

# Fast-slow asymptotics for a Markov chain model of fast sodium current

Tomáš Starý, Vadim N. Biktashev<sup>1</sup>

*College of Engineering, Mathematics and Physical Sciences, University of Exeter, Harrison Building,  
North Park Road, Exeter EX4 4QF, United Kingdom*

(Dated: 23 August 2017)

We explore the feasibility of using fast-slow asymptotic to eliminate the computational stiffness of the discrete-state, continuous-time deterministic Markov chain models of ionic channels underlying cardiac excitability. We focus on a Markov chain model of the fast sodium current, and investigate its asymptotic behaviour with respect to small parameters identified in different ways.

Modern models of excitability of cardiac cells include description of ionic channels in terms of deterministic Markov chains. The transition rates in these chains often vary by several orders of magnitude, which makes numerical simulations more difficult and necessitates development of specialized numerical approaches. We follow the usual wisdom that the small parameters in a mathematical model can be turned from an impediment into an advantage by development of asymptotic description exploiting those small parameters. The usual problem with experiment-derived (rather than postulated) models is that small parameters in them are not defined a priori but need to be identified, and sometimes this can be done equally plausibly in more than one way. In this paper, we show how the standard fast-slow asymptotic theory can be applied to the Markov chain models of ionic channels using one important model of this class as an example. We explore three selected ways of identifying small parameters in this model, and investigate the factors on which the utility of resulting asymptotics depends.

## I. INTRODUCTION

The bioelectricity is one of the driving forces of our life. Our mind and body is a manifestation of complex dynamics of electric impulses that carry information and trigger reactions in different organs of our body. The pulses of electrical excitation in heart are responsible for starting a chain of reactions, resulting in contraction of the cardiac muscle, causing the blood circulation. Understanding the detailed mechanisms of formation and propagation of electrical excitation can help in treatment and prevention of cardiac diseases.

The direct experimentation with living systems is difficult and rises many ethical issues, hence a mathematical description provides a valuable tool to gain insight and understanding of the internal working of the heart.

The elementary part of a cardiac excitation model are models of the ion-specific channel in the membrane, that close or open in response to the change in the transmem-

brane voltage. On the molecular scale, functioning of a single channel is an inherently stochastic process, which is adequately described as continuous time Markov chain (or Markov processes, as they are sometimes called). For most applications, however, the stochastic component is not essential, and it is sufficient to describe the behaviour of the channel in terms of the deterministic “master equation” for the probabilities of the channel to be in certain states, as functions of time. Simulation of resulting excitation models for single cells does not create problems; but when scaled to the tissue or whole-organ level, this becomes computationally expensive<sup>1</sup>. The computer technology is constantly improving, and recent publications describe results that would be unthinkable a few years before: *e.g.* Richards *et al.*<sup>2</sup> simulated excitation in whole human heart with spatial resolution of 0.13 mm in nearly real time. However, that was done on a system with 1.6 million CPUs and a peak speed of 20 petaflops. For the moment, such systems are far from ubiquitous. It is therefore important to try and improve the computational methods for simulation of cardiac excitation models.

One significant factor of computational complexity is that the Markov chain models of ionic channels often involve processes on time scales differing by several orders of magnitude, i.e. are stiff. So a direct approach using explicit time steppers requires very small time steps, hence high computational demands. For instance, Bondarenko<sup>3</sup>, for a model involving a number of Markov chain channels, used time steps as short as 2 picoseconds—this is to be compared to the duration of the onset of an action potential of the order of 1 millisecond, and duration of the cardiac pulse of the order of 1 second. The natural alternative is generic implicit time steppers. However, approaches based on exploiting specific properties of cardiac excitation model present an attractive third possibility. In this paper, we explore one possible way to exploit specific properties of the Markov chain models of ionic channels. This is based on the traditional idea that a small parameter in the model can be turned from a hindrance into an advantage, by finding asymptotics in this parameter. For small parameters responsible for numerical stiffness, the adequate approach is singular perturbation theory, of fast-slow asymptotics.

The problem of stiffness of the description of the ion gates’ dynamics is not new for the Markov chain models, and was there already for the Hodgkin-Huxley (HH)

type “gate” models, starting from the Hodgkin and Huxley<sup>4</sup>. Out of many approaches to overcome stiffness in these models, arguably the most popular one follows the work by Rush and Larsen<sup>5</sup>. In some respects, it offers an “ideal” solution for the HH-type models, combining accuracy and stability. This approach exploits the fact that the equation controlling an HH-type gate is quasi-linear, in the sense that it is linear with respect to the gate variable, although it depends on other, “control”, variables in a non-linear way. Typically, the control variable is the transmembrane voltage, but in some channels it is concentration of Ca ions in addition or instead of the transmembrane voltage. If the control variables change negligibly during one time step, then “freezing” these variables allows one to write an “exact” solution. In the simplest formulation, the resulting computational scheme is first-order accurate, but the coefficient in the leading order term depends on the rate of the voltage (or whatever the control variable is) and does not depend on the stiffness of the equation with respect to the gate variable. This approach may be formalized as fast-slow approach, where the control variable is slow and gate variable is fast. In that case, the leading-order solution for the gate variable, the “instant equilibrium”, corresponds to the limit when the time step is much longer than the characteristic time constant defined by the current values of the transition rate. However, the Rush-Larsen scheme in fact retains its accuracy when the time step is much less than, or is comparable to this characteristic time.

The Rush-Larsen approach has been so popular, it inevitably spawned a number of attempts to improve and/or extend it. For instance, Perego and Veneziani<sup>6</sup> proposed how to increase its accuracy from first to second order, and Marsh, Ziaratgahi, and Spiteri<sup>7</sup> suggested how to apply it to equations which are not quasi-linear, by linearising them for the duration of the time step. The Markov chain description poses a different sort of challenge: the equations are quasi-linear with respect to the Markov state occupancies so there is no need in linearization, but instead of one equation for a gate variable, there is a system of simultaneous linear equations. The “straightforward” Rush-Larsen-type approach, when each of the equations of the system is considered in turn, by freezing all other variables together with the control variables, has been proven effective for a number of examples (see Ref. [1] and references therein). A more radical approach for Markov chain generalization of the Rush-Larsen scheme, which utilized the exact solution for the whole linear system, thus avoiding extra errors caused by freezing some Markov states while updating others, was described in Refs. 8 and 9.

The methods discussed above offer efficient numerical schemes, but do not exploit the fact that the different processes within the same Markov chain may, and often do, have vastly different speeds. Taking these into account can, at least theoretically, offer further advantages. There have been a number of inspiring examples

of this kind in literature. For instance, Hinch *et al.*<sup>10</sup> and Plank *et al.*<sup>1</sup> used the concept of “rapid equilibrium”, exploiting rapid transition rates between some of the states of a Markov chain, to effectively “merge” the closely connected states into one “combined” state, for a number of Markov chain models. This approach can be formalized as a leading-order asymptotic in the classical fast-slow perturbation theory descending from the works by Tikhonov<sup>11</sup> and Fenichel<sup>12</sup>, for a particular form in which a small parameter  $\varepsilon$  appears in the equations: as a factor  $1/\varepsilon$  in front of the transition rates between two selected Markov states.

In the present paper, we seek to analyse this sort of asymptotics in more detail, following the full formalism, rather than immediately getting to the answer by following the rather obvious, but still only intuitive “rapid equilibrium” argument. The motivation for a more detailed analysis includes possibilities of generalization of the asymptotic approach to the fast/slow separation cases other than pairs of fast reciprocal transition rates, and getting higher-order terms in asymptotics. Moreover, an important theoretical question is whether stiff Markov chain formulations of ionic channels can always be well described by the standard singular perturbation theory. The intrigue here comes from the fact that the Hodgkin-Huxley description of some cardiac channels, specifically the fast sodium current channel, may be only partly described by the Tikhonov asymptotics, in regards of the activation, ‘ $m$ ’-gate; whereas the attempts to treat the inactivation, ‘ $h$ ’-gate as fast or slow compared to the transmembrane voltage are ineffective, and at the very least fail to describe some essential qualitative features, such as the maximum of the action potential which is lower than the reversal potential for the sodium ions and which is different in a single cell in a propagating wave in tissue, to name the simplest example<sup>13–16</sup>. Hence the answer for the Markov chain model of the same channel is far from obvious *a priori* and requires investigation.

To address the theoretical question posed above, we have chosen the Markov chain model of the fast sodium current developed by Clancy and Rudy<sup>17</sup>. This is not the stiffest model of existing models of the kind, but it is one of the most popular ones. As our study was methodological rather than practical, it was important for our choice that the transition rates, and related characteristic times, in this model are varied in wide ranges which makes identification of small parameters a nontrivial issue. In other words, our aim was not to identify examples when the standard asymptotics can successfully treat the problem of stiffness (such examples do exist, see above), but rather analyse cases when the standard approach fails, as a necessary step towards developing more adequate, non-standard approaches. An extra motive for the choice of the fast Na current in this context was that the Hodgkin-Huxley description of this current is known to require non-Tikhonov asymptotics, as discussed above.

The structure of the paper is as follows. Section II introduces the notation and main principles of the sin-

gular perturbation approach we are using. Section III discusses amendments required of this approach with account of the specifics of Markov chain models. Section IV presents a formalization of the process of identification of small parameters in experiment-based models, which we call parametric embedding. Section V introduces the Markov chain model of the fast sodium current which we use to apply the singular perturbations. Section VI presents the main results, coming out of a few different parametric embeddings of this model. This is concluded by discussion in Section VII. We also present an Appendix containing technical material which is required for reproducing the main results but not for their understanding.

## II. GENERAL THEORY FOR DIMENSIONALITY REDUCTION

The singular perturbation theory is well known in a variety of different formulations. We mostly follow the terminology and notation used e.g. in Refs. 18–22, adjusting where necessary for our present purposes.

We consider an autonomous system of ordinary differential equations

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}) + \varepsilon \mathbf{h}(\mathbf{u}) \quad (1)$$

where  $\mathbf{u}, \mathbf{f}, \mathbf{h} \in \mathbb{R}^n$ , and  $\varepsilon$  is a small positive parameter. We assume existence of a stable  $m$ -dimensional manifold  $\{\mathbf{U}\}$  of equilibria of the unperturbed system,  $\varepsilon = 0$ , i.e.  $\mathbf{f}(\mathbf{U}) = \mathbf{0}$ , where  $\mathbf{0}$  stands for the null vector, with coordinates  $\mathbf{a} \in \mathbb{R}^m$ ,  $1 \leq m < n$ , and looking for solutions of the perturbed system,  $\varepsilon > 0$ , in the form

$$\mathbf{u} = \mathbf{U}(\mathbf{a}) + \varepsilon \mathbf{v}(t) \quad (2)$$

where the perturbation of the solution  $\mathbf{b} \in \mathbb{R}^n$  is orthogonal to the manifold, in the sense that

$$\mathbf{v}(t) = \sum_{\ell} b_{\ell} \mathbf{V}_{\ell}(\mathbf{a}), \quad (3)$$

where the vectors  $\mathbf{V}_i(\mathbf{a})$  are right eigenvectors of a Jacobian matrix  $\hat{F}(\mathbf{U}) = \partial \mathbf{f} / \partial \mathbf{u}|_{\mathbf{u}=\mathbf{U}}$ ,

$$\hat{F} \mathbf{V}_i = \Lambda_i \mathbf{V}_i, \quad (4)$$

and the summation index  $\ell$  runs through the stable eigenvalues,  $\text{Re}(\Lambda_{\ell}) < 0$ ,  $\ell = m+1, \dots, n$ , skipping the zero eigenvalues,  $\Lambda_k$ ,  $k = 1, \dots, m$ , corresponding to the directions tangent to the manifold. Table II summarises the meaning of these and other index conventions as used throughout the text, subject to a small amendment in the next section.

