

## Synaptic Information Storage Capacity Measured With Information Theory

**Mohammad Samavat**

*msamavat@ucsd.edu*

*Department of Electrical and Computer Engineering, Jacobs School of Engineering, University of California, San Diego, and Computational Neurobiology Laboratory, Salk Institute for Biological Sciences, La Jolla, CA 92037, U.S.A.*

**Thomas M. Bartol**

*bartol@salk.edu*

*Computational Neurobiology Laboratory, Salk Institute for Biological Sciences, La Jolla, CA 92037, U.S.A.*

**Kristen M. Harris**

*kharris@utexas.edu*

*Center for Learning and Memory and Department of Neuroscience, University of Texas at Austin, Austin, TX 78712, U.S.A.*

**Terrence J. Sejnowski**

*terry@salk.edu*

*Computational Neurobiology Laboratory, Salk Institute for Biological Sciences, La Jolla, CA 92037, U.S.A., and Department of Neurobiology, University of California, San Diego, La Jolla, CA 92093, U.S.A.*

Variation in the strength of synapses can be quantified by measuring the anatomical properties of synapses. Quantifying precision of synaptic plasticity is fundamental to understanding information storage and retrieval in neural circuits. Synapses from the same axon onto the same dendrite have a common history of coactivation, making them ideal candidates for determining the precision of synaptic plasticity based on the similarity of their physical dimensions. Here, the precision and amount of information stored in synapse dimensions were quantified with Shannon information theory, expanding prior analysis that used signal detection theory (Bartol et al., 2015). The two methods were compared using dendritic spine head volumes in the middle of the stratum radiatum of hippocampal area CA1 as well-defined measures of synaptic strength. Information theory delineated the number of distinguishable

---

Terrence Sejnowski, Kristen Harris, and Mohammad Samavat are corresponding authors.

**synaptic strengths based on nonoverlapping bins of dendritic spine head volumes. Shannon entropy was applied to measure synaptic information storage capacity (SISC) and resulted in a lower bound of 4.1 bits and upper bound of 4.59 bits of information based on 24 distinguishable sizes. We further compared the distribution of distinguishable sizes and a uniform distribution using Kullback-Leibler divergence and discovered that there was a nearly uniform distribution of spine head volumes across the sizes, suggesting optimal use of the distinguishable values. Thus, SISC provides a new analytical measure that can be generalized to probe synaptic strengths and capacity for plasticity in different brain regions of different species and among animals raised in different conditions or during learning. How brain diseases and disorders affect the precision of synaptic plasticity can also be probed.**

## 1 Introduction

---

In the late 19th century, Santiago Ramón y Cajal (1894) proposed that memories are stored at synapses and not through the generation of new neurons. Since then, there has been an extensive search for synaptic mechanisms responsible for learning and memory. Although synaptic plasticity is well established as an experience-dependent mechanism for modifying synaptic features, the precision of this mechanism is unknown. Connectomic studies are generally concerned with whether synapses exist between pairs of neurons, an important starting point; we are concerned here with quantifying synaptic strengths, the precision of synaptic plasticity, and the amount of information that synapses can store.

The existence of both intrinsic and extrinsic origins of variability and dysfunction of structural modulation (Kasai et al., 2021) motivates further exploration of the potential precision with which synaptic strengths can be adjusted. From an information theory perspective, there can be no information stored without precision. Indeed, the more precise that synaptic plasticity is, the more distinguishable synaptic strengths are possible and the greater the amount of information that can be stored at synapses. The synaptic strength is itself the information that is stored at a synapse, and this information is retrieved with subsequent synaptic transmission. Synapses are complex dynamical structures and “synaptic strength” encompasses intrinsic pre- and postsynaptic variables that can be modulated by synaptic plasticity. Among the variables are probability of presynaptic vesicular release, number of docked vesicles, number of postsynaptic receptors, and degree of short-term facilitation/depression, to name but a few. Thus, when a particular synapse is activated repeatedly, the resulting responses are expected to vary in amplitude in a way that is consistent with the synaptic strength (Kandaswamy et al., 2010; Klyachko et al., 2006).

Several studies have shown that synaptic strength is highly correlated with dendritic spine head volume (SHV; Harvey & Svoboda, 2007; Matsuzaki et al., 2004; Harris, 2020, reviewed in Yang & Liu, 2022). (Note that

the abbreviations used in the article are listed at the end of section 1.) Pairs of dendritic spines on the same dendrite that receive input from the same axon (SDSA pairs) occur naturally in the brain and are expected to have experienced the same activation histories (Harris & Sorra, 1998). Hence, the precision of synaptic strength can be estimated by measuring the difference between the spine head volumes of individual SDSA pairs. These measurements can then be used to calculate the number of distinguishable synaptic strengths in a particular population of synapses.

Signal detection theory was used in previous studies to estimate the number of distinguishable synaptic strengths (Bartol et al., 2015, Bromer et al., 2018). Here, Shannon information theory was applied to overcome the limitations of using signal detection theory. Information theory is based on the distinguishability of messages sent from a transmitter to a receiver in the presence of noise. In the case of synaptic strength, the postsynaptic dendrite or soma is the receiver that must distinguish the strengths of messages coming from discrete synaptic inputs. The distinguishability of the received messages depends on the precision with which synaptic plasticity sets the strength of individual synapses. The new approach calculates the number of bits of Shannon information stored in synaptic strengths and thus quantifies empirically the synaptic information storage capacity (SISC).

The analysis of precision begins by measuring the coefficient of variation (CV) of SDSA pairs, as in Bartol et al. (2015). The new method differs in that nonoverlapping binning is performed across the entire range of sampled dendritic spines with the bin width, or precision level, equal to the median CV of the SDSA pairs. 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. Finally, in the new SISC analysis, the number of distinct synaptic strengths converges toward the actual (population) number of strengths as the number of spine head volumes increases and the true shape and extent of the distribution are sampled. Statistics and information theory allowed us to quantify the precision and information capacity of the message. Earlier versions of this new method appeared in Samavat et al., 2022a, 2022b; Samavat, Bartol, Harris et al., 2022; Samavat, 2023; Samavat et al., 2024).

