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

Tomáš Štárý, Vadim N. Biktashev

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

(Dated: 29 March 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 transmembrane 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. 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. Hence a direct approach using explicit time steppers requires very small time steps, hence high computational demands. 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 structure of the paper is as follows. Section II introduces the notation and main principles of the singular 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. [1]–[5], adjusting where necessary for our present purposes.

We consider an autonomous system of ordinary differential equations

$$\frac{du}{dt} = f(u) + \varepsilon h(u)$$  \hspace{1cm} (1)

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

$$u = U(a) + \varepsilon v(t)$$  \hspace{1cm} (2)

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

$$v(t) = \sum_{\ell} b_\ell V_\ell(a),$$  \hspace{1cm} (3)

where the vectors $V_\ell(a)$ are right eigenvectors of a Jacobian matrix $\hat{F}(U) = \partial f/\partial u|_{u=U}$, $\hat{F}V_\ell = \Lambda_\ell V_\ell,$  \hspace{1cm} (4)

and the summation index $\ell$ runs through the stable eigenvalues, $\text{Re}(\Lambda_\ell) < 0$, $\ell = m + 1, \ldots, n$, skipping the zero eigenvalues, $\Lambda_k$, $k = 1, \ldots, m$, corresponding to the directions tangent to the manifold. Table I 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

$$V_k = \frac{\partial U}{\partial a_k}. \hspace{1cm} (5)$$

We substitute (2) and (3) into (1), expand the non-linear functions into their Taylor series and separate the components using left eigenvectors $W^T_k$ as projectors. The detailed derivation is presented in the Appendix A. The final result reads as the following system of ODEs

$$\frac{1}{\varepsilon} \frac{db_k}{dt} = W^T_k h(U) + \varepsilon F_k(a, b) + \mathcal{O}(\varepsilon^2), \hspace{1cm} (6a)$$

$$\frac{db}{dt} = \Lambda b + W^T_k h(U) + \mathcal{O}(\varepsilon), \hspace{1cm} (6b)$$

where in the right-hand side of the first equation we have kept the leading order term, $W^T_k h(U)$, and the first-order correction $F_k$, which works out as

$$F_k(a, b) = W^T_k \hat{H}(U)v + W^T_k \sum_{j_1, j_2} \frac{\partial^2 f}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} + \frac{1}{\varepsilon} \sum_k \frac{\partial W^T_k}{\partial a_k} \frac{da_k}{dt} v,$$  \hspace{1cm} (7)

where

$$\hat{H}(U) = \frac{\partial h}{\partial u} \bigg|_{u=U}. \hspace{1cm} (8)$$

Equations (6a) and (6b) are coupled through higher-order terms, and to complete the reduction, we need to eliminate $b$. For the solution of the manifold coordinates $a$ up to $\mathcal{O}(\varepsilon^N)$ it is sufficient to find the correction term $b$ up to $\mathcal{O}(\varepsilon^{N-1})$. The leading order term for the correction $b$ in terms of $a$ 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

$$b_\ell = - \frac{W^T_k h(U)}{\Lambda_\ell} + \mathcal{O}(\varepsilon), \hspace{1cm} (9)$$

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

III. DIMENSIONALITY REDUCTION FOR TIME-INHOMOGENEOUS MARKOV CHAINS

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

$$\frac{dx}{dt} = \hat{A}(t)x. \hspace{1cm} (10)$$

Entries in the vector of dynamical variables $x \in \mathbb{R}^n$ 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

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

$^a$ Here $\ell = 1$ for Section II and $\ell = 0$ from the next section on.
$^b$ This is used starting from Section III.
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 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 autonomization, 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

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

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

\[
\hat{A} = \frac{1}{\varepsilon} \hat{A}_f + \hat{A}_s. \tag{12}
\]

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 \( \kappa_q(\sigma) \):

\[
\hat{A}_f(\sigma)\kappa_q(\sigma) = \lambda_q(\sigma)\kappa_q(\sigma) \tag{13}
\]

(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).

