Metastability in a Stochastic Neural Network Modeled as a Velocity Jump Markov Process

Paul C Bressloff

Department of Mathematics, University of Utah

June 26, 2014
**Multiscale Dynamics**

Brain dynamics is noisy at the single cell level...but often observe coherent states at the macroscopic level.

- Noisy spike trains
- Coherent waves and oscillations at network level
SINGLE CELL RECORDINGS

Single cell recordings in vivo suggest that individual cortical neurons are noisy with inter-spike intervals (ISIs) close to Poisson (Softy and Koch 1993)

Electrophysiological Data

For our analysis, we used a subset of these trials. We rejected single units in area MT (or V5), a region of extrastriate visual cortex concerned with motion processing (M. muhta; Newsome and Pare, 1988). The amount of motion coherency, as well as the direction of the influence of the nonclassical receptive field (RF) on single-unit activity in the primary visual cortex of two alert and behaving monkeys, whose heads were restrained, had to fixate a cross. If fixation was broken—as monitored by a search coil—the trial was terminated.

Care was taken to record only single-unit activity. Extracellular spike trains were recorded from cells in visual cortices of awake adult macaques. Our primary interest was neither in the nature of the stimuli used nor in the cells' selective responses to these stimuli, but only in the statistical properties of neuronal firing.

In general, we did not find any significant difference between the degree of variability of V1 or MT neurons. Therefore, except when otherwise explicitly noted, we will lump these two sets of neuron data together.

The two-threshold window discriminator produced pulses corresponding to single action potentials whose time of arrival was recorded with 1 msec resolution. Care was taken to ensure that the two-threshold window pulses were not spaced by less than 0.5 msec. Only one of these cells showed any bursting activity (as defined below), and was rejected.

The spikes following the stimulus onset arrived at times $t_i$. Thus, the interspike interval (ISI) is $S_i = t_i - t_{i-1}$, where $S_i$ is the number of spikes in the train. The other parameters are the mean of the histogram (the average interspike interval), and the standard deviation about that mean, which is calculated as $\sigma = \sqrt{\frac{1}{N-1} \sum (S_i - \bar{S})^2}$, where $N$ is the number of spikes in the train.

We will analyze histograms of these ISIs through two of their parameters. One is the mean of the histogram (the average interspike interval). The other parameter is the standard deviation about that mean, which is calculated as the square root of the variance $\sigma = \sqrt{\frac{1}{N-1} \sum (S_i - \bar{S})^2}$, where $N$ is the number of spikes in the train.

The third parameter is the coefficient of variation $\frac{\sigma}{\bar{S}}$, ranging between 0.5 and 1.0. We attempt to understand the origin of these values by two different theoretical methods: modified scale spaces and models, compartmental models, and simulations of detailed compartmental models of cortical pyramidal cells. Our analysis revealed a strong contradiction between the large observed interspike interval variability and the predictions of well-accepted analytical and biophysical single-cell models, including the models of Wilbur and Rinzel (1983) and Bugmann (1990).

The second set of data (referred to in the following simply as MT data) was recorded during an investigation into the relationship between motion discrimination and the behavior of single units in area MT (or V5), a region of extrastriate visual cortex concerned with motion processing (M. muhta; Newsome and Pare, 1988). Data were only accepted for trials during which the monkey fixated or performed a fixation-related task. Some et al., 1989b; Britten et al., 1992). In brief, three monkeys were trained to report the direction of motion of a random dot display in which a fixed fraction of dots (the amount of "motion coherency") is moving in a specific direction and the remaining dots are static. The two-threshold window discriminator produced pulses corresponding to single action potentials whose time of arrival was recorded with 1 msec resolution. Care was taken to ensure that the two-threshold window pulses were not spaced by less than 0.5 msec. Only one of these cells showed any bursting activity (as defined below), and was rejected.

All spike trains that contained any dominant ISI characteristic of "bursting" behavior. A "bursting" neuron frequently fires a burst of spikes in the 2 msec bin as in the 5 msec bin. These criteria yielded a subset of 233 nonbursting neurons.

We used data from two different laboratories. In both cases, standard electrophysiological procedures were used to identify single-unit activity in the 2 msec bin as in the 5 msec bin. These criteria yielded a subset of 233 nonbursting neurons.

