ORIGINAL PAPER

Biological Cybernetics

Predicting single spikes and spike patterns with the Hindmarsh–Rose model

Enno de Lange · Martin Hasler

Abstract Most simple neuron models are only able to model traditional spiking behavior. As physiologists discover and classify different electrical phenotypes, computational neuroscientists become interested in using simple phenomenological models that can exhibit these different types of spiking patterns. The Hindmarsh–Rose model is a three-dimensional relaxation oscillator which can show both spiking and bursting patterns and has a chaotic regime. We test the predictive powers of the Hindmarsh–Rose model on two different test databases. We show that the Hindmarsh–Rose model can predict the spiking response of rat layer 5 neocortical pyramidal neurons on a stochastic input signal with a precision comparable to the best known spiking models. We also show that the Hindmarsh–Rose model can capture qualitatively the electrical footprints in a database of different types of neocortical interneurons. When the model parameters are fit from sub-threshold measurements only, the model still captures well the electrical phenotype, which suggests that the sub-threshold signals contain information about the firing patterns of the different neurons.

Keywords Hindmarsh–Rose model · Nonlinear dynamics · Bifurcation analysis · Parameter fitting · Nonlinear optimization · Quantitative neuron modeling · Spike-timing · Neocortical neurons

E. de Lange · M. Hasler
Laboratory of Nonlinear Systems, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Present address:
E. de Lange (✉)
Department of Physiology, University of Bern, Bern, Switzerland
e-mail: delange@pyl.unibe.ch

1 Introduction

In vitro electrical recordings of neocortical neurons display a wide range of different intrinsic firing patterns. In the last two decades many of these firing patterns have been identified and different classification schemes were proposed. Important contributions are early ones of McCormick et al. (1985) and Connors and Gutnick (1990) and the more recent classification schemes by Kawaguchi and Kubota (1997), Gupta et al. (2000), and Markram et al. (2004). To date, the exact number of neocortical cell types, and the criteria for distinguishing them are not universally agreed upon, especially in the case of inhibitory interneurons.

The diversity in firing patterns results from a variability in types and density of ion channels (Toledo-Rodriguez et al. 2002; Bargas and Galarraga 2002). Ion channels are state dependent, stochastic units and their properties may change according to different conditions (Steriade 2004). Under standardized stimulation conditions in vitro, neuronal responses are stable for long periods of time (Rauch et al. 2003), which makes in vitro experiments the tool by excellence to study the different types of electrical behavior neurons may exhibit in vivo.

In recent years there has been a trend towards studying ever larger networks of neurons. To make the link between microscopic and macroscopic behavior it is necessary to combine experimental studies with numerical simulation and analytical study of mathematical models in an interdisciplinary effort. Finding appropriate mathematical models of electrical behavior is one of the main problems because the experimental, numerical and analytical requirements on neuron models are often irreconcilable, especially if a quantitative model is required.

Generic bifurcation models, such as the Hindmarsh–Rose (HR) model and the FitzHugh–Rinzel model are models

based on a phenomenological, i.e., qualitative, rather than quantitative, description of neuron behavior. They are usually capable of both spiking and (different types of) bursting behavior, in contrast to simple integrate-and-fire models and relaxation oscillators, such as the FitzHugh–Nagumo model, but generic bifurcation models are slower in simulation than these spiking-only models. Another advantage of generic bifurcation models is that they are designed from a bifurcation theory perspective and their structure and behavior in networks can be analytically treated using the entire range of techniques and theory from the field of nonlinear dynamics and complex system theory (Belykh et al. 2005; Izhikevich 2000).

The difficulty in fitting generic bifurcation models to experimental data is that the loss surface, which mathematically describes the quality of the prediction as a function of the fit parameters, is non-smooth due to the discontinuities at the bifurcation points. We propose a two-step method to facilitate automatic parameter fitting for these models. In this paper we illustrate this approach for the HR model. In principle it can be used for any model of the generic bifurcation type.

Steur (2006) and Steur et al. (2007) show that a slightly adapted version of the HR model cannot reproduce all recordings form single neurons form a mouse hippocampus. But they show that with a slight modification the model can reproduce details in spike trains. Selverston et al. (2000) show that the HR model, when extended with a fourth differential equation, can reproduce the complex dynamical (chaotic bursting) behavior of pyloric central pattern generator neurons from a lobster stomatogastric system.

We test the predictive powers of the original HR model on two sets of single-cell patch clamp recordings. The first set consists of four rat neocortical pyramidal cells to which a protocol of in vivo like stimulation signals has been applied (Rauch et al. 2003; Destexhe et al. 2001). The spiking behavior observed with these type of stimuli can be quantitatively predicted by integrate-and-fire neurons (Jolivet et al. 2006b; Clopath et al. 2007) and therefore presents an ideal test bench for the HR model and the fitting method. The second set consists of a set of patch clamp recordings, applying a traditional stimulation protocol, to different types of rat neocortical cells. This data set contains recordings of neurons with a wide variety in dynamical behavior (Toledo-Rodriguez et al. 2004). Therefore it represents a good set to test the advantages the HR model presents with respect to simple spiking models.

This paper is organized as follows. In Sect. 2 we present brief analytical treatment and a two-parameter numerical bifurcation analysis of the HR model for the fit parameters used in the rest of the paper. The bifurcation analysis serves as a guideline for the identification approach presented in Sect. 3. The technical details of how we use the approach on

the data in the first data set are described in Sect. 3.4. For the fitting of the second data set, an exact quantitative fit is unrealistic. The fit criteria we used for the second data set are described in Sect. 3.5. Sections 4 and 5 describe the results of the fit of the HR model to the first and second data set, respectively.

2 Model analysis

The HR model is given by the following set of three ordinary differential equations

$$
\begin{aligned}
\dot{x} &= y - x^3 + bx^2 + I - z, \\
\dot{y} &= 1 - 5x^2 - y, \\
\dot{z} &= \mu\left(s\left(x + 1.6\right) - z\right)
\end{aligned}
\tag{1}
$$

where we used the standard parameter values given by Hindmarsh and Rose (1984). Variable x represents the membrane potential and we will refer to it as the "output". Of the remaining free parameters, I is the external input current, μ the time constant of the slow recovery current, and b and s are design parameters. The input current I is probably the most important model parameter since it is linked to experiments (current injection) and in vivo function (synaptic current). It is the only parameter that can vary on a short timescale in experiments and is therefore an obvious, almost imperative, choice as a bifurcation parameter. Indeed all one-dimensional bifurcation analyses in literature take I as the bifurcation parameter. González-Miranda (2003) takes $\mu = 0.0021$ and $b = 3$ and Innocenti et al. (2007) take the same values and also $\mu = 0.001$.

To access the full range of possible behaviors of the HR model one needs to have at least two free parameters. González-Miranda (2007) presents a two-dimensional bifurcation analysis with μ as the second bifurcation parameter. Storace et al. (2008) describe a two-dimensional bifurcation analysis with b and I as bifurcation parameters and argue that any nontrivial combination of two bifurcation parameters will yield a qualitatively identical bifurcation structure. As a starting point to describe our method, we use the results described in Storace et al. (2008). Note however in some situations μ or s (c.f. González-Miranda (2007)) could be better choices for the second bifurcation parameter.

