To reverse engineer an entire nervous system

Gal Haspel, Edward S Boyden, Jeffrey Brown, George Church, Netta Cohen, Christopher Fang-Yen, Steven Flavell, Miriam B Goodman, Anne C. Hart, Oliver Hobert, Konstantinos Kagias, Shawn Lockery, Yangning Lu, Adam Marblestone, Jordan Matelsky, Hanspeter Pfister, Horacio G. Rotstein, Monika Scholz, Eli Shlizerman, Vivek Venkatachalam, Guangyu Robert Yang, Quilee Simeon, Eviatar Yemini, Manuel Zimmer, Konrad Paul Kording

Abstract:
There are many theories of how behavior may be controlled by neurons. Testing and refining these theories would be greatly facilitated if we could correctly simulate an entire nervous system so we could replicate the brain dynamics in response to any stimuli or contexts. Besides, simulating a nervous system is in itself one of the big dreams in systems neuroscience. However, doing so requires us to identify how each neuron's output depends on its inputs, a process we call reverse engineering. Current efforts at this focus on the mammalian nervous system, but these brains are mind-bogglingly complex, allowing only recordings of tiny subsystems. Here we argue that the time is ripe for systems neuroscience to embark on a concerted effort to reverse engineer a smaller system and that *Caenorhabditis elegans* is the ideal candidate system as the established optophysiology techniques can capture and control each neuron's activity and scale to hundreds of thousands of experiments. Data across populations and behaviors can be combined because across individuals the nervous system is largely conserved in form and function. Modern machine-learning based modeling should then enable a simulation of *C. elegans'* impressive breadth of brain states and behaviors. The ability to reverse engineer an entire nervous system will benefit the design of artificial intelligence systems and all of systems neuroscience, enabling fundamental insights as well as new approaches for investigations of progressively larger nervous systems.

Why should we reverse engineer a nervous system?
Let us first be clear about what we mean by successful reverse engineering. Independent of variations in cellular biophysics, if we consider each neuron and muscle cell as computing its output as a function of the activity of its input cells and spontaneous activity. As such, it is characterized by a spatiotemporal function mapping input traces into an output trace (i.e. an input-output function), which represents both synaptic and non-synaptic effects (e.g. through peptides). In our framing, reverse engineering consists of figuring out the input-output mapping for all neurons and muscle cells as well as the inputs from the world, and reassembling the collection of input-output functions into a reliable model of the entire nervous and muscle system that we can simulate and connect to a simulated body in an outside world (Fig. 1). To be successful, the model should recapitulate behavior under a range of conditions, stimuli, and perturbations.

"What I cannot create I cannot understand," one of Richard Feynman's famous dictums, nicely highlights the need to build systems to drive their understanding. With a working model of a nervous system and its interactions with the body and world, we could test all the hypothesized

neuroscientific models and principles *in silico* inexpensively and rapidly, and we could develop candidates for new therapeutic approaches by first testing them on this simulation. Such a model may also catalyze the design of new information processing systems, intelligent signal processing systems that solve well-defined goals. The results may inspire a new generation of artificial intelligence systems that are orders of magnitude more efficient than current ones. We need reverse engineering to enable such a simulation.

A central goal of systems and computational neuroscience is thus to reverse engineer how brains convert stimuli, spontaneous activity, and internal states into behaviors (also see (Harel 2003)). Indeed, the National Institute of Health (NIH) and other funders started the BRAIN Initiative (Insel, Landis, and Collins 2013) with a multi-billion dollar investment to develop new large-scale neurotechnologies. Major initiatives by other funders include the Human Brain Project (Markram 2012), MICrONS (funded by the Intelligence Advanced Research Projects Activity), and the Simons Global Brain Collaborations. The investment of significant resources into reverse engineering the brain reflects the value of this endeavor.

What will we learn by reverse engineering a nervous system?
If we could decompose a simulation of the nervous system into neural input-output functions, we could predict behavior in response to any sensory signals (past and present) under any experimental manipulations and internal states. We could predict the full behavioral repertoire. And we could predict what every single neuron does as a function of brain state and stimulation. This ability to build a simulation is one of the definitions used for understanding in the systems neuroscience community (Kording et al. 2020). But prediction is only the first step, a means of validating our *in-silico* model. Ultimately, reverse engineering a nervous system aims to be able to build an explanatory model of the dynamics of a complete nervous system that captures adaptation and plasticity at cellular, circuit, whole-brain, and behavior levels. By running in-silico experiments, it may be possible to translate these neural dynamics to interpretable concepts such as decision-making, memory, sensory integration, attention, and coordination. This link between neuronal dynamics (both in the form of input-output relations and whole-brain states) and interpretable function is, arguably, the holy grail of systems and computational neuroscience.

By reverse engineering any entire nervous system, we would gain important insights about the scientific process of biological reverse engineering that could generalize to larger systems, discovering which information matters, and which shortcuts are possible. We now know synaptic patterns of connectivity (White et al. 1986; Cook et al. 2019; Witvliet et al. 2021; Brittin et al. 2021) poorly predict interactive patterns of neural activity (Bentley et al. 2016; Yemini et al. 2021; Susoy et al. 2021; Uzel, Kato, and Zimmer 2022; Beets et al. 2022; Ripoll-Sánchez et al. 2022), so what are we missing to understand dynamic operations in these circuits? We know that as in other animals so in *C. elegans*, individual neurons can compartmentalize signals and thus perform multiplex computations (Hendricks et al. 2012), so what resolution do we need to distinguish these compartmentalized signals? Do millisecond timescales matter or is it enough to have lower frequency signals? Can we solve these problems with optical imaging only? Can the information describing the dynamics of nervous systems be compressed into a small

number of principles? Is it enough to reverse-engineer using data on only some parts from each individual? We know that the biophysical properties of different cell types matter (Dag et al. 2023), but to what extent must these be understood and modeled? We now know that even in *C. elegans* with its invariant cell lineage, the wiring of neural circuits varies across individuals (Brittin et al. 2021). How can individual variability be taken into account and might data from many animals produce a meaningful general model? We cannot currently answer these questions because we have not performed the required experiments and nervous system modeling. Pioneering reverse engineering in *C. elegans* can point out favorable approaches when we try to reverse engineer the more complex nervous systems of rodents and other animals. Demonstrating reverse engineering of an entire nervous system would clarify what kind of data we may need for future, increasingly ambitious endeavors.

By trying to reverse engineer a nervous system, we may learn about failure modes. Can we easily be misled and believe we understand how it works from partial recording, casting doubt on current approaches to reverse engineer mammalian brains? How probable is it that the models we fit get the correlations right and the causality wrong (Tremblay et al. 2022)? How much data of what kind is too little to reverse engineer systems? Answers to these questions could guide research in all areas where the goal is to understand nervous systems. Reverse engineering a system may also lead us to discover misleading "principles" in past neuroscience research.

By reverse engineering a nervous system, we will galvanize and motivate the enrichment and expansion of technologies that are critical for research across neuroscience. Neuroscience can scale: we record from thousands of neurons where we used to record from a handful (Stevenson and Kording 2011). We would optically image a million neurons where we previously imaged hundreds (Abdelfattah et al. 2022). Similar progress happened in molecular techniques where, for example, driving the human genome project ultimately made sequencing a cheap tool used by virtually all labs for countless objectives, and analogously reverse engineering nervous systems should produce broadly useful techniques. Reverse engineering any nervous system would require the development of hardware automation, data handling, and sharing, as well as algorithms and software to extract and analyze neural activity from dense whole-brain recordings, then incorporate these innovations into a tech stack that can then be scaled up. One can only imagine the value of scaling up the tech stack enabled by fully reverse engineering a brain: connectomics, activity, behavior, and models would have to be shared and scaled up in unprecedented ways. The existence and fact-checking of all these resources would jump-start a drive toward scaling entire nervous systems reverse engineering and simulation.

In contrast to artificial intelligence systems, the nervous system of *C. elegans* is compact and incredibly low-energy consuming and yet able to allow the species to thrive all over the planet. As such, we may hope to find design principles that we can generalize to AI systems. Of particular interest here are the basic building blocks, figuring out which computational elements are used in *C.elegans,* e.g. in terms of nonlinearities, promises to inform the design of new low-energy AI systems as well as to make them more resilient in an adversarial world (Agarwal

et al. 2017). In analogy, being able to simulate the full nervous system of *C. elegans* may allow us to design new biological systems to solve technical problems.

Why have we not yet reverse engineered a nervous system?
Experimental limitations and theoretical realities mean that, despite early efforts (Sarma et al. 2018) we have yet to simulate the entire nervous system of *C. elegans*. We need many parameters to describe *C. elegans* and astronomically many to describe mammals. Our experiments so far have been too limited in breadth but also, focusing on correlations, have not measured the causal parameters needed for simulations. Indeed, it remains an open and hotly debated question whether, and how theories can supply the missing parameters.

