Correlation-based model of artificially induced plasticity in motor cortex by a bilateral Brain-Machine Interface

Guillaume Lajoie*,[1] Nedialko I. Krouchev*,[2] John F. Kalaska,[2] Adrienne Fairhall,[1,3,4] and Eberhard E. Fetz[1,3]

[1] *University of Washington Institute for Neuroengineering, Seattle, Washington, USA*
[2] *Groupe de recherche sur le système nerveux central, Département de physiologie, Université de Montreal, Montreal, Quebec, Canada*
[3] *University of Washington, Dept. of Physiology and Biophysics, Seattle, Washington, USA*
[4] *University of Washington, Dept. of Physics, Seattle, Washington, USA*

ABSTRACT GOES HERE
We do many research things, very excellent. Bravo!

INTRODUCTION

G: Below is the original text from Ned's draft. To be adapted later on.

MODEL DESCRIPTION

As mentioned earlier, we adopt an approach that aims at using modelling mechanisms that are as simple as possible while still capturing the main features of experiments involving bilateral brain-computer interfaces [1] [All the relevant Fetz citations]. The goal here was to design a simple theoretical framework which would serve as a predictive tool for the design of next generation experiments involving artificially induced plasticity in cortex. As such, we opted to use a framework based on statistics of neural populations present in motor cortex. The hope here is that experimental measurements would be able to establish some basic statistics that would be used to fit the model.

- Say what model needs to capture:
  - Basic firing rates and activations.
  - Basic cross-correlations between MC neurons.
  - IMPORTANTLY: need to have dynamics synapses, evolving under a common plasticity rule
- Discuss the challenges of having recurrent connectivity with plastic synapses: homeostatic stability is hard to achieve, especially with different assemblies and internal correlations.
  - Add a bunch of citation and say how they got to stability.
  - Say that a bunch of model use unrealistic rescaling and that the approaches are often very sensitive and hard to reproduce.
- Now describe the work from VonHemmen, Gerstner, Burkit and Gilson in establishing a correlation-based framework.
- Say that we adapt from them.
- Describe that in what follows, a specific stochastic model for spike generation of neurons is not necessary.
- However, we describe and implement a spiking network that contains Poisson neurons.
- Briefly discuss that in principle, one could implement everything we discuss below using IF neurons while using Transfer functions derived from a Fokker-Planck formalism. Cite Doiron and Longtin, Ocker et al.

Network dynamics

Plastic synaptic connectivity

We consider a network of N model neurons whose underlying connectivity is stored in an $N \times N$ adjacency matrix J^0. Here, $J^0_{ij} = 1$ if there is a connection from neuron j to neuron i and J^0_{ij} otherwise. The matrix J^0 determines the presence of synapses in the network but their strength is stored in the $N \times N$ matrix J — J_{ij} will determine how strongly the activity of neuron j affect that of neuron i and $J^0_{ij} = 0 \Rightarrow J_{ij} = 0$. As we will see below, the weights J_{ij} will be allowed to change according to plasticity rules but the matrix J^0 remains immutable. In effect, J^0 determines the network's skeleton: only existing synapses are allowed to change and no new synapses can be formed. We consider sparsely connected random networks with every connection J^0_{ij} selected independently with probability $p = 0.3$. For all other forthcoming parameters, specific values used throughout this article are listed in Table X for reference.

As will be establish in more detail below, all neurons in the network emit spike trains that interact with each other according to synaptic strengths J_{ij}. In addition to spiking dynamics, the synapses themselves are allowed to change over time. Indeed, for each pair of pre and post-synaptic spike times, the synapse J_{ij} will be incrementally changed according to the Spike-time Dependent Plasticity (STDP) rule W:

$$W(\Delta t, J_{ij}) = \begin{cases} f^+(J_{ij})W^+(\Delta t) & \Delta t < 0 \\ -f^-(J_{ij})W^-(\Delta t) & \Delta t > 0 \end{cases} \tag{1}$$

