LOCALIST NEURAL PLASTICITY IDENTIFIED BY MUTUAL INFORMATION

Gabriele Scheler
Carl Correns Foundation for Mathematical Biology
Mountain View, CA 94040
gscheler@gmail.com

Johann M. Schumann
Carl Correns Foundation for Mathematical Biology
Mountain View, CA 94040
johann.schumann@gmail.com

June 7, 2024

ABSTRACT

We present a model of pattern memory and retrieval with novel, technically useful and biologically realistic properties. Specifically, we enter n variations of k pattern classes ($n * k$ patterns) onto a cortex-like balanced inhibitory-excitatory network with heterogeneous neurons, and let the pattern spread within the recurrent network. We show that we can identify high mutual-information (MI) neurons as the major information-bearing elements within each pattern representation. We employ a simple one-shot adaptive (learning) process focusing on MI neurons, as well as groups of inhibitory neurons. We can then reproduce each pattern class by activating only the high MI neurons. Such 'localist plasticity' has high efficiency, because it requires only few selected adaptations for each pattern. Specifically we store $k = 10$ pattern classes of size $s = 400$ in a 1000/1200 neuron network. We retrieve patterns one at a time, such that the whole network represents this pattern. We assess the quality of the representation (a) after entering the full pattern into a naive network and (b) after adaptation (learning) and stimulation of MI neurons (retrieval). The pattern 'unfolds' over the recurrent network with high similarity, albeit compressed, with respect to the original input pattern. We discuss the distribution of neuron properties in the network, and find that an initial Gaussian or uniform distribution changes into a more heavy-tailed, lognormal distribution during the adaptation process. The remarkable result of our experiments is that we are able to achieve reliable pattern recall by stimulating only high information neurons. This work has interesting technical applications, but its main motivation is to provide a biologically-inspired model of cortical memory.

***K*eywords** Neural Plasticity · Information Theory · Intrinsic Plasticity · Hebbian learning · Sparse Coding · Distributed Coding · Memory · Recall · Extinction · Reconsolidation · Engram · Synaptic Plasticity

1 Introduction

Storing patterns and achieving sequence-independent recall is a problem for neural network models. We are interested in exploring the concept of localist ensemble memory in cortical networks. From biology we know that there is both synaptic and intrinsic memory, which have a 'hidden' component in cell-internal memory [14]. Therefore we want to explore the concept of storing memory by local plasticity in neurons within ensemble -like neuronal groups. For this purpose we define a realistic excitatory-inhibitory network with recurrent interactions and use simple, visually-defined patterns ([7]) as inputs to the network. We then examine the representations that develop on a naive network, where 'naive' means without previously stored patterns, and in the absence of plasticity. Representations are classified by a downstream machine learning mechanism. This machine learning mechanism stands for a, probably subcortical, interpretation, which is not explicitly modelled (cf. Fig. 1).

This construction is a suitable model for cortical networks [6, 13], it is also reminiscent of LSN [8] and ESN [4] models. To implement plasticity, we perform a information-theoretic analysis over neurons and patterns to find the neurons with the highest mutual information (MI) for each class of patterns. Such high MI response neurons form spontaneously in the representations. We analyze high MI neurons. It turns out that while absolute MI values for neurons differ across