One reason why these challenges appear insurmountable is that we focus on reverse engineering rodents, with most of the circuit work in mice. However, we can only ever observe a tiny part of their nervous system; neither can we record all inputs to a single neuron. There is no doubt that the relevant models for mice are exceptionally complicated, both in terms of neuronal and neural processing and in terms of learning during experimental procedures. Because neuroscientific research takes place mostly on mammals, it is not known how good modern techniques may be at reverse engineering these circuits (Jonas and Kording 2017). After all, individual neurons are often weakly correlated with behavior and the enormous numbers of neurons in mammals preclude a population-level view with the single-neuron resolution necessary to understand their in-depth circuit dynamics. Human nervous systems have about 86 billion neurons; and even mice have 70 million (Herculano-Houzel, Mota, and Lent 2006), compared to *C. elegans*' 302 neurons (White et al. 1986). Moreover, every cortical neuron in the rodent tends to receive input from thousands of neurons, on the same order of magnitude as all the neural connections in the nematode nervous system. In our quest to understand complex nervous systems we have started with some of the most complicated ones; maybe to get close to human physiology as soon as possible. Not only do we not know how to simulate such nervous systems, we do not know what we do not know on the path to simulate them.

We have had much of *C. elegans* connectome since 1986 (White et al. 1986); further annotated (Varshney et al. 2011; Chen, Hall, and Chklovskii 2006; Cook et al. 2019); and new datasets were added (Brittin et al. 2021; Witvliet et al. 2021; Mulcahy et al. 2022). But connectivity alone is insufficient because without knowing the strength and temporal properties of all connections and neuronal properties, "... it would not be possible to simply go from the wiring diagram to the dynamics of even two neurons" (Marder and Bargmann; Nature Methods, 2013). Despite impressive progress(Einevoll et al. 2019; Eliasmith and Trujillo 2014), we cannot yet simulate any entire nervous system.

No discussion of reverse engineering *C. elegans* can be complete without careful consideration of the OpenWorm project (Szigeti et al. 2014)(https://openworm.org/). The OpenWorm project impressively used relatively solid knowledge about the physical environment of the nematode, the physics of swimming, and muscle properties (Boyle and Cohen 2008; Boyle, Berri, and Cohen 2012). They used the very partial relevant published information (White et al. 1986; Varshney et al. 2011) as well as a two-dimensional atlas(Hall and Altun 2008) of nematode

nuclei. Impressively, OpenWorm even began compiling a list of ion channel inventories for neurons and models of their biophysical properties (Sarma et al. 2018). However, estimating neuronal interactions from behavior, or even from ongoing neuronal activity, is essentially impossible, as many different neuronal properties can produce the same behavior (Prinz, Bucher, and Marder 2004). Moreover, in the absence of standardized data formats and norms for data sharing, the transfer of neural data into published papers is equally an irreversible process. By starting the simulation project, the OpenWorm project took a massive step in the right direction. But missing are experiments explicitly aimed at producing the kind of data that would allow a faithful simulation. The available data was the limiting factor all along.

Why a small nervous system? Why *C. elegans*?
C. elegans has far fewer neurons and neuron-neuron connections making it a great starting point to measure the causal dynamics captured by its input-output functions. It has established optical techniques for recording and stimulating at scale (Nguyen et al. 2016; M. Liu et al. 2022; Bergs et al. 2023; Randi et al. 2022a). Importantly, *C. elegans* allows us to record and combine information across animals. Until recently, variations among nematodes were thought to be relatively unimportant, and physiological parameters relatively conserved (Randi et al. 2022a), supporting the pooling of data across individuals. More recent evidence of variability in the neuronal wiring (Brittin et al. 2021) and of neuronal encoding (Atanas et al. 2023 in press) across individuals suggests a role for individuality. Put together, *C. elegans* offers a unique system to address these questions: by analyzing many individuals, we can determine the extent of variation at the cellular, circuit and behavioral levels, and we can take these variations into account in bulding computational models. However, there is still considerable disagreement about the remaining variation across animals (Brittin et al. 2021). It may also be easier as electrical signals common in *C. elegans* require lower temporal resolution of both stimulation and recording (M. B. Goodman et al. 1998; Jiang et al. 2022; Q. Liu et al. 2018; Lockery and Goodman 2009). As a historical example of this process, sequencing the *C. elegans* genome also significantly advanced biology and laid the groundwork for sequencing the human genome (C. elegans Sequencing Consortium 1998). Most of the benefits of reverse engineering a nervous system listed above can be realized by reverse engineering a relatively small nervous system.

C. elegans is small enough to perturb and record calcium signals from each neuron, or even groups of neurons. Continuous recording and stimulation on a single setup would reach about 50,000 specimens a year, and can be parallelized across setups (Fig. 1AB). Undoubtedly this is a large number, but it is only an order of magnitude larger than the number of animals whose behavior was recorded for recent publications (Dag et al. 2023). A very large number of experiments are possible, clearly enough for at least measuring all pairwise effects (~90k pairs, e.g., neuromodulator by neurotransmitter).
In *C. elegans* and only in *C. elegans*, this approach can sample a massive number of combinations of presynaptic neurons. In other words, within a single year, we could densely sample the space of inputs to all the neurons. Within a single year, a single targeted *C. elegans* experimental paradigm that is run on multiple setups may obtain more data than all mouse experiments taken together.

What should be the deliverables of reverse engineering the *C. elegans* nervous system?
The result of reverse engineering a brain is a dynamical model from which one could generate testable predictions for all possible manipulations and experiments. In other words, we want a causally correct 'digital twin' of the biological system.This twin should react to any kind of stimulation in the same way as the real system. So what is the minimal aspect of what we mean by wanting to reverse engineer it? Let us go through the deliverables we want to have.

The model should be able to describe the nervous system and the motor output (i.e. muscle contractions that are behavior). This description will have to include the influence of each neuron on other neurons and ultimately on behavior to predict stimulation effects. Importantly, neurons interact through synapses with neurotransmitters, but also through gap junctions, neuromodulators, and non-synaptic and non-neuronal paths such as glia (Raiders et al. 2021; Mu et al. 2019) as well as muscle activations and body shape(Cohen and Denham 2019; Zhan et al. 2023) . We thus need a model that captures interactions of these kind. It should be able to simulate the entire trajectory of neural activities and thus predict activities and tuning curves at all times and contexts (Hallinen et al. 2021).

While our conceptualization is built on there being little structured natural variability in wiring, neural activity, and behavior among individuals, assays should be included in the model as the measured variability in input-output functions and should replicate the variability in behavior. Importantly, if there are enough experiments, we can not just measure average properties but gain relevant insights into the overall distribution, or indeed identify multiple solutions. The behavioral variability among hundreds of simulations drawn from the measured distribution should be similar to that among hundreds of animals.

For behavior, we should be able to produce all the basic behaviors the animal performs, including spontaneous, responsive, and coupled behaviors. The hermaphrodite *C. elegans* has an impressively rich behavioral repertoire (Atanas et al. 2022; Hart 2006), constructed of first-order building blocks including: 1) locomotion behaviors, including forward, backward, change of speed, steer, turn, and halt; 2) feeding behaviors, more specifically occurrence, rate, and coordination of pharyngeal pumping (used to collect and crush organic material such as bacteria); 3) egg laying; 4) defecation; 5) fixed action patterns for mating. This list of behaviors excludes any kind of behavior that has a timescale long relative to the doable experiments (e.g. learning). We envision a simulation that replicate these building blocks of all behaviors in a single model. This means we need a simulation of the whole body, the inputs from the world around it, and how its effectors in turn affect the world that is around them, at least at a level that allows interpretation of motor output to behavior. These are the minimal deliverables, in our opinion, for producing acceptable models of the nervous system of *C. elegans.* Our expectation is that such a simulation should replicate the more complex behaviors that are made out of the first-order ones(Ghosh et al. 2016; Dekkers et al. 2021). Examples of these more complex behaviors include finding and attracting mates, fleeing from predators, avoiding problematic chemicals or temperatures, collective behavior, or avoiding parasites. In other words, a good

simulation should cover all the ethologically relevant high-order behaviors that are described for the nematode.

