On the Hamiltonian structure of large deviations in stochastic hybrid systems

J. Stat. Mech. (2017) 033206

(http://iopscience.iop.org/1742-5468/2017/3/033206)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

On the Hamiltonian structure of large deviations in stochastic hybrid systems

Paul C Bressloff[1] **and Olivier Faugeras**[2]

[1] Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, Utah 84112, United States of America
[2] NeuroMathComp Group, INRIA, Sophia Antipolis, France
E-mail: bressloff@math.utah.edu and olivier.faugeras@inria.fr

Received 19 October 2016, revised 3 January 2017
Accepted for publication 24 February 2017
Published 23 March 2017

Online at stacks.iop.org/JSTAT/2017/033206
https://doi.org/10.1088/1742-5468/aa64f3

Abstract. We present a new derivation of the classical action underlying a large deviation principle (LDP) for a stochastic hybrid system, which couples a piecewise deterministic dynamical system in $\mathbb{R}^d$ with a time-homogeneous Markov chain on some discrete space Γ. We assume that the Markov chain on Γ is ergodic, and that the discrete dynamics is much faster than the piecewise deterministic dynamics (separation of time-scales). Using the Perron–Frobenius theorem and the calculus-of-variations, we show that the resulting action Hamiltonian is given by the Perron eigenvalue of a $|\Gamma|$-dimensional linear equation. The corresponding linear operator depends on the transition rates of the Markov chain and the nonlinear functions of the piecewise deterministic system. We compare the Hamiltonian to one derived using WKB methods, and show that the latter is a reduction of the former. We also indicate how the analysis can be extended to a multi-scale stochastic process, in which the continuous dynamics is described by a piecewise stochastic differential equations (SDE). Finally, we illustrate the theory by considering applications to conductance-based models of membrane voltage fluctuations in the presence of stochastic ion channels.

Keywords: dynamical processes, fluctuation phenomena, large deviations in non-equilibrium systems, noise models

1742-5468/17/033206+30$33.00

Contents

1. Introduction

There are a growing number of problems in biology that involve the coupling between a piecewise deterministic dynamical system in $\mathbb{R}^d$ and a time-homogeneous Markov chain on some discrete space Γ [4], resulting in a type of stochastic hybrid system (SHS) known as a piecewise deterministic Markov process (PDMP) [10][3]. One important example at the single-cell level is the occurrence of membrane voltage fluctuations in neurons due to the stochastic opening and closing of ion channels [1, 5, 8, 9, 16, 19, 22, 33, 35, 37, 42]. Here the discrete states of the ion channels evolve according to a continuous-time Markov process with voltage-dependent transition rates and, in-between discrete jumps in the ion channel states, the membrane voltage evolves according to a deterministic equation that depends on the current state of the ion channels. In the limit that the number of ion channels goes to infinity, one can apply the

[3] In this paper we use the term SHS mainly within the restricted sense of a PDMP. However, we do briefly consider a more general class of SHS in section 3, in which the piecewise deterministic continuous process is replaced by a piecewise stochastic continuous process.

J. Stat. Mech. (2017) 033206

law of large numbers and recover classical Hodgkin–Huxley type equations. However, finite-size effects can result in the noise-induced spontaneous firing of a neuron due to channel fluctuations. Another major example is a gene regulatory network, where the continuous variable is the concentration of a protein product and the discrete variable represents the activation state of the gene [2, 23, 30, 34, 36]. A third example concerns a stochastic formulation of synaptically-coupled neural networks that has a mathematical structure analogous to gene networks [3, 4].

In many of the above examples, one finds that the transition rates between the discrete states $n \in \Gamma$ are much faster than the relaxation rates of the piecewise deterministic dynamics for $x \in \mathbb{R}^d$. Thus there is a separation of time scales between the discrete and continuous processes, so that if t is the characteristic time-scale of the relaxation dynamics then t/ϵ is the characteristic time-scale of the Markov chain for some small positive parameter ϵ. Assuming that the Markov chain is ergodic, in the limit $\epsilon \to 0$ one obtains a deterministic dynamical system in which one averages the piecewise dynamics with respect to the corresponding unique stationary measure. This then raises the important problem of characterizing how the law of the underlying stochastic process approaches this deterministic limit in the case of weak noise, $0 < \epsilon \ll 1$.

A rigorous mathematical approach to addressing the above issue is *large deviation theory*, which has been developed extensively within the context of stochastic differential equations (SDEs) [11, 17, 40]. In particular, consider some random dynamical system in $\mathbb{R}^d$ for which there exists a well defined probability density functional or law $P_\epsilon[x]$ over the different sample trajectories $\{x(t)\}_0^T$ in a given time interval $[0,\ T]$. Here ϵ is a small parameter that characterizes the noise level, with $x(t)$ given by the solution $x^*(t)$ of some ODE $\dot{x} = F(x)$ in the limit $\epsilon \to 0$. A large deviation principle (LDP) for the random paths of the SDE over some time interval $[0,\ T]$ is

$$P_\epsilon[x] \sim \mathrm{e}^{-J_T[x]/\epsilon}, \quad \epsilon \to 0,$$

where $J_T[x]$ is known as the rate function and $J_T[x^*] = 0$. In the case of SDEs, the rate function can be interpreted as a classical action with corresponding Lagrangian L [17],

$$J_T[x] = \int_0^T L(x, \dot{x})\mathrm{d}t.$$

Such a Lagrangian formulation is more amenable to explicit calculations. In particular, it can be used to solve various first passage time problems associated with the escape from a fixed point attractor of the underlying deterministic system in the weak noise limit. This involves finding the most probable paths of escape, which minimize the action with respect to the set of all trajectories emanating from the fixed point. Evaluating the action along a most probable path from the fixed point to another point x generates a corresponding *quasipotential* $\Phi(x)$. From classical variational analysis, it can be shown that the quasipotential satisfies a Hamilton–Jacobi equation $H(x, \partial_x\Phi) = 0$, where H is the Hamiltonian obtained from L [17] via a Fenchel–Legendre transformation:

$$H(x, p) = \sup_y \{ py - L(x, y)\}.$$

J. Stat. Mech. (2017) 033206

The optimal paths of escape correspond to solutions of Hamilton's equations on the zero energy surface ($H = 0$). This is a consequence of the fact that the paths of escape are constant energy solutions that converge to a stable fixed point of the underlying deterministic system in the limit $t \to -\infty$, and the Hamiltonian is zero at fixed points. Interestingly, the same Hamilton–Jacobi equation is obtained by considering a Wentzel–Kramers–Brillouin (WKB) approximation of the stationary state of the continuous process $x(t)$ in the weak-noise limit [27, 28, 39]. Analogous connections between large deviation theory and WKB methods have also been established for continuous-time Markov chains [12, 21, 26, 41].

More recently, rigorous large deviation theory has been applied to PDMPs [13, 14, 24]. Independently of these developments in large deviation theory, a variety of techniques in applied mathematics and mathematical physics have been used to solve first passage time problems in biological applications of stochastic hybrid systems. These include WKB approximations and matched asymptotics [5, 22, 30, 32, 33], and path-integrals [4, 6]. Although such approaches are less rigorous than large deviation theory, they are more amenable to explicit calculations. In particular, they allow one to calculate the prefactor in Arrhenius-like expressions for mean first passage times, rather than just the leading order exponential behavior governed by the quasipotential. A major aim of this paper is to make explicit the connection between large deviation theory and more applied approaches to stochastic hybrid systems, by highlighting the common underlying Hamiltonian structure. Consistent with this aim, we present a new derivation of the classical action arising from an LDP for stochastic hybrid systems. We take as our starting point the LDP due to Faggionato *et al* [13, 14]. Using the Perron–Frobenius theorem and the calculus-of-variations, we evaluate the LDP rate function in terms of the classical action, whose equations of motion along the most probable paths of escape are given by a Hamiltonian dynamical system. We show that one major difference between hybrid and non-hybrid stochastic processes is that in the former case the WKB Hamiltonian tends to be a reduced version of the LDP Hamiltonian. However, the corresponding Hamiltonian dynamical systems yield the same most probable paths of escape.

The structure of the paper is as follows. In section 2 we define a stochastic hybrid system corresponding to a one-dimensional PDMP, and specify our various mathematical assumptions. We then present our detailed derivation of the classical LDP action in section 3 (theorem 3.1 and its proof). Although we focus on a one-dimensional PDMP, we also indicate how to extend our results to higher-dimensional PDMPs and multi-scale processes. The relationship between the classical action of large deviation theory and the quasipotential of WKB theory is developed in section 4. In particular, we show how the WKB quasipotential satisfies a Hamilton–Jacobi equation, whose associated Hamiltonian is consistent with the Hamiltonian obtained from the classical LDP action. This result is consistent with alternative formulations based either on more abstract probability theory [24] or formal path-integral methods [4, 6]. Finally, we illustrate our analysis of LDPs for stochastic hybrid systems by considering some conductance-based models of membrane voltage fluctuations in the presence of stochastic ion channels (section 5).

J. Stat. Mech. (2017) 033206

2. Stochastic hybrid systems

Consider a one-dimensional SHS corresponding to a PDMP with continuous variable $x \in \Omega \subset \mathbb{R}$ and a discrete variable $n \in \Gamma \equiv \{0, \cdots, N\}$ [10, 24]. (Note that one could extend the analysis to higher-dimensions, $x \in \mathbb{R}^d$. In this case Ω is taken to be a connected, bounded domain with a regular boundary $\partial\Omega$, see section 3.4. It is also possible to have more than one discrete variable, but one can always relabel the discrete states so that they are effectively indexed by a single integer.) When the internal state is n, the system evolves according to the ordinary differential equation (ODE)

$$\dot{x} = F_n(x), \tag{2.1}$$

where the vector field $F_n : \Omega \to \mathbb{R}$ is a continuous function, locally Lipschitz. That is, for any compact subset $\mathcal{K}$ of Ω, there exists a positive constant K_n such that

$$|F_n(x) - F_n(y)| \leqslant K_n |x - y|, \quad \forall\, x, y \in \mathcal{K}. \tag{2.2}$$

We assume that the dynamics of x is confined to the domain Ω so that we have existence and uniqueness of a trajectory for each n. One final mild constraint is that the vector field does not have identical components anywhere in Ω, that is, for any $x \in \Omega$, there exists at least one pair $(n, m) \in \Gamma \times \Gamma$, $n \neq m$ for which $F_m(x) \neq F_n(x)$. For fixed x, the discrete stochastic variable evolves according to a homogeneous, continuous-time Markov chain with transition matrix $\mathbf{W}(x)$. We make the further assumption that the chain is irreducible for all $x \in \Omega$, that is, for fixed x there is a non-zero probability of transitioning, possibly in more than one step, from any state to any other state of the Markov chain. This implies the existence of a unique invariant probability distribution on Γ for fixed $x \in \Omega$, denoted by $\rho(x, n)$, such that

$$\sum_{n \in \Gamma} \rho(x, n) W_{nm}(x) = \rho(x, m), \quad \forall\, m \in \Gamma, \quad \forall\, x \in \Omega. \tag{2.3}$$

The existence of the unique invariant measure is a consequence of the well known Perron–Frobenius Theorem[4].

The above stochastic model defines a one-dimensional PDMP. It is also possible to consider generalizations of the continuous process, in which the ODE (2.1) is replaced by a stochastic differential equation (SDE), see section 3.5, or even a partial differential equation (PDE). In order to allow for such possibilities we will refer to all of these processes as examples of a stochastic hybrid system.

