, k \in \mathbb{N}$) with different parameters b, where we consider the evolution of the discrepancy functions defined in Equations (4.10) and (4.11) as δ goes to 0. The convergence of the density functions and firing rate functions are roughly linear when δ is moderately small, while the rates of convergence of the cases with different connectivity parameters b vary.

Then we define the convergence rates for the discrepancy functions as $\delta \to 0$ as follows:

$$D^f(\delta) = A_f \delta^{R^f} + o(\delta^{R^f}), \delta \to 0, \tag{4.12}$$

where R^f denotes the convergence rate of density function f^δ and A_f is a fixed parameter. We can define convergence rates R^{f_0}, R^N, R^{N_0} for the discrepancy functions in Equations (4.10) and (4.11) in the similar way to Equation (4.12).

In Table 1 we show the convergence rates of the functions with different b. The convergence rates are computed through linear fitting after eliminating the data when δ is too large or too small in order to avoid inaccuracy. More specifically, we retain only the data when $\delta = \frac{1}{2^k}$ where $k = 4, 5, 6, 7$ for the linear fitting.

We remark that the connectivity parameter b is chosen to be moderately small such that the solution of (1.4) with non-linear terms (1.8) does not blow up, and we have observed that $f^\delta(x,t)$ converges to $f(x,t)$ as $\delta \to 0$. However, when $b > 0$ increases, the solution to (1.4) with non-linearity (1.8) may blow up in finite time while the solution of (4.1) remains globally well posed. In fact, time periodic solutions have been shown to exist or been numerically observed for variant regularized models. The interested reader may refer to [12] and [32] for detailed discussions.

4.3 Self-similar structure

As $\delta \to 0$, the jump-diffusion process X_t^δ converges to X_t, which takes values on the half space $(-\infty, 1]$ rather than the whole space. The exact process of how X_t^δ vanishes on $(1, +\infty)$ remains an open question. In this subsection, we aim to study the self-similar profile of the density function f^δ of X_t^δ on $(1, +\infty)$ through numerical experiments.

Jump-diffusion process X_t^δ	$b = 1$	$b = 0.5$	$b = 0$	$b = -0.5$	$b = -1$
Convergence rate R^f	0.3716	0.3187	0.3505	0.3766	0.3961
Convergence rate R^{f_0}	0.3365	0.3832	0.4092	0.4262	0.4307
Killing process $\widetilde{X_t^\delta}$	$b = 1$	$b = 0.5$	$b = 0$	$b = -0.5$	$b = -1$
Convergence rate R^N	0.3416	0.3856	0.4122	0.4219	0.4228
Convergence rate R^{N_0}	0.4166	0.4302	0.4309	0.4322	0.4302

Table 1: Convergence rates for density functions and firing rate functions of the Fokker-Planck equations of two processes with different connectivity parameter b. See Equation (4.12) for the definition of the convergence rates of the functions. See Equations (2.4), (3.47) and (4.1) for the Fokker-Planck equations of the jump-diffusion process X_t^δ and the killing process $\widetilde{X_t^\delta}$.

We assume an ansatz for f^δ when $\delta \to 0$ as follows:

$$f^\delta(x,t) = \delta^\alpha \psi\left(\delta^\beta(x-1)\right) + o\left(\delta^\alpha\right), \quad \forall x \in [1, +\infty), \tag{4.13}$$

where ψ is defined on $\mathbb{R}^+$ and α, β are two fixed parameters. In this subsection, we aim to explore the self-similar structure of f^δ (defined in Equation (2.4)) with such an ansatz and find the fixed parameters α and β numerically. Moreover, we also make a similar ansatz for the density f_0^δ of the killing process $\widetilde{X_t^\delta}$ (defined in Equation (3.47)) as a reference, since the f_0^δ display a similar vanishing structure to f^δ.

Numerical examinations involve the cases of $\delta = \frac{1}{2^k}$ ($0 \le k \le 7, k \in \mathbb{N}$) with different parameters b. Through similar data choices and linear fitting, numerical results for α and β are shown in Table 2.

