RESEARCH

CrossMark

Effect of Neuromodulation of Short-term Plasticity on Information Processing in Hippocampal Interneuron Synapses

**Elham Bayat Mokhtari[1] · J. Josh Lawrence[2] ·
Emily F. Stone[1]**

Abstract Neurons in a micro-circuit connected by chemical synapses can have their connectivity affected by the prior activity of the cells. The number of synapses available for releasing neurotransmitter can be decreased by repetitive activation through depletion of readily releasable neurotransmitter (NT), or increased through facilitation, where the probability of release of NT is increased by prior activation. These competing effects can create a complicated and subtle range of time-dependent connectivity. Here we investigate the probabilistic properties of facilitation and depression (FD) for a presynaptic neuron that is receiving a Poisson spike train of input. We use a model of FD that is parameterized with experimental data from a hippocampal basket cell and pyramidal cell connection, for fixed frequency input spikes at frequencies in the range of theta (3–8 Hz) and gamma (20–100 Hz) oscillations. Hence our results will apply to micro-circuits in the hippocampus that are responsible for the interaction of theta and gamma rhythms associated with learning and memory. A control situation is compared with one in which a pharmaceutical neuromodulator (muscarine) is employed. We apply standard information-theoretic measures such as entropy and mutual information, and find a closed form approximate expression for the probability distribution of release probability. We also use techniques that measure the dependence of the response on the exact history of stimulation the synapse has received, which uncovers some unexpected differences between control and muscarine-added cases.

✉ E.F. Stone
stone@mso.umt.edu

E. Bayat Mokhtari
Elham.bayatmokhtari@umontana.edu

J.J. Lawrence
john.lawrence@ttuhsc.edu

[1] Department of Mathematical Sciences, The University of Montana, Missoula, USA

[2] Department of Pharmacology and Neuroscience, Texas Tech University Health Sciences Center, Lubbock, USA

Keywords Short-term synaptic plasticity · Mutual information · Cholinergic modulation · Hippocampal GABAergic synapses

Abbreviations

BC	: basket cell
CA1	: Cornu Ammonis, earlier name for hippocampus
FD	: facilitation and depression
IPSC	: inhibitory postsynaptic current
ISI	: inter-spike interval
KL	: Kozachenko and Leonenko
KSG	: Kraskov, Stögbauer, and Grassberger
mAChR	: muscarinic acetylcholine receptors
MCMC	: Monte Carlo Markov Chain
MI	: Mutual Information
NT	: neurotransmitter
PSR	: postsynaptic response
PV	: parvalbumin-positive

1 Introduction

Neuronal activity can have profound effects on functional connectivity at the level of synapses. Through repetitive activation, the strength, or efficacy, of synaptic release of neurotransmitter (NT) can be decreased through depletion, or increased through facilitation. Both of these competing processes involve intracellular calcium and may occur within a single synapse. Different time scales of facilitation and depression enable temporally complex functional connectivity. Here we investigate the probabilistic properties of facilitation and depression (FD) for a presynaptic neuron that is receiving repetitive in vivo like Poisson spike train of input, e.g. the inter-spike interval follows an exponential distribution. We use a model of FD that was parameterized with experimental data from dual whole-cell recordings from a presynaptic parvalbumin-positive (PV) basket cell (BC) connected to a postsynaptic CA1 (Cornu Ammonis 1 subregion) pyramidal cell, for family of fixed frequency input spikes into the presynaptic PV BC [1]. Our results will thus apply to micro-circuits in the hippocampus that participate in the generation of theta and gamma rhythms associated with learning and memory.

The role of synaptic plasticity and computation has been analyzed and reported on in numerous papers over the past 30 years. A review of feed-forward synaptic mechanisms and their implications can be found in [2]. In this paper Abbott and Regehr state "The potential computational power of synapses is large because their basic signal transmission properties can be affected by the history of presynaptic and postsynaptic firing in so many different ways." They also outline the basic function of a synapse as a signal filter as follows: Synapses with an initial low probability of release act as high pass filters through facilitation, while synapses with an initially high probability of release exhibit depression and subsequently serve as low pass filters. Intermediate cases, in which the synapse can act as a band-pass filter, exist. Furthermore, short-term plasticity can influence the availability of postsynaptic receptors

to bind neurotransmitter. For example, reducing neurotransmitter release probability can reduce postsynaptic receptor desensitization, effectively increasing the efficacy of synaptic transmission during high-frequency stimulation.

Other functional roles of short-term synaptic plasticity can be found in [3]. There they discuss signal processing capabilities in the auditory system, visual processing in the retina, olfactory processing and even electrosensory processing in weakly electric fish. In the hippocampus facilitation of excitatory synapses combined with depression of inhibitory synapses can amplify high-frequency inputs, which selects for the high-frequency output seen in place cells. Inhibitory interneuron synapses in the hippocampus, such as the kind studied here, show a dynamic range of reaction when recruited via different pathways. In a related synopsis [4], the plasticity of inhibitory connections is studied in the context of spike timing-dependent plasticity and network function. Taking network function with plasticity a step further, [5] analyzes simple neural networks in cortex with Hebbian-type learning rules affected by plasticity, modeling memory storage in networks with inhibitory synapses. Results are largely general and any conclusions drawn are not immediately applicable to actual brain function. Furthermore, the distinction between long- and short-term plasticity in this analysis is not made clear.

Synaptic depression as a regulator of synaptic transmission is discussed thoroughly in [6]. Here it is supposed that synaptic depression can keep synaptic efficacy constant relative to changes in probability of release. Also, depressing synapses can serve as a gain control in the face of rapidly firing presynaptic cells. It is also established that synaptic depression favors temporal encoding of information because the steady state is reached quickly and maintains no information as regards the absolute firing rate. However, these synapses are sensitive to a sudden change in firing rate, detecting rate changes in low- and high-frequency inputs with similar sensitivity.

In cortex [7] draws the distinction between what they call synaptic vs. structural plasticity, focusing on structural plasticity, which directly effects the synaptic weights in a neural network model. They find that certain characteristics arise directly from the interaction of structural plasticity and synaptic plasticity rules. These characteristics in turn create a variety of stable synaptic weight distributions which could support information storage mechanisms. Neuromodulation can further alter the time-dependent characteristics of a synapse. Acting on the presynaptic side, many neuromodulators reduce the probability of release, protecting the synapse from depletion and therefore extending the duration or frequency sensitivity of the synapse overall.

The calculation of various kinds of measures of information transfer at the synapse level has been explored in many papers. For instance, in [8] information transmission is studied through a master equation based stochastic model of presynaptic release of vesicles (which depends on intracellular calcium concentration), combined with a low dimensional model of membrane charging at the postsynaptic side. The model itself has an advantage over models of average quantities, in that it captures fluctuations of the dynamic variables. In [9] they consider synaptic transmission in neocortex, and compute the amount of information conveyed by a single response to a specific sequence of spike stimulation, as it is effected by short-term synaptic plasticity. They determine that for any given dynamic synapse there is an optimal frequency of input stimulation for which the information transfer is maximal. A mathematical model

of the calyx of Held was used in [10] to study synaptic depression due to repeated stimulation seen in vitro. They quantify the information contained in the postsynaptic current amplitude about preceding inter-spike intervals using a mutual information calculation. Both [9] and [10] have directly inspired the work we present in this paper. We add here that Tsodyks and Markram also included such dynamic synapses in a numerical network model [11].

Other work focuses on mutual information in spike trains, measuring the transfer of information between input spikes and output spike response, with the goal of addressing reliability in rate coding [12]. Another study [13] used in vivo spike trains recorded from monkeys to derive a synapse model with activity-dependent depression, showing a decrease in redundancy from input to output spike train. Also concerning rate coding, in [14] a functional measure called the Synaptic Information Efficacy (SIE) is devised to measure mutual information between input and output spike trains in a noisy environment.

In [15] we parameterize a simple model of presynaptic plasticity from work by Lee and colleagues [16] with experimental data from cholinergic neuromodulation of GABAergic transmission in the hippocampus. The model is based upon a calcium-dependent enhancement of probability of release and recovery of signaling resources. (For a review of these mechanisms see [17].) It is one of a long sequence of models developed from 1998 to the present, with notable contributions by Markram, [18], and Dittman, Kreitzer and Regehr [19]. The latter is a good exposition of the model as it pertains to various types of short-term plasticity seen in the central nervous system, and the underlying dependence of the plasticity is based on physiologically relevant dynamics of calcium influx and decay within the presynaptic terminal. In our work, we use the Lee model to create a two dimensional discrete dynamical system in variables for calcium concentration in the presynaptic area and the fraction of sites that are ready to release neurotransmitter into the synaptic cleft. The map is parameterized with experimental results from paired whole-cell recordings at CA1 PV basket cell-pyramidal cell synapses. The PV basket cell was presented with current pulses at fixed frequencies, resulting in trains of presynaptic action potentials evoking GABA transmission onto the postsynaptic pyramidal cell. Synaptic transmission is manifested as trains of GABAA receptor-mediated inhibitory postsynaptic currents (IPSCs). Experiments were run in control and with muscarine added, which acts upon presynaptic muscarinic acetylcholine receptors (mAChRs) to cause a reduction in the observed IPSCs.

Various parameterizations and hidden parameter dependencies were investigated using Monte Carlo Markov Chain (MCMC) parameter estimation techniques. This analysis reveals that a frequency dependence of cholinergic modulation requires both calcium-dependent recovery from depression and mAChR-induced inhibition of presynaptic calcium channels. A reduction in calcium entry into the presynaptic terminal in the kinetic model accounted for the frequency-dependent effects of the mAChR activation.

We now use our model to investigate the information processing properties of this synapse, in control and neuromodulation conditions. We possess these two parameterizations; one from experiments in control conditions, the other from synapses undergoing activation of mAChR receptors by muscarine. Hence we can analyze the

effect of this cholinergic modulation on information processing at the synapse level. Because network oscillations associated with learning and memory feature PC basket cells firing in the gamma range [20], we examine a range of frequencies from near zero to 100 Hz in what follows.

As mentioned previously, this analysis is motivated and guided by the work of Markram et al. in [9]. In this paper it is analyzed how much information as regards previous inter-spike intervals is contained in the size of a single response of a dynamic synapses, i.e. a synapse which is affected by the history of presynaptic activity. They derive expressions for the optimal frequency of the input Poisson spike train in terms of coding temporal information in depressing and facilitating neocortical synapses. It was found that depressing synapses are optimal in this sense at low firing rates (0.5–5 Hz) while facilitating synapses are optimal at higher rates (9–70 Hz). This is not surprising, given that the average postsynaptic response is larger at low frequencies in depressing synapses, and larger at higher frequencies in facilitating synapses. The more interesting result is the presence of a peak in the information transferred at a set frequency, an actual local maximum that is particular to each specific synapse. We examine how cholinergic neuromodulation (the addition of muscarine) effects this result in what follows, and an alternative way to measure the information transferred from a sequence of preceding inter-spike intervals.