where $\Delta t = t_{pre} - t_{post}$. For each pair of spikes, the synapse is updated: $J_{ij} \to J_{ij} + \Delta J_{ij}$ where $\Delta J_{ij} = \eta W(t_{pre} - t_{post}, J_{ij})$ where $\eta \ll 1$ is a synaptic change rate. The functions $W^+(\Delta t)$ and $W^-(\Delta t)$ describe the potentiation and depression of a synapse respectively, as a function of spike-timing. We assume Hebbian plasticity and require that these functions be both positive and decay to zero as $|\Delta t|$ grows, so that a pre-synaptic spike preceding a post-synaptic one leads to synaptic potentiation while the opposite leads to depression, in accordance to experimental estimates for excitatory STDP in cortex [CITATIONS - Bi and Poo perhaps (?)]. We note that many choice of $W^\pm$ may lead to similar results but we choose

$$\begin{aligned} W^+(\Delta t) &= A^+ \frac{|\Delta t|}{\tau^+} \exp(-|\Delta t|/\tau^+) \\ W^-(\Delta t) &= A^- \frac{|\Delta t|}{\tau^-} \exp(-|\Delta t|/\tau^-) \end{aligned} \tag{2}$$

where A^+ and A^- indicate maximal increments and τ^+ and τ^- are decay time constants. Experimental evidence suggests that W^+ and W^- are not identical so that STDP is not symmetric. Indeed, it has been observed that the maximal potentiation increment is often larger than its depression counterpart ($A^+ > A^-$) while depression has a longer decay than potentiation ($\tau^- > \tau^+$) [CITE]. We hence choose these parameters accordingly and refer the reader to TABLE X in the appendix for all model parameter values.

Finally, the weight-dependent multipliers $f^+(J_{ij})$ and $f^-(J_{ij})$ control the size of plastic increments as a function of synaptic strength and decay to zero as J_{ij} approaches set bounds J_{max} and J_{min} respectively. This way, for η small enough, synaptic weights can only asymptotically approach these bounds under Equation (1). Here, we choose these functions to be

$$\begin{aligned} f^+(J_{ij}) &= \left(1 - \frac{J_{ij}}{J_{max}}\right)^\gamma \\ f^-(J_{ij}) &= \left(\frac{J_{ij}}{J_{max}}\right)^\gamma \end{aligned} \tag{3}$$

where the exponent γ controls the strength of the weight-dependence. We note that rather than introducing unrealistic rate dependent plasticity terms and/or imposing periodic rescaling of synapses, as is often done in the modelling literature, the use of weight-dependent terms in Equation (1) will give rise to stable connectivity equilibria akin to homeostasis via a biologically plausible mechanism. Indeed, it is evident that synapses cannot be infinitely potentiated and although the experimental evidence is slim, it is reasonable to believe that as a synapse potentiates, it becomes harder to potentiate it further as resources to do so saturate (e.g. availability of neurotransmitters and vesicles, synaptic cleft volume, etc.).

In summary, the plasticity rule given by Equation (1) is the only rule governing synaptic changes throughout the network and will be sufficient to capture key mechanisms leading to stable connectivity structure.

Spiking model

Formally, if neuron i fires spikes at times $\{t_i^s\}$, then its spike train can be represented by a function of time: $S_i(t) = \sum_{t_i^s} \delta(t - t_i^s)$ where $\delta(t)$ is the Dirac-δ function. The advantage of $S_i(t)$ is that it can be treated as a local probability mass over small time intervals δt. That is, the probability of observing a spike from neuron i in the time window $(t - \delta t, t)$ is given by $\int_{t-\delta t}^{t} ds S_i(s)$ (as long as two spikes do not lie in the same interval). In the case where spike times are known, $S_i(t)$ is well defined and such probabilities are either 1 or 0. However, if our goal is to generate spike times that obey certain interaction rules and statistics, it is convenient to think as spike trains as true probability densities over time. To borrow notation from Physics, we are interested in the *ensemble averages* $\langle S_i(t) \rangle$ which can be interpreted as the expected probability over repeated trials of the same probabilistic process. From here on out, we assume that the ensemble averages $\langle S_i(t) \rangle$ can be written in the form of a density $\lambda_i(t)$ and that this density represents an instantaneous Poisson firing rate.

