Evolution of moments and correlations in non-renewal escape-time processes

Wilhelm Braun,[1, *] Rüdiger Thul,[2] and André Longtin[1]

[1]*Department of Physics and Centre for Neural Dynamics,*
University of Ottawa, 598 King Edward, Ottawa K1N 6N5, Canada
[2]*Centre for Mathematical Medicine and Biology, School of Mathematical Sciences,*
University of Nottingham, Nottingham, NG7 2RD, UK
(Dated: April 27, 2017)

The theoretical description of non-renewal stochastic systems is a challenge. Analytical results are often not available or can only be obtained under strong conditions, limiting their applicability. Also, numerical results have mostly been obtained by ad-hoc Monte–Carlo simulations, which are usually computationally expensive when a high degree of accuracy is needed. To gain quantitative insight into these systems under general conditions, we here introduce a numerical iterated first-passage time approach based on solving the time-dependent Fokker–Planck equation (FPE) to describe the statistics of non-renewal stochastic systems. We illustrate the approach using spike-triggered neuronal adaptation in the leaky and perfect integrate-and-fire model, respectively. The transition to stationarity of first-passage time moments and their sequential correlations occur on a non-trivial timescale that depends on all system parameters. Surprisingly this is so for both single exponential and scale-free power-law adaptation. The method works beyond the small noise and timescale separation approximations. It shows excellent agreement with direct Monte Carlo simulations, which allows for the computation of transient and stationary distributions. We compare different methods to compute the evolution of the moments and serial correlation coefficients (SCC), and discuss the challenge of reliably computing the SCC which we find to be very sensitive to numerical inaccuracies for both the leaky and perfect integrate-and-fire models. In conclusion, our methods provide a general picture of non-renewal dynamics in a wide range of stochastic systems exhibiting short and long-range correlations.

I. INTRODUCTION

A general property of diverse systems, ranging from superconducting quantum interference devices (SQUIDs) [1], to lasers [2] to excitable cells [3–6] is that time intervals between specific events are not statistically independent. The theoretical description of such non-renewal stochastic processes [7] poses a significant challenge, as it implies that the present state of the system depends, in general, on the whole past evolution or parts of it, and not just on the previous state. Analytical approximations to tackle such memory effects have included the assumption of stationarity [8], small stochasticity [9] and time-scale separation [10–12] between stochastic and deterministic parts of the dynamics.

Even if these approximations allow for some insight into the parameter dependence of e.g. serial correlations and can be used to understand experimental data, as exemplified in [9, 13–16] in the context of excitable systems, it is desirable to understand the statistics of model systems without making simplifying assumptions. Regarding stationarity, real systems rarely operate in a stationary state due to transients that arise from deterministic or random perturbations. A prominent example are cortical neurons. An average cortical neuron receives random inputs from approximately 10^4 other neurons, whose activity is modulated by non-stationary sensory and other inputs, resulting in transient neuronal dynamics [17] that

only become stationary after a certain time. It is therefore important to understand how statistical properties of inter-event times evolve and become invariant following a transient regime due to internal dynamics and external inputs. Keeping with illustrations from neural dynamics, it is well known that physiologically relevant processes underlying neural coding rarely only have one well-defined timescale [18, 19]. This has lead researchers in various theoretical fields to consider multiple time-scale dynamics [20]. An important example of a system with multiple time scales is neuronal adaptation, where a neuron's firing rate adjusts in response to a stimulus. Adaptation with multiple timescales, or even no time scale as in the case of power-law adaptation [21–23], is now known to be biophysically relevant, and even optimal for some tasks [24]. Recently, it was also shown that a neuron model with adaptive firing thresholds exhibiting multiple timescales is the optimal choice for the prediction of spike times in cortical neurons [25, 26]. Therefore, a theoretical description of adaptation without a single well-defined timescale is an important goal.

In this paper, we show how to describe two-dimensional non-renewal dynamics by an iterated first-passage time (iFPT) approach. This approach allows us to determine stationary statistical properties of the system as well as providing a description of the transition to stationarity. We furthermore show how to compute serial correlations in the time series generated by the firing times of the system. While our approach is general and applicable to any system where first-passage times [27] play a role, we illustrate its versatility with two important examples, namely spike-triggered neuronal adaptation with a sin-

* wilhelm.braun@cantab.net

gle exponential current and a power-law current without an intrinsic timescale, respectively. Using the underlying time-dependent FPE to describe the system, we only need to apply mathematically convenient standard absorbing boundary conditions to obtain stationary distributions, e.g. that of the adaptation current upon firing. Moreover, the methods developed here can easily be extended to models with correlation-generating deterministic input currents as recently considered in [28].

II. MODEL

We consider a stochastic differential equation (SDE) driven by an external signal $s(t)$:

$$\mathrm{d}X(t) = \mu(X(t))dt + \phi(X(t))\mathrm{d}W(t) - s(t)\mathrm{d}t. \quad (1)$$

X is defined on the domain $(-\infty, x_{\mathrm{th}}]$. If X reaches x_{th}, the system is said to have generated and event, and X is instantaneously reset to 0. For all examples in this study, we chose the Ornstein–Uhlenbeck process (OUP) given its prominence in the field of stochastic systems. For the OUP, we fix the correlation time $\tau_m = \frac{1}{\gamma}$, bias current I_0 and noise intensity σ as follows: $\mu(X(t)) = \gamma(I_0 - X(t))$, $\phi(X(t)) = \sigma\gamma$. $W(t)$ is a standard Brownian motion and we set $x_{\mathrm{th}} = 1$. Given that the OUP is the basis for integrate-and-fire (IF) neuron models, which are among the most popular neuron descriptions [29], we refer to events as spikes and to $s(t)$ as a time-dependent adaptation current in the present study. The general dynamics of $s(t)$ obeys a single autonomous ordinary differential equation (ODE)

$$\dot{s} = \omega(s), \quad (2)$$

and s is increased by a fixed amount κ when $X = x_{\mathrm{th}}$: $s \to s + \kappa$, which is the mechanism for spike-triggered adaptation [30]. When $s(t)$ is also reset to its starting value $s(0)$, the model is a renewal model and its firing statistics may be studied using standard techniques, see e.g. [31] for a recent review.
Here we focus on two forms of the adaptation current. The first one is power-law adaptation, for which

$$\omega(s) = -\frac{1}{\alpha}s^2(t). \quad (3)$$

This ODE has the general solution $s(t) = \left(\frac{t}{\alpha} + \frac{1}{s(0)}\right)^{-1}$. Therefore, the current s in this case has a power-law time dependence with no intrinsic time scale [21].
The second adaptation current is given by a single exponential decay with time scale τ_a:

$$\omega(s) = -\frac{1}{\tau_a}s(t), \quad (4)$$

which has the general solution $s(t) = s(0)e^{-\frac{t}{\tau_a}}$.
The time to the first spike event is the following first-passage time (FPT):

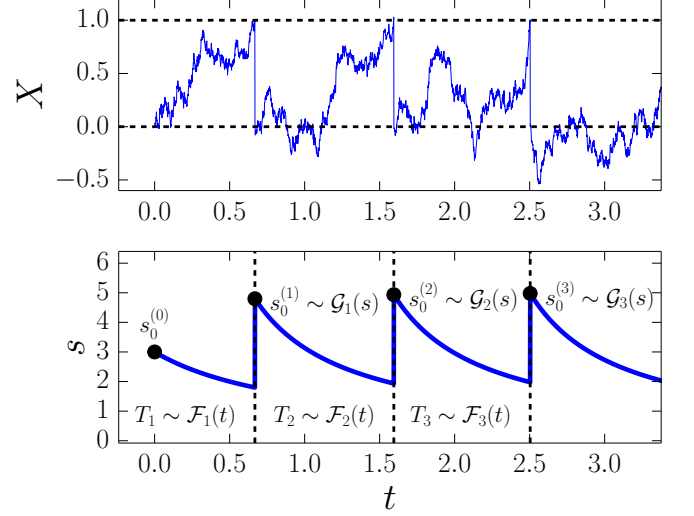$$T_1 = \inf(t > 0 : X(t) > x_{\mathrm{th}} | X(0) = 0, s(t = 0) = s(0)).$$

FIG. 1. Sample paths of the model for a power-law adaptation current given by Eq. 3. **Top:** $X(t)$ (Eq. 1), the horizontal dashed lines are at $X = 0$ and $x_{\mathrm{th}} = 1$. **Bottom**: $s(t)$ (Eq. 2) When X reaches x_{th}, s undergoes a jump of size κ. The subsequent ISIs T_k have distributions $\mathcal{F}_k(t)$, and the starting values $s_0^{(k)}$ have distributions $\mathcal{G}_k(s)$ for $k \geq 1$. Parameter values are $\alpha = 3.0$, $\gamma = 1.0$, $\sigma = 0.8$, $I_0 = 4.0$, $\kappa = 3.0$.

In the non-renewal case we are studying here, subsequent firing times will in general not have the same distribution as T_1. We define the kth interspike interval (ISI) as

$$T_k = \inf\left(t - \sum_{i=1}^{k-1}T_i \ : \ X(t) \geq x_{\mathrm{th}}, t > \sum_{i=1}^{k-1}T_i\right). \quad (5)$$

The first moment of the kth ISI is given by $\tau_k^1 = \mathbb{E}(T_k)$. The second moment of the kth ISI will be denoted by $\tau_k^2 = \mathbb{E}((T_k)^2)$ and the kth firing rate is given by the inverse of the corresponding mean ISI $r_k = \frac{1}{\tau_k^1}$. The kth standard deviation $m_2(k)$ is then given by

$$m_2(k) = \sqrt{\tau_k^2 - (\tau_k^1)^2}. \quad (6)$$

The values of the peak adaptation current after the kth firing are defined for $k \geq 1$ as

$$s_0^{(k)} = \left(s(t^-) + \kappa \ : \ t = \sum_{i=1}^{k}T_i\right), \quad (7)$$

where t^- indicates that we take the left-sided limit. For simplicity, we choose s to be started from a point $(s_0^{(0)} = \kappa)$, instead of from a biophysically more realistic initial distribution. However, the methods we are going to describe in the following are also valid when s is initially started from a distribution.