Correspondingly, we introduce also the left eigenvectors \( \rho_q \),

\[
\hat{A}_f^T \rho_q = \lambda_q \rho_q, \quad \rho_q^T \kappa_q' = \delta_{q,q'}. \tag{14}
\]

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

\[
\frac{\partial \rho_q^T}{\partial \sigma} \kappa_j = -\rho_q^T \frac{\partial \kappa_j}{\partial \sigma}. \tag{14}
\]

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

\[
\begin{align*}
\frac{d\sigma}{d\tau} &= \varepsilon, \\
\frac{d\mathbf{x}}{d\tau} &= \hat{A}_f(\sigma)\mathbf{x} + \varepsilon \hat{A}_s(\sigma)\mathbf{x},
\end{align*} \tag{15}
\]

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

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

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 \( \iota \) in Table III 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 \( \hat{A}_f \):

\[
\mathbf{U}(\sigma, \mathbf{a}) = \begin{bmatrix} a_0 \\ \mathbf{a}^T \end{bmatrix} \tag{17}
\]

where \( a_0 = \sigma, \mathbf{a} = [a_1, \ldots, a_m]^T \) and

\[
\mathbf{U}^\sigma(\mathbf{a}) = \sum_r a_r \kappa_r. \tag{18}
\]

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

\[
\hat{F} \mathbf{U} = \begin{bmatrix} 0 \\ \hat{A}_f(\sigma) \end{bmatrix} \mathbf{U}^\sigma. \tag{19}
\]

Let us denote the components of the eigenvectors as

\[
\mathbf{V}_i = \begin{bmatrix} \mu_i \\ \mathbf{V}^\sigma_i \end{bmatrix}. \tag{20}
\]

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

\[
\begin{align*}
\Lambda_i \mu_i &= 0, \tag{21a} \\
\Lambda_i \mathbf{V}^\sigma_i &= \frac{\partial \hat{A}_f}{\partial \sigma} \mathbf{U}^\sigma_i + \hat{A}_f \mathbf{V}^\sigma_i. \tag{21b}
\end{align*}
\]
Let us consider separately the cases $\mu_i = 0$ and $\mu_i \neq 0$. For $\mu_i = 0$, equation (21b) becomes

$$\dot{A}_f V^x_q = \lambda_q V^x_q, \quad (22)$$

which is the definition of an eigenvalue problem for matrix $\dot{A}_f$, so we can take $V^x_q = \kappa_q$ and $\lambda_q = \Lambda_q$, for $q = 1, \ldots, n$, out of which the first $m$ are zero eigenvalues. For $r = 1, \ldots, m$, we have $\Lambda_r = 0$, and differentiation of $[13]$ with respect to $\sigma$ gives

$$\frac{\partial \dot{A}_f}{\partial \sigma} \kappa_r = -\dot{A}_f \frac{\partial \kappa_r}{\partial \sigma}. \quad (23)$$

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

$$\dot{A}_f V^x_0 = -\frac{\partial \dot{A}_f}{\partial \sigma} U^x. \quad (24)$$

If we substitute $[18]$ into $[24]$ and use $[23]$, we get

$$\dot{A}_f V^x_0 = \dot{A}_f \sum_r a_r \frac{\partial \kappa_r}{\partial \sigma} \quad (25)$$

so we can choose

$$\kappa_0 = \sum_r a_r \frac{\partial \kappa_r}{\partial \sigma} \quad (26)$$

(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 $V_i, W_i$ of the Jacobian in the time-extended system are related to those $\lambda_q, \kappa_q, \rho_q$ of the transition rate matrix via the following relationships:

$$V_0 = \begin{bmatrix} 1 & \sum_r a_r \frac{\partial \kappa_r}{\partial \sigma} \end{bmatrix}, \quad V_q = \begin{bmatrix} 0 \\ \kappa_q \end{bmatrix}, \quad (27)$$

$$W_0 = \begin{bmatrix} 1 \\ \rho_q \end{bmatrix}, \quad W_q = \begin{bmatrix} -\rho_q T \sum_r a_r \frac{\partial \kappa_r}{\partial \sigma} \\ \rho_q \end{bmatrix},$$

$$\Lambda_0 = 0, \quad \Lambda_q = \lambda_q. \quad (28)$$

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

$$W_q^T \dot{h}(U) = \begin{bmatrix} \frac{\partial \rho_q T}{\partial \sigma} U^x, & \rho_q T \end{bmatrix} \begin{bmatrix} 1 \\ \dot{A}_s(\sigma) \kappa_q \end{bmatrix}$$

$$= -\rho_q T \sum_r a_r \left( \frac{\partial \kappa_r}{\partial \sigma} - \dot{A}_s(\sigma) \kappa_r \right),$$