#### Abbreviations of Terms

Term	Abbreviation
Synaptic information storage capacity	SISC
Postsynaptic density	PSD
Spine head volume	SHV
Same dendrite same axon	SDSA
Coefficient of variation	CV
Kullback-Leibler divergence	KL divergence

## 2 Results

---

**2.1 Precision Analysis.** Precision is defined as the degree of reproducibility of a measurement; it is often mistaken for accuracy, which is defined as the deviation of the average measurement from a reference value (see Figure 4 in the appendix). The CV shown in equation 2.1 is a statistic that measures variation within a sample, defined by the standard deviation ( $\sigma$ ) (see equation 2.2, normalized by the mean of the sample ( $\mu$ ), making it a useful metric for measuring precision. Here we used ( $\sigma$ ) = 2 in equation 2.2 because we analyzed SDSA pairs.

$$CV = \frac{\sigma}{\mu} \quad (2.1)$$

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \quad (2.2)$$

Precision is a key factor for discovering the number of distinguishable synaptic strengths. Spine head volume provides a convenient estimate of synaptic efficacy as it correlates well with other measures of synaptic strength (Bartol et al., 2015). The precision is estimated from the CV of spine head volumes among the SDSA pairs (see Figure 1A). This value is then used in algorithm 2 in section 5.5 to assign each spine head to its appropriate bin (algorithm 2 provides the steps and details). To accomplish this, we first determined that the measurement error among four investigators was smaller than the intrinsic variability of the measured spine head volumes among the SDSA spine pairs (see Figure 1B). Then we could calculate the median CV, after calculating the CV of all SDSA pairs for the CA1 data, as an estimate of the precision of synaptic plasticity (see Figure 1D). The CVs among the SDSA pairs varied from pair to pair by over an order of magnitude within the data set, but there was no significant trend from the smallest to the largest spine head volumes. These outcomes suggest that the synaptic plasticity based on the coactivation history among small spines is equally precise for small and large spines. This median CV value establishes the precision level of SDSA pairs and applies to the other synapses of similar type (i.e., spines here) sampled from a particular brain region, not just spines in the SDSA pairs. Thus, the median CV is used for binning and calculation of the number of distinguishable synaptic strengths as described below.

**2.2 Measurement of the Distinguishable Strength Distribution.** The performance of the new method for measuring the distribution of distinguishable synaptic strengths was tested on the same CA1 data set that was previously analyzed with signal detection theory (Bartol et al., 2015). A total of 288 spine head volumes were fully contained within a  $6 \times 6 \times 5 \mu\text{m}^3$  CA1 neuropil volume (see Figure 2A). Previous signal detection theory revealed

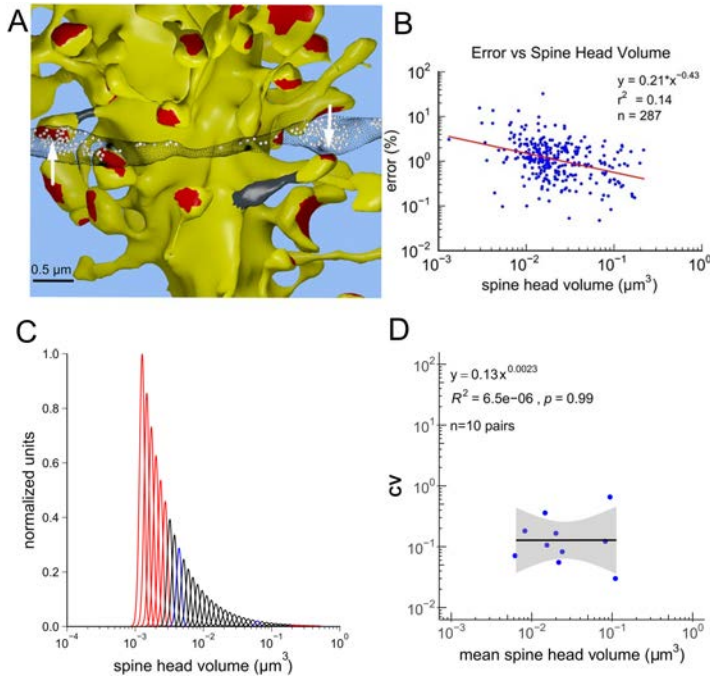


Figure 1: Same dendrite, same axon pairs and implication for information theory. (A) Visualization of a pair of spines (an SDSA pair) (with gray necks) from the same dendrite (yellow) and the postsynaptic density (PSD) associated area (red, indicated by white arrows) formed by the same axon (black stippling) with presynaptic vesicles (white spheres). (B) The 3D reconstructions of the spine head volumes of the five data sets were performed with the same protocol as that used for the CA1 data set in Bartol et al. (2015). Four individuals made hand tracings from the two-dimensional electron micrographs; after alignment, the automatic 3D reconstruction was made. The average measurement error is about 0.01 as shown in the figure. (C) Bartol et al. (2015) showed that 26 distinguishable gaussian distributions constructed with equal CV and overlap of  $\sim 31\%$ , corresponding to SNR of 1, spanned the range of spine head volumes of SDSA pairs (black and blue gaussians). Extension of this analysis to include in the data set with the complete range of measured spine head volumes (plus including a  $0.55 \mu\text{m}^3$  spine head found in another data set (Harris & Stevens, 1989)) is illustrated by red gaussian distributions. The blue gaussians have means located in gaps that lack sample data. Gaps are identified by methods developed in this article (see Figure 2B). All the gaussian distributions are constructed by using the median CV (fixed ratio for standard deviation over the mean for each gaussian distribution), a fixed amount of overlap between consecutive gaussians, a fixed range (min, max), and all with the same area under the curve (Bartol et al., 2015). (D) Uniform CV across the range of SDSA spine head volumes. The gray region is the 95% confidence interval for the regression line.

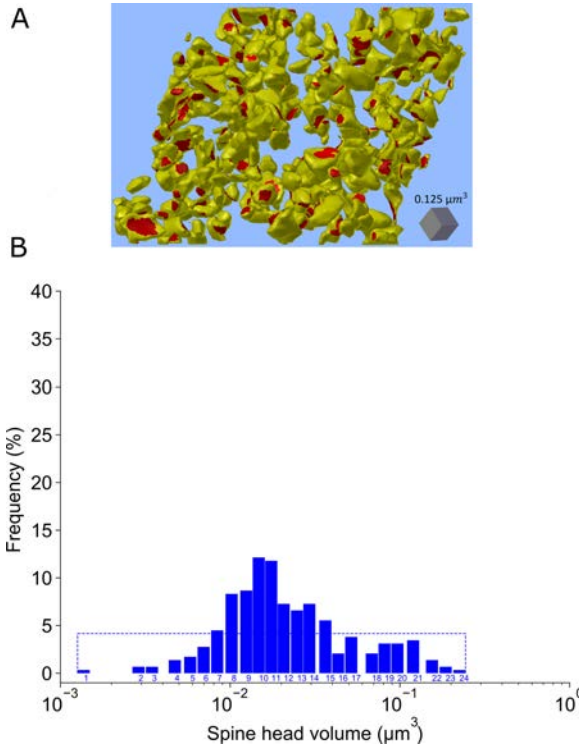


Figure 2: Binning the spine head volumes from the area CA1 data set. (A) Illustration of all 288 spine heads that are fully contained within the reconstructed volume. The postsynaptic density (PSD) associated area is displayed in red and spine head membranes in yellow. (B) In this histogram (one-dimensional bin diagram), the  $y$ -axis indicates the frequency of spine head volumes within each bin, and the  $x$ -axis indicates the value of spine head volumes in log scale. The width of each bar is equal to the median CV, which appears constant on a log scale. The starting value of each bar along the  $x$ -axis is derived from the smallest SHV of the given bin (see section 5 for details of Figure 2B generation.) The dashed rectangular box around the histogram is plotted to demonstrate the frequency of spine head volumes if the 288 spine head volumes were distributed uniformly among the 24 bins,  $pr = 1/N_S$ , where  $N_S$  is the number of sizes (number of bins).