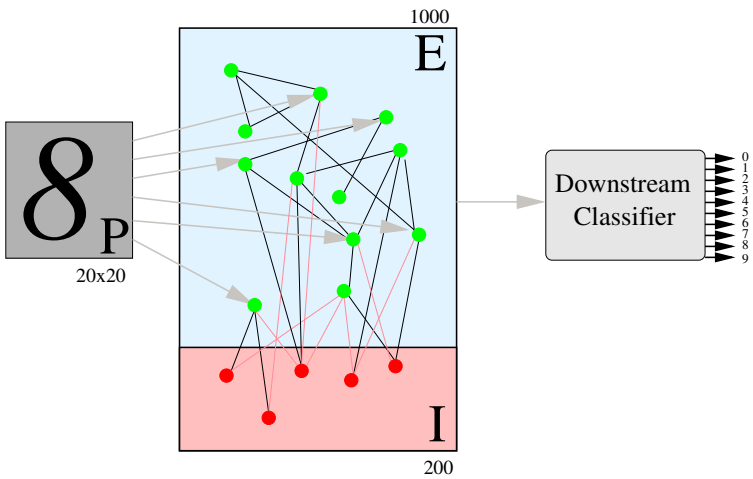

Figure 1: Architecture of the system: patterns from a field of 20x20 excitatory P neurons are 'loaded' into a network of E neurons. The network of E (N=1000) and I neurons (N=200) generates a representation. The representation is identified by a downstream (e.g. subcortical) classifier.

classes, and the number of high MI neurons above a threshold is different for each class, we can rank the highest MI neurons for each pattern class. We select only the high MI neurons to use for plasticity. The idea is to 'compress' the pattern information into the high MI neurons' intrinsic properties and adjacency network, i.e. perform localist plasticity. Then by stimulating only these neurons, we can recall neural representations and we achieve high precision pattern recall [11].

We also look at inhibitory neurons which play an important role in defining ensembles in real cortical networks. When an ensemble is activated, inhibitory plasticity guarantees that the activated ensemble is less inhibited, while inhibition remains strong for all other neurons. When there is low or no overlap in the high MI neurons for each pattern class, this localist learning means that there is a disjunctive set of adaptations for each pattern set. Dissimilar pattern classes are stored free of interference. The goal of our experiments is to ensure that pattern information can be retrieved by stimulation of a few, high MI neurons alone - even though the patterns were learned from a full input representation.

2 Methods

2.1 Neurons and Network Structure

The network model is initialized as a fully connected model with E (=1000) excitatory neurons and I (=200) inhibitory neurons, plus P (=400) pattern input neurons, which are directly linked to P of E excitatory neurons, in order to 'load' patterns of length P into the network. The connections between the P neurons and corresponding E neurons is strong to ensure solid transfer of activity. Synaptic connections within the network are of type NMDA, AMPA and GABA-A and are modeled as in [13].

Neurons are modeled as spiking neural models as in [3] with an equation for the membrane model v and an equation for the gating variable u (Eq 1), such that v is set back to a low membrane potential $v := c$; and the gating variable u is increased by an amount d ($u := u + d$), when a neuron spikes (at $v(t) = \theta$ and $\theta = 30mV$). For excitatory neurons, parameters are variable in order to capture different types of neurons and generate different distributions of intrinsic excitability, cf. Table 1. For excitatory neurons, parameters a and b are varied in this model [13], which results in different intrinsic excitability or gain of a neuron. The gain is captured by $g = 25,000a$. We measure the spontaneous neural activity with a background input of around $0.15nA$ in order to obtain baseline rates. These parameters also allow for intrinsic plasticity.

$$\begin{aligned} \dot{v} &= 0.04v^2 + 5v + 140 - u - I_{syn} \\ \dot{u} &= a(bv - u) \end{aligned} \tag{1}$$

Table 1: Parameter ranges for excitatory and inhibitory neurons

Parameter	E (lognormal)	E (normal)	E (uniform)	I	P
a	$0.001 \ldots 0.039$	$0.006 \ldots 0.034$	0.015	0.3	0.1
b	$0.01 \ldots 0.38$	$0.06 \ldots 0.34$	0.15	0.1	0.1
c	-70	-70	-70	-70	-70
d	3	3	3	2	3
$gain$	25-785	215-795	375	15000	2000
$baseline\ rate\ (Hz\ at\ 0.15nA)$	0.4-8	3-7	5	20	6

For inhibitory neurons, the parameters are $a = 0.3, b = 0.1, c = -70mV, d = 2, \theta = 30mV$. This results in neurons firing around 20Hz. For the input pattern neurons P, the values are similar with a lower gain and longer delay ($a = 0.1, b = 0.1, c = -70mV, d = 3$).