Feasibility: What do we need to do to reverse engineer a nervous system?
To close the loop we need a model of how the nematode's sensors translate the state of the nematode's body and the environment into neural activity, how neurons interact, and how neural activities through muscle contraction (effectors) influence the nematode's body and the environment. We also need a model for the world because the world is an important part of the solution to many behaviors. So we need a model of the world, the sensors, and the effectors (Fig. 1C). With a model of the body and environment in hand, we should be able to calculate the information going into the animal at all times and how the action of the effectors produces its future dynamics (Fig. 1D). To get there, we primarily need a sequence of experiments and modeling that require the components below. Because each component can be achieved by multiple technologies and approaches, we will describe the need and give a concrete example.
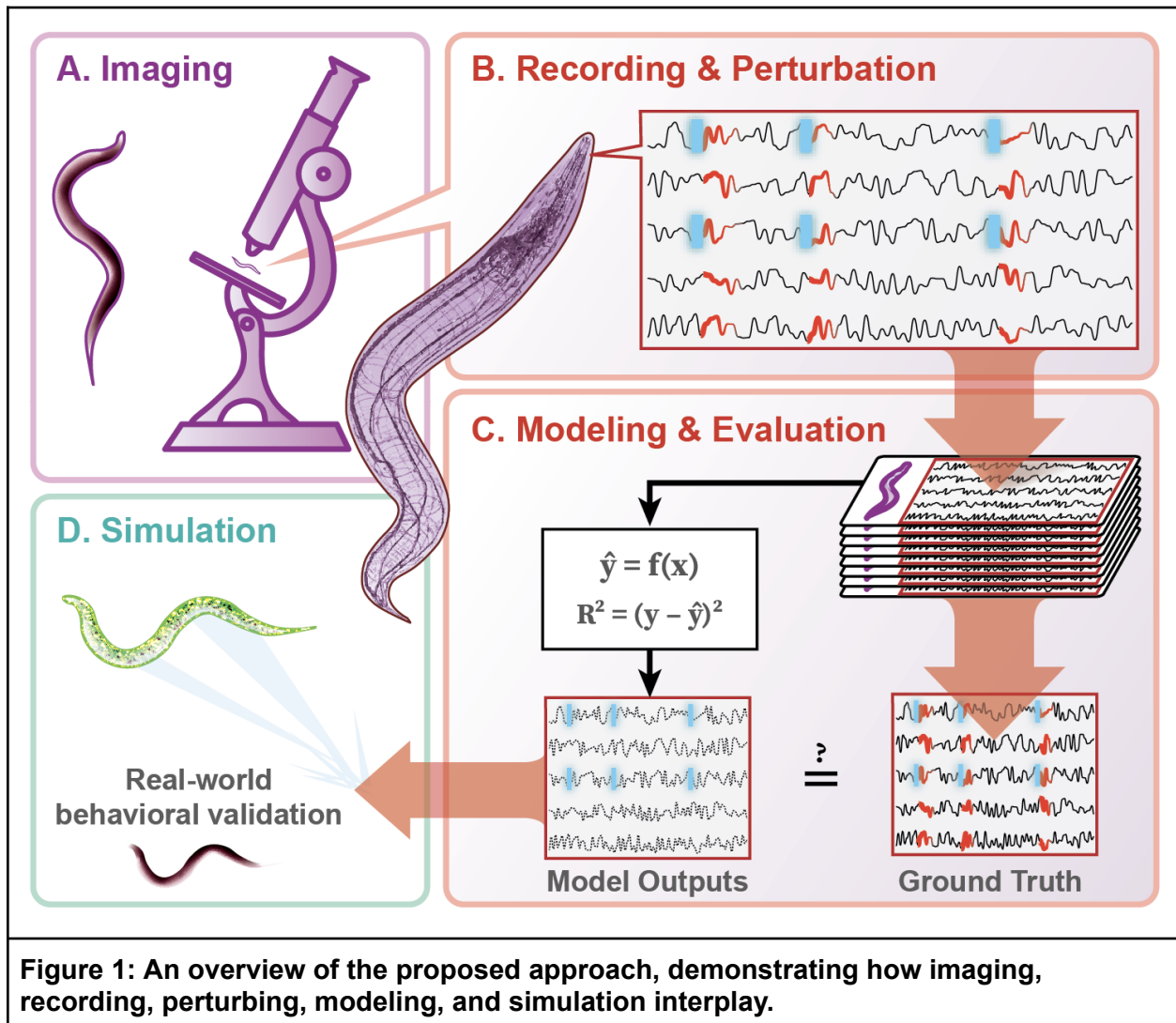

Figure 1: An overview of the proposed approach, demonstrating how imaging, recording, perturbing, modeling, and simulation interplay.

(1) Staging

Needed: Any approach we can think of will require the use of many animals, as we will need many thousands of hours of experiments to obtain the statistical power needed to meaningfully simulate all neurons. We thus need a way of getting nematodes in and out of set-ups with high throughput. If we were to do that manually, we would require large amounts of human labor. Another complication would be to standardize a manual workflow, so automatic staging is a better option.

Potential approaches: Depending on the experimental design, animals could be immobilized, restricted, or freely moving. Each option introduces challenges in collecting and interpreting the data but all past experiments have used extensive human labor. Automation could include robotic manipulation (Li et al. 2022), microfluidics approaches that either move nematodes in an aqueous solution or small bubbles of oil (San-Miguel and Lu 2013), and machine-vision-based focusing and positioning (Li et al. 2022). Neuronal calcium imaging data recorded in freely-moving animals (Atanas et al. 2022; Nguyen et al. 2016; Venkatachalam et al. 2016; Susoy et al. 2021), by their nature, do not reveal causality. The publicly available data (Atanas et al. 2022) can be used to test and refine simulated models constructed from input-output functions.

A concrete potential solution: Microfluidics can be used to get nematodes onto a microscope stage. We would have a reservoir of age-synchronized animals, microfluidically introduce animals to the chemogenetics reagents, and sequentially automatically collect each animal to be post-processed or assayed. We would develop a microfluidic system that places ten chemogenetically paralyzed adult hermaphrodites next to each other to fill a rectangular field of view [REF Adela Ben-Yakar]. Either way, microfluidic technology (San-Miguel and Lu 2013) and automation (Li et al. 2022) are well-established in *C. elegans* and appear to pose no relevant risk to the project.

(2) Microscopy

Needed: The nervous system of *C. elegans* is distributed along its body. About two-thirds of the 302 neuronal cell bodies are in a head ganglion and the rest split between a tail ganglion and a ventral nerve cord. Animals are about 1 mm in length and the maximal thickness of the nervous system is about 50 microns. This volume should be imaged at a sampling rate of at least 10 and preferably 100 Hz, requiring some progress relative to today's typically slower rates (Atanas et al. 2022). We will need one microscope for each set-up (although we could place multiple animals in the same field of view) with high enough resolution to be able to resolve individual neurons as we need to identify and match them across experiments. Importantly, we need to overcome field of view limitations, where typical microscopes have a field of view that is too small for fitting the entire worm. Fortunately, neuronal cell bodies are large enough that diffraction-limited imaging should be sufficient to resolve the relevant signals, in particular, if the relevant indicators can be localized to the soma. Importantly, we will have to design the setup specifically for the goal of automated, large-scale systems identification.

Potential approaches: Because diffraction-limited microscopy of fluorescent indicators is expected to be sufficient, there are many possible configurations (Lemon and McDole 2020).

This includes various light-sheet and spinning disk confocal designs that can scan this volume at the required sampling rate. It also includes the more sophisticated two-photon setups that allow exceptional precision in space, e.g., for stimulation.

A concrete potential solution: An inverted SCAPE2.0 inverted microscope (Voleti et al. 2019) with a microfluidic stage would be suitable for our approach. The inverted design provides access to liquid handling and stimuli and the SCAPE2 can scan the required volume (1 x 1 x 0.05 mm for ten immobilized animals) at about 10 Hz. This volume is sufficient to image ten animals.

(3) Neuron Identification

Needed: For each neuron that we record from we need to be able to assign it reliably to one of the 302 neuron identities. To do so, we will need to automatically identify each neuron from optical stacks after recording.

Potential approaches: A large number of solutions are possible. High-resolution, potential super-resolution, stacks, e.g. from expansion, would allow solving the problem by machine learning from large datasets. NeuroPAL (Yemini et al. 2021), a multicolor transgene allows for nervous-system-wide neuronal identification using a combination of reporters and colors to generate an invariant color map across individuals radically simplifying the identification process. Alternatively, machine learning may enable identification from large sets of multimodal data (Kirillov et al. 2023). This field of identification is quickly maturing.

A concrete potential solution: We would build on NeuroPAL (Yemini et al. 2021) animals. For these approaches, techniques are quickly getting standard but will still require additional software development (Skuhersky et al. 2022) and the development of novel methods for aligning neurons across thousands of animals. Importantly, it is sufficient to solve the identification problem in most animals.

