Using Shannon Information to Probe the Precision of Synaptic Strengths

Mohammad Samavat
Department of Electrical and Computer Engineering
UC San Diego
Computational Neurobiology laboratory
The Salk Institute
10010 N Torrey Pines Rd, La Jolla, CA 92037
msamavat@ucsd.edu

Thomas M Bartol*
Computational Neurobiology laboratory
The Salk Institute
10010 N Torrey Pines Rd, La Jolla, CA 92037
bartol@salk.edu

Kristen M Harris†
Center for Learning and Memory
UT Austin
100 E 24TH ST AUSTIN, TX 78712
kharris@utexas.edu

Terrence J Sejnowski‡
Computational Neurobiology laboratory
The Salk Institute
10010 N Torrey Pines Rd, La Jolla, CA 92037
Terry@salk.edu

Abstract

Synapses between neurons control the the strengths of neuronal communication in neural circuits and their strengths are in turn dynamically regulated by experience. Because dendritic spine head volumes are highly correlated synaptic strength [1], anatomical reconstructions can probe the distributions of synaptic strengths. Synapses from the same axon onto the same dendrite (SDSA pairs) have a common history of coactivation and have nearly the same spine head volumes, suggesting that synapse function precisely modulates structure. We have applied Shannon information theory to obtain a new analysis of synaptic information storage capacity (SISC) using non-overlapping clusters of dendritic spine head volumes as a measure of synaptic strengths with distinct states based on the synaptic precision level calculated from 10 SDSA pairs. SISC analysis revealed spine head volumes in the stratum radiatum of hippocampal area CA1 occupied 24 distinct states (4.1 bits). This finding indicates an unexpected degree of precision that has implications for learning algorithms in artificial neural network models.

*http://www.mcell.cnl.salk.edu/
†https://3dem.org/
‡https://cnl.salk.edu/

Information-Theoretic Principles in Cognitive Systems Workshop at the 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

1 Introduction

Synapses are the fundamental units of storage of information in the brain and 20% of the human genome consists of genes related to synapse formation/modification. Almost every known neurological disorder can be attributed to synaptic dysfunction. Consequently, studying them is crucial and fundamental to multiple disciplines including neuroscience, cognitive science, psychiatry and the brain inspired intelligent systems.

Since the late 19th century that Santiago Ramón y Cajal proposed that memories are stored at synapses, and not through the generation of new neurons [2] there has been an extensive search for synaptic mechanisms responsible for learning and memory. While synaptic plasticity is a known experience-dependent mechanism for modifying these and other synaptic features, the precision of this mechanism has been unknown. The existence of both intrinsic and extrinsic origins of variability and dysfunction of structural modulation [3] urges scientists to further explore the potential precision of synaptic strength adjustments. From an information theory point of view, without precise synaptic plasticity there could be no information stored at the synapses of neural circuits. Consequently, no cognitive function could be available without precise adjustment of synaptic structure during learning and memory.

It has been long understood that LTP and LTD stands for long lasting changes in synaptic strength and the test pulses show that the information is there and retrieved. The weight itself is the information that is being stored. We hypothesize that synapses are not simple on/off switches in neural circuits but rather have distinguishable synaptic strength levels due to a high precision for adjustment of synaptic strength. Here we used information theory to develop a new method using reconstructed dendrites that quantifies empirically the SISC – the number of bits of Shannon information stored per synapse. In the new method, the precision analysis is based on the coefficient of variation (CV) of SDSA pairs, the same starting point as in [1]. The new method, however, performs non-overlapping cluster analysis (Algorithm 1) to obtain the number of distinguishable categories (N_c) of spine head volumes of whole populations of reconstructed spines in CA1 using the precision level estimated from the CV. The clustering analysis yields the Shannon Information capacity per synapse in CA1.