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

$$b_\ell = \frac{\rho_\ell^T}{\lambda_\ell} \sum_r a_r \left( \frac{\partial \kappa_r}{\partial \sigma} - \dot{A}_s(\sigma) \kappa_r \right) + O(\varepsilon), \quad (29)$$

and the transversal correction itself as

$$v^x = \sum_\ell b_\ell \kappa_\ell = \sum_\ell \left[ \frac{a_\ell}{\lambda_\ell} \rho_\ell^T \left( \frac{\partial \kappa_\ell}{\partial \sigma} - \dot{A}_s(\sigma) \kappa_\ell \right) \right] \kappa_\ell. \quad (30)$$

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

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

$$\frac{1}{\varepsilon} \frac{da_\ell}{dt} = \left( \frac{\partial \rho_\ell T}{\partial \sigma} + \rho_\ell T \dot{A}_s(\sigma) \right) \sum_r a_r \kappa_r \quad (30b)$$

$$+ \varepsilon \left( \rho_\ell T \dot{A}_s(\sigma) + \frac{\partial \rho_\ell T}{\partial \sigma} \right) \sum_r \left[ \frac{a_r}{\lambda_\ell} \rho_\ell^T \left( \frac{\partial \kappa_r}{\partial \sigma} - \dot{A}_s(\sigma) \kappa_r \right) \right] \kappa_\ell + O(\varepsilon^2).$$

This result can be written in the matrix form as

$$\frac{1}{\varepsilon} \frac{dA_0}{dt} = \left( M_0(\sigma) + \varepsilon M_1(\sigma) \right) A + O(\varepsilon^2), \quad (31)$$

where

$$M_0 = \hat{P}^T + \hat{P} \dot{A}_s \hat{K}, \quad (32)$$

$$M_1 = \hat{P}^T + \hat{P} \dot{A}_s \hat{L} \left[ \hat{K}' - \dot{A}_s \hat{K} \right], \quad (33)$$

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. [2] [4] and [7]. 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 $\varepsilon \to 0$ 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 $\varepsilon \to 0$ 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 $\varepsilon \to 0$ 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 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 [8] (we consider the wild-type version). 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. [9]: these are single-letter names, as opposed to the original names in Ref. [8] 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

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

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

$$\hat{A} = \begin{pmatrix}
D_O & -P & 0 & 0 & 0 & 0 & U & 0 & 0 \\
O & D_P & Q & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & P & Q & R & 0 & T & 0 & 0 & 0 \\
0 & 0 & Q & R & D_R & SR & 0 & 0 & 0 \\
0 & 0 & 0 & RS & D_S & TS & 0 & 0 & 0 \\
0 & 0 & QT & ST & D_T & UT & 0 & 0 & 0 \\
OU & PU & 0 & 0 & 0 & TU & D_U & VU & 0 \\
0 & 0 & 0 & 0 & 0 & UV & D_V & WV & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & VW & D_W & 0 \\
\end{pmatrix},$$

where we use the notation $\hat{ij}$ for the transition rate from state $i$ to state $j$, which is shorter than the $a_{ij}$ used traditionally for this purpose (however we maintain the traditional notations in the figures).

As always in transition rate matrices, each of the diagonal entries of $\hat{A}$ is a negative of the sum of the other entries in the same column, so

$$D_O = -(OP + OU),$$

$$D_P = -(PQ + PU + PO),$$

$$D_Q = -(QR + QT + QP),$$

$$D_R = -(RS + RQ),$$

$$D_S = -(ST + SR),$$

$$D_T = -(TQ + TS + TU),$$

(35)
of their connection, i.e., $i \leftrightarrow j \triangleq ij + ji$. Indeed, it this 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. EMBDDINGS 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 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 either into $\hat{A}_f$ or into $\hat{A}_s$, which guarantees that $\hat{A}_f$ and $\hat{A}_s$ are left-stochastic.

We have tested a number of different combinations of reciprocal transition rates for $\hat{A}_f$. Figure 2 shows results of simulation of some of those combinations. The $I_{\text{Na}}$ model was extracted from the authors code. The simulation of the model were driven by a recorded values of $V_m(t)$ during a standard action potential from a single-cell simulation. The time step in the simulation of $I_{\text{Na}}$ 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 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.

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 $OP$ and $PO$. 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 slow transition rates.
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 ($\varepsilon = 1$), the green lines show the embeddings of transition rates between states $O \leftrightarrow P$ (i.e. $\varepsilon = 0.1$), etc., as specified in the legend of panel $W$.