Let us decompose the transition matrix of the Markov chain as

$$W_{nm}(x) = P_{nm}(x)\lambda_n(x),$$

with $\sum_{m, m \neq n} P_{nm}(x) = 1$ for all x. That is, for a given x, the jump times from state n are exponentially distributed with rate $\lambda_n(x)$ and $P_{nm}(x)$ is the probability distribution that when a jump occurs the new state is m for some $m \neq n$. The hybrid evolution of the system with respect to $x(t)$ and $n(t)$ is then described as follows. Suppose the system starts at time zero in the state (x_0, n_0). Call $x_0(t)$ the solution of (2.1) with $n = n_0$ such that $x_0(0) = x_0$. Let θ_1 be the random variable such that

[4] A finite-dimensional, real square matrix with positive entries has a unique largest real eigenvalue (the Perron eigenvalue) and the corresponding eigenvector has strictly positive components [20].

$$\mathbb{P}(\theta_1 < t) = 1 - \exp\left(-\int_0^t \lambda_{n_0}(x_0(t'))\mathrm{d}t'\right).$$

Then in the random time interval $s \in [0, \theta_1)$ the state of the system is $(x_0(s), n_0)$. We draw a value of θ_1 from $\mathbb{P}(\theta_1 < t)$, choose an internal state $n_1 \in \Gamma$ with probability $P_{n_1 n_0}(x_0(\theta_1))$, and call $x_1(t)$ the solution of the following Cauchy problem on $[\theta_1, \infty)$:

$$\begin{cases} \dot{x}_1(t) = F_{n_1}(x_1(t)), & t \geqslant \theta_1 \\ x_1(\theta_1) = x_0(\theta_1). \end{cases}$$

Iterating this procedure, we construct a sequence of increasing jumping times $(\theta_k)_{k\geqslant 0}$ (setting $\theta_0 = 0$) and a corresponding sequence of internal states $(n_k)_{k\geqslant 0}$. The evolution $(x(t),\ n(t))$ is then defined as

$$(x(t), n(t)) = (x_k(t), n_k) \quad \text{if } \theta_k \leqslant t < \theta_{k+1}. \tag{2.4}$$

Note that the path $x(t)$ is continuous and piecewise C^1.

Given the above iterative definition of the stochastic hybrid process, let $X(t)$ and $N(t)$ denote the stochastic continuous and discrete variables, respectively, at time t, $t > 0$, given the initial conditions $X(0) = x_0$, $N(0) = n_0$. Although the evolution of the continuous variable $X(t)$ or the discrete variable $N(t)$ is non-Markovian, it can be proven that the joint evolution $(X(t),\ N(t))$ is a strong Markov process [10]. Introduce the probability density $\rho(x, n, t|x_0, n_0, 0)$ with

$$\mathbb{P}\{X(t) \in (x, x + \mathrm{d}x),\ N(t) = n|x_0, n_0) = \rho(x, n, t|x_0, n_0, 0)\mathrm{d}x.$$

It follows that ρ evolves according to the forward differential Chapman–Kolmogorov (CK) equation [4, 18]

$$\frac{\partial \rho}{\partial t} = -\mathbb{L}\rho, \tag{2.5}$$

with the adjoint generator $-\mathbb{L}$ (dropping the explicit dependence on initial conditions) defined according to

$$\mathbb{L}\rho(x, n, t) = \frac{\partial F_n(x)\rho(x, n, t)}{\partial x} - \frac{1}{\epsilon}\sum_{m\in\Gamma} A^{\top}_{nm}(x)\rho(x, m, t), \tag{2.6}$$

with

$$A_{nm}(x) = W_{nm}(x) - \delta_{n,m}\sum_{k\in\Gamma} W_{nk}(x), \tag{2.7}$$

such that $\sum_{m\in\Gamma} A_{nm}(x) = 0 \quad \forall\, n \in \Gamma$ and $\forall\, x \in \Omega$, and $\sum_{n\in\Gamma}\rho(x, n)A_{nm}(x) = 0$. The first term on the right-hand side of equation (2.6) represents the probability flow associated with the piecewise deterministic dynamics for a given n, whereas the second term represents jumps in the discrete state. Note that we have rescaled the transition matrix $\mathbf{W}$ (and hence $\mathbf{A}$) by introducing the dimensionless parameter ϵ, $\epsilon > 0$. This is motivated by the observation that in many biological applications there a separation of time-scales between the relaxation time for the dynamics of the continuous variable x and the rate of switching between the different discrete states n [4].

J. Stat. Mech. (2017) 033206

Let us now define the averaged vector field $\overline{F}:\mathbb{R}\to\mathbb{R}$ by

$$\overline{F}(x)=\sum_{n\in\Gamma}\rho(x,n)F_n(x).$$

It can be shown [14] that, given the assumptions on the matrix $\mathbf{W}$, the functions $\rho(x,n)$ on Ω belong to $C^1(\mathbb{R})$ for all $n\in\Gamma$ and that this implies that $\overline{F}(x)$ is locally Lipschitz. Hence, for all $t\in[0,T]$, the Cauchy problem

$$\begin{cases}\dot{x}(t) &= \overline{F}(x(t))\\ x(0) &= x_0\end{cases}\tag{2.8}$$

has a unique solution. Intuitively speaking, one would expect the stochastic hybrid system (2.1) to reduce to the deterministic dynamical system (2.8) in the limit $\epsilon\to 0$. That is, for sufficiently small ϵ, the Markov chain undergoes many jumps over a small time interval Δt during which $\Delta x\approx 0$, and thus the relative frequency of each discrete state n is approximately $\rho(x,n)$. This can be made precise in terms of a Law of Large Numbers for stochastic hybrid systems proven in [14].

3. Classical action from a large deviation principle

We now turn to the major part of the paper, namely, a new derivation of the classical action corresponding to the abstract rate function of the LDP for a one-dimensional PDMP introduced by Faggionato *et al* [13, 14]. The advantage of our approach is that it avoids many of the technical difficulties associated with probabilistic approaches to large deviation theory [24]. We begin by introducing some notation in order to state the LDP of Faggionato *et al* [13, 14] (section 3.1). The associated rate function J_T is expressed as a variational principle. In section 3.2 we summarize the main result of our paper in the form of theorem 3.1, which expresses the equivalence of the rate function J_T with a classical action, whose associated Lagrangian is related to the Perron eigenvalue of a certain linear operator. The latter depends on the transition matrix $\mathbf{W}$ of the Markov chain and the nonlinear functions $F_n(x)$, $n\in\Gamma$ of the piecewise deterministic system (2.1). The proof of theorem 3.1 is presented in section 3.3, where we explicitly derive the classical action by solving the variational principle using the Perron–Frobenius theorem and the calculus-of-variations. Extensions of the classical action to higher dimensional PDMPs and multi-scale processes are presented in sections 3.4 and 3.5, respectively.

3.1. Large deviation principle of Faggionato *et al* [13, 14]

In order to write down the LDP of Faggionato *et al* [13, 14], it is first necessary to introduce some notation. Let $\mathcal{M}_+(\Gamma)$ denote the space of all probability measures on Γ and introduce the product space $\mathcal{M}_+(\Gamma)^{[0,T]}$. For each $t\in[0,T]$, let $\psi(t)=(\psi_0(t),\ldots,\psi_N(t))\in\mathcal{M}_+(\Gamma)$ such that

$$\psi_n(t)\geqslant 0,\quad \sum_{n\in\Gamma}\psi_n(t)=1.$$

A particular realization of the stochastic process, $\{(x(t), n(t))\}_{t\in[0,T]}$, then lies in the product space $C([0,T])\times\mathcal{M}_+(\Gamma)^{[0,T]}$ with

$$\psi_n(t) = 1_{\{n(t)=n\}} \equiv \begin{cases} 1, & \text{if } n(t)=n, \\ 0, & \text{if } n(t)\neq n \end{cases} \tag{3.1}$$

and

$$x(t) = x_0 + \int_0^t \sum_{n\in\Gamma} \psi_n(s) F_n(x(s))\mathrm{d}s. \tag{3.2}$$

Let $\mathcal{Y}_{x_0}$ denote the subspace of $C([0,T])\times\mathcal{M}_+(\Gamma)^{[0,T]}$ for which equation (3.2) holds but ψ is now a general element of $\mathcal{M}_+(\Gamma)^{[0,T]}$. Such a space contains both the set of trajectories of the stochastic hybrid system with $\psi_n(t)$ given by equation (3.1) and $n(t)$ evolving according to the Markov chain, and the solution $x^*(t)$ of the averaged system (2.8) for which $\psi_n = \psi_n^*$ with $\psi_n^*(t) = \rho(n, x^*(t))$. It can be proven that $\mathcal{Y}_{x_0}$ is a compact subspace of $C([0,T])\times\mathcal{M}_+(\Gamma)^{[0,T]}$ with topology defined by the metric [14]

$$\begin{aligned} & d(\{(x(t),\psi(t))\}_{t\in[0,T]}, \{(\tilde{x}(t),\widetilde{\psi}(t))\}_{t\in[0,T]}) \\ &\quad = \sup_{t\in[0,T]} |x(t)-\tilde{x}(t)| + \sum_{n\in\Gamma} \sup_{0\leqslant t\leqslant T} \left| \int_0^t [\psi_n(s)-\widetilde{\psi}_n(s)]\mathrm{d}s \right|. \end{aligned}$$

Finally, we take $P^{\epsilon}_{x_0}$ to be the probability density functional or law of the set of trajectories $\{x(t)\}_{t\in[0,T]}\in C([0,T],\Omega)$.

The following large deviation principle then holds [13, 14]: Given an element $\{x(t)\}_{t\in[0,T]}\in C([0,T],\Omega)$,

$$\mathbb{P}^{\epsilon}_{x_0}[\{x(t)\}_{t\in[0,T]}] \sim \mathrm{e}^{-J_T(\{x(t)\}_{t\in[0,T]})/\epsilon}, \tag{3.3}$$

where the rate function $J_T : C([0,T],\Omega)\to[0,\infty)$ is given by

$$J_T(\{x(t)\}_{t\in[0,T]}) = \inf_{\{\psi(t)\}_{t\in[0,T]}:\{x(t),\psi(t)\}_{t\in[0,T]}\in\mathcal{Y}_{x_0}} \int_0^T j(x(t),\psi(t))\,\mathrm{d}t, \tag{3.4}$$

and

$$j(x,\psi) = \sup_{z\in(0,\infty)^{\Gamma}} \sum_{(n,m)\in\Gamma\times\Gamma} \psi_n W_{nm}(x) \left[1 - \frac{z_m}{z_n}\right]. \tag{3.5}$$

Here the symbol $\sim$ means asymptotic logarithmic equivalence in the limit $\epsilon\to 0$.

We summarize a few useful properties of $j(x,\psi)$ defined by equation (3.5). Since the double sum in equation (3.5) excludes diagonal terms, we introduce the set $\Gamma_\Delta = \Gamma\times\Gamma\setminus\Delta$ where Δ is the diagonal of $\Gamma\times\Gamma$. Let $c_{nm} = \psi_n W_{nm}(x)$. Suppose that ψ is a strictly positive measure, $\psi_n > 0$ for all $n\in\Gamma$. It then follows from the properties of the transition matrix $\mathbf{W}$ that the mapping $c:\Gamma_\Delta\to[0,\infty)$ is *irreducible* in the sense that, for all $n\neq m\in\Gamma$, there exists a finite sequence $n_1, n_2, \cdots, n_k$ such that $n_1 = n$, $n_k = m$ and $c_{n_i n_{i+1}} > 0$ for $i = 1,\cdots,k-1$. Let $\mathcal{R}_N$ denote the set of $(N+1)\times(N+1)$ positive matrices with diagonal zero. Define the mapping $\mathcal{J}:\mathcal{R}_N\to\mathbb{R}$ as

J. Stat. Mech. (2017) 033206

$$\mathcal{J}(c) = \sup_{z \in [0,\infty)^\Gamma} \widehat{\mathcal{J}}(c,z), \quad \widehat{\mathcal{J}}(c,z) = \sum_{(n,m)\in\Gamma_\Delta} c_{nm}(1 - z_m/z_n). \tag{3.6}$$

The following lemma is proven in [14, 24] and follows largely from material found in [38].

Lemma 1. *The function $\mathcal{J}$ is convex and continuous and takes its values in $[0,\infty)$. Moreover, for each irreducible c, the supremum on $[0,\infty)^\Gamma$ of the function $\widehat{\mathcal{J}}(c,\cdot)$ is the unique solution of the set of equations*

$$\sum_{m\in\Gamma} c_{nm}\frac{z_m}{z_n} = \sum_{m\in\Gamma} c_{mn}\frac{z_n}{z_m}, \quad n\in\Gamma, \tag{3.7}$$

under the normalization $\sum_{n\in\Gamma} z_n = 1$.