Finally, in Figure 3, we take the values α and β in Table 2 in ansatz (4.13) for X_t^δ and plot the profiles of ψ for each δ with connectivity parameter $b = 0, 1$. Numerically, ψ is nearly independent of δ and decays exponentially in y. Therefore, we conclude that it is very likely that $f^\delta(x)$ exhibits the self-similar profile in Equation (4.13) when $x \ge 1$.
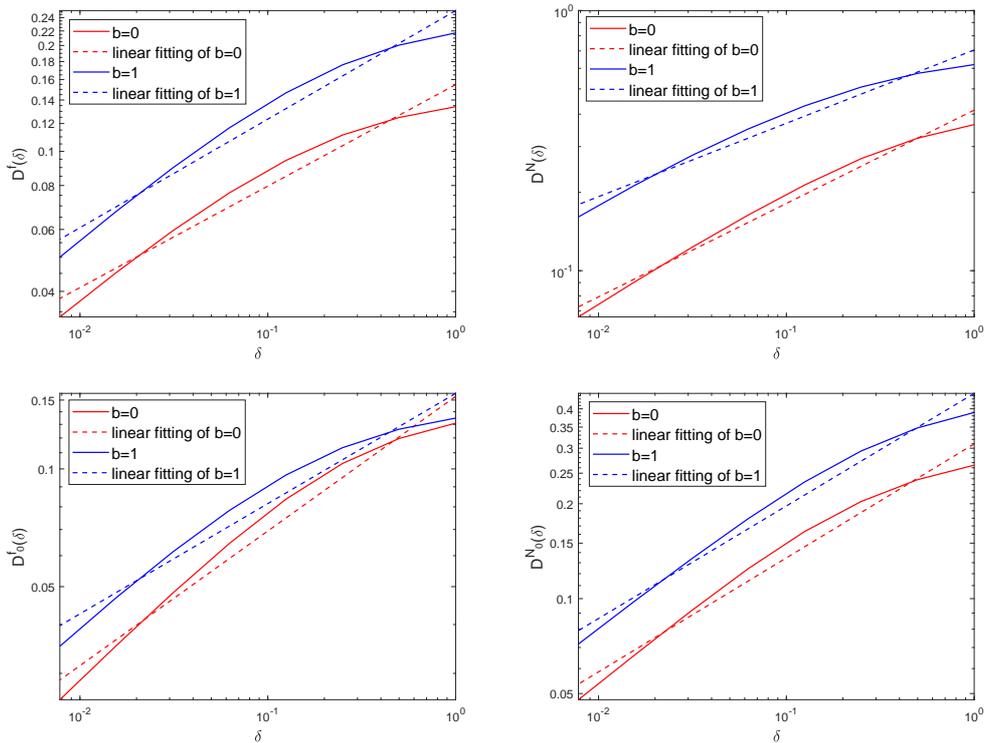
Figure 2: (Convergence of the density functions and the firing rate functions of the jump-diffusion process X_t^δ and killing process $\widetilde{X_t^\delta}$ with parameter $b = 0$ and $b = 1$.) These figures show the evolution of the discrepancy functions defined in Equations (4.10) and (4.11) as $\delta \to 0$. Top: the density discrepancy $D^f(\delta)$ and firing rate discrepancy $D^N(\delta)$ of the jump-diffusion process X_t^δ with connectivity parameter $b = 0$ and $b = 1$. See Equations (2.4) and (4.1) for the Fokker-Planck equations of the process X_t^δ. Bottom: the density discrepancy $D^{f_0}(\delta)$ and firing rate discrepancy $D^{N_0}(\delta)$ of the killing process $\widetilde{X_t^\delta}$ with connectivity parameter $b = 0$ and $b = 1$. See Equations (3.47) and (4.1) for the Fokker-Planck equations of the process $\widetilde{X_t^\delta}$

.

Jump-diffusion process X_t^δ	$b = 1$	$b = 0.5$	$b = 0$	$b = -0.5$	$b = -1$
Values of α	0.2713	0.3187	0.3505	0.3766	0.3961
Values of β	-0.4256	-0.4363	-0.4283	-0.4317	-0.4268
Killing process $\widetilde{X_t^\delta}$	$b = 1$	$b = 0.5$	$b = 0$	$b = -0.5$	$b = -1$
Values of α	0.3448	0.3856	0.4122	0.4344	0.4507
Values of β	-0.4307	-0.4148	-0.4136	-0.4317	-0.4317