2 FD Model

Many mathematical models have been developed to describe short-term plasticity over the past 20 years [18, 19, 21, 22]. More flexible models include intracellular calcium dynamics [23], because probability of release is thought to be directly dependent upon calcium concentration in the presynaptic terminal. Recovery from synaptic depression is also thought to be accelerated by the presence of calcium, thereby unifying the underlying molecular mechanisms of facilitation and depression [19, 24–26].

In our model we take the probability of release (P_{rel}) to be the fraction of a pool of synapses that will release a vesicle upon the arrival of an action potential at the terminal. Following the work of Lee et al. [16], we postulate that P_{rel} increases monotonically as a function of calcium concentration in a sigmoidal fashion to asymptote at some P_{max}. The kinetics of the synaptotagmin-1 receptors that binds the incoming calcium suggests a Hill equation with coefficient 4 for this function. The half-height concentration value, K, and P_{max} are parameters determined from the data.

After releasing vesicles upon stimulation, some portion of the pool of synapses will not be able to release vesicles again if stimulated within some time interval, i.e. they are in a refractory state. This causes "depression"; a monotonic decay of the amplitude of the response upon repeated stimulation. The rate of recovery from the refractory state is thought to depend on the calcium concentration in the presynaptic terminal [24, 25, 27]. Following Lee et al., [16], we assume a simple monotonic dependence of rate of recovery on calcium concentration, a Hill equation with coefficient of 1, starting at some k_{min}, increasing to k_{max} asymptotically as the concentration increases, with a half height of K_r. Muscarine, binding with muscarinic acetylcholine presynaptic receptors (mAChR), is thought to cause inhibition of presynaptic
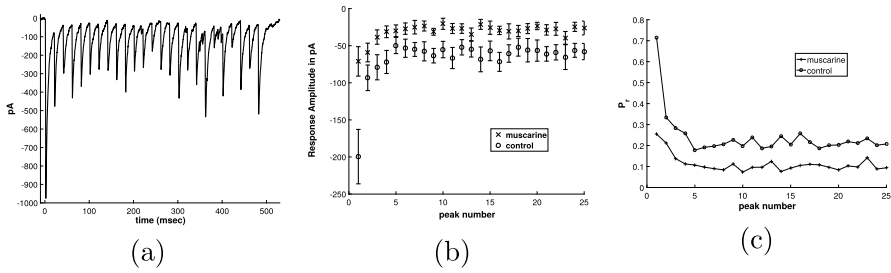
(a)   (b)   (c)

Fig. 1 (**a**) Response to pulse train stimulus from one cell in control conditions, in pA, picoamps. (**b**) Comparison of response in control and muscarine conditions, averaged over 7 cells, plotted with error bars at one standard deviation of the mean. The baseline for each pulse has been subtracted from each peak, to capture the *change* in the current upon stimulation. (**c**) Absolute value of the response in control and muscarine conditions, baseline removed as described in (**b**), averaged over 7 cells and normalized by Nq (N number of synapses, q quantal amplitude of single synapse) to have units of "probability of release", Pr

calcium channels, thereby decreasing the amount of calcium that floods the terminal when it receives an action potential [28].

An example of this process is seen in the set of experiments illustrated in Fig. 1. Here whole-cell recordings were performed from synaptically connected pairs of neurons in mouse hippocampal slices from PV-GFP mice [1]. The presynaptic neuron was a PV basket cell (BC) and the postsynaptic neuron was a CA1 pyramidal cell. Using short, 1–2 ms duration suprathreshold current steps to evoke action potentials in the PV BC from a resting potential of -60 mV and trains of 25 of action potentials are evoked at 5, 50, and 100 Hz from the presynaptic basket cell. The result in the postsynaptic neuron is the activation of $GABA_A$-mediated inhibitory postsynaptic currents (IPSCs). Upon repetitive stimulation, the amplitude of the synaptically evoked IPSC declines to a steady-state level. These experiments were conducted with 5, 50 and 100 Hz stimulation pulse trains, in order to test frequency-dependent short-term plasticity effects. We note that oscillations in neural networks in the "gamma" range are associated with learning and memory, for a review see [20]. Muscarine activates presynaptic metabotropic/muscarinic acetylcholine receptors (mAcHRs) which cause a reduction in the response overall, and subsequently the amount of depression in the train.

The peak of the measured postsynaptic IPSC is presumed to be proportional to the total number of synapses that receive stimulation N_{tot}, which are also ready to release (R_{rel}), e.g. $N_{tot}R_{rel}$, multiplied by the probability of release P_{rel}. That is, the peak IPSC $\sim N_{tot}R_{rel}P_{rel}$. P_{rel} and R_{rel} are both fractions of the total, and thus range between 0 and 1. Without loss of generality, we consider the peak IPSC to be proportional to $R_{rel}P_{rel}$.

The presynaptic calcium concentration itself, $[Ca]$, is assumed to follow first order decay kinetics to a base concentration, $[Ca]_{base}$. At this point we choose that $[Ca]_{base} = 0$, since locally (near the synaptotagmin-1 receptors) the concentration of calcium will be quite low in the absence of an action potential. The evolution equation for $[Ca]$ then is simply $\tau_{ca}\frac{d[Ca]}{dt} = -[Ca]$ where τ_{ca} is the calcium decay time constant, measured in ms. Upon pulse stimulation, presynaptic voltage-gated calcium channels open, and the concentration of calcium at the terminal increases rapidly by

an amount δ (measured in μm) : $[Ca] \to [Ca] + \delta$ at the time of the pulse. Note that calcium build-up is possible over a train of pulses if the decay time is long enough relative to the inter-pulse interval.

We non-dimensionalize the calcium concentration, rescaling it by the value of δ in the control case, δ_c, and defining $C = \frac{[Ca]}{\delta_c}$. After a stimulus occurring at a time $t = 0$, which results in an increase in C by an amount $\Delta = \frac{\delta}{\delta_c}$, the concentration of calcium is $C = C_0 e^{-t/\tau_{ca}} + \Delta$. In the control case this further simplifies to $C = C_0 e^{-t/\tau_{ca}} + 1$.

As mentioned previously, the probability of release P_{rel} and the rate of recovery, k_{recov}, depend monotonically on the instantaneous calcium concentration via Hill equations with coefficients of 4 and 1, respectively; e.g. $P_{rel} = P_{max} \frac{C^4}{C^4 + K^4}$, and $k_{recov} = k_{min} + \Delta k \frac{C}{C + K_r}$, where $\Delta k = k_{max} - k_{min}$. The variable R_{rel} is governed by the ordinary differential equation $\frac{dR_{rel}}{dt} = k_{recov}(1 - R_{rel})$, which can be solved exactly for $R_{rel}(t)$. $R_{rel}(t) = 1 - (1 - R_0)\left(\frac{C_0 e^{-t/\tau_{ca}} + K_r}{K_r + C_0}\right)^{\Delta k} e^{-k_{min}t}$. P_{rel} is also a function of time as it follows the concentration of calcium after a stimulus.

We are interested in capturing the peak value of the IPSC (or Pr), so we construct a discrete dynamical system (or "map") that describes $P_{rel} R_{rel}$ upon repetitive stimulation. Given an inter-spike interval of T, the calcium concentration after a stimulus is $C(T) + \Delta$, and the peak IPSC is proportional to $P_{rel}(T) R_{rel}(T)$, which depends upon C. After the release, R_{rel} is reduced to the fraction of synapses that fired, e.g. $R_{rel} \to R_{rel} - P_{rel} R_{rel} = R_{rel}(1 - P_{rel})$. This value is used as the initial condition in the solution to the ODE for $R_{rel}(t)$. A two dimensional map (in C and R) from one peak value to the next is thus constructed. To simplify the formulas we let $P = P_{rel}$ and $R = R_{rel}$. The map is

$$C_{n+1} = C_n e^{-T/\tau_{ca}} + \Delta,$$

$$P_{n+1} = P_{max} \frac{C_{n+1}^4}{C_{n+1}^4 + K^4},$$

$$R_{n+1} = 1 - \left(1 - (1 - P_n)R_n\right)\left(\frac{C_n e^{-T/\tau_{ca}} + K_r}{K_r + C_n}\right)^{\Delta k} e^{-k_{min}T}.$$

The peak value upon the nth stimulus is $Pr_n = R_n P_n$, where R_n is the value of the reserve pool before the release reduces it to the fraction $(1 - P_n)$.

2.1 Parameter Estimation

The parameter values for the model are summarized in Table 1. The rescaled data presented in a previous section were fitted to the map using the Matlab package lsqnonlin, and with MCMC techniques ([29, 30]). The value of P_{max} was determined by variance-mean analysis, and it is 0.85 for the control data and 0.27 for the muscarine data. The common fitted parameter values for both data sets are shown in Table 2.

The control data set was assigned $\Delta = 1$ which can be done without loss of generality if the concentration of calcium is scaled by that amount, and the muscarine data set has the fitted value of $\Delta = 0.17$. From this result it is clear that the size of

Table 1 Parameters in the map

Parameter	Description
Δ	Increase in the amount of calcium relative to that seen under control conditions
P_{max}	Maximum probability of release
K	Half calcium concentration value for probability of release function
k_{min}	Minimum rate of recovery of synapses
k_{max}	Maximum rate of recovery of synapses
K_r	Half calcium concentration value for rate of recovery function
τ_{ca}	Decay constant for calcium

Table 2 Parameter values

Parameter	Fitted value
K	0.2
k_{min}	0.0017 1/ms
k_{max}	0.0517 1/ms
K_r	0.1
τ_{ca}	1.5 ms
P_{max}	0.85
Δ	1
μ	0.5
σ	0.1

the spike in calcium during a stimulation event must be much reduced to fit the data from the muscarine experiments. This is in accordance with the idea that mAChR activation reduces calcium ion influx at the terminal.

2.2 Discussion of the Model

It is common in the experimental literature to classify a synapse as being "depressing" or "facilitating", depending upon its response to a pulse train at some relevant frequency. Simple models have been built that create each effect individually. The model here (developed originally by Lee [16], recall) combines both mechanisms through the introduction of the calcium variable. Depending upon the parameters, both facilitation and depression and a mixture of the two can be represented, as is shown in Lee [16]. Note that facilitation is built into this model through the calcium-dependent P value and rate of recovery.

This interplay of the presynaptic probability of release and the rate of the recovery creates a nonlinear filter of an incoming stimulus train. Adding muscarine modifies the properties of this filter. To investigate this idea, we consider the distribution of values of Pr created by exponentially distributed random ISIs for varying rates λ, or mean ISI, denoted $\langle T \rangle = 1/\lambda$. Doing so explores the filtering properties of the

synapse when presented with a Poisson process spike train. To develop our intuition, we first present results from numerical studies to determine of the effect of varying frequency of the pulse train, or mean rate, on the muscarine and control cases. Then we create analytic expressions for the distribution of calcium and the *Pr* values, which are corroborated by the numerical results. Finally, the information processing properties of the synapse in the control and the muscarine cases at physiological frequencies are compared.