The central challenge is to obtain the distributions $\mathcal{G}_k$ and $\mathcal{F}_k$ for $k \geq 1$, which are the distributions of $s_0^{(k)}$ and T_k defined by Eqs. 7 and 5, respectively. An example realization for the case of power-law adaptation is shown in Fig. 1. The knowledge of these distributions is key to understanding the non-renewal dynamics, as they form a hidden Markov model of the underlying non-Markovian dynamics [15, 16, 32]. Therefore, once these distributions are known, the non-renewal dynamical system breaks up into coupled renewal dynamical systems, which are much more tractable mathematically. This gives rise to the iFPT approach which we now explain.

III.   THE IFPT APPROACH

Being a diffusion process, the system given by Eqs. 1 and 2 is governed by a two-dimensional time-dependent FPE [33]. The FPE for the probability density function $p(t; x, s)\mathrm{d}x\mathrm{d}s = \mathbb{P}(X(t) \in (x, x + \mathrm{d}x), s(t) \in (s, s + \mathrm{d}s)|X(0) = x_0, s(0) = s_0)$ reads (we use $\mathbf{x}^\top = (x, s)$ for brevity)

$$\partial_t p(t; \mathbf{x}) = \nabla \cdot (\mathbf{A}(\mathbf{x})\nabla p(t; \mathbf{x})) - \nabla \cdot (\mathbf{F}(\mathbf{x})p(t; \mathbf{x})), \quad (8)$$

where $\mathbf{A}$ is the diffusion matrix and $\mathbf{F}$ the drift vector, which can be obtained in a straightforward way from the SDE for X, Eq. 1, and the ODE for s, Eq. 2. Explicitly, we have

$$\mathbf{A}(\mathbf{x}) = \begin{pmatrix} \frac{\phi(x)^2}{2} & 0 \\ 0 & 0 \end{pmatrix}, \quad (9)$$

and

$$\mathbf{F}(\mathbf{x}) = (\mu(x) - s, \omega(s))^\top.$$

The IF property of X entails that we have an absorbing boundary at $X = x_{\mathrm{th}}$ for all times t: $p(t; x = x_{\mathrm{th}}, s) = 0$. With this boundary condition, we can compute the cumulative distribution function (CDF) of the first-passage time T_1, $\mathrm{CDF}_1(t) = \int_0^t \mathcal{F}_1(\lambda)\mathrm{d}\lambda$, by time evolution of the FPE on a computational domain $\Omega \subset \mathbb{R}^2$, which we choose to be a rectangle extending to sufficiently negative values in the x-direction [34]:

$$\mathrm{CDF}_1(t) = 1 - \int_\Omega p_1(t; \mathbf{z})\mathrm{d}\mathbf{z}, \quad (10)$$

where p_1 is the solution to Eq. 8 with the initial condition $p_1(0; x, s) = \delta(\mathbf{x} - \mathbf{x}_0)$, where $\mathbf{x}_0 = (0, \kappa)^\top$. For the computation of $\mathcal{F}_1$, we then only need to differentiate Eq. 10.

To describe adaptation, one needs to compute the statistics of the peak adaptation currents, as defined by Eq. 7. Hence, we need to characterize the hidden Markov model generated by the ISIs T_k and the peak adaptation currents $s_0^{(k)}$. Whereas the dynamics of s and X as a whole is non-Markovian, the distribution of the ISI T_k is completely determined by the distribution $\mathcal{G}_{k-1}$, and the sequence $\{T_k, s_0^{(k-1)}\}_{k=1}^\infty$ is therefore Markovian. Knowing the values of T_k and $s_0^{(k-1)}$, the value of $s_0^{(k)}$ is fixed, and the distribution of the ISI T_{k+1} can be obtained by solving an FPT problem with $s_0^{(k)}$ as initial condition for s.

The central observation now is that for the second ISI, X again starts from 0, whereas s starts from a distribution $\mathcal{G}_1$. This is because X evolves stochastically, and therefore reaches the threshold x_{th} at different times T_1, corresponding to different values of $s(t = T_1) + \kappa$ immediately after the firing event. To compute the second ISI, we therefore need to know $\mathcal{G}_1$. This can be iterated: to compute the distribution $\mathcal{F}_k$ of the kth ISI, we need the distribution $\mathcal{G}_{k-1}$. This is the central idea of the iFPT approach. To set up the iFPT approach, we first observe that between threshold crossings of X, s evolves deterministically. Therefore, when we know the PDF of the first FPT, by conservation of probability, we also know the distribution of s after the first firing event:

$$\mathcal{G}_1(s - \kappa) = \left| \frac{\mathrm{d}t(s)}{\mathrm{d}s} \right| \mathcal{F}_1(t(s)), \quad (11)$$

where $t(s)$ is the inverse function of s. The support of $\mathcal{G}_1$ is shifted, because of the jump of size κ that s undergoes when X reaches its threshold x_{th}. For the second ISI T_2, s is started from the distribution $\mathcal{G}_1(s)$ instead of a point, whereas X is started from a point again (Fig. 1). This means that to obtain $\mathcal{F}_2$, the FPE is started from a distribution: $p_2(0; \mathbf{x}) \propto \mathcal{G}_1(s)\delta(x)$. This generalizes to values of k larger than 1. For the kth ISI distribution $\mathcal{F}_k(t)$, we therefore must choose

$$p_k(0; \mathbf{x}) \propto \mathcal{G}_{k-1}(s)\delta(x). \quad (12)$$

We show how to obtain the distributions $\mathcal{G}_k$ for $k > 1$ in Section V. Linear splines are used to create a mesh function approximating Eq. 12 on the computational domain Ω. Due to this approximation, Eq. 12 then has to be normalized appropriately, so that $\int_\Omega p_k(0; \mathbf{z})\mathrm{d}\mathbf{z} = 1$. The FPE is then solved again, and the distributions $\mathcal{F}_k(t)$ are obtained analogously to Eq. 10: $\mathrm{CDF}_k(t) = 1 - \int_\Omega p_k(t; \mathbf{z})\mathrm{d}\mathbf{z}$, i.e. by timestepping the FPE to obtain the CDF of the kth ISI followed by a numerical differentiation. This constitutes the iFPT approach.

To quantify the accuracy of our numerical methods, we also compute the relative disagreement Δ between results obtained by the iFPT approach and direct MC simulations. It is defined for a quantity Z by

$$\Delta(Z) = \frac{|Z_{\mathrm{iFPT}} - Z_{\mathrm{MC}}|}{Z_{\mathrm{iFPT}}} . \tag{13}$$

We performed MC simulations for two different simulation setups, the first one without any boundary correction (plain MC), and the second one with a boundary correction according to Giraudo and Sacerdote (MC-GS) [31, 35, 36]. This boundary correction is applied to reduce the systematic overestimation of FPTs when using the Euler–Maruyama scheme. We compute the relative disagreement given by Eq. 13 using either MC simulations with or without boundary correction; we observed that the order of the relative disagreement is unchanged, but in general, the plain MC algorithm gives rise to larger disagreements than MC-GS. A decrease in the relative disagreement is expected, because the GS correction method should yield an improved weak error of $\mathcal{O}(h)$ [37], in contrast to the plain MC simulation, which has a weak error of $\mathcal{O}(h^{\frac{1}{2}})$ [38], where h is the time step for the discretization of the SDE, Eq. 1. In the following, the timestep for MC simulations is chosen to be $h = 10^{-3}$ and we choose $M = 10^6$ independent realizations. The plain Euler-Maruyama scheme then gives rise to a weak error of $\mathcal{O}(h^{\frac{1}{2}}) \approx 3 \cdot 10^{-2}$ when estimating moments of first passage times, which is one order of magnitude larger than the MC error proportional to $\frac{1}{\sqrt{M}} = 10^{-3}$. Therefore, in our simulations, the plain MC error is negligible in comparison to the error introduced by the finite-time discretization of the SDE (Eq. 1).

For the numerical solution of the FPE (Eq. 8), we choose a finite element discretization method [39] and evolve the system using either a stabilized Crank–Nicolson (CN) scheme [40] in Fig. 2 or an Euler timestepping scheme [39] in Fig. 3.

The relative disagreement between MC simulations and finite-element solutions stays largely constant across different lags k when we use the CN scheme instead of the Euler scheme as can be seen by comparing the lower panels of Figs. 2 and 3; the sizes of the disagreement are comparable in magnitude. This suggests that the remaining small discrepancy between MC simulations and PDE results can be largely explained with the errors associated with the MC simulation method. In particular, note that for the examples we show, the MC-GS weak error of size $\mathcal{O}(h)$ is comparable in magnitude to the numerator of $\Delta(Z)$ (Eq. 13), i.e. the absolute disagreement. We will see what effects this has on the computation of correlations in Section VI.
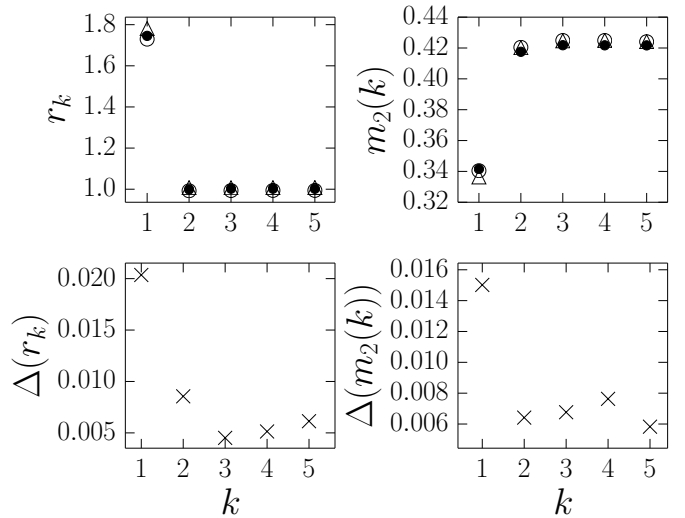

FIG. 2. **Top:** Evolution of the rate $r_k = \frac{1}{\tau_k^1}$ (left) and standard deviation given by Eq. 6 (right) of ISIs as a function of k. Empty circles: Plain MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Filled circles: moments obtained from numerical solution of FPE using a CN timestepping scheme. $M = 10^6$ independent MC realizations for each value of k. Power-law adaptation (Eq. 3) with $\alpha = 5.5$, $I_0 = 6.0$, $\sigma = 1.3$, $\gamma = 1.0$, $\kappa = 5.5$. **Bottom:** Relative disagreements defined by Eq. 13, where an MC-GS algorithm was used.