A key idea behind the above LDP is that a slow dynamical process coupled to the fast Markov chain on Γ rapidly samples the different discrete states of Γ according to some non-negative measure ψ. In the limit $\epsilon\to 0$, one has $\psi\to\rho$, where ρ is the ergodic measure of the Markov chain. On the other hand, for small but non-zero ϵ, ψ is itself distributed according to a LDP, whereby one averages the different functions $F_n(x)$ over the measure ψ to determine the dynamics of the slow system. In most biological applications, one is not interested in the internal discrete state of the system, that is, one only observes the statistical behavior of the continuous variable $x(t)$. For example, $x(t)$ could represent the membrane voltage of a single neuron [33] or the synaptic current in a population of neurons [3]. Faggionato *et al* [13, 14] explicitly calculated the rate function (3.4) for a restricted class of stochastic hybrid systems, whose stationary density is exactly solvable. One major constraint on this class of model is that the vector field of the piecewise deterministic system has non-vanishing components within a given confinement domain. However, this constraint does not hold for biological systems such as ion channels [5, 22, 33], gene networks [23, 30], and neural networks [3]. (Faggionato *et al* were motivated by a model of molecular motors that is exactly solvable. Such a solvability condition also breaks down for molecular motors when local chemical signaling is taken into account [31].)

3.2. Statement of main theorem

We now state the main result of our paper in the form of a theorem, which will then be proven in section 3.3.

Theorem 3.1. *Let $\lambda = \lambda(x,\mu)$ for fixed x,μ be the Perron eigenvalue of the linear equation*

$$\sum_{m\in\Gamma} A_{nm}(x)z_m(x,\mu) + \mu F_n(x)z_n(x,\mu) = \lambda(x,\mu)z_n(x,\mu), \tag{3.8}$$

where $\mathbf{z}(x,\mu) = (z_n(x,\mu), n\in\Gamma)$ is the unique positive eigenvector of the linear equation under the normalization $\sum_n z_n(x,\mu) = 1$. Similarly, let $\mathbf{R}(x,\mu) = (R_n(x,\mu), n\in\Gamma)$ be the positive eigenvector of the adjoint equation

$$\sum_{m\in\Gamma} A^\top_{nm}(x)R_m(x,\mu) + \mu F_n(x)R_n(x,\mu) = \lambda(x,\mu)R_n(x,\mu) \tag{3.9}$$

J. Stat. Mech. (2017) 033206

under the normalization $\sum_m z_m(x,\mu)R_m(x,\mu)=1$*. The rate function* $J_T(\{x(t)\}_{t\in[0,T]})$ *defined in equation (3.4) can then be written in the form of the classical action*

$$J_T(\{x(t)\}_{t\in[0,T]})=\int_0^T L(x,\dot{x})\mathrm{d}t, \tag{3.10}$$

with Lagrangian given by

$$L(x,\dot{x})=\mu(x,\dot{x})\dot{x}-\lambda(x(t),\mu(x,\dot{x})), \tag{3.11}$$

and $\mu=\mu(x,\dot{x})$ *is given by the unique solution of the invertible equation*

$$\dot{x}=\sum_{m\in\Gamma}\psi_m(x,\mu)F_m(x), \tag{3.12}$$

with

$$\psi_m(x,\mu)=z_m(x,\mu)R_m(x,\mu). \tag{3.13}$$

3.3. Proof of theorem 3.1

The proof proceeds in two main steps.

3.3.1. Evaluating the supremum of equation (3.5). We begin by introducing the following ansatz regarding the solution $\mathbf{z}=(z_n,n\in\Gamma)$ of the variational problem (3.5) for fixed x and strictly positive measure ψ, namely, that it is an eigenvector of the following matrix equation:

$$\mathbf{A}(x)\mathbf{z}+\mathbf{q}\circ\mathbf{z}=\lambda\mathbf{z} \tag{3.14}$$

for some bounded vector $\mathbf{q}=(q_n,n\in\Gamma)$. Here, for any $\mathbf{a},\mathbf{b}\in\mathbb{R}^{N+1}$,

$$[\mathbf{a}\circ\mathbf{b}]_n\equiv[\mathrm{diag}(\mathbf{a})\mathbf{b}]_n=a_nb_n.$$

Note that we are free to shift the vector $\mathbf{q}$ by a constant since, under the transformation $q_n\to q_n-c$, the eigenvalue shifts by $\lambda\to\lambda-c$ and the eigenvector is unchanged. That is, for fixed x,

$$\lambda(x,\mathbf{q}-c\mathbf{1}_{N+1})=\lambda(x,\mathbf{q})-c. \tag{3.15}$$

For simplicity, we choose $c=q_N$ and take $\mathbf{Q}=(q_0,q_1,\ldots,q_{N-1},0)$ so that $\lambda(x,\mathbf{Q})=\lambda(x,\mathbf{q})-q_N$ and $\mathbf{z}=\mathbf{z}(x,\mathbf{Q})$ are solutions of the matrix equation

$$\mathbf{A}(x)\mathbf{z}(x,\mathbf{Q})+\mathbf{Q}\circ\mathbf{z}(x,\mathbf{Q})=\lambda(x,\mathbf{Q})\mathbf{z}(x,\mathbf{Q}). \tag{3.16}$$

There are N independent variables, q_n, $n=0,\ldots,N-1$. One of the crucial features of the above ansatz is that we can then ensure $z_n=z_n(x,\mathbf{Q})$ is strictly positive by using the Perron–Frobenius theorem and taking $\lambda=\lambda(x,\mathbf{Q})$ to be the Perron eigenvalue. Indeed, choosing κ such that

$$\kappa>\max_{n=0,\ldots,N-1}\{|q_n|\}, \tag{3.17}$$

it is clear that the new matrix $\mathbf{A}(x)+\mathrm{diag}(\mathbf{Q})+\kappa\mathbf{I}_{N+1}$ is irreducible and positive. According to the Perron–Frobenius theorem, it has a unique strictly positive eigenvector

J. Stat. Mech. (2017) 033206

$\mathbf{z}(x, \mathbf{Q})$ with the normalization $\sum_m z_m(x, \mathbf{Q}) = 1$, and this eigenvector is also an eigenvalue of $\mathbf{A} + \mathrm{diag}(\mathbf{Q})$ (but with a shifted eigenvalue).

We now show that for a given $\mathbf{Q}$ and fixed x, the Perron eigenvector $\mathbf{z}(x, \mathbf{Q})$ is the solution to the variational problem (3.5) provided that the measure ψ takes the specific form

$$\psi_n = \psi_n(x, \mathbf{Q}) \equiv R_n(x, \mathbf{Q}) z_n(x, \mathbf{Q}) \tag{3.18}$$

for each $n = 0, \ldots, N-1$, where $\mathbf{R}(x, \mathbf{Q}) = (R_n(x, \mathbf{Q}), n \in \Gamma)$ is the corresponding unique strictly positive eigenvector (up to scalar multiplication) of the adjoint linear equation

$$\mathbf{A}^{\top}(x)\mathbf{R}(x, \mathbf{Q}) + \mathbf{Q} \circ \mathbf{R}(x, \mathbf{Q}) = \lambda(x, \mathbf{Q})\mathbf{R}(x, \mathbf{Q}), \tag{3.19}$$

such that $\mathbf{z}^{\top}\mathbf{R} = 1$. Equation (3.18) ensures that ψ is a strictly positive measure and that $\sum_{m=0}^{N} \psi_m = 1$. We proceed by establishing that $\mathbf{z}(x, \mathbf{Q})$ satisfies equation (3.7) for $c_{nm} = \psi_n(x, \mathbf{Q}) W_{nm}(x)$ and $\psi_n(x, \mathbf{Q})$ given by equation (3.18). The left-hand side of (3.7) becomes

$$\begin{aligned}
\sum_{m=0}^{N} \psi_n W_{nm}(x) \frac{z_m}{z_n} &= \frac{\psi_n}{z_n} \sum_{m=0}^{N} W_{nm}(x) z_m = \frac{\psi_n}{z_n} \left(\lambda(x, \mathbf{Q}) - Q_n + \sum_k W_{nk} \right) z_n \\
&= \psi_n \left(\lambda(x, \mathbf{Q}) - Q_n + \sum_k W_{nk} \right),
\end{aligned} \tag{3.20}$$

whereas the right-hand side of (3.7) reads

$$z_n \sum_{m=0}^{N} W_{mn}(x) \frac{\psi_m}{z_m} = z_n \sum_{m=0}^{N} \left(A_{mn}(x) + \delta_{mn} \sum_{k=0}^{N} W_{nk} \right) \frac{\psi_m}{z_m}.$$

Cancelling the factors of z_n, setting $\psi_m / z_m = R_m$ and using equation (3.19) we recover (3.20). Hence, equation (3.7) holds. Finally, setting $\psi_n = \psi_n(x, \mathbf{Q})$ in equation (3.5) we can evaluate the supremum to obtain

$$\begin{aligned}
j(x, \psi) &= \sum_{m,n=0}^{N} \psi_n W_{nm}(x) \left[1 - \frac{z_m}{z_n} \right] \\
&= \sum_{n=0}^{N} \psi_n \left[\sum_{m=0}^{N} W_{nm}(x) - \frac{1}{z_n} \sum_{m=0}^{N} \left(A_{nm}(x) + \delta_{nm} \sum_{k=0}^{N} W_{nk}(x) \right) z_m \right] \\
&= \sum_{n=0}^{N} \psi_n [Q_n - \lambda(x, \mathbf{Q})] \\
&= \sum_{n=0}^{N-1} q_n \psi_n - \lambda(x, \mathbf{Q}),
\end{aligned} \tag{3.21}$$

since $\sum_{m=0}^{N} \psi_m = 1$ and $Q_N = 0$. For ease of notation, we have suppressed the explicit dependence of $\mathbf{z}$ and ψ on $x, \mathbf{Q}$.

In the given variational problem for fixed x, there are N independent variables ψ_n, $n = 0, \ldots, N-1$ with $\psi_N = 1 - \sum_{n=0}^{N-1} \psi_n$. Similarly, there are N independent variables

J. Stat. Mech. (2017) 033206

q_n, $n = 0, \ldots, N-1$. Therefore, equation (3.18) determines a mapping between the sets $\{q_n,\ n = 0, \ldots, N-1\}$ and $\{\psi_n,\ n = 0, \ldots, N-1\}$. It remains to show that there exists a unique solution $\mathbf{Q}$ for each $\psi \in \mathcal{M}_+(\Gamma)$, that is, the mapping is invertible. For then we can set $q_n = q_n(x, \psi)$ in equation (3.21) to obtain the value of the supremum in equation (3.5) for any positive measure $\psi \in \mathcal{M}_+(\Gamma)$. Differentiating equation (3.19) with respect to q_m, $m = 0, \ldots, N-1$, yields the inhomogeneous linear equation

$$\mathbf{L}(x, \mathbf{Q})\frac{\partial \mathbf{R}}{\partial q_m} \equiv [\mathbf{A}^\top(x) + \mathrm{diag}(\mathbf{Q}) - \lambda \mathbf{I}_{N+1}]\frac{\partial \mathbf{R}}{\partial q_m} = \frac{\partial \lambda}{\partial q_m}\mathbf{R} - R_m \mathbf{e}_m, \tag{3.22}$$

where $(\mathbf{e}_m)_n = \delta_{m,n}$. Multiplying both sides of equation (3.22) on the left with $\mathbf{z}^\top$ and using (3.16) we obtain

$$\frac{\partial \lambda(x, \mathbf{Q})}{\partial q_m} = R_m(x, \mathbf{Q}) z_m(x, \mathbf{Q}), \quad m = 0, \ldots, N-1. \tag{3.23}$$

Since $\mathbf{R}$ and $\mathbf{z}$ are strictly positive, $\lambda(x, \mathbf{Q})$ is a monotonically increasing function of the q_m. Moreover, equations (3.14) and (3.19) imply that, in the limit $q_l \to \infty$ with all other q_m fixed, $R_l, z_l \to 1$ and $\partial\lambda/\partial q_l \to 1$. On the other hand, if $q_l \to -\infty$ then $R_l, z_l \to 0$ and the Perron eigenvalue becomes independent of q_l with $\partial\lambda/\partial q_l \to 0$. Hence, by continuity, for each $\psi \in \mathcal{M}_+(\Gamma)$ there exists a vector $\mathbf{Q}$ such that $\psi_n = R_n(x, \mathbf{Q}) z_n(x, \mathbf{Q})$ for all $n = 0, \ldots, N-1$. For such a solution to be unique, the inverse function theorem implies that the Jacobian of the transformation must be invertible. Differentiating equation (3.23) with respect to q_n shows that the Jacobian is equivalent to the Hessian of λ with respect $\mathbf{Q}$, since

$$D_{mn}(x, \mathbf{Q}) \equiv \frac{\partial R_m(x, \mathbf{Q}) z_m(x, \mathbf{Q})}{\partial q_n} = \frac{\partial^2 \lambda(x, \mathbf{Q})}{\partial q_m \partial q_n} \tag{3.24}$$

for all $m, n = 0, \ldots, N-1$. This also establishes that the Jacobian is a symmetric matrix with real eigenvalues. Invertibility follows from the convexity of the function $\mathcal{J}(c)$ defined by equation (3.6). That is, differentiating

$$j(x, \psi) = \mathcal{J}(c), \quad c_{nm} = \psi_n W_{nm}(x),$$

with respect to ψ for fixed x gives

$$W_{nm}(x) W_{mm'}(x) \frac{\partial^2 \mathcal{J}(c)}{\partial c_{nm} \partial c_{mm'}} = \frac{\partial^2 j(x, \psi)}{\partial \psi_n \partial \psi_m}, \tag{3.25}$$

with $n, m \neq m'$. On the other hand, differentiating equation (3.21) for $j(x, \psi)$ with respect to ψ_n gives

$$\frac{\partial j(x, \psi)}{\partial \psi_n} = q_n + \sum_{j=0}^{N-1}\left[\psi_j - \frac{\partial \lambda}{\partial q_j}\right]\frac{\partial q_j}{\partial \psi_n} = q_n,$$

and so

$$\frac{\partial^2 j(x, \psi)}{\partial \psi_n \partial \psi_m} = \frac{\partial q_n}{\partial \psi_m} = [\mathbf{D}^{-1}]_{nm}.$$

Hence, convexity of $\mathcal{J}(c)$ together with irreducibility of the non-negative transition matrix $\mathbf{W}$ means that the Jacobian is invertible and positive definite.

