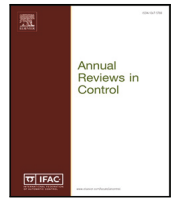




Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Full Length Article

Developing control-theoretic objectives for large-scale brain dynamics and cognitive enhancement

Matthew F. Singh^{a,b,c}, Michael W. Cole^b, Todd S. Braver^c, ShiNung Ching^{a,*}^a Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, 63130, MO, USA^b Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, 07102, NJ, USA^c Psychological and Brain Science, Washington University in St. Louis, St. Louis, 63130, MO, USA

ARTICLE INFO

Keywords:

Brain stimulation
Neurocontrol theory
Brain dynamics
Neural reachability

ABSTRACT

The development of technologies for brain stimulation provides a means for scientists and clinicians to directly actuate the brain and nervous system. Brain stimulation has shown intriguing potential in terms of modifying particular symptom clusters in patients and behavioral characteristics of subjects. The stage is thus set for optimization of these techniques and the pursuit of more nuanced stimulation objectives, including the modification of complex cognitive functions such as memory and attention. Control theory and engineering will play a key role in the development of these methods, guiding computational and algorithmic strategies for stimulation. In particular, realizing this goal will require new development of frameworks that allow for controlling not only brain activity, but also latent dynamics that underlie neural computation and information processing. In the current opinion, we review recent progress in brain stimulation and outline challenges and potential research pathways associated with exogenous control of cognitive function.

1. Introduction

An overt point of intersection between control engineering and clinical brain science lies in the development of neurostimulation technologies for modifying brain activity and consequent behavior. Such methods have been in use for many decades for a broad spectrum of neurological and neuropsychiatric illnesses, and have demonstrated high clinical efficacy (Husain, Lisanby, & Kay, 2015; Perlmutter & Mink, 2006; Read & Bentall, 2010). However, the invasiveness of some stimulation methods, including electroconvulsive therapy and deep brain stimulation, have limited application to exceptional cases, typically in individuals with very advanced forms of disease (Lange et al., 2017). An emerging goal in neurology and psychiatry is the development of methods and technology that can be used in first-line treatments, particularly to address cognitive deficits, such as impairments in attention and memory (Goodman & Insel, 2009; Hoy & Fitzgerald, 2010). Typical first-line clinical strategies, including behavioral and systemic pharmacological interventions are effective with respect to specific symptom clusters (e.g., mood) but are less effective with respect to cognitive effects, though exceptions exist (Herrera-Guzmán et al., 2009). Moreover, these strategies are often mechanistically opaque and, for unknown reasons, are ineffective in many patients (Stassen, Angst, Hell, Scharfetter, & Szege, 2007). Thus, a challenge and opportunity exists to extend the synergy between

control engineering and clinical brain science to develop principled and interpretable strategies for brain stimulation that can enhance cognitive function in patients and perhaps, eventually, in healthy individuals.

A promising avenue towards the above goal involves use of non-invasive brain stimulation technologies, such as transcranial Electrical Stimulation (tES) (Fertonani & Miniussi, 2017; Paulus, 2011), which involves applying weak current to the brain using two or more electrodes positioned on the scalp. Most tES devices are current-controlled, to account for variation in skin/skull conductance (see Antonakakis et al. (2020), Antonenko et al. (2021)); the strength of stimulation is therefore described in terms of current (typically 1–4 mA).

Contemporary approaches to tES are (almost always) open-loop in nature, and have taken the form of either (i) DC stimulation (tDCS), theorized to increase/decrease neural excitability (Nitsche et al., 2003, 2004) or (ii) single-frequency AC stimulation (tACS), theorized to entrain brain activity to the applied sinusoidal signal (Huang et al., 2021). These techniques largely share the same motivation: to increase the strength of activity in a brain area (and/or frequency) correlated with a target cognitive function. The efficacy of the approach has been debated (particularly tDCS Horvath, Forte, & Carter, 2015a, 2015b; Meron, Hedger, Garner, & Baldwin, 2015; Roche, Geiger, & Bussel, 2015) and appears to exhibit individual variability (Emonson, Fitzgerald, Rogasch, & Hoy, 2019; Hsu, Juan, & Tseng, 2016; Katz

* Corresponding author.

E-mail address: shinung@wustl.edu (S. Ching).<https://doi.org/10.1016/j.arcontrol.2022.05.001>

Received 4 January 2022; Accepted 14 May 2022

1367-5788/© 2022 Elsevier Ltd. All rights reserved.

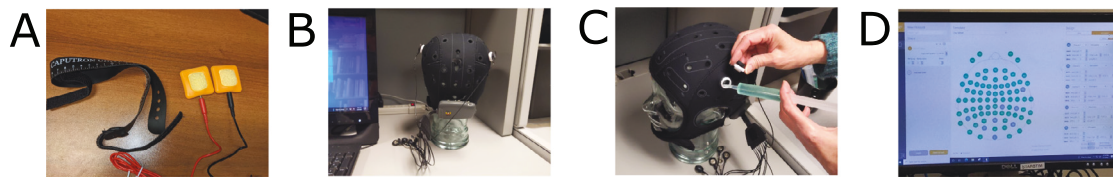


Fig. 1. Common tES devices. (A) Example of a single-channel device with medium-sized sponge electrodes. Electrodes are attached to the head using an elastic band (left side). (B) Reverse view of a high-density device mounted on an adjustable head-cap. This particular device is mobile with the battery/control-unit on the back. (C) An electrolytic solution is used to improve contact between the scalp and electrodes for the device in (B). (D) High-density devices typically include a software suite for controlling the device and monitoring impedance (screen shows impedance monitoring used during device setup).

et al., 2017; López-Alonso, Cheeran, Río-Rodríguez, & Fernández-del Olmo, 2014), though some early studies indicate intriguing potential for affecting complex brain functions (Green et al., 2017; Miller, Berger, & Sauseng, 2015; Reinhart & Nguyen, 2019; Zhu et al., 2015). Current approaches are, however, potentially limiting, since they are largely based on trial-and-error, without directly leveraging the role of brain dynamics in neural computation. Hence, there is a clear opportunity at the nexus of engineering and brain science to develop control-theoretic approaches that can positively impact the utilization of these emerging technologies.

In the current paper, we explore the challenges and potential of a control-systems framework for designing controllers to enhance human cognitive function, with a focus on the development of conceptual and mathematical objective functions. The intent of this paper is to identify current technical and theoretical challenges in human neural control across spatial scales and suggest promising pathways forward. In particular, we focus on the formulation of objectives and system identification paradigms as a precursor to the eventual synthesis of control strategies for brain stimulation.

The paper is organized with a brief introduction to tES application, followed by a general discussion of the state-space description and theoretical constraints of closed-loop tES. In Section 3 we review previous approaches to brain modeling and control and highlight recent advances in neural system-identification. In Section 4 we critically-review current tES frameworks and argue against their underlying logic. Our novel contribution is a control-theoretic reframing of neurostimulation objectives in Section 5. Lastly, we identify promising pathways and fundamental limitations in using macroscale tES to influence the cellular-computations that underlie cognition (Section 6).

1.1. tES hardware

Similar to EEG hardware, tES devices come in three forms: single-channel, multi-channel, and high-density (see Fig. 1). Single-channel devices typically use two unmounted electrodes and are unique in continuing to support large surface-area electrodes in the form of either “paddles” or saline-soaked sponges (Fig. 1A). These first-generation devices continue to be popular, particularly for DC stimulation. By contrast, multi-channel devices use slightly larger variants of EEG electrodes with contact formed by a small sponge-pedestal or electrolytic gel. The multiple channels afforded by such devices have been used to either concentrate the electric field (e.g. by surrounding a cathode by a ring of anodes) or to stimulate separate parts of the brain. Most recently, high-density devices have been developed in which electrodes are mounted in an EEG-style net/cap (Fig. 1B–D). These devices are often designed to be used alongside high-density EEG and some integrated systems allow the same electrodes to be assigned either stimulation (tES) or sensing (EEG) roles.

1.2. Previous attempts at closing the loop for tES

As previously mentioned, the preponderance of tES studies have employed open-loop stimulation. However, there has been increasing

interest in developing closed-loop brain-stimulation approaches. Several emerging lines of research have begun exploring these avenues with considerable variation in the measurements used to construct feedback. Example approaches include coupling tES and EEG spectrograms (this did not work for Stecher, Notbohm, Kasten, & Herrmann, 2021) and approaches which attempt to empirically maximize stimulation parameters (amplitude, frequency) by measuring their effect on behavior (automated trial-and-error; Lorenz et al., 2019). Other studies have explored synchronizing tACS onset to ongoing EEG measurements (i.e., starting tACS when EEG is at 0-phase; Ketz, Jones, Bryant, Clark, & Pilly, 2018; Raco, Bauer, Tharsan, & Gharabaghi, 2016; Reinhart & Nguyen, 2019; Schwab, Misselhorn, & Engel, 2019). The latter case represents the current state-of-the-art, but has only been explored recently and in a small number of studies such as one concerning the repair of memory function in older adults (Reinhart & Nguyen, 2019), and another in which tES was applied during sleep (Ketzer et al., 2018).

1.2.1. Closed-loop processing

As discussed above, tES has clearly advanced beyond pure open-loop protocols. However, these advances have largely been driven by heuristics which may be improved upon by formal control-synthesis. For the present, we solely consider the coupling of tES with brain measurements (as opposed to behavior). As with other control-systems, there are several design considerations in establishing closed-loop tES. On the measurement side, neurophysiological effects are often superimposed with other sources of bioelectricity such as eye movements (the eye is polarized), muscle tension, and cardiac activity (Mumtaz, Rasheed, & Irfan, 2021; Zhou, Johnson, & Muller, 2018). These artifacts are typically removed by Independent Component Analysis (ICA) which is very successful at identifying linear artifacts, such as heart-activity (Mennes, Wouters, Vanrumste, Lagae, & Stiers, 2010; Subasi & Gursoy, 2010), and easy to implement online as a spatial filter using public software packages (e.g., Mammone, La Foresta, and Morabito (2011), Pion-Tonachini, Hsu, Chang, Jung, and Makeig (2018)). This process is further augmented by regressing out other physiological signals which can be recorded simultaneously (EOG, EMG, ECG, etc.). However, other artifacts, such as motion, may be too large and variable to fully correct online (such timepoints are normally discarded offline).

Timing is also an important consideration for closed-loop design. Unmodeled temporal delays compromise the performance of a closed-loop system and can lead to qualitative changes such as a loss of stability (Agrawal, Fujino, & Bhartia, 1993; Wischert, Wunderlin, Pelster, Olivier, & Gros Lambert, 1994). A significant literature now considers delayed systems, particularly in the context of linear regulators (e.g., Cloosterman, van de Wouw, Heemels, and Nijmeijer (2006), Hmamed, Benzaouia, and Bensalah (1995)). However, for nonlinear stochastic systems, delays are particularly bothersome. For a deterministic system, model-predictive methods can be used to estimate the system’s current state from lagged measurements. In a stochastic system, however, the action of a control signal ($\delta x/\delta u$) will depend upon unknowable disturbances (process noise) that occur during the delay. Minimizing such delays should therefore be considered in the control-design process. Since the lag between a control-command and electrical control-delivery can be made negligible, we focus upon delays generated by computing devices. We note, however, that many

commercial tES systems do not allow direct control of the amplifier by external devices which can generate insurmountable lags (i.e., seconds) when using proprietary software.

At rest, human EEG is dominated by fluctuations in the alpha-band (8-12 Hz). In principal, M/EEG data are thus digitized at an adequate rate (typically 250 Hz-2 KHz) to observe this process in real-time. However, additional time is lost during the buffer process and subsequent data processing. Although some of these concerns are related to processing-speed, online digital filters inherently generate output-lags. This feature leads to a fundamental trade-off between filter-order and temporal lag which is not solved by faster hardware. Lastly, significant processing delays can result from the control-algorithm. Brain dynamics are nonlinear, noisy, and high-dimensional. Such features generally necessitate numerical techniques and some algorithms may be too expensive to compute online. For example, solving the Riccati equations associated with a 100-dimensional system can take tens of ms per case, but 100-dimensional matrix inversion is easily performed at several KHz (tested in MATLAB). Thus, control algorithms involving an inverse-Jacobian may be tenable, whereas regulator-style solutions may be too expensive to apply directly. We also note that direct approaches are not always necessary, or recommended. Substantial advances have been made in model-order reduction (Bai, 2002; Benner, Gugercin, & Willcox, 2015; Breiten & Damm, 2010) and statistical/machine-learning approaches to solving control-problems from a parameterized model of the underlying system (Gomi & Kawato, 1993; Vrabie & Lewis, 2009). We expect that such approaches will prove instrumental in filling this gap.

