Extraction of Synaptic Input Properties *in Vivo*

Paolo Puggioni
p.paolo321@gmail.com
Neuroinformatics Doctoral Training Centre and Institute for Adaptive
and Neural Computation, School of Informatics, University of Edinburgh,
Edinburgh EH8 9AB, U.K.

Marta Jelitai
jelitai@koki.hu
Ian Duguid
ian.duguid@ed.ac.uk
Centre for Integrative Physiology, University of Edinburgh,
Edinburgh EH8 9XD, U.K.

Mark C.W. van Rossum
mvanross@inf.ed.ac.uk
Institute for Adaptive and Neural Computation, School of Informatics,
University of Edinburgh, Edinburgh EH8 9AB, U.K.

Knowledge of synaptic input is crucial for understanding synaptic integration and ultimately neural function. However, *in vivo*, the rates at which synaptic inputs arrive are high, so that it is typically impossible to detect single events. We show here that it is nevertheless possible to extract the properties of the events and, in particular, to extract the event rate, the synaptic time constants, and the properties of the event size distribution from *in vivo* voltage-clamp recordings. Applied to cerebellar interneurons, our method reveals that the synaptic input rate increases from 600 Hz during rest to 1000 Hz during locomotion, while the amplitude and shape of the synaptic events are unaffected by this state change. This method thus complements existing methods to measure neural function *in vivo*.

1 Introduction

Neurons typically receive a barrage of thousands of excitatory and inhibitory events per second. As these inputs determine to a large extent the spiking activity of the neuron, it is important to know the properties

---

M. J. is now at the Institute of Experimental Medicine of the Hungarian Academy of Sciences, Budapest, Hungary.
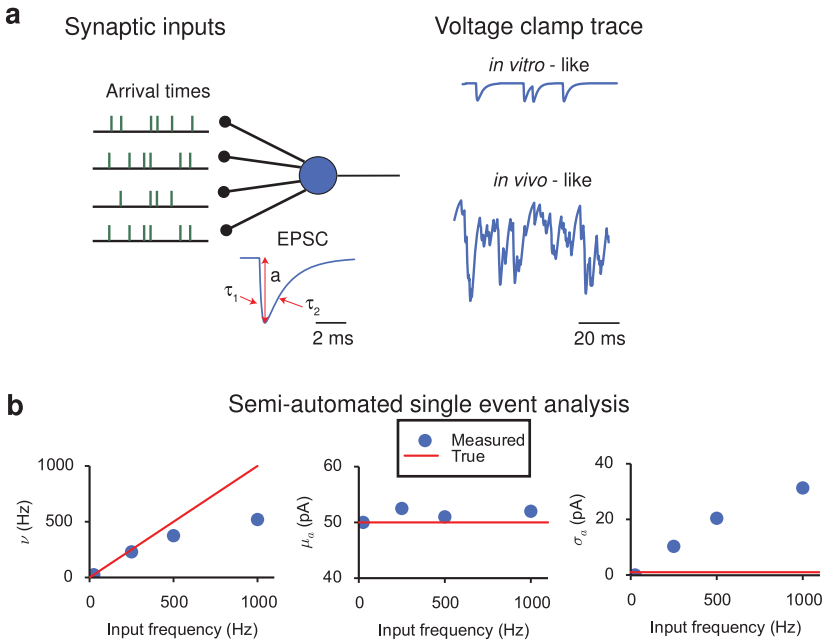
Figure 1: Inference of synaptic properties. (a) A neuron receives input from a number of synapses under a Poisson rate assumption. The events have identical shape, but the amplitude a can vary between events. Right: For *in vitro* experiments, synaptic events rates are typically low, and the individual events can be extracted and quantified. However, for *in vivo* experiments, rates are high, and individual events are not distinguishable. (b) Analysis based on semiautomated single-event extraction produces incorrect results when the total rate exceeds 500 Hz. From left to right: estimated event frequency, estimated mean event amplitude, estimated standard deviation of the event amplitude. Model parameters: $a_k = 50$ pA. rise time $\tau_1 = 0.3$ ms and decay time $\tau_2 = 2$ ms.

of synaptic input and how it changes, for example, with behavioral state (e.g., sleep, attention, locomotion), plasticity, or homoeostasis. Consider a neuron receiving synaptic input while the total current is being measured in voltage-clamp (see Figure 1a). While *in vitro*, or in cases where activity is artificially lowered, individual excitatory and inhibitory inputs can be resolved (see Figure 1a, top); *in vivo* the rates are typically so high that this is impossible. Instead, the total synaptic current trace is wildly fluctuating, and single-event extraction methods will fail.

Nevertheless, information can still be extracted from the statistical properties of the recorded *in vivo* currents. Although individual synaptic events might not be distinguishable in the observed current trace, the trace will still bear signatures of the underlying events. Intuitively, the mean current

should be proportional to the product of the synaptic event size and the total event frequency. But it is possible to extract other information as well. For instance, when the synaptic events have short time constants, the observed current trace will have more high-frequency content than when the synaptic time constants are slow. Similarly, when the input is composed of many small events, the variance of the current trace will be smaller than when it is composed of a few large events.

An early application of these ideas was used at the neuromuscular junction (Katz & Miledi, 1972) and in the retina to measure visually evoked transmitter release (Ashmore & Falk, 1982). Other earlier methods have estimated both the excitatory and inhibitory conductances using across-trial averages of current injections of different magnitude (Borg-Graham, Monier, & Frégnac, 1996; Häusser & Roth, 1997; Anderson, Carandini, & Ferster, 2000; Wehr & Zador, 2003; Rudolph, Piwkowska, Badoual, Bal, & Destexhe, 2004; Greenhill & Jones, 2007). More recently, conductances have been estimated from a single trace by applying a diverse range of probabilistic inference methods. In some of those studies, the size of the excitatory and inhibitory inputs is assumed to be identical, fixed, and known a priori, while the synaptic inputs were assumed to be δ-functions, with instantaneous rise and decay time and Poisson statistics (Kobayashi, Shinomoto, & Lansky, 2011). Some of the assumptions were relaxed in Paninski, Vidne, Depasquale, and Ferreira (2012), where the number of inputs in a time window followed either an exponential or a truncated gaussian distribution, but the synaptic decay time constant had to be known a priori. Finally, Lankarany, Zhu, Swamy, and Toyoizumi (2013) further generalized the distribution of the number of inputs in a time window by making use of a mixture of gaussians. This method allowed a good estimation of the conductance traces even when the distribution of synaptic amplitudes has long tails.

These methods typically attempt to recover the global excitatory and inhibitory conductances. Here instead we focus on recovering the statistics of the constituent synaptic events. Specifically, we introduce a method that aims to infer the event rate, synaptic time constants, and distribution of synaptic event amplitudes from the power spectral density and statistical moments of the observed current trace. We applied our method to voltage-clamp traces of electrotonically compact interneurons recorded in the cerebellum of awake mice. We find that during voluntary locomotion, the excitatory input rate increases from 600 to 1000 Hz, while the synaptic event amplitudes remain the same. Our method thus provides a novel way to resolve synaptic event properties *in vivo*.

2 Results

A common method to extract synaptic properties is to identify and analyze isolated events from current traces, but *in vivo* this fails because the

events will overlap (see Figure 1a). To demonstrate the problem explicitly, we simulated a neuron randomly receiving excitatory synaptic events (see below for model details). For illustration purposes, we assume momentarily that the amplitude of all events is identical (50 pA). From the total current recorded in voltage clamp, we attempt to reconstruct the frequency of events and the distribution of their amplitudes.