(4) Stimulation and Recording

Needed: The core of reverse engineering a nervous system is figuring out how interactions among components (here neurons) shape the dynamics of information processing; we need to know how a neuron's output is caused by all its input cells (including those acting through neuromodulators) and how a neuron's output affects other neurons and muscle cells. In other words, we need the **interactome** which is a generalization of the connectome. This interactome is complicated because there are both synaptic interactions and neuropeptide interactions (Ripoll-Sánchez et al. 2022; Beets et al. 2022), but there have been recent attempts at reverse engineering a larger number of interactions (Randi et al. 2022b; Uzel, Kato, and Zimmer 2022). These approaches have, however, only revealed the average linear neuron-to-neuron interactions instead of the full nonlinear interactions. We need to know how stimulating (or inactivating) combinations of neurons affect the activity of each neuron. If we had this information, we should be able to calculate each neuron's output given its inputs.

Potential approaches: There are many ways of stimulating nematode neurons. We can stimulate them through direct physical effects (Suzuki et al. 2003; O'Hagan, Chalfie, and Goodman 2005; Ramot, MacInnis, and Goodman 2008; Miriam B. Goodman and Sengupta 2019), stimulate them electrically (M. B. Goodman et al. 1998), or stimulate them through optogenetic means (Husson, Gottschalk, and Leifer 2013; Nagel et al. 2005). We can do so with single-photon or

multi-photon approaches, trading of price and precision. All of these approaches are well established, but optogenetic methods are particularly scalable. There are also many ways of recording from nematode neurons. We can record from them electrically (M. B. Goodman et al. 1998), with calcium (Kerr et al. 2000), or by voltage imaging with genetically encoded sensors (Azimi Hashemi et al. 2019). Neurons are the most obvious target for stimulation but a complete interactome might require stimulating other cells such as glia (Perea et al. 2014). There are a large number of interactions to measure. And the main question is which kinds of approaches will best scale to very large numbers.

A concrete potential solution: We would generate transgenic nematodes that stochastically express an optogenetic activator in a random subset of neurons or muscle cells (Aoki et al. 2018). Given our experience with Brainbow, this is doable and presents minimal risk. For example, if each neuron has a 1/150 chance of expressing the activator, we expect most (~80%) of the animals to express the activator in one, two, or three neurons. A few animals will express it in more than three neurons, while 10-15% of animals will not express the activator in any neuron and will serve as an internal control group. Light activation should be as brief as possible and then as low as possible. We envision optogenetic stimulation on only one channel, as this makes experiments far cheaper, with the potential to add 2-photon stimulation as our knowledge about the system matures. We would use genetically encoded calcium indicators such as GCaMP as a reasonable proxy for neuronal activity. To reduce intrinsic activity and multisynaptic activation, we can subtly inactivate all neurons and muscle by making use of transgenic animals expressing a histamine-gated chloride channel (Pokala et al. 2014). Standard machine-learning approaches would then be used to identify the interactions.

(5) Quality control

Needed: Countless variables affect the results of the kinds of experiments we sketch here. This ranges from the genetic background of the nematodes (cites), to the properties of the microscope (cites), to the expression patterns of the indicators (cites). As such, careful monitoring of the overall pipeline is essential. To be able to pool results, the results must be all comparable.

Potential Approaches: There are two kinds of well-established approaches. One, as in the case of the International Brain Lab (cites) is to establish standards that are held up in many labs working in parallel. The other, as in the Allen Institute (cite), is to rather establish one central location where large-scale experiments are made with a team that is very much centered around quality control.

(6) Annotated Connectomes

Needed: To further understand the interactome data and possibly predict neuronal function, we want to have a specific molecularly annotated connectome (Bhattacharya et al. 2019; Taylor et al. 2021). In other words, we want to know the whole structure of the nervous system of animals from which we collected input-output data, but also where key molecules are expressed. This kind of information can significantly cut down on the set of models we need to consider.

Potential approaches: Serial electron microscopy and several super-resolution approaches can be adapted to image the morphology of the nervous system of collected and fixed animals. These are mostly low-throughput methods. Expansion microscopy is the only approach we are

aware of that could provide molecular annotation in addition to morphological data and at a relatively high throughput (Alon et al. 2021; Shen et al. 2020).

A concrete solution: For a fraction of the animals, after stimulation and recording, we would extrude animals from the microfluidic device, expand (Yu et al. 2020), and identify a host of molecules in 3 dimensions, along with the full connectome, which, in turn, will require advances in computer vision such as automatic proofreading.

(7) Model fitting

Needed: We need to be able to fit models to the data that are good enough so that they can enable a faithful simulation that is good enough that it generalizes to other situations and conditions.

Potential approaches: There are two major branches for such fitting. There is a school focusing on detailed models (in the sense of cell properties and cable equations, (Hines and Carnevale 1997; Kim, Leahy, and Shlizerman 2019; Kunert, Shlizerman, and Kutz 2014) and embodiment (Kim et al. 2019)). There is another school focusing on setting up machine learning models, e.g. by fitting deep learning models(Tanaka et al. 2019; Benjamin et al. 2018; Yamins and DiCarlo 2016). In all cases we would have to fit models where we estimate neuron outputs $y_j(t)$ as a function of neuron inputs $x_i(t)$. Subsequently, we build in biological prior knowledge, e.g. the stronger interactions between nearby synapses, e.g. by adapting the standard attention function used in transformer models.

A concrete solution: One could set up a deep learning system based on transformers (Vaswani et al. 2017) for every neuron. We may tokenize the input state as a set of discrete descriptors for each input neuron's state $x_i(t)$ learned using a Vector Quantised-Variational AutoEncoder (VQ-VAE) (Van Den Oord, Vinyals, and Others 2017). This is a standard strategy to get the continuous states (e.g. of the neurons) with minimal loss into a discrete format that is easy to model. An alternative is to model the relevant vectors directly without tokenization.

(8) Open Science

Needed: As we are proposing a project that will require a concentration of effort and resources, we need an open community. And we need that community for quality assurance - the data may have problems that we do not anticipate.

Potential solutions: We can try and produce a governance structure that minimizes that risk. Another possibility is to be absolutely open.

A concrete solution: We suggest doing both. We want to engage the relevant community to guide the approaches and open science has been the standard approach in the *C. elegans* field. We insist on radically open science in which code and data are made available within three months of the time at which they are developed or collected.

(9) Diverse Science

Needed: This project will require integration of scientists with expertise in instrumentation, genetic analysis, molecular tools, behavioral neuroscience, cellular biophysics, computation, data science, and theory.
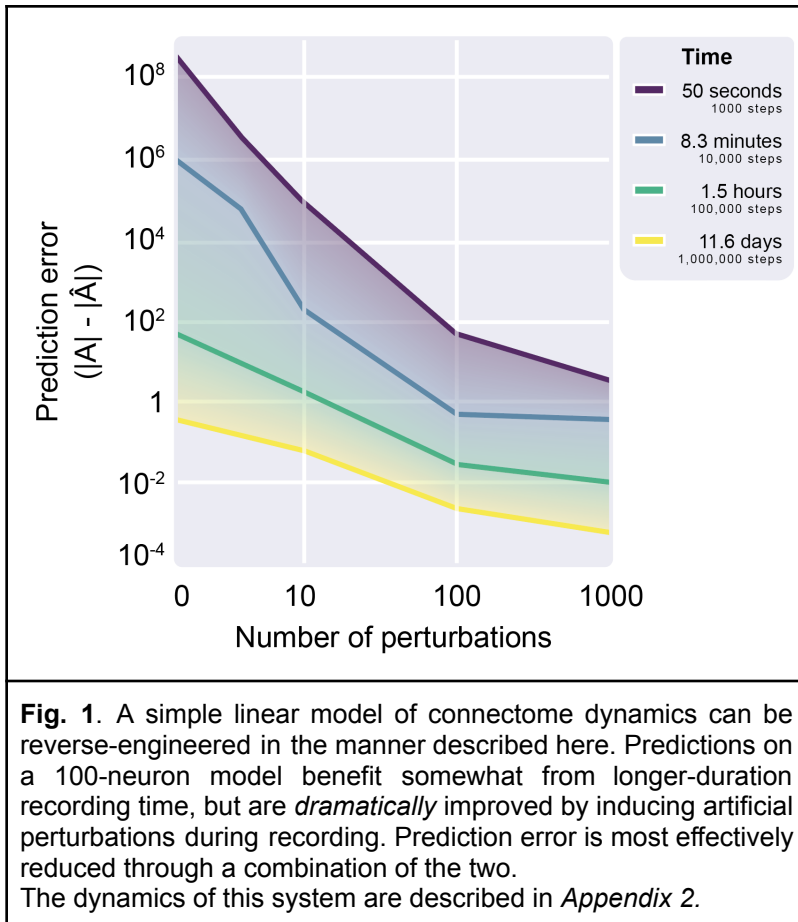
Potential solutions: Recruit scientists broadly and empower team members to cross-train in multiple disciplines. Intentionally build an inclusive community of scientists and embed a tradition of design and data review into teamwork.

