J Comput Neurosci (2007) 23:251–264
DOI 10.1007/s10827-007-0035-9

Reconstructing parameters of the FitzHugh–Nagumo system from boundary potential measurements

Yuan He · David E. Keyes

Received: 6 July 2006 / Revised: 13 March 2007 / Accepted: 27 March 2007 / Published online: 10 May 2007

Abstract We consider distributed parameter identification problems for the FitzHugh–Nagumo model of electrocardiology. The model describes the evolution of electrical potentials in heart tissues. The mathematical problem is to reconstruct physical parameters of the system through partial knowledge of its solutions on the boundary of the domain. We present a parallel algorithm of Newton–Krylov type that combines Newton's method for numerical optimization with Krylov subspace solvers for the resulting Karush–Kuhn–Tucker system. We show by numerical simulations that parameter reconstruction can be performed from measurements taken on the boundary of the domain only. We discuss the effects of various model parameters on the quality of reconstructions.

Keywords FitzHugh–Nagumo model · Electrocardiology · Parameter identification · PDE-constrained optimization · KKT system · Newton–Krylov method · Inverse problems

1 Introduction

Parameter identification problems for electrocardiology models have received wide attention in the last decade (Brooks et al. 1999; Cheng et al. 2003; Cox and Griffith 2001; Cox and Ji 2003; Cox and Wagner 2004; MacLeod and Brooks 1998; Moreau-Villéger et al. 2006; Shahidi et al. 1994; Vanier and Bower 1999; Willms et al. 1999). By parameter identification we mean the estimation or reconstruction of coefficients in an electrocardiology model (usually a system of differential equations) from partial observations of the behavior of the model (e.g., the solution of the differential equation system). The knowledge of such coefficients can be very useful for diagnostic purposes. However, they are not directly measurable. Instead, they have to be inferred from measurements of other types of information. We consider in this work a parameter identification problem for the FitzHugh–Nagumo system of equations (FitzHugh 1961; Murray 1993; Nagumo et al. 1962). This system models the evolution of electrical potentials in hearts and other excitable media, such as systems of neurons. Our objective is to reconstruct the reactive coefficient that controls the excitability of the medium. Two special features of our study are as follows. First, instead of identifying only spatially uniform coefficients, which is the focus of much previous research on parameter estimation problems for electrocardiology models (Cox and Wagner 2004; Pernarowski 2001), we try to reconstruct the reactive coefficient as a function of space, i.e., we consider a distributed parameter identification problem (Banks and Kunisch 1989). Second, we assume that only surface measurements are available, instead of requiring measurements from the interior of the heart. Reliable interior measurements are hardly likely to be available in practice.

Our ultimate aim is to develop a numerical method that exploit large-scale parallelism to solve the

Action Editor: David Terman

Y. He (✉) · D. E. Keyes
Department of Applied Physics & Applied Mathematics,
Columbia University, New York, NY 10027, USA
e-mail: yh2030@columbia.edu

D. E. Keyes
e-mail: kd2112@columbia.edu

reconstruction problem. Here, we perform numerical simulations to demonstrate the algorithm in a modestly parallel environment, leaving detailed scaling and performance studies to a forthcoming paper. We emphasize herein the effect of various physical parameters and algorithmic parameters on the quality of the reconstructions.

The paper is structured as follows. In Section 2 we briefly recall some basic results concerning the FitzHugh–Nagumo model and formulate the parameter identification problem for the model. We then propose in detail a numerical procedure to realize the reconstruction in Section 3. Section 4 is devoted to illustrative examples of numerical reconstructions from synthetic data. Conclusions are offered in Section 5.

2 Problem description

We begin with a multidimensional, diffusive version of the classic FitzHugh–Nagumo system of semilinear reaction-diffusion equations.

2.1 The FitzHugh–Nagumo system

Let $\Omega \subset \mathbb{R}^n$ be the spatial domain of interest and $\partial\Omega$ its boundary. We then define $Q \equiv \Omega \times (0, T)$ and $\partial Q \equiv \partial\Omega \times (0, T)$. The FitzHugh–Nagumo model in which we are interested is a two-component reaction-diffusion model

$$\begin{aligned} \partial_t u &= \mu \Delta u + u(u-\alpha)(1-u) - v, \ \text{in } Q, \\ \partial_t v &= \kappa \Delta v + \epsilon(\vartheta u - \gamma v), \qquad \text{in } Q, \end{aligned} \tag{1}$$

where Δ denotes the Laplacian operator and $\partial_t \equiv \partial/\partial t$. Physically, the function $u(\mathbf{x}, t)$ denotes the transmembrane electrical potential at $(\mathbf{x}, t)$ and $v(\mathbf{x}, t)$ denotes a likelihood that a particular class of ion channel is open (Hoffman et al. 1997). The diffusion coefficients satisfy $\kappa \ll \mu$. $\epsilon \ll 1$ and $\vartheta, \gamma > 0$ are nondimensionalized parameters. The nonlinear function $f(u, \alpha) = u(u-\alpha)(1-u)$ represents the reactive properties of the medium. The reactive coefficient $\alpha(\mathbf{x})$, taking values in $(0, 1)$, is a function of space. We assume that the system is at rest at time $t = 0$. In other words, we consider the following initial condition

$$u(\mathbf{x}, 0) = 0, \qquad v(\mathbf{x}, 0) = 0 \quad \text{in } \Omega. \tag{2}$$

A current stimulus $I(\mathbf{x}, t)$ is applied on the boundary of the domain, expressed as an inhomogeneous Neumann boundary condition

$$\mathbf{n} \cdot \nabla u(\mathbf{x}, t) = I(\mathbf{x}, t), \quad \mathbf{n} \cdot \nabla v(\mathbf{x}, t) = 0 \quad \text{on } \partial Q. \tag{3}$$

There are many variants of system (1) in the literature. For example, in some versions with the conductivity coefficient μ a function of space, $\mu\Delta u$ is left in the more general form $\nabla \cdot \mu \nabla u$. In many other versions, the term $\kappa \Delta v$ is dropped. The analysis and algorithms described in the following sections can be adapted to those variants straightforwardly.

FitzHugh–Nagumo models are simplifications of the more complicated bidomain models (Franzone et al. 2005; Patel and Roth 2005; Roth 2004). These have been extensively studied in the past a few decades from both physical point of view (Argentina et al. 2000; Bernus et al. 2002; Dauby et al. 2006; Scott 1975; Suckley and Biktashev 2003; Yamada and Nozaki 1990) and via rigorous mathematical analysis (Chen and Oshita 2006; Colli Franzone and Pavarino 2004; Elmer and Van Vleck 2005; Gao and Wang 2004; Krupa et al. 1997; Murillo and Cai 2004; Nii 1997; Pennacchio et al. 2006; Petersson 2005; Sneyd et al. 1998; Tsai and Sneyd 2005). It has been shown that there exist many types of fast and slow wave solutions in the system. In fact, these models have served as the basis for models of formulation of spiral waves in excitable media such as hearts and systems of neurons which have been of interest to physicists and physicians during last two decades (Bub et al. 2002; Courtemanche et al. 1990; Winfree 1990). Spiral waves are integral to the study of certain diseases of the heart (Davidenko et al. 1992; Riccio et al. 1999; Weiss et al. 1999).

2.2 The reconstruction problem

The parameter identification problem (inverse problem) for the FitzHugh–Nagumo model is to reconstruct the coefficient $\alpha(\mathbf{x})$ from the knowledge of the solution of the equations. Natural and important questions are how much data is available in practice and how well we can do with relatively little data. It turns out that v is not measurable in physical experiments. Only u can be directly measured, via voltage-sensitive probes. Furthermore, it is not convenient to measure u at arbitrary locations. It is realistic to measure u only on the boundary of the domain (surface of the medium). Thus, the data to be measured is

$$h = u(\mathbf{x}, t), \quad \text{at some finite subset of points on } \partial Q. \tag{4}$$

One can measure the boundary potential for various applied current stimuli I. The objective is thus to reconstruct the function $\alpha(\mathbf{x})$ through the knowledge of all available pairs (I, h). Mathematically, the above measurement process is often described by the so called

time-dependent Neumann-to-Dirichlet map (Isakov 1998)

$$\Lambda : I(\mathbf{x}, t) \to h. \tag{5}$$

Note that since we do not have access to $v(\partial\Omega, t)$, $u(\mathbf{x}, T)$ and $v(\mathbf{x}, T)$, the Neumann-to-Dirichlet map (5) is not complete. We have only a partial Neumann-to-Dirichlet map. With the help of this notation, we can now formulate the parameter identification problem as

Identification problem Reconstruct the reactive function $\alpha(\mathbf{x})$ in Eq. (1) from full or partial knowledge of the incomplete Neumann-to-Dirichlet map Λ.

There appear to be no theoretical uniqueness or stability results on this identification problem so far. There are only partial results on much simplified versions. For example, in the case α is assumed to be a constant instead of a function of space (which significantly simplifies the reconstruction problem), it is shown in Cox and Wagner (2004) that one can reconstruct the term $f(u) \equiv f(u, \alpha) = u(u - \alpha)(1 - u)$ from boundary measurements. In the case when a full Neumann-to-Dirichlet map is available, it is shown (Isakov 1998, Theorem 4.2) that $\alpha(\mathbf{x})$ can be uniquely reconstructed. However, full measurements of $u(\mathbf{x}, T)$, $v(\mathbf{x}, T)$ and $v(\partial\Omega, t)$ are unavailable in practice.

The lack of theoretical results on the identifiability and stability of the identification problem is part of the motivation of the current work, in which we seek computational experience and intuition about the reconstruction problem.