Although seemingly superfluous, these mathematical details are mentioned here in order to stress that the probabilistic framework we derive need not necessarily be restricted to a Poisson process. We chose to do so for the sake of clarity and numerical simulations but most of the derivations we make below are generalizable by replacing $\lambda_i(t)$ by $\langle S_i(t) \rangle$. In addition, other probabilistic or dynamical models of network spiking activity could be well described by our findings, such as Intergate-and-fire with the use of stochastic transfer functions derived from Fokker-Planck Equations as is done for small network motifs in [Doiron et Longtin, Ocker et Doiron, etc...]. We refer the reader to [Kempter, Burkit, Gilson] for more details.

In the network, the spiking probability density of neuron i ($\lambda_i(t)$) is influenced by two factors: (i) a time-dependent external driving rate $\nu_i(t)$ (ii) every preceding spike from pre-synaptic neurons j. We also introduce two types of delays associated to each synapse that are independently sampled from uniform distributions: an axonal delay $d_{ij}^a \sim \bar{d}^a \pm \delta d^a$ and a dendritic delay $d_{ij}^d \sim \bar{d}^d \pm \delta d^d$. A pre-synaptic spike fired from neuron j at t_j^s will reach the synapse J_{ij} at $t_{pre} = t_j^s + d_{ij}^a$ and likewise, a post-synaptic spike from neuron i at t_i^s will be delayed in it's effect on the upstream synapses J_{ij} so that $t_{post} = t_i^s + d_{ij}^d$. As a result the dynamics of the network are completely determined by:

$$\lambda_i(t) = \nu_i(t) + \sum_j \sum_{t_j^s} J_{ij}(t) \epsilon(t - t_j^s - d_{ij}^a) \tag{4}$$

where $\epsilon(t) = H(t)\frac{1}{\tau}e^{-t/\tau}$ is a synaptic filter with H denoting the Heaviside function and τ a synaptic time constant ($\tau = 5$ms). Importantly, note that this filter is causal ($\epsilon(t) = 0$ for $t < 0$) and that it is normalized so that $\int_{-\infty}^{\infty} \epsilon(t) dt = 1$. We simulate (4) numerically by discretizing time in small bin of size δt and independently drawing spikes for neuron i in bin $[t, t + \delta t]$ according to the probability $\lambda_i(t)\delta t$. See Appendix: numerical methods for more details.

The synapses $J_{ij}(t)$ evolve as follows. For every spike t_j^s fired by neuron j, for all post-synaptic neurons i (i.e. all i such that $J_{ij}^0 = 1$) and all pre-synaptic neurons k (i.e. all k such that $J_{jk}^0 = 1$), the following updates are made:

$$\begin{aligned} J_{ij} &\to J_{ij} + \eta \sum_{t_i^s < t_j^s + (d_{ij}^a - d_{ij}^d)} W([t_j^s + d_{ij}^a] - [t_i^s + d_{ij}^d], J_{ij}) \\ J_{jk} &\to J_{jk} + \eta \sum_{t_k^s < t_j^s + (d_{jk}^d - d_{jk}^a))} W([t_k^s + d_{jk}^a] - [t_j^s + d_{jk}^d], J_{jk}). \end{aligned} \tag{5}$$

Together, Equations (4) and (5) describe the evolution of both the spiking and the synapses in the network. It is clear that both are inter-dependent, although on different time-scales, and will greatly depend on external rates $\nu_i(t)$.

Functional groups and closed-loop stimulation

We divide the neurons in our network in three groups of equal size $N/3$: group $a = \{1, \dots, \frac{N}{3}\}$, group $b = \{\frac{N}{3}+1, \dots, \frac{2N}{3}\}$ and group $c = \{\frac{2N}{3}+1, \dots, N\}$. Every neuron i in a given group $\alpha_i = a, b, c$ is subject the the same external modulating rate so that $\nu_i(t) = \nu_{\alpha_i}(t)$. For our poisson network, this means that in the absence of network effects, neurons from the same group still fire independently but do so according to the same modulated rate $\nu_\alpha(t)$.