3 Numerical Study of the Distributions in *Pr*

To create approximations to the distribution of *Pr* values we computed 2^{14} samples from the stochastic map, after discarding a brief initial transient. The values, ranging between 0 and 1, were placed into evenly spaced bins. The histograms, normalized to be frequency distributions, were computed for a range of mean frequencies (or rates) in the theta range, gamma range and higher (non-physiological, for comparison). The parameters used in the following simulations are from fitting the model to the control and muscarine data set (see Table 2).

3.1 Frequency Distributions of *Pr* in Control and Muscarine Cases

There is a similar evolution of the frequency distribution for *Pr* as the rate of the incoming Poisson spike train is increased, in both cases. For very small rates (between almost 0 and 1 Hz) the distribution is peaked near $P_{\max}$. The peak is necessarily skewed to the left, as it is restricted to the support $[0, 1]$ and the exponential distribution of ISIs always contributes some very small values, which generate lower *Pr* values. For very large rates (non-physical, 200 Hz and larger) the distribution is peaked near 0, reflecting the fact that the synapse does not have time to recover between spikes. Again the peak is skewed, this time to the right, necessarily. In between the two extremes the distribution must spread out between 0 and 1, and it does so in a very particular way.

   As the rate is increased from 0.5 to 10 Hz the distribution spreads quickly out over the entire interval; see Fig. 2. This might be expected, since the range of frequencies present in the exponential distribution will cause a wide range of *Pr* responses at these lower mean firing rate values. First the skewness to the left becomes more pronounced, then the left half of the distribution grows up to match the right, becoming pretty much flat at around 1.8 Hz. It then begins to drop on the right side, becoming almost triangular around 3 Hz. From there the peak on the left sharpens, while maintaining a shoulder for the very lowest values of *Pr*. Note that this covers the range of rates generally thought to be theta frequency. Here the synapse could be said to be the most sensitive and allow for widely varying responses.

   For rates larger than 10 Hz the peak on the left sharpens as the rate is increased, but very slowly. The shoulder on the left of the peak remains in place. In contrast with the theta range, the gamma range is marked by a nearly steady response. Having these two distinct "tunings" of the synapse at physiological frequencies is quite interesting. For even larger frequencies the shoulder on the left grows up to meet the peak (data
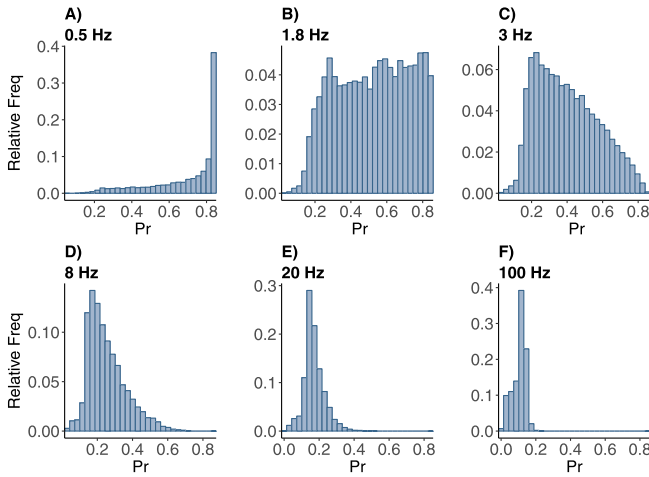
Fig. 2 Frequency distributions of normalized postsynaptic response *Pr* with control parameter set under stimulation at (**A**) 0.5, (**B**) 1.8, (**C**) 3, (**D**) 8, (**E**) 20, and (**F**) 100 Hz Poisson spike train. The horizontal axis is for the *Pr* values, the vertical axis is for the relative frequency of the *Pr* values

not shown). The peak near 0.1 persists until the rate is greater than 300 Hz, after which it is subsumed into the peak near zero. The synaptic dynamics thus creates a small probability response that is stable over a wide range of frequencies.

As mentioned previously, for the control case the influx of calcium (Δ) was set to unity, because the variable in the map for calcium was rescaled by this amount. In order to fit the muscarine data a much smaller value of (Δ) was needed (0.17). This is consistent with the hypothesis that muscarine shuts down the influx of calcium ions to the presynaptic cell. This reduces the size of the response, but it was also shown to reduce the relative amount of depression seen with repeated spikes, at least at intermediate frequencies around gamma. This could have important implications for the effect of neuromodulation at these frequencies. How is this manifested in the *Pr* distributions? In Fig. 3 are shown the *Pr* frequency distributions for varying mean rate. The range on the *x*-axis is set to [0, 0.3] because $P_{max} = 0.27$. With this rescaling the distributions look much like those in the control case: The distribution is skewed left with a peak near P_{max} for low frequencies, and shifts to have a peak in the middle of the range for intermediate frequencies while spreading out. The peak in the middle of the range gradually moves to the left as the frequency is increased further. In the control case the peak is more triangular skewed right than in the muscarine case at low frequencies, at gamma frequencies the control distribution has a shoulder on the left, while the muscarine case distribution is more symmetrical. Therefore the muscarine treated synapse focuses the response in a narrow range around a small *Pr* for all but the lowest frequencies. At high frequencies the peak of the muscarine distribution does not go to zero as fast as the control distribution does, a somewhat non-intuitive result, though we must point out that this range of frequencies is not physiological.

In both cases the dynamics of the map create a low pass filter, which is the hallmark of a depressing synapse. The mid-range peak makes it something more complex
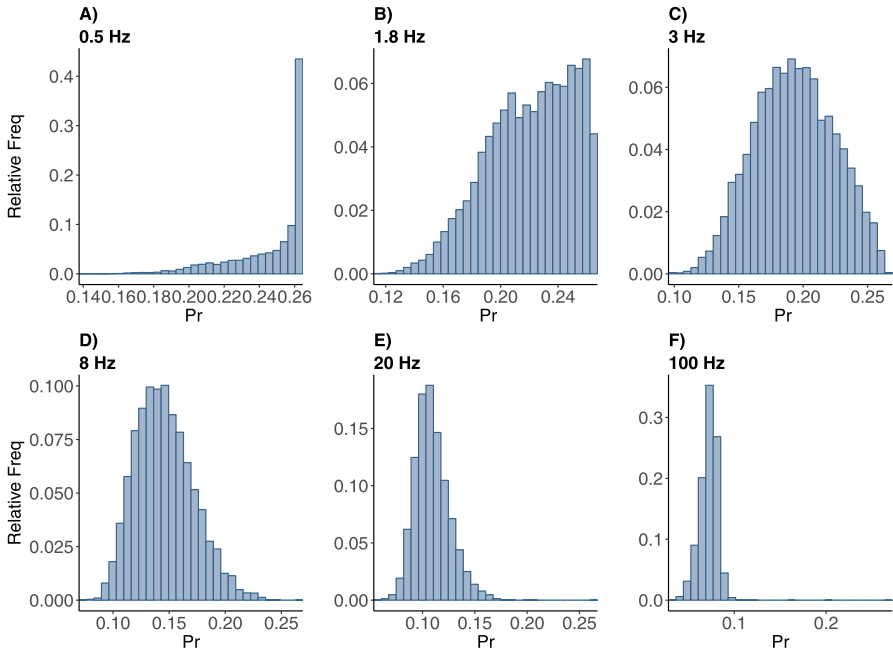
Fig. 3 Frequency distributions of normalized postsynaptic response *Pr* with muscarine parameter set under stimulation at (**A**) 0.5, (**B**) 1.8, (**C**) 3, (**D**) 8, (**E**) 20, and (**F**) 100 Hz Poisson spike train. The horizontal axis is for the *Pr* values, the vertical axis is for the relative frequency of the *Pr* values
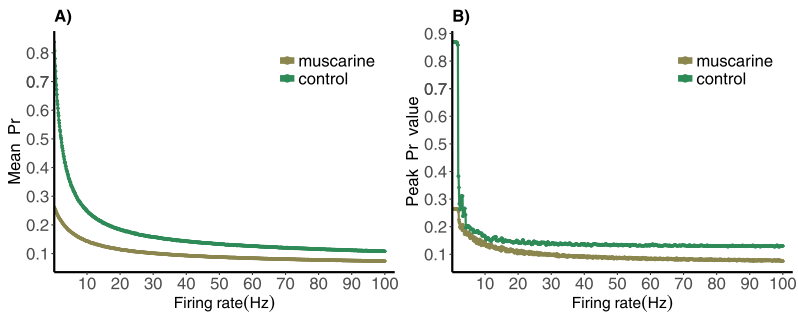

Fig. 4 (**A**) Mean and (**B**) peak of normalized postsynaptic response values for control and muscarine parameter sets stimulated by Poisson spike trains with mean frequencies ranging from 0.1 to 100 Hz

than a linear low pass filter, as it has a typical response size (the peak or the mean value) for each frequency. However, it shows an interesting uniformity: the peak (or most likely) response is fairly insensitive to changes in frequency over the physiological range of 3–70 Hz.

We can then compare the mean and the peak of the distributions as the input frequency is varied. See Fig. 4. As expected, the mean and the peak of the muscarine distributions are much lower than that of the control, except in the case of the peak
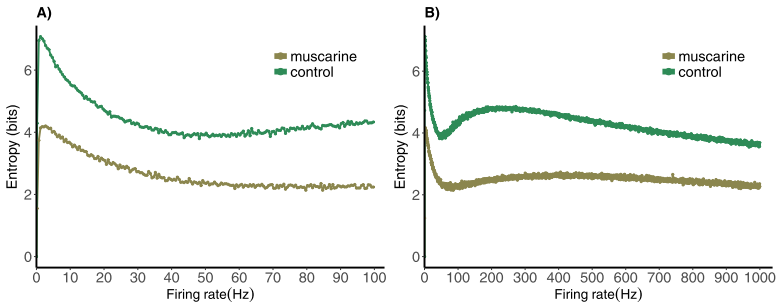
Fig. 5 Measure of information transmission for deterministic model of synapses stimulated by Poisson spike train with mean frequencies ranging from (**A**) 0.1 to 100 Hz and (**B**) 0.1 to 1000 Hz for control and muscarine parameter sets

within the theta range. This could mean that even under depression caused by pharmaceutical applications the synapse response remains consistent, a kind of built-in stability to theta frequencies. We also note that the amount of depression relative to the initial response of the synapse is less for the muscarine case than the control: the mean ranges from over 0.8 to between 0.1 and 0.2 for the control case, a change of about 0.6-0.7 overall, and from 0.3 to near 0.1 in the muscarine case, a much smaller change of about 0.2. This confirms the assertion in [1].

3.2 Entropy of *Pr* Distribution vs. Mean Rate

The entropy of a distribution is a convenient scalar measure for comparing the overall structure in distributions as a parameter is varied. We therefore compute the entropy of the *Pr* distributions in the control and muscarine case for varying mean rate and compare it with the distributions themselves. Finally, we draw some physiological conclusions.

