Describing Spike-Trains

Maneesh Sahani

Gatsby Computational Neuroscience Unit
University College London

Term 2, Winter 2017

Neural Coding

- The brain manipulates information by combining and generating action potentials (or spikes).

- It seems natural to ask how information (about sensory variables; inferences about the world; action plans; cognitive states . . .) is represented in spike trains.

- Experimental evidence comes largely from sensory settings
  - ability to repeat the same stimulus (although this does not actually guarantee that all information represented is identical, but some is likely to be shared across trials).
  - Computational methods are needed to characterise and quantify these results.

- Theory can tell us what representations *should* look like.

- Theory also suggests what internal variables might need to be represented:
  - categorical variables
  - uncertainty
  - reward predictions and errors

Spikes

- The timecourse of every action potential (AP) in a cell measured at the soma may vary slightly, due to differences in the open channel configuration.

- However, axons tend to contain only the Na^+ and K^+ channels needed for AP propagation, and therefore exhibit little or no AP shape variation.

- No experimental evidence (as far as I know) that AP shape affects vesicle release.

- Thus, from the point of view of inter-neuron communication, it seems that the only thing that matters about an AP or spike is its time of occurance.

Notation for spike trains

A **spike train** is the sequence of times at which a cell spikes:

$$\mathcal{S} = \{t_1, t_2, \ldots t_N\}.$$

It is often useful to write this as a function in time using the Dirac-delta form,

$$s(t) = \sum_{i=1}^{N} \delta(t - t_i) \qquad \text{(D\&A call this } \rho(t))$$

or using a (cumulative) counting function for the number of spikes to time t,

$$N(t) = \int_0^{\to t} d\xi \; s(\xi)$$
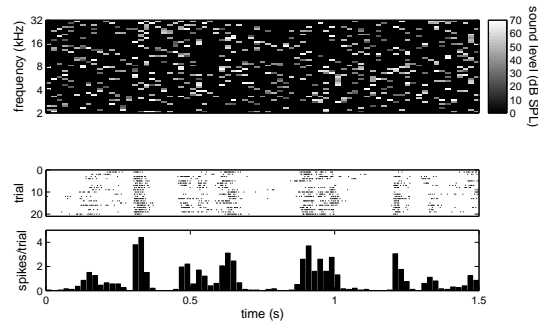
$(\to t$ means that t is not included in the integral)

or as a vector by discretizing with time step Δt

$$\mathbf{s} = (s_1 \ldots s_{T/\Delta t}); \qquad s_t = \int_{t-\Delta t}^{\to t} d\xi \; s(\xi)$$

Note that the neural refractory period means that for $\Delta t \approx 1\mathrm{ms}$, s_t is binary.

Variability

Empirically, spike train responses to a repeated stimulus are (very) variable. This is particularly true in the cortex, but might be less so at earlier stages.

This variability probably arises in more than one way.

- ▶ **Noise**. Perhaps due to vesicle release; or thermal noise in conductances.
- ▶ **Ongoing processes**. The brain doesn't just react to sensory input. Ongoing processing is likely to affect firing, particularly in cortex; and there is experimental evidence for this. This might lead to variability on a slower time-scale than noise.

We do not know the relative sizes of these two contributions.

Point Processes

A probabilistic process that produces events of the type

$$\mathcal{S} = \{t_1, \ldots, t_N\} \subset \mathcal{T}$$

is called a **point process**.

This is the statistical object best suited for the description of spike trains. We take $\mathcal{T} = [0, T]$ to be an interval of time.

