Dynamical models in neuroscience from a closed-loop control perspective

Sebastián Martínez, Demián García-Violini, Mariano Belluscio, Joaquín Piriz and Ricardo Sánchez-Peña

***Abstract*— Modifying neural activity is a substantial goal in neuroscience that facilitates the understanding of brain functions and the development of medical therapies. Neurobiological models play an essential role, contributing to the understanding of the underlying brain dynamics. In this context, control systems represent a fundamental tool to provide a correct articulation between model stimulus (system inputs) and outcomes (system outputs). However, throughout the literature there is a lack of discussions on neurobiological models, from the formal control perspective. In general, existing control proposals applied to this family of systems, are developed empirically, without theoretical and rigorous framework. Thus, the existing control solutions, present clear and significant limitations. The focus of this work is to survey dynamical neurobiological models that could serve for closed-loop control schemes or for simulation analysis. Consequently, this paper provides a comprehensive guide to discuss and analyze control-oriented neurobiological models. It also provides a potential framework to adequately tackle control problems that could modify the behavior of single neurons or networks. Thus, this study constitutes a key element in the upcoming discussions and studies regarding control methodologies applied to neurobiological systems, to extend the present research and understanding horizon for this field.**

***Index Terms*— Closed-loop, control systems, dynamical systems, feedback, neurobiological models, neuron**

I. Introduction

Neurobiological models have been actively developed during the past decades. In general, these models aim to describe the dynamical behavior of different nervous system structures at cell or circuit-level.

In this sense, the study of these complex systems and processes requires that the proposed models capture the fundamental dynamics for simulation, prediction, or control purposes either *in silico*, *in vitro*, or *in vivo* monitored environments. Such models range from highly detailed descriptions, involving thousands of coupled differential equations for large networks, to greatly simplified structures, which are useful for studying the mean population response in a computationally efficient manner. Given the broad variety of approaches and models, choosing the most appropriate level of description for a particular application (e.g. research, interfacing or disorder treatment), requires a careful assessment of the experimental information available as well as the technical specifications.

Oversimplified models can produce misleading results, whereas excessively detailed models, with a massive degree of complexity, can mask essential results and/or produce computational and numerical problems. By way of example, the renowned Hodgkin-Huxley model [1] describes the dynamics of a spiking neuron in a biologically detailed way that could be unsuitable for network-level closed-loop applications due to its complexity, numerical problems, and associated delays.

Neurobiological models can be essentially separated into two main categories: 1) single neuron models; and 2) global population models. On the one hand, the former characterise the behavior of a neuron, usually describing the interplay of ionic currents, that affect the cell potential. Next, in order to increase its range and applicability, these single models consider different functional interactions (excitatory or inhibitory) and connectivity, so that they can be used to construct network models to describe large neurobiological populations. Beyond the fact that a single neuron could be modeled as multi-compartmental, and thus requiring multiple equations on its own (e.g. pyramidal cells in [2]), using simple point neurons to model networks, by itself, can scale the equations (generally non-linear) to a huge number imposing a significant computational burden [3]. On the other hand, population models are generally employed to describe the behavior of massive arrangement of neurons. Thus, simplified descriptions, resulting on a reduced number of differential equations, can be used to model the dynamics of the entire ensemble. Temporal evolution and/or spatial variation of states, like local field potentials (LFPs), or global firing rates, are typically described by this family of models.

During the past decades, new technological developments have opened new opportunities related to experimental research [4], [5]. In particular, temporo-spatial precision and accuracy have been markedly improved, due to new acquisition, sensing, and actuation systems. This is the case of optogenetic approaches, which extend the possibility to perform control of neuronal activity with high temporo-spatial accuracy and

selectivity. Analogously, these light-based techniques have inspired new strategies to measure the synaptic/electrical activity of neurons, at both single-cell or circuit scales. Traditional systems, based on electrical actuation or sensing mechanisms, have been also improved during the past years. Efficient, micro-electrode arrays have been adapted to a number of different applications using several geometrical structures, materials, and also new and advanced manufacturing processes [6], [7].

Within this context, a correct articulation between high-precision actuation and sensing systems in neurobiological systems, can consequently provide an effective tool to enrich the understanding of complex neural processes and mechanisms. Similarly, as it has been shown throughout the literature, correct interaction of actuation and sensing systems can result in promising treatments of a wide range of neurological diseases [8]–[11]. Thus, control systems, mainly feedback control systems, can play a significant role, systematizing this interaction. By way of example, recent studies have presented results implementing control routines in neurobiological structures [12], [13]. However, in most cases, the proposed control systems are designed, either using simple assumptions and control laws not based in the theoretical background about the process, or without any a-priori knowledge of the system dynamics, i.e. lacking a model that describes the underlying dynamical behavior of the process. It is worth mentioning that, as a general rule in control systems, the more knowledge there is about the system, the more accurate the results can be. To this end, a model that captures the essential dynamical behavior of the neurobiological system, directly contributes to improving the resulting performance. However, to the best of the authors knowledge, to date there is no literature review that analyses neurobiological models from the control perspective, covering its dynamical features, and describing the inputs and outputs of these systems. In particular, the translation between model inputs and outputs into a practical, feasible technology, represents another key challenge for an effective model consideration regarding any neurobiological control problem.

The motivation of this study is to provide an analysis of the most relevant neurobiological models available in the literature, from a control-oriented point of view. Thus, to standardise the discussion, the main features of the most representative models are presented, highlighting those elements that are relevant for control purposes. Taking into consideration the vast existing theory of control systems, and the increasing interest in closed-loop applications in the area of neurosciences, this work aims to fulfill the existing gap in the literature regarding the rigorous approach to control-oriented neurobiological models and its applications. It goes a step further, and also describes implementation characteristics as sensors and actuators.

The remainder of this paper is organised as follows. Section II discusses the main features related to single neuron models and Section III provides a general perspective related to population models. The control perspective vision in these previous sections is further explained and deepened in Section IV, along with references from *state-of-the-art* applications. Finally, in Section V the main features of the presented models, are discussed from an overall perspective.

II. Single neuron models

From the fundamental observation that a neuron performs a summation (also interchangeably referred to as integration) of incoming inputs, and then a comparison with a certain threshold, a process that determines if the neuron fires a spike, several models for this computational property has been proposed. From an engineering point of view, as in many areas of application, there is an inherent trade-off between simplicity and level of description; when both qualities are balanced, the model is said to be parsimonious. For example, a simple model to predict this all-or-none response is the McCulloch-Pitts model [14], seminal in the development of Artificial Neural Networks [15]. Although this model is computationally efficient in its discrete-time, binary-output formulation, it does not account for the neuron dynamics (e.g the capacitive property of the membrane), nor the time-course of the input post synaptic potentials (PSPs) or their relative timing.

Towards a time description of a neuron potential (voltage), and based on several observations including the net current flow through ion channels (conductances), and the aforementioned dielectric nature of the lipidic bi-layer that conforms the membrane (capacitance), an R-C circuit analogy can be proposed. The resultant model is often referred to as Leaky integrate-and-fire (LIF) model [16], [17] and describes, with a linear differential equation, the relation between these three electrical properties according to Ohm's law. Here, the system is characterized by a one dimensional physically-meaningful state, i.e. the voltage

$$C_M \dot{V} = I - g_{leak}(V - E_{leak}), \tag{1}$$

where C_M is the capacitance, I the current, g_{leak} the conductance, and E_{leak} the equilibrium potential given by Nernst equation for that ion species. This differential equation can be rescaled or normalized as

$$\dot{v} = a - v. \tag{2}$$

Whereas the sub-threshold variation can obey this simple form, the spike is modeled with a reset condition:

$$\text{if } v = 1, \text{ then } v \leftarrow 0, \tag{3}$$

that is, when the voltage reaches a threshold (usually 1 if the parameters are normalized), the spike occurs and the membrane voltage is instantaneously set to an hyperpolarized value associated with the repolarizing effect of the outward potassium current. It should be noted that this model does not produce spikes per se, since it lacks a mechanism to that end, but it can handle excitatory and inhibitory inputs for positive or negative values for I, with a spike frequency response based on the level of that stimulus (class I excitability)[1]. A simple model capable of producing spikes is the quadratic integrate-and-fire model [19], simplified as

$$\dot{V} = I + V^2. \tag{4}$$

[1]For inputs that surpass the threshold, the spiking frequency grows continuously with the applied intensity [18].

Now, the rate of change of the voltage depends on squared voltage, not linearly as before. This leads to interesting properties, such as the fact that $\dot{V} \geq I$ for $I > 0$, where the quadratic term makes the rate increase in a positive-feedback fashion that resembles the regenerative mechanism of the sodium current in charge of the spike upstroke. When the voltage reaches a peak level, its value is reset to an appropriate base value

$$\text{if } V \geq V_{peak}, \text{ then } V \leftarrow V_{reset}. \tag{5}$$

Without these conditional step, the voltage would tend to infinity (an exclusive nonlinear-system characteristic known as finite-time escape) which, of course, it has no biological plausibility [20]. An extension that explicitly incorporates the dynamics for persistent potassium-like currents, that could be modulating the membrane potential at rest, is the resonate-and-fire (RF) or generalized integrate-and-fire model (GIF) [21], [22]. With the addition of a state variable that represents the magnitude of that current, the two-dimensional system can be described as

$$\begin{aligned} C\dot{V} &= I - g_{leak}(V - E_{leak}) - W, \\ \dot{W} &= (V - V_{1/2})/k - W, \end{aligned} \tag{6}$$

with $V_{1/2}$ and k a shifting and a scaling constant parameters, respectively. Some important aspects of this model include the ability to show sub-threshold damped oscillations (although it cannot produce sustained oscillations), and a preferential frequency of the input, which resonates with the neuron own sub-threshold oscillation, encoding spectral content more than strength or input level, and also, a relatively narrow frequency response (class II excitability)[2]. The Adaptive exponential integrate-and-fire model (AdEx) [24],

$$\begin{aligned} C\dot{V} = I &- g_{leak}(V - E_{leak}) + \\ &g_{leak}\Delta_T \exp(\frac{V - V_{thresh}}{\Delta_T}) - w \\ \tau_w \dot{w} &= a(V - E_{leak}) - w, \end{aligned} \tag{7}$$

with Δ_T, the slope factor, and V_{thresh}, the threshold potential, introduces an adaptation variable w, an exponential voltage dependence, and a reset condition if $t = t^{fire}$ then $V \leftarrow V_{reset}$, that provides a plethora of electrophysiological features such as spike-frequency adaptation, regular/fast spiking, phasic spiking, and damped oscillations to name a few.