2 Results

2.1 Number of distinguishable categories for synaptic weight

To introduce and compare the performance of our proposed method, we reanalyzed the CA1 dataset that was previously analyzed with signal detection theory (Bartol et al., 2015). A total of 288 spine head volumes were fully contained within a 6 × 6 × 5 μm^3 CA1 neuropil volume (Fig. 1A). Signal detection theory revealed 26 distinguishable Gaussian distributions with equal CV of 0.12 ± 0.046 (inset, Fig. 1B), and assuming an overlap of 31% (Fig. 1B). This amount of overlap is equivalent to assuming a signal-to-noise ratio = 1 and a 69% discrimination threshold common in psychophysics (Schultz, 2007). Our new clustering method (Algorithms 1) based upon the median CV of the SDSA pairs, without any assumptions regarding the arbitrary parameter of signal-to-noise ratio, placed the CA1 spine head volumes into 24 distinguishable categories (Fig. 1C). The upper left inset contains 3D reconstructions of the smallest and largest spine head volumes. The largest spine in each cluster is illustrated beneath each bin. The frequency of spine head volumes in each category reveals that the spine head volumes are distributed across distinguishable states non-uniformly. The highest frequency occurs in cluster 10, which contains 36 spine head volumes (Fig. 1C). Interestingly, there appears to be a second peak at around cluster 21.