26 distinguishable gaussian distributions with equal CV when assuming an overlap of 31% (see Figure 1C; the black represents gaussians). This amount of overlap is equivalent to assuming a signal-to-noise ratio (SNR) equal to 1 and a 69% discrimination threshold common in psychophysics (Schultz, 2007). Here, in the new method, we derived discrete entities by binning

the distribution of spine head volumes, where the width of the bins for the new analysis was determined by the precision analysis protocol with the median CV of  $0.12 \pm 0.046$  for the CA1 SDSA pairs (see Figure 1D), without any assumptions regarding the signal-to-noise ratio (see algorithms 1 and 2 in, respectively, sections 5.4 and 5.5). The outcome resulted in 24 distinguishable strengths (shown by SHV on log scale in Figure 2B and by bin number in appendix Figure 5). The number of spines in each bin is readily visualized (see appendix Figure 5), as are vacancies in the distribution (see Figure 2B). Since the histogram of spine head volumes is shown on a log scale in Figure 2B, the bins appear to be equal in width because of the constant value of CV, and the shape of the CV-width binned distribution replaces the skewed histogram of absolute sizes of spine-head volumes. The highest frequency occurs in bin #10, which contains 36 spine head volumes, and there appears to be a second peak at around bin 21 (see Figure 2B and appendix Figure 5). It is critical to note that the 24 bins found by the new method are conceptually different from the potential multiple modes derived from the distribution of spine head volumes.

**2.3 Shannon Information Storage Capacity of Synapses.** The concept of entropy,  $H(P)$ , comes from the field of thermodynamics and measures the number of possible configurations in a system. Shannon entropy is defined as the average of Shannon information (Shannon, 1948). Shannon entropy measures the amount of information in a set of distinguishable strengths, each of which has a certain probability of occurrence. With more information, it is possible to distinguish more strengths. The Shannon entropy of a discrete random variable is defined as follows:

$$H(X) = \mathbb{E} \left[ \log_2 \frac{1}{P_X(X)} \right] = \sum_{x \in \mathcal{X}} P_X(x) \log_2 \frac{1}{P_X(x)}. \quad (2.3)$$

The Shannon information per synapse was calculated from the frequency of spine head volumes in the bins and using equation 2.3, where each bin is considered one of the possible synaptic strengths of the system, that is, one of the possible “messages” that can be reliably sent and received. The observed Shannon entropy  $H(P)$  (expressed in  $\log_2$  “binary units” or “bits” of information) for the CA1 data set is listed in column 4 of Table 2, setting a lower bound for SISC at  $4.1 \pm 0.39$  bits in this data set. These data clearly demonstrate that the synapses are not simple one bit on/off switches as was widely assumed.

The closeness of the observed value of SISC to the theoretical upper bound on SISC is defined as the benchmark distribution having maximum entropy, which occurs when there is a uniform distribution across the possible sizes. The KL divergence between the two discrete probability distributions does not depend on the choice of scale. Under these circumstances



Table 1: Number of Distinguishable Strength, or Sizes, ( $N_S$ ) of Spine Head Volumes.

Data Set Type	Number of SDSA pairs	Median CV of SDSA set ( $\pm$ SEM)	Number of spine head volumes	Median spine head volumes ( $\mu\text{m}^3$ ) ( $\pm$ SEM)	SRF	$N_S$
CA1	10	$0.12 \pm 0.046$	288	$0.018 \pm 0.00091$	163	$24 \pm 5.40$

Notes: For columns 3 and 5, the term, SEM stands for standard error of median calculated using algorithm 1. SRF = scale range factor. For calculating the standard error on  $N_S$ , bootstrapping is used. For each of the 10,000 bootstrap samples, algorithms 1 and 2 are used with resampling the spine head volumes and resampling 10 CV values of the 10 observed SDSA pairs to calculate the binning threshold.

Table 2: Calculating the Entropy of Synaptic Strengths Based on the Calculated Frequency of Spine Head Volumes in Distinguishable Sizes.

Data Set	Mean of SHVs	CV of all SHVs	Shannon Entropy $H(P)$	Maximum Entropy $H(U)$	KL divergence $\text{KL}(P    U)$	Efficiency ratio $\eta$
CA1	$0.031 \pm 0.0021$	$1.045 \pm 0.066$	$4.1 \pm 0.39$	$4.59 \pm 0.37$	$0.49 \pm 0.068$	89%

Notes: For columns 2 to 6, standard error calculated with bootstrapping using algorithms 1 and 2 concurrently. SHV = spine head volume. Note that  $\eta = (1 - \text{KL}/\text{KL}_{\text{MAX}}) \times 100$  and stands for efficiency ratio.

equation 2.3 reduces to  $\log_2(N_S)$ , where  $N_S$  is the number of bins. Thus, the maximum number of bits of Shannon information is calculated as  $H(U) = \log_2(N_S)$ , the entropy of a uniform distribution among  $N_S$  sizes. In this CA1 data set,  $N_S = 24 \pm 5.40$  (see Table 1) and sets an upper bound of  $4.59 \pm 0.37$  bits for SISC (see Table 2, column 5), a value that is within the error range of the observed lower bound of  $4.1 \pm 0.39$  bits.

**2.4 KL Divergence Analysis.** Measurement of the distance between an observed distribution with a reference probability distribution can be done by KL divergence. A uniform distribution is the discrete probability distribution with maximum entropy when there is no constraint on the distribution except having the sum of the probabilities equal 1 for a fixed number of sizes. Formally, the KL divergence between the distribution of spine head volume bins ( $P$ ) and the uniform distribution of sizes ( $U$ ) is the difference between cross entropy of ( $P$ ) and ( $U$ ) and the entropy of ( $P$ ):  $[H(P, U) - H(P)]$ . The closeness of fit between the distribution of the distinguishable values to the maximum entropy distribution was determined with the fixed number of sizes (bins of spine head volumes) as the only known constraint on the distribution. When the distribution of the distinguishable values approaches the uniform distribution, the KL



divergence decreases and the Shannon entropy will approach maximum (see Figure 3).