Perhaps the most widespread model of a neuron is the Hodgkin Huxley model (HH) [1]. Interestingly, a feedback control strategy, the now ubiquitous voltage-clamp [25], [26], was what enabled Hodgkin and Huxley to record, study and later isolate and identify the ionic currents without fluctuations on membrane potential. In this approach, by maintaining a constant voltage (variable to be controlled) at different desired levels (input or reference), they ruled out the capacitive currents across the membrane, since $\dot{V} = 0$, and were able to investigate the voltage sensibility and kinetics of activation/inactivation of ion channels. Introducing auxiliary gating variables n, m, and h, which were fitted to experimental data (and later associated with open channel probabilities, since these mechanisms were not directly observed at that moment), they derived the equations:

$$\begin{aligned} C_M\dot{V} = I &- g_K n^4(V - V_K) - \\ &g_{Na} m^3 h(V - V_{Na}) - g_l(V - V_l), \\ \dot{n} &= (n_\infty(V) - n)/\tau_n(V), \\ \dot{m} &= (m_\infty(V) - m)/\tau_m(V), \\ \dot{h} &= (h_\infty(V) - h)/\tau_h(V). \end{aligned} \tag{8}$$

with, τ_i, the variable time constants, n_∞, m_∞, the steady-state (voltage dependent, or sensitive) activation functions, and h_∞ the inactivation function. In this formulation, the total current I is described on a space-clamped preparation, i.e. the potential does not propagate as only depends on time. To describe the action potential propagation along axons, the axial currents along the membrane need to be accounted for, which, using cable theory [27], leads to the partial differential equation

$$C\frac{\partial V}{\partial t} = \frac{r}{2R}\frac{\partial^2 V}{\partial x^2} + I - I_K - I_{Na} - I_{leak}, \tag{9}$$

where r is, usually, the radius of the axon, R is the resistance of the axon medium, and x is the position along the nerve fiber. This conductance-based, biologically detailed formalism can be extended to include another ion species as well as external current sources, and predicts the principal aspects of spike generation [28]. The high interpretability and detail achieved comes at the expense of complexity: the system (8) evolves in a four-dimensional state space described by a set of coupled nonlinear differential equations. Firstly, several phase-portrait based graphical analyses for up to three-dimensional systems are no longer available (typically carried out in offline computing) [19], [29] , and secondly, the increased computational cost can hinder real-time, large-scale simulations in a closed-loop application (online computing) [30]. These drawbacks are addressed in many simplified, reduced-order models [31]. By focusing on a voltage-like variable and a recovery variable, and based on the relaxation oscillator proposed by Van der Pol [32], FitzHugh was able to capture the basic excitability properties of neurons with only two state variables:

$$\begin{aligned} \dot{V} &= I - V(a - V)(V - 1) - W, \\ \dot{W} &= bV - cW, \end{aligned} \tag{10}$$

where I is the injected current, the constant a shapes the cubic V nullcline, i.e. the trajectory where $\dot{V} = 0$, shifting the system equilibrium points, and the non-negative constants b and c control the kinetics of the recovery variable W. The solutions of this planar system, known as FitzHugh-Nagumo (FHN) model [33], [34], considered in the complex network syncronisation control problem presented in [35], can be analyzed with geometrical tools, which give an insight on biologically observed phenomena such as spike accommodation, anodal break excitation, and the absence of a fixed threshold. The Morris-Lecar model [36], also makes use of a recovery variable and assumes the existence of two non-

[2]Neurons of this type present a discontinuous F-I curve, i.e. they cannot fire at arbitrarily slow rates, and when they fire, the frequency is relatively insensitive to increases in the applied stimulus intensity [23].

inactivating voltage-sensitive conductances

$$\begin{aligned} C\dot{V} &= -g_{Ca}m(V-E_{Ca}) - g_K n(V-E_K) - \\ & g_{leak}(V-V_{leak}) + I, \\ \dot{m} &= (m_\infty(V)-m)/T_m(V), \\ \dot{n} &= (n_\infty(V)-n)/T_n(V), \end{aligned} \tag{11}$$

where the variables m and n are analogous to the same parameters in the HH model, and symbolize the instantaneous fractions of open ion channels. The reduced order compared to the HH model, and the feasibility of fitting parameters based on measurements, allows the description of several excitable systems (e.g. muscle fibers) as well as networks of neural oscillators [37]. Based upon the FHN model, Hindmarsh and Rose [38] introduced a model of neuronal activity composed of three coupled differential equations

$$\begin{aligned} \dot{x} &= y - ax^3 + bx^2 - z + I, \\ \dot{y} &= c - dx^2 - y, \\ \dot{z} &= r(s(x-x_R) - z), \end{aligned} \tag{12}$$

where I is the external current, x is the membrane potential and x_R its resting value, y a recovery current variable used to model the spiking behavior, and z the adaptation variable used to induce bursting. The constants a, b, c, d are tuned to match observed trajectories, and the constants r and s are used to control the number of bursts. These equations are used, primarily, as a reference model for spiking-bursting behavior of membrane potential, since the addition of the third state equation allows for a number of dynamic behaviors (including chaotic dynamics) that has been observed experimentally.

Most of the current dynamical models used today in single-cell simulations and experiments are based upon, or share similarities with the models mentioned above. Some particular features are key for a closed-loop approach. When the objective is to modulate voltage, for example, not only an appropriate stimulus is needed (by potential means described in subsection IV-B), but also a setup capable of measure or estimate that signal of interest (common approaches are summarized in subsection IV-A). The model, if correctly fitted, will predict the preparation change in response to an exogenous input.

Another application of single-cell models is to use them as building blocks of neural circuits. From the simple case of two coupled neurons to the construction of much more complex networks, these models can be interconnected aiming to describe population dynamics. The large-scale end of the approach, requires extremely detailed physiological knowledge in terms of parameters and connections, and also, high computing power. The models presented in Section III could, in an eventual closed-loop application, circumvent the mentioned drawbacks.

Table I, shows a selection of modeled regions/applications where the presented models were used.

Model	Applications/Used to study	Refs	Eqn.
Leaky integrate-and-fire	Excitation of sciatic nerve on the frog.	[16]	(1)
	Firing in *Limulus polyphemus* visual cell.	[39]	
	Optimal control of spike timing.	[40]	
	Plasticity in closed-loop stimulated networks.	[13]	
Quadratic integrate-and-fire	Neurons with low firing rates.	[41]	(4)
Resonate-and-fire	Cells of rat entorhinal cortex (layer II and III).	[42]	(6)
Adaptive exponential integrate-and-fire	Layer V pyramidal neurons of rat neocortex.	[43]	(7)
	Inter-neuronal gamma oscillations in hippocampus.	[44]	
	Anesthesia-like slow waves in spiking networks.	[45]	
Hodgkin-Huxley	Squid giant axon.	[1]	(8)
	Control of repetitive firing in bistable modes.	[46]	
	Proportional control of spiking dynamics.	[47]	
FitzHugh-Nagumo	Cardiac cells action potentials and abnormal myocardial activity.	[48]	(10)
	Robust control of neuronal chaotic behaviour with electrical stimulation.	[49]	
	Predictive-control suppression of periodic regimes in pathological heart and brain conditions.	[12]	
Morris-Lecar	Barnacle giant muscle fiber.	[36]	(11)
	Open- and closed-loop control of firing rate.	[50]	
	Developing starburst amacrine cells.	[51]	
Hindmarsh-Rose	Control of synchronicity in coupled neurons, via 1) delayed feedback and 2) sliding mode control.	[52], [53]	(12)
	Control of neuronal behavior by Takagi-Sugeno fuzzy observer-controller.	[54]	

TABLE I
SUMMARY OF COMMONLY USED SINGLE NEURON, OR COUPLED NEURON MODELS.

III. POPULATION MODELS

The understanding of the collective properties of neurons, in terms of functional interaction, is of paramount importance in neuroscience. For instance, the firing pattern of a given population can offer more physiological insight of the processes involved than the time-course of a single neuron voltage [55]. This properties can arise, or *emerge*[3], from models of populations or ensembles of single neuron models. Whereas the interconnection of individual neuron models is possible, the problem addressed by connectomics [56], the computational cost associated can be huge and thus not suitable for closed-loop applications. In general terms, the population models

[3]Emergence occurs when a system shows properties its constituent parts do not have on their own, this characteristics appear only as a collective behavior.

aim to describe statistical properties of dynamical states that characterize the whole ensemble, such as mean firing rate or mean soma membrane potential. Typically, this description is circumscribed to a physical region in which, the function of the neurons is assumed almost identical. In this sense, the principal appeal/advantage and possible weakness/disadvantage of the approach, lies on the same feature, the volume simplification. This way of characterize the dynamics, from a many-body problem (many individual neurons whose correlations are neglected) to a one-body problem (averaged effective interaction) is within the framework of mean-field techniques. The strategy is analogous to apply statistical thermodynamics to link the complex Brownian motion of particles to a mean macroscopic description of a gas volume [57].

Within the homogeneous mass under study, different types and proportions of neurons can be considered. Most significant distinction is the excitatory and inhibitory type. The resulting interaction then, modeled as dynamical *activity*, could be interpreted as: rate of change of the proportion of active (i.e. spiking) cells, mean firing rate, or even a voltage-like activity depending on the particular model and formulation.

Regarding the feasibility of driving a population with a external stimulus (which it will ultimately depend on the effective internal coupling of the ensemble), the prospect of regulate its activity enables exploratory and treatmental approaches. Moreover, this population models can be used in conjunction with widespread non-invasive neuroimaging techniques such as electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) [58], [59].

A. Neural mass models

The term neural mass refers to an ensemble of neurons belonging to a certain class, depending on function, location and/or morphology. Describing the dynamical behavior of the whole with properties such as mean firing rate, for example, comprises a mesoscopic modeling scheme focused on the reproduction of measurable signals or levels. In this time-series description, there is no need for large-scale simulations of complex individual models in the network. The first attempt to characterize population activity is probably attributed to Beurle [60], who proposed a simple dynamical model for a mass of neurons. Although, in that formulation, only excitatory neurons having a fixed firing threshold were assumed, several concepts introduced were pivotal for the later development of more realistic models, like mean integrated inputs and synaptic feedback and delays.

In most neural mass models, the system is characterized by activity variables or rates, i.e. the proportion of neurons activated per unit time in certain compact tissue volume. The inputs of the system mimic the synaptic inputs of the individual neurons that conform the ensemble. Additionally, a nonlinear function models the input-output relationship, analogously to a gain function (FI curve) of single neurons.

The first mean-field theory of neural activity is attributed to Wilson and Cowan [61]. They derived coupled nonlinear differential equations for spatially localized excitatory an inhibitory sub-populations. The local redundancy supports that similar properties and responses can be found in small volumes. Applying a process of *coarse-graining* [62], which consisted in replacing the time integrals in the original formulation by moving time averages led to the system

$$\begin{aligned} \tau_e \dot{a_E} &= -a_E + (1 - r_E a_E) f_E\{c_1 a_E - c_2 a_I + P\}, \\ \tau_i \dot{a_I} &= -a_I + (1 - r_I a_I) f_I\{c_2 a_E - c_3 a_I + Q\}. \end{aligned} \tag{13}$$

Each subpopulation receives inputs from the other, weighted by coefficients $c_i,\ i \in \{1, 2, 3, 4\}$ (the average number of synapses per cell of the E-E, I-E, E-I, or I-I type respectively), that drive the rate of change of each activity, $\dot{a_E}$ and $\dot{a_I}$. The response function $f\{\cdot\}$ is typically sigmoidal and implies monotonic growth of activation given sufficient stimulus level. In fact, this smooth approximation can be seen as the superposition of many Heaviside step functions for individual neurons whose firing thresholds distribute normally [63]. The term $1 - r_i a_i$ denotes the proportion of cells that can fire since it is used to account for the spiking refractory period (with r_i, a constant that results from the averaging), and P, Q are external current-like inputs. The model exhibits properties like multistability and hysteresis, which could be used to address memory [64].

The Jansen-Rit model [65], [66], introduces a third class of neurons, and thus it comprises three different coupled sub-populations. One represents the pyramidal projection neurons and the other two, the excitatory and inhibitory interneurons, functioning in a feedback-loop fashion. For each sub-population, mean-voltage and mean firing rate states variables are used. Also, there are two operators that relate these relevant variables: a rate-to-potential, second-order ordinary differential operator, which maps the mean input firing rate into the mean membrane potential, and a potential-to-rate nonlinear function which transforms this potential into output firing rate. Thus, the system is described by

$$\begin{aligned} \ddot{x_0} &= Aaf\{x_1 - x_2\} - 2a\dot{x_0} - a^2 x_0, \\ \ddot{x_1} &= Aa(P + C_2 f\{C1 x_0\}) - 2a\dot{x_1} - a^2 x_1, \\ \ddot{x_2} &= BbC_4 f\{C_3 x_0\} - 2b\dot{x_2} - b^2 x_2, \end{aligned} \tag{14}$$

or alternatively, as the equivalent six-dimensional first-order system

$$\begin{aligned} \dot{x_0} &= x_3, \quad \dot{x_1} = x_4, \quad \dot{x_2} = x_5, \\ \dot{x_3} &= Aaf\{x_1 - x_2\} - 2ax_3 - a^2 x_0, \\ \dot{x_4} &= Aa(P + C_2 f\{C1 x_0\}) - 2ax_4 - a^2 x_1, \\ \dot{x_5} &= BbC_4 f\{C_3 x_0\} - 2bx_5 - b^2 x_2. \end{aligned} \tag{15}$$

The x_0, x_1 and x_2 variables describe the mean PSPs of each neuron sub-population. The input P could be used to model external stimulus to the system, -or inputs from other active neural populations-, constants A and B determine the maximum amplitude of the EPSP and IPSP, respectively, and constants a, b are used to account for delays associated with the passive membrane and dendritic network. The constant weights $C_i,\ i \in \{1, 2, 3, 4\}$ represent the number of synapses between interneurons and the neurons of the cortical column.