Below follows a brief analytical treatment, which gives bounds on parameters b and I as well as a brief description of the bifurcation structure of the model as a function of b and I and comments on the influence of the other parameters. This gives an example how a bifurcation analysis can be used in the identification process.

2.1 Analytical considerations

Some relatively simple analytical observations, taking into account our practical use as a model for neuronal activity, allow us to give bounds on the values of b and I to use in the exhaustive numerical bifurcation analysis.

The model equations (1) can be separated in a fast subsystem, consisting of the first two equations, and a slow subsystem, consisting of the third equation. For $\mu = 0$, the system reduces to the z-parameterized fast subsystem with equilibrium points on the cubic $z = f(x) \equiv -x^3 - (5 - b) x^2 + I + 1$. The equilibrium points $E = (x_e, y_e, z_e)$ of the full system are then defined by the intersection of $f(x)$ with the line $z = g(x) \equiv s(x + 1.6)$.

For ease of notation we introduce the parameters $\alpha = 5 - b$ and $i = I - 1.6s + 1$. The cubic $f(x)$ has its knee points in $x_1^k = -2\alpha/3$ and $x_2^k = 0$. The system has a single equilibrium for $s > \alpha^2/3$, or $b \in \left(5 - \sqrt{3s}, 5 + \sqrt{3s}\right)$. This equilibrium lies between the knee points, that is, on the middle branch of $f(x)$, if i lies between the points $i_1 = -2/3\alpha \left(s - 2/9\alpha^2\right)$ and $i_2 = 0$. For $\alpha < 0$, $i_1 > i_2$ and for $\alpha > 0$, $i_1 < i_2$. For $s < \alpha^2/3$, the system has either one or three equilibrium points.

The equilibrium points of the fast subsystem are the intersections of the line $z = 0$ with $f(x)$; we will call them E_1, E_2 and E_3, respectively, with E_1 through E_3 having increasing x_e. For $\alpha > 0$, E_2 will have $x_e < 0$ and for $\alpha < 0$ it will have $x_e > 0$. Equilibrium E_1 always has negative x_e and E_3 always positive. Equilibrium E_3 can undergo an Andronov–Hopf bifurcation for $x_H = \left(b \pm \sqrt{b^3 - 3}\right)/3$

A bifurcation analysis of the fast subsystem shows that E_1 is a stable node or a focus, E_2 is always a saddle and E_3 is a node or a focus. Furthermore, the divergence of the fast subsystem is negative everywhere for $b < \sqrt{3}$, excluding cycles.

The bifurcation structure of the full system can now be separated in four different cases according to the value of b (see Fig. 1). We look at the change of the locus of the x-value of E. It is easily seen from $f(x)$ how x_e is modulated with I.

(a) If $b < \sqrt{3}$, the fast subsystem has no cycles and E_1 and E_3 are asymptotically stable nodes (when they exist). When $x_e < x_1^k$, the full system has an asymptotically stable equilibrium corresponding to E_1 in the fast subsystem. For $x_e = x_1^k$, a homoclinic orbit is created. For $x_1^k < x_e < x_2^k$, E is a saddle and there exists a bursting solution generated from the homoclinic orbit. For $x_e > x_2^k$, E corresponds to E_3 and the system is asymptotically stable again.

(b) If $\sqrt{3} < b < 2$, the Andronov–Hopf bifurcation of the reduced system at $x_H = \left(b \mp \sqrt{b^3 - 3}\right)/3$ is supercritical. The other Andronov–Hopf bifurcation (positive
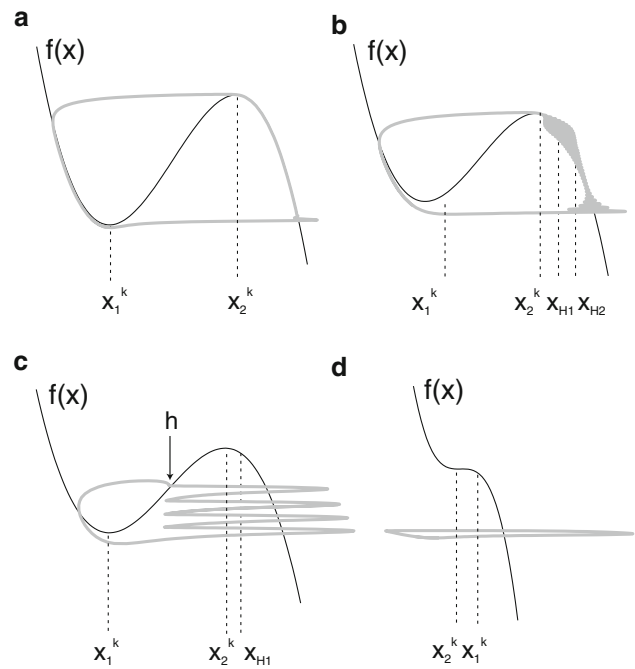

Fig. 1 Nullclines (in *black*) and examples of attractors (in *gray*) of the HR model in the x-z plane. **a** $b = 1$; **b** $b = 1.8$; **c** $b = 2.3$; **d** $b = 5.5$

sign) lies far to the right on $f(x)$, so it is of no concern. For $b < 2$ the left knee point of $f(x)$ has lower z value than the bifurcation. The full system will jump from this left knee point onto the other branch, upon which z will increase and the fast subsystem will undergo the bifurcation, generating the spikes within the burst. For $\mu \to 0$ all periodic solutions are of a conic bursting form (Izhikevich 2000). When μ increases, the jump is not strictly with constant z, so the value of b where the conic bursting starts is slightly lower than 2. The border $b = 2$ is only exact for $\mu = 0$, but gives an upper bound for $\mu > 0$.

(c) If $2 < b < 5$, the second Andronov–Hopf bifurcation lies below the left knee point: the solution will immediately jump onto the periodic cycle around E_3 of the fast subsystem created in the bifurcation. If z varies enough to let the system occasionally jump back to the left branch of $f(x)$, this creates a bursting solution. If z does not vary enough to reach the jump point, the solution stays on the limit cycle around E_3, which gives a regular spiking solution. Whether z can reach the right knee point depends in a complex way on the parameter values. This case is the most physiologically plausible one.

(d) If $b > 5$, the Andronov–Hopf bifurcation is subcritical. The inversion of the criticality separates the state space in two parts, creating different asymptotic solutions each with their basin of attraction.

2.2 Bifurcation analysis

To get an assessment of the behavior of the HR model with respect to parameters b and I we use the two-dimensional bifurcation analysis presented in Storace et al. (2008). This analysis gives two-dimensional plots of the main bifurcation structure of the model. These diagrams make a practical tool in understanding the behavior of the HR model. Also when combining the two-parameter analysis for different values of μ, the bounds found in the analysis from Sect. 2.1 can be extended to give estimates of the regions of different qualitative behavior in the space of the three most important parameters.