Table 2: Numerical values of the parameters α and β in ansatz (4.13). Parameters of the two processes with different connectivity b are shown in the table. See Equations (2.4), (3.47) and (4.1) for the Fokker-Planck equations.
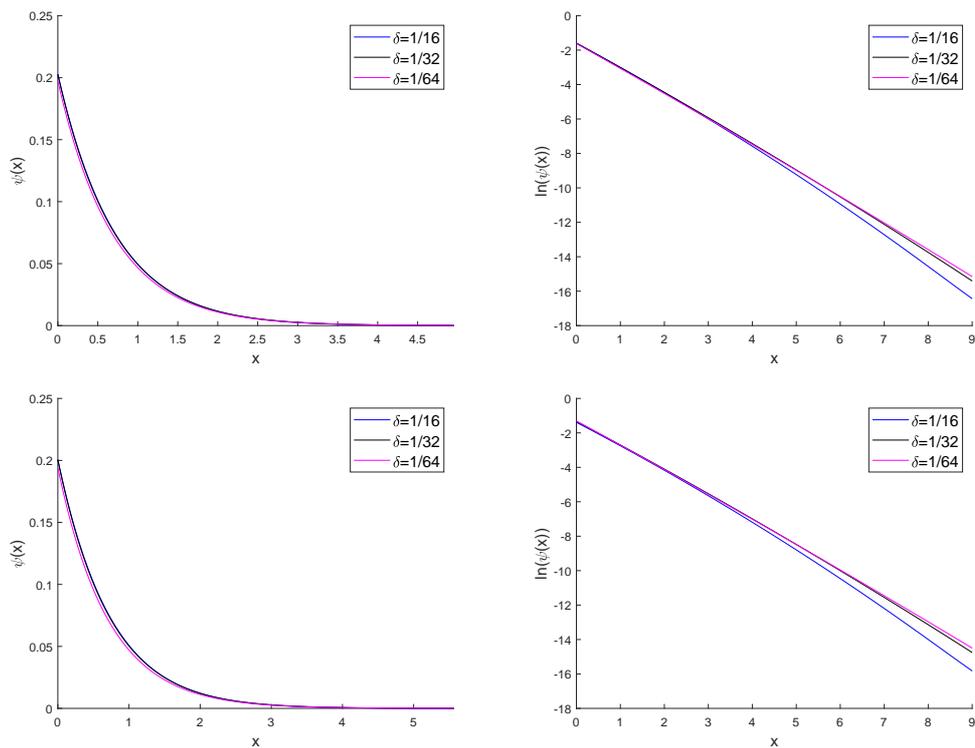
Figure 3: (Image for ψ and $\ln(\psi)$ with connectivity parameter $b = 0, 1$.) In this figure, we put the values of α and β in Table 2 into the ansatz (4.13) for the density function f^δ with different δ. Numerically, we see the profile of ψ is independent of δ, which indicates that f^δ is likely to exhibit self-similar structure when $x \geq 1$ as $\delta \to 0$. See Equations (2.4) and (4.1) for the Fokker-Planck equations of the process X_t^δ. Left: Image for ψ. Right: Image for $\ln(\psi)$. Top: $b = 0$. Bottom: $b = 1$.

5 Conclusion and Discussion

In this work we aim to reduce the gap in understanding between the mean-field integrate-and-fire model as a stochastic process and the PDE model as an evolving density function. As shown in [5], it is possible to find an initial probability distribution such that the solution of the nonlinear Fokker-Planck equation with a deterministic firing potential must blow up in finite time, which is conjectured to be linked to the multiple firing events (synchronization) of neuronal networks. The random discharge mechanism is introduced to prevent the blow up of the solution of the PDE model such that the synchronized state becomes possible on the macroscopic level. In this paper, we have rigorously justified that the regularized solution is indeed an approximation to the original one, which confirms the scientific intuition behind the random discharge mechanism. As the continuation of [29], we only focus on the linear cases and show that the relevant random discharge model converges to the original integrate-and-fire model in distribution as the regularization parameter goes to 0. Mathematically, the iterated scheme can effectively reduce the difficulties of analyzing the problems with the firing-and-resetting mechanism, and gives more intuitive stochastic interpretations of the macroscopic quantities of the PDE, which are otherwise obscure. Using specifically designed numerical experiments, we have observed evidence for the convergence rate and the asymptotic behavior for both the linear cases and the more sophisticated nonlinear cases, which motivates us to carry out a rigorous asymptotic analysis in subsequent work. It is worth noting that we have not yet incorporated the dependence on the mean firing rate in the drift velocity and in the diffusion coefficient and we shall investigate those directions in later work. However, there are still additional challenges due to the interacting nature and the nonlinearity within the model.