We used single-event dectection software (see the appendix) to find putative postsynaptic currents (PSCs). At low-input frequencies (50 Hz), most PSCs were correctly identified, and the resulting estimation of the synaptic input amplitude distribution was correct. However, at higher frequencies, when the event interval became shorter than the synaptic decay time, the event frequency was grossly underestimated and reached a plateau (see Figure 1b, left). At this point the individual excitatory postsynaptic currents (EPSCs) overlapped and became indistinguishable. The reason is that the most probable intertime interval of a Poisson process (a common model for the inputs received by a neuron, but see Lindner, 2006) is zero. In addition, as a result of the overlap, the estimated PSC amplitude distribution had peaks at multiples of the original amplitude and the variance of the event amplitude was highly overestimated (see Figure 1b, right). Finally, at high-input frequencies, the traces had to be manually postprocessed to correct mistakes in event detection. This manual processing is time-consuming; even an experienced researcher spent more than 1 hour to analyze a 10 second trace. Thus, at high-input frequencies, single-event analysis is not only incorrect; it is also time-consuming. While a somewhat better result might be achievable when the PSCs are identified by their rising phase, such methods will still fail at high rates.

2.1 Generative Model for the Observed Current Trace.

Unlike the *in vitro* situation, the synaptic properties are not directly accessible from *in vivo* recordings. Instead, the data indirectly and stochastically reflect the synaptic properties. We therefore use a generative model to couple the data, in particular, the statistics of the current trace, to the underlying synaptic properties. We define the generative model as follows. The synaptic inputs are assumed to arrive according to a Poisson process with a rate ν (see Figure 1a and section 3). The synaptic events are modeled with a biexponential time course, as this can accurately fit most fast synapses (e.g., Roth & van Rossum, 2009),

$$f(t) = (1 - e^{-t/\tau_1})e^{-t/\tau_2} \quad (t > 0) \tag{2.1}$$

with rise time τ_1 and decay time τ_2. While we initially assumed that all PSCs have the same time constants, the effect of heterogeneous time constants is studied below. The total current is

$$I(t) = \sum_{k=1}^{K} a_k f(t - t_k),$$ (2.2)

where t_k denotes the time of event k and a_k is the amplitude of that event. Unlike the schematic example above, the event amplitudes are drawn from a synaptic amplitude distribution $P(a)$ (with $a \geq 0$). This distribution captures the spread of amplitudes across the population of synapses, as well as variation in single-synapse-event amplitudes due to randomness and nonstationarities such as short-term plasticity.

Although our method is general and not restricted to any specific distribution of synaptic amplitudes, we consider for concreteness the amplitudes to be distributed as either (1) a log-normal distribution (*LN*),

$$P(a) = \frac{1}{p_2 \sqrt{2\pi} a} e^{-\frac{(\ln a - p_1)^2}{2 p_2^2}},$$ (2.3)

with raw moments $\overline{a^n} \equiv \int_0^\infty P(a) a^n \mathrm{d}a = e^{n p_1 + n^2 p_2^2 / 2}$; or (2) a stretched exponential distribution (*SE*),

$$P(a) = \frac{1}{p_1 \Gamma(1 + 1/p_2)} e^{-(a/p_1)^{p_2}},$$ (2.4)

with moments $\overline{a^n} = \frac{p_1^n}{p_2} \Gamma\left((1 + n)/p_2\right) / \Gamma(1 + 1/p_2)$ where $\Gamma(\cdot)$ is the gamma function; or (3) a zero-truncated-normal distribution (*TN*),

$$P(a) = \frac{\phi(a/p_2 + h)}{p_2 [1 - \Phi(h)]},$$ (2.5)

where $h = -p_1/p_2$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the density of a normal distribution with zero-mean and unit variance and its cumulative. The mean $\mu_a = p_1 + p_2 \rho$ and variance $\sigma_a^2 = p_2^2(1 - \rho(\rho - h))$, where $\rho = \phi(h)/[1 - \Phi(h)]$ (for higher moments see, e.g., Horrace, 2013).

These three probability distributions (examples are shown in Figure 2) are commonly used in the experimental and theoretical literature (Song, Sjöström, Reigl, Nelson, & Chklovskii, 2005; Barbour, Brunel, Hakim, & Nadal, 2007; Buzsáki & Mizuseki, 2014). The stretched exponential distribution has a maximum at zero amplitude, while the two other distributions have an adjustable mode but differ in the heaviness of their tails. The *LN* and the *SE* distribution (for $p_2 < 1$) are heavy tailed, while the *TN* distribution is not. Conveniently, all of these distributions are characterized by the two parameters p_1 and p_2. These parameters determine the mean and variance of the events, but as can be seen from the equations for the

moments above, the exact relations are different for each of the three candidate distributions.

2.1.1 Moments of the Synaptic Current. Next, we calculated the current trace $I(t)$ that results from the random inputs. The statistics of the current follow from the distribution of synaptic event amplitudes and the time course of the events according to Campbell's theorem (Rice, 1954; Bendat & Piersol, 1966; Ashmore & Falk, 1982). The cumulants κ_n of the current probability distribution $P(I)$ follow from the event distribution and the synaptic time course as

$$\kappa_n = \nu \overline{a^n} \int_0^\infty [f(t)]^n dt. \tag{2.6}$$

In this equation, the raw moments $\overline{a^n}$ of the synaptic event amplitude distribution $P(a)$ are given above for the different candidate distributions. Furthermore, for the biexponential synaptic kernel $f(t)$ (see equation 2.1) the integrals are $\int_0^\infty [f(t)]^n dt = n!\tau_1\Gamma(n\frac{\tau_1}{\tau_2})/\Gamma(1 + n + n\frac{\tau_1}{\tau_2})$. Finally, the moments of the current trace M_I are expressed in the cumulants κ_n. In practice, we use the first four moments of the current distribution,

$$\mu_I = \kappa_1,$$
$$\sigma_I = \sqrt{\kappa_2},$$
$$\text{skew}(I) = \kappa_3/\kappa_2^{3/2},$$
$$\text{kurtosis}(I) = (\kappa_4 + 3\kappa_2)/\kappa_2^2 - 3. \tag{2.7}$$

We can thus express the statistical moments of the distribution of the observed current trace, equation 2.7, in the underlying model.

2.1.2 Power Spectrum of the Synaptic Current. Also the power spectral density (PSD) of the current $I(t)$ can be expressed in the model parameters. The current is the convolution of a Poisson process, which has a flat power spectrum, with the synaptic kernel. As a result, the PSD is the magnitude of the Fourier transform of the synaptic kernel, equation 2.1, and for nonzero frequencies equals (Puggioni, 2015)

$$\text{PSD}(f)$$
$$= 2\nu(\mu_a^2 + \sigma_a^2)\frac{\tau_2^4}{(\tau_1 + \tau_2)^2 + (2\pi f\tau_2)^2(2\tau_1^2 + 2\tau_1\tau_2 + \tau_2^2) + (2\pi f\tau_2)^4\tau_1^2}. \tag{2.8}$$
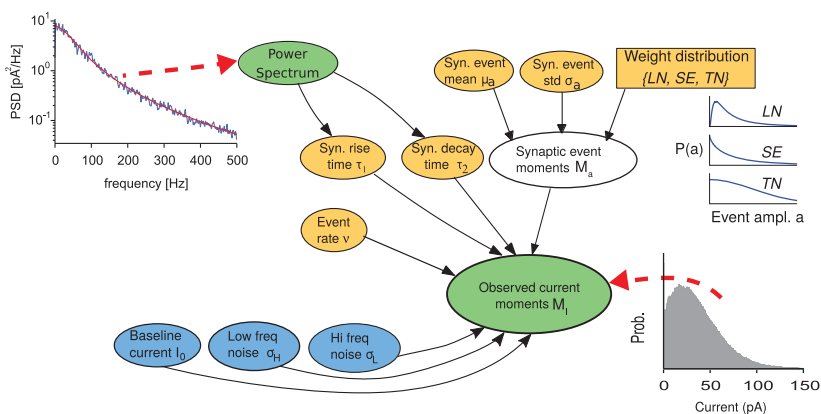
Figure 2: Bayesian network representing the dependencies between the variables. Orange nodes represent variables that have to be inferred from the data, and green nodes stand for variables that are measured directly from the data. The blue nodes are additional contributions to the current in typical experiments. The top left graph shows the PSD fit (the red line is the fit with equation 2.8), and the bottom right graph is the probability distribution of $I(t)$, used to calculate the observed moments M_I. All variables are described in Table 1.

