A proof-of-concept process for validated brain emulation boot-strapped on an in-silico fully known ground-truth neural circuit

Randal A. Koene
Thomas Liao
Carboncopies Foundation for Brain Research (Should this say 'Substrate Independent Mind' instead?)

Abstract: (This will need to be re-written... the text here is temporary.) There is presently not a single published proof-of-concept example where a process of brain emulation has been carried out at some scale and where a validation was performed to quantify the degree to which an emulation satisfies necessary success criteria. Scientific motivation to make progress in brain emulation depends on having a touch-stone implementation and reference evaluation to quantitatively improve upon. The only way to evaluate claims of brain emulation is through the use of metrics that compare an emulation with an original system. Similarity metrics must measure similarity in ways that matter to the goal of whole brain emulation, i.e. that satisfy cognitive success criteria. E.g. spike train timing is not duplicated, but the probability of spiking is modulated sensibly and the evolution of system attractors is plausible. Similarity metrics are needed at multiple levels and while some can be used only with fully known ground-truth systems, others will carry over to whole brain systems. The principle of brain emulation depends on the ability to satisfy cognitive success criteria while replacing implementation details at some scale. In analog, potentially chaotic systems, scale separation is achieved through the application of operational constraints. E.g. rhythms (brain) or clock cycles (computer), neural population activity (brain) or parity bits (computer), action potentials (brain) or binary thresholds (computer). The application of constraints at consecutive levels limits the size of each black box in system identification.

Corresponding author: Randal A. Koene, CSO, Carboncopies Foundation for Brain Research, Sacramento, CA, and Nextup Neuro, LLC, Email: rkoene@carboncopies.org.

ORCID IDs: https://orcid.org/0000-1234-5678-9999 ;

Introduction

(Here: tiny paragraph about "brain emulation".)

(Here: a paragraph about how the process of model building for emulation is different than the typical computational neuroscience model building process.) field of neuroscience that is aimed at neural circuit function reconstruction in a way that extends beyond the requirements explored in systems neuroscience, cognitive neuroscience, or neuroinformatics. Existing specializations have emphasized
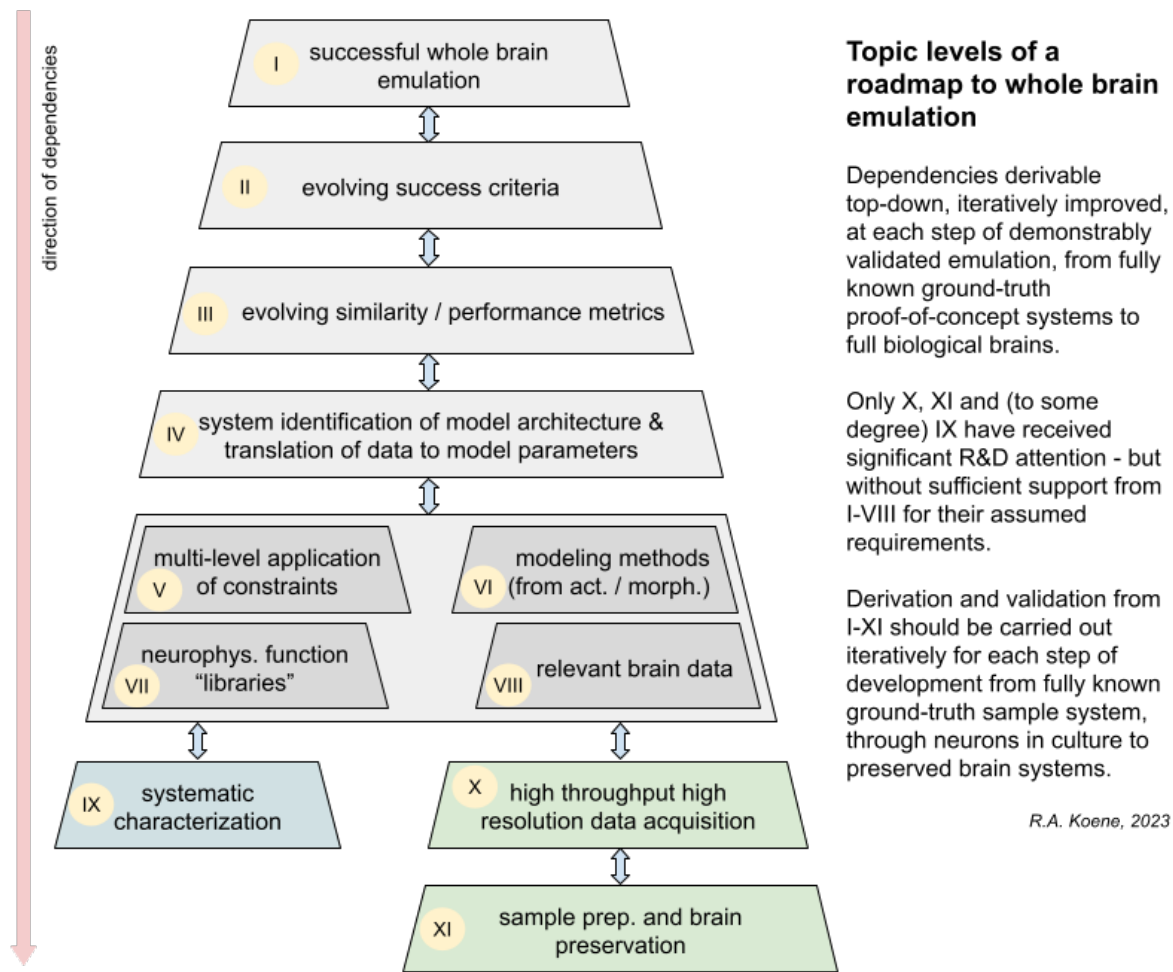
Topic levels of a roadmap to whole brain emulation

