Simple data-driven models of intracellular calcium dynamics with predictive power

Alejandra C. Ventura, Luciana Bruno, and Silvina Ponce Dawson
Departamento de Física, Facultad de Ciencias Exactas y Naturales, U.B.A., Ciudad Universitaria, Pabellón I, (1428) Buenos Aires, Argentina

(Received 6 November 2005; revised manuscript received 22 May 2006; published 26 July 2006)

Biology is complex. However, it is not clear how much of this complexity must necessarily translate into complicated mathematical models of biological processes. Simple models can be appealing to physicists but are usually deceiving for biologists. Complicated models, on the other hand, depend on too many parameters whose values are frequently unknown. Therefore, complicated models, although in principle more realistic, can lead to erroneous results if they are sensitive to these unknown parameter values. Intracellular calcium signals provide an example of utmost biological importance in which the issue of "simple *vs* complex" can be explored. In this paper we show that simple models describing the dynamics of intracellular calcium can be directly inferred from experimental data, without no *a priori* information on unknown parameters. A similar approach can be followed to study other reaction-diffusion systems. In spite of their simplicity, these models can provide quantitative information on some of the processes that shape calcium signals, such as the calcium current that underlies an experimental observation. This shows that simple models of biological systems are not limited to qualitative descriptions.

DOI: 10.1103/PhysRevE.74.011917 PACS number(s): 87.16.Ac, 87.16.Uv

I. INTRODUCTION

Mathematical models of complex phenomena can be as complicated as the system under study or can be much simpler. The typical physicist's approach has a tendency to simplicity. The typical biologist's approach, on the other hand, pays much attention to the enormous diversity that small details give rise to. It is hard to reconcile these two points of view. Many of the simple models of biological systems that physicists develop are commonly limited to a qualitative description. Such descriptions, which are at the realm of the theory of dynamical systems, are very appealing to physicists and applied mathematicians. However, they can be very deceiving for biologists who are seeking more quantitative answers. Is there any room for simple models to give quantitative answers to biological questions? There have been very remarkable examples that this is possible, such as the Hodgkin-Huxley model of action potentials [1]. However, the enormous amount of data that biological experiments are producing makes the elaboration of simple but predictive models even more difficult. How can we then bridge the gap between a simplifying approach and the search for quantitative results? In this paper we present a possible approach that is applicable to the case of intracellular calcium (Ca^{2+}) signals.

Intracellular calcium (Ca^{2+}) dynamics provides a very interesting example in which the issue of "simple *vs* complex" can be investigated. Calcium ions are used for signaling purposes by virtually all cell types, regulating functions as diverse as secretion, contraction, proliferation, and cell death [2]. The variety of end responses produced by the different spatiotemporal distributions of the Ca^{2+} concentration, $[Ca^{2+}]$, requires tight control of this distribution [3,4]. Since prolonged elevations of the free cytosolic Ca^{2+} concentration leads to cell death, cells have various mechanisms to trap or remove free Ca^{2+}. Ca^{2+} enters the cytosol through specialized ion channels located on the plasma membrane or on the membrane of internal stores. Upon its entrance, Ca^{2+} diffuses, binds to various mobile and immobile cytosolic buffers (usually, proteins), and is removed into internal stores or extruded into the extracellular medium by pumps and exchangers [5]. The presence of this large variety of removing mechanisms complicates the description of the intracellular $[Ca^{2+}]$ dynamics. Namely, if one wants to describe all these processes in detail, one is faced with a set of many coupled reaction-diffusion equations [6]. This description gets even more complicated if one wants to include the description of the ion channels, whose conformational changes depend stochastically on cytosolic $[Ca^{2+}]$ [7]. Furthermore, Ca^{2+} transport between the cytosol and the other "compartments" (such as the endoplasmic reticulum or the extracellular medium) requires, in principle, a separate description of the dynamics of $[Ca^{2+}]$ in each of them [8]. Clearly, the mathematical model can become endlessly complicated. Ca^{2+} removal by pumps and exchangers occurs on a slower time scale than buffering [9,10]. Thus it is possible not to include them in the description if one is interested in phenomena that occur within a certain time scale. But even with this simplification, the description still remains quite complicated. The main question is then: how much detail is necessary to extract quantitative information from experiments?

There are different types of experiments that provide partial pictures of the intracellular $[Ca^{2+}]$ dynamics [11,12]. The ones that give the most direct link to this dynamics are optical techniques, in which fluorescent indicators are used to visualize the spatiotemporal distribution of the Ca^{2+}-dye bound complex [13,14]. Although, inferring the $[Ca^{2+}]$ distribution from these observations may not be so direct, it can be done within certain assumptions [15]. But still, even if one is able to deduce how the free $[Ca^{2+}]$ is distributed within the cell, it is hard to infer from these experiments how much Ca^{2+} entered the cytosol, due to the presence of the various removal mechanisms mentioned before. The combination of models and experiments is thus unavoidable to es-

timate this Ca^{2+} current. Simple models might suffice to describe the observed distribution in a qualitative way. However, the question arises of whether such simple models are also able to provide quantitative information, such as the amplitude and kinetics of the Ca^{2+} current, which is of most interest for biologists. We think that simple models are indeed capable of providing quantitative answers to specific questions. In particular, the algorithm we presented in [16] is based on this assumption. The question is then how to obtain these simple models.

One possible way to obtain a simple dynamical model is to start from a complete (complicated) one and then reduce it under certain assumptions [10,17–22]. We discuss in the next section some approaches along this line. In the case of intracellular Ca^{2+}, however, detailed knowledge of a full model, including quantitative information on the kinetics of the various processes at work, may be unavailable. Therefore, this step-by-step reduction is not always an option. In [23] we explored the possibility of obtaining a dynamical model from experimental observations. We discuss this "data-driven" approach in the next section too. The drawback of "data-driven" models is that they could just be descriptive, if a different model is obtained for each different observation. Models with these characteristics would not be comprehensive, but would be in an almost one-to-one correspondence with a particular solution. Based on the theory of dynamical systems, comprehensive models have been built from data in the case of low dimensional dissipative systems [24–27]. In [23] we applied a similar approach to the case of reaction-diffusion systems. In this paper, we advance on the idea of data-driven models and we show that in fact these models may have predictive power and the ability to provide quantitative information in the case of Ca^{2+} signals.

Ca^{2+} signals involve Ca^{2+} entry in the cytosol, its subsequent transport in the presence of buffers, and its recapture [5]. The buffer types, their concentrations and kinetics, and the typical dynamics of the pumps and other recapturing mechanisms may vary depending on the cell type. Different amounts of buffers can also be expressed during different stages of the same cell. Therefore a different mathematical model or at least, a model characterized by different parameter values, should correspond to each cell type in a particular stage. On the other hand, the same cell can be subjected to different patterns of Ca^{2+} entry which lead, in turn, to different Ca^{2+} signals and end responses [3]. In oocytes, Ca^{2+} signals that are due to Ca^{2+} entrance through one or more IP_3 receptor/Ca^{2+} channels (IP_3R's) have been observed to coexist. Among them, *blips, puffs*, and global waves [28]. Ca^{2+} *puffs* themselves are highly variable, depending on the number of Ca^{2+} channels that open during the puff [29]. A mathematical model of Ca^{2+} *puffs* would then involve one set of dynamical equations with some stochastic variable(s) determining Ca^{2+} entry. In this way, each puff would correspond to a different solution of the same model [30]. The main assumption that underlies our "data-driven" approach is that the mathematical equations of the model, including the various parameters that characterize them, can be inferred from individual solutions. The immediate practical application of this approach to Ca^{2+} signals is to obtain the amplitude and kinetics of the Ca^{2+} entry that corresponds to each experimental observation. In this type of application, Ca^{2+} entry is an unknown (stochastic) function of time to be determined from the data, which varies from signal to signal. The mathematical model that describes Ca^{2+} transport, buffering, and recapture, on the other hand, is also unknown and to be determined from the data, but should remain the same from signal to signal. The time course of the Ca^{2+} entry is solution-dependent. The model that describes the subsequent Ca^{2+} dynamics is not. In this paper we show that a very simple model, whose parameters can be inferred from data, can in fact describe various observed Ca^{2+} signals, whenever Ca^{2+} enters through one or several distinguishable spatially localized regions (channels or clusters of channels). We discuss why this simple model, which has analytic solution, is suitable to describe these signals and what its limitations are. One of the main features is that it does not require any *a priori* information on the underlying processes that shape Ca^{2+} signals, but extracts it from the data itself. On the other hand, it is not merely descriptive, since it can be used to infer Ca^{2+} currents. It is a simple model that yet provides quantitative information, bridging the gap between the typical simplifying approach of physicists and the needs of biologists.