Appendix

A Proof of Proposition 3.10

Proof. To prove the case when $n = 2$, by the Markov property, we first have:

$$\mathbf{P}^0(X_{t_1}^\delta \in (a_1, b_1], X_{t_2}^\delta \in (a_2, b_2]) = \int_{a_1}^{b_1} \mathbf{P}^y(X_{t_2-t_1}^\delta \in (a_2, b_2]) f^\delta(t_1, y) dy.$$

and

$$\mathbf{P}^0(X_{t_1} \in (a_1, b_1], X_{t_2} \in (a_2, b_2]) = \int_{a_1}^{b_1} \mathbf{P}^y(X_{t_2-t_1} \in (a_2, b_2]) f(t_1, y) dy.$$

Thus

$$\begin{aligned}
&\mathbf{P}^0(X_{t_1}^\delta \in (a_1, b_1], X_{t_2}^\delta \in (a_2, b_2]) - \mathbf{P}^0(X_{t_1} \in (a_1, b_1], X_{t_2} \in (a_2, b_2]) \\
&\leq \left| \int_{a_1}^{b_1} \mathbf{P}^y(X_{t_2-t_1} \in (a_2, b_2]) \left[f^\delta(t_1, y) - f(t_1, y) \right] dy \right| \\
&\quad + \int_{a_1}^{b_1} \left| \mathbf{P}^y(X_{t_2-t_1}^\delta \in (a_2, b_2]) - \mathbf{P}^y(X_{t_2-t_1} \in (a_2, b_2]) \right| f^\delta(t_1, y) dy \\
&=: I_1 + I_2
\end{aligned}$$

(A.1)

For I_1, consider the function

$$\varphi(y) = \begin{cases} \mathbf{P}^y(X_{t_2-t_1} \in (a_2, b_2]) & \text{if } y \in (a_1, b_1], \\ 0 & \text{otherwise.} \end{cases}$$

(A.2)

According to the strong Feller property of X_t, $\mathbf{P}^y(X_{t_2-t_1} \in (a_2, b_2])$ is a continuous function with respect to $y \in (-\infty, 1)$. Thus $\varphi(y)$ is a bounded measurable function with

discontinuities at a_1 and b_1. Moreover, by Proposition 3.1 of [29], $F(t, y)$ is continuous in y and thus puts 0 measure on $\{a_1, b_1\}$. Thus by Theorem 3.2.10 of [19] and the fact that $\mathbf{E}^0[\varphi(X_{t_1}^\delta)] \to \mathbf{E}^0[\varphi(X_{t_1})]$ as $\delta \to 0^+$, we get $I_1 \to 0$.

For I_2, by Corollary 3.11, for all $\delta \le \eta_0$ we have

$$I_2 \le \epsilon \int_{a_1}^{b_1} f^\delta(t, y) dy \le \epsilon. \tag{A.3}$$

Thus we have proved Proposition 1 for $n = 2$. In general, suppose Proposition 3.10 holds for all $k \le n$. Now for $k = n+1$, we can similarly define $\varphi^{(n)}(y)$ for $y = (y_1, \cdots, y_n)$ as

$$\varphi^{(n)}(y) = \begin{cases} \mathbf{P}^y(X_{t_{n+1}-t_n} \in (a_{n+1}, b_{n+1})) & \text{if } y \in (a_1, b_1] \times \cdots \times (a_n, b_n], \\ 0 & \text{otherwise.} \end{cases} \tag{A.4}$$

and have

$$\mathbf{P}^0(X_{t_1}^\delta \in (a_1, b_1], X_{t_2}^\delta \in (a_2, b_2], \cdots, X_{t_n}^\delta \in (a_n, b_n], X_{t_{n+1}}^\delta \in (a_{n+1}, b_{n+1}])$$
$$-\mathbf{P}^0(X_{t_1} \in (a_1, b_1], X_{t_2} \in (a_2, b_2], \cdots, X_{t_n} \in (a_n, b_n], X_{t_{n+1}} \in (a_{n+1}, b_{n+1}])$$
$$\le \left| \mathbf{E}^0\left[\varphi^{(n)}(X_{t_1}, \cdots, X_{t_1}) \right] - \mathbf{E}^0\left[\varphi^{(n)}(X_{t_1}^\delta, \cdots, X_{t_1}^\delta) \right] \right|$$
$$+ \int_{(a_1, b_1] \times \cdots \times (a_n, b_n]} \left| \mathbf{P}^y(X_{t_{n+1}-t_n}^\delta \in (a_{n+1}, b_{n+1}]) - \mathbf{P}^y(X_{t_{n+1}-t_n} \in (a_{n+1}, b_{n+1}]) \right| dF^{(n)}(y). \tag{A.5}$$

where $F^{(n)}(y)$ is the distribution of $(X_{t_1}, \cdots, X_{t_n})$. Note that $\varphi^{(n)}(\cdot)$ is a bounded measurable function on $\mathbb{R}^n$ whose discontinuities are given by

$$D_{\varphi^{(n)}} = \{y : \{y_1 = a_1 \text{ or } b_1\} \cup \cdots \cup \{y_n = a_n \text{ or } b_n\}\}$$