Dependencies derivable top-down, iteratively improved, at each step of demonstrably validated emulation, from fully known ground-truth proof-of-concept systems to full biological brains.

Only X, XI and (to some degree) IX have received significant R&D attention - but without sufficient support from I-VIII for their assumed requirements.

Derivation and validation from I-XI should be carried out iteratively for each step of development from fully known ground-truth sample system, through neurons in culture to preserved brain systems.

R.A. Koene, 2023

Figure 1. *The dependencies pyramid from goals of successful whole brain emulation to brain data requirements.*

the use of computation for hypothesis testing in the context of fundamental but abstracted concepts. A neuroscience of reconstruction must emphasize not just shared operational fundamentals but individually unique functional structure, such as that which enables retrieval of personal memories.

(Here: Rewrite the sentences below into proper journal style and language.) There is presently not a single published proof-of-concept example where a process of brain emulation has been carried out at some scale and where a validation was performed to quantify the degree to which an emulation satisfies necessary success criteria. Scientific motivation to make progress in brain emulation depends on having a touch-stone implementation and reference evaluation to quantitatively improve upon. The only way to evaluate claims of brain emulation is through the use of metrics that compare an emulation with an original system. Similarity metrics must measure similarity in ways that matter to the goal of whole brain emulation, i.e. that satisfy cognitive success criteria. E.g. spike train timing is not duplicated, but the probability of spiking is modulated sensibly and the evolution of system attractors is plausible. Similarity metrics are needed at multiple levels and while some can be used only with fully known ground-truth systems, others will carry over to whole brain systems. The principle of brain emulation depends on the ability to satisfy cognitive success criteria while replacing implementation details at some scale. In analog, potentially chaotic systems, scale separation is achieved through the application of operational constraints. E.g. rhythms (brain) or clock cycles (computer), neural population activity (brain) or parity bits (computer), action potentials (brain) or binary thresholds (computer). The application of constraints at consecutive levels limits the size of each black box in system identification.

(Here: A paragraph that explains why the first example uses two imaginary ball-and-stick neurons.) The ball-and-stick example is intended to provide the simplest in-silico case with the smallest number of variables to address while still demonstrating the full process and dependencies chain for whole brain emulation (see Fig. 1). This is an opportunity to anchor the development of useful similarity metrics for brain emulation.

(Here: Describe the I-XI requirements in the pyramid figure. Remove the topic levels text from the figure and put that description in here instead.)

In the following sections, we describe the process (parts I through XI) with the aid of an example experiment, *e0_bs* , that is designed as a simplest case in-silico with the minimum number of variables that can demonstrate the process. The abbreviation "bs" stands for ball-and-stick, because the neurons have only a spherical soma and a single axon component.

Known ground-truth: In-silico sample preparation (XI) and characterized physiology (IX and VII)

(Here: A paragraph explaining how we equate characterized ephys and biological brain samples with VBP architecture and components.)

Requirement IX: Establishing a set of virtual physiological components with well-characterized dynamics

For the purposes of testing methods and metrics, within this virtual brain laboratory, we define both the **morphology** and **physiology** of virtual components that we declare to be our **ground-truth** about which we know everything. The components of the *e0_bs* ground-truth model are as follows:

- The system comprises a set of brain regions.