The first set of data ("V1" data) came from an investigation into the relationship between motion discrimination and the behavior of single units in area MT (or V5), a region of extrastriate visual cortex concerned with motion processing (M. muhta; Newsome and Pare, 1988). The amount of motion coherency, as well as the direction of the influence of the nonclassical receptive field (RF) on single-unit activity in the primary visual cortex of two alert and behaving monkeys, whose heads were restrained, had to fixate a cross. If fixation was broken—as monitored by a search coil—the trial was terminated.

Care was taken to record only single-unit activity. Extracellular spike trains were recorded from cells in visual cortices of awake adult macaques. Our primary interest was neither in the nature of the stimuli used nor in the cells' selective responses to these stimuli, but only in the statistical properties of neuronal firing.

In general, we did not find any significant difference between the degree of variability of V1 or MT neurons. Therefore, except when otherwise explicitly noted, we will lump these two sets of neuron data together.

The two-threshold window discriminator produced pulses corresponding to single action potentials whose time of arrival was recorded with 1 msec resolution. Care was taken to ensure that the two-threshold window pulses were not spaced by less than 0.5 msec. Only one of these cells showed any bursting activity (as defined below), and was rejected.

All spike trains that contained any dominant ISI characteristic of "bursting" behavior. A "bursting" neuron frequently fires a burst of spikes in the 2 msec bin as in the 5 msec bin. These criteria yielded a subset of 233 nonbursting neurons.

We used data from two different laboratories. In both cases, standard electrophysiological procedures were used to identify single-unit activity in the 2 msec bin as in the 5 msec bin. These criteria yielded a subset of 233 nonbursting neurons.
LOCAL FIELD POTENTIAL (LFP)

“1/f” noise in LFP recordings of parietal cortex of awake cats (Bedard et al Destexhe 2006)

Multisite bipolar LFP recordings (Destexhe et al 1999)

![Diagram of brain regions and LFP recordings](image)

- **LFPs**
  - 1, 2, 3, 4, 5, 6, 7, 8

- **Units**
  - 1, 2, 3, 4, 5, 6, 7, 8

- **ISI** (ms)
  - Wake: 500 sec, SWS: 200 sec

- **Power spectra**
  - Log-log scale (dashed line represents Poisson process)
  - **log P** vs **log f**

- **Avalanche analysis**
  - Avalanche sizes and durations
  - 1/f scaling at high frequencies
  - **α=1** for exponential distribution fits
  - **α=3** for power-law distribution fits

- **Sleep states**
  - Wake: 0.5 mV, 5 sec
  - SWS: 5 sec

- **Spectral density**
  - LFPs at 8 locations separated by 1 mm
  - Log-linear scale (Fig. 2; log-log scale in insets)

- **EEG slow waves**
  - Scales exponentially during SWS
  - Power-law behavior during waking

- **Graphs and plots**
  - Log N vs log ISI for Wake and SWS
  - Power-law distribution in SWS
  - Exponential distribution in Wake

- **Log-log plots**
  - log N vs log ISI
  - Log-log scale for ISIs, log-linear scale for frequency bands

- **Statistics**
  - Ramos et al, 2005

- **References**
  - Bedard et al, 2006
  - Destexhe et al, 1999
Neuronal avalanches (Beggs and Plenz 2003, 2004)

(a) Slice of rat somatosensory cortex. LFP measured at multiple sites (in superficial layers) by an 8 × 8 multi-electrode array with spacing 200 μm.

(b) Examples of LFP population spikes

(c) Each LFP spike represents the synchronous activity of multiple neurons in a local population
Each point represents time of occurrence of an LFP spike. Grouped into an avalanche when peaks occur in contiguous time bins of width $\Delta t_{\text{avg}}$. Avalanche terminated when there is an empty time bin.

Size of an avalanche $s$ is either the number of active electrodes or the sum of participating LFP spike amplitudes.
Distribution $P(s)$ of avalanche sizes $s$ is a heavy-tail distribution that exhibits a power law over several orders of magnitude

$$P(s) \propto s^\alpha$$

Find that $\alpha = -1.5$ irrespective of value of $\Delta t_{avg}$ (2-6 m sec)
Part I. Neural master equation
Consider $M$ homogeneous networks labelled $k = 1, \ldots, M$, each containing $N$ identical neurons.

Suppose that in the interval $[t, t + \Delta t)$, $n_k(t)$ neurons in the $k$th population fire an action potential or spike.