The distribution of CA1 spine head volume bins was compared with a uniform distribution having 24 sizes (see Figure 2B and Table 2). The outcomes provide an upper bound on SISC for these sampled synapses but not necessarily for other synapses. We also calculated the ratio of the KL divergence values over the maximum value that KL divergence can possibly have ( $KL/KL_{MAX}$ ) where  $KL_{MAX}$  equals the entropy of a uniform distribution,  $H(U)$ . This ratio was then used to define the efficiency ratio  $\eta = (1 - KL/KL_{MAX}) \times 100$ . It measures the efficiency of the measured KL divergence value with respect to the maximum value it could hypothetically have (see Table 2, column 7). The higher the efficiency ratio, the more efficient is the use of distinguishable values for the storage of synaptic strength values across the population of synapses (see Figure 3). The KL divergence for this CA1 data set is  $0.49 \pm 0.068$ , a value suggesting that the number of distinguishable values approaches maximum entropy with an efficiency ratio of 89%. Thus, the reported distinguishable values are close to the maximum entropy.

This outcome provides a new statistic to assess information storage efficiency. This analysis for CV-width binned distribution of synaptic strength is more useful for comparisons between experimental conditions or between regions in diseased or normal brains. The precision level has a high impact on the efficiency ratio (on both KL in the numerator and  $KL_{MAX}$  in the denominator of  $\eta$ ). Namely, cases with similar precision levels may still result in an insensitivity of the efficiency ratio to reflect deviations from a uniformly binned distribution. The efficiency ratio of binned distributions or the information storage capacity can be the same, while the shapes of the spine-head size distributions are different. Thus, when evaluating the information storage capacity of synaptic strength, both the entropy of synaptic strength and the efficiency ratio should be determined.

### 3 Discussion

---

This article introduces a new analytical approach for determining SISC that has several advantages over our prior approach (Bartol et al., 2015). SDSA pairs are independent synapses, but because they are driven by the same input, they respond in a similar way and arrive at the same size due to the precision of similar underlying mechanisms in each synapse. Thus, the CV in the dimensions of these coactive pairs provides the basis for determining the range over which synaptic strength is essentially equal. There are more distinguishable strengths with greater Shannon information as the CV decreases due to higher precision. Bootstrapping analysis was used to estimate the standard error of the precision level, number of distinguishable strengths, and Shannon entropy and established SISC as a robust measure.

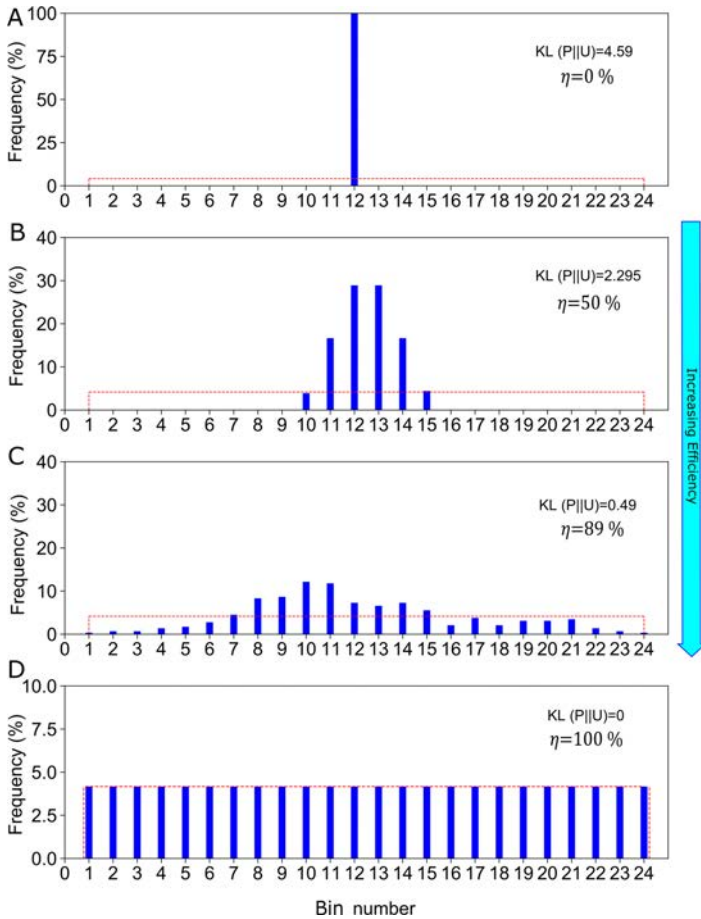


Figure 3: KL divergence of the sample spine head volumes approaches maximum entropy. (A) Hypothetical minimal KL divergence showing all spines in one bin (synaptic strength). (B) Hypothetical distribution with 50% KL divergence produces a distribution with six bins. (C) Actual spine head volumes of the CA1 data set distributed using algorithm 2 shows efficiency ratio  $\eta$  equals 89%. (D) Uniform distribution of the spines across the 24 sizes produces the minimal KL divergence and efficiency ratio equals 100%. The  $y$ -axis indicates the frequency of spine head volumes within each bin, and the maximum decreases from panels A to D to reflect the proportional decrease of spine frequency in each bin. Note the change in scale of the  $y$ -axis in panels A to D. The  $x$ -axis indicates the bin number that reflects a unique synaptic strength. The dashed rectangular box around the distributions of sizes in each graph illustrates the frequency of spine head volumes if the samples were uniformly distributed among the 24 bins as in panel D. KL divergence value and efficiency ratio  $\eta$  are shown in all scenarios.

The new method was tested on the same data from area CA1 of the hippocampus to compare it with the prior approach.