Initial membrane potential	$V_m = -60mV$
Resting membrane potential	$V_{rest} = -60mV$
Action potential firing threshold	$V_{act} = -50mV$
Spike potential during refractory period	$V_{spike} = 60mV$
Time span of absolute refractory period	$\tau_{abs} = 1ms$
After-hyperpolarization potential	$V_{AHP} = -20mV$
After-hyperpolarization decay time constant	$\tau_{AHP} = 30ms$
Rise time constant of the post-synaptic potential	$\tau_{PSPr} = 5ms$
Decay time constant of the post-synaptic potential	$\tau_{PSPd} = 25ms$
Amplitude of the post-synaptic potential	$V_{PSP} = 20mV$

Table 1. *Default parameter values of ball-and-stick neuron class BS_Neuron and BS_Receptor.*

- A brain region (*BrainRegion*) has a specified geometric shape and specified physiological content in the form of neural circuits.

- The type of neural circuit defined in *e0_bs* is a linearly aligned ball-and-stick neural circuit (*BS_Aligned_NC*), so called, because the morphology of the neurons is essentially a ball with a stick.

- A neural circuit consists of some number of cells. Each of the cells has a specified morphology, characteristic dynamic functions of the neuron, and connectivity established by specifying receptors between a pre- and post-synaptic cell.

In *e0_bs* , the morphology is specified as a spherical soma (*BS_Soma*) and a cylindrical axon (*BS_Axon*) of particular dimensions (Fig. 2). A location is specified on the morphology for each receptor (*BS_Receptor*).

Functional characteristics of each cell in *e0_bs* are specified by a neuron definition (*BS_Neuron*) and by weights associated with the synaptic receptors and their characteristic functions.

The *e0_bs* example is purposely extremely simple:

- There is 1 brain region with a box shape, $20\mu m$ on each side.

- The brain region contains a single neural circuit.

- That neural circuit has 2 neurons, one pre-synaptic, one post-synaptic.

- The neurons are identical, using the default parameter values for the ball-and-stick neuron class (*BS_Neuron*), see Table 1.

- The weight of the synaptic connection formed by a receptor has a value of $w_{syn} = 1.0$.

A simulation of activity in the ground-truth model carried out in small time increments (e.g. $1ms$). For
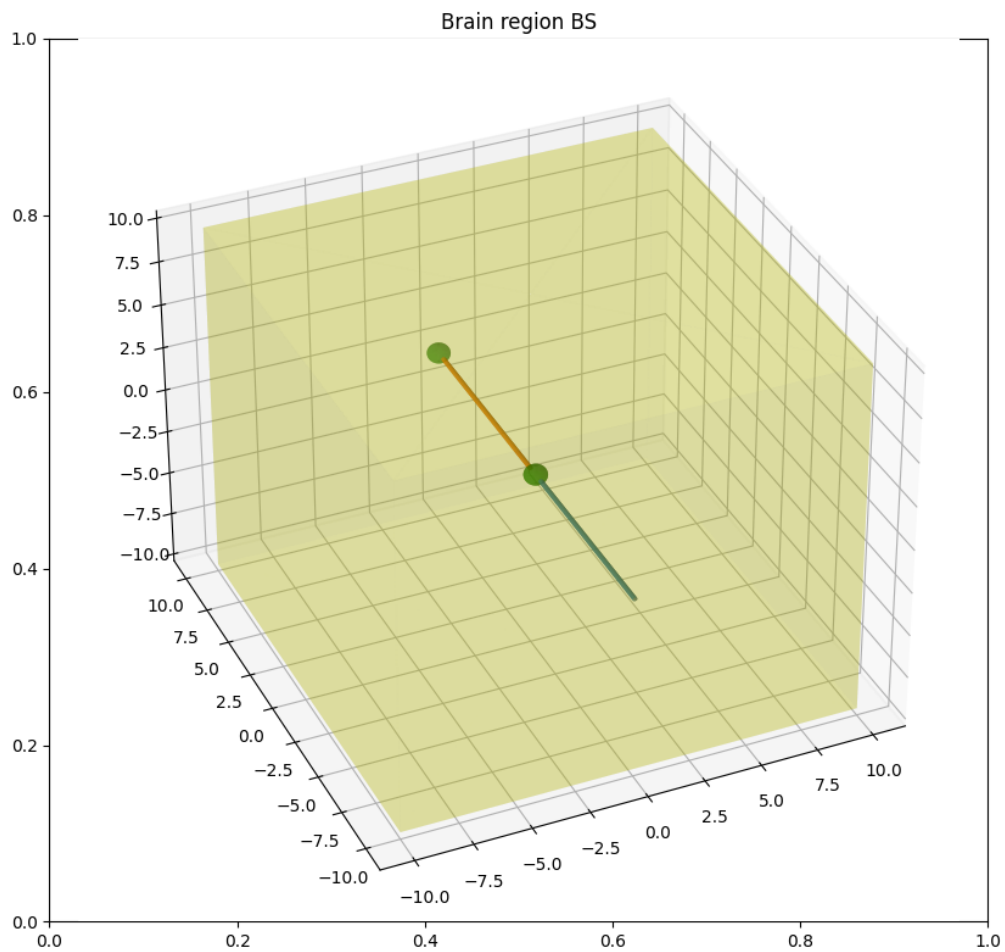
Figure 2. *Diagram of the known ground-truth ball-and-stick neural circuit architecture within in-silico brain region.*