The entropy of a distribution, $p(x)$, of a random variable, X, is defined to be

$$H(X) = - \sum_{x \in \mathcal{X}} p(X = x) \log_2 p(X = x). \tag{1}$$

It measures the number of states the variable can assume, along with the probability of each state. Note that for continuous random variables it is necessarily dependent upon the exact partition of the distribution into a histogram of values that can be summed. In what follows we keep this partition constant across different cases, comparing the entropies of the resulting histograms. All other things being equal, the entropy of a distribution with more "spread" is higher than one that is more focused on a single value. If a random variable is tightly constrained, its distribution will have a lower entropy, i.e. it is much more certain what state the variable will be in for any given sample.

In Fig. 5 we plot the entropy as a function of the mean rate of the driving exponential distribution for both control and muscarine parameter sets. Figure 5(b) shows a range of frequencies from near zero to 1000 Hz to see the limit for large frequencies.

Figure 5(a) shows a zoom in on the physiological range of frequencies between zero and 100 Hz.

Studying the entropy vs. frequency from 0.1 to 100 Hz, we see a local maximum for low values of frequency (at approximately 4 Hz). In this frequency range the distribution spreads out between a peak at high Pr values to a peak at lower Pr values. At these frequencies there would be a large variability in the size of the response, which could be a useful characteristic in a communication channel. However, if the criterion is a stable level of connectivity, lower entropy is desirable. For larger frequencies (not physiological) the entropy increases again to a local maximum near 260 Hz, after which it decays as the distribution narrows near almost zero Pr values. For the muscarine case the second local maximum is less pronounced and occurs near 388 Hz. The maximum in both cases is the result of the distribution shifting from one peaked near $Pr > 0$ to one peaked at zero skewed to the right. This occurs through the widening out of the peak, causing the increase in entropy.

Note that the size of response (Pr) is a measure of the "strength" of the synaptic connection created by the "pool" of synapses. The advantage of having a peaked distribution with low entropy over a range of frequencies is that a stable connection is created, even when presented with a stochastic signal of the Poisson type. Alternatively, when the distribution is switching from a peak at high Pr values to low, as the mean rate is increased, the entropy is at a maximum and there is a greater range of coupling strengths. In that case the exact value of the strength of the connection will depend on the past synaptic signaling history.

4 Stochastic Recurrence Relationships for Calcium and Pr

The map for C, P and R driven by a Poisson spike train is a stochastic recurrence relation with noise added in an exponent. We shall show that while it is not possible to create a closed form for the complete map, an approximation can be created that relies on the properties of the deterministic map for low input rates. Only the map for the calcium concentration allows for direct analysis, but it involves the introduction of a random variable for Δ. We provide an outline of the work required to create this, for completeness.

4.1 Calcium Concentration

Suppose an independent, identically distributed increase in the amount of calcium $\{\Delta_n\}_{n \geq 1}$ occurs at times $\{t_n\}_{n \geq 1}$ and we wish to find the distribution of the concentration of calcium accumulated in the presynaptic terminal following this supposition.

The concentration is governed by a random recurrence equation given by

$$C_n = A_n C_{n-1} + \Delta_n, \quad n \geq 1, \tag{2}$$

where $A_n = e^{-T_n/\tau_{ca}}$, and the waiting times $T_1 = t_1$, $T_n = t_n - t_{n-1}$, $n \geq 2$, are independent and identically distributed, i.i.d., making $\{T_n\}$ a renewal process. Moreover, (A_n, Δ_n) are assumed to be i.i.d. vectors with initial condition $C(0) = C_0$. C_0 is a base concentration of calcium which is considered to be zero in the absence of a
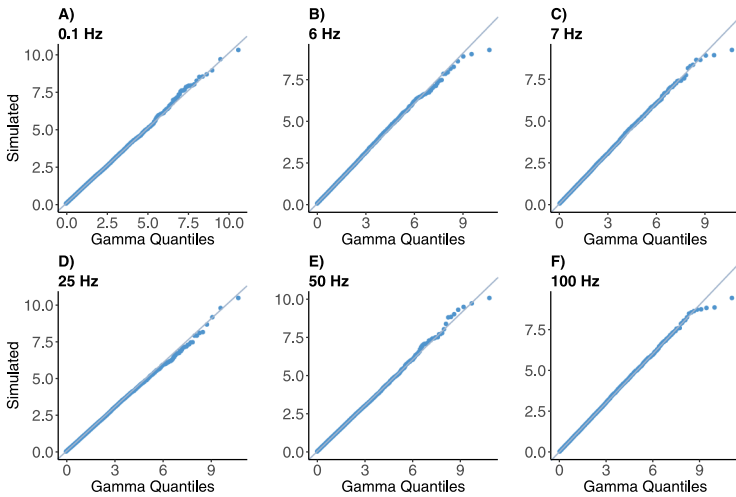
Fig. 6 Fitted percentiles of gamma-distributed data to simulated calcium concentration data from physical model for frequencies (**A**) 0.1, (**B**) 6, (**C**) 7, (**D**) 25, (**E**) 50, (**F**) 100 Hz. The line of identity is also plotted. The alignment of the line and quantile plot indicates the adequacy of the fit of the gamma model for calcium concentration

stimulus. After some manipulation it can be shown that the concentration of calcium follows a gamma distribution with a shape parameter $\lambda \tau_{\mathrm{ca}} + 1$ and a scale parameter 1. Thus,

$$f_C(c) = \frac{1}{\Gamma(\lambda \tau_{\mathrm{ca}} + 1)} c^{\lambda \tau_{\mathrm{ca}}} e^{-c}, \quad \lambda \tau_{\mathrm{ca}} + 1 > 0,$$

where c is a realization of random variable C. The distribution of calcium concentration C has mean around $\lambda \tau_{\mathrm{ca}} + 1$. The coefficient of variation of C is $1/(\sqrt{\lambda \tau_{\mathrm{ca}} + 1})$, and hence, the shape of distribution depends on the value of $\lambda \tau_{\mathrm{ca}} + 1$. The details of the derivation can be found in Appendix B.

We compare this result with the distribution obtained by numerical simulation of the recurrence relation for C by creating quantile plots. Figure 6 displays quantile plots for the map with input frequencies 0.1, 6, 7, 25, 50, 100 Hz, with the theoretical quantiles based upon the gamma distribution. This type of graphical display shows whether the simulated data can reasonably be described by a gamma distribution. Plots show adherence to a linear relationship between the observed and theoretical quantiles, confirming our analytic results.

4.2 Distribution of *Pr* for Small τ_{ca} and Large *T*

We conclude from the previous subsection that the random variable describing the calcium concentration does have a parametric distribution. However, this is not the case for the variable R due to the complexity of the map, and so a closed form for the distribution of $Pr = PR$ is not possible. However, we can understand it partially by considering the mechanisms involved. We can also use some information from the deterministic map. The map has a single attracting fixed point, and the collapse to

Fig. 7 Fixed point normalized postsynaptic response values for the deterministic map stimulated by Poisson spike train with mean frequencies ranging from 0.1 to 100 Hz, compared with the mean of the normalized postsynaptic response values for the computed distribution
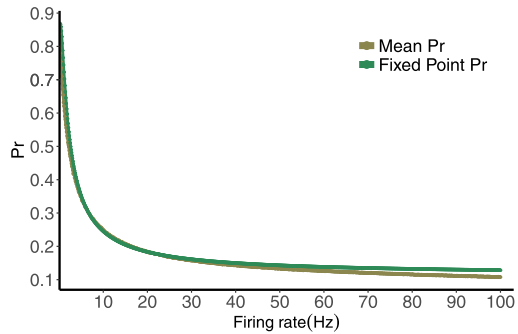
this fixed point from physiological initial conditions is very rapid [15]. The value of the fixed point depends on the frequency, with a smaller value for larger frequency in general. In Fig. 7 we plot the expression for the fixed point of the deterministic map vs. rate, along with the mean of the distribution of *Pr* for varying frequencies. The values decrease with increasing frequency, as expected, and are remarkably close.

This motivates the idea that if the *Pr* value was directly determined by the fixed point value for the ISI value preceding it, we would be able to convert the distribution of the ISIs into that of the *Pr*s by using composition rules for distributions of random variables. We examine this when the calcium decay time (τ_{ca}) is notably smaller than the inter-spike interval (T). With this approximation C, P have time in between pulses to decay to their steady-state value before another pulse. This means that the fixed point value for a rate given by $1/T$ where T is the preceding inter-spike interval is more likely to give a good estimate of the actual value or $Pr = PR$.

It was shown in [15] that in this case $\overline{C} \to \Delta$ as T increases and hence $\overline{P} \to P_{\max}$. Therefore, the fixed point $\overline{R}$ is then

$$\overline{R} = \frac{1 - e^{-k_{\min}T}}{1 - (1 - P_{\max})e^{-k_{\min}T}}.$$

From this we can compute the probability density function of $\overline{R}$, given an Exponential distribution of the variable T. Details of this calculation are provided in Appendix C. For ease of notation we let $X = \overline{R}$ and $Y = \overline{PR}$ be random variables, then an analytic expression for probability density function (PDF) of X exists and is given by

$$f(x|\lambda, a, k_{\min}) = \frac{\lambda(1-a)}{k_{\min}}(1-x)^{-(1-\lambda/k_{\min})}(1-ax)^{-(1+\lambda/k_{\min})}, \qquad (3)$$

where $a = 1 - P_{\max}$, $\lambda > 0$ is the mean Poisson rate and $k_{\min} > 0$ is the baseline recovery rate. The distribution is supported on the interval [0, 1].

From Eq. (3), the mean or expected value of random variable $X = \overline{R}$ is given by

$$E(X) = (1-a)\lambda \left(\frac{1}{\lambda(1-a)} - \frac{{}_2F_1(1, \frac{k_{\min}+\lambda}{k_{\min}}; 2 + \frac{\lambda}{k_{\min}}; a)}{k_{\min} + \lambda} \right), \qquad (4)$$

where ${}_2F_1(1, \frac{k_{\min}+\lambda}{k_{\min}}; 2 + \frac{\lambda}{k_{\min}}; a)$ is the hypergeometric function.

Fig. 8 Probability density function of the stochastic fixed point of normalized postsynaptic response (or $\overline{Pr} = \overline{PR}$) for varying input ISI in ms. Parameters $k_{\min} = 0.0013$ and $p_{\max} = 0.85$, are from the control set
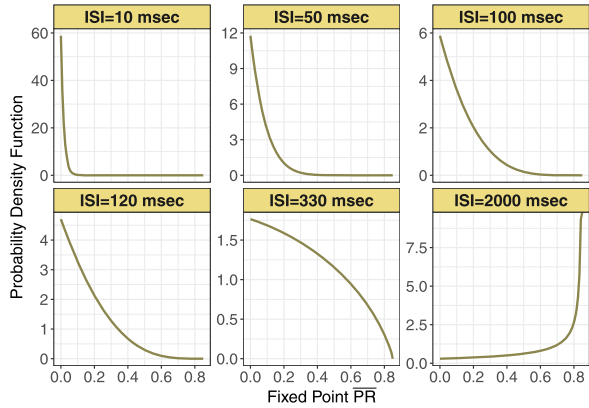
Similarly, we can compute the analytical expression of the probability density function of fixed point $Y = \overline{PR}$. We will refer to this in what follows as the *stochastic fixed point*. Hence, the probability density function for the stochastic fixed point is

$$f(y|\lambda, a, k_{\min}) = \frac{\lambda P_{\max}(1-a)}{k_{\min}}(P_{\max} - y)^{-(1-\lambda/k_{\min})}(P_{\max} - ay)^{-(1+\lambda/k_{\min})}. \quad (5)$$

This distribution is supported on the interval $[0, P_{\max}]$. Figure 8 shows this expression for different mean input inter-spike interval, in ms.