Every point process (on an ordered set) is associated with a dual **counting process** which produces events of the type
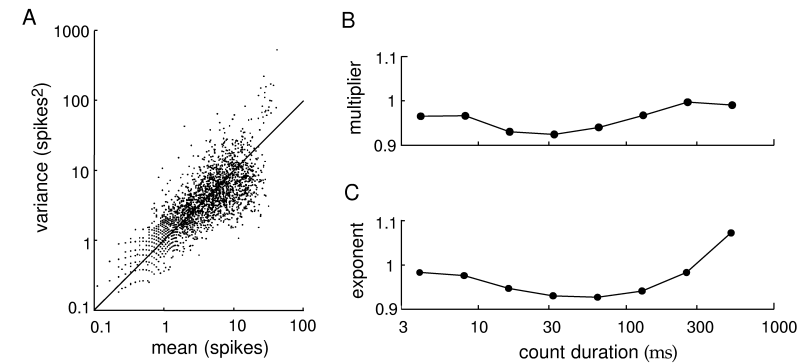
$$N(t) \text{ such that } N(t) \geq 0$$
$$N(t') \geq N(t) \text{ if } t' > t$$
$$N(t) - N(s) = N[s, t) \in \mathbb{Z}$$

$N(t)$ gives the number of events with $t_i < t$.

Count variability

Everything about the spike train can be variable, even the spike count on the ith repetition (or "trial") $N_i = \int_0^T d\xi \, s_i(\xi)$

Variability in N_i is on order of the mean.

Fits of the form $\mathrm{Var}[N_i] = A \cdot \mathrm{E}[N_i]^B$ yield values of A and B between about 1 and 1.5.

Homogeneous Poisson Process: $N_\lambda(t)$

In the simplest point process, events are all **independent** and occur at a fixed **rate** λ.

Independence is defined formally:

1. **Independence.** For all disjoint intervals $[s, t)$ and $[s', t')$, $N_\lambda[s, t) \perp N_\lambda[s', t')$.

Knowing the number (or times) of one or more events tells us nothing about other possible events.

Homogeneous Poisson Process: $N_\lambda(t)$

The rate condition can be defined in two ways.

If we assume that $\lim_{ds \to 0} N_\lambda[s, s+ds] \in \{0, 1\}$ (technically **conditional orderliness** – at most one event occurs at one time) then it is sufficient to assume that

   2. **Mean event rate.** $\mathbb{E}[N_\lambda[s, t]] = (t-s)\lambda$.

Without assuming conditional orderliness, we could instead define the process by giving the whole distribution of $N_\lambda[s, t]$. Instead, we will use the more restrictive defining assumption to derive the distribution.

Homogeneous Poisson Process: $N_\lambda(t)$

So a Poisson **process** produces event counts which follow the Poisson **distribution**. As we mentioned above, we could instead have dispensed with the conditional orderliness assumption and instead made this a defining property of the process:

   $2'$. **Count distribution.** $N_\lambda[s, t] \sim \text{Poiss}[(t-s)\lambda]$.

We will now derive a number of properties of the homogeneous Poisson process. These are good to know. We will also employ some tricks in the derivations that can be applied more generally.

Homogeneous Poisson Process: $N_\lambda(t)$

Divide $[s, t]$ into M bins of length Δ (i.e. $M = (t-s)/\Delta$). If $\Delta \ll 1/\lambda$ conditional orderliness implies that spike count per bin is binary.

For a binary random variable, the expectation is the same as the probability of event, so $\lambda\Delta \approx P(N[t, t+\Delta] = 1)$.

Thus, distribution of $N[s, t]$ binomial:

$$P[N_\lambda[s, t] = n] = \binom{M}{n}(\lambda\Delta)^n(1 - \lambda\Delta)^{M-n}$$

$$= \frac{M!}{n!(M-n)!}\left(\frac{\lambda(t-s)}{M}\right)^n\left(1 - \frac{\lambda(t-s)}{M}\right)^{M-n}$$

write $\mu = \lambda(t-s)$

$$= \frac{\mu^n}{n!}\frac{M(M-1)\cdots(M-n+1)}{M^n}\left(1 - \frac{\mu}{M}\right)^{-n}\left(1 - \frac{\mu}{M}\right)^M$$

now take the limit $\Delta \to 0$ or, equivalently, $M \to \infty$

$$= \frac{\mu^n}{n!}1^n 1^n e^{-\mu} = \frac{\mu^n e^{-\mu}}{n!}$$

So the spike count in any interval is Poisson distributed.