We introduce a closed-loop, spike-triggered stimulation protocol with the following mechanism. We assign group a to be the "Recording" site, group b to be the "Stimulation" site and c to be the "Control" site. We choose neuron $i = 1$ in group a, and trigger the stimulation of every neuron in group b on spikes from that neuron. For a given stimulation delay $d^\dagger$, we impose the following dynamics:

$$\begin{cases} \lambda_i^\dagger(t) = \nu_b(t) + \sum_j J_{ij}(t) \sum_{t_j^s} \epsilon(t - t_j^s - d_{ij}^a) + \sum_{t_1^s} \delta(t - t_1^s - d^\dagger) & ; \quad i \in b \\ \lambda_i^\dagger(t) = \nu_{\alpha_i}(t) + \sum_j J_{ij}(t) \sum_{t_j^s} \epsilon(t - t_j^s - d_{ij}^a) & ; \quad i \notin b \end{cases} \tag{6}$$

where † indicates the presence of stimulation.

RESULTS

Below, we investigate how the dynamics of our network, both on the spiking and synaptic level, depend on statistics of $\nu_\alpha(t)$ and stimulation delay $d^\dagger$. To do so, we develop analytical expressions describing synaptic dynamics and validate our findings with numerical simulations.

Stable synaptic equilibria determined by correlations

Separation of timescales

Following the framework developed by Gilson et al., we assume that the timescale on which impactful plastic changes occurs is much longer than the spiking activity timescale, hence the requirement that $\eta \ll 1$. This allows to effectively separate the spiking dynamics from the synaptic dynamics by assuming that over a reasonably long time period of length T, synaptic weights $J_{ij}(t)$ are approximately constant ($\approx J_{ij}$). As a result, it is possible to derive dynamic equations for the synaptic connectivity matrix J on the timescale given by T. The main ingredients that will affect connectivity changes are statistical dependencies between the firing of pairs of cells; namely their relative firing rates and cross-correlation.

As defined in Equation (4), $\lambda_i(t)$ is composed of a instantaneous rate $\nu_i(t)$ and the additive effect of synaptically filtered spikes from other neurons in the network. The latter come from specific realizations of processes λ_j up to time t. In what follows, we consider the ensemble average over all these realizations to get the expected rate $\langle \lambda_i(t) \rangle$. For the sake of convenience, we drop the brackets and write

$$\lambda_i(t) = \nu_i(t) + \sum_j J_{ij}(\lambda_j * \epsilon_{ij}^a)(t) \tag{7}$$

where we dropped the time-dependence of J_{ij} and we define the delayed convolution as

$$(\lambda_j * \epsilon_{ij}^a)(t) \equiv \int_0^t ds \lambda_i(s) \epsilon(t - d_{ij}^a - s).$$

Calling on the separation of timescales described earlier, we assume for $\eta \ll 1$ that synaptic weights J_{ij} stay constant over *plasticity epochs* of length T. This way, we use Equation (7) defined over such epochs ($[t-T, t]$) to compute the accumulated synaptic weight changes.

The density count of spike pairs separated by $u = t_j^s - t_i^s$ over the epoch $[t-T, t]$ is given by the (un-normalized) cross-correlation between neurons i and j: $\int_{t-T}^t ds \lambda_i(s) \lambda_j(s+u)$. As such quantities will be called upon often in what

follows, we denote the cross-correlation matrices for both neural and external firing intensities for an epoch ending at time t as

$$\begin{aligned} C(u;t) &= \int_{t-T}^{t} ds\lambda(s)\lambda(s+u)^T \\ \hat{C}(u;t) &= \int_{t-T}^{t} ds\nu(s)\nu(s+u)^T \end{aligned} \tag{8}$$

where $\lambda(t) = (\lambda_1(t), \ldots, \lambda_N(t))^T$ with the "T" denoting matrix transposition and where $\nu(t)$ is similarly defined. Therefore, the expected change of synaptic weight J_{ij} over the epoch $[t-T, t]$ is given by

$$\langle \Delta J_{ij}(t) \rangle = \int_{-\infty}^{\infty} du C_{ij}(u + d_{ij}^d; t) W(u - d_{ij}^a, J_{ij}(t-T)). \tag{9}$$

We will see later how this expression becomes a crucial component of our results.