**3.1 Advantages of the New SISC Analysis.** There are several advantages to the new SISC method for assessing the storage capacity of synapses. Signal detection theory (Bartol et al., 2015) assumed that the width of the gaussian curves, based on the CV of the SDSA pairs, was distributed equally along the full range of sampled spine head volumes, without accounting for gaps in the distribution. Thus, the signal detection theoretical approach tended to overestimate the true number of distinct synaptic strengths because the distribution of the population was assumed to be continuous. In the new SISC analysis, the number of distinct synaptic strengths defined by the individual bins converges toward the true number of sizes as the number of spine head volumes increases and the true shape of the distribution is sampled. A second advantage is that the full population of spine head volumes, not just the SDSA pairs, is included in the analysis, greatly improving the statistical power of the estimate. A third advantage is that there are no free parameters in the estimate, unlike signal detection theory where the degree of overlap of the gaussians is a free parameter. A fourth advantage is that the new method is robust to the outliers. The largest spine head volume we found in an earlier data set in rat hippocampal area CA1 was 0.55 cubic microns (Harris & Stevens, 1989). It is worth mentioning that by adding this one value to the 288 CA1 spine head volumes would result in 25 distinguishable strengths, an increment of 1 size using SISC. In contrast, using the prior signal detection approach results in 39 distinguishable gaussians spanning the range, a 50% increase in the number of sizes. Thus, the previous method is not robust to outliers. Finally, the new method, based on information theory, allows access to the frequency of the bins, making it possible to compute the entropy of the distinguishable synaptic strengths, the number of modes in the CV-width distribution, and any gaps in the range of functionally distinguishable synaptic strengths.

**3.2 Information Theory of Synapses.** Information theory has been applied to analysis of spike trains (Dayan & Abbott, 2005) but has not been used at the level of synaptic strength. We have shown that the amount of information represented by synaptic strengths in neural circuits can be quantified by the distinguishability of synaptic dimensions. Here distinguishability fundamentally depends on the precision of the synaptic strength measurements and histories of activation. When the precision of synaptic strengths is low, the amount of information that can be stored in the ensemble of the neurons will also be low. Complete absence of precision implies a random process for setting synaptic strength and no information being stored at synapses. Because spine head volume is highly correlated with synapse size (Bartol et al., 2015), the precision of spine head volumes can be used to assess the distinguishability of the synaptic strengths. High

precision yields a greater number of distinguishable spine head volume bins and hence higher information storage capacity. The number of distinguishable strengths is not static but varies with the history of synaptic plasticity and is different in different parts of the brain. Thus, the amount of information that a population of synapses can store is not fixed but can be changed.

Comparison with the uniform distribution was made because it is the most conservative assumption when anatomical constraints on the spine head volume distribution are not known. When no constraint is applied (except the number of sizes being fixed), the uniform distribution has maximal entropy among every discrete distribution. This is why the uniform distribution for a fixed number of sizes is a lower bound on optimality and an upper bound for SISC. For example, the entropy of the observed synaptic strength distribution in area CA1 based on the probability of the 24 distinguishable sizes was 4.1 bits, while the maximal entropy for the same number of sizes is the uniform distribution with 4.6 bits of information. Thus, the KL divergence analysis provided a reliable measure of the closeness of the CV-width binned distributions of spine head volume to the maximum entropy distribution with an efficiency ratio of 89%. The amount of information stored across 24 sizes is equivalent to the maximum information that could hypothetically be stored across 17 sizes. The closer these two numbers are, the more efficient is the use of the distinguishable synaptic strengths.

If biological constraints were found at some point in the future, the maximal entropy distribution could be calculated under those constraints, and the distance between the empirical distribution (such as log normal) and maximal entropy distribution (with optimized parameters) could be determined. For example, if one calculates the maximal entropy distribution as a log normal distribution with well-constrained shape parameters (i.e., biological constraints on  $\mu$  and  $\sigma$ ), under those conditions the KL divergence would be different.

It is worth noting that synaptic activity, and hence SISC, is highly variable and changes during an animal's behavior. This article is concerned only with the optimality of information storage capacity based on proxies for synaptic strength as the unit of storage of information and the distribution of strengths at a single moment of time. This analysis sets a lower bound on the information storage capacity of synapses. Other sources of storage capacity that include biochemical and structural substrates across timescales could augment our estimate of long-term plasticity based on spine head sizes. Another question is how working memories are stored in neural circuits and how efficiently the synaptic strength is used in those codes. These issues are beyond the scope of this study and are open for future research.

#### 4 Conclusion

---

Information storage and coding in neural circuits have multiple substrates over a wide range of spatial and temporal scales. How information is coded

and its potential efficiency depend on how the information is stored. The efficiency of information storage was analyzed by comparing the distribution of synaptic strengths, as measured in spine head volumes, with the uniform strength distribution, the maximum entropy distribution. The outcomes reveal a near maximal efficiency in storing information across the distinguishable sizes of spine head volumes. These new measures of precision and efficiency can be used to assess the impact on SISC of other factors, such as pre- or postsynaptic dimensions and composition and variation across brain regions and species. It will be possible to determine effects on SISC of intrinsic differences or changes in the location or dimensions of key subcellular components, such as mitochondria, smooth endoplasmic reticulum, polyribosomes, endosomes, and perisynaptic astroglia. These opportunities to generalize SISC will result in greater understanding of the role of each component individually, and in concert, in determining the precision and efficiency of synaptic strength and plasticity. Ultimately, the outcomes should give new insight into how disrupted synapses result in cognitive decline during aging or as a consequence of neurological disorders.

## 5 Methods and Materials

---

**5.1 Summary of Methods.** The CA1 data set was published in Harris et al. (2015) and Kinney et al. (2013), and reanalyzed in Bartol et al. (2015) and this article.

**5.2 Reconstruction of a  $6 \times 6 \times 5 \mu\text{m}^3$  Volume of Hippocampal Area CA1 Neuropil.** The 3DEM data set that was used in this study was reconstructed from serial thin sections in the middle of the stratum radiatum of hippocampal area CA1 from an adult male rat (55–65 days old) (Mishchenko et al., 2010; Bourne et al., 2013). The dense reconstruction of  $6 \times 6 \times 5 \mu\text{m}^3$  of hippocampal neuropil reanalyzed in this study was originally processed as previously described in a study of the extracellular space (Kinney et al., 2013). In this data set, authors in Bartol et al. (2015) identified 10 axon-coupled synaptic pairs. To perform an accurate and robust geometric analysis of the dendrites, dendritic spines, and axons, the reconstructed surface meshes for artifacts were corrected by Bartol et al. (2015) to be computational-quality meshes as described elsewhere (Kinney et al., 2013; Edwards et al., 2014). The postsynaptic densities (PSDs) and presynaptic active zones (AZs) were identified in the serial section transmission electron microscopy (ssTEM) images by their electron density and the presence of closely apposed presynaptic vesicles. A total of 449 synaptic contacts were found in the dense reconstructed volume of neuropil. Bartol et al. (2015) excluded a number of synapses from the analysis if they were partially clipped by the edge of the data set (142) or were shaft synapses (20), leaving 287 valid synapses on dendritic spines in the dense model.