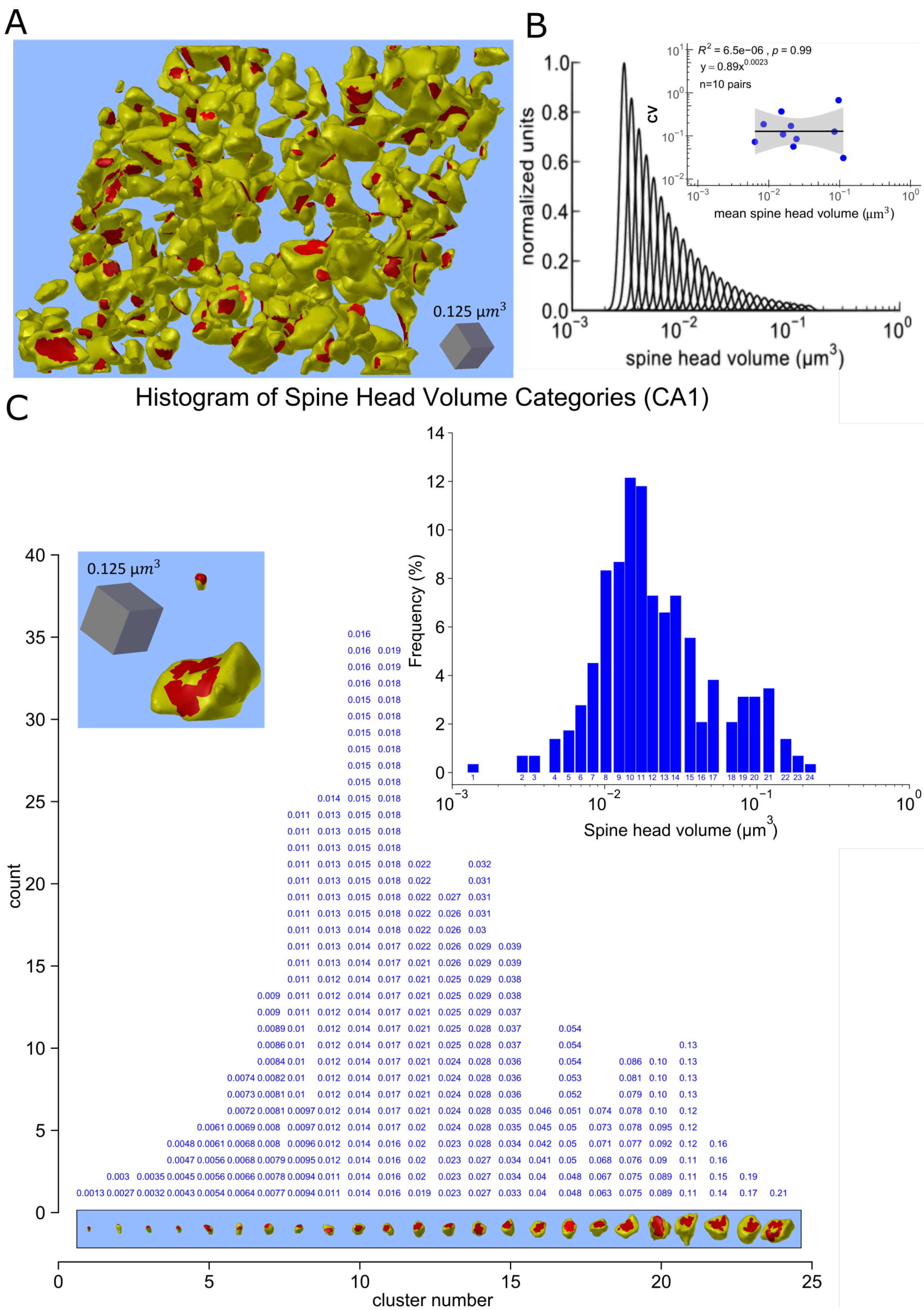

Figure 1: (A) The 288 spine heads fully captured in the reconstructed volume, displaying the PSD (red) and spine head membrane (yellow). (B) Authors in [1] using assumptions from signal detection theory showed that 26 distinguishable Gaussian distributions with equal CV (see inset) and overlap of 31% can span the range of spine head volumes of SDSA pairs equivalent to signal to noise ratio of 1 and 69% discrimination threshold common in psychophysics.

Figure 1: (C) Our new clustering algorithm (see Algorithm 2, methods) obtains 24 distinguishable categories of all 288 spine heads in the dataset based on the median CV value. The histogram of spine head volumes in log scale is depicted in the panel C inset. The Y axis shows the number of spine head volumes within each category. The actual spine head volumes of the individual spine heads of a given category are stacked vertically in sorted order for that category. The 3D object shown below each category (vertical column) is the actual 3D reconstructed spine head of the largest head volume in the category. The X axis shows the distinguishable category numbers. All spine head volumes are rounded to two significant digits.

Algorithm 1 Clustering Algorithm

function PRECISION CALCULATION((Same Dendrite Same Axon pairs (SDSA), N pairs (a,b) of spine head volumes))
  for $a, b \in SDSA[i]$ **do**
$cv = \sigma/\mu$
CV[i]=cv
  end for
  return $\{Median(CV)\}$
end function
function CLUSTERING SPINE HEAD VOLUMES(SHV vector)
  Sort SHV s.t. $SHV[i] < SHV[i+1]$
  $Listofshcluster = NULL$
  while $Length(SHV) \neq 0$ **do** ▷ Here we do the clustering with the median value of SDSA pairs calculated with the above function.
    a= SHV[1]
    for any $b \in SHV$ **do**
      Cluster=NULL
      if $cv(a,b) < Median(CV)$ **then**
        $Cluster \leftarrow b$
      end if
    end for
    $Listofshcluster[j] \leftarrow Cluster$
    $SHV = SHV[-Cluster]$(deleting the spine head volumes stored in cluster j from the SHV vector)
    j=j+1
  end while
  return $\{Listofshcluster\}$ and $N_c = j - 1$
end function

2.2 Shannon Information Storage Capacity of Synapses (CA1)

Shannon Information is a measurable quantity [4]. It can be defined as the amount of reduction in a recipient's uncertainty upon receipt of a message or observation of an event. For example, when observing a coin toss we are uncertain about the outcome of this probabilistic event. Is the coin biased or unbiased?. How many flips does one have to make to measure the amount of bias to some degree of certainty? How much decrease in uncertainty about the bias is obtained on each successive flip of the coin? If the coin is fair with uniform distribution:

$$P(Heads) = P(Tails) = 0.5 \tag{1}$$

then the outcome of each flip will remove the maximum uncertainty and hence carry the maximum possible information. However, if the coin is highly biased with distribution:

$$P(Heads) = 0.99, P(Tails) = 0.01 \tag{2}$$

then the outcome of each flip will remove very little uncertainty about the true result of coin flip experiment and thus carry very little information as we are almost sure that the outcome of the coin flip is heads most of the time. In this formulation, the amount of information in a message or an event is inversely proportional to its probability of occurrence. And formally, the Shannon Information content I of an event or message is defined as the base-2 logarithm of its reciprocal probability p (it uses the ~logarithm of base-2 and hence the unit of measure of I is called a "binary unit", or "bit" for short.):

$$I = \log_2 \frac{1}{p} \tag{3}$$

The concept of entropy, H, comes from the field of thermodynamics and measures the amount of uncertainty or disorder, or number of possible states of a system. Shannon entropy is defined as the average of Shannon information I. The Shannon entropy of a discrete random variable is defined as follows:

$$H(X) = \mathbb{E}\left[\log_2 \frac{1}{P_X(X)}\right] = \sum_{x\in\mathcal{X}} P_X(x) \log_2 \frac{1}{P_X(x)} \tag{4}$$

Hence we can say Shannon Information (or entropy) measures the number of distinguishable states. When you have higher information you can distinguish more things from one another.

The Shannon information per synapse was calculated from the frequency of spine head volumes in the distinguishable categories where each category is considered a different message. Shannon entropy (bits of information) for the hippocampal area CA1 dataset is measured with the value of 4.1 ± 0.39 bits. Results demonstrate that the synapses are not on/off switches. (For further detail please see the Synaptic Information Storage Capacity' section in Appendix)

3 Discussion

This paper introduces a new analytical approach for determining SISC that has several advantages over a prior method published in [1]. The new method is applied to the same data in [1] to compare it with the prior approach. The analyses revealed that synaptic precision, based on covariance of spine head volume in SDSA pairs, was not altered. The Shannon information per synapse was calculated from the frequency of spine head volumes in the distinguishable categories where each category is considered a different message.

1. Authors in [1] have not used the total 288 spine head volumes for calculating the number of distinguishable states. If the full range of spine head volumes were used the total number of Gaussian's for the 69% threshold with the signal detection method would be 33. (Here are clustering method reveals 24 distinguishable synaptic states)
2. The proposed method in [1] utilized an arbitrary threshold for the amount of overlap between consecutive Gaussian's (31% overlap). No evidence exists for what would be the right threshold and whether that differs between brain regions.
3. The novel method presented in the current manuscript is robust to outliers. For example, if we add a large spine head volume with value of 0.55 μm^3 (previously found in CA1 in other data sets) to the 288 spine head volumes of this study, the number of clusters with our method will be 25 instead of 24. However, the previous method in [1] reveals 39 distinguishable Gaussian's spanning the range of the augmented data set.

Comparison of the new SISC measurements with the previous results demonstrates that the new method is more robust to outliers and, importantly, can reveal gaps and variation in the shape of the distribution. In contrast, with signal detection theory gaps were filled in with Gaussians in the absence of any data. The new method provides the opportunity for neuroscientists to explore various synaptic features such as spine head volume, post synaptic density area, spine neck diameter, spine neck length and number of docked vesicles among the distinguishable bins and investigate potential role of sub cellular resources such as spine apparatus, smooth endoplasmic reticulum, polyribosomes and mitochondria on them.

Moreover, it is worth noting that our new study with empirical data without using any discrimination threshold other than utilizing the precision level can provide an estimate of the SNR for synaptic weights. It is approximately 0.51 in hippocampal area CA1 and calculated using the signal detection theory method presented in [1] and the setting number of Gaussian's to be equal 24. (Larger data sets are needed to have the robust estimation of SNR across various brain regions.)

Statistical Inference [7],[8] is one of the major disciplines that can assist us to further explore large 3DEM datasets in control [5], and LTP [6] conditions and infer the underlying mechanisms of learning and memory at synapse resolution.