Note that being a second-order statistic, the PSD depends on the mean and variance of the amplitude distribution $P(a)$ only.

2.2 Inference Procedure. Now that we have expressed both the PSD and the moments of the current distribution in the model parameters, one could proceed using classical fitting techniques, such as least square fitting, to find the synaptic parameters that best fit the data. However, we use a probabilistic approach that yields the distribution of parameters that best fit the data. A probabilistic approach is advantageous for three reasons. First, we expect strong correlations between the model parameters, which can cause traditional fitting to fail (see, e.g., Costa, Sjostrom, & van Rossum, 2013). Second, the probabilistic approach naturally yields the probability distribution of possible fit parameters. Third, the probabilistic approach offers a natural way to perform model selection.

We first present an idealized model, which ignores some distortions typical of *in vivo* recordings. Figure 2 shows the Bayesian network and the dependencies among the variables (nodes). The green nodes stand for variables that are measured directly from the data: the PSD and the first four moments of the current $M_I = [\mu_I, \sigma_I, \text{skew}_I, \text{kurtosis}_I]$. Together the data are succinctly denoted **D**. The orange nodes represent variables that are to be inferred. The five parameters of the model are the rate ν, the mean synaptic amplitude μ_a, its variance σ_a, synaptic risetime τ_1, and decay time τ_2,

Table 1: Description of the Parameters and Variables of the Model.

Parameter Name	Description
Measured data D	
$M_I = (\mu_I, \sigma_I, \text{skew}_I, \text{kurtosis}_I)$	Observed first four moments of the current $I(t)$
PSD	Power spectral density of $I(t)$
Parameters of idealized model	
τ_1, τ_2	Rise and decay time of the EPSCs
$S_a = \{LN, SE, TN\}$	Synaptic amplitude distribution = {log-normal, stretched exponential, truncated normal}
μ_a, σ_a	Mean and SD of the amplitude distribution
M_a	Moments of the synaptic amplitude distribution
ν	Frequency of synaptic inputs
Additional parameters of full model	
i_0	Voltage-clamp baseline current
σ_H	SD of high-frequency noise
σ_L	SD of low-frequency fluctuations

as well as the type of distribution S_a (see Table 1). The set of parameters is denoted θ.

Written formally, the joint probability of the Bayesian network in Figure 2 is

$$P_{joint}(\theta, \mathbf{D}) = P(\tau_1|\text{PSD})P(\tau_2|\text{PSD})f(M_a|\mu_a, \sigma_a, S_a)$$
$$\times P(\mu_a)P(\sigma_a)P(\nu)P(M_I|\tau_1, \tau_2, M_a, \nu), \tag{2.9}$$

where we introduced the vector of moments of the synaptic amplitude distribution, M_a (see below).

From this equation the parameter distribution given the data $P(\theta|\mathbf{D})$ follows as $P(\theta|\mathbf{D}) \propto P_{joint}(\theta, \mathbf{D})$. We now describe P_{joint} and the probabilistic dependencies among the nodes term by term. The first two terms infer the values for the synaptic time constants from the PSD. Since we cannot obtain an analytic expression of the likelihood of the PSD, we use empirical Bayes to set the prior on the time constants of the postsynaptic current (Casella, 1985). We fit the shape of equation 2.8 with a least square method to the PSD to find τ_1 and τ_2 (see the top-left inset in Figure 2; note that the values ν, μ_a, σ_a are not needed to perform this fit). Since we found the relative cross-terms of the Hessian matrix between τ_1 and τ_2 to be very small (< 0.005), we model the time constants with independent gaussian distributions with mean and variance given by the mean and the Hessian of the PSD fit. A common criticism of empirical Bayes is that it uses data for both prior and inference, thus double-counting the data. Here, however, the PSD data are used to set the prior on the time constants, but the data are not used as evidence in the inference process (see Figure 2).

The next term in equation 2.9 is $f(M_a|\mu_a, \sigma_a, S_a)$. This is a deterministic function, because the moments of the synaptic amplitude distribution M_a are fully determined by μ_a, σ_a and the type of amplitude distribution type (see equation 2.3–2.5). The parameters μ_a, σ_a, and ν are given uninformative uniform priors spanning a reasonable and positive range of values.

The final term in equation 2.9, the likelihood of the moments of the current $P(M_I|\tau_1, \tau_2, M_a, \nu)$, cannot be calculated analytically. Although equation 2.6 gives the expected value, M_I is a stochastic quantity that due to the Poisson process is different on each run and thus requires simulation. Below, we present a method to speed up its calculation.

2.2.1 Inclusion of In Vivo Variability and Other Experimental Confounds. In *vivo* voltage-clamp recordings show a number of effects that need to be included in the model with additional parameters. The first additional feature is the baseline current (i_0) of the voltage clamp that has to be subtracted from the current. In *in vitro* situations, one can estimate it by finding the baseline of the current trace, but due to the high input rates, this is challenging for *in vivo* recordings. Instead a prior probability of $P(i_0)$ was included. It was normally distributed with mean and variance estimated with an informed guess, reflecting the uncertainty in the value of i_0.

The second feature is the inclusion of high-frequency noise coming, for instance, from the recording setup or the stochastic opening and closing of ion channels. Its standard deviation σ_H is measured experimentally, and we model it as a zero mean Ornstein-Uhlenbeck (OU) process,

$$dU_{H_t} = -\tau_H U_{H_t} dt + \sigma_H \sqrt{2/\tau_H} dB_t, \tag{2.10}$$

where B_t is a Wiener process and the cutoff frequency is $1/(2\pi \tau_H) = 600$ Hz.

Finally we include low-frequency fluctuations typically present in *in vivo* synaptic activity (e.g., Schiemann et al., 2015). We relax the constant rate assumption by adding a modulation term to the Poisson rate, which is modeled as an OU process with power σ_L^2 and cutoff frequency $f_L = 1/(2\pi \tau_L)$ of 5 Hz:

$$dU_{L_t} = -\tau_L U_{L_t} dt + \sigma_L \sqrt{2/\tau_L} dB_t. \tag{2.11}$$

As a result, in the expression for the PSD, equation 2.8, the rate ν is replaced by $(\nu + PSD_{OU}(f))$, where the power spectrum of the OU process is given by $PSD_{OU}(f) = \sigma_L^2 \tau_L / [1 + (2\pi f \tau_L)^2]$. To find the variance of this slow noise, we fit the PSD with equation 2.8 in a range above f_L and calculate its integral σ_{th} (the theoretical standard deviation of the modulation-free trace). The value of f_L was set heuristically as the minimal value that resulted in a good fit. Since the observed variance of the signal σ_{obs}^2 is the sum of σ_{th}^2, σ_H^2, and σ_L^2

(the slow component is independent of the underlying process), it follows that $\sigma_L^2 = \sigma_{obs}^2 - \sigma_H^2 - \sigma_{th}^2$.