3 Matrix-free Newton–Krylov inversion method

In absence of analytical reconstruction formulas, parameter identification problems are usually solved by numerical optimization algorithms. The idea is to minimize the discrepancy between observed data and the data produced by the forward model over the space of admissible parameters. In the case when the forward model is a set of partial differential equations (PDE), the problem can often be formulated as a PDE-constrained optimization problem (Akcelik 2002; Akcelik et al. 2006; Biegler et al. 2003; Biros and Ghattas 2005a,b).

3.1 PDE-constrained optimization

In practical applications, only a finite number of sources and detectors can be used in experimental measurements. Let us assume that we have N_s sources and for each source we have N_d measurement points. We then define the functional to be minimized as

$$\mathcal{F}(\alpha, u) := \frac{1}{2}\sum_{s=1}^{N_s}\sum_{j=1}^{N_d}\int_0^T\int_{\partial\Omega}(u_s - h_s)^2\delta(\mathbf{x}-\mathbf{x}_j) \times d\sigma(\mathbf{x})dt + \rho\mathcal{R}(\alpha), \tag{6}$$

where u_s and h_s are model solutions and measurements corresponding to source I_s, respectively. $\mathbf{x}_j$, $j = 1, ..., N_d$, are detector positions. To simplify notation, we have used $u = (u_1, ..., u_s, ..., u_{N_s})$. $d\sigma$ denotes the surface measure on $\partial\Omega$. $\mathcal{R}(\alpha)$ is a regularization functional. It is used to put additional constraints on the solution α. ρ is the regularization parameter used to establish the strength of the regularization. The regularization used in this paper is the Tikhonov functional. An analytical expression of the functional is given in Section 4.

The reconstruction problem can now be defined as the following PDE-constrained optimization problem

$$\min_{\alpha,u}\mathcal{F}(\alpha, u)$$

subject to

$$\begin{aligned}
&\partial_t u_s = \mu\Delta u_s + f(u_s, \alpha) - v_s, && \text{in } Q,\\
&\partial_t v_s = \kappa\Delta v_s + \epsilon(\vartheta u_s - \gamma v_s), && \text{in } Q,\\
&u_s(\mathbf{x}, 0) = 0, \quad v_s(\mathbf{x}, 0) = 0, && \text{in } \Omega,\\
&\mathbf{n}\cdot\nabla u_s = I_s(\mathbf{x}, t), \quad \mathbf{n}\cdot\nabla v_s = 0, && \text{on } \partial Q,
\end{aligned} \tag{7}$$

for $s = 1, 2, ..., N_s$, where there are N_s source scenarios producing detectable measurements, all starting from rest.

To solve the above constrained minimization problem, it is convenient to use the machinery of distributed Lagrange multipliers. Let us introduce the Lagrangian functional to convert the problem to unconstrained minimization in a higher dimensional space.

$$\mathcal{L}(u, v, \alpha, \lambda, \eta) = \mathcal{F}(\alpha, u) + L_1 + L_2 + L_3 + L_4, \tag{8}$$

where again full sets of variables corresponding to different sources, such as $v = (v_1, ..., v_s, ..., v_{N_s})$, are used. The terms L_1, L_2, L_3 and L_4 are integrals over Q or its boundary:

$$\begin{aligned}
L_1 &= \sum_{s=1}^{N_s}\int_0^T\int_\Omega \lambda_s\left[\partial_t u_s - \mu\Delta u_s - f(u_s, \alpha) + v_s\right]d\mathbf{x}dt,\\
L_2 &= \sum_{s=1}^{N_s}\int_\Omega \lambda_s u_s(\mathbf{x}, 0)d\mathbf{x} + \sum_{s=1}^{N_s}\int_0^T\int_{\partial\Omega}\lambda_s(\mathbf{n}\cdot\nabla u_s - I_s)d\sigma dt,\\
L_3 &= \sum_{s=1}^{N_s}\int_0^T\int_\Omega \eta_s\left[\partial_t v_s - \kappa\Delta v_s - \epsilon(\vartheta u_s - \gamma v_s)\right]d\mathbf{x}dt,\\
L_4 &= \sum_{s=1}^{N_s}\int_\Omega \eta_s v_s(\mathbf{x}, 0)d\mathbf{x} + \sum_{s=1}^{N_s}\int_0^T\int_{\partial\Omega}\eta_s\mathbf{n}\cdot\nabla v_s d\sigma(\mathbf{x})dt,
\end{aligned}$$

with λ_s and η_s denoting the Lagrangian multipliers (adjoint variables) corresponding to u_s and v_s, respectively.

3.2 First-order optimality conditions

The solution to the constrained minimization problem satisfies the first-order optimality condition of the Lagrangian functional, which states that at a local minimum the gradient of the Lagrangian functional must vanish. This condition is also called the the first-order Karush–Kuhn–Tucker (KKT) condition. It results in the nonlinear system of equations:

$$\begin{aligned} &\mathcal{L}_\lambda(u, v, \alpha, \lambda, \eta) = 0, \qquad \mathcal{L}_\eta(u, v, \alpha, \lambda, \eta) = 0, \\ &\mathcal{L}_u(u, v, \alpha, \lambda, \eta) = 0, \qquad \mathcal{L}_v(u, v, \alpha, \lambda, \eta) = 0, \\ &\mathcal{L}_\alpha(u, v, \alpha, \lambda, \eta) = 0. \end{aligned} \tag{9}$$

Here the subscript u (respectively v, α, λ and η) denotes differentiation. Considering the Lagrangian functional Eq. (8), we can work out the explicit form of the optimality conditions in Eq. (9) by the method of calculus of variations. In fact, the first two optimality conditions reproduce the constraints in the optimization problem, i.e., the FitzHugh–Nagumo model (7). The third and the fourth optimality conditions form the adjoint equations for the constraints. We refer readers to Cox (2006) for more details about the adjoint method. These adjoint equations and their side conditions are

$$\begin{aligned} &\partial_t\lambda_s + \mu\Delta\lambda_s - f'(u_s, \alpha)\lambda_s + \epsilon\vartheta\eta_s = 0, && \text{in } Q, \\ &\partial_t\eta_s + \kappa\Delta\eta_s - \epsilon\gamma\eta_s - \lambda_s = 0, && \text{in } Q, \\ &\lambda_s(\mathbf{x}, T) = 0, \qquad \eta_s(\mathbf{x}, T) = 0, && \text{in } \Omega, \\ &\mathbf{n}\cdot\nabla\lambda_s = \textstyle\sum_{j=1}^{N_d}(u_s - h_s)\delta(\mathbf{x} - \mathbf{x}_j), \ \mathbf{n}\cdot\nabla\eta_s = 0, && \text{on } \partial Q, \end{aligned} \tag{10}$$

for $s = 1, \ldots, N_s$. Here $f'(u_s, \alpha) \equiv \partial_{u_s} f = 3u_s^2 - 2(\alpha + 1)u_s + \alpha$. The adjoint equations are similar to the state equations with the important difference that they march backwards in time and are driven by the mismatch in the objective function. Note that just as with the forward problems, the adjoint equations for different s are independent of each other. However, each of them depends on the solution of the corresponding forward problem.

The last optimality condition is often referred to as the control problem. It reads

$$\rho\mathcal{R}'(\alpha) + \sum_{s=1}^{N_s}\int_0^T \lambda_s u_s(1 - u_s)dt = 0 \tag{11}$$

in our case. $\mathcal{R}'(\alpha)$ denotes the derivative of $\mathcal{R}(\alpha)$ with respect to α. Note that unlike the forward and adjoint problems, the control problem is not a differential equation.

In summary, the first-order optimality conditions of the Lagrangian functional consist of the set of multicomponent forward problems (7), the set of multicomponent adjoint problems (10), and the control problem (11).

3.3 The Newton–Krylov method

The optimality conditions make up a set of quasilinear partial differential equations of reaction-diffusion type and algebraic equations. The solution of this set of equations requires not only the discretization of the PDEs, but also algebraic solvers for the resulting nonlinear algebraic systems. The Newton–Krylov family of methods provides an efficient way to solve such PDE systems (Knoll and Keyes 2004).

Let us briefly recall Newton's method for nonlinear problems. To compress the notation, we denote $(u, v, \alpha, \lambda, \eta)$ by $\mathbf{u}$, and rewrite the first-order optimality condition (9) as:

$$\mathcal{L}_\mathbf{u}(\mathbf{u}) = 0. \tag{12}$$

To find a solution of Eq. (12), Newton's method iteratively improves a given value of $\mathbf{u}$, from an initial guess $\mathbf{u}_0$, updating $\mathbf{u}$ according to

$$\mathbf{u}_{k+1} = \mathbf{u}_k + l_k\delta\mathbf{u}_k, \tag{13}$$

until certain stopping criteria are satisfied. Here, the update direction $\delta\mathbf{u}_k$ at Newton iteration k is given by solving the saddle point problem

$$\mathcal{L}_{\mathbf{uu}}(\mathbf{u}_k)\delta\mathbf{u}_k = -\mathcal{L}_\mathbf{u}(\mathbf{u}_k), \tag{14}$$

and the step length l_k is given by a line search method or other globalization technique. Equation (14) is often referred to as the KKT system. We need to solve one KKT system per Newton iteration. The operator $\mathcal{L}_{\mathbf{uu}}$ is the full-space Hessian operator (a matrix, in discretized form).

For the discretized version of a distributed parameter identification problem, the KKT system (14) is a large sparse linear system. It can be solved by any of several numerical iterative solvers. When a Krylov subspace method is used to solve the system, we call the nested iteration method Newton–Krylov method (Knoll and Keyes 2004). The method has received wide attention in recent years; see the references cited in Knoll and Keyes (2004).