Self-consistent dynamics

In order to derive equilibrium properties for our network, we would like an expression for the firing intensity vector $\lambda(t)$ which only involves to knowledge of synaptic weights and external rates $\nu(t)$. In other words, we would like to solve for $\lambda(t)$ in Equation (7). The problem arises because of the convolution with the synaptic filter ϵ which prevents a direct isolation of rates $\lambda_j(t)$ from the right-hand side. In Gilson et al., the authors circumvent this problem by taking a Laplace transform and performing calculations in Laplace space. This enables a self-consistent equation for $\lambda(t)$. However, important assumptions about the shape of external cross-correlations $\hat{C}$ were assumed so that inverse transforms were possible. In our case, we need $\hat{C}$ to remain as general as possible which makes solving a self-consistent equation exactly very challenging.

Nevertheless, we find that a simple first-order expansion of Equation (7) is an adequate estimate. Substituting $\lambda_j(t)$ by it's own expression in (7) yields

$$\begin{aligned} \lambda_i(t) &= \nu_i(t) + \sum_j J_{ij} \left[\nu_j(t) + \sum_k J_{jk} \lambda_k(t) * \epsilon(t - d_{jk}^a) \right] * \epsilon(t - d_{ij}^a) \\ &= \nu_i(t) + \sum_j J_{ij} \nu_j(t) * \epsilon(t - d_{ij}^a) + \sum_j J_{ij} \cancel{\left[\sum_k J_{jk} \lambda_k(t) * \epsilon(t - d_{jk}^a) \right] * \epsilon(t - d_{ij}^a)} \end{aligned}$$

where we truncate to first order in J_{ij} for two reasons: (i) the sparsity of J along with $J_{ij} < 1$ enforce that terms involving powers of J will decay to zero fast (ii) the multiple convolution with delayed synaptic filters attenuates the effect spiking in the past. It follows that

$$\lambda_i(t) \simeq \nu_i(t) + \sum_j J_{ij} (\nu_j * \epsilon_{ij}^a)(t). \tag{10}$$

We now have an expression for spiking probabilities throughout the network in terms of external rates only. Evaluating synaptic changes for given external rates $\nu(t)$ becomes as simple as computing the cross-correlation matrix $C(u;t)$ over a given epoch using expression (10), then substituting this in Equation (9) to get estimates for $\langle \Delta J_{ij} \rangle$. Reference figure and simulation that shows that this is a valid estimate. Introduce the function $F(x) = \int_{-\infty}^{\infty} du C_{ij}(u + d_{ij}^d; t) W(u - d_{ij}^a, x)$ which gives a landscape of the ΔJ_{ij}.

Correlation-based dynamics and equilibria

So far, we derived an expression for firing intensities of connected neurons $\lambda(t)$ based on known external rates $\nu(t)$. This is enough to predict both the dynamics and equilibria of the synaptic matrix $J(t)$. However, experiments often only inform us about activity statistics and it would be ideal to obtain an expression for ΔJ_{ij}, and thus the evolution of $J(t)$, as a function of the input $\nu(t)$'s cross-correlation only: $\hat{C}(u;t)$.

Consider the cross-correlation between pre-synaptic neuron j and post-synaptic neuron i: $C_{ij}(u;t)$. Substituting (10) into (8), we get

$$\begin{aligned} C_{ij}(u;t) &= \int_{T-t}^{t} ds \left[\nu_i(s) + \sum_k J_{ik}(\nu_k * \epsilon^a_{ik})(s)\right]\left[\nu_j(s+u) + \sum_l J_{jl}(\nu_l * \epsilon^a_{jl})(s+u)\right] \\ &= \hat{C}_{ij}(u;t) + \int_{t-T}^{t} ds \left[\left\{\nu_i(s)\sum_l J_{jl}(\nu_l * \epsilon^a_{jl})(s+u)\right\} + \left\{\nu_j(s+u)\sum_k J_{ik}(\nu_k * \epsilon^a_{ik})(s)\right\} \ldots \right. \\ &\left. + \left\{\left(\sum_k J_{ik}(\nu_k * \epsilon^a_{ik})(s)\right)\left(\sum_l J_{jl}(\nu_l * \epsilon^a_{jl})(s+u)\right)\right\}\right]. \end{aligned}$$