2. Problem setting and constraints

2.1. State-space description of closed-loop tES

We consider a prototypical tES scenario wherein current is administered at one set of scalp electrodes and brain activity is recorded via another (possibly overlapping) set of scalp electrodes. Due to the low conductance of the skull, a large portion of applied current shunts across the scalp while, conversely, the scalp potentials (e.g., recorded via electroencephalography, EEG) generated by brain activity are relatively weak (Rice, Rorden, Little, & Parra, 2013). Electrical fields primarily derive from/interact with pyramidal neurons, which are oriented approximately normal to the cortical surface, whereas the contribution of non-oriented cells is typically assumed negligible (Buzsáki, Anastassiou, & Koch, 2012; Kirschstein & Köhling, 2009). The temporal resolution of measurement/stimulation is on the order of milliseconds and hence it is reasonable to assume that the dynamics can be well-approximated by a discrete-time system with time-step equal to the sampling resolution. Likewise, a quasistatic approximation of induced electromagnetic fields is justified, so that a model can be written in the form:

$$x_{t+1} = f_t(x_t, Bu_t, \omega_t) \quad (1)$$

$$y_t = Cx_t + \eta_t \text{ for } u_t = 0, \text{ else undef.}, \quad (2)$$

in which f_t is a state-transition function, u_t is the applied current (i.e., stimulation), x_t is a vector of neural state variables and B, C denote the control and measurement matrices (gain), respectively. We denote the process and measurement noise as ω_t, η_t , respectively. For the present purposes, we do not require any assumptions (e.g., independence) for ω_t, η_t . This framing describes a closed system in the absence of stimulation. In reality, the brain is not a closed system, hence an extension of ω_t to include sensory input may also be useful (in which case ω_t will be autocorrelated). Often, in practice, the influence of ω_t is modeled as additive.

2.2. Constraints on tES as a control system

Despite the potential of the above modeling paradigm, several nontrivial challenges for identification and control permeate this setup. First are basic but important pragmatic issues surrounding tES technology. Since stimulation saturates recording electrodes (a phenomenon known as stimulation artifact Noury & Siegel, 2017, 2018), generic instantaneous-feedback approaches are technically impractical. Similarly, this issue presents a potential challenge to identifying input-output system models. The latter is compounded because only a limited amount of stimulation data can be collected from humans, on account of both safety constraints and practical issues such as maximum amount of time individuals can maintain attention while engaged with a stimulation apparatus (Rossi et al., 2009).

Further, the measurement process involves both spatial mixing of measured outputs (from the different neural generators/brain regions; Grech et al., 2008) and a latent-variable problem as the majority of state-variables (other cell-types, subcortex, molecular mechanisms etc.) are not directly measurable. This results in a challenging dual-estimation problem (Singh, Wang, Cole, & Ching, In Press; Tronarp, Subramaniam, Särkkä, & Parkkonen, 2018) for identification (i.e., unknown state and parameter values) as well as state-estimation (e.g. filtering problem) for closed-loop control design. Together, these constraints provide significant challenges for black-box identification/control paradigms.

2.3. Limitations of identification-free frameworks

Due to this complexity, a natural question is whether system-identification is truly necessary. Control approaches that do not require explicit system identification have been used in neuroscience and typically involve tuning a prototype control module, e.g., PID, based upon input-output experiments. Such approaches have previously been employed in other neuroscience applications. At the cellular-level, identification-free feedback control of single-cells has been a staple of neurophysiology for the past century: the voltage clamp, current clamp, patch-clamp etc. techniques (Hamill, Marty, Neher, Sakmann, & Sigworth, 1981; Neher & Sakmann, 1992; Verkhratsky & Párpura, 2014). Feedback control has also been applied to controlling the firing rate of neurons using PID designs (Miranda-Dominguez, Gonía, & Netoff, 2010). However, a common feature of these identification-free scenarios is that: (a) they are amenable to tuning a small number of feedback-control parameters and, (b) the control objectives are fully defined by measurements (i.e., a neuron's membrane potential) as opposed to latent variables (i.e., an underlying neural 'code' or unmeasured state-variables). Moreover, controls of this class are generally designed within a linear range of operation and are tolerant of input constraints. By contrast, human brain stimulation at the whole-brain scale involves strong input constraints. Current technical constraints also prevent continuous real-time feedback control due to the previously-mentioned inability to (noninvasively) measure and stimulate the brain at the same time. These features further suggest that a challenging system-identification stage is necessary for control.

2.4. The difficulty of formalizing cognitive enhancement as a control objective

More significantly, neurostimulation faces ambiguity in the formulation of control objectives. Whereas the framing of objectives is comparatively straight-forward in other control domains (e.g., tracking, disturbance rejection, etc.), the ultimate objective of neurostimulation is cognitive enhancement or, more generally, altering some aspect of cognitive processing as reflected in human behavior. Framing these objectives in quantifiable state-space formulations is nontrivial due, in

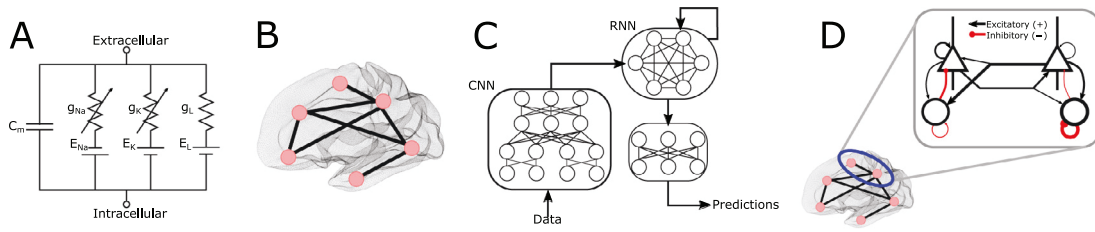


Fig. 2. Levels of modeling in neuroscience. (A) The cellular Hodgkin–Huxley model (Hodgkin & Huxley, 1952) accurately describes the major currents that impact neural membranes. (B) Whole-brain “neural-mass” models describe macroscale (as opposed to cellular) brain activity. Brain connections are typically parameterized from diffusion imaging data which can estimate the amount of wiring between regions, but not direction or sign of influence. Circles indicate different brain regions. (C) Black-box models typically employ artificial neural networks, including combinations of convolutional “encoding” layers (CNN) and recurrent networks (RNN) to generate dynamics. (D) Grey-box models leverage brain-activity timeseries to directly parameterize detailed, mechanistic models of local brain circuitry and brain connections (including direction/sign). Inset shows the microcircuitry model used in Singh et al. (In Press) which models the interaction of pyramidal cells (triangles) and interneurons (circles).

part, to the abstractness of the problem. In particular, there is an obvious but formidable distinction between a motor-output (e.g. pressing a button) and the set of neural computations which led to this output. Whereas the former may have a straightforward conceptualization in terms of state-space targets (exciting the relevant portions of motor-cortex), the latter describes a mapping from the cognitive-task context onto the appropriate responses. In this context, there are two inputs to the system: the task environment and electrical stimulation.

From this perspective, the functional significance of neurostimulation is to imbue favorable properties to the vector-field that support the implementation of a given computation (i.e., produce the correct response given environmental input). Therefore, we propose that the objective of neurostimulation should be to alter the brain’s control properties such that environmental ‘inputs’ will produce brain activity that will push the system towards the correct output. Such a concept is distinct from imposing a prescribed pattern of activity determined *a priori* (Ching, 2018). We proceed to formalize these notions with an emphasis upon altering input–output relationships and control properties, including reachability, within relevant brain circuits and areas.

3. Controlling models of brain dynamics

As a precursor, it is useful to first discuss different ways in which brain dynamics can be described mathematically.

3.1. Modeling and control at the cellular scale

For the purposes of discussion, we dichotomize traditional control design as being either based upon a model of the underlying system, or using model-free approaches. As in other sciences, models have a long history in neurophysiology, with some of the earliest models, Hodgkin–Huxley (Abbott & Kepler, 1990; Hodgkin & Huxley, 1952) and the cable theory (Koch, 2004; Rall, 1962), continuing to be used in cutting-edge research (Fig. 2A). These models are currently viewed as deriving from biophysical first-principles, although the Hodgkin–Huxley equations were originally derived through system-identification. These models, coupled with later descriptions of other voltage-gated channels, give highly accurate predictions for many single-cell phenomena. Significant attention has been directed at analysis and control of neural activity at this spatial scale, including work describing control strategies for one or more Hodgkin–Huxley model neurons (Zlotnik, Chen, Kiss, Tanaka, & Li, 2013; Zlotnik & Li, 2012) and controlling the timing and patterning of spiking activity (Ching & Ritt, 2013; Nandi, Schattler, & Ching, 2019; Nandi, Schattler, Ritt, & Ching, 2017; Narayanan, Ritt, Li, & Ching, 2019). Reduced forms of these models have been used for real-world control of brain cells (Nabi, Stigen, Moehlis, & Netoff, 2013).

3.2. Large-scale (whole-brain) modeling frameworks

At the whole-brain scale, such canonical models do not yet exist. The closest analogues, in this case, consist of ‘neural-mass’ and neural-field models in which mean-field approximations are used to reduce all of the neurons within a brain region into a single state-variable per cell-type (Fig. 2B), e.g., Bojak, Oostendorp, Reid, and Kötter (2010), David and Friston (2003), Dayan and Abbott (2001), Destexhe and Sejnowski (2009), Jansen and Rit (1995), Wang (2002), sometimes coupled with additional population-average molecular/channel dynamics. These models often take one of two forms:

$$x_{t+1} = W\psi(x_t) + g(x_t) + \omega_t; \quad \text{or} \quad (3)$$

$$x_{t+1} = f(x_t)\psi(Wx_t) + g(x_t) + \omega_t. \quad (4)$$

The first form models x_t as a voltage-type variable which is transformed into neuronal “spiking” via the activation function ψ . Neuronal “spiking” is multiplied by connection matrix W and local dynamics are represented by g . In the latter form x_t directly represents the neuronal “spiking” (i.e., x_t in Eq. (4) is analogous to $\psi(x_t)$ in Eq. (3)), so the activation function ψ is applied after multiplication by W . The optional function f applies state-dependent gain. All indicated functions (f, ψ, g) are “stacks” of univariate functions (i.e. $\partial f_i(x)/\partial x_{j \neq i} = 0$). Some variants of these models also include lags, although their necessity is unknown.

Previous work has concerned controlling these models in small networks (sub-circuits of the brain) (Muldoon et al., 2016), including applications in epilepsy (Ching, Brown, & Kramer, 2012; Yaffe et al., 2015) and motor disorders, primarily in Parkinson’s Disease (Duchet et al., 2020; Gorzelic, Schiff, & Sinha, 2013; Santaniello, Gale, & Sarma, 2018; Schiff, 2010). In these applications, model parameters are typically borrowed from previous anatomical and neurophysiological studies.

The analogous whole-brain models also use prior literature to parameterize the internal dynamics of each brain region, while the connection strengths between brain regions are often assumed to be proportional to the number/volume of connections (white-matter) between brain regions (Deco, Tononi, Boly, & Kringelbach, 2015; Honey, Kötter, Breakspear, & Sporns, 2007). These models have achieved some successes in reproducing long-term statistics that are characteristic of brain activity (power spectrum or covariance between brain regions; Hansen, Battaglia, Spiegler, Deco, & Jirsa, 2015; Honey et al., 2007, 2009; Nakagawa, Jirsa, Spiegler, McIntosh, & Deco, 2013) including in tES contexts (Stiso et al., 2019).

However, the development of models that forecast or predict neural activity trajectories at whole-brain scales without identification (e.g., using white matter connectivity) remains an elusive goal. Clearly, such capability to approximate well the underlying vector fields is an important precursor to model-based control synthesis. Indeed, while some control-strategies derived from existing models may prove useful,

significant concerns may arise regarding the accuracy of the ensuing vector fields, and consequently could undermine confidence in the reliability of the synthesized solution as well. As such, there is an unmet need for methods to identify large-scale brain models accurately from brain-activity recordings.

3.3. Specifying modeling and identification frameworks for whole-brain dynamics

Neural systems identification methods can be grouped into black-box and grey-box frameworks. These approaches differ in that black-box models emphasize predictive power (in terms of measurements) whereas grey-box models must balance prediction power with interpretability and model-complexity. From a scientific standpoint, one advantage of grey-box models is the ability to interpret system parameters for basic scientific inferences. For the current discussion, however, we limit our comparison to their potential efficacy for whole-brain control. By ‘grey-box’ model we will refer to model forms with parameters that are explicitly driven by physiology. To make this distinction concrete, we will consider models to be grey-box if at least some of the model parameters (known or unknown) have units other than Hz (i.e., they represent physical or physiological quantities). Importantly, our description of a model as ‘black-box’ does not imply that it is physically uninterpretable, and we begin with a brief review of such techniques.