In summary, we have shown that for a strictly positive measure ψ and fixed x, a unique solution $q_n = q_n(x, \psi)$ exists for all $n = 0, \ldots, N-1$, and we have solved the first variational problem by identifying $\mathbf{z}$ with the unique (up to scalar multiplication), strictly positive eigenfunction of the matrix equation (3.16). Now suppose ψ is a non-negative rather than a strictly positive measure, that is, $\psi_m = 0$ for at least one state $m \in \Gamma$. In this case $c_{nm} = \psi_n W_{nm}(x)$, $n \neq m$, is not irreducible and lemma 1 no longer applies. However, as proven by Faggionato *et al* [14], the function $\mathcal{J}(c)$ is continuous with respect to c. Hence, assuming that the form of the rate function (3.3) still holds (which isn't necessarily true), we can take a sequence of strictly positive measures $\psi^{(l)}$ on Γ such that $\psi_n^{(l)} \to \psi_n$ for each $n \in \Gamma$. This implies that (for fixed x)

$$c[\psi^{(l)}] \to c[\psi]$$

and

$$j(x, \psi^{(l)}) = \mathcal{J}(c[\psi^{(l)}]) \to \mathcal{J}(c[\psi]) = j(x, \psi),$$

so that one can extend equation (3.21) to non-negative measures by taking

$$q_n(x, \psi) = \lim_{l \to \infty} q_n(x, \psi^{(l)}).$$

Example. We will illustrate the Perron eigenvalue solution to the supremum variational problem by considering an example for $N = 1$. Let us take the transition matrix to be

$$\mathbf{W} = \begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}.$$

Consider the eigenvalue equation

$$\mathbf{W}\mathbf{z} + \mathbf{q} \circ \mathbf{z} = \lambda \mathbf{z},$$

where we have absorbed the diagonal terms $\sum_{k=1,2} W_{km}$ into the definition of q_m. The resulting characteristic equation is quadratic in λ and the leading or Perron eigenvalue is given by

$$\lambda = \frac{q_1 + q_2}{2} + \frac{7}{12} + \frac{1}{2}\sqrt{(q_1 - q_2)^2 - (q_1 - q_2)/3 + 25/36}\,.$$

It follows that

$$\psi_1 \equiv \frac{\partial \lambda}{\partial q_1} = \frac{1}{2} + f(q_1 - q_2), \quad \psi_2 \equiv \frac{\partial \lambda}{\partial q_2} = \frac{1}{2} - f(q_1 - q_2),$$

with

$$f(q) = \frac{1}{4}\frac{2q - 1/3}{\sqrt{q^2 - q/3 + 25/36}}.$$

J. Stat. Mech. (2017) 033206

Note that $\psi_1 + \psi_2 = 1$ as required. The function $f(q)$ is a monotonically increasing function of q with $f(-\infty) = -1/2$ and $f(\infty) = 1/2$. Thus, one can find a unique, finite value of $q = q_1 - q_2$ for all $\psi_1 \in (0, 1)$, that is, for all strictly positive ψ. In the case of a non-negative ψ with $\psi_1 = 0$ or $\psi_2 = 0$, we have $q \to \pm\infty$.

3.3.2. Evaluating the infimum of equation (3.4). The next step in the proof is to substitute for $j(x, \psi)$ in the rate function (3.4) using equation (3.21), which gives

$$J_T(\{x(t)\}_{t\in[0,T]}) = \inf_{\psi:\dot{x}=\sum_{n=0}^{N}\psi_n F_n(x)} \int_0^T \left[\sum_{n=0}^{N-1} q_n(t)\psi_n(t) - \lambda(x(t), \mathbf{Q}(t))\right] \mathrm{d}t, \tag{3.26}$$

with $q_n(t) = q_n(x(t), \psi(t))$ and $\sum_{m=0}^{N} \psi_m = 1$. In order to solve this variational problem, we introduce a Lagrange multiplier $\mu(t)$ and set

$$J_T(\{x(t)\}_{t\in[0,T]}) = \inf_{\psi,\mu} S[x, \mu, \psi], \tag{3.27}$$

where

$$S[x, \mu, \psi] = \int_0^T \left[\sum_{n=0}^{N-1} q_n(t)\psi_n(t) - \lambda(x(t), \mathbf{Q}(t)) \right. \\ \left. + \mu(t)\left(\dot{x} - \sum_{n=0}^{N-1}[F_n(x) - F_N(x)]\psi_n(t) - F_N(x)\right)\right] \mathrm{d}t, \tag{3.28}$$

and we have imposed the constraint $\sum_{m=0}^{N} \psi_m = 1$. The variational problem can now be expressed in terms of functional derivatives of S:

$$\frac{\delta S}{\delta \mu(s)} = \dot{x}(s) - \sum_{n=0}^{N} F_n(x(s))\psi_n(s) = 0, \tag{3.29a}$$

and

$$\frac{\delta S}{\delta \psi_m(s)} = \sum_{n=0}^{N-1} \frac{\partial q_n}{\partial \psi_m}\psi_n(s) + q_m(s) - \sum_{n=0}^{N-1} \frac{\partial \lambda}{\partial q_n}\frac{\partial q_n}{\partial \psi_m} \\ - \mu(s)[F_m(x(s)) - F_N(x(s))] = 0 \tag{3.29b}$$

for $m = 0, \ldots N - 1$. Combining with equations (3.18) and (3.23), we obtain the following solution to the variational problem in terms of μ:

$$q_m = \mu[F_m(x) - F_N(x)], \tag{3.30a}$$

$$\psi_m(x, \mu) = z_m(x, \mu)R_m(x, \mu), \tag{3.30b}$$

for all $m = 0, \ldots N - 1$, with $R_n(x, \mu), z_n(x, \mu)$ the positive eigenvectors of the matrix equation

$$\mathbf{A}^\top(x)\mathbf{R}(x, \mu) + \mu\mathbf{F}(x) \circ \mathbf{R}(x, \mu) = \lambda(x, \mu)\mathbf{R}(x, \mu), \tag{3.31}$$

and its adjoint, respectively. We have thus established equations (3.8) and (3.9). Here

$$\lambda = \lambda(x, \mu) \equiv F_N(x) + \lambda(x, \mathbf{Q} | q_m = \mu(F_m - F_N), m = 0, \ldots, N-1).$$

Substituting equation (3.30b) into (3.29a) shows that $\mu = \mu(t)$ where $\mu(t)$ is the solution to the equation

$$\dot{x}(t) = \sum_{n=0}^{N} F_n(x(t))\psi_n(x(t), \mu(t)), \tag{3.32}$$

provided that the equation is invertible for the given trajectory $\{x(t)\}_{t\in[0,T]}$. Hence, evaluating the action at the infimum we have

$$J_T(\{x(t)\}_{t\in[0,T]}) = \int_0^T \left[\mu(t)\sum_{n=0}^{N} F_n(x(t))\psi_n(\mu(t), x(t)) - \lambda(x(t), \mu(t))\right]\mathrm{d}t,$$

The final step is to show that equation (3.32) is invertible so that the function $\mu(t) = \mu(x(t), \dot{x}(t))$ exists and, hence, the rate function J_T has the required Lagrangian form (3.10). From the inverse function theorem we require that

$$\sum_{m=0}^{N} \frac{\partial \psi_m(x, \mu)}{\partial \mu} F_m(x) \neq 0$$

for all $x \in \Omega$. Following along identical lines to the analysis of equation (3.22), we differentiate the linear equation (3.31) with respect to μ to give

$$\mathbf{L}(x, \mu)\frac{\partial \mathbf{R}}{\partial \mu} = \frac{\partial \lambda}{\partial \mu}\mathbf{R} - \mathbf{F} \circ \mathbf{R}, \tag{3.33}$$

with $\mathbf{L}(x, \mu) = \mathbf{L}(x, \mathbf{Q})|_{q_m = \mu[F_m - F_N], m=0,\ldots,N-1}$. Using the same arguments as previously, we obtain

$$\frac{\partial \lambda(x, \mu)}{\partial \mu} = \sum_{n=0}^{N} F_n(x) z_n(x, \mu) R_n(x, \mu). \tag{3.34}$$

Hence, we require

$$\frac{\partial^2 \lambda(x, \mu)}{\partial \mu^2} \equiv \sum_{m,n=0}^{N-1} \left.\frac{\partial^2 \lambda(x, \mathbf{Q})}{\partial q_m \partial q_n}\right|_{q_m = \mu(F_m(x) - F_N(x))} \times (F_m(x) - F_N(x))(F_n(x) - F_N(x)) \neq 0.$$

This holds since the Jacobian $\mathbf{D}$ of equation (3.24) is invertible and $F_m(x) \neq F_N(x)$ for at least one $m \neq N$. Finally, from equation (3.17), we require μ to be bounded, that is, there exists a κ for which

$$\kappa > \mu \max_{m=0,\ldots,N} \{|F_m(x)|\}$$

for all $x \in \Omega$. Hence, we have obtained the classical action (3.10) from the LDP rate function (3.4) and the proof of theorem 3.1 is complete. □

J. Stat. Mech. (2017) 033206

3.4. Extension to higher dimensions, $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{d}}$, $\boldsymbol{d} > 1$

In theorem 3.1 we considered the case of one-dimensional piecewise deterministic dynamics by taking $x \in \mathbb{R}$. However, it is relatively straightforward to derive the corresponding classical action for $\mathbf{x} \in \mathbb{R}^d$. When the internal state is n, the system now evolves according to the ODE

$$\dot{\mathbf{x}} = \mathbf{F}_n(\mathbf{x}), \tag{3.35}$$

where the vector field $\mathbf{F}_n : \Omega \to \mathbb{R}^d$ is a continuous function, locally Lipschitz. That is, given a compact subset $\mathcal{K}$ of Ω, there exists a positive constant $K_n(\mathcal{K})$ such that

$$|\mathbf{F}_n(\mathbf{x}) - \mathbf{F}_n(\mathbf{y})| \leqslant K_n(\mathcal{K})|\mathbf{x} - \mathbf{y}|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{K}. \tag{3.36}$$