The organization of the paper is as follows. In Sec. II we briefly describe the complete set of evolution equations that describe intracellular Ca^{2+} dynamics. In Sec. III we summarize some of the previously reduced descriptions, the ones that have been derived from the complete dynamical model in different limiting regimes and the reduced description obtained from data introduced in [23]. In Sec. IV we discuss in more detail our "data-driven" approach, describing the simpler description that we introduce in this paper. In Sec. V we test this reduction by means of numerical simulations. The reduction is tested with experimental data in Sec. VI. Finally, a discussion is presented in Sec. VII, where also the conclusions are summarized.

II. COMPLETE MATHEMATICAL MODEL

Upon its release in the cytosol, Ca^{2+} diffuses, reacts with various buffers (among them, the dye that is used in optical experiments) and is recaptured by different pumps. The reactions with the buffers are assumed to be of the form

$$Ca^{2+} + B_i \underset{k_{\mathrm{off},i}}{\overset{k_{\mathrm{on},i}}{\rightleftharpoons}} CaB_i, \tag{1}$$

where $k_{\mathrm{on},i}$ and $k_{\mathrm{off},i}$ are the kinetic rate constants for the particular buffer B_i. It is further assumed that the free and bound forms of the buffer, B_i, and CaB_i, respectively, diffuse at the same rate, $D_{B,i}$, and that their total concentration, $[B]_{T,i}$, is uniformly distributed. Thus the evolution equations for this system can be written as

$$\frac{\partial[Ca^{2+}]}{\partial t} = \sum_{i=1}^{N} R_i + D_{Ca}\nabla^2[Ca^{2+}] + M + Q_{Ca}, \tag{2}$$

$$\frac{\partial[CaB_i]}{\partial t} = -R_i + D_{B,i}\nabla^2[CaB_i], \quad 1 \leqslant i \leqslant N, \tag{3}$$

where

$$R_i = -k_{\mathrm{on},i}[\mathrm{Ca}^{2+}]([B]_{T,i} - [\mathrm{CaB}_i]) + k_{\mathrm{off},i}[\mathrm{CaB}_i], \quad 1 \leqslant i \leqslant N \tag{4}$$

is the reaction term between Ca^{2+} and the buffer B_i, and the sum goes over all the (N) buffers that affect the cytosolic calcium dynamics; D_{Ca} is the free diffusion coefficient of Ca^{2+} ($\sim 220\ \mu\mathrm{m}^2/\mathrm{s}$ [31]); and Q_{Ca} is the Ca^{2+} source. All other Ca^{2+} sequestration mechanisms (among them, the Ca^{2+}-ATPase that pumps Ca^{2+} back into internal stores) are included in M. Usually, M is regarded as a function of $[Ca^{2+}]$ exclusively and we do so in this paper.

III. PREVIOUSLY EXPLORED REDUCED DESCRIPTIONS

A. Reduced descriptions obtained from models

In this section we briefly describe some of the reductions of Eqs. (2) and (3) that have been derived in different limiting regimes.

The rapid buffering approximation (RBA) [10,17–20] is based on the assumption that the reactions with the buffers occur much faster than any other process, so that Ca^{2+} and the buffers are locally in chemical equilibrium. In this way, Eqs. (2) and (3) can be approximated by a single equation for $[Ca^{2+}]$ of the form

$$\frac{\partial[\mathrm{Ca}^{2+}]}{\partial t} = \beta\left\{\left(D_{\mathrm{Ca}} + \sum_i \kappa_i D_{B,i}\right)\nabla^2[\mathrm{Ca}^{2+}] - 2\left(\sum_i \frac{\kappa_i D_{B,i}}{K_i + [\mathrm{Ca}^{2+}]}\right)\nabla[\mathrm{Ca}^{2+}]\cdot\nabla[\mathrm{Ca}^{2+}] + M + Q_{\mathrm{Ca}}\right\}, \tag{5}$$

$$[\mathrm{CaB}_i] = \frac{[B]_{T,i}[\mathrm{Ca}^{2+}]}{K_i + [\mathrm{Ca}^{2+}]}, \tag{6}$$

where

$$\beta = \left(1 + \sum_i \frac{K_i[B]_{T,i}}{(K_i + [\mathrm{Ca}^{2+}])^2}\right)^{-1}, \tag{7}$$

$$\kappa_i = \frac{K_i[B]_{T,i}}{(K_i + [\mathrm{Ca}^{2+}])^2}. \tag{8}$$

Equation (5) can be rewritten as a reaction-diffusion equation with a concentration-dependent diffusion coefficient [32].

Another reduction, first introduced in [21] and revisited in [18], is the so-called excess buffer approximation (EBA). In the EBA it is assumed that there is always enough free buffer available at any given point so that $[B]_{T,i} - [\mathrm{CaB}_i] \approx [B_i]_{\mathrm{rest}} = [B]_{T,i}/([\mathrm{Ca}^{2+}]_{\mathrm{rest}}/K_i + 1) \approx [B]_{T,i}$. This approximation has been shown to be valid when mobile buffers are in high concentration and/or when the source amplitude is small [10,18,21]. Under the EBA, Eqs. (2) and (3) can be approximated by a single equation for $[Ca^{2+}]$ of the form

$$\frac{\partial[\mathrm{Ca}^{2+}]}{\partial t} = -\left(\sum_i k_{\mathrm{on},i}[B]_{T,i}\right)([\mathrm{Ca}^{2+}] - [\mathrm{Ca}^{2+}]_{\mathrm{rest}}) + D_{\mathrm{Ca}}\nabla^2[\mathrm{Ca}^{2+}] + M + Q_{\mathrm{Ca}}. \tag{9}$$

In this approximation the $[Ca^{2+}]$ dynamics is described by a single evolution equation.

Under similar assumptions as those of the EBA, a better approximation was derived in [22]. The equations, in this case, are obtained by linearization around the resting solution. We will refer to this reduction as the Naraghi and Neher approximation (NNA). Defining $\delta[Ca^{2+}] = [Ca^{2+}] - [Ca^{2+}]_{\mathrm{rest}}$ and $\delta[\mathrm{CaB}_i] = [\mathrm{CaB}_i] - [\mathrm{CaB}_i]_{\mathrm{rest}}$, the linearized version of Eqs. (2) and (3) is

$$\frac{\partial\delta[\mathrm{Ca}^{2+}]}{\partial t} = -\left(\sum_i \frac{\kappa_i}{\tau_i}\right)\delta[\mathrm{Ca}^{2+}] + \sum_i\left(\frac{1}{\tau_i}\delta[\mathrm{CaB}_i]\right) + D_{\mathrm{Ca}}\nabla^2\delta[\mathrm{Ca}^{2+}] + \frac{\partial M}{\partial[\mathrm{Ca}^{2+}]}\delta[\mathrm{Ca}^{2+}] + Q_{\mathrm{Ca}}, \tag{10}$$

$$\frac{\partial\delta[\mathrm{CaB}_i]}{\partial t} = \frac{\kappa_i}{\tau_i}\delta[\mathrm{Ca}^{2+}] - \frac{1}{\tau_i}\delta[\mathrm{CaB}_i] + D_{B,i}\nabla^2\delta[\mathrm{CaB}_i], \tag{11}$$

where $\tau_i = 1/(k_{\mathrm{off},i} + k_{\mathrm{on},i}[\mathrm{Ca}^{2+}])$ and κ_i was defined in Eq. (8). Both parameters are now evaluated at $[Ca^{2+}]_{\mathrm{rest}}$. In this approximation the dynamics is described by a system of linear equations that can be solved analytically for the case of a point source.