each neuron, the update function has two main steps:

1. Update the momentary membrane potential, V_m.

2. Detect threshold crossing and possibly generate a spike. This step is ignored if the neuron is still within the absolute refractory period, τ_{abs}, of its most recent action potential.

The momentary membrane potential is the sum of contributing potentials:

$$v_m(t) = V_{rest} + v_{spike}(t) + v_{AHP}(t) + \sum_i v_{PSPi}(t), \tag{1}$$

where $v_{spike}(t) = V_{spike}$ during the absolute refractory period, and $v_{PSPi}(t)$ is the momentary post-synaptic potential contributed by the i-th receptor.

The refractory contribution of the spike:

$$
\begin{aligned}
v_{spike}(t) \quad &= V_{spike} \quad \text{if } \Delta t_{act} \leq \tau_{abs} \\
&= 0mV \quad \text{otherwise,}
\end{aligned}
\tag{2}
$$

where Δt_{act} is the time since onset of the most recent action potential.

Modulation of the membrane potential by after-hyperpolarization:

$$
\begin{aligned}
v_{AHP}(t) \quad &= V_{AHP} \exp(\tfrac{-\Delta t_{act}}{\tau_{AHP}}) \quad \text{if } \Delta t_{act} > \tau_{abs} \\
&= 0mV \quad \text{otherwise.}
\end{aligned}
\tag{3}
$$

Post-synaptic contributions to the membrane potential caused by the propagation of pre-synaptic action potentials through input receptors:

$$
\begin{aligned}
v_{PSPi}(t) \quad &= w_{syn} V_{PSP}(-\exp(\tfrac{-\Delta t_{act,i}}{\tau_{PSPr}}) + \exp(\tfrac{-\Delta t_{act,i}}{\tau_{PSPd}})) \quad \text{if the pre-synaptic neuron has spiked.} \\
&= 0mV \quad \text{otherwise,}
\end{aligned}
\tag{4}
$$

where $\Delta t_{act,i}$ is the time since onset of the most recent action potential at the pre-synaptic neuron connected through receptor i.

A threshold crossing occurs if the new momentary membrane potential is at or beyond the firing threshold, $V_m \geq V_{act}$. If so, then a new spike onset time, t_{act} is appended to the list of the neuron's spike times. That list is consulted by adjacent post-synaptic neurons that receive receptor input from this spiking neuron.

(Here: Mention how one would proceed on to the next more sophisticated step and towards working with real brain tissue for which there is no known ground-truth.)