In Fig. 9 are histograms of Pr values obtained from the map with very small τ_{ca}, and other parameters from the control set, as in Fig. 8. The similarity between the two is evident. Apparently this approximation captures not only the mean value of the numerical distribution, but also the shape of the distribution and how it changes with varying input spike train rate.

Figure 10 compares these two distributions via quantile–quantile plots, for mean ISI of 10, 50, 100, 120, 330, and 2000 ms. In every QQ-plot the quantiles of all $\overline{PR}$ are plotted against the quantiles of all Pr values. If the values of the two different data sets have the same distribution, the points in the plot should form a straight line. From these plots it is clear that when the mean ISI is significantly larger than calcium decay time, the distribution of the stochastic fixed point is similar to that of Pr. However, for smaller mean ISI < 10 ms the approximation becomes less exact, as does the similarity between the two distributions.

5 The Stochastic Model of the Postsynaptic Response

To compute mutual information between the stimulus and the response of a synapse we must capture the variability in the postsynaptic response. A certain probability of release will generate a distribution of responses that has an analytical description given certain fundamental assumptions about stochastic processes. We provide a short overview of these here.
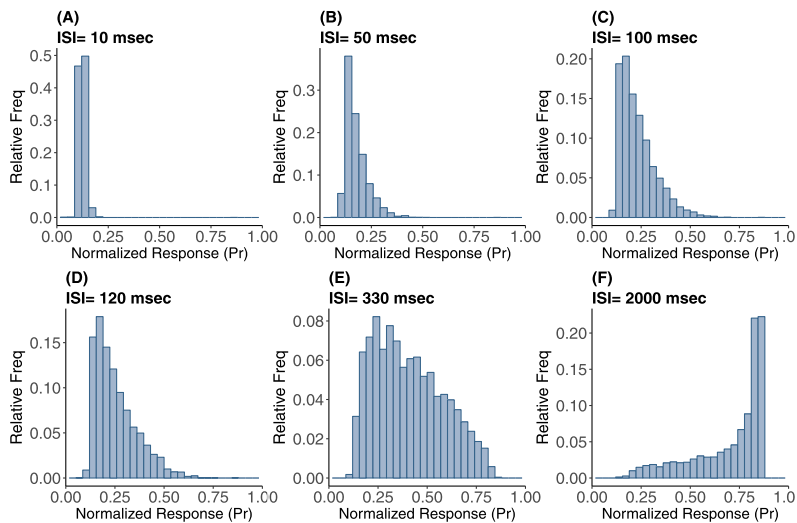
Fig. 9 Frequency distributions of normalized postsynaptic response Pr for varying input ISI (**A**) 10, (**B**) 50, (**C**) 100, (**D**) 120, (**E**) 330, and (**F**) 2000 in ms

Fig. 10 Quantile plots comparing the distributions of simulated normalized postsynaptic response values Pr and fixed point postsynaptic response $\overline{PR}$ for varying input ISI (**A**) 10, (**B**) 50, (**C**) 100, (**D**) 120, (**E**) 330, and (**F**) 2000 in ms. The plots show adherence to a linear relationship between the observed quantiles of both data sets
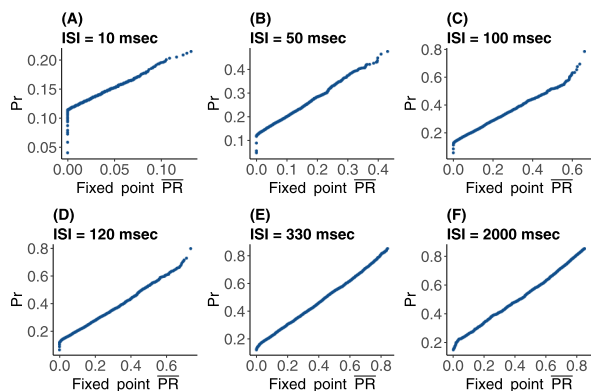
5.1 Model Description

Noise is a fundamental constraint to information transmission, and such variability is inherent in individual neurons in the nervous system. To account for the variability across identical stimulation trials, we use stochastic model of the synapse. Here, we follow the Katz model of neurotransmitter release; see [31] and [9].

Upon the arrival of an action potential at a presynaptic terminal, calcium influx through calcium channels causes some of the synaptic vesicles to fuse to the terminal bouton membrane at special release sites and neurotransmitter diffuses across the synaptic cleft. Each site can release either one or zero vesicles and we assume there are N release sites. Each release event is independent, and we assume that release of K vesicles from N release sites follows a *binomial* distribution with two parameters

(N, Pr) and is written

$$K \sim \mathcal{B}(N, Pr).$$

Pr is the release probability for each release site following an action potential obtained from the map. Note that failure of release results in a zero amplitude response from the postsynaptic neuron, so it cannot be informative. In addition, there is not only variability in the number of sites activated and hence the number of vesicles actually released, but also in the postsynaptic response to a single vesicle, due to inherent stochasticity in the postsynaptic receptors. Therefore, we assume that the size of the postsynaptic response ($\mathcal{S}_{\text{resp}}$) at the time of spike is not a constant, but instead it follows a *normal* distribution with mean μ and variance σ^2, with a two-sided truncation which is written

$$\mathcal{S}_{\text{resp}_i} \sim \mathcal{N}(\mu, \sigma).$$

Therefore, the amplitude of postsynaptic response following each action potential is obtained by combining the binomial model of vesicle release with the Normal model of a single response. The summation of responses evoked by each release can be written

$$\mathcal{S} = \begin{cases} 0 & \text{if } K = 0, \\ \sum_{i=1}^{k} \mathcal{S}_{\text{resp}_i} & \text{if } K > 0. \end{cases}$$

Note that, for $K > 0$, $\mathcal{S} \sim N(k\mu, \sqrt{k}\sigma)$. The probability density of the postsynaptic response to release of single vesicle is thus

$$f(\mathcal{S} = s | \mu, \sigma^2) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2 \mathcal{N}_{cts}}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} & \text{if } 0 < s < 2\mu, \\ 0 & \text{if OW,} \end{cases}$$

where

$$\mathcal{N}_{cts} = \int_0^{2\mu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} \, ds.$$

We will use this formulation in what follows.
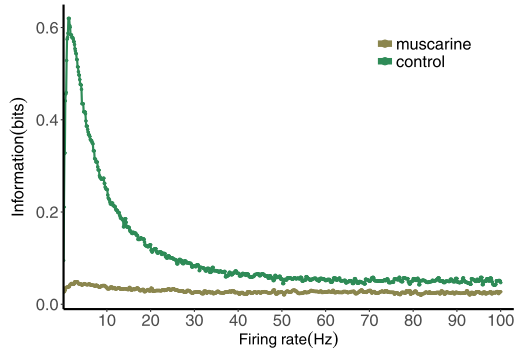
5.2 Mutual Information Calculations

In this section we apply information-theoretic measures introduced by Shannon [32] to quantify the ability of synapses to transmit information. One important concept in information theory is *mutual information*. It measures the expected reduction in uncertainty about x that results from learning y, or vice versa. This quantity can be formulated

$$I(X; Y) = H(X) + H(Y) - H(X, Y),$$

where the *entropy* was defined earlier and

$$H(X, Y) = -\sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(X = x, Y = y) \log_2 p(X = x, Y = y), \quad (6)$$

Fig. 11 Mutual information between normalized postsynaptic response and preceding ISIs for the control and muscarine parameter sets stimulated by Poisson spike train with mean frequencies ranging from 0.1 to 100 Hz

is the *joint entropy* of two random variables X and Y quantifies the uncertainty of their joint distribution. Using Eqs. (1) and (6), the mutual information can be rewritten

$$I(X; Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2 \frac{p(x|y)}{p(x)}. \tag{7}$$

The mutual information is symmetric in the variables X and Y, $I(X; Y) = I(Y; X)$, and is zero if the random variables are independent. Note that if the relation between them is deterministic, it is equal to the entropy in either variable. In order to compute this from data one is faced with the basic statistical problem of estimating the entropies: selecting the correct bin width in the histograms of the random variables. Here, we use the *Freedman–Diaconis* rule [33] for finding the optimal number of bins. According to this rule, the optimal number of bins can be calculated based on the interquartile range ($Iqr = Q_3 - Q_1$) and the number of data points n. It is defined as

$$nbins = \frac{\max(x) - \min(x)}{2 * Iqr * n^{-1/3}}.$$

One of the advantages of the *Freedman–Diaconis rule* is that the variability in the data is taken into account using *Iqr*, which is resistant to outliers in the data.

We stimulated synapses with input Poisson spike trains with different mean frequencies and computed the mutual information between the postsynaptic response and the preceding inter-spike interval. The mutual information then is obtained by

$$I(\mathcal{S}; T) = H(\mathcal{S}) + H(T) - H(\mathcal{S}, T).$$

Figure 11 shows the result of this calculation for the control and muscarine cases. In both we observe a rapid decline in mutual information at frequencies above 7 Hz. The peak mutual information occurs between 0.1 to 2 Hz, all of which demonstrate the frequency dependence of temporal information encoding. Large inter-spike intervals (low frequency) allow enough time for recovery of all release sites, leading to a tight distribution about $Pr = P_{\max}$, and very little communication of information. Very short inter-spike intervals give no time for recovery and Pr is confined to a small region near 0, also leading to low information transmission. Between these extremes,

variable inter-spike intervals have a significant effect on the amplitude of *Pr*. The main difference between the muscarine and control cases is the absolute value of the mutual information, which is significantly lower in muscarine case. This can be understood by considering that the range of *Pr* values in the muscarine case is much smaller, so variations in *Pr* are less distinguishable, and contribute less to the mutual information.

This coincides with results in [9], where depressing synapses are found to have a peak mutual information at very low frequencies. However, here we also see this in the distributions themselves, where the amount of variability is represented by the overall width or entropy of the distribution. It is not surprising that the mutual information will reflect this, especially considering that these entropies are the foundation of the calculation.

6 Information "Stored" in the Postsynaptic Response

We showed that a distribution of postsynaptic responses (PSRs) carries information about the distribution of the ISIs (inter-spike intervals). However, the exact sequence of ISIs determines a given postsynaptic response. The length of the sequence that effects a response will depend on the parameters of the model, in that the calcium can accumulate in certain situations. We call this a kind of "memory" for an exact sequence of ISIs. We measure this by computing the mutual information between a prior sequence and the PSR. This can be done in a coarse grained way by adding up a sum of N previous ISIs and using that as a random variable, which we consider first. This, however, ignores the subtleties of the exact sequence; for instance, a long ISI followed by a short ISI will give a PSR that is smaller than the reverse, given that the sum of the two is the same. A more complete approach is to compute the mutual information between a PSR and an N-tuple of preceding spikes, in order. The former method has been used in previous work so we begin with that approach to demonstrate some of these aforementioned subtleties.