These three additional features are depicted by the blue nodes in Figure 2. The joint probability for the full model becomes

$$P_{joint} = P(\tau_1|\text{PSD})P(\tau_2|\text{PSD})P(M_a|\mu_a, \sigma_a, \text{S}_a)P(\mu_a)P(\sigma_a)P(\nu)$$

$$\times P(i_0)P(\sigma_H)P(\sigma_L|\text{PSD})P(M_I|\tau_1, \tau_2, M_a, \nu, i_0, \sigma_H, \sigma_L). \quad (2.12)$$

2.2.2 Description of the Sampling Algorithm. In the Bayesian framework, the posterior probabilities of the parameters of the model can be estimated by sampling from P_{joint}, for instance, using a suitable Markov chain Monte Carlo (MCMC) algorithm. However, this approach is very slow, because the likelihood does not have a closed form and has to be estimated with multiple simulations after each MCMC sample. As the estimation of the likelihood takes about 1 minute on a standard PC, a typical MCMC run of about 100,000 samples would take approximately two months.

We introduce a speed-up that can be used whenever a likelihood can be obtained only by sampling from the generative model, but its means can be calculated analytically. The idea is to fit the likelihood with a kernel density estimate (KDE). Assuming that the shape of the likelihood does not depend much on the parameter values, the same KDE can be exploited to approximate the likelihood for different parameter values. As a result, we can keep the shape fixed, but we translate it to a new location determined by the analytically calculated average moments of the likelihood. A thorough validation shows that the method correctly infers the parameters across a wide range of biologically plausible values (see below).

To perform the inference, we first initialize the parameters $\{\tau_1, \tau_2, M_a, \nu\}$ by least-square-fitting equation 2.7 to the observed moments and equation 2.8 to the observed PSD. Next, we run the generative model multiple times to calculate the shape of $P(M_I|\tau_1, \tau_2, M_a, \nu)$ using an exponential KDE. Finally, during the main MCMC run where we sample P_{joint}, we keep the shape fixed, but at each step we translate it to the location of the analytically calculated average moments (see equation 2.6 and the red crosses in Figure 3).

2.3 Validation on Simulated Data. To validate the method, we simulated 10 s current traces with known parameters and apply our inference method to recover their values. One parameter at a time was varied, while the other parameters were set to a default value ($\mu_a = 50$ pA, $\sigma_a = 40$ pA, $\nu = 700$ Hz). The rise time $\tau_1 = 0.3$ ms and the decay time $\tau_2 = 2$ ms are based on *in vivo* PSCs in cerebellar interneurons (Szapiro & Barbour, 2007, and below). We first assumed that the shape of the synaptic amplitude distribution (*LN*,*SE*, or *TN*) is known a priori. Figure 4a compares the estimated parameters versus their true value. The inference works well in a
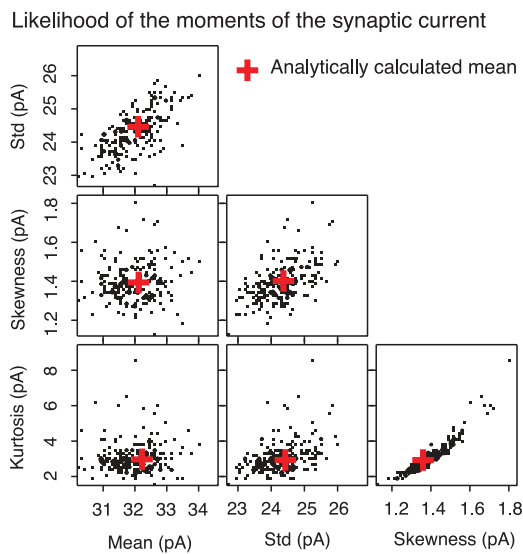
Likelihood of the moments of the synaptic current

Figure 3: Likelihood of the moments of the synaptic current $P(M_I | \tau_1, \tau_2, M_a, \nu)$ as estimated by simulating the generative model multiple times with fixed parameters τ_1, τ_2, M_a, ν. Red crosses represent the analytic predictions of the expected values of the M_I.

physiologically plausible range, and the true value is almost always within the confidence interval. The largest error bars occur when either the mean event amplitude is small or the standard deviation is large (i.e., the CoV is large).

The approach also yields the inferred joint distribution for a given parameter setting. The posterior distribution of the parameters contains the true values in the region of maximal density (see Figure 4b). Unlike single-point estimates (e.g., maximum a posteriori MAP estimates), one can also evaluate the dependencies between the parameters. In particular we observe a strong anticorrelation between event rate and event size (bottom left panel). In other words, the model compensates for changes in the rate by changing the estimate for the event size; their product is approximately invariant.

2.3.1 Model Selection. Next, we tested whether the method is able to recover the correct amplitude distribution (*LN, SE,* or *TN*) when it is not known a priori. The Bayesian framework offers straightforward tools to assess the likelihood of a model, such as the deviance information criterion (DIC) (Spiegelhalter, Best, Carlin, & Linde, 2002). The higher the DIC is, the less likely is the model suitable to describe the data, and this would be
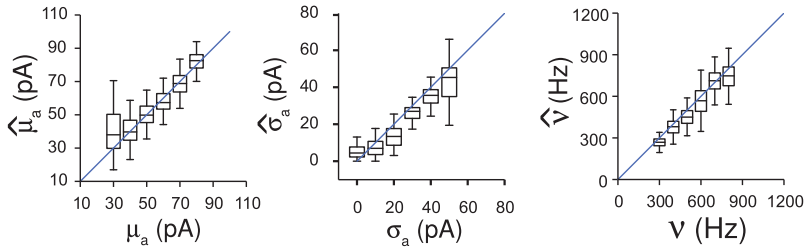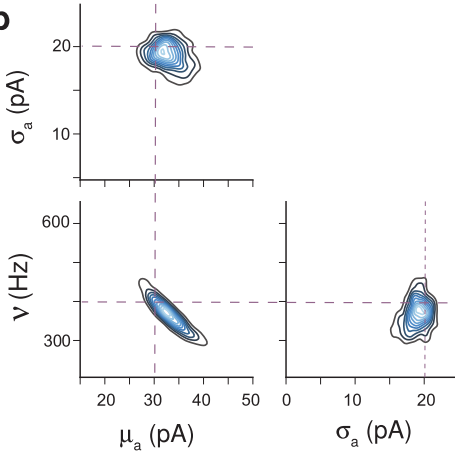
a

b

Figure 4: Validation of the method on simulated data. (a) The inference gives good results in a physiological range of parameters. True (*x*-axis) versus estimated parameters. The boxes represent the 33rd and 66th quantile of the distributions, the whiskers indicate the full range. The synaptic inputs were drawn from a log-normal distribution. (b) Inferred distribution of model parameters across multiple trials generated with fixed parameters (indicated with dashed lines). The contour lines are linearly spaced between maximal and minimal probability.

