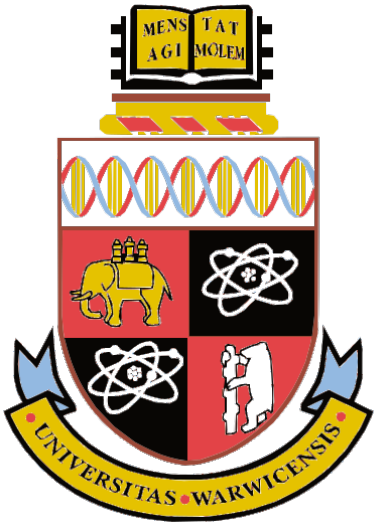

Splitting integrators for spiking neuronal models

by

Ramón Dineth Nartallo-Kaluarachchi

U1804657

MA4K9 Dissertation

Submitted to The University of Warwick

Supervised by **Dr Massimiliano Tamborrino**

Administrated by **Dr Martin Lotz**

Mathematics Institute

April, 2022

Contents

Acknowledgements

I would like to thank Dr Massimiliano Tamborrino for supervising this project. Your time, patience and kindness has been most appreciated. You have taught me lessons I will carry through a lifetime of mathematical research.

I would like thank Dr Martin Lotz for administrating this project. Your time and your attention is also dearly appreciated.

This dissertation is dedicated to Eona, for never getting tired of hearing about my spikes.

"There is no scientific study more vital to man than the study of his own brain. Our entire view of the universe depends on it." - Francis Crick

1 Introduction

In this dissertation, we aim to construct and analyse splitting integrators for spiking neuronal models. These are methods from geometric numerical analysis for the approximate solution of systems of ordinary or stochastic differential equations. The aim is to study whether these numerical methods could preserve the spiking dynamics of the models more robustly than traditional schemes. We will focus on two splitting approaches and four neuronal models. The first is the "canonical splitting" proposed by Buckwar et al for the stochastic FitzHugh-Nagumo model [2]. The second is an approach for "conditionally linear systems of ODEs" proposed by Chen et al, and illustrated on the deterministic Van der Pol and Hodgkin-Huxley models [3]. We extend the latter approach to systems of SDEs and prove that the 1-step transition probability of the resulting follows a normal distribution under specific conditions, a desirable property for the method to be used in parameter estimation [2]. Furthermore, we show that the order of composition can affect the hypoellipticity of the method. The four models considered here are the deterministic and stochastic Van der Pol (VDP) [4], the Fitz-Hugh Nagumo (FHN) ([2]), the Izhikevich (IZ) [12] and the adaptive exponential integrate-and-fire models (AdEX) [1]. By generalising a coordinate change proposed by León and Samson [14] for the stochastic FHN, we also extend Stern's approach to models that are not originally in this "conditionally linear" form. We compare these splitting integrators with classical finite difference methods such as Forward Euler and Runge-Kutta (fourth-order), for the deterministic models, as well as Euler-Maruyama and tamed Euler-Maruyama, for the stochastic models.

In the numerical experiments, Section 5, we reinforce existing results and present new results showing that the splitting methods for the deterministic and stochastic oscillatory models preserve the dynamics of the spiking behaviour (i.e. amplitudes and frequencies) more robustly than classical integrators. We analyse the preservation of limit cycles, the distributions of the inter-spike-intervals and the quadratic errors of all these methods. The canonical splitting proved to be the most robust in almost all cases outperforming the Stern approach and the classical integrators, except for in very "stiff" deterministic settings. We perform similar analysis for the IZ and AdEx models, noting that these "spike-reset" models appear to be more sensitive to the coordinate changes, limiting the performance of Stern's approach. We observe that all numerical methods are also less robust to changes in time-step for these models due to error in the first spiking time - a studied difficulty when integrating these models [19].

2 Splitting integrators

2.1 Lie-Trotter and Strang compositions

Splitting methods are a geometric approach to numerical integration. The approach consists of decomposing a vector field into exactly integrable components.

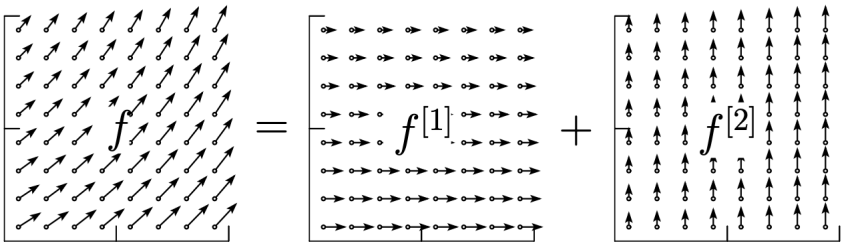

Figure 1: Diagram showing the splitting of a vector field from [6]

Consider an initial value problem (IVP) given by the dynamical system:

$$dx(t) = f(x(t))\, dt; \quad x(0) = x_0; \quad t \in [0, T] \tag{1}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ and $x : \mathbb{R} \to \mathbb{R}^d$. In general, for a Lipschitz-continuous, non-linear, vector field, f, this problem has a unique solution that cannot be written in a closed form. As such, we aim to numerically approximate the solution, $x(t)$, at discrete time points t_i i.e. $x_i \approx x(t_i)$. In order to approximate the solution to this problem, we discretise time using a constant step size $h > 0$. Thus we define, for $i \in \mathbb{N}$,

$$t_i = ih. \tag{2}$$

We will focus on **explicit** methods meaning that our approximation x_i depends only on x_j with $0 \leq j < i$. Therefore, our approximations will form a recursive sequence $(x_i)_{i \in \mathbb{N}}$. To this end, we adopt a splitting approach. We can split the vector field into two (or more) sub-components,

$$dx(t) = f_1(x(t)) + f_2(x(t))\, dt; \quad x(0) = x_0; \quad t \in [0, T] \tag{3}$$

with $f_i : \mathbb{R} \to \mathbb{R}$ for $i = 1, 2$. We use the sub-components to generate two separate IVPs, given by,

$$dx_1(t) = f_1(x_1(t))\, dt; \quad x_1(0) = x_0^{[1]}; \quad t \in [0, T] \tag{4}$$

$$dx_2(t) = f_2(x_2(t))\, dt; \quad x_2(0) = x_0^{[2]}; \quad t \in [0, T] \tag{5}$$

where $x_1(t) : \mathbb{R} \to \mathbb{R}^d$ and $x_2(t) : \mathbb{R} \to \mathbb{R}^d$ are solutions to the corresponding problem. The aim is to choose the f_i in such a way that the exact solution to the sub-problems can be written in a explicit closed form. Assuming the sub-systems can be solved explicitly, let $\varphi_t^{[k]}(x_0)$ denote the flow of the k-th sub-equation at time t starting from x_0. As such, we would like to take $x_0^{[1]} = x_0$ in the first step, and $x_0^{[2]} = \varphi_h^{[1]}(x_0)$ for some small time step $h > 0$.

Starting from a point, x_0, we can solve according to one sub-vector field and then the other. As in Figure 2, with $x(0)$ acting as y_0, this composition of the exact flows of the sub-vector fields gives an approximation of the flow of the total vector field, and so an approximation of the solution to the IVP. We can use this composition to

$$\boldsymbol{\Phi}_h^* = \varphi_h^{[2]} \circ \varphi_h^{[1]}$$
$$\boldsymbol{\Phi}_h = \varphi_h^{[1]} \circ \varphi_h^{[2]}$$

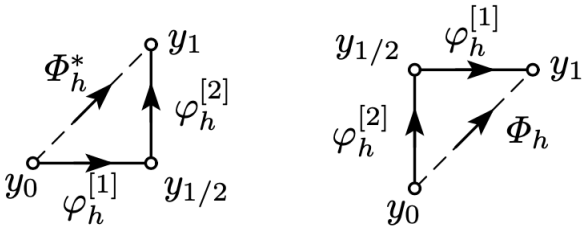

Figure 2: Diagram showing the Lie-Trotter approach from [6]

define an explicit numerical scheme. This is the so-called *Lie-Trotter approach*,

$$\tilde{x}^{LT}(t_i) = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(\tilde{x}^{LT}(t_{i-1})), \tag{6}$$

as in the right hand scheme in Figure 2. Using the same idea, we can flow according to $\varphi^{[1]}$ for time $h/2$, flow according to $\varphi^{[2]}$ for time h and, then flow again according to $\varphi^{[1]}$ for time $h/2$, as in Figure 3 defining another explicit numerical scheme. This

$$\boldsymbol{\Phi}_h^{[S]} = \varphi_{h/2}^{[1]} \circ \varphi_h^{[2]} \circ \varphi_{h/2}^{[1]}\,,$$

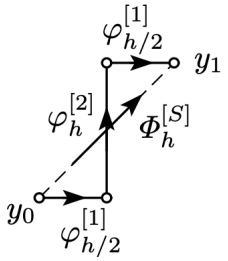

Figure 3: Diagram showing the Strang approach from [6]

approach is the so-called *Strang approach*,

$$\tilde{x}^{S}(t_i) = (\varphi_{h/2}^{[1]} \circ \varphi_h^{[2]} \circ \varphi_{h/2}^{[1]})(\tilde{x}^{S}(t_{i-1})). \tag{7}$$

Remark 1. *Switching $\varphi^{[1]}$ and $\varphi^{[2]}$ in both the Lie-Trotter and Strang approaches*

yields two further numerical schemes.

2.2 The canonical splitting

One potential approach to split vector fields for semi-linear equations is presented by Buckwar et al [2]. The idea is to split the vector field, f, into a linear and a non-linear part, thus generating two IVPs to solve. In this case, the IVP with a linear vector field will be exactly integrable and the IVP with the non-linear vector field may or may not be solvable depending on the chosen splitting and the considered model. This approach will henceforth be referred to as *the canonical splitting*.

Consider again the IVP (1) with $f(x(t)) = Ax(t) + N(x(t))$ where $A \in \mathbb{R}^{d\times d}$ is a matrix and $N : \mathbb{R}^d \to \mathbb{R}^d$ is a non-linear function. In the canonical splitting, we pick

$$f^{[1]}(x(t)) = Ax(t), \tag{8}$$
$$f^{[2]}(x(t)) = N(x(t)). \tag{9}$$

Remark 2. *This splitting is not necessarily unique as linear terms can also be included in $N(x)$.*

This gives sub-systems

$$dx^{[1]}(t) = Ax^{[1]}(t)\, dt; \quad x^{[1]}(0) = x_0^{[1]}; \quad t \in [0, T], \tag{10}$$
$$dx^{[2]}(t) = N(x^{[2]}(t))\, dt; \quad x^{[2]}(0) = x_0^{[2]}; \quad t \in [0, T]. \tag{11}$$

Integrating the first system exactly gives us

$$x^{[1]}(t) = e^{At}x_0^{[1]}. \tag{12}$$

In particular, this gives us the $h-$time flow

$$\varphi_h^{[1]}(x^{[1]}(t_{i-1})) := x^{[1]}(t_i) = e^{Ah}x^{[1]}(t_{i-1}), \tag{13}$$

where e^{Ah} is an exponential matrix. We assume that N is such that the exact solution to the IVP in (11) can be derived exactly. Therefore, we can write down the second flow,

$$\varphi_h^{[2]}(x^{[2]}(t_{i-1})) := x^{[2]}(t_i) = g(x^{[2]}(t_{i-1}), h), \tag{14}$$

where $g : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ and $g(x_0^{[2]}, t)$ solves the second IVP (11).

Applying the Lie-Trotter and Strang compositions, we obtain the following explicit numerical schemes,

$$\tilde{x}^{LT}(t_i) = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(\tilde{x}^{LT}(t_{i-1})) = e^{Ah} g(\tilde{x}^{LT}(t_{i-1}), h), \tag{15}$$

$$\tilde{x}^{S}(t_i) = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(\tilde{x}^{S}(t_{i-1})) = g(e^{Ah} g(\tilde{x}^{LT}(t_{i-1}), h/2), h/2). \tag{16}$$

This method generalises to a stochastic setting. Consider a stochastic differential equation (SDE) of additive noise type,

$$dx(t) = f(x(t))\, dt + \Sigma\, dW(t); \quad x(0) = x_0; \quad t \in [0, T], \tag{17}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$, $f(x(t)) = Ax(t) + N(x(t))$, $x : \mathbb{R} \to \mathbb{R}^d$ and $\Sigma : \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is a symmetric positive semi-definite matrix. W is the standard $d-$dimensional Wiener process.
We split the vector field f as before but including the noise term in the linear sub-equation. This gives us two sub-systems,

$$dx^{[1]}(t) = Ax^{[1]}(t)\, dt + \Sigma\, dW(t); \quad x^{[1]}(0) = x_0^{[1]}; \quad t \in [0, T] \tag{18}$$

$$dx^{[2]}(t) = N(x^{[2]}(t))\, dt; \quad x^{[2]}(0) = x_0^{[2]}; \quad t \in [0, T]. \tag{19}$$

The second sub-equation and its flow remain unchanged. The first sub-equation now has exact solution,

$$x^{[1]}(t) = e^{At} x_0^{[1]} + \int_0^t e^{A(t-s)} \Sigma\, dW(s). \tag{20}$$

Following the analysis in [2], we can see that the stochastic integral is normally distributed with mean 0 and covariance matrix,

$$C(t) = \int_0^t e^{A(t-s)} \Sigma\Sigma^\top (e^{A(t-s)})^\top\, ds, \tag{21}$$

where $A^\top$ represents the transpose of the matrix i.e. $A = (a_{ij})$ has transpose $A^\top = (a_{ji})$. Hence the first flow becomes

$$\varphi_h^{[1]}(x^{[1]}(t_{i-1})) := x^{[1]}(t_i) = e^{Ah} x^{[1]}(t_{i-1}) + \xi_{i-1}, \tag{22}$$

where the ξ_i are independent, identically distributed $d-$dimensional Gaussian vectors with mean 0 and covariance matrix $C(h)$. This gives us new numerical schemes,

$$\tilde{x}^{LT}(t_i) = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(\tilde{x}^{LT}(t_{i-1})) = e^{Ah} g(\tilde{x}^{LT}(t_{i-1}), h) + \xi_{i-1}, \tag{23}$$

$$\tilde{x}^{S}(t_i) = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(\tilde{x}^{S}(t_{i-1})) = g(e^{Ah} g(\tilde{x}^{LT}(t_{i-1}), h/2) + \xi_{i-1}, h/2). \tag{24}$$

2.3 Conditional linearity and Stern's approach

A different approach to splitting the vector field is presented by Chen et al in [3]. This splitting is for conditionally linear systems. It will be henceforth referred to as *Stern's approach* after Ari Stern who presented the method at the Warwick Seminar for Algorithms & Computationally Intensive Inference [18].

Definition 1. *A **conditionally linear ODE system** is a system of ODEs of the form*

$$dx_i(t) = a_i(x(t))x_i(t) + b_i(x(t))\ dt \quad i = 1, ..., d, \tag{25}$$

*where $x(t) = (x_1(t), x_2(t), ..., x_d(t))$ and the a_i, b_i are (real-valued) functions that **do not depend on** x_i.*

Notice that if the x_j with $j \neq i$ are constant then $a_i(x)$ and $b_i(x)$ are constants and each ODE in the system reduces to the following first order linear ODE with constant coefficients

$$dx_i(t) = a_i x_i(t) + b_i\ dt. \tag{26}$$

Following the analysis from [3], the exact t-time flow starting from x_0 has form,

$$x_i(t) = \begin{cases} \exp(ta_i)x_0 - \frac{b_i}{a_i}(1 - \exp(ta_i)) & \text{if } a_i \neq 0 \\ x_0 + b_i t & \text{if } a_i = 0 \end{cases}. \tag{27}$$

Chen et al use this property to define the following splitting approach [3].
Consider a conditionally linear system of ODEs where $f : \mathbb{R}^d \to \mathbb{R}^d$ is the vector field with components $f_i : \mathbb{R}^d \to \mathbb{R}$ given by,

$$f_i(x(t)) = a_i(x(t))x_i(t) + b_i(x(t)) \quad i = 1, ..., d, \tag{28}$$

where $a_i(x(t)), b_i(x(t))$ depend only on components $x_j(t)$ with $j \neq i$. Then, for $i = 1, ..., d$, we define sub-vector fields $f^{(i)} : \mathbb{R}^d \to \mathbb{R}^d$ with j^{th} component given by,

$$f_j^{(i)} = \begin{cases} f_i(x) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \tag{29}$$

Clearly, we have

$$f = f^{(1)} + f^{(2)} + ... + f^{(d)}, \tag{30}$$

and so the vector field has been decomposed into sub-vector fields that are explicitly

solvable. The exact solution to each sub-problem is known because, for each vector field $f^{(i)}$, all $x_j(t)$ with $j \neq i$ are kept constant whilst $x_i(t)$ evolves according to a linear ODE with constant coefficients. Let $\varphi_h^{(i)} : \mathbb{R}^d \to \mathbb{R}^d$ denote the exact h-time flow of $f^{(i)}$. We can directly integrate to get the j^{th} component,

$$\varphi_{h,j}^{(i)}(x) = \begin{cases} \exp(ha_i(x))x_i - \frac{b_i(x)}{a_i(x)}(1 - \exp(ha_i(x))) & \text{if } i = j \text{ and } a_i(x) \neq 0 \\ x_i + hb_i(x) & \text{if } i = j \text{ and } a_i(x) = 0 \\ x_j & \text{if } i \neq j \end{cases} \tag{31}$$

2.4 Generalisation to Ornstein-Uhlenbeck equations

Whilst Chen et al only consider conditionally linear ODEs, we extend the approach to the systems of Orstein-Uhlenbeck type.

Definition 2. *An **Ornstein-Uhlenbeck equation** is an SDE of the form,*

$$dx(t) = (ax(t) + b)\, dt + c\, dW(t) \tag{32}$$

for constants $a, b \in \mathbb{R}, c > 0$, where $x : \mathbb{R} \to \mathbb{R}$ and W is the standard 1-d Wiener process.

In order to extend Stern's method, we need to be able to write down the exact solution of an SDE of this form, as we could for the linear ODE with constant coefficients. First we define the following family of equations,

Definition 3. *An SDE is called **linear in the narrow sense** if it has the form,*

$$dx(t) = (a(t)x(t) + b(t))\, dt + c(t)\, dW(t) \tag{33}$$

where $x, a, b : \mathbb{R} \to \mathbb{R}$, $c : \mathbb{R} \to \mathbb{R}^+$ and W is the standard 1-d Wiener process.

Kloeden and Platen give the exact solution to this more general case in [13]. The exact solution, starting from x_0, has the form

$$x(t) = \Phi_{t,0} \left(x_0 + \int_0^t b(s)\Phi_{s,0}^{-1}\, ds + \int_0^t c(s)\Phi_{s,0}^{-1}\, dW(s) \right), \tag{34}$$

where

$$\Phi_{t,0} = \exp\left(\int_0^t a(s)\, ds \right). \tag{35}$$

We can now substitute the functions $a(t), b(t), c(t)$ for the constants a, b, c and com-

pute the integrals, yielding

$$x(t) = x_0 e^{at} - \frac{b}{a}(1 - e^{at}) + ce^{at} \int_0^t e^{-as}\, dW(s), \tag{36}$$

for $a \neq 0$. This expression is normally distributed with mean $x_0 e^{at} - \frac{b}{a}(1 - e^{at})$. We can calculate the variance,

$$\mathrm{Var}(x(t)) = c^2 e^{2at} \mathbb{E}\left[\left(\int_0^t e^{-au}\, dW(u)\right)^2\right]. \tag{37}$$

We apply Itô's isometry to get a deterministic integral.

$$\mathrm{Var}(x) = c^2 e^{2at} \mathbb{E}\left[\int_0^t e^{-2au}\, du\right] = c^2 e^{2at} \mathbb{E}\left[\frac{-1}{2a}(e^{-2at} - 1)\right] \tag{38}$$

$$= c^2 e^{2at} \cdot \frac{-1}{2a}(e^{-2at} - 1) = \frac{c^2}{2a}(e^{2at} - 1) \tag{39}$$

For the case $a = 0$, the SDE reduces to,

$$dx(t) = b\, dt + c\, dW(t), \tag{40}$$

which means $x(t)$ is normally distributed with mean $x_0 + bt$ and variance $c^2 t$.
We can now extend Stern's approach to SDEs.