IV. TRANSITION TO STATIONARITY

We show the evolution of the rate and standard deviation of ISIs in Figs. 2 and 3. The rates decreases, whereas the standard deviation increases until both quantities reach a stationary value. Given that these quantities are derived from moments of the ISI distributions, the distributions also converge towards a stationary form. The convergence towards stationarity of ISI and peak adaptation current distributions ($\mathcal{F}_k$ and $\mathcal{G}_k$, respectively) is shown in Figs. 4 and 5. As expected for adapting models, the ISIs (whose distributions are shown in Fig. 4) increase with higher k, which means that the rate r_k decreases. This is well captured by the iFPT approach, with a maximal relative disagreement smaller than 3% in Fig. 2 and smaller than 2% in Fig. 3. Also, the width of both the ISI distributions and the peak adaptation current distributions (Fig. 5) increases, which is reflected by the increase of the variances of $\mathcal{F}_k$ and $\mathcal{G}_k$ shown in Fig. 2 and 3. The mean of the peak adaptation currents shifts to the right as stationarity is reached. Moreover, stationarity is reached with varying speed, i.e. for different values of the lag k (compare Fig. 2 with Fig. 3). Generally, the speed of adaptation can be controlled by adjusting the bias current I_0 and the noise level σ as well as the adaptation strength (size of the kick κ and, in the case of single exponential adaptation, the timescale τ_a). We have carried out additional MC simulations (not shown)
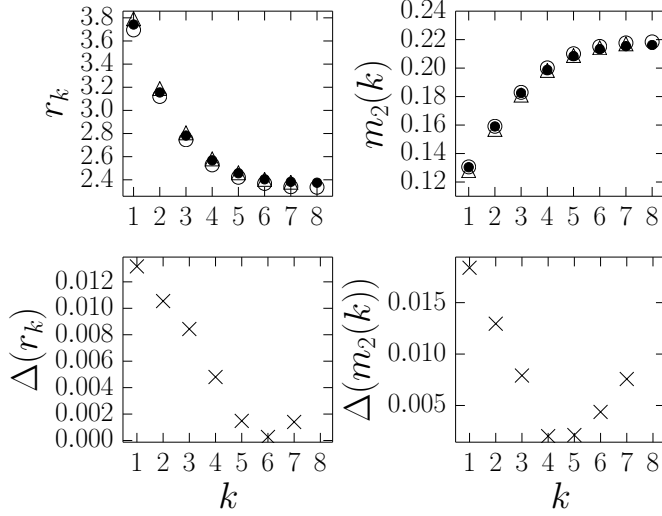
FIG. 3. **Top:** Evolution of the rate $r_k = \frac{1}{\tau_k^1}$ (left) and standard deviation given by Eq. 6 (right) of ISIs as a function of k. Empty circles: Plain MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Filled circles: moments obtained from numerical solution of FPE using an Euler timestepping scheme. $M = 10^6$ independent MC realizations for each value of k. Single exponential adaptation (Eq. 4) with $\tau_a = 1.0$, $I_0 = 5.0$, $\sigma = 1.0$, $\gamma = 1.0$, $\kappa = 1.0$. **Bottom:** Relative disagreements defined by Eq. 13, where an MC-GS algorithm was used.

FIG. 4. PDF $\mathcal{F}_k$ of the kth ISI T_k (Eq. 5). The symbols are MC simulations ($M = 10^6$ MC realizations) as indicated in the legends. Solid black lines: PDF obtained by numerical solution of the FPE, Eq. 8. In the left panel, the distributions are practically indistinguishable after $k = 2$. **Top:** Power-law adaptation (Eq. 3). **Bottom:** Single exponential adaptation (Eq. 4). Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation. Both panels show results for plain MC simulations.

to obtain insight into how these parameters influence the speed of the transition to stationarity. A higher bias current and a higher noise level will in general lead to a less rapid transition to stationarity. This can be understood as follows: X is driven to threshold more rapidly, causing the inhibition to build up quickly, reaching values that are higher than those typically found around the peak of the stationary distribution. This slows down the transition to stationarity, because s needs to decay first. Moreover, a large kick size κ and a rapidly decaying adaptation current will cause a quick transition. For the latter case, this is easily understood as we are then nearly dealing with a renewal system: after a short initial period, the effect of the adaptation current on the firing time statistics is negligible. For the former case, we note that the larger the kick size κ, the more pronounced the inhibitory effect of adaptation within one ISI, which means that X takes longer to reach threshold before a large out-of-equilibrium average value of the current s (a value that is larger than those typically found around the mode of the stationary distribution) can build up. The system reaches stationarity rapidly because it is quasi-deterministic as the dynamics of s dominates the system, and the stochastic fluctuations of X will only cause small perturbations. For both power-law and single exponential adaptation currents, it is possible to reach the stationary regime already after one or two firing events as in Fig. 2, or to have a long transient regime as in Fig. 3.
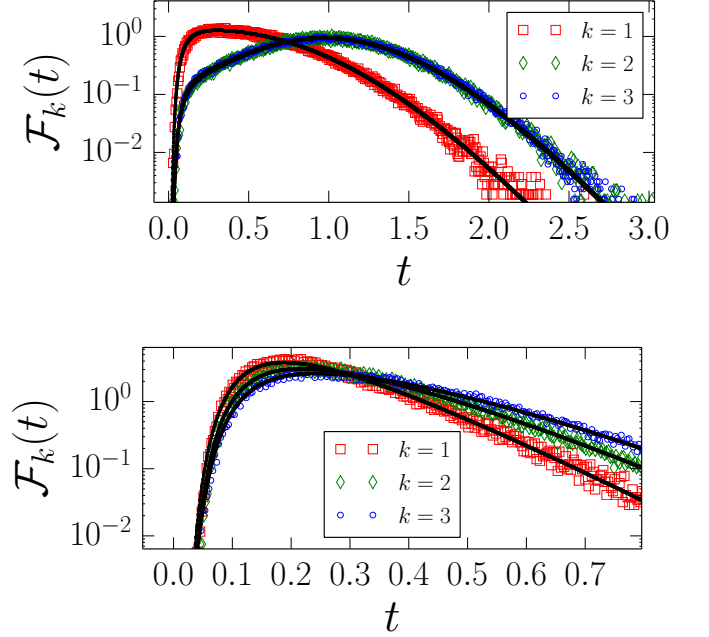
The initial condition $s_0^{(0)}$ for the adaptation current can also be chosen to control the speed of the transition. If it is placed far away from the mean of the stationary distribution, the transition will take a longer time; also, it is possible to obtain a non-monotonic behaviour of the rate as a function of the interval number when $s_0^{(0)}$ is placed far above the aforementioned mean. The first mean ISI will then be the longest statistically, in contrast to the examples we show in Figs. 2 and 3.

V. STATISTICS OF THE kTH ISI

The iFPT approach can be iterated beyond the first two firing events to obtain the distribution for the third ISI, $\mathcal{F}_3(t)$. However, for the computation of the third ISI, no equation similar to Eq. 11 can be used to obtain $\mathcal{G}_2$ because s was started from a distribution to obtain $\mathcal{F}_2$. Indeed, for one fixed time T_2, there are many different starting values $s_0^{(1)}$ due to the stochastic dynamics of X. Importantly, the ISI T_2 and the initial condition $s_0^{(1)}$ are not independent random variables (for a large value of $s_0^{(1)}$, a large ISI T_2 is more probable and vice versa) so
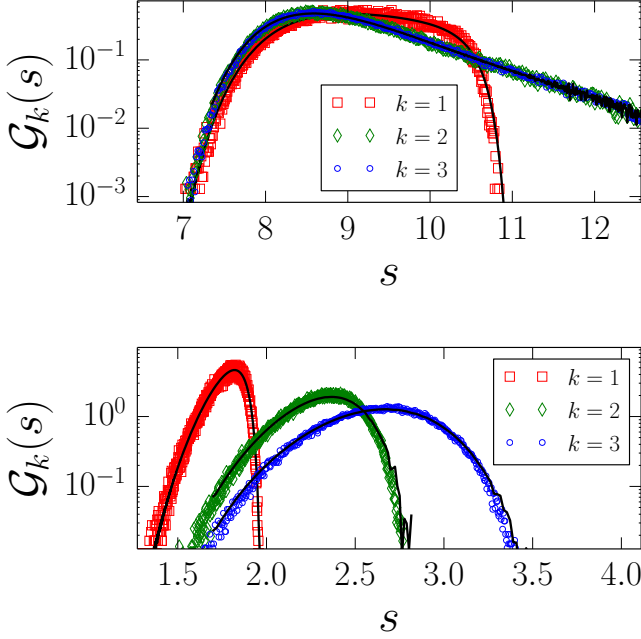
FIG. 5. PDF $\mathcal{G}_k$ of $s_0^{(k)}$ (Eq. 7). The symbols are MC simulations ($M = 10^6$ MC realizations) as indicated in the legends. The black solid line is $\mathcal{G}_1$ obtained from Eq. 11 or $\mathcal{G}_2$ ($\mathcal{G}_3$) obtained using Eq. 18. In the left panel, the distributions are practically indistinguishable after $k = 2$. **Top:** Power-law adaptation (Eq. 3). **Bottom:** Single exponential adaptation (Eq. 4). Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation. Both panels show results for plain MC simulations.

that we can obtain the value that s reaches after the second firing by the following observation (focusing on power-law adaptation): given that $T_2 = \lambda$ and $s_0^{(1)} = \nu$, we have $s_0^{(2)} = \frac{1}{\frac{\lambda}{\alpha} + \frac{1}{\nu}} + \kappa$. We have included the jump of size κ due to the definition of $s_0^{(2)}$ (see Fig. 1). This emphasizes that once two values in the triplet $(T_2, s_0^{(1)}, s_0^{(2)})$ are fixed, the third one is determined. In the following, we again use λ to denote a fixed FPT and ν to denote a fixed initial value of the adaptation current s. Analogously, we generally have for $s_0^{(k)}$: given $T_k = \lambda$ and $s_0^{(k-1)} = \nu$, $s_0^{(k)} = f(\lambda, \nu)$ is determined. The function f determining the subsequent value of the peak adaptation current given the previous ISI λ and the previous peak value of the adaptation current ν reads for power-law adaptation (Eq. 3):

$$ f(\lambda, \nu) = \kappa + \frac{1}{\frac{\lambda}{\alpha} + \frac{1}{\nu}} \,, \qquad (14) $$

whereas for exponential adaptation (Eq. 4), we have

$$ f(\lambda, \nu) = \kappa + \nu \exp\left(-\frac{\lambda}{\tau_a}\right) \,. \qquad (15) $$