We briefly present the method used to obtain the bifurcation plots. For further detail, see Storace et al. (2008). First a two-dimensional plot of the model behavior as a function of the bifurcation parameters b and I is obtained by an exhaustive simulation of the relevant part of the parameter space. The parameter space is discretized to a grid of size n by m. The model is simulated for every point (b_i, I_j), $i = 1 \ldots n$, $j = 1 \ldots m$ on this grid using the advanced solver for stiff systems from the SUNDIALS CVODE package Hindmarsh et al. (2005). This simulation is used to classify the behavior of the model for this parameter combination. The classification of the behavior is based on the number of intersections of the trajectory with the Poincaré section $y - x^3 + bx^2 + I - z = 0$. This corresponds to the local maxima of x.

The exhaustive simulation diagram already gives a good idea of the model behavior for different parameter values and different regions of qualitatively similar behavior can be recognized. For additional clarity, the exact boundaries of these regions are then determined by continuation using the numerical continuation software AUTO 2000 (Doedel et al. 2001).

A strongly simplified version of the bifurcation diagram for $\mu = 0.01$ and $s = 4$, with regions of different types of behavior highlighted, is shown in Fig. 2a. Going into details on the different bifurcations of the model is out of the scope of this paper. We limit the discussion to mention that the plectrum-shaped region is bordered by an Andronov–Hopf bifurcation and that the horizontal separation between the spiking and bursting regions is formed by a fold bifurcation of cycles. For more details, the interested reader is referred to Storace et al. (2008), which also contains comments the organizing principles of the model and the influence of μ and s.

These diagrams, as well as those reported elsewhere in the literature (Innocenti et al. 2007; González-Miranda 2003, 2007) can be combined to use as orientation for the fitting process. As an example Fig. 2b, shows how two-parameter diagrams with b and I as bifurcation parameters, taken for different values of μ, and generating bifurcation lines in the
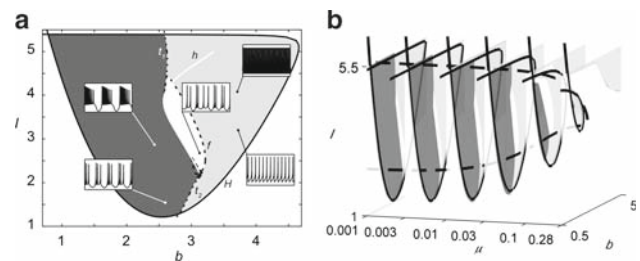

Fig. 2 Classification of the behavior of the HR model. **a** Classification for parameters I and b. Insets show examples of plots of x versus time for different regions. *Large white region* at *bottom*: quiescence; *light gray region*: spiking; *dark gray region*: regular bursting; *white region at center*: chaotic spiking and bursting. The *solid line* labeled H is an Andronov–Hopf (AH) bifurcation, the *dotted line* labeled f a flip bifurcation, the *dotted lines* labeled t are folds of cycles and the *white line* labeled h a homoclinic bifurcation. **b** Classification for of I, b and μ. Coding is the same as in **a**. *Solid lines* are the AH bifurcations

(I, μ)-space can be combined to form an impression of the behavior in the three-dimensional parameter space.

2.3 Scaling parameters for fitting to experimental data

The HR model is a phenomenological model with dimensionless parameters and variables. To fit it to experimental data with physical units for the input and output and time (pA and mV and ms, respectively), we introduce three additional parameters. We call these parameters R, which converts input current in pA into the units of x; τ_s, which converts physical time into model time; and G, which converts the units of x into millivolts. So, if I_n (in pA) is the physical input current used in the experiment, V (in mV) is the physical membrane potential measured in the experiments, t (in s) is the (real) time, and t_m is the model's unit of time, then we have the following relations between physical quantities and the variables from (1):

$$I = RI_n, \quad x = V/G, \quad t_m = \tau_s t. \tag{2}$$

Note that these new parameters do not directly influence the dynamics of the model.

3 Identification of model parameters from experimental data

We used two different databases of electrophysiological measurements to assess the quantitative prediction power of the HR model in two different situations. The first data set is a homogeneous data set of layer 5 pyramidal neurons. The signals applied to the neurons in this data set were designed to mimic in vivo like input signals and therefore this data set forms a good testbed to test the spike-prediction power of the model. The second data set is a heterogeneous set of
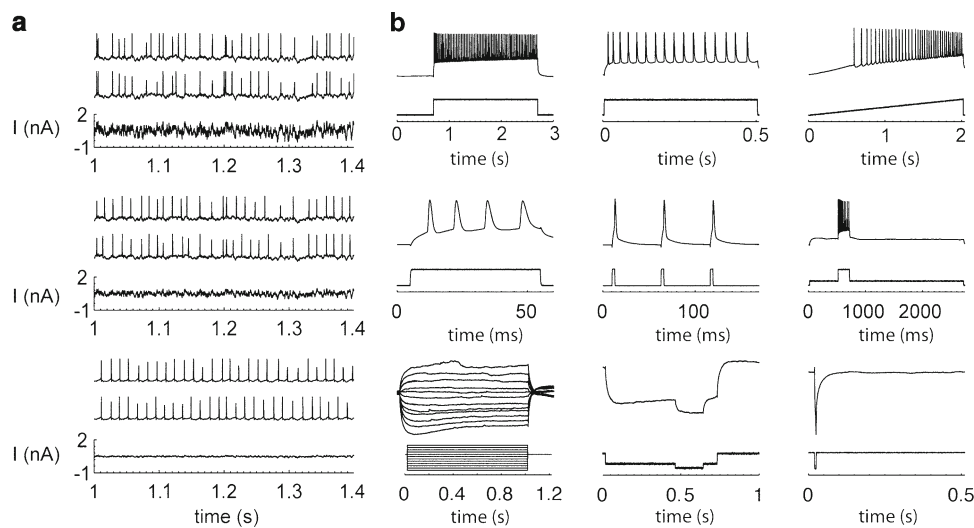
Fig. 3 Stimulation protocols used in the data sets. **a** Examples from the three classes of input currents from the first data set: from top to bottom high, medium and low variance frozen noise signals. Above each input are shown two membrane potential recordings representing two of the four trial repetitions of the same neuron to the input. **b** The nine stimuli in the protocol of the second data set (*bottom part* of each plot) along with an example of a neuron's membrane potential in response to the stimulus (*top parts*). Each row corresponds to a category as defined in the text. *Top row* macroscopic spiking features (group A); stimulus names from left to right are ID_rest, AP_drop and AP_threshold. *Middle row* microscopic spiking features (group B); stimulus names from left to right are AP_waveform, ADHP_rest and sAHP. *Bottom row* sub-threshold features (group C); stimulus names from left to right are IV, IR_hyperpol and Delta

different neocortical interneurons. The electrophysiological protocol applied to the neurons in this set contains a set of standard step and ramp-like signals. This data set was used to test the capability of the HR neuron to model different types of neurons and thus capture the diversity observed in neocortical neuron behavior.