The matrix of the fast transition rates

$$
\hat{A}_f = \begin{bmatrix}
* & \vec{PO} & 0 & 0 & 0 & 0 & 0 & 0 \\
\vec{OP} & * & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \vec{PQ} & * & 0 & 0 & \vec{UP} & 0 \\
0 & 0 & \vec{PQ} & * & \vec{RQ} & 0 & \vec{TQ} & 0 \\
0 & 0 & 0 & 0 & \vec{QR} & * & \vec{SR} & 0 \\
0 & 0 & 0 & 0 & \vec{QR} & * & \vec{SR} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \vec{RS} & * & \vec{TS} \\
0 & 0 & 0 & 0 & 0 & 0 & \vec{RS} & * & \vec{TS} \\
\vec{OU} & \vec{PU} & 0 & 0 & 0 & 0 & \vec{TU} & * & \vec{VU} \\
0 & 0 & 0 & 0 & 0 & 0 & \vec{UV} & * & \vec{WV} \\
0 & 0 & 0 & 0 & 0 & 0 & \vec{VW} & * & \vec{WV}
\end{bmatrix}
$$

and

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

(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
\[ 1^T \hat{A} = 0. \]  
(40)

In fact, we have $\lambda_r = 0$ for $r = 1, \ldots, m$, where $m = 8$, and just one non-zero eigenvalue $\lambda_9 = - (\hat{P} \hat{O} + \hat{O} \hat{P})$. The corresponding right eigenvectors are
\[ \kappa_1 = (\hat{P} \hat{O} + \hat{O} \hat{P})^{-1} \left[ \hat{P} \hat{O}, \hat{O} \hat{P}, 0, 0, 0, 0, 0, 0 \right]^T, \]
\[ \kappa_2 = [0, 0, 1, 0, 0, 0, 0, 0]^T, \]
\[ \kappa_3 = [0, 0, 0, 1, 0, 0, 0, 0]^T, \]
\[ \kappa_4 = [0, 0, 0, 0, 1, 0, 0, 0]^T, \]
\[ \kappa_5 = [0, 0, 0, 0, 0, 1, 0, 0]^T, \]
\[ \kappa_6 = [0, 0, 0, 0, 0, 0, 1, 0]^T, \]
\[ \kappa_7 = [0, 0, 0, 0, 0, 0, 0, 1]^T, \]
\[ \kappa_8 = [-1, 1, 0, 0, 0, 0, 0, 0]^T, \]
(41)

and the left eigenvectors are
\[ \rho_1 = [1, 1, 0, 0, 0, 0, 0, 0]^T, \]  
\[ \rho_i = \kappa_i, \quad i = 2, \ldots, 8, \]
\[ \rho_9 = (\hat{P} \hat{O} + \hat{O} \hat{P})^{-1} \times \left[ -\hat{O} \hat{P}, \hat{P} \hat{O}, 0, 0, 0, 0, 0, 0 \right]^T. \]

We note that the left eigenvector 1 asserted by the identity (40) 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 names 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
\[ \frac{1}{\varepsilon} \frac{da_r}{dt} = \rho_r^T \hat{A}_r(\sigma) \sum_{r'} a_{r'} \kappa_{r'} \]
\[ + \varepsilon \left[ \rho_r^T \hat{A}_r(\sigma) \kappa_9 \beta_9 \right] + \mathcal{O}(\varepsilon^2). \]  
(43)

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 (35). This is because the first-order term vanishes as $\rho_r^T \hat{A}_r(\sigma) \kappa_9 = 0$ 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
\[ \mathbf{a} = [\hat{N}, \hat{Q}, R, S, T, \hat{U}, V, W]^T, \]  
(44)

where $\hat{N} \equiv a_1$, $\hat{Q} \equiv a_2$ and $\hat{U} \equiv a_6$.

The components $r = 1, 2, 6$ in (43), that is differential equations for $a_1 = \hat{N}$, $a_2 = \hat{Q}$ and $a_6 = \hat{U}$, will have nonzero first-order terms. According to $\rho_1$ as given by (42), the new variable $\hat{N}$ 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 $Q$ and $U$ are motivated by the fact that according to (42) 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 (45) then defines the relationship between the original and the reduced variables in the leading order, which in our case is
\[ \mathbf{U} = \left[ \beta_{PO} \hat{N}, \beta_{OP} \hat{N}, \hat{Q}, R, S, T, \hat{U}, V, W \right], \]
(45)

where we define the fractions of the transition rates as
\[ \beta_{\ell} = \frac{\hat{\gamma}_{\ell}}{\hat{P}O + \hat{O}P}. \]
(46)