Power calculations
It is notoriously difficult to estimate statistical power in the context of machine learning applications. However, there are two things that we can do. (1) We can produce a simulation that models the relevant problem and see how well we can reverse engineer that simulation to obtain useful intuition. (2) We can use special cases, such as linear models with nonlinear basis functions, to obtain closed-form estimates.

(1) One simulation strategy is to produce virtual connectomes of varying scale (different numbers of neurons and different densities of connections between them) and varying complexity. There are of course many ways to vary simulated neuron complexity, but one valuable approach is to change the number of input parameters that a neuron uses in its activation transfer function. A second valuable dimension is how much of its own and its synaptic partners' history a neuron takes as input — where a perfectly Markovian neuron takes only the latest state of the network as input and an unrealistically complex neuron takes as input all of the complete history of the entire network. Of course, real neurons fall in the middle of this complexity scale, though we do not know where. It is not hard to empirically evaluate how much recording time is required to analyze the complexity of a neuron (Fig. **2**).

(2) We find with considerable amounts of uncertainty (because we do not yet know how complex *C elegans* neurons are) that we will need to continuously record from the entire *C. elegans* nervous systems for the equivalent of about 250 days (see Appendix 1: analytical power calculations). As such, the proposed experiments should be well-doable, even if we use conservative calculations.

Fig. 1. A simple linear model of connectome dynamics can be reverse-engineered in the manner described here. Predictions on a 100-neuron model benefit somewhat from longer-duration recording time, but are *dramatically* improved by inducing artificial perturbations during recording. Prediction error is most effectively reduced through a combination of the two.
The dynamics of this system are described in *Appendix 2.*

Going to many nervous systems

Ultimately, the value of reverse engineering a nervous system is that it will make reverse engineering nervous systems accessible. Being able to compare nervous systems is, arguably, much more interesting. What goes wrong with disease? How can we scale things up? How do things differ across species? Just like the human genome project enabled sequencing the genomes of many humans, adding much more value, reverse engineering a nervous system promises to open the floodgates to reverse engineering many other nervous systems (Vogelstein et al. 2016).

Scaling to larger nervous systems

Reverse engineering *C. elegans* should just be a step in the path toward understanding nervous systems that are more like our own. There are of course many differences, in terms of size (human brains with about 86 billion neurons vs 302 in *C.elegans*), in terms of computational primitives (most cells in *C. elegans* do not spike at the millisecond timescale; (M. B. Goodman et al. 1998; Jiang et al. 2022; Q. Liu et al. 2018; Lockery and Goodman 2009), and in terms of complexity of brain organization. However, we can sketch what a path toward larger brain understanding can look like. Power calculations make the identification of the input-output function of cortical neurons in humans, that have of the order of 10,000 inputs from stimulation, seem infeasible. However, through the work that we argue needs to happen in this paper, we

will not merely establish causal interactions between *C. elegans* cells, we will also establish how to predict them from annotated connectomes. Annotated connectomes can readily be scaled up to much bigger nervous systems. We also expect that there are some principles (some known, most probably still unknown) that may carry over from small nervous systems to larger ones. The important thing is that the work on reverse engineering *C.elegans* will provide a rosetta-stone-like translation between the language of annotated connectomes and functional properties. This insight into the causal interactions and this ground truth-based approach can then be scaled up. For example, we could analyze the kinds of interactions in human brain organoids producing ground truth causality and then scale up from there. But we will need the whole nervous system scaling to know how to put together such local information with other factors to get at the causal flow of information in larger nervous systems.

Conclusions

Here we have outlined how reverse engineering the entire sensory-neuro-motor system of *C. elegans* may be possible, producing, for the first time, a real understanding of how an entire brain computes. Reverse engineering the causal interactions in the nervous system of *C. elegans* promises to establish *in-silico* simulations as a way of accelerating neuroscience. It promises to teach us how to run experiments that reveal the causal ways of how circuits work. We would learn about the many potential failure modes of reverse engineering. We may learn how to build powerful, error-robust, energetically-economical, synaptically-compact AI systems. A concerted, open effort would galvanize the development of approaches to help understand the brains of more complicated animals, like our own, promising a new era of causal neuroscience. The time has come to reverse engineer the *C. elegans* nervous system.

References

Abdelfattah, Ahmed S., Sapna Ahuja, Taner Akkin, Srinivasa Rao Allu, Joshua Brake, David A. Boas, Erin M. Buckley, et al. 2022. "Neurophotonic Tools for Microscopic Measurements and Manipulation: Status Report." *Neurophotonics* 9 (Suppl 1): 013001.

Agarwal, Nikita, Neil Mehta, Alice C. Parker, and Karam Ashouri. 2017. "C. Elegans Neuromorphic Neural Network Exhibiting Undulating Locomotion." In *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE. https://doi.org/10.1109/ijcnn.2017.7966349.

Alon, Shahar, Daniel R. Goodwin, Anubhav Sinha, Asmamaw T. Wassie, Fei Chen, Evan R. Daugharthy, Yosuke Bando, et al. 2021. "Expansion Sequencing: Spatially Precise in Situ Transcriptomics in Intact Biological Systems." *Science* 371 (6528). https://doi.org/10.1126/science.aax2656.

Aoki, Wataru, Hidenori Matsukura, Yuji Yamauchi, Haruki Yokoyama, Koichi Hasegawa, Ryoji Shinya, and Mitsuyoshi Ueda. 2018. "Cellomics Approach for High-Throughput Functional Annotation of Caenorhabditis Elegans Neural Network." *Scientific Reports* 8 (1): 10380.

Atanas, Adam A., Jungsoo Kim, Ziyu Wang, Eric Bueno, Mccoy Becker, Di Kang, Jungyeon Park, et al. 2022. "Brain-Wide Representations of Behavior Spanning Multiple Timescales and States in C. Elegans." *bioRxiv*. https://doi.org/10.1101/2022.11.11.516186.

Azimi Hashemi, Negin, Amelie C. F. Bergs, Christina Schüler, Anna Rebecca Scheiwe, Wagner Steuer Costa, Maximilian Bach, Jana F. Liewald, and Alexander Gottschalk. 2019. "Rhodopsin-Based Voltage Imaging Tools for Use in Muscles and Neurons of Caenorhabditis Elegans." *Proceedings of the National Academy of Sciences of the United States of America* 116 (34): 17051–60.

Beets, Isabel, Sven Zels, Elke Vandewyer, Jonas Demeulemeester, Jelle Caers, Esra Baytemur, William R. Schafer, Petra E. Vértes, Olivier Mirabeau, and Liliane Schoofs. 2022. "System-Wide Mapping of Neuropeptide-GPCR Interactions in C. Elegans." *bioRxiv*. https://doi.org/10.1101/2022.10.30.514428.

Benjamin, Ari S., Hugo L. Fernandes, Tucker Tomlinson, Pavan Ramkumar, Chris VerSteeg, Raeed H. Chowdhury, Lee E. Miller, and Konrad P. Kording. 2018. "Modern Machine Learning as a Benchmark for Fitting Neural Responses." *Frontiers in Computational Neuroscience* 12 (July): 56.

Bentley, Barry, Robyn Branicky, Christopher L. Barnes, Yee Lian Chew, Eviatar Yemini, Edward T. Bullmore, Petra E. Vértes, and William R. Schafer. 2016. "The Multilayer Connectome of Caenorhabditis Elegans." *PLoS Computational Biology* 12 (12): e1005283.

Bergs, Amelie C. F., Jana F. Liewald, Silvia Rodriguez-Rozada, Qiang Liu, Christin Wirt, Artur Bessel, Nadja Zeitzschel, et al. 2023. "All-Optical Closed-Loop Voltage Clamp for Precise Control of Muscles and Neurons in Live Animals." *Nature Communications* 14 (1): 1939.

Bhattacharya, Abhishek, Ulkar Aghayeva, Emily G. Berghoff, and Oliver Hobert. 2019. "Plasticity of the Electrical Connectome of C. Elegans." *Cell* 176 (5): 1174–89.e16.

Boyle, Jordan H., Stefano Berri, and Netta Cohen. 2012. "Gait Modulation in C. Elegans: An Integrated Neuromechanical Model." *Frontiers in Computational Neuroscience* 6 (March): 10.

Boyle, Jordan H., and Netta Cohen. 2008. "Caenorhabditis Elegans Body Wall Muscles Are Simple Actuators." *Bio Systems* 94 (1-2): 170–81.