3.1 First data set: L5 pyramidal cells

The first data set contained recordings from four different rat L5 pyramidal cells. The stimulus protocol used in this data set consisted of a series of frozen Ornstein–Uhlenbeck (Tuckwel 1988) noise signals designed to imitate the in vivo input to a neuron's soma (Destexhe et al. 2001). Between 10 and 14 of these frozen noise signals of 6.8 s duration with varying mean and standard deviation were applied four times to each cell (four so-called "trial repetitions"). The mean of the signals was chosen to have both sub- and supra-rheobase values. It is known form literature that the reliability of the neuron's response, i.e., how similar the spike timing of the neuron is upon repeated representation of the same frozen noise signal, depends on the standard deviation of the input signals for L5 rat pyramidal cells (Mainen and Sejnowski 1995). To assess the model's capacity off reproducing this behavior, the data set consists of frozen noise signals with standard deviation around three different values; the lowest ($\sigma(I) = 83.1 \pm 22.7$ pA) yielding low reliability, the middle ($\sigma(I) = 181.9 \pm 16.8$ pA) yielding intermediate reliability and the

highest variance ($\sigma(I) = 333.8 \pm 10.1$ pA) yielding high reliability (see Sect. 4 for a quantification of reliability for this data set). From each of the three category between three and five signals were available for each cell. An example of an input signal from each of these categories and of two different responses of the same neuron to this same signal is shown in Fig. 3a. For further details on the data, see Jolivet et al. (2006b).

It has been shown that the spike-timing of these frozen-noise signals is well modeled by simple nonlinear integrators (Rauch et al. 2003; Jolivet et al. 2006b), so if the more complex HR model is to be used to model spike-timing, it has to be capable to predict the spikes in this data set with at least the same accuracy as these non-linear Integrate-and-Fire models.

3.2 Second data set: neocortical interneurons

The HR model can exhibit qualitatively different types of behavior (Sect. 2.2; Fig. 2). This raises the question whether the model can be used to capture the different classes of behavior observed in neocortical interneurons. This question was addressed using a second data set, which contained whole-cell patch clamp recordings from around 50 interneurons from layers 2–6 of the rat somatosensory cortex. The exact details of the experimental protocol have been described by Toledo-Rodriguez et al. (2004). In brief, nine different stimuli, shown in Fig. 3b , have been applied to each neuron in the data set. The stimuli can be divided in three categories

of three stimuli each, according to which features of the neuron's electrical phenotype the stimuli aim to generate information on. Group A contains stimuli focussing on macroscopic spiking features (discharge and step responses), group B contains stimuli focussing on microscopic spiking features (spike shape and after hyper/depolarization) and group C contains sub-threshold stimuli. For a detailed description of the stimuli and their electrophysiological purpose we refer to Toledo-Rodriguez et al. (2004).

The data set contained many different neuronal responses but had not been formally classified according to a electrophysiological cell-type classification. For a quantitative classification of the electrophysiological diversity in the data set see Toledo-Rodriguez et al. (2004); Markram et al. (2004).

3.3 Fitting method

To approximate the best possible fit to the parameters we used state-of-the-art nonlinear fitting techniques. To use automatic parameter fitting techniques it is essential to be able to define a single quantitative optimization criterion, a so-called loss function. Let the parameter set be a vector θ. The loss function is then a function $L(\theta)$ mapping vector θ to a scalar and quantifying the goodness of the fit, with $L = 0$ meaning a perfect fit. The fitting algorithm tries to find the θ that minimizes the value of L.

3.4 Pyramidal cells: fitting quantitatively on spike timing

For the first data set, the loss function was an adapted version of the coincidence factor Γ defined by Jolivet et al. (2006b). In short, the coincidence factor gives a measure of the quality of spike timing prediction, by comparing the number of spikes of the model falling into a time-window of 2 ms within the spikes of the real neuron—i.e., correctly predicted spikes—to the number of spikes that would be correctly predicted by a spike train generated by a random Poisson process. This measure falls between zero (not better than random Poisson process) and one (perfect prediction of every spike, no spurious spikes), although it can theoretically also become negative (worse prediction than a random Poisson process).

To assess the goodness of the fits on the data by comparing the spike timing of the model's output with the membrane potential recordings of the neuron, one needs to know how well the neuron reproduces its own spike timing when represented multiple times with the same input signal. This is called the intrinsic reliability of the cell and gives an upper bound to the predictive power of the model. The reliability of the cell, as a function of the variance of the input current, can be assessed by computing the coincidence factor Γ for a comparison between pairs of the recordings of the same frozen-noise stochastic input current. This gives the intrinsic coincidence factor $\Gamma_{n \to n}$ (Jolivet et al. 2006b).

The low-variance group had an average intrinsic coincidence factor of $\Gamma_{n \to n} = 0.0952 \pm 0.0773$, which means that it contains practically no information. Therefore the data from this group were not used in our fits. The medium-variance group had an intrinsic coincidence factor of $\Gamma_{n \to n} = 0.5977 \pm 0.1810$ and the high-variance group had $\Gamma_{n \to n} = 0.8389 \pm 0.0428$.

For the medium- and high-variance each group 2/3 of the signals were used as training data for each cell and 1/3 were kept for validation. If the number of signals could not be divided by three, the remainder of the signals were used as training data. The coincidence factor of the model output and the neuron was defined as the average of the coincidence factor of the model output with each of the four recordings of the neuron: $\Gamma_{m \to n} = 1/4 \sum_{i=1}^{4} \Gamma_{m \to n(i)}$; with m the model's output and $n(i)$ the neuron's output for the ith repetition of the input signal.

Nonlinear optimization theory distinguishes two types of optimization: global and local. Local methods begin with a given initial parameter vector and then try to converge from this point in n-dimensional parameter space to the point that minimizes the loss function. Critical for the success of local methods is that the loss function be more or less smooth and monotonic as a function of the parameters; otherwise they quickly get stuck in local minima.

In the case of the HR model, the loss function is everything but smooth and monotonous. Figure 2 clearly shows why: at each bifurcation the model changes its behavior radically (by definition a bifurcation is a discontinuous change in the model behavior). Global optimization strategies consider the entire parameter space and therefore rarely get stuck in a local minimum. The use of a global optimization strategy is therefore unavoidable for this problem.

The disadvantage of global strategies is that they are much slower. To minimize the execution time of the optimization it is of importance to minimize the size of the global optimization. This can be achieved by minimizing the number of parameters in the global optimization. Note that the model behavior is very irregular with respect to bifurcation parameters b, s and μ. The three other parameters, τ_s, G and R (the scaling parameters, see Sect. 2.3) however are normalization constants and therefore their influence on the optimum is much smoother. Note furthermore that G does not play a role for this problem, since we are looking at spike timing, which is independent of the magnitude of the output signal.

Figure 4 shows the dependence of Γ on R around an optimum found below. This optimum can efficiently be found by local optimization of R. Similar behavior can be observed for τ_s.