Definition 4. *A system of SDEs is said to be* ***conditionally linear*** *if every component of the system has the form,*

$$dx_i(t) = (a_i(x(t))x_i(t) + b_i(x(t)))\, dt + c_i(x(t))\, dW_i(t), \tag{41}$$

where $a_i(x(t)), b_i(x(t)), c_i(x(t))$ depend only on the components $x_j(t)$ for $j \neq i$.

We define $f_i(x(t))$ and $f^{(i)}(x(t))$ as before and we denote the exact h-time flow of $f^{(i)}$ by $\varphi_h^{(i)}$. The j-th component is therefore given by,

$$\varphi_{h,j}^{(i)}(x) = \begin{cases} \psi_i & \text{if } i = j \text{ and } a_i(x) \neq 0 \\ \nu_i & \text{if } i = j \text{ and } a_i(x) = 0\,, \\ x_j & \text{if } i \neq j \end{cases} \tag{42}$$

where

$$\psi_i \sim N(x_i e^{a_i(x)h} - \frac{b_i(x)}{a_i(x)}(1 - e^{a_i(x)h}), \frac{c_i^2(x)}{2a_i(x)}(e^{2a_i(x)h} - 1)) \tag{43}$$

$$\nu_i \sim N(x_i + b_i(x)h, c_i^2(x)h) \tag{44}$$

2.5 Coordinate transforms for conditional linearity

Whilst the VDP is already conditionally linear, the remaining models of interest are not conditionally linear. As such, we will perform a coordinate change which will result in a conditionally linear system where Stern's approach can be applied. All three of the remaining models of interest are of the (deterministic) form,

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} g(v(t)) + ku(t) \\ a(v(t))u(t) + b(v(t)) \end{pmatrix} dt, \tag{45}$$

where $g, a, b : \mathbb{R} \to \mathbb{R}$ are non-linear functions and $g \in C^2(\mathbb{R})$. In other words, only the second coordinate fulfills the criteria for being conditionally linear. León and Samson proposed a coordinate change for the stochastic FHN in order to transform the model into a stochastic damped oscillator [14]. We can apply a generalised version of this coordinate change to transform this general model into a conditionally linear form. Consider a new coordinate

$$y(t) := g(v(t)) + ku(t). \tag{46}$$

Applying the chain rule, we have that

$$\dot{y} = \dot{v}g'(v) + \dot{u}k = yg'(v) + (a(v)u + b(v))k \tag{47}$$

$$= yg'(v) + (a(v)\left(\frac{y - g(v)}{k}\right) + b(v))k = (g'(v) + a(v))y + (b(v)k - g(v)a(v)). \tag{48}$$

The means that the resulting system in (v, y) is

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ (g'(v(t)) + a(v(t)))y(t) + (b(v(t))k - g(v(t))a(v(t))) \end{pmatrix} dt, \tag{49}$$

which is conditionally linear. For a stochastic model of additive noise type, we can proceed similarly. Our model is of the form

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} g(v(t)) + ku(t) \\ a(v(t))u(t) + b(v(t)) \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t). \tag{50}$$

We perform the same coordinate change $y(t) = g(v(t)) + ku(t)$, but now proceed using Itô's lemma [10]. Itô's lemma states that given an SDE of the form,

$$dX(t) = \mu(t)dt + \Sigma(t)dW(t), \tag{51}$$

where $X, \mu : [0, \infty) \to \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{m \times n}$ positive semi-definite and W is the $m-$dimensional Wiener process, then given $y(t, X)$, with $y : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$, we have

$$dy(t, X) = \left(\frac{\partial y}{\partial t} + (\nabla_X y)^\top \mu(t) + \frac{1}{2} \mathbf{Tr}[\Sigma_t^\top (H_X y) \Sigma(t)] \right) dt + (\nabla_X y)^\top \Sigma_t dW(t), \tag{52}$$

where ∇ represents the gradient operator, H represents the Hessian operator and **Tr** represents the trace operator. Applying Itô's lemma to $x(t) = (v(t), u(t))$ and $y(t) := y(x(t)) = g(v(t)) + ku(t)$, we have that

$$dy(t) = \left\{ \begin{pmatrix} g'(v(t)) \\ k \end{pmatrix}^\top \begin{pmatrix} y(t) \\ a(v(t))u(t) + b(v(t)) \end{pmatrix} + \frac{1}{2} \mathbf{Tr} \left[\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}^\top \begin{pmatrix} g''(v(t)) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \right] \right\} dt \tag{53}$$

$$+ \begin{pmatrix} g'(v(t)) \\ k \end{pmatrix}^\top \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t).$$

This simplifies down to

$$dy(t) = \left\{ [g'(v(t)) + a(v(t))]y(t) + [kb(v(t)) - a(v(t))g(v(t)) + \frac{1}{2}\sigma_1^2 g''(v(t))] \right\} dt \tag{54}$$

$$+ g'(v(t))\sigma_1 \, dW_1(t) + k\sigma_2 \, dW_2(t)$$

where W_1, W_2 are independent 1-dimensional Wiener processes. Using their independence, we can combine these processes into a single term, leaving our expression as

$$dy(t) = (g'(v(t)) + a(v(t)))y(t) + (kb(v(t)) - a(v(t))g(v(t)) + \frac{1}{2}\sigma_1^2 g''(v(t))) \, dt \tag{55}$$
$$+ \sqrt{(g'(v(t))\sigma_1)^2 + (k\sigma_2)^2} \, dW_3(t),$$

where W_3 is a 1-dimensional Wiener process. In our new coordinate system (v, y), omitting t to simplify notation, the model has the conditionally linear form

$$d \begin{pmatrix} v \\ y \end{pmatrix} = \begin{pmatrix} y \\ (g'(v) + a(v))y + (kb(v) - a(v)g(v) + \frac{1}{2}\sigma_1^2 g''(v)) \end{pmatrix} dt \tag{56}$$
$$+ \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sqrt{(g'(v)\sigma_1)^2 + (k\sigma_2)^2} \end{pmatrix} dW(t),$$

as in Definition 4, where now $W = (W_1, W_3)$ is a 2-dimensional Wiener process and its components are no longer independent.

2.6 Applications to parameter estimation

Definition 5. *An SDE with additive noise of the form,*

$$dX(t) = f(X(t))\, dt + \Sigma\, dW(t), \tag{57}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d\times d}$, is called ***hypoelliptic*** *if $\Sigma\Sigma^\top$ is not of full rank yet f is such that the transition probability admits a smooth density [2].*

A component of $X(t)$ is called *smooth* if the noise term does not enter directly into its dynamics. A component is called *rough* if the noise term does enter directly. A common scenario of interest occurs when one component of the state $X(t)$ is smooth, therefore yielding a degenerate matrix Σ, but a rough component enters into the smooth component. In this scenario, the SDE is often hypoelliptic. We extend the definition of transition probability to a discrete case.

Definition 6. *The $k-$**step transition probability** of a numerical solution $\tilde{X}(t_i)$ of the SDE above is,*

$$\tilde{P}_{t_k}(\mathcal{A}, x) := \mathbb{P}(\tilde{X}(t_k) \in \mathcal{A} | \tilde{X}(0) = x) \tag{58}$$

*If $\kappa \in \mathbb{N}$ is the smallest natural such that the transition probability above has smooth density, we say the method is $\kappa-$**step hypoelliptic** [2].*

The 1-step transition probability plays an important role in the field of likelihood-based parameter estimation [2]. A particular interest lies in the situation where (58) corresponds to a multivariate normal distribution [2]. More explicitly, when $\tilde{X}(t_i)|\tilde{X}(t_{i-1})$ is normally distributed. Buckwar et al proved that the canonical splitting was 1-step hypoelliptic and therefore generated a well-defined parameter estimator [2]. Euler-Maruyama has been proven to be 2-step hypoelliptic i.e. the covariance matrix of $\tilde{X}(t_i)|\tilde{X}(t_{i-1})$ is degenerate [15]. Here we prove that under the stochastic Stern approach, defined in Section 2.4, $\tilde{X}(t_i)|\tilde{X}(t_{i-1})$ only follows a normal distribution under specific conditions [2].

Lemma 1. *Consider a 2-d conditionally linear SDE of the form,*

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \alpha_1(u(t))v(t) + \beta_1(u(t)) \\ \alpha_2(v(t))u(t) + \beta_2(v(t)) \end{pmatrix} dt + \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix} dW(t) \tag{59}$$

Let $\tilde{X}(t_i)$ be the approximate solution at time-points t_i under the stochastic Stern ap-

proach (see Section 2.4) composed with Lie-Trotter. Then $\tilde{X}(t_i)|\tilde{X}(t_{i-1})$ is normally distributed if $\alpha_1 \equiv 0$ and $\beta_1(u)$ is linear.

Proof. Given $\tilde{X}(t_{i-1}) = (v_{i-1}, u_{i-1})$, then, following the flows given in Section 2.4, we have that,

$$\tilde{X}(t_i) = \left(\varphi_h^{[1]} \circ \varphi_h^{[2]}\right) \begin{pmatrix} v_{i-1} \\ u_{i-1} \end{pmatrix} = \varphi_h^{[1]} \begin{pmatrix} v_{i-1} \\ \xi \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \tag{60}$$

where $\xi \sim N(\mu, \sigma^2)$ with

$$\mu = \begin{cases} u_{i-1}e^{\alpha_2(v_{i-1})h} - \frac{\beta_2(v_{i-1})}{\alpha_2(v_{i-1})}(1 - e^{\alpha_2(v_{i-1})h}) & \text{if } \alpha_2(v_{i-1}) \neq 0 \\ u_{i-1} + \beta_2(v_{i-1})h & \text{if } \alpha_2(v_{i-1}) = 0 \end{cases} \tag{61}$$

$$\sigma^2 = \begin{cases} \frac{\gamma_2^2}{2\alpha_2(v_{i-1})}(e^{2\alpha_2(v_{i-1})h} - 1) & \text{if } \alpha_2(v_{i-1}) \neq 0 \\ \gamma_2^2 h & \text{if } \alpha_2(v_{i-1}) = 0 \end{cases} \tag{62}$$

and $\eta|\xi, v_{i-1} \sim N(\pi(\xi), \kappa^2(\xi))$ where,

$$\pi(\xi) = \begin{cases} v_{i-1}e^{\alpha_1(\xi)h} - \frac{\beta_1(\xi)}{\alpha_1(\xi)}(1 - e^{\alpha_1(\xi)h}) & \text{if } \alpha_1(\xi) \neq 0 \\ v_{i-1} + \beta_1(\xi)h & \text{if } \alpha_1(\xi) = 0 \end{cases} \tag{63}$$

$$\kappa^2(\xi) = \begin{cases} \frac{\gamma_1^2}{2\alpha_1(\xi)}(e^{2\alpha_1(\xi)h} - 1) & \text{if } \alpha_1(\xi) \neq 0 \\ \gamma_1^2 h & \text{if } \alpha_1(\xi) = 0 \end{cases} \tag{64}$$

We show now that η is only normally distributed under specific conditions. Recall that the characteristic function of the normal distribution is given by,

$$\Phi_{X \sim N(\mu,\sigma^2)}(t) = e^{\mathbf{i}t\mu - \frac{1}{2}\sigma^2 t^2}. \tag{65}$$

Consider the characteristic function of η,

$$\Phi_\eta(t) = \mathbb{E}[e^{\mathbf{i}t\eta}] = \mathbb{E}[\mathbb{E}[e^{\mathbf{i}t\eta}|\xi]] = \mathbb{E}[e^{\mathbf{i}t\pi(\xi) - \frac{1}{2}\kappa^2(\xi)t^2}] \tag{66}$$

In the case that $\alpha_1(\xi) \neq 0$, $\pi(\xi)$ and $\kappa^2(\xi)$ are non-linear functions of ξ so Φ_η cannot have the form given in (65) thus η does not follow a normal distribution. If $\alpha_1(\xi) = 0$, then we have that $\kappa^2(\xi) = \gamma_1^2 h$,

$$\Phi_\eta(t) = e^{-\frac{1}{2}\gamma_1^2 h t^2}\mathbb{E}[e^{\mathbf{i}t\pi(\xi)}], \tag{67}$$

and therefore, η is normally distributed in the case $\pi(\xi)$ is linear i.e. $\beta_1(\xi)$ is a linear function. □

Furthermore consider the hypoelliptic case $\gamma_2 = 0$, where $u(t)$ is now a smooth component. In this case, we have,

$$\tilde{X}(t_i) = \left(\varphi_h^{[1]} \circ \varphi_h^{[2]}\right) \begin{pmatrix} v_{i-1} \\ u_{i-1} \end{pmatrix} = \begin{pmatrix} \eta \\ \mu \end{pmatrix}, \tag{68}$$

where $\eta|\mu \sim N(\pi(\mu), \kappa^2(\mu))$. In other words, the second component has variance 0 even though, in the true solution, noise would propagate through the rough component $v(t)$ into $u(t)$. Thus, composing the flows in this order yields a method that is not 1-step hypoelliptic.

3 Classical integrators

The most common class of integrators for ODEs are the so-called **finite difference methods**. These methods work by approximating derivatives with finite differences, thereby discretising the problem. The convergence of these methods is well studied, but the method often have stability and convergence limitations when it comes to **stiff** systems. For such systems, explicit finite difference methods often require very small time-steps in order to maintain stability or to preserve the dynamics of the systems in question, whilst implicit schemes are computational intractable in the context of parameter estimation [2]. Due to their place as the standard in numerical schemes of ODEs, we will use the **forward Euler** and **fourth order Runge-Kutta** methods as the comparison point for the splitting integrators on the deterministic models [6]. We will use the **Euler-Maruyama** scheme and its tamed version for the stochastic variants of the models [13].
Consider again the initial value problem presented in Section 2,

$$dx(t) = f(x(t))\, dt; \quad x(0) = x_0; \quad t \in [0, T], \tag{69}$$

we define the following numerical schemes.

3.1 Forward Euler

One such explicit finite difference method is the so-called **forward Euler** (FE) scheme. The method is given by the recursive formula

$$x_{n+1} = x_n + hf(x_n, t_n), \tag{70}$$

$$x_0 = x(0). \tag{71}$$

When $f \in C^1$ and $f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial t}$ are bounded in a neighbourhood of the solution, then this method has order of convergence 1 [7].

3.2 Fourth order Runge-Kutta

The **"classical" Runge-Kutta method** (RK4) is the most commonly used method from the Runge-Kutta class of methods. It is an explicit method given by the formula

$$x_{n+1} = x_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \tag{72}$$

$$x_0 = x(0), \tag{73}$$

where

$$k_1 = f(x_n, t_n), \tag{74}$$

$$k_2 = f(x_n + h\frac{k_1}{2}, t_n + \frac{h}{2}) \tag{75}$$

$$k_3 = f(x_n + h\frac{k_2}{2}, t_n + \frac{h}{2}) \tag{76}$$

$$k_4 = f(x_n + hk_3, t_n + h). \tag{77}$$

If $f \in C^4$ and all its partial derivatives up to order 4 exist and are continuous, then this method has order of convergence 4 [7].

3.3 Euler-Maruyama

Consider instead a SDE of the form,

$$dX(t) = a(X(t), t)\ dt\ + b(X(t), t)\ dW(t), \tag{78}$$

where $X : [0, \infty) \to \mathbb{R}^d$ is an Itô process, $a : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^d$, $b : \mathbb{R}^d_{\geq 0} \times [0, \infty) \to \mathbb{R}^d$ and W is the standard $d-$dimensional Wiener process. Consider also an initial condition $X(0) = x_0$. We discretise time identically as before, such that $t_n = nh$ for $h > 0$ and $n \in \mathbb{N}$. We then recursively define our approximation as follows,

$$X_{n+1} = X_n + a(X_n, t_n)h + b(X_n, t_n)\xi_n, \tag{79}$$

$$X_0 = x_0, \tag{80}$$

where the $\xi_n = (\xi_{n_1}, \xi_{n_2}, ..., \xi_{n_d}) \in \mathbb{R}^d$ are vectors of independent, identically distributed random variables ξ_{n_i} taken from the normal distribution with mean 0 and variance h. If a, b are measurable, Lipschitz and satisfy a linear growth bound, then the Euler-Maruyama scheme is weakly convergent with order 1 and strongly convergent with order 0.5 [13].

3.4 Tamed Euler-Maruyama

Hutzenthaler et al proved that the Euler-Maruyama scheme does not converge for super-linearly growing coefficients [8]. This includes some of the models discussed in the project. For this reason, we also consider the **tamed Euler-Maruyama scheme**, given by

$$X_{n+1} = X_n + \frac{a(X_n, t_n)h}{1 + h||a(X_n, t_n)||} + b(X_n, t_n)\xi_n, \tag{81}$$

$$X_0 = x_0. \tag{82}$$

where the ξ_n are as in the Euler-Maruyama scheme above. If a, b are measurable, globally one-side Lipschitz and satisfy a polynomial growth bound, then the tamed Euler-Maruyama scheme converges strongly with order 0.5 [9].

4 Application to neuronal models

4.1 Membrane potential and spiking

In spiking neuronal models, one component represents the membrane potential of the neuron. When the membrane potential reaches a given threshold, a 'spike' or action potential is released, which is the release of neural information. After the action potential is released, the neuron returns to its resting potential, hence the characteristic spiking or oscillatory behaviour expected in neuronal models.

4.2 The Van der Pol model

The VDP oscillator is given by the 2-dimensional system of first order ODEs,

$$d\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \epsilon(1 - x_1^2(t))x_2(t) - x_1(t) \end{pmatrix} dt, \tag{83}$$

where $\epsilon \geq 0$ is a scalar parameter representing the strength of the damping in the oscillator. The VDP oscillator has been used to model a range of phenomena in biology and physics, most relevantly action potentials in the brain. The model was extended to the FHN model of neurons. We consider the component $x_1(t)$ to represent the action potential of a neuron at time t.