Count Variance

$$
\begin{aligned}
\mathbb{V}[N_\lambda[s, t]] &= \langle(n - \mu)^2\rangle \\
&= \langle n^2\rangle - \mu^2 \\
&= \langle n(n-1) + n\rangle - \mu^2 \\
&= \sum_{n=0}^{\infty} n(n-1)\frac{e^{-\mu}\mu^n}{n!} + \mu - \mu^2 \\
&= \sum_{n=0}^{\infty} \frac{e^{-\mu}\mu^{n-2}}{(n-2)!}\mu^2 + \mu - \mu^2 \\
&= 0 + 0 + \sum_{(n-2)=0}^{\infty} \frac{e^{-\mu}\mu^{n-2}}{(n-2)!}\mu^2 + \mu - \mu^2 \\
&= \mu^2 + \mu - \mu^2 = \mu
\end{aligned}
$$

Thus:

   3. **Fano factor**[1]. $\dfrac{\mathbb{V}[N_\lambda[s, t]]}{\mathbb{E}[N_\lambda[s, t]]} = 1$.

---

[1] Note that this ratio (unlike the CV that we will encounter later) is only dimensionless for counting processes, or other dimensionless random variables.

ISI distribution

The next few properties relate to the inter-spike interval (ISI) statistics.

First, it is fairly straightforward that, since the counting processes before and after event t_i are independent, the times to the previous and following spikes are independent.

4. **ISI independence.** $\forall i > 1, \quad t_i - t_{i-1} \perp t_{i+1} - t_i.$

The full distribution of ISIs can be found from the count distribution:

$$P\left[t_{i+1} - t_i \in [\tau, \tau + d\tau]\right] = P\left[N_\lambda[t_i, t_i + \tau) = 0\right] \times P\left[N_\lambda[t_i + \tau, t_i + \tau + d\tau) = 1\right]$$
$$= e^{-\lambda\tau}\lambda d\tau e^{-\lambda d\tau}$$

taking $d\tau \to 0$

$$= \lambda e^{-\lambda\tau} d\tau$$

5. **ISI distribution.** $\forall i \geq 1, \quad t_{i+1} - t_i \sim \text{ iid Exponential}[\lambda^{-1}].$

ISI distribution

5. **ISI distribution.** $\forall i \geq 1, \quad t_{i+1} - t_i \sim \text{ iid Exponential}[\lambda^{-1}].$

From this it follows that

6. **Mean ISI.** $\mathbb{E}\left[t_{i+1} - t_i\right] = \lambda^{-1}$
7. **Variance ISI.** $\mathbb{V}\left[t_{i+1} - t_i\right] = \lambda^{-2}$

These two properties imply that the **coefficient of variation** (CV), defined as the ratio of the standard deviation to the mean, of the ISIs generated by an homogeneous Poisson process is 1.

Joint density

Finally, consider the probability density of observing spike train $\{t_1 \ldots t_N\}$ in interval $\mathcal{T}$.

Spike times are independent of one another and arrive at a uniform rate. So:

$$p(t_1 \ldots t_N)dt_1 \ldots dt_N = P\left[N \text{ spikes in } \mathcal{T}\right] \times P\left[i\text{th spike} \in [t_i, t_i + dt_i)\right] \times$$
$$[\# \text{ of equivalent spike orderings}]$$

The first term is given by the Poisson distribution, the second by the uniform distribution of spike times conditioned on N, and the third is $N!$, giving us

$$p(t_1 \ldots t_N)dt_1 \ldots dt_N = \frac{(\lambda T)^N e^{-\lambda T}}{N!} \cdot \frac{dt_1}{T} \ldots \frac{dt_N}{T} \cdot N!$$
$$= \lambda^N e^{-\lambda T} dt_1 \ldots dt_N$$

We will see another way to write down this same expression while considering the inhomogeneous Poisson process below.