An alternative way is to fix the value of $s_0^{(k)}$ and then put a constraint on the time T_k when $s_0^{(k-1)}$ is fixed: given $s_0^{(k-1)} = \nu$ and $s_0^{(k)} = \theta$, $T_k = h(\nu, \theta)$ is determined, where, for power-law adaptation, we have the ISI as a function of the previous and subsequent adaptation values:

$$ h(\nu, \theta) = \alpha \left(\frac{1}{\theta - \kappa} - \frac{1}{\nu}\right) \,. \qquad (16) $$

The function h is defined by solving the equation $f(\lambda, \nu) = \theta$ for λ. To actually compute the density $\mathcal{G}_k$, we need to relate the above observations to densities that we can compute with the iFPT approach. To that end, we now define the conditional density $\mathcal{H}$:

$$ \mathcal{H}(\lambda, \nu)\mathrm{d}\lambda \equiv \mathbb{P}\left(T_1 \in (\lambda, \lambda + \mathrm{d}\lambda)|s_0^{(0)} = \nu\right) \,. \qquad (17) $$

We have used T_1 and $s_0^{(0)}$ in the definition Eq. 17 to stress that, for the purpose of the computation of $\mathcal{H}$, we only need to solve the FPT problem for T_1 using different values of the initial condition $s_0^{(0)}$. It will become apparent below that we only need to compute $\mathcal{H}$ once, because it does not depend on the firing index k. For a fixed value of ν, $\mathcal{H}(\lambda, \nu)$ is an FPT probability density. Our notation emphasizes that $\mathcal{H}$ is a function of two variables. ν sets the level of initial inhibition, i.e. the starting value of s. We show the function $\mathcal{H}$ for both power-law and exponential adaptation in Fig. 6. We see that with increasing starting value ν, the mode of the FPT distribution shifts to larger times. For power-law adaptation, the shape of the FPT distributions does not change much, whereas for exponential adaptation, the distributions become broader with increasing ν.

With this at hand, we now show how to practically compute the distributions $\mathcal{G}_k$ for $k > 1$. We can obtain the CDF of $s_0^{(k)}$ by observing that:

$$ \mathbb{P}\left(s_0^{(k)} \leq \theta\right) = \int_{\mathcal{D}^{(k-1)}(\theta)} \mathcal{H}(\lambda, \nu)\mathcal{G}_{k-1}(\nu)\mathrm{d}\lambda\mathrm{d}\nu \,, \qquad (18) $$

with

$$ \begin{aligned} \mathcal{D}^{(k-1)}(\theta) = \\ (\lambda, \nu > 0| \, \nu \in \mathrm{supp}\left(\mathcal{G}_{k-1}\right), \, h(\nu, \theta) \leq \lambda \leq T_{\max}) \,. \end{aligned} \qquad (19) $$

The function h defined in Eq. 16 ensures that for a fixed value of ν, we collect all times λ so that $s_0^{(k)} \leq \theta$, which ensures that $f(\lambda, \nu) \leq \theta$ for fixed values of θ and ν. $T_{\max}$ is chosen so that $\mathcal{H}(T_{\max}, \nu) \approx 0 \ \forall \nu \in \mathrm{supp}\left(\mathcal{G}_{k-1}\right)$. This means that $T_{\max}$ should be chosen in the tail of the FPT distribution. Note that for the iFPT approach, one only has to compute $\mathcal{H}(\lambda, \nu)$ over the support of $\mathcal{G}_k$ for $k \geq 1$ once [41], and then multiply it with the adaptation current distribution of the previous iteration $\mathcal{G}_{k-1}$. This
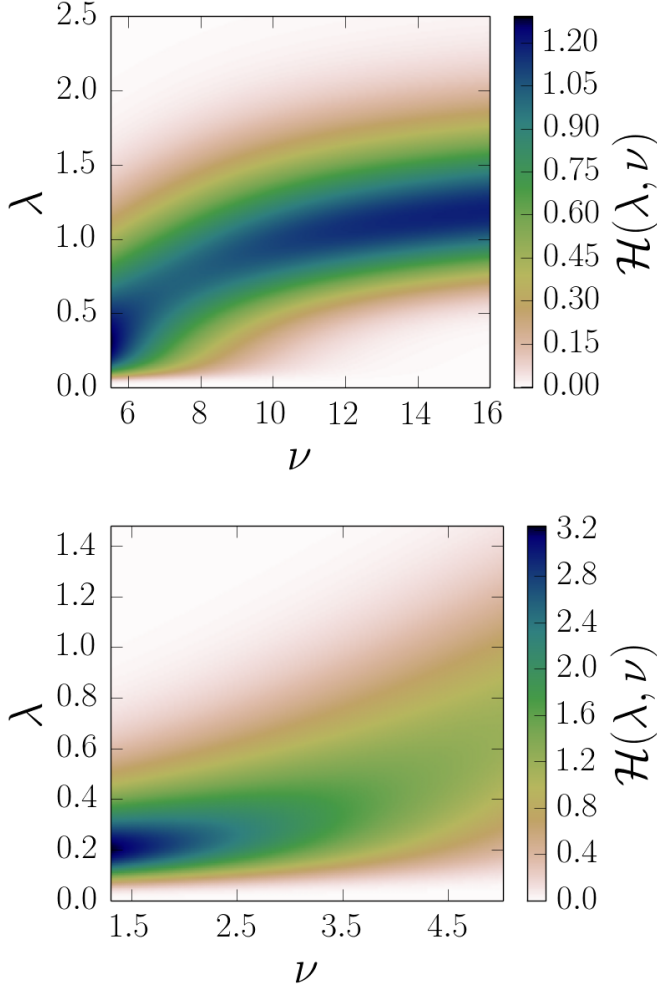
$$\mathcal{Q}(\theta) = \int_{\mathcal{D}(\theta)} \mathcal{H}(\lambda,\nu)\mathcal{Q}'(\nu)\mathrm{d}\lambda\mathrm{d}\nu\,, \qquad (20)$$

where $\mathcal{Q}(\theta) \equiv \mathbb{P}\left(s_0^{(k^*)} \leq \theta\right)$ denotes the CDF of the stationary peak value for the adaptation current s (Eq. 2). We have checked that this equation is indeed satisfied by the stationary distributions for the peak adaptation current obtained from MC simulations (data not shown). Therefore, Eq. 20 can serve as a tool to check whether a given distribution for the peak adaptation current is stationary, or alternatively as a way to compute $\mathcal{Q}(\theta)$ directly if the function $\mathcal{H}$ is known.

VI. CORRELATIONS BETWEEN INTERSPIKE INTERVALS

We now show how to compute serial correlations with the iFPT approach. We define the SCC [15, 42] between the nth ISI T_n and the $(n+k)$th ISI T_{n+k} according to

$$\mathrm{SCC}(n,k) = \frac{\mathbb{E}\left(T_n T_{n+k}\right) - \mathcal{Q}_1(n,k)}{\mathcal{Q}_2(n,k)}\,, \qquad (21)$$

where

$$\mathcal{Q}_1(n,k) = \mathbb{E}(T_n)\mathbb{E}(T_{n+k})\,, \qquad (22)$$

and

$$\mathcal{Q}_2(n,k) = m_2(k)m_2(n+k) = \sqrt{\mathrm{Var}(T_n)\mathrm{Var}(T_{n+k})}\,. \qquad (23)$$

Here, $\mathrm{Var}(T_n)$ denotes the variance of the nth ISI distribution, and m_2 is the standard deviation given by Eq. 6. Note that the definition Eq. 21 does not make use of the notion of stationarity, so that the SCC depends on both the position n of the ISI in the spike train as well as on the lag k between ISIs. Since we have already computed the distributions of the kth ISI, we can readily compute the variances and means in Eq. 21, i.e. the terms given by Eqs. 22 and 23. It is slightly more complicated to compute the first term in the numerator, $\mathbb{E}\left(T_n T_{n+k}\right)$, because we need the joint density $p_2(T_n, T_{n+k})$ of T_n and T_{n+k}. In the present study, we focus on $k=1$. By definition, we have

$$\mathbb{E}(T_n T_{n+1}) = \int \mathrm{d}T_n \mathrm{d}T_{n+1}\, T_n\, T_{n+1}\, p_2(T_n, T_{n+1}),$$
$$= \int \mathrm{d}T_n\, \mathrm{d}T_{n+1}\, T_n\, T_{n+1} p_1(T_{n+1}|T_n)\mathcal{F}_n(T_n),$$

where as previously $\mathcal{F}_n(T_n)$ is the density of the nth ISI and $p_1(T_{n+1}|T_n)$ is the conditional density of T_{n+1} given T_n. Because T_{n+1} is statistically determined only by $s_0^{(n)}$, we can define this conditional density as

FIG. 6. Conditional FPT densities $\mathcal{H}(\lambda,\nu)$ given by Eq. 17 relating the initial value of the adaptation current to the distribution of the following ISI. Computed using numerical solutions to the FPE Eq. 8. **Top:** Power-law adaptation, computed using a CN timestepping scheme. **Bottom:** Exponential adaptation, computed using an Euler timestepping scheme. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation.

function then needs to be integrated according to Eq. 18, and the PDF $\mathcal{G}_k$ can be obtained by numerical differentiation. We show results for $\mathcal{G}_2$ and $\mathcal{G}_3$ using Eq. 18 in Fig. 5. The agreement between MC simulations and Eq. 18 is excellent.