4.2.1 The canonical splitting

We can apply the canonical splitting to the VDP model. The splitting is not unique, but one example would be,

$$d\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \left[\begin{pmatrix} 0 & 1 \\ -1 & \epsilon \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0 \\ -\epsilon x_1(t)^2 x_2(t) \end{pmatrix} \right] dt. \tag{84}$$

We therefore have

$$A = \begin{pmatrix} 0 & 1 \\ -1 & \epsilon \end{pmatrix}; \quad N(x_1, x_2) = \begin{pmatrix} 0 \\ -\epsilon x_1^2(t) x_2(t) \end{pmatrix}. \tag{85}$$

This therefore yields sub-problems,

$$dx^{[1]}(t) = \begin{pmatrix} 0 & 1 \\ -1 & \epsilon \end{pmatrix} x^{[1]}(t)\, dt = Ax^{[1]}(t)\, dt \tag{86}$$

$$dx^{[2]}(t) = \begin{pmatrix} 0 \\ -\epsilon x_1^{[2]}(t)^2 x_2^{[2]}(t) \end{pmatrix} dt \tag{87}$$

where $x^{[1]}(t) = (x_1^{[1]}(t), x_2^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (x_1^{[2]}(t), x_2^{[2]}(t)) \in \mathbb{R}^2$. This is a suitable splitting as the non-linear sub-equation is separable and therefore exactly integrable. Given an initial condition $x(0) = y = (y_1, y_2) \in \mathbb{R}^2$, we can write down the exact $t-$time flow of each of the sub-problems, starting from this initial condition

$$\begin{cases} \varphi_t^{[1]}(y) & = ye^{At} \\ \varphi_t^{[2]}(y) & = \begin{pmatrix} y_1 \\ y_2 e^{-\epsilon y_1^2 t} \end{pmatrix}. \end{cases} \tag{88}$$

For some step size $h > 0$, these flows can be recomposed using the Lie-Trotter and Strang compositions as in Section 3. Labelling $\tilde{x}^{LT}(t_{i-1}) = x_{i-1} = (x_{i-1}^{[1]}, x_{i-1}^{[2]})$, we obtain explicit numerical schemes,

$$\tilde{x}_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} x_{i-1}^{[1]} \\ x_{i-1}^{[2]} e^{-\epsilon (x_{i-1}^{[1]})^2 h} \end{pmatrix}, \tag{89}$$

$$\tilde{x}_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} x_{i-1}^{[1]} \\ x_{i-1}^{[2]} e^{-\epsilon (x_{i-1}^{[1]})^2 \frac{h}{2}} \end{pmatrix} \right). \tag{90}$$

4.2.2 Stern's approach

Notice that the VDP system is conditionally linear as defined in Definition 1.

$$d\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} 0 \cdot x_1(t) + x_2(t) \\ \epsilon(1 - x_1^2(t)) \cdot x_2(t) + (-x_1(t)) \end{pmatrix}. \tag{91}$$

We can write down the functions a_i, b_i, from the definition of conditionally linear, for this system

$$a_1(x) = 0; \quad b_1(x) = x_2, \tag{92}$$

$$a_2(x) = \epsilon(1 - x_1^2); \quad b_2(x) = -x_1. \tag{93}$$

Again, considering an initial condition $x(0) = y = (y_1, y_2) \in \mathbb{R}^2$, we can write down the exact $t-$time flow of each of the sub-problems, starting from this initial condition, using the expressions from Section 3, being careful to apply the special case as $a_1(x) = 0$. This gives us $t-$time flows from y

$$\begin{cases} \varphi_t^{[1]}(y) &= \begin{pmatrix} y_1 + ty_2 \\ y_2 \end{pmatrix}, \\ \varphi_t^{[2]}(y) &= \begin{pmatrix} y_1 \\ \exp(t\epsilon(1 - y_1^2))y_2 - \frac{\exp(t\epsilon(1-y_1^2))-1}{\epsilon(1-y_1^2)} y_1 \end{pmatrix}. \end{cases} \tag{94}$$

Again, these flows can be recomposed with the Lie-Trotter and Strang compositions to generating explicit numerical schemes

$$\tilde{x}_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) \tag{95}$$

$$= \begin{pmatrix} x_{i-1}^{[1]} + h(\exp(h\epsilon(1 - (x_{i-1}^{[1]})^2))x_{i-1}^{[2]} - \frac{\exp(h\epsilon(1-(x_{i-1}^{[1]})^2))-1}{\epsilon(1-(x_{i-1}^{[1]})^2)} x_{i-1}^{[1]}) \\ \exp(h\epsilon(1 - (x_{i-1}^{[1]})^2))x_{i-1}^{[2]} - \frac{\exp(h\epsilon(1-(x_{i-1}^{[1]})^2))-1}{\epsilon(1-(x_{i-1}^{[1]})^2)} x_{i-1}^{[1]} \end{pmatrix}, \tag{96}$$

$$\tilde{x}_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \begin{pmatrix} \varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h/2,2}^{[2]}(x_{i-1})) \\ \varphi_{h/2,2}^{[2]}(\varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h,2}^{[2]}(x_{i-1})), (\varphi_{h/2,2}^{[2]})(x_{i-1})) \end{pmatrix}. \tag{97}$$

4.2.3 Van der Pol with additive noise

In the stochastic setting, we include an additional term which converts our ODE into an SDE,

$$d\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \epsilon(1-x_1(t)^2)x_2(t) - x_1(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t), \tag{98}$$

where $\sigma_1, \sigma_2 \geq 0$ and W is the standard 2-d Wiener process. We can use a similar splitting to the deterministic case, including the additive noise in the linear sub-equation,

$$dx^{[1]}(t) = \begin{pmatrix} 0 & 1 \\ -1 & \epsilon \end{pmatrix} x^{[1]}(t)\, dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t) \tag{99}$$

$$dx^{[2]}(t) = \begin{pmatrix} 0 \\ -\epsilon x_1^{[2]}(t)^2 x_2^{[2]}(t) \end{pmatrix} dt \tag{100}$$

The $t-$time flow of the linear sub-equation is now given by the stochastic harmonic oscillator, following the analysis in Section 2.2. The $t-$time flow of the second sub-equation remains unchanged. Therefore, we have

$$\varphi_t^{[1]}(y) = ye^{At} + \xi_t, \tag{101}$$

$$\varphi_t^{[2]}(y) = \begin{pmatrix} y_1 \\ y_2 e^{-\epsilon y_1^2 t} \end{pmatrix}, \tag{102}$$

with

$$A = \begin{pmatrix} 0 & 1 \\ -1 & \epsilon \end{pmatrix}, \tag{103}$$

and $\xi_t \sim \mathcal{N}(0, C(t))$. The expression for the covariance matrix $C(t)$ is given in Appendix C. Using these flows and a time-step $h > 0$, we can define the following numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} x_{i-1}^{[1]} \\ x_{i-1}^{[2]} e^{-\epsilon (x_{i-1}^{[1]})^2 h} \end{pmatrix} + \xi_{i-1}, \tag{104}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} x_{i-1}^{[1]} \\ x_{i-1}^{[2]} e^{-\epsilon (x_{i-1}^{[1]})^2 \frac{h}{2}} \end{pmatrix} + \xi_{i-1} \right), \tag{105}$$

where $\xi_{i-1} \sim \mathcal{N}(0, C(h))$. The Van der Pol oscillator with additive noise is a conditionally linear SDE, as in Definition 4, with functions,

$$a_1(x) = 0;\ b_1(x) = x_2 \tag{106}$$
$$a_2(x) = \epsilon(1 - x_1^2);\ b_2(x) = -x_1 \tag{107}$$
$$c_2(x) = \sigma_1;\ c_2(x) = \sigma_2 \tag{108}$$

We can substitute these functions into the $t-$time flows given in Section 2.4 to obtain flows,

$$\begin{cases} \varphi_t^{[1]}(y) &= \begin{pmatrix} \xi_t \\ y_2 \end{pmatrix}, \\ \\ \varphi_t^{[2]}(y) &= \begin{pmatrix} y_1 \\ \nu_t \end{pmatrix}, \end{cases} \tag{109}$$

where $\xi_t \sim N(y_1 + y_2 t, \sigma_1^2 t)$ and
$\nu_t \sim N(y_2 \exp(\epsilon(1-y_1^2)t) + \frac{y_1}{\epsilon(1-y_1^2)}(1-\exp(\epsilon(1-y_1)^2 t)), \frac{\sigma_2^2}{2\epsilon(1-y_1)^2}(\exp(2\epsilon(1-y_1^2)t)-1))$.

4.3 The FitzHugh-Nagumo model

The FHN model is given by the following 2-d system of first-order ODEs

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{\epsilon}(v(t) - v(t)^3 - u(t)) \\ \gamma v(t) - u(t) + \beta \end{pmatrix} dt. \tag{110}$$

The first component, $v(t)$, models the membrane potential of the neuron at time t whilst the second component $u(t)$ is a recovery variable modelling ion channel kinetics [2]. The parameter $\epsilon > 0$ represents the timescale separation of the components. When the membrane potential, $v(t)$, crosses a certain threshold, a spike is generated. $\beta \geq 0$ controls the position of the spike whilst $\gamma > 0$ controls the duration [2].

4.3.1 The canonical splitting

We can apply the canonical splitting to the FHN model as in Buckwar et al [2].

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix} \begin{pmatrix} v(t) \\ u(t) \end{pmatrix} + \begin{pmatrix} \frac{1}{\epsilon}(v(t) - v^3(t)) \\ \beta \end{pmatrix} dt. \tag{111}$$

We therefore have

$$A = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix}; \quad N(v,u) = \begin{pmatrix} \frac{1}{\epsilon}(v(t) - v^3(t)) \\ \beta \end{pmatrix}. \tag{112}$$

This yields sub-problems

$$dx^{[1]}(t) = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix} x^{[1]}(t)\, dt \tag{113}$$

$$dx^{[2]}(t) = \begin{pmatrix} \frac{1}{\epsilon}(v^{[2]}(t) - {v^{[2]}}^3(t)) \\ \beta \end{pmatrix} dt \tag{114}$$

where $x^{[1]}(t) = (v^{[1]}(t), u^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), u^{[2]}(t)) \in \mathbb{R}^2$. As in the VDP model, the non-linear sub-equation is separable and therefore integrable. Given an initial condition $x(0) = x_0 = (v_0, u_0) \in \mathbb{R}^2$, we can write down the exact $t-$time flow of each of the sub-problems, starting from x_0.

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} \frac{v_0}{\sqrt{e^{-2t/\epsilon} + v_0^2(1-e^{-2t/\epsilon})}} \\ \beta t + u_0 \end{pmatrix}, \end{cases} \tag{115}$$

where,

$$A = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix}. \tag{116}$$

Using the flows and a time-step $h > 0$, we apply the Lie-Trotter and Strang compositions to give the following explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} \frac{v_{i-1}}{\sqrt{e^{-2h/\epsilon} + v_{i-1}^2(1-e^{-2h/\epsilon})}} \\ \beta h + u_{i-1} \end{pmatrix}, \tag{117}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} \frac{v_{i-1}}{\sqrt{e^{-h/\epsilon} + v_{i-1}^2(1-e^{-h/\epsilon})}} \\ \frac{\beta h}{2} + u_{i-1} \end{pmatrix} \right), \tag{118}$$

where $x_{i-1} = (v_{i-1}, u_{i-1})$.

4.3.2 Coordinate change and Stern's approach

The FHN model is not conditionally linear due to the cubic term in the first component. However, it is in the general form discussed in Section 2.5 with functions

and constant given by,

$$g(v) = \frac{1}{\epsilon}(v(t) - v^3(t)), \qquad k = -\frac{1}{\epsilon}, \qquad a(v) = -1, \qquad b(v) = \gamma v(t) + \beta. \quad (119)$$

Applying the coordinate change proposed in Section 2.5, we get the resulting conditionally linear system in (v, y) where $y(t)$ is our new coordinate $y(t) = g(v(t)) + ku(t)$:.

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ (\frac{1}{\epsilon} + 1 - \frac{3}{\epsilon}v^2(t))y - \frac{1}{\epsilon}(\gamma + 1)v(t) + \frac{1}{\epsilon}v^3(t) - \frac{\beta}{\epsilon} \end{pmatrix} dt \quad (120)$$

Whilst we perform this linearisation in order to apply Stern's method, we can also apply the canonical splitting or classical integrators to the resulting system as well. This helps isolate any error in the numerical solution that may arise from the coordinate change. First, we apply the canonical splitting to this linearised system. A suitable splitting would be

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{\epsilon}(\gamma + 1) & (\frac{1}{\epsilon} + 1) \end{pmatrix} \begin{pmatrix} v(t) \\ y(t) \end{pmatrix} \quad (121)$$

$$+ \begin{pmatrix} 0 \\ -\frac{3}{\epsilon}v^2(t)y(t) + \frac{1}{\epsilon}v^3(t) - \frac{\beta}{\epsilon} \end{pmatrix} dt. \quad (122)$$

Again, the non-linear part yields a separable sub-equation. Proceeding as before, from an initial condition $x(0) = x_0 = (v_0, y_0) \in \mathbb{R}^2$, we get $t-$time flows

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} v_0 \\ (y_0 + \frac{\beta}{3v_0^2} + v_0)e^{-\frac{3}{\epsilon}v_0^2 t} + \frac{\beta}{3v_0^2} - \frac{v_0}{3} \end{pmatrix}, \end{cases} \quad (123)$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{1}{\epsilon}(\gamma + 1) & (\frac{1}{\epsilon} + 1) \end{pmatrix}. \quad (124)$$

We recompose these flows with the Lie-Trotter and Strang compositions to obtain explicit numerical schemes

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} v_{i-1} \\ (y_{i-1} + \frac{\beta}{3v_{i-1}^2} + v_0)e^{-\frac{3}{\epsilon}v_{i-1}^2 h} + \frac{\beta}{3v_{i-1}^2} - \frac{v_{i-1}}{3} \end{pmatrix}, \quad (125)$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} v_{i-1} \\ (y_{i-1} + \frac{\beta}{3v_{i-1}^2} + v_0)e^{-\frac{3}{\epsilon}v_{i-1}^2 \frac{h}{2}} + \frac{\beta}{3v_{i-1}^2} - \frac{v_{i-1}}{3} \end{pmatrix} \right), \quad (126)$$

where $x_{i-1} = (v_{i-1}, y_{i-1})$.
We can now use the Stern approach to obtain numerical schemes for this conditionally linear version of the FHN. The linearised FHN has functions

$$a_1(y) = 0;\ b_1(y) = y \quad (127)$$

$$a_2(v) = \frac{1}{\epsilon} + 1 - \frac{3}{\epsilon}v^2;\ b_2(v) = -\frac{1}{\epsilon}(\gamma + 1)v + \frac{1}{\epsilon}v^3 - \frac{\beta}{\epsilon} \quad (128)$$

as in Definition 1. The formulas in Section 2.3 give us $t-$time flows from an initial condition $x_0 = (v_0, y_0)$,

$$\begin{cases} \varphi_t^{[1]}(x_0) &= \begin{pmatrix} v_0 + ty_0 \\ y_0 \end{pmatrix}, \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} v_0 \\ \exp(t(\frac{1}{\epsilon} + 1 - \frac{3}{\epsilon}v_0^2))y_0 + \frac{\exp(t(\frac{1}{\epsilon}+1-\frac{3}{\epsilon}v_0^2))-1}{(\frac{1}{\epsilon}+1-\frac{3}{\epsilon}v_0^2)}(-\frac{1}{\epsilon}(\gamma+1)v_0 + \frac{1}{\epsilon}v_0^3 - \frac{\beta}{\epsilon}) \end{pmatrix}. \end{cases} \quad (129)$$

Again, these flows can be recomposed using the Lie-Trotter and Strang compositions and a time-step $h > 0$ to give explicit numerical schemes

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = \begin{pmatrix} v_{i-1} + h\varphi_{h,2}^{[2]}(x_{i-1}) \\ \varphi_{h,2}^{[2]}(x_{i-1}) \end{pmatrix}, \quad (130)$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \begin{pmatrix} \varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h/2,2}^{[2]}(x_{i-1})) \\ \varphi_{h/2,2}^{[2]}(\varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h,2}^{[2]}(x_{i-1})), (\varphi_{h/2,2}^{[2]})(x_{i-1})) \end{pmatrix}. \quad (131)$$

4.3.3 FitzHugh-Nagumo with additive noise

In the stochastic setting, we again include an additive noise term which yields the SDE

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{\epsilon}(v(t) - v(t)^3 - u(t)) \\ \gamma v(t) - u(t) + \beta \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t), \tag{132}$$

where $\sigma_1, \sigma_2 \geq 0$ and W is the standard 2-d Wiener process. Repeating the analysis from Buckwar et al [2], we use a similar splitting to the deterministic case but again including the additive noise in the linear sub-equation yielding sub-problems

$$dx^{[1]}(t) = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix} x^{[1]}(t)\, dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t), \tag{133}$$

$$dx^{[2]}(t) = \begin{pmatrix} \frac{1}{\epsilon}(v^{[2]}(t) - {v^{[2]}}^3(t)) \\ \beta \end{pmatrix} dt \tag{134}$$

where $x^{[1]}(t) = (v^{[1]}(t), u^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), u^{[2]}(t)) \in \mathbb{R}^2$. Following the analysis in Section 2.2, we get flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At} + \xi_t \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} \frac{v_0}{\sqrt{e^{-2t/\epsilon} + v_0^2(1 - e^{-2t/\epsilon})}} \\ \beta t + u_0 \end{pmatrix}. \end{cases} \tag{135}$$

where $\xi_t \sim \mathcal{N}(0, C(t))$, $x_0 = (v_0, u_0)$ and

$$A = \begin{pmatrix} 0 & -\frac{1}{\epsilon} \\ \gamma & -1 \end{pmatrix}. \tag{136}$$

The expression for the covariance matrix $C(t)$ is given in Appendix C. Using these flows and a time-step $h > 0$, we can define the following numerical schemes

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} \frac{v_{i-1}}{\sqrt{e^{-2h/\epsilon} + v_{i-1}^2(1 - e^{-2h/\epsilon})}} \\ \beta h + u_{i-1} \end{pmatrix} + \xi_{i-1}, \tag{137}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} \frac{v_{i-1}}{\sqrt{e^{-h/\epsilon} + v_{i-1}^2(1 - e^{-h/\epsilon})}} \\ \beta \frac{h}{2} + u_{i-1} \end{pmatrix} + \xi_{i-1} \right), \tag{138}$$

where $\xi_{i-1} \sim \mathcal{N}(0, C(h))$. Using the stochastic version of the coordinate change proposed in Section 2.5 yields the SDE

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ \frac{1}{\epsilon}(y(t)(1-\epsilon-3v^2(t)) - v(t)(\gamma - 1 + 3\sigma_1^2) - v(t)^3 - \beta) \end{pmatrix} dt \quad (139)$$
$$+ \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sqrt{\frac{\sigma_1^2}{\epsilon^2}(1-3v^2(t))^2 + \frac{\sigma_2^2}{\epsilon^2}} \end{pmatrix} dW(t).$$

Following Definition 4, the conditionally linearised FHN with additive noise is a conditionally linear SDE in (v, y) with functions,

$$a_1(y) = 0;\ a_2(v) = \frac{1}{\epsilon}(1 - \epsilon - 3v^2) \quad (140)$$

$$b_1(y) = y;\ b_2(v) = -\frac{1}{\epsilon}((\gamma - 1 + 3\sigma_1^2)v + v^3 + \beta) \quad (141)$$

$$c_1(y) = \sigma_1;\ c_2(v) = \sqrt{\frac{\sigma_1^2}{\epsilon^2}(1-3v^2)^2 + \frac{\sigma_2^2}{\epsilon^2}} \quad (142)$$

We can substitute these functions into the $t-$time flows given in Section 2.4, starting at $x_0 = (v_0, y_0)$ to obtain flows

$$\begin{cases} \varphi_t^{[1]}(x_0) &= \begin{pmatrix} \xi_t \\ y_0 \end{pmatrix}, \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} v_0 \\ \nu_t \end{pmatrix}, \end{cases} \quad (143)$$

where $\xi_t \sim N(v_0 + y_0 t, \sigma_1^2 t)$ and
$\nu_t \sim N(y_0 \exp(\frac{1}{\epsilon}(1-\epsilon-3v_0^2)t) + \frac{v_0(\gamma-1+3\sigma_1^2)+v_0^3+\beta}{1-\epsilon-3v_0^2}(1-\exp(\frac{1}{\epsilon}(1-\epsilon-3v_0^2)t)), \frac{\sigma_1^2(1-3v_0^2)^2+\sigma_2^2}{2\epsilon(1-\epsilon-3v_0^2)}(\exp(\frac{2t}{\epsilon}(1-\epsilon-3v_0^2)-1)$.