Brittin, Christopher A., Steven J. Cook, David H. Hall, Scott W. Emmons, and Netta Cohen. 2021. "A Multi-Scale Brain Map Derived from Whole-Brain Volumetric Reconstructions." *Nature* 591 (7848): 105–10.

C. elegans Sequencing Consortium. 1998. "Genome Sequence of the Nematode C. Elegans: A Platform for Investigating Biology." *Science* 282 (5396): 2012–18.

Chen, Beth L., David H. Hall, and Dmitri B. Chklovskii. 2006. "Wiring Optimization Can Relate

Neuronal Structure and Function." *Proceedings of the National Academy of Sciences of the United States of America* 103 (12): 4723–28.

Cohen, Netta, and Jack E. Denham. 2019. "Whole Animal Modeling: Piecing Together Nematode Locomotion." *Current Opinion in Systems Biology* 13 (February): 150–60.

Cook, Steven J., Travis A. Jarrell, Christopher A. Brittin, Yi Wang, Adam E. Bloniarz, Maksim A. Yakovlev, Ken C. Q. Nguyen, et al. 2019. "Whole-Animal Connectomes of Both Caenorhabditis Elegans Sexes." *Nature* 571 (7763): 63–71.

Dag, Ugur, Ijeoma Nwabudike, Di Kang, Matthew A. Gomes, Jungsoo Kim, Adam A. Atanas, Eric Bueno, et al. 2023. "Dissecting the Functional Organization of the C. Elegans Serotonergic System at Whole-Brain Scale." *Cell*, May. https://doi.org/10.1016/j.cell.2023.04.023.

Einevoll, Gaute T., Alain Destexhe, Markus Diesmann, Sonja Grün, Viktor Jirsa, Marc de Kamps, Michele Migliore, Torbjørn V. Ness, Hans E. Plesser, and Felix Schürmann. 2019. "The Scientific Case for Brain Simulations." *Neuron* 102 (4): 735–44.

Eliasmith, Chris, and Oliver Trujillo. 2014. "The Use and Abuse of Large-Scale Brain Models." *Current Opinion in Neurobiology* 25 (April): 1–6.

Farhoodi, Roozbeh, Khashayar Filom, Ilenna Simone Jones, and Konrad Paul Kording. 2019. "On Functions Computed on Trees." *Neural Computation* 31 (11): 2075–2137.

Goodman, M. B., D. H. Hall, L. Avery, and S. R. Lockery. 1998. "Active Currents Regulate Sensitivity and Dynamic Range in C. Elegans Neurons." *Neuron* 20 (4): 763–72.

Goodman, Miriam B., and Piali Sengupta. 2019. "How Senses Mechanical Stress, Temperature, and Other Physical Stimuli." *Genetics* 212 (1): 25–51.

Hall, David H., and Zeynep F. Altun. 2008. *C. Elegans Atlas*. CSHL Press.

Hallinen, Kelsey M., Ross Dempsey, Monika Scholz, Xinwei Yu, Ashley Linder, Francesco Randi, Anuj K. Sharma, Joshua W. Shaevitz, and Andrew M. Leifer. 2021. "Decoding Locomotion from Population Neural Activity in Moving C. Elegans." *eLife* 10 (July). https://doi.org/10.7554/eLife.66135.

Harel, David. 2003. "A Grand Challenge: Full Reactive Modeling of a Multi-Cellular Animal." In *Hybrid Systems: Computation and Control*, 2–2. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hart, Anne. 2006. "Behavior." *WormBook: The Online Review of C. Elegans Biology*. https://doi.org/10.1895/wormbook.1.87.1.

Hendricks, Michael, Heonick Ha, Nicolas Maffey, and Yun Zhang. 2012. "Compartmentalized Calcium Dynamics in a C. Elegans Interneuron Encode Head Movement." *Nature* 487 (7405): 99–103.

Herculano-Houzel, Suzana, Bruno Mota, and Roberto Lent. 2006. "Cellular Scaling Rules for Rodent Brains." *Proceedings of the National Academy of Sciences of the United States of America* 103 (32): 12138–43.

Hines, M. L., and N. T. Carnevale. 1997. "The NEURON Simulation Environment." *Neural Computation* 9 (6): 1179–1209.

Husson, Steven J., Alexander Gottschalk, and Andrew M. Leifer. 2013. "Optogenetic Manipulation of Neural Activity in C. Elegans: From Synapse to Circuits and Behaviour." *Biology of the Cell / under the Auspices of the European Cell Biology Organization* 105 (6): 235–50.

Insel, Thomas R., Story C. Landis, and Francis S. Collins. 2013. "Research Priorities. The NIH BRAIN Initiative." *Science* 340 (6133): 687–88.

Jiang, Jingyuan, Yifan Su, Ruilin Zhang, Haiwen Li, Louis Tao, and Qiang Liu. 2022. "C. Elegans Enteric Motor Neurons Fire Synchronized Action Potentials Underlying the Defecation Motor Program." *Nature Communications* 13 (1): 2783.

Jonas, Eric, and Konrad Paul Kording. 2017. "Could a Neuroscientist Understand a Microprocessor?" *PLoS Computational Biology* 13 (1): e1005268.

Kato, Saul, Yifan Xu, Christine E. Cho, L. F. Abbott, and Cornelia I. Bargmann. 2014. "Temporal Responses of C. Elegans Chemosensory Neurons Are Preserved in Behavioral Dynamics." *Neuron* 81 (3): 616–28.

Kerr, R., V. Lev-Ram, G. Baird, P. Vincent, R. Y. Tsien, and W. R. Schafer. 2000. "Optical Imaging of Calcium Transients in Neurons and Pharyngeal Muscle of C. Elegans." *Neuron* 26 (3): 583–94.

Kim, Jimin, William Leahy, and Eli Shlizerman. 2019. "Neural Interactome: Interactive Simulation of a Neuronal System." *Frontiers in Computational Neuroscience* 13 (March): 8.

Kim, Jimin, Julia A. Santos, Mark J. Alkema, and Eli Shlizerman. 2019. "Whole Integration of Neural Connectomics, Dynamics and Bio-Mechanics for Identification of Behavioral Sensorimotor Pathways in Caenorhabditis Elegans." *bioRxiv*. bioRxiv. https://doi.org/10.1101/724328.

Kirillov, Alexander, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, et al. 2023. "Segment Anything." *arXiv [cs.CV]*. arXiv. http://arxiv.org/abs/2304.02643.

Kording, Konrad P., Gunnar Blohm, Paul Schrater, and Kendrick Kay. 2020. "Appreciating the Variety of Goals in Computational Neuroscience." *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/2002.03211.

Kunert, James, Eli Shlizerman, and J. Nathan Kutz. 2014. "Low-Dimensional Functionality of Complex Network Dynamics: Neurosensory Integration in the Caenorhabditis Elegans Connectome." *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 89 (5): 052805.

Lemon, William C., and Katie McDole. 2020. "Live-Cell Imaging in the Era of Too Many Microscopes." *Current Opinion in Cell Biology* 66 (October): 34–42.

Liu, Mochi, Sandeep Kumar, Anuj K. Sharma, and Andrew M. Leifer. 2022. "A High-Throughput Method to Deliver Targeted Optogenetic Stimulation to Moving C. Elegans Populations." *PLoS Biology* 20 (1): e3001524.

Liu, Qiang, Philip B. Kidd, May Dobosiewicz, and Cornelia I. Bargmann. 2018. "C. Elegans AWA Olfactory Neurons Fire Calcium-Mediated All-or-None Action Potentials." *Cell* 175 (1): 57–70.e17.

Li, Zihao, Anthony D. Fouad, Peter Bowlin, Yuying Fan, Siming He, Meng-Chuan Chang, Angelica Du, et al. 2022. "A Robotic System for Automated Genetic Manipulation and Analysis of Caenorhabditis Elegans on Agar Media." *bioRxiv*. https://doi.org/10.1101/2022.11.18.517134.

Lockery, Shawn R., and Miriam B. Goodman. 2009. "The Quest for Action Potentials in C. Elegans Neurons Hits a Plateau." *Nature Neuroscience* 12 (4): 377–78.

Markram, Henry. 2012. "The Human Brain Project." *Scientific American* 306 (6): 50–55.

Matelsky, Jordan K., Elizabeth P. Reilly, Erik C. Johnson, Jennifer Stiso, Danielle S. Bassett, Brock A. Wester, and William Gray-Roncal. 2021. "DotMotif: An Open-Source Tool for Connectome Subgraph Isomorphism Search and Graph Queries." *Scientific Reports* 11 (1): 13045.

Mulcahy, Ben, Daniel K. Witvliet, James Mitchell, Richard Schalek, Daniel R. Berger, Yuelong Wu, Doug Holmyard, et al. 2022. "Post-Embryonic Remodeling of the C. Elegans Motor Circuit." *Current Biology: CB* 32 (21): 4645–59.e3.