Iterating these ideas into the stationary regime, the ideas of the previous paragraph can be used to directly compute the stationary density of the peak adaptation current. We assume that stationarity is reached after $k^* - 1$ firing events. Then $s_0^{(k^*)}$ and $s_0^{(k^*+1)}$ have the same distribution. Consequently, the stationary density of the peak adaptation current after firing satisfies the two-dimensional integral equation (cf. Eq. 18)

$$p_1(T_{n+1}|T_n) = \int \mathrm{d}y\, p(T_{n+1}, y|T_n),$$

where $p(T_{n+1}, y|T_n)$ denotes the joint density of T_{n+1} and $s_0^{(n)} = y$ conditioned on the previous ISI T_n. We can rewrite this as follows:

$$\begin{aligned}
p(T_{n+1}, y|T_n) &= \frac{p_3(T_{n+1}, y, T_n)}{p(T_n)}, \\
&= \frac{p(T_{n+1}|y, T_n)p(y, T_n)}{p(T_n)}, \\
&= \frac{p(T_{n+1}|y, T_n)p(y|T_n)p(T_n)}{p(T_n)}, \\
&= p(T_{n+1}|y, T_n)p(y|T_n).
\end{aligned}$$

Now, as we have previously shown, the statistics of T_{n+1} is completely determined when $s_0^{(n)} = y$ is fixed, hence $p(T_{n+1}|y, T_n) = p(T_{n+1}|y) \equiv \mathcal{H}(T_{n+1}, y)$. Therefore, we have

$$\begin{aligned}
\mathbb{E}(T_n T_{n+1}) = \\
\int \mathrm{d}T_n\, \mathrm{d}T_{n+1}\, \mathrm{d}y\, T_n\, T_{n+1} \mathcal{H}(T_{n+1}, y)p(y|T_n)\mathcal{F}_n(T_n).
\end{aligned}$$

This can be further simplified by noting that $p(y|T_n) = \frac{p(y, T_n)}{\mathcal{F}_n(T_n)}$ and therefore

$$\mathbb{E}(T_n T_{n+1}) = \int \mathrm{d}T_n\, \mathrm{d}T_{n+1}\, \mathrm{d}y\, T_n\, T_{n+1}\, \mathcal{H}(T_{n+1}, y)p(y, T_n).$$
$$(24)$$

For $n = 1$, Eq. 24 can be simplified because $s_0^{(1)}$ is a deterministic function of T_1 (see Eq. 11), so that $p(s_0^{(1)} = y, T_1 = x) = \delta\left(y - f(x, s_0^{(0)})\right)\mathcal{F}_1(x)$, where f is defined by Eqs. 14 and 15. Hence, we have for $n = 1$

$$\mathbb{E}(T_1 T_2) = \int \mathrm{d}T_1\, \mathrm{d}T_2\, T_1\, T_2\, \mathcal{H}(T_2, f(T_1, s_0^{(0)}))\mathcal{F}_1(T_1).$$

We have shown in Section V how to obtain the conditional FPT density $\mathcal{H}$. To evaluate Eq. 24 for general n, we still need to compute the joint density $p\left(s_0^{(n)} = y, T_n\right)$. This can be achieved by means of an MC simulation, where we fix a value of n and then record the frequency with which pairs of $s_0^{(n)}$ and T_n are generated by the system. We show an example of these densities in Fig. 7. The most notable feature is an inverse proportionality between $s_0^{(n)}$ and T_n. The longer e.g. T_2, the less likely it is for the value of s after the second firing, $s_0^{(2)}$, to attain a high value.
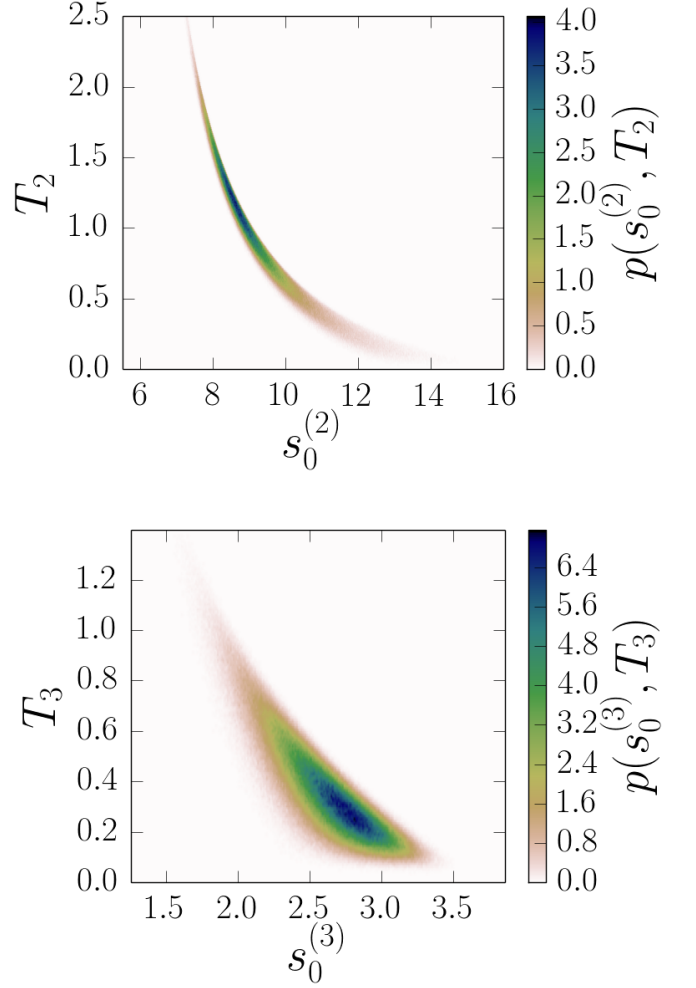

FIG. 7. MC simulations with GS boundary correction for the joint density of $s_0^{(n)}$ (i.e. nth peak value of adaptation current following the nth FPT) and T_n for different values of n. **Top:** Power-law adaptation, $n = 2$. **Bottom:** Exponential adaptation, $n = 3$. $M = 10^6$ MC realizations. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation.

Eq. 24 is formally correct, but not very practical for actual computations. This is because to apply the iFPT approach, it is desirable to obtain all quantities needed for the SCC using solutions of the FPE only, and no MC simulations. These, however, are required to obtain an approximation for the joint density $p(s_0^{(n)}, T_n)$ in Eq. 24. We therefore propose an approximation to compute $\mathbb{E}(T_n T_{n+1})$ using the available densities $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ only. To that end, we note that

$$p_3(T_{n+1}, T_n, y) = p(T_{n+1}, T_n|y)p(y). \qquad (25)$$

If we now assume that T_n and $s_0^{(n)}$ are independent, we can approximate this as follows:

$$p_3(T_{n+1}, T_n, y) \approx p(T_{n+1}|y)p(y)p(T_n) = \mathcal{H}(T_{n+1}, y)\mathcal{G}_n(y)\mathcal{F}_n(T_n). \tag{26}$$

This results in an alternative, approximative expression for the expectation $\mathbb{E}(T_n T_{n+1})$:

$$\mathbb{E}(T_n T_{n+1}) = \int dT_n dT_{n+1} dy\, T_n\, T_{n+1}\, \mathcal{H}(T_{n+1}, y)\mathcal{G}_n(y)\mathcal{F}_n(T_n). \tag{27}$$

Eq. 27 is therefore equivalent to Eq. 24 if $p(y, T_n) = \mathcal{G}_n(y)\mathcal{F}_n(T_n)$. Although Fig. 7 demonstates that p does not factorise (the joint density is negatively sloped), we show below that Eq. 27 approximates Eq. 24 very well. Given that Eq. 27 does not require additional MC simulations, the small error introduced by Eq. 27 is well offset by the large reduction in computational cost.

There exists a third alternative expression for the expectation of the product of ISIs suggested for a different, but related, model, in [43]. It reads in our notation [44]

$$\mathbb{E}(T_n T_{n+1}) = \int dT_n\, dT_{n+1}\, dy\, T_n\, T_{n+1}\, \mathcal{H}(T_{n+1}, f(T_n, y))\mathcal{H}(T_n, y)\mathcal{G}_{n-1}(y), \tag{28}$$

where f is given by Eq. 14 or Eq. 15. The term $\mathcal{H}(T_n, y)\mathcal{G}_{n-1}(y)$ that appears in Eq. 28 is the same as the one on the right-hand side in Eq. 18, which we used to obtain the CDF of $s_0^{(n)}$.

For $n = 1$, Eq. 28 reads

$$\mathbb{E}(T_1 T_2) = \int dT_1\, dT_2\, T_1\, T_2\, \mathcal{H}\left(T_2, f\left(T_1, s_0^{(0)}\right)\right)\underbrace{\mathcal{H}\left(T_1, s_0^{(0)}\right)}_{=\mathcal{F}_1(T_1)},$$

because s is started from a point $s_0^{(0)}$, so that formally $\mathcal{G}_0(y) = \delta\left(y - s_0^{(0)}\right)$, which collapses the integration over y in Eq. 28. Thus, for $n = 1$, Eq. 24 and 28 coincide. This is also true for higher values of n. A proof for this equivalence is presented in Appendix A. Eq. 28 only makes use of the quantities $\mathcal{H}$ and $\mathcal{G}$, which can be computed using the iFPT approach as explained in the previous section. We show comparisons between MC simulations and the three expressions for $\mathbb{E}(T_n T_{n+1})$, Eq. 24, Eq. 27 and Eq. 28, in Fig. 8. The results presented in Fig. 8 are in agreement with the observation that the two expressions Eq. 24 and Eq. 28 are equivalent. We find that the agreement of Eq. 28 with MC simulations is comparable to Eq. 27, particularly for exponential adaptation. Interestingly, the formally correct Eq. 24 and the approximate Eq. 27 give comparable results; Eq. 24 slightly deviates

from MC simulations and Eq. 27 when n gets larger. The maximal relative disagreement between MC and iFPT results is less than 2% (Fig. 8, bottom panels). We will see below that the SCC is best approximated by using the exact result Eq. 24 (or equivalently Eq. 28), as we expect. We attribute the discrepancy between MC simulations and the exact result Eq. 24 to the error caused by the numerical integration over the MC approximation of the joint density $p(y, T_n)$. We checked that applying a kernel density estimation [45] to the MC results for $p(y, T_n)$ did not alter these results.