References

[1] Bartol TM, Bromer C, Kinney J, Chirillo MA, Bourne JN, Harris KM, Sejnowski TJ. Nanoconnectomic upper bound on the variability of synaptic plasticity. Elife. 2015; 4, p.e10778.

[2] Ramón y Cajal S. The Croonian lecture.—La fine structure des centres nerveux. Proceedings of the Royal Society of London. 1894; 55(331-335):444–468.

[3] Kasai H, Ziv NE, Okazaki H, Yagishita S, Toyoizumi T. Spine dynamics in the brain, mental disorders and artificial neural networks. Nature Reviews Neuroscience. 2021; 22(7):407–422.

[4] Shannon, C.E., 1948. A mathematical theory of communication. The Bell system technical journal, 27(3), pp.379-423.

[5] Motta A, Berning M, Boergens KM, Staffler B, Beining M, Loomba S, Hennig P, Wissler H, Helmstaedter M. Dense connectomic reconstruction in layer 4 of the somatosensory cortex. Science. 2019; 366(6469).

[6] Samavat, M., Bartol, T.M., Bromer, C., Bowden, J.B., Hubbard, D.D., Hanka, D.C., Kuwajima, M., Mendenhall, J.M., Parker, P.H., Abraham, W.C. and Harris, K.M., 2022. Regional and LTP-Dependent Variation of Synaptic Information Storage Capacity in Rat Hippocampus. bioRxiv.

[7] Efron, B. and Hastie, T., 2021. Computer Age Statistical Inference, Student Edition: Algorithms, Evidence, and Data Science (Vol. 6). Cambridge University Press.

[8] Samavat, M., Luli, D. and Crook, S., 2016, November. Neuronal network models for sensory discrimination. In 2016 50th Asilomar Conference on Signals, Systems and Computers (pp. 1066-1073). IEEE.

A Appendix

A.1 Synaptic Information Storage Capacity

Spine morphology has substantial variation across the population and lifetime of synapses. Hebbian plasticity puts forth a causal relationship and transformation of information from the presynaptic site to the postsynaptic site by the adjustment of efficacy of synaptic transmission, or "synaptic weight." The pattern of synaptic weights in the ensemble of neural circuits allows us to define both information and the recipient of the message in the context of synaptic plasticity. The recipient of the message is the neural ensemble or the pattern of synaptic weights that store the message and read the message during the recall process, which is the reactivation of the synaptic weights in the memory trace. The amount of information is quantified by the distinguishability of synaptic weights which comprise the memory trace. Here "distinguishability" implies that the precision of synaptic weights play a significant role.

The synapse is the unit of information storage in an ensemble of neurons, and if the precision level of synaptic weights is low then the amount of information that can be stored per synapse and in the ensemble of the neurons will also be low. Because the spine head volume is highly correlated with synapse size, the precision of spine head volumes can be used to measure the distinguishability of the synaptic weights. High precision yields a greater number of distinguishable categories (i.e. states or clusters) for spine head volumes and hence higher information storage capacity.

Algorithm 2 Bootstrap Algorithm for Estimating the Standard Error of Median

Require: $n \geq 1$

Let $X_1, \ldots, X_n$ be some data and $\hat{\theta}_n = t\left(X_1, \ldots, X_n\right)$

For $b = 1, \ldots, B$

Simulate $X_1^{*(b)}, \ldots, X_n^{*(b)} \overset{iid}{\sim} F_n$ by sampling with replacement from $\{X_1, \ldots, X_n\}$

Evaluate $\hat{\theta}_n^{*(b)} = t\left(X_1^{*(b)}, \ldots, X_n^{*(b)}\right)$

$$\hat{\sigma}_{n,B}^2 = \frac{1}{B} \sum_{b=1}^{B} \left(\hat{\theta}_n^{*(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_n^{*(b)} \right)^2$$

Return the bootstrap estimate of standard error of median

$$\hat{\sigma}_{n,B}$$