The right eigenvectors corresponding to zero eigenvalues  $\Lambda_k = 0$  are tangent to the invariant manifold and can be found as

$$\mathbf{V}_k = \frac{\partial \mathbf{U}}{\partial a_k}. \quad (5)$$

We substitute (2) and (3) into (1), expand the nonlinear functions into their Taylor series and separate the components using left eigenvectors  $\mathbf{W}_i^T$  as projectors. The detailed derivation is presented in the Appendix A. The final result reads as the following system of ODEs

$$\frac{1}{\varepsilon} \frac{da_k}{dt} = \mathbf{W}_k^T \mathbf{h}(\mathbf{U}) + \varepsilon \mathcal{F}_k(\mathbf{a}, \mathbf{b}) + \mathcal{O}(\varepsilon^2), \quad (6a)$$

$$\frac{db_{\ell}}{dt} = \Lambda_{\ell} b_{\ell} + \mathbf{W}_{\ell}^T \mathbf{h}(\mathbf{U}) + \mathcal{O}(\varepsilon), \quad (6b)$$

where in the right-hand side of the first equation we have kept the leading order term,  $\mathbf{W}_k^T \mathbf{h}(\mathbf{U})$ , and the first-order correction  $\mathcal{F}_i$ , which works out as

$$\begin{aligned} \mathcal{F}_i(\mathbf{a}, \mathbf{b}) = & \mathbf{W}_i^T \hat{H}(\mathbf{U}) \mathbf{v} + \mathbf{W}_i^T \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} \\ & + \frac{1}{\varepsilon} \sum_k \frac{\partial \mathbf{W}_i^T}{\partial a_k} \frac{da_k}{dt} \mathbf{v}, \end{aligned} \quad (7)$$

where

$$\hat{H}(\mathbf{U}) = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right|_{\mathbf{u}=\mathbf{U}}. \quad (8)$$

Equations (6a) and (6b) are coupled through higher-order terms, and to complete the reduction, we need to eliminate  $\mathbf{b}$ . For the solution of the manifold coordinates  $\mathbf{a}$  up to  $\mathcal{O}(\varepsilon^N)$  it is sufficient to find the correction term  $\mathbf{b}$  up to  $\mathcal{O}(\varepsilon^{N-1})$ . The leading order term for the correction  $\mathbf{b}$  in terms of  $\mathbf{a}$  can be found by solving (6b) using the integrating factor method. The solution also requires Taylor expansion of the integrating factor and of the non-homogeneous term. This leads to

$$b_{\ell} = -\frac{\mathbf{W}_{\ell}^T \mathbf{h}(\mathbf{U})}{\Lambda_{\ell}} + \mathcal{O}(\varepsilon), \quad (9)$$

which is to be substituted into the first-order term into (6a), which then becomes a closed equation for  $\mathbf{a}$ .

TABLE I. Ranges of indices used in the text, unless explicitly stated otherwise.

index	values	corresponds to
$i, j, j_1, j_2$	$\iota, \dots, n^a$	all eigenvalues of Jacobian (J.)
$k, p$	$\iota, \dots, m^a$	zero eigenvalues of J.
	$0^b$	autonomous time direction
$q, q'$	$1, \dots, n^b$	all eigenvalues of Markov chain (M.C.)
$r, r'$	$1, \dots, m^b$	zero eigenvalues of M.C. $\lambda_r = 0$
$\ell$	$m+1, \dots, n$	non-zero eigenvalues of J. or M.C.

<sup>a</sup> Here  $\iota = 1$  for Section II and  $\iota = 0$  from the next section on.

<sup>b</sup> This is used starting from Section III.

### III. DIMENSIONALITY REDUCTION FOR TIME-INHOMOGENEOUS MARKOV CHAINS

The master equation for Markov chain models of ionic channels can be written in the form

$$\frac{d\mathbf{x}}{dt} = \hat{A}(t)\mathbf{x}. \quad (10)$$

Entries in the vector of dynamical variables  $\mathbf{x} \in \mathbb{R}^n$  represent the probabilities, that an ion channel resides in a particular state. Entries of the transition matrix  $\hat{A} \in \mathbb{R}^{n \times n}$  describe the conditional probabilities of a channel in one given state to transit to another state per unit of time, i.e. transition rates. In reality, the matrix  $\hat{A}$  depends on other dynamic variables of the model, e.g. the transmembrane voltage, which in turn are affected by the dynamics of the Markov chain; however this is not essential for the formalism we describe here and we assume that  $\hat{A}$  is an explicit function of time, just for simplicity of notation. The sum of the entries in the vector of dynamical variables is equal to 1, i.e. it is a *stochastic vector*. This implies that the sum of the entries in each column of the transition matrix  $\hat{A}$  has to be equal to 0. This is achieved as the entries on the diagonal of the transition matrix are a sum of the entries out of the diagonal for each column of the matrix. This property together with the fact that the non-diagonal elements are non-negative constitutes the definition of  $\hat{A}$  as a *left-stochastic matrix*.

To use the theory described in Section II we have to take into account one simplifying fact and two complications. The simplifying fact is that the system (10) is linear. The complications are, firstly, that the theory described in the previous section applies to an autonomous system, but the Markov chain in (10) has an explicit time dependence of the transition matrix  $\hat{A}(t)$ . Secondly, the theory requires a small parameter, however the Markov chain models contain transition rates determined experimentally, and identification of any small parameters in such a case is a separate task, sometimes nontrivial.

The first complication is dealt with using autonomisation, which means that we introduce an additional dynamical variable  $\sigma$  to represent time (henceforth referred to as “*autonomous time*”). Then the vector of dynamical variables is

$$\mathbf{u} = \begin{bmatrix} \sigma \\ \mathbf{x} \end{bmatrix}$$

and the dynamic equation is

$$\frac{d}{dt} \begin{bmatrix} \sigma \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} 1 \\ \hat{A}(\sigma)\mathbf{x} \end{bmatrix}. \quad (11)$$

Note that system (11) is no longer linear unless the function  $\hat{A}(\sigma)$  is a constant.

To address the second complication, we introduce the small parameters artificially in an empirical procedure we call parametric embedding, which is discussed in detail in the next section. For now it is important that as a result,

we can split the transition rates matrix  $\hat{A}$  into a fast part  $\hat{A}_f$  and a slow part  $\hat{A}_s$ , and the difference between them is identified by the small parameter  $\varepsilon$  appearing as

$$\hat{A} = \frac{1}{\varepsilon} \hat{A}_f + \hat{A}_s. \quad (12)$$

We restrict consideration to the embeddings in which  $\hat{A}_f$  and  $\hat{A}_s$  are left-stochastic matrices. We assume that the fast matrix  $\hat{A}_f(\sigma)$  is diagonalizable, and introduce the eigenvalues  $\lambda_q(\sigma)$  and the right eigenvectors  $\boldsymbol{\kappa}_q(\sigma)$ :

$$\hat{A}_f(\sigma)\boldsymbol{\kappa}_q(\sigma) = \lambda_q(\sigma)\boldsymbol{\kappa}_q(\sigma) \quad (13)$$

(and drop from now on the dependence on  $\sigma$ , for brevity). We assume that for all  $\sigma$ , matrix  $\hat{M}_1$  has a full set of eigenvectors, the first  $m \geq 1$  of the eigenvalues are zero, and the remaining are all real (and of course negative)<sup>23</sup>. Correspondingly, we introduce also the left eigenvectors  $\boldsymbol{\rho}_q$ ,

$$\hat{A}_f^T \boldsymbol{\rho}_q = \lambda_q \boldsymbol{\rho}_q, \quad \boldsymbol{\rho}_q^T \boldsymbol{\kappa}_{q'} = \delta_{q,q'}.$$

Differentiation of the last identity with respect to  $\sigma$  yields a relationship that will be useful:

$$\frac{d\boldsymbol{\rho}_i^T}{d\sigma} \boldsymbol{\kappa}_j = -\boldsymbol{\rho}_i^T \frac{d\boldsymbol{\kappa}_j}{d\sigma}. \quad (14)$$

We transform the system (11) into fast time  $\tau = t/\varepsilon$  to get a system

$$\begin{aligned} \frac{d\sigma}{d\tau} &= \varepsilon, \\ \frac{d\mathbf{x}}{d\tau} &= \hat{A}_f(\sigma)\mathbf{x} + \varepsilon \hat{A}_s(\sigma)\mathbf{x}, \end{aligned} \quad (15)$$

This can be considered in the format of (1) with

$$\mathbf{u} = \begin{bmatrix} \sigma \\ \mathbf{x} \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} 0 \\ \hat{A}_f(\sigma)\mathbf{x} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 1 \\ \hat{A}_s(\sigma)\mathbf{x} \end{bmatrix}. \quad (16)$$

The dimensionality of the autonomized system (15) is  $n+1$ ; we keep the upper value of the corresponding indices as  $n$  but reserve the value 0 for the time variable  $\sigma$ ; this is where parameter  $\iota$ , designating the minimal value of the eigenvalues' indices in Table II, becomes 0.

The manifold of equilibria in this case is in fact a linear subspace of  $\mathbb{R}^{n+1}$  which is the hull of the one-dimensional subspace corresponding to the time coordinate  $\sigma$  and the kernel of the fast matrix  $\hat{A}_f$ :

$$\mathbf{U}(\sigma, \mathbf{a}) = \begin{bmatrix} a_0 \\ \mathbf{U}^x(\mathbf{a}) \end{bmatrix} \quad (17)$$

where  $a_0 = \sigma$ ,  $\mathbf{a} = [a_1, \dots, a_m]^T$  and

$$\mathbf{U}^x(\mathbf{a}) = \sum_r a_r \boldsymbol{\kappa}_r. \quad (18)$$

To construct the reduced system, we need to find the Jacobian of function  $\mathbf{f}$  and solve the eigenvalue problem. The Jacobian is easily found as

$$\hat{F}(\mathbf{U}) = \begin{bmatrix} 0 & \mathbf{0}^T \\ \frac{d\hat{A}_f}{d\sigma} \mathbf{U}^x & \hat{A}_f \end{bmatrix}. \quad (19)$$

Let us denote the components of the eigenvectors as

$$\mathbf{V}_i = \begin{bmatrix} \mu_i \\ \mathbf{V}_i^x \end{bmatrix}. \quad (20)$$

Substituting (19) and (20) into  $\Lambda_i \mathbf{V}_i = \hat{F} \mathbf{V}_i$ , we get

$$\Lambda_i \mu_i = 0, \quad (21a)$$

$$\Lambda_i \mathbf{V}_i^x = \frac{d\hat{A}_f}{d\sigma} \mathbf{U}^x \mu_i + \hat{A}_f \mathbf{V}_i^x. \quad (21b)$$

Let us consider separately the cases  $\mu_i = 0$  and  $\mu_i \neq 0$ . For  $\mu_i = 0$ , equation (21b) becomes

$$\hat{A}_f \mathbf{V}_q^x = \Lambda_q \mathbf{V}_q^x, \quad (22)$$

which is the definition of an eigenvalue problem for matrix  $\hat{A}_f$ , so we can take  $\mathbf{V}_q^x = \boldsymbol{\kappa}_q$  and  $\lambda_q = \Lambda_q$ , for  $q = 1, \dots, n$ , out of which the first  $m$  are zero eigenvalues. For  $r = 1, \dots, m$ , we have  $\Lambda_r = 0$ , and differentiation of (13) with respect to  $\sigma$  gives

$$\frac{d\hat{A}_f}{d\sigma} \boldsymbol{\kappa}_r = -\hat{A}_f \frac{d\boldsymbol{\kappa}_r}{d\sigma}. \quad (23)$$

We find one more eigenpair for the case  $\mu_0 \neq 0$ . Let us normalise the corresponding eigenvector so that  $\mu_0 = 1$ . Then to satisfy equation (21a) we must have  $\Lambda_0 = 0$ , and (21b) becomes

$$\hat{A}_f \mathbf{V}_0^x = -\frac{d\hat{A}_f}{d\sigma} \mathbf{U}^x. \quad (24)$$

If we substitute (18) into (24) and use (23), we get

$$\hat{A}_f \mathbf{V}_0^x = \hat{A}_f \sum_r a_r \frac{d\boldsymbol{\kappa}_r}{d\sigma} \quad (25)$$

so we can choose

$$\boldsymbol{\kappa}_0 = \sum_r a_r \frac{d\boldsymbol{\kappa}_r}{d\sigma} \quad (26)$$

(this is choice is of course non-unique because the zero eigenvalue has multiplicity  $m + 1$ ).