Therefore, a nested optimization technique can be applied. The idea is to apply a global optimization strategy on a subset of the model parameters, namely only on those which need to be optimized globally: $\theta_g = \{b, s, \mu\}$ (incidentally, those
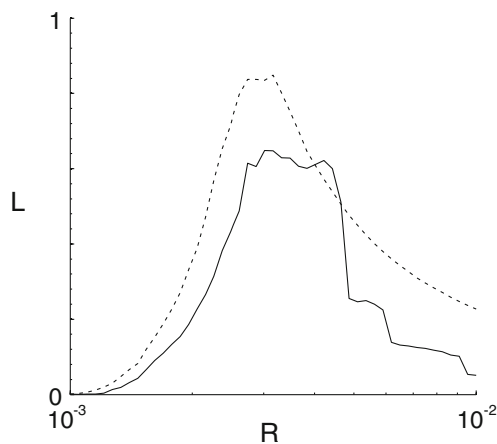
Fig. 4 Comparison of loss function L (*solid line*) and the spike rate-difference, defined by $1 - |N_m - N_n| / (N_m + N_n)$, with N_m and N_n the number of spikes in the model and the neuron, respectively, (*dashed line*) for parameter R

which can cause bifurcations). The global strategy generates vectors θ_g for which it evaluates the loss function. Then, before evaluating the loss function, a local strategy is used to find the optimal value of the local parameter vector $\theta_l = \{\tau_s, R\}$ for every θ_g.

A further simplification to the local optimization can be done by observing that the influence of R is to smoothly modify the spike rate. Obviously, for a good fit, the spike rate of the model will be close to the spike rate of the neuron. From Fig. 4 we conclude that this is indeed the case for the $L(R)$. The loss function shows similar behavior as a function of τ_s. The advantage of using the spike rate is that the model only has to be simulated for part of the input signal, since the mean of the input current is constant and therefore the spike rate of the neuron remains approximately constant over the entire recording.

So, in summary, we adopted a nested optimization approach. The global parameter vector θ_g was optimized with a Genetic Algorithm (Conn et al. 1997). For every point considered in the global strategy, the optimum value of the local parameter vector θ_l was determined by using a local line search (Fletcher 1987) with the spike rate as optimization criterion. This loss function for local optimization was defined as

$$L_l(\theta_l, \theta_g) = 1 - \frac{|N_m - N_n|}{N_m + N_n}, \tag{3}$$

with N_m and N_n the number of spikes in the model and the neuron, respectively. We chose to simulate one fifth of the entire duration of the recording. After local optimization, the system was simulated once for the entire recording using the optimum of θ_l found to compute the loss function $L_g = 1 - \langle \Gamma \rangle$, with $\langle \Gamma \rangle$ the mean of Γ over all trials for the cell.

Since the cells were all pyramidal cells, in theory they are all simply spiking, which is a specific region in parameter space (the conic region, bordered by the Andronov–Hopf and fold of cycles bifurcations in Fig. 2b). We started the optimization with values of θ_g lying in this region. Since random mutations of the parameter vector occur during the Genetic Algorithm, it will still generate points lying outside this region, and in the event where the optimum would lie outside the region, this would be found.

3.5 Neocortical interneurons: fitting on a mix of quantitative and qualitative criteria

It is quite difficult to quantitatively characterize the different classes of electrophysiogical behavior observed in the interneurons in the second data set. As a matter of fact the classification of neocortical interneurons itself is still open for discussion. Toledo-Rodriguez et al. (2004) extract up to 64 features from the experimental data to characterize the electrical behavior of the cell. To keep the fitting process manageable, we extracted nine different features: one from each of the nine stimuli in the protocol (see Fig. 3b).

Many of these features are hard to quantify. Even if one would succeed in doing this (the approach take by Toledo-Rodriguez et al. (2004) could serve as a guide) it would be hard to objectively assign a weight to each feature to give a scalar loss function (an exception forms the sub-threshold group; see below). Finally, an exact quantitative reproduction of the traces is neither realistic, nor necessary for the context where the model would be used; the goal would rather be the knowledge that the model can reproduce the qualitative features and diversity of the interneurons and typical parameter values to do so. Therefore, this data set was partially fitted by hand.

We fitted the data separately on each group (c.f. Sect. 3.2) of stimuli. The stimuli of group A and group B were fitted by hand, using a mix of visual an numerical criteria. The stimuli of group C can all be quantified efficiently and were fitted automatically using a simplex search method (Fletcher 1987). See Table 1) for an overview of the fit criteria.

To facilitate the hand fitting we developed a MATLAB tool which uses the MATLAB mex interface to simulate the HR model using C code and the SUNDIALS CVODE package Hindmarsh et al. (2005) for the different input stimuli for given parameter values. This integration is practically instantaneous on a laptop computer, thus allowing an efficient trial-and-error approach: the user immediately sees the effect a change in a certain parameter has on the fit of the model.

In both hand-fitting and automatic fitting, the results from the bifurcation analysis were used to first make an educated guess of the region in which the parameter values b, and μ could lie. To restrict the degrees of freedom parameter s was chosen to be either 1, 2 or 4; the values one encounters in

Table 1 Overview of the fit criteria used to fit the HR model to data set 2

Stimulus	Extracted property	Type
Group A		
ID_rest	Frequency and discharge type[a]	N/V
AP_drop	Variability in ISI	N/V
AP_threshold	Onset of spiking	N/V
Group B		
AP_waveform	Spike form and width	V
ADHP_rest	After hyper/depolarization	V
sAHP	hyper/Depolarization after burst	V
Group C		
IV	Input resistance	N
IR_hyperpol	Mean square error	N
Delta	First time constant from exp fit	N

In the third column, 'N' stands for 'numerical' and 'V' for 'visual' and N/V means that for that criterion both numerical and visual criteria were combined. ISI means inter-spike interval
[a] According to classification by Markram et al. (2004)

literature. Only then the result was fine-tuned by hand by choosing an appropriate value of τ_s to get the correct time scale and R to get the correct threshold and spike frequency

(group A and B). For group C, parameters b, s, R and τ_s were considered in the automatic fit, which was performed for all three values of s.

4 Quantitative prediction power of the Hindmarsh–Rose model

As described in Sect. 3.4, the model is fit to each cell using a nested optimization approach in which the lossfunction $1 - \langle \Gamma \rangle$, with $\langle \Gamma \rangle$ the mean of the coincidence factor for all training data of that cell. Figure 5 shows the results of the model optimization. All results in the figure are taken on the validation data. A comparison between the response of the model and a neuron for an input signal in the high-variance group is shown in Fig. 5a. The main difference is in the spike onset, which is faster for the real neuron. Nevertheless, most of the spikes are predicted accurately. Also, the subthreshold behavior of the model is very similar to that of the real neuron.