The nonlinear function is usually sigmoid,

$$f\{x\} = \frac{\nu_{max}}{1 + \exp(r(v_0 - x))}, \tag{16}$$

where ν_{max} is the maximum firing rate, and v_0, r, denote the input for which half of the maximum firing rate is reached, and the slope of the sigmoid, respectively. This model can reproduce several dynamic behaviors including, constant outputs, harmonic oscillations and non-harmonic oscillations of large amplitude. Also, is capable of describe multistability and quasiperiodic responses depending on amplitude and frequency of the input [67].

B. Neural field models

Since the neurons are distributed in brain tissue, the population activity unfolds in a spatial domain. Neural field models are a spatiotempral description of this evolution under the assumption of certain level of coarse graining, i.e. the ensemble is characterized by the mean dynamics and the fluctuations around the mean are neglected. The added complexity of the spatial variable to the already complex behavior of a neural mass, naturally calls for a simplification scheme, where, hopefully, the relevant dynamics are retained by arguments of self-consistency and redundancy. The resulting models of interconnected masses or populations avoid the treatment of individual neurons and usually take the form of integro-differential equations. Moreover, as pointed out by Sholl [68] and Beurle [60], the description of single neurons probably cannot encompass higher-level phenomena such as sensory integration, memory and learning.

After Griffith established the effect of inhibition [69] and proposed his reaction-diffusion equation for the population activity [70], [71], several models followed. Wilson and Cowan introduced their dynamical model for two-dimensional sheets of tissue [72], under the assumptions of homogeneous distribution of excitatory/inhibitory type neurons with lateral connections. As in their neural mass model, the redundancy and dense interconnection of neurons supports the description of the ensemble net activity. The system is described by the partial-differential equations

$$\begin{aligned}\tau_e \frac{\partial a_E}{\partial t}(x,t) &= -a_E(x,t) + \\ &(1 - r_E a_E(x,t)) f_E\{w_{EE} * a_E - w_{IE} * a_I + P\}, \\ \tau_i \frac{\partial a_I}{\partial t}(x,t) &= -a_I(x,t) + \\ &(1 - r_I a_I(x,t)) f_I\{w_{EI} * a_E - w_{II} * a_I + Q\},\end{aligned} \tag{17}$$

where $w_{jk} * a_k$ denotes a convolution given by the operator

$$w_{jk} * a_k = \int_D w_{jk}(x - y) a_k(y,t) dy \tag{18}$$

that is the net drive to the sub-population j at location x originated from sub-population k, and D is the physical domain where the activity evolves. This operation models the spatially distributed synaptic connections, and the weighting functions (which depend only on the distance $x-y$) are usually Gaussians

$$w_{jk}(x - y) = \alpha_{jk} \exp(\frac{-(x-y)^2}{\sigma_k^2}), \tag{19}$$

where α_{jk}, σ_k^2 denote the amplitude and the area of influence of the kernel, repectively, accounting for the distance-decaying sub-population impact.

The Amari model [73], [74], is a neural field description that includes excitatory and inhibitory effects, similarly to the Wilson-Cowan model, played by the corresponding-type neurons and coupled with typical cortical connectivities. The single-layer, one-dimensional form, i.e. the average membrane potential of the neurons located at position x is described by

$$\frac{\partial u}{\partial t}(x,t) = -u(x,t) + \int_D w(x,y) f\{u(y)\} dy + P(x,t), \tag{20}$$

where $f\{\cdot\}$ is the activation function mapping voltage to firing rate, and w is the connectivity function weighting the strength of propagation from point y to point x, in the domain D and P is an exogenous input. This model typically considers local excitation and distal (lateral) inhibition, as weighted by the "Mexican hat" connectivity function, i.e.

$$w(x - y) = K e^{-k(x-y)^2} - M e^{-m(x-y)^2}, \tag{21}$$

where the positive constants K, k, M, m determine the shape of the kernel. Using a two-layer distinct excitatory/inhibitory version of the model, oscillatory and traveling wave patterns can be predicted.

A model that is used to describe, particularly, the dynamics of cortical macrocolumns is the Liley model [63]. Inhibitory and excitatory sub-populations are considered, being locally coupled with each other in feedforward and feedback local connections. Then, macrocolumn to macrocolumn interaction is exclusively excitatory and is carried out by long-range axonal fibers. The mean-field description is obtained from spatially-averaged neurons (individual conductance-based models)

$$\tau_k \frac{\partial h_k}{\partial t}(\mathbf{x},t) = h_k^r - h_k(\mathbf{x},t) + \sum_{l=e,i} \frac{h_{lk}^{eq} - h_k(\mathbf{x},t)}{|h_{lk}^{eq} - h_k^r|} I_{lk}(\mathbf{x},t), \tag{22}$$

where $\mathbf{x}$ is the position on the sheet of tissue, $l, k \in \{e, i\}$ denote excitatory and inhibitory sub-populations, respectively, the first subscript being the presynaptic source and the second the postsynaptic target. The parameter h_k^r is the mean resting membrane potential, and I_{lk}, the inputs to the system weighted by the fraction shown. These postsynaptic potentials have dynamics described by damped oscillators driven by the mean firing rate of input excitatory/inhibitory axonal pulses A_{lk}, according to

$$\begin{aligned}\left(\frac{1}{\gamma_{lk}} \frac{\partial}{\partial t} + 1\right)^2 I_{lk} &= \frac{e \Gamma_{lk}}{\gamma_{lk}} A_{lk}, \\ A_{ek} &= N_{ek}^{\beta} S_e h_e + N_{ek}^{\alpha} \Phi_{ek} + p_{ek}, \\ A_{ik} &= N_{ik}^{\beta} S_i h_i.\end{aligned} \tag{23}$$

For excitatory postsynaptic conductances, three sources of axonal pulses can be present, local (S_e), cortico-cortical (Φ_{ek}),

Model	Applications/Used to study	Refs	Eqn.
Wilson-Cowan	Kalman-filter based control of cortical neuronal activity.	[75]	(13)
	Firing rate regulation of subthalamo-pallidal loop via P controller.	[76]	
	Disruption of pathological beta band oscillations.	[77]	
	Deep brain stimulation treatment for essential tremor.	[78]	
Spatially distributed Wilson-Cowan	Closed-loop induced Gamma rhythm in visual cortex.	[79]	(17)
Jansen-Rit	Epileptic activity suppression via closed-loop PI controller.	[80]	(14)
	Validation of model predictive control applied for seizure suppression.	[81]	
Amari	Language processing and brain waves.	[82]	(20)
	Dynamic cognitive modeling.	[83]	
Liley	Description of electroencephalogram patterns associated with anesthesia.	[84]	(22)
	Closed-loop suppression of epileptiform activity via optogenetic stimulation.	[85]	

TABLE II
SUMMARY OF COMMONLY USED NEURAL MASS/FIELD MODELS.

and sub-cortical (p_{ek}). In the inhibitory case, only local source is considered (S_i). The firing rates S_e and S_i are a function of the the potential h_k, and are usually ruled by sigmoidal activation. N_{ek}^{α} and N_{lk}^{β} represent the number of synaptic connections, and the propagation of the pulses is described by the damped wave equation

$$\left[\left(\frac{1}{v_{ek}}\frac{\partial}{\partial t}+\frac{1}{\lambda_{ek}}\right)^2-\nabla^2\right]\Phi_{ek}=\frac{1}{\lambda_{ek}^2}Se\{h_e\}. \qquad (24)$$

Table II summarizes several studies of neurobiological regions and closed-loop applications, including their corresponding equations.

IV. CONTROL FRAMEWORK

The main focus of this work, besides presenting a state-of-the-art of different single and population neuron models, is to insert these models in the structure of identification and automatic control. The purpose of these techniques is to provide significant and relevant dynamics to produce a control-oriented model (COM). The latter can be tuned to the particular neurobiological system under study by means of identification and (in)validation[4]. Next, a controller (see Fig. 2, bottom) could be designed and analysed based on this COM, to control the behaviour of the neurobiological system. The complete procedure, from the theoretical model to the experimental control implementation is depicted in Fig. 3. Particularly, in Fig. 3 (d), a description of the system components is presented from a general perspective. A control system implementation in a neurobiological system comprises: 1) biological unit; 2) acquisition interface; 3) acquisition conditioning; 4) computer hardware; 5) actuation conditioning; and 6) actuation interface over the biological unit. The biological unit (1), defines the neurobiological system under study, such as a single neuron, a circuit of interconnected neurons, or even a complex neurobiological structure as, for example, the hippocampus. The acquisition interface (2) is essentially composed of the acquisition technology (sensors). The acquisition conditioning (3) is simply the signal adaptation stage, to prepare the measured signal to be acquired by the computer hardware. The computer hardware (4) is where the control signal is computed, based on the particular control methodology. Once the control signal is computed, it is conditioned in the actuation conditioning stage (5). Finally, the control actuation is applied to the system using the actuation interface (6). By way of example, in [87] a closed-loop system applied to epilepsy treatment in humans is described, detailing these complete required instrumentation.

The closed-loop methodology has been used in recent years ([46], [47], [49], [50], [52]–[54]) in particular cases in a very basic way. The purpose of this work is to provide a structured framework in which to apply identification and control to modify the behavior of brain regions in order to modulate its activity in a principled way.This has an enormous potential field of applications in the treatment of several brain diseases, e.g. Parkinson, epilepsy, and essential tremor. Nonetheless, two fundamental aspects must be kept in mind. Firstly, it must be noted that there are legal regulations and formalities that strongly distinguish applications allowed for animals or humans, regardless of whether they are research or clinical applications. For animal applications, in compliance with legal formalities and ethical protocols, there are virtually no limits for allowed experimental interventions, which includes sensing or actuation techniques. As a matter of fact, all the allowed experimental interventions already tested in humans (clinical stage), have been previously studied and tested using animal models (pre-clinical stage). Thus, general human applications have been first thoroughly studied using animal models, even in those cases of testing trials using volunteer human subjects. By way of example, the cases in [87] and [88] can be cited as cases of human experimental interventions. Secondly, when considering research applications using animal models, where some cognitive hypotheses, e.g. memory mechanisms, are subjected to validation, the conclusions can be relatively straightforwardly extrapolated to humans. This is the case, for example, of primitive structures, such as the hippocampus. Techniques such as optogenetics, which application is currently forbidden in humans mainly for ethical reasons, are generally used in research, for validating biological hypotheses. By way of example, the results presented in [89] can be mentioned, where cognitive hypotheses are studied, which

[4]The term (in)validation refers to the fact that a theory, or in this case a model representing a physical or biological phenomena, can only be invalidated or falsified with certainty, but not validated [86], because future data could prove the theory or model is incomplete.

results can be extrapolated to humans, even though still under analysis.

In the following subsections a brief presentation is made of the main issues to be considered when applying this framework to a neurobiological system: sensors, actuators, model identification (to compute model parameters), (in)-validation (to compute uncertainty bounds), and model-based control analysis and design.

A. Sensors

Being able to perform appropriate measurements, and therefore knowing the status of the system, is the first step to modify its behavior in a closed-loop approach. The devices in charge of providing information, in the form of representative signals related to the system, are the sensors [90]. Several sensor characteristics are cross-cutting the areas of application, and they need to be accounted for in the process of its selection, and the closed-loop design.

Regardless the setting, any sensor is subject to measurement noise. Some situations can be more or less hampering, because the signal level is too low (e.g. voltage measurement of 10–100 μV in EEG [91]) and/or the sensor is receiving mixed/filtered signals from the environment [92]. In any case, the noise introduced will impose constraints on the feedback design. Moreover, the sensors themselves often have non-negligible dynamics, i.e. they do not respond instantaneously to the change of the variable of interest. For example, there are optical voltage sensors in the form of dyes that offer a slow response (e.g. merocyanine 540), making them inadequate to track fast transients as action potentials [93].