The left eigenvectors are treated similarly. To summarise the results, the eigenvalues  $\Lambda_i$  and eigenvectors  $\mathbf{V}_i$ ,  $\mathbf{W}_i$  of the Jacobian in the time-extended system are related to those  $\lambda_q$ ,  $\boldsymbol{\kappa}_q$ ,  $\boldsymbol{\rho}_q$  of the transition rate matrix via the following relationships:

$$\begin{aligned} \mathbf{V}_0 &= \begin{bmatrix} 1 \\ \sum_r a_r \frac{d\boldsymbol{\kappa}_r}{d\sigma} \end{bmatrix}, & \mathbf{V}_q &= \begin{bmatrix} 0 \\ \boldsymbol{\kappa}_q \end{bmatrix}, \\ \mathbf{W}_0 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, & \mathbf{W}_q &= \begin{bmatrix} -\boldsymbol{\rho}_q^T \sum_r a_r \frac{d\boldsymbol{\kappa}_r}{d\sigma} \\ \boldsymbol{\rho}_q \end{bmatrix}, \\ \Lambda_0 &= 0, & \Lambda_q &= \lambda_q. \end{aligned} \quad (27)$$

With these the time-component of the “leading-order term” works out as  $\mathbf{W}_0^T \mathbf{h}(\mathbf{U}) = 1$ , as should be expected. For the Markov chain subspace, we use (14), (16), (18) and (27) to get

$$\begin{aligned} \mathbf{W}_q^T \mathbf{h}(\mathbf{U}) &= \left[ \frac{d\boldsymbol{\rho}_q^T}{d\sigma} \mathbf{U}^x, \boldsymbol{\rho}_q^T \right] \begin{bmatrix} 1 \\ \hat{A}_s(\sigma) \mathbf{x} \end{bmatrix} \\ &= -\boldsymbol{\rho}_q^T \sum_r a_r \left( \frac{d\boldsymbol{\kappa}_r}{d\sigma} - \hat{A}_s(\sigma) \boldsymbol{\kappa}_r \right), \end{aligned}$$

and then (9) gives the formula for the components of the transversal correction,

$$b_\ell = \frac{\boldsymbol{\rho}_\ell^T}{\lambda_\ell} \sum_r a_r \left( \frac{d\boldsymbol{\kappa}_r}{d\sigma} - \hat{A}_s(\sigma) \boldsymbol{\kappa}_r \right) + \mathcal{O}(\varepsilon), \quad (28)$$

and the transversal correction itself as

$$\mathbf{v}^x = \sum_\ell b_\ell \boldsymbol{\kappa}_\ell = \sum_{\ell, r} \left[ \frac{a_r}{\lambda_\ell} \boldsymbol{\rho}_\ell^T \left( \frac{d\boldsymbol{\kappa}_r}{d\sigma} - \hat{A}_s(\sigma) \boldsymbol{\kappa}_r \right) \right] \boldsymbol{\kappa}_\ell. \quad (29)$$

Finally, the first-order accurate reduced system of ODEs is given by (6a), leading to

$$\frac{1}{\varepsilon} \frac{da_0}{dt} = 1 + \mathcal{O}(\varepsilon^2), \quad (30a)$$

$$\frac{1}{\varepsilon} \frac{da_r}{dt} = \left( \frac{d\boldsymbol{\rho}_r^T}{d\sigma} + \boldsymbol{\rho}_r^T \hat{A}_s(\sigma) \right) \sum_{r'} a_{r'} \boldsymbol{\kappa}_{r'} \quad (30b)$$

$$\begin{aligned} &+ \varepsilon \left( \boldsymbol{\rho}_r^T \hat{A}_s(\sigma) + \frac{d\boldsymbol{\rho}_r^T}{d\sigma} \right) \\ &\times \sum_{\ell, r'} \left[ \frac{a_{r'}}{\lambda_\ell} \boldsymbol{\rho}_\ell^T \left( \frac{d\boldsymbol{\kappa}_{r'}}{d\sigma} - \hat{A}_s(\sigma) \boldsymbol{\kappa}_{r'} \right) \right] \boldsymbol{\kappa}_\ell + \mathcal{O}(\varepsilon^2). \end{aligned}$$

This result can be written in the matrix form as

$$\frac{1}{\varepsilon} \frac{d\mathbf{a}}{dt} = \left( \hat{M}_0(\sigma) + \varepsilon \hat{M}_1(\sigma) \right) \mathbf{a} + \mathcal{O}(\varepsilon^2), \quad (31)$$

where

$$\hat{M}_0 = \left[ \hat{P}' + \hat{P} \hat{A}_s \right] \hat{K}, \quad (32)$$

$$\hat{M}_1 = \left[ \hat{P}' + \hat{P} \hat{A}_s \right] \hat{L} \left[ \hat{K}' - \hat{A}_s \hat{K} \right], \quad (33)$$

$$\begin{aligned} \hat{K}(\sigma) &= [\boldsymbol{\kappa}_1 | \dots | \boldsymbol{\kappa}_m] = \begin{bmatrix} \kappa_1^1 & \dots & \kappa_m^1 \\ \vdots & \ddots & \vdots \\ \kappa_1^m & \dots & \kappa_m^m \end{bmatrix} \in \mathbb{R}^{n \times m}, \\ \hat{P}(\sigma) &= \begin{bmatrix} \boldsymbol{\rho}_1^T \\ \vdots \\ \boldsymbol{\rho}_m^T \end{bmatrix} = \begin{bmatrix} \rho_1^1 & \dots & \rho_1^m \\ \vdots & \ddots & \vdots \\ \rho_m^1 & \dots & \rho_m^m \end{bmatrix} \in \mathbb{R}^{m \times n}, \\ \hat{L}(\sigma) &= \sum_\ell \boldsymbol{\kappa}_\ell (\lambda_\ell)^{-1} \boldsymbol{\rho}_\ell^T \in \mathbb{R}^{n \times n}. \end{aligned} \quad (34)$$

and dash ' stands for differentiation with respect to  $\sigma$ .

#### IV. PARAMETRIC EMBEDDING

To address the second complication, we introduce the small parameters artificially in a procedure known as parametric embedding, previously introduced in Refs. 19, 21, and 24. This procedure is a formalization of the replacement of a small constant with a small parameter.

**Definition 1** *We will call a system*

$$\dot{u} = F(u; \varepsilon), \quad u \in \mathbb{R}^d,$$

*depending on parameter  $\varepsilon$ , a one-parametric embedding of a system*

$$\dot{u} = f(u), \quad u \in \mathbb{R}^d,$$

*if  $f(u) \equiv F(u, 1)$  for all  $u \in \text{dom}(f)$ . If the limit  $\varepsilon \rightarrow 0$  is concerned then we call it an asymptotic embedding.*

The typical use of this procedure has the form of a replacement of a small constant with a small parameter. If a system contains a dimensionless constant  $a$  which is “much smaller than 1”, then replacement of  $a$  with  $\varepsilon a$  constitutes a 1-parametric embedding; and then the limit  $\varepsilon \rightarrow 0$  can be considered. In practice, constant  $a$  would more often be replaced with parameter  $\varepsilon$  rather than multiplied by it, but mathematically speaking, in the context of  $\varepsilon \rightarrow 0$  and  $a = \text{const} \neq 0$ , these two ways are formally equivalent. *This explains the paradoxical use of a zero limit for a parameter whose true value is one.*

In some applications, the “small parameters” appear naturally and are readily identified. However, this is not always the case, and in complex systems identification of adequate small parameters may be a task in itself, which is where the formalization of this procedure can be helpful. In the context of the definition above, it is important to understand that there are infinitely many ways a given system can be parametrically embedded, as there are infinitely many ways to draw a curve  $F(u; \varepsilon)$  in the functional space given the only constraint that it passes through a given point,  $F(u; 1) = f(u)$ . In terms of asymptotics, which of the embeddings is “better” depends on the qualitative features of the original systems that need to be represented, or classes of solutions that need to be approximated.

If a numerical solution of the system can be found easily, then there is a simple practical recipe: to look at the solutions of the embedding at different, progressively decreasing values of the artificial small parameter  $\varepsilon$ , and see when the features of interest will start to converge. If the convergent behaviour is satisfactorily similar to the original system with  $\varepsilon = 1$ , the embedding is adequate for these features.

To summarize, we claim that identification of small parameters in a given mathematical model with experimentally measured functions and constants will, from the formal mathematical viewpoint, always be arbitrary, even though in the simplest cases the choice may be so

natural that that this ambiguity is not even realized by the modeller, and that “validity” of such identification can be defined only empirically: if the asymptotics describe the required class of solutions sufficiently well. The exceptions may be when the asymptotic series are in fact convergent, the approximation errors can be estimated *a priori*, but this is rare.

In the subsequent text, slightly abusing the above definition for the sake of brevity, we refer as “embedding” to particular instances of one-parametric embedding of a given system for a selected value of the parameter  $\varepsilon$ . The overall structure of the embeddings is always like in (12), and the difference is in the choice of the matrices  $\hat{A}_f$  and  $\hat{A}_s$ .

#### V. DEFINITION OF THE MARKOV CHAIN MODEL OF THE FAST SODIUM CURRENT (INA)

We apply the asymptotic theory described above to a Markov chain model of the fast sodium current developed by Clancy and Rudy<sup>17</sup> (we consider the wild-type version). The relevant part of the model has the form

$$\frac{dV_m}{dt} = -g_{\text{Na}} [V_m - E_{\text{Na}}(\mathbf{X})] O - \sum_{\ell} I_{\ell}(\mathbf{X}), \quad (35a)$$

$$\frac{d\mathbf{x}}{dt} = \hat{A}(V_m)\mathbf{x}, \quad (35b)$$

$$\frac{d\mathbf{X}}{dt} = \dots \quad (35c)$$

where  $V_m$  is the transmembrane voltage,  $g_{\text{Na}}$  is the maximal conductance of the fast Na current,  $E_{\text{Na}}$  is the reversal potential of the Na ions due to the transmembrane difference in the concentration of these ions,  $O$  is one of the components of the vector  $\mathbf{x}$  representing the fraction of open channels, corresponding to the open state of the fast Na current channels,  $I_{\ell}$  represent all other transmembrane currents, and the vector  $\mathbf{X}$  comprises all other dynamic variables of the model, such as other ionic channels, concentrations etc.