Mu, Yu, Davis V. Bennett, Mikail Rubinov, Sujatha Narayan, Chao-Tsung Yang, Masashi Tanimoto, Brett D. Mensh, Loren L. Looger, and Misha B. Ahrens. 2019. "Glia Accumulate Evidence That Actions Are Futile and Suppress Unsuccessful Behavior." *Cell* 178 (1): 27–43.e19.

Nagel, Georg, Martin Brauner, Jana F. Liewald, Nona Adeishvili, Ernst Bamberg, and Alexander Gottschalk. 2005. "Light Activation of Channelrhodopsin-2 in Excitable Cells of Caenorhabditis Elegans Triggers Rapid Behavioral Responses." *Current Biology: CB* 15 (24): 2279–84.

Nguyen, Jeffrey P., Frederick B. Shipley, Ashley N. Linder, George S. Plummer, Mochi Liu, Sagar U. Setru, Joshua W. Shaevitz, and Andrew M. Leifer. 2016. "Whole-Brain Calcium Imaging with Cellular Resolution in Freely Behaving Caenorhabditis Elegans." *Proceedings of the National Academy of Sciences of the United States of America* 113 (8): E1074–81.

O'Hagan, Robert, Martin Chalfie, and Miriam B. Goodman. 2005. "The MEC-4 DEG/ENaC Channel of Caenorhabditis Elegans Touch Receptor Neurons Transduces Mechanical Signals." *Nature Neuroscience* 8 (1): 43–50.

Perea, Gertrudis, Aimei Yang, Edward S. Boyden, and Mriganka Sur. 2014. "Optogenetic Astrocyte Activation Modulates Response Selectivity of Visual Cortex Neurons in Vivo." *Nature Communications* 5: 3262.

Pokala, Navin, Qiang Liu, Andrew Gordus, and Cornelia I. Bargmann. 2014. "Inducible and Titratable Silencing of Caenorhabditis Elegans Neurons in Vivo with Histamine-Gated Chloride Channels." *Proceedings of the National Academy of Sciences of the United States of America* 111 (7): 2770–75.

Prinz, Astrid A., Dirk Bucher, and Eve Marder. 2004. "Similar Network Activity from Disparate Circuit Parameters." *Nature Neuroscience* 7 (12): 1345–52.

Raiders, Stephan, Erik Calvin Black, Andrea Bae, Stephen MacFarlane, Mason Klein, Shai Shaham, and Aakanksha Singhvi. 2021. "Glia Actively Sculpt Sensory Neurons by Controlled Phagocytosis to Tune Animal Behavior." *eLife* 10 (March). https://doi.org/10.7554/eLife.63532.

Ramot, Daniel, Bronwyn L. MacInnis, and Miriam B. Goodman. 2008. "Bidirectional Temperature-Sensing by a Single Thermosensory Neuron in C. Elegans." *Nature Neuroscience* 11 (8): 908–15.

Randi, Francesco, Anuj K. Sharma, Sophie Dvali, and Andrew M. Leifer. 2022a. "Neural Signal Propagation Atlas of C. Elegans." https://doi.org/10.48550/ARXIV.2208.04790.

———. 2022b. "Neural Signal Propagation Atlas of $\textit{C. elegans}$." *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/2208.04790.

Ripoll-Sánchez, Lidia, Jan Watteyne, Haosheng Sun, Robert Fernandez, Seth R. Taylor, Alexis Weinreb, Mark Hammarlund, et al. 2022. "The Neuropeptidergic Connectome of C. Elegans." *bioRxiv*. https://doi.org/10.1101/2022.10.30.514396.

San-Miguel, Adriana, and Hang Lu. 2013. "Microfluidics as a Tool for C. Elegans Research." *WormBook: The Online Review of C. Elegans Biology*, September, 1–19.

Sarma, Gopal P., Chee Wai Lee, Tom Portegys, Vahid Ghayoomie, Travis Jacobs, Bradly Alicea, Matteo Cantarelli, et al. 2018. "OpenWorm: Overview and Recent Advances in Integrative Biological Simulation of." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 373 (1758). https://doi.org/10.1098/rstb.2017.0382.

Shen, Fred Y., Margaret M. Harrington, Logan A. Walker, Hon Pong Jimmy Cheng, Edward S. Boyden, and Dawen Cai. 2020. "Light Microscopy Based Approach for Mapping Connectivity with Molecular Specificity." *Nature Communications* 11 (1): 4632.

Skuhersky, Michael, Tailin Wu, Eviatar Yemini, Amin Nejatbakhsh, Edward Boyden, and Max Tegmark. 2022. "Toward a More Accurate 3D Atlas of C. Elegans Neurons." *BMC Bioinformatics* 23 (1): 195.

Stevenson, Ian H., and Konrad P. Kording. 2011. "How Advances in Neural Recording Affect Data Analysis." *Nature Neuroscience* 14 (2): 139–42.

Susoy, Vladislav, Wesley Hung, Daniel Witvliet, Joshua E. Whitener, Min Wu, Core Francisco Park, Brett J. Graham, Mei Zhen, Vivek Venkatachalam, and Aravinthan D. T. Samuel. 2021. "Natural Sensory Context Drives Diverse Brain-Wide Activity during C. Elegans Mating." *Cell* 184 (20): 5122–37.e17.

Suzuki, Hiroshi, Rex Kerr, Laura Bianchi, Christian Frøkjaer-Jensen, Dan Slone, Jian Xue, Beate Gerstbrein, Monica Driscoll, and William R. Schafer. 2003. "In Vivo Imaging of C. Elegans Mechanosensory Neurons Demonstrates a Specific Role for the MEC-4 Channel in

the Process of Gentle Touch Sensation." *Neuron* 39 (6): 1005–17.

Szigeti, Balázs, Padraig Gleeson, Michael Vella, Sergey Khayrulin, Andrey Palyanov, Jim Hokanson, Michael Currie, Matteo Cantarelli, Giovanni Idili, and Stephen Larson. 2014. "OpenWorm: An Open-Science Approach to Modeling Caenorhabditis Elegans." *Frontiers in Computational Neuroscience* 8 (November): 137.

Tanaka, Hidenori, Aran Nayebi, Niru Maheswaranathan, Lane McIntosh, Stephen Baccus, and Surya Ganguli. 2019. "From Deep Learning to Mechanistic Understanding in Neuroscience: The Structure of Retinal Prediction." *Advances in Neural Information Processing Systems* 32. https://proceedings.neurips.cc/paper/2019/hash/eeaebbffb5d29ff62799637fc51adb7b-Abstract.html.

Taylor, Seth R., Gabriel Santpere, Alexis Weinreb, Alec Barrett, Molly B. Reilly, Chuan Xu, Erdem Varol, et al. 2021. "Molecular Topography of an Entire Nervous System." *Cell* 184 (16): 4329–47.e23.

Tremblay, Sébastien, Camille Testard, Jeanne Inchauspé, and Michael Petrides. 2022. "Non-Necessary Neural Activity in the Primate Cortex." *bioRxiv*. https://doi.org/10.1101/2022.09.12.506984.

Uzel, Kerem, Saul Kato, and Manuel Zimmer. 2022. "A Set of Hub Neurons and Non-Local Connectivity Features Support Global Brain Dynamics in C. Elegans." *Current Biology: CB* 32 (16): 3443–59.e8.

Van Den Oord, Aaron, Oriol Vinyals, and Others. 2017. "Neural Discrete Representation Learning." *Advances in Neural Information Processing Systems* 30. https://proceedings.neurips.cc/paper/2017/hash/7a98af17e63a0ac09ce2e96d03992fbc-Abstract.html.

Varshney, Lav R., Beth L. Chen, Eric Paniagua, David H. Hall, and Dmitri B. Chklovskii. 2011. "Structural Properties of the Caenorhabditis Elegans Neuronal Network." *PLoS Computational Biology* 7 (2): e1001066.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention Is All You Need." *Advances in Neural Information Processing Systems* 30. https://proceedings.neurips.cc/paper/7181-attention-is-all.

Venkatachalam, Vivek, Ni Ji, Xian Wang, Christopher Clark, James Kameron Mitchell, Mason Klein, Christopher J. Tabone, et al. 2016. "Pan-Neuronal Imaging in Roaming Caenorhabditis Elegans." *Proceedings of the National Academy of Sciences of the United States of America* 113 (8): E1082–88.

Vogelstein, Joshua T., Katrin Amunts, Andreas Andreou, Dora Angelaki, Giorgio Ascoli, Cori Bargmann, Randal Burns, et al. 2016. "Grand Challenges for Global Brain Sciences." *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/1608.06548.