For the FitzHugh–Nagumo model we consider here, the KKT system has the form

$$\begin{pmatrix} \mathcal{L}_{uu} & 0 & \mathcal{L}_{u\alpha} & \mathcal{L}_{u\lambda} & \mathcal{L}_{u\eta} \\ 0 & 0 & 0 & \mathcal{L}_{v\lambda} & \mathcal{L}_{v\eta} \\ \mathcal{L}_{\alpha u} & 0 & \mathcal{L}_{\alpha\alpha} & \mathcal{L}_{\alpha\lambda} & 0 \\ \mathcal{L}_{\lambda u} & \mathcal{L}_{\lambda v} & \mathcal{L}_{\lambda\alpha} & 0 & 0 \\ \mathcal{L}_{\eta u} & \mathcal{L}_{\eta v} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \\ \delta\alpha \\ \delta\lambda \\ \delta\eta \end{pmatrix} = - \begin{pmatrix} \mathcal{L}_u \\ \mathcal{L}_v \\ \mathcal{L}_\alpha \\ \mathcal{L}_\lambda \\ \mathcal{L}_\eta \end{pmatrix}. \tag{15}$$

where

$$\delta u = \begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \vdots \\ \delta u_{N_s} \end{bmatrix}, \delta v = \begin{bmatrix} \delta v_1 \\ \delta v_2 \\ \vdots \\ \delta v_{N_s} \end{bmatrix}, \delta \lambda = \begin{bmatrix} \delta \lambda_1 \\ \delta \lambda_2 \\ \vdots \\ \delta \lambda_{N_s} \end{bmatrix}, \delta \eta = \begin{bmatrix} \delta \eta_1 \\ \delta \eta_2 \\ \vdots \\ \delta \eta_{N_s} \end{bmatrix}, \tag{16}$$

and, the $\mathcal{L}_u$, $\mathcal{L}_v$, $\mathcal{L}_\lambda$, $\mathcal{L}_\eta$ are similarly defined. Because the forward problems for different sources are decoupled, the operator $\mathcal{L}_{uu}$ has diagonal structure:

$$\mathcal{L}_{uu} = \begin{pmatrix} \mathcal{L}_{u_1 u_1} & & & \\ & \mathcal{L}_{u_2 u_2} & & \\ & & \ddots & \\ & & & \mathcal{L}_{u_{N_s} u_{N_s}} \end{pmatrix}, \tag{17}$$

and similarly for operators $\mathcal{L}_{u\lambda}$, $\mathcal{L}_{u\eta}$, $\mathcal{L}_{v\lambda}$ and $\mathcal{L}_{v\eta}$ and their adjoint operators, $\mathcal{L}_{\lambda u}$, $\mathcal{L}_{\eta u}$, $\mathcal{L}_{\lambda v}$ and $\mathcal{L}_{\eta v}$.

Operators $\mathcal{L}_{u\alpha}$ and $\mathcal{L}_{\lambda\alpha}$ have the structure

$$\mathcal{L}_{u\alpha} = \begin{bmatrix} \mathcal{L}_{u_1\alpha} \\ \mathcal{L}_{u_2\alpha} \\ \vdots \\ \mathcal{L}_{u_{N_s}\alpha} \end{bmatrix}, \quad \text{and} \quad \mathcal{L}_{\lambda\alpha} = \begin{bmatrix} \mathcal{L}_{\lambda_1\alpha} \\ \mathcal{L}_{\lambda_2\alpha} \\ \vdots \\ \mathcal{L}_{\lambda_{N_s}\alpha} \end{bmatrix}. \tag{18}$$

$\mathcal{L}_{\alpha u}$ and $\mathcal{L}_{\alpha\lambda}$ are their adjoint operators, respectively.

3.4 The Hessian-free reduced-space algorithm

In each Newton iteration, one can, in principle, solve the the KKT system as one large system and update u, v, α, λ and η simultaneously. In this way, u_s, v_s and α do not necessarily satisfy the forward equation (with source I_s) during each Newton iteration but only at convergence. In other words, this method progresses through a PDE-infeasible set of u_s,v_s and α (for $s = 1, ..., N_s$). This approach is used in (Biros and Ghattas 2005a,b). For time-dependent problems, this full-space method seems impractical because of the huge storage requirement (Akcelik 2002).

We adopt a reduced-space method in this work. In each Newton iteration, for a given α, we first solve the FitzHugh–Nagumo model (7). We then solve the adjoint problem (10). After this, the terms $\mathcal{L}_u$, $\mathcal{L}_v$, $\mathcal{L}_\lambda$ and $\mathcal{L}_\eta$ vanish in the KKT system. The KKT system thus becomes

$$\begin{pmatrix} \mathcal{L}_{uu} & 0 & \mathcal{L}_{u\alpha} & \mathcal{L}_{u\lambda} & \mathcal{L}_{u\eta} \\ 0 & 0 & 0 & \mathcal{L}_{v\lambda} & \mathcal{L}_{v\eta} \\ \mathcal{L}_{\alpha u} & 0 & \mathcal{L}_{\alpha\alpha} & \mathcal{L}_{\alpha\lambda} & 0 \\ \mathcal{L}_{\lambda u} & \mathcal{L}_{\lambda v} & \mathcal{L}_{\lambda\alpha} & 0 & 0 \\ \mathcal{L}_{\eta u} & \mathcal{L}_{\eta v} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \\ \delta\alpha \\ \delta\lambda \\ \delta\eta \end{pmatrix} = - \begin{pmatrix} 0 \\ 0 \\ \mathcal{L}_\alpha \\ 0 \\ 0 \end{pmatrix}. \tag{19}$$

We can now perform a Gauss elimination on Eq. (19) to eliminate the δu, δv, $\delta\lambda$ and $\delta\eta$ to obtain

$$\mathcal{H}_{red}\delta\alpha = -\mathcal{L}_\alpha, \tag{20}$$

where the reduced gradient $\mathcal{L}_\alpha$ is given by

$$\mathcal{L}_\alpha = \sum_{s=1}^{N_s} \int_0^T \lambda_s u_s (1 - u_s) dt + \rho \mathcal{R}'(\alpha) \tag{21}$$

and the reduced Hessian $\mathcal{H}_{red}$ is given by

$$\mathcal{H}_{red} = \mathcal{L}_{\alpha\alpha} - \mathcal{L}_{\alpha u} W - W^* \mathcal{L}_{u\alpha} + W^* \mathcal{L}_{uu} W, \tag{22}$$

with W defined as

$$W = [\mathcal{L}_{\lambda u}^{-1} + \mathcal{L}_{\lambda u}^{-1}(\mathcal{L}_{\eta v} - \mathcal{L}_{\eta u}\mathcal{L}_{\lambda u}^{-1})^{-1}\mathcal{L}_{\eta u}\mathcal{L}_{\lambda u}^{-1}]\mathcal{L}_{\lambda\alpha}. \tag{23}$$

Here W^* denotes the adjoint of W. The reduced Hessian $\mathcal{H}_{red}$ has a much smaller size (and is much denser) than the original Hessian $\mathcal{L}_{\mathbf{uu}}$. It is easy to see that $\mathcal{H}_{red} = \mathcal{H}_{red}^*$, that is, $\mathcal{H}_{red}$ is a self-adjoint operator.

We can now solve the reduced KKT system (20) to find an update for α, $\delta\alpha$. The overall process is repeated until we converge to α. Since u_s, v_s and α satisfy the forward equation in each Newton iteration, it is apparent that we are solving the problem on a feasible set of u_s, v_s and α. By doing this, we actually convert the constrained full-space problem into an unconstrained problem for α.

Though the reduced Hessian operator is self-adjoint, it is not guaranteed to be positive definite, and it requires special care. Here, we try to obtain positive definiteness by using the Gauss–Newton approximation (Haber et al. 2000; Nocedal and Wright 1999), where we drop the second derivative information in the $\mathcal{L}_{\alpha u}$ and $\mathcal{L}_{u\alpha}$ terms. The reduced Hessian can now be replaced with:

$$\mathcal{H}_{red}^{GN} = \mathcal{L}_{\alpha\alpha} + W^* \mathcal{L}_{uu} W. \tag{24}$$

For those cases in which the forward model is linear, the operator $\mathcal{L}_{uu}$ is positive definite, and so is $\mathcal{H}_{red}^{GN}$. Here, because of the cubic nonlinear term in the forward problem, instead of being a positive definite Boolean operator as usual, $\mathcal{L}_{uu}$ in this problem is defined by the action:

$$(\mathcal{L}_{uu}\hat{u})\tilde{u} = \sum_{s=1}^{N_s}\sum_{j=1}^{N_d} \int_0^T \int_{\partial\Omega} \hat{u}_s \tilde{u}_s \delta(\mathbf{x} - \mathbf{x}_j) d\sigma(\mathbf{x}) dt + \sum_{s=1}^{N_s} \int_0^T \int_\Omega \hat{u}_s \tilde{u}_s \lambda_s (6u_s - 2\alpha - 2) d\mathbf{x} dt, \tag{25}$$

where again $\hat{u} \equiv (\hat{u}_1, ..., \hat{u}_{N_s})$ and $\tilde{u} \equiv (\tilde{u}_1, ..., \tilde{u}_{N_s})$. One can see that $\mathcal{L}_{uu}$ is *not* in general positive, since the second term in Eq. (25) is not necessarily positive. To obtain positive definiteness for the reduced Hessian

operator, we also need to suppress the second term in Eq. (25); then Eq. (24) becomes

$$\mathcal{H}_{red}^{GN-PD} = \mathcal{L}_{\alpha\alpha} + W^* \mathcal{L}_{uu}^B W, \tag{26}$$

where $(\mathcal{L}_{uu}^B \hat{u})\tilde{u} \equiv \sum_{s=1}^{N_s} \sum_{j=1}^{N_d} \int_0^T \int_{\partial\Omega} \hat{u}_s \tilde{u}_s \delta(\mathbf{x} - \mathbf{x}_j) d\sigma(\mathbf{x}) dt.$

We are able to use the conjugate gradient (CG) method (Saad 2003) to solve the positive definite form of KKT operator. However, the positive definiteness of the KKT operator is not necessary if we are willing to use the generalized minimal residual (GMRES) method (Saad 2003) to solve the KKT system.