Since $F(\cdot, t_i)$, $i = 1, 2, \cdots n$ are all continuous on $(-\infty, 1]$, the joint distribution $F^{(n)}(\cdot)$ of $(X_{t_1}, \cdots, X_{t_n})$ puts 0 mass on $D_{\varphi^{(n)}}$. Thus by (vi) of Theorem 3.10.1 in [19], we have

$$\left| \mathbf{E}^0\left[\varphi^{(n)}(X_{t_1}, \cdots, X_{t_1}) \right] - \mathbf{E}^0\left[\varphi^{(n)}(X_{t_1}^\delta, \cdots, X_{t_1}^\delta) \right] \right| \to 0.$$

At the same time, by a similar argument as in (A.3), we have

$$\int_{(a_1, b_1] \times \cdots \times (a_n, b_n]} \left| \mathbf{P}^y(X_{t_{n+1}-t_n}^\delta \in (a_{n+1}, b_{n+1}]) - \mathbf{P}^y(X_{t_{n+1}-t_n} \in (a_{n+1}, b_{n+1}]) \right| dF^{(n)}(y) \le \epsilon.$$

Thus by induction, the proof of Proposition 3.10 is complete. $\square$

B Proof of Lemma 3.20

Proof. Recall Lemma 3.19, the definition of E^{ϵ_n} and the OU process

$$Z_t^{1+\epsilon_n} = 1 + \epsilon_n - \int_0^t Z_s^{1+\epsilon_n} ds + B_t.$$

In event E^{ϵ_n} there is

$$Z_{\Gamma_{4\epsilon_n}}^{1+\epsilon_n} \ge 1 + \epsilon_n - 2 \cdot 4\epsilon_n^2 + 4\epsilon_n > 1 + 2\epsilon_n = 1 + \epsilon_{n-1}. \tag{B.1}$$

Thus we know that $\tau_{n-1} < \Gamma_{4\epsilon_n}$ in event E^{ϵ_n}. And for any $t < \Gamma_{4\epsilon_n} < 4\epsilon_n^2$ with $\Gamma_{4\epsilon_n} < \Gamma_{-\frac{1}{2}\epsilon_n}$,

$$Z_t^{1+\epsilon_n} \ge 1 + \epsilon_n - 2 \cdot 4\epsilon_n^2 - \frac{1}{2}\epsilon_n > 1 - 2\epsilon_n = 1 - \epsilon_{n-1}.$$

Thus we know that $Z_t^{1+\epsilon_n}$ does not hit $1 - \epsilon_{n-1}$ before $\Gamma_{4\epsilon_n}$. Thus it must hit $1 + \epsilon_{n-1}$ before $\Gamma_{4\epsilon_n}$. However, note that $\forall\, t < \Gamma_{\frac{2}{3}\epsilon_n} < \Gamma_{4\epsilon_n}$,

$$Z_t^{1+\epsilon_n} \leq 1 + \epsilon_n - 0 - \frac{2}{3}\epsilon_n < 1 + 2\epsilon_n = 1 + \epsilon_{n-1}.$$

Thus we have $\Gamma_{\frac{2}{3}\epsilon_n} < \tau_{n-1} < \Gamma_{4\epsilon_n}$ in E^{ϵ_n}. And in E^{ϵ_n} for any $t < \Gamma_{4\epsilon_n}$ such that $B_t \in [\frac{1}{4}\epsilon_n, \frac{1}{2}\epsilon_n]$,

$$Z_t^{1+\epsilon_n} \geq 1 + \epsilon_n - 2 \cdot 4\epsilon_n^2 + \frac{1}{4}\epsilon_n > 1 + \epsilon_n.$$

Thus in event E^{ϵ_n}

$$\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1+\epsilon_n} > 1+\epsilon_n\}} dt \geq \int_0^{\Gamma_{\frac{2}{3}\epsilon_n}} \mathbb{1}_{\{B_t \in [\frac{1}{4}\epsilon_n, \frac{1}{2}\epsilon_n]\}} dt \geq \epsilon_n^2,$$

and then

$$\mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1+\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right) \geq \mathbf{P}(E^{\epsilon_n}).$$