B. Reduced descriptions obtained from data

In [16,23] we introduced a different approach which consists of deriving simple dynamical equations directly from the data, instead of obtaining them from a complete mathematical model under various assumptions. Let us define $R = \Sigma_i R_i + M$. While M is a function of $[Ca^{2+}]$ exclusively, each term R_i is a nonlinear function of $[Ca^{2+}]$ and $[\mathrm{CaB}_i]$. We argued in [23] that $R([Ca^{2+}](r,t),[\mathrm{CaB}_i](r,t))$ could be approximated by a function of $[Ca^{2+}]$ and its space and time derivatives. Due to the dissipative nature of the processes involved, we assumed that the lowest order derivatives compatible with the symmetry of the problem suffice to provide a good enough description. Namely, we assumed that R could be approximated by

$$R = g\frac{\partial[\mathrm{Ca}^{2+}]}{\partial t} + f\nabla^2[\mathrm{Ca}^{2+}] + h|\nabla[\mathrm{Ca}^{2+}]|^2 + k, \tag{12}$$

where g, f, h, and k are unknown functions of the free calcium concentration, $[Ca^{2+}]$. The approach then consists of determining the unknown functions directly from the data. Inserting Eq. (12) into Eq. (2), we then obtain a single evolution equation for Ca^{2+} to describe the dynamics. The main assumption that underlies the use of this approach to infer Ca^{2+} currents as done in [16] is that the functional form of R (which is deduced from one data set that corresponds to a

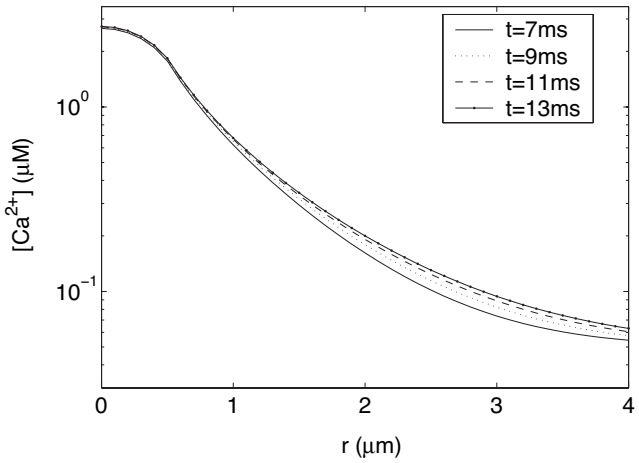

FIG. 1. Calcium concentration in the presence of one buffer and a localized source obtained with numerical simulations of Eqs. (2) and (3). Spatial profiles are shown at four different times, as indicated. Notice that the vertical axis is in logarithmic scale.

particular Ca^{2+} source) remains the same regardless of the source term, Q_{Ca}. Namely, we assume that R describes the dynamics of the removal processes in all circumstances. In this paper, we simplify the functional form of R, taking into account that the sources that we deal with are highly localized in space, and analyze the limits of validity of this assumption.

IV. DATA DRIVEN REDUCTION IN THE PRESENCE OF LOCALIZED SOURCES

Many intracellular Ca^{2+} signals involve Ca^{2+} entry through isolated channels or highly localized clusters of channels [33,34]. Thus it is of relevance to study the dynamics described by Eqs. (2) and (3) when Q_{Ca} is a point source (or a spherical source of very small radius). For this type of sources, the solutions of Eqs. (2) and (3) become almost stationary soon after the source turns on in the region around the source, while the source strength remains constant, as explained in [17,18,35–38]. This is clear, for example, in the case of the diffusion equation in the presence of a point source. In such a case, for each distance, r, there is a time, $t^*(r)$, after which $[Ca^{2+}](r,t)$ is approximately given by the stationary solution, $[Ca^{2+}]_s$. For example, if $r/(\sqrt{4Dt*}) = 0.05$, $[Ca^{2+}](r,t)$ and $[Ca^{2+}]_s$ differ by about 5% for $t > t^*$. If $r=0.6$ nm, which is approximately the size of the pore of an IP_3R, it is $t^*=0.00072$ ms for $D=50$ $\mu m^2/s$ (some typical effective Ca^{2+} diffusion coefficient in the cytosol [31]) while $t^*=0.00012$ ms for $D=300$ $\mu m^2/s$ (the free Ca^{2+} diffusion coefficient in the cytosol [31]). If we take r as the typical size of a single cluster of IP_3R's in the oocyte, $r=50$ nm [39], we obtain $t^*=0.005$ ms for $D=50$ $\mu m^2/s$ and $t^*=0.83$ ms for $D=300$ $\mu m^2/s$. The same type of behavior also occurs with the solutions of Eqs. (2) and (3) in the presence of localized sources, as discussed in [17,18,35–38] and illustrated in Fig. 1. These curves were obtained via numerical simulations of Eqs. (2) and (3) in the presence of one buffer with $k_{on} = 1$ μM^{-1} s^{-1}, $k_{off}=10$ s^{-1}, $[B]_T=100$ μM, $D_B=50$ $\mu m^2/s$ and with a 0.5 pA source that remained on for all $t \geqslant 0$ and that was uniformly distributed over a sphere of radius $R_f = 0.5$ μm and zero outside. We observe that, for each r, there is a time, t^*, after which $[Ca^{2+}]$ remains practically constant. Furthermore, the whole solution becomes time-independent, including the various buffer concentrations.

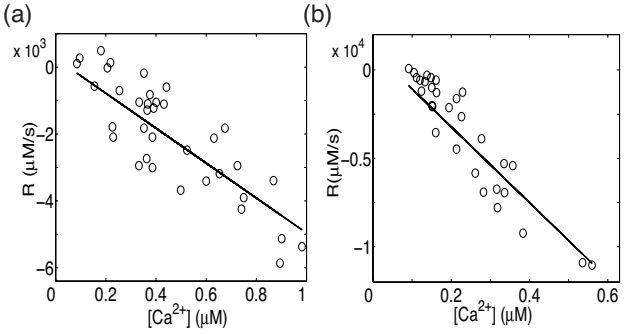

FIG. 2. $R=\Sigma_i R_i+M$ as a function of the free calcium concentration, $[Ca^{2+}]$, for different data sets (circles) and its linear fittings (solid line). (a) Experimental data from calcium sparklets in *Xenopus* oocytes. (b) Experimental data from IP_3-mediated Ca^{2+} puffs in *Xenopus* oocytes.

We observe in Fig. 1 that the stationary solution, $[Ca^{2+}]_s(r)$, that is approached as time goes by, is a monotonically decreasing function of the position, r. Thus the function $[Ca^{2+}]_s(r)$ can be inverted to obtain $r([Ca^{2+}]_s)$. Given that $[Ca^{2+}](r,t) \approx [Ca^{2+}]_s(r)$ for $t > t^*$ and that t^* is relatively small, at least close enough to the source, we can assume that, after a short transient, there is a unique function, $r([Ca^{2+}])$, with $[Ca^{2+}]$ the concentration given by Eqs. (2) and (3). Given that the various buffer concentrations become also approximately constant, we can assume that $R(r,t) \equiv \Sigma_i R_i + M$ in Eq. (2) can be approximated by a function of r or, equivalently, by a function of $[Ca^{2+}]$.

Assuming that the functional form of $R([Ca^{2+}])$ is relatively insensitive to the amplitude and temporal course of the Ca^{2+} source, then we may use the function $R([Ca^{2+}])$ in Eq. (2) to obtain a single dynamical equation that describes the evolution of $[Ca^{2+}]$:

$$\frac{\partial [Ca^{2+}]}{\partial t} = R([Ca^{2+}]) + D_{Ca}\nabla^2[Ca^{2+}] + Q_{Ca}. \qquad (13)$$

Assuming that R is only a function of $[Ca^{2+}]$ is a simplification of the previously proposed ansatz [Eq. (12)] which approximately holds in real experiments where localized Ca^{2+} signals are observed, as illustrated in Fig. 2.

We show in Fig. 2 $R(r,t)$ vs $[Ca^{2+}(r,t)]$, for two sets of experimental data. The free $[Ca^{2+}]$ in these cases and $R(r,t)$ are obtained as done in [16] and briefly explained in the Appendix. Figure 2(a) corresponds to experiments in which Ca^{2+} enters the cytosol from the extracellular medium through voltage-gated Ca^{2+} channels that are expressed in the plasma membrane [40]. Figure 2(b) corresponds to IP_3-evoked Ca^{2+} puffs obtained in *Xenopus Laevis* oocytes with the same protocol as in [41,42], in which Ca^{2+} enters the cytosol from the endoplasmic reticulum through IP_3R's.