6.1 MI Between PSR and Sum of Preceding ISIs

In [9], one computes the mutual information between the postsynaptic responses a sum of the preceding ISIs, increasing the number in the sum to go further back in time. Assuming that first spike occur at $t_0 = 0$ ms, this can be formulated as follows. Let $t_1 = T_1$, $t_2 = T_1 + T_2$, ..., $t_k = T_1 + \cdots + T_k$ be a vector of the sums of the preceding ISIs. It can be simplified by

$$t_k = \sum_{i=1}^{k} T_i \quad \text{for } k = 1, \ldots, M,$$

where $M > 0$ is a natural number. The mutual information between the postsynaptic response and the sum of preceding spike ISIs is then

$$I(\mathcal{S}; t_k) = H(\mathcal{S}) - H(\mathcal{S}|t_k) \quad \text{for } k = 1, \ldots, M.$$

Fig. 12 Mutual information between postsynaptic response and number of sum of preceding inter-spike intervals in the control and muscarine cases under 5 Hz Poisson stimulation
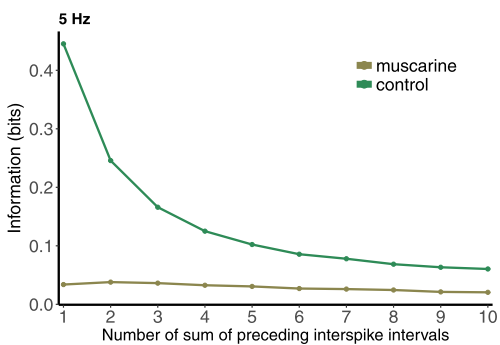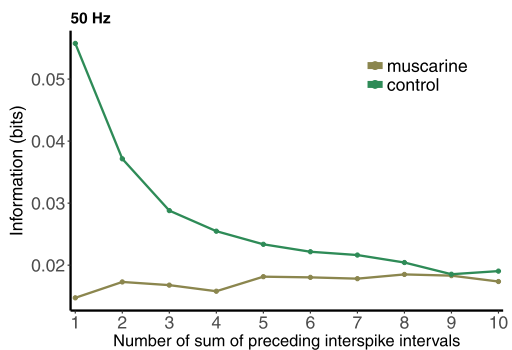

Fig. 13 Information between postsynaptic response and number of sum of preceding inter-spike intervals in control and muscarine cases under 50 Hz Poisson stimulation

Figures 12 and 13 illustrate the results for both the control and the muscarine cases, at 5 and 50 Hz, respectively. The information contained in the sum is significantly lower for muscarine compared to the control case, which is not surprising, given the results from the preceding section. For both firing rates, 5 and 50 Hz, in the control case the MI decreases as more ISIs are added to the sum. The further back in time the sum goes, the less the ISIs in the sum are directly involved in determining the PSR. The MI curve becomes almost flat past five preceding ISIs for 50 Hz, indicating an effect that can be measured only that far back. In the control 5 Hz case the decay begins to level off at $N = 5$ as well, but continues to decline for $N > 5$. This shows a frequency dependence of the measure that makes sense mechanistically. The muscarine MI is much less dependent on the cumulative history of spikes, showing very little change as N is increased. This also makes sense mechanistically, because in the muscarine case the response range is very narrow, and cannot carry much information forward from one preceding ISI, let alone any ISIs preceding that.

6.2 Mutual Information (MI) Between Postsynaptic Response $\mathcal{S}$ and n-Tuple Inter-spike Intervals $T_1, T_2, \ldots, T_n$

Our second method is much more computationally intense, but preserves the exact structure of the sequence of preceding ISIs.

Consider a single-input and single-output channel with inter-spike input T and postsynaptic response output $\mathcal{S}$. The mutual information between T and $\mathcal{S}$ in this

notation is defined

$$I(T; \mathcal{S}) = H(T) + H(\mathcal{S}) - H(T, \mathcal{S}).$$

Consider now a channel with two inter-spike interval inputs T_1 and T_2 and a single output $\mathcal{S}$. The mutual information between the inputs and the output of two-way channel is commonly known as the 2-tuple information and can be defined as the amount of reduction of uncertainty in $\mathcal{S}$ knowing (T_1, T_2). The 2-tuple mutual information is written

$$I(\langle T_1, T_2 \rangle; \mathcal{S}) = H(T_1, T_2) + H(\mathcal{S}) - H(T_1, T_2, \mathcal{S}).$$

Following McGill [34] we can extend the definition for mutual information to include more than two sources $T_1, T_2, \ldots, T_n$ that transmit to $\mathcal{S}$, arriving at

$$I(\langle T_1, T_2, \ldots, T_n \rangle; \mathcal{S}) = H(\mathcal{S}) + H(T_1, T_2, \ldots, T_n) - H(T_1, T_2, \ldots, T_n, \mathcal{S}). \quad (8)$$

The histogram construction of probability density functions in higher dimensions suffers from the "curse of dimensionality" [35]. Kozachenko and Leonenko [36] instead proposed an improved nonparametric estimator for entropy: the KL estimator based on k-nearest neighbor distances. In [37], Kraskov et al. reported on a modified KL estimator to compute mutual information with improved performance and applicability in higher dimensional spaces (the result is referred to as the KSG algorithm). Compared to other methods for multivariate data sets, estimators obtained by the KSG algorithm have minimal bias when applied to data sets of limited size, as is common in real world problems.

The KL estimator demonstrates that it is not necessary to estimate the probability density function in order to compute information theoretic functionals. Instead, the estimator is based on the metric properties of nearest neighbor balls in both the joint and the marginal spaces. The general form of the KL estimator for the differential entropy is written as

$$\hat{H}(X) = \psi(N) - \psi(k) + \log c_d + \frac{d}{N} \sum_{i=1}^{N} \log \epsilon(i),$$

where $\psi : \mathbb{N} \to \mathbb{R}$ denotes the digamma function, $\epsilon(i)$ is twice the distance from point x_i to its kth neighbor, d is the dimension of x and c_d is the volume of the d-dimensional unit ball. Similarly, KL estimator for the joint entropy is written as

$$\hat{H}(X, Y) = -\psi(k) + \psi(N) + \log c_{d_X} c_{d_Y} + \frac{d_X + d_Y}{N} \sum_{i=1}^{N} \log \epsilon(i). \quad (9)$$

Therefore, for any fixed k, the mutual information can be computed by subtracting the joint entropy estimate from $H(X)$ and $H(Y)$. However, this introduces biases due to the different distance scales in different dimensions. The KSG algorithm corrects this by instituting a random choice of the nearest neighbor parameter k. For a more detailed derivation see [37].
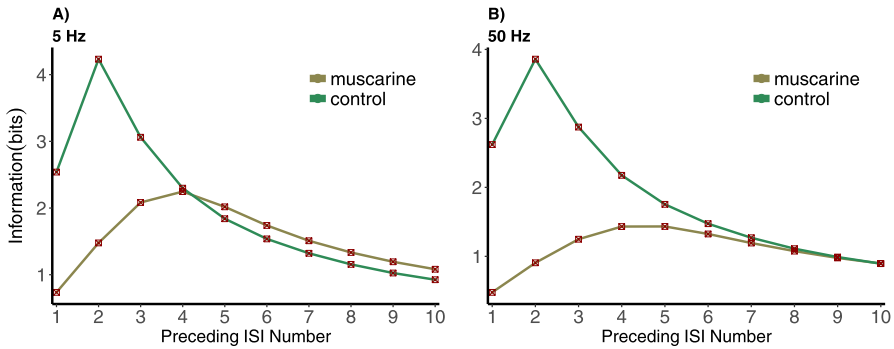
Fig. 14 *N*-tuple information between ISIs and normalized postsynaptic response *Pr* for control and muscarine cases under (**A**) 5 and (**B**) 50 Hz Poisson stimulation

We now employ it to estimate mutual information between the probability of release (*Pr*) as single output and preceding ISIs as a multivariate input. Note that we are able to use the deterministic *Pr* distributions in these calculations because *Pr* is not directly dependent upon the *n*-tuple of the preceding ISIs.

Figure 14 shows the increase in mutual information (or reduction in uncertainty) between the *Pr* and an *n*-tuple ISI, with increasing *n*. Mean rates in the gamma and theta ranges, 5 and 50 Hz, respectively, are plotted for the control and muscarine cases. At 5 Hz in the control case the mutual information increases from 1 to 2 preceding ISIs, after which it decreases, tending toward some limiting value. Because the MI is still greater than for one ISI for three preceding ISIs, we could say that this synapse "remembers" up to three preceding ISIs. For the muscarine case the increase happens over the first four preceding ISIs, meaning the uncertainty is reduced by adding in previous ISIs, and in this parameter regime the synapse "remembers" somewhat further back in the ISI train than in the control case. The mutual information is even greater in muscarine than control for more than 4 ISIs at 50 Hz. This is not at all obvious from the other results we have presented in this paper, and could not be uncovered without these statistics on sequences of ISIs. At 50 Hz the muscarine MI is always smaller than the control MI, but we see almost the same history dependence for each synapse-independent of frequency.

7 Discussion and Conclusion

When presented with a Poisson spike train input, our depressing synapse model acts as a nonlinear filter of the exponential distribution of inter-spike intervals. For high and low input frequencies, the result is as expected, the output *Pr* distribution is peaked at values near zero and $P_{\max}$, respectively, with very small variance (see Fig. 2). In between, for frequencies near theta, the distribution is spread across the entire interval between zero and $P_{\max}$. This creates a peak in the entropy of the distribution near this value, which demonstrates a wide dynamic range in the response of the synapse (Fig. 5). The range in the control case is larger than in muscarine, necessarily, because $P_{\max}$ is larger.

Over the gamma frequency range we see that the peak and mean of the distribution remain more or less constant, indicating a stable response size when presented with a Poisson spike train. This would create a stable connection and is the advantage of having a peaked distribution with low entropy. The addition of muscarine reduces the strength of this connection.

These are the results of a numerical investigation of the properties of the synapse. Creating closed form expressions for these distributions is not possible, the map is too complex. However, the stochastic recurrence relation for the calcium concentration alone is simple enough to be analyzed and we discovered that it follows a gamma distribution with shape parameter $\lambda \tau_{\mathrm{ca}} + 1$. For the Pr distribution we had to rely on an approximation motivated by numerical results. We found the mean of the distribution followed the frequency in the same way that the fixed point of the map did. The collapse to the fixed point is very rapid, one or two iterations at most, which justifies our assumption that the Pr values are determined by the fixed point value associated with the instantaneous rate associated with the preceding inter-spike interval, e.g. $\lambda = T$. Then the formula for the fixed point as a function of frequency could be used to generate a distribution, using the exponential distribution of the Ts. The results confirm the validity of this approximation, even in ranges outside the other approximations necessary to arrive at a closed form. Most importantly, the "sloshing" of the distribution between zero and $P_{\max}$ as the frequency is decreased through the physiological range is captured by this form.