So far, disturbances and lags pertaining to sensors were described. In some cases, the state of the system cannot be directly determined. Whether these states do not have a physical meaning, and therefore are impossible to measure, or the number/cost of the associated sensors are prohibitive, they can be inferred from the knowledge of the system inputs and outputs. This process is encompassed in the concept of *observability* [94], and is carried out with estimators known as observers (see Fig. 2, bottom). An ubiquitous state observer, that has been widely used in many engineering applications, is the Kalman filter [95], [96]. Briefly, model output and available measurement are recursively combined in a prediction-correction manner, in order to obtain an estimate of the states that result more accurate than any of the approaches alone (since they have associated model uncertainty and measurement noise, respectively). Several extensions have been developed specifically for application to nonlinear systems, most notably, the unscented Kalman filter [97] can achieve superior performance and efficiency in some circumstances [98]. Whether it is a single spiking neuron dynamics or a network-of-neurons state to be observed, the estimation problem is nontrivial. Some progress was made in this area and its potential application to closed-loop approaches is promising. In [99], the authors reconstruct the dynamics of the HH model gating variables describing a pyramidal cell, using a Kalman filter and membrane potential measurements. In [100], an ensemble Kalman filter [101] is applied to an oscillator network of neurons described by a grid of Wilson-Cowan equations, in order to track its states, using voltage-sensitive dye measurements.

Regarding sensors, it has to be kept in mind that the goal is to measure neuronal activity. Neurons, express in a constantly interacting loop of transmembrane potential and metabolic changes. Both electrical and metabolic changes could be measured and taken as proxies for neuronal activity but the nature of the sensors required in each case is different. Overall, electrical signals could be measured by electrodes and depending on the type of measurements different considerations should be accounted for. Metabolic changes could be expressed as changes in intracelullar concentration of calcium, glucose consumption or blood-flow. Each of them require a different type of sensor. Because of technical simplicity, electrical sensors are the most implemented when it comes to measure neuronal activity in most clinical or interfacing applications. A detailed discussion of all these aspects is described next.

- **Electroencephalography** (EEG). EEG measures electric activity mainly evoked during synaptic excitation. The procedure is carried out with noninvasive electrodes placed on the scalp and can be used for online analysis and stimulus triggering. Since the signals have to traverse through several layers including skull and scalp, they are heavy filtered versions at electrode level showing weak amplitude, and thus, low signal-to-noise ratio (SNR). The method offers a temporal resolution in the 1-5ms range, and a spatial resolution constrained in part by inter-electrode distances of $\approx$2.5cm, in the case of over 100 multichannel configurations (multiple outputs). Due to these characteristics, EEG is probably the most common sensing technique applied to humans. EEG is being used in several closed-loop applications from sleep waves reinforcement [102], motion decoding [103], to epilepsy treatment [104].
- **Electrocorticography** (ECoG). ECoG is an invasive technique, with consequently health risk, for measuring cortical activity that employs electrodes placed on the surface of the brain. Compared to EEG, the resulting signal is more robust to artifacts whereas spatio-temporal resolution is improved. Current high-density array configurations present shorter inter-electrode distances as 700 μm. An ECoG-reading based closed-loop approach is presented in [105], and a biomedical setup in [106].
- **Intracortical Neuron Recording**. This technique measures electrical activity at gray-matter level of the brain. Performed by means of implanted microelectrode arrays, it is being capable of the acquisition of three types of signals: discrete-like single-unit activity (SUA) and multi-unit activity (MUA), by high-pass filtering (>300Hz) the responses of nearby single of multiple neurons respectively. The third type is LFPs, which are electric potentials recorded in the extracellular space in tissue from relatively localized populations of neurons, extracted by low-pass filtering (<300Hz) the measurement. The spatial resolution ranges from $\approx$0.1mm to $\approx$0.5mm in MUA and

LFP modalities, respectively. ECoG is being currently used in both animal and human applications. A closed-loop stimulation setup based on recorded LFP beta-band power, used to Parkinson disease treatment, is presented in [107]. In [108], LFP readings are used to predict and terminate seizures by delayed stimulation.

- **Magnetoencephalography** (MEG). MEG register the magnetic activity of the brain. The intracellular currents flowing through neural circuits induce magnetic fields that are collected in a noninvasive way with external probes. The signals measured via MEG are less prone to distortion (although severely affected by electromagnetic interference) and tissue filtering compared to EEG acquisition, and offer a better spatial resolution (≈5mm). This technique is commonly used in humans, and to a lesser extent, in animal models. In [109], nonparalyzed subjects MEG readings are used to control a neuroprosthetic hand, as a brain–machine interface assessment approach.
- **Functional Magnetic Resonance Imaging** (fMRI). fMRI is based on electromagnetic detection of changes in localized blood volumes in the brain, since the blood flow and its oxygenation varies its levels upon neural activation. This correlation is captured by the blood oxygen level dependent (BOLD) signal, which metabolic effect introduces a delay of several seconds, and thus must be accounted for in an eventual closed-loop design [110]. fMRI is being currently used in humans. In [111], fMRI signals are used as neurofeedback in order to correct, via visual stimulus, attentional-related biases in depressed patients. A pilot study for fMRI-based closed-loop modulation, in conjunction to transcranial electrical stimulation (tES) is presented in [112].
- **Functional near-infrared spectroscopy** (fNIRS). Similarly to fMRI, detecting changes in blood properties can be performed by the noninvasive method fNIRS. Applying light in the near-infrared portion of the spectrum, and based on the differential absorption of two blood chromophores, relative concentrations can be determined and thus, indirectly, fluctuations in brain activity. This technique is being currently used in humans, e.g., in [113], a device capable of perform fNIRS for imaging in conjuction to tES for closed-loop stimulation is presented.

Some optical sensing approaches are a growing alternative for neural activity imaging, and can overcome some of its electrical counterparts limitations in terms of measurement specificity, since they can be chemically or genetically tuned to a specific target population. Common optogenetic-based reporters include

- **Genetically encoded calcium indicators** (GECIs). GECIs allow the measurement of the activity of large ensembles of neurons indirectly, using the intracellular calcium concentration as indicator. They can be used for *in vivo* imaging with a high SNR, at varying resolutions, e.g. for sub-cellular scale monitoring and to track potential changes slower than spikes, due to its relatively slow kinetics. This technique has been applied in animal models only. In [114], the authors propose a closed-loop setup for neural activity modulation based on GECI (GCaMP6) real-time readout.
- **Genetically encoded voltage indicators** (GEVIs). GEVIs are capable of report spikes and sub-threshold fluctuations of membrane potential. Compared to electrophysiology methods, GEVIs still offer lower temporal resolution but improved spatial resolution. Due to sensor progressive degrading process known as photobleaching, the technique requires calibration and correction for repetitive use. This technique has been applied in animal models only. In [115], a closed-loop real-time protocol for cell membrane potential control, based on GEVI (ArcLight Q239) is presented.

In any case, the choosing of the measurement/estimation method is usually multifaceted. Single electrode methods may work well for single neuron or neural masses whereas multi-electrode arrays are more appropriate for space-distributed neural field models. In both cases, the relevant state is assumed to be the electric potential. Optical sensors, e.g. GECIs, can provide information about the neural activity, whether lumped or distributed, as described by population rate in the selected models of Section III.

B. Actuators

The term actuator refers to instruments or devices that have the ability to affect the system in order to drive the process from a current state to a desired state [116]. Before choosing or designing actuators, a goal must be stated in order to formulate proper control objectives. This typically includes what is to be achieved in the first place, what variables need to be controlled to that end, and what are the performance specifications (in terms of tracking error or speed for example). The assessment of actuators properties also have importance in the modeling process, as it will be highlighted below.

One potential source of limitation in a closed-loop performance can be attributed to the response of actuators. This response can be highly nonlinear, imposing constraints in the form of saturation limits on actuation level. For instance, a reference input too far from the actual state of the system can induce a *control signal* beyond the actuator handling capabilities. Moreover, the effect of a given actuator can only drive the system in one direction. For example, considering an excitatory light-activated channel as actuator, the manipulated input to the system can only depolarize the membrane, lacking the possibility of steering the potential toward hyperpolarized values.

Additionally, actuator dynamics can dominate over certain system features. Specifically, when a closed-loop application is said to operate on a *millisecond scale*, for example, it is implicitly assumed that the actuators respond much faster than that. This temporal scale characteristic will impose upper bounds to the achievable closed-loop bandwidth. In any case, these characteristics can be addressed in the modeling stage, and/or be included in the closed-loop simulation framework.

A fundamental question that naturally arises from the stimulation of a system is whether or not we can steer it from a given

initial state, via a control input, to any location in the phase space. This issue can be formally analyzed (at least in LTI or linearized models) and it regards to the *controllability* of the system [90]. Of course, in biological applications, as in many others, some states have natural bounds usually imposed by physical interactions. For example, a neuron firing rate cannot surpass ≈1KHz, since it is constrained by a ≈1 ms absolute refractory period. So any desired state must be interpreted as any underlying-process abiding, reasonable state.

Any stimulation approach can be used in conjunction with others, including pharmacological, thermal, sensory, and electrical/optical technologies. The actuators resolution will impose practical constraints in terms of the minimum targeted area and fastest achievable response, operating at different spatial and temporal scales, respectively.

The most widespread methods of stimulation are based on electric or electromagnetic principles and date back as early as the 18th century with the famous Galvani experiments on frog nerves [117]. In most cases, the process involves electrode(s) delivering current or voltage stimulation at neuron or neural-network level, exciting directly or indirectly the cells membranes [118]. A large number of stimuli parameters can be modulated, e.g. amplitude, pulse width and frequency, signal waveform, etc. and thus, qualitatively and quantitatively different effects can be achieved [119].

One way of categorize the actuation techniques is in invasive or non-invasive procedures. Some non-invasive approaches, currently used in humans, include the following forms of stimulation:

- **Transcranial magnetic stimulation** (TMS). This technique delivers a magnetic pulse that induces stimulating currents in nerve cells in large regions. TMS has a limited depth influence (≈30mm) and a spatial resolution of ≈5mm. Some closed-loop applications can be found in [120] and [121].
- **Transcranial Direct Current Stimulation** (tDCS). With low intensity current flowing through electrodes placed over the head, two modes of operation are used anodal and cathodal, exciting or inhibiting neural activity respectively. A proof-of-concept of a potential implementation in humans is found in [122], and in [123], a closed-loop modulation for a network in the prefrontal cortex is presented.
- **Transcranial Alternating Current Stimulation** (tACS). Low frequency (<100 Hz) alternating current is applied, tuned to interact with cortical brain oscillations. A seizure triggered closed-loop application of tACS is presented in [87], and a potential setup for feedback control is detailed in [124]. For a comprehensive review of the method see [125].

Invasive techniques require physical contact between the device and the tissue, typically by microelectrode arrays (MEAs), that could serve also as recording devices, and through implantation, can be used in *in vivo* applications. The following modalities are currently applied to human patients:

- **Cortical electrical stimulation** (CES). Besides the use of cortical stimulation for mapping functions and pre-surgical aid, CES, or ICMS (intra-cortical micro stimulation), has also been used for therapeutic and neuromodulating applications. To this end, electrodes are implanted in epidural or subdural places inducing currents through the distribution of the electrical field, with horizontal fibers being excited preferentially by cathode tip and perpendicular fibers by the anode tip [126].
- **Deep brain stimulation** (DBS). Currently used as treatment for neurological disorders in humans including Parkinson's disease and dystonia, most DBS systems are operated in open-loop manner, with efficacy subjected to the choosing of stimulation settings in contrast to adaptive parameters. Usually comprising stimulation and recording electrodes, the actuation is achieved upon the generation of electrical fields in the surrounding contact areas, which could lead to the generation of spikes and/or neurotransmitter release [127].