3.3.1. Black-box system identification

To organize our discussion of black-box models in neuroscience, we consider three classes: machine-learning, input–output models, and statistical descriptions (e.g., auto-regressive models). As in other fields, machine learning approaches for neuroscience are developing rapidly and are dominated by artificial neural networks (Fig. 2C). These approaches have been used for latent factor identification (Pandarinath et al., 2018) and to forecast future brain activity patterns, e.g., predicting epileptic seizures (Truong et al., 2018). Input–output characterizations include the calculation of transfer-functions using electrical stimulation, such as in the estimation of phase-response curves to analyze (Gutkin, Ermentrout, & Reyes, 2005) and design (Dasanayake & Li, 2011; Moehlis, Shea-Brown, & Rabinitz, 2006; Stigen, Danzl, Moehlis, & Netoff, 2009) control of individual neurons. Statistical models (which can also involve input–output descriptions) (Nandi, Kafashan, & Ching, 2017; Saxena, Schieber, Thakor, & Sarma, 2012), involving one or more autoregressive terms, have been used for state-estimation with EEG and fMRI data among other applications.

Black-box models are also useful in predicting nonstationary phenomena such as changes in spectral power or synchronization patterns. Successful applications in this domain typically use optimization methods and include the prediction of changes in (region-specific) spectral power (Sikka et al., 2020) and seizure components from EEG data (Zhang, Guo, Yang, Chen, & Lo, 2019). Due to the long timescales involved, the timing and nature of these phenomena are difficult to predict online using state–space methods (e.g., forward simulation). They may also involve slower molecular mechanisms or other physiological details that are excluded from state–space brain models, such as neural mass models.

In contrast to neural networks, statistical “black-box” models are easy to interrogate and flexible in terms of how “states” are defined. Hidden Markov Models (HMMs; Rabiner & Juang, 1986), for instance, are a class of unsupervised models in which the underlying system obeys a discrete Markov process. Such frameworks are intriguing as they often deal with temporally-extended states which could be defined in the spectral domain or in terms of (auto)covariance (Stevner et al., 2019; Vidaurre, Smith, & Woolrich, 2017; Zhang, Cai, et al., 2019). As such, they are amenable to predicting the probable long-term outcomes of moving the brain into such a state. However, since these models do not operate in state–space they cannot predict *de-novo* how control

signals interact with the system dynamics. As such, HMMs may be useful for considering multiple timescales in the control objective, but cannot, by themselves, identify tES control solutions.

Moreover, from the standpoint of control design, there are key considerations with respect to generalizability of black-box models, and in terms of their compatibility with pertinent control objectives. In considering the generalizability of black-box models, we do not refer to whether the model is over/underfit with respect to the training-data’s domain (which can be assessed via cross-validation). Rather, we refer to the ability of a given model (fit to a subject’s output-only data) to describe the brain’s vector-field in novel regions of state–space. This aspect may be especially important if the objective is to steer brain activity from a pathological state or regime, to one that is benign. In some cases (e.g., epilepsy), patients may be theorized to switch between pathological and non-pathological dynamical regimes. In this case, data may be available for the target state–space regions, although a designed controller may nonetheless steer activity through an unnatural and hence novel region of the state space in attempting to reach the previously-observed non-pathological regime. On the other hand, degenerative brain disorders result in a more persistent pathology so data in the ‘target’ state–space regions may be unavailable. In this case, grey-box models which were less accurate in describing the baseline regions of state–space may still outperform black-box methods in generalizing to novel regions of state–space, purely by virtue of their physiological constraints and interpretability.

Lastly, black-box methods are inherently tied to the measurement-space. In other words, for the measurement matrix C (as in Eq. (1)) the state–space models of such systems are restricted to the form:

$$z_{t+1} = f(z_t); \quad z_t \in \text{span}(C), \quad \forall t \quad (5)$$

as opposed to engaging with the full state–space as in Eq. (1). This aspect means that it is harder to treat objectives that are not defined in the measurement space, such as those featuring latent state variables or properties of the underlying vector-field. In Sections 4 and 5 we proceed to discuss such objectives in more detail and describe their potential importance for emerging applications in brain stimulation for cognitive enhancement.

3.3.2. Grey-box models and numerical advances

Grey-box system identification methods have a long history in cellular neuroscience, starting with the aforementioned work of Hodgkin and Huxley using input–output identification on the squid giant axon (Hodgkin & Huxley, 1952). Current efforts in cellular grey-box identification emphasize joint-estimation, typically with only the membrane voltage or firing rate directly measurable (Doi, Onoda, & Kumagai, 2002; Lankarany, Zhu, & Swamy, 2014). Grey-box identification has also been performed with small network models (Cullen, Rey, Guitton, & Galiana, 1996; Friston, Harrison, & Penny, 2003) and in larger-scale brain models. In the latter case, efforts have been divided into approaches which seek to refine parameterized models and those which perform identification with little or no assumptions outside of a general mathematical form. In the former case, neural-mass models have had a very small set of parameters tuned in order to better replicate long-term statistics such as correlation patterns (Demirtaş et al., 2019; Jirsa et al., 2017; Schirner, Rothmeier, Jirsa, McIntosh, & Ritter, 2015; Wang et al., 2019). However, since these models are fit to long-term statistics, as opposed to directly forecasting brain activity, it is not yet clear whether they are sufficiently accurate for control synthesis, particularly over short timescales. Other approaches, however, have been developed to fit single-subject brain models directly to brain activity using neuroimaging data (fMRI, EEG, MEG, etc.; Fig. 2D). Dynamic Causal Modeling (DCM; Friston et al., 2003) is a prominent case of using Bayesian methods for the latter approach. DCM models consist of linear state–space models for fMRI (Frässle et al., 2017, 2021; Friston et al., 2003; Friston, Kahan, Biswal, & Razi, 2014; Razi, Seghier, Zhou, & McColgan, 2017), while faster-timescale data (MEG and EEG)

is estimated using a linear approximation of an underlying nonlinear neural-mass model (David et al., 2006; Kiebel, Moran, & Friston, 2008). The DCM methodology is an example of joint-estimation for latent variables in that DCM models estimate latent neural activity from fMRI data and the activity of non-pyramidal cell-types from EEG data (which primarily derives from pyramidal cells).

A limitation of the existing DCM methodology, however, has been high computational complexity which has limited its ability to scale to whole brain models (Singh, Braver, Cole, & Ching, 2020), although this barrier has been recently reduced under assumptions of linearity in the dynamics (Frässle et al., 2021). For models with high temporal resolution, however, analyses have been limited to a small number of brain regions (usually less than 10) with the assumption that spatially-localized (source-level) signals are available (i.e., it does not directly treat the issue of spatial mixing; David et al., 2006). To bridge this gap, our group has recently developed a number of techniques for highly-scalable joint-estimation of neural systems (i.e., unknown states and parameters). These include the ability to quickly estimate high-dimensional (419) brain models (which take the form of Eq. (3)) using output-only fMRI data (Singh, Braver, et al., 2020; Singh, Wang, Braver, & Ching, 2020) and to identify environmental and endogenous drivers of brain activity as reflected in fMRI (Singh, Wang, Cole, Ching, & Braver, 2021). Most recently, however, we have developed a general, scalable framework for joint-estimation that also treats the issue of signal-mixing via a Kalman-Filtering technique. This advance has enabled us to estimate whole-brain neural-mass models (100 regions $\times$ 2 cell-types) directly from single-subject MEG data (Fig. 2D). We anticipate that these advances in systems identification will facilitate corresponding improvements in neurocontrol technology.

4. Current frameworks for brain stimulation and their limitations

Traditional control objectives include control-to-point/set, tracking, stabilization, and synchronization, potentially including additional penalties (minimum energy, time etc.). By contrast, the formulation of objectives for human brain stimulation is less straightforward. While the desired outcomes can be concisely described in terms of psychological constructs (e.g., improving working memory) and quantified in terms of behavior (maximize the number of items that can be recalled from a list with some level of accuracy) formulations in terms of brain activity are more challenging. A key feature of this problem is that while noninvasive brain stimulation induces anatomically broad effects (relative to sub-millimeter micro-circuitry), the neuronal microcircuits underlying a given computation are often spatially diffuse and less well understood. Linking these scales formally results in a control-problem of the form:

$$\hat{x}_{t+1} = f(\hat{x}_t, V^T B u_t, \hat{\omega}_t) \quad (6)$$

$$y_t = C V \hat{x}_t + \eta_t \quad (7)$$

in which the matrix V adds together the contribution of cells (x_t) in each brain area (i.e., $x_t = V \hat{x}_t$), while the macroscale matrices C, B are the same as in Eq. (1). The symmetries generated by V obviously impede treatment of the underlying microscale system $\hat{x}$ by macroscale controller u_t .

4.1. Case study I: Random dot motion

4.1.1. Increasing activity is not sufficient to augment performance

This mismatch has significant impact on how objectives should be defined. To illustrate this point, consider a popular test of perceptual decision making: the Random Dot Motion task (Born & Tootell, 1992; Van Maanen et al., 2012). In this task, subjects are visually presented with a patch of particles/dots moving according to a random walk with slight drift in one direction (i.e., the dots' motion is mostly random, but weighted towards one direction). Subjects must decide which of

two options (e.g., left/right) is the correct direction. Empirical and computational studies of this task have converged around a model in which each option (direction) is represented by a particular set of neurons which integrate noisy input ('evidence') elicited by objects moving in that direction (Purcell et al., 2010; Ratcliff, 1978; Ratcliff, Smith, Brown, & McKoon, 2016; Usher & McClelland, 2001). The subject's decision corresponds to the first option (direction) to pass a threshold level of activity (the phenomenology of such integration is often modeled as an Ornstein-Uhlenbeck, or drift-diffusion, process).

Under a conventional tDCS perspective, the stimulation objective should be to increase the brain activity within that region since gross neural activity in the region (which contains the neurons for all directional choices) increases when performing the task. However, the model predicts that when the input to both options (neuronal populations) is increased, decisions will be faster but less accurate (the speed-accuracy tradeoff). Thus, direct-current stimulation is not predicted to enhance the subject's perceptual decision making *per se*, but rather effects would be expected to alter the degree to which they weight speed vs. accuracy in making a decision (which can already be accomplished by telling the subjects to wait longer/shorter; de Hollander et al., 2016). Analogous arguments can be made for other brain circuits. Thus, we argue that increasing neural excitability within a given brain area is not, in general, a sufficient strategy to improve cognition. This is not to say that direct-current stimulation is totally ineffective, as it may produce desirable effects related to increases in the spread/duration of neuronal activity or promoting synaptic plasticity (learning). However, the key point is that some cognitive functions may require a more nuanced objective formulation.

4.2. Using dynamic stimulation to modify information propagation

A natural alternative to tDCS is tACS which instead uses alternating current (typically sinusoidal). The objective of tACS stimulation is to increase the amplitude of a particular frequency band in the targeted brain area or to entrain neural activity to a desired phase, such as synchronizing two regions. There are many theories regarding the role of neural oscillations in computation (Kopell, Gritton, Whittington, & Kramer, 2014; Sejnowski & Paulsen, 2006). One premise is that brain oscillations act as carrier waves (Ainsworth et al., 2012; Freeman & Barrie, 1994; Fries, 2005). In this regard, information is transmitted via phase, amplitude, frequency-modulation or, in other scenarios, by the relationship between different frequency-bands or between the macroscale oscillations and cellular ('spiking') inputs (Jacobs, Kahana, Ekstrom, & Fried, 2007; Smith et al., 2019; Tort, Komorowski, Eichenbaum, & Kopell, 2010). These mechanisms selectively engage micro-circuits which are especially resonant to the oscillation frequency or to the phase-offset (delay) between events (e.g., the crest/trough of inputs from different sources, Bastos, Vezoli, and Fries (2015), Fries (2005)). Under this framework, AC stimulation is thought to act by increasing the oscillation's signal strength without modifying the information it conveys (Ali, Sellers, & Fröhlich, 2013; Huang et al., 2021; Reato, Rahman, Bikson, & Parra, 2010), such as by increasing the amplitude of a reference wave when information is coded by phase-coupling. Returning to the Random Dot Motion task (see previous section), a recent study (Gong & Fang, 2018) found that participant's accuracy increases when stimulated at 10 Hz (the dominant frequency in visual areas), which is interpreted as increasing the quality of information sent to decision making circuits, since non-selectively increasing signal strength (or quantity) would decrease accuracy (see previous section).