Inhomogeneous Poisson Process: $N_{\lambda(t)}(t)$

The inhomogeneous Poisson process generalizes the constant event-arrival rate λ to a time-dependent one, $\lambda(t)$, while preserving the assumption of independent spike arrival times.

We can quickly summarize the properties of the inhomogeneous process by reference to the homogeneous one.

To begin, the two defining properties (this time we just state the Poisson distribution property directly.)

1. **Independence.** For all disjoint intervals $[s, t)$ and $[s', t')$, $N_{\lambda(t)}[s, t) \perp N_{\lambda(t)}[s', t')$.
2. **Count distribution.** $N_{\lambda(t)}[s, t) \sim \text{Poiss}[\int_s^t d\xi \, \lambda(\xi)]$.

The variance in the counts is simply a consequence of the Poisson counting distribution, and so the next property follows directly.

3. **Fano factor.** $\dfrac{\mathbb{V}\left[N_{\lambda(t)}[s, t)\right]}{\mathbb{E}\left[N_{\lambda(t)}[s, t)\right]} = 1.$

ISI distribution

Independence of counting in disjoint intervals means that ISIs remain independent.

4. **ISI independence.** $\forall i > 1, \quad t_i - t_{i-1} \perp t_{i+1} - t_i$.

The full distribution of ISIs is found in a similar manner to that of the homogeneous process distribution:

$$P\left[t_{i+1} - t_i \in [\tau, \tau + d\tau]\right] = P\left[N_{\lambda(t)}(t_i, t_i + \tau) = 0\right] P\left[N_{\lambda(t)}[t_i + \tau, t_i + \tau + d\tau) = 1\right]$$

$$= e^{-\int_{t_i}^{t_i+\tau} \lambda(\xi)d\xi} e^{-\int_{t_i+\tau}^{t_i+\tau+d\tau} \lambda(\xi)d\xi} \int_{t_i+\tau}^{t_i+\tau+d\tau} \lambda(\xi)d\xi$$

taking $d\tau \to 0$

$$= e^{-\int_{t_i}^{t_i+\tau} \lambda(\xi)d\xi} e^{-\lambda(t_i+\tau)d\tau} \lambda(t_i + \tau)d\tau$$

$$= e^{-\int_{t_i}^{t_i+\tau} \lambda(\xi)d\xi} \lambda(t_i + \tau)d\tau$$

5. **ISI distribution.** $\forall i \geq 1, \quad p(t_{i+1} - t_i) = e^{-\int_{t_i}^{t_{i+1}} \lambda(\xi)d\xi} \lambda(t_{i+1})$.

As the ISI distribution is not *iid* it is not very useful to consider its mean or variance.

Time rescaling

Finally, we derive an additional important property of the inhomogeneous process. Let us rewrite the density above, by changing variables from t to u according to

$$u(t) = \int_0^t \lambda(\xi)d\xi \qquad \text{i.e.} \quad u_i = \int_0^{t_i} \lambda(\xi)d\xi$$

Then

$$p(u_1 \ldots u_n) = p(t_1 \ldots t_n) / \prod_i \frac{du_i}{dt_i}$$

$$= e^{-u(T)} \prod_{i=1}^N \lambda(t_i) / \prod_{i=1}^N \lambda(t_i)$$

$$= e^{-u(T)}$$

Comparing this to the density for a homogeneous Poisson process shows that the variables u_i are distributed according to a homogeneous Poisson process with mean rate $\lambda = 1$.

This result is called **time rescaling**, and is central to the study of point processes in time.