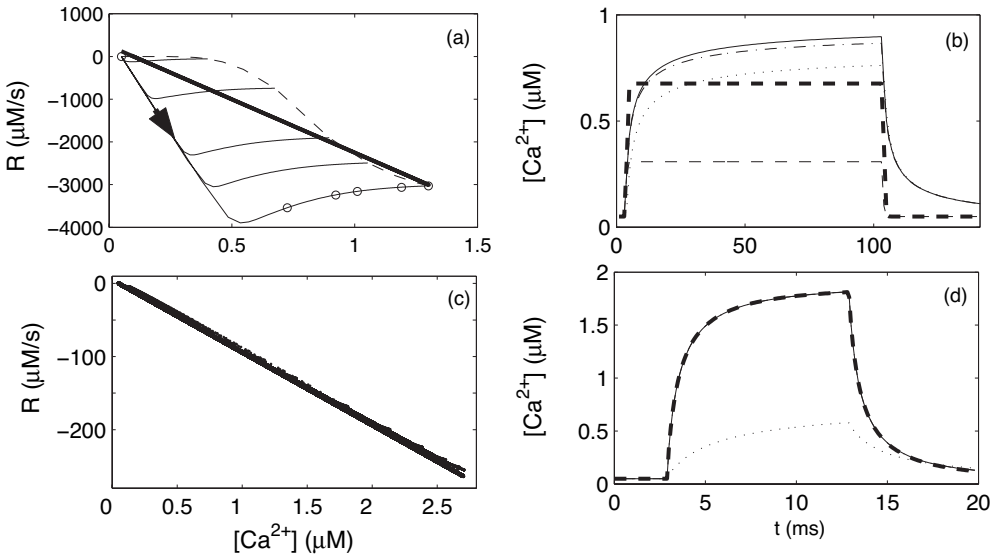

FIG. 3. Sequence of comparative plots for different sets of numerical simulations with a single buffer and a source that is uniformly distributed over a sphere of radius 0.5 μm and is on during a finite time interval. (a) and (b) $k_{on}=100\ \mu M^{-1}\ s^{-1}$, $k_{off}=1000\ s^{-1}$; (c) and (d) $k_{on}=1\ \mu M^{-1}\ s^{-1}$, $k_{off}=10\ s^{-1}$. The other parameter values are given in Table I. (a) and (c): $R(r,t)$ vs $[Ca^{2+}](r,t)$ at times during which the source is on [3 ms$\leqslant t\leqslant$100 ms in (a) and 3 ms$\leqslant t\leqslant$13 ms in (b)] and at five values of r (0, 0.5, 0.6, 0.7, and 1 μm). The points obtained at the each value of r are connected with solid curves. The arrow added to one of them in (a) shows the way in which the values are "visited" as time goes by. The dashed curve [indistinguishable in (c)] corresponds to the relationship between R and $[Ca^{2+}]$ at the stationary solution that the full solution approaches with time. The open circles in (a) correspond to times $t=3$, 4, 6, 8, 23, and 100 ms and are plotted to illustrate how the stationary relationship is approached with time. The straight line corresponds to the best linear fit of all the points, $R(r,t)$ vs $[Ca^{2+}](r,t)$. (b) and (d) Temporal profiles of $[Ca^{2+}]$ at $r=0.6$ μm (right outside the source) obtained with numerical simulations of the full set of reaction-diffusion equations (solid line); of Eq. (14) with μ and $[Ca^{2+}]_0$ determined from the linear fit displayed in (a) and (b) (thick dashed line); of the evolution equations in the EBA (thin dashed line), in the RBA (dotted line), and in the NNA (dashed-dotted lines).

All data sets correspond to fluorescence images obtained with a confocal microscope in linescan mode. Figure 2 displays the results for five signals in (a) and four signals in (b). We see from Fig. 2 that, within the resolution of these experiments, R can be very well approximated by a function of $[Ca^{2+}]$. Furthermore, this function is approximately linear, as shown by the straight lines that correspond to best linear fits to the experimental points.

Assuming that $R([Ca^{2+}])$ can be approximated by a linear function of $[Ca^{2+}]$, Eq. (13) reduces to

$$\frac{\partial[Ca^{2+}]}{\partial t} = -\mu([Ca^{2+}]-[Ca^{2+}]_0)+D_{Ca}\nabla^2[Ca^{2+}]+Q_{Ca}, \tag{14}$$

where μ and $[Ca^{2+}]_0$ are constant parameters to be determined from the data.

An appealing feature of the linear approximation Eq. (14) is that it can be solved analytically for very localized sources Q_{Ca}. In particular, its solution in an infinite medium and in the presence of a point source located at $r=0$ that turns on at $t=0$ is (e.g., see [43,44])

$$[Ca^{2+}]=[Ca^{2+}]_0+\frac{\sigma}{4\pi D_{Ca}r}\exp(-r/\lambda)\mathcal{F}(r,t), \tag{15}$$

where $\lambda=\sqrt{D_{Ca}/\mu}$ and

$$\mathcal{F}(r,t)=\frac{1}{2}\left[\operatorname{erfc}\left(\frac{r}{\sqrt{4D_{Ca}t}}-\sqrt{\mu t}\right)+\exp\left(\frac{2r}{\lambda}\right)\times\operatorname{erfc}\left(\frac{r}{\sqrt{4D_{Ca}t}}+\sqrt{\mu t}\right)\right]. \tag{16}$$

The solution of Eq. (14) in the case of a source that is uniformly distributed over a sphere can be found, for example, in [45].

Depending on how fast the stationary solution is reached, the data-driven linear reduction can be more or less accurate, as we show in Fig. 3. The curves in these figures correspond to two sets of numerical simulations of Eqs. (2) and (3), both of them with a single buffer ($N=1$), no other recapturing mechanism ($M=0$), and a source of the form:

$$Q_{Ca}=\begin{cases}\alpha\mathbf{I} & \text{if } r<R_f \text{ and } t_s<t<t_s+t_d,\\ 0 & \text{otherwise},\end{cases} \tag{17}$$

where $\alpha=5.2\times10^3\ \mu M\ \mu m^3/(\frac{4}{3}\pi R_f^3 pC)$. The two sets have different kinetic rate constants, k_{on} and k_{off}, but the same dissociation constant, $K_d=k_{off}/k_{on}$. The simulations were started from a uniform equilibrium distribution of all species at a resting Ca^{2+} concentration of 50 nM. The model equations were solved numerically using a finite difference scheme on a spherically symmetric grid of length $R_t=6$ μm, with spatial and temporal steps of 0.1 μm and 10^{-5} s, respectively. The total integration time was $T_t=150$ ms. All param-

TABLE I. Parameters used in the simulations with only one buffer.

D_{Ca}	220 $\mu m^2/s$	D_{CaB}	50 $\mu m^2/s$
$[B]_T$	100 μM	$[Ca^{2+}]_{rest}$	0.05 μM
k_{on}^B	100.0–1.0 $\mu M^{-1}\ s^{-1}$	k_{off}^B	1000.0–10.0 s^{-1}
R_f	0.5 μm	t_s	3 ms
t_d	10–80 ms	I	0.5 pA
R_t	1.5 μm	T_t	20–100 ms
Δr	0.1 μm	Δt	10^{-5} s

eter values (from which we have dropped the superscript identifying the buffer for simplicity) are listed in Table I.

We plot in Figs. 3(a) and 3(c) $R(r,t)$ *vs* $[Ca^{2+}(r,t)]$ for five values of r (the distance from the center of the source) and all times, t, during which the source is on. Figure 3(a) corresponds to $k_{on}=100.0\ \mu M^{-1}\ s^{-1}$ and Fig. 3(c) to $k_{on}=1.0\ \mu M^{-1}\ s^{-1}$. We observe that there is a well-defined function $R([Ca^{2+}])$ for all times in Fig. 3(c). Furthermore, $R([Ca^{2+}])$ is very well approximated by a linear function. The assumption that there is a unique function $R([Ca^{2+}])$ for all times is not that good in Fig. 3(a). As expected, it does hold at the stationary solution (shown with dashed lines in the figure). However, the numerical solution differs from the stationary one by less than 30% at $t=8$ ms and by less than 12% at $t=23$ ms, as illustrated by the open circles that correspond to a discrete set of times. Therefore, after a transient that lasts ~10 ms (which in several experiments is not resolved), we can assume that the removal term, R, can be approximated by the function $R([Ca^{2+}])$ that characterizes the stationary solution with at most a 20% error. In this case, however, the function is nonlinear. In any case, we decided to approximate $R(r,t)$ vs $[Ca^{2+}(r,t)]$ by a linear function and obtain the parameters μ and $[Ca^{2+}]_0$ that provided the best linear fit [shown in Figs. 3(a) and 3(c) with a solid line]. With these parameters, we then obtained the solution of Eq. (14) in the presence of the same source and compared it with the numerical solution of the full set of reaction-diffusion equations (RDE). We show a comparison of these solutions in Figs. 3(b) (for $k_{on}=100.0\ \mu M^{-1}\ s^{-1}$) and (d) (for $k_{on}=1.0\ \mu M^{-1}\ s^{-1}$).