Complimentary to electrical interfacing, optical methods have been actively developed and offer improved recording and stimulation capabilities. In optical actuation methods, the application of light modifies properties of a portion of the cell membrane in order to induce currents and therefore, shifting its potential. This approach is based on optogenetic techniques [128]. In its most common form, some alterations are introduced via genetic manipulation in targeted neurons that express light-sensitive channels (opsins) as mechanism for stimulation [129]. The high selectivity of the method should be highlighted, since it is capable to solely alter some sub-populations of neurons based on genetic parameters that correlate with specific cell type. Additionally, several sub-populations can be stimulated from different constructs, based on wavelength selectivity. Although, in more complex combinations, considerations must be taken in terms of their action spectra to avoid overlapping, similarly for the use of optical sensors and actuators simultaneously. Since the opsins photocycle dynamics could introduce restrictions to the loop, their model can be included in the general dynamical model used for simulations, in order to achieve more detailed descriptions [130]. A key feature of the technique is that the actuation can be performed bidirectionally, since there are certain classes of optogenetic tools that enable outward currents, promoting hyperpolarization. This property can drastically improve the performance of a closed-loop approach. On the other hand, the method introduces several challenges. In terms of the delivered input, the targeted tissue can scatter and absorb the applied light, broadening the bounds of the region of interest. Also, this region can exhibit differential protein expression that could result in heterogeneous activation, and more important, the density of the deployed actuators may not be high enough to effectively drive the entire network to a desired state. This controllability-related issue is explored in [131].

C. Model-based control

The analysis and design of closed-loop control systems can be performed in basically two different ways, according to the use or not of a model representing the dynamical system under study.

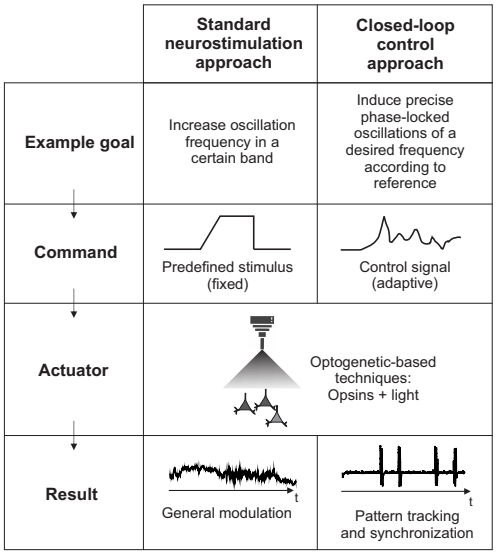

Fig. 1. Primary difference between pre-designed stimulus application, often an empirically derived fixed-level input, and a graded control signal, that is based on both system and controller, to achieve a desired behavior.

Among the model-less methods, the popular proportional (P), proportional-integral (PI) and proportional-integral-differential (PID) controllers are based usually on time responses of the system dynamics and implicitly assume a Linear Time-invariant (LTI) model that represents it. They come from a long tradition of control systems successfully used in industrial applications since the 1940's [132]. At the end of the 90's, other methods not based on models appeared, known as *unfalsified* controllers [133], [134], based on Popper's falsification theory [86]. The latter performs an online optimization and selection of the best controller based solely in real-time input/output data, but has still certain difficulties for practical applications [135] and exceeds the scope of this work.

On the other hand, the best use of the information relative to the system is achieved by model-based controllers. First, a model usually originated in the full mathematical description of the system or in an identification procedure, is obtained. The latter will be described in subsection IV-D. The full model, which is usually described by nonlinear differential equations, could be used as a simulator to test the performance of the closed-loop system. Next, a simplified version, defined as *control-oriented* model, is used to design a controller. This simplification is performed by a linearization of the full model in order to obtain LTI dynamics and possibly a model order reduction of the latter [136]. This model is used to design the controller which will be tested in conjunction with the simulator in closed-loop.

More recent developments in the area of control systems describes the dynamics as a *set* of models so that model uncertainty can be accommodated. This provides a more realistic description of the system and is known as robust control theory [137], [138]. In this framework, a more adequate procedure to compute the set of models is by means of (in)validation [139], possibly combined with identification. Here, a nominal model is obtained from *apriori* information or from an identification procedure, and the bounds on the sets of external perturbations and model uncertainty are minimized in order to reproduce the experimental data. Some guidelines to apply (in)validation to biological systems can be found in [140] and [141].

In addition, due to difficulties associated to complex systems and under-actuation, other approaches based on neural networks, could be useful in this stage, as discussed in [142] and [143], and also, modern strategies based on fuzzy logic and adaptive control as in [144] and [145].

In any case, the first step in order to obtain a model or set of models of the system dynamics is the location of their inputs and outputs (I/O). These will be excited and measured by the actuators and sensors, respectively, as described in the previous sections. The resulting model can be classified upon the definition of relevant I/O and, consequently, the possible control strategies can be elucidated. The simplest case is single-input single-output (SISO) configuration. Much of the existing control theory is developed for SISO systems. Intermediate complex setups include single-input multiple-output (SIMO) and multiple-input single-output (MISO) models. The most general type is multiple-input multiple-output (MIMO), which contemplates the multiple interfacing variables that the system might possess.

It should be highlighted that many applications, where biological phenomena such as habituation and plasticity are common, could benefit from adaptive automatic control, as opposed to pre-defined, manually-tuned input stimulus. In Fig. 1, a side by side comparison between a modulation problem and a more complex, control specific task in presented. In both cases the same interfacing is assumed, i.e. electrical sensing and optical actuation, in order to standardize the comparison. Here, the adaptation of the control signal, based on the system state (as predicted by the chosen model) and the controller (as resulting from the chosen control strategy) is key. In this way, this system input achieves the desired behavior of the system, in the example, a specific pattern tracking. Even though much of the pursued laboratory or clinical outcomes are empirically achieved by simplistic (On-Off control), brute force (stimulating at maximum level) or trial and error basis (tedious and expensive tests), the model-based paradigms are gaining ground, becoming a systematic implementable procedure [146].

This general control framework presents several advantages, particularly in these type of complex systems. Firstly, complicated dynamics can be modeled as nonlinear models. These can in turn be linearized at a particular working condition, or dynamically linearized, e.g. feedback linearization [147]–[149]. Very seldom, complex systems can be modelled as linear. Nevertheless in some cases they can be modeled as linear and time-variant, also known as Linear Parameter Varying (LPV) models [150]–[152]. Different methods are specifically suited for all these dynamics, which provides more flexibility to the control framework.

Secondly, uncertainty in the models and external perturbations can also be dealt with. Model uncertainty comes from unmodeled internal dynamics and/or parametric uncertainty. Nonlinear models can also be presented as a linear (possibly

time-varying) model with (bounded) model uncertainty. In addition, external unknown signals, unwanted and/or hard to describe inputs to the system can be considered. These can be quantified in terms of their energy bounds and/or bandwidth (spectral characterization). Bounded disturbances and model uncertainty can be dealt with by adding dynamical weights in the design and analysis of the closed-loop in the framework of robust control [137], [138]. For some fully known disturbances, e.g. power grid noise of fixed frequency, or DC biases, there is the possibility of total rejection in the closed-loop as another merit.

There exist, however, some difficulties that affect the control design. When dealing with complex neurobiological systems, the relatively high levels of both uncertainty and measurement noise hinder the achievement of performance levels. Other phenomena as qualitative changes in the system, for example, changes in the type of spiking behavior or the triggering of internal regulation processes could be impossible to address with a single controller, adding considerable complexity to the design.

Depending on the chosen control strategy, the design parameters will affect different aspects of the loop. For instance, for the tuneable parameters of the PI technique, as implemented in [153], the gain associated with the P term (proportional to the error signal) will drive the system to react more quickly at the expense of risking overshoot (unwanted drift from reference). The gain of the I term (that integrates the error) will tend to decrease the steady-state error, but introduces an unwanted effect of accumulating the error, slowing down the response, if a large change in the reference level takes place. Another well-established control technique is the Linear Quadratic Regulator (LQR), as implemented in [154]. In this approach, a state-feedback control law $u(t) = -Fx(t)$ is applied to the system. The F gain is such that a quadratic cost function is minimized. This optimal-control technique can be tuned with weighting matrices Q, R in the cost function, that penalize the state $x(t)$ and control energy $u(t)$ vectors, respectively.

Prior to experiments, the proposed control strategy must be validated in simulation. In this stage, multiple scenarios could be considered in order to strain the control design, and strengthen its robustness. For example, model parameters can suffer variations within certain range, according to the biological experiment that this model represents, i.e. these fluctuations must keep physical sense. Also, reasonable levels of noise, in compliance with the existing setup, should be added to the simulated measurements to check for specifications fulfilment. Once data is available, another way to test and verify the model is performed through a technique that is explained in next subsection, known as (in)validation.

In Fig. 2 (top), a potential closed-loop system is depicted, particularly, based on common electrical sensing (measurements stem from electrodes) and optical actuation methods (the actuators are the opsins in the cells along with the light-emitting equipment). In Fig. 2 (bottom), an in-depth detail of the control block is provided, depicting a standard combination of estimation and regulation, achieved by the observer and regulation blocks respectively. Intuitively, when the actual state is far from the desired behavior, the error

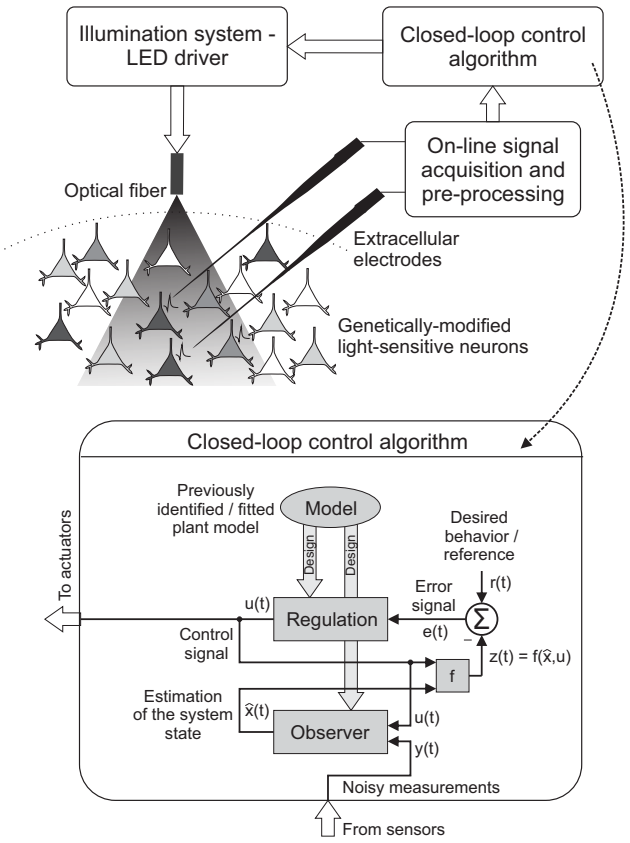

Fig. 2. **(top)** Typical setup, using electrical sensors and optical actuators, capable of varying closed-loop control strategies embodied in the algorithm block, i.e. the controller. **(bottom)** Detail of the control block. The chosen model is used to design a regulator and an observer, handling uncertainties and performance specifications in the process. The actual state of the system is obtained with the estimator, as indicated with the observer block, affected with a possibly nonlinear function $\boldsymbol{f}$, and then it is compared with the reference. The controller delivers a control signal accordingly to the difference, i.e. the error signal.

signal $e(t) = r(t) - z(t)$ increases, pushing the response of the controller $u(t)$, which in turn injects an appropriate command to the actuators, in pursuit of minimizing $e(t)$.

D. Identification and (in)-validation

The connection between the theory (model) and the application (experiment) should be performed by means of an identification and/or an (in)validation routines, when experimental data is available. Both procedures relate the model parameters and its uncertainty bounds with the actual experimental data.