Joint density

The probability density of the event $\{t_1 \ldots t_N\}$ can be derived by setting the count in intervals **between** spikes to 0, and the count in an interval **around** t_i to 1. This gives

$$p(t_1 \ldots t_N)dt_1 \ldots dt_N = P\left[N[0, t_1) = 0\right] \times P\left[N[t_1, t_1 + dt_1) = 1\right] \times \cdots \times P\left[N(t_N, T) = 0\right]$$

$$= e^{\int_0^{t_1} \lambda(\xi)d\xi} \times \lambda(t_1)dt_1 \times \cdots \times e^{\int_{t_N}^T \lambda(\xi)d\xi}$$

$$= e^{-\int_0^T \lambda(\xi)d\xi} \prod_{i=1}^N \lambda(t_i)dt_1 \ldots dt_N$$

Setting $\lambda(t) = \lambda$ gives us the result for the homogeneous process.

Self-exciting point processes

A self-exciting process has an intensity function which is conditioned on past events

$$\lambda(t) \to \lambda(t | N(t), t_1 \ldots t_{N(t)})$$

It will be helpful to define the notation $H(t)$ to represent the event *history* at time t—representing both $N(t)$ and the times of the corresponding events. Then the self-exciting intensity function can be written $\lambda(t | H(t))$.

This is actually the most general form of a point process – we can re-express any (conditionally orderly) point process in this form. To see this, consider the point process to be the limit as $\Delta \to 0$ of a binary time series $\{b_1, b_2, \ldots b_{T/\Delta}\}$ and note that

$$P(b_1, b_2, \ldots b_{T/\Delta}) = \prod_i P(b_i | b_{i-1} \ldots b_1)$$

Renewal processes

If the intensity of a self-exciting process depends only the time since the last spike, i.e.

$$\lambda(t|H(t)) = \lambda(t - t_{N(t)})$$

then the process is called a **renewal** process. ISIs from a renewal process are iid and so we could equivalently have defined the process by its ISI density. This gives an (almost) easy way to write the probability of observing the event $\{t_1 \ldots t_N\}$ in T. Suppose, for simplicity, that there was an event at $t_0 = 0$. Then, if the ISI density is $p(\tau)$:

$$p(t_1 \ldots t_N) \, dt_1 \ldots dt_N = \prod_{i=1}^{N} p(t_i - t_{i-1})\Big(1 - \int_0^{T-t_N} d\tau \; p(\tau)\Big)$$

The last term gives the probability that no more spikes are observed after t_N. If had not assumed that there was a spike at 0 we would have needed a similar term at the front.

The conditional intensity (sometimes called **hazard function**) for the renewal process defined by its ISI density $p(\tau)$ is

$$\lambda(t|t_{N(t)}) \, dt = \frac{p(t - t_{N(t)})}{1 - \int_0^{t-t_{N(t)}} d\tau \; p(\tau)} \, dt$$

which is indeed a function only of $t - t_{N(t)}$.

Inhomogeneous renewal processes

In an inhomogeneous renewal processes, the rate depends both on time since the last spike and on the current time.

$$\lambda(t) \rightarrow \lambda(t, t - t_{N(t)})$$

Called "inhomogeneous Markov interval" processes by Kass and Ventura (2000).

Two popular ways to construct an inhomogeneous renewal process:

1. **Time-rescaling**. Given unit-mean ISI density $p(\tau)$, and time-varying intensity $\rho(t)$, define:

$$p(t_1 \ldots t_N) \, dt_1 \ldots dt_N = \prod_{i=1}^{N} p\Big(\int_{t_{i-1}}^{t_i} \rho(\xi)d\xi \Big)\Big(1 - \int_0^{\int_{t_N}^{T} d\xi \; \rho(\xi)} d\tau \; p(\tau)\Big)$$

2. **Spike-response**.

$$\lambda(t, t - t_{N(t)}) = f(\rho(t), h(t - t_{N(t)}))$$

for a simple f. Often, f just multiplies the two functions (or, equivalently, adds log intensities).

The term "spike-response" comes from Gerstner, who uses such spike-triggered currents to create a potentially more tractable approximation to an integrate-and-fire neuron.

These definitions differ in how ISI density depends on ρ. **Rescaling**: higher rates make time pass faster, so ISI interactions are rescaled. **Spike-response**: a refractory h may not suppress spikes as well at higher rates, but the duration of influence does not change.