We initialize excitatory neurons with a Gaussian distribution over its gain and baseline rate, such that we use a mean of $\mu = 5$ and distribution $\sigma^2 = 0.96$ for the rate (Fig. 2B,D). Attested values for cortical neurons in mice, which are for a lognormal distribution of excitability over E neurons in the network, are rate distributions with $\mu = 4.96$ and $\sigma^2 = 0.31$ ([12]) (Fig. 2A,C).

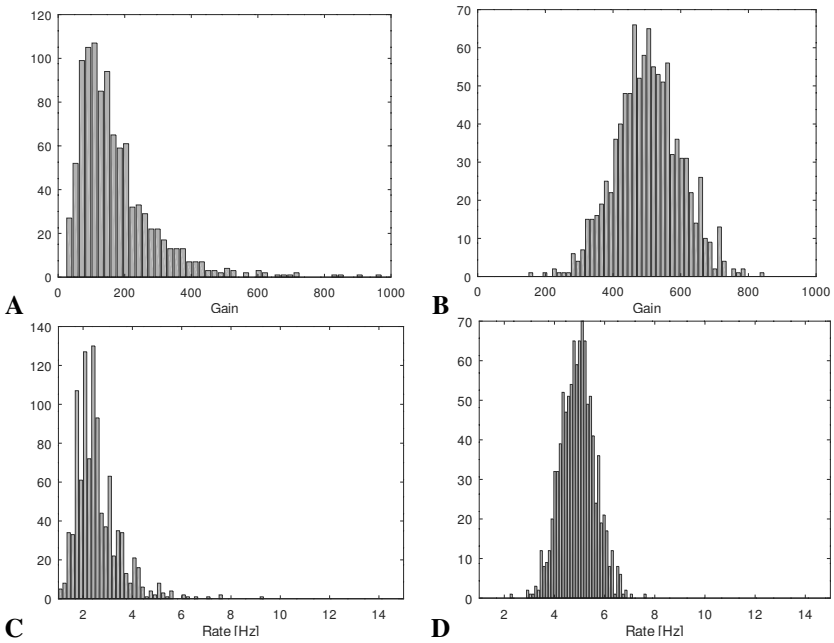

Figure 2: Lognormal and Gaussian distribution of gains (A, B) and baseline rates (C, D) for excitatory neurons as shown in Table 1.

Initially, the network has full connectivity for E-E, E-I and I-E synapses. There are no I-I connections. Synaptic strength was set to achieve baseline spontaneous spike rates in the network with a background input ($\sim 0.15nA$). Accordingly, the network is initialized with a Gaussian distribution of synaptic strength for both AMPA and NMDA connections between E-E ($\mu = 0.0015, \sigma = 0.00027$) and E-I ($\mu = 0.014, \sigma = 0.00024$). For GABA-A connections, we use a Gaussian distribution with $\mu = -0.5$ and $\sigma^2 = 0.1$.

There is a signal delay between neurons, which is randomly distributed:

- AMPA: $10 \ldots 30ms$
- NMDA[1],$10 \ldots 15ms$
- GABA-A: $5 \ldots 15ms$

[1]NMDA has intrinsic delays of about 120ms

We do not use this parameter during plasticity.

2.2 Patterns

In order to use a simple set of variations of patterns which can be easily classified into classes, we used the MNIST database [7] for 10 handwritten digits, with 50 variations for each digit. The format was an integer vector (for the grayscale values) of length 400 (10x50=500 patterns).

We load patterns into the E+I (=1200) processing network via P (=400) pattern neurons which have a single excitatory connection (AMPA) to P excitatory neurons in the processing network. We apply one input pattern of length 400 at a time for 300ms. We measure the resulting neural representation for the pattern after end of application. Spike rates recorded from excitatory neurons (E=1000 neurons) for a length of about 300 ms are regarded as the neural representation of the pattern. After that time, the neural representation fades away in our set-up.