We show in Figs. 3(b) and 3(d) the time course of $[Ca^{2+}]$ right outside the source ($r=0.6\ \mu m$) obtained with the numerical simulations of Eqs. (2) and (3) (solid lines) and using the solution of Eq. (14) with the linear fit determined in Figs. 3(a) and 3(c) (thick dashed lines). We observe that both solutions are indistinguishable in the case of Fig. 3(d), for which the linear approximation of $R(r,t)$ vs $[Ca^{2+}(r,t)]$ is good at all times and points [as shown in Fig. 3(c)]. In Fig. 3(b), on the other hand, in spite of the apparently poor quality of the linear fit shown in Fig. 3(a), the solution obtained within this linear approximation differs from the real solution by less than 16% for $3 \leqslant t \leqslant 28$ ms (a typical sparklet duration) and by less than 23% for $3 \leqslant t \leqslant 63$ ms (a typical "puff" duration). We can observe that the solution of the linear approximation reaches its stationary state earlier than the full solution. The stationary solution, on the other hand, is underestimated by the linear fit. This last problem can be partially solved by using the nonlinear function $R([Ca^{2+}])$ that describes the stationary solution, but that results in a larger difference between the solution of the RDE and of the linear model during the transient behavior (data not shown). In spite of the discrepancies, the linear "data-driven" model provides better descriptions than those of model-derived reduced descriptions which result in single evolution equations for $[Ca^{2+}]$, such as the RBA and the EBA described in Sec. III A. This may be observed in Figs. 3(b) and 3(d), where we show the solutions obtained using the RBA (dotted lines) and the EBA (thin dashed lines). In particular, we can observe in Fig. 3(b) that the EBA underestimates the stationary solution by a larger amount than the linear "data-driven" model. This occurs because, in this case, the EBA overestimates the fraction of $[Ca^{2+}]$ that is removed by buffers. The solutions of the EBA, of the linear "data-driven" model, and of the full system are indistinguishable in Fig. 3(d), where the rate of $[Ca^{2+}]$ removal by buffers occurs at a relatively slow pace. The RBA, on the other hand, performs much worse in this case because the buffer is not fast enough. We can also observe in Figs. 3(b) and 3(d) that the NNA (dashed-dotted lines) provides the best description in all cases. Although this model can be solved analytically, it is not characterized by a single evolution equation for $[Ca^{2+}]$ and depends on many parameters. Therefore its practical implementation requires a fairly detailed knowledge of the processes at work.

Figure 3(d) shows that Eq. (14), with the parameters μ and $[Ca^{2+}]_0$ obtained from the fitting in Fig. 3(c), can also describe the Ca^{2+} dynamics after the source is turned off ($t \geqslant 13$ ms). For this reason, we decided to investigate which data points provide parameters, μ and $[Ca^{2+}]_0$, that best describe the $[Ca^{2+}]$ distribution of the real system and how this solution compares with the one that is obtained if a nolinear fit of $R([Ca^{2+}])$ is used instead. To this end, we approximated the R vs $[Ca^{2+}]$ relationship of the data points of Fig. 3(a) in various ways: we used a linear fit of all the data points, a linear fit of the data points that correspond to $63\ ms \leqslant t \leqslant 100$ ms, so that the solution is relatively close to the stationary one, and a nonlinear fit of the stationary solution using a fifth degree polynomial. For the linear fittings we obtained the parameters μ and $[Ca^{2+}]_0$ in Eq. (13) and integrated it numerically, in the presence of the same source as the original simulation. We could observe that, after the source is turned off, all these approximate solutions decay faster than the solution of the full system of equations (data not shown). This is an indication that the data-driven linear model can describe the dynamics after the source is turned off only when Ca^{2+} buffering does not occur too fast. There are no significant differences between using all the data points and using only points near the stationary solution to do the fitting. This is good for the application to real data for which it is not always possible to distinguish when the release of Ca^{2+} has stopped. In the following section, we use linear fittings of all the available points.

We also studied what happened in the case of time-dependent sources. If a sinusoidal source is used in the simulations instead of the steplike one displayed in Eq. (17), R is still well fitted by a linear function of $[Ca^{2+}]$ if the same parameters as the ones in Fig. 3(c) are used (not shown). The

error of the reconstructed calcium signal varies between 2% and 14% depending on whether $[Ca^{2+}]$ is high or low at the corresponding r and t. We conclude that the linear approximation is valid even if the source varies with time. The validity of the lineal reduction approximation is mainly limited by kinetics of the buffers reactions.

V. TESTING THE DATA-DRIVEN LINEAR REDUCTION WITH NUMERICAL SIMULATIONS

We now perform a systematic analysis of the differences between the solutions of the data-driven linear model, Eq. (14), and of the full set of equations (2) and (3) for a variety of parameter values. To this end, we have done numerical simulations of Eqs. (2) and (3), with two or more buffers, one of which, $B \equiv B_1$, corresponds to the dye that is used in fluorescent experiments, with a recapturing mechanism of the form:

$$M = -\frac{k_1[Ca^{2+}]^m}{k_2^m + [Ca^{2+}]^m}, \tag{18}$$

and a source of the form Eq. (17).

The results of these simulations were then compared with the corresponding solution of Eq. (14), in which μ and $[Ca^{2+}]_0$ were obtained from a linear fit of R vs $[Ca^{2+}]$, as done in the previous section. To quantify the differences between the full calculation, RDE, and the reduced *linear* one, we defined the error:

$$\epsilon(t) = \frac{\int |[Ca^{2+}]_{RDE} - [Ca^{2+}]_{linear}| dr}{\int [Ca^{2+}]_{RDE} dr}, \tag{19}$$

in which the domain of integration is chosen as either the whole space or the region with $(r < R_f)$ or without $(r > R_f)$ a source.

Figures 4 and 5 correspond to simulations with only one additional buffer, $E \equiv B_2$, which has the properties of the exogenous Ca^{2+} buffer, EGTA. In these simulations the source is on from $t=3$ to 13 ms, unless otherwise noted. The other parameters are listed in Table II.

In Fig. 4 we analyze how the error depends on the dye reaction constants, k_{on}^B and k_{off}^B. These two parameters are varied from 100 to 1000, in the corresponding units ($\mu M^{-1}\,s^{-1}$ for k_{on}^B and s^{-1} for k_{off}^B). The other simulation parameters are kept as in Table II. Figures 4(a) and 4(b) show error contour lines at $t=12$ ms and $t=13.5$ ms, respectively, which correspond to times before and after the source turns off. For a given value of k_{on}^B, the error increases with k_{off}^B, while it decreases if k_{off}^B is fixed and k_{on}^B is increased. In terms of $K_D^B = k_{off}^B / k_{on}^B$, K_D^B we observe that the error mainly increases with K_D^B. In Fig. 4(a), the error is less than 0.1 (which corresponds to a 10% error), while in Fig. 4(b) it goes up to 0.4. Figures 4(c) and 4(d) show error contour lines at $t=12$ ms but performing the integral in Eq. (19) inside (c) and outside (d) the source ($0 \leqslant r \leqslant = R_f = 0.05$ μm and $0.05 \leqslant r \leqslant = R_t = 1.5$ μm, respectively). We obtain errors that are less than 2% integrating only inside the source, and get up to 16% integrating only outside. We may conclude that the approximation is more accurate while the source is on, and gets worse as we move out of the source, with an error that is always smaller than 16%, for this particular model and choice of parameters.

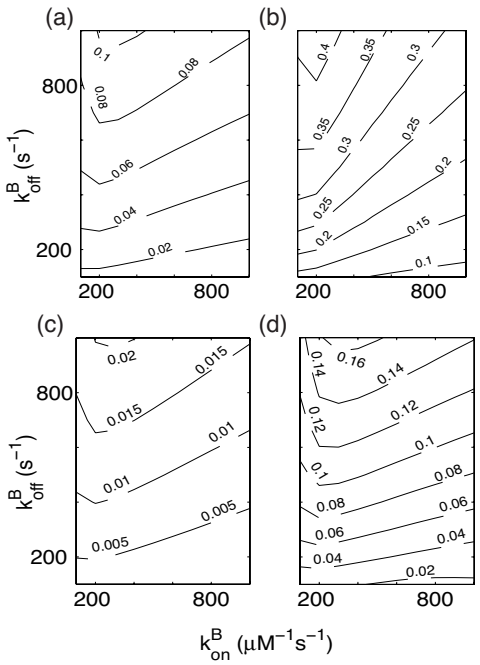

FIG. 4. Error dependence on the dye reaction constants k_{on}^B and k_{off}^B, for the model with Ca^{2+}, dye, and one extra buffer. Simulation parameters as in Table II. (a) and (b) Error contour lines for $\epsilon(t_1 = 12$ ms) and $\epsilon(t_2 = 13.5$ ms), respectively. The source is on from $t = 3$ to 13 ms. (c) and (d) Error contour lines for $t_2 = 12$ ms but integrating: (c) only inside the source, from $r=0$ to $r = R_f = 0.05$ μm; and (d) only outside the source, form $r = R_f = 0.05$ μm to $r = R_t = 1.5$ μm.

In Fig. 5(a) we analyze the error dependence on the Ca^{2+} and dye diffusion coefficients, D_{Ca} and D_{dye}, respectively, varying them between 50 and 320 $\mu m^2/s$. The other simulation parameters are kept as in Table II. Error contour lines are shown for $t=12$ ms. The error is not affected much by changes in D_{Ca}, while it decreases when D_{dye} is increased. This is a positive result since in real experiments one has no control on D_{Ca} but, conversely, one can choose different dyes to improve the approximation. Even in the worst case, the error is less than 5%.