Fig. 1(a) shows the diagram of the Markov chain. We find it more convenient to rename the dynamic variables, *i.e.* the names of the states of the Markov chain, as reported in Ref. [8]: these are single-letter names, as opposed to the original names in Ref. [17] which use up to three symbols. The only state in the model that corresponds to the the channel being open is  $O$ , and this name coincides with the nomenclature used by Clancy and Rudy. So, for this Markov chain we have  $n = 9$  and the state vector

$$\mathbf{x} = [O, P, Q, R, S, T, U, V, W]^T.$$

According to the diagram of fig. 1(a), the transition rate

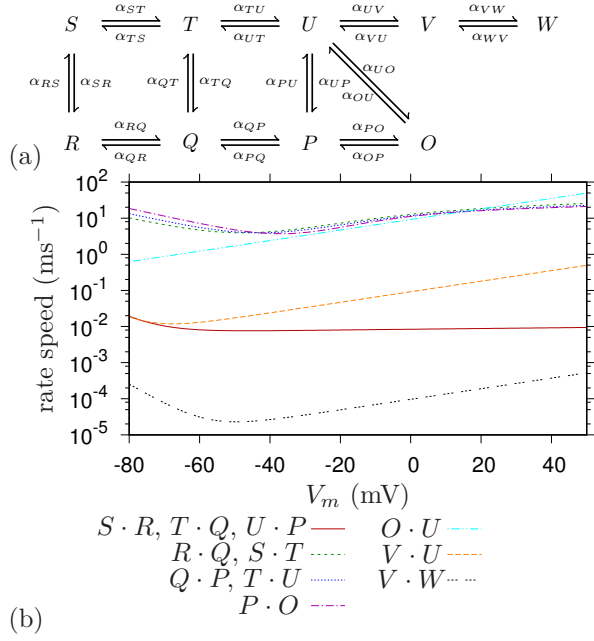


FIG. 1. (a) Diagram of  $I_{Na}$  channel and (b) speed of transition rates in the range of physiological cell membrane voltages. Here  $S \cdot R = \bar{S}R + R\bar{S}$  etc are reciprocal transition rates, defining the speed with which the two given states tend to equilibrate with each other.

matrix has the structure

$$\hat{A} = \begin{bmatrix} * & \vec{PO} & 0 & 0 & 0 & 0 & \vec{UO} & 0 & 0 \\ \vec{OP} & * & \vec{QP} & 0 & 0 & 0 & \vec{UP} & 0 & 0 \\ 0 & \vec{PQ} & * & \vec{RQ} & 0 & \vec{TQ} & 0 & 0 & 0 \\ 0 & 0 & \vec{QR} & * & \vec{SR} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \vec{RS} & * & \vec{TS} & 0 & 0 & 0 \\ 0 & 0 & \vec{QT} & 0 & \vec{ST} & * & \vec{UT} & 0 & 0 \\ \vec{OU} & \vec{PU} & 0 & 0 & 0 & \vec{TU} & * & \vec{VU} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vec{UV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix}. \quad (36)$$

Here and elsewhere in transition rates matrices, in the interests of saving space, we do not show diagonal elements and replace them with \*: they are uniquely defined by the condition that the sum of elements in each row should vanish. So *e.g.* the top left diagonal element in the above matrix is  $-\vec{PO} - \vec{UO}$ , and the bottom right element is  $-\vec{VW}$ .

All the transition rates in  $\hat{A}$  are functions of the transmembrane voltage  $V_m$ . Their exact definitions can be found in the original publication<sup>17</sup> (see also Refs. 8 and 25) and we do not present them here; however fig. 1(b) gives a graphical illustration of the magnitudes of these rates in the physiological range of  $V_m$ . In that figure, we

use the sum of the transition rates between two states as the measure of the speed of their connection, *i.e.*  $i \cdot j \triangleq \vec{ij} + \vec{ji}$ . Indeed, it is this quantity that determines the speed with which the dynamic equilibrium between the two states is reached if occupancies of all other states are fixed.

## VI. EMBEDDINGS OF THE INA MODEL

Fig. 1(b) allows one to see what transition rates may be considered “fast” and thus included into  $\hat{A}_f$ . For instance, connections  $PO$ ,  $QP$ ,  $TU$  are relatively fast in the whole range of voltages, connection  $VW$  is always slow, whereas connection  $OU$  is fast at high  $V_m$  but not so high at low  $V_m$ , and connection  $VU$  is somewhat intermediate between the group of clearly fast connection and the group of clearly slow connections. In accordance with the above discussed formal definition and informal semantics of the concept of embedding, we intend to treat the question of which connections can or should be considered fast as strictly empirical, so Fig. 1(b) does not provide the ultimate answer to this question, but merely the possible directions of search. For simplicity, we always construct  $\hat{A}_f$  by including into it reciprocal pairs of transition rates, and completing the diagonal elements to ensure  $\hat{A}_f$  is left-stochastic. As a result, any nonzero non-diagonal element of  $\hat{A}$  is always included into  $\hat{A}_f$  and/or into  $\hat{A}_s$ , which guarantees that  $\hat{A}_f$  and  $\hat{A}_s$  are left-stochastic. We assess the quality of an embedding by how well it approximates the transients of the Markov states in a typical solution, a standard action potential; and of course of all the states the most important is the open state  $O$ .

We have tried a number of different combinations of reciprocal transition rates for  $\hat{A}_f$ . Not all such combinations pass the embedding test, *i.e.* give reasonable approximation of the original solution in the limit  $\varepsilon \rightarrow 0$ . In particular, the would-be “straightforward” solution to consider as “fast” all the transition rates that appear so in fig. 1(b), does not work<sup>25</sup>. Figure 2 shows results of simulation of some of the more successful of those combinations. The  $I_{Na}$  model was extracted from the authors code<sup>17</sup>. The simulation of the model were driven by recorded values of  $V_m(t)$  during a standard action potential from a single-cell simulation. That means, we have performed a simulation of the original full model (35) once, and the resulting function  $V_m(t)$  was then used for computations of the embedded version of only the subsystem (35b), in which  $V_m(t)$  was considered given and fixed; in other words, performed “virtual voltage clamp” experiments. The time step in the simulation of  $I_{Na}$  was  $\Delta_t = 1 \mu s$ . The original model is shown with red lines, the transition rates embeddings are shown for a value of  $\varepsilon = 0.1$ .

As can be deduced from the figure, although the transition rates included in the embeddings have roughly the

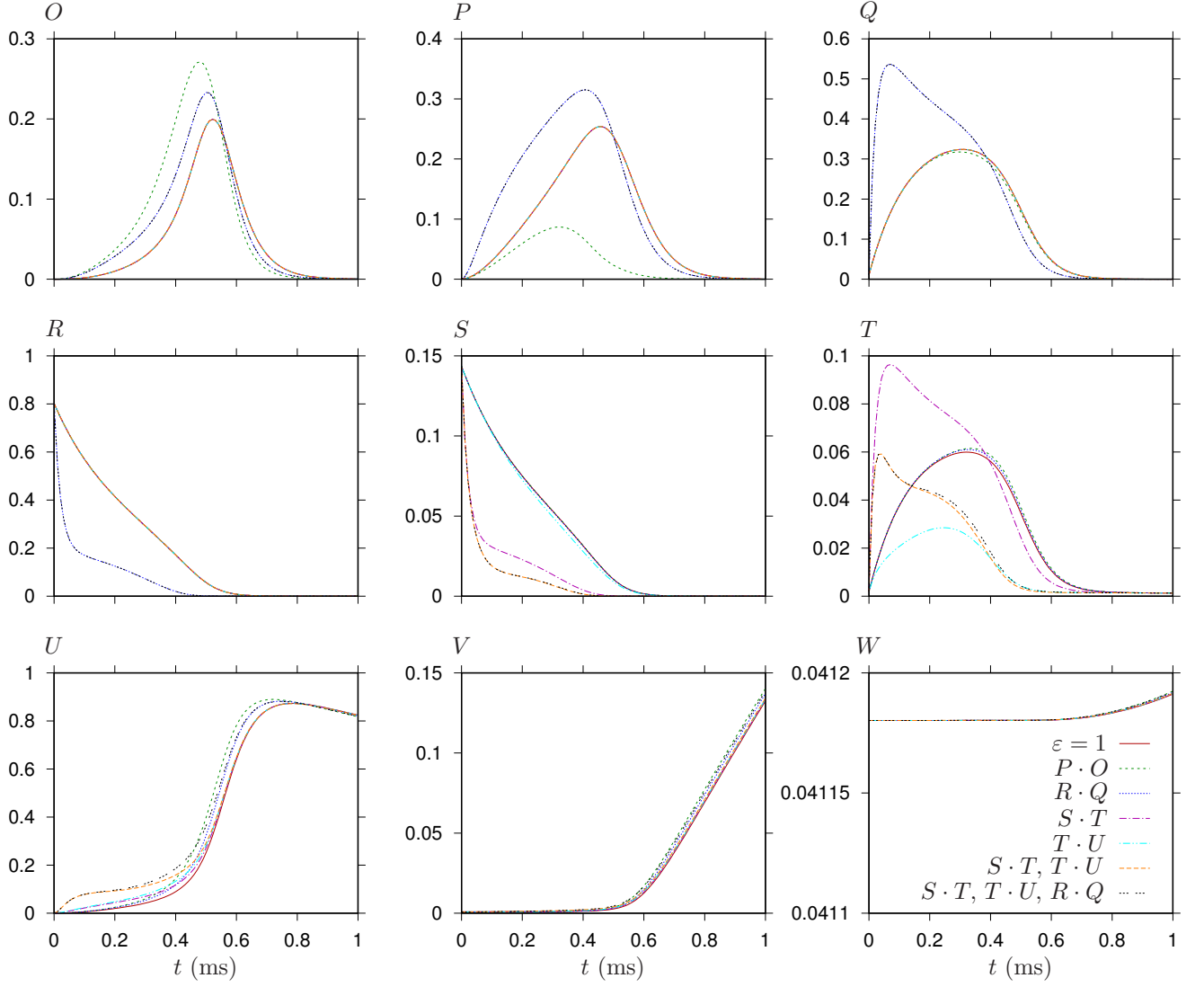


FIG. 2. Time evolution of state occupancy of  $I_{Na}$  with transition rates embeddings. The vertical axis shows the occupancy of the states from  $O$ – $W$  in alphabetical order as specified in top left corner. The red lines show the original model ( $\varepsilon = 1$ ), other lines show the embeddings as specified in the legend of panel  $W$ , *e.g.* the green lines show the embeddings of transition rates between states  $OP$ , *i.e.* both  $\vec{OP}$  and  $\vec{PO}$  for  $\varepsilon = 0.1$ , the grey line shows the embedding of the reciprocal transition rates between  $ST$ ,  $TU$  and  $RQ$ , etc.

same orders of magnitude, their expected effect on the accuracy of approximation of the  $O$  transient by asymptotic methods is rather different: the  $OP$  embedding is relatively poor, the  $RQ$  is somewhat better, whereas embeddings involving transitions between  $S$ ,  $T$  and  $U$ , any pair or all three, promises very good accuracy: the corresponding graphs are indistinguishable in the plot resolution. Note that this assessment heavily depends on the special role of the  $O$  state, and would be completely different if we were more interested in another Markov state. For instance, for the  $S(t)$  transient, the  $OP$  embedding promises good accuracy, and  $STU$  embedding is very poor. Obviously, it matters how close are the embedded rates to the state in question.

### A. OP-embedding

In this section we develop an example of a particular embedding of the transition rates between the states  $O$  and  $P$ , *i.e.* rates  $\vec{OP}$  and  $\vec{PO}$ . As seen from the above discussion, the empirical evidence suggests that the asymptotics of this embedding is not likely to give a good approximation, so the purpose of this exercise is mainly didactic, to demonstrate in detail the application of the general theory, including the first-order correction, on a simple example.