The rate function of the LDP (3.3) becomes

$$J_T(\{\mathbf{x}(t)\}_{t\in[0,T]}) = \int_0^T L(\mathbf{x}, \dot{\mathbf{x}})\mathrm{d}t, \tag{3.37}$$

where L is the Lagrangian

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \langle \boldsymbol{\mu}(\mathbf{x}, \dot{\mathbf{x}}),\, \dot{\mathbf{x}} \rangle - \lambda(\mathbf{x}(t), \boldsymbol{\mu}(\mathbf{x}, \dot{\mathbf{x}})),$$

and $\lambda(\mathbf{x}, \boldsymbol{\mu})$ is the Perron eigenvalue of the linear equation

$$\mathbf{A}(x)\mathbf{z}(\mathbf{x}, \boldsymbol{\mu}) + (\mathbf{F}(\mathbf{x})\boldsymbol{\mu}) \circ \mathbf{z}(\mathbf{x}, \boldsymbol{\mu}) = \lambda(\mathbf{x}, \boldsymbol{\mu})\mathbf{z}(\mathbf{x}, \boldsymbol{\mu}), \tag{3.38}$$

where $\mathbf{z}(\mathbf{x}, \boldsymbol{\mu})$ is the positive eigenvector of the linear equation under the normalization $\sum_n z_n(\mathbf{x}, \boldsymbol{\mu}) = 1$. Here

$$[\mathbf{a} \circ \mathbf{b}]_n = a_n b_n,\ n = 0, \cdots, N \tag{3.39}$$

for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{N+1}$ and $\mathbf{F}(\mathbf{x})$ is the $(N+1) \times d$ matrix whose $N+1$ rows are the d-dimensional vectors $\mathbf{F}_m(\mathbf{x})$, $m = 0, \cdots, N$. Finally, the d-dimensional vector $\boldsymbol{\mu} = \boldsymbol{\mu}(\mathbf{x}, \dot{\mathbf{x}})$ is the solution of the invertible equation

$$\dot{\mathbf{x}} = \sum_{m\in\Gamma} \psi_m(\mathbf{x}, \boldsymbol{\mu})\mathbf{F}_m(\mathbf{x}), \tag{3.40}$$

with

$$\psi_m(\mathbf{x}, \boldsymbol{\mu}) = z_m(\mathbf{x}, \boldsymbol{\mu})R_m(\mathbf{x}, \boldsymbol{\mu}), \tag{3.41}$$

where $\mathbf{R}$ is the positive eigenvector of the adjoint equation

$$\mathbf{A}^\top(\mathbf{x})\mathbf{R}(\mathbf{x}, \boldsymbol{\mu}) + (\mathbf{F}(\mathbf{x})\boldsymbol{\mu}) \circ \mathbf{R}(\mathbf{x}, \boldsymbol{\mu}) = \lambda(\mathbf{x}, \boldsymbol{\mu})\mathbf{R}(\mathbf{x}, \boldsymbol{\mu}) \tag{3.42}$$

under the normalization $\sum_m z_m(\mathbf{x}, \boldsymbol{\mu})R_m(\mathbf{x}, \boldsymbol{\mu}) = 1$.

3.5. Extension to a multi-scale process on $\mathbb{R}$

So far we have assumed that the slow process is piecewise deterministic. However, one of the useful features of taking the Lagrangian LDP [13, 14] as our starting point is that it is relatively straightforward to extend our analysis to the case where the slow process is a piecewise SDE. First, recall that the key idea behind the Faggionato *et al* Lagrangian construction is that the slow dynamical process coupled to the fast Markov chain on Γ rapidly samples the different discrete states of Γ according to some non-negative measure ψ. In order to extend this construction to a piecewise SDE, it is necessary to take account of the fact that there are now two levels of stochasticity. That is, after averaging the transition rates of the drift and variance of the SDE with respect to a given measure ψ, the resulting system is still stochastic. Since the slow system operates in a weak noise regime, it follows that one can apply an LDP to the slow system for a given ψ. The LDP for the full system is then obtained by combining the rate function of the slow system with the infimum rate function for ψ. Here we sketch how to extend the analysis to a piecewise SDE. (Approaching large deviation theory for multi-scale stochastic processes in terms of solutions to an eigenvalue problem has also been considered by Feng and Kurtz [15].)

Consider the piecewise Ito SDE

$$\mathrm{d}X(t) = F_n(X) + \sqrt{\epsilon}\sigma_n(X)\mathrm{d}W(t), \tag{3.43}$$

where $n \in \Gamma$ and $W(t)$ is a Wiener process. The drift term $F_n(x)$ and diffusion term $\sigma_n(x)$ are both taken to be Lipschitz. When the piecewise SDE is coupled to the fast discrete process on Γ, the stochastic dynamics is described by a differential Chapman–Kolmogorov equation of the form (see also equation (2.6))

$$\begin{aligned}\frac{\partial \rho(x,n,t)}{\partial t} &= -\frac{\partial}{\partial x}[F_n(x)\rho(x,n,t)] + \frac{\epsilon}{2}\frac{\partial^2}{\partial x^2}[\sigma_n^2(x)\rho(x,n,t)] \\ &\quad + \frac{1}{\epsilon}\sum_m A_{nm}^\top(x)\rho(x,m,t),\end{aligned} \tag{3.44}$$

where

$$\mathbb{P}\{X(t) \in (x, x+\mathrm{d}x),\ n(t) = n | x_0, n_0) = \rho(x,n,t)\mathrm{d}x.$$

We define the measure space $\mathcal{M}_+(\Gamma)$ as before, but now modify the definition of the subspace $\mathcal{Y}_{x_0} \subset C([0,T]) \times \mathcal{M}_+(\Gamma)^{[0,T]}$ by taking it to be the set of stochastic trajectories satisfying

$$\begin{aligned}\mathrm{d}X(t) &= \sum_{n=0}^{N}\psi_n(t)F_n(X)\mathrm{d}t + \sqrt{\epsilon}\sqrt{\sum_{n=0}^{N}\psi_n(t)\sigma_n^2(X)}\ \mathrm{d}W(t) \\ &\equiv F(X,\psi)\mathrm{d}t + \sqrt{\epsilon}\sigma(X,\psi)\mathrm{d}W(t)\end{aligned} \tag{3.45}$$

for $\psi \in \mathcal{M}_+(\Gamma)^{[0,T]}$. Such a space includes the set of trajectories of the piecewise SDE (3.43) with $\psi_n(t)$ given by equation (3.1) and $n(t)$ evolving according to the Markov chain on Γ. Consider a particular realization of the Wiener process $W(t)$ on $[0, T]$, which is independent of $\{x(t), \psi(t)\}_{[0,T]}$. For a given realization of W, one can write down an LDP along identical lines to the case of a piecewise deterministic system, see equation (3.3), yielding

$$J_T(\{X(t)\}_{t\in[0,T]}) = \inf_{\{\psi(t)\}_{t\in[0,T]}:\{X(t),\psi(t)\}_{t\in[0,T]}\in\mathcal{Y}_{x_0}} \int_0^T j(X(t),\psi(t))\,\mathrm{d}t, \tag{3.46}$$

where $X(t)$ is the sample path generated by the particular realization of $W(t)$, and $j(X,\psi)$ is given by equation (3.21). If we now formally average with respect to the Wiener process we find that

J. Stat. Mech. (2017) 033206

$$J_T(\{x(t)\}_{t\in[0,T]}) = \inf_{\{\psi(t)\}_{t\in[0,T]}} S[x,\psi], \tag{3.47}$$

where

$$S[x,\psi] = \int_0^T \left[\sum_{n=0}^{N-1} q_n(t)\psi_n(t) - \lambda(x(t),\mathbf{Q}(t)) + \frac{(\dot{x} - \sum_{n=0}^N \psi_n F_n(x))^2}{2\sum_{n=0}^N \psi_n\sigma_n^2(x)}\right]\mathrm{d}t, \tag{3.48}$$

with $q_n(t) = q_n(x(t),\psi(t))$, $\sum_{m=0}^N \psi_m(t) = 1$ and λ the Perron eigenvalue of equation (3.16). Note that the final term on the right-hand side is the well-known Freidlin–Wentzell action for SDEs [17]. Taking the infimum with respect to ψ_k, $k = 0,\ldots,N-1$, gives

$$\begin{aligned}0 = \frac{\delta S}{\delta\psi_k} &= q_k - \frac{(\dot{x} - \sum_{n=0}^N \psi_n F_n(x))[F_k(x) - F_N(x)]}{\sum_{n=0}^N \psi_n\sigma_n^2}\\ &\quad - \frac{(\dot{x} - \sum_{n=0}^N \psi_n F_n(x))^2[\sigma_k^2(x) - \sigma_N^2(x)]}{2[\sum_{n=0}^N \psi_n\sigma_n^2]^2} + \sum_{n=0}^{N-1}\left(\psi_n - \frac{\partial\lambda}{\partial q_n}\right)\frac{\partial q_n}{\partial\psi_k}.\end{aligned} \tag{3.49}$$

Introducing the new variables

$$\mu = \sum_{n=0}^N \psi_n F_n(x), \quad \sigma^2 = \sum_{n=0}^N \psi_n\sigma_n^2(x), \quad \nu = \frac{\dot{x}-\mu}{\sigma^2}, \tag{3.50}$$

we have

$$q_k = \nu[F_k(x) - F_N(x)] + \frac{\nu^2}{2}[\sigma_k^2(x) - \sigma_N^2(x)], \tag{3.51}$$

and

$$\psi_m = \partial\lambda/\partial q_m = R_m(x,\nu)z_m(x,\nu), \tag{3.52}$$

with $R_n(x,\nu)$, $z_n(x,\nu)$ the unique positive eigenvectors of the matrix equation

$$\sum_n A_{mn}^\top(x)R_n(x,\nu) + (\nu F_m(x) + \nu^2\sigma_m^2(x)/2)R_m(x,\nu) = \lambda(x,\nu)R_m(x,\nu). \tag{3.53}$$

and its adjoint. Moreover, equation (3.50) implies that $\nu = \nu(x(t),\dot{x}(t)) \equiv \nu(t)$ where $\nu(t)$ is the solution to the equation

$$\dot{x}(t) = \nu(t)\sum_{n=0}^N \sigma_n^2(x(t))\psi_n(x(t),\nu(t)) + \sum_{n=0}^N F_n(x(t))\psi_n(x(t),\nu(t)), \tag{3.54}$$

provided that the equation is invertible for the given trajectory $\{x(t)\}_{t\in[0,T]}$. Finally, evaluating the action at the infimum yields the corresponding classical action:

$$\begin{aligned}J_T(\{x(t)\}_{t\in[0,T]}) &= \int_0^T \left[\nu(t)\sum_{n=0}^N F_n(x(t))\psi_n(x(t),\nu(t))\right.\\ &\quad \left. + \frac{\nu(t)^2}{2}\sum_{n=0}^N \psi_n(x(t),\nu(t))\sigma_n^2(x(t)) - \lambda(x(t),\nu(t)) + \frac{\nu(t)^2}{2}\right]\mathrm{d}t\\ &= \int_0^T L(x,\dot{x})\mathrm{d}t.\end{aligned} \tag{3.55}$$

4. Classical Hamiltonian and the WKB approximation of the stationary state

As indicated in the introduction, for non-hybrid stochastic processes the quasipotential of WKB theory satisfies a Hamilton–Jacobi equation $H(x, \partial_x \Phi) = 0$, where H is the Hamiltonian obtained from the classical action of large deviation theory. It turns out that such a connection is more subtle in the case of a PDMP, as we now show. We will focus on the one-dimensional case of theorem 3.1. Given the action (3.10), we can determine the Hamiltonian H from the Lagrangian L according to the Fenchel–Legendre transformation:

$$H(x,p) = \sup_y \left[(p - \mu(x,y))y + \lambda(x, \mu(x,y)) \right]. \tag{4.1}$$

Evaluating the right-hand side yields the equation

$$p - \mu(x,y) + \left[\frac{\partial \lambda}{\partial \mu} - y \right] \frac{\partial \mu}{\partial y} = 0 \tag{4.2}$$

with

$$y = \sum_{m \in \Gamma} \psi_m(x,\mu) F_m(x).$$

Differentiating equation (3.8) with respect to μ gives

$$\sum_{m\in\Gamma} A_{nm}(x) \frac{\partial z_m(x,\mu)}{\partial \mu} + [\mu F_n(x) - \lambda(x,\mu)] \frac{\partial z_n(x,\mu)}{\partial \mu} \\ = \left[\frac{\partial \lambda(x,\mu)}{\partial \mu} - F_n(x) \right] z_n(x,\mu). \tag{4.3}$$

Since the adjoint of the linear operator on the left-hand side has a one-dimensional null space spanned by R_n, it follows from the normalization $\sum_m z_m R_m = 1$ that

$$\frac{\partial \lambda(x,\mu)}{\partial \mu} = \sum_{m\in\Gamma} \psi_m(x,\mu) F_m(x) = y.$$

Equation (4.2) thus shows that $p = \mu$, and we can identify p as the 'conjugate momentum' of the Hamiltonian

$$H = \lambda(x,p), \tag{4.4}$$

where $\lambda(x,p)$ is the Perron eigenvalue of the linear equation (3.8) and its adjoint (3.9).