Model identification is a procedure that computes the particular parameters of a mathematical model that fits the experimental data from the real-system dynamics. It consists in the optimization of a vector of n parameters $\mathbf{p} \in \mathbb{R}^n$ that minimizes the difference between the input/output experiment on the system provided by input/output vectors (u, y) and the mathematical model $g(\cdot, \cdot)$. The latter is the differential equation, which for simplicity has been represented as an LTI model in terms of its Laplace transform variable s, but can be straightforwardly generalised with a nonlinear description. The resulting (single) model equation is $g(\mathbf{p}_\star, s)$ for the optimal vector of n parameters $\mathbf{p}_\star$. The procedure is

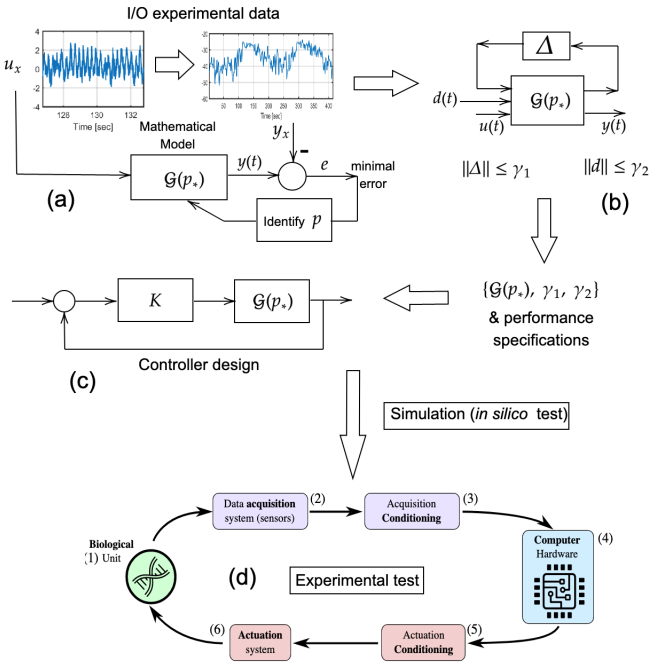

Fig. 3. **(a)** Model identification: $\mathbf{min_p}\ \|e\|$, **(b)** model set (in)validation: $\mathbf{min}\ \{\|d\|, \|\Delta\|\}$, **(c)** controller design, **(d)** experimental stage: potential implementation

depicted with a block-diagram in Fig. 3 (a), where the real system is symbolized as $\mathcal{G}$. A deeper treatment can be found in [155] and a more recent framework based on subspace theory is described in [156]. In [157] a general identification framework for closed-loop stimulation is described, and in [154], [158], subspace identification is used for learning input-output dynamical models.

(In)validation instead proposes a nominal model $G(s)$ of the system, that could originate in *apriori* information or an identification process, as the one described before. Next, it minimizes the two possible causes of differences between the model and the experimental set (u, y). These differences are due to (1) external disturbances d and (2) model uncertainty Δ. Therefore the method minimizes the bounds (γ_1, γ_2) so that the input/output data (u, y) fit the nominal model values within the sets $\|d\| < \gamma_1$ and $\|\Delta\| < \gamma_2$. The resulting set of models is the combination $\{F_u[G(s), \Delta], \|\Delta\| < \gamma_2\}$, where $F_u[a, b]$ is the lineal fractional combination between a and b. This has been illustrated in Fig. 3 (b) and a complete description can be found in [139].

Both processes can be used in conjunction, prior to the experimental application, first to fit the parameters of the chosen model, and latter to refine and account for model uncertainty for a closed-loop implementation.

V. Conclusions and Future Research

A comprehensive review of neurobiological models, from the automatic control perspective, is provided in this work for both, single neuron and populations. It is important to highlight that this work arises in a context where, even though to date the required technologies are available, throughout the literature most control strategies approaches applied to neurobiological systems are developed with no *a-priori* information on the dynamical behavior of the system. Within this framework, the main relevant models in the neurobiological field have been clearly presented and discussed in this study. Consequently, the existing opportunities related to the use of a theoretical framework for dynamical models of neurobiological systems have been presented, always indicating the available technological platform required to pursue this goal.

Recent actuation and sensing technologies, which are key components in any closed-loop control implementation, have been indicated here. Similarly, the connection between the models presented here with those technologies, has been precisely discussed. Thus, the feasibility of realistic model based control routines is shown. This work contributes in the path towards more efficient control analysis, design and implementation applied to neurobiological systems, exploring the existing technologies and models in the field.

As a future aspect to be taken into account, a "gas law" model should be considered rather than a combination of millions of particle models, when stated as a thermodynamic metaphor. In addition, model identification and model-data (in)validation could be applied, as well as numerical model order reduction techniques to produce a dynamically significant control-oriented model. Furthermore, control methodologies that consider uncertain nonlinear and time-varying systems could also be applied to these complex problems, e.g. neural networks, fuzzy logic, LPV control.

References

[1] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, pp. 500–544, 1952.

[2] N. L. Golding, W. L. Kath, and N. Spruston, "Dichotomy of action-potential backpropagation in ca1 pyramidal neuron dendrites," *Journal of neurophysiology*, vol. 86, no. 6, pp. 2998–3010, 2001.

[3] J. C. Crone, M. M. Vindiola, A. B. Yu, D. L. Boothe, D. Beeman, K. S. Oie, and P. J. Franaszczuk, "Enabling large-scale simulations with the genesis neuronal simulator," *Frontiers in neuroinformatics*, vol. 13, p. 69, 2019.

[4] A. Bolotsky, D. Butler, C. Dong, K. Gerace, N. R. Glavin, C. Muratore, J. A. Robinson, and A. Ebrahimi, "Two-dimensional materials in biosensing and healthcare: from in vitro diagnostics to optogenetics and beyond," *ACS nano*, vol. 13, no. 9, pp. 9781–9810, 2019.

[5] C. Carrasco-López, S. A. García-Echauri, T. Kichuk, and J. L. Avalos, "Optogenetics and biosensors set the stage for metabolic cybergenetics," *Current Opinion in Biotechnology*, vol. 65, pp. 296–309, 2020.

[6] K. Woeppel, Q. Yang, and X. T. Cui, "Recent advances in neural electrode–tissue interfaces," *Current opinion in biomedical engineering*, vol. 4, pp. 21–31, 2017.

[7] J. S. Choi, H. J. Lee, S. Rajaraman, and D.-H. Kim, "Recent advances in three-dimensional microelectrode array technologies for in vitro and in vivo cardiac and neuronal interfaces," *Biosensors and Bioelectronics*, vol. 171, p. 112687, 2021.

[8] C. A. Edwards, A. Kouzani, K. H. Lee, and E. K. Ross, "Neurostimulation devices for the treatment of neurologic disorders," in *Mayo Clinic Proceedings*, vol. 92, no. 9. Elsevier, 2017, pp. 1427–1444.

[9] A. S. Widge, D. A. Malone Jr, and D. D. Dougherty, "Closing the loop on deep brain stimulation for treatment-resistant depression," *Frontiers in neuroscience*, vol. 12, p. 175, 2018.

[10] N. D. Sisterson, T. A. Wozny, V. Kokkinos, A. Constantino, and R. M. Richardson, "Closed-loop brain stimulation for drug-resistant epilepsy: towards an evidence-based approach to personalized medicine," *Neurotherapeutics*, vol. 16, no. 1, pp. 119–127, 2019.

[11] R. Molina, C. J. Hass, S. Cernera, K. Sowalsky, A. C. Schmitt, J. A. Roper, D. Martinez-Ramirez, E. Opri, C. W. Hess, R. S.

Eisinger *et al.*, "Closed-loop deep brain stimulation to treat medication-refractory freezing of gait in parkinson's disease," *Frontiers in human neuroscience*, vol. 15, p. 48, 2021.

[12] M. Kesmia, S. Boughaba, and S. Jacquir, "Control of continuous dynamical systems modeling physiological states," *Chaos, Solitons & Fractals*, vol. 136, p. 109805, 2020.

[13] L. Shupe and E. Fetz, "An integrate-and-fire spiking neural network model simulating artificially induced cortical plasticity," *Eneuro*, vol. 8, no. 2, 2021.

[14] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[15] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[16] L. Lapique, "Recherches quantitatives sur l'excitation electrique des nerfs traitee comme une polarization." *Journal of Physiology and Pathololgy*, vol. 9, pp. 620–635, 1907.

[17] L. F. Abbott, "Lapicque's introduction of the integrate-and-fire model neuron (1907)," *Brain research bulletin*, vol. 50, no. 5-6, pp. 303–304, 1999.

[18] E. M. Izhikevich, "Neural excitability, spiking and bursting," *International journal of bifurcation and chaos*, vol. 10, no. 06, pp. 1171–1266, 2000.

[19] ——, *Dynamical systems in neuroscience*. MIT press, 2007.

[20] S. H. Wright, "Generation of resting membrane potential," *Advances in physiology education*, vol. 28, no. 4, pp. 139–142, 2004.

[21] G. Young, "Note on excitation theories," *Psychometrika*, vol. 2, no. 2, pp. 103–106, 1937.

[22] Ş. Mihalaş and E. Niebur, "A generalized linear integrate-and-fire neural model produces diverse spiking behaviors," *Neural computation*, vol. 21, no. 3, pp. 704–718, 2009.

[23] S. A. Prescott, Y. De Koninck, and T. J. Sejnowski, "Biophysical basis for three distinct dynamical mechanisms of action potential initiation," *PLoS computational biology*, vol. 4, no. 10, p. e1000198, 2008.

[24] R. Brette and W. Gerstner, "Adaptive exponential integrate-and-fire model as an effective description of neuronal activity," *Journal of neurophysiology*, vol. 94, no. 5, pp. 3637–3642, 2005.

[25] G. Marmont, "Studies on the axon membrane. i. a new method," *Journal of cellular and comparative physiology*, vol. 34, no. 3, pp. 351–382, 1949.

[26] K. S. Cole and J. W. Moore, "Ionic current measurements in the squid giant axon membrane," *The Journal of general physiology*, vol. 44, no. 1, pp. 123–167, 1960.

[27] A. C. Scott, "The electrophysics of a nerve fiber," *Reviews of Modern Physics*, vol. 47, no. 2, p. 487, 1975.

[28] R. Brette, "What is the most realistic single-compartment model of spike initiation?" *PLoS computational biology*, vol. 11, no. 4, p. e1004114, 2015.

[29] H. Poincaré, *New methods of celestial mechanics*. National Aeronautics and Space Administration, 1967, vol. 2.

[30] H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever *et al.*, "Reconstruction and simulation of neocortical microcircuitry," *Cell*, vol. 163, no. 2, pp. 456–492, 2015.

[31] T. B. Kepler, L. Abbott, and E. Marder, "Reduction of conductance-based neuron models," *Biological cybernetics*, vol. 66, no. 5, pp. 381–387, 1992.

[32] B. Van der Pol, "Lxxxviii. on "relaxation-oscillations"," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 978–992, 1926.

[33] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical journal*, vol. 1, no. 6, pp. 445–466, 1961.

[34] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proceedings of the IRE*, vol. 50, no. 10, pp. 2061–2070, 1962.

[35] N. Faedo, D. García-Violini, and J. V. Ringwood, "Controlling synchronization in a complex network of nonlinear oscillators via feedback linearisation and $\mathcal{H}_\infty$-control," *Chaos, Solitons & Fractals*, vol. 144, p. 110722, 2021.

[36] C. Morris and H. Lecar, "Voltage oscillations in the barnacle giant muscle fiber," *Biophysical journal*, vol. 35, no. 1, pp. 193–213, 1981.

[37] P.-L. Buono and M. Golubitsky, "Models of central pattern generators for quadruped locomotion i. primary gaits," *Journal of mathematical biology*, vol. 42, no. 4, pp. 291–326, 2001.

[38] J. L. Hindmarsh and R. Rose, "A model of neuronal bursting using three coupled first order differential equations," *Proceedings of the Royal society of London. Series B. Biological sciences*, vol. 221, no. 1222, pp. 87–102, 1984.

[39] B. W. Knight, "The relationship between the firing rate of a single neuron and the level of activity in a population of neurons: Experimental evidence for resonant enhancement in the population response," *The Journal of general physiology*, vol. 59, no. 6, pp. 767–778, 1972.

[40] Y. Ahmadian, A. M. Packer, R. Yuste, and L. Paninski, "Designing optimal stimuli to control neuronal spike timing," *Journal of neurophysiology*, vol. 106, no. 2, pp. 1038–1053, 2011.