Define population firing rate in terms of the number of neurons that spike in the interval $\Delta t$.

\[
a_k(t) = \frac{n_k(t)}{N \Delta t}.
\]

Treat the number of active neurons $n_k(t)$ as a stochastic variable that evolves according to a one-step jump Markov process.
Rates of state transitions $n_k \rightarrow n_k \pm 1$ are chosen so that under a mean-field approximation one obtains deterministic Wilson-Cowan equations – transition rates not unique!
Master equation III

- Let $P(n, t)$ with $n = (n_1, \ldots, n_M)$ denote probability that $m_i(t) = n_i$ for all $i$.

- Probability distribution evolves according to birth-death master equation

$$\frac{dP(n, t)}{dt} = \sum_k \left[ (\mathbb{T}_k - 1) (\Omega_k^-(n)p(n, t)) + (\mathbb{T}_k^{-1} - 1) (\Omega_k^+(n)p(n, t)) \right]$$

where $\mathbb{T}_k^{\pm 1} F(\ldots, n_k, \ldots) = F(\ldots, n_k \pm 1, \ldots)$

- Transition rates are (for sigmoid function $F$)

$$\Omega_k^-(n) = \frac{n_k}{\tau_a}, \quad \Omega_k^+(n) = \frac{N \Delta t}{\tau a} F\left( \sum_l w_{kl} n_l/N \Delta t \right)$$
Mean–field approximation

- Multiply both sides of master equation by $n_k$ and sum over all states $n$. This gives

$$\frac{d}{dt} \langle n_k \rangle = \sum_{r=\pm 1} r \langle T_{k,r}(n) \rangle$$

where $\langle f(n) \rangle = \sum_n P(n, t)f(n)$ for any function of state $f(n)$.

- Assume all statistical correlations can be neglected so that $\langle T_{k,r}(n) \rangle \approx T_{k,r}(\langle n \rangle)$

- Setting $a_k = (N\Delta t)^{-1} \langle n_k \rangle$ leads to the mean–field equation

$$\tau_a \frac{d}{dt} a_k = -a_k + F \left( \sum_l \omega_{kl} a_l \right)$$
COMPARISON OF PCB AND BCC MASTER EQUATIONS

- It’s all about the bin size $\Delta t$!

- Master equation keeps track of changes in spiking activity.

- PCB model assumes that network operates in a Gaussian-like regime close to an asynchronous state for large $N$. Thus changes in population activity could be slow ie can set $\Delta t = 1$.

- BCC model assumes that the network operates in a Poisson-like regime for large $N$. Therefore, necessary to take $\Delta t \to 0$ as $N \to \infty$ with $N\Delta t = 1$ fixed.
Effects of Fluctuations

- Can adapt methods from chemical master equations in PCB model: system-size expansion in $N^{-1}$, Langevin approximation, WKB or path integral methods for metastable states.

- No small parameter ($1/N$) in BCC model. However, one can analyze the moment hierarchy using factorial moments or a loop expansion of a path integral representation.

- Both models determine how higher order correlations couple to mean-field dynamics (see also Touboul, Ermentrout...).

- BCC model has been used to analyze power law behavior in terms of directed percolation theory.
**Neural Langevin Equation I**

- Set $\Delta t = 1$ and introduce the rescaled variables $x_k = n_k / N$ and corresponding transition rates

\[
\Omega_{k,-1}(x) = \frac{x_k}{\tau_a}, \quad \Omega_{k,1}(x) = \frac{1}{\tau_a} F \left( \sum_l w_{kl} x_l \right).
\]

- Carrying out a Kramers–Moyal expansion to second order in $\epsilon = N^{-1/2}$ then leads to the multivariate FP equation

\[
\frac{\partial P(x, t)}{\partial t} = - \sum_{k=1}^{M} \frac{\partial}{\partial x_k} [V_k(x)P(x, t)] + \frac{\epsilon^2}{2} \sum_{k=1}^{M} \frac{\partial^2}{\partial x_k^2} [B_k(x)P(x, t)]
\]

with

\[
V_k(x) = \Omega_{k,1}(x) - \Omega_{k,-1}(x), \quad B_k(x) = \Omega_{k,1}(x) + \Omega_{k,-1}(x)
\]
The FP equation determines the probability density function for a corresponding stochastic process $X(t) = (X_1(t), \ldots, X_M(t))$, which evolves according to the neural Langevin equation

$$dX_k = V_k(X)dt + \epsilon b_k(X)dW_k(t).$$

with $b_k(x)^2 = B_k(x)$.