In order to relate the Hamiltonian (4.4) with the quasipotential of WKB methods, consider the WKB approximation of the stationary state ρ_{ss} of the CK equation (2.5) (assuming it exists), with

$$\mathbb{L}\rho_{ss}(x,n) \equiv \frac{\partial F_n(x) \rho_{ss}(x,n)}{\partial x} - \frac{1}{\epsilon} \sum_{m\in\Gamma} A^{\top}_{nm}(x) \rho_{ss}(x,m) = 0. \tag{4.5}$$

The WKB approximation of ρ_{ss} takes the form

J. Stat. Mech. (2017) 033206

$$\rho_{ss}(x,n) \sim \eta(x,n)\exp\left(-\frac{\Phi(x)}{\epsilon}\right). \tag{4.6}$$

Substituting into equation (4.5) yields

$$\sum_{m\in\Gamma}(A^{\top}_{nm}(x)+\Phi'(x)\delta_{n,m}F_m(x))\eta(x,m) = \epsilon\frac{\mathrm{d}F_n(x)\eta(x,n)}{\mathrm{d}x}, \tag{4.7}$$

where $\Phi' = \mathrm{d}\Phi/\mathrm{d}x$. Introducing the asymptotic expansions $\eta \sim \eta^{(0)} + \epsilon\eta^{(1)} + \ldots$ and $\Phi \sim \Phi_0 + \epsilon\Phi_1 + \ldots$, the leading order equation is

$$\sum_{m\in\Gamma} A^{\top}_{nm}(x)\eta^{(0)}(x,m) = -\Phi_0'(x)F_n(x)\eta^{(0)}(x,n). \tag{4.8}$$

As it stands, it is not clear that (4.8) has a solution for which $\eta^{(0)}(x,m)$ is positive for all x, m, and the relationship to the Hamiltonian structure of large deviation theory is not explicit. However, the structure of equation (4.8) is a reduced version of equation (3.9). This suggests introducing the family of eigenvalue equations

$$\sum_{m\in\Gamma} A^{\top}_{nm}(x)R_m(x,p)+pF_n(x)R_n(x,p) = \lambda(x,p)R_n(x,p), \tag{4.9}$$

which are parameterized by the pair (x, p) with p an auxiliary variable and $\lambda(x,p)$ the Perron eigenvalue. Comparison of equation (4.8) with (4.9) then shows that we can make the identifications $\Phi_0'(x) = p$, $\eta^{(0)}(x,m) = R_m(x,p)$ and $\lambda(x,p) = 0$. It immediately follows that $\eta^{(0)}$ is positive. The quasipotential is then the solution of the Hamilton–Jacobi equation

$$\lambda(x,\Phi_0'(x)) = 0. \tag{4.10}$$

This is equivalent to finding zero energy solutions of Hamilton's equations

$$\dot{x} = \frac{\partial\lambda(x,p)}{\partial p}, \quad \dot{p} = -\frac{\partial\lambda(x,p)}{\partial x}, \tag{4.11}$$

and identifying Φ_0 as the action along the resulting solution curve $(x(t), p(t))$:

$$\Phi_0(x) = \int_{-\infty}^{T} p(t)\dot{x}(t)\mathrm{d}t, \tag{4.12}$$

with $x(T) = x$. Note that here t is the parameter of a curve rather than physical time.

One of the major applications of WKB methods is to solving escape problems for stochastic process in the weak-noise limit, both for non-hybrid systems [12, 21, 26–28, 39], and hybrid systems [5, 22, 30, 32, 33]. For example, suppose that the mean-field equation (2.8) is bistable with two stable fixed-points $x_{\pm}$ separated by an unstable fixed point x_0, see also figure 1 of section 5. Given the quasipotential Φ_0, the mean first passage time τ to escape from x_- via x_0, say, can be calculated by considering higher order terms in the WKB approximation, and using matched asymptotics to deal with an imposed absorbing boundary at x_0. One finds that τ takes the general Arrhenius form [22, 32]

$$\tau \sim \frac{\chi(x_0,x_-)}{\sqrt{|\Phi_0''(x_0)|\Phi_0''(x_-)}}\mathrm{e}^{[\Phi_0(x_0)-\Phi_0(x_-)]/\epsilon},$$

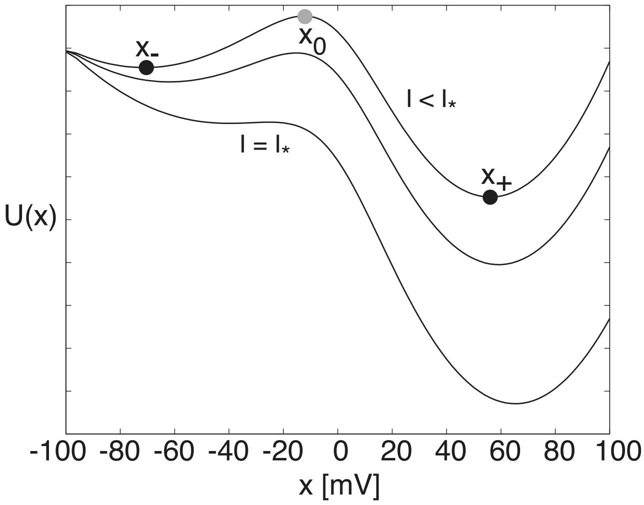

Figure 1. Plot of deterministic potential $U(x)$ as a function of voltage x for different values of the external stimulus current I. Parameter values are $V_{\mathrm{Na}} = 120$ mV, $V_{\mathrm{L}} = -62.3$ mV, $g_{\mathrm{Na}} = 4.4$ mS cm^{-2}, $g_{\mathrm{L}} = 2.2$ mS cm^{-2}, and $\alpha(x) = \beta \exp[(x - v_1)/v_2]$ with $\beta = 0.8$ s^{-1}, $v_1 = -1.2$ mV, $v_2 = 18$ mV.

where χ is an appropriate prefactor. Hence, the WKB method provides a powerful calculational tool. On the other hand, there is no *a priori* justification for interpreting the quasipotential and its associated Hamiltonian in terms of an underlying variational principle for most probable paths in the space of stochastic trajectories. This becomes crucial when solving escape problems in higher dimensions, since a metastable state is now surrounded by a non-trivial boundary (rather than a single point) and one needs to determine the relative weighting of optimal paths crossing different points on the boundary. Establishing the connection between WKB analysis and large deviation theory provides such a variational principle.

5. Applications to stochastic ion channels

In this final section we illustrate the Hamiltonian structure of stochastic hybrid systems by considering a few explicit models taken from neuroscience.

5.1. Binary model

Before considering conductance-based models of membrane voltage fluctuations, we begin with the simple example of a binary stochastic hybrid process (two discrete states $n = 0, 1$), which was analyzed in some detail by Faggionato *et al* [13] using a different method. The latter authors exploited the fact that the model is exactly solvable, in the sense that the stationary density of the corresponding Chapman–Kolmogorov equation can be computed explicitly, and used a fluctuation-dissipation theorem to determine the Hamiltonian and quasipotential. Here we will obtain the same results more

J. Stat. Mech. (2017) 033206

directly by calculating the Perron eigenvalue. One biological application of the binary model is to the bidirectional transport of a molecular motor along a one-dimensional microtubular track, in which x represents the spatial location of the motor on the track and the two discrete states represent the motor moving either towards the + end or − end of the track. (In more complex models, the discrete space Γ represents multiple internal conformational states of the motor, each of which has an associated velocity on the track [3].)

Suppose that the continuous variable evolves according to piecewise dynamics on some finite interval (a, b),

$$\dot{x} = F_n(x), \quad n = 0, 1, \tag{5.1}$$

with F_0, F_1 continuous and locally Lipschitz. Suppose that F_0, F_1 are non-vanishing within the interval (a, b), and $F_n(a) \geqslant 0$, $F_n(b) \leqslant 0$ for $n = 0, 1$; the dynamics is then confined to (a, b). Denote the transition rates of the two-state Markov chain by $\omega_\pm(x)$ with

$$\{n = 0\} \underset{\omega_-(x)}{\overset{\omega_+(x)}{\rightleftharpoons}} \{n = 1\}.$$

The stationary measure of the Markov chain is given by

$$\rho(x, 0) = \frac{\omega_-(x)}{\omega_-(x) + \omega_-(x)}, \quad \rho(x, 1) = \frac{\omega_+(x)}{\omega_-(x) + \omega_-(x)}. \tag{5.2}$$

The linear equation (3.9) can be written as the two-dimensional system

$$\begin{pmatrix} -\omega_+(x) + pF_0(x) & \omega_-(x) \\ \omega_+(x) & -\omega_-(x) + pF_1(x) \end{pmatrix} \begin{pmatrix} R_0 \\ R_1 \end{pmatrix} = \lambda \begin{pmatrix} R_0 \\ R_1 \end{pmatrix}. \tag{5.3}$$

The corresponding characteristic equation is

$$\begin{aligned} 0 = {} & \lambda^2 + \lambda[\omega_+(x) + \omega_-(x) - p(F_0(x) + F_1(x))] \\ & + (pF_1(x) - \omega_-(x))(pF_0(x) - \omega_+(x)) - \omega_-(x)\omega_+(x). \end{aligned}$$

It follows that the Perron eigenvalue is given by

$$\lambda(x, p) = \frac{1}{2}\left[\Sigma(x, p) + \sqrt{\Sigma(x, p)^2 - 4\gamma(x, p)}\right], \tag{5.4}$$

where

$$\Sigma(x, p) = p(F_0(x) + F_1(x)) - [\omega_+(x) + \omega_-(x)],$$

and

$$\gamma(x, p) = (pF_1(x) - \omega_-(x))(pF_0(x) - \omega_+(x)) - \omega_-(x)\omega_+(x).$$

A little algebra shows that

$$D(x, p) \equiv \Sigma(x, p)^2 - 4\gamma(x, p) = [p(F_0 - F_1) - (\omega_+ - \omega_-)]^2 + \omega_+\omega_- > 0,$$

so that as expected λ is real. Hence, from Hamilton's equations

$$\begin{aligned}\dot{x} &= \frac{\partial \lambda(x,p)}{\partial p} \\ &= \frac{F_0(x)+F_1(x)}{2} + \frac{\partial D(x,p)}{\partial p}\frac{1}{2\sqrt{D(x,p)}} \\ &= \frac{F_0(x)+F_1(x)}{2} + \frac{F_0(x)-F_1(x)}{2}\frac{p(F_0-F_1)-(\omega_+-\omega_-)}{\sqrt{[\,p(F_0-F_1)-(\omega_+-\omega_-)]^2+\omega_+\omega_-}},\end{aligned} \tag{5.5}$$

which is the same result as obtained in example 10.4 of Faggionato *et al* [13]. Moreover, writing

$$\dot{x} = F_0(x)\psi_0(x) + F_1(x)\psi_1(x),$$

we see that

$$\psi_0(x) = \frac{1}{2}\left[1 + \frac{p(F_0-F_1)-(\omega_+-\omega_-)}{\sqrt{[\,p(F_0-F_1)-(\omega_+-\omega_-)]^2+\omega_+\omega_-}}\right], \tag{5.6}$$

and

$$\psi_1(x) = \frac{1}{2}\left[1 - \frac{p(F_0-F_1)-(\omega_+-\omega_-)}{\sqrt{[\,p(F_0-F_1)-(\omega_+-\omega_-)]^2+\omega_+\omega_-}}\right], \tag{5.7}$$

so that $\psi_{0,1}(x) \geqslant 0$ with $\psi_0(x)+\psi_1(x)=1$.