Voleti, Venkatakaushik, Kripa B. Patel, Wenze Li, Citlali Perez Campos, Srinidhi Bharadwaj, Hang Yu, Caitlin Ford, et al. 2019. "Real-Time Volumetric Microscopy of in Vivo Dynamics and Large-Scale Samples with SCAPE 2.0." *Nature Methods* 16 (10): 1054–62.

White, J. G., E. Southgate, J. N. Thomson, and S. Brenner. 1986. "The Structure of the Nervous System of the Nematode Caenorhabditis Elegans." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 314 (1165): 1–340.

Witvliet, Daniel, Ben Mulcahy, James K. Mitchell, Yaron Meirovitch, Daniel R. Berger, Yuelong Wu, Yufang Liu, et al. 2021. "Connectomes across Development Reveal Principles of Brain Maturation." *Nature* 596 (7871): 257–61.

Yamins, Daniel L. K., and James J. DiCarlo. 2016. "Using Goal-Driven Deep Learning Models to Understand Sensory Cortex." *Nature Neuroscience* 19 (3): 356–65.

Yemini, Eviatar, Albert Lin, Amin Nejatbakhsh, Erdem Varol, Ruoxi Sun, Gonzalo E. Mena, Aravinthan D. T. Samuel, Liam Paninski, Vivek Venkatachalam, and Oliver Hobert. 2021.

"NeuroPAL: A Multicolor Atlas for Whole-Brain Neuronal Identification in C. Elegans." *Cell* 184 (1): 272–88.e11.

Yu, Chih-Chieh Jay, Nicholas C. Barry, Asmamaw T. Wassie, Anubhav Sinha, Abhishek Bhattacharya, Shoh Asano, Chi Zhang, et al. 2020. "Expansion Microscopy of." *eLife* 9 (May). https://doi.org/10.7554/eLife.46249.

Zhan, Xu, Chao Chen, Longgang Niu, Xinran Du, Ying Lei, Rui Dan, Zhao-Wen Wang, and Ping Liu. 2023. "Locomotion Modulates Olfactory Learning through Proprioception in C. Elegans." *Nature Communications* 14 (1): 4534.

Appendix 1: analytical power calculations

General setting. We need to describe the dependency of the neuron's output on its inputs which we can formalize as: $y = f(x)$. Let there be L noisy observations, abstracting away time, of the form $y_i = f(x_i) + \eta_i$ with isotropic Gaussian noise $\eta_i \sim N(0, \sigma^2 I)$ as neural observations generally have channel noise. In this context, I represents the identity matrix and σ^2 denotes the noise variance. In this setting, we want to ask how well we can approximate $f(x)$ by function fitting on limited amounts of data. We want to describe this function so that we can predict it for all possible inputs x. This thus covers all possible behaviors of the nervous system for any stimuli and behaviors as well as the response to any internal perturbations. This setting protects us from having a fragile model that only performs well inside the specific contexts studied in an experiment.

Analytical power calculations

Linear setting
Instead of the general setting, let us operate in one in which we can describe the function as a sum of K terms of relevant basis functions:

$$f(x) = \sum_{k=1}^{K} W_k g_k(x)$$

with a set of appropriate basis functions $g_k(x)$, which may, e.g. describe synapse-synapse interactions or local dendrite interactions, and weights W_k which describe how important each of the basis functions is. Now, to be clear, real neuron transfer functions can not meaningfully be written in this form, for example because they have an output nonlinearity. However, while output nonlinearities can e.g., set half of the outputs to be zero, they are unlikely to massively affect the power calculations because real neural output nonlinearities tend to be relatively simple smooth functions(Kato et al. 2014). Importantly, in this scenario, due to the linear nature of the identification problem, we can use the well-established theory for linear systems identification to obtain solid intuitions.

How many such basis functions should we need for a neuron? No one knows. In a super simplistic world, neurons could mostly be linear, in which case we would just have one basis function for each synapse and we could get away with *K=30.* In a super complicated world, every synapse could be multiplied together into a basis function and we would need *K>>1 million.* However, in reality, complexity will be way higher than linear as we know that the neuromodulators have major modulating effects. But clearly we do not expect all combinations to meaningfully interact, e.g. because many synapses are far away from one another. As such, we may believe that the right *K* will be somewhere between *K=1,000*, allowing interactions between any pair of synapses, and *K=100,000* allowing many three-way interactions. We will thus use *K=10,000* as our estimate, knowing full well that uncertainty about this number of parameters needed to describe neurons is high.

Because noise is isotropic in our system, we can use whatever basis function system we like to do our analysis. Importantly, in the basis function of the singular vectors (components) of the $E[xx']$ system (~the system discovered by PCA), the dimensions stop being coupled to one another, and we can view the identification problem as K uncoupled estimates. Instead of estimating the weights in the inconvenient original coordinate system of X where all dimensions have the same noise but complex covariance structure, we will thus estimate all our weights in the convenient coordinate system of the relevant components, where all dimensions are uncorrelated but have different noise levels. We will now call the transformed input activities with means subtracted x_i and whenever we use indices i we imply that we are in this transformed coordinate system. Importantly they are now whitened, have unit variance and zero covariance with one another.

In that coordinate system we have our weight estimates:

$$\hat{\beta}_i \; = \; < x_i y > \; / \; < x_i x_i >$$

Now, in this estimate, errors come from misestimations of the first and the second term. However, when whitening the signals, the x_i are being divided by PC_i. And for the difficult to estimate dimensions with $k >> 1$, our error estimate is entirely dominated by the second term in the standard error expansion equation. This term has noise level of

$$\sigma_i \approx \frac{\sigma}{PC_i \sqrt{L}}$$

If all PCs are of the same size (say 1), then we obtain the well known result:

$$||\beta - \hat{\beta}||_2 \; = \; \sigma \frac{\sqrt{K}}{\sqrt{L}}$$

But if the PCs are distinct as they always are in practice, we instead obtain a variance correction

In practice, PC spectra are extremely heavy-tailed. If we record even just a few hundred neurons, we generally find that the smallest PCs are smaller than the largest PCs by a factor of thousands. As such, estimation becomes very difficult, in particular given that $\sigma \approx PC_1$ in most systems. The intuition for all this is simple, small singular values have only a tiny bit of associated variance. For example, if we have two convergent input neurons that are strongly correlated with one another then the weight associated with the first PC (the average of the two neurons) is easy to estimate while the weight associated with the second PC (the difference of their contributions to their shared target) is hard to estimate as it hard to assign credit to either of the neurons.

This is where stimulation comes in. If we randomly stimulate in the space of the g_i then we are effectively adding the identity matrix to the covariance matrix. Stimulation thus makes all singular values become similar to one another, and reduces the difference between the large and the small singular values. In relevant simulations stimulation can typically get this correction factor (condition number *c*) to be roughly 1 (or at least not larger than 3), see below. Let us go back to the example of the two correlated input neurons. The stimulation will make them less correlated, making it easier to assign credit to either of the inputs.

So the calculations of power are relatively simple from a mathematical perspective. Suppose an analysis of one neuron with just one input will take, say, 10 seconds (enough for order 100 observations, which we will call Δt_0) to identify the transfer function. If a neuron has K parameters we will instead need $c^2 K \Delta t_o$. We assume a worst-case c=3 and hence that we would need 900,000 seconds = 15,000 minutes ≈ 250 days. In other words, we would need massive but tractable amounts of data.

Appendix 2: empirical power calculations

Here we compute empirical results of a timeseries linear dynamical model on a large sample (N≈1000) of randomly generated connectomes. Dynamics at time-step *t+1* are defined as:

$$Y_t = logistic((A \times X_t + \sigma_t) - \mu)$$

Where A is the (static) connectome adjacency matrix, X_t is the input-state at time t, σ is a variability (standard deviation) which is generated anew at each timestep, and μ, a threshold / offset value which remains constant during the simulation.

Taking from biology the understanding that such dynamics operate at the millisecond timescale, we can calculate that one timestep in our simulation corresponds roughly to 20ms of real-world wallclock time, and therefore a simulation of 1,000 timesteps corresponds to 50 seconds (the purple line in **Figure X**).

We randomly sample the number of perturbations and compute the recording time to evenly space these perturbations across the duration. Connectivity (density of A) is sampled as a uniform value between .01 and 0.5, broadly accommodating the synaptic density of *C. elegans* (0.06) (Matelsky et al. 2021). The σ parameter is randomly sampled along the log distribution from [0.01-10.0). μ was set to 0.5 for these simulations, and the scale of the perturbations, as a ratio of the number of neurons randomly influenced in each simulation (one value per simulation) was randomly uniformly sampled on the interval [0.01-0.5].