Here $W_k(t)$ denotes an independent Wiener process such that

$$\langle W_k(t) \rangle = 0, \quad \langle W_k(t)W_l(s) \rangle = \delta_{k,l} \min(t, s).$$

Langevin equation captures the relatively fast stochastic dynamics within the basin of an attraction of a stable fixed point (or limit cycle) of the corresponding deterministic rate equations.

Rigorous analysis of Langevin approximation can be carried out by extending work of Kurtz on chemical master equations (Buckwar and Riedler 2012).
Part II. WKB approximation and rare event statistics
\[
\frac{du}{dt} = -u + F(wu), \quad F(u) = \frac{f_0}{1 + e^{-\gamma(u-\theta)}}.
\]
**Birth-death master equation**

\[ \frac{dP(n,t)}{dt} = T_+(n-1)P(n-1,t) + T_-(n+1)P(n+1,t) - (T_+(n)+T_-(n))P(n,t) \]

- Birth and death rates
  \[ T_+(n) = NF(n/N), \quad T_-(n) = n, \]

- Write steady-state solution in terms of probability current \( J(n) \):
  \[ J(n + 1) = J(n), \quad J(n) = T_-(n)P_S(n) - T_+(n - 1)P_S(n - 1). \]

- \( J(0) = 0 \implies J(n) = 0 \) for all \( n \geq 0 \) so steady-state solution is
  \[ P_S(n) = \frac{T_+(n - 1)}{T_-(n)}P_S(n - 1) = P_S(0) \prod_{m=1}^{n} \frac{T_+(m - 1)}{T_-(m)} \]

  with \( P_S(0) = 1 - \sum_{m \geq 1} P_S(m) \).
For large $N$, steady-state solution of master eqn is $(x = n/N)$

$$P_S(x) = \mathcal{A} \exp \left( N \int_{x}^{x'} \ln \frac{\Omega_{+}(x')}{\Omega_{-}(x')} dx' \right)$$

whereas soln of FP eqn is

$$P_S(x) = \mathcal{A}' \exp \left( 2N \int_{x}^{x'} \frac{\Omega_{+}(x') - \Omega_{-}(x')}{\Omega_{+}(x') + \Omega_{-}(x')} dx' \right)$$
**Noise-induced switching**

- Bistable deterministic network with stable fixed points $x = x_{\pm}$ and saddle $x = x_0$.

- Noise induces switching between basins of attraction for finite $N$ – exponentially small escape rates $r_{\pm} \sim e^{-N\tau_{\pm}}$.

- Rare transitions allow network to approach steady state PDF in limit $t \to \infty$
WKB APPROXIMATION

- Set $x = n/N$. Place an absorbing BC at saddle $u_0$. Eigenvalue expansion of PDF:

$$ P(x, t) = c_0 P_0(x) e^{-\lambda_0 t} + c_1 P_1(x) e^{-\lambda_1 t} + \cdots. $$

- For large $N$, $\lambda_0 \sim e^{-NE_0}$ with $E_0 = \mathcal{O}(1)$. Can identify $\lambda_0$ with MFPT to escape basin of attraction of metastable state $u_-$, say.

- Can approximate $P_0(x)$ by a solution to the stationary master equation with a reflecting BC at $u_0$—quasistationary solution.

- Take $P_0(x)$ to have the WKB form

$$ P_0(x) \sim K(x) e^{-NW(x)}, \quad K(S) = 1, W(S) = 0, $$

- Asymptotic expression for $\lambda_0$ (large-deviation theory)

$$ \log \lambda_0 \sim N[W(x_0) - W(x_-)] $$
HAMILTON–JACOBI EQUATION FOR $W$

- Expansion in powers of $N^{-1}$ yields Hamilton-Jacobi eqn for $W$:

$$H(x, p) = \sum_{r=\pm 1} \Omega_r(x) [e^{rp} - 1] = 0, \quad p = \frac{\partial W}{\partial x}$$

with $\Omega_+(x) = F(wx)$, $\Omega_-(x) = x$

- Classical mechanical interpretation: $H$ determines the motion of a “particle” with position $x$ and conjugate momentum $p$

$$\dot{x} = \frac{\partial H}{\partial p} = -xe^{-p} + F(wx)e^p$$

$$\dot{p} = -\frac{\partial H}{\partial x} = [e^{-p} - 1] + \omega F'(wx) [e^p - 1]$$

$t$ is a parameterization of paths rather than time.