We have only one stable eigenvalue in the present case, so equation (9) reduces to
\[ \mathbf{v} = b_9 \mathbf{k}_9, \]
(47)

and equation (28), with account of $\kappa_r' = 0$, $r' \neq 1$, gives
\[ b_9 = \frac{\rho_9^T}{\lambda_9} \left( a_1 \kappa_1' - \hat{A}_1(\sigma) \mathbf{U} \right). \]  
(48)

Then the leading-order transition matrix, according to (32, 33) is
\[ \hat{M}_0 = \left[ \begin{array}{cccccccc} * & QP & 0 & 0 & 0 & 0 & (\hat{U}O + \hat{U}P) & 0 & 0 \\ \hat{N} & * & RQ & 0 & TQ & 0 & 0 & 0 \\ 0 & QR & * & SR & 0 & 0 & 0 & 0 \\ 0 & 0 & RS & * & TS & 0 & 0 & 0 \\ 0 & QT & 0 & ST & * & UT & 0 & 0 \\ \hat{N}U & 0 & 0 & 0 & \hat{U}U & * & VU & 0 \\ 0 & 0 & 0 & 0 & \hat{U}V & * & \hat{W}V & 0 \\ 0 & 0 & 0 & 0 & \hat{V}V & * & \hat{W}V & 0 \end{array} \right], \]
(49)

where the new transition rates are defined as
\[ \hat{N}Q = \beta_{OP} \hat{P}Q, \]
\[ \hat{N}U = \beta_{PO} \hat{O}U + \beta_{OP} \hat{P}U, \]
(50)

and the first-order correction to the transition matrix defined by (43, 44) works out as
where the nonzero entries are split over two lines, and horizontal lines between matrix rows are added for clarity.

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

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

where $U^{\tau}$ is given by (45) and $v^{\tau}$ is given by (47), with (48) giving $b_{g}$. This leads to

$$ O = \frac{\tilde{P}O}{P + \beta O P} \tilde{N} - \varepsilon b_{g}, \quad (51) $$

where

$$ b_{g} = \left( \frac{\tilde{P}O}{P + \beta O P} \right)^{-1} \frac{\partial \beta P O}{\partial \sigma} \tilde{N} - \beta O P \beta_{PQ} (\beta_{PQ} + \beta_{PU} - \beta_{OU}) \tilde{N} + \beta O P \beta_{QP} \tilde{Q} + (\beta_{PU} - \beta_{OP} \beta_{UO}) \tilde{U}. \quad (52) $$

Matrix $\tilde{M}_{1}$ and coordinate $b_{g}$ 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.

$$ \frac{d}{d\sigma} \left( \frac{\tilde{P}O(V_{m}(\sigma))}{d\sigma} \right) = \frac{d}{d\sigma} \left( \tilde{P}O(V_{m}(\sigma)) \right) \frac{dV_{m}}{dt}, $$

$$ \frac{d}{d\sigma} \left( \frac{\tilde{P}O(V_{m}(\sigma))}{d\sigma} \right) = \frac{d}{d\sigma} \left( \tilde{P}O(V_{m}(\sigma)) \right) \frac{dV_{m}}{dt}. $$

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 (51,52). For comparison of the reduced model with the embedding, the occupancy of state $N$ 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

$$ \left\| \tilde{N} - \tilde{N}_{ref} \right\| = \left[ \int_{0}^{t_{max}} (\tilde{N}(t) - \tilde{N}_{ref}(t))^{2} dt \right]^{1/2} \quad (53) $$

where $\tilde{N}_{ref}$ is the reference solution obtained for a very small time step, and comparison is done for the interval $t_{max} = 2$ 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.

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

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 $M$. This choice is supported by the empirical embedding procedure as described in Section 4, details can be found in Ref. [10].
Now the matrix of fast transition rates is

\[ \hat{A}_f = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \]  \hspace{1cm} \text{(54)}

and the remaining, slow rates constitute the matrix

\[ \hat{A}_s = \begin{bmatrix}
* & PO & 0 & 0 & 0 & 0 & \underline{UO} & 0 & 0 \\
OP & * & QP & 0 & 0 & 0 & \underline{UP} & 0 & 0 \\
0 & PQ & * & \underline{RQ} & 0 & \underline{TQ} & 0 & 0 & 0 \\
0 & 0 & QR & * & SR & 0 & 0 & 0 & 0 \\
0 & 0 & RS & * & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & QT & 0 & 0 & * & 0 & 0 & 0 \\
0 & OU & PU & 0 & 0 & 0 & * & \underline{VU} & 0 \\
0 & 0 & 0 & UV & \underline{*} & WV & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & WV & * & \underline{WV}
\end{bmatrix} \]  \hspace{1cm} \text{(55)}