[41] P. Latham, B. Richmond, S. Nirenberg, and P. Nelson, "Intrinsic dynamics in neuronal networks. ii. experiment," *Journal of neurophysiology*, vol. 83, no. 2, pp. 828–835, 2000.

[42] I. Erchova, G. Kreck, U. Heinemann, and A. V. Herz, "Dynamics of rat entorhinal cortex layer ii and iii cells: characteristics of membrane potential resonance at rest predict oscillation properties near threshold," *The Journal of physiology*, vol. 560, no. 1, pp. 89–110, 2004.

[43] C. Clopath, R. Jolivet, A. Rauch, H.-R. Lüscher, and W. Gerstner, "Predicting neuronal activity with simple models of the threshold type: Adaptive exponential integrate-and-fire model with two compartments," *Neurocomputing*, vol. 70, no. 10-12, pp. 1668–1673, 2007.

[44] L. Shiau and L. Buhry, "Interneuronal gamma oscillations in hippocampus via adaptive exponential integrate-and-fire neurons," *Neurocomputing*, vol. 331, pp. 220–234, 2019.

[45] T.-A. E. Nghiem, N. Tort-Colet, T. Górski, U. Ferrari, S. Moghimyfiroozabad, J. S. Goldman, B. Teleńczuk, C. Capone, T. Bal, M. Di Volo *et al.*, "Cholinergic switch between two types of slow waves in cerebral cortex," *Cerebral Cortex*, vol. 30, no. 6, pp. 3451–3466, 2020.

[46] R. Guttman, S. Lewis, and J. Rinzel, "Control of repetitive firing in squid axon membrane as a model for a neuroneoscillator." *The Journal of physiology*, vol. 305, no. 1, pp. 377–395, 1980.

[47] G. Ullah and S. J. Schiff, "Tracking and control of neuronal hodgkin-huxley dynamics," *Physical Review E*, vol. 79, no. 4, p. 040901, 2009.

[48] O. Berenfeld and S. Abboud, "Simulation of cardiac activity and the ecg using a heart model with a reaction-diffusion action potential," *Medical engineering & physics*, vol. 18, no. 8, pp. 615–625, 1996.

[49] Y. Jiang and J. Dai, "Robust control of chaos in modified fitzhugh-nagumo neuron model under external electrical stimulation based on internal model principle," *Kybernetika*, vol. 47, no. 4, pp. 612–629, 2011.

[50] A. Iolov, S. Ditlevsen, and A. Longtin, "Stochastic optimal control of single neuron spike trains," *Journal of neural engineering*, vol. 11, no. 4, p. 046004, 2014.

[51] D. Matzakos-Karvouniari, L. Gil, E. Orendorff, O. Marre, S. Picaud, and B. Cessac, "A biophysical model explains the spontaneous bursting behavior in the developing retina," *Scientific reports*, vol. 9, no. 1, pp. 1–23, 2019.

[52] M. G. Rosenblum and A. S. Pikovsky, "Controlling synchronization in an ensemble of globally coupled oscillators," *Physical Review Letters*, vol. 92, no. 11, p. 114102, 2004.

[53] Y. Che, S. Zhang, J. Wang, S. Cui, C. Han, B. Deng, and X. Wei, "Synchronization of inhibitory coupled hindmarsh-rose neurons via adaptive sliding mode control," in *2011 2nd International Conference on Intelligent Control and Information Processing*, vol. 2. IEEE, 2011, pp. 1134–1139.

[54] S. Beyhan, "Affine ts fuzzy model-based estimation and control of hindmarsh–rose neuronal model," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 8, pp. 2342–2350, 2017.

[55] J. S. Montijn, G. T. Meijer, C. S. Lansink, and C. M. Pennartz, "Population-level neural codes are robust to single-neuron variability from a multidimensional coding perspective," *Cell reports*, vol. 16, no. 9, pp. 2486–2498, 2016.

[56] J. K. Wong, E. H. Middlebrooks, S. S. Grewal, L. Almeida, C. W. Hess, and M. S. Okun, "A comprehensive review of brain connectomics and imaging to improve deep brain stimulation outcomes," *Movement Disorders*, vol. 35, no. 5, pp. 741–751, 2020.

[57] L. Reichl, *A modern course in statistical physics*. American Association of Physics Teachers, 1980.

[58] M. Breakspear, J. R. Terry, and K. J. Friston, "Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics," *Network: Computation in Neural Systems*, vol. 14, no. 4, p. 703, 2003.

[59] A. C. Marreiros, J. Daunizeau, S. J. Kiebel, and K. J. Friston, "Population dynamics: variance and the sigmoid activation function," *Neuroimage*, vol. 42, no. 1, pp. 147–157, 2008.

[60] R. L. Beurle, "Properties of a mass of cells capable of regenerating pulses," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, pp. 55–94, 1956.

[61] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical journal*, vol. 12, no. 1, pp. 1–24, 1972.

[62] J. G. Kirkwood, "The statistical mechanical theory of transport processes i. general theory," *The Journal of Chemical Physics*, vol. 14, no. 3, pp. 180–201, 1946.

[63] D. T. Liley, P. J. Cadusch, and M. P. Dafilis, "A spatially continuous mean field theory of electrocortical activity," *Network: Computation in Neural Systems*, vol. 13, no. 1, p. 67, 2001.

[64] B. G. Cragg and H. N. V. Temperley, "Memory: the analogy with ferromagnetic hysteresis," *Brain*, vol. 78, no. 2, pp. 304–316, 1955.

[65] B. H. Jansen, G. Zouridakis, and M. E. Brandt, "A neurophysiologically-based mathematical model of flash visual evoked potentials," *Biological cybernetics*, vol. 68, no. 3, pp. 275–283, 1993.

[66] B. H. Jansen and V. G. Rit, "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biological cybernetics*, vol. 73, no. 4, pp. 357–366, 1995.

[67] A. Spiegler, S. J. Kiebel, F. M. Atay, and T. R. Knösche, "Bifurcation analysis of neural mass models: Impact of extrinsic inputs and dendritic time constants," *NeuroImage*, vol. 52, no. 3, pp. 1041–1058, 2010.

[68] D. Sholl, "The organization of the visual cortex in the cat," *Journal of anatomy*, vol. 89, no. Pt 1, p. 33, 1955.

[69] J. Griffith, "On the stability of brain-like structures," *Biophysical journal*, vol. 3, no. 4, pp. 299–308, 1963.

[70] ——, "A field theory of neural nets: I: Derivation of field equations," *The bulletin of mathematical biophysics*, vol. 25, no. 1, pp. 111–120, 1963.

[71] J. S. Griffith, "A field theory of neural nets: Ii. properties of the field equations," *The Bulletin of mathematical biophysics*, vol. 27, no. 2, pp. 187–195, 1965.

[72] H. R. Wilson and J. D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetik*, vol. 13, no. 2, pp. 55–80, 1973.

[73] S.-I. Amari, "Homogeneous nets of neuron-like elements," *Biological cybernetics*, vol. 17, no. 4, pp. 211–220, 1975.

[74] S.-i. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biological cybernetics*, vol. 27, no. 2, pp. 77–87, 1977.

[75] S. J. Schiff and T. Sauer, "Kalman filter control of a model of spatiotemporal cortical dynamics," *Journal of neural engineering*, vol. 5, no. 1, p. 1, 2007.

[76] W. Pasillas-Lépine, I. Haidar, A. Chaillet, and E. Panteley, "Closed-loop deep brain stimulation based on firing-rate regulation," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 166–169.

[77] J. Orłowski, A. Chaillet, M. Sigalotti, and A. Destexhe, "Adaptive stimulation strategy for selective brain oscillations disruption in a neuronal population model with delays," *Trends in Neurosciences*, vol. 30, no. 7, pp. 357–364, 2019.

[78] B. Duchet, G. Weerasinghe, C. Bick, and R. Bogacz, "Optimizing deep brain stimulation based on isostable amplitude in essential tremor patient models," *Journal of neural engineering*, vol. 18, no. 4, p. 046023, 2021.

[79] F. Arab, S. Rostami, M. Dehghani-Habibabadi, V. Salari, and M.-S. Safari, "Optogenetically induced spatiotemporal gamma oscillations in visual cortex," *arXiv preprint arXiv:2103.05635*, 2021.

[80] J. Wang, E. Niebur, J. Hu, and X. Li, "Suppressing epileptic activity in a neural mass model using a closed-loop proportional-integral controller," *Scientific reports*, vol. 6, no. 1, pp. 1–12, 2016.

[81] Z. Liang, Z. Luo, K. Liu, J. Qiu, and Q. Liu, "Deep koopman-operator based model predictive control for closed-loop electrical neurostimulation in epilepsy," *arXiv preprint arXiv:2103.14321*, 2021.

[82] P. Beim Graben, D. Pinotsis, D. Saddy, and R. Potthast, "Language processing with dynamic fields," *Cognitive Neurodynamics*, vol. 2, no. 2, p. 79, 2008.

[83] P. Beim Graben and R. Potthast, "Inverse problems in dynamic cognitive modeling," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 1, p. 015103, 2009.

[84] B. L. Foster, I. Bojak, and D. T. Liley, "Population based models of cortical drug response: insights from anaesthesia," *Cognitive Neurodynamics*, vol. 2, no. 4, pp. 283–296, 2008.

[85] P. Selvaraj, J. W. Sleigh, H. E. Kirsch, and A. J. Szeri, "Closed-loop feedback control and bifurcation analysis of epileptiform activity via optogenetic stimulation in a mathematical model of human cortex," *Physical Review E*, vol. 93, no. 1, p. 012416, 2016.

[86] K. R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge, 1963.

[87] A. Berényi, M. Belluscio, D. Mao, and G. Buzsáki, "Closed-loop control of epilepsy by transcranial electrical stimulation," *Science*, vol. 337, no. 6095, pp. 735–737, 2012.

[88] A. Nowacki, T.-K. Nguyen, G. Tinkhauser, K. Petermann, I. Debove, R. Wiest, and C. Pollo, "Accuracy of different three-dimensional subcortical human brain atlases for dbs–lead localisation," *NeuroImage: Clinical*, vol. 20, pp. 868–874, 2018.

[89] N. I. Bertone-Cueto, J. Makarova, A. Mosqueira, D. García-Violini, R. Sánchez-Peña, O. Herreras, M. Belluscio, and J. Piriz, "Volume-conducted origin of the field potential at the lateral habenula," *Frontiers in systems neuroscience*, p. 78, 2020.

[90] K. J. Åström and B. Wittenmark, *Computer-controlled systems: theory and design*. Courier Corporation, 2013.

[91] J. C. Henry, "Electroencephalography: basic principles, clinical applications, and related fields," *Neurology*, vol. 67, no. 11, pp. 2092–2092, 2006.

[92] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents—eeg, ecog, lfp and spikes," *Nature reviews neuroscience*, vol. 13, no. 6, pp. 407–420, 2012.

[93] E. W. Miller, "Small molecule fluorescent voltage indicators for studying membrane potential," *Current opinion in chemical biology*, vol. 33, pp. 74–80, 2016.

[94] C.-T. Chen, *Linear System Theory and Design*. Oxford University Press, 1999.

[95] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, 1960.

[96] G. Welch, G. Bishop *et al.*, *An introduction to the Kalman filter*. Chapel Hill, NC, USA, 1995.

[97] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–193.

[98] Y. Hao, Z. Xiong, F. Sun, and X. Wang, "Comparison of unscented kalman filters," in *2007 International Conference on Mechatronics and Automation*. IEEE, 2007, pp. 895–899.

[99] Y. Wei, G. Ullah, R. Parekh, J. Ziburkus, and S. J. Schiff, "Kalman filter tracking of intracellular neuronal voltage and current," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 5844–5849.

[100] T. D. Sauer and S. J. Schiff, "Data assimilation for heterogeneous networks: The consensus set," *Physical Review E*, vol. 79, no. 5, p. 051909, 2009.

[101] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans*, vol. 99, no. C5, pp. 10 143–10 162, 1994.