- $W(x)$ with $W(S) = 0$ determined by action along zero energy trajectories of Hamiltonian system: most probable fluctuational path from $S$ to $x$ (in the large–$N$ limit)
HAMilton–JAcoBi eQUATION FOR $W$

Since HJ equation is a quadratic in $e^p$, there are two classes of zero–energy solution

$$p = 0, \quad p = p_*(x) \equiv \ln \frac{\Omega_-(x)}{\Omega_+(x)}.$$
**Matched Asymptotics**

- Can calculate pre factor using matched asymptotics - need to match absorbing BC at $u_0$

- Along an activation trajectory ($p = p_*(x)$)

  $$P_0(x) = \frac{A}{\sqrt{xF(wx)}} e^{-NW(x)}, \quad W(x) = \int^x \ln \left[ \frac{y}{F(\omega y)} \right] dy$$

- Along a relaxation trajectory ($p = 0$)

  $$P_0(x) = \frac{B}{F(wx) - x}$$

- Find that exit time from metastable state around $x_-$ is

  $$\lambda_0 = \frac{2\pi}{\sqrt{-1 + \omega F'(wx_0)}} \frac{1}{\sqrt{-1 + \omega F'(wx_-)}} \sqrt{\frac{x_0}{x}} e^{N[W(x_0) - W(x_-)]}$$
Part III. Path-integral formulation of neural master equation
Consider the representation of the joint probability density for the fields $\Phi_i = \{\Phi_i(s), 0 \leq s \leq t\}$, with $\Phi_i = Nu_i$ and $u_i$ satisfying the deterministic rate equation.

Rewrite as an infinite product of Dirac delta functions:

$$P[\Phi] = \mathcal{N} \prod_{s \leq t} \prod_i \delta \left( \partial_t \Phi_i + \alpha \Phi_i - NF \left( \sum_j W_{ij} \Phi_j / N \right) \right),$$

Introduce the Fourier representation of the Dirac delta function:

$$P[\Phi] = \int \prod_i D\tilde{\Phi}_i e^{-S[\Phi, \tilde{\Phi}]} , \quad D\tilde{\Phi}_i \sim \mathcal{N} \prod_{s \leq t} d\tilde{\Phi}_i(s)$$

where each $\tilde{\Phi}_i(s)$ is integrated along the imaginary axis.
Path Integral Formulation II

- Deterministic action is

\[ S[\Phi, \tilde{\Phi}] = \int dt \sum_i \tilde{\Phi}_i(t) \left[ \partial_t \Phi_i + \alpha \Phi_i - NF \left( \sum_j W_{ij}\Phi_j/N \right) \right] \]

- Path integral representation persists when fluctuations are taken into account, with modified action

\[ S[\Phi, \tilde{\Phi}] = \int dt \sum_i \tilde{\Phi}_i \left[ \partial_t \Phi_i + \alpha \Phi_i - NF \left( \sum_j W_{ij}\Psi_j/N \right) \right], \]

where \( \Psi_j = \tilde{\Phi}_j\Phi_j + \Phi_j \).
**Moment equations**

- Given $P[\Phi]$, we can calculate mean–fields according to

$$\langle\langle \Phi_k(t) \rangle\rangle = \int \prod_i D\Phi_i \Phi_k(t) P[\Phi] = \int \prod_i D\Phi_i \int \prod_i D\tilde{\Phi}_i \Phi_k(t)e^{-S[\Phi,\tilde{\Phi}]}$$

- Similarly, two–point correlations are given by

$$\langle\langle \Phi_k(t_1)\Phi_l(t_2) \rangle\rangle = \int \prod_i D\Phi_i \int \prod_i D\tilde{\Phi}_i \Phi_k(t_1)\Phi_l(t_2)e^{-S[\Phi,\tilde{\Phi}]}$$

- In terms of the physical activity variables $m_i(t)$,

$$\langle m_k(t) \rangle \equiv \sum_n n_k P(n, t) = \langle\langle \Phi_k(t) \rangle\rangle,$$

$$\langle m_k(t)m_l(t) \rangle - \langle m_k(t) \rangle \langle m_l(t) \rangle$$

$$= \langle\langle \Phi_k(t)\Phi_l(t) \rangle\rangle - \langle\langle \Phi_k(t) \rangle\rangle \langle\langle \Phi_l(t) \rangle\rangle + \langle\langle \Phi_k(t) \rangle\rangle \delta_{k,l}.$$
**Large Deviations I**