To simplify this expression, we make two assumptions. First,

$$\int_{t-T}^{t} ds \nu_i(s)(\nu_j * \epsilon^a_{ij})(s) \simeq \int_{t-T}^{t} ds \nu_i(s)\nu_j(s+u-d^a_{ij})$$

which is justified for T large enough and $\nu(t)$ fluctuating on longer timescales than τ — the time-constant of ϵ. Second, we replace individual delays d^a_{ij} by their mean $\bar{d}^a$. We will see later that this simplification is justified once we take averages over large groups of neurons. These two substitutions enable a matrix notation for network cross-correlations:

$$C(u;t) \simeq \hat{C}(u;t) + J\hat{C}(u+\bar{d}^a;t) + \hat{C}(u-\bar{d}^a,t)J^T + J\hat{C}(u;t)^T J^T \tag{11}$$

It follows from (9) and (11) that the expected change of synaptic strengths matrix is given by the self-consistent expression:

$$\langle \Delta J(t) \rangle \simeq \int_{-\infty}^{\infty} du C(u+\bar{d}^d;t) \cdot W\left(u-\bar{d}^a, J(t-T)\right) \tag{12}$$

where $\cdot$ denotes element-wise multiplication. Expressions (11) and (12) together imply that we can compute the expected synaptic increment of any matrix J, given the cross-correlation structure of external driving rates. We now describe how this is used to find stable synaptic equilibria.

Consider the following function

$$F(M,J;\hat{C}) \equiv \int_{-\infty}^{\infty} du \left[\hat{C}(u) + J\hat{C}(u+\bar{d}^a) + J\hat{C}(u-\bar{d}^a)^T + J\hat{C}(u)^T J^T\right] \cdot W\left(u-\bar{d}^a, M\right) \tag{13}$$

derived from Equations (11) and (12), giving plastic increments of all entries a the synaptic matrix M (also $N \times N$), given an epoch of network activity with external cross-correlations $\hat{C}(u)$ and connectivity J. For convenience, we assume that external inputs have stationary statistics that are known to us so that $\hat{C}(u)$ is fixed. We use F to define a discrete, state-dependent dynamical system on the space of $N \times N$ matrices describing the evolution of our network's synapses from one activity epoch to the next:

$$J^{n+1} = J^n + \eta F(J^n, J^n; \hat{C})$$

where J^n denotes J after the n^{th} epoch. Thus, a synaptic equilibrium is a matrix J^* for which $F(J^*, J^*; \hat{C}) = \mathbf{0}$, the $N \times N$ zero matrix.It is unclear if such a matrix exists but we can define $M(J)$ such that it works and show that it is close to diag.

Averaging over groups

We now restrict ourselves to group-averaged quantities as this greatly simplifies calculations and gives remarkably precise estimates for network where N is large enough. Consider the 3×3 matrix $\bar{J}$ whose entries $\bar{J}_{\alpha\beta}$ represent the average strength of a non-zero synapse from a neuron in group β to a neuron in group α. For ease of notation, we write $\bar{J}_{ab}$ instead of $\bar{J}_{12}$, etc. Recall that p is the probability that any two cells are connected which means that $\frac{pN}{3}\bar{J}_{\alpha\beta}$ is the average strength of total synaptic inputs from group β to a neuron in group α. Following our previous

derivations using $\bar{\lambda}_\alpha(t) =\simeq \nu_\alpha(t) + \sum_\beta \frac{pN}{3}\bar{J}_{\alpha\beta}(\nu_\beta * \epsilon^a)(t)$, we have