Similarly, given E^{ϵ_n}, for all $t < \Gamma_{4\epsilon_n} < 4\epsilon_n^2$, since $\Gamma_{4\epsilon_n} < \Gamma_{-\frac{1}{2}\epsilon_n}$, we have

$$Z_t^{1-\epsilon_n} \geq 1 - \epsilon_n - 2 \cdot 4\epsilon_n^2 - \frac{1}{2}\epsilon_n > 1 - 2\epsilon_n = 1 - \epsilon_{n-1}$$

and

$$Z_{\Gamma_{4\epsilon_n}}^{1-\epsilon_n} \geq 1 - \epsilon_n - 2 \cdot 4\epsilon_n^2 + 4\epsilon_n > 1 + 2\epsilon_n = 1 + \epsilon_{n-1}.$$

At the same time, $\forall t < \Gamma_{2.8\epsilon_n} < \Gamma_{4\epsilon_n}$, we have

$$Z_t^{1-\epsilon_n} \leq 1 - \epsilon_n - 0 + 2.8\epsilon_n < 1 + 2\epsilon_n = 1 + \epsilon_{n-1}.$$

Thus $\Gamma_{2.8\epsilon_n} < \tau_{n-1} < \Gamma_{4\epsilon_n}$. And for any $t < \Gamma_{4\epsilon_n}$, $B_t \in [2.2\epsilon_n, 2.5\epsilon_n]$, we also have

$$Z_t^{1-\epsilon_n} \geq 1 - \epsilon_n - 2 \cdot 4\epsilon_n^2 + 2.2\epsilon_n > 1 + \epsilon_n.$$

So again in E^{ϵ_n} we have

$$\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1-\epsilon_n} > 1+\epsilon_n\}} dt \geq \int_0^{\Gamma_{2.8\epsilon_n}} \mathbb{1}_{\{B_t \in [2.2\epsilon_n, 2.5\epsilon_n]\}} dt \geq \epsilon_n^2,$$

and then

$$\mathbf{P}\left(\int_0^{\tau_{n-1}} \mathbb{1}_{\{Z_t^{1-\epsilon_n} > 1+\epsilon_n\}} dt > \epsilon_n^2\right) \geq \mathbf{P}(E^{\epsilon_n}).$$

$\square$

References

[1] Iddo Ben-Ari and Ross G. Pinsky. Spectral analysis of a family of second-order elliptic operators with nonlocal boundary condition indexed by a probability measure. *J. Funct. Anal.*, 251(1):122–140, 2007. https://doi.org/10.1016/j.jfa.2007.05.019

[2] Iddo Ben-Ari and Ross G. Pinsky. Ergodic behavior of diffusions with random jumps from the boundary. *Stochastic Process. Appl.*, 119(3):864–881, 2009. https://doi.org/10.1016/j.spa.2008.05.002

[3] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.*, 8(3):183–208, 2000. https://doi.org/10.1023/A:1008925309027

[4] Nicolas Brunel and Vincent Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.*, 11(7):1621–1671, 1999. https://doi.org/10.1162/089976699300016179

[5] María J Cáceres, José A Carrillo, and Benoît Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *J. Math. Neurosci.*, 1(1):7, 2011. https://doi.org/10.1186/2190-8567-1-7

[6] María J Cáceres, José A Carrillo, and Louis Tao. A numerical solver for a nonlinear fokker–planck equation representation of neuronal network dynamics. *J. Comput. Phys.*, 230(4):1084–1099, 2011. https://doi.org/10.1016/j.jcp.2010.10.027

[7] María J. Cáceres and Benoît Perthame. Beyond blow-up in excitatory integrate and fire neuronal networks: refractory period and spontaneous activity. *J. Theoret. Biol.*, 350:81–89, 2014. https://doi.org/10.1016/j.jtbi.2014.02.005

[8] María J. Cáceres and Ricarda Schneider. Analysis and numerical solver for excitatory-inhibitory networks with delay and refractory periods. *ESAIM Math. Model. Numer. Anal.*, 52(5):1733–1761, 2018. https://doi.org/10.1051/m2an/2018014

[9] José A Carrillo, María D. M. González, Maria P Gualdani, and Maria E Schonbek. Classical solutions for a nonlinear Fokker-Planck equation arising in computational neuroscience. *Comm. Partial Differential Equations*, 38(3):385–409, 2013. https://doi.org/10.1080/03605302.2012.747536

[10] Carlo Cercignani. *Rarefied gas dynamics: From basic concepts to actual calculations.* Cambridge University Press, 2000.