- Perform rescaling $\Phi_i \rightarrow \phi_i = \Phi_i/N$ so that we have a path-integral of the form

$$ P \sim \int \prod_i D\phi_i \int \prod_i D\tilde{\phi}_i e^{-NS[\phi, \tilde{\phi}]} $$

- Rescaled action is

$$ S[\phi, \tilde{\phi}] = \int dt \left[ \sum_i \tilde{\phi}_i \partial_t \phi_i + \mathcal{H}(\phi, \tilde{\phi}) \right] $$

and

$$ \mathcal{H}(\phi, \tilde{\phi}) = \sum_i \tilde{\phi}_i \left[ \alpha \phi_i - F \left( \sum_j W_{ij} \psi_j \right) \right] $$
In the limit $N \to \infty$, the path integral is dominated by the "classical" solutions $u(t), \tilde{u}(t)$:

$$\frac{\delta S[\phi, \tilde{\phi}]}{\delta \phi_i(t)} \bigg|_{\tilde{\phi}=\tilde{u}, \phi=u} = 0, \quad \frac{\delta S[\phi, \tilde{\phi}]}{\delta \tilde{\phi}_i(t)} \bigg|_{\tilde{\phi}=\tilde{u}, \phi=u} = 0.$$

These equations reduce to

$$\frac{\partial u_i}{\partial t} = -\frac{\partial \mathcal{H}(u, \tilde{u})}{\partial \tilde{u}_i}, \quad \frac{\partial \tilde{u}_i}{\partial t} = \frac{\partial \mathcal{H}(u, \tilde{u})}{\partial u_i}.$$

Hamiltonian dynamical system in which $u_i$ is a "coordinate" variable, $\tilde{u}_i$ is its "conjugate momentum"

Equivalent to WKB Hamiltonian under a canonical transformation
Part IV. Beyond the neural master equation
LIMITATIONS OF THE NEURAL MASTER EQUATION

- Transition rates are not uniquely determined
- What is $\tau_a$?
- Mean-field dynamics given by an activity-based rate equation. What about a current or voltage-based equation? Unlike number of active neurons, current is not a discrete variable.
- Neglects synaptic dynamics ie assumes that time-scale $\tau$ of synaptic dynamics smaller than $\Delta t$. But $\Delta t \to 0$ in Poisson regime and could be small in Gaussian regime.
**Velocity-Jump Markov Model I (PCB/Newby)**

- $U_k(t)$ is a population averaged synaptic current evolving as
  
  $$\tau dU_k(t) = \left[ -U_k(t) + \sum_{k=1}^{M} w_{kl} A_l(t) \right] dt.$$  

- The stochastic population firing rate is given by
  
  $$A_k(t) = \frac{N_k(t)}{N \Delta t}$$

- $N_k(t)$ evolves according to a one-step jump Markov process
  
  $$N_k(t) \rightarrow N_k(t) \pm 1$$

  with transition rates
  
  $$\Omega_+ = \frac{N \Delta t F(U_k)}{\tau_a}, \quad \Omega_- = \frac{n_k}{\tau_a}.$$  

- Take limit $N \rightarrow \infty$, $\Delta t \rightarrow 0$ with $N \Delta t = 1$.  

Example of a stochastic hybrid system

$U_k(t)$ is a piecewise deterministic variable coupled to the discrete jump Markov processes $N_i(t)$

The Markov processes are coupled to $U_k(t)$ via the transition rates

Introduce the probability density

$$
Pr\{U_k(t) \in (u_k, u_k + du), N_k(t) = n_k; k = 1, \ldots, M\} = p(u, n, t | u_0, n_0, 0)du
$$

with $u = (u_1, \ldots, u_M), n = (n_1, \ldots, n_M)$. 
Neural Chapman-Kolmogorov Equation

- $p$ evolves according to the Chapman-Kolmogorov (CK) equation

$$
\frac{\partial p}{\partial t} + \frac{1}{\tau} \sum_{k} \frac{\partial[v_k(x)p(x, t)]}{\partial u_k}
= \frac{1}{\tau_a} \sum_{k} \left[ (T_k - 1) (\omega_-(n_k)p(x, t)) + (T_k^{-1} - 1) (\omega_+(u_k)p(x, t)) \right],
$$

with $x = (u, n)$, and

$$
\omega_+(u_k) = F(u_k), \quad \omega_-(n_k) = n_k, \quad v_k(x) = -u_k + \sum_{k} w_{kl}n_l.
$$