4.4 The Izhikevich model

The Izhikevich (IZ) model is another 2-d spiking neuron model but with an auxiliary after-spike reset [12]. The model is given by the 2-d system of ODEs

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} av^2(t) + bv(t) + c + du(t) + fI \\ \alpha(\beta v(t) - u(t)) \end{pmatrix} dt \quad (144)$$

with a reset condition,

$$\text{If } v \geq 30$$
$$v \to \theta \tag{145}$$
$$u \to u + \delta. \tag{146}$$

As in the FHN, $v(t)$ represents the membrane potential and $u(t)$ is a recovery variable modelling ion channel kinetics [12]. Here a, b, c, d, f are considered as parameters. Izhikevich sets these parameters to constant values that keep the membrane potential and spike timescale in the correct order of magnitude but we consider them parameters for greater clarity in the mathematical analysis [11, 12]. α controls the timescale of oscillations in $u(t)$ whilst β controls the coupling between $v(t)$ and $u(t)$ i.e. β controls the sensitivity of u to fluctuations in v. θ represents the after-spike resting potential of the neuron, δ models the increment in the after-spike reset in u and I represents an input current [12]. We can apply the canonical splitting to this model following the analysis in Section 2.2. We can also perform the coordinate change proposed in Section 2.5 to obtain a conditionally linearised version of the IZ model. As this analysis is somewhat similar to that performed in Sections 4.2 & 4.3, and for the sake of space, it is included in the Appendix A.1.

4.5 The adaptive exponential integrate-and-fire model

The AdEx model is another 2-d spiking neuron model with an auxiliary after-spike reset [1]. The model is given by the 2-d system of ODEs

$$d\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} a\exp(\frac{v(t)}{c}) + dv(t) + fw(t) + gI \\ jv(t) + lw(t) + k \end{pmatrix} dt \tag{147}$$

with a reset condition,

$$\text{If } v \geq 18$$
$$v \to \theta \tag{148}$$
$$w \to w + \delta. \tag{149}$$

Again $v(t)$ represents the membrane potential whilst $w(t)$ represents the adaptation current [1]. In this project, we have rearranged the model into its most rudimentary mathematical form to support clarity in the mathematical operations and to facilitate comparison to the other models. Due to this, the new parameters are composites of the original biologically relevant parameters. One can see Brette and Gerstner's original paper for a full description of the original parameters and their

biological relevance [1]. The analysis of this model is included in Appendix A.2.

5 Numerical experiments

5.1 Deterministic models

5.1.1 The "stiff" Van der Pol oscillator

When $\epsilon \gg 1$ in the VDP, it is known as "stiff" [4]. Typically of explicit schemes a very small time-step is needed to preserve the oscillatory dynamics [3]. Here we investigate the performance of the numerical schemes with $\epsilon = 50$. Figure 4 shows the x_1 trajectories under the various numerical schemes against a reference solution calculated using RK4 with $h = 10^{-5}$. We focus on the first component as it models the membrane potential and has characteristic spiking behaviour. Figure 4 shows

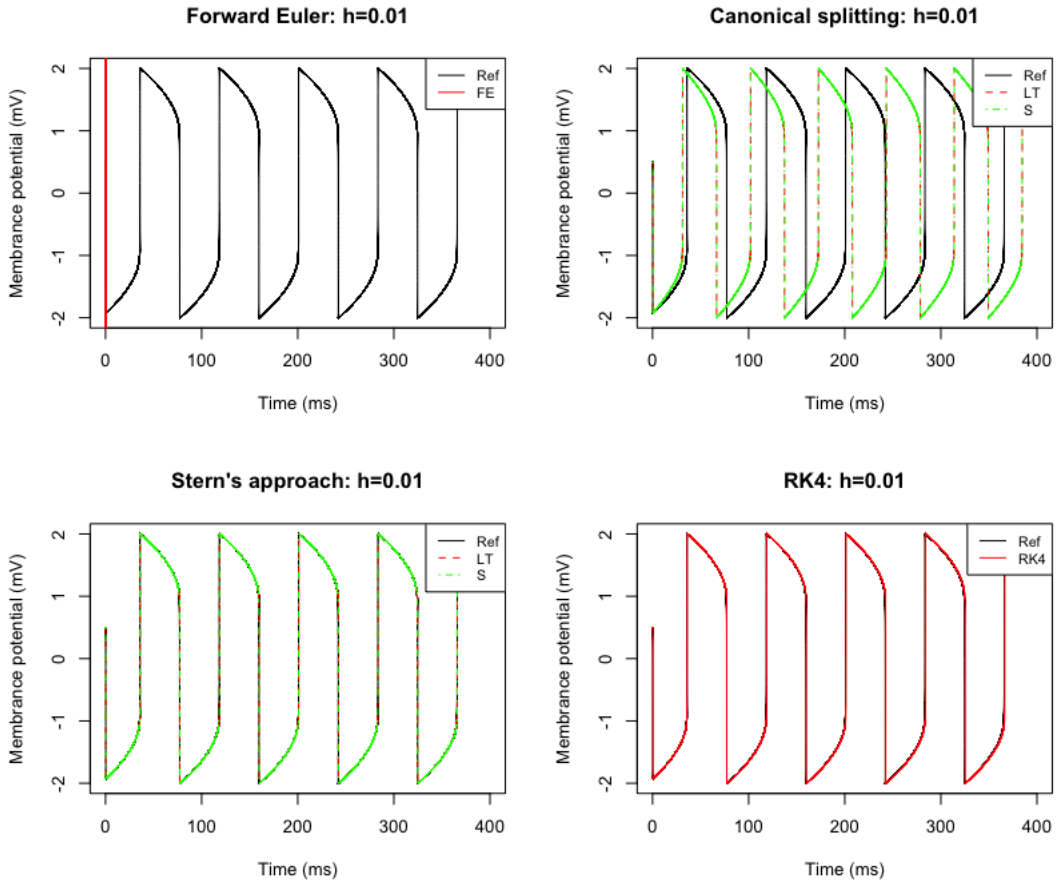

Figure 4: Stiff VDP: Plot of the x_1 component calculated with numerical schemes against reference solution with $\epsilon = 50$, time-step $h = 0.01$ and initial condition $(x_0, y_0) = (0.5, 0)$

that for this relatively large time step both RK4, defined in Section 3.2, and Stern's approach, for both Lie-Trotter (LT) and Strang (S) compositions, exhibit the correct

oscillatory dynamics for the x_1 component. The canonical splitting shows spiking behaviour but out of phase with the reference solution. Forward Euler, defined in Section 3.1, fails to converge.
To investigate the behaviour for smaller time-steps, we compute the **quadratic loss** between a numerical solution and the reference solution.

Definition 7. *Given the reference solution $X(t)$ and an approximate solution $\tilde{X}(t_i)$ at discrete time points t_i for $i = 1, ..., n$, we define the* **quadratic loss (QL) to** *$X(t)$ as,*

$$QL(\tilde{X}) := \frac{1}{n}\sqrt{\Sigma_{i=1}^{n}(X(t_i) - \tilde{X}(t_i))^2} \tag{150}$$

The quadratic loss is $\frac{1}{n}L_2$ norm of the difference between the solutions as a vector. Figure 5 shows this quadratic loss for the various numerical schemes at a range of decreasing time-steps for the x_1 component. It shows that for smaller time-steps, the

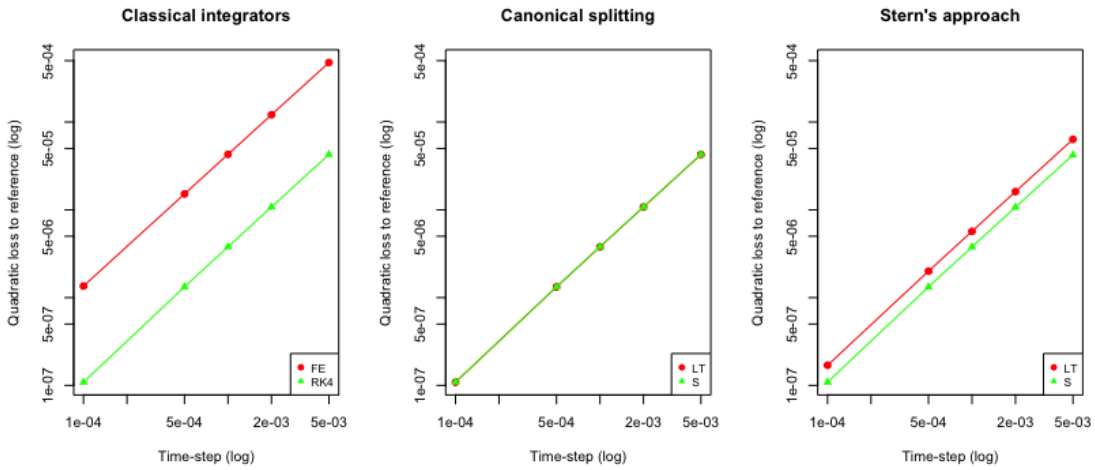

Figure 5: Stiff VDP: Quadratic loss to reference solution for various time-steps

Lie-Trotter and Strang compositions have the same performance for the canonical splitting but there is a disparity between the two composition methods in Stern's approach. RK4, the canonical splitting methods and Stern's approach with Strang all have similar QL values. Again, the Forward Euler performs poorly, having the highest QL at each time-step.

5.1.2 The FitzHugh-Nagumo model

Under suitable parameter conditions, the FHN has a stable limit cycle [16]. We will investigate the performance of the numerical schemes in preserving this limit cycle mirroring some of the analysis Chen et al performed for VDP [3]. Figure 6 shows

the (v, u) plane under the different numerical schemes for increasing time-steps. For each of the splitting methods, we use the Strang composition.

Remark 3. *In the following plots, Canonical Linearised Splitting (CLS) will refer to the conditionally linearised version of a model being solved with the canonical splitting, as applied to FHN in Section 4.3.2.*

Clearly, RK4 and the canonical splitting for the original system best preserve the limit cycles for the larger time-steps. The two schemes that use the linearised system, Stern's approach and the CLS, preserve some of the structure, but not for the larger time-steps. Forward Euler again performs the worst, only preserving oscillatory dynamics for small time steps.
Figure 7 show that RK4 and the splitting approaches have similar QL for a range of small time steps. We examine these well-performing methods for larger time-steps to see when a disparity occurs. Figure 8 shows that, of these well-performing methods, when composed with the Strang composition, the canonical splitting has the lowest QL.

5.1.3 The Izhikevich model

The IZ model varies drastically from the two "oscillatory models" studied earlier. It has a spike-reset condition that can create issues with traditional numerical integration [19]. There are also no existing results of splitting integrators for models of this type. Figure 9 shows IZ trajectories computed with a relatively large time step of $h = 5 \times 10^{-2}$. Clearly, we can see that the classical integrators perform the best, preserving the phase of the spikes, but not perfect amplitude. The Strang approach with the canonical splitting has the correct frequency but is out of phase. The Lie-Trotter approach exhibits "chattering" behaviour (a feature of the model not associated with this parameter setting) [12]. For the conditionally linearised version of the system, the Lie-Trotter composition, for both the Stern and canonical splitting approaches, exploded in finite time and hence is not displayed - this can occur for weak numerical methods due to bifurcations in the model [11]. To investigate differences between the splitting approaches, Figure 10 shows the same schemes but for a time step of $h = 5 \times 10^{-4}$. In this figure, we can see that the classical integrators and canonical splitting both preserve the frequency and amplitude of the spikes. For the linearised system, only the Strang composition show the correct behaviour. This suggests that some numerical approaches may be more robust to the coordinate change in the presence of boundary conditions. Figure 11 shows the trajectories grouped according to which method is used to compose solutions rather than which splitting. It shows that only the Strang approach yields the same, and correct, trajectory for both the original and the linearised system.

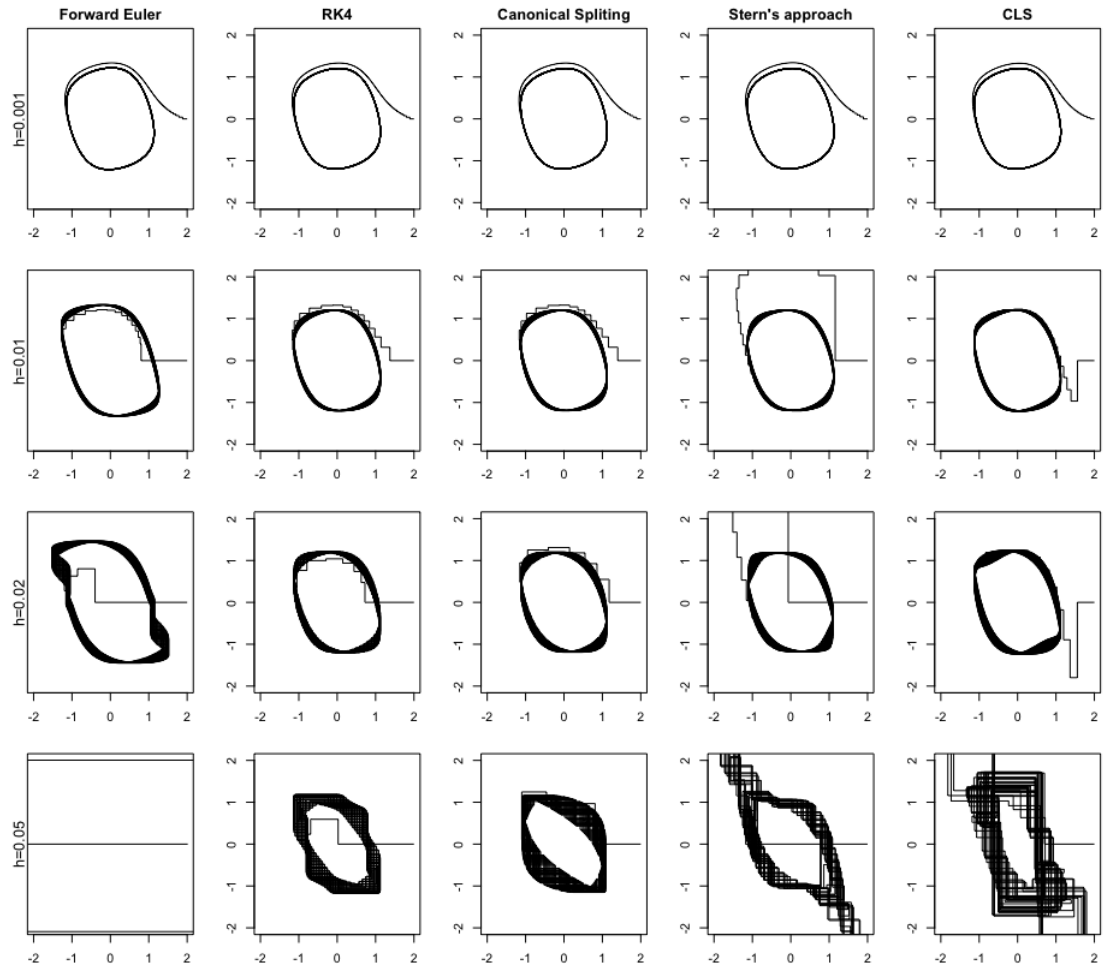

Figure 6: FHN limit cycles in (v, u) with $\epsilon = 0.05; \beta = 0.1; \gamma = 20; (v_0, u_0) = (2, 0)$.

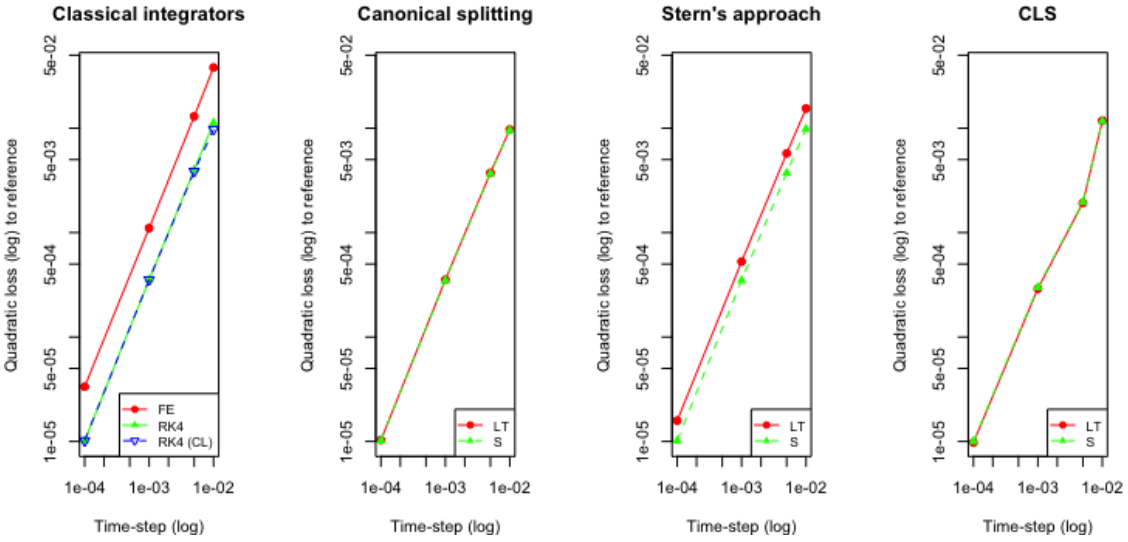

Figure 7: FHN: Quadratic loss to reference solution for various time-steps

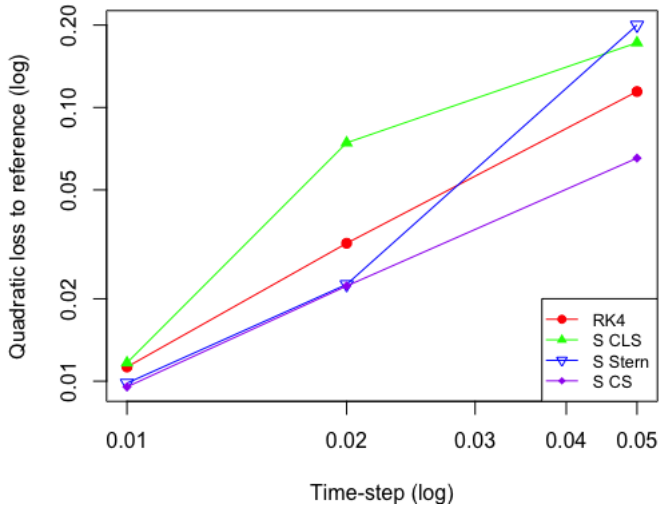

Figure 8: FHN: Quadratic loss to reference solution for large time steps

This provides some evidence of the advantages of splitting approaches even for the spike-reset models.

Comparisons of trajectories for different time-steps are difficult for spike-reset models due to the **first passage time problem**. The first passage time (FPT) is the first time t at which the trajectory crosses the threshold and gets reset. In a numerical solution, the FPT depends on the time-step, and as such the entire trajectory is effected. An error is therefore introduced into the numerical solution that is not due to the approximate integration. In order to combat this issue, more complex algorithms are needed such as those presented by Stewart and Bair [19]. These algorithms involve numerically calculating the exact spike time rather than resetting at the discrete time-steps [19]. Further work could be done to combine this approach with the splitting methods, but the resulting method could be computational intractable for inference purposes.