Similar results for $\mathcal{Q}_1$ and $\mathcal{Q}_2$ (Eqs. 22 and 23) are shown in Figs. 9 and 10. The agreement is good, with the maximal relative disagreement always less than 5%. The relative disagreement for the statistics of the product of two adjacent ISIs, $\mathcal{Q}_1(n, 1)$, is in general larger than the error for the moments, as can be seen by comparing Figs. 2 and 3 with Fig. 9. Indeed, for the case of power-law adaptation, we observe an increase of roughly one order of magnitude in the relative error even when the more accurate CN scheme is used (see e.g. left panels of Fig. 9). An exception is the computation of the joint expectation, shown in Fig. 8, where, depending on which methods are compared, the relative disagreement is comparable in size to the one for the computation of the moments shown in Figs. 2 and 3. This increase of the relative disagreement makes the computation of the SCC using the iFPT approach hard, because the two expressions in the numerator of Eq. 21 are quite close to one another for the parameter values we have chosen here, meaning that the numerator is small and indeed of the same magnitude or even smaller as the relative disagreement, e.g. $-2.6 \cdot 10^{-2}$ in the left panel and $-8.6 \cdot 10^{-4}$ in the right panel of Fig. 8 for $n = 1$ for the MC-GS simulation method. The increase in the relative discrepancy is caused by error propagation, because for the second-order statistics, one has to multiply two quantities that both come with an individual error. Therefore, whereas the iFPT approach can in principle also be used to compute serial correlations present in the spike train, obtaining reliable results can in general be a computational challenge. When the negative serial correlations are stronger, so that the difference in the numerator of Eq. 21 is larger, the iFPT approach should give more accurate results. We stress that the dominant source of error is not the computation of the joint expectation $\mathbb{E}(T_n T_{n+1})$ of ISIs, but the product of the expectation of ISIs and the variances, which can be seen by comparing the lower panels of Fig. 8 with those of Figs. 9 and 10.

Finally, we show the SCC at lag 1 obtained by MC simulations and PDE numerics in Fig. 11. The agreement is worse than for all previously considered quantities, but still reasonable. To verify the MC simulations, we checked that our MC simulation setup was able to reproduce known analytical results for the SCC obtained in [15] for certain limiting cases. The anti-correlations between adjacent ISIs ($\mathrm{SCC}(n, 1) < 0$) strengthen until they reach a stationary value.
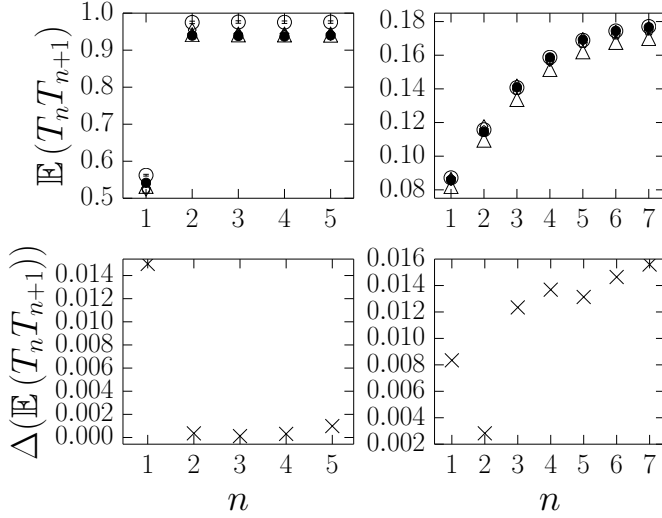
FIG. 8. **Top:** Expectation $\mathbb{E}(T_n T_{n+1})$ of adjacent ISIs for different values of n. Left: power-law adaptation. Right: exponential adaptation. $M = 10^6$ MC realizations. Empty circles: Plain MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Pentagons: Eq. 24. For power-law adaptation, the pentagons are on top of the empty triangles and filled circles and hence not visible. Filled circles: Eq. 27. Diamonds: Eq. 28. The diamonds are nearly on top of the filled circles and hence not visible. The PDE results were obtained using a CN scheme (power-law adaptation) or an Euler timestepping scheme (exponential adaptation). The vertical error bars show the MC error for a $> 99.99\%$ confidence interval. **Bottom:** Relative disagreements defined by Eq. 13, where for power-law adapation, the GS boundary corrected MC algorithm was used, and for exponential adaptation, the plain MC algorithm was used. For the iFPT quantity, Eq. 24 was used. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation.

FIG. 9. **Top:** $\mathcal{Q}_1(n,1)$ (defined in Eq. 22) of adjacent ISIs for different values of n. Left: power-law adaptation. Right: exponential adaptation. Empty circles: MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Filled circles: Eq. 22, CN timestepping scheme for power-law adaptation and Euler timestepping scheme for exponential adaptation. **Bottom:** Relative disagreement defined by Eq. 13, a plain MC algorithm was used to obtain the relative disagreement. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation.

Thus, we see that MC and PDE results for the SCC in general do not agree as well as one would expect from the good agreement of the expecations $\mathbb{E}(T_n T_{n+1})$ in Fig. 8. The deviation is likely more pronounced for parameters that lead to small negative SCCs, which we have for both models considered in this section. In the next Section, we will compare this with results for the perfect integrate-and-fire model, where parameter values are chosen so that the SCCs are more negative and hence the agreement is better. This is because the two terms in the numerator of Eq. 21 are close to each other for small SCCs, and hence a small error in them impacts the accuracy of the SCC computation quite dramatically.

We show in the next section that our methods reproduce known stationary analytical results for the SCC when we consider the perfect integrate-and-fire model with single exponential adaptation in a parameter regime where we have large negative correlations, thus demonstrating that our methods are sound, but SCC calculations are very sensitive to numerical inaccuracies.
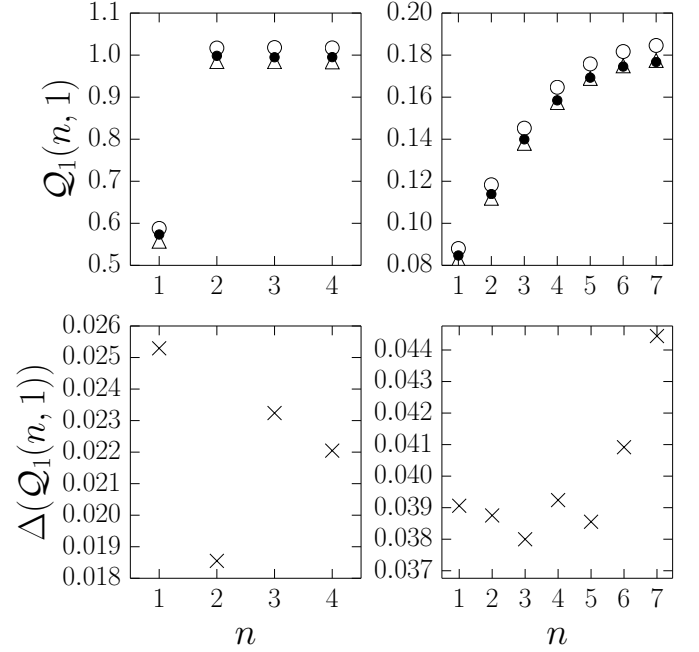
A. The perfect integrate-and-fire model

The adapting perfect integrate-and-fire (PIF) model driven by white Gaussian noise and a single exponential adaptation current is one of the simplest models for spike-triggered adaptation. For small noise intensity, analytical expressions for the stationary SCC exist. We here study this stationary limit case and compare analytical formulas to results obtained with the iFPT approach.

The model reads (we follow the notation of [46] and [47])

$$\mathrm{d}X = (I_0 - s)\mathrm{d}t + \sqrt{2D}\mathrm{d}W(t)\,, \qquad (29)$$

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\frac{s}{\tau_a}\,. \qquad (30)$$

The adaptation mechanism works in analogy to the previous model (Eq. 2): whenever X reaches the threshold $X = 1$, s receives a kick of size $\Delta \equiv \frac{\tilde{\Delta}}{\tau_a}$ and X is instantaneously reset to 0.

The stationary SCC at lag 1 for this model under the assumption of small noise (i.e. $D \ll 1$) reads [46]
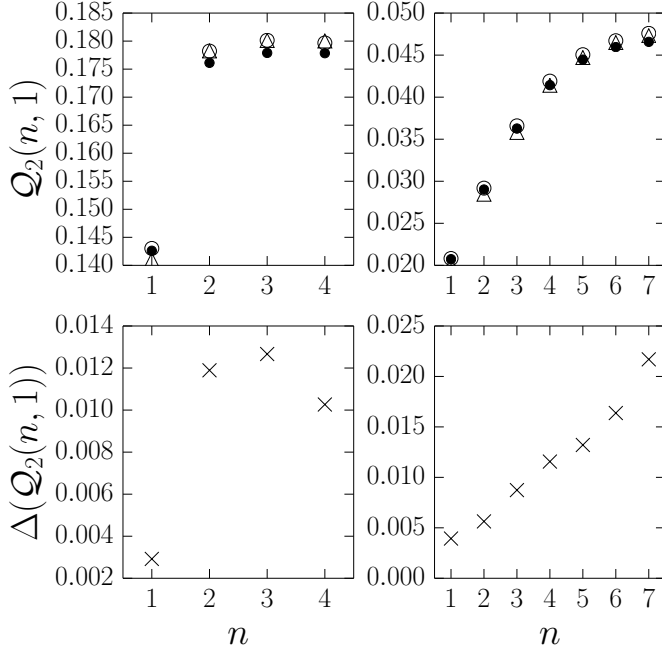
FIG. 10. **Top:** $\mathcal{Q}_2(n,1)$ (defined in Eq. 23) for adjacent ISIs for different values of n. Left: power-law adaptation. Right: exponential adaptation. Empty circles: MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Filled circles: Eq. 23, CN timestepping scheme for power-law adaptation and Euler timestepping scheme for exponential adaptation. **Bottom:** Relative disagreement defined by Eq. 13, a plain MC algorithm was used to obtain the relative disagreement. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation.