- In the limit $\tau \to 0$ for $\tau_a > 0$ fixed, we recover the BCC neural master equation with $u = u(n)$ such that

$$
v_k(u(n), n) = 0.
$$
LIMITING CASE $\tau_a \ll \tau$

- In the limit $\tau_a \to 0$ for $\tau > 0$ fixed we obtain the deterministic voltage or current-based equation

$$\tau \frac{du_k(t)}{dt} = \left[ -u_k(t) + \sum_{k=1}^{M} w_{kl} \langle N_k(t) \rangle \right] dt.$$

where

$$\langle N_k(t) \rangle = \sum_n n_k \rho(n, u(t))$$

- $\rho(n, u(t))$ is the steady-state density of the birth-death process and is given by $M$ independent Poisson processes with rates $F(u_k)$:

$$\langle N_k(t) \rangle = F(u_k(t)).$$

- What about the regime $0 < \epsilon \ll 1$ with $\epsilon = \tau_a / \tau$?
Path integral

- Path-integral representation of stochastic dynamics

\[ p(x, \tau | x_0, 0) = \int_{x(0)=x_0}^{x(\tau)=x} \mathcal{D}[p] \mathcal{D}[x] \exp \left( -\frac{1}{\epsilon} S[x, p] \right) \]

with action

\[ S[x, p] = \int_0^\tau \left[ \sum_{\alpha=1}^M p_\alpha \dot{x}_\alpha - \lambda_0(x, p) \right] dt. \]

- \( \lambda_0 \) is the Perron eigenvalue of the following linear operator equation

\[ \sum_m A(n, m; x) R^{(0)}(x, p, m) = [\lambda_0(x, p) - \sum_{\alpha=1}^M p_\alpha v_\alpha(x, n)] R^{(0)}(x, p, n), \]
PERRON EIGENVALUE I

- Use the ansatz

\[ R^{(0)}(x, p, n) = \prod_{\alpha=1}^{M} \frac{\Lambda_{\alpha}(x, p)^{n_{\alpha}}}{n_{\alpha}!}. \]

- Using the explicit expressions for \( A \) and \( v_{\alpha} \), we find that

\[
\begin{align*}
\sum_{\alpha=1}^{M} \left( \left[ \frac{F(x_{\alpha})}{\Lambda_{\alpha}} - 1 \right] n_{\alpha} + \Lambda_{\alpha} - F(x_{\alpha}) \right) - \lambda_{0} &= - \sum_{\alpha=1}^{M} p_{\alpha} \left[ -x_{\alpha} + \sum_{\beta} w_{\alpha \beta} n_{\beta} \right].
\end{align*}
\]

- Collecting terms in \( n_{\alpha} \) for each \( \alpha \) yields

\[
\frac{F(x_{\alpha})}{\Lambda_{\alpha}} - 1 = - \sum_{\beta=1}^{M} p_{\beta} w_{\beta \alpha},
\]
Collecting terms independent of all $n_{\alpha}$ gives

$$\lambda_0 = \sum_{\alpha=1}^{M} [\Lambda_{\alpha} - F(x_{\alpha}) - x_{\alpha}p_{\alpha}].$$

Solving for each $\Lambda_{\alpha}$ in terms of $p$, we have

$$\lambda_0(x, p) \equiv \sum_{\alpha=1}^{M} \left[ \frac{F(x_{\alpha})}{1 - \sum_{\beta=1}^{M} p_{\beta}w_{\beta\alpha}} - x_{\alpha}p_{\alpha} - F(x_{\alpha}) \right]$$
GAUSSIAN APPROXIMATION I

- Performing the rescaling $p \rightarrow iP/\epsilon$ in path-integral

$$P(x, t) = \int_{x(0)=x_0}^{} D[x]D[p] \times \exp \left( - \int_0^\tau i \sum_\alpha p_\alpha \left[ \dot{x}_\alpha + x_\alpha - \sum_\beta \frac{w_{\alpha\beta}F(x_\beta)}{1 - i\epsilon \sum_\gamma w_{\gamma\beta}p_\gamma} \right] dt \right)$$