In Fig. 5(b) we analyze the error dependence on the magnitude, I, and duration, t_d, of the source. The current is in the range 0.1–1.9 pA, which is typical for localized Ca^{2+} puffs in oocytes [46], and the duration, in the range 1–10 ms, is somewhat smaller than typical puff durations [29]. The other simulation parameters are kept as in Table II. Error contour lines are shown, in this case, at $t=4$ ms, a time at which the source is on for all the values of t_d considered. In this case the error is not affected by changes in t_d, but they do increase as I increases. As before, even in the worst case, the error is less than 5%. We have also studied how the errors change when R_f (the source radius) or $[B_T]$ (the total dye concentra-

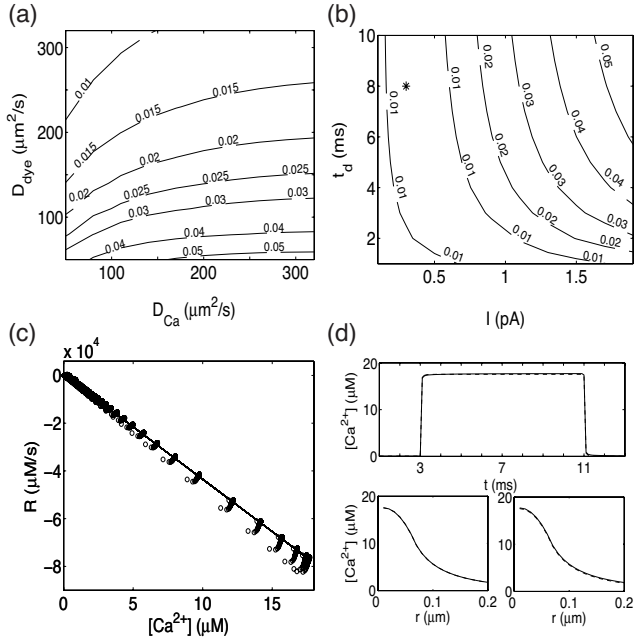

FIG. 5. (a) Error dependence on free calcium and dye diffusion constants, D_{Ca} and D_{dye}, respectively, for the model with Ca^{2+}, dye, and one extra buffer. Simulation parameters as in Table II. Error contour lines for $\epsilon(t=12$ ms) (the source is on from $t=3$ to $t=13$ ms). (b) Error dependence on the magnitude, I, and duration, t_d, of the source, for the model with Ca^{2+}, dye, and one extra buffer. Simulation parameters as in Table II. Error contour lines for $\epsilon(t=4$ ms) (the source is on from $t=3$ ms to $t=3$ ms$+t_d$). (c) R vs $[Ca^{2+}]$ (circles) and its best linear fitting (solid line) for the simulation indicated with an asterisk in (b). (d) RDE (solid line) and *linear* (dashed line) free calcium profiles for the same simulation. Upper panel, temporal profile for $r=0$; lower panels, spatial profiles for $t_1=4$ ms (left) and $t_2=10$ ms (right).

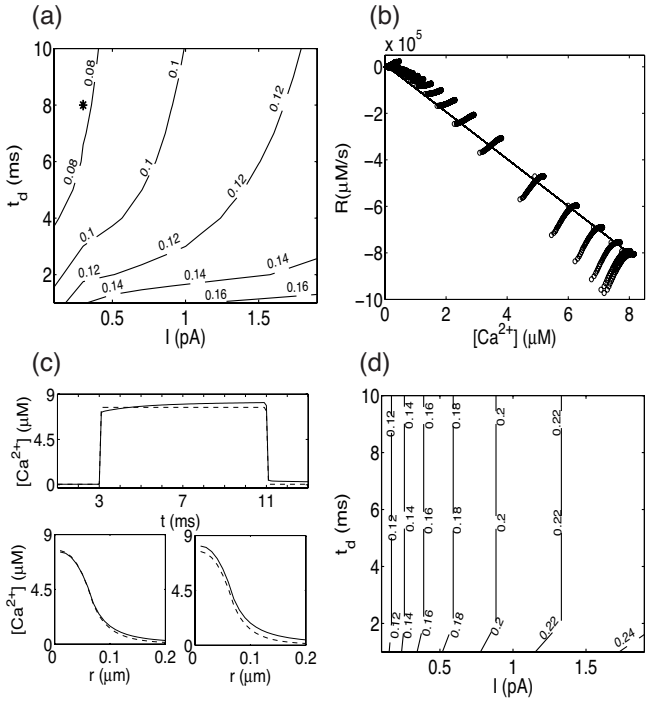

FIG. 6. (a) Error dependence on the magnitude, I, and duration, t_d, of the source, for the model with various buffers. Simulation parameters as in [16]. Error contour lines for $\epsilon(t=4$ ms). (b) R vs $[Ca^{2+}]$ (circles) and its best linear fitting (solid line) for the simulation indicated with an asterisk in (a). (c) RDE (solid line) and *linear* (dashed line) free calcium profiles for the same simulation. Upper panel, temporal profile for $r=0$; lower panels, spatial profiles for $t_1=4$ ms (left) and $t_2=10$ ms (right). (d) The same as (a) but using only one pair of values, μ and $[Ca^{2+}]_0$, for the whole set of simulations. μ and $[Ca^{2+}]_0$ were obtained from the linear fitting of the numerical simulation done with $I=1.9$ pA and $t_d=10$ ms. Notice that the maximum value for ϵ in this case is 0.24 while it is 0.16 in (a).

tion) are changed. Varying R_f from its value in Fig. 5(b), 0.05 to 0.1 μm, we obtained errors in the range 0.01–0.06. Varying $[B_T]$ from 40 to 100 μM does not change the errors obtained for $[B_T]=40$ μM (data not shown). We show in Figs. 5(c) and 5(d) plots of $R(r,t)$ vs $[Ca^{2+}](r,t)$ and of $[Ca^{2+}]$ as a function of time and space for a particular set of parameter values [the one indicated with an asterisk in Fig. 5(b)]. We show in Fig. 5(c) the points obtained from the numerical simulation of the full set of equations (circles) and its best linear fit (solid line). Figure 5(d) compares the RDE (solid line) and *linear* (dashed line) $[Ca^{2+}]$ profiles. The upper panel is a temporal profile for $r=0$; lower panels are spatial profiles at $t=4$ ms (left) and $t=10$ ms (right). The differences between both solutions are indistinguishable in the scale of the figure.

TABLE II. Parameters used in the simulations with dye and one extra buffer.

D_{Ca}	220 $\mu m^2/s$	$D_{\mathrm{Ca}B}$	50 $\mu m^2/s$
$D_{\mathrm{Ca}E}$	113 $\mu m^2/s$	$[Ca^{2+}]_{\mathrm{rest}}$	0.05 μM
$[B]_T$	40 μM	$[E]_T$	1000 μM
k_{on}^B	100.0 $\mu M^{-1}\ s^{-1}$	k_{off}^B	400.0 s^{-1}
k_{on}^E	1.5 $\mu M^{-1}\ s^{-1}$	k_{off}^E	0.3 s^{-1}
R_f	0.05 μm	t_d	10 ms
I	0.1 pA	k_1	50.0 $\mu M\ s^{-1}$
k_2	0.184 μM	m	3.9
R_t	1.5 μm	T_t	20 ms
Δr	0.01 μm	Δt	10^{-7} s

We have also analyzed the behavior of R in a more realistic model of intracellular Ca^{2+} dynamics, more specifically, in the modified version of the model of Ca^{2+} sparks in frog skeletal muscle [6] studied in [16]. All parameters of the model are described in [16]. In this case we used a source of the form Eq. (17) with $R_f=0.05$ μm, and different values of t_d and I. The parameters R_t, T_t, Δr, and Δt are as in Table II.

In Fig. 6(a) we analyze the error dependence on the magnitude, I, and duration, t_d, of the source, as done in Fig. 5(b), but for this new model. The errors, in this case, are more sensitive to changes in t_d than in Fig. 5(b). While in that case the errors were in the range 0.01–0.05, they are now in the range 0.08–0.16 (which means less than 16%). Figures 6(b) and 6(c) are similar to Figs. 5(c) and 5(d) but obtained for this new model and for the parameter values indicated with

an asterisk in Fig. 6(a). The differences between the RDE (solid line) and the *linear* solutions are more noticeable in this case than in that of Fig. 5(d), but the errors never exceed 16%.