5.1.4 The adaptive exponential integrate-and-fire model

The AdEx model is structured similarly to the IZ model with a spike-reset. As such, some similar behaviour can be seen and similar problems arise. Figure 12 shows some trajectories of the AdEx model calculated with various numerical methods. The classical integrators and canonical splitting methods are run with a time-step of $h = 5 \times 10^{-4}$ and exhibit the correct spiking behaviour with the same frequency and amplitude as the reference solution (calculated with RK4 and $h = 5 \times 10^{-5}$). The splitting approaches on the conditionally linear version of the system do not perform as well for the larger time-step and so the figure shows trajectories calculated

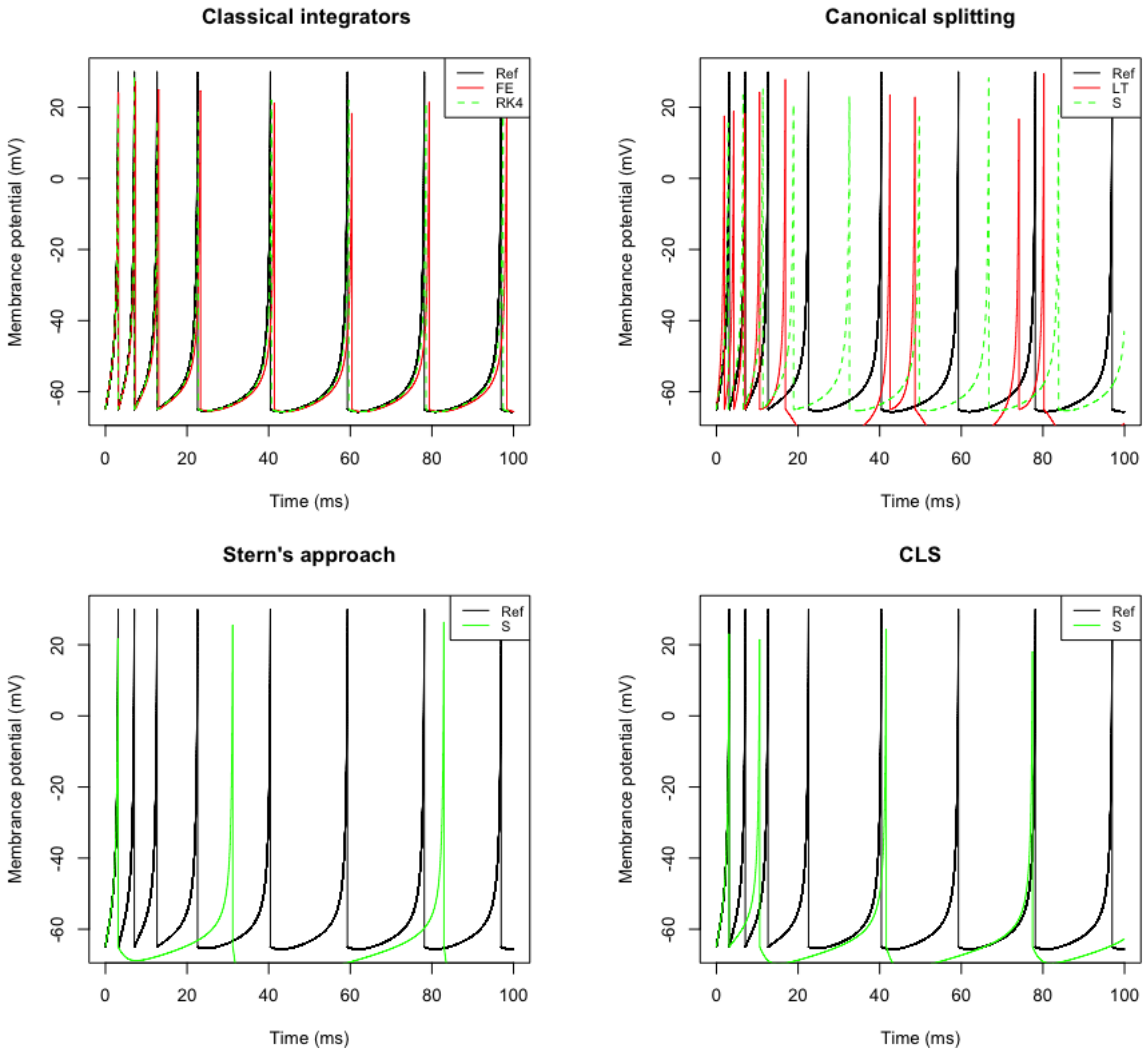

Figure 9: IZ: trajectories compared with reference solution for various numerical methods run with time-step $h = 0.05$ and parameters $a = 0.04; b = 5; c = 140; d = -1; \alpha = 0.02; \beta = 0.2; I = 10; f = 1; (v_0, u_0) = (-65, -13)$.

using a time-step of $h = 5 \times 10^{-5}$. The Lie-Trotter compositions either exploded in finite time or did not converge. The Strang compositions are out of phase with the reference solution. From Figure 12, one can see that the trajectories line up closely before the first spike-reset, indicating that the FPT problem also plays a role here. The disparity between the schemes that use the original system and those that use the linearised system suggest that the AdEx model, most probably due to its reset condition, is highly sensitive to the coordinate transformation.

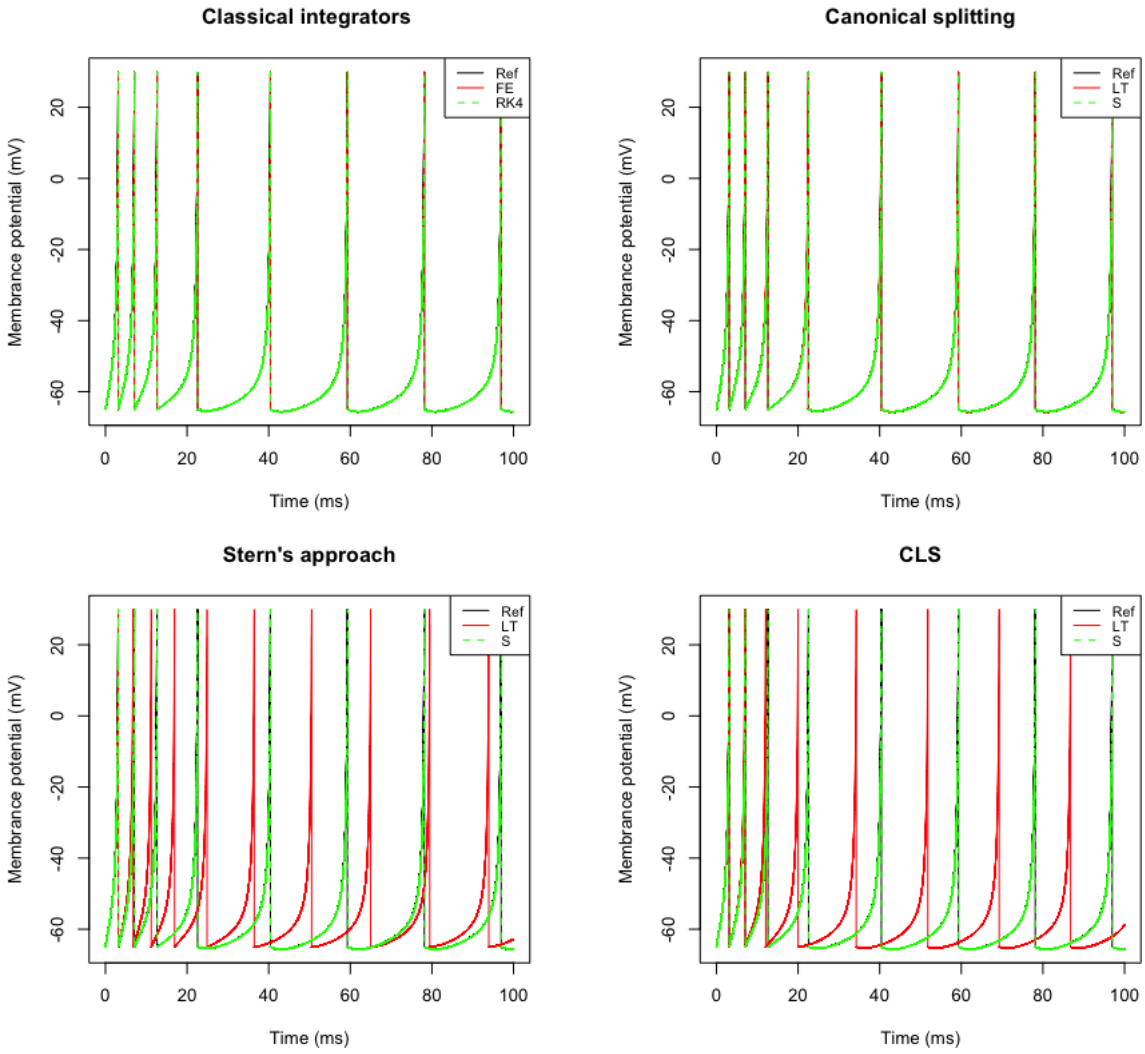

Figure 10: IZ: trajectories compared with reference solution for time-step $h = 0.0005$

5.2 Stochastic models

For a stochastic model, the time between two adjacent spikes varies from path to path. To compare the performance of the numerical methods for the stochastic models, we will consider the distribution of this inter-spike interval (ISI) for a given interval. Over 500 iterations, we use the normal kernel density approximation (see appendix B), to calculate an approximate distribution of the ISI_j for $j \in \{1, 2, 3, 4, 5\}$, and then calculate the **absolute integral loss** between this density curve and a reference distribution. The **absolute integral loss** is the area between the two curves in absolute value. A smaller absolute integral loss means that the distribution of the ISI matches more closely with the reference solution and therefore that the numerical method has better preserved a key feature of the model. The reference solution in each case is computed using the mean-square convergent

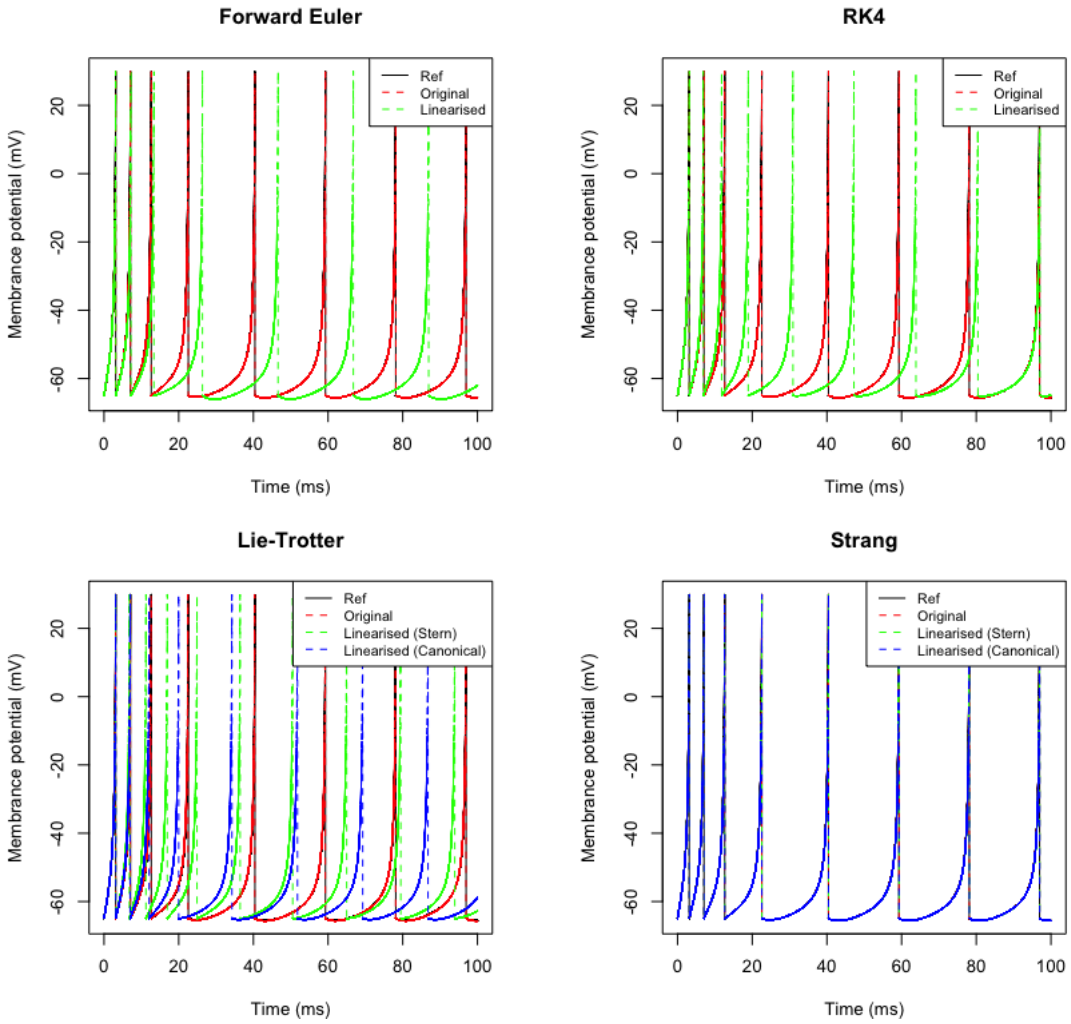

Figure 11: IZ: numerical schemes with time-step $h = 0.0005$ comparing the original system to the linearised system.

tamed Euler-Maruyama scheme, as defined in Section 3.4, and a small time-step.

5.2.1 The "non-stiff" Van der Pol oscillator

If $0 < \epsilon \ll 1$, the VDP model is called "non-stiff". For this range of ϵ values, the VDP has a circular limit cycle of radius 2 centered at the origin. This is a key feature that we wish to preserve in a numerical solution. Figure 13 shows paths simulated using different methods for the stochastic VDP with $\epsilon = 0.05$ and starting position $(0, 0)$. A well-performing numerical method will preserve the radius of the limit cycle for larger time steps. There is some qualitative evidence from Figure 13 that the splitting approaches, particularly the canonical splitting, preserve this amplitude more robustly than the Euler-Maruyama schemes, defined in Section 3.3, as the time-step increases.

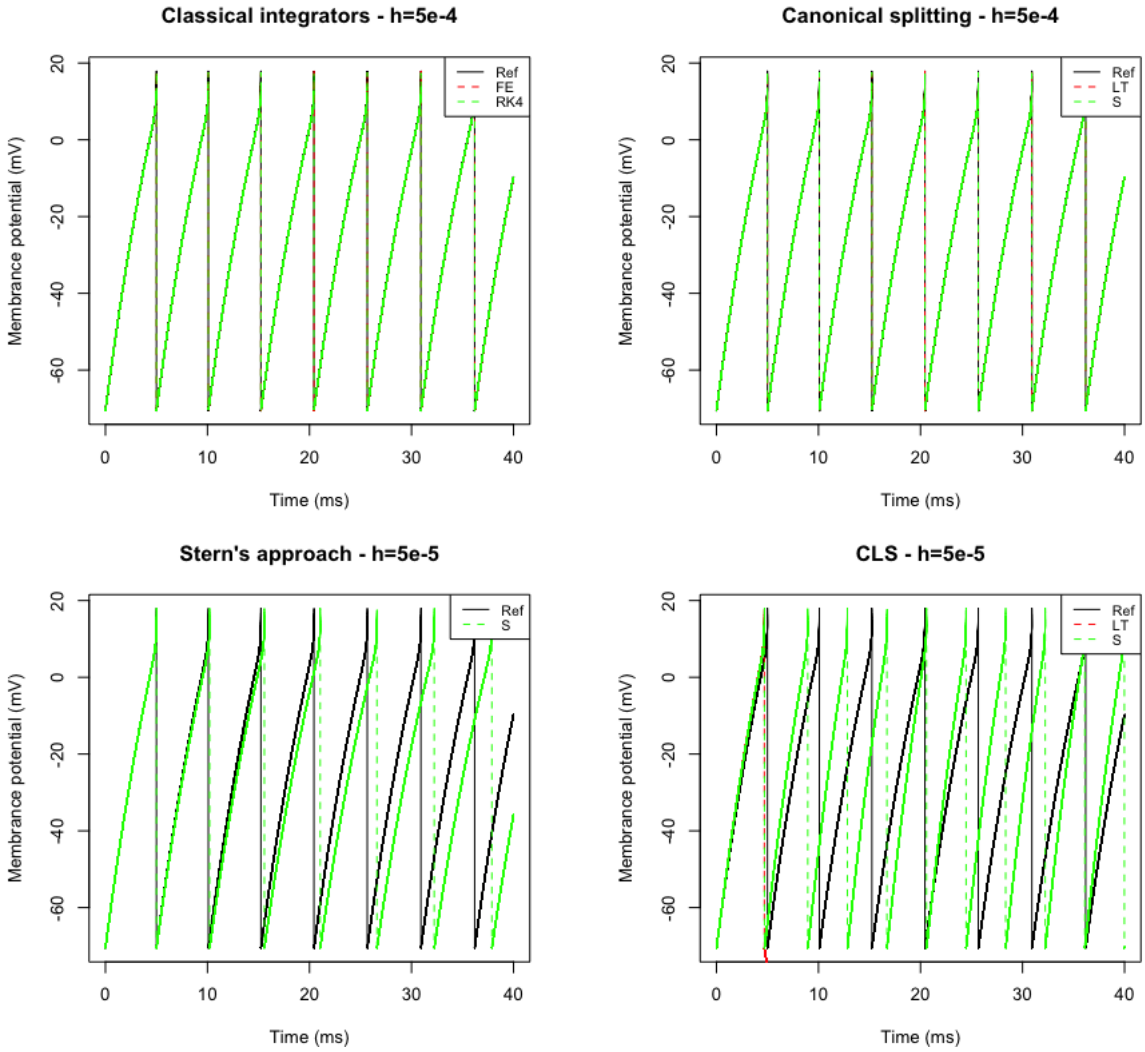

Figure 12: AdEx:trajectories compared with reference solution for various numerical methods for parameters $a = 0.2135; c = 2; d = -0.1068; e = -7.5374; f = -0.0356; g = 0.0356; j = 0.02778; k = 1.9611; l = -0.0694; I = 350 (v_0, u_0) = (-70.6, -13)$

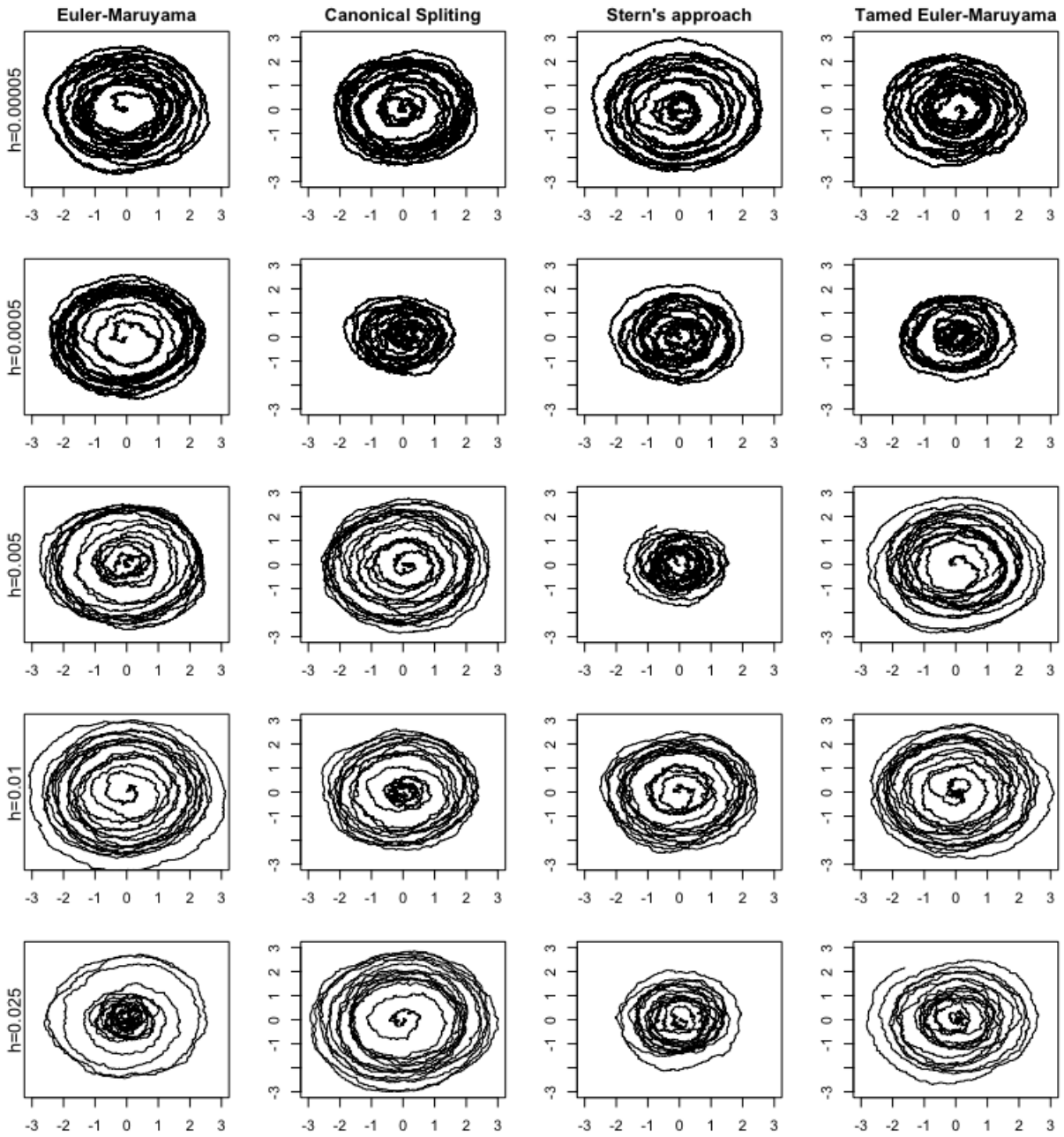

Figure 13: VDP: phase portraits for the stochastic VDP model in (x, y) space showing limit cycle of true radius 2. The splitting methods are composed with Strang.