Multiple studies have indicated that tACS can be efficacious in partially restoring cognitive function when applied in a personalized, phase-locked manner. An especially impressive study, by Reinhart and colleagues (Reinhart & Nguyen, 2019), demonstrated that memory performance in older adults could be improved to levels comparable to younger counterparts by employing theta-frequency (i.e., 7–10 Hz) tACS phase-locked to brain activity. This study further indicated that

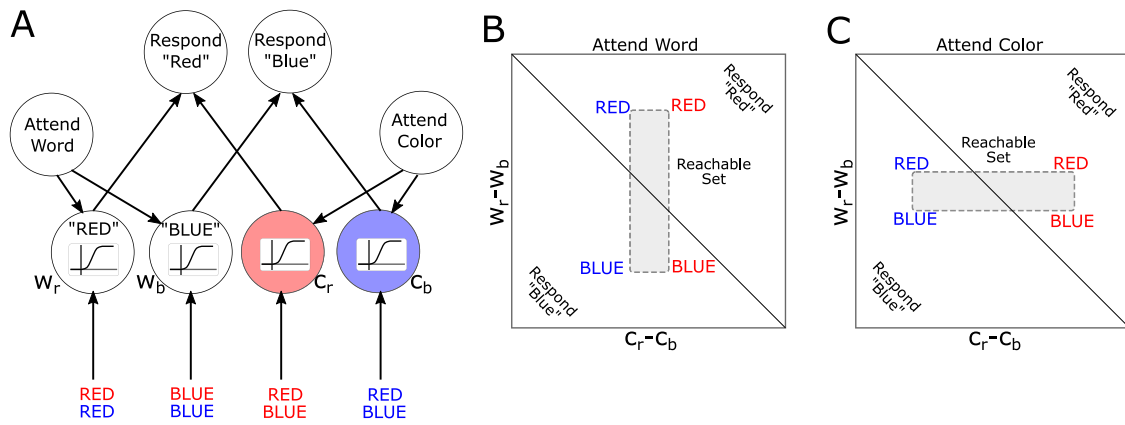


Fig. 3. Schematic of reachability framework for cognitive control in the Stroop Task. (A) The Guided-Activation Theory model of the Stroop-Task as a feedforward neural network (Cohen, Dunbar, & McClelland, 1990; MacLeod & Dunbar, 1988). Task instructions (attend to word vs. color) promote specific pathways through which inputs propagate. (B) Activation of the latent ‘attend word’ state shapes the reachable set so that the response-boundary separates inputs by word. (C) Analogous figure for the ‘attend color’ condition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

applying tACS in an individualized manner, at each subject’s spontaneous peak-frequency, further increased the magnitude of neural effects. However, despite these successes, there are several theoretical reasons to believe that these results could be further improved. Recent tACS studies, such as the aforementioned, have demonstrated significant potential to partially restore function in under-performing groups, which have been characterized by altered frequency-domain properties. However, the comparatively weak evidence for improvement in healthy subjects may imply that tACS methods primarily serve to restore degraded features of signal transmission (e.g., carrier/reference signals) rather than improving the brain’s ability to operate upon these features. This approach is also based upon the assumption that sinusoidal stimulation does not alter the content of signals. This assumption is likely violated if information is coded in signal frequency. Likewise, when information is coded in signal phase, this assumption requires that the stimulation input does not alter the phase (which explains the advantage of phase-dependent stimulation in Reinhart and Nguyen (2019)). More generally, these stimulation frameworks are based on neural computation models in which temporal dynamics are used to encode/decode information as it is propagated between brain areas, but does not enable a general framework in which neural dynamics are the underlying mechanism of computation.

5. A proposal for formulating control-theoretic objectives

We now explore the potential of control-theoretic methods to reformulate the optimization objective. Our framing involves a change of viewpoint to emphasize facilitating the endogenous ability of the brain to reach target cognitive states as opposed to forcing the brain along a particular time-varying trajectory. In other words, we propose an approach to exogenously modifying brain activity that is premised on leveraging endogenous neural control mechanisms.

5.1. Case study II: Stroop task

5.1.1. Lessons from endogenous control mechanisms

In neuroscience and psychology, ‘cognitive control’ refers to the diverse set of processes that modify human behavior based upon the environmental context and goals (Braver, 2015). Impairment of cognitive control is a core component of many psychiatric disorders (Cohen, Braver, & O’Reilly, 1996). Prominent theories of cognitive control stress its modulatory role in determining how information is processed, interacts with internal states, and affects actions (Braver, 2012). This ability is thought to be accomplished by altering how the brain interacts with inputs from the periphery. Brain regions (e.g., prefrontal cortex,

PFC), which maintain rule representations, are believed to proactively bias the neural pathways along which information flows (Guided Activation Theory; Cohen, Aston-Jones, & Gilzenrat, 2004; Miller & Cohen, 2001).

The above theory has been used to develop neural network models of classic cognitive tasks such as the Stroop Task (Stroop, 1935), in which subjects are asked to report either the ink-color or word of colored text (e.g., ‘Red’ in blue font; see Fig. 3A). Over short timescales, these instructions are believed to be maintained in a stable ‘read-only’ state associated with the modulatory neurotransmitter Dopamine (Braver & Cohen, 1999, 2000; Durstewitz, Seamans, & Sejnowski, 2000). When the current behavioral strategy is ineffective, neural activity in PFC is thought to destabilize, enabling new information to be encoded (Fallon, Williams-Gray, Barker, Owen, & Hampshire, 2013; Rougier, Noelle, Braver, Cohen, & O’Reilly, 2005). Once an effective approach is found, PFC dynamics stabilize around this activity pattern, thereby storing the associated behavioral strategy (Rougier et al., 2005). More generally, brain activity has been found to stabilize along a more limited set of patterns in response to sensory stimuli and cognitive tasks (Churchland et al., 2010; Ito et al., 2020; Ponce-Alvarez, He, Hagmann, & Deco, 2015). These conceptual models suggest an alternative route towards cognitive enhancement. Rather than altering the frequency-profile of brain activity, a framework emerges that seeks to modulate how environmental inputs interact with the brain by altering the stability properties and reachable-sets (see Fig. 3B,C) of local brain dynamics.

5.2. Case study III: The attentional blink

5.2.1. Relevant control-theoretic measures

There are thus two levels to the control problem: first identifying the local control properties that are advantageous to cognition and then designing a stimulation protocol that will (locally) promote these properties on the augmented stimulator-brain system. To be concrete, we will present these issues in terms of another well-studied cognitive task and phenomenon: the attentional blink.

The ‘attentional blink’ refers to the fact that when two visual images are rapidly presented in sequence (200–500 ms delay between images presented for 100 ms each), most subjects (>50%) are unable to report the second image (Shapiro, Raymond, & Arnell, 1997). Interestingly, this effect is non-monotonic, as subjects can process both stimuli for extremely short delays (e.g., 100 ms) and subconscious traces of information present in the second image (e.g., the meaning of word images) are evident in later behavior. These effects suggest that the attentional blink reflects changes in the system’s (i.e., the brain’s) internal states as opposed to limitations in the processing-speed for visual input.

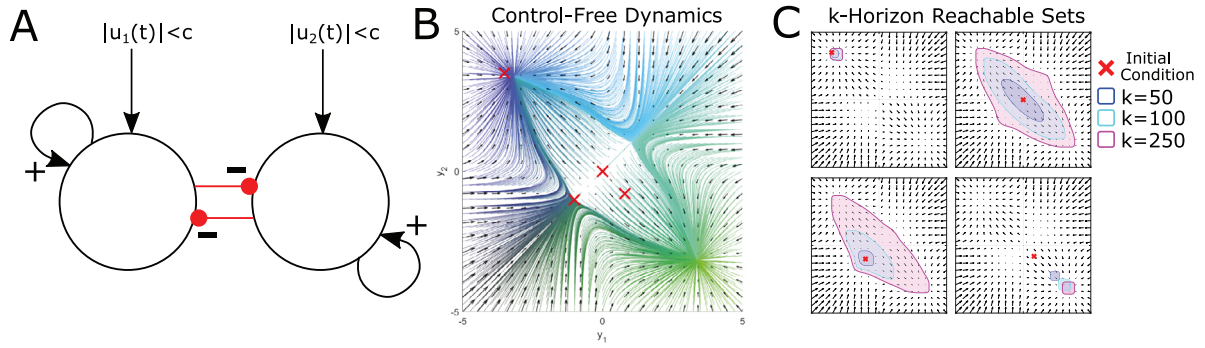


Fig. 4. Illustration of state-dependent reachability. (A) Schematic of recurrent two-neuron network with self-excitation, reciprocal inhibition, and two independent, bounded, inputs. (B) Uncontrolled vector field and integral curves (color = initial condition for each curve). (C) Approximate reachable sets for different initial conditions and time-steps (horizons). Note the strong dependence of initial condition and that reachable sets are larger when starting near a separatrix (e.g., top right) than a fixed point (e.g., top left). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

From a stimulation design perspective, we are interested in how a control paradigm could be designed to eliminate or attenuate the attentional blink, as a means to improve visual attention. In this regard, several attempts have been made to attenuate the attentional blink using DC brain stimulation, yet with inconsistent findings (London & Slagter, 2015; Reteig, Newman, Ridderinkhof, & Slagter, 2021). Limited prior study is currently available concerning AC stimulation. We hypothesize that the attentional blink and similar cognitive bottlenecks manifest changes in the brain's local reachability properties for visual input. More mathematically, for an initial state-space condition x_0 , a class of admissible control-inputs (i.e., sensory stimuli) $\mathcal{U}$, and a realization of the noise-process ω_t , we define the k -horizon reachable set as a set-valued function of the initial condition x_0 and random process noise ω_t :

$$\mathcal{R}(f, x_0, \omega_t, \mathcal{U}) := \bigcup_{v_i \in \mathcal{U}} \phi_f(x_0, \omega_t, v_i, k) \quad (8)$$

Here, ϕ denotes the evolution function for vector field f , which could be formulated in continuous or discrete time. As a deterministic function of the process ω_t , the reachable-set given ω_t is well-defined, even if the distribution of all $\mathcal{R}$ is unbounded due to an unbounded distribution of ω . Returning to the attentional-blink example, we propose formalizing brain sensitivity to incoming visual information using the 'size' of the reachable space according to a quantity, denoted μ , which captures biases in the vector field toward movement in task-relevant directions (e.g., measuring the degree to which the second image can elicit responses in visual brain areas). Direct measures based upon well-established geometric nonlinear control analyses (Brockett, 1978; Sussmann, 1985) are natural here, but bring about concerns of analytical or computational tractability for high dimensional models. Indeed, reachability and similar control-theoretic assessments are exceptionally challenging, particularly in the presence of multiple inputs. Several metrics have been developed for linear systems (Pasqualetti, Zampieri, & Bullo, 2014), which can be applied to infer properties of linear whole-brain models, including structural brain networks (Gu et al., 2015). Such approaches could potentially be used in a state-dependent manner to perform local reachability estimates for nonlinear systems in the vicinity of an attractor under sufficiently weak input and short control horizons. We later discuss more general techniques (see Section 5.3).

For non-local characterizations, however, the reachable set is not convex, and admits multiple ways to describe its breadth. However, a simple description of the reachable set may still be attainable via a metric induced by a quadratic form (i.e., $\mu[X] = \mu_L[AX]$ for a matrix A and the Lebesgue metric μ_L) with the corresponding objective to minimize/maximize $\mathbb{E}_\omega[\mu(\mathcal{R}(x_0, \omega, \mathcal{U}))]$. By a quadratic-norm we are only referring to how the "size" of a reachable set is measured as opposed to implying a quadratic (e.g., ellipsoidal) approximation of the reachable set.

The matrix A should be constructed to have large singular-values in directions in which movement is desirable and small singular-values for directions which contribute noise (or, alternatively, subtract reachability in these directions). Changes in neural activity within the default-mode network (Fox et al., 2005; Kucyi & Davis, 2014; Raichle, 2015), for instance, are associated with mind-wandering (Kajimura, Kochiyama, Nakai, Abe, & Nomura, 2016), so one construction of the A matrix for the attentional blink would be to have heavily weighted singular-vectors spanning visual-attention pathways, and weakly-weighted (right) singular vectors spanning state-coordinates associated with the default-mode network. This would allow measurement of whether the second image is able to differentially induce representations in task-relevant (i.e., visual) vs. irrelevant (i.e., DMN) areas. In our formulation, the matrix A is specified a-priori based upon existing knowledge of the underlying neuroanatomy. Another option, may be to refine pre-existing assumptions on A based upon which directions are easiest to reach in-practice using an empirical estimate of the Gramian (Himpe & Ohlberger, 2014). The empirical Gramian can be useful in identifying directions which are (effectively) inaccessible, in which case they should be removed from A .