In this embedding, the transition matrix  $\hat{A}$  is split according to (12) into the matrix of the slow transition



rates

$$\hat{A}_s = \begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 & \vec{UO} & 0 & 0 \\ 0 & * & \vec{QP} & 0 & 0 & 0 & \vec{UP} & 0 & 0 \\ 0 & \vec{PQ} & * & \vec{RQ} & 0 & \vec{TQ} & 0 & 0 & 0 \\ 0 & 0 & \vec{QR} & * & \vec{SR} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \vec{RS} & * & \vec{TS} & 0 & 0 & 0 \\ 0 & 0 & \vec{QT} & 0 & \vec{ST} & * & \vec{UT} & 0 & 0 \\ \vec{OU} & \vec{PU} & 0 & 0 & 0 & \vec{TU} & * & \vec{VU} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vec{UV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix} \quad (37)$$

and the matrix of the fast transition rates

$$\hat{A}_f = \begin{bmatrix} * & \vec{PO} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vec{OP} & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (38)$$

For the dimensionality reduction we need to calculate the eigenvalues and eigenvectors of the fast matrix  $\hat{A}_f$ . This will result to a number of zero eigenvalues corresponding to the zero part of the matrix. There will be also at least one zero eigenvalue  $\lambda_1 = 0$  corresponding to the Markov chain since

$$\mathbf{1}^T \hat{A} = 0. \quad (39)$$

In fact, we have  $\lambda_r = 0$  for  $r = 1, \dots, m$ , where  $m = 8$ , and just one non-zero eigenvalue  $\lambda_9 = -(\vec{PO} + \vec{OP})$ . The corresponding right eigenvectors are

$$\begin{aligned} \kappa_1 &= (\vec{PO} + \vec{OP})^{-1} \left( \vec{PO} \mathbf{e}_1 + \vec{OP} \mathbf{e}_2 \right), \\ \kappa_i &= \mathbf{e}_{i+1}, \quad i = 2, \dots, 8, \\ \kappa_9 &= -\mathbf{e}_1 + \mathbf{e}_2, \end{aligned} \quad (40)$$

where  $\mathbf{e}_i$  is the standard notation for the column-vector which has  $i$ -th component equal to one and all other components equal to zero, and the left eigenvectors are

$$\begin{aligned} \rho_1 &= \mathbf{e}_1 + \mathbf{e}_2, \\ \rho_i &= \mathbf{e}_{i+1}, \quad i = 2, \dots, 8, \\ \rho_9 &= (\vec{PO} + \vec{OP})^{-1} \left( -\vec{OP} \mathbf{e}_1 + \vec{PO} \mathbf{e}_2 \right). \end{aligned} \quad (41)$$

We note that the left eigenvector  $\mathbf{1}$  asserted by the identity (39) is a linear combination of these, namely

$\mathbf{1} = \sum_{i=1}^8 \rho_i$ . The choice of normalization for  $\kappa_1$  and  $\rho_1$  is motivated by the ease of interpretation of the slow variable  $a_1$ , which will transpire shortly below.

Now we are ready to substitute the specifics of the selected embedding into the equation (30b) describing the reduced model. The left eigenvectors are constant for all  $r$ , so their derivatives are zero, and  $\{\ell\} = \{9\}$ . Then upon substituting (28) into (30b) we get

$$\begin{aligned} \frac{1}{\varepsilon} \frac{da_r}{dt} &= \rho_r^T \hat{A}_s(\sigma) \sum_{r'} a_{r'} \kappa_{r'} \\ &+ \varepsilon \left[ \rho_r^T \hat{A}_s(\sigma) \kappa_9 b_9 \right] + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (42)$$

The differential equation for  $a_r$  for  $r = 3, 4, 5, 7, 8$  come out identical to the equations for the states  $R, S, T, V, W$  from (36). This is because the first-order term vanishes as  $\rho_r^T \hat{A}_s(\sigma) \kappa_9 = 0$  for these  $r$ . Hence we retain the same names for the corresponding components of the reduced model, as they had in the original model, and the vector of dynamic variables in the reduced system has the form

$$\mathbf{a} = [\tilde{N}, \tilde{Q}, R, S, T, \tilde{U}, V, W]^T, \quad (43)$$

where  $\tilde{N} \triangleq a_1$ ,  $\tilde{Q} \triangleq a_2$  and  $\tilde{U} \triangleq a_6$ .

The components  $r = 1, 2, 6$  in (42), that is differential equations for  $a_1 = \tilde{N}$ ,  $a_2 = \tilde{Q}$  and  $a_6 = \tilde{U}$ , will have nonzero first-order terms. According to  $\rho_1$  as given by (41), the new variable  $\tilde{N}$  is just a sum of the old states occupancies  $O$  and  $P$ ; this is where the chosen normalization for  $\rho_1$  comes helpful. The names of the slow variables  $\tilde{Q}$  and  $\tilde{U}$  are motivated by the fact that according to (41) they map exactly to  $Q$  and  $U$  respectively, and the difference from the old variables is only in the first-order corrections in the reduced differential equations they obey.

Equation (18) then defines the relationship between the original and the reduced variables in the leading order, which in our case is

$$\mathbf{U}^x = [\beta_{PO} \tilde{N}, \beta_{OP} \tilde{N}, \tilde{Q}, R, S, T, \tilde{U}, V, W], \quad (44)$$

where we define the fractions of the transition rates as

$$\beta_{ij} = \frac{\vec{ij}}{\vec{PO} + \vec{OP}}. \quad (45)$$

We have only one stable eigenvalue in the present case, so equation (3) reduces to

$$\mathbf{v}^x = b_9 \kappa_9, \quad (46)$$

and equation (28), with account of  $\kappa_{r'} = 0$ ,  $r' \neq 1$ , gives

$$b_9 = \frac{\rho_9^T}{\lambda_9} \left( a_1 \kappa_1' - \hat{A}_s(\sigma) \mathbf{U}^x \right). \quad (47)$$

Then the leading-order transition matrix, according to (32,34) is

$$\hat{M}_0 = \begin{pmatrix} * & \vec{Q}\vec{P} & 0 & 0 & 0 & (\vec{U}\vec{O} + \vec{U}\vec{P}) & 0 & 0 \\ \vec{N}\vec{Q} & * & \vec{R}\vec{Q} & 0 & \vec{T}\vec{Q} & 0 & 0 & 0 \\ 0 & \vec{Q}\vec{R} & * & \vec{S}\vec{R} & 0 & 0 & 0 & 0 \\ 0 & 0 & \vec{R}\vec{S} & * & \vec{T}\vec{S} & 0 & 0 & 0 \\ 0 & \vec{Q}\vec{T} & 0 & \vec{S}\vec{T} & * & \vec{U}\vec{T} & 0 & 0 \\ \vec{N}\vec{U} & 0 & 0 & 0 & \vec{T}\vec{U} & * & \vec{V}\vec{U} & 0 \\ 0 & 0 & 0 & 0 & 0 & \vec{U}\vec{V} & * & \vec{W}\vec{V} \\ 0 & 0 & 0 & 0 & 0 & 0 & \vec{V}\vec{W} & * \end{pmatrix}, \quad (48)$$

where the new transition rates are defined as

$$\begin{aligned} \vec{N}\vec{Q} &= \beta_{OP} \vec{P}\vec{Q}, \\ \vec{N}\vec{U} &= \beta_{PO} \vec{O}\vec{U} + \beta_{OP} \vec{P}\vec{U}, \end{aligned} \quad (49)$$

and the first-order correction to the transition matrix defined by (33,34) works out as

$$\begin{aligned} \hat{M}_1 &= \left[ \beta_{OP} \beta_{PO} (\beta_{PQ} + \beta_{PU} - \beta_{OU}) - \frac{d\beta_{PO}/d\sigma}{\vec{P}\vec{O} + \vec{O}\vec{P}} \right] \\ &\times (\vec{P}\vec{Q} + \vec{P}\vec{U} - \vec{O}\vec{U}) \mathbf{e}_1 \mathbf{e}_1^T \\ &- (\vec{P}\vec{Q} + \vec{P}\vec{U} - \vec{O}\vec{U}) \beta_{PO} \beta_{QP} \mathbf{e}_1 \mathbf{e}_2^T \\ &- (\vec{P}\vec{Q} + \vec{P}\vec{U} - \vec{O}\vec{U}) (\beta_{PO} \beta_{UP} - \beta_{OP} \beta_{UO}) \mathbf{e}_1 \mathbf{e}_6^T \\ &+ \left[ \frac{d\beta_{PO}/d\sigma}{\vec{P}\vec{O} + \vec{O}\vec{P}} - \beta_{OP} \beta_{PO} (\beta_{PQ} + \beta_{PU} - \beta_{OU}) \right] \\ &\times \vec{P}\vec{Q} \mathbf{e}_2 \mathbf{e}_1^T \\ &+ \vec{P}\vec{Q} \beta_{PO} \beta_{QP} \mathbf{e}_2 \mathbf{e}_2^T \\ &+ \vec{P}\vec{Q} (\beta_{PO} \beta_{UP} - \beta_{OP} \beta_{UO}) \mathbf{e}_2 \mathbf{e}_6^T \\ &+ \left[ \frac{d\beta_{PO}/d\sigma}{\vec{P}\vec{O} + \vec{O}\vec{P}} - \beta_{OP} \beta_{PO} (\beta_{PQ} + \beta_{PU} - \beta_{OU}) \right] \\ &\times (\vec{P}\vec{U} - \vec{O}\vec{U}) \mathbf{e}_6 \mathbf{e}_1^T \\ &+ (\vec{P}\vec{U} - \vec{O}\vec{U}) \beta_{PO} \beta_{QP} \mathbf{e}_6 \mathbf{e}_2^T \\ &+ (\vec{P}\vec{U} - \vec{O}\vec{U}) (\beta_{PO} \beta_{UP} - \beta_{OP} \beta_{UO}) \mathbf{e}_6 \mathbf{e}_6^T. \end{aligned}$$

The Markov chain of the  $I_{Na}$  channel is linked to the rest of the cell excitability model via the state  $O$  which is the probability of the channel being open, so we need to compute  $O$  in terms of the new dynamic variables. This is obtained from

$$\mathbf{x} = \mathbf{U}^x + \varepsilon \mathbf{v}^x,$$

where  $\mathbf{U}^x$  is given by (44) and  $\mathbf{v}^x$  is given by (46), with (47) giving  $b_9$ . This leads to

$$O = \frac{\vec{P}\vec{O}}{\vec{P}\vec{O} + \vec{O}\vec{P}} \tilde{N} - \varepsilon b_9, \quad (50)$$

where

$$\begin{aligned} b_9 &= \left( \vec{P}\vec{O} + \vec{O}\vec{P} \right)^{-1} \frac{d\beta_{PO}}{d\sigma} \tilde{N} \\ &- \beta_{OP} \beta_{PO} (\beta_{PQ} + \beta_{PU} - \beta_{OU}) \tilde{N} \\ &+ \beta_{PO} \beta_{QP} \tilde{Q} + (\beta_{PO} \beta_{UP} - \beta_{OP} \beta_{UO}) \tilde{U}. \end{aligned} \quad (51)$$

Matrix  $\hat{M}_1$  and coordinate  $b_9$  depend on time derivatives of the transition rates, which in fact depend on the transmembrane voltage, hence the time derivative are to be calculated by the chain rule, *e.g.*

$$\begin{aligned} \frac{d}{d\sigma} \left( \vec{O}\vec{P} \right) &= \frac{d}{dV_m} \left( \vec{O}\vec{P} \right) \frac{dV_m}{dt}, \\ \frac{d}{d\sigma} \left( \vec{P}\vec{O} \right) &= \frac{d}{dV_m} \left( \vec{P}\vec{O} \right) \frac{dV_m}{dt}. \end{aligned}$$

Figure 3 shows the simulation results in the  $OP$ -embeddings and corresponding  $OP$ -reduction. The results should be compared against the original model shown by red lines. The simulations were done using extracted Markov chain model of  $I_{Na}$  driven by recordings of membrane voltage from whole cell simulations saved every 0.01 ms and interpolated as necessary. The state  $O$  in reduced model was computed using (50,51). For comparison of the reduced model with the embedding, the occupancy of state  $\tilde{N}$  in the embedded model was found as  $\tilde{N} = O + P$ .