Gamma-interval process

The specific choice of the **gamma-interval process** with

$$t_{i+1} - t_i \overset{\text{iid}}{\sim} \text{Gamma}[\alpha, \beta]$$

where

$$\tau \sim \text{Gamma}[\alpha, \beta] \Rightarrow p(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)}\tau^{\alpha-1}e^{-\beta\tau}$$

is an important renewal process in theoretical neuroscience, because the ISI distribution has a refractory-like component.

A homogeneous Poisson process is a gamma-interval process (and therefore a renewal process) with $\alpha = 1$. The parameter α is sometimes called the order or the shape-parameter of the gamma-interval process. Larger values of α shape the polynomial rising part of the Gamma density, thus implementing a relative refractory period. The long-time behaviour is dominated by the exponential decay with coefficient β.

You might wish to show that a gamma-interval process of integral order α can be constructed by selecting every αth event from a homogeneous Poisson process.

General Spike-Response processes

This category of processes has come to be used with increasing frequency recently, particularly in a generalised-linear form.

The product form of spike-response renewal process can be written:

$$\lambda(t, t - t_{N(t)}) = e^{\rho(t) + h(t - t_{N(t)})}$$

and then generalised to include influence from all (or > 1) past spikes:

$$\lambda(t|H(t)) = e^{\rho(t) + \sum_j h(t - t_{(N(t)-j)})}$$

Often, we wish to estimate the parameters of a point-process model from spike data. Assuming a generalised linear form makes this easier. Write history influence h in terms of a basis of fixed functions $h_i(\tau)$:

$$\lambda(t|H(t)) = e^{\rho(t) + \sum_{ij} \alpha_i h_i(t - t_{(N(t)-j)})}$$

If $\rho(t)$ is also written as linear function of external covariates, then the complete model can be fit by the standard methods used for generalised linear models (GLMs: note, a different use of this abbreviation to the commonly used models for fMRI data).

Doubly stochastic Poisson (or Cox) process

In the doubly stochastic process, or Cox process, $\lambda(t)$ is itself a random variable; or depends on another random process $x(t)$. An example is the randomly scaled IHPP:

$$\lambda(t) = s \cdot \rho(t) \qquad \text{with } \rho(t) \text{ fixed and } s \sim \text{Gamma}(\alpha, \beta)$$

These models have been the subject of some recent attention, as a way to model a stimulus-dependent response $\rho(t)$ which is modulated by cortical excitability. The counting process for such a DSPP has a Fano factor > 1. DSPP models also provide a useful way to introduce dependencies between two or more point processes, through correlations in the intensity functions.

Joint Models

▶ 2D point process is not correct model

▶ superimposed processes

▶ infinitely divisible Poisson process

▶ multivariate self-exciting process

▶ log-linear spike-response models with interaction terms

▶ doubly stochastic processes
  ▶ Common input to log-linear spike-response models
  ▶ Gaussian process factor analysis

Measuring point processes

Event data $\Rightarrow$ point process generative models. What about measurement?

Consider spike trains from repeated experiments under (as far as possible) constant experimental conditions.

$$s^{(k)}(t) = \sum_{i=1}^{N^{(k)}} \delta(t - t_i^{(k)}) \qquad \text{(trials) } k = 1 \ldots K$$

How to characterise $s^{(k)}(t)$, and relationship to stimulus (or task)?

▶ Parametric point-process model, possibly dependent on stimulus $a(t)$:

$$s^{(k)}(t) \sim \lambda \left(t, a[0, t), N^{(k)}(t), t_1^{(k)}, \ldots, t_{N^{(k)}(t)}^{(k)}, \theta \right).$$

  Encoding: stimulus-response function. Discussed later.

▶ Construct an algorithm to estimate $a(t)$ from $s^{(k)}(t)$:

$$\widehat{a}(t) = F[s^{(k)}[0, t)].$$

  Decoding: what does the neuron say about the world? (Not always causal). Also discussed later.