We thus come to the following reduced-space Newton's algorithm. We will denote by $\mathcal{F}(\alpha)$ the objective function $\mathcal{F}(\alpha, u)$ introduced in Eq. (6).

Algorithm 1 Reduced-space Newton algorithm

set k_{max}, ε_1, ε_2
guess $\alpha_0(\mathbf{x})$; set $k = 0$
evaluate $\mathcal{F}(\alpha_0)$
while $(k < k_{max}$ & $\frac{\|\mathcal{L}_{\alpha_k}\|}{\|1+\mathcal{F}(\alpha_k)\|} > \varepsilon_1$ & $\frac{\mathcal{F}(\alpha_k)}{\mathcal{F}(\alpha_0)} > \varepsilon_2)$ **do**
  evaluate $\mathcal{L}_{\alpha_k}$ by Eq. (21)
  compute $\delta\alpha_k$ by Eq. (20)
  compute l_k by a line search
  $\alpha_{k+1} = \alpha_k + l_k \delta\alpha_k$
  evaluate $\mathcal{F}(\alpha_{k+1})$
  $k = k + 1$
end while

The reduced gradient $\mathcal{L}_{\alpha_k}$ can be computed using the following algorithm.

Algorithm 2 Computing reduced gradient

set $s = 1$, $\mathcal{L}_{\alpha_k} = 0$
while $(s \le N_s)$ **do**
  compute u_s
  compute λ_s
  compute $\mathcal{L}_{\alpha_k} = \mathcal{L}_{\alpha_k} + \int_0^T \lambda_s u_s (1 - u_s) dt$
  $s = s + 1$
end while
compute $\mathcal{R}'(\alpha_k)$
$\mathcal{L}_{\alpha_k} = \mathcal{L}_{\alpha_k} + \rho \mathcal{R}'(\alpha_k)$

As seen above, the reduced Hessian has a rather complicated form. In practice, it is hard to build this Hessian operator (a matrix in discrete form) explicitly. Fortunately, there is a well-known way around building it explicitly in solving the reduced KKT system (20). The idea is to use Krylov subspace methods to solve the linear system. Instead of requiring explicitly the reduced Hessian operator, the Krylov subspace methods require only the action of the Hessian on functions. In discrete form, they require only matrix-vector products without reference to the full matrix. This property allows one to build the so called matrix-free method (Knoll and Keyes 2004) to solve the system (20).

To see how we can calculate the action of $\mathcal{H}_{red}$ on an arbitrary vector ζ, let us write

$$\mathcal{H}_{red}^{GN-PD} \zeta = \mathcal{L}_{\alpha\alpha}\zeta + W^* \mathcal{L}_{uu}^B W\zeta. \tag{27}$$

Let us now introduce the auxiliary variables w_1^s, w_2^s and w_3^s $(s = 1, 2, \cdots, N_s)$ that solve the systems

$$\begin{cases} \mathcal{L}_{\lambda_s u_s} w_1^s = \mathcal{L}_{\lambda_s \alpha}\zeta \\ \mathcal{L}_{\eta_s v_s} w_2^s - \mathcal{L}_{\eta_s u_s} w_3^s = \mathcal{L}_{\eta_s u_s} w_1^s \\ \mathcal{L}_{\lambda_s u_s} w_3^s = w_2^s \end{cases}. \tag{28}$$

We then define the variable $Y_1^s(\mathbf{x}, t) = -\mathcal{L}_{u_s u_s}^B (w_1^s + w_3^s)$ and introduce further the auxiliary variables w_4^s, w_5^s and w_6^s $(s = 1, 2, \cdots, N_s)$ that solve the systems

$$\begin{cases} \mathcal{L}_{u_s \lambda_s} w_4^s = Y_1^s \\ \mathcal{L}_{v_s \eta_s} w_5^s - w_6^s = w_4^s \\ \mathcal{L}_{u_s \lambda_s} w_6^s = \mathcal{L}_{u_s \eta_s} w_5^s \end{cases}. \tag{29}$$

We then define $Y_2^s(\mathbf{x}) = \mathcal{L}_{\alpha\lambda_s}(w_4^s + w_6^s)$. We can then show that

$$\mathcal{H}_{red}^{GN-PD} \zeta = \mathcal{L}_{\alpha\alpha}\zeta - \sum_{s=1}^{N_s} Y_2^s. \tag{30}$$

We finally arrive at the following algorithm for computing Hessian-function product in the k^{th} Newton iteration.

Algorithm 3 Hessian-vector product

set $s = 1$, $H = 0$
while $(s \le N_s)$ **do**
  solve Eq. (28), store Y_1^s
  solve Eq. (29), store Y_2^s
  $H = H\text{-}Y_2^s$
  $s = s + 1$
end while
compute $\mathcal{R}''(\alpha_k)\zeta$
$\mathcal{H}_{red}^{GN-PD} \zeta = H + \rho \mathcal{R}''(\alpha_k)\zeta$

The explicit expressions for the various operators introduced in this section can be found in the Appendix.

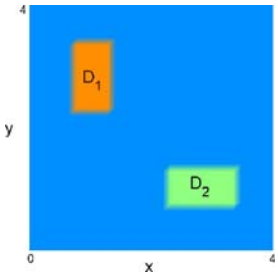

Fig. 1 Geometrical configuration for the reconstructions

3.5 Numerical implementations

We present here some details about the numerical implementation of the algorithms introduced above.

3.5.1 Discretization of the PDEs

All the partial differential equations involved in the above methods have been discretized using elementary finite differences for the spatial and time variables. For the adjoint equations, we first perform a change of variable $t = T - t'$. Then the adjoint equations become initial boundary value problems for time variable t'. These new equations have structure similar to the forward problems and are discretized using the same methods as those forward problems.

3.5.2 Line search for optimization

We implemented both cubic line search and quadratic line search (Dennis et al. 1996; Nocedal and Wright 1999) methods to find the step length l_k in Newton iteration k. We did not observe much difference between the performance of the two methods in our simulations. We enforce the Wolfe conditions (Nocedal and Wright 1999) on both line search methods. That is, we look for an l_k that solves the one-dimensional minimization problem

$$\min_{l_k>0} \mathcal{F}(\alpha_k + l_k \delta\alpha_k), \tag{31}$$

and satisfies:

$$\mathcal{F}(\alpha_k + l_k \delta\alpha_k) \leq \mathcal{F}(\alpha_k) + c_1 l_k \nabla\mathcal{F}^T(\alpha_k)\delta\alpha_k, \tag{32}$$

$$\nabla\mathcal{F}^T(\alpha_k + l_k \delta\alpha_k)\delta\alpha_k \geq c_2 \nabla\mathcal{F}^T(\alpha_k)\delta\alpha_k. \tag{33}$$

Here c_1 and c_2 are two small positive constants. In our simulations, we take the values suggested in Nocedal and Wright (1999), these are, $c_1 = 10^{-4}$ and $c_2 = 0.1$.

3.5.3 Regularization

We adopt the Tikhonov regularization (Engl et al. 1996) in this study. The Tikhonov functional we take here is

$$\mathcal{R}(\alpha) = \frac{1}{2}\int_\Omega \nabla\alpha \cdot \nabla\alpha \, d\mathbf{x}. \tag{34}$$

It is straightforward to check that the first- and second-order derivatives of this functional with respect to α are defined by

$$\mathcal{R}'\hat{\alpha} = -\int_\Omega \hat{\alpha}\nabla \cdot \nabla\alpha d\mathbf{x} + \int_{\partial\Omega} \hat{\alpha}\mathbf{n} \cdot \nabla\alpha d\sigma, \tag{35}$$

and

$$(\mathcal{R}''\tilde{\alpha})\hat{\alpha} = -\int_\Omega \hat{\alpha}\nabla \cdot \nabla\tilde{\alpha} d\mathbf{x} + \int_{\partial\Omega} \hat{\alpha}\mathbf{n} \cdot \nabla\tilde{\alpha} d\sigma, \tag{36}$$

respectively. The regularization parameter, ρ, is chosen by the L-curve method (Vogel 2002). A simple *continuation* method is adopted to reduce the computational cost of choosing this parameter; see the discussion in Section 4.4.

3.5.4 Linear solvers

We use the GMRES algorithm (Saad 2003) to solve all the linear systems encountered in this paper. Because the reduced Hessian $\mathcal{H}_{red}$ is symmetric and positive-definite after the Gauss–Newton approximation, we also consider using the conjugate gradient method. The choice of linear solvers does not have much impact on the quality of the reconstructions, as observed in our simulations. The rates of the convergence for the two methods vary only slightly from case to case.