For a given train of inter-spike intervals the map for $Pr = PR$ is completely deterministic, it captures the mean of the response behavior. Therefore the mutual information between the Pr distribution and the ISI distribution should be equal to the entropy of Pr. In order to introduce stochasticity to the response we modeled the noise in the release of vesicles of neurotransmitter as a binomial process, and the noisy postsynaptic response to the vesicles with a truncated Normal distribution. With a distribution of the resulting PSR values, we calculated mutual information. The mutual information as a function of firing rate was found to have a peak around 3 Hz for both the muscarine and the control cases, though the overall mutual information was much lower for muscarine, in part because of the lower entropy of the Pr distribution in this case (Fig. 11). The peak in the mutual information in the control case is 6 times that of the baseline value for large firing rates. For the muscarine case it was only about 1.25 times larger. Therefore the synapses with muscarine added as a neuromodulator are much less sensitive to changes in frequency overall. This can be viewed as a good thing with regard to stability of the response, but creates a much less dynamic filter than in the control case.

For comparison, we note here that in [38] they determine that changes in response amplitude with presynaptic short-term plasticity are balanced out by postsynaptic conductance noise at higher firing rates. Interestingly, at lower rates, the effects of short-term plasticity are more evident. For our experimentally fit synaptic model, maximum entropy and maximum mutual information occur a very low frequencies, around 3–5 Hz. This means that these effects would not be washed out by fluctuations in membrane conductance on the postsynaptic side. Additionally, Rotman et al. [39], in a rate coding type study, conclude that depressing interneuron synapses in hippocampus possess optimal information transfer characteristics at lower frequencies,

indeed, single spikes. Finally, it was found in [40] that models that include presynaptic depression and stochastic variation in vesicle release have lower information transfer properties at lower frequencies than high frequencies, while their model without stochasticity showed no such frequency dependence. Clearly the effects of short-term plasticity on information transfer in these different studies need to be compared very carefully to arrive at definitive conclusions.

We then took these calculations a step further by attempting to determine how far back in a spike train the synapses "remembers", in that information from previous ISIs is carried forward into a certain measurement of the Prs. This is a non-trivial problem, with technical challenges in the calculations. We performed a mutual information calculation with a sum of the previous ISIs, successively adding more times. This sum collapses much of the information in the sequence, measuring only if it was overall a long time or a short time, compared to other samples. Some results are obtainable, however, and we saw an overall decline in mutual information as more ISIs were added to the sum, at least in the control case (Figs. 12 and 13). In the muscarine case the mutual information is seemingly insensitive to adding in more ISIs, being relatively flat. The absolute value of the mutual information was much less in the two cases for 50 Hz than for 5 Hz. Again, the larger entropy of the Pr distribution at 5 Hz creates the potential for larger mutual information. We did see that the MI leveled out more slowly at 5 Hz than 50 Hz, too, for the control case, indicating a longer "memory" at 5 Hz than at 50 Hz.

In an attempt to keep all the structure in a preceding inter-spike interval sequence we used a multivariate version of the mutual information calculation. Because of the high dimensionality of the data this is more challenging. A k-nearest neighbor approach to estimating the entropy is appropriate in this case and we used the KSG algorithm to do the calculation. This measures the mutual information between the response and the sequence of ISIs as the number in the sequence is increased. If it increases with more ISIs in the sequence we can assume that the reduction of uncertainty is greater as longer histories are considered. It is seen to increase initially and then decrease with history length as the memory effect is washed out (Fig. 14). Somewhat non-intuitively the muscarine case increases over longer histories than in control, though the overall MI is again smaller. Adding up to 4 ISIs to the sequence improves the prediction of the Pr vs. two–three ISIs in control. This result is more or less the same at gamma and theta frequencies. Using the sum of the ISIs in the mutual information calculation hides this dependence in the muscarine case, as anticipated.

Through this analysis we have gained insight into the information processing characteristics of PV BCs. Within a physiological context, PV BCs receive synaptic input from intrinsic and extrinsic sources, effectively tuning them to fire specifically at theta frequency [41]. It is clear from analysis that firing of PV BCs at theta frequency optimizes the information content of PV BC to pyramidal cell synaptic transmission. Moreover, PV BCs in vivo often burst more than one spike per theta cycle [41]. This short "burst" of more than one action potential in vivo serves to further optimize the information content, providing a stronger "memory" of PV BC activity.

Cholinergic neuromodulation of PV BCs occurs both pre- and postsynaptically ([1, 42]) and is associated with the generation of population-level gamma rhythms. By reducing Pr and protecting the synapse from vesicular depletion, we now show that

presynaptic neuromodulation reduces entropy and mutual information. While this effect may reduce the information content of the synapse, it will promote stability and regularity of the synaptic response that might be advantageous in the generation of population-level gamma rhythms. Therefore, we hypothesize two modes for PV BC synapses. In the absence of cholinergic neuromodulation, PV BCs are optimally tuned to transfer information at theta frequency, which might be advantageous in encoding the onset of salient stimuli [43]. In the presence of cholinergic neuromodulation, when PV BCs may become depolarized [42] and are more likely to fire at gamma frequency, the information processing capability is reduced, gaining synapse stability, which would promote population-level network synchronization. Future studies that employ "in vivo" spike trains could corroborate these hypotheses.

We are continuing our quest to quantify the information processing properties of synapses using more novel techniques than mutual information. Computational mechanics ([44, 45]) gives us a theory and a basis for understanding time series from a process as a hidden Markov model with multiple states. How the model changes as input frequency is varied, or neuromodulation is applied, would give us an even better categorization of the synaptic processes. Furthermore, these measures can be applied to output processes alone, without any information about the input stimuli, allowing a much more robust classification of signals measured from synapses in vivo or in vitro.

Understanding the interaction of the synaptic dynamics with voltage oscillations in the hippocampus is the next step to connecting the dots between synaptic plasticity and it effect on the mechanisms involved in learning and memory. We have preliminary results in a simple deterministic model and will be extending these to stochastic models of micro-circuit rhythms. We believe that a careful piecing together of the model and behavior will guide our intuition much more than a large scale numerical approach, and it will more easily interface with electrophysiological experiments on actual neurons in the hippocampus.

Availability of data and materials   Please contact author for data requests.

Ethics approval and consent to participate   Not applicable.

Competing interests   The authors declare that they have no competing interests.

Consent for publication   Not applicable.

Authors' contributions   EBM did all the statistical analysis in the study. JJL carried out all experiments and preliminary data analysis. EFS conceived of the study, developed the design and analyzed the results. All authors read and approved the final manuscript.

Appendix A

Let $\theta_1, \theta_2, \ldots$ be i.i.d. $\mathbb{R}^2$-valued ($d \geq 1$) random variables with generic copy θ and independent of X_0. Suppose that $X_n = \Psi(X_{n-1}, \theta_n)$ for all $n \geq 1$ and a continuous function $\Psi : \mathbb{R}^{d+1} \to \mathbb{R}$. If X_n converges in distribution to X_∞, then

$$\Psi(X_{n-1}, \theta_n) \xrightarrow{d} \Psi(X_\infty, \theta) \quad \text{and} \quad X_\infty \overset{d}{=} \Psi(X_\infty, \theta).$$

Appendix B

Suppose independent, identically distributed increases in the amount of calcium $\{\Delta_n\}_{n \geq 1}$ occur at times $\{t_n\}_{n \geq 1}$, and we are interested in the distribution of the amount of calcium accumulated. In one dimension, the random recurrence equation for the calcium concentration is given by

$$C_n = A_n C_{n-1} + \Delta_n, \quad n \geq 1, \tag{10}$$

where $A_n = e^{-T_n/\tau_{\mathrm{ca}}}$, and the waiting times $T_1 = t_1$, $T_n = t_n - t_{n-1}$, $n \geq 2$, are i.i.d., making the $\{T_n\}$ a renewal process. Moreover, (A_n, Δ_n) are assumed to be i.i.d. vectors. C_0 is the base calcium concentration which is assumed to be zero.

Iterating (10) leads to

$$\begin{aligned}
C_n &= A_n C_{n-1} + \Delta_n \\
&= A_n A_{n-1} C_{n-2} + A_n \Delta_{n-1} + \Delta_n \\
&= A_n A_{n-1} A_{n-2} C_{n-3} + A_n A_{n-1} \Delta_{n-2} + A_n \Delta_{n-1} + \Delta_n \\
&\;\;\vdots \\
&= A_n A_{n-1} \cdot \ldots \cdot A_1 C_0 + \sum_{k=1}^{n} A_n \cdot \ldots \cdot A_{k+1} \Delta_k
\end{aligned}$$

for each $n \geq 1$. Using the independence assumptions and replacing $(A_k, \Delta_k)_{1 \leq k \leq n}$ with the copy $(A_{n+1-k}, \Delta_{n+1-k})_{1 \leq k \leq n}$ we observe that

$$C_n \overset{d}{=} A_n A_{n-1} \cdot \ldots \cdot A_1 C_0 + \sum_{k=1}^{n} A_1 A_2 \cdot \ldots \cdot A_{k-1} \Delta_k,$$

where $\overset{d}{=}$ denotes equality in distribution.

A fundamental theoretical result shown in [46] asserts that if

$$E\left(\ln |A|\right) < 0 \quad \text{and} \quad E\left(\ln |\Delta|\right) < \infty$$

then the series

$$C = \sum_{k=1}^{\infty} A_1 A_2 \cdot \ldots \cdot A_{k-1} \Delta_k$$

will converge w.p.1 and the distribution of C_n converges to that of C, independently of C_0. Also, by the continuous mapping theorem (see Appendix A), we infer from (10) that if $C_n \overset{d}{\to} C$, then C satisfies the distributional identity

$$C \overset{d}{=} AC + \Delta, \quad C \text{ and } (A, \Delta) \text{ independent.}$$

Following [47] let $X := AC$, where A, C and X are independent. Then we can write

$$X \overset{d}{=} A(X + \Delta), \tag{11}$$

and hence

$$C \overset{d}{=} X + \Delta.$$

Iterating (11) results in

$$X \overset{d}{=} \sum_{n=1}^{\infty} A_1 A_2 \cdot \ldots \cdot A_n \Delta_n,$$

where $A_n = e^{-T_n/\tau_{\text{ca}}}$.