5.2. Stochastic Na^+ ion channels and the initiation of spontaneous action potentials

An important example of stochastic hybrid systems at the single-cell level concerns a conductance-based model of a neuron, in which the stochastic opening of membrane ion channels generates a stochastic ionic current that drives the membrane voltage. It is then possible that ion channel noise induces spontaneous action potentials (SAPs), which can have a large effect on a neuron's function [5, 8, 9, 16, 19, 22, 33]. If SAPs are too frequent, a neuron cannot reliably perform its computational role. Hence, ion channel noise imposes a fundamental limit on the density of neural tissue. Smaller neurons must function with fewer ion channels, making ion channel fluctuations more significant and more likely to cause a SAP. Here we will consider the simple case of a single type of ion channel, namely, a fast sodium (Na) channel, which was previously analyzed using WKB methods [22]. Let $x(t)$ denote the membrane voltage of the neuron at time t and N be the fixed number of sodium channels. We assume that each channel can either be open (O) or closed (C), and can switch between each state according to the kinetic scheme

$$C \underset{\beta(x)}{\overset{\alpha(x)}{\rightleftharpoons}} O, \tag{5.8}$$

with voltage-dependent transition rates. (A more detailed biophysical model would need to treat each ion channel as a cluster of subunits rather than a single unit. In other words, the Markov chain of events associated with opening and closing of an ion channel would involve transitions between more than two internal states,.) The stochastic membrane voltage is taken evolves according to the piecewise deterministic equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F_n(x) \equiv \frac{n}{N} f(x) - g(x), \tag{5.9}$$

where n is the number of open ion channels at time t, and

$$f(x) = g_{\mathrm{Na}}(V_{\mathrm{Na}} - x), \quad g(x) = -g_{\mathrm{L}}[V_{\mathrm{L}} - x] - I.$$

Here g_{Na} is the maximal conductance of a sodium channel and V_{Na} is the corresponding membrane reversal potential. Similarly, g_L and V_l are the effective maximal conductance and reversal potential of any other currents, which are assumed to be independent of the opening and closing of ion channels, and I is an external current. The four quantities (g_{Na}, g_L, V_{Na}, V_L) are taken to be constants. Since the right-hand side of (5.9) is negative for large x and positive for small x, it follows that the voltage x is confined to some interval $\Omega = [x_L, x_R]$. The function $F_n(x)$ is clearly continuous and locally Lipschitz.

In this example the space Γ of discrete states is the set of integers $\{n = 0, 1, \ldots, N\}$ and the Markov chain is given by a birth-death process:

$$n \underset{\omega_+(n,x)/\epsilon}{\rightarrow} n + 1, \quad n \underset{\omega_-(n,x)/\epsilon}{\rightarrow} n - 1, \tag{5.10}$$

with transition rates

$$\omega_+(x, n) = \alpha(x)(N - n), \quad \omega_-(x, n) = \beta(x)n. \tag{5.11}$$

The small parameter ϵ reflects the fact that sodium channels open at a much faster rate than the relaxation dynamics of the voltage [22]. It follows that the matrix $\mathbf{A}(x)$ for fixed x is tridiagonal matrix with

$$A_{n-1,n}(x) = \omega_+(x, n - 1), \ A_{n+1,n}(x) = \omega_-(x, n + 1), \tag{5.12a}$$

$$A_{nn}(x) = -\omega_+(x, n) - \omega_-(n) \tag{5.12b}$$

for $n = 0, 1, \ldots, N$. It is straightforward to show that the Markov chain is ergodic with unique invariant measure (for fixed n) given by

$$\rho(x, n) = \frac{N!}{(N - n)!\, n!} a(x)^n b(x)^{N-n}, \tag{5.13}$$

with

$$a(x) = \frac{\alpha(x)}{\alpha(x) + \beta(x)}, \quad b(x) = \frac{\beta(x)}{\alpha(x) + \beta(x)}. \tag{5.14}$$

The above stochastic hybrid system satisfies all of the conditions specified in section 2. Hence, the law of large numbers implies that in the mean-field limit $\epsilon \to 0$, we obtain the deterministic kinetic equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \overline{F}(x) \equiv a(x)f(x) - g(x), \tag{5.15}$$

J. Stat. Mech. (2017) 033206

where

$$a(x) = \langle n \rangle / N, \quad \langle n \rangle = \sum_{n=1}^{N} n\rho(x, n),$$

and ρ is the stationary density (5.13). One of the features of the averaged model is that it can exhibit bistability for a range of physiologically reasonable parameter values. This is illustrated in figure 1, where we plot the deterministic potential $U(x) = -\mathrm{d}\overline{F}/\mathrm{d}x$ as a function of x. Here x_- represents a resting state of the neuron, whereas x_+ represents an active state; noise-induced transitions from x_- to x_+ can be interpreted in terms of the initiation of a spontaneous action potential. Elsewhere, WKB methods and matched asymptotics have been used to calculate the MFPT to escape from x_-. Here, we will focus on the quasipotential and its relation to the Perron eigenvalue

Substituting the explicit expressions for $\mathbf{A}$ and $F_n(x)$ into equation (3.9), yields the following equation for the Perron eigenvalue λ and the right eigenvector $\mathbf{R}$:

$$\begin{aligned}(N-n+1)\alpha(x)R_{n-1} - [\lambda + n\beta(x) + (N-n)\alpha(x)]R_n \\ +(n+1)\beta(x)R_{n+1} = -p\left(\frac{n}{N}f(x) - g(x)\right)R_n.\end{aligned} \tag{5.16}$$

Consider the trial solution [5]

$$R_n(x, p) = \frac{\Gamma(x, p)^n}{(N-n)!\,n!}, \tag{5.17}$$

which yields the following equation relating Γ and λ:

$$\frac{n\alpha}{\Gamma} + \Gamma\beta(N-n) - \lambda - n\beta - (N-n)\alpha = -p\left(\frac{n}{N}f - g\right).$$

Collecting terms independent of n and terms linear in n yields the pair of equations

$$p = -\frac{N}{f(x)}\left(\frac{1}{\Gamma(x, p)} + 1\right)(\alpha(x) - \beta(x)\Gamma(x, p)), \tag{5.18}$$

and

$$\lambda(x, p) = -N(\alpha(x) - \Gamma(x, p)\beta(x)) - pg(x). \tag{5.19}$$

Eliminating Γ from these equation gives

$$p = \frac{1}{f(x)}\left(\frac{N\beta(x)}{\lambda(x, p) + N\alpha(x) + pg(x)} + 1\right)(\lambda(x, p) + pg(x)).$$

This yields a quadratic equation for λ of the form

$$\lambda^2 + \sigma(x)\lambda - h(x, p) = 0, \tag{5.20}$$

with

$$\begin{aligned}\sigma(x) &= (2g(x) - f(x)) + N(\alpha(x) + \beta(x)), \\ h(x, p) &= p[-N\beta(x)g(x) + (N\alpha(x) + pg(x))(f(x) - g(x))].\end{aligned}$$

J. Stat. Mech. (2017) 033206

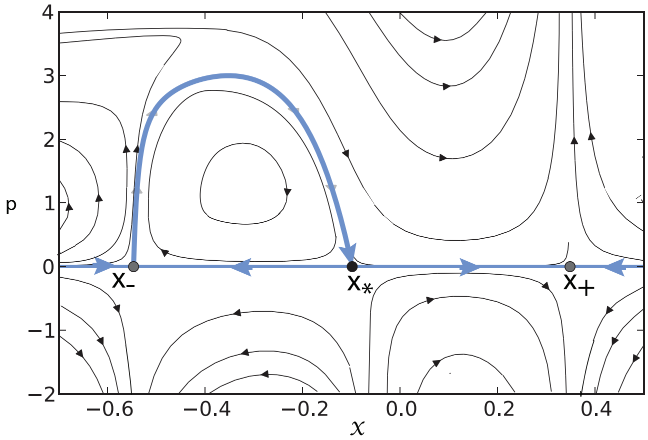

Figure 2. Phase portrait of Hamilton's equations of motion for the ion channel model with Hamiltonian given by the Perron eigenvalue. (x and p are taken to be dimensionless.) The zero energy solution representing the maximum likelihood path of escape from x_- is shown as the gray curve. (The corresponding path from x_+ is not shown.) Same parameter values as figure 1 and $I = 0$.

Given the Perron eigenvalue, we can determine the quasipotential $\Phi(x)$ by solving the Hamilton–Jacobi equation $\lambda(x, \partial_x\Phi) = 0$. Equation (5.20) then yields the reduced Hamilton–Jacobi equation

$$h(x, \partial_x\Phi_0) = 0. \tag{5.21}$$

The latter is precisely the equation for the quasipotential previously derived using WKB methods [33]. It has the following pair of solutions for $\Phi_0' = \partial_x\Phi_0$:

$$\Phi_0' = 0 \text{ and } \Phi_0'(x) = -N\frac{\alpha(x)f(x) - (\alpha(x) + \beta)g(x)}{g(x)(f(x) - g(x))}. \tag{5.22}$$

The trivial solution Φ_0 = constant occurs along deterministic trajectories, which converge to the fixed point, whereas the non-trivial solution for $\Phi_0(x)$ occurs along the most likely escape trajectories. In figure 2 we show solutions to Hamilton's equations in the (x, p)-plane, highlighting the zero energy maximum likelihood curve linking x_- and x_0.

5.3. Stochastic Morris–Lecar model

One of the major simplifications of the above model is to assume that the slow potassium channels are frozen. If one now incorporates the slow opening and closing of these channels, then the underlying deterministic system becomes excitable rather than bistable. That is, there is a single stable fixed point such that for small stimuli the voltage returns directly to rest, whereas for stronger stimuli the voltage makes a large detour before returning to rest, which corresponds to an action potential. There is no longer a well-defined, unique firing threshold. A simple deterministic model of neural excitability is the Morris–Lecar (ML) model [29]:

$$\dot{x} = a(x)f_{\mathrm{Na}}(x) + yf_{\mathrm{K}}(x) - g(x), \tag{5.23}$$

J. Stat. Mech. (2017) 033206

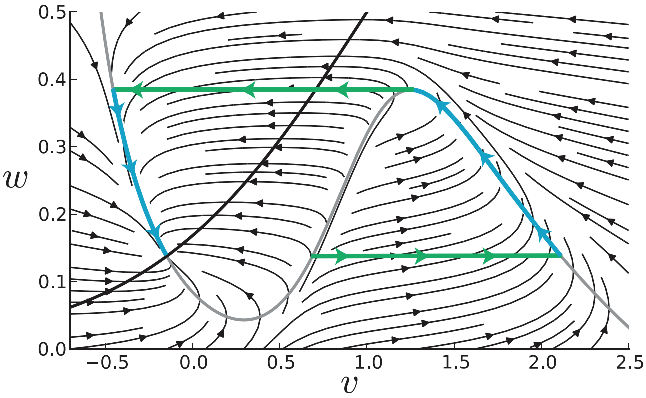

Figure 3. Deterministic phase plane dynamics. Thick curves show the nullclines: $\dot{x}=0$ as grey and $\dot{y}=0$ as black. Black stream lines represent deterministic trajectories. Green/blue curves represent an action potential trajectory in the limit of slow y.

$$\dot{y} = \frac{y_\infty(x) - y}{\tau_y(x)}, \tag{5.24}$$

where x is voltage and y represents the fraction of open K^+ channels. The dynamics of this system can be explored using phase-plane analysis as illustrated in figure 3. A slow/fast analysis of the deterministic system suggests that the initiation of an action potential occurs without any change in w, thus motivating the analysis of Keener and Newby [22]. However, it turns out that this adiabatic approximation breaks down when stochastic fluctuations in the opening and closing of K^+ channels are taken into account. This can be established by extending the WKB analysis outlined in section 4 to a stochastic version of the ML model [33]. Since one now has two continuous variables x and y in the deterministic limit, it follows that stochastic trajectories in the phase-plane correspond to characteristic projections of an underlying Hamiltonian dynamical system. In general, it is difficult to solve FPT problems in more than one dimension. In the case of a metastable state with a well-defined basin of attraction, one has to calculate the MFPT to cross the separatrices forming the boundary of the basin of attraction. There is an additional level of complexity for an excitable system, due to the fact that there is no well-defined deterministic separatrix. Interestingly, one finds that the stochastic ML model has an effective separatrix that any stochastic trajectory has to cross in order to generate a stochastic action potential [33], see also [25].