[11] Albert Compte, Nicolas Brunel, Patricia S. Goldman-Rakic, and Xiao-Jing Wang. Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model. *Cereb. Cortex*, 10(9):910–923, 2000. https://doi.org/10.1093/cercor/10.9.910

[12] Quentin Cormier, Etienne Tanré, and Romain Veltz. Hopf bifurcation in a Mean-Field model of spiking neurons. *Electron. J. Probab.*, 26(art. 121):1–40, 2021. https://doi.org/10.1214/21-ejp688

[13] Quentin Cormier, Etienne Tanré, and Romain Veltz. Long time behavior of a mean-field model of interacting neurons. *Stochastic Process. Appl.*, 130(5):2553–2595, 2020. https://doi.org/10.1016/j.spa.2019.07.010

[14] Robert Dautray and Jacques Louis Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 2. Functional and variational methods.* Springer-Verlag, Berlin, 1988. https://doi.org/10.1007/978-3-642-61566-5

[15] François Delarue, James Inglis, Sylvain Rubenthaler, and Etienne Tanré. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.*, 125(6):2451–2492, 2015. https://doi.org/10.1016/j.spa.2015.01.007

[16] François Delarue, James Inglis, Sylvain Rubenthaler, and Etienne Tanré. First hitting times for general non-homogeneous 1d diffusion processes: Density estimates in small time. *Tech. Report.*, 2013. http://hal.archives-ouvertes.fr/hal-00870991

[17] Laurent Desvillettes and Mario Pulvirenti. The linear Boltzmann equation for long-range forces: a derivation from particle systems. *Math. Models Methods Appl. Sci.*, 9(8):1123–1145, 1999. https://doi.org/10.1142/S0218202599000506

[18] Richard Durrett. *Stochastic calculus: a practical introduction.* Probab. Stochastics Ser. CRC press, 1996.

[19] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019. https://doi.org/10.1017/9781108591034

[20] László Erdos and Horng-Tzer Yau. Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation. *Comm. Pure Appl. Math.*, 53(6):667–735, 2000. https://doi.org/10.1002/(SICI)1097-0312(200006)53:6<667::AID-CPA1>3.0.CO;2-5

[21] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and convergence*, volume 282. John Wiley & Sons, 1986. https://doi.org/10.1002/9780470316658

[22] William Feller. The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math. (2)*, 55:468–519, 1952. https://doi.org/10.2307/1969644

[23] William Feller. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.*, 77:1–31, 1954. https://doi.org/10.2307/1990677

[24] Avner Friedman. *Partial differential equations of parabolic type.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1964.

[25] Maria Giovanna Garroni and José Luis Menaldi. *Green functions for second order parabolic integro-differential problems*, Pitman Research Notes in Mathematics Series **275**. Harlow: Longman Scientific & Technical; copublished in the United States with Wiley, New York, 1992.

[26] Wulfram Gerstner and Werner M Kistler. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press, 2002. https://doi.org/10.1017/CBO9780511815706

[27] Ĭosip Ĭllich Gīhman and Anatolii Volodimirovich Skorohod. *Stochastic differential equations*. Springer-Verlag, New York-Heidelberg, 1972.

[28] Toni Guillamon. An introduction to the mathematics of neural activity. *Butl. Soc. Catalana Mat*, 19(2):25–45, 2004.

[29] Jian Guo Liu, Ziheng Wang, Yuan Zhang, and Zhennan Zhou. Rigorous justification of the Fokker-Planck equations of neural networks based on an iteration perspective, arXiv:2005.08285, 2021.