FIG. 11. Serial correlation coefficient at lag $n=1$ (defined in Eq. 21) for adjacent ISIs for different values of n for the LIF model (Eq. 1). **Left:** power-law adaptation. **Right:** exponential adaptation. Empty circles: Plain MC simulations of Eqs. 1 and 2. Triangles: MC simulations with GS boundary correction. Filled circles: PDE results, SCC computed using Eq. 27. Pentagons: PDE results, SCC computed using Eq. 24. CN timestepping scheme for power-law adaptation and Euler timestepping scheme for exponential adaptation. Diamonds: PDE results, SCC computed using Eq. 28. Parameter values as in Fig. 2 for power-law adaptation and as in Fig. 3 for exponential adaptation. Even if the joint expectations shown in Fig. 8 agree well, this does not imply that the SCC will be well approximated; the small correlation values (i.e. the difference between $\mathbb{E}(T_nT_{n+1})$ and $\mathcal{Q}_1(n,1)$) for the two examples lead to large discrepancies in the SCCs, which are especially significant for the case of power-law adaptation.

$$\text{SCC}(k=1) = -\frac{\alpha(1-\theta)(1-\alpha^2\theta)}{1+\alpha^2-2\alpha^2\theta}, \qquad (31)$$

where

$$\alpha = \frac{s^*-\Delta}{s^*}, \qquad \theta = \frac{I_0-s^*}{I_0-s^*+\Delta},$$
$$T^* = \frac{1+\widetilde{\Delta}}{I_0}, \qquad s^* = \frac{\Delta}{1-\exp\left(-\frac{T^*}{\tau_a}\right)}.$$

Thus, we can compute the SCC in closed analytical form as a function of the system parameters. This formula serves as an important benchmark for our numerical results. In particular, we expect that after the described transition to stationarity, the SCC given by Eq. 21 will approach the stationary SCC given by Eq. 31. This is confirmed in Fig. 12. In particular, the agreement between MC simulations and the exact formula Eq. 24 is very good (the relative disagreement between PDE numerics and the analytical result is less than 6% for the
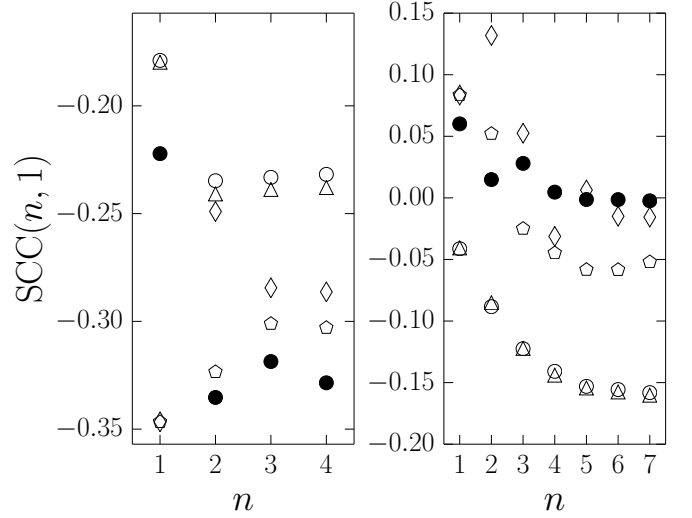
stationary value); the agreement of MC simulations with the approximation Eq. 27 is a bit worse, but still reasonable. Thus, we conclude that our methodology can be used more generally to compute the evolution of the moments and SCCs. However, as seen in the previous section, to obtain a good agreement between MC simulations and PDE numerics, the computational effort might be rather large. In particular, we note that the PIF example shown in Fig. 12 gives rise to stronger negative SCCs, which means that the error propagation has less of an effect, but is still present, even when moments of firing times between MC and PDE numerics disagree by less than 1% (data not shown). We finally note that it is also possible to analytically compute the stationary SCC at higher lags and for different models (e.g. the leaky integrate-and-fire model in the presence of weak noise or for small adaptation currents) using the approach described in [47], or, using a different approach, in [15].
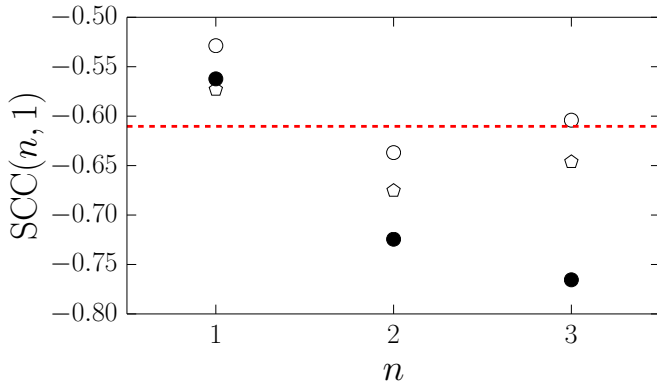
FIG. 12. Serial correlation coefficient at lag $n = 1$ (defined in Eq. 21) for different values of n for the PIF model (Eq. 30). The dashed horizontal line is the stationary SCC given by Eq. 31. Empty circles: Plain MC simulations of Eq. 30. Pentagons: Eq. 24. Filled circles: Eq. 27. The PDE results were obtained using a CN timestepping scheme. $M = 10^6$ MC realizations. Timestep $h = 10^{-3}$. Parameter values: $D = 0.1$, $\tau_a = 5.0$, $\widetilde{\Delta} = 10$, $I_0 = 5.5$, $s_0^{(0)} = 5.0$.

VII. SUMMARY AND CONCLUSIONS

In this paper, we have developed a numerical method for the computation of moments and correlations in general two-dimensional non-renewal escape time processes. Our approach relies on the numerical solution of a two-dimensional time-dependent FPE with initial conditions obtained from marginal distributions of previous states of the system. Crucially, the computation scheme presented in this study is general insofar as it can be applied to any stochastic process with a known reset condition (Eq. 1) and any deterministic signal (Eq. 2). As an important application, we have described the transition to stationarity in a stochastic IF neuron model with spike-triggered adaptation, which causes non-trivial ISI correlations. A different mechanism for introducing positive correlations between ISIs has recently been reported in [28] and can equally well be analyzed with the presented methodology. Moreover, our approach enables us to determine the non-trivial timescale of transition to a stationary adapted state by counting the number of intervals needed for this transition.

Experimentally, the transition to stationarity is often characterized by the behaviour of the instantaneous firing rate [17, 18] [48]. The instantaneous firing rate is usually obtained by averaging the neuronal activity bin-wise for a fixed time. This differs from the firing rate used here as given by the inverse of the mean ISI (Eq. 5). In other words, while the instantaneous firing rate is measured in real time, our firing rate relates to interval numbers. This entails that for a given time t, the firing rate contains contributions from, in general, past firing events that may have occurred at any point k in the spike train. Knowing the joint distributions of all ISIs T_k, it is at least in princi-

ple possible to reconstruct the instantaneous firing rate, whereas given the instantaneous firing rate, we cannot reconstruct the joint distributions of the individual ISIs T_k. Despite the difference in the definition of the firing rate, it might be an interesting topic for further study to classify the time scales of the transition to stationarity both experimentally and based on the theory presented here.

The computation of ISI moments using the iFPT approach is computationally inexpensive, giving rise to small relative disagreements between solutions of the FPE and direct MC simulations. In contrast, the computation of correlations is harder. We observed that we lost one order of magnitude in accuracy compared to the simulation of the moments for the quantities $\mathcal{Q}_1$ (Eq. 22) and $\mathcal{Q}_2$ (Eq. 23), which makes the reliable computation of SCCs a computationally challenging task. We conclude that even a relative disagreement of ISI moments between Monte Carlo estimations and PDE solutions of the order 10^{-3} is not enough to reliably estimate the SCC using PDE numerics only (but this might be specific for the examples we have considered), indicating that more refined numerical methods or larger computational ressources, or indeed both, are needed. When the difference between the two terms in the numerator of Eq. 21 is large, the small error made by the numerical solution of the PDE should have a less detrimental influence on the final result. The need for more refined numerical methods is further substantiated by the fact that the more accurate asymptotically stable CN timestepping scheme did not result in a significant decrease in the relative disagreement between PDE results and both plain MC and MC-GS simulations, for both moments of firing intervals and the SCC. In this paper, we have only discussed the error associated with MC simulations, because it is readily available. The numerical solution of the FPE is of course also subject to numerical errors and future work will likely benefit from a discussion about how to systematically reduce these errors. In this context, it might be beneficial to compare the finite-element methods used here to other methods for solving PDEs, such as finite difference and finite volume methods [49]. A systematic error estimation study might be made more difficult by the fact that the diffusion matrix (Eq. 9) is not positive definite [50, 51].

There is an alternative method to compute the ISI distributions given the distributions of the peak adaptation currents using the formula

$$\mathcal{F}_k(t) = \int_{\text{supp}(\mathcal{G}_{k-1})} \mathcal{H}(t, y)\mathcal{G}_{k-1}(y)\mathrm{d}y. \qquad (32)$$

This is an integral equation frequently used in the context of randomized FPT problems [52, 53], where usually $\mathcal{F}_k$ and the kernel $\mathcal{H}$ are given, and one tries to find a matching distribution $\mathcal{G}_{k-1}$ of starting points. Using Eq. 32, we do not have to solve a time-dependent PDE for each ISI, but must compute $\mathcal{H}$ once as the solution

of a time-dependent FPE with varying initial conditions for s, similar to the computation of $\mathcal{F}_1$. The averaging that the iFPT approach amounts to is particularly clear in this formulation. The densities $\mathcal{G}_k$ are obtained as discussed above (see Eq. 18). The approach relying on Eq. 32 might be computationally less expensive, but we found that it is not as exact as solving a time-dependent FPE for each ISI, especially at larger times. This is likely caused by errors when computing $\mathcal{H}$, as the numerical integration in Eq. 32 can be performed accurately and efficiently. However, Eq. 32 could be useful for analytical explorations when $\mathcal{H}$ is known.

We finally emphasize that our approach did not use the complicated boundary conditions for stationary IF models, where the probability *flux* at threshold gives rise to a discontinuity of the probability flux at reset [9, 54]. In contrast, our approach allows for the computation of transient and stationary distributions of the adaptation dynamics in an iterative fashion, requiring the solution of a two-dimensional time-dependent PDE. The only boundary condition that has to be taken into account is an absorbing boundary condition for the probability *density* at the threshold x_{th}. This makes the problem tractable using finite-element approximation methods for time-dependent PDEs, resulting in a general description of two-dimensional IF models with spike-triggered adaptation. The approach we have described in this paper can in principle also be used to gain analytical insight into these system, however, quantities such as $\mathcal{H}$ and the solution of a two-dimensional time-dependent PDE seem to be unavailable in closed analytical form except in the most simple cases.

ACKNOWLEDGMENTS

Appendix A: Equivalence of Eqs. 24 and 28

We here show that Eq. 24 and Eq. 28 are equivalent. We recall Eq. 24:

$$\mathbb{E}(T_n T_{n+1}) = \tag{A1}$$
$$\int \mathrm{d}T_n \, \mathrm{d}T_{n+1} \, \mathrm{d}s_0^{(n)} \, T_n \, T_{n+1} \, \mathcal{H}(T_{n+1}, s_0^{(n)}) p(s_0^{(n)}, T_n).$$

We re-write Eq. 28 as follows:

$$\mathbb{E}(T_n T_{n+1}) =$$
$$\int \mathrm{d}T_n \, \mathrm{d}T_{n+1} \, \mathrm{d}y \, T_n \, T_{n+1} \, \mathcal{H}(T_{n+1}, f(T_n, y)) \, p(y, T_n),$$
$$\tag{A2}$$

where $y = s_0^{(n-1)}$ and we have replaced $\mathcal{H}(T_n, y)\mathcal{G}_{n-1}(y) = p(y, T_n)$. Note that in Eq. A1, p is the joint density of $s_0^{(n)}$ and T_n, whereas p is the joint density of $s_0^{(n-1)}$ and T_n in Eq. A2.