The reachable set also depends upon the space of admissible controls ($\mathcal{U}$). For example, only a small subset of potential sensory input signals to visual cortex are likely. Further study will be needed to identify this set (Singh, Wang, Cole, Ching, & Braver, 2021). However, a preliminary estimate for $\mathcal{U}$ could be formed via a linear-in-control assumption with corresponding bounds ($[a, b]$) on input, i.e., mathematically,

$$\mathcal{U} := \{Mv(t)|v_i : [0, k] \rightarrow [a_i, b_i]\} \quad (9)$$

We note that the control matrix for modeling endogeneous control (M) is not, in general, the same matrix used for exogeneous control (denoted B in Eq. (1)). In the case that the exogeneous (u_t) and endogeneous ($v_t \in \mathcal{U}$) controls add, the reachability-optimal control u is the (constrained) solution to:

$$\arg \max_{u \in \mathcal{F}_y} \mathbb{E}_\omega \left[\mu_L \left(A \left[\mathcal{R}_f(x_0, \omega_t, Bu + M\{v\}) \right] \right) \right] \quad (10)$$

with $M\{v\}$ denoting the set of admissible endogenous inputs and $u \in \mathcal{F}_y$ indicating that the control-solution can, potentially, be a policy on the filtration of y (i.e., past measurements). Importantly, because reachability properties depend upon a model of the brain as a control-system, these objectives are inherently causal in nature: they require specification of which neural populations are the subject of endogenous/environmental control and how changes in their activity under such control relates to cognition. As such, reachability-based control objectives are specific to the cognitive function being targeted and likely involve trade-offs such that there is likely not a universal reachability objective that benefits all aspects of cognition.

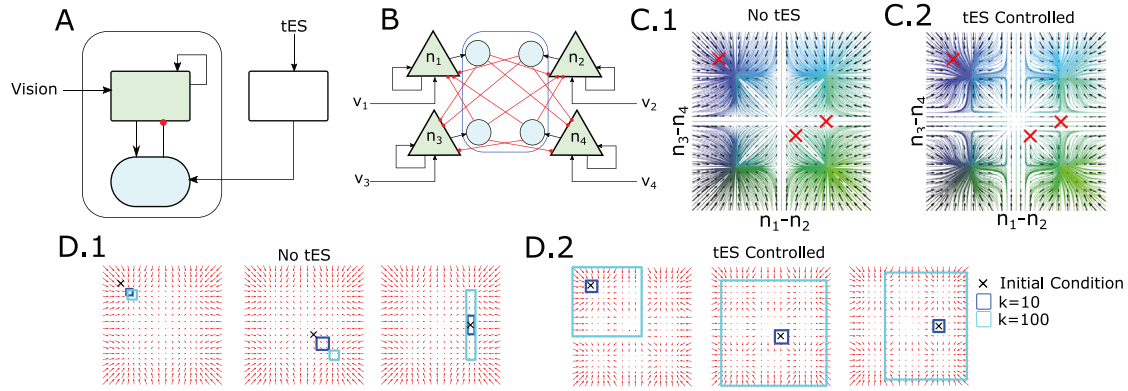


Fig. 5. Schematic for controlling computation-relevant (microscale) reachability via macroscale stimulation. (A) Macroscale system schematic: visual input enters a brain region containing excitatory (green) and inhibitory (blue) neurons. Electrical stimulation (tES) is applied to a separate brain area which connects to inhibitory cells. (B) Microscale architecture consists of self-excitation and lateral inhibition (e.g., representative of sensory cortex). Black connections are excitatory and red connections are inhibitory. (C.1) Conceptual 2D representation of the vector field without additional input (C.1) and with tES-induced macroscale disinhibition (C.2) leading to reduced activation of inhibitory (blue) cells. (D) Reachable sets from 3 initial conditions due to sensory input: $v_1, \dots, v_4$ in (B). (D.1) K-horizon reachable sets Without tES. (D.2) With tES induced disinhibition. Decreasing the inhibition between excitatory (green) cells results in larger reachable sets, hence visual inputs can more easily switch activity between attractors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.3. Identifying reachable sets for brain dynamic models

As alluded to above, for a general nonlinear control system, the calculation of finite-time reachable sets is highly nontrivial, although a growing number of numerical techniques and freely-available tool-boxes (Althoff, 2015; Meyer, Devonport, & Arcak, 2019) offer some potential for relevant approximations. When no assumptions are placed upon the system vector field, numerical methods typically need to simulate a very large number of potential inputs to estimate the reachable set or using iterated geometric methods, such as zonotope-based analyses (Althoff & Krogh, 2011; Girard, 2005).

Fortunately, canonical models of macroscale neural activity generally feature monotone relations (e.g., a region's output is a bounded, monotone function of its current state), which enable much more efficient approaches such as mixed-monotonicity algorithms (Abate & Coogan, 2021; Meyer, 2021). A primary challenge in computing reachable sets is the ability to efficiently represent non-convex sets, particularly those which are not simply-connected. These considerations are important for the reachability analysis of neural systems which often generate limit-cycles or quasi-periodic attractors that under small control perturbations generate infinite-horizon reachable sets that are not simply-connected.

Mutual competition is a fundamental principle of the nervous system which promotes winner-take-all dynamics (Maass, 2000; Usher & McClelland, 2001). This feature generates a large number of hyperbolic fixed points (Burylko, Kazanovich, & Borisyuk, 2018), which correspond to when competing populations are perfectly balanced, and the associated unstable manifolds that arise under these conditions. This case brings an important consideration as to whether reachable sets for the brain's internal control laws should take into account the natural imprecision of biological systems. Whereas the analytic reachable set in this case may be topologically-connected (e.g., by traversing an unstable manifold), a small deviation from this manifold could be impossible to later correct with a bounded, delayed control signal. As a result, the set of achievable outcomes, in practice, would be disconnected and significantly smaller than the reachable set. Further study is therefore needed to explicate the properties of endogenous brain control, and exogenous control design may benefit from further development in the theory of reachable sets that consider issues of robustness and uncertainty in constraining the set of admissible endogenous controls.

5.4. Potential control strategies for reachability properties

The second level of formulation is designing the exogenous controller with the aim of modifying (and optimizing) reachability properties for endogenous control. We postulate two pathways in this

regard: either (i) by preemptively shifting brain activity to a desirable point/set in the state space, or (ii) by altering the vector field itself. The first case corresponds to an boundary value type problem, but with a new criterion being used to define the desired state (i.e., the state that optimizes reachability properties; Fig. 4). Future work will be needed to determine efficient means of performing the reachability optimization, although naïve first estimates may be possible by assuming that endogenous inputs are sufficiently small to admit local (time-varying) linearization about the uncontrolled orbits. This assumption would lead to an optimization problem in terms of the sequence of Jacobians (maximizing/minimizing the rate of expansion along certain directions).

An alternative to this approach is modifying the brain's vector-field (Fig. 5), which has more in common with stabilization techniques. This pathway is potentially more interesting as it allows switching the reachability objectives mid-task, such as shifting between periods of high reachability (flexible brain activity) and low reachability (more robust activity; Fig. 5B,C). However, design strategies in this space of problem are usually predicated on the use of feedback. As previously mentioned, instantaneous feedback is not possible for tES as stimulation artifact greatly limits the ability to collect simultaneous measurements, although technological developments may eventually alleviate this problem by enabling sufficiently fast switching between stimulation and recording states (approximating simultaneous stimulation and measurement). Since the present technology is limited to, at most, intermittent measurements during the stimulation period, another pathway is to model the controller as a deterministic, autonomous dynamical system and analyze reachability (with respect to endogenous control) in the augmented brain-tES space (combining state-variables). This formulation is compatible with open-loop control, or using intermittent measurements for 'resetting' the initial conditions for the tES controller's state-space. In either case, these discussions make clear that there is substantial room for control-theory to co-evolve with neurostimulation, as optimizing non-standard control objectives (reachability properties) may require new forays in basic control theory.

5.5. Evaluating potential control strategies in large-scale models

The combination of the control-theoretic premises outlined above, with aforementioned large-scale brain models, provides an intriguing platform for theoretical and computational research. For example, our recent work (Singh & Ching, 2018) has explored how exogenous inputs can be used to shape network vector fields, which can in turn impact

the input–output properties. The numerical advances in enabling grey-box models such as MINDy (Singh, Braver, et al., 2020; Singh, Wang, et al., 2020; Singh et al., In Press; Singh, Wang, Cole, Ching, & Braver, 2021) will provide an *in silico* test-bed for exploring and evaluating the efficacy of control designs, as a precursor to implementation in actual experiments. Reciprocally, these experiments will provide more data to refine models, which in turn will validate hypotheses regarding the dynamical substrates of cognitive function and iterate the development of relevant control objectives.

6. Multiscale implications of macroscale control

In previous sections we have emphasized stimulation for cognitive enhancement which, by definition, involves improving the ability of the brain to solve tasks as opposed to only transmitting information into the brain. However, brain stimulation also finds use in neuroprosthetics and brain–machine interfaces. These applications will especially benefit from the development of controllers capable of manipulating neural microcircuitry due to the information-rich content being transmitted (e.g., for sensory prosthesis to generate percepts Lee et al., 2018). The application of formal control methods to such problems will require consideration of how information is stored/communicated within the target brain areas.

6.1. Brain neurophysiological considerations

The degree and manner of neuronal organization varies dramatically between brain areas and this may also have significant implications for control design. Neurons in the human auditory cortex are anatomically organized according to their maximally-sensitive auditory frequency (tonotopy) (Romani, Williamson, & Kaufman, 1982) whereas neurons in early visual cortex are anatomically organized according to positions in visual space (retinotopy) (Tootell et al., 1998) and those in another set of other brain areas are organized according to body parts (somatotopy). Analogous anatomical organizations exist for certain motor responses (e.g., eye movements/saccades; Schall, Morel, & Kaas, 1993), positions in space (O’Keefe & Dostrovsky, 1971), head orientations (Knierim, Kudrimoti, & McNaughton, 1995) etc. These organizations arise due to both pre-programmed neural and biochemical factors (concentration gradients) and environmental structure (spatial autocovariance). By contrast, neuronal sensitivity to more abstract concepts, such as different object-categories, do not exhibit analogous spatial structure.

We also note that neuronal organization can occur in multiple ways. In visual areas, the sensitivity of neurons is largely driven by the specificity of brain wiring such that regions sensitive to line orientations within a region of visual-space primarily receive input from lower brain areas which represent light intensity at different ‘pixels’ within that region of visual space (Hubel & Wiesel, 1962). By contrast, evidence indicates that neurons also separate information based upon the temporal-content of incoming signals potentially due to heterogeneity in their integration time-constants. This feature has been described as a ‘Laplace transform’ of input to those cells (Howards & Hasselmo, 2020), which has been hypothesized to underlie the representation of temporal contingencies (Howards & Eichenbaum, 2013). This operation maps the time-domain of input to the (complex) frequency-domain for storage with cells sensitive to different frequencies.

6.2. Heterogeneity of neural dynamics

Regional heterogeneity and structure are thus important considerations for brain-control synthesis. The control of under-actuated systems (e.g., the brain) is intimately tied to the ability to exploit asymmetries in dynamics, states, and/or input strength (Whalen, Brennan, Sauer, & Schiff, 2015). Noninvasive methods for measuring brain activity are spatially coarse which presents limited ability to differentiate the activity of different cell-groups. Fine-grained brain control

will thus depend upon the ability to leverage neuronal heterogeneity and/or differential stimulation of cell-groups. Neuronal heterogeneity in temporal-sensitivity offers a pathway to neurocontrol using current technology (Ching & Ritt, 2013; Ritt & Ching, 2015). In this case, it may be possible to selectively recruit cell groups without spatially-precise stimulation. By contrast, groups of homogeneous cells will require differential input for control, which is most easily exploited for brain regions with spatial organization.

One further possibility may be the coupling of anatomically-coarse brain stimulation with sensory inputs. For instance, the brain’s representation of semantic categories, e.g. ‘mammal’, are thought to involve distributed patterns of neuronal activity as opposed to a single, concentrated group of cells. It is likely to be impractical to attempt activating the ‘mammal’ concept purely using transcranial stimulation alone. The task is trivial, however, using sensory input (‘display the word ‘mammal’’). Neurocontrol methods may therefore benefit from a hybrid approach, in which electrical stimulation is paired with sensory inputs. Primitive variations of this approach have long been used in attempting to shape brain plasticity to improve learning and are now being explored for clinical desensitization in treating phobias (Cobb et al., 2021) and PTSD (Cassani, Novak, Falk, & Oliveira, 2020; van’t Wout-Frank, Shea, Larson, Greenberg, & Philip, 2019). Recent studies have demonstrated that acoustic stimuli can be used to entrain brain oscillations to desired frequencies with associated improvements in task performance (Beauchene, Abaid, Moran, Diana, & Leonessa, 2016; Stanyer et al., 2021; Will & Berg, 2007). Approaches which leverage the neuronal specificity of sensory stimulation with the temporal resolution of tES (e.g., Kar and Krekelberg (2014)) may thus be especially fruitful.