**Algorithm 1:** Bootstrap Algorithm for Estimating the Standard Error of Median.**Require:**  $n \geq 1$ Let  $X_1, \dots, X_n$  be some data and  $\hat{\theta}_n = t(X_1, \dots, X_n)$ For  $b = 1, \dots, B$ Simulate  $X_1^{*(b)}, \dots, X_n^{*(b)} \stackrel{\text{iid}}{\sim} F_n$  by sampling with replacement from  $\{X_1, \dots, X_n\}$ Evaluate  $\hat{\theta}_n^{*(b)} = t(X_1^{*(b)}, \dots, X_n^{*(b)})$ 

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

Return the bootstrap estimate of standard error of median

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

**5.3 Segmentation of the Dendritic Spines.** Bartol et al. (2015) used Blender's functionality, which is user extensible via a Python interface for creating add-ons. Bartol et al. created a Python add-on for Blender that enabled the selection of the mesh triangles of the dendrite corresponding to the spine head and whole spine of each individual spine. The add-on tagged each selected set of triangles with metadata for the spine name and geometric attributes of the head, whole spine, and neck as described below. The selection of the spine head was made by hand based on a standardized procedure in which the junction between the head and neck was visually identified as halfway along the concave arc as the head narrows to form the neck (see Figure 3 and the figure supplement 1A in Bartol et al., 2015). To select the whole spine, a similar visual judgment was made to locate the junction where the neck widens as it joins the dendritic shaft. Once the appropriate area was selected, the tool was designed to automatically create the convex hull of the selected region. The closed mesh formed by the Boolean intersection of the convex hull and the dendrite was used to determine the measured volume of the spine head or whole spine. The volume of the neck was calculated by taking the difference between these two measurements. Areas were computed from the selected regions for spine head and whole spine.

**5.4 Code Availability.** The codes for algorithms used in this study will be available in the following github link upon publication: <https://github.com/MohammadSamavat>.

**5.5 Standard Error of Median.** The standard error of median for the precision levels of each of the CA1 data sets' SDSA pairs is calculated with algorithm 1. The idea is to generate 1000 bootstrap samples of length  $n$ , each sampled from the  $n$  SDSA pairs with replacement, to estimate the standard error of median for the  $n$  SDSA pairs (see Table 1, column 3). The standard error of the median of spine head volumes follows the same procedure using algorithm 1.

**5.6 Binning Algorithm.** To construct the bins, spine head volumes are sorted from smallest to the largest. The first value (smallest value) is selected and the CV of that value and the remaining head volumes are calculated in a pairwise manner. The head volumes for which the calculated CV is below the threshold (the median value of the SDSA pairs CV) are assigned to the first bin and deleted from the pool of  $N$  spine head volumes. This procedure is repeated until the CV exceeds the median SDSA pairs CV and a new bin is formed. New categories are formed until all the remaining spine head volumes are assigned to a bin and the original vector of spine head volumes is empty (see algorithm 2 for details). It is guaranteed that the coefficient of variation between each random pair within each bin is less than the threshold value measured from the reconstructed tissue SDSA pairs. All spine head volumes are rounded to two significant digits for the display.

For Figure 2B, the  $y$ -axis shows the percentage of spine head volume counts in the respective bin. The  $x$ -axis shows the spine head volumes in  $\mu\text{m}^3$  on the log scale. The width of each bin is exactly the median value of the set of CV values plotted in Figure 1D (example: for bin-1 [ $x_1, x_2$ ],  $\text{CV}(x_1, x_2) = 0.12$ , where  $x_1$  is the smallest spine head volume in CA1 and  $x_2$  is a larger hypothetical head volume that has a CV of 0.12 with  $x_1$ ). The height of bin-1 is the number of spine head volumes in that range normalized to the total number of spine head volumes in this data set (288).

**5.7 Information and Entropy in Synaptic Strengths.** Shannon's information theory is the rigorous basis for a new method to quantify empirically SISC—that is, the number of bits of Shannon entropy stored per synapse. For the new method, first the precision of SDSA pairs was measured from the coefficient of variation (CV) of the pairs (illustrated in Figure 1D, also used in Bartol et al., 2015). The measured precision was then used as a binning threshold for nonoverlapping bin analysis (see Figure 2B and appendix Figure 2) using algorithm 2. This analysis yielded the number of distinguishable sizes of spine head volumes in the sample of reconstructed dendrites.

Next, the Shannon information per synapse is calculated by using the frequency of spine head volumes in the distinguishable values where each bin is a different message for the calculation of Shannon information. The maximum number of bits is calculated as the  $\log_2(N_S)$  where  $N_S$  is the number of sizes or categories that set an upper bound for the SISC.

When calculating the amount of entropy per synapse, the random variable is the synapse size, and the number of distinguishable values is the realization of a random variable for the occurrence of each size. The probability of the occurrences of each size is calculated by the fraction of the number of spine head volumes in each of the bins over the total number of spine head volumes in the reconstructed volume.



---

**Algorithm 2:** Binning Algorithm.
 

---

1: **function** Precision Calculation (Same Dendrite Same Axon pairs (SDSA),  $N$  pairs (a,b) of spine head volumes)

2:       **for**  $a, b \in \text{SDSA}[i]$   $i=1:N$  **do**

3:  $cv = \sigma(a,b)/\mu(a,b)$

4:  $cv[i] = cv$

5:       **end for**

6:       **return** {Precision\_Level= $Median(cv)$ }

7: **end function**

8: **function** binning spine head volumes ( $shv$  vector)

9:       Sort  $shv$  s.t.  $shv[i] < shv[i + 1]$

10:

11:       List\_of\_shcluster = NULL

12:       **while** Length( $shv$ )  $\neq 0$  **do**    $\triangleright$  Here we do the binning with the median value of SDSA pairs as a binning threshold calculated with the above function.

13:            $a = shv[1]$

14:       **for** every  $b \in shv[k]$  where  $K=1:Length(shv)$  **do**

15:

16:           **if**  $cv(a, b) < Median(CV)$  **then**

17:               Cluster  $\leftarrow b$

18:           **end if**

19:       **end for**

20:       List\_of\_shcluster[j]  $\leftarrow$  Cluster

21:        $shv = shv[-Cluster]$  (deleting the spine head volumes of bin  $j$  from the  $shv$  vector)

22:        $j=j+1$

23:       **end while**

24:       **return** {List\_of\_shcluster,  $N_s = j-1$ } (List of bins and number of bins, respectively)

25: **end function**

---

The information coding efficiency at synapses is measured by KL divergence to quantify the difference between two probability distributions—one from the categories of spine head volumes and the other from a corresponding uniform distribution with the same number of categories.

**5.8 Statistical Analysis.** The statistical analysis and plots were generated using Python 3.4 with NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), and Matplotlib (Hunter, 2007) and R package ggplot2 (Wickham, H., 2011).