When comparing the coincidence factor of the model with the intrinsic reliability of the cells (Fig. 5b), it can be noted that the quality of the fit seems to vary quite a lot. The average ratio of the coincidence factor of the model and the intrin-

Fig. 5 **a** Example of two membrane potential recordings of a neuron in the first data set upon presentation of the same frozen-noise signal (*black traces*) and of the output of a fitted model neuron (*gray traces*) for the same signal. The *y*-scale of the plots is normalized to the maximum voltage amplitude. **b** Comparison between prediction coincidence factor and intrinsic coincidence factor for each of the validation traces. The *solid line* $\Gamma_{n \to n} = \Gamma_{m \to n}$, represents the theoretical maximum performance of the model. The *dotted line* $\Gamma_{n \to n} = 0.7\Gamma_{m \to n}$, is the linear regression of all data points. **c** Mean quality of the fits, given by $\Gamma_{m \to n}/\Gamma_{n \to n}$, averaged for all validation traces for each cell. The total mean of the ensemble is 0.70 ± 0.12
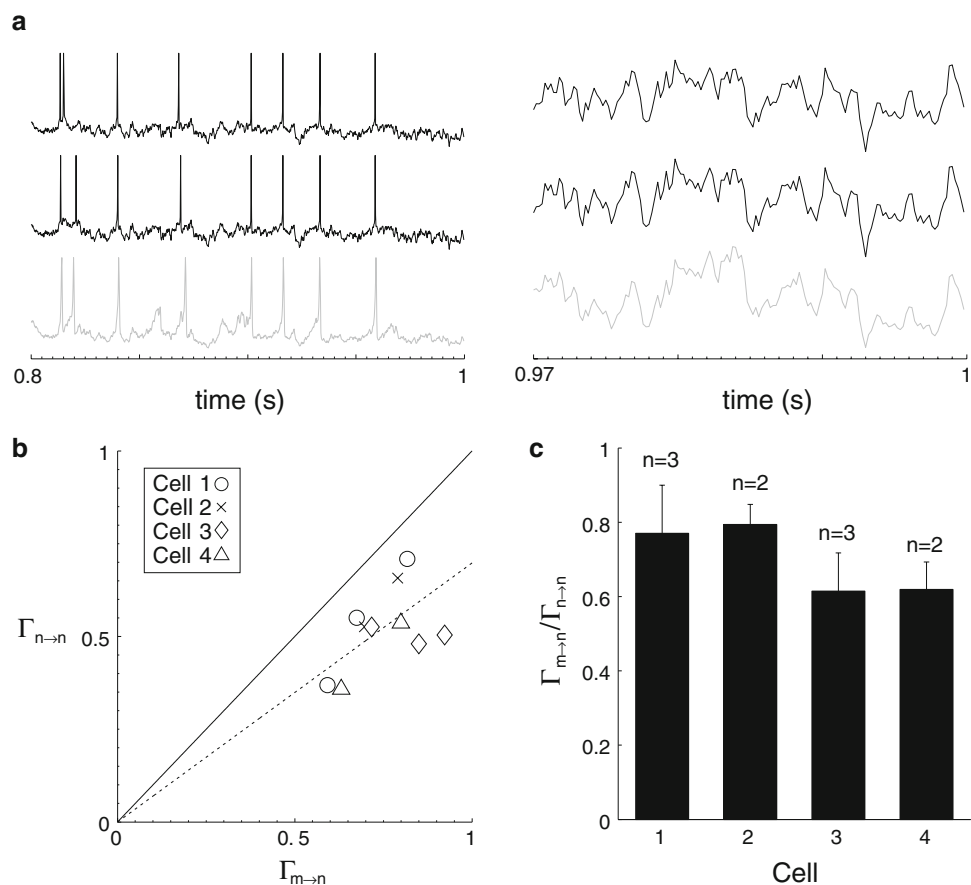
Table 2 Parameter values for the fittings of the HR model to the four cells in data set 1

Cell	parameter				
	b	s	μ	R	τ_s
1	3.20	1.91	0.098	0.0040	1,460
2	2.82	2.02	0.22	0.0029	1,691
3	3.88	3.36	0.016	0.0038	3,697
4	2.84	1.76	0.096	0.0030	2,107

sic coincidence factor, $\Gamma_{m \to n}/\Gamma_{n \to n}$, over all validation data, was 0.70 ± 0.12; this result is very similar to that reported by Jolivet et al. (2006a) for a threshold model (0.69 ± 0.7), but with almost twice as high variance. This could indicate that the HR model could, in principle, perform better, but that since it is more complex and has more parameters, more training data is needed to attain this performance. Since the variance of $\Gamma_{m \to n}/\Gamma_{n \to n}$ is high for both the high-variance input signals and medium-variance input signals, it is unlikely that the model was over-fitted.

A comparison of the results for the four different cells (Fig. 5c) shows that the fits of cell 1 and cell 2 is considerably better than those of cell 3 and cell 4. A study of the experimental data reveals that for some of the recordings of these cell, the rheobase varies and the responses vary qualitatively between trial repetitions. This could indicated that the physiological characteristics of these cells fluctuated during the recording, making a fit on a unique set of model parameters unrealistic.

The parameter values of the four fitted models are reported in Table 2,

5 Modeling neocortical diversity with the Hindmarsh–Rose model

The cells in the second data set were divided in three different groups: 1. non-accommodating (NAC), 2. accommodating (AC) and 3. Other. One cell was selected randomly from each group for identification. The cell from the third group was classified as irregular spiking (IS) cell based on the spike pattern observed in the response to the ID_rest stimulus (c.f. Fig. 6a, bottom left trace). Each cell was fitted to each of the three groups of stimuli, described in Sect. 3.2, separately.

For those curves that were fit by hand, first a point in parameter space was chosen using the bifurcation diagrams described in Sect. 2. The parameter values were adapted blindly by slide-rulers, without information about the precise value, such as not to bias subsequent fits for different criteria on the same cell by prior knowledge. In addition the fits were performed in a random order. The fits shown were the second cells

for each type after having gained experience and collected heuristics on the first series of three cells.

5.1 Fit on individual types of stimuli

The HR model was fit to each of the three group of stimuli, for each of the three cells. Below a discussion of the fit results shown in Fig. 6 ordered by the stimulus group (A, B and C) on which the fit was done.

(A) The spiking frequency of NAC cells could be faithfully mimicked by the HR model (Fig. 6a). The contribution of accommodation (the numerical evaluation of the property extracted from AP_drop, c.f. Table 1), since the normalization of this contribution becomes ill-defined when the accommodation tends to zero. Accommodating discharge patterns can be reproduced faithfully as well. As can be expected in view of the mathematical limitations of the HR model, the irregular spiking behavior is very hard to capture. Using parameter values in the chaotic regime we were able to reproduce irregular behavior, although not similar to the one observed in the cell.

(B) The spike shape is for a large part fixed in the HR model. To approach different shape displayed in the data, unrealistic parameter values had to be taken (for very low b, the shape changes considerably), which had too large an impact on the quality of the fit on the second and the third stimulus from this group. The spike width and distance could be approximated in a satisfactory way as well as the shape of the spikes on delta pulses and the after hyperpolarization. The fit on the NAC cell was especially good, reproducing perfectly the hyperpolarization after the burst and the detail in the delta response.