7. Conclusion

Emerging neural technology and advances in numerical optimization are presenting opportunities for neural engineering. High-resolution brain models (100s of neural populations) are now being identified from noninvasive functional brain data (e.g., EEG, MEG, fMRI) (Frässle et al., 2017, 2021; Singh, Braver, et al., 2020; Singh et al., In Press), enabling the precise study of personalized brain dynamics. This new generation of brain modeling provides fertile ground in which control theoretic methods will provide a foundation for brain stimulation design and delivery. This intersection also provides opportunity to rethink the objectives of brain stimulation in terms of the underlying latent dynamics, as opposed to operating only within the space of measurements. We have suggested several ways in which control-theoretic methods may be used in the framing of neurostimulation objectives and in particular, to optimize reachability properties of the brain’s endogenous control properties, which reflect how environmental information impacts brain activity. This perspective, coupled with the brain’s complexity, provides new challenges for formal and numerical control theory. This scenario offers opportunities for mutual growth and therapeutic success in treating neurological and psychiatric disorders.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially supported by grants 1653589 (to SC) and 1835209 (to SC and TB) from the National Science Foundation, and grant T32 DA007261-29 (to MS) from the National Institutes of Health.

References

- Abate, M., & Coogan, S. (2021). Improving the fidelity of mixed-monotone reachable set approximations via state transformations. In *2021 American control conference* (pp. 4674–4679). IEEE.
- Abbott, L. F., & Kepler, T. B. (1990). Model neurons: from Hodgkin-Huxley to Hopfield. In *Statistical mechanics of neural networks* (pp. 5–18). Springer.
- Agrawal, A. K., Fujino, Y., & Bhartiya, B. K. (1993). Instability due to time delay and its compensation in active control of structures. *Earthquake Engineering & Structural Dynamics*, 22(3), 211–224.
- Ainsworth, M., Lee, S., Cunningham, M. O., Traub, R. D., Kopell, N. J., & Whittington, M. A. (2012). Rates and rhythms: A synergistic view of frequency and temporal coding in neuronal networks. *Neuron*, 75(4), 572–583.
- Ali, M. M., Sellers, K. K., & Fröhlich, F. (2013). Transcranial alternating current stimulation modulates large-scale cortical network activity by network resonance. *Journal of Neuroscience*, 33(27), 11262–11275.
- Althoff, M. (2015). An introduction to CORA 201. In *Proc. of the workshop on applied verification for continuous and hybrid systems*.
- Althoff, M., & Krogh, B. H. (2011). Zonotope bundles for the efficient computation of reachable sets. In *2011 50th IEEE conference on decision and control and European control conference* (pp. 6814–6821). IEEE.
- Antonakakis, M., Schrader, S., Aydin, U., Khan, A., Gross, J., Zervakis, M., et al. (2020). Inter-subject variability of skull conductivity and thickness in calibrated realistic head models. *Neuroimage*, 223, Article 117353.
- Antonenko, D., Grittner, U., Saturnino, G., Nierhaus, T., Thielscher, A., & Flöel, A. (2021). Inter-individual and age-dependent variability in simulated electric fields induced by conventional transcranial electrical stimulation. *NeuroImage*, 224, Article 117413.
- Bai, Z. (2002). Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics*, 43(1–2), 9–44.
- Bastos, A. M., Vezoli, J., & Fries, P. (2015). Communication through coherence with inter-areal delays. *Current Opinion in Neurobiology*, 31, 173–180.
- Beauchene, C., Abaid, N., Moran, R., Diana, R. A., & Leonessa, A. (2016). The effect of binaural beats on visuospatial working memory and cortical connectivity. *PLoS One*, 11(11), Article e0166630.
- Benner, P., Gugercin, S., & Willcox, K. (2015). A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4), 483–531.
- Bojak, I., Oostendorp, T. F., Reid, A. T., & Köster, R. (2010). Connecting mean field models of neural activity to EEG and fMRI data. *Brain Topography*, 23(2), 139–149.
- Born, R. T., & Tootell, R. B. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature*, 357(6378), 497–499.
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, 16(2), 106–113.
- Braver, T. S. (2015). *Motivation and cognitive control*. Routledge.
- Braver, T. S., & Cohen, J. D. (1999). Dopamine, cognitive control, and schizophrenia: The gating model. In *Progress in brain research*, vol. 121 (pp. 327–349). Elsevier.
- Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In *Control of cognitive processes: attention and performance XVIII*, no. 2000. Citeseer.
- Breiten, T., & Damm, T. (2010). Krylov subspace methods for model order reduction of bilinear control systems. *Systems & Control Letters*, 59(8), 443–450.
- Brockett, R. W. (1978). Feedback invariants for nonlinear systems. *IFAC Proceedings Volumes*, 11(1), 1115–1120.
- Burylko, O., Kazanovich, Y., & Borisjuk, R. (2018). Winner-take-all in a phase oscillator system with adaptation. *Scientific Reports*, 8(1), 1–24.
- Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13(6), 407–420.
- Cassani, R., Novak, G. S., Falk, T. H., & Oliveira, A. A. (2020). Virtual reality and non-invasive brain stimulation for rehabilitation applications: A systematic review. *Journal of Neuroengineering and Rehabilitation*, 17(1), 1–16.
- Ching, S. (2018). Control-theoretic approaches for modeling, analyzing, and manipulating neuronal (In) activity. In *Dynamic neuroscience* (pp. 219–238). Springer.
- Ching, S., Brown, E. N., & Kramer, M. A. (2012). Distributed control in a mean-field cortical network model: implications for seizure suppression. *Physical Review E*, 86(2), Article 021920.
- Ching, S., & Ritt, J. (2013). Control strategies for underactuated neural ensembles driven by optogenetic stimulation. *Frontiers in Neural Circuits*, 7, 54.
- Churchland, M., Yu, B., Cunningham, J., et al. (2010). Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nature Neuroscience*, 13, 369–378. <http://dx.doi.org/10.1038/nn.2501>.
- Cloosterman, M., van de Wouw, N., Heemels, M., & Nijmeijer, H. (2006). Robust stability of networked control systems with time-varying network-induced delays. In *Proceedings of the 45th IEEE conference on decision and control* (pp. 4980–4985). IEEE.
- Cobb, A. R., O'Connor, P., Zaizar, E., Caulfield, K., Gonzalez-Lima, F., & Telch, M. J. (2021). tDCS-Augmented in vivo exposure therapy for specific fears: A randomized clinical trial. *Journal of Anxiety Disorders*, 78, Article 102344.
- Cohen, J. D., Aston-Jones, G., & Gilzenrat, M. S. (2004). A systems-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation versus exploration. The Guilford Press.
- Cohen, J. D., Braver, T. S., & O'Reilly, R. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1346), 1515–1527.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97(3), 332.
- Cullen, K. E., Rey, C. G., Guitton, D., & Galiana, H. L. (1996). The use of system identification techniques in the analysis of oculomotor burst neuron spike train dynamics. *Journal of Computational Neuroscience*, 3(4), 347–368.
- Dasanayake, I., & Li, J.-S. (2011). Optimal design of minimum-power stimuli for phase models of neuron oscillators. *Physical Review E*, 83(6), Article 061916.
- David, O., & Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage*, 20(3), 1743–1755.
- David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., & Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage*, 30(4), 1255–1272.
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Computational Neuroscience Series.
- de Hollander, G., Labruna, L., Sellaro, R., Trutti, A., Colzato, L. S., Ratcliff, R., et al. (2016). Transcranial direct current stimulation does not influence the speed-accuracy tradeoff in perceptual decision-making: evidence from three independent studies. *Journal of Cognitive Neuroscience*, 28(9), 1283–1294.
- Deco, G., Tononi, G., Boly, M., & Kringelbach, M. L. (2015). Rethinking segregation and integration: contributions of whole-brain modelling. *Nature Reviews: Neuroscience*, 16, 430–439. <http://dx.doi.org/10.1038/nrn3963>.
- Demirtaş, M., Burt, J. B., Helmer, M., Ji, J. L., Adkinson, B. D., Glasser, M. F., et al. (2019). Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron*, 101(6), 1181–1194. <http://dx.doi.org/10.1016/j.neuron.2019.01.017>.
- Destexhe, A., & Sejnowski, T. J. (2009). The Wilson-Cowan model, 36 years later. *Biological Cybernetics*, 101(1), 1–2.
- Doi, S., Onoda, Y., & Kumagai, S. (2002). Parameter estimation of various Hodgkin-Huxley-type neuronal models using a gradient-descent learning method. In *Proceedings of the 41st SICE annual conference*, vol. 3 (pp. 1685–1688). IEEE.
- Duchet, B., Weerasinghe, G., Cagnan, H., Brown, P., Bick, C., & Bogacz, R. (2020). Phase-dependence of response curves to deep brain stimulation and their relationship: From essential tremor patient data to a Wilson-Cowan model. *The Journal of Mathematical Neuroscience*, 10(1), 1–39.
- Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *Journal of Neurophysiology*, 83(3), 1733–1750. <http://dx.doi.org/10.1152/jn.2000.83.3.1733>.
- Emonson, M., Fitzgerald, P., Rogasch, N., & Hoy, K. (2019). Neurobiological effects of transcranial direct current stimulation in younger adults, older adults and mild cognitive impairment. *Neuropsychologia*, 125, 51–61.
- Fallon, S. J., Williams-Gray, C. H., Barker, R. A., Owen, A. M., & Hampshire, A. (2013). Prefrontal dopamine levels determine the balance between cognitive stability and flexibility. *Cerebral Cortex*, 23(2), 361–369. <http://dx.doi.org/10.1093/cercor/bhs025>.
- Fertonani, A., & Miniussi, C. (2017). Transcranial electrical stimulation: what we know and do not know about mechanisms. *The Neuroscientist*, 23(2), 109–123.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673–9678.
- Frässle, S., Lomakina, E. I., Razi, A., Friston, K. J., Buhmann, J. M., & Stephan, K. E. (2017). Regression DCM for fMRI. *Neuroimage*, 155, 406–421.
- Frässle, S., Manjaly, Z. M., Do, C. T., Kasper, L., Pruessmann, K. P., & Stephan, K. E. (2021). Whole-brain estimates of directed connectivity for human connectomics. *NeuroImage*, 225, Article 117491.
- Freeman, W. J., & Barrie, J. M. (1994). Chaotic oscillations and the genesis of meaning in cerebral cortex. In *Temporal coding in the brain* (pp. 13–37). Springer.
- Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480.
- Friston, K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273–1302. [http://dx.doi.org/10.1016/S1053-8119\(03\)00202-7](http://dx.doi.org/10.1016/S1053-8119(03)00202-7).
- Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). A DCM for resting state fMRI. *NeuroImage*, 94, 396–407. <http://dx.doi.org/10.1016/j.neuroimage.2013.12.009>.
- Girard, A. (2005). Reachability of uncertain linear systems using zonotopes. In *International workshop on hybrid systems: computation and control* (pp. 291–305). Springer.
- Gomi, H., & Kawato, M. (1993). Neural network control for a closed-loop system using feedback-error-learning. *Neural Networks*, 6(7), 933–946. [http://dx.doi.org/10.1016/S0893-6080\(93\)80004-X](http://dx.doi.org/10.1016/S0893-6080(93)80004-X).
- Gong, X., & Fang, F. (2018). Effect of 10Hz transcranial alternating current stimulation (tACS) on motion direction identification. *Journal of Vision*, 18(10), 342.
- Goodman, W. K., & Insel, T. R. (2009). Deep brain stimulation in psychiatry: concentrating on the road ahead. *Biological psychiatry*, 65(4), 263–266.
- Gorzelic, P., Schiff, S., & Sinha, A. (2013). Model-based rational feedback controller design for closed-loop deep brain stimulation of Parkinson's disease. *Journal of Neural Engineering*, 10(2), Article 026016.

- Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., et al. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of Neuroengineering and Rehabilitation*, 5(1), 1–33.
- Green, A. E., Spiegel, K. A., Giangrande, E. J., Weinberger, A. B., Gallagher, N. M., & Turkeltaub, P. E. (2017). Thinking cap plus thinking zap: tDCS of frontopolar cortex improves creative analogical reasoning and facilitates conscious augmentation of state creativity in verb generation. *Cerebral Cortex*, 27(4), 2628–2639.
- Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Alfred, B. Y., Kahn, A. E., et al. (2015). Controllability of structural brain networks. *Nature communications*, 6(1), 1–10.
- Gutkin, B. S., Ermentrout, G. B., & Reyes, A. D. (2005). Phase-response curves give the responses of neurons to transient inputs. *Journal of Neurophysiology*, 94(2), 1623–1635.
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B., & Sigworth, F. J. (1981). Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv*, 391(2), 85–100.
- Hansen, E. C., Battaglia, D., Spiegler, A., Deco, G., & Jirsa, V. K. (2015). Functional connectivity dynamics: Modeling the switching behavior of the resting state. *NeuroImage*, 105, 525–535. <http://dx.doi.org/10.1016/j.neuroimage.2014.11.001>.
- Herrera-Guzmán, I., Gudayol-Ferré, E., Herrera-Guzmán, D., Guàrdia-Olmos, J., Hinojosa-Calvo, E., & Herrera-Abarca, J. E. (2009). Effects of selective serotonin reuptake and dual serotonergic-noradrenergic reuptake treatments on memory and mental processing speed in patients with major depressive disorder. *Journal of Psychiatric Research*, 43(9), 855–863.
- Himpe, C., & Ohlberger, M. (2014). Cross-gramian-based combined state and parameter reduction for large-scale control systems. *Mathematical Problems in Engineering*, 2014.
- Hmamed, A., Benzaouia, A., & Bensalah, H. (1995). Regulator problem for linear continuous-time delay systems with nonsymmetrical constrained control. *IEEE Transactions on Automatic Control*, 40(9), 1615–1619.
- Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4), 500–544.
- Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24), 10240–10245. <http://dx.doi.org/10.1073/pnas.0701519104>.
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., et al. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6), 2035–2040. <http://dx.doi.org/10.1073/pnas.0811168106>.
- Horvath, J. C., Forte, J. D., & Carter, O. (2015a). Evidence that transcranial direct current stimulation (tDCS) generates little-to-no reliable neurophysiologic effect beyond MEP amplitude modulation in healthy human subjects: A systematic review. *Neuropsychologia*, 66, 213–236.
- Horvath, J. C., Forte, J. D., & Carter, O. (2015b). Quantitative review finds no evidence of cognitive effects in healthy populations from single-session transcranial direct current stimulation (tDCS). *Brain Stimulation*, 8(3), 535–550.
- Howards, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, 142(4), 1211.
- Howards, M. W., & Hasselmo, M. E. (2020). Cognitive computation using neural representations of time and space in the Laplace domain. *arXiv preprint arXiv:2003.11668*.
- Hoy, K. E., & Fitzgerald, P. B. (2010). Brain stimulation in psychiatry and its effects on cognition. *Nature Reviews Neurology*, 6(5), 267–275.
- Hsu, T.-Y., Juan, C.-H., & Tseng, P. (2016). Individual differences and state-dependent responses in transcranial direct current stimulation. *Frontiers in Human Neuroscience*, 10, 643.
- Huang, W. A., Stitt, I. M., Negahbani, E., Passey, D., Ahn, S., Davey, M., et al. (2021). Transcranial alternating current stimulation entrains alpha oscillations by preferential phase synchronization of fast-spiking cortical neurons to stimulation waveform. *Nature communications*, 12(1), 1–20.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106–154.
- Husain, M. M., Lisanby, S. H., & Kay, J. (2015). Brain stimulation in psychiatry. *Psychiatry*, 1, 2291–2310.
- Ito, T., Brincat, S. L., Siegel, M., Mill, R. D., He, B. J., Miller, E. K., et al. (2020). Task-evoked activity quenches neural correlations and variability across cortical areas. *PLoS Computational Biology*, 16(8), Article e1007983.
- Jacobs, J., Kahana, M. J., Ekstrom, A. D., & Fried, I. (2007). Brain oscillations control timing of single-neuron activity in humans. *Journal of Neuroscience*, 27(14), 3839–3844.
- Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4), 357–366.
- Jirsa, V., Proix, T., Perdikis, D., Woodman, M., Wang, H., Gonzalez-Martinez, J., et al. (2017). The virtual epileptic patient: Individualized whole-brain models of epilepsy spread. *NeuroImage*, 145, 377–388. <http://dx.doi.org/10.1016/j.neuroimage.2016.04.049>.
- Kajimura, S., Kochiyama, T., Nakai, R., Abe, N., & Nomura, M. (2016). Causal relationship between effective connectivity within the default mode network and mind-wandering regulation and facilitation. *NeuroImage*, 133, 21–30.
- Kar, K., & Krekelberg, B. (2014). Transcranial alternating current stimulation attenuates visual motion adaptation. *Journal of Neuroscience*, 34(21), 7334–7340.
- Katz, B., Au, J., Buschkuhl, M., Abagis, T., Zabel, C., Jaeggi, S. M., et al. (2017). Individual differences and long-term consequences of tDCS-augmented cognitive training. *Journal of Cognitive Neuroscience*, 29(9), 1498–1508.
- Ketz, N., Jones, A. P., Bryant, N. B., Clark, V. P., & Pilly, P. K. (2018). Closed-loop slow-wave tACS improves sleep-dependent long-term memory generalization by modulating endogenous oscillations. *Journal of Neuroscience*, 38(33), 7314–7326. <http://dx.doi.org/10.1523/JNEUROSCI.0273-18.2018>.
- Kiebel, M. I., Moran, R. J., & Friston, K. J. (2008). Dynamic causal modelling for EEG and MEG. *Cognitive Neurodynamics*, 2(2), 121–136.
- Kirschstein, T., & Köhling, R. (2009). What is the source of the EEG? *Clinical EEG and Neuroscience*, 40(3), 146–149.
- Knierim, J. J., Kudrimoti, H. S., & McNoughton, B. L. (1995). Place cells, head direction cells, and the learning of landmark stability. *Journal of Neuroscience*, 15(3), 1648–1659.
- Koch, C. (2004). *Biophysics of computation: information processing in single neurons*. Oxford University Press.
- Kopell, N. J., Gritton, H. J., Whittington, M. A., & Kramer, M. A. (2014). Beyond the connectome: The dynamo. *Neuron*, 83(6), 1319–1328.
- Kucyi, A., & Davis, K. D. (2014). Dynamic functional connectivity of the default mode network tracks daydreaming. *NeuroImage*, 100, 471–480. <http://dx.doi.org/10.1016/j.neuroimage.2014.06.044>.
- Lange, M., Mauerer, J., Schlaier, J., Janzen, A., Zeman, F., Bogdahn, U., et al. (2017). Underutilization of deep brain stimulation for Parkinson's disease? A survey on possible clinical reasons. *Acta Neurochirurgica*, 159(5), 771.
- Lankarany, M., Zhu, W.-P., & Swamy, M. (2014). Joint estimation of states and parameters of Hodgkin-Huxley neuronal model using kalman filtering. *Neurocomputing*, 136, 289–299.
- Lee, B., Kramer, D., Armenta Salas, M., Kellis, S., Brown, D., Dobrev, T., et al. (2018). Engineering artificial somatosensation through cortical stimulation in humans. *Frontiers in Systems Neuroscience*, 12, 24.
- London, R. E., & Slagter, H. A. (2015). Effects of transcranial direct current stimulation over left dorsolateral pFC on the attentional blink depend on individual baseline performance. *Journal of Cognitive Neuroscience*, 27(12), 2382–2393.
- López-Alonso, V., Cheeran, B., Río-Rodríguez, D., & Fernández-del Olmo, M. (2014). Inter-individual variability in response to non-invasive brain stimulation paradigms. *Brain Stimulation*, 7(3), 372–380.
- Lorenz, R., Simmons, L. E., Monti, R. P., Arthur, J. L., Limal, S., Laakso, I., et al. (2019). Efficiently searching through large tACS parameter spaces using closed-loop Bayesian optimization. *Brain Stimulation*, 12(6), 1484–1489.
- Maass, W. (2000). On the computational power of winner-take-all. *Neural Computation*, 12(11), 2519–2535.
- MacLeod, C. M., & Dunbar, K. (1988). Training and stroop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(1), 126.
- Mammone, N., La Foresta, F., & Morabito, F. C. (2011). Automatic artifact rejection from multichannel scalp EEG by wavelet ICA. *IEEE Sensors Journal*, 12(3), 533–542.
- Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology*, 47(6), 1142–1150.
- Meron, D., Hedger, N., Garner, M., & Baldwin, D. S. (2015). Transcranial direct current stimulation (tDCS) in the treatment of depression: Systematic review and meta-analysis of efficacy and tolerability. *Neuroscience & Biobehavioral Reviews*, 57, 46–62.
- Meyer, P.-J. (2021). Reachability analysis of neural networks using mixed monotonicity. *arXiv preprint arXiv:2111.07683*.
- Meyer, P.-J., Devnport, A., & Arcak, M. (2019). TIRA: Toolbox for interval reachability analysis. In *Proceedings of the 22nd ACM international conference on hybrid systems: computation and control* (pp. 224–229).
- Miller, J., Berger, B., & Sauseng, P. (2015). Anodal transcranial direct current stimulation (tDCS) increases frontal-midline theta activity in the human EEG: A preliminary investigation of non-invasive stimulation. *Neuroscience Letters*, 588, 114–119.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167–202.
- Miranda-Dominguez, O., Gonia, J., & Netoff, T. (2010). Firing rate control of a neuron using a linear proportional-integral controller. *Journal of Neural Engineering*, 7(6), Article 066004.
- Moehlis, J., Shea-Brown, E., & Rabitz, H. (2006). Optimal inputs for phase models of spiking neurons.
- Muldoon, S. F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S. T., Vettel, J. M., et al. (2016). Stimulation-based control of dynamic brain networks. *PLoS Computational Biology*, 12(9), Article e1005076.
- Mumtaz, W., Rasheed, S., & Irfan, A. (2021). Review of challenges associated with the EEG artifact removal methods. *Biomedical Signal Processing and Control*, 68, Article 102741. <http://dx.doi.org/10.1016/j.bspc.2021.102741>.