3.5.5 Parallelization

For the time-dependent FitzHugh–Nagumo model we consider in this paper, solving KKT system requires

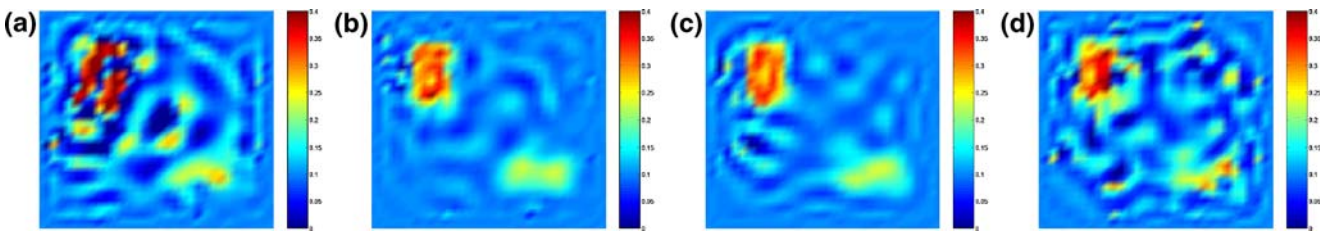

Fig. 2 Reconstructed $\alpha(\mathbf{x})$ with different initial guesses (**a**)-(**d**) reconstructed $\alpha(\mathbf{x})$ with initial guess 0.075, 0.1, 0.125, 0.15, respectively

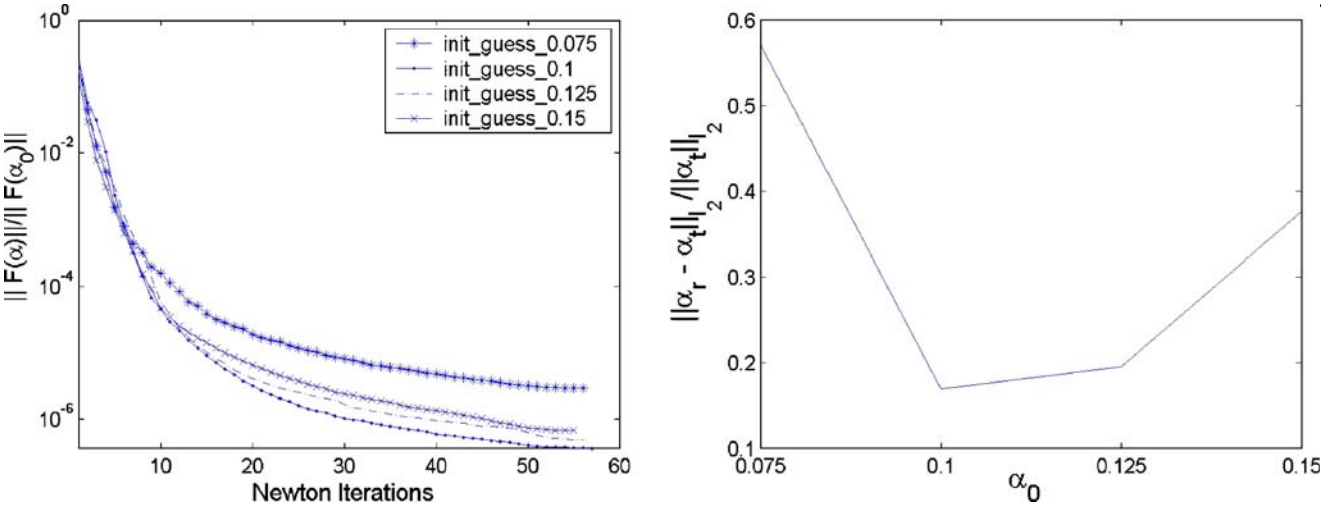

Fig. 3 *Left*: Convergence history of the reconstructions in Fig. 2; *Right*: Relative error as function of the values of homogeneous initial guesses

a significant amount of memory and CPU resources; thus parallel solution techniques are unavoidable. Both the PDEs (after discretization) and the KKT systems are solved on parallel processors in our simulations. The algorithms are implemented in the PETSc software (Balay et al. 2007).

We finally remark that the algorithm we have developed here is an "optimization first and discretization second" type of method. We formulate everything on a continuous level. Alternatively, one can discretize the FitzHugh–Nagumo equation first and then perform minimization on the discretized system.

4 Reconstructions with synthetic data

In this section, we test our numerical algorithms by performing reconstructions under various situations. We will focus on two-dimensional ($n = 2$) reconstructions, although our algorithms are designed for both 2D and 3D cases. The goal is to reconstruct the distributed parameter $\alpha(\mathbf{x})$ with only boundary measurements of time-dependent electrical potential $u(\mathbf{x}, t)$. We use eight sources ($N_s = 8$) in the experiments of this section except Section 4.2. All the potential "measurements" used in the following simulations are synthetic data.

The configuration of the reconstructions is as follows. We try to reconstruct two small inhomogeneities in a homogeneous background. More precisely, the "true" α is given by

$$\alpha(\mathbf{x}) = \begin{cases} 0.3 & \mathbf{x} \in D_1, \\ 0.2 & \mathbf{x} \in D_2, \\ 0.1 \text{ otherwise}, \end{cases} \tag{37}$$

where D_1 and D_2 are small rectangles as shown in Fig. 1.

4.1 Effect of initial guess

We first consider the reconstruction of $\alpha(\mathbf{x})$ with different initial guesses. Since this is a highly nonlinear inverse problem, the objective functional might have multiple local minima. We expect that when the initial guess is too far from the true value, we will have no guarantee of being able to recover the objects.

We present in Fig. 2 the reconstruction of two rectangular inclusions starting from different values of homogeneous initial guess of α.

We find that experimentally the initial guess must be relatively close to the true solution to obtain reasonable reconstructions. This is not very surprising because the problem is highly nonlinear. There exist various local

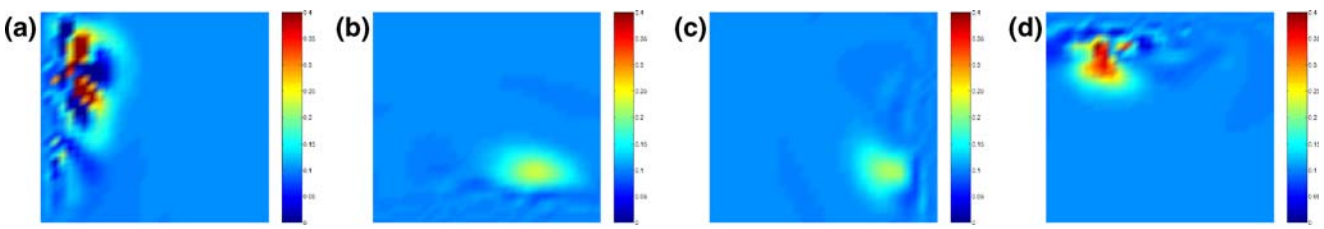

Fig. 4 Reconstructed $\alpha(\mathbf{x})$ with data only on the (**a**) *left*, (**b**) *bottom*, (**c**) *right* and (**d**) *top* part of the boundary

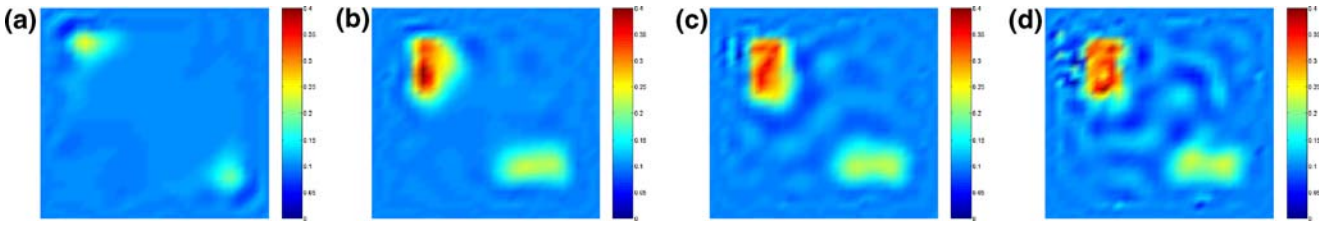

Fig. 5 Reconstructed $\alpha(\mathbf{x})$ with data in different time intervals. (**a**) $t \in [0, T/4]$; (**b**) $t \in [0, T/2]$; (**c**) $t \in [0, 3T/4]$; (**d**) $t \in [0, T]$

minima of the objective functional. Multilevel techniques have recently been designed to deal with this problem (Akcelik 2002; Benzi et al. 2006), and should be further investigated in this context. In practice, *a priori* information on the distribution of α may be available. One can always guess around the real distribution.

Figure 3 shows the convergence history of the Newton's method and relative l_2 error for the reconstructions in Fig. 2. If we denote by α_t the true value of α and by α_r the reconstructed α, then the relative l_2 error is defined here as $\|\alpha_r - \alpha_t\|_{l_2}/\|\alpha_t\|_{l_2}$.

The inverse problem is in some sense initial guess-dependent.

4.2 Effect of measurement locations

We now consider the reconstruction problem with data on only part of the boundary. This is useful when some parts of the boundary are practically unavailable. We consider four different cases: data from left, bottom, right and top part of the boundary only. Note that in all the cases, the localized sources are also placed on the part of boundary where measurements are taken.