VI. ABILITY OF THE LINEAR DATA-DRIVEN MODEL TO BE INDEPENDENT OF Ca^{2+} CURRENT PROPERTIES

Figures 4, 5, and 6(a)–6(c) were done comparing the numerical solution of the full model with the one of Eq. (14) with parameters μ and $[Ca^{2+}]_0$ that were determined from the linear fitting of R vs $[Ca^{2+}]$ for each simulation separately. In other words, each point in Figs. 4, 5(a), 5(b), and 6(a) was obtained using a different pair of parameters $(\mu,[Ca^{2+}]_0)$. It is reasonable that μ and $[Ca^{2+}]_0$ change with buffer parameters. However, if μ and $[Ca^{2+}]_0$ (and, therefore, R) depend on the current properties (like I and t_d), then the "data-driven" linear model is only descriptive: a different model is obtained for each different observation. As we have mentioned in the Introduction, our goal is to obtain a single data-driven model that can describe Ca^{2+} dynamics in the presence of different Ca^{2+} sources. A reduced model with this property can be used to infer the $[Ca^{2+}]$ current from a given fluorescence image. Furthermore, it allows the simultaneous use of data coming from several Ca^{2+} signals, which is highly desirable given the relatively small amount of information that each isolated signal usually provides. In this section we investigate the ability of the model defined by Eq. (14) with a single pair of values, μ and $[Ca^{2+}]_0$, to reproduce solutions corresponding to $[Ca^{2+}]$ currents of various kinetics and amplitudes.

To this end we integrated Eq. (14) with the currents, I, and time durations, t_d, explored in Fig. 5(b) but using always the same set of parameter values, μ and $[Ca^{2+}]_0$, obtained from one of the simulations of this figure. Actually, we repeated this test choosing the values $(\mu,[Ca^{2+}]_0)$ obtained for the largest and smallest values of I and t_d [i.e., for $v_1=(I=0.1$ pA; $t_d=1$ ms), $v_2=(0.1$ pA; 10 ms), $v_3=(1.9$ pA; 1 ms), and $v_4=(1.9$ pA; 10 ms)]. We did the same with the simulations of Fig. 6(a). We show in Fig. 6(d) the error contour lines obtained for this second model using the values $(\mu,[Ca^{2+}]_0)$ determined from the simulation with $I=0.1$ pA and $t_d=1$ ms, which gives the largest errors. Comparing Figs. 6(d) and 6(a) we observe that, in the worst case, the error increases from 0.16 to 0.24. For the model of Fig. 5(b), the error increases from 0.05 to 0.1 in the worst case (that corresponds to using v_4 as the fitting simulation). From these results we conclude that the linear data-driven model (14) with a single set of parameters, $(\mu,[Ca^{2+}]_0)$, is able to reproduce the $[Ca^{2+}]$ distribution that arises due to currents of different intensities and durations, with errors that are less than 25% in the worst case, at least for the range of currents and durations that characterize localized Ca^{2+} signals in oocytes.

VII. TESTING THE DATA-DRIVEN LINEAR REDUCTION WITH REAL DATA

We now test the data-driven linear model with experimental signals that arise due to Ca^{2+} entry through voltage gated channels expressed in the plasma membrane of *Xenopus laevis* oocytes [40]. The signals were obtained with optical techniques using a confocal microscope in line-scan mode and the fluorescent Ca^{2+} indicator fluo-4 dextran (low affinity). The fluorescent images provide information on the spatiotemporal distribution of the Ca^{2+}-bound dye, $[CaB]$, which is assumed to be proportional to the fluorescence intensity [16]. The distribution of free $[Ca^{2+}]$ can be obtained from the reaction term in Eq. (4), as done in [16] and briefly explained in the Appendix. The data analyzed in this section corresponds to the one displayed in Fig. 2(b).

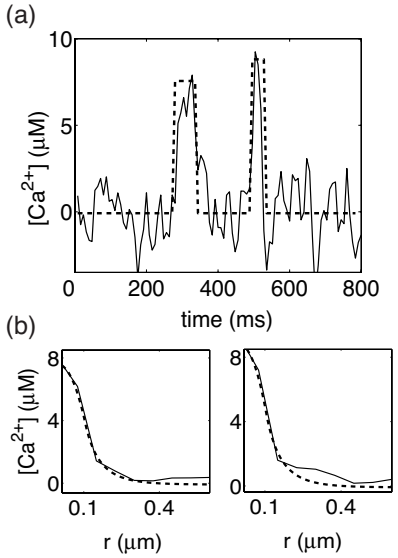

FIG. 7. Results of the forward integration of Eq. (14) for the experimental data in Fig. 2. Simulated profiles with dashed lines, experimental profiles with solid lines. (a) Temporal profiles for $r=0$. (b) Spatial profiles for the two temporal peaks in (a).

We first analyzed the data as done in [16] to estimate the Ca^{2+} source, Q_{Ca}, that gave rise to each signal. We then obtained the parameters, μ and $[Ca^{2+}]_0$, that characterized the best linear fit of the $R([Ca^{2+}])$ relationship [the straight line in Fig. 2(b)]. We then chose one particular (localized) signal and approximated the Ca^{2+} source, Q_{Ca}, by a piecewise constant (in space and time) one, $\overline{Q}_{Ca}$, chosen so as to give the same mean current as Q_{Ca}. Finally, we integrated Eq. (14) using the source, $\overline{Q}_{Ca}$, and the values μ and $[Ca^{2+}]_0$ of the linear fit. We show in Fig. 7 the temporal and spatial profiles of $[Ca^{2+}]$ obtained in this way (dashed lines) comparing them with the ones determined directly from the experimental image (solid lines). The temporal profiles in (a) correspond to $r=0$ and the spatial profiles in (b) correspond to the two temporal peaks in (a), with errors of ~14% and ~6%, respectively. Similar results were obtained for the other signals.

The results in this section show that a single set of parameter values, μ and $[Ca^{2+}]_0$, characterizes a variety of experimental Ca^{2+} signals [see Fig. 2(b)] and that the linear data-driven model that uses these values is able to reproduce the experimentally determined $[Ca^{2+}]$ distributions. This holds as long as the signals arise from sufficiently localized Ca^{2+} sources. We thus expect the linear data-driven model to pro-

vide an accurate means to infer Ca^{2+} currents from fluorescent images that is easier and simpler to automatize than the algorithm discussed in [16].

VIII. CONCLUSIONS

In this paper we have shown that the dynamics of intracellular Ca^{2+} in the presence of localized sources can be accurately described by a single evolution equation for $[Ca^{2+}]$ whose parameters can be directly inferred from individual solutions and are relatively insensitive to the Ca^{2+} source properties. We have argued that the success of this approach relies on the localized nature of the Ca^{2+} sources. Therefore we expect similar "data-driven models" to be applicable in reaction-diffusion systems in which the entrance of the reactants into the reaction region is spatially localized to almost a point.

In the presence of very localized ($\sim$ point) sources, the solutions of reaction-diffusion systems become almost stationary very quickly in the region around the source. For typical Ca^{2+} signals, this time is $\lesssim 1$ ms within a radius of ~ 1.5 μm around the Ca^{2+} point of entry. In these stationary solutions, $[Ca^{2+}]$ is a decreasing function of the distance, r, to the point of release. Thus when the solution of the system becomes almost stationary, it is possible to invert the relation between r and $[Ca^{2+}]$ and write any function of r (among them, the terms that characterize Ca^{2+} removal in the cytosol) as a function of $[Ca^{2+}]$. In real Ca^{2+} signals observed in *Xenopus laevis* oocytes the relationship between the sum of all the Ca^{2+} removal terms and $[Ca^{2+}]$ is approximately linear. Thus the dynamics of $[Ca^{2+}]$ in the presence of a very localized source, many buffers and pumps, can be described in terms of a single linear evolution equation for $[Ca^{2+}]$.

In this paper we have shown that the two parameters that characterize the sum of all the Ca^{2+} removal terms can be obtained from individual solutions. We have probed this "linear data-driven model" both with numerical simulations and with experimental Ca^{2+} signals observed in *Xenopus Laevis* oocytes. In the case of the simulations, we determined that the solution of the "linear data-driven model" reproduced the $[Ca^{2+}]$ spatiotemporal distribution observed with less than 25%. The ability to reproduce the experimentally observed Ca^{2+} signals (with less than 14%) implies an experimental validation for the reduced description that we introduce in this paper. The simplicity of the model, which is characterized by very few parameters and does not require an *a priori* detailed knowledge of the processes that affect Ca^{2+} dynamics, makes it of wide applicability for the analysis of real Ca^{2+} signals.

ACKNOWLEDGMENTS

We are grateful to Ian Parker, Angelo Demuro, and Sheila Dargan for having provided the experimental data that were analyzed in this paper. This research was supported by NIH BRP Grant R01GM65830-01, PICT 03-08133 (ANPCyT, Argentina), UBACyT X099, and Fundación Antorchas. S.P.D. is a member of the Carrera del Investigator Científico (CONICET).