2.3 Classification of Representations

For the analysis, we use machine learning classification algorithms which are autoselected by AutoGluon [1]. The classifier is presented with the neural representations for the $50 * 10 = 500$ patterns. The classifier's task is to recognize the correct digit for each pattern. The original neural representations for patterns are labeled, and the algorithm is trained to classify them ('supervised learning'). The test consists in recognizing representations which are derived from local stimulation of a few neurons after neural plasticity. These 'recalled' neural representations are derived from stimulation of m high MI neuron sets where $m = 5, 10, 20$. These representations are recognized by the classifier without any labels, based on their similarity with the original representations in an 'unsupervised' manner.

2.4 Information Theoretical Analysis

Each pattern class s is presented with equal probability $p(s) = 1/N_{digits})$, where $N_{digits} = 10$.

The response for each neuron is the rate response r. To simplify, we distinguish 3 rate responses (A,B,C), with f^s_{rel} as the neuron's firing rate for digit pattern s relative to its baseline rate.

$$r = \begin{cases} A & 80\% < f^s_{rel} < 120\% \\ B & f^s_{rel} \leq 80\% \\ C & f^s_{rel} \geq 120\% \end{cases} \tag{2}$$

Now for each collection of 50 variations of a digit, each neuron has a probability for each of the three responses. For instance, for digit 5, neuron j may have $A = 0.8, B = 0.1, C = 0.1$, or 40 times A, 5 times B and 5 times C. Assume $p(A) = 0.5, p(B) = 0.25, p(C) = 0.25$. (p(A)=0.5 means 50% A for $N_{patterns}$). Assume $p(digit5) = 0.1$. Since A is the 'neutral' option of a fluctuating value in an intermediate range, neuron j would not be very specific in recognizing pattern 5. Its mutual information (MI^5_j) would be $(0.054 + (-0.01322) + (-0.01322) = 0.0278\, bit)$.

We calculate the mutual information (MI) between each digit s (10 with 50 variations) and rate response r for each neuron n as in (Eq. 3):

$$\begin{aligned} MI^s_n &= \textstyle\sum_{r \in \{A,B,C\}} p(s,r) * \log_2 \frac{p(s|r)}{p(s)} = \\ & p(s,A) * \log_2 \frac{p(s|A)}{p(s)} \\ & + p(s,B) * \log_2 \frac{p(s|B)}{p(s)} \\ & + p(s,C) * \log_2 \frac{p(s|C)}{p(s)} \end{aligned} \tag{3}$$

Figure 4 shows the overall MI for each excitatory neuron of the network after processing 500 patterns without any plasticity. The neurons are arranged according to their intrinsic baseline firing rate (right-to-left, and top-to-bottom).

We can also calculate the mutual information for each neuron with respect to each digit pattern. For each neuron, in this case, we have 10 values for MI^s_n, since there are $s = 1..10$ digit patterns. We then rank for each digit pattern the m highest MI^s_n values. This is shown in Fig. 5A-C. If a neuron appears as a high MI neuron for more than one pattern, it is marked by a black dot.

2.5 Plasticity

The learning rule is based on a neuron's information content (MI). The system has access to the mutual information of a neuron for patterns. A neuron must have high MI (for some pattern) in order to undergo permanent plasticity. For each

pattern, we rank neurons by mutual information, and select the m highest neurons which are active for this pattern, for $m = 5, 10, 20$ (Fig. 5). We will discuss that the specific size of m shows little difference in the behavior of the system.

We use two types of single-shot update rules, for intrinsic plasticity, and for synaptic plasticity.

- Intrinsic plasticity: For the adaptation of high MI neuronal parameters a, b, we use an exponential update rule:

$$a := a \cdot e^{(a-a_{min})/(\lambda(a_{max}-a_{min})))}$$

$$b := b \cdot e^{(b-b_{min})/(\kappa(b_{max}-b_{min})))}$$

  With this rule, we strengthen the excitability in a non-linear fashion with stretch factors $\lambda = 3$ and $\kappa = 3$. It turned out that a proportional update rule $(a \cdot (1 + \lambda))$ did not lead to desired results (s. Results).