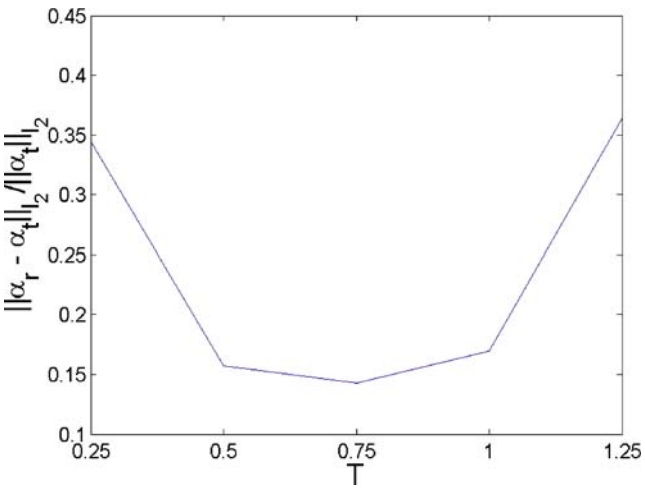

Fig. 6 Convergence of relative l_2 error with respect to the length of measurement time

The simulation results are presented in Fig. 4. The relative l_2 errors in the reconstructions are 0.69, 0.35, 0.36 and 0.30, respectively.

We observe that only the inclusion close to the detector positions can be reconstructed. This is because the forward solutions are traveling waves and these traveling waves collect information of the underlying medium when they propagate through the medium. The detectors on only part of boundary (where the source is) cannot get enough information of the medium since the traveling waves propagate away. There is only small wave energy, scattering back to and being received by the detectors.

4.3 Effect of propagation time

It is obvious that one needs the potential measurements on a sufficiently long time interval to obtain reasonable reconstructions. We present in Fig. 5 reconstructions with full boundary measurements on different time intervals.

The reconstructions measured over too short a time interval are not as accurate as those reconstructions over longer time. We also observe from our numerical simulation that the quality of the reconstructions does not improve anymore when we take measurements over period longer than T, see Fig. 6, because the traveling waves in the forward model starting on one side of the domain propagate cross the opposite side of the rectangular domain at time T. After time T, these traveling waves just propagate away without reflecting back.

4.4 Effect of noise

We now consider the effect of noise on reconstructions, by polluting our synthetic data with uniformly distributed noise. We consider here only multiplicative noise. In other words, if h is our measurement without noise, we then use $h(1 + \varsigma * \text{random})$ as our noisy data, where "random" is a random number in $[-1, 1]$ with uniform distribution.

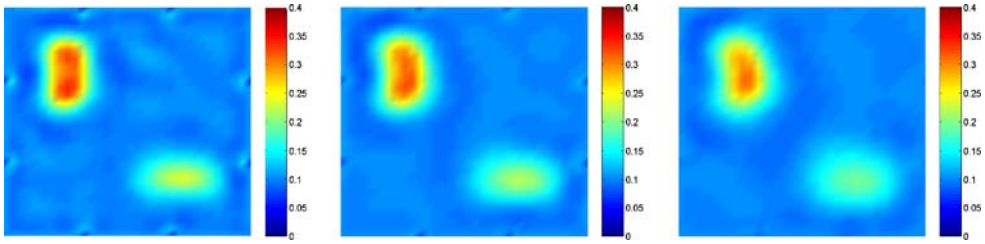

Fig. 7 Reconstructions of $\alpha(\mathbf{x})$ with data of various noise levels. From left to right: ς =1, 3 and 5%, respectively

We present in Fig. 7 the reconstructions with different noise levels. Those reconstructions are done with the optimal regularization parameters which are chosen with the L-curve method (Vogel 2002). Although there exist proofs that the L-curve method fails to converge for certain classes of inverse problems (Vogel 1996), we have observed satisfactory results in our applications. We plot the log of the regularization functional against the squared norm of the regularized residual for a range of values of the regularization parameter. The optimal parameter ρ is the one at which the L-curve reaches the maximum of its curvature (Hansen and O'Leary 1993; Vogel 2002).

Since the L-curve method requires several reconstructions for any single problem, it is very time-consuming. Here we adopt a simple *continuation* method proposed in Haber et al. (2000) to reduce the computational cost of regularization parameter selection.

We start the first reconstruction with a relatively large ρ. The result of this reconstruction is then taken to be the initial guess of next reconstruction with a smaller ρ. If the two ρ are not dramatically different from each other, then the two reconstructions should converge to similar results. Thus, the reconstruction with smaller ρ is supposed to converge fast since its initial guess is chosen to be close enough to its true solution. The process can be repeated to perform reconstructions with several values of ρ. In practice, this continuation method saves tremendous computational time for finding optimal regularization parameters. We present in Fig. 8 the L-curve we have used in the reconstructions of Fig. 7 to choose the optimal regularization parameter ρ.

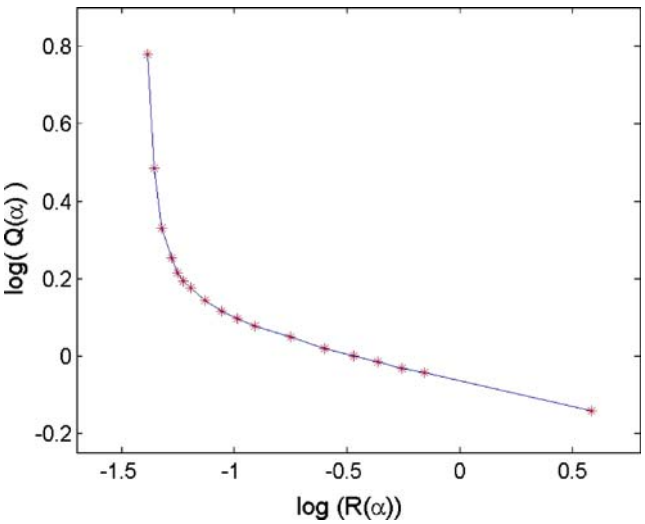

Fig. 8 L-curve used to choose the optimal regularization parameter for the reconstruction with data of 1% noise in Fig. 7. Here $\mathcal{Q}(\alpha) \equiv \mathcal{F}(\alpha, u) - \rho\mathcal{R}(\alpha)$

4.5 Reconstructions in two states

As we know, the FitzHugh–Nagumo model has two states depending on the injection current stimuli: a stable rest state and a stable traveling wave. We now discuss the effect of source strength on the quality of reconstruction.

We find that the reconstruction qualities are very different when the forward problems belong two different states. The forward solution cannot be a stable traveling wave when the injection current stimulus is too small, which means that the detectors on the boundaries cannot receive any information of the underlying medium. Thus the inclusions in the α domain can not be reconstructed, see Fig. 9(a). On the contrary, the reconstruction quality is good when the forward solution is a stable traveling wave, see Fig. 9(b),(c). In this case the information of the medium can be propagated to the detectors through the traveling wave. However, for those cases that forward solutions are stable traveling waves, the reconstruction qualities are similar independent of amplitude.

4.6 Effect of inexactness

We now consider the reconstructed α with different maximum GMRES iterations, that is, when the inner linear iterations are truncated well short of the solution. This is inspired by the observation that very accurate solution of the update $\delta\alpha(\mathbf{x})$ requires a large number

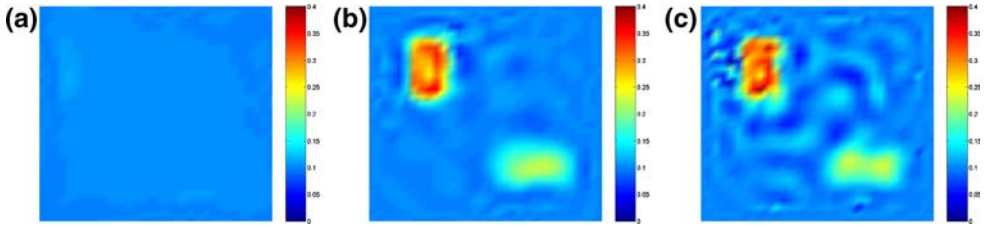

Fig. 9 Reconstruction with different source strengths; (**a**) $I_s = 0.5$; (**b**) $I_s = 0.8$; (**c**) $I_s = 100$

of Krylov iterations. However, with a much smaller number of iterations, we can achieve similar accuracy already as we can see from the left plot in Fig. 10.

We perform reconstructions with two different Krylov solvers, one with maximum iteration number 30 and the other with maximum iteration number 60. The results are presented in Fig. 11. The evolution of the objective function in the reconstruction process for the two reconstructions are shown in the right plot of Fig. 10.

There are no major differences observed between these reconstructions. This is important in practice since it allows one to do reconstruction with similar quality using much less computational resources (by saving extra Krylov iterations, which, in the case of GMRES, cost both additional work and memory). For example, in the two reconstructions here, the one with large Krylov iteration bound takes almost twice as much CPU time as the one with small Krylov iteration bound. Of course, if the maximum Krylov iteration number is too small, the algorithm will not converge at all. We observed this in our numerical experience. Some directional information must be carried back to the nonlinear iteration. Some theoretical results concerning the convergence of the inexact Newton method can be found in Eisenstat and Walker (1994).

4.7 Reconstructions with varying excitability contrast

The contrast of α between the inclusions and the homogeneous background affects the quality of the reconstructions considerably. Here we consider the case that α in the inclusions is much larger than it in the background compare to the numerical setting used in all former numerical experiments. We observe in Fig. 12(a) that the reconstruction is better when the contrast of α is greater.

Figure 12(b) shows the reconstruction in the case that the value of α in the inclusions is smaller than it in the background. In Section 4.5, we showed that nothing can be reconstructed if there is no electrical potential wave propagating in the forward models, therefore, one must ensure that the electrical potential waves in the forward models are not shut down when tuning the parameters.