FIG. 3. Evolution of state occupancy in OP-embedded (E.) and OP-reduced (R.) model (A-C) and error analysis of OP-reduced model (D). 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 \( \varepsilon = 0.5 \) is shown with green lines, the reduced model without correction term \( \varepsilon = 0 \) is shown with blue lines, the OP-reduced model with correction term for \( \varepsilon = 0.5 \) is shown with magenta lines, and the OP-reduced model with correction term for \( \varepsilon = 1.0 \) is shown with orange lines. Panel (D) shows the order of approximation in \( \varepsilon \) 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 \( \Delta t = 0.01 \text{ ms} \) and simulations of \( \tilde{N}_{\text{ref}} = O + P \) computed with a time step of \( \Delta t = 5 \cdot 10^{-5} \text{ ms} \) in the original model using the same value of \( \varepsilon \). The magenta line represents first-order dependence on \( \varepsilon \) (as a function of \( 2 \cdot 0.5 \cdot \varepsilon^2 \)), blue line represents second-order dependence on \( \varepsilon \) (as a function of \( 2 \cdot 2 \cdot \varepsilon^2 \)). The data are shown on double logarithmic scale.
The right eigenvectors corresponding to zero eigenvalue of this system can be chosen as
\[ \mathbf{\kappa}_1 = \begin{bmatrix} 1, 0, 0, 0, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_2 = \begin{bmatrix} 0, 1, 0, 0, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_3 = \begin{bmatrix} 0, 0, 1, 0, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_4 = \begin{bmatrix} 0, 0, 0, 1, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_5 = (\mathbf{UTTS + UTST + TUST})^{-1} \times \begin{bmatrix} 0, 0, 0, 0, \mathbf{UTTS}, \mathbf{UTST}, \mathbf{TUST}, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_6 = \begin{bmatrix} 0, 0, 0, 0, 0, 0, 1, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_7 = \begin{bmatrix} 0, 0, 0, 0, 0, 0, 0, 1 \end{bmatrix}^T. \]
The corresponding left eigenvectors are
\[ \mathbf{\rho}_1 = \mathbf{\kappa}_i, \quad i = 1, 2, 3, 4, 6, 7, \]
\[ \mathbf{\rho}_5 = \begin{bmatrix} 0, 0, 0, 0, 1, 1, 0, 0 \end{bmatrix}^T. \] (57)
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
\[ \mathbf{a} = [O, P, Q, R, \tilde{M}, V, W]^T. \] (58)
These are all the ingredients needed for the derivation of the leading-order approximation. We have \( \mathbf{\rho}_i' = 0 \) for all \( i = 1, \ldots, 7 \) so \( \mathbf{\rho}_2 \) gives the leading-order transition rate matrix for the reduced model as
\[ \tilde{\mathbf{M}}_0 = \begin{bmatrix} \ast & \mathbf{PO} & 0 & 0 & \mathbf{MO} & 0 & 0 \\ \mathbf{OP} & \ast & \mathbf{QP} & 0 & \mathbf{MP} & 0 & 0 \\ 0 & \mathbf{PQ} & \ast & \mathbf{RQ} & \mathbf{MQ} & 0 & 0 \\ 0 & 0 & \mathbf{QR} & \ast & \mathbf{MR} & 0 & 0 \\ \mathbf{OM} & \mathbf{PM} & \mathbf{QM} & \mathbf{RM} & \ast & \mathbf{VM} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{MV} & \ast & \mathbf{WV} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{WV} & \ast \end{bmatrix}, \] (59)
with the new transition rates defined as
\[ \mathbf{MO} = \mathbf{UO} \gamma_{STTU}, \quad \mathbf{OM} = \mathbf{OU}, \]
\[ \mathbf{MP} = \mathbf{UP} \gamma_{STTU}, \quad \mathbf{PM} = \mathbf{PU}, \]
\[ \mathbf{MQ} = \mathbf{TQ} \gamma_{USTT}, \quad \mathbf{QM} = \mathbf{QT}, \]
\[ \mathbf{MR} = \mathbf{SR} \gamma_{UTTS}, \quad \mathbf{RM} = \mathbf{RS}, \]
\[ \mathbf{MV} = \mathbf{UV} \gamma_{STTU}, \quad \mathbf{VM} = \mathbf{VU}. \]
These expression use the notation \( \gamma_{ijkl} \) as an abbreviation for
\[ \gamma_{ijkl} = \frac{\tilde{\gamma}_{ijkl}}{\mathbf{UTTS + UTST + STTU}}. \] (61)
The original coordinates are recovered from the reduced one by
\[ S = \gamma_{UTTS} \tilde{M}, \] (62a)
\[ T = \gamma_{USTT} \tilde{M}, \] (62b)
\[ U = \gamma_{STTU} \tilde{M}. \] (62c)
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 [59], which will now be called \( M \), while the fast matrix in the new embedding is
\[ \tilde{\mathbf{M}}_f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ast & \mathbf{RQ} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ast & \mathbf{RQ} & 0 \end{bmatrix}, \] (63)
and the slow matrix is
\[ \tilde{\mathbf{M}}_s = \begin{bmatrix} \ast & \mathbf{PO} & 0 & 0 & \mathbf{MO} & 0 & 0 \\ \mathbf{OP} & \ast & \mathbf{QP} & 0 & \mathbf{MP} & 0 & 0 \\ 0 & \mathbf{PQ} & \ast & \mathbf{MR} & 0 & 0 & 0 \\ \mathbf{OM} & \mathbf{PM} & \mathbf{QM} & \mathbf{RM} & \ast & \mathbf{VM} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{MV} & \ast & \mathbf{WV} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{WV} & \ast \end{bmatrix}. \] (64)
Acting as before, we find the right eigenvectors of \( \tilde{\mathbf{M}}_f \) corresponding to zero eigenvalue as
\[ \mathbf{\kappa}_1 = \begin{bmatrix} 1, 0, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_2 = \begin{bmatrix} 0, 1, 0, 0, 0, 0 \end{bmatrix}^T, \]
\[ \mathbf{\kappa}_3 = (\mathbf{QR} + \mathbf{RQ})^{-1} \begin{bmatrix} 0, 0, \mathbf{QR}, \mathbf{RQ}, 0, 0 \end{bmatrix}^T, \]
the corresponding left eigenvectors as