the simplest way to choose the most likely distribution. However, the DIC value is a random variable that varies from trial to trial. Thus, rather than selecting the lowest DIC, we use Bayesian model comparison based on the distribution of the DIC values. We generated 100 traces using a given amplitude distribution and run the inference algorithm assuming either *LN*, *SE*, or *TN* amplitude distribution and calculate the DIC for each mode (see Figure 5a). From the three DIC values of the three models DIC_{LN}, DIC_{SE}, and DIC_{TN} (corresponding to the *LN*, *SE*, and *TN* model, respectively), we calculate two quantities: $\Delta_{LT} = DIC_{LN} - DIC_{TN}$ and $\Delta_{LE} = DIC_{LN} - DIC_{SE}$.
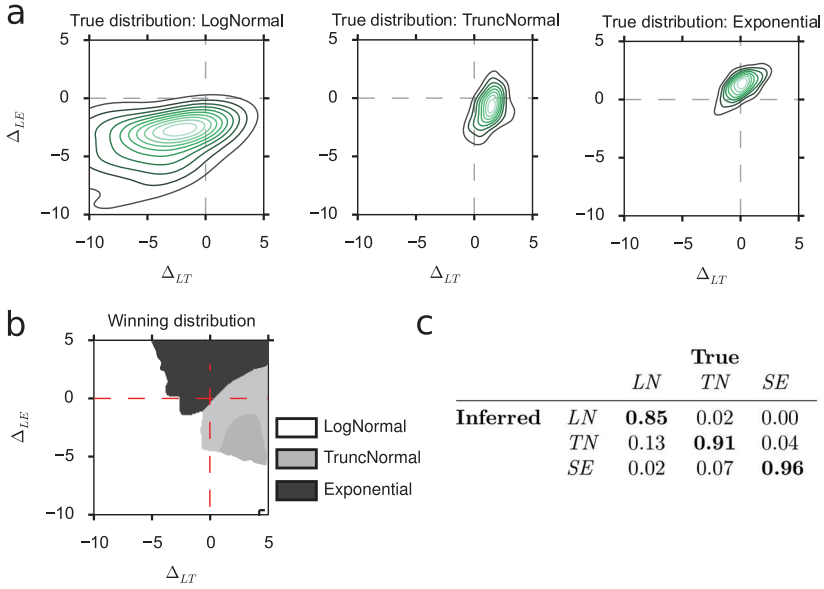
Figure 5: Inference of the underlying weight distribution of simulated data. (a) The distribution of DIC differences for the three simulation weight distributions. As the shapes of the distributions differ, we used Bayesian model selection. The contour lines are linearly spaced. (b) The resulting maximum likelihood solution that tells which underlying distribution is most likely. (c) Performance of the algorithm to recover the correct weight distribution (expressed as fraction correct, based on 100 runs).

To find the most likely amplitude distribution, we apply Bayes theorem and calculate

$$P(X|\Delta_{LE}, \Delta_{LT}) = \frac{P(\Delta_{LE}, \Delta_{LT}|X)P(X)}{\Sigma_{Y \in [LN, TN, SE]}P(\Delta_{LE}, \Delta_{LT}|Y)P(Y)}$$

$$= \frac{P(\Delta_{LE}, \Delta_{LT}|X)}{\Sigma_{Y \in [LN TN, SE]}P(\Delta_{LE}, \Delta_{LT}|Y)}, \qquad (2.13)$$

where, in the second line, we assumed that each amplitude distribution is a priori equiprobable. Thus, for each point in the space (Δ_{LE}, Δ_{LT}), we select the distribution that has the highest probability according to equation 2.13 (see Figure 5b). This method is able to correctly identify the amplitude distribution with approximately 90% accuracy (see Figure 5c).

2.3.2 Robustness of Method. We examined the robustness of the method in a number of ways. First, we explored how the posterior of the parameters
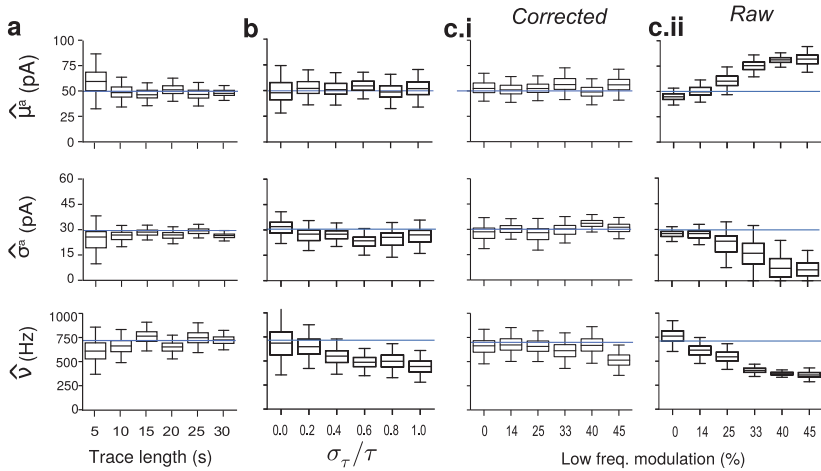
Figure 6: Robustness of inference demonstrated on simulated data. The traces were simulated with input amplitude drawn from a log-normal distribution ($\mu_a = 50$ pA, $\sigma_a = 30$ pA) and $\nu = 700$ Hz. (a) The estimates are robust using current traces of about 10 s or longer. For shorter traces, the inference is based on too few data and deteriorates. Top: Estimate of mean single-event amplitude μ_a; middle: standard deviation of amplitude; bottom: event rate. (b) Robustness to heterogeneity in the synaptic time constants as expressed in the CV of the rise- and decay-time constants, both independently drawn from a truncated normal distribution. (c) Robustness to *in vivo* variability when an inhomogeneous low-frequency (< 5 Hz) component is added to the Poisson rate. The parameters' estimation is plotted against the contribution (in percentage) of the low-frequency modulation to the total standard deviation. i: Result after correcting the PSD at low frequency (text). ii: Without the correction, substantial biases arise.

depends on the length of the trace. Longer traces should lead to less uncertainty and yield narrower, more precise distributions because more statistics are collected. However, short intervals are preferable, because they allow the analysis of shorter periods in *in vivo* traces and one can see more rapid modulation in the synaptic inputs. Indeed, longer traces lead to less uncertainty on the parameter (see Figure 6a). The analysis shows that 10 second recordings in general are enough to obtain a reasonable estimation of the parameters.

Next, *in vivo* PSCs rise and decay times might vary across synapses as different synapses may have different kinetic properties and may be subject to different amounts of dendritic filtering (Williams & Mitchell, 2008). To test whether our model performs well when the shape of the PSCs varies, both time constants that determine the PSC shape were independently drawn from truncated normal distributions for each PSC. When the

heterogeneity of the time constants was modest (CV $\lesssim$ 0.3), the model correctly extracted μ_a, σ_a, and ν, but at larger values, the frequency estimate in particular diverges (see Figure 6b). One reason might be that our model assumes fixed time constants, but in this case, the cumulants of equation 2.6 should actually be averaged over the distribution of time constants.

Finally, we tested what happens when *in vivo* activity breaks the stationary assumption of the homogeneous Poisson model and inputs typically fluctuate on a slow timescale. To test the robustness of our model, we generate *in vivo*–like traces by adding an inhomogeneous component to the Poisson rate, modeled as an OU process with 5 Hz cutoff frequency. When we applied the correction described above, the model performs well even in the presence of considerable fluctuations in the synaptic input rate (see Figure 6c.i). Including the correction is important; without it, large biases arise (see Figure 6c.ii).

2.4 Inference Method Applied to Cerebellar in Vivo Data. We applied our inference method to *in vivo* recordings obtained from cerebellar interneurons in the molecular layer (basket and/or stellate cells). These neurons are ideal to test our method as they are electronically compact (Kondo & Marty, 1998), although some cable filtering can be observed in older animals (Abrahamsson, Cathala, Matsui, Shigemoto, & DiGregorio, 2012). The voltage clamp held neurons at -70 mV to isolate excitatory inputs. The head-restrained mice displayed bouts of self-paced voluntary locomotion on a cylindrical treadmill (see Figure 7a). All traces ($n = 8$) were 90 seconds and contained at least 10 seconds of movement. Locomotion modulates subthreshold and spiking activity in a large number of brain regions (Dombeck, Graziano, & Tank, 2009; Polack, Friedman, & Golshani, 2013; Schiemann et al., 2015). In cerebellar interneurons, locomotion is associated with increased excitatory input drive (see Figure 7b). In particular we were interested in what underlies this increased drive. For instance, it could be caused by increased frequency, increased amplitude as an effect of neuromodulation, or recruitment of a distinct set of synapses.