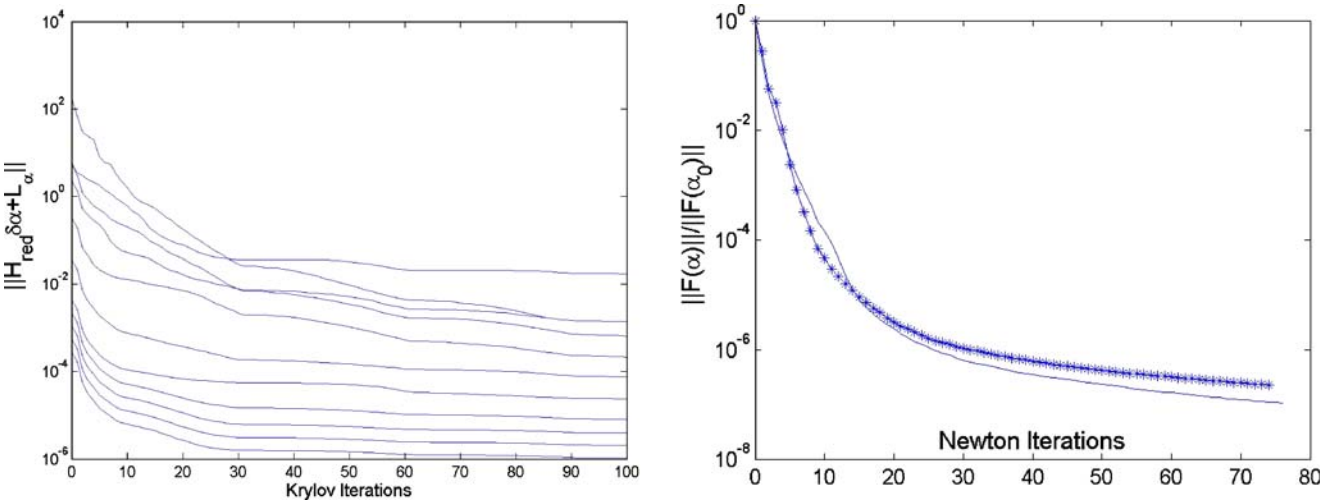

Fig. 10 *Left*: Convergence history for the Krylov solver at different Newton iterations. *Right*: Evolution of objective functions in reconstructions with accurate (*line*) and inaccurate (*symbol*) Krylov solver

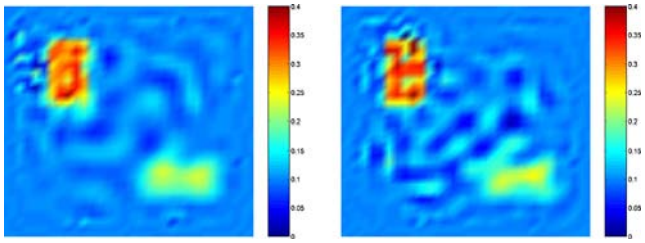

Fig. 11 Reconstructed α with different maximum GMRES iterations

5 Conclusions and discussion

We have considered distributed parameter identification problems for the FitzHugh–Nagumo system that models the propagation of electrical potentials in excitable media such as systems of neurons and hearts.

The mathematical problem consists of reconstructing physical parameters in the system through partial knowledge of its solutions on the boundary of the domain. We have developed a parallel algorithm of Newton–Krylov type that combines Newton's method for numerical optimization with Krylov subspace solvers for the resulting Karush–Kuhn–Tucker system. We show by numerical simulations that the parameter reconstructions can be done with measurement data only on the boundary of the domain.

We discuss the effect of data quality, data amount and some algorithmic parameters on the quality of reconstructions. Future research will be focus on accelerating the current code, extending it to three dimensions on much larger numbers of processors, and comparing simulations on more realistic geometries with experimental measurements.

Finally, we remark that although we consider specifically the reconstruction of the reactive parameter in this paper, the methodology can also be used to reconstruct other parameters in the model, such as the conductivity.

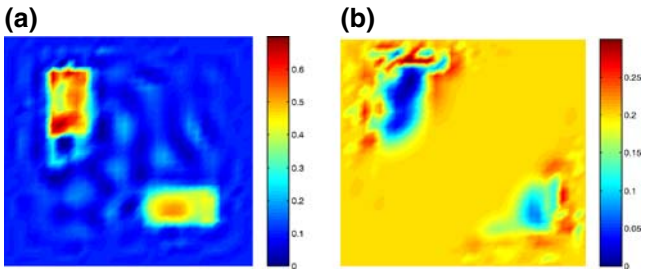

Fig. 12 Reconstructions with varying excitability contrast; (**a**) the new objective α is given by $\alpha(\mathbf{x}) = 0.6$ ($\mathbf{x} \in D_1$), 0.4 ($\mathbf{x} \in D_2$), 0.1 (otherwise); (**b**) the new objective α is given by $\alpha(\mathbf{x}) = 0.05$ ($\mathbf{x} \in D_1$), 0.1 ($\mathbf{x} \in D_2$), 0.2 (otherwise)

Acknowledgements The authors would like to thank Professor George Biros of the University of Pennsylvania for stimulating discussions. Our algorithm is implemented using PETSc citePETSc-web. The simulations were performed on cluster *System X* at the Virginia Polytechnic Institute and State University. We are very grateful for the technical support of Professor Calvin J. Ribbens in using *System X*. This work was supported in p art by the National Science Foundation under NSF CCF-03-52334 and by Columbia University under an AQF grant for cluster computing.

Appendix: Operators in the KKT system

We record here the explicit expressions for the various operators introduced in the KKT system (15). Note that since the operators for the problems with different sources are the same, we will drop their dependence on s, the source index.

Operators for the first equation in KKT system

$$\mathcal{L}_{uu}(u)\delta u := \left(\sum_{j=1}^{N_d} \delta(\mathbf{x} - \mathbf{x}_j) + \lambda f''(u, \alpha)\right)\delta u \tag{38}$$

$$\mathcal{L}_{u\alpha}(u)\delta\alpha := \lambda(1 - 2u)\delta\alpha \tag{39}$$

$$\mathcal{L}_{u\lambda}(u)\delta\lambda := \begin{cases} -\partial_t(\delta\lambda) - \mu\Delta\delta\lambda + f'(u, \alpha)\delta\lambda \\ \delta\lambda(\mathbf{x}, T) = 0 \\ \mathbf{n} \cdot \nabla\delta\lambda = 0 \end{cases} \tag{40}$$

$$\mathcal{L}_{u\eta}(u)\delta\eta := -\vartheta\epsilon\delta\eta \tag{41}$$

Operators for the second equation in KKT system

$$\mathcal{L}_{v\lambda}(v)\delta\lambda := \delta\lambda \tag{42}$$

$$\mathcal{L}_{v\eta}(v)\delta\eta := \begin{cases} -\partial_t(\delta\eta) - \kappa\Delta\delta\eta + \epsilon\gamma\delta\eta \\ \delta\eta(\mathbf{x}, T) = 0 \\ \mathbf{n} \cdot \nabla\delta\eta = 0 \end{cases} \tag{43}$$

Operators for the third equation in KKT system

$$\mathcal{L}_{\alpha u}(\alpha)\delta u := \int_0^T \lambda(1 - 2u)\delta u dt \tag{44}$$

$$\mathcal{L}_{\alpha\alpha}(\alpha)\delta\alpha := \rho\mathcal{R}''(\alpha)\delta\alpha \tag{45}$$

and

$$\mathcal{L}_{\alpha\lambda}(\alpha)\delta\lambda := \int_0^T u(1 - u)\delta\lambda dt \tag{46}$$

Note that the boundary condition of the operator in Eq. (45) is determined by the regularization functional.

Operators for the fourth equation in KKT system

$$\mathcal{L}_{\lambda u}(\lambda)\delta u := \begin{cases} \partial_t \delta u - \mu \Delta \delta u + f'(u,\alpha)\delta u \\ \delta u(\mathbf{x},0) = 0 \\ \mathbf{n}\cdot\nabla \delta u = 0 \end{cases} \tag{47}$$

$$\mathcal{L}_{\lambda v}(\lambda)\delta v := \delta v \tag{48}$$

$$\mathcal{L}_{\lambda \alpha}(\lambda)\delta \alpha := u(1-u)\delta\alpha \tag{49}$$

Operators for the fifth equation in KKT system

$$\mathcal{L}_{\eta u}(\eta)\delta u := -\vartheta \epsilon \delta u \tag{50}$$

$$\mathcal{L}_{\eta v}(\eta)\delta v := \begin{cases} \partial_t \delta v - \kappa \Delta \delta v + \epsilon\gamma\delta v \\ \delta v(\mathbf{x},0) = 0 \\ \mathbf{n}\cdot\nabla \delta v = 0 \end{cases} \tag{51}$$

References

Akcelik, V. (2002). Multiscale Newton–Krylov Methods for inverse acoustic wave propagation. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Akcelik, V., Biros, G., Ghattas, O., Hill, J., Keyes, D., & van Bloemen Waanders, B. (2006). Parallel algorithms for PDE-constrained optimization. In: M. Heroux, P. Raghaven, & H. Simon (Eds.), *Frontiers of parallel computing*. SIAM.

Argentina, M., Coullet, P., & Krinsky, V. (2000). Head-on collisions of waves in an excitable FitzHugh–Nagumo system: a transition from wave annihilation to classical wave behavior. *Journal of Theoretical Biology, 205*, 47–52.

Balay, S., Buschelman, K., Gropp, W. D., Kaushik, D., Knepley, M. G., McInnes, L. C. & et al. (2007). *PETSc Homepage*. http://www.mcs.anl.gov/petsc.

Banks, H. T., & Kunisch, K. (1989). *Estimation techniques for distributed parameter systems*. Boston, MA: Birkhäuser.

Benzi, M., Haber, E., & Hanson, L. (2006). *Multilevel algorithms for large-scale interior point methods in bound constrained optimization*. Technical Report TR-2006-002-A, Department of Mathematics and Computer Science, Emory University. 16p.