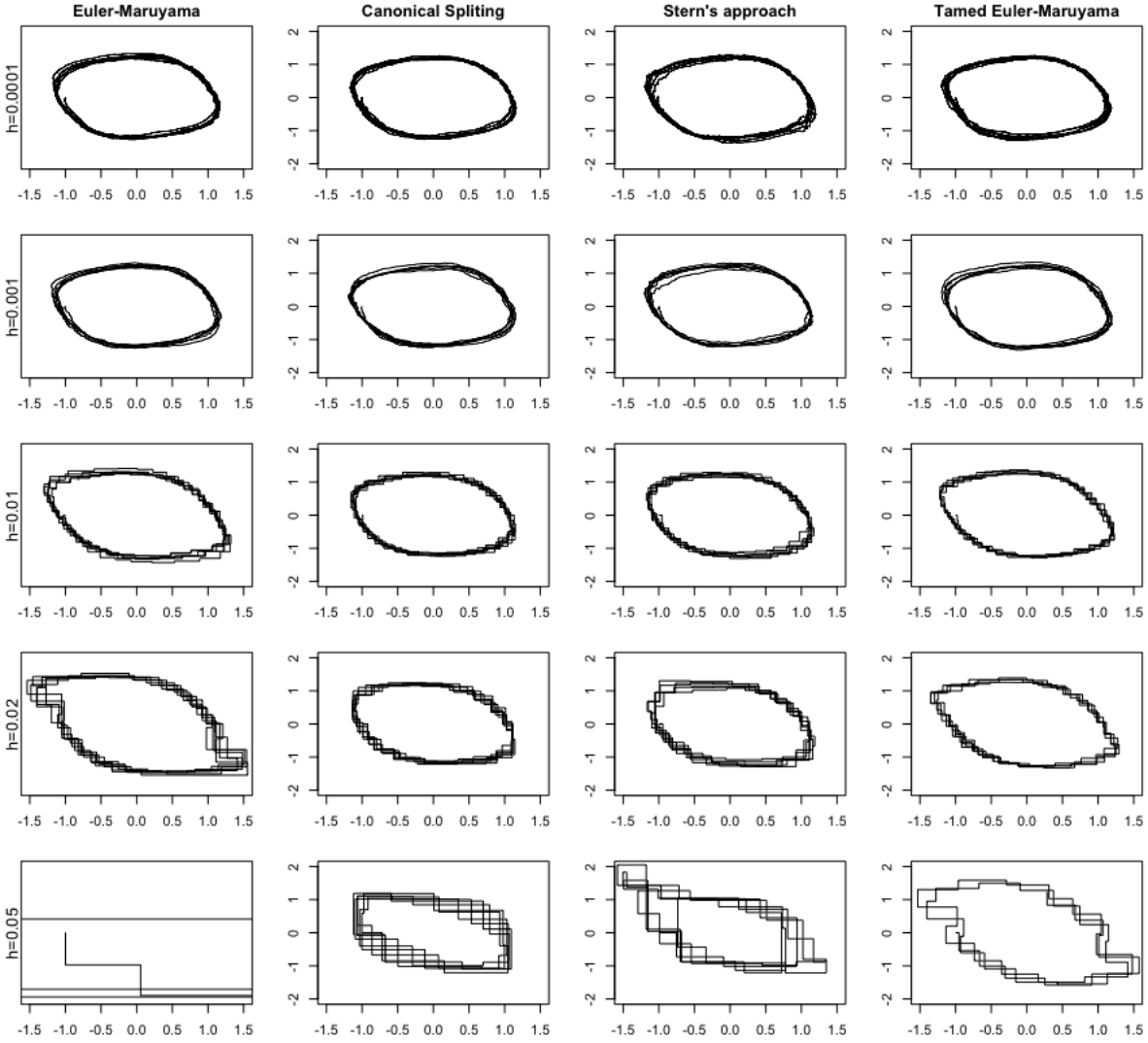

Figure 14: FHN: phase portraits in (v, u) for various numerical schemes, showing the relative preservation of the limit cycle. The Strang composition is used for the splitting methods.

5.2.2 The Fitz-Hugh Nagumo model

As seen in the deterministic case, the FHN has limit cycle behaviour [16]. Figure 14 shows phase portraits of paths simulated from the stochastic FHN with various numerical schemes. The result here is even clearer than for the stochastic VDP model. The splitting approaches preserve the limit cycle behavior better than the Euler-Maruyama schemes for larger time-steps. Furthermore, the canonical splitting out-performs Stern's approach for these time-steps. Figure 15 shows the absolute integral loss between the inter-spike-interval (ISI) distribution densities, for the first 5 spikes. These are calculated with 500 samples and compared against a reference distribution, obtained by using the tamed Euler Maruyama scheme with a time step of $h = 2 \times 10^{-4}$ for 500 iterations. The spike threshold was chosen to be $v(t) \geq 1$.

Figure 15 shows the absolute integral loss for each of the first five spikes across a range of time-steps. The Stern approach has a significantly higher loss than both the canonical splitting and the Euler-Maruyama scheme. The Euler-Maruyama method was not convergent for the larger time-steps, hence the integral loss is only plotted for the smaller time-steps. The integral loss is comparable for the canonical splitting and the Euler-Maruyama scheme for small time-steps, but as the canonical splitting is the only method with low loss for all time-steps, it can be regarded as the most robust method.

5.2.3 The Izhikevich model

For the stochastic IZ model, the distribution of the ISI is an important property we wish to preserve. Moreover, this is an interesting experiment as the FPT problem does not play a role, as we are interested in the *interval* between spikes as opposed to the exact spiking time. Again, we compute the absolute integral loss between the various methods for the IZ model with a reference density. The reference density was computed with tamed Euler-Maruyama and a time-step of $h = 2 \times 10^{-5}$. Figure 16 shows the absolute integral loss between reference and the approximate densities for the IZ model at a range of different time-steps. Here it is clear that the Stern approach does not behave as expected or desired. The loss does not decrease with time-step meaning that the distributions of the ISI under this numerical scheme are not approaching the reference distributions. A problem with the initial spike meant that data for the Stern approach was not included for Spike 1. This weak performance is most likely due to the problems with changing coordinates in a spike-reset system as seen with the results for the deterministic AdEx model in Figure 12. The Euler-Maruyama and canonical splitting methods seem to have comparable loss values.

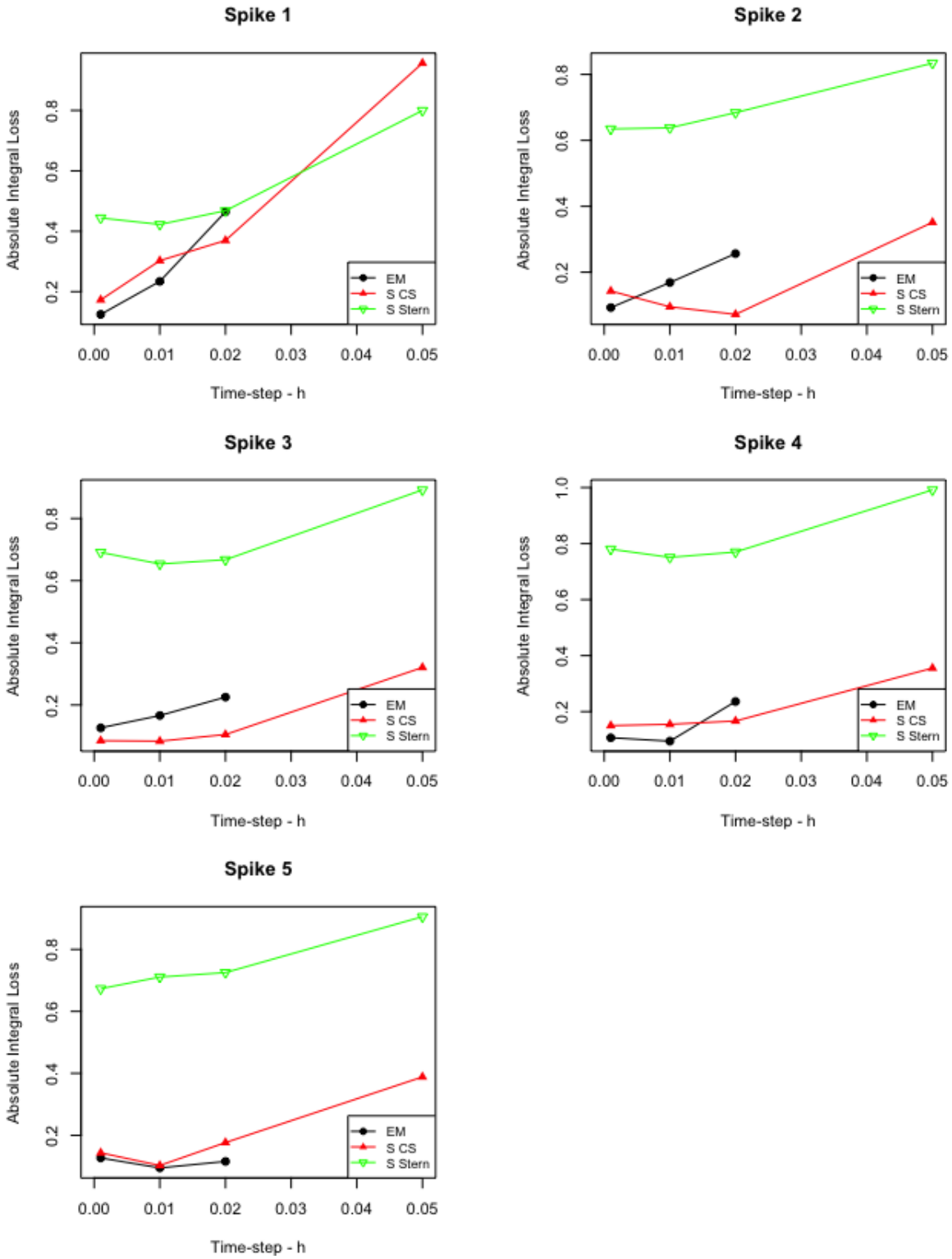

Figure 15: FHN: Absolute integral loss between reference ISI density and the density obtained from the studied methods for the first 5 spikes.

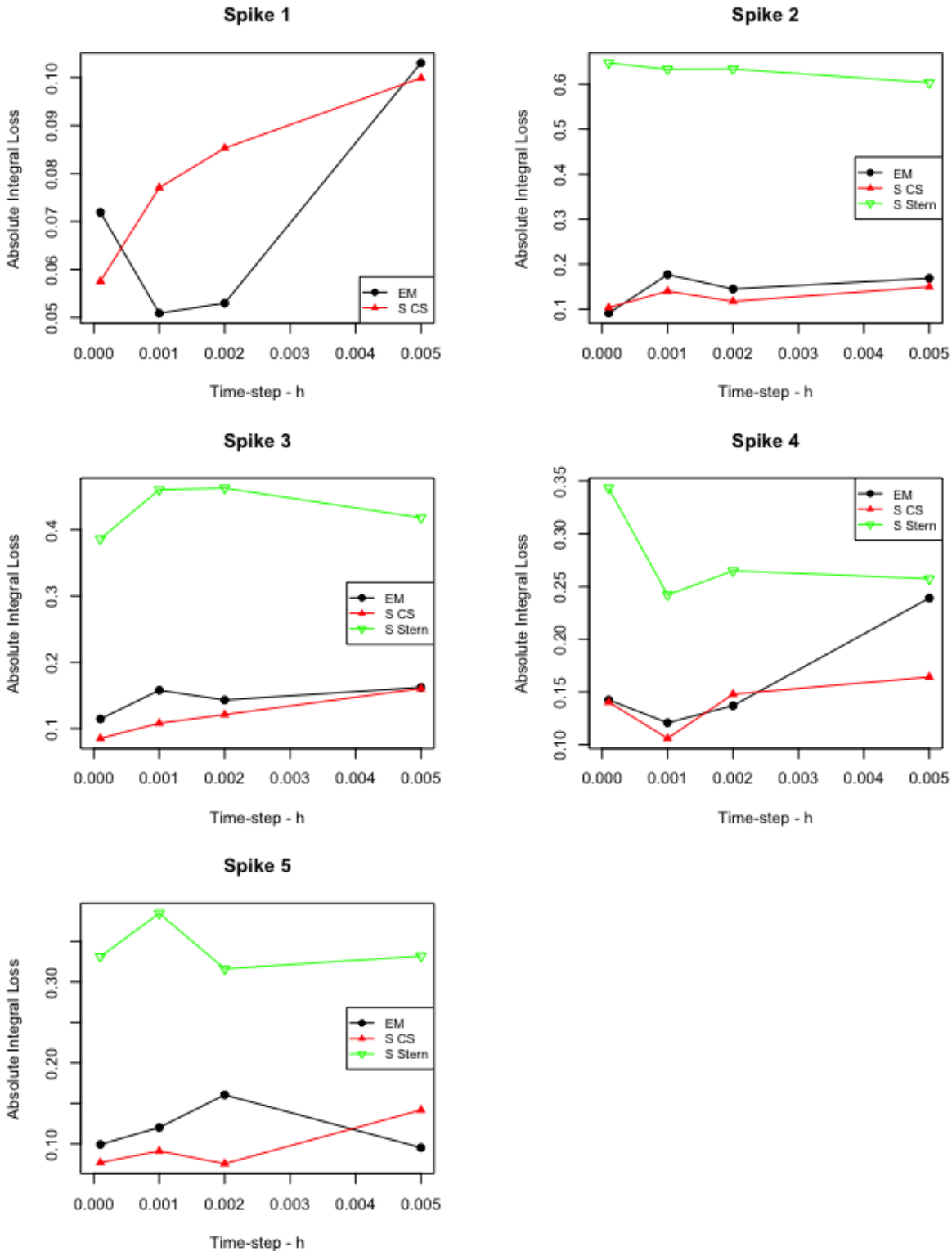

Figure 16: IZ: Absolute integral loss between reference ISI density and the density obtained from the studied methods for the first 5 spikes.

6 Discussion

In summary, this project has studied the application of splitting methods to four deterministic and stochastic neuronal models (VDP, FHN, IZ, AdEx). Building off of work done by Chen et al for conditionally linear ODEs and Buckwar et al for semi-linear SDEs, we have considered the "canonical" approach of splitting the model into a linear and a non-linear part, as well as a splitting for conditionally linear systems [2][3]. Generalising a coordinate change proposed by León and Samson [14], we have been able to extend Stern's approach to models that are not conditionally linear. We have also proposed an extension to Stern's approach so that it can be applied to SDEs. This method was proved to have a normally distributed 1-step transition probability under specific conditions, a property useful for applications in parameter inference [2]. Furthermore, it was shown that the 1-step hypoellipticity can depend on the order of composition of the flows. The ergodicity and mean-square convergence of these methods remain unproved. The numerical results provided some evidence that the canonical splitting preserved spiking dynamics for larger time steps for the FHN model better than classical integrators and the alternative splitting approaches. For the stiff VDP model, Stern's approach proved to be the most robust to time-step changes, reinforcing previous results from Chen et al [3]. For the spike-reset models (IZ, AdEx) only the Strang composition was resistant to coordinate changes in the IZ, and the canonical splitting preserved the spiking dynamics of the AdEx. However, all methods proved not robust for larger time-steps due to the first passage time problem. Stewart and Bair propose a method of reducing this error by numerically approximating the exact spike-time and integrating this information to reset more accurately [19]. Further work could be done to combine this approach with the splitting methods to better analyse the properties of the splitting integrators in the absense of the reset error and to produce high accuracy schemes. However, this could increase computational cost making the methods intractable for parameter inference. Furthermore, numerical results showed that the coordinate change produced some unexpected behaviour in the spike-reset models. The resets in these numerical implementations were performed in the original (non-transformed) coordinates as the unexpected behaviours were worse when the reset conditions were transformed along with the system. This phenomenon could be investigated theoretically to see if there is a mathematical basis for this effect. Finally, this project neglects to study the computational efficiency of the various methods. Preliminary investigations suggest a large disparity between the splitting methods, with Stern's approach being faster than the canonical splitting. Numerical experiments could obtain results about how the error of these methods scales with machine time, an important metric for spiking network simulations.

7 References

[1] Romain Brette and Wulfram Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 2005.

[2] Evelyn Buckwar, Adeline Samson, Massimiliano Tamborrino, and Irene Tubikanec. Splitting methods for SDEs with locally Lipschitz drift. an illustration on the FitzHugh-Nagumo model. *arXiv:2101.01027*, January 2021.

[3] Zhengdao Chen, Baranidharan Raman, and Ari Stern. Structure-preserving numerical integrators for Hodgkin–Huxley-type systems. *SIAM Journal on Scientific Computing*, 2020.

[4] Balth Van der Pol. On "relaxation-oscillations". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1926.

[5] Artur Gramacki. *Nonparametric Kernel Density Estimation and Its Computational Aspects*, volume 37, pages 27–39. Springer Series in Big Data, First edition, 2018.

[6] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration*, volume 31, pages 47–50. Springer Series in Computational Mathematics, Second edition, 2002.

[7] Ernst Hairer, Gerhard Wanner, and Syvert P. Nørsett. *Solving Ordinary Differential Equations I*, volume 8, pages 156–162. Springer Series in Computational Mathematics, Second edition, 1993.

[8] Martin Hutzenthaler, Arnulf Jentzen, and Peter Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A*, 2010.

[9] Martin Hutzenthaler, Arnulf Jentzen, and Peter Kloeden. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Annals of Applied Probability*, 2012.

[10] Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 1944.

[11] Eugene Izhikevich. Neural excitability, spiking and bursting. *International Journal of Bifurcation and Chaos*, 2000.

[12] Eugene Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 2003.

[13] Peter Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23, pages 23–111. Springer Stochastic Modelling and Applied Probability, Third edition, 1999.

[14] José R. León and Adeline Samson. Hypoelliptic stochastic FitzHugh–Nagumo neuronal model: Mixing, up-crossing and estimation of the spike rate. *The Annals of Applied Probability*, 2018.