To apply our method we extracted the PSD and distribution from the current trace (see Figure 7b). We corrected for the low-frequency modulation as described above, while the high-frequency noise was measured directly from the system. The subsequent inference showed that the increase in excitatory synaptic current is associated with an increased input frequency, shown for a representative trace in Figure 7c (bottom panel). However, in this cell movement did not lead to changes in the mean amplitude or the standard deviation of the synaptic amplitudes (see Figure 7c, top and middle panels). During movement, the input frequency roughly doubled, from 484 to 1233 Hz. The synaptic time constants found by fitting the power spectrum of the current were $\tau_1 = 0.25 \pm 0.04$ ms and $\tau_2 = 1.56 \pm 0.21$ ms (mean $\pm$ standard error), comparable to the 20% to 80% rise time of 0.41 ± 0.14 ms and the 1.85 ± 0.52 ms decay reported in slice (Szapiro & Barbour, 2007).
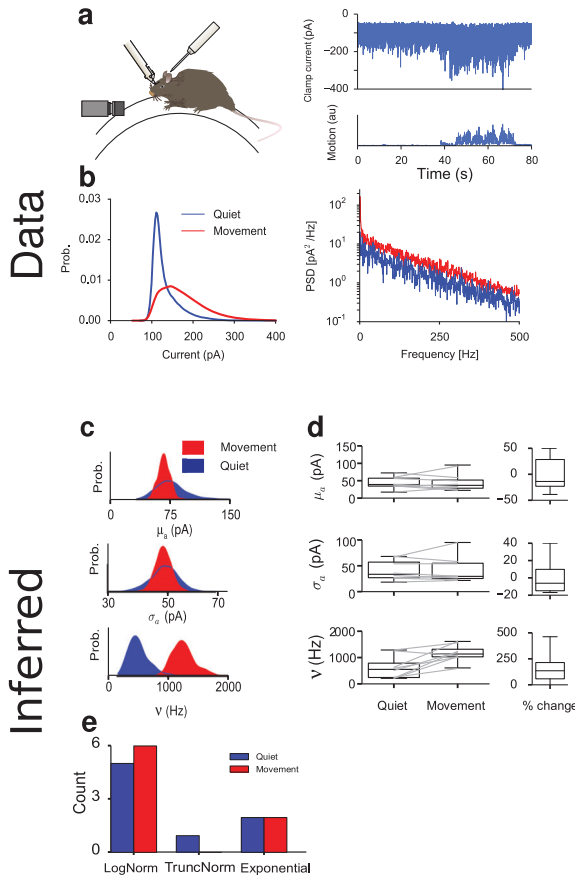
Figure 7: Analysis of *in vivo* voltage-clamp recordings. (a) Experimental setup: head-fixed awake mice, walking voluntarily on a wheel. Right: Voltage-clamp current in cerebellar interneurons (top) and simultaneously recorded animal movement (bottom). Periods of movement are accompanied by an increased excitatory current in the neuron. (b) Left: Observed current distribution in the moving and quiet periods. Note that due to the high input frequency, periods with zero current are very rare. Right: Samples of the recorded power spectral density. (c) Posterior distribution of the input parameters of a representative interneuron (under log-normal assumption, the most likely distribution for this neuron). While the distribution of μ_a and σ_a during movement is narrower, their maximum likelihood value is hardly changed. (d) Inference of the synaptic input parameters across eight recordings displaying an increase in the input frequency during movement but not in the mean or variance of the event amplitude. (e) Classification of the synaptic event amplitude distribution. In both conditions both log-normal and stretched exponential distributions were observed. The truncated normal was inferred only once. Error bars denote the (min, max) range, boxes the 25th to 75th percentile, horizontal bar the median.

Table 2: Summary of the MAP Values of the Parameters Estimated from $n = 8$ *in Vivo* Recordings.

	Quiet		Movement			Power	Power to Detect
	Mean	SE	Mean	SE	*p*-Value	of Data	10% Change
τ_1	0.27 ms	0.03 ms	0.28 ms	0.03 ms	0.24	0.15	0.61
τ_2	1.68 ms	0.22 ms	1.65 ms	0.19 ms	0.61	0.1	0.74
μ_a	42.8 pA	8.7 pA	43.2 pA	7.9 pA	1.00	0.04	0.83
σ_a	31.3 pA	6.2 pA	31.0 pA	4.9 pA	0.86	0.04	0.42
ν	585 Hz	153 Hz	1006 Hz	80 Hz	0.03	0.93	0.07

Across the population, the MAP estimates of μ_a, σ_a, and ν during quiet wakefulness and movement show a similar pattern (see Figure 7d and Table 2). Note that given the small changes between quiet and moving state, the power of the test calculated from the data is low, but 10% changes would be detected with high probability.

Next, we applied our inference method to each trace using the *LN* (log-normal), *SE* (stretched exponential), and *TN* (truncated normal) distribution and determined which synaptic amplitude distribution was the most likely, where it should be kept in mind that the statistical power of the data is limited. In general, we found that during both quiet periods and movement, the most likely distributions were heavy-tailed, being either LN or SE (with exponent on average 0.8, range 0.7–1.2; see Figure 7e). In particular, during active periods, the *LN* distribution (the most common) was significantly more likely than the *TN* ($p = 0.046$), but the *SE* distribution was not significantly less likely ($p = 0.37$). Thus, while this suggests that the distribution is stretched, the current data cannot distinguish between the *LN* and *SE* types. Next, we wondered if the shape changed using a binomial test. For instance, if a recording yielded five times the LN distribution out of eight data traces during the quiet period, and did so six out of eight times during the active period, there is no evidence for a change. We found no evidence for a change in the distribution shape between quiet and active period (*LN*, $p = 0.78$; *SE*, $p = 0.96$; *TN*, $p = 0.71$).

Finally, we compared our estimates to a standard single-event extraction method (see the appendix). Because the event extraction method fails at frequencies higher than about 500 inputs per second, the frequency of the synaptic inputs is underestimated by a factor two due to the misclassification of overlapping events.

3 Discussion

In the past decade, numerous studies have been published using voltage-clamp data from anesthetized animals to investigate the contribution of

excitation and inhibition to the V_m dynamics, with recordings from auditory cortex (Wehr & Zador, 2003; Poo & Isaacson, 2009; Liu et al., 2010), visual cortex (Liu et al., 2010; Haider, Hausser, & Carandini, 2012), and prefrontal cortex (Haider, Duque, Hasenstaub, & McCormick, 2006). However, in these experiments, only the total excitatory or inhibitory contributions can be extracted; therefore, they are unable to distinguish properties of single synapses and changes therein. We proposed a novel probabilistic method to infer the synaptic time constants, the mean and variance of the synaptic event amplitude distribution, and the synaptic event rate from *in vivo* voltage-clamp traces. Moreover, the method accurately recovers the shape of the distribution of synaptic inputs. The inference is robust to slow fluctuations of synaptic input rate, experimental noise, and heterogeneity in the time constants of the PSCs.

The extracted distribution reflects the amplitude of events received by the neuron. It therefore includes not only variations across synapses but also variation due to synaptic unreliability and heterogeneity from effects like short-term synaptic plasticity (Szapiro & Barbour, 2007; Abrahamsson et al., 2012). Furthermore, the contribution of each synapse is weighted by its own input rate: synapses receiving inputs at higher rates will contribute more to the estimated amplitude distribution than synapses receiving low rates. Our method thus captures the effective distribution of synaptic inputs in an *in vivo* recording and thereby complements techniques that infer the amplitude distribution either anatomically from spine size or from paired recordings *in vitro* and that are not weighted by the input rate.

Applied to voltage-clamp recordings from cerebellar interneurons of awake mice, we found that the excitatory synaptic amplitude distribution is either a stretched exponential or log normal. This means that the probability for large events is larger than for a gaussian with the same mean and variance. Such heavy-tailed distributions have been observed in a number of systems (Sayer, Friedlander, & Redman, 1990; Song, Sjöström, Reigl, Nelson, & Chklovskii, 2005; Barbour et al., 2007; Ikegaya et al., 2013) and are believed to be an important characteristic of neural processing (Koulakov, Hromádka, & Zador, 2009; Roxin, Brunel, Hansel, Mongillo, & Vreeswijk, 2011; Teramae, Tsubo, & Fukai, 2012). While any distribution can be tested (although for efficiency reasons the moments should ideally be available analytically), a future goal is to reconstruct the amplitude distribution directly, for instance, by reconstructing it from its moments. However, there are currently no fully satisfactory mathematical methods to achieve this.