Requirement VII: In-silico representation of virtual brain components

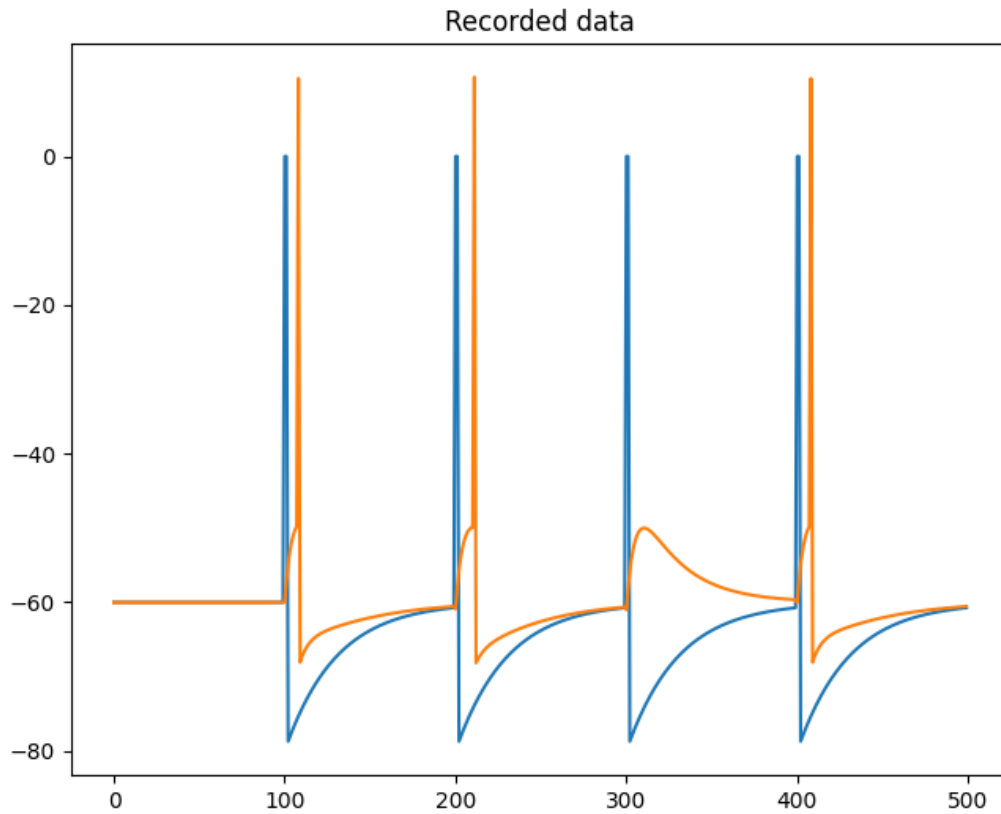(Here: Describe preparation and use of the components library. Add reference to SW.)

Figure 3. *Plot of ball-and-stick neuron membrane potentials as recorded in "God's eye" mode during an experiment run.*

Requirement XI: Preparation of the virtual brain architecture of the known ground-truth system

(Here: Describe how the architecture is set up. Add reference to SW. I.e. describe the steps that the model script takes to get to a defined known ground-truth system.)

A virtual brain ground-truth system provides a "God's eye" record

(Here: Describe how every calculated variable can be recorded for analysis and how that aids the development of similarity validation metrics.)

(Here: Make it clear that this is not the same as experimental data acquisition from the origin system.)

Data acquisition: Double-blind experimentation in a virtual brain laboratory (X and VIII)

(Here: Describe that even though the ground-truth model is fully known to the designer, the experimenter is blind to this and can use only data obtained through in-silico data acquisition that mirrors the process of data acquisition from biological brain systems.)

Simulated data acquisition is not permitted the God's eye view. Instead, it is constrained to the use of simulated data recording devices and simulated stimulation devices.

Activity observed is the result of either spontaneous simulated activity in the ground-truth system or activity caused by simulated stimulation methods. These may be simulated stimulation electrodes, simulated optogenetic stimulation, or others.

Typical simulated functional data recording devices include simulated recording electrodes of various types, simulated calcium imaging in a number of variants, or even simulated fNIRS or fMRI. In each case, detectable contributions of neuronal activity are combined with confouding factors (including noise, unreliability, defects) in simulation of the physics involved in measurements.

Typical simulated structural data recording devices include simulated microscopy of various types (e.g. electronmicroscopy, two-photon, light-sheet, etc). Again, the resulting images combine information about the ground-truth system morphology with confounding factors and effects of the physics involved.

Simulated data acquisition can be sophisticated and can attempt to model realistic results closely. Alternatively, simulated data acquisition can purposely apply simplifications for the following reasons:

1. Ease and rapidity of implementation.

2. Reduced computational cost.

3. Producing simplified laboratory condition examples with fewer variables to consider.

The simplified laboratory conditions reason is particularly useful, because this allows methods to be tested first in vastly simplified conditions where the capabilities and limitations of the method can be identified, demonstrated and evaluated in their most obvious form, without distracting contributors.

Requirement X: In-silico experimental data acquisition

(Here: Describe the data acquisition set up with the previously prepared KGT system and running data acquisition simulations. Add reference to SW.)

The steps of simulated data acquisition are:

1. Initialize the ground-truth model (loading or rerunning the previous stage).

2. Initialize simulated functional data acquisition by placing simulated electrodes or by setting up simulated calcium imaging.