- The Gaussian approximation involves Taylor expanding the Lagrangian to first order in $\epsilon$, which yields a quadratic in $p$:

$$P(x, t) = \int_{x(0)=x_0}^{} D[x]D[p] \exp \left( \int_0^\tau \left[ i \sum_\alpha p_\alpha \left( \dot{x}_\alpha + x_\alpha - \sum_\beta w_{\alpha\beta}F(x_\beta) \right) \right] dt \right)$$

where $Q_{\alpha\gamma}(x) = \sum_\beta w_{\alpha\beta}F(x_\beta)w_{\gamma\beta}$
Gaussian Approximation II

- Performing the Gaussian integration along similar lines to the one-population model yields the multi-variate Onsager-Machlup path-integral

\[ P(x, t) = \int D[x] e^{-A[x]/\epsilon}, \]

with action functional

\[ A[x] = \frac{1}{4} \int_0^\tau \sum_{\alpha, \beta} (\dot{x}_\alpha(t) - V_\alpha(x(t)))Q_{\alpha\beta}^{-1}(x)(\dot{x}_\beta(t) - V_\beta(x(t)))dt, \]

where \( V_\alpha(x) = -x_\alpha + \sum_\beta w_{\alpha\beta}F(x_\beta). \)

- The corresponding Ito Langevin equation is

\[ dX_\alpha(t) = V_\alpha(X)dt + \sqrt{2\epsilon} \sum_\beta w_{\alpha\beta} \sqrt{F(x_\beta)}dW_\beta(t), \]

where \( W_\alpha(t) \) are independent Wiener processes.
Part V. Metastability in a two population model
Consider an E-I network with mean-field equations

\[
\begin{align*}
\frac{dx}{dt} &= -x + w_{EE}F(x) - w_{EI}F(y) \\
\frac{dy}{dt} &= -y + w_{IE}F(x) - w_{II}F(y),
\end{align*}
\]

E-I network (PCB/Newby)
**Phase-plane trajectories**

- Red curves show the $x$-nullclines, and blue curve show the $y$-nullcline.

- The red nullcline through the saddle is its stable manifold and acts as the separatrix $\Sigma$ between the two stable fixed points.

- Two deterministic trajectories are shown (black curves), starting from either side of the unstable saddle and ending at a stable fixed point.
**Quasi-Potential**

- The quasi-potential can be obtained by finding zero energy solutions of Hamilton’s equations

\[
\dot{x} = \nabla_p \mathcal{H}(x, p), \quad \dot{p} = -\nabla_x \mathcal{H}(x, p),
\]

with \( x = (x, y), p = (p_x, p_y) \) and \( \mathcal{H} = \lambda_0 \).

- Substituting for \( \mathcal{H} \), Hamilton’s equations have the explicit form

\[
\begin{align*}
\frac{dx_\alpha}{dt} &= -x_\alpha + \sum_\beta \frac{w_{\alpha\beta}F(x_\alpha)}{1 - \sum_{\gamma=1}^M p_\gamma w_{\gamma\alpha}} \\
\frac{dp_\alpha}{dt} &= p_\alpha - \frac{F'(x_\alpha)}{1 - \sum_{\gamma=1}^M p_\gamma w_{\gamma\alpha}} + F'(x_\alpha)
\end{align*}
\]

- The quasi-potential \( \Phi \) is the action along a zero energy solution curve \( x(t) \):

\[
\frac{d\Phi}{dt} \equiv \sum_{\alpha=1}^M \frac{\partial \Phi}{\partial x_\alpha} \frac{dx_\alpha}{dt} = \sum_{\alpha=1}^M p_\alpha \frac{dx_\alpha}{dt},
\]
CHARACTERISTIC PATHS OF MAXIMUM LIKELIHOOD

- Rays originating from the left (right) stable fixed point are shown in orange (cyan).
- The ray connecting to the saddle shown in red (blue).
- The grey curve is the separatrix $\Gamma$. 
STOCHASTIC TRAJECTORIES

- Sample trajectories of the two-population model
CROSSING THE SEPARATRIX
Mean FPT

- Good agreement between analytical results and Monte Carlo simulations
FUTURE DIRECTIONS

- Apply stochastic phase reduction method to noise driven synchrony of coupled E-I networks
- Extend master equation framework to continuum neural fields (see eg. path integral methods of Buice and Cowan)
- Derivation of a master equation from first principles using multi-scale analysis
- Incorporation of other biophysical processes such as synaptic depression, channel noise etc.
REFERENCES