[102] L. Besedovsky, H.-V. V. Ngo, S. Dimitrov, C. Gassenmaier, R. Lehmann, and J. Born, "Auditory closed-loop stimulation of eeg slow oscillations strengthens sleep and signs of its immune-supportive function," *Nature communications*, vol. 8, no. 1, pp. 1–8, 2017.

[103] V. Mondini, R. J. Kobler, A. I. Sburlea, and G. R. Müller-Putz, "Continuous low-frequency eeg decoding of arm movement for closed-loop, natural control of a robotic arm," *Journal of Neural Engineering*, vol. 17, no. 4, p. 046031, 2020.

[104] Y. Zhong, Y. Wang, Z. He, Z. Lin, N. Pang, L. Niu, Y. Guo, M. Pan, and L. Meng, "Closed-loop wearable ultrasound deep brain stimulation system based on eeg in mice," *Journal of Neural Engineering*, vol. 18, no. 4, p. 0460e8, 2021.

[105] R. Pashaie, R. Baumgartner, T. J. Richner, S. K. Brodnick, M. Azimipour, K. W. Eliceiri, and J. C. Williams, "Closed-loop optogenetic brain interface," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 10, pp. 2327–2337, 2015.

[106] C.-Y. Wu, C.-H. Cheng, and Z.-X. Chen, "A 16-channel cmos chopper-stabilized analog front-end ecog acquisition circuit for a closed-loop epileptic seizure control system," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 3, pp. 543–553, 2018.

[107] A. Velisar, J. Syrkin-Nikolau, Z. Blumenfeld, M. Trager, M. Afzal, V. Prabhakar, and H. Bronte-Stewart, "Dual threshold neural closed loop deep brain stimulation in parkinson disease patients," *Brain stimulation*, vol. 12, no. 4, pp. 868–876, 2019.

[108] Y. Takeuchi, M. Harangozó, L. Pedraza, T. Földi, G. Kozák, Q. Li, and A. Berényi, "Closed-loop stimulation of the medial septum terminates epileptic seizures," *Brain*, vol. 144, no. 3, pp. 885–908, 2021.

[109] R. Fukuma, T. Yanagisawa, S. Yorifuji, R. Kato, H. Yokoi, M. Hirata, Y. Saitoh, H. Kishima, Y. Kamitani, and T. Yoshimine, "Closed-loop control of a neuroprosthetic hand by magnetoencephalographic signals," *PLoS One*, vol. 10, no. 7, p. e0131547, 2015.

[110] A. V. Kim and A. V. Ivanov, *Systems with delays: analysis, control, and computations*. John Wiley & Sons, 2015.

[111] A. C. Mennen, N. B. Turk-Browne, G. Wallace, D. Seok, A. Jaganjac, J. Stock, M. T. deBettencourt, J. D. Cohen, K. A. Norman, and Y. I. Sheline, "Closed-loop fmri neurofeedback to reduce negative attentional bias in depression," *bioRxiv*, 2020.

[112] B. Mulyana, A. Tsuchiyagaito, J. Smith, M. Misaki, R. Kuplicki, G. Soleimani, A. Rashedi, D. Shereen, T. O. Bergman, S. Cheng *et al.*, "Online closed-loop real-time tes-fmri for brain modulation: Feasibility, noise/safety and pilot study," *bioRxiv*, 2021.

[113] R. McKendrick, R. Parasuraman, and H. Ayaz, "Wearable functional near infrared spectroscopy (fnirs) and transcranial direct current stimulation (tdcs): expanding vistas for neurocognitive augmentation," *Frontiers in systems neuroscience*, vol. 9, p. 27, 2015.

[114] Z. Zhang, L. E. Russell, A. M. Packer, O. M. Gauld, and M. Häusser, "Closed-loop all-optical interrogation of neural circuits in vivo," *Nature methods*, vol. 15, no. 12, pp. 1037–1040, 2018.

[115] J. Selberg, M. Jafari, J. Mathews, M. Jia, P. Pansodtee, H. Dechiraju, C. Wu, S. Cordero, A. Flora, N. Yonas *et al.*, "Machine learning-driven bioelectronics for closed-loop control of cells," *Advanced Intelligent Systems*, vol. 2, no. 12, p. 2000140, 2020.

[116] G. C. Goodwin, S. F. Graebe, M. E. Salgado *et al.*, *Control system design*. Prentice Hall New Jersey, 2001, vol. 240.

[117] M. Piccolino, "Luigi galvani's path to animal electricity," *Comptes rendus biologies*, vol. 329, no. 5-6, pp. 303–318, 2006.

[118] F. Rattay, "The basic mechanism for the electrical stimulation of the nervous system," *Neuroscience*, vol. 89, no. 2, pp. 335–346, 1999.

[119] N. J. Michelson, J. R. Eles, A. L. Vazquez, K. A. Ludwig, and T. D. Kozai, "Calcium activation of cortical neurons by continuous electrical stimulation: Frequency dependence, temporal fidelity, and activation density," *Journal of neuroscience research*, vol. 97, no. 5, pp. 620–638, 2019.

[120] D. Kraus, G. Naros, R. Bauer, F. Khademi, M. T. Leão, U. Ziemann, and A. Gharabaghi, "Brain state-dependent transcranial magnetic closed-loop stimulation controlled by sensorimotor desynchronization induces robust increase of corticospinal excitability," *Brain Stimulation*, vol. 9, no. 3, pp. 415–424, 2016.

[121] A. E. Tervo, J. Metsomaa, J. O. Nieminen, J. Sarvas, and R. J. Ilmoniemi, "Automated search of stimulation targets with closed-loop transcranial magnetic stimulation," *NeuroImage*, vol. 220, p. 117082, 2020.

[122] J. Leite, L. Morales-Quezada, S. Carvalho, A. Thibaut, D. Doruk, C.-F. Chen, S. C. Schachter, A. Rotenberg, and F. Fregni, "Surface eeg-transcranial direct current stimulation (tdcs) closed-loop system," *International journal of neural systems*, vol. 27, no. 06, p. 1750026, 2017.

[123] M. A. Yaqub, K.-S. Hong, A. Zafar, and C.-S. Kim, "Control of transcranial direct current stimulation duration by assessing functional connectivity of near-infrared spectroscopy signals," *International Journal of Neural Systems*, p. 2150050, 2021.

[124] M. R. Boyle and F. Fröhlich, "Eeg feedback-controlled transcranial alternating current stimulation," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 140–143.

[125] F. Fröhlich and L. Townsend, "Closed-loop transcranial alternating current stimulation: Towards personalized non-invasive brain stimulation for the treatment of psychiatric illnesses," *Current Behavioral Neuroscience Reports*, pp. 1–7, 2021.

[126] J.-P. Lefaucheur, "Methods of therapeutic cortical stimulation," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 39, no. 1, pp. 1–14, 2009.

[127] T. Hashimoto, C. M. Elder, M. S. Okun, S. K. Patrick, and J. L. Vitek, "Stimulation of the subthalamic nucleus changes the firing pattern of pallidal neurons," *Journal of neuroscience*, vol. 23, no. 5, pp. 1916–1923, 2003.

[128] K. Deisseroth, "Optogenetics," *Nature methods*, vol. 8, no. 1, pp. 26–29, 2011.

[129] N. Grossman, K. Nikolic, C. Toumazou, and P. Degenaar, "Modeling study of the light stimulation of a neuron cell with channelrhodopsin-2 mutants," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1742–1751, 2011.

[130] J. C. Williams, J. Xu, Z. Lu, A. Klimas, X. Chen, C. M. Ambrosi, I. S. Cohen, and E. Entcheva, "Computational optogenetics: empirically-derived voltage-and light-sensitive channelrhodopsin-2 model," *PLoS computational biology*, vol. 9, no. 9, p. e1003220, 2013.

[131] S. Ching and J. Ritt, "Control strategies for underactuated neural ensembles driven by optogenetic stimulation," *Frontiers in neural circuits*, vol. 7, p. 54, 2013.

[132] K. J. Åström and T. Hagglund, *PID Controllers: Theory, Design and Tuning*. International Society of Automation, 1995.

[133] M. Safonov and T. Tsao, "The unfalsified control concept and learning," *IEEE Transsactions on Automatic Control*, vol. 42, no. 6, pp. 843–847, 1997.

[134] M. Stefanovic and M. G. Safonov, *Safe Adaptive Control*. London: Springer, 2011.

[135] F. D. Bianchi, C. Ocampo-Martinez, C. Kunusch, and R. S. Sánchez-Peña, "Fault-tolerant unfalsified control for PEM fuel cell systems," *IEEE Transactions on Energy Conversion*, vol. 30, no. 1, pp. 307–315, 2015.

[136] K. Glover, "All optimal hankel-norm approximations of linear multivariable systems and their L_∞-error bounds," *International journal of control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[137] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice–Hall, 1996.

[138] R. S. Sánchez Peña and M. Sznaier, *Robust Systems Theory and Applications*. John Wiley & Sons, Inc., 1998.

[139] R. Smith, G. Dullerud, S. Rangan, and K. Poolla, "Model validation for dynamically uncertain systems," *Mathematical Modelling of Systems*, vol. 3, no. 1, pp. 43–58, 1997.

[140] J. Anderson and A. Papachristodoulou, *Model Invalidation*. New York, NY: Springer New York, 2013, pp. 1395–1398.

[141] W. Dubitzky, O. Wolkenhauer, K.-H. Cho, and H. Yokota, Eds., *Model Falsification*. New York, NY: Springer New York, 2013, pp. 1391–1391.

[142] T. Yang, N. Sun, and Y. Fang, "Neuroadaptive control for complicated underactuated systems with simultaneous output and velocity constraints exerted on both actuated and unactuated states," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[143] T. Yang, H. Chen, N. Sun, and Y. Fang, "Adaptive neural network output feedback control of uncertain underactuated systems with actuated and unactuated state constraints," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021.

[144] Y.-m. Li, X. Min, and S. Tong, "Adaptive fuzzy inverse optimal control for uncertain strict-feedback nonlinear systems," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 10, pp. 2363–2374, 2019.

[145] Y. Li, X. Min, and S. Tong, "Observer-based fuzzy adaptive inverse optimal output feedback control for uncertain nonlinear systems," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 6, pp. 1484–1495, 2020.

[146] M. Bolus, "Closed loop optogenetic control and thalamic state," Ph.D. dissertation, Georgia Institute of Technology, 2020.

[147] A. Isidori, *Nonlinear Control Systems*. London: Springer Verlag, 1995.

[148] H. K. Khalil, *Nonlinear Systems*. New Jersey: Prentice Hall, 2002.

[149] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

[150] A. Packard, "Gain scheduling via LFTs," *Systems and Control Letters*, vol. 22, pp. 79–92, 1993.

[151] G. S. Becker and A. Packard, "Robust performance of LPV systems using parametrically-dependent linear feedback," *Systems and Control Letters*, vol. 23, pp. 205–215, 1994.

[152] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*. Heidelberg: Springer, LNCIS, 2010.

[153] J. P. Newman, M.-f. Fong, D. C. Millard, C. J. Whitmire, G. B. Stanley, and S. M. Potter, "Optogenetic feedback control of neural activity," *Elife*, vol. 4, p. e07192, 2015.

[154] M. F. Bolus, A. A. Willats, C. J. Rozell, and G. B. Stanley, "State-space optimal feedback control of optogenetically driven neural activity," *Journal of Neural Engineering*, vol. 18, no. 3, p. 036006, 2021.

[155] A. K. Tangirala, *Principles of system identification: theory and practice*. Crc Press, 2018.

[156] T. Katayama *et al.*, *Subspace methods for system identification*. Springer, 2005, vol. 1.

[157] Y. Yang, A. T. Connolly, and M. M. Shanechi, "A control-theoretic system identification framework and a real-time closed-loop clinical simulation testbed for electrical brain stimulation," *Journal of neural engineering*, vol. 15, no. 6, p. 066007, 2018.

[158] O. G. Sani, H. Abbaspourazad, Y. T. Wong, B. Pesaran, and M. M. Shanechi, "Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification," *Nature Neuroscience*, vol. 24, no. 1, pp. 140–149, 2021.