[15] J.C. Mattingly, A.M. Stuart, and D.J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Processes and their Applications*, 2002.

[16] Matthias Ringqvist. On dynamical behaviour of FitzHugh-Nagumo systems. *Stockholm University Research Reports in Mathematics Number 5*, 2006.

[17] Bernard Silverman. *Density Estimation for Statistics and Data Analysis*, page 45. Chapman and Hall, 1986.

[18] Ari Stern. Talk: Structure-preserving numerical integrators: classic methods, new perspectives. *Warwick Seminar for Algorithms and Computationally Intensive Inference*, 2021.

[19] Robert Stewart and Wyeth Bair. Spiking neural network simulation: numerical integration with the Parker-Sochacki method. *Journal of Computational Neuroscience*, 2009.

A Application of the splitting methods to the spike-reset models

A.1 The Izhikevich Model

The Izhikevich (IZ) model is is given by the 2-d system of ODEs,

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} av^2(t) + bv(t) + c + du(t) + fI \\ \alpha(\beta v(t) - u(t)) \end{pmatrix} dt, \tag{151}$$

with a reset condition,

$$\begin{aligned} &\text{If } v \geq 30 \\ &\qquad v \to \theta \qquad (152) \\ &\qquad u \to u + \delta. \qquad (153) \end{aligned}$$

A.1.1 The canonical splitting

We can apply the canonical splitting to the IZ model with the following, non-unique, approach,

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} b & d \\ \alpha\beta & -\alpha \end{pmatrix} \begin{pmatrix} v(t) \\ u(t) \end{pmatrix} + \begin{pmatrix} av(t)^2 + c + fI \\ 0 \end{pmatrix} dt. \tag{154}$$

Following the analysis in Section 2.2, we therefore have,

$$A = \begin{pmatrix} b & d \\ \alpha\beta & -\alpha \end{pmatrix} \;\; ; \;\; N(v,u) = \begin{pmatrix} av(t)^2 + c + fI \\ 0 \end{pmatrix}. \tag{155}$$

This therefore yields sub-problems,

$$dx^{[1]}(t) = \begin{pmatrix} b & d \\ \alpha\beta & -\alpha \end{pmatrix} x^{[1]}(t)\, dt \tag{156}$$

$$dx^{[2]}(t) = \begin{pmatrix} av^{[2]}(t)^2 + c + fI \\ 0 \end{pmatrix} dt \tag{157}$$

where $x^{[1]}(t) = (v^{[1]}(t), u^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), u^{[2]}(t)) \in \mathbb{R}^2$. The non-linear sub-equation is separable and therefore integrable. Proceeding as before we

get $t-$time flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a} t + \arctan(\sqrt{\frac{a}{\eta}} v_0)) \\ u_0 \end{pmatrix}, \end{cases} \tag{158}$$

where $\eta = c + fI$. We recompose these flows with Lie-Trotter and Strang compositions to obtain explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a} h + \arctan(\sqrt{\frac{a}{\eta}} v_{i-1})) \\ u_{i-1} \end{pmatrix}, \tag{159}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) \tag{160}$$

$$= \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a} \frac{h}{2} + \arctan(\sqrt{\frac{a}{\eta}} v_{i-1})) \\ u_{i-1} \end{pmatrix} \right), \tag{161}$$

where $x_{i-1} = (v_{i-1}, u_{i-1})$.

A.1.2 Coordinate change and Stern's approach

Again this model is not conditionally linear but is in the form described in Section 2.5. In this case the functions and constants are,

$$g(v) = av^2 + bv + c + fI, \tag{162}$$

$$k = d, \tag{163}$$

$$a(v) = -\alpha, \tag{164}$$

$$b(v) = \alpha\beta v. \tag{165}$$

Applying the coordinate change $y(t) = g(v(t)) + ku(t)$, yields the conditionally linear system,

$$d \begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ (2av(t) + b - \alpha)y(t) + (d\alpha\beta v(t) + \alpha(av^2(t) + bv(t) + c + fI)) \end{pmatrix} dt \tag{166}$$

The spike reset condition can either be transformed under the coordinate change such that the reset is defined in the coordinates (v, y),

$$v \to \theta \tag{167}$$

$$y \to y + d\delta, \tag{168}$$

or a coordinate switch can be performed in each step of the numerical implementation to perform the reset in the original coordinates (v, u). The (v, u) reset proved to be far more effective at the implementation level.
Again, we apply the canonical splitting to the conditionally linear version of the system. A suitable splitting would be,

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ d\alpha\beta + \alpha b & b - \alpha \end{pmatrix} \begin{pmatrix} v(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 2av(t)y(t) + \alpha a v^2(t) + \alpha c + \alpha f I \end{pmatrix} dt. \tag{169}$$

Again, the non-linear part yields a separable sub-equation. Proceeding as before, from an initial condition $x(0) = x_0 = (v_0, y_0) \in \mathbb{R}^2$, we get $t-$time flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} v_0 \\ \frac{1}{2av_0}\left[\exp(2av_0t)(2av_0y_0) + (\exp(2av_0t) - 1)(\alpha a v_0^2 + \alpha c + \alpha f I)\right] \end{pmatrix}, \end{cases} \tag{170}$$

with,

$$A = \begin{pmatrix} 0 & 1 \\ d\alpha\beta + \alpha b & b - \alpha \end{pmatrix} \tag{171}$$

We can then obtain explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) \tag{172}$$

$$= e^{Ah} \begin{pmatrix} v_{i-1} \\ \frac{1}{2av_{i-1}}\left[\exp(2av_{i-1}h)(2av_{i-1}y_{i-1}) + (\exp(2av_{i-1}h) - 1)(\alpha a v_{i-1}^2 + \alpha c + \alpha f I)\right] \end{pmatrix}, \tag{173}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) \tag{174}$$

$$= \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} v_{i-1} \\ \frac{1}{2av_{i-1}}\left[\exp(2av_{i-1}\frac{h}{2})(2av_{i-1}y_{i-1}) + (\exp(2av_{i-1}\frac{h}{2}) - 1)(\alpha a v_{i-1}^2 + \alpha c + \alpha f I)\right] \end{pmatrix} \right), \tag{175}$$

where $x_{i-1} = (v_{i-1}, y_{i-1})$. This conditionally linear system has functions

$$a_1(y) = 0;\ b_1(y) = y \tag{176}$$

$$a_2(v) = 2av + b - \alpha;\ b_2(v) = d\alpha\beta v + \alpha(av^2 + bv + c + fI) \tag{177}$$

from Definition 1. These give us $t-$time flows from x_0,

$$\begin{cases} \varphi_t^{[1]}(x_0) &= \begin{pmatrix} v_0 + ty_0 \\ y_0 \end{pmatrix}, \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} v_0 \\ \exp(t(2av_0 + b - \alpha))y_0 + \frac{\exp(t(2av_0+b-\alpha))-1}{t(2av_0+b-\alpha)} t(av_0^2 + (\alpha\beta d + b)v_0 + c + fI) \end{pmatrix}. \end{cases} \tag{178}$$

Again, these flows can be recomposed using the Lie-Trotter and Strang compositions and a time-step $h > 0$ to give explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = \begin{pmatrix} v_{i-1} + h\varphi_{h,2}^{[2]}(x_{i-1}) \\ \varphi_{h,2}^{[2]}(x_{i-1}) \end{pmatrix}, \tag{179}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \begin{pmatrix} \varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h/2,2}^{[2]}(x_{i-1})) \\ \varphi_{h/2,2}^{[2]}(\varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h,2}^{[2]}(x_{i-1})), (\varphi_{h/2,2}^{[2]})(x_{i-1})) \end{pmatrix}. \tag{180}$$

A.1.3 Izhikevich model with additive noise

As before, we now include an additive noise term which yields the SDE,

$$d\begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} av^2(t) + bv(t) + c + du(t) + fI \\ \alpha(\beta v(t) - u(t)) \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t), \tag{181}$$

where $\sigma_1, \sigma_2 \geq 0$ and W is the 2-d Wiener process. Proceeding as before, we generate sub-problems,

$$dx^{[1]}(t) = \begin{pmatrix} b & d \\ \alpha\beta & -\alpha \end{pmatrix} x^{[1]}(t)dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t), \tag{182}$$

$$dx^{[2]}(t) = \begin{pmatrix} av^{[2]}(t)^2 + c + fI \\ 0 \end{pmatrix} dt \tag{183}$$

where $x^{[1]}(t) = (v^{[1]}(t), u^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), u^{[2]}(t)) \in \mathbb{R}^2$. Proceeding as before we get $t-$time flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) &= x_0 e^{At} + \xi_t, \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a}t + \arctan(\sqrt{\frac{a}{\eta}} v_0)) \\ y_0 \end{pmatrix}, \end{cases} \tag{184}$$

where $\eta = c + fI$, $\xi_t \sim \mathcal{N}(0, C(t))$ and,

$$A = \begin{pmatrix} b & d \\ \alpha\beta & -\alpha \end{pmatrix}. \tag{185}$$

$C(t)$ is defined in Section 2.2 and explained further in Appendix C, but the expression is very long and not included. Using these flows and a time-step $h > 0$, we can define the following numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a} h + \arctan(\sqrt{\frac{a}{\eta}} v_{i-1})) \\ y_{i-1} \end{pmatrix} + \xi_{i-1}, \tag{186}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} \sqrt{\frac{\eta}{a}} \tan(\sqrt{\eta a} \frac{h}{2} + \arctan(\sqrt{\frac{a}{\eta}} v_{i-1})) \\ y_{i-1} \end{pmatrix} + \xi_{i-1} \right), \tag{187}$$

where $\xi_{i-1} \sim \mathcal{N}(0, C(h))$. Again, following the analysis in Section 2.5, we can perform the stochastic version of the coordinate change to derive the conditionally linear SDE below,

$$d \begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ y(t)(2av(t) + b - \alpha) + \alpha(d\beta v(t) + av^2(t) + bv(t) + c + fI) + a\sigma_1^2 \end{pmatrix} dt \tag{188}$$

$$+ \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sqrt{(2av + b)^2 \sigma_1^2 + d^2 \sigma_2^2} \end{pmatrix} dW(t).$$

As in Definition 4, the conditionally linearised IZ with additive noise is a conditionally linear SDE in (v, y) with functions,

$$a_1(y) = 0;\ a_2(v) = 2av + b - \alpha \tag{189}$$

$$b_1(y) = y;\ b_2(v) = \alpha((d\beta + b)v + av^2 + c + fI) + a\sigma_1^2 \tag{190}$$

$$c_1(y) = \sigma_1;\ c_2(v) = \sqrt{(2av + b)^2 \sigma_1^2 + d^2 \sigma_2^2}. \tag{191}$$

We can substitute these functions into the $t-$time flows given in Section 2.4, starting at $x_0 = (v_0, y_0)$ to give flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) &= \begin{pmatrix} \xi_t \\ y_0 \end{pmatrix} \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} v_0 \\ \nu_t \end{pmatrix}, \end{cases} \tag{192}$$

where

$$\xi_t \sim N(v_0 + y_0 t, \sigma_1^2 t), \tag{193}$$

$$\nu_t \sim N(\hat{\mu}, \hat{\sigma}), \tag{194}$$

with

$$\hat{\mu} = y_0 \exp((2av_0 + b - \alpha)t) - \frac{\alpha(d\beta v_0 + av_0^2 + bv_0 + c + fI) + a\sigma_1^2}{2av_0 + b - \alpha}(1 - \exp((2av_0 + b - \alpha)t)), \tag{195}$$

$$\hat{\sigma} = \frac{(2av + b)^2\sigma_1^2 + d^2\sigma_2^2}{2(2av_0 + b - \alpha)}(\exp(2(2av_0 + b - \alpha)) - 1). \tag{196}$$

A.2 The adaptive exponential integrate-and-fire model

The adaptive exponential integrate-and-fire model is given by the 2-d system of ODEs

$$d\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} a\exp(\frac{v(t)}{c}) + dv(t) + fw(t) + gI \\ jv(t) + lw(t) + k \end{pmatrix} dt \tag{197}$$

With a reset condition,

$$\text{If } v \geq 18$$

$$v \to \theta \tag{198}$$

$$w \to w + \delta \tag{199}$$

A.2.1 The canonical splitting

We can apply the canonical splitting to this ODE in the following way,

$$d\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} d & f \\ j & l \end{pmatrix}\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} + \begin{pmatrix} a\exp(\frac{v(t)}{c}) + gI \\ k \end{pmatrix} dt. \tag{200}$$

We therefore have,

$$A = \begin{pmatrix} d & f \\ j & l \end{pmatrix} \quad ; \quad N(v,w) = \begin{pmatrix} a\exp(\frac{v(t)}{c}) + gI \\ k \end{pmatrix}. \tag{201}$$

This therefore yields sub-problems,

$$dx^{[1]}(t) = \begin{pmatrix} d & f \\ j & l \end{pmatrix} x^{[1]}(t)dt \tag{202}$$

$$dx^{[2]}(t) = \begin{pmatrix} a\exp(\frac{v(t)}{c}) + gI \\ k \end{pmatrix} dt \tag{203}$$

where $x^{[1]}(t) = (v^{[1]}(t), w^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), w^{[2]}(t)) \in \mathbb{R}^2$. Proceeding as before we get $t-$time flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_0/c)}{gI\exp(gIt/c)} - \frac{a}{gI}) \\ w_0 + tk \end{pmatrix}. \end{cases} \tag{204}$$

We recompose these flows with Lie-Trotter and Strang compositions to obtain explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_{i-1}/c)}{gI\exp(gIh/c)} - \frac{a}{gI}) \\ w_{i-1} + hk \end{pmatrix}, \tag{205}$$

$$x_i^S = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) \tag{206}$$

$$= \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_{i-1}/c)}{gI\exp(gIh/2c)} - \frac{a}{gI}) \\ w_{i-1} + \frac{hk}{2} \end{pmatrix} \right), \tag{207}$$

where $x_{i-1} = (v_{i-1}, w_{i-1})$.

A.2.2 Coordinate change and the Stern approach

Again, for this final model, we have a system which is not conditionally linear but instead of the form described in Section 2.5 with functions and constants,

$$g(v) = a\exp\left(\frac{v}{c}\right) + dv + gI, \tag{208}$$

$$k = f, \tag{209}$$

$$a(v) = l, \tag{210}$$

$$b(v) = jv + k. \tag{211}$$

Note that the first k refers to the coefficient of u in the first component whereas the second k is a parameter in the AdEx model - similar with g, a. Applying the coordinate change $y(t) = g(v(t)) + ku(t)$ as usual, we derive a conditionally linear version of the AdEx model. Recall that, as in the Izhikevich model, the spike rest condition can either be transformed into (v, y) under the coordinate change, or maintained in (v, w) with coordinate change every time a spike reset must occur. We obtain the following conditionally linear version of the AdEx model,

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ y(t)(\frac{a}{c}\exp(\frac{v(t)}{c}) + d + l) + ((fj - dl)v(t) - al\exp(\frac{v(t)}{c}) - gIl + fk) \end{pmatrix} dt \tag{212}$$

Applying the canonical splitting to this conditional linear version of the model we get,

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ (fj - dl) & (d + l) \end{pmatrix} \begin{pmatrix} v(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{a}{c}\exp(\frac{v(t)}{c})y + fk - al\exp(\frac{v(t)}{c}) - gIl \end{pmatrix} dt. \tag{213}$$

Proceeding as before we get $t-$time flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) & = x_0 e^{At}, \\ \varphi_t^{[2]}(x_0) & = \begin{pmatrix} v_0 \\ (y_0 + \frac{\beta}{\alpha})e^{\alpha t} - \frac{\beta}{\alpha} \end{pmatrix}, \end{cases} \tag{214}$$

where $\alpha = \frac{a}{c}\exp(\frac{v_0}{c})$, $\beta = fk - al\exp(\frac{v_0}{c}) - gIl$ and,

$$A = \begin{pmatrix} 0 & 1 \\ (fj - dl) & (d + l) \end{pmatrix}. \tag{215}$$

Applying the Lie-Trotter and Strang compositions, we obtain explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) \tag{216}$$

$$= e^{Ah} \begin{pmatrix} v_{i-1} \\ (y_{i-1} + \frac{\beta_{i-1}}{\alpha_{i-1}})e^{\alpha_{i-1}h} - \frac{\beta_{i-1}}{\alpha_{i-1}} \end{pmatrix}, \tag{217}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) \tag{218}$$

$$= \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} v_{i-1} \\ (y_{i-1} + \frac{\beta_{i-1}}{\alpha_{i-1}})e^{\alpha_{i-1}\frac{h}{2}} - \frac{\beta_{i-1}}{\alpha_{i-1}} \end{pmatrix} \right), \tag{219}$$

where $x_{i-1} = (v_{i-1}, y_{i-1})$, $\alpha_{i-1} = \frac{a}{c}\exp(\frac{v_{i-1}}{c})$ and $\beta = fk - al\exp(\frac{v_{i-1}}{c}) - gIl$. Following from Definition 1, we have functions,

$$a_1(y) = 0;\ b_1(y) = y, \tag{220}$$

$$a_2(v) = \frac{a}{c}\exp(\frac{v}{c}) + d + l;\ b_2(v) = (fj - dl)v + fk - al\exp(\frac{v}{c}) - gIl, \tag{221}$$

which we can substitute into the formulas given in Section 2.3 to obtain flows,

$$\varphi_t^{[1]}(x_0) = \begin{pmatrix} v_0 + ty_0 \\ y_0 \end{pmatrix}, \tag{222}$$

$$\varphi_t^{[2]}(x_0) = \begin{pmatrix} v_0 \\ \exp(t(\frac{a}{c}\exp(\frac{v_0}{c}) + d + l))y_0 + \frac{\exp(t(\frac{a}{c}\exp(\frac{v_0}{c})+d+l))-1}{t(\frac{a}{c}\exp(\frac{v_0}{c})+d+l)}t((fj - dl)v_0 + fk - ale^{\frac{v_0}{c}} - gIl) \end{pmatrix}. \tag{223}$$

We recompose these flows with Lie-Trotter and Strang to obtain explicit numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = \begin{pmatrix} v_{i-1} + h\varphi_{h,2}^{[2]}(x_{i-1}) \\ \varphi_{h,2}^{[2]}(x_{i-1}) \end{pmatrix}, \tag{224}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \begin{pmatrix} \varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h/2,2}^{[2]}(x_{i-1})) \\ \varphi_{h/2,2}^{[2]}(\varphi_{h,1}^{[1]}(x_{i-1}^{[1]}, \varphi_{h,2}^{[2]}(x_{i-1})), (\varphi_{h/2,2}^{[2]})(x_{i-1})) \end{pmatrix}. \tag{225}$$

A.2.3 Adaptive exponential integrate-and-fire model with additive noise

As before, we consider an additive noise term, yielding an SDE model,

$$d\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} a\exp(\frac{v(t)}{c}) + dv(t) + fw(t) + gI \\ jv(t) + lw(t) + k \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} dW(t) \tag{226}$$

where $\sigma_1, \sigma_2 \geq 0$ and W is a 2-d Wiener process. Again, we apply the canonical splitting to this model to generate sub-problems,

$$dx^{[1]}(t) = \begin{pmatrix} d & f \\ j & l \end{pmatrix} x^{[1]}(t)\ dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}\ dW(t), \tag{227}$$

$$dx^{[2]}(t) = \begin{pmatrix} a\exp(\frac{v^{[2]}(t)}{c}) + gI \\ k \end{pmatrix}\ dt, \tag{228}$$

where $x^{[1]}(t) = (v^{[1]}(t), w^{[1]}(t)) \in \mathbb{R}^2$ and $x^{[2]}(t) = (v^{[2]}(t), w^{[2]}(t)) \in \mathbb{R}^2$. Proceeding as before we get $t-$time flows,

$$\varphi_t^{[1]}(x_0) = x_0 e^{At} + \xi_t, \tag{229}$$

$$\varphi_t^{[2]}(x_0) = \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_0/c)}{gI\exp(gIt/c)} - \frac{a}{gI}) \\ w_0 + tk \end{pmatrix}, \tag{230}$$

where $\xi_t \sim \mathcal{N}(0, C(t))$ and

$$A = \begin{pmatrix} d & f \\ j & l \end{pmatrix} \tag{231}$$

Using these flows and a time-step $h > 0$, we can define the following numerical schemes,

$$x_i^{LT} = (\varphi_h^{[1]} \circ \varphi_h^{[2]})(x_{i-1}) = e^{Ah} \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_{i-1}/c)}{gI\exp(gIh/c)} - \frac{a}{gI}) \\ w_{i-1} + hk \end{pmatrix} + \xi_{i-1}, \tag{232}$$

$$x_i^{S} = (\varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]})(x_{i-1}) = \varphi_{h/2}^{[2]} \left(e^{Ah} \begin{pmatrix} -c\log(\frac{a+gI\exp(-v_{i-1}/c)}{gI\exp(gIh/2c)} - \frac{a}{gI}) \\ w_{i-1} + \frac{hk}{2} \end{pmatrix} + \xi_{i-1} \right), \tag{233}$$

where $\xi_{i-1} \sim \mathcal{N}(0, C(h))$.