[30] Floyd B. Hanson. *Applied stochastic processes and control for jump-diffusions: Modeling, analysis and computation.* SIAM, 2007. https://doi.org/10.1137/1.9780898718638

[31] Jingwei Hu, Jian-Guo Liu, Yantong Xie, and Zhennan Zhou. A structure preserving numerical scheme for Fokker-Planck equations of neuron networks: Numerical analysis and exploration *J. Comput. Phys.*, 433 (pap.23), 2021 https://doi.org/10.1016/j.jcp.2021.110195

[32] Kota Ikeda, Pierre Roux, Delphine Salort, and Didier Smets. Theoretical study of the emergence of periodic solutions for the inhibitory NNLIF neuron model with synaptic delay. 2021. https://hal.archives-ouvertes.fr/hal-03157218/

[33] James Inglis and Denis Talay. Mean-field limit of a stochastic particle system smoothly interacting through threshold hitting-times and applications to neural networks with dendritic component. *SIAM J. Math. Anal.*, 47(5):3884–3916, 2015. https://doi.org/10.1137/140989042

[34] Patrick Jahn, Rune W. Berg, Jørn Hounsgaard, and Susanne Ditlevsen. Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J. Comput. Neurosci.*, 31(3):563–579, 2011. https://doi.org/10.1007/s10827-011-0326-z

[35] Louis Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation *Journal de Physiologie et de Pathologie Générale* 9: 620–635, 1907.

[36] Thomas M. Liggett. *Continuous Time Markov Processes: An Introduction.* Graduate Studies in Mathematics, vol. 113., American Mathematical Society, 2010. https://doi.org/10.1090/gsm/113

[37] Maurizio Mattia and Paolo Del Giudice. Population dynamics of interacting spiking neurons. *Phys. Rev. E (3)*, 66(5):051917, 2002. https://doi.org/10.1103/PhysRevE.66.051917

[38] Katherine A. Newhall, Gregor Kovačič, Peter R. Kramer, and David Cai. Cascade-induced synchrony in stochastically driven neuronal networks. *Phys. Rev. E (3)*, 82(4):041903, 2010. https://doi.org/10.1103/PhysRevE.82.041903

[39] Katherine A. Newhall, Gregor Kovačič, Peter R. Kramer, Douglas Zhou, Aaditya V. Rangan, and David Cai. Dynamics of current-based, Poisson driven, integrate-and-fire neuronal networks. *Commun. Math. Sci.*, 8(2):541–600, 2010. http://projecteuclid.org/euclid.cms/1274816894

[40] Duane Quinn Nykamp. *A population density approach that facilitates large-scale modeling of neural networks*. Thesis (Ph.D.), Courant Institute of Mathematical Sciences, New York University, 2000.

[41] Bernt Oksendal. *Stochastic differential equations: An introduction with applications*. Universitext, Springer-Verlag, Berlin, 2998. https://doi.org/10.1007/978-3-662-03620-4

[42] Ahmet Omurtag, Bruce W. Knight, and Lawrence Sirovich. On the simulation of large populations of neurons. *J. Comput. Neurosci.*, 8(1):51–63, 2000. https://doi.org/10.1023/A:1008964915724

[43] Jun Peng. A note on the first passage time of diffusions with holding and jumping boundary. *Statist. Probab. Lett.*, 93:58–64, 2014. https://doi.org/10.1016/j.spl.2014.06.012

[44] Jun Peng and WenBo V. Li. Diffusions with holding and jumping boundary. *Sci. China Math.*, 56(1):161–176, 2013. https://doi.org/10.1007/s11425-012-4416-9

[45] Philip Protter. Stochastic differential equations with jump reflection at the boundary. *Stochastics*, 3(3):193–201, 1980. https://doi.org/10.1080/17442508008833144

[46] Alfonso Renart, Nicolas Brunel, and Xiao-Jing Wang. Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks. *in Computational neuroscience: A comprehensive approach*, pp.431–490, ed. by J. Feng, CRC Press, 2004. https://doi.org/10.1201/9780203494462.

[47] Lawrence Sirovich, Ahmet Omurtag, and Kip Lubliner. Dynamics of neural populations: Stability and synchrony. *Network*, 17(1):3–29, 2006. https://doi.org/10.1080/09548980500421154

[48] Leszek Sł ominski and Tomasz Wojciechowski. Stochastic differential equations with jump reflection at time-dependent barriers. *Stochastic Process. Appl.*, 120(9):1701–1721, 2010. https://doi.org/10.1016/j.spa.2010.04.008

[49] Jonathan Touboul, Geoffroy Hermann, and Olivier Faugeras. Noise-induced behaviors in neural mean field dynamics. *SIAM J. Appl. Dyn. Syst.*, 11(1):49–81, 2012. https://doi.org/10.1137/110832392

[50] Henry C Tuckwell. *Introduction to theoretical neurobiology: volume 2, Nonlinear and Stochastic Theories*, Cambridge Studies in Mathematical Biology, vol. 8. Cambridge University Press, 1988. https://doi.org/10.1017/CBO9780511623202