APPENDIX: HOW TO OBTAIN $[Ca^{2+}]$ AND R FROM EXPERIMENTAL DATA.

The steps that must be followed in order to obtain $[Ca^{2+}]$ and R from experimental data are largely described in [16]. Here we present a brief summary of the whole process.

Fluorescent images provide information on the distribution of Ca^{2+}-bound dye, $[CaD]$, whose evolution equation can be assumed to be (under some standard hypothesis)

$$\frac{\partial [CaD]}{\partial t} = -R_{dye} + D_D \nabla^2 [CaD], \tag{A1}$$

where D_D is the dye diffusion coefficient and R_{dye} is the reaction term:

$$R_{dye} = -k_{on}[Ca^{2+}]([D]_T - [CaD]) + k_{off}[CaD]. \tag{A2}$$

We can compute the time and spatial derivatives of $[CaD]$ from the experimental records. Line scan images only contain data along one spatial dimension, x. Assuming that the release of Ca^{2+} has spherical symmetry (this is satisfied, in particular, if the release occurs over a very small region that can be treated as a point and that there is nothing else around that can break this symmetry), the solution only depends on time and the distance, r, from the site of release. $\nabla^2[CaD]$ can, therefore, be computed using only the information along the direction x. Calculating $\nabla^2[CaD]$ and $\partial[CaD]/\partial t$ we can obtain from Eq. (A1) the value of R_{dye} at every spatial point and time:

$$R_{dye} = D_D \nabla^2 [CaD] - \frac{\partial [CaD]}{\partial t}. \tag{A3}$$

Using Eq. (A2) we can then obtain a spatio-temporal series of the variable of interest, $[Ca^{2+}]$:

$$[Ca^{2+}] = \frac{k_{off}[CaD] - R_{dye}}{k_{on}([D]_T - [CaD])}. \tag{A4}$$

Equation (13) has two unknowns, R (R_{dye} is one of the terms in the sum inside R) and the Ca^{2+} source, Q_{Ca}. The fact that Ca^{2+} release is localized in space and time implies that Q_{Ca} vanishes away from the source and everywhere in space when Ca^{2+} release ceases. Thus there are points of the experimental record where $Q_{Ca}=0$ and R is then the only unknown term in Eq. (13). R is then reexpressed as a function of $[Ca^{2+}]$. Now, in a given experiment, $[Ca^{2+}]$ will be larger in the region with $Q_{Ca} \neq 0$ than in the region where there is no source. However, the "table" of R values as a function of $[Ca^{2+}]$ is obtained in the region with $Q_{Ca}=0$. This means that we will not have direct information on the values of R for the larger $[Ca^{2+}]$ values that occur in the region with $Q_{Ca} \neq 0$. One possibility to overcome this difficulty is to extrapolate the fitting of R vs $[Ca^{2+}]$, in order to obtain an estimate of its value for these larger $[Ca^{2+}]$, using the information for the lower concentrations. In the experiments analyzed in this paper R vs $[Ca^{2+}]$ has an approximately linear behavior, so the extrapolation is a reliable procedure.

[1] A. L. Hodgkin and A. F. Huxley, J. Physiol. (London) **117**, 500 (1952).
[2] M. J. Berridge, M. D. Bootman, and P. Lipp, Nature (London) **395**, 645 (1998).
[3] M. Nishiyama, K. Hong, K. Mikoshiba, M. Poo, and K. Kato, Nature (London) **408**, 584 (2000).
[4] M. T. Nelson, H. Cheng, M. Rubart, L. F. Santana, A. D. Bonev, H. J. Knot, and W. J. Lederer, Science **270**, 633 (1995).
[5] M. D. Bootman, P. Lipp, and M. J. Berridge, J. Cell. Sci. **114**, 2213 (2001).
[6] S. M. Baylor, S. Hollingworth, and W. K. Chandler, J. Gen. Physiol. **120**, 349 (2002).
[7] D. Fraiman and S. Ponce Dawson, Cell Calcium **35**, 403 (2004).
[8] M. Falcke, J. L. Hudson, P. Camacho, and J. D. Lechleiter, Biophys. J. **77**, 37 (1999).
[9] G. W. De Young and J. Keizer, Proc. Natl. Acad. Sci. U.S.A. **89**, 9895 (1992).
[10] G. D. Smith, J. Wagner, and J. Keizer, Biophys. J. **70**, 2527 (1996).
[11] J. H. Horne and T. Meyer, Science **276**, 1690 (1997).
[12] N. Callamaras and I. Parker, J. Gen. Physiol. **113**, 199 (1999).
[13] A. Takahashi, P. Camacho, J. D. Lechleiter, and B. Herman, Physiol. Rev. **79**(4), 1089 (1999).
[14] I. Parker, N. Callamaras, and W. G. Wier, Cell Calcium **21**, 441 (1997).
[15] The dye acts as a Ca^{2+} buffer and, if its reaction rates and diffusion coefficient are known and if it only binds Ca^{2+} then the distribution of Ca^{2+} can be obtained as done in [16].
[16] A. C. Ventura, L. Bruno, A. Demuro, I. Parker, and S. Ponce Dawson, Biophys. J. **88**, 2403 (2005).
[17] G. D. Smith, Biophys. J. **71**, 3064 (1996).
[18] G. D. Smith, L. Dai, R. M. Miura, and A. Sherman, Surv. Math. Ind. **61**, 1816 (2001).
[19] D. E. Strier and S. P. Dawson, J. Chem. Phys. **112**, 825 (2000).
[20] J. Wagner and J. Keizer, Biophys. J. **67**, 447 (1994).
[21] E. Neher, in *Calcium Electrogenesis and Neuronal Functioning, Exp. Brain Res. 14*, edited by U. Heinemann, M. Klee, E. Neher, and W. Singer (Springer-Verlag, Berlin, 1986), pp. 80–96.
[22] M. Naraghi and E. Neher, J. Neurosci. **17**, 6961 (1997).
[23] A. C. Ventura, L. Bruno, and S. Ponce Dawson, Physica A **342**, 281 (2004).
[24] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
[25] G. B. Mindlin, X. Hou, H. G. Solari, R. Gilmore, and N. B. Tufillaro, Phys. Rev. Lett. **64**, 2350 (1990).
[26] G. Gouesbet, Phys. Rev. A **46**, 1784 (1992).
[27] G. B. Mindlin, N. Merener, and P. T. Boyd, Europhys. Lett. **42**, 31 (1998).
[28] N. Callamaras, J. S. Marchant, X.-P. Sun, and I. Parker, J. Physiol. (London) **509**, 81 (1998).
[29] X.-P. Sun, N. Callamaras, J. S. Marchant, and I. Parker, J. Physiol. (London) **509**, 67 (1998).
[30] J. Sneyd, A. LeBeau, and D. Yule, Physica D **145**, 158 (2000).
[31] N. L. Allbritton, T. Meyer, and L. Stryer, Science **258**, 1812 (1992).
[32] J. Sneyd, P. Dale, and A. Duffy, SIAM J. Appl. Math. **58**, 1178 (1998).
[33] I. Parker, J. Choi, and Y. Yao, Cell Calcium **20**, 105 (1996).
[34] M. J. Berridge, Neuron **21**, 13 (1998).
[35] M. D. Stern, Cell Calcium **13**, 183 (1992).
[36] G. D. Smith, J. E. Keizer, M. D. Stern, W. J. Lederer, and H. Cheng, Biophys. J. **75**, 15 (1998).
[37] R. Thul and M. Falcke, Biophys. J. **86**, 2660 (2004).
[38] M. Falcke, Adv. Phys. **53**, 255 (2004).
[39] S. Swillens, G. Dupont, L. Combettes, and P. Champeil, Proc. Natl. Acad. Sci. U.S.A. **96**, 13750 (1999).
[40] A. Demuro and I. Parker, Cell Calcium **34**, 499 (2003).
[41] S. L. Dargan and I. Parker, J. Physiol. (London) **553**, 775 (2003).
[42] S. L. Dargan and I. Parker, J. Physiol. (London) **556**, 447 (2004).
[43] P. C. Pape, D. S. Jong, and W. K. Chandler, J. Gen. Physiol. **106**, 259 (1995).
[44] P. C. Pape, D. S. Jong, and W. K. Chandler, J. Gen. Physiol. **112**, 263 (1998).
[45] W. Cong, L. V. Wang, and G. Wang, Biomed. Eng. Online **3**, 1 (2004).
[46] L. Bruno, A. C. Ventura, S. Dargan, I. Parker, and S. Ponce Dawson (in preparation).