Here we focus on calculating the Hamiltonian associated with the LDP for the stochastic Morris–Lecar model. Suppose that at time t there are $n = 0, 1, \cdots, N$ open Na channels and $m = 0, 1, \cdots, M$ open K channels. The membrane voltage x then evolves as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x, m, n) \equiv \frac{n}{N} f_{\mathrm{Na}}(x) + \frac{m}{M} f_{\mathrm{K}}(x) + f_{\mathrm{L}}(x) + I. \tag{5.25}$$

We assume that each channel is either open or closed and switches between each state according to

$$O \underset{\beta_i(x)}{\overset{\alpha_i(x)}{\rightleftharpoons}} C, \quad i = \mathrm{Na},\ \mathrm{K}. \tag{5.26}$$

The space Γ of discrete states is the set of integer pairs (n, m) with $0 \leqslant n \leqslant wN, 0 \leqslant m \leqslant M$ and the Markov chain is given by a birth-death process

$$(n, m) \underset{\omega^{\pm}_{Na}(n,x)}{\rightarrow} (n \pm 1, m), \quad (n, m) \underset{\omega^{\pm}_{\mathrm{K}}(m,x)}{\rightarrow} (n, m \pm 1), \tag{5.27}$$

with transition rates

$$\omega^{-}_{\mathrm{Na}}(n, x) = n\beta_{\mathrm{Na}}(x), \quad \omega^{+}_{\mathrm{Na}}(n, x) = (N - n)\alpha_{\mathrm{Na}}(x), \tag{5.28a}$$

$$\omega^{-}_{\mathrm{K}}(m, x) = m\beta_{\mathrm{K}}(x), \quad \omega^{+}_{\mathrm{K}}(m, x) = (M - m)\alpha_{\mathrm{K}}(x). \tag{5.28b}$$

In contrast to the bistable sodium ion channel model of section 5.2, we cannot treat both the Na and K channel kinetics as fast, and therefore we cannot develop a variational problem by scaling all transition rates in terms of a small parameter ϵ and applying the analysis of section 5. In fact, rather than a piecewise deterministic system, we now have a multi-scale stochastic system, in which both fast and slow processes are intrinsically stochastic. Multi-scale stochastic processes also arise in models of gene regulatory networks [30]. We will proceed along similar lines to [7] by treating $n(t)$ as a fast variable with $\alpha_{\mathrm{Na}}, \beta_{\mathrm{Na}} = \mathcal{O}(1/\epsilon)$, and treating $y(t) = m(t)/M$ as a continuous (recovery) variable with $M = 1/\epsilon$. We can then derive a piecewise hybrid SDE by carrying out a system size expansion with respect to y. Setting $M\Omega_{\pm}(x, y) = \omega^{\pm}_{m}(My, x)$, we obtain the SDE [7]

$$\mathrm{d}X(t) = F_n(X, Y)\mathrm{d}t \tag{5.29a}$$

$$\mathrm{d}Y(t) = [\Omega_{+}(X, Y) - \Omega_{-}(X, Y)]\mathrm{d}t + \sqrt{\epsilon\sigma^2(X, Y)}\,\mathrm{d}W(t), \tag{5.29b}$$

with

$$F_n(x, y) = \frac{n}{N}f_{\mathrm{Na}}(x) + yf_{\mathrm{K}}(x) + f_{\mathrm{L}}(x) + I, \quad \sigma^2(x, y) = \Omega_{+}(x, y) + \Omega_{-}(x, y). \tag{5.30}$$

We can now determine the Hamiltonian of the associated LDP by combining our analysis in sections 3.4 and 3.5. That is, $H(\mathbf{x}, \mathbf{p}) = \lambda(\mathbf{x}, \mathbf{p})$ with $\mathbf{x} = (x, y), \mathbf{p} = (p_x, p_y)$, and λ is the Perron eigenvalue of the linear equation

$$\begin{aligned}\lambda(\mathbf{x}, \mathbf{p})R_n(\mathbf{x}, \mathbf{p}) = &\sum_m A^{\top}_{nm}(\mathbf{x})R_m(\mathbf{x}, \mathbf{p}) \\ &+ \{p_xF_n(\mathbf{x}) + p_y[\Omega_{+}(x, y) - \Omega_{-}(x, y)]\}R_n(\mathbf{x}, \mathbf{p}) \\ &+ \frac{1}{2}p_y^2\sigma^2(x, y)R_n(\mathbf{x}, \mathbf{p}),\end{aligned} \tag{5.31}$$

where the matrix $\mathbf{A}$ is the same as the Na model of section 5.2. Equation (5.31) can be solved along similar lines to (5.16) using the Ansatz $R_n(\mathbf{x}, \mathbf{p}) = \Gamma(\mathbf{x}, \mathbf{p})^n/(n!\,(N - n)!)$. Collecting terms linear in n gives

$$\Gamma(\mathbf{x}, \mathbf{p}) = \alpha_{\mathrm{Na}}(x) - \frac{1}{N}(p_xg(x, w) + h(x, y, p_y) - \lambda(\mathbf{x}, \mathbf{p})),$$

where $g(x, y) = yf_{\mathrm{K}}(x) + f_L(x)$ and

J. Stat. Mech. (2017) 033206

$$h(x, y, p_y) = p_y[\Omega_+(x, y) - \Omega_-(x, y)] + \frac{1}{2}p_y^2\sigma^2(x, y). \tag{5.32}$$

On the other hand, collecting terms independent of n and substituting for $\Gamma(\mathbf{x}, \mathbf{p})$ gives the following quadratic equation for λ:

$$\lambda^2 - (2h(x, y, p_y) + k(x, y, p_x))\lambda + h_1(\mathbf{x}, \mathbf{p}) = 0, \tag{5.33}$$

with

$$k(x, y, p_x) = (2g(x, y) + f_{\mathrm{Na}}(x))p_x - N/(1 - y_\infty(x)),$$

and

$$\begin{aligned} h_1(\mathbf{x}, \mathbf{p}) = {} & (2g(x, y) + f_{\mathrm{Na}}(x))p_x h(x, y, p_y) \\ & + (f_{\mathrm{Na}}(x) + g(x, y))g(x, y)p_x^2 + h(x, y, p_y)^2 \\ & - \frac{N}{1 - y_\infty(x)}([y_\infty(x)f_{\mathrm{Na}}(x) + g(x, y)]p_x + h(x, y, p_y)), \end{aligned} \tag{5.34}$$

with $y_\infty(x) = \alpha_{\mathrm{K}}(x)/(\alpha_{\mathrm{K}}(x) + \beta_{\mathrm{K}}(x))$. Note, in particular, for escape problems we are interested in zero energy solutions of Hamilton's equations, which reduce to solutions of $h_1(\mathbf{x}, \mathbf{p}) = 0$. Our derivation of the Hamiltonian based on an LDP is equivalent to one obtained previously using formal WKB methods [33].

Acknowledgments

We thank Jay Newby for drawing our attention to the work of Yuri Kifer [24] and helpful discussion. The work was conducted while PCB was visiting the NeuroMathComp group where he holds an INRIA International Chair. OF was partially supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 269921 (BrainScaleS), no. 318723 (Mathematics), and by the ERC advanced grant NerVi no. 227747.

References

[1] Anderson D, Ermentrout G B and Thomas P J 2015 Stochastic representations of ion channel kinetics and exact stochastic simulation of neuronal dynamics *J. Comput. Neurosci.* **38** 67–82
[2] Assaf M and Meerson B 2006 Spectral theory of metastability and extinction in birth-death systems *Phys. Rev. Lett.* **97** 200602
[3] Bressloff P C and Newby J M 2013 Metastability in a stochastic neural network modeled as a velocity jump Markov process *SIAM Appl. Dyn. Syst.* **12** 1394–435
[4] Bressloff P C and Newby J M 2014 Path-integrals and large-deviations in stochastic hybrid systems *Phys. Rev.* E **89** 042701
[5] Bressloff P C and Newby J M 2014 Stochastic hybrid model of spontaneous dendritic NMDA spikes *Phys. Biol.* **11** 016006
[6] Bressloff P C 2015 Path-integral methods for analyzing the effects of fluctuations in stochastic hybrid neural networks *J. Math. Neurosci.* **5** 1–33
[7] Brooks H A and Bressloff P C 2015 Quasicycles in the stochastic hybrid Morris–Lecar neural model *Phys. Rev.* E **92** 012704
[8] Buckwar E and Riedler M G 2011 An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution *J. Math. Biol.* **63** 1051–93
[9] Chow C C and White J A 1996 Spontaneous action potentials due to channel fluctuations *Biophys. J.* **71** 3013–21

[10] Davis M H A 1984 Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models *J. R. Soc.* B **46** 353–88

[11] Dembo A and Zeitouni O 1998 *Large Deviations Techniques and Applications* (*Applications of Mathematics* vol 38) (Berlin: Springer)

[12] Dykman M I, Mori E, Ross J and Hunt P M 1994 Large fluctuations and optimal paths in chemical kinetics *J. Chem. Phys.* A **100** 5735–50

[13] Faggionato A, Gabrielli D and Crivellari M R 2009 Non-equilibrium thermodynamics of piecewise deterministic Markov Processes *J. Stat. Phys.* **137** 259–304

[14] Faggionato A, Gabrielli D and Crivellari M R 2010 Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors *Markov Process. Relat. Fields* **16** 497–548

[15] Feng J and Kurtz T G 2006 *Large Deviations for Stochastic Processes* (Providence, RI: American Mathematical Society)

[16] Fox R F and Lu Y N 1994 Emergent collective behavior in large numbers of globally coupled independent stochastic ion channels *Phys. Rev.* E **49** 3421–31

[17] Freidlin M I and Wentzell A D 1998 *Random Perturbations of Dynamical Systems* (New York: Springer)

[18] Gardiner C W 2009 *Handbook of Stochastic Methods* 4th edn (Berlin: Springer)

[19] Goldwyn J H and Shea-Brown E 2011 The what and where of adding channel noise to the Hodgkin–Huxley equations *PLoS Comput. Biol* **7** e1002247

[20] Grimmett G R and Stirzaker D R 2001 *Probability and Random Processes* 3rd edn (Oxford: Oxford University Press)

[21] Hanggi P, Grabert H, Talkner P and Thomas H 1984 Bistable systems: master equation versus Fokker–Planck modeling *Phys. Rev. A* **29** 371–8

[22] Keener J P and Newby J M 2011 Perturbation analysis of spontaneous action potential initiation by stochastic ion channels *Phy. Rev.* E **84** 011918

[23] Kepler T B and Elston T C 2001 *Biophys. J.* **81** 3116–36

[24] Kifer Y 2009 Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging *Mem. AMS* **201** 944

[25] Khovanov I A, Polovinkin A V, Luchinsky D G and McClintock P V E 2013 Noise-induced escape in an excitable system *Phys. Rev.* E **87** 032116

[26] Knessl C, Matkowsky B J, Schuss Z and Tier C 1985 An asymptotic theory of large deviations for Markov jump processes *SIAM J. Appl. Math.* **46** 1006–28

[27] Matkowsky B J and Schuss Z 1977 The exit problem for randomly perturbed dynamical systems *SIAM J. Appl. Math.* **33** 365–82

[28] Naeh T, Klosek M M, Matkowsky B J and Schuss Z 1990 A direct approach to the exit problem *SIAM J. Appl. Math.* **50** 595–627

[29] Morris C and Lecar H 1981 Voltage oscillations in the barnacle giant muscle fiber *Biophys. J.* **35** 193–213

[30] Newby J M 2012 Isolating intrinsic noise sources in a stochastic genetic switch *Phys. Biol.* **9** 026002

[31] Newby J M and Bressloff P C 2010 Local synaptic signalling enhances the stochastic transport of motor-driven cargo in neurons *Phys. Biol.* **7** 036004

[32] Newby J M and Keener J P 2011 An asymptotic analysis of the spatially inhomogeneous velocity-jump process *Multiscale Model. Simul.* **9** 735–65

[33] Newby J M, Bressloff P C and Keener J P 2013 Breakdown of fast-slow analysis in an excitable system with channel noise *Phys. Rev. Lett.* **111** 128101

[34] Newby J M and Chapman S J 2013 Metastable behavior in Markov processes with internal states *J. Math Biol.* **69** 941–76

[35] Newby J M 2014 Spontaneous excitability in the Morris–Lecar model with ion channel noise *SIAM J. Appl. Dyn. Syst.* **13** 1756–91

[36] Newby J M 2015 Bistable switching asymptotics for the self regulating gene *J. Phys. A: Math. Theor.* **48** 185001

[37] Pakdaman K, Thieullen M and Wainrib G 2010 Fluid limit theorems for stochastic hybrid systems with application to neuron models *Adv. Appl. Probab.* **42** 761–94

[38] Rockafellar R T 1970 *Convex Analysis* (Princeton, NJ: Princeton University Press)

[39] Maier R S and Stein D L 1997 Limiting exit location distribution in the stochastic exit problem *SIAM J. Appl. Math.* **57** 752–90

[40] Touchette H 2009 The large deviation approach to statistical mechanics *Phys. Rep.* **478** 1–69

[41] Vellela M and Qian H 2009 Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited *J. R. Soc. Interface* **6** 925–40

[42] Wainrib G, Thieullen M and Pakdaman K 2012 Reduction of stochastic conductance-based neuron models with time-scales separation *J. Comput. Neurosci.* **32** 327–46