\[ \rho_i = \kappa_i, \quad i = 1, 2, 4, 5, 6, \]
\[ \rho_3 = [0, 0, 0, 0, 0, 1]^T, \]

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

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

The resulting leading-order reduced matrix works out as

\[
\hat{M}_0 = \begin{bmatrix}
    * & \tilde{P}O & 0 & \tilde{M}O & 0 & 0 \\
    \tilde{O}P & * & \tilde{L}P & M\tilde{P} & 0 & 0 \\
    0 & \tilde{P}L & * & \tilde{M}L & 0 & 0 \\
    \tilde{O}M & \tilde{P}M & \tilde{L}M & * & \tilde{V}M & 0 \\
    0 & 0 & 0 & \tilde{M}V & * & \tilde{WV} \\
    0 & 0 & 0 & 0 & \tilde{VW} & *
\end{bmatrix}
\]

with the new transition rates defined by

\[
\tilde{L}P = \tilde{Q}P\delta_{QR}, \quad \tilde{P}L = \tilde{P}Q \quad (66)
\]
\[
\tilde{L}\tilde{M} = \tilde{Q}\tilde{M}\delta_{QR} + \tilde{R}\tilde{M}\delta_{RQ}, \quad \tilde{M}\tilde{L} = \tilde{M}\tilde{Q} + \tilde{M}\tilde{R},
\]

where

\[ \delta_{ij} = \frac{i_j}{QR + RQ}. \quad (67) \]

Fig. 4 presents the results of the QR-STU in addition to the previously considered OP and STU reductions. This was done in simulations where 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 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, our main purpose 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 important phenomena such as conduction block, wavebreaks etc. Hence further increase of the number of the reduced degrees of freedom does not seem to be an answer (although of course further research is needed to establish that with certainty).