▶ Estimate nonparametric features (usually **moments**) of the distribution of $s^{(k)}(t)$.

Mean intensity and PSTH

Simplest non-parametric characterisation of spike process is with **mean intensity**:

$$\overline{\lambda}(t) = \langle s(t) \rangle = \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} s^{(k)}(t)$$

Not the intensity function for the point process (unless Poisson) – *marginalised* over history.

$$\overline{\lambda}(t, a(\cdot)) \equiv \int dN(t) \int dt_1 \ldots dt_{N(t)} \, p\left(t_1 \ldots t_{N(t)}\right) \lambda\left(t, a(\cdot), N(t), t_1, \ldots, t_{N(t)}\right)$$

For finite K, estimating $\overline{\lambda}$ by summing δ-functions yields spiky results. Instead, histogram:

$$\overline{N[t, \widehat{t + \Delta t}]} = \frac{1}{K} \sum_{k=1}^{K} N^{(k)}[t, t + \Delta t]$$

This is called the Post- (or Peri-) Stimulus-Time Histogram $\Rightarrow$ PSTH.

Smoothing the PSTH

If we expect $\overline{\lambda}(t)$ to be smooth, could use kernel $\phi(\tau)$:

$$\widehat{\overline{\lambda}(t)} = \frac{1}{K} \sum_{k=1}^{K} \int d\tau \phi(\tau) s^{(k)}(t - \tau)$$

Resembles kernel density estimation (without normalisation).

Width of ϕ could be chosen adaptively, depending on local density of spikes. Note: sampling a smoothed function makes *much* more sense than smoothing a binned histogram!

Alternatively, can impose a smooth prior (e.g. GP) on time-varying aspect of intensity: $\lambda(t)$ for inhomog Poisson, or (say) $\rho(t)$ for inhomog gamma-interval of order γ:

$$\boldsymbol{\rho} \sim \mathcal{N}(\mu\mathbf{1}, \mathsf{K}_\theta)$$

$$p(t_1 \ldots t_N | \boldsymbol{\rho}) = \prod_{i=1}^{N} \left[\frac{\gamma x_{t_i}}{\Gamma(\gamma)} \left(\gamma \sum_{j=t_{i-1}}^{t_i - 1} \rho_j \Delta \right)^{\gamma - 1} \exp\left\{ -\gamma \sum_{j=t_{i-1}}^{t_i - 1} \rho_j \Delta \right\} \right]$$

Poterior on $\rho(t)$ can be found by approximate inference (Laplace/EP).

Autocorrelation

The autocorrelation function for a process that generates spike trains $s(t)$ is:

$$R_{ss}(\tau) = \left\langle \frac{1}{T} \int dt\, s(t) s(t - \tau) \right\rangle$$

where $\langle \cdot \rangle$ is expectation wrt to random draws of $s(t)$ from the process.

Time-averaged local second moment of joint on $s(t)$; $\overline{\lambda}(t)$ was the (non-time-averaged) first moment.

Note that, since $s(t)$ is a sum δ functions, $R_{ss}(0) = \infty$ under this definition.

Alternatively, could define R_{ss} as time-averaged conditional first moment $\Rightarrow$ mean intensity at $t + \tau$, conditioned on event at t, averaged over t.

$$R_{ss}^{alt}(\tau) = \frac{1}{T} \int dt\, \langle \lambda(t + \tau | \exists i : t_i = t) \rangle,$$

where $\langle \cdot \rangle$ is expectation with respect to $N(T)$ and $t_{j \neq i}$. Now $R_{ss}^{alt}(0) = 0$.

We will stick to the first (i.e., second moment) definition.