The simulations with leading order approximation (blue lines) show relatively large deviation from the original model. The first-order accurate asymptotic model computed for  $\varepsilon = 0.5$  (magenta lines) provides better approximation than only the leading order term, however the state  $O$  in this approximation goes below zero, which does not make sense physically, as it represents a probability, so should be in the interval  $[0, 1]$ : note that the generic asymptotic theory does not take into account these specifics.

The panel (d) shows the error norms computed using the following formula

$$\| \tilde{N} - \tilde{N}_{\text{ref}} \| = \left[ \int_0^{t_{\text{max}}} (\tilde{N}(t) - \tilde{N}_{\text{ref}}(t))^2 dt \right]^{1/2} \quad (52)$$

where  $\tilde{N}_{\text{ref}}$  is the reference solution obtained for a very small time step, and comparison is done for the interval of  $t_{\text{max}} = 2$  ms of time-evolution. The error norms increase monotonically with  $\varepsilon$  and show the convergence for the leading-order and first-order approximations as expected, which confirms the correctness of the formulas.

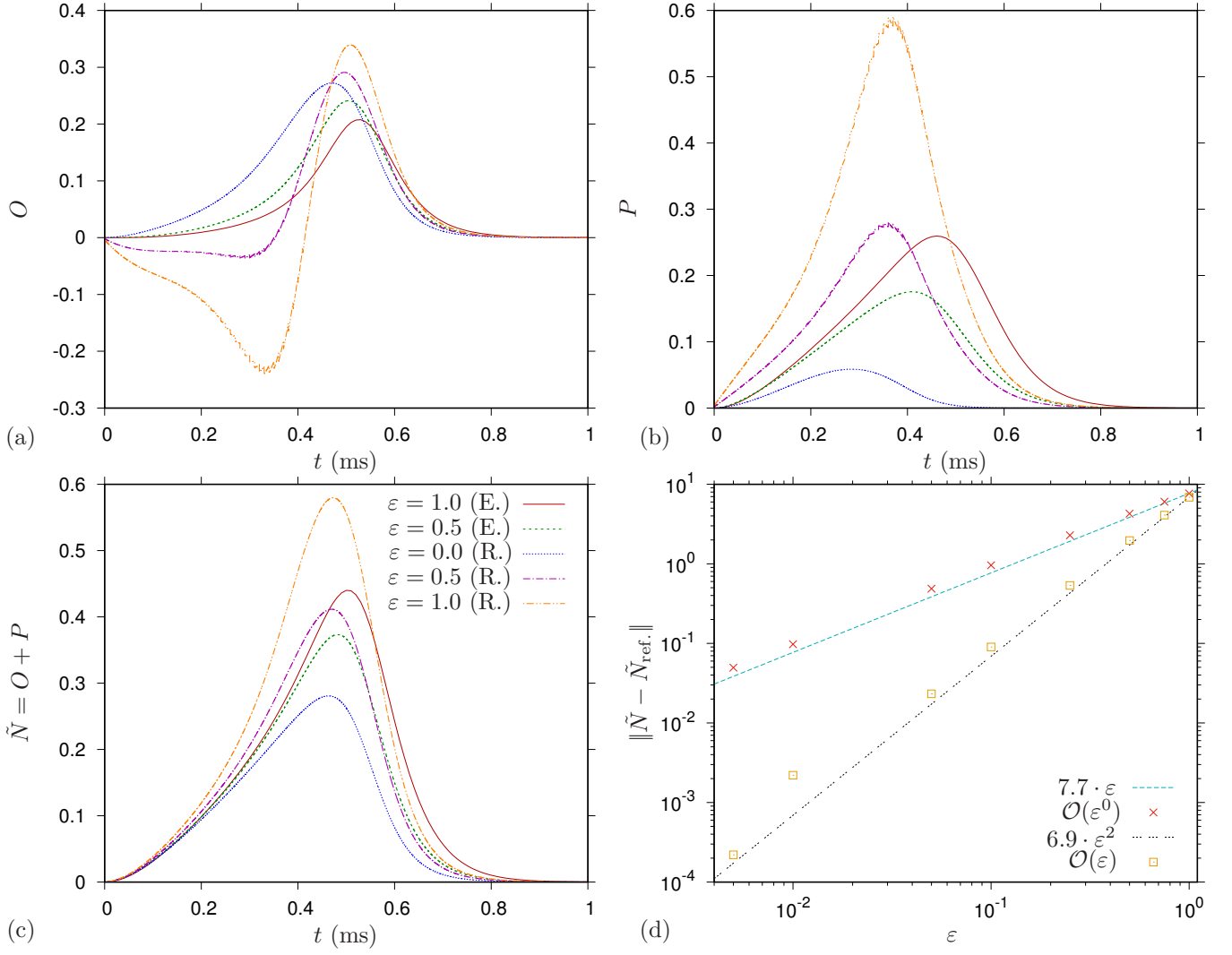


FIG. 3. (a-c) Evolution of state occupancy in  $OP$ -embedded (E.) and  $OP$ -reduced (R.) model, and (d) error analysis of  $OP$ -reduced model. State  $O$  occupancy (a), state  $P$  occupancy (b), and state  $\tilde{N}$  occupancy (c). The key in (c) applied to plots (a-c): the original model is denoted by red lines, the  $OP$ -embedded model  $\epsilon = 0.5$  is shown with green lines, the reduced model without correction term ( $\epsilon = 0$ ) is shown with blue lines, the  $OP$ -reduced model with correction term for  $\epsilon = 0.5$  is shown with magenta lines, and the  $OP$ -reduced model with correction term for  $\epsilon = 1.0$  is shown with orange lines. Panel (D) shows the order of approximation in  $\epsilon$  for the leading-order reduced model (red crosses), first-order reduced model (yellow squares). The norms were computed as a difference between the simulations of  $\tilde{N}$  at time step  $\Delta t = 0.01$  ms and simulations of  $\tilde{N}_{\text{ref}} = O + P$  computed with a time step of  $\Delta t = 5 \cdot 10^{-5}$  ms in the original model using the same value of  $\epsilon$ . The cyan and grey straight lines are best fits by the corresponding powers of  $\epsilon$ . The data are shown on double logarithmic scale.

## B. STU-embedding and reduction of S, T and U into M

Now the matrix of fast transition rates is

$$\hat{A}_f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & \vec{TS} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \vec{ST} & * & \vec{UT} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \vec{TU} & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (53)$$

In this section we develop another approximation of the original system, which considers the transitions between states  $S$ ,  $T$  and  $U$  as fast, which in asymptotics leads to their merger into a new state  $\tilde{M}$ . This choice is supported by the empirical embedding procedure as described in Section IV, details can be found in Ref. [25].

and the remaining, slow rates constitute the matrix

$$\hat{A}_s = \begin{bmatrix} * & \vec{PO} & 0 & 0 & 0 & 0 & \vec{UO} & 0 & 0 \\ \vec{OP} & * & \vec{QP} & 0 & 0 & 0 & \vec{UP} & 0 & 0 \\ 0 & \vec{PQ} & * & \vec{RQ} & 0 & \vec{TQ} & 0 & 0 & 0 \\ 0 & 0 & \vec{QR} & * & \vec{SR} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \vec{RS} & * & 0 & 0 & 0 & 0 \\ 0 & 0 & \vec{QT} & 0 & 0 & * & 0 & 0 & 0 \\ \vec{OU} & \vec{PU} & 0 & 0 & 0 & 0 & * & \vec{VU} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vec{UV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix} \quad (54)$$

The right eigenvectors corresponding to zero eigenvalue of this system can be chosen as

$$\begin{aligned} \kappa_i &= \mathbf{e}_i, \quad i = 1, \dots, 4, \\ \kappa_5 &= (\vec{UTTS} + \vec{UTST} + \vec{TUST})^{-1} \\ &\quad \times \left( \vec{UTTS} \mathbf{e}_5 + \vec{UTST} \mathbf{e}_6 + \vec{TUST} \mathbf{e}_7 \right), \\ \kappa_i &= \mathbf{e}_{i+2}, \quad i = 6, 7. \end{aligned} \quad (55)$$

The corresponding left eigenvectors are

$$\begin{aligned} \rho_i &= \kappa_i, \quad i = 1, 2, 3, 4, 6, 7, \\ \rho_5 &= \mathbf{e}_5 + \mathbf{e}_6 + \mathbf{e}_7. \end{aligned} \quad (56)$$

With account of these, we can keep the names of the original dynamic variables for all states except  $S$ ,  $T$ ,  $U$ , so the vector of states of the reduced system is

$$\mathbf{a} = [O, P, Q, R, \tilde{M}, V, W]^T. \quad (57)$$

These are all the ingredients needed for the derivation of the leading-order approximation. We have  $\rho_i' = 0$  for all  $i = 1, \dots, 7$  so (32) gives the leading-order transition rate matrix for the reduced model as

$$\hat{M}_0 = \begin{bmatrix} * & \vec{PO} & 0 & 0 & \vec{MO} & 0 & 0 \\ \vec{OP} & * & \vec{QP} & 0 & \vec{MP} & 0 & 0 \\ 0 & \vec{PQ} & * & \vec{RQ} & \vec{MQ} & 0 & 0 \\ 0 & 0 & \vec{QR} & * & \vec{MR} & 0 & 0 \\ \vec{OM} & \vec{PM} & \vec{QM} & \vec{RM} & * & \vec{VM} & 0 \\ 0 & 0 & 0 & 0 & \vec{MV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix}, \quad (58)$$

with the new transition rates defined as

$$\begin{aligned} \vec{MO} &= \vec{UO} \gamma_{STTU}, & \vec{OM} &= \vec{OU}, \\ \vec{MP} &= \vec{UP} \gamma_{STTU}, & \vec{PM} &= \vec{PU}, \\ \vec{MQ} &= \vec{TQ} \gamma_{UTST}, & \vec{QM} &= \vec{QT}, \\ \vec{MR} &= \vec{SR} \gamma_{UTTS}, & \vec{RM} &= \vec{RS}, \\ \vec{MV} &= \vec{UV} \gamma_{STTU}, & \vec{VM} &= \vec{VU}. \end{aligned} \quad (59)$$

These expression use the notation  $\gamma_{ijkl}$  as an abbreviation for

$$\gamma_{ijkl} = \frac{\vec{ij} \vec{kl}}{\vec{UTTS} + \vec{UTST} + \vec{STTU}}. \quad (60)$$

The original coordinates are recovered from the reduced one by

$$S = \gamma_{UTTS} \tilde{M}, \quad (61a)$$

$$T = \gamma_{UTST} \tilde{M}, \quad (61b)$$

$$U = \gamma_{STTU} \tilde{M}. \quad (61c)$$

As can be seen in fig. 4 below, the quality of the approximation obtained with these asymptotics, is very good. This was of course to be expected based on the results of the empirical embedding study, as discussed above.