Furthermore, we found no evidence that the synaptic amplitude distribution changes in these neurons when the animal is moving. Instead, the increase of the excitatory current during movement is due to the higher frequency of the inputs. The most parsimonious explanation is that all inputs, big and small, increase their rates similarly during movement. However, it is important to remember that the method is based on the ensemble of inputs. While our findings are inconsistent with a case where only large

inputs become more active and a case where all single synaptic events become stronger by, say, neuromodulation, we cannot rule out that, for instance, a second population of inputs with an identical amplitude distribution becomes active during movement.

We summarize restrictions of the method. First, as in most other methods, the *in vivo* traces need to be stationary over a period long enough to accumulate sufficient statistics. The second assumption is that the neurons are electronically compact such that a good "space clamp" can be achieved, which is problematic for Purkinje and pyramidal neurons (Williams & Mitchell, 2008). It would be of interest to examine how robust our method is toward deviations from this (Abrahamsson et al., 2012), for example, using compartmental simulations.

The third assumption is that synaptic inputs are uncorrelated and follow a Poisson distribution. Experimental measurements of correlations in the brain are contradictory and largely depend on what timescale is considered, reviewed in Cohen and Kohn (2011). Notably, slow correlations are visible in the PSD, adding a component with a different time constant (Moreno-Bote, Renart, & Parga, 2008). When fitting the PSD of *in vivo* data, we observed a bump in activity in the low frequencies ($f < 10$ Hz) that could correspond to spike correlations on timescales longer than about 15 ms. Such correlations are included in our model. The method would not be able to identify spike correlations on the order of the synaptic time constants (τ_1 and τ_2) because they would contribute to the PSD in the same frequency range. However, it is generally believed that spike count correlations on a short timescale (about 2 ms) are small, normally below 0.03 (Smith & Kohn, 2008; Helias, Tetzlaff, & Diesmann, 2014; Grytskyy et al., 2013; Renart et al., 2010; Ecker et al., 2010), and thus the inference would likely still give correct results. Recent experimental evidence shows that high-frequency firing Purkinje cells contact interneurons directly (Witter, Rudolph, Pressler, Lahlaf, & Regehr, 2016), which could lead to strongly correlated input. We did not observe bumps in the interneuron powerspectum at around 200 Hz to 250 Hz, the typical cerebellar oscillation frequency (de Solages et al., 2008). Nevertheless, when applying this method, one should be aware of the possibility of such effects. Finally, in these population measurements, truly instantaneous correlations, where multiple events arrive simultaneously, in principle can never be distinguished from altered distributions. However, the error associated with this effect is likely limited. Consider a neuron that receives inputs of equal amplitude a at a rate v. If the inputs have correlation $c = 0.05$, every 100 events, as a first approximation, one will observe on average only 95 events: 90 of size a and 5 of size $2a$. In general, for a given correlation c, the observed frequency is $v_{obs} = v_{true}(1 - c)$ and the observed average amplitude $a_{obs} = a_{true}/(1 - c)$. Thus, even assuming $c = 0.05$, the error in the estimate would be 5% or less.

While the robustness of the method can be tested further using (multicompartment) simulations, physiological validation is much harder; even

with optogenetics, there is no obvious way to generate high-frequency Poisson-like input trains.

The method used the first four moments of the measured current. While in order to infer the distribution shape, one needs at least three moments, we found that using only the first three led to a consistent overestimate of the event amplitude with some 10 pA (not shown). In contrast, higher moments are difficult to estimate with brief recordings. Thus, four seems a good middle ground for the recordings analyzed here.

In summary, commonly used methods to analyze *in vivo* voltage-clamp data cannot infer the single-event statistics at all or introduce large errors. Instead the proposed method represents an important step in extracting such information from *in vivo* intracellular recordings.

Appendix: Implementation Details

We implemented the model in PyMC2, a Python package to perform Bayesian computation (Patil, Huard, & Fonnesbeck, 2010), using a Metropolis Hastings sampler, with normal proposal distribution and standard deviation in each dimension equal to 1 over the absolute value of the parameters. Usually the autocorrelation of the chains was about 300 to 500 samples, and the burn-in phase was about 10 effective samples. To construct the posterior, we generated 150,000 samples yielding approximately 400 effective samples and assessed the mixing by using the Geweke method provided by the PyMC package. The computational analysis tools and data are available at https://github.com/ppuggioni/invivoinfer.

To compare our method to traditional single-event detection methods, we employed TaroTools, a freely available IgorPro package (see sites.google.com/site/tarotoolsregister/) to detect putative postsynaptic currents (PSCs).

The experimental data are described in detail elsewhere (Jelitai, Puggioni, Ishikawa, Rinaldi, & Duguid, 2016). Briefly, whole-cell patch clamp recordings of molecular layer interneurons (basket and stellate cells) were obtained from awake, behaving, but head-restrained, mice at a depth of 100 to 300 μm from the pial surface of the cerebellum, using a Multiclamp 700B amplifier (Molecular Devices, USA). The signal was filtered at 6–10 kHz and acquired at 10–20 kHz using PClamp 10 software in conjunction with a DigiData 1440 DAC interface (Molecular Devices). Patch pipettes (5–8 MΩ) were filled with internal solution (285–295 mOsm) containing (in mM): 135 K-gluconate, 7 KCl, 10 HEPES, 10 sodium phosphocreatine, 2 MgATP, 2 Na2ATP, 0.5 Na2GTP and 1 mg/ml biocytin (pH adjusted to 7.2 with KOH). External solution contained (in mM): 150 NaCl, 2.5 KCl, 10 HEPES, 1.5 CaCl2, 1 MgCl2 (adjusted to pH 7.3 with NaOH). While biocytin was included in the pipette for histological identification, to allow for offline classification interneuron type, our recovery rate was relatively low (<10%)

Thus, we were unable to differentiate further between different interneuron subtypes in this study.

To detect movement, the animals were filmed using a moderate frame rate digital camera (60 fps) synchronized with the electrophysiological recording. We defined a region of interest (ROI) covering the forepaws, trunk and face and calculated a motion index between successive frames (as in Schiemann et al., 2015). All movements (positioning, grooming and locomotion) were included.

Acknowledgments

References

Abrahamsson, T., Cathala, L., Matsui, K., Shigemoto, R., & DiGregorio, D. A. (2012). Thin dendrites of cerebellar interneurons confer sublinear synaptic integration and a gradient of short-term plasticity. *Neuron*, *73*(6), 1159–1172.

Anderson, J. S., Carandini, M., & Ferster, D. (2000). Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J. Neurophysiol.*, *84*(2), 909–926.

Ashmore, J., & Falk, G. (1982). An analysis of voltage noise in rod bipolar cells of the dogfish retina. *Journal of Physiology*, *332*, 273–297.

Barbour, B., Brunel, N., Hakim, V., & Nadal, J. P. (2007). What can we learn from synaptic weight distributions? *Trends Neurosci.*, *30*(12), 622–629.

Bendat, J., & Piersol, A. (1966). *Measurement and analysis of random data*. Hoboken, NJ: Wiley.

Borg-Graham, L. J., Monier, C., & Frégnac, Y. (1996). Voltage-clamp measurement of visually-evoked conductances with whole-cell patch recordings in primary visual cortex. *J. Physiol. Paris*, *90*(3–4), 185–188.

Buzsáki, G., & Mizuseki, K. (2014). The log-dynamic brain: How skewed distributions affect network operations. *Nat. Rev. Neurosci.*, *15*(4), 264–278.