An alternative approach, which has proved to be more practical than the one considered here, has been described in our previous works, 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 a 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 $f(u)$ and $h(u)$, such that, after substitution of the sought solution (2), we get the expression on the right hand side of (1) as

$$f(U) + \varepsilon \sum_j \frac{\partial f}{\partial u_j} v_j + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 f}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2}$$

(A1)$$+ \varepsilon h(U) + \varepsilon^2 \sum_{j_1} \frac{\partial h}{\partial u_{j_1}} v_{j_1} + O(\varepsilon^3).$$

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

\[
\varepsilon \hat{F}(\mathbf{U}) \sum_{\ell} b_{\ell} \mathbf{V}_\ell + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} + (A2)
\]

\[
\varepsilon \mathbf{h}(\mathbf{U}) + \varepsilon^2 \sum_{j_1} \frac{\partial \mathbf{h}}{\partial u_{j_1}} v_{j_1} + \mathcal{O}(\varepsilon^3).
\]

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

\[
\sum_k \frac{\partial \mathbf{U}}{\partial a_k} \frac{da_k}{dt} + \varepsilon \sum_{\ell} \left( \frac{d\mathbf{V}_\ell}{dt} + b_{\ell} \frac{d\mathbf{V}_\ell}{da_k} \right) = \sum_k \mathbf{V}_k \frac{da_k}{dt} + (A3)
\]

Combining the right-hand side given by \[\text{[A1]}\] and the left-hand side given by \[\text{[A3]}\], we rewrite \[\text{[1]}\] as

\[
\sum_k \mathbf{V}_k \frac{da_k}{dt} + \varepsilon \sum_{\ell} \left( \frac{d\mathbf{V}_\ell}{dt} + b_{\ell} \frac{d\mathbf{V}_\ell}{da_k} \right) = \varepsilon \hat{F} \sum_{\ell} b_{\ell} \mathbf{V}_\ell + \varepsilon^2 \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2}
\]

\[
+ \varepsilon \mathbf{h}(\mathbf{U}) + \varepsilon^2 \sum_{j_1} \frac{\partial \mathbf{h}}{\partial u_{j_1}} v_{j_1} + \mathcal{O}(\varepsilon^3).
\]

Multiplying the equation by the adjoint eigenvectors \(\mathbf{W}_\ell^T\) gives

\[
\frac{1}{\varepsilon} \frac{da_i}{dt} + \frac{db_i}{dt} = \Lambda_i b_i + \mathbf{W}_i^T \mathbf{h}(\mathbf{U}) \quad (A5)
\]

\[
+ \varepsilon \mathbf{W}_i^T \left[ \sum_j \frac{\partial \mathbf{h}}{\partial u_j} v_j + \sum_{j_1, j_2} \frac{\partial^2 \mathbf{f}}{\partial u_{j_1} \partial u_{j_2}} v_{j_1} v_{j_2} \right]
\]

\[
- \frac{1}{\varepsilon} \sum_{\ell, k} \frac{\partial \mathbf{V}_\ell}{\partial a_k} \frac{da_k}{dt} + \mathcal{O}(\varepsilon^2).
\]

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

1. V. Biktashev, “Envelope equations for modulated non-conservative waves,” IUTAM Symposium Asymptotics, Singularities and Homogenisation in Problems of Mechanics 5(1), 11 (2003).
2. R. Suckley and V. N. Biktashev, “Comparison of asymptotics of heart and nerve excitability,” Phys Rev E Stat Nonlin Soft Matter Phys 68, 011902 (2003).
3. V. N. Biktashev and R. Suckley, “Non-Tikhonov asymptotic properties of cardiac excitability,” Phys Rev Lett 93, 168103 (2004).
4. I. V. Biktasheva, R. D. Simitev, R. Suckley, and V. N. Biktashev, “Asymptotic properties of mathematical models of excitability,” Philos Transact A Math Phys Eng Sci 364, 1283–1298 (2006).
5. 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).
6. 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 and \(\hat{A} = \lim_{\varepsilon \to 0} (\varepsilon \hat{A})\).
7. 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 NP-kappaB system,” J. Math. Biol. 70, 391–620 (2015).
8. C. E. Clancy and Y. Rudy, “Na^+ channel mutation that causes both Brugada and long-QT syndrome phenotypes: a simulation study of mechanism.” Circulation 105, 1208–1213 (2002).
9. T. Stary and V. N. Biktashev, “Exponential integrators for a markov chain model of the fast sodium channel of cardiomyocytes.” IEEE Trans Biomed Eng 62, 1070–1076 (2015).
10. T. Starý, Mathematical and Computational Study of Markovian Models of Ion Channels in Cardiac Excitation, Ph.D. thesis, University of Exeter (2016).
11. T. Stary and V. Biktashev, “Evaluating exponential integrators for Markov chain ion channel models,” Computing in Cardiology 42, 885–888 (2015).
12. W. J. Anderson, Continuous Time Markov Chains. An Application-Oriented Approach (Springer, New York etc, 1991).