Bernus, O., Verschelde, H., & Panfilov, A. V. (2002). Modified ionic models of cardiac tissue for efficient large scale computations. *Physics in Medicine and Biology, 47*, 1947–1959.

Biegler, L., Ghattas, O., Heinkenschloss, M., & Bloemen-Waanders, B. v. (Eds.) (2003). Large-scale PDE-constrained optimization. In *Lecture Notes in Computational Science and Engineering*. Berlin: Springer-Verlag.

Biros, G., & Ghattas, O. (2005a). Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver. *SIAM Journal on Scientific Computing, 27*, 687–713.

Biros, G., & Ghattas, O. (2005b). Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part II: The Lagrange–Newton solver and its application to optimal control of steady viscous flows. *SIAM Journal on Scientific Computing, 27*, 714–739.

Brooks, D. H., Ahmad, G. F., MacLeod, R. S., & Maratos, G. M. (1999). Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE Transactions on Biomedical Engineering, 46*, 3–18.

Bub, G., Shrier, A., & Glass, L. (2002). Spiral wave generation in heterogeneous excitable media. *Physical Review Letters, 88*, 058101.

Chen, X., & Oshita, Y. (2006). Periodicity and uniqueness of global minimizers of an energy functional containing a long-range interaction. *SIAM Journal on Mathematical Analysis, 37*, 1299–1332.

Cheng, L. K., Bodley, J. M., & Pullan, A. J. (2003). Comparison of potential- and activation-based formulations for the inverse problem of electrocardiology. *IEEE Transactions on Biomedical Engineering, 50*, 11–22.

Colli Franzone P., & Pavarino, L. F. (2004). A parallel solver for reaction-diffusion systems in computational electrocardiology. *Mathematical Models and Methods in Applied Sciences, 14*, 883–911.

Courtemanche, M., Skaggs, W., & Winfree, A. T. (1990). Stable 3-dimensional action-potential circulation in the FitzHugh–Nagumo model. *Physica D, 41*, 173–182.

Cox, S. J. (2006). An adjoint method for channel localization. *Mathematical Medicine and Biology, 23*, 139–152.

Cox, S. J., & Griffith, B. E. (2001). Recovering quasi-active properties of dendritic neurons from dual potential recordings. *Journal of Computational Neuroscience, 11*, 95–110.

Cox S. J., & Ji, L. (2003) Discerning ionic currents and their kinetics from input impedance data. *Bulletin of Mathematical Biology, 63*, 909–932.

Cox, S. J., & Wagner, A. (2004). Lateral overdetermination of the FitzHugh–Nagumo system. *Inverse Problems, 20*, 1639–1647.

Dauby, P. C., Desaive, T., & Croisier, H. (2006). Standing waves in the FitzHugh–Nagumo model of cardiac electrical activity. *Physical Review E, 73*, 021908.

Davidenko, J. M., Pertsov, A. V., Salomonsz, R., Baxter, W., & Jalefe, J. (1992). Stationary and drifting spiral waves of excitation in isolated cardiac-muscle. *Nature, 355*, 349–351.

Dennis, Jr J. E., & Schnabel, R. B. (1996). Numerical methods for unconstrained optimization and nonlinear equations. In *Classics in Applied Mathematics*. Philadelphia: SIAM.

Eisenstat S. C., & Walker, H. F. (1994). *Globally convergent inexact Newton methods*, *4*, 393–422.

Elmer, C. E., & Van Vleck, E. S. (2005). Spatially discrete FitzHugh–Nagumo equations. *SIAM Journal on Applied Mathematics, 65*, 1153–1174.

Engl, H. W., Hanke, M., & Neubauer, A. (1996). *Regularization of inverse problems*. Dordrecht: Kluwer.

FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal, 1*, 445–466.

Franzone, P. C., Pavarino, L. F., & Taccardi, B. (2005). Simulating patterns of excitation, repolarization and action potential duration with cardiac bidomain and monodomain models. *Mathematical Biosciences, 197*, 35–66.

Gao, W., & Wang, J. (2004). Existence of wavefronts and impulses to FitzHugh–Nagumo equations. *Nonlinear Analysis, 57*, 667–676.

Haber, E., Ascher, U., & Oldenburg, D. (2000). On optimization techniques for solving nonlinear inverse problems. *Inverse Problems, 16*, 1263–1280.

Hansen, P. C., & O'Leary, D. P. (1993). The use of L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing, 14*, 1487–1503.

Hoffman, D. A., Magee, J. C., Colbert, C. M., & Johnston, D. (1997). K^+ Channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature, 387*, 869–875.

Isakov, V. (1998). *Inverse problems for partial differential equations*. New York: Springer-Verlag.

Knoll, D. A., & Keyes, D. E. (2004). Jacobian-free Newton–Krylov methods: A survey of approaches and applications. *Journal of Comparative Physiology, 193*, 357–397.

Krupa, M., Sandstede, B., & Szmolyan, P. (1997). Fast and slow waves in the FitzHugh–Nagumo equation. *Journal of Differential Equations, 133*, 49–97.

MacLeod, R. S., & Brooks, D. H. (1998). Recent progress in inverse problems in electrocardiology. *IEEE Engineering in Medicine and Biology Magazine, 17*, 73–83.

Moreau-Villéger, V., Delingette, H., Sermesant, M., Ashikaga, H., Faris, O., & McVeigh, E., et al. (2006). Building maps of local apparent conductivity of the epicardium with a 2D electrophysiological model of the heart. *IEEE Transactions on Biomedical Engineering, 53*(8), 1457–1466.

Murillo, M., & Cai, X.-C. (2004). A fully implicit parallel algorithm for simulating the non-linear electrical activity of the heart. *Numerical Linear Algebra with Applications, 11*, 261–277.

Murray, J. D. (1993). *Mathematical biology*. Berlin: Springer

Nagumo, J., Arimoto, S., & Yoshizawa, S. (1962). An active pulse transmission line simulating nerve axon. *Proceedings of the Institute of Radio Engineers, 50*, 2061–2070.

Nii, S. (1997). Stability of traveling multiple-front (multiple-back) wave solutions of the FitzHugh–Nagumo equations. *SIAM Journal on Mathematical Analysis, 28*, 1094–1112.

Nocedal, J. & Wright, S. J. (1999). *Numerical optimization*. New York: Springer-Verlag.

Patel, S. G., & Roth, B. J. (2005). Approximate solution to the bidomain equations for electrocardiogram problems. *Physical Review E, 72*, 051931.

Pennacchio, M., Savare, G., & Franzone, P. C. (2006). Multiscale modeling for the bioelectric activity of the heart. *SIAM Journal on Mathematical Analysis, 37*, 1333–1370.

Pernarowski, M. (2001). Controllability of excitable systems. *Bulletin of Mathematical Biology, 63*, 167–184.

Petersson, J. H. (2005). On global existence for semilinear parabolic systems. *Nonlinear Analysis, 60*, 337–347.

Riccio, M. L., Koller, M. L., & Gilmour, P. F. (1999). Electrical restitution and spatiotemporal organization during ventricular fibrillation. *Circulation Research, 84*, 955– 963.

Roth, B. J. (2004). Art Winfree and the bidomain model of cardiac tissue. *Journal of Theoretical Biology, 230*, 445–449.

Saad, Y. (2003). *Iterative methods for sparse linear systems*, 2nd ed. Philadelphia: SIAM.

Scott, A. C. (1975). The electrophysics of a nerve fiber. *Reviews of Modern Physics, 47*, 487–533.

Shahidi, V., Savard, P., & Nadeau, R. (1994). Forward and inverse problem of electrocardiography: Modeling and recovery of epicardial potentials in humans. *IEEE Transactions on Biomedical Engineering, 41*, 249–256.

Sneyd, J., Dale, P. D., & Duffy, A. (1998). Traveling waves in buffered systems: applications to calcium waves. *SIAM Journal on Applied Mathematics, 58*, 1178–1192.

Suckley, R., & Biktashev, V. N. (2003). Comparison of asymptotics of heart and nerve excitability. *Physical Review E, 68*, 011902.

Tsai, J.-C., & Sneyd, J. (2005). Existence and stability of traveling waves in buffered systems. *SIAM Journal on Applied Mathematics, 66*, 237–265.

Vanier, M. C., & Bower, J.-M. (1999). A comparative survey of automated parameter-search methods for compartmental neural models. *Journal of Computational Neuroscience, 7*, 149–171.

Vogel, C. R. (1996). Non-convergence of the L-curve regularization parameter selection method. *Inverse Problems, 12*, 535–547.

Vogel, C. R. (2002). *Computational methods for inverse problems. In Frontiers in applied mathematics*. Philadelphia: SIAM.

Weiss, J. N., Garfinkel, A., Karagueuzian, H. S., Qu, Z., & Chen, P. S. (1999). Chaos and the transition to ventricular fibrillation—A new approach to antiarrhythmic drug evaluation. *Circulation, 99*, 2819–2826.

Willms, A. R., Baro, D. J., Harris-Warrick, R. M., & Guckenheimer, J. (1999). An improved parameter estimation method for Hodgkin–Huxley models. *Journal of Computational Neuroscience, 6*, 145–168.

Winfree, A. T. (1990). Stable particle-like solutions to the nonlinear-wave equations of 3-dimensional excitable media. *SIAM Review, 32*, 1–53.

Yamada, H., & Nozaki, K. (1990). Interaction of pulses in dissipative systems. FitzHugh–Nagumo equations. *Progress of Theoretical Physics, 84*, 801–809.