- Synaptic connections:

  – For high MI neurons only, their postsynaptic connections are updated according to the response of the input neurons i: if the response is of type C (Eq 2), the connection will be strengthened

$$w_{ni} := w_{ni}(1 + \alpha_1)$$

    α_1 is a value of ~ 0.4, which derives from the fact that a single update of synaptic value is sufficient, and each neuron receives many unstrengthened synapses. The parameter was set manually such that recall was optimized. Here a proportional update was sufficient. For responses A and B, w_{ni} is left unchanged.

  – Presynaptic connections are updated according to the response of the output neurons o: if its response is of type C, its input synaptic connection is strengthened:

$$w_{ni} := w_{ni}(1 + \beta_1)$$

    Here $\beta_1 \sim 400$ is similar to α_1 and shifts the weight to a much higher value.
    In this case, synaptic connections for output neurons with responses A or B are weakened: $w_{no} = w_{no}(1 - \beta_2)$ with $\beta_2 = 0.9$. However, the updates for weakened synapses are not crucial, they mainly serve to better calibrate the overall distributions

- Tuning inhibition:

  - For each pattern, a few inhibitory neurons i are selected to receive strengthened connections from pattern-specific high MI neurons n. ($\gamma_1 = 200$):

$$w_{ni} := w_{ni} * (1 + \gamma_1)$$

  - The GABA-A connections to excitatory neurons o with no or negative response to the pattern (responses A or B) are also strengthened in order to suppress neurons which do not participate in coding for the pattern ($\gamma_2 = 8000$):

$$w_{io} := w_{io} * (1 + \gamma_2)$$

  For inhibition, proportional strengthening is sufficient. As before, a high parameter value selectively strengthens a few synapses and in this way 'carves' a trace of the neural representation into the network.

In this case, all parameters were hand-tuned to achieve appropriate results for localist recall. Automatic setting and re-calibration of parameters could be added and might uncover more parameter combinations with acceptable solutions ('regimes'). It is most likely that calibration by homeostatic plasticity, which occurs from time to time, would also stabilize the parameter regimes, and make some weakening rules unnecessary.

The intrinsic plasticity rules mean that a neuron which carries much information for a pattern will increase its intrinsic frequency. The pattern will leave a trace at that neuron. We have seen that intrinsic frequency and high MI are initially not correlated (Fig. 4). We have also seen that for small sets of high MI neurons ($m/P < 0.05$), there is limited overlap in the identity of the high MI 'control' neurons for each pattern (Fig. 5), and therefore the update rules are mostly disjoint (separate). A high MI neuron's input synapses are updated only when they receive positive pattern input from neurons i. All of its output synapses are updated. We also update inhibitory parameters in a restricted way. This means that we have few, local updates which furthermore are applied only once. This makes plasticity highly efficient.

3 Results

3.1 Neural Representations and Classification

We modeled the network similar to a cortical network, which receives pattern information from thalamic input neurons as well as from other cortical and subcortical sources. Accordingly, we used a vector representation of visual patterns for perceptual input and added weak fluctuating background noise (of $1.5nA$). We initialized the network with random (Gaussian) distributions of neural parameters such that the network had heterogeneous intrinsic strengths and synaptic weights. This naive network, without any plasticity, developed neural representations for, in this case, 500 pictures from 10 classes. The neural representations, which lasted $\sim 300ms$, could be analyzed for information content of individual neurons using the tool of mutual information (R+EF). It turned out, like we noticed before [12], that a few neurons with high mutual information develop for each input pattern class. These neurons represent pattern classes in the manner of 'symbolic' or 'control' neurons. They stand for the whole pattern. We developed forms of plasticity to underscore this form of representation. By stimulating these "high-MI" neurons, we aim to recall pattern information.

Neural representations in each case were analyzed by a downstream classifier. The classifier was obtained from a machine learning library ([1]), which selected the optimal classifier (here: the 'Weighted Ensemble L2' algorithm). We assume that the interpretation and identification of a cortical neural representation in the brain is also handled by downstream, presumably sub-cortical, mechanisms. For the naive representations, without any plasticity applied, we could show that the projection of the 400 pixel input pattern onto the 1200 recurrent cortical network model resulted in recognizable neural representations. Figure 3 shows neural representations for digit "6".