$$\begin{aligned}\bar{C}(u,t) &\simeq \hat{\bar{C}}(u;t) + \frac{pN}{3}\left[\bar{J}\hat{\bar{C}}(u+\bar{d}^a;t) + \bar{J}\hat{\bar{C}}(u-\bar{d}^a,t)^T\right] + \frac{p^2N^2}{9}\bar{J}\hat{\bar{C}}(u;t)^T\bar{J}^T \\ \langle\bar{J}(t)\rangle &\simeq \int_{-\infty}^{\infty} du\bar{C}(u+\bar{d}^d;t)\cdot W\left(u-\bar{d}^a,\bar{J}(t-T)\right)\end{aligned} \tag{14}$$

Describe using figures that this is a good approximation.

Synaptic equilibria

Talk about different iteration schemes.

Closed-loop stimulation: drift rates and new stable equilbria

Averaging over groups

Equation (6) describes the effect of spike-triggered stimulation of neural group b. It is easily implementable for numerical simulations but a conceptual difficulty arises when attempting to transform it into an expression for the densities $\lambda_i^\dagger(t)$ similar to (10). The last term involves a particular realization of spike times from neuron 1 and must be fixed across all neurons i in group b. This is akin to a Hawkes process [REF] and cannot be replace by a convolution involving $\lambda_1(t)$. Fortunately, it is possible to directly modify the cross-correlation functions to accurately capture stimulated interactions.

Similarly to Equation (14), we wish to derive an expression for the network's reduced cross-correlation under spike-triggered stimulation: $\bar{C}^\dagger(u;t)$. We derive first order modifications applied directly to $\bar{C}(u;t)$ (i.e. assuming the network-induced effects due to extra added spikes do not affect the correlations). Consider a specific spike train realization from neuron 1, $S_1(t) = \sum_{t_1^s}\delta(t-t_1^s)$ where $\{t_1^s\}\sim\lambda_1(t)$, and let $\bar{\lambda}_b^\dagger(t) = \bar{\lambda}_b(t) + S_1(t-d^\dagger)$. It follows that

$$\int_{t-T}^t ds\bar{\lambda}_b^\dagger(s)\bar{\lambda}_\alpha(s+u) = \bar{C}_{b\alpha}(u;t) + \int_{t-T}^t dsS_1(s-d^\dagger)\bar{\lambda}_\alpha(s+u).$$

To get a general expression for cross-correlations when stimulation is turned on, the last term in this expression needs to be averaged over all possible realizations of $S_1(t)$, hence taking a second ensemble average. To avoid cumbersome notation, we describe three general case that need to be considered.

If $\alpha = c$: there are no direct copies of $S_1(t)$ present in the group so that

$$\int_{t-T}^t dsS_1(s-d^\dagger)\bar{\lambda}_c(s+u) = \bar{C}_{ac}(u+d^\dagger;t).$$

If $\alpha = a$: $S_1(t)$ is itself one of $N/3$ processes from that group, thus

$$\int_{t-T}^t dsS_1(s-d^\dagger)\bar{\lambda}_a(s+u) = \left(1-\frac{3}{N}\right)\bar{C}_{aa}(u+d^\dagger) + \frac{3}{N}\left[\int_{t-T}^t dsS_1(s)\right]\delta(u+d^\dagger).$$

If $\alpha = b$: $S_1(t)$ is present in every spike train from group b and there is therefore an additional factorization to consider:

$$\begin{aligned}\int_{t-T}^t ds\bar{\lambda}_b^\dagger(s)\bar{\lambda}_b^\dagger(s+u) &= \bar{C}_{bb}(u;t) + \int_{t-T}^t ds\left[S_1(s-d^\dagger)\bar{\lambda}_b(s+u) + \bar{\lambda}_b(s)S_1(s-d^\dagger+u)\right] + \int_{t-T}^t dsS_1(s-d^\dagger)S_1(s-d^\dagger+u) \\ &= \bar{C}_{bb}(u;t) + \bar{C}_{ab}(u+d^\dagger;t) + \bar{C}_{ba}(u-d^\dagger,t) + \left[\int_{t-T}^t dsS_1(s)\right]\delta(u).\end{aligned}$$