- Nabi, A., Stigen, T., Moehlis, J., & Netoff, T. (2013). Minimum energy control for in vitro neurons. *Journal of Neural Engineering*, 10(3), Article 036005.
- Nakagawa, T. T., Jirsa, V. K., Spiegler, A., McIntosh, A. R., & Deco, G. (2013). Bottom up modeling of the connectome: Linking structure and function in the resting brain and their changes in aging. *Neuroimage*, 80, 318–329.
- Nandi, A., Kafashan, M. M., & Ching, S. (2017). Control analysis and design for statistical models of spiking networks. *IEEE Transactions on Control of Network Systems*, 5(3), 1146–1156.
- Nandi, A., Schattler, H., & Ching, S. (2019). Optimal control for fast, accurate threshold-hitting. *SIAM Journal on Control and Optimization*, 57(4), 2269–2291.
- Nandi, A., Schattler, H., Ritt, J. T., & Ching, S. (2017). Fundamental limits of forced asynchronous spiking with integrate and fire dynamics. *The Journal of Mathematical Neuroscience*, 7(1), 1–37.
- Narayanan, V., Ritt, J. T., Li, J.-S., & Ching, S. (2019). A learning framework for controlling spiking neural networks. In *2019 American control conference* (pp. 211–216). IEEE.
- Neher, E., & Sakmann, B. (1992). The patch clamp technique. *Scientific American*, 266(3), 44–51.
- Nitsche, M. A., Fricke, K., Henschke, U., Schlitterlau, A., Liebetanz, D., Lang, N., et al. (2003). Pharmacological modulation of cortical excitability shifts induced by transcranial direct current stimulation in humans. *The Journal of Physiology*, 553(1), 293–301.
- Nitsche, M. A., Liebetanz, D., Schlitterlau, A., Henschke, U., Fricke, K., Frommann, K., et al. (2004). Gabaergic modulation of DC stimulation-induced motor cortex excitability shifts in humans. *European Journal of Neuroscience*, 19(10), 2720–2726.
- Noury, N., & Siegel, M. (2017). Phase properties of transcranial electrical stimulation artifacts in electrophysiological recordings. *Neuroimage*, 158, 406–416.
- Noury, N., & Siegel, M. (2018). Analyzing EEG and MEG signals recorded during tES, |2a reply. *Neuroimage*, 167, 53–61.
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain Research*.
- Pandarinath, C., O'Shea, D. J., Collins, J., Jozefowicz, R., Stavisky, S. D., Kao, J. C., et al. (2018). Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, 15(10), 805–815.
- Pasqualetti, F., Zampieri, S., & Bullo, F. (2014). Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1), 40–52.
- Paulus, W. (2011). Transcranial electrical stimulation (tES–tDCS; tRNS, tACS) methods. *Neuropsychological Rehabilitation*, 21(5), 602–617.
- Perlmutter, J. S., & Mink, J. W. (2006). Deep brain stimulation. *Annual Review of Neuroscience*, 29, 229–257.
- Pion-Tonachini, L., Hsu, S.-H., Chang, C.-Y., Jung, T.-P., & Makeig, S. (2018). Online automatic artifact rejection using the real-time eeg source-mapping toolbox (REST). In *2018 40th annual international conference of the IEEE engineering in medicine and biology society* (pp. 106–109). IEEE.
- Ponce-Alvarez, A., He, B. J., Hagmann, P., & Deco, G. (2015). Task-driven activity reduces the cortical activity space of the brain: experiment and whole-brain modeling. *PLoS Computational Biology*, 11(8), Article e1004445. <http://dx.doi.org/10.1371/journal.pcbi.1004445>.
- Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, 117(4), 1113.
- Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *Ieee Assp Magazine*, 3(1), 4–16.
- Raco, V., Bauer, R., Tharsan, S., & Gharabaghi, A. (2016). Combining TMS and tACS for closed-loop phase-dependent modulation of corticospinal excitability: A feasibility study. *Frontiers in Cellular Neuroscience*, 10, <http://dx.doi.org/10.3389/fncel.2016.00143>.
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, 38, 433–447.
- Rall, W. (1962). Theory of physiological properties of dendrites. *Annals of the New York Academy of Sciences*, 96(4), 1071–1092.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281.
- Razi, A., Seghier, M. L., Zhou, Y., & McColgan, P. (2017). Large-scale DCMs for resting-state fMRI. *Network Neuroscience*, 1(3), http://dx.doi.org/10.1162/NETN_a.00015.
- Read, J., & Bentall, R. (2010). The effectiveness of electroconvulsive therapy: A literature review. *Epidemiology and Psychiatric Sciences*, 19(4), 333–347.
- Reato, D., Rahman, A., Bikson, M., & Parra, L. C. (2010). Low-intensity electrical stimulation affects network dynamics by modulating population rate and spike timing. *Journal of Neuroscience*, 30(45), 15067–15079.
- Reinhart, R. M., & Nguyen, J. A. (2019). Working memory revived in older adults by synchronizing rhythmic brain circuits. *Nature Neuroscience*, 22(5), 820–827.
- Reteig, L. C., Newman, L. A., Ridderinkhof, K. R., & Slagter, H. A. (2021). Effects of tDCS on the attentional blink revisited: A statistical evaluation of a replication attempt. *BioRxiv*.
- Rice, J. K., Rorden, C., Little, J. S., & Parra, L. C. (2013). Subject position affects EEG magnitudes. *NeuroImage*, 64, 476–484.
- Ritt, J. T., & Ching, S. (2015). Neurocontrol: Methods, models and technologies for manipulating dynamics in the brain. In *2015 American control conference* (pp. 3765–3780). IEEE.
- Roche, N., Geiger, M., & Bussel, B. (2015). Mechanisms underlying transcranial direct current stimulation in rehabilitation. *Annals of Physical and Rehabilitation Medicine*, 58(4), 214–219.
- Romani, G. L., Williamson, S. J., & Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. *Science*, 216(4552), 1339–1340.
- Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., of T.M.S. Consensus Group, S., et al. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, 120(12), 2008–2039.
- Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences*, 102(20), 7338–7343. <http://dx.doi.org/10.1073/pnas.0502455102>.
- Santaniello, S., Gale, J. T., & Sarma, S. V. (2018). Systems approaches to optimizing deep brain stimulation therapies in Parkinson's disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 10(5), Article e1421.
- Saxena, S., Schieber, M. H., Thakor, N. V., & Sarma, S. V. (2012). Aggregate input-output models of neuronal populations. *IEEE Transactions on Biomedical Engineering*, 59(7), 2030–2039.
- Schall, J. D., Morel, A., & Kaas, J. H. (1993). Topography of supplementary eye field afferents to frontal eye field in macaque: implications for mapping between saccade coordinate systems. *Visual Neuroscience*, 10(2), 385–393.
- Schiff, S. J. (2010). Towards model-based control of Parkinson's disease. *Philosophical Transactions of the Royal Society of London A (Mathematical and Physical Sciences)*, 368(1918), 2269–2308.
- Schirner, M., Rothmeier, S., Jirsa, V. K., McIntosh, A. R., & Ritter, P. (2015). An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. *NeuroImage*, 117, 343–357. <http://dx.doi.org/10.1016/j.neuroimage.2015.03.055>.
- Schwab, B. C., Misselhorn, J., & Engel, A. K. (2019). Modulation of large-scale cortical coupling by transcranial alternating current stimulation. *Brain Stimulation*, 12(5), 1187–1196.
- Sejnowski, T. J., & Paulsen, O. (2006). Network oscillations: Emerging computational principles. *Journal of Neuroscience*, 26(6), 1673–1676.
- Shapiro, K. L., Raymond, J., & Arnell, K. (1997). The attentional blink. *Trends in Cognitive Sciences*, 1(8), 291–296.
- Sikka, A., Jamalabadi, H., Krylova, M., Alizadeh, S., van der Meer, J. N., Danyeli, L., et al. (2020). Investigating the temporal dynamics of electroencephalogram (EEG) microstates using recurrent neural networks. *Human Brain Mapping*, 41(9), 2334–2346.
- Singh, M. F., Braver, T. S., Cole, M. W., & Ching, S. (2020). Estimation and validation of individualized dynamic brain models with resting state fMRI. *NeuroImage*, 221, Article 117046.
- Singh, M. F., & Ching, S. (2018). Network restructuring control for conic invariance with application to neural networks. In *2018 IEEE conference on decision and control* (pp. 2704–2709). IEEE.
- Singh, M. F., Wang, A., Braver, T. S., & Ching, S. (2020). Scalable surrogate deconvolution for identification of partially-observable systems and brain modeling. *Journal of Neural Engineering*, 17(4), Article 046025.
- Singh, M. F., Wang, C., Cole, M. W., & Ching, S. (2022). Efficient identification for modeling high-dimensional brain dynamics. *2022 American Control Conference (ACC)*, In Press.
- Singh, M. F., Wang, A., Cole, M., Ching, S., & Braver, T. S. (2021). Enhancing task fMRI preprocessing via individualized model-based filtering of intrinsic activity dynamics. *NeuroImage*, Article 118836.
- Smith, E., Horga, G., Yates, M., et al. (2019). Widespread temporal coding of cognitive control in the human prefrontal cortex. *Nature Neuroscience*, 22, 1883–1891. <http://dx.doi.org/10.1038/s41593-019-0494-0>.
- Stanyer, E. C., Baniqued, P. D. E., Awais, M., Kouara, L., Davies, A. G., Killan, E. C., et al. (2021). The impact of acoustic stimulation during sleep on memory and sleep architecture: A meta-analysis. *Journal of Sleep Research*, Article e13385.
- Stassen, H. H., Angst, J., Hell, D., Scharfetter, C., & Szegedi, A. (2007). Is there a common resilience mechanism underlying antidepressant drug response? Evidence from 2848 patients. *The Journal of Clinical Psychiatry*, 68(8), 9882.
- Stecher, H. I., Notbohm, A., Kasten, F. H., & Herrmann, C. S. (2021). A comparison of closed loop vs. Fixed frequency tACS on modulating brain oscillations and visual detection. *Frontiers in Human Neuroscience*, 15.
- Stevner, A., Vidaurre, D., Cabral, J., Rapuano, K., Nielsen, S. F. V., Tagliazucchi, E., et al. (2019). Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. *Nature communications*, 10(1), 1–14.
- Stigen, T., Danzl, P., Moehlis, J., & Netoff, T. (2009). Linear control of neuronal spike timing using phase response curves. In *2009 annual international conference of the IEEE engineering in medicine and biology society* (pp. 1541–1544). IEEE.
- Stiso, J., Khambhati, A. N., Menara, T., Kahn, A. E., Stein, J. M., Das, S. R., et al. (2019). White matter network architecture guides direct electrical stimulation through optimal state transitions. *Cell Reports*, 28(10), 2554–2566.

- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Subasi, A., & Gursay, M. I. (2010). EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications*, 37(12), 8659–8666.
- Sussmann, H. J. (1985). Lie brackets and real analyticity in control theory. *Banach Center Publications*, 14(1), 515–542.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409–1422.
- Tort, A. B., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of Neurophysiology*, 104(2), 1195–1210.
- Tronarp, F., Subramaniam, N. P., Särkkä, S., & Parkkonen, L. (2018). Tracking of dynamic functional connectivity from MEG data with Kalman filtering. In *2018 40th annual international conference of the IEEE engineering in medicine and biology society* (pp. 1003–1006). IEEE.
- Truong, N. D., Nguyen, A. D., Kuhlmann, L., Bonyadi, M. R., Yang, J., Ippolito, S., et al. (2018). Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105, 104–111.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550.
- Van Maanen, L., Grasman, R. P., Forstmann, B. U., Keuken, M. C., Brown, S. D., & Wagenmakers, E.-J. (2012). Similarity and number of alternatives in the random-dot motion paradigm. *Attention, Perception, & Psychophysics*, 74(4), 739–753.
- van't Wout-Frank, M., Shea, M. T., Larson, V. C., Greenberg, B. D., & Philip, N. S. (2019). Combined transcranial direct current stimulation with virtual reality exposure for posttraumatic stress disorder: Feasibility and pilot results. *Brain Stimulation*, 12(1), 41–43.
- Verkhatsky, A., & Parpura, V. (2014). History of electrophysiology and the patch clamp. In *Patch-clamp methods and protocols* (pp. 1–19). Springer.
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48), 12827–12832.
- Vrabie, D., & Lewis, F. (2009). Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems. *Neural Networks*, 22(3), 237–246. <http://dx.doi.org/10.1016/j.neunet.2009.03.008>.
- Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5), 955–968.
- Wang, P., Kong, R., Kong, X., Liégeois, R., Orban, C., Deco, G., et al. (2019). Inversion of a large-scale circuit model reveals a cortical hierarchy in the dynamic resting human brain. *Science Advances*, 5(1), <http://dx.doi.org/10.1126/sciadv.aat7854>.
- Whalen, A. J., Brennan, S. N., Sauer, T. D., & Schiff, S. J. (2015). Observability and controllability of nonlinear networks: The role of symmetry. *Physical Review X*, 5(1), Article 011005.
- Will, U., & Berg, E. (2007). Brain wave synchronization and entrainment to periodic acoustic stimuli. *Neuroscience Letters*, 424(1), 55–60.
- Wischert, W., Wunderlin, A., Pelster, A., Olivier, M., & Grosblum, J. (1994). Delay-induced instabilities in nonlinear feedback systems. *Physical Review E*, 49(1), 203.
- Yaffe, R. B., Borger, P., Megevand, P., Groppe, D. M., Kramer, M. A., Chu, C. J., et al. (2015). Physiology of functional and effective networks in epilepsy. *Clinical Neurophysiology*, 126(2), 227–236.
- Zhang, G., Cai, B., Zhang, A., Stephen, J. M., Wilson, T. W., Calhoun, V. D., et al. (2019). Estimating dynamic functional brain connectivity with a sparse hidden Markov model. *IEEE Transactions on Medical Imaging*, 39(2), 488–498.
- Zhang, Y., Guo, Y., Yang, P., Chen, W., & Lo, B. (2019). Epilepsy seizure prediction on EEG using common spatial pattern and convolutional neural network. *IEEE Journal of Biomedical and Health Informatics*, 24(2), 465–474.
- Zhou, A., Johnson, B. C., & Muller, R. (2018). Towards true closed-loop neuromodulation: Artifact-free recording during stimulation. *Current Opinion in Neurobiology*, 50, 119–127. <http://dx.doi.org/10.1016/j.conb.2018.01.012>.
- Zhu, F. F., Yeung, A. Y., Poolton, J. M., Lee, T. M., Leung, G. K., & Masters, R. S. (2015). Cathodal transcranial direct current stimulation over left dorsolateral prefrontal cortex area promotes implicit motor learning in a golf putting task. *Brain Stimulation*, 8(4), 784–786.
- Zlotnik, A., Chen, Y., Kiss, I. Z., Tanaka, H.-A., & Li, J.-S. (2013). Optimal waveform for fast entrainment of weakly forced nonlinear oscillators. *Physical Review Letters*, 111(2), Article 024102.
- Zlotnik, A., & Li, J.-S. (2012). Optimal entrainment of neural oscillator ensembles. *Journal of Neural Engineering*, 9(4), Article 046015.