Remark 4. *There are three expressions for each covariance matrix $C(t)$ (for each model) depending on the sign of an auxiliary variable (a function of the parameters). In the case of the AdEx, one of these expressions was not possible to obtain in a closed form - this case coincides with the standard parameter setting used, as such, an implementation of the canonical splitting for the AdEx with additive noise was not achieved.*

Performing the coordinate change defined in Section 2.5, we obtain a conditionally linear form of the stochastic AdEx model,

$$d\begin{pmatrix} v(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ y(t)(\frac{a}{c}\exp(\frac{v(t)}{c}) + l) + (d - gIl + fk + (\frac{a\sigma_1^2}{2c^2} - al)\exp(\frac{v(t)}{c}) + (fj - dl)v(t)) \end{pmatrix} dt \tag{234}$$

$$+ \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sqrt{(\frac{a}{c}\exp(\frac{v(t)}{c})\sigma_1)^2 + (f\sigma_2)^2} \end{pmatrix} dW(t).$$

As in Definition 4, the conditionally linearised AdEx with additive noise is a condi-

tionally linear SDE in (v, y) with functions,

$$a_1(y) = 0;\ a_2(v) = \frac{a}{c}\exp\left(\frac{v}{c}\right) + l, \tag{235}$$

$$b_1(y) = y;\ b_2(v) = d - gIl + fk + \left(\frac{a\sigma_1^2}{2c^2} - al\right)\exp\left(\frac{v}{c}\right) + (fj - dl)v, \tag{236}$$

$$c_1(y) = \sigma_1\ c_2(v) = \sqrt{\frac{a}{c}\exp\left(\frac{v(t)}{c}\right)\sigma_1)^2 + (f\sigma_2)^2}. \tag{237}$$

We can substitute these functions into the $t-$time flows given in Section 2.4, starting at $x_0 = (v_0, y_0)$ to give flows,

$$\begin{cases} \varphi_t^{[1]}(x_0) &= \begin{pmatrix} \xi_t \\ y_0 \end{pmatrix} \\ \varphi_t^{[2]}(x_0) &= \begin{pmatrix} v_0 \\ \nu_t \end{pmatrix}, \end{cases} \tag{238}$$

where $\xi_t \sim N(v_0 + y_0 t, \sigma_1^2 t)$ and $\nu_t \sim N(y_0 \exp((\frac{a}{c}\exp(\frac{v(t)}{c}) + l)t) - \frac{d-gIl+fk+(\frac{a\sigma_1^2}{2c^2}-al)\exp(\frac{v(t)}{c})+(fj-dl)v(t)}{\frac{a}{c}\exp(\frac{v(t)}{c})+l}(1-\exp((\frac{a}{c}\exp(\frac{v(t)}{c})+l)t)), \frac{(\frac{a}{c}\exp(\frac{v(t)}{c})\sigma_1)^2+(f\sigma_2)^2}{2(\frac{a}{c}\exp(\frac{v(t)}{c})+l)}(\exp(2(\frac{a}{c}\exp(\frac{v(t)}{c})+l)t) - 1))$.

B Normal kernel density estimation

In Section 5.2, we aim to compute an estimate of the probability density function of the ISI in order to compare to the reference distribution obtained with tamed Euler-Maruyama. We use **kernel density estimation** with a **normal** kernel.
Using a numerical scheme, we obtain samples $(x_1, x_2, ..., x_n)$ from the distribution ISI_j. Because we are considering each spike j separately, these samples are independent and identically distributed.

Definition 8. *The* **kernel density estimator** *is given by,*

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i), \tag{239}$$

$$= \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{240}$$

where K is a non-negative function called the **kernel**. *$h > 0$ is called the* **band-**

width *[5]. We choose the* **normal kernel**,

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{241}$$

For choosing h, we apply **Silverman's rule of thumb** *[17],*

$$h = 0.9 \min\left(\sigma, \frac{IQR}{1.34}\right) n^{-\frac{1}{5}}, \tag{242}$$

where σ is the standard deviation of the samples and IQR is the inter-quartile range.

C Covariance matrices for stochastic harmonic oscillators

When applying the canonical splitting to a model with additive noise, we must exactly solve a linear SDE of the form,

$$dX(t) = AX(t)\, dt + \Sigma\, dW(t). \tag{243}$$

Following the analysis from Tamborrino et al [2], we know the exact solution is given by,

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}\Sigma\, dW(s) \tag{244}$$

The stochastic integral is normally distributed with mean 0 and covariance matrix,

$$C(h) = \int_0^h e^{A(h-s)}\Sigma\Sigma^\top (e^{A(h-s)})^\top\, ds \tag{245}$$

The form of the covariance matrix $C(h)$ often depends on the sign of a function of the parameters (a latent parameter). These covariance matrices are given for the VDP and FHN models. We will label the entries of the symmetric matrix as follows,

$$C(h) = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}. \tag{246}$$

C.1 Van der Pol

For the VDP model, the covariance matrix changes its form based upon the sign of the latent parameter $\epsilon^2 - 4$. In the case $\epsilon^2 - 4 < 0$, we have

$$c_{11} = \frac{1}{2\epsilon\left(\epsilon^2-4\right)}(\sinh(\epsilon h)\left(\sqrt{4-\epsilon^2}\epsilon\sin\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-1\right)\sigma_1^2+\sigma_2^2\right)+\epsilon^2\cos\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-3\right)\sigma_1^2+\sigma_2^2\right)-4\left(\sigma_1^2+\sigma_2^2\right)\right) \tag{247}$$
$$+\cosh(\epsilon h)\left(\sqrt{4-\epsilon^2}\epsilon\sin\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-1\right)\sigma_1^2+\sigma_2^2\right)+\epsilon^2\cos\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-3\right)\sigma_1^2+\sigma_2^2\right)-4\left(\sigma_1^2+\sigma_2^2\right)\right)-\left(\left(\epsilon^2-4\right)\left(\left(\epsilon^2+1\right)\sigma_1^2+\sigma_2^2\right)\right))$$

$$c_{12} = \frac{1}{2\left(\epsilon^2-4\right)}(\sinh(\epsilon h)\left(\cos\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-2\right)\sigma_1^2+2\sigma_2^2\right)+\epsilon\sqrt{4-\epsilon^2}\sigma_1^2\sin\left(\sqrt{4-\epsilon^2}h\right)-2\left(\sigma_1^2+\sigma_2^2\right)\right) \tag{248}$$
$$+\cosh(\epsilon h)\left(\cos\left(\sqrt{4-\epsilon^2}h\right)\left(\left(\epsilon^2-2\right)\sigma_1^2+2\sigma_2^2\right)+\epsilon\sqrt{4-\epsilon^2}\sigma_1^2\sin\left(\sqrt{4-\epsilon^2}h\right)-2\left(\sigma_1^2+\sigma_2^2\right)\right)-\left(\left(\epsilon^2-4\right)\sigma_1^2\right))$$

$$c_{22} = \frac{1}{2\epsilon\left(\epsilon^2-4\right)}(\sinh(\epsilon h)\left(\sqrt{4-\epsilon^2}\epsilon\left(\sigma_1^2-\sigma_2^2\right)\sin\left(\sqrt{4-\epsilon^2}h\right)+\epsilon^2\left(\sigma_1^2+\sigma_2^2\right)\cos\left(\sqrt{4-\epsilon^2}h\right)-4\left(\sigma_1^2+\sigma_2^2\right)\right) \tag{249}$$
$$+\cosh(\epsilon h)\left(\sqrt{4-\epsilon^2}\epsilon\left(\sigma_1^2-\sigma_2^2\right)\sin\left(\sqrt{4-\epsilon^2}h\right)+\epsilon^2\left(\sigma_1^2+\sigma_2^2\right)\cos\left(\sqrt{4-\epsilon^2}h\right)-4\left(\sigma_1^2+\sigma_2^2\right)\right)-\left(\left(\epsilon^2-4\right)\left(\sigma_1^2+\sigma_2^2\right)\right))$$

If $\epsilon^2 - 4 > 0$, we have,

$$c_{11} = -\frac{1}{2\epsilon\left(\epsilon^2-4\right)}(\epsilon^4\sigma_1^2+\sqrt{\epsilon^2-4}\epsilon e^{\epsilon h}\sinh\left(\sqrt{\epsilon^2-4}h\right)\left(\left(\epsilon^2-1\right)\sigma_1^2+\sigma_2^2\right)-\epsilon^2 e^{\epsilon h}\cosh\left(\sqrt{\epsilon^2-4}h\right)\left(\left(\epsilon^2-3\right)\sigma_1^2+\sigma_2^2\right) \tag{250}$$
$$-3\epsilon^2\sigma_1^2+\epsilon^2\sigma_2^2+4\sigma_1^2e^{\epsilon h}+4\sigma_2^2e^{\epsilon h}-4\sigma_1^2-4\sigma_2^2)$$

$$c_{12} = -\frac{1}{2\left(\epsilon^2-4\right)}(\cosh(\epsilon h)\left(-\cosh\left(\sqrt{\epsilon^2-4}h\right)\left(\left(\epsilon^2-2\right)\sigma_1^2+2\sigma_2^2\right)+\epsilon\sqrt{\epsilon^2-4}\sigma_1^2\sinh\left(\sqrt{\epsilon^2-4}h\right)+2\left(\sigma_1^2+\sigma_2^2\right)\right) \tag{251}$$
$$+\sinh(\epsilon h)\left(-\cosh\left(\sqrt{\epsilon^2-4}h\right)\left(\left(\epsilon^2-2\right)\sigma_1^2+2\sigma_2^2\right)+\epsilon\sqrt{\epsilon^2-4}\sigma_1^2\sinh\left(\sqrt{\epsilon^2-4}h\right)+2\left(\sigma_1^2+\sigma_2^2\right)\right)+\left(\epsilon^2-4\right)\sigma_1^2)$$

$$c_{22} = \frac{1}{2\epsilon\left(\epsilon^2-4\right)}(\epsilon e^{\epsilon h}\left(\sqrt{\epsilon^2-4}\left(\sigma_2^2-\sigma_1^2\right)\sinh\left(\sqrt{\epsilon^2-4}h\right)\right. \tag{252}$$
$$\left.+\epsilon\left(\sigma_1^2+\sigma_2^2\right)\cosh\left(\sqrt{\epsilon^2-4}h\right)-\left(\epsilon^2+4e^{\epsilon h}-4\right)\left(\sigma_1^2+\sigma_2^2\right)\right))$$

If $\epsilon = 2$, we have,

$$c_{11} = \frac{1}{4}\left(\left(e^{2h}(2(h-3)h+5)-5\right)\sigma_1^2+\left(e^{2h}(2(h-1)h+1)-1\right)\sigma_2^2\right) \tag{253}$$

$$c_{12} = \frac{1}{2}\left(e^{2h}h^2\sigma_2^2+\left(e^{2h}(h-1)^2-1\right)\sigma_1^2\right) \tag{254}$$

$$c_{22} = \frac{1}{4}\left(\left(e^{2h}(2(h-1)h+1)-1\right)\sigma_1^2+\left(e^{2h}(2h(h+1)+1)-1\right)\sigma_2^2\right) \tag{255}$$

If $\epsilon = -2$, we have,

$$c_{11} = \frac{1}{4}e^{-2h}\left(\left(-2h(h+3)+5e^{2h}-5\right)\sigma_1^2+\left(-2h(h+1)+e^{2h}-1\right)\sigma_2^2\right) \tag{256}$$

$$c_{12} = \frac{1}{2}\left(e^{-2h}\left(h^2\sigma_2^2+(h+1)^2\sigma_1^2\right)-\sigma_1^2\right) \tag{257}$$

$$c_{22} = \frac{1}{4}\left(-e^{-2h}\left((2h(h+1)+1)\sigma_1^2+2(h-1)h\sigma_2^2+\sigma_2^2\right)+\sigma_1^2+\sigma_2^2\right) \tag{258}$$

C.2 Fitz-Hugh Nagumo

For the FHN model, the latent parameter governing $C(h)$ is $\epsilon - 4\gamma$. If $\epsilon - 4\gamma < 0$,

$$c_{11} = \frac{1}{2\epsilon\gamma(\epsilon-4\gamma)}\left((\epsilon-4\gamma)\left(\epsilon^2\sigma_1^2+\epsilon\gamma\sigma_1^2+\sigma_2^2\right)+\sinh(h)\left(-\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^{5/2}\sigma_1^2\sqrt{4\gamma-\epsilon}-\epsilon\gamma\sigma_1^2\sqrt{-\epsilon(\epsilon-4\gamma)}+\sigma_2^2\sqrt{-\epsilon(\epsilon-4\gamma)}\right)+\right.\right. \tag{259}$$

$$\left.\left.\epsilon\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^2\sigma_1^2-3\epsilon\gamma\sigma_1^2+\sigma_2^2\right)-4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right.\right.$$

$$\left.\left.+\cosh(h)\left(\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^{5/2}\sigma_1^2\sqrt{4\gamma-\epsilon}-\epsilon\gamma\sigma_1^2\sqrt{-\epsilon(\epsilon-4\gamma)}+\sigma_2^2\sqrt{-\epsilon(\epsilon-4\gamma)}\right)-\epsilon\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^2\sigma_1^2-3\epsilon\gamma\sigma_1^2+\sigma_2^2\right)+4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right)\right.$$

$$c_{12} = \frac{1}{2(\epsilon-4\gamma)}(-\sinh(h)\left(\epsilon^{3/2}\sigma_1^2\sqrt{4\gamma-\epsilon}\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)-\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^2\sigma_1^2-2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right)+2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right) \tag{260}$$

$$+\cosh(h)\left(\epsilon^{3/2}\sigma_1^2\sqrt{4\gamma-\epsilon}\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)-\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon^2\sigma_1^2-2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right)+2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right)+\epsilon\sigma_1^2(\epsilon-4\gamma))$$

$$c_{22} = \frac{1}{2(\epsilon-4\gamma)}\left(\epsilon^2\gamma\sigma_1^2-4\epsilon\gamma^2\sigma_1^2+\sinh(h)\left(-\sqrt{-\epsilon(\epsilon-4\gamma)}\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon\gamma\sigma_1^2-\sigma_2^2\right)+\epsilon\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right.\right. \tag{261}$$

$$\left.\left.-4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)+\cosh(h)\left(\sqrt{-\epsilon(\epsilon-4\gamma)}\sin\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon\gamma\sigma_1^2-\sigma_2^2\right)-\epsilon\cos\left(\frac{h}{\sqrt{-\frac{\epsilon}{\epsilon-4\gamma}}}\right)\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)+4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right)+\epsilon\sigma_2^2-4\gamma\sigma_2^2\right. \tag{262}$$

If $\epsilon - 4\gamma > 0$,

$$c_{11} = \frac{1}{2\epsilon\gamma(\epsilon-4\gamma)}\left((\epsilon-4\gamma)\left(\epsilon^2\sigma_1^2+\epsilon\gamma\sigma_1^2+\sigma_2^2\right)-\cosh(h)\left(\sinh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^{5/2}\sigma_1^2\sqrt{\epsilon-4\gamma}-\epsilon\gamma\sigma_1^2\sqrt{\epsilon(\epsilon-4\gamma)}+\sigma_2^2\sqrt{\epsilon(\epsilon-4\gamma)}\right)\right.\right. \tag{263}$$

$$\left.\left.+\epsilon\cosh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^2\sigma_1^2-3\epsilon\gamma\sigma_1^2+\sigma_2^2\right)-4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)+\sinh(h)\left(\sinh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^{5/2}\sigma_1^2\sqrt{\epsilon-4\gamma}-\epsilon\gamma\sigma_1^2\sqrt{\epsilon(\epsilon-4\gamma)}+\sigma_2^2\sqrt{\epsilon(\epsilon-4\gamma)}\right)\right.\right.$$

$$\left.\left.+\epsilon\cosh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^2\sigma_1^2-3\epsilon\gamma\sigma_1^2+\sigma_2^2\right)-4\gamma\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right.\right.$$

$$c_{12} = \frac{1}{2(\epsilon-4\gamma)}\left(-\cosh(h)\left(\epsilon^{3/2}\sigma_1^2\sqrt{\epsilon-4\gamma}\sinh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)+\cosh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^2\sigma_1^2-2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right)-2\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right)\right. \tag{264}$$

$$\left.+\sinh(h)\left(\epsilon^{3/2}\sigma_1^2\sqrt{\epsilon-4\gamma}\sinh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)+\cosh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon^2\sigma_1^2-2\epsilon\gamma\sigma_1^2+2\sigma_2^2\right)-2\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right)+\epsilon\sigma_1^2(\epsilon-4\gamma)\right.$$

$$c_{22} = \frac{1}{2(\epsilon-4\gamma)}\left(e^{-h}\left(\left(e^h(\epsilon-4\gamma)+4\gamma\right)\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)-\sqrt{\epsilon(\epsilon-4\gamma)}\sinh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon\gamma\sigma_1^2-\sigma_2^2\right)-\epsilon\cosh\left(\frac{h\sqrt{\epsilon-4\gamma}}{\sqrt{\epsilon}}\right)\left(\epsilon\gamma\sigma_1^2+\sigma_2^2\right)\right)\right) \tag{265}$$

Finally, if $\epsilon = 4\gamma$,

$$c_{11} = \frac{e^{-h}\left(4\gamma^2\left(-h(h+6)+10e^h-10\right)\sigma_1^2+\left(-h(h+2)+2e^h-2\right)\sigma_2^2\right)}{16\gamma^2} \tag{266}$$

$$c_{12} = \frac{e^{-h}\left(4\gamma^2\left(4e^h-(h+2)^2\right)\sigma_1^2-h^2\sigma_2^2\right)}{8\gamma} \tag{267}$$

$$c_{22} = \frac{1}{4}e^{-h}\left(4\gamma^2\left(-h(h+2)+2e^h-2\right)\sigma_1^2+\left(-((h-2)h)+2e^h-2\right)\sigma_2^2\right) \tag{268}$$

For the IZ model the latent parameter is $b^2+2bf+f^2+4dfg$, however, as mentioned in Appendix 4.4, these parameters are actually fixed in the model [12], and this latent

parameter takes only positive values. The expression for $C(h)$ is extremely long and is available online in the code. No closed form expression was found for the covariance for the AdEx model.