Note that if T_n is exponentially distributed with rate parameter λ, then random variable A_n has a beta distribution with shape parameters $(\lambda\tau_{\text{ca}}, 1)$ and is denoted by $A_n \sim Beta(\lambda\tau_{\text{ca}}, 1)$. We then apply beta–gamma algebraic identities ([48]) which state that

$$Beta(a, b) \odot Gamma(a + b, 1) = Gamma(a, 1).$$

Alternatively,

$$Beta(a, b) \odot \big(Gamma(a, 1) + Gamma(b, 1)\big) \overset{d}{=} Gamma(a, 1). \tag{12}$$

Applying (12) in (11), and considering that $A_n \sim Beta(\lambda\tau_{\text{ca}}, 1)$, then

$$X \sim Gamma(\lambda\tau_{\text{ca}}, 1),$$
$$\Delta \sim Gamma(1, 1).$$

Note that $C \overset{d}{=} X + \Delta$ implies

$$C \overset{d}{=} Gamma(\lambda\tau_{\text{ca}}, 1) + Gamma(1, 1).$$

Finally,

$$C \sim Gamma(\lambda\tau_{\text{ca}} + 1, 1).$$

Note that a random variable X that is gamma-distributed with shape α and scale θ parameters is denoted by $Gamma(\alpha, \theta)$ and the corresponding probability density function is

$$f(x; \alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} \quad \text{for } x > 0 \text{ and } \alpha, \theta > 0.$$

Also, a random variable X that is beta-distributed with shape parameters α and β is denoted by $Beta(\alpha, \beta)$ and the probability density function of the beta distribution is

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1 - x)^{\beta-1} \quad \text{for } 0 < x < 1 \text{ and } \alpha, \beta > 0,$$

where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

Appendix C

We address computing an approximate distribution for Pr. For ease of notation, let $X = \overline{R}$ be a random variable defined as follows:

$$X = \frac{1 - e^{-k_{\min}T}}{1 - (1 - P_{\max})e^{-k_{\min}T}}.$$

Let the random variable T have the exponential distribution with probability density function

$$f_T(t) = \lambda e^{-\lambda t}, \quad t > 0.$$

We can compute an analytical expression for probability density function (PDF) of fixed point $\overline{R}$ using the distribution for T.

The transformation $X = g(T) = \frac{1-e^{-k_{\min}T}}{1-(1-P_{\max})e^{-k_{\min}T}}$ is a 1–1 transformation from $\mathcal{T} = \{t|t > 0\}$ to $\mathcal{X} = \{x|0 < x < 1\}$ with inverse $T = g^{-1}(X) = \frac{1}{k_{\min}} \log(\frac{1-ax}{1-x})$ and Jacobian

$$\frac{dT}{dX} = \frac{1 - a}{k_{\min}(1 - x)(1 - ax)},$$

where $a = 1 - P_{\max}$. By the rule for functions of random variables, the probability density function of X is

$$f_X(x) = f_T\left(g^{-1}(x)\right)\left|\frac{dt}{dx}\right|$$

$$= \frac{\lambda(1 - a)}{k_{\min}}(1 - x)^{-(1-\lambda/k_{\min})}(1 - ax)^{-(1+\lambda/k_{\min})}.$$

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Lawrence JJ, Haario H, Stone EF. Presynaptic cholinergic neuromodulation alters the temporal dynamics of short-term depression at parvalbumin-positive basket cell synapses from juvenile CA1 mouse hippocampus. J Neurophysiol. 2015;113(7):2408–19.
2. Abbott LF, Regehr WG. Synaptic computation. Nature. 2004;431:796–803.
3. Anwar H, Li X, Bucher D, Nadim F. Functional roles of short-term synaptic plasticity with an emphasis on inhibition. Curr Opin Neurobiol. 2017;43(Supplement C):71–8. Neurobiology of Learning and Plasticity.
4. Vogels T, Froemke R, Doyon N, Gilson M, Haas J, Liu R, Maffei A, Miller P, Wierenga C, Woodin M, Zenke F, Sprekeler H. Inhibitory synaptic plasticity: spike timing-dependence and putative network function. Front Neural Circuits. 2013;7:119.
5. Destexhe A, Marder E. Plasticity in single neuron and circuit computations. Nature. 2004;431:789–95.
6. Abbott LF, Varela JA, Sen K, Nelson SB. Synaptic depression and cortical gain control. Science. 1997;275(5297):221–4.
7. Fauth M, Wörgötter F, Tetzlaff C. The formation of multi-synaptic connections by the interaction of synaptic and structural plasticity and their functional consequences. PLoS Comput Biol. 2015;11(1):e1004031.
8. Cartling B. Control of neural information transmission by synaptic dynamics. J Theor Biol. 2002;214(2):275–92.
9. Fuhrmann G, Segev I, Markram H, Tsodyks M. Coding of temporal information by activity-dependent synapses. J Neurophysiol. 2002;87(1):140–8.
10. Yang Z, Hennig MH, Postlethwaite M, Forsythe ID, Graham BP. Wide-band information transmission at the calyx of held. Neural Comput. 2009;21(4):991–1017. PMID: 19018705.
11. Tsodyks M, Pawelzik K, Markram H. Neural networks with dynamic synapses. Neural Comput. 1998;10(4):821–35. https://doi.org/10.1162/089976698300017502.
12. Zador A. Impact of synaptic unreliability on the information transmitted by spiking neurons. J Neurophysiol. 1998;79(3):1219–29. https://doi.org/10.1152/jn.1998.79.3.1219. PMID: 9497403.
13. Goldman MS, Maldonado P, Abbott LF. Redundancy reduction and sustained firing with stochastic depressing synapses. J Neurosci. 2002;22(2):584–91. https://doi.org/10.1523/JNEUROSCI.22-02-00584.2002. http://www.jneurosci.org/content/22/2/584.
14. London M, Schreibman A, Häusser M, Larkum ME, Segev I. The information efficacy of a synapse. Nat Neurosci. 2002;5:332–40.
15. Stone E, Haario H, Lawrence JJ. A kinetic model for the frequency dependence of cholinergic modulation at hippocampal GABAergic synapses. Math Biosci. 2014;258:162–75.
16. Lee C-CJ, Anton M, Poon C-S, McRae GJ. A kinetic model unifying presynaptic short-term facilitation and depression. J Comput Neurosci. 2008;26(3):459–73.
17. Khanin R, Parnas I, Parnas H. On the feedback between theory and experiment in elucidating the molecular mechanisms underlying neurotransmitter release. Bull Math Biol. 2006;68(5):997–1009.
18. Markram H, Wang Y, Tsodyks M. Differential signaling via the same axon of neocortical pyramidal neurons. Proc Natl Acad Sci USA. 1998;95(9):5323–8.
19. Dittman JS, Kreitzer AC, Regehr WG. Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. J Neurosci. 2000;20(4):1374–85.
20. Buzsáki G, Wang X-J. Mechanisms of gamma oscillations. Annu Rev Neurosci. 2012;35(1):203–25. PMID: 22443509.
21. Tsodyks MV, Markram H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci USA. 1997;94(2):719–23.
22. Lu T, Trussell LO. Inhibitory transmission mediated by asynchronous transmitter release. Neuron. 2000;26(3):683–94.
23. Zucker RS, Regehr WG. Short-term synaptic plasticity. Annu Rev Physiol. 2002;64(1):355–405. PMID: 11826273.
24. Dittman JS, Regehr WG. Calcium dependence and recovery kinetics of presynaptic depression at the climbing fiber to Purkinje cell synapse. J Neurosci. 1998;18(16):6147–62.
25. Stevens CF, Wesseling JF. Activity-dependent modulation of the rate at which synaptic vesicles become available to undergo exocytosis. Neuron. 1998;21(2):415–24.
26. von Gersdorff H, Schneggenburger R, Weis S, Neher E. Presynaptic depression at a calyx synapse: the small contribution of metabotropic glutamate receptors. J Neurosci. 1997;17(21):8137–46.

27. Wang L-Y, Kaczmarek LK. High-frequency firing helps replenish the readily releasable pool of synaptic vesicles. Nature. 1998;394:384–8.
28. González JC, Lignani G, Maroto M, Baldelli P, Hernández-Guijo JM. Presynaptic muscarinic receptors reduce synaptic depression and facilitate its recovery at hippocampal GABAergic synapses. Cereb Cortex. 2014;24(7):1818–31.
29. Haario H, Saksman E, Tamminen J. An adaptive Metropolis algorithm. Bernoulli. 2001;7(2):223–42.
30. Haario H, Laine M, Mira A, Saksman E. DRAM: efficient adaptive MCMC. Stat Comput. 2006;16(4):339–54.
31. Stevens CF. Quantal release of neurotransmitter and long-term potentiation. Cell. 1993;72(Supplement):55–63.
32. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27(3):379–423.
33. Freedman D, Diaconis P. On the histogram as a density estimator: L2 theory. Z Wahrscheinlichkeitstheor Verw Geb. 1981;57(4):453–76.
34. McGill WJ. Multivariate information transmission. Psychometrika. 1954;19(2):97–116.
35. Kwak N, Choi C-H. Input feature selection by mutual information based on Parzen window. IEEE Trans Pattern Anal Mach Intell. 2002;24(12):1667–71.
36. Kozachenko LF, Leonenko NN. Sample estimate of the entropy of a random vector. Probl Inf Transm. 1987;23(1–2):95–101.
37. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E. 2004;69:066138.
38. Lindner B, Gangloff D, Longtin A, Lewis JE. Broadband coding with dynamic synapses. J Neurosci. 2009;29(7):2076–87. https://doi.org/10.1523/JNEUROSCI.3702-08.2009. http://www.jneurosci.org/content/29/7/2076.
39. Rotman Z, Deng P-Y, Klyachko VA. Short-term plasticity optimizes synaptic information transmission. J Neurosci. 2011;31(41):14800–9. https://doi.org/10.1523/JNEUROSCI.3231-11.2011. http://www.jneurosci.org/content/31/41/14800.
40. Rosenbaum R, Rubin J, Doiron B. Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. PLoS Comput Biol. 2012;8(6):e1002557. https://doi.org/10.1371/journal.pcbi.1002557.
41. Varga C, Golshani P, Soltesz I. Frequency-invariant temporal ordering of interneuronal discharges during hippocampal oscillations in awake mice. Proc Natl Acad Sci USA. 2012;109(40):E2726–34.
42. Yi F, Ball J, Stoll KE, Satpute VC, Mitchell SM, Pauli JL, Holloway BB, Johnston AD, Nathanson NM, Deisseroth K, Gerber DJ, Tonegawa S, Lawrence JJ. Direct excitation of parvalbumin-positive interneurons by M1 muscarinic acetylcholine receptors: roles in cellular excitability, inhibitory transmission and cognition. J Physiol. 2014;592(16):3463–94.
43. Pouille F, Scanziani M. Routing of spike series by dynamic circuits in the hippocampus. Nature. 2004;429:717–23.
44. Crutchfield JP. The calculi of emergence: computation, dynamics, and induction. Physica D. 1994;75:11–54.
45. Shalizi CR, Crutchfield JP. Computational mechanics: pattern and prediction, structure and simplicity. J Stat Phys. 2001;104(3):817–79.
46. Kesten H. Random difference equations and renewal theory for products of random matrices. Acta Math. 1973;131:207–48.
47. Vervaat W. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. Adv Appl Probab. 1979;11(4):750–83.
48. Dufresne D. Algebraic properties of beta and gamma distributions, and applications. Adv Appl Math. 1998;20(3):285–99.