For the precision analysis, we used the coefficient of variation (see equation 2.1) as a metric to show the precision level by calculating the ratio of standard deviation (see equation 2.2) over the mean of  $N$  joint synapses. Here  $N$  is 2 but can take higher values as up to 7 have been detected in previous studies. Since this is a sample from the unknown population of joint synapses, we used the corrected standard deviation formula with  $1/(N - 1)$  factor.

Bootstrapping was done by combining algorithms 1 and 2 to calculate the standard errors as explained in sections 5.4 and 5.5.

The standard errors of the entropy, efficiency ratio, maximum entropy for uniform distribution, and KL divergence (see Table 2, columns 2–5) are all calculated similarly using the bootstrapping technique explained in algorithm 1. (See Efron et al., 2021, for further information regarding bootstrapping for the calculation of standard error.)

## Appendix

## Accuracy and Precision

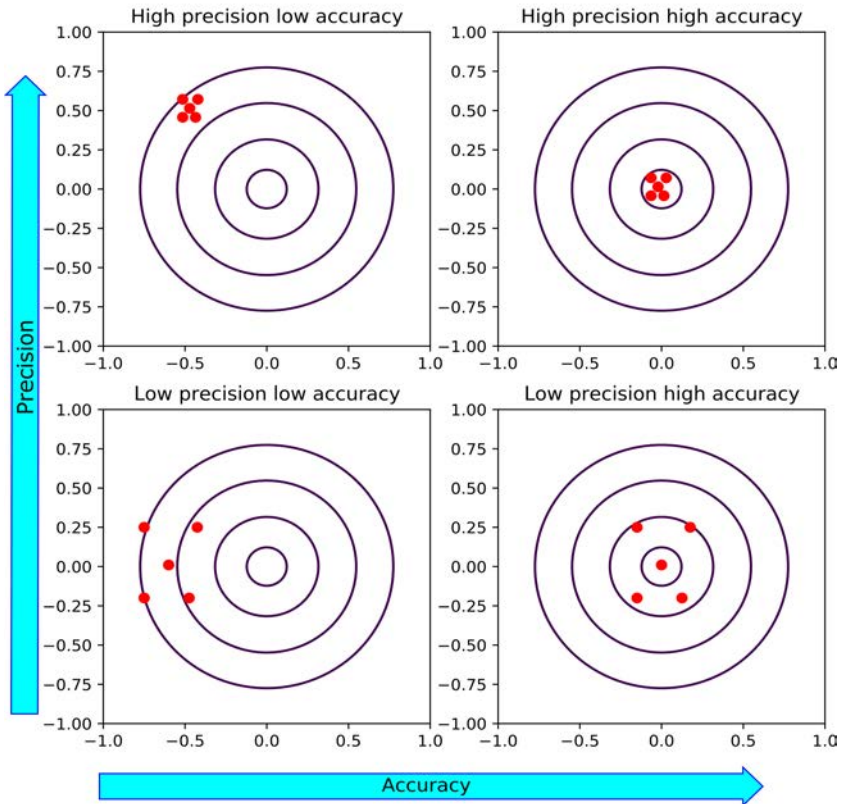
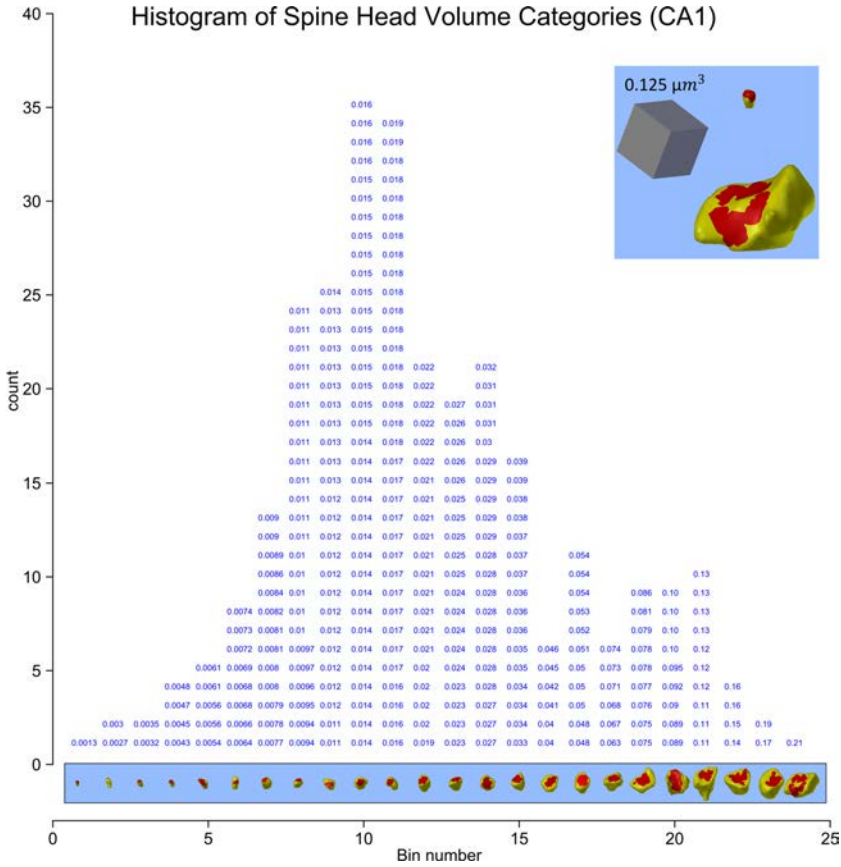


Figure 4: Precision and accuracy are different measures. When a process or system is repeated with the same input, the amount of variation in the output shows the precision level of the process. For accuracy, there is a reference frame with which the average value of measurements is compared. The graphs illustrate a low-precision and low-accuracy outcome (top left), low-precision and high accuracy (top right; the average of the positions is almost on the bull's eye), high precision and low accuracy (bottom left), and high precision and high accuracy (bottom right).



T.J.S., M.S. and T.J.S. wrote the article with contributions from T.M.B. and K.M.H.

## Acknowledgments

---

We thank Adel Aghajan and Wenxin Zhou for helpful discussion regarding information theory and bootstrapping analysis and Patrick Parker for editorial support. Also, we thank Cailey Bromer for her help in determining some of the dendritic spine head volumes. Feedback from the reviewer greatly helped us clarify the terminology and the conceptual framework in this article. This research was supported by NSF DBI-1707356; DBI-2014862; NSF 2219894 NSF IIS-2219979; NIH P41GM103712; NIH MH095980-07; NIH MH115556; NIH MH129066.