The problem with the modeling of the shape of the action potential becomes apparent when comparing the fit of the AC and the IS cells: the fit concentrated too much on capturing exactly the timing and shape of the action potential, losing quality in the fit of the other two stimuli. This could possibly be improved by redefining the fitness criteria, or adjusting the weights of the different criteria.

(C) The HR model is capable of representing the diversity in subthreshold dynamics very faithfully. This is in sharp contrast with results obtained for simpler bifurcation models, like the FitzHugh–Nagumo model (de Lange 2006). In the fit for the NAC cell, no specific attention was paid to reproduce the fact that the highest of the steps elicited a spike (accidentally). This trace was taken out as an outlier, but included in the validation, and indeed reproduced. This could have been a coincidence however and further investigation is

Fig. 6 Fits of HR model (*gray traces*) on electrophysiological recordings of three different neocortical cells; a non-accommodating (*NAC*), accommodating (*AC*) and intrinsically spiking (*IS*) cell (*black traces*). **a** Fits on the spiking behavior (group A). **b** Fits on the spike shapes and characteristics (group B) of the same three cells. **c** Fits on the subthreshold behavior (group C) of the three cells. **d** Response for the entire protocol of a HR model that was fit only on the subthreshold behavior (group C)
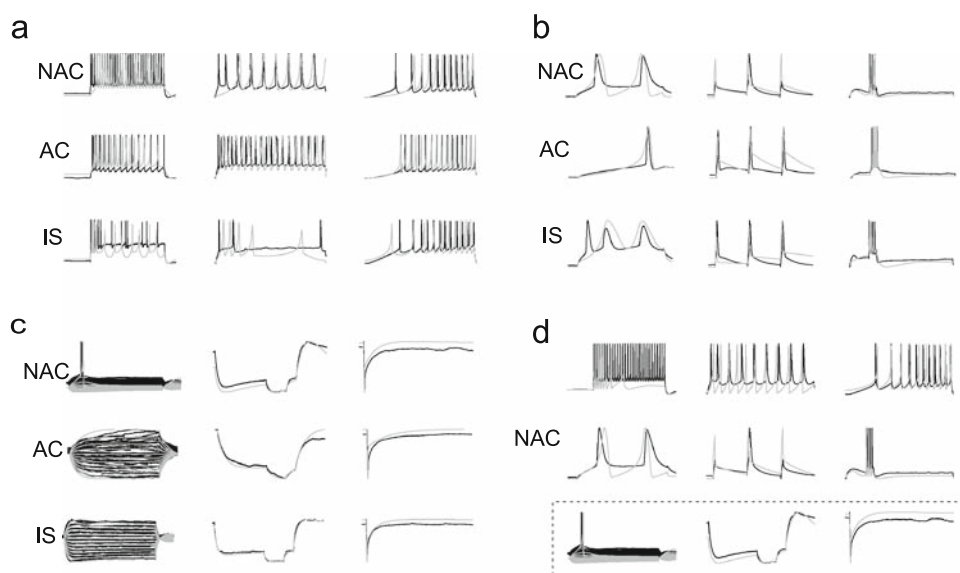
necessary on other cells in the data set that contain this "accident".

The parameter values for each of the cells and each of the fitted groups are given in Table 3.

The results presented above show that many, but not all, of the features of the in vitro protocol can be reproduced by the HR model. Since the fits were done on specific features at a time this does not prove that the HR model is a good representation of neocortical neurons on the whole. Especially the quality of the fit on the subthreshold data could seem trivial, considering the general consensus that the subthreshold behavior of neurons does not determine the complex nonlinear dynamics that govern the super-threshold behavior. To be able to assess the quality of the HR model in representing neuronal behavior it is therefore important to look at the results of the fits on the entire protocol.

Table 3 Parameter values for the fittings of the HR model to three cells in data set 2

Cell	Group	Parameter				
		b	s	μ	R	τ_s
NAC	A	3	1	0.002	0.004	100
	B	3.6	4	0.01	0.01	150
	C	3.6	4	0.005	0.007	200
AC	A	3	4	0.03	0.009	150
	B	2.8	4	0.01	0.01	100
	C	2.8	1	0.01	0.005	200
IS	A	3	2	0.005	0.007	100
	B	3	2	0.005	0.007	100
	C	3.4	2	0.002	0.0035	100

It seems very difficult to formulate unambiguous criteria for the fitting of the entire protocol at once, but it is possible to compare the fits presented above on the other stimuli. Not only this allows the assessment of the general quality of the predictions, but additionally the results give insight in which features are decisive for the quality of the fit. This will help formulate a protocol for tuning the HR model to specific cell-type behavior.

5.2 Comparison of the fits on the entire profile

In this section we look at the response of the neocortical cells that were fitted on one group to the stimuli of the two other groups. For brevity, we discuss the results for one cell (NAC). While not identical, the results for the AC cell allow the same general conclusions, whereas results for the IS cell are harder to compare since the model is not capable of faithfully reproducing the discharge response.

The fit based on the discharge responses is not very good. The main problem is that there are too many degrees of freedom when a fit is done on the spike frequency and adaptation alone. The spike frequency is controlled both by the time scaling factor and by the input scaling factor. For a given sustained input current, there exist pareto-optimal curves of identical frequency in (τ_s, R). In principle, if the frequency were fit for steps of different amplitude and if the threshold were known, it should be possible to locate the exact optimal point on these curves, because the spike threshold will only depend on R. In practice it is hard to determine the spike threshold from the ramp current.

When looking at the values in Table 3 it seems that this caused a misfit of parameters b and s, which were identified better from the other two groups of stimuli. This shows that the choice of the type of behavior and the accommodation

from discharge responses only is a dangerous one, although to induce statistics much more cells have to be fit.

The fit that resulted from focussing on the second group of stimuli was good for the subthreshold signals. This can be expected, since both groups focus on details of dynamic behavior, rather than the more large-scale phenomena taken into consideration by the first group. The frequency of the discharge responses is underestimated which is probably due to an underestimation of the spike threshold that can also be seen from the IV characteristic in the bottom row.

Finally, the model fit on subthreshold stimuli from group 3 yielded the best results (see Fig. 6d). Especially surprising is the good fit of the AP_threshold and AP_drop stimuli. The value of b is probably too high, which explains the cessation of spiking after sustained current injection and what looks like adaptation in the AP_drop response.

6 Conclusions

This paper shows how the combination of standard techniques from nonlinear system identification, heuristics and a priori knowledge from neuroscience permit the quick identification of model parameters from measurement data. We outlined an approach to the parameter fitting of generic bifurcation models for diverse types of in vitro input–output recordings of neurons. Using this approach, we automatically fitted the HR model to recordings from L5 neocortical pyramidal neurons. The HR model could faithfully reproduce around 70% of the spike timing in these recordings. Also, the subthreshold fluctuations in the neurons are faithfully reproduced.