### C. Embedding and reduction of R and Q states of STU-reduction into L

In this section, we investigate how one can build on the success of the  $STU$  embedding and achieve further reduction. As we have already considered the  $OP$  reduction above, we now consider  $RQ$  reduction. That is, we consider the transition rates between  $R$  and  $Q$  as fast, which will lead to the merger of these two states into a new state  $\tilde{L}$ . So in the context of the present section, the “original model” is defined by the matrix (58), which will now be called  $\hat{M}$ , while the fast matrix in the new embedding is

$$\hat{M}_f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & \vec{RQ} & 0 & 0 & 0 \\ 0 & 0 & \vec{QR} & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (62)$$

and the slow matrix is

$$\hat{M}_s = \begin{bmatrix} * & \vec{PO} & 0 & 0 & \vec{MO} & 0 & 0 \\ \vec{OP} & * & \vec{QP} & 0 & \vec{MP} & 0 & 0 \\ 0 & \vec{PQ} & * & 0 & \vec{MQ} & 0 & 0 \\ 0 & 0 & 0 & * & \vec{MR} & 0 & 0 \\ \vec{OM} & \vec{PM} & \vec{QM} & \vec{RM} & * & \vec{VM} & 0 \\ 0 & 0 & 0 & 0 & \vec{MV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix} \quad (63)$$

Acting as before, we find the right eigenvectors of  $\hat{M}_f$  corresponding to zero eigenvalue as

$$\begin{aligned} \kappa_i &= \mathbf{e}_i, & i &= 1, 2, \\ \kappa_3 &= (\vec{QR} + \vec{RQ})^{-1} \left( \vec{QR} \mathbf{e}_3 + \vec{RQ} \mathbf{e}_4 \right), \\ \kappa_i &= \mathbf{e}_{i+1}, & i &= 4, 5, 6, \end{aligned}$$

the corresponding left eigenvectors as

$$\begin{aligned} \rho_i &= \kappa_i, & i &= 1, 2, 4, 5, 6, \\ \rho_3 &= \mathbf{e}_3 + \mathbf{e}_4, \end{aligned}$$

and we set the names of the components of the reduced vector as

$$\mathbf{a} = [O, P, \tilde{L}, \tilde{M}, V, W]^T.$$

The resulting leading-order reduced matrix works out as

$$\hat{M}_0 = \begin{bmatrix} * & \vec{PO} & 0 & \vec{MO} & 0 & 0 \\ \vec{OP} & * & \vec{LP} & \vec{MP} & 0 & 0 \\ 0 & \vec{PL} & * & \vec{ML} & 0 & 0 \\ \vec{OM} & \vec{PM} & \vec{LM} & * & \vec{VM} & 0 \\ 0 & 0 & 0 & \vec{MV} & * & \vec{WV} \\ 0 & 0 & 0 & 0 & \vec{VW} & * \end{bmatrix} \quad (64)$$

with the new transition rates defined by

$$\begin{aligned} \vec{LP} &= \vec{QP} \delta_{QR}, & \vec{PL} &= \vec{PQ} \\ \vec{LM} &= \vec{QM} \delta_{QR} + \vec{RM} \delta_{RQ}, & \vec{ML} &= \vec{MQ} + \vec{MR}, \end{aligned} \quad (65)$$

where

$$\delta_{ij} = \frac{\vec{ij}}{\vec{QR} + \vec{RQ}}. \quad (66)$$

It is easily seen that the resulting reduced model (65) will be the same if, instead, we do the  $QR$  reduction first and  $STU$  reduction second, or do them simultaneously, *i.e.* include all of  $S \cdot T$ ,  $T \cdot U$  and  $Q \cdot R$  in  $\hat{A}_f$  in the first place.

Fig. 4 presents the results of the  $QR$ - $STU$ , together with the previously considered  $OP$  and  $STU$  reductions. In these simulations, the  $I_{Na}$  channel model was not driven by the recorded  $V_m(t)$  as before, but rather was part of the full cell model (35). The full original model and the three reduced version were run in the same protocol, which included stimulation with a period of one second, starting from  $t = 1$  ms (this was done in order to be able to show the time in panel (a) in the logarithmic scale). One can see that the reduced models are indistinguishable from the full model except for the upstroke of the action potential. The upstroke of the fifth action potential is shown in detail in panels (b) and (c), for the probability of the  $I_{Na}$  channel being open, and the resulting value of this current. We see that the results generally agree with what could be expected from the empirical embedding studies illustrated in fig. 2. Namely, the  $OP$  embedding gives a rather poor approximation, the  $QR$ - $STU$  embedding is slightly better, while  $STU$  is very good.

## VII. DISCUSSION

Fig. 5 summarises the Markov chain models occurring as a result of the three asymptotics we have considered: this is to be compared with the original scheme shown in fig. 1.

Asymptotic reduction based on time scale separation can pursue at least two different goals: reducing the number of dynamic equations, and reducing stiffness of those equations. The reductions considered in this paper are not particularly impressive in terms of reducing the number of equations: we have reduced by maximum of three out of nine, which is even less significant in comparison with the number of other dynamic equations in a typical model of an excitable cell, beyond the Markov chain of the  $I_{Na}$  channel. However, in practical applications the main goal is the other one: reducing the stiffness. To achieve a simple practical estimate of this characteristic, we measured the stiffness of the model by the maximum time step size  $\Delta_t$  which provides a stable solution using the forward Euler solver for the isolated  $I_{Na}$  model driven by a recorded action potential. The original full model allows the time step of about  $\Delta_t \approx 0.04$  ms for stable computations; an increase above that leads to numerical instability. In comparison to that, all three models considered allow  $\Delta_t \approx 0.044$  ms, *i.e.* a rather modest improvement. The limited progress in this is due to the fact that in all three examples considered, we have included in the embedding only some of the fastest transition rates. And even in these cases, we have seen that asymptotic removal of some of the fast processes affects the accuracy of computations. Even though these effects are seen only during the upstrokes of the action potential, these upstrokes are of principal significance as they determine the conduction velocity in spatially-distributed simulation, and therefore also the more delicate and more

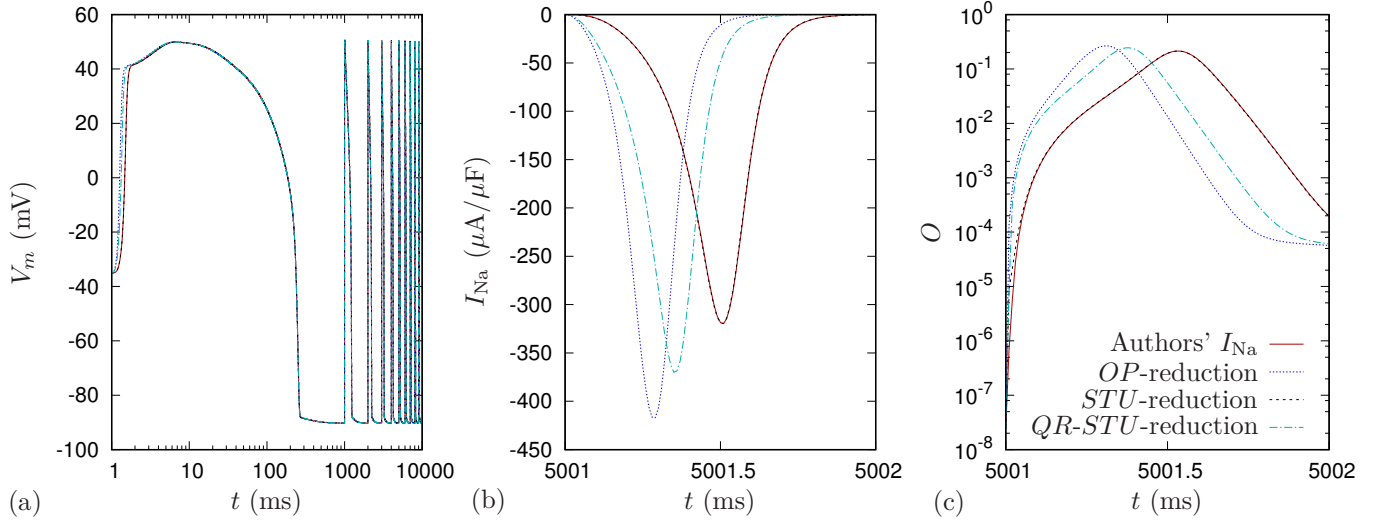


FIG. 4. Comparison of action potential solutions produced with different asymptotics of the  $I_{Na}$  channel model. (a) Transmembrane voltage during 10 action potential vs time (log scale). (b) The  $I_{Na}$  current at the onset of the fifth action potential. (c) The state occupancy of  $O$  (log scale) during the fifth action potential. Red solid line: the original model. Blue dotted line: the  $OP$ -reduction. Black dashed line: the  $STU$ -reduction. Cyan dash-dotted line:  $QR$ - $STU$  reduction.

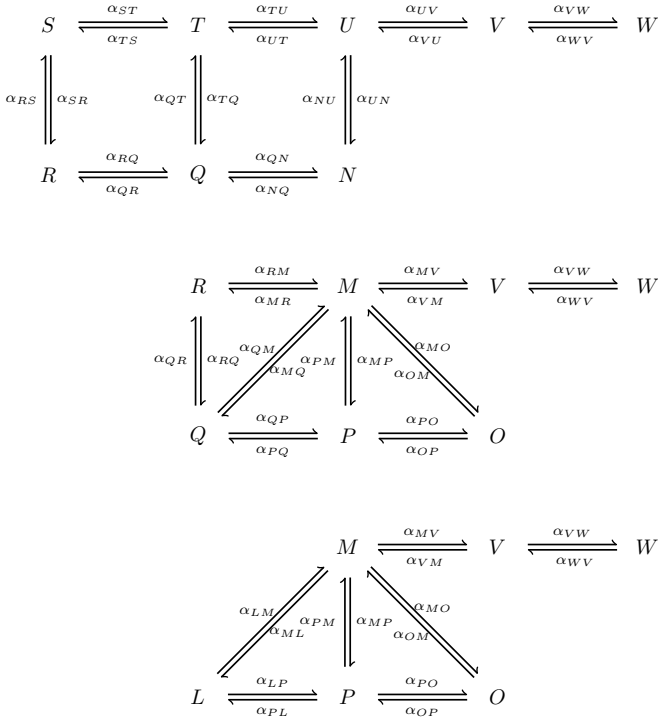


FIG. 5. Diagram of reduced Markov chain models of  $I_{Na}$  channel. Top diagram shows the  $OP$ -reduced model, middle diagram shows  $STU$ -reduced model, and bottom diagram shows  $RQ$ - $STU$ -reduced model.

important phenomena such as conduction block, wave-breaks etc. Hence further increase of the number of the reduced degrees of freedom does not seem to be an answer. Further research is of course needed to establish that with certainty, but, as already noted above, *e.g.* the

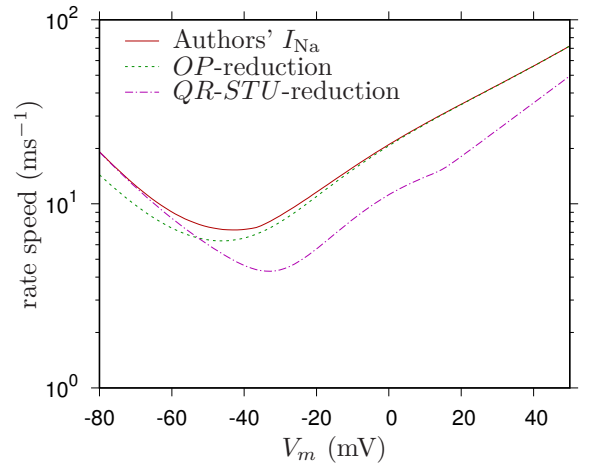


FIG. 6. Largest absolute values of the eigenvalues of the transition matrices for the original Markov chain and the two selected reduced models.

“straightforward” approach embedding *all* the transition rates that appear “fast” in fig. 1 does not yield a satisfactory approximation<sup>25</sup>.

From a more theoretical viewpoint, stiffness can sometimes be characterized by the eigenvalues of the system; in particular, the upper limit of the integration step is mainly affected by the eigenvalue with the largest absolute value. In fig. 6 we plot those absolute values for the original model and the reduced models, as functions of the transmembrane voltage. We see that whereas  $OP$  reduction somewhat reduces stiffness at the lower end of the  $V_m$  scale, it has virtually no effect at the upper end. On the contrary,  $QR$ - $STU$  reduction noticeably reduces stiffness at the upper end, without changing it at the

lower end.