Casella, G. (1985). An introduction to empirical Bayes data analysis. *American Statistician*, *39*(2), 83–87.

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.*, *14*(7), 811–819.

Costa, R. P., Sjostrom, P. J., & Rossum, M. C. W. (2013). Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Front. Comput. Neurosci., 7*, 75.

de Solages, C., Szapiro, G., Brunel, N., Hakim, V., Isope, Buisseret, P., . . . Léna, C. (2008). High-frequency organization and synchrony of activity in the Purkinje cell layer of the cerebellum. *Neuron*, *58*(5), 775–788.

Dombeck, D. A., Graziano, M. S. A., & Tank, D. W. (2009). Functional clustering of neurons in motor cortex determined by cellular resolution imaging in awake behaving mice. *J. Neurosci.*, *29*(44), 13751–13760.

Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., & Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, *327*(5965), 584–587.

Greenhill, S. D., & Jones, R. S. G. (2007). Simultaneous estimation of global background synaptic inhibition and excitation from membrane potential fluctuations in layer III neurons of the rat entorhinal cortex in vitro. *Neuroscience*, *147*(4), 884–892.

Grytskyy, D., Tetzlaff, T., Diesmann, M., & Helias, M. (2013). A unified view on weakly correlated recurrent networks. *Front. Comput. Neurosci.*, *7*, 131.

Haider, B., Duque, A., Hasenstaub, A. R., & McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *Journal of Neuroscience*, *26*(17), 4535–4545.

Haider, B., Hausser, M., & Carandini, M. (2012). Inhibition dominates sensory responses in the awake cortex. *Nature*, *493*(7430), 97–100.

Häusser, M., & Roth, A. (1997). Estimating the time course of the excitatory synaptic conductance in neocortical pyramidal cells using a novel voltage jump method. *Journal of Neuroscience*, *17*(20), 7606–7625.

Helias, M., Tetzlaff, T., & Diesmann, M. (2014). The correlation structure of local neuronal networks intrinsically results from recurrent dynamics. *PLoS Comput. Biol.*, *10*(1), e1003428.

Horrace, W. (2013). Moments of the truncated normal distribution. *J. Prod. Anal.*, *43*, 1–6.

Ikegaya, Y., Sasaki, T., Ishikawa, D., Honma, N., Tao, K., Takahashi, N., . . . Matsuki, N. (2013). Interpyramid spike transmission stabilizes the sparseness of recurrent network activity. *Cerebral Cortex*, *23*(2), 293–304.

Jelitai, M., Puggioni, P., Ishikawa, T., Rinaldi, A., & Duguid, I. (2016). Dendritic excitation–inhibition balance shapes cerebellar output during motor behaviour. *Nature Communications*, *7*, 13722.

Katz, B., & Miledi, R. (1972). The statistical nature of the acetylcholine potential and its molecular components. *Journal of Physiology*, *224*(3), 665–699.

Kobayashi, R., Shinomoto, S., & Lansky, P. (2011). Estimation of time-dependent input from neuronal membrane potential. *Neural Comput.*, *23*(12), 3070–3093.

Kondo, S., & Marty, A. (1998). Synaptic currents at individual connections among stellate cells in rat cerebellar slices. *J. Physiol.*, *509*(pt. 1), 221–232.

Koulakov, A. A., Hromádka, T., & Zador, A. M. (2009). Correlated connectivity and the distribution of firing rates in the neocortex. *J. Neurosci.*, *29*(12), 3685–3694.

Lankarany, M., Zhu, W.-P., Swamy, M. N. S., and Toyoizumi, T. (2013). Inferring trial-to-trial excitatory and inhibitory synaptic inputs from membrane potential using gaussian mixture Kalman filtering. *Front. Comput. Neurosci.*, *7*, 109.

Lindner, B. (2006). Superposition of many independent spike trains is generally not a Poisson process. *Physical Review E*, *73*(2), 022901.

Liu, B., Li, P., Sun, Y. J., Li, Y., Zhang, I., & Tao, H. W. (2010). Intervening inhibition underlies simple-cell receptive field structure in visual cortex. *Nat. Neurosci.*, *13*(1), 89–96.

Moreno-Bote, R., Renart, A., & Parga, N. (2008). Theory of input spike auto- and cross-correlations and their effect on the response of spiking neurons. *Neural Comput.*, *20*(7), 1651–1705.

Paninski, L., Vidne, M., Depasquale, B., & Ferreira, D. G. (2012). Inferring synaptic inputs given a noisy voltage trace via sequential Monte Carlo methods. *J. Comput. Neurosci.*, *33*(1), 1–19.

Patil, A., Huard, D., & Fonnesbeck, C. (2010). PYMC: Bayesian stochastic modelling in Python. *Journal of Statistical Software*, *35*(4), 1.

Polack, P.-O., Friedman, J., & Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nature Neuroscience*, *16*(9), 1331–1339.

Poo, C., & Isaacson, J. S. (2009). Odor representations in olfactory cortex: "Sparse" coding, global inhibition, and oscillations. *Neuron*, *62*(6), 850–861.

Puggioni, P. (2015). *Input-output transformations in the awake mouse brain using whole-cell recordings and probabilistic analysis*. Ph.D. diss., University of Edinburgh.

Renart, A., Rocha, J. D. L., Bartho, P., Hollender, L., Parga, N., Reyes, N., & Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science*, *327*(5965), 587–590.

Rice, S. O. (1954). Mathematical analysis of random noise. In N. Wax (Ed.), *Selected papers on noise and random processes*. New York: Dover.

Roth, A., & van Rossum, M. (2009). Modeling synapses. In E. D. Schutter (Ed.), *Computational modeling methods for neuroscientists*. Cambridge, MA: MIT Press.

Roxin, A., Brunel, N., Hansel, D., Mongillo, G. J., & Vreeswijk, C. V. (2011). On the distribution of firing rates in networks of cortical neurons. *J. Neurosci.*, *31*(45), 16217–16226.

Rudolph, M., Piwkowska, Z., Badoual, M., Bal, T., & Destexhe, A. (2004). A method to estimate synaptic conductances from membrane potential fluctuations. *J. Neurophysiol.*, *91*(6), 2884–2896.

Sayer, R. J., Friedlander, M. J., & Redman, S. J. (1990). The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice. *J. Neurosci.*, *10*(3), 826–836.

Schiemann, J., Puggioni, P., Dacre, J., Pelko, M., Domanski, A., van Rossum, M. C. W., & Duguid, I. (2015). Cellular mechanisms underlying behavioral state-dependent bidirectional modulation of motor cortex output. *Cell Rep.*, *11*(8), 1319–1330.

Smith, M. A., & Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *Journal of Neuroscience*, *28*(48), 12591–12603.

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLOS. Biol.*, *3*(3), e68.

Spiegelhalter, D., Best, N., Carlin, B. J., & Linde, A. V. D. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639.

Szapiro, G., & Barbour, B. (2007). Multiple climbing fibers signal to molecular layer interneurons exclusively via glutamate spillover. *Nat. Neurosci.*, *10*(6), 735–742.

Teramae, J.-N., Tsubo, Y., & Fukai, T. (2012). Optimal spike-based communication in excitable networks with strong-sparse and weak-dense links. *Sci. Rep.*, *2*, 485.

Wehr, M., & Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, *426*(6965), 442–446.

Williams, S. R., & Mitchell, S. J. (2008). Direct measurement of somatic voltage clamp errors in central neurons. *Nat. Neurosci.*, *11*(7), 790–798.

Witter, L., Rudolph, S., Pressler, R. T., Lahlaf, S. I., & Regehr, W. G. (2016). Purkinje cell collaterals enable output signals from the cerebellar cortex to feed back to Purkinje cells and interneurons. *Neuron*, *91*(2), 312–319.