By inspection, the two expressions are identical if we can show that

$$\int \mathrm{d}s_0^{(n)} \mathcal{H}(T_{n+1}, s_0^{(n)}) p(s_0^{(n)}, T_n) =$$
$$\int \mathrm{d}s_0^{(n-1)} \mathcal{H}(T_{n+1}, f(T_n, s_0^{(n-1)})) p(s_0^{(n-1)}, T_n),$$
$$\tag{A3}$$

for T_n and T_{n+1} fixed.

Starting from the second line in Eq. A3, we change the integration variable from $s_0^{(n-1)}$ to $s_0^{(n)}$ by observing that from $s_0^{(n)} = f(T_n, s_0^{(n-1)})$, we have $\frac{\mathrm{d}s_0^{(n)}}{\mathrm{d}s_0^{(n-1)}} = \frac{\partial f}{\partial s_0^{(n-1)}}$ and therefore $\mathrm{d}s_0^{(n-1)} = \mathrm{d}s_0^{(n)} \left(\frac{\partial f}{\partial s_0^{(n-1)}} \right)^{-1}$. We need to assume that f is invertible with respect to the second argument, which is the case for both power-law (Eq. 14) and exponential adaptation (Eq. 15) considered in this paper. The integral then becomes

$$\int \mathrm{d}T_n \, \mathrm{d}T_{n+1} \, \mathrm{d}s_0^{(n)} \, T_n \, T_{n+1} \, \mathcal{H}\left(T_{n+1}, s_0^{(n)}\right) p \left[f^{-1}(T_n, s_0^{(n)}), T_n \right] \left(\frac{\partial f}{\partial s_0^{(n-1)}} \right)^{-1}. \tag{A4}$$

But $p\left[f^{-1}(T_n, s_0^{(n)}), T_n \right] \left(\frac{\partial f}{\partial s_0^{(n-1)}} \right)^{-1}$ is nothing but the transformation from $p(s_0^{(n-1)}, T_n)$ to $p(s_0^{(n)}, T_n)$. Indeed, we have (fixing T_n)

$$p(s_0^{(n)}, T_n) \frac{\partial s_0^{(n)}}{\partial s_0^{(n-1)}} = p(s_0^{(n-1)}, T_n), \tag{A5}$$

so that

$$p(s_0^{(n)}, T_n) = p\left[f^{-1}(T_n, s_0^{(n)}), T_n\right] \left(\frac{\partial f}{\partial s_0^{(n-1)}}\right)^{-1}.$$

(A6)

Therefore, Eq. 24 and Eq. 28 are equivalent.

[1] A. Palacios, J. Aven, P. Longhini, V. In, and A. R. Bulsara, Phys. Rev. E **74**, 021122 (2006).

[2] A. Aragoneses, L. Carpi, N. Tarasov, D. V. Churkin, M. C. Torrent, C. Masoller, and S. K. Turitsyn, Phys. Rev. Lett. **116**, 033902 (2016).

[3] M. J. Chacron, B. Lindner, and A. Longtin, Phys. Rev. Lett. **92** (2004).

[4] J. A. Reinoso, M. C. Torrent, and C. Masoller, Phys. Rev. E **94**, 032218 (2016).

[5] S. Coombes, R. Thul, J. Laudanski, A. R. Palmer, and C. J. Sumner, Frontiers in Life Science **5**, 91 (2011).

[6] W. H. Nesse, L. Maler, and A. Longtin, Proc. Nat. Acad. Sci. USA **107**, 21973 (2010).

[7] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Methuen's Monographs on Applied Probability and Statistics (John Wiley, London, 1966).

[8] R. Rosenbaum, Front. Comput. Neurosci. **10** (2016).

[9] T. Schwalger and B. Lindner, Phys. Rev. E **92**, 062703 (2015).

[10] Y.-H. Liu and X.-J. Wang, Journal of Computational Neuroscience **10**, 25 (2001).

[11] E. Muller, L. Buesing, J. Schemmel, and K. Meier, Neural Computation **19**, 2958 (2007).

[12] L. Hertäg, D. Durstewitz, and N. Brunel, Frontiers in Computational Neuroscience **8** (2014), 10.3389/fncom.2014.00116.

[13] T. Schwalger, K. Fisch, J. Benda, and B. Lindner, PLoS Comput. Biol. **6** (2010).

[14] K. Fisch, T. Schwalger, B. Lindner, A. Herz, and J. Benda, J Neurosci. **32**, 17332 (2012).

[15] E. Urdapilleta, Phys. Rev. E **84**, 041904 (2011).

[16] E. Urdapilleta, Europhysics Letters **115** (2016).

[17] R. Naud and W. Gerstner, PLOS Computational Biology **8**, 1 (2012).

[18] G. La Camera, A. Rauch, D. Thurbon, H.-R. Lüscher, W. Senn, and S. Fusi, Journal of Neurophysiology **96**, 3448 (2006).

[19] N. Ulanovsky, L. Las, D. Farkas, and I. Nelken, Journal of Neuroscience **24**, 10440 (2004).

[20] C. Kuehn, *Multiple Time Scale Dynamics*, 1st ed., Applied Mathematical Sciences No. 191 (Springer-Verlag, 2015).

[21] P. J. Drew and L. Abbott, J Neurophysiol **96**, 826 (2006).

[22] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall, Nat. Neurosci. **11**, 1335 (2008).

[23] C. Pozzorini, R. Naud, S. Mensi, and W. Gerstner, Nat. Neurosci. **16**, 942 (2013).

[24] S. E. Clarke, R. Naud, A. Longtin, and L. Maler, Proc. Nat. Acad. Sci. USA **110**, 13624 (2013).

[25] R. Kobayashi, Y. Tsubo, and S. Shinomoto, Front. Comput. Neurosci. **3**, 121 (2009).

[26] W. Gerstner and R. Naud, Science **326**, 379 (2009).

[27] S. Redner, *A Guide to First-Passage Processes*, 1st ed. (Cambridge University Press, 2010).

[28] G. D'Onofrio, E. Pirozzi, and M. Magnasco, in *EUROCAST 2015, LNCS 9520*, edited by R. M.-D. et al. (Springer-Verlag, Berlin Heidelberg, 2015) pp. 166–173.

[29] A. N. Burkitt, Biol. Cybern. **95**, 1 (2006).

[30] J. Benda and A. V. M. Herz, Neural Computation **15**, 2523 (2003).

[31] L. Sacerdote and M. T. Giraudo, in Stochastic Biomathematical Models, (M.Bachar et al. (eds.)), Lecture Notes in Mathematics 2058, Chapter 5. , 99 (2013).

[32] C. van Vreeswijk, in *Analysis of Parallel Spike Trains*, Springer Series in Computational Neuroscience, Vol. 7, edited by S. Grün and S. Rotter (Springer Verlag, 2010) pp. 3–20.

[33] H. Risken, *The Fokker-Planck equation: Methods of solution and applications*, 2nd ed. (Springer, 1989).

[34] B. F. Spencer Jr. and L. Bergman, Nonlinear Dynamics **4**, 357 (1993).

[35] M. T. Giraudo and L. Sacerdote, Communications in Statistics - Simulation and Computation **28**, 1135 (1999).

[36] B. Lindner and A. Longtin, Journal of Theoretical Biology **232**, 505 (2005).

[37] E. Gobet, Stochastic Processes and their Applications **87**, 167 (2000).

[38] D. Higham, X. Mao, M. Roj, Q. Song, and G. Yin, SIAM/ASA Journal on Uncertainty Quantification **1**, 2 (2013).

[39] A. Logg, K.-A. Mardal, and G. Wells, eds., *Automated Solution of Differential Equations by the Finite Element Method*, Lecture Notes in Computational Science and Engineering, Vol. 84 (Springer Verlag, 2012).

[40] P. Knabner and L. Angermann, *Numerical methods for elliptic and parabolic partial differential equations*, Texts in Applied Mathematics, Vol. 44 (Springer Verlag, New York Berlin Heidelberg, 2003).

[41] The support of $\mathcal{G}_k$ is the open interval $(0, s_0^{(0)} + k \cdot \kappa)$.

[42] R. Mankin and A. Rekker, Phys. Rev. E **94**, 062103 (2016).

[43] M. J. Chacron, K. Pakdaman, and A. Longtin, Neural Computation **15**, 253 (2003).

[44] Note that Eq. 3.17 in [43] is in the stationary state: $\mathbb{E}(T_n T_{n+1}) = \int dT_n dT_{n+1} dy\, T_n T_{n+1} \mathcal{H}(T_{n+1}, f(T_n, y)) \mathcal{H}(T_n, y) \mathcal{G}^*(y)$.

[45] "Scipy v0.18.1 reference guide for Gaussian kernel density estimation," http://docs.scipy.org/doc/scipy-0.18.1/reference/generated/scipy.stats.gaussian_kde.html, accessed: 2017-01-31.

[46] T. Schwalger, *The interspike-interval statistics of non-renewal neuron models*, Ph.D. thesis, Humboldt-Universität zu Berlin (2013).

[47] T. Schwalger and B. Lindner, Frontiers in Computational

Neuroscience **7** (2013).

[48] Note that in [18], the timescale for single exponential adaptation was *inferred* from the time course of the numerically obtained instantaneous firing rate, showing that the time course of the rate is not described by the same single exponential time course of the adaptation current. A similar observation is made in [30].

[49] R. Rosenbaum, F. Marpeau, J. Ma, A. Barua, and K. Josić, Journal of Mathematical Biology **65**, 1 (2012).

[50] S. Salsa, *Partial Differential Equations in Action*, 1st ed.

(Springer-Verlag, 2008).

[51] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, 1st ed., Applied Mathematical Sciences, Vol. 159 (Springer-Verlag New York, 2004) p. 526.

[52] K. Jackson, A. Kreinin, and W. Zhang, Statistics and Probability Letters **79**, 2422 (2009).

[53] S. Jaimungal, A. Kreinin, and A. Valov, SIAM Theory Probabl. Appl. **58**, 493 (2014).

[54] N. Brunel and S. Sergi, J. Theor. Biol. **195**, 87 (1998).