So, as far as the question posed in the introduction is concerned, the results obtained here seem to suggest that Tikhonov asymptotic structure, which implies fixed distribution of the roles of “fast” and “slow” variables, or, in this particular class of applications, rather “fast” and “slow” transition rates, may not be quite adequate for the this particular model of the fast sodium current, and some non-Tikhonov parametric embedding may be more fruitful, say when transition rates are considered fast in one range of  $V_m$  and slow in the complementary range, possibly with the asymmetry between the reciprocal rates taken into account.

An alternative approach, which has proved to be more practical than the one considered here, has been described in our previous works<sup>8,9</sup>, dubbed “exponential solvers”. However, that approach is purely numerical and does not explicitly take into account the fast-slow structure of the model, hence an asymptotic approach seems to have an *a priori* advantage, which ought to have been explored. We hope that the present study fills this gap to a certain extent.

An attractive possibility to improve the accuracy of the asymptotics and hence to open the way to further decrease the number of equations and reduce the stiffness, seems to be using higher-order asymptotics. We have explored this only in one of the three examples, but it already shows that (i) the algebraic complexity of the resulting formulas increases considerably, (ii) more significantly, some improvement in accuracy is devalued by the fact that the resulting model, unlike the leading-order asymptotics, no longer behaves as a “proper” Markov chain: the vector of dynamic variables is not guaranteed to remain stochastic, in particular, it can easily lead to negative values of the state occupancies. This happens because we have used the asymptotic theory which was designed for generic systems and is not tailored for the specific requirements of Markov chains. Hence another possible way for improvement may be in developing higher-order asymptotics strictly within the class of Markov chains.

## ACKNOWLEDGMENTS

VNB gratefully acknowledges the current financial support of the EPSRC via grant EP/N014391/1 (UK) TS acknowledges financial support of the University of Exeter via PhD Studentship and of the EPSRC via grant EP/N024508/1.

## Appendix A: Derivation of Reduced System

We use the Taylor expansion for the functions  $\mathbf{f}(\mathbf{u})$  and  $\mathbf{h}(\mathbf{u})$ , such that, after substitution of the sought solution (2), we get the expression on the right hand side of (1)

as

$$\mathbf{f}(\mathbf{U}) + \varepsilon \sum_j \frac{\partial \mathbf{f}}{\partial u_j} v_j + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} + \varepsilon \mathbf{h}(\mathbf{U}) + \varepsilon^2 \sum_{j_1} \frac{\partial \mathbf{h}}{\partial u_{j_1}} v_{j_1} + \mathcal{O}(\varepsilon^3). \quad (\text{A1})$$

The first term  $\mathbf{f}(\mathbf{U}) = 0$  by assumption, and in the second term we note that the derivatives constitute the Jacobian matrix and expand the  $\mathbf{v}$  according to (3). Then the previous expression (A1) rewrites as

$$\varepsilon \hat{F}(\mathbf{U}) \sum_\ell b_\ell \mathbf{V}_\ell + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} + \varepsilon \mathbf{h}(\mathbf{U}) + \varepsilon^2 \sum_{j_1} \frac{\partial \mathbf{h}}{\partial u_{j_1}} v_{j_1} + \mathcal{O}(\varepsilon^3). \quad (\text{A2})$$

We substitute the sought solution also to the left hand side of (1). We use the knowledge of eigenvectors corresponding to zero eigenvalues from (5) and expand the perturbed term (3). Then using a chain rule for the derivative of  $\mathbf{U}(\mathbf{a})$  and  $\mathbf{V}_\ell(\mathbf{a})$  we get

$$\sum_k \frac{\partial \mathbf{U}}{\partial a_k} \frac{da_k}{dt} + \varepsilon \frac{d\mathbf{v}}{dt} = \sum_k \mathbf{V}_k \frac{da_k}{dt} + \varepsilon \sum_\ell \left( \frac{db_\ell}{dt} \mathbf{V}_\ell + b_\ell \frac{d\mathbf{V}_\ell}{da_k} \frac{da_k}{dt} \right). \quad (\text{A3})$$

Combining the right-hand side given by (A1) and the left-hand side given by (A3), we rewrite (1) as

$$\begin{aligned} & \sum_k \mathbf{V}_k \frac{da_k}{dt} + \varepsilon \sum_\ell \left( \frac{db_\ell}{dt} \mathbf{V}_\ell + b_\ell \frac{d\mathbf{V}_\ell}{da_k} \frac{da_k}{dt} \right) \\ &= \varepsilon \hat{F} \sum_\ell b_\ell \mathbf{V}_\ell + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} \\ &+ \varepsilon \mathbf{h}(\mathbf{U}) + \varepsilon^2 \sum_{j_1} \frac{\partial \mathbf{h}}{\partial u_{j_1}} v_{j_1} + \mathcal{O}(\varepsilon^3). \end{aligned} \quad (\text{A4})$$

Multiplying the equation by the adjoint eigenvectors  $\mathbf{W}_i^T$  gives

$$\begin{aligned} \frac{1}{\varepsilon} \frac{da_i}{dt} + \frac{db_i}{dt} &= \Lambda_i b_i + \mathbf{W}_i^T \mathbf{h}(\mathbf{U}) \\ &+ \varepsilon \mathbf{W}_i^T \left[ \sum_j \frac{\partial \mathbf{h}}{\partial u_j} v_j + \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} \right. \\ &\left. - \frac{1}{\varepsilon} \sum_{\ell, k} b_\ell \frac{\partial \mathbf{V}_\ell}{\partial a_k} \frac{da_k}{dt} \right] + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (\text{A5})$$

Considering separately the zero and the stable eigenvalues then yields equations (6a) and (6b) respectively.

- <sup>1</sup>G. Plank, L. Zhou, J. L. Greenstein, S. Cortassa, R. L. Winslow, B. O'Rourke, and N. A. Trayanova, "From mitochondrial ion channels to arrhythmias in the heart: computational techniques to bridge the spatio-temporal scales," *Philos Transact Roy Soc A* **366**, 3381–3409 (2008).
- <sup>2</sup>D. F. Richards, J. N. Glosli, E. W. Draeger, A. A. Mirin, B. Chan, J. Fattebert, W. D. Krauss, T. Oppelstrup, C. J. Butler, J. A. Gunnels, V. Gurev, C. Kim, J. Magerlein, M. Reumann, H. F. Wen, and J. J. Rice, "Towards real-time simulation of cardiac electrophysiology in a human heart at high resolution," *Computer Methods in Biomechanics and Biomedical Engineering* **16**, 802–805 (2013).
- <sup>3</sup>V. E. Bondarenko, "A compartmentalized mathematical model of the  $\beta_1$ -adrenergic signaling system in mouse ventricular myocytes," *PLOS ONE* **9**, e89113 (2014).
- <sup>4</sup>A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J Physiol Lond* **117**, 500–544 (1952).
- <sup>5</sup>S. Rush and H. Larsen, "A practical algorithm for solving dynamic membrane equations," *IEEE Trans BME* **25**, 389–392 (1978).
- <sup>6</sup>M. Perego and A. Veneziani, "An efficient generalization of the Rush-Larsen method for solving electro-physiology membrane equations," *Electronic Transactions on Numerical Analysis* **35**, 234–256 (2009).
- <sup>7</sup>M. E. Marsh, S. T. Ziaratgahi, and R. J. Spiteri, "The secrets to the success of the Rush-Larsen method and its generalizations," *IEEE Trans BME* **59**, 2506–2515 (2012).
- <sup>8</sup>T. Stary and V. N. Biktashev, "Exponential integrators for a Markov chain model of the fast sodium channel of cardiomyocytes," *IEEE Trans BME* **62**, 1070–1076 (2015).
- <sup>9</sup>T. Stary and V. Biktashev, "Evaluating exponential integrators for Markov chain ion channel models," *Computing in Cardiology* **42**, 885–888 (2015).
- <sup>10</sup>R. Hinch, J. L. Greenstein, A. J. Tanskanen, L. Xu, and R. L. Winslow, "A simplified local control model of calcium-induced calcium release in cardiac ventricular myocytes," *Biophys J* **87**, 3723–3736 (2004).
- <sup>11</sup>A. N. Tikhonov, "Systems of differential equations containing small parameters in the derivatives," *Mat. Sb. (N.S.)* **31(73)**, 575–586 (1952).
- <sup>12</sup>N. Fenichel, "Geometric singular perturbation theory for ordinary differential equations," *Journal of Differential Equations* **31**, 53–98 (1979).
- <sup>13</sup>V. N. Biktashev, "Dissipation of the excitation wavefronts," *Phys Rev Lett* **89**, 168102 (2002).
- <sup>14</sup>V. N. Biktashev and R. Suckley, "Non-Tikhonov asymptotic properties of cardiac excitability," *Phys Rev Lett* **93**, 168103 (2004).
- <sup>15</sup>I. V. Biktasheva, R. D. Simitev, R. S. Suckley, and V. N. Biktashev, "Asymptotic properties of mathematical models of excitability," *Philos Transact Roy Soc A* **364**, 1283–1298 (2006).
- <sup>16</sup>R. D. Simitev and V. N. Biktashev, "Asymptotics of conduction velocity restitution in models of electrical excitation in the heart," *Bull Math Biol* **73**, 72–115 (2011).
- <sup>17</sup>C. E. Clancy and Y. Rudy, "Na<sup>+</sup> channel mutation that causes both Brugada and long-QT syndrome phenotypes: a simulation study of mechanism," *Circulation* **105**, 1208–1213 (2002).
- <sup>18</sup>V. Biktashev, "Envelope equations for modulated non-conservative waves," *IUTAM Symposium Asymptotics, Singularities and Homogenisation in Problems of Mechanics* **5(1)**, 11 (2003).
- <sup>19</sup>R. Suckley and V. N. Biktashev, "Comparison of asymptotics of heart and nerve excitability," *Phys Rev E* **68**, 011902 (2003).
- <sup>20</sup>V. N. Biktashev and R. Suckley, "Non-Tikhonov asymptotic properties of cardiac excitability," *Phys Rev Lett* **93**, 168103 (2004).
- <sup>21</sup>I. V. Biktasheva, R. D. Simitev, R. Suckley, and V. N. Biktashev, "Asymptotic properties of mathematical models of excitability," *Philos Transact Roy Soc A* **364**, 1283–1298 (2006).
- <sup>22</sup>V. N. Biktashev, R. Suckley, Y. E. Elkin, and R. D. Simitev, "Asymptotic analysis and analytical solutions of a model of cardiac excitation," *Bull Math Biol* **70**, 517–554 (2008).
- <sup>23</sup>Note that diagonalizability and reality of the eigenvalues of the full transition rate matrix  $\hat{A}$  can be guaranteed under the assumption of detailed balance<sup>26</sup>, and  $\hat{A}_f = \lim_{\varepsilon \rightarrow 0} (\varepsilon \hat{A})$ .
- <sup>24</sup>S. West, L. J. Bridge, M. R. H. White, P. Paszek, and V. N. Biktashev, "A method of 'speed coefficients' for biochemical model reduction applied to the NF-kappaB system," *J Math Biol* **70**, 591–620 (2015).
- <sup>25</sup>T. Stary, *Mathematical and Computational Study of Markovian Models of Ion Channels in Cardiac Excitation*, Ph.D. thesis, University of Exeter (2016).
- <sup>26</sup>W. J. Anderson, *Continuous Time Markov Chains. An Application-Oriented Approach* (Springer, New York etc, 1991).