3. Run simulated data acquisition and store that functional data.

4. Run simulated imaging by obtaining 2D projections of the 3D model and store that structural data.

Initializing simulated functional data acquisition

In our example experiment, *e0_bs* , preparation of simulated functional data acquisition involves these steps:

1. Specify expected spontaneous activity of the neurons in the system.

2. Set up a simulated single recording electrode in a location approximately between the two neurons.

3. Set up a simulated calcium imaging microscope that sees both neurons.

To specify the expected spontaneous activity, we pick a mean spontaneous firing interval, $280ms$, and its standard deviation, $140ms$. This will be the same for both neurons in the example system. We call the *set_spontaneous_activity()* member function with a list that associates the mean-stdev pair with each of the enumerated neurons.

A call to the *attach_recording_electrodes()* member function is used to set up any number of simulated electrodes. In our example, we provide a list with the specifications for only one electrode. We provide the following specifications:

• The position of the tip of the electrode, at the geometric center of the system.

• The positions of recording sites on the electrode, in this case one site at the very tip.

Similarly, a call to the *attach_calcium_imaging()* member function is used to set up a calcium imaging device. In *e0_bs* , we specify:

• Both neurons will fluoresce and show up during imaging.

• We will use a simulated jGCaMP8 calicium indicator ([?]) with relatively fast and short response dynamics, specified by an indicator rise time of $2ms$ and an indicator interval of $20ms$.

• The lens front position is $(0, 20, 0)$, $20\mu m$ above the simulated sample, and is positioned vertically, as indicated by a rear position $(0, 40, 0)$.

Simulating functional data acquisition

Simulated electrode and calcium imaging devices record data while we run the model for a specified number of simulated milliseconds.

(Here: put more about how those simulated devices generate the recorded data by using simulated physics.)

Simulating structural data acquisition

We provide high-throughput microscope specifications:

- Obtain images from the 'full' sample.

- The full sample is carved into sample sections $6\mu m$ wide and long.

- Each image has 6000 by 6000 pixels.

- The voxel resolution (represented by each pixel) is $4nm$ in x and y and $30nm$ in the z dimension.

(Here: put more about how a simulated high-throughput EM image stack is generated using simulated optics.)

Requirement VIII: Collected data and post-processing

(Here: Describe what it means for the collected data to be the "relevant" brain data.)

(Here: Describe the format of data obtained and how it may be post-processed for this simple experiment. Add reference to SW.)

System Identification and Translation to emulation model parameters (IV-VI)

(Here: Explain what system identification is, model selection and structuring. Explain what Translation means, model fitting. Point out the importance of constraints and their application. Explain that this process can involve multiple concurrent attempts or repeated attempts, guided by the validation step and error identification.)

Requirement VI: Model selection and structure derivation from activity data and morphological data (VI)

(Here: Describe the process. Add reference to SW.)

Requirement V: Application of constraints at multiple levels (V)

(Here: Describe the process. Add reference to SW.)

Requirement IV: Completing a process of system identification and translation for model architecture and parameters

(Here: Describe the process. Add reference to SW.)

Validation of candidate emulated systems using similarity and performance metrics based on success criteria for successful whole brain emulation (I-III)

(Here: Describe similarity metrics that can be used in known ground-truth system and those that can be used in a broader category of systems, even biological brains. Explain that these will evolve as research proceeds from these most basic in-silico systems to more sophisticated systems.)

Requirement III: Measuring similarity and performance

(Here: Describe the application of metrics and the evaluation of results.)

Using known ground-truth systems to develop methods for the identification of error sources and their correction

(Here: Describe an example of an error and how its cause is determined. Describe how the system identification and translation is adjusted and the outcome improved. Add reference to SW.)

Requirement II: Meeting success criteria

(Here: Describe this important relationship.)

Requirement I: Achieving a successful whole brain emulation for the ball-and-stick neural system

(Here: Describe the outcome.)

Discussion

(Here: Discuss the main takeaways and important insights about the process for this simple system.)

(Here: Point to the follow-up research and the general procedure of step-wise advancement. Add a reference to the project and company.)

Data, Code and Materials Availability Statement

(Here: Add links and DOIs to data, code and related materials.)

Authorship and Contributorship Statement

(Here: List who conceived the study, who designed the study and wrote the first draft of the manuscript. List other contributions. Mention if someone analysed data and revised the manuscript.)

All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

Appendices etc.

(Here: Only if applicable.)