The HR model was also fit on three different types of neocortical interneurons. The HR model can model much of the specific dynamical features present in neocortical neurons. When the model was fitted on sub-threshold recordings only, it could still reproduce qualitative aspects of the spiking patterns, which indicates that the sub-threshold responses of neurons might contain information about the firing pattern and cell type. This result indicates that possibly, neurons could be classified based on their sub-threshold responses. If this is indeed the case, this identification could be automated much more efficiently using standard techniques form system identification and efficiency of the protocols could be greatly improved using, for instance, filtered noise, as is common in system identification. However, these claims have to be verified by further study, since the sub-threshold stimuli in the data set used in this work were not diverse enough to optimally identify the neurons.

The clear limitations of the model lie in the capacity to quantitatively model the irregular spiking behavior observed in neocortical interneurons. If these complex dynamics are governed by a slow intracellular calcium oscillations only

(Chay 1996), fitting the extended version of the HR model used by Selverston et al. (2000), should give better results. The method described above could still be applied in this case.

The fact that in both cases the quality of the fit seems to a large extend independent of the global parameters (parameters that control the qualitative behavior of the model) could imply that the spike timing and firing patterns are robust to fluctuations in environment parameters and depend mostly on characteristics like cell geometry and input resistance. Another indication for this claim is that the performance of the HR model in reproducing spike timing is similar to that of a simple threshold model, which does not model different qualitative regimes.

References

Arbib M (ed) (2002) The handbook of brain theory and neural networks, 2nd edn. MIT Press, London

Bargas J, Galarraga E (2002) Ion channels: keys to neuronal specialization. In: Arbib M (ed) The handbook of brain theory and neural networks, 2nd edn. MIT Press, London, pp 496–501

Belykh I, de Lange E, Hasler M (2005) Synchronization of bursting neurons: what matters in the network topology. Phys Rev Lett 94(18):8101

Chay TR (1996) Electrical bursting and luminal calcium oscillation in excitable cell models. Biol Cybern 75:419–431

Clopath C, Jolivet R, Rauch A, Lüscher HR, Gerstner W (2007) Predicting neuronal activity with simple models of the threshold type: Adaptive exponential integrate-and-fire model with two compartments. Neurocomputing 70:1668–1673

Conn AR, Gould NIM, Toint PL (1997) A globally convergent augmented Lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds. Math Comput 66: 261–288

Connors B, Gutnick M (1990) Intrinsic firing patterns of diverse neocortical neurons. Trends Neurosci 17:3894–3906

Destexhe A, Rudolph M, Fellous JM, TS (2001) Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. Neuroscience 107:13–24

Doedel E, Paffenroth R, Champneys A, Fairgrieve T, Kuznetsov Y, Sandstede B, Wang X (2001) Auto 2000: continuation and bifurcation software for ordinary differential equations (with homcont). Technical Reportr, Caltech

Fletcher R (1987) Practical methods of optimization. Wiley, New York

González-Miranda JM (2003) Observation of a continuous interior crisis in the Hindmarsh–Rose neuron model. Chaos 13(3):845

González-Miranda JM (2007) Complex bifurcation structures in the Hindmarsh–Rose neuron model. Int J Bifurcations Chaos 17(9):3071–3083

Gupta A, Wang Y, Markram H (2000) Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. Science 287(5451):273–278

Hindmarsh J, Rose R (1984) A model of neuronal bursting using three coupled first order differential equations. Proc R Soc Lond Ser B 221(1222):87–102

Hindmarsh AC, Brown P, Grant KE, Lee SL, Serban R, Shumaker DE, Woodward CS (2005) Sundials, suite of nonlinear and differential/algebraic equation solvers.. ACM Trans Math Softw 31: 363–396

Innocenti G, Morelli A, Genesio R, Torcini A (2007) Dynamical phases of the Hindmarsh–Rose neuronal model: studies of the transition from bursting to spiking chaos. Chaos 17:043,128

Izhikevich E (2000) Neural excitability, spiking and bursting. Int J Bifurcation Chaos 10(6):1171–1266

Jolivet R, Rauch A, Lüscher HR, Gerstner W (2006a) Integrate-and-fire models with adaptation are good enough. In: Weiss Y, Schölkopf B, Platt J (eds) (2006) Advances in neural information processing systems, vol 18. MIT Press, Cambridge, pp 595–602

Jolivet R, Rauch A, Lüscher HR, Gerstner W (2006b) Predicting spike timing of neocortical pyramidal neurons by simple threshold models. J Comput Neurosc 21(1):35–49

Kawaguchi Y, Kubota Y (1997) GABAergic cell subtypes and their synaptic connections in rat frontal cortex. Cereb Cortex 7: 476–486

de Lange E (2006) Neuron models of the generic bifurcation type: Network analysis and data modeling. PhD thesis, Ecole Polytechnique Fédérale de Lausanne

Mainen Z, Sejnowski T (1995) Reliability of spike timing in neocortical neurons. Science 268:1503–1506

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C (2004) Interneurons of the neocortical inhibitory system. Nat Rev Neurosci 5:793–807

McCormick D, Connors B, Lighthall J, Prince D (1985) Comparative electrophysiology of pyramidal and sparsely spiny neurons of the neocortex. J Neurophysiol 54:782–806

Rauch A, La Camera G, Lüscher H, Senn W, Fusi S (2003) Neocortical pyramidal cells respond as Integrate-and-Fire neurons to in vivo-like input currents. J Neurophysiol 90:1598–1612

Selverston AI, Rabinovich MI, Abarbanel HDI, Elson R, Szücs A, Pinto RD, Huerta R, Varona P (2000) Reliable circuits from irregular neurons: a dynamical approach to understanding central pattern generators. J Physiol (Paris) 94:357–374

Steriade M (2004) Neocortical cell classes are flexible entities. Nat Rev Neurosci 5:121–134

Steur E (2006) Parameter estimation in Hindmarsh–Rose neurons. Technical Report DCT 2006.073, Technische Universiteit Eindhoven

Steur E, Tyukin I, Nijmeijer H, van Leeuwen C (2007) Reconstructing dynamics of spiking neurons form input–output measurements in vitro. In: Third international IEEE scientific conference on physics and control, Potsdam

Storace M, Linaro D, de Lange E (2008) The Hindmars–Rose neuron model: bifurcation analysis and piecewise-linear approximations. Chaos 18:033128

Toledo–Rodriguez M, Gupta A, Wang Y, Wu C, Markram H (2002) Neocortex: basic neuron types. In: Arbib M (ed) The handbook of brain theory and neural networks, 2nd edn. MIT Press, London, pp 791–725

Toledo-Rodriguez M, Blumenfeld B, Wu C, Luo J, Attali B, Goodman P, Markram H (2004) Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. Cereb Cortex 14(12):1310–1327

Tuckwel H (1988) Introduction to theoretical neurobiology. Linear cable theory and dendritic structure, vol 1. Cambridge University Press, Cambridge

Weiss Y, Schölkopf B, Platt J (eds) (2006) Advances in neural information processing systems, vol 18. MIT Press, Cambridge