## References

---

- Bartol Jr., T. M., Bromer, C., Kinney, J., Chirillo, M. A., Bourne, J. N., Harris, K. M., & Sejnowski, T. J. (2015). Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife*, 4, e10778. 10.7554/eLife.10778
- Bourne, J. N., Chirillo, M. A., & Harris, K. M. (2013). Presynaptic ultrastructural plasticity along CA3→CA1 axons during long-term potentiation in mature hippocampus. *Journal of Comparative Neurology*, 521(17), 3898–3912. 10.1002/cne.23384
- Bromer, C., Bartol, T. M., Bowden, J. B., Hubbard, D. D., Hanka, D. C., Gonzalez, P. V., . . . Harris, K. M. (2018). Long-term potentiation expands information content of hippocampal dentate gyrus synapses. *Proceedings of the National Academy of Sciences*, 115(10), E2410–E2418. 10.1073/pnas.1716189115
- Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.
- Edwards, J., Daniel, E., Kinney, J., Bartol, T., Sejnowski, T., Johnston, D., . . . Bajaj, C. (2014). VolRoverN: Enhancing surface and volumetric reconstruction for realistic dynamical simulation of cellular and subcellular function. *Neuroinformatics*, 12, 277–289. 10.1007/s12021-013-9205-2
- Efron, B., & Hastie, T. (2021). *Computer age statistical inference, student edition: Algorithms, evidence, and data science* (Vol. 6). Cambridge University Press.
- Harris, K. M. (2020). Structural LTP: From synaptogenesis to regulated synapse enlargement and clustering. *Current Opinion in Neurobiology*, 63, 189–197. 10.1016/j.conb.2020.04.009
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., . . . Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. 10.1038/s41586-020-2649-2
- Harris, K. M., Spacek, J., Bell, M. E., Parker, P. H., Lindsey, L. F., Baden, A. D., . . . Burns, R. (2015). A resource from 3D electron microscopy of hippocampal neuropil for user training and tool development. *Scientific Data*, 2(1), 1–19. 10.1038/sdata.2015.46

- Harris, K. M., & Stevens, J. K. (1989). Dendritic spines of CA 1 pyramidal cells in the rat hippocampus: Serial electron microscopy with reference to their biophysical characteristics. *Journal of Neuroscience*, 9(8), 2982–2997. 10.1523/JNEUROSCI.09-08-02982.1989
- Harvey, C. D., & Svoboda, K. (2007). Locally dynamic synaptic learning rules in pyramidal neuron dendrites. *Nature*, 450(7173), 1195–1200. 10.1038/nature06416
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(03), 90–95. 10.1109/MCSE.2007.55
- Kandaswamy, U., Deng, P. Y., Stevens, C. F., & Klyachko, V. A. (2010). The role of presynaptic dynamics in processing of natural spike trains in hippocampal synapses. *Journal of Neuroscience*, 30(47), 15904–15914. 10.1523/JNEUROSCI.4050-10.2010
- Kasai, H., Ziv, N. E., Okazaki, H., Yagishita, S., & Toyoizumi, T. (2021). Spine dynamics in the brain, mental disorders and artificial neural networks. *Nature Reviews Neuroscience*, 22(7), 407–422. 10.1038/s41583-021-00467-3
- Kinney, J. P., Spacek, J., Bartol, T. M., Bajaj, C. L., Harris, K. M., & Sejnowski, T. J. (2013). Extracellular sheets and tunnels modulate glutamate diffusion in hippocampal neuropil. *Journal of Comparative Neurology*, 521(2), 448–464. 10.1002/cne.23181
- Klyachko, V. A., & Stevens, C. F. (2006). Temperature-dependent shift of balance among the components of short-term plasticity in hippocampal synapses. *Journal of Neuroscience*, 26(26), 6945–6957. 10.1523/JNEUROSCI.1382-06.2006
- Matsuzaki, M., Honkura, N., Ellis-Davies, G. C., & Kasai, H. (2004). Structural basis of long-term potentiation in single dendritic spines. *Nature*, 429(6993), 761–766. 10.1038/nature02617
- Mishchenko, Y., Hu, T., Spacek, J., Mendenhall, J., Harris, K. M., & Chklovskii, D. B. (2010). Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron*, 67(6), 1009–1020. 10.1016/j.neuron.2010.08.014
- Ramón y Cajal, S. (1894). The Croonian Lecture: La fine structure des centres nerveux. *Proceedings of the Royal Society of London*, 55(331–335), 444–468.
- Samavat, M. (2023). *A mathematical theory of synaptic information storage capacity*. PhD diss., University of California, San Diego.
- Samavat, M., Bartol, T. M., Bromer, C., Bowden, J. B., Hubbard, D. D., Hanka, D. C., . . . Sejnowski, T. (2022a, November). Shannon information of synaptic weights post induction of long-term potentiation (learning) is nearly maximized. In *Proceedings of the NeurIPS 2022 Workshop on Information-Theoretic Principles in Cognitive Systems*.
- Samavat, M., Bartol, T. M., Bromer, C., Bowden, J. B., Hubbard, D. D., Hanka, D. C., . . . Sejnowski, T. J. (2022b). *Regional and LTP-dependent variation of synaptic information storage capacity in rat hippocampus*. bioRxiv:2022-08.
- Samavat, M., Bartol, T. M., Bromer, C., Hubbard, D. D., Hanka, D. C., Kuwajima, M., . . . Harris, K. M. (2024). *Long-term potentiation produces a sustained expansion of synaptic information storage capacity in adult rat hippocampus*. bioRxiv:2024-01.
- Samavat, M., Bartol, T. M., Harris, K., & Sejnowski, T. (2022, November). Using Shannon Information to Probe the Precision of Synaptic Strengths. In *Proceedings of the NeurIPS 2022 Workshop on Information-Theoretic Principles in Cognitive Systems*.
- Schultz, S. (2007). Signal-to-noise ratio in neuroscience. *Scholarpedia*, 2, 2046. 10.4249/scholarpedia.2046

- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. 10.1002/j.1538-7305.1948.tb01338.x
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . van Mulbregt, P. (2020). Fundamental algorithms for scientific computing in Python and SciPy 1.0 contributors. *SciPy 1.0. Nat. Methods*, 17, 261–272. 10.1038/s41592-019-0686-2
- Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2), 180–185. 10.1002/wics.147
- Yang, Y., & Liu, J. J. (2022). Structural LTP: Signal transduction, actin cytoskeleton reorganization, and membrane remodeling of dendritic spines. *Current Opinion in Neurobiology*, 74, 102534. 10.1016/j.conb.2022.102534

---

Received July 31, 2023; accepted January 2, 2024.