We can replace the spike counts of $S_1(t)$ over the t-epoch by $\rho(t) \simeq \int_{t-T}^{t} dsS_1(s)$ where $\rho(t) = \bar{\nu}(t) + \frac{pN}{3}\bar{J}(t)\bar{\nu}(t)$ with $\bar{\nu}(t) = \int_{t-T}^{t} ds\nu(s)$. We drop the dependence on t for clarity and get the following expression for $\bar{C}^{\dagger}(u)$:

$$
\begin{array}{|l|}
\hline
\bar{C}^{\dagger}_{ba}(u) = \bar{C}_{ba}(u) + (1-\frac{3}{N})\bar{C}_{aa}(u+d^{\dagger}) + \frac{3}{N}\rho_a T\delta(u+d^{\dagger}) \\
\bar{C}^{\dagger}_{ab}(u) = \bar{C}_{ab}(u) + (1-\frac{3}{N})\bar{C}_{aa}(u-d^{\dagger}) + \frac{3}{N}\rho_a T\delta(u-d^{\dagger}) \\
\bar{C}^{\dagger}_{bb}(u) = \bar{C}_{bb}(u) + \bar{C}_{aa}(u) + \bar{C}_{ba}(u-d^{\dagger}) + \bar{C}_{ab}(u+d^{\dagger}) + \delta(u)(\rho_a T - \bar{C}_{aa}(0)) \\
\bar{C}^{\dagger}_{cb}(u) = \bar{C}_{cb}(u) + \bar{C}_{ca}(u-d^{\dagger}) \\
\bar{C}^{\dagger}_{bc}(u) = \bar{C}_{bc}(u) + \bar{C}_{ac}(u+d^{\dagger}) \\
\bar{C}^{\dagger}_{aa}(u) = \bar{C}_{aa}(u) \\
\bar{C}^{\dagger}_{cc}(u) = \bar{C}_{cc}(u) \\
\bar{C}^{\dagger}_{ca}(u) = \bar{C}_{ca}(u) \\
\bar{C}^{\dagger}_{ac}(u) = \bar{C}_{ac}(u) \\
\hline
\end{array}
$$

The above expression combined with Equation (14) implies that we can compute the expected synaptic increment of any matrix $\bar{J}$, with and without stimulation, given the cross-correlation structure and mean of external driving rates. We note that a similar derivation is straightforward for the entire matrix J but that the notation is quite cumbersome. In summary, this is a powerful tool to investigate the effect of different stimulation schemes under varied activity regimes as we now demonstrate.

Better table description of BCI xcorss

$\hat{C}(u)$	$\mathbb{I} + A \cdot A^T$
$\hat{C}(u-d^{\dagger})$	
$\hat{C}(u-\bar{d}^{a})$	
$\hat{C}(u+d^{\dagger})$	
$\hat{C}(u+\bar{d}^{a})$	
$\hat{C}(u-\bar{d}^{a}-d^{\dagger})$	
$\hat{C}(u+\bar{d}^{a}+d^{\dagger})$	
$\hat{C}(u-\bar{d}^{a}-2d^{\dagger})$	
$\hat{C}(u+\bar{d}^{a}+2d^{\dagger})$	
$\hat{C}(u-2d^{\dagger})$	
$\hat{C}(u+2d^{\dagger})$	
$\hat{C}(u-\bar{d}^{a}+d^{\dagger})$	
$\hat{C}(u+\bar{d}^{a}-d^{\dagger})$	

Bring in figures and compariosn of X corrs to make a point

Convergence to and from artificial equilibria

Describe dynamics when stim is turned on or off

Effectiveness of stimulation as a function of delay and activation statistics

- Discuss the effect of varying the delay $d^{\dagger}$.
- Discuss the effect of varying the Xcorrs of activation patterns themselves

DISCUSSION

BLABLABLA

ACKNOWLEDGMENTS

People we like and helped a bunch go here.

[1] A. Jackson, S. N. Baker, and E. E. Fetz, Journal of Physiology-London **573**, 107 (2006).