Autocovariance

Using the identity $\langle x^2 \rangle = \langle (x - \mu)^2 \rangle + \mu^2$, we can decompose the autocorrelation function:

$$R_{ss}(\tau) = \overline{\Lambda}^2 + \frac{1}{T} \int dt\, (\overline{\lambda}(t) - \overline{\Lambda})(\overline{\lambda}(t - \tau) - \overline{\Lambda})$$

$$+ \underbrace{\left\langle \frac{1}{T} \int dt\, (s(t) - \overline{\lambda}(t))(s(t - \tau) - \overline{\lambda}(t - \tau)) \right\rangle}_{Q_{ss}(\tau)}$$

where $\overline{\Lambda}$ is the time-averaged mean rate. $Q_{ss}(\tau)$ is called the **autocovariance** function.

[D&A call Q_{ss} the autocorrelation function; in the experimental literature, estimates of Q_{ss} are usually called "shift-" or "shuffle-corrected autocorrelograms".]

- For an (inhomogeneous) Poisson process $Q_{ss}(\tau) = \delta(\tau)$, by independence.
- For a general self-exciting process, $Q_{ss}(\tau)$ gives (to second order) dependence on nearby spike times.
- Often used to look for oscillatory structure in spike trains (where spikes tend to repeat around fixed intervals, but at random phase wrt stimulus) or similar spike-timing relationships.
- But, as any point process is self-exciting, *any* non-Poisson process will have non-δ autocovariance, even if nearby spike-timing relationships are not the most natural (or causal) way to describe the generative process. (Think about effects of random (but slow) variations in a non-constant $\lambda(t)$, as in a DSPP).

Estimating correlation functions

- Correlation functions are typically estimated by constructing **correlograms**: histograms of time *differences* between (not necessarily adjacent) spikes.
- Covariance function is estimated by subtracting an estimate of the correlation of the mean intensity:

$$\frac{1}{T} \int dt\, \widehat{\overline{\lambda}}(t) \widehat{\overline{\lambda}}(t - \tau)$$

$$= \frac{1}{TK^2} \int dt\, \sum_k s^{(k)}(t) \sum_{k'} s^{(kp)}(t - \tau) = \frac{1}{TK^2} \sum_{kk'} \int dt\, s^{(k)}(t) s^{(kp)}(t - \tau)$$

'Shift' or 'shuffle' correction.

- May also be constructed in frequency domain: power spectrum, spectrogram, coherence (for multiple processes). Usually based on FT of binary-binned spike trains.

Multiple spike trains

Often, we may be interested in simultaneously modelling responses from many cells.

If no two processes can generate events at precisely the same time (a form of conditional orderliness), or if simultaneous spiking events are independent, then dependences between the processes generally by dependence on all previous events in all cells:

$$\lambda^{(c)}(t) \rightarrow \lambda^{(c)}\left(t \mid N^{(c)}(t), t_1^{(c)}, \ldots, t_{N^{(c)}(t)}^{(c)}, \{N^{(c')}(t), t_1^{(c')}, \ldots, t_{N^{(c')}(t)}^{(c')}\}\right)$$

This is analogous to the self-exciting point process intensity function.

Dependencies can also be expressed by other forms, for example by DSPPs with the latent random process shared (or correlated) between cells. Such representations may often be more natural or causally accurate.

Cross-correlations

Techniques for measuring relationships between cells, analogous to those for single processes—cross-correlogram estimates of the cross-correlation function:

$$R_{s^{(c)}s^{(c')}}(\tau) = \left\langle \frac{1}{T} \int dt\; s^{(c)}(t)s^{(c')}(t-\tau) \right\rangle;$$

shift- or shuffle-corrected correlogram estimates of the cross-covariance function:

$$Q_{s^{(c)}s^{(c')}}(\tau) = \left\langle \frac{1}{T} \int dt\; (s^{(c)}(t) - \overline{\lambda}^{(c)}(t))(s^{(c')}(t-\tau) - \overline{\lambda}^{(c')}(t-\tau)) \right\rangle;$$

or by cross-spectra or empirical coherences.

As for autocovariograms, structure in a cross-covariogram needn't imply that dependencies between individual spike times are the most natural way to think about the interaction between the processes – DSPPs with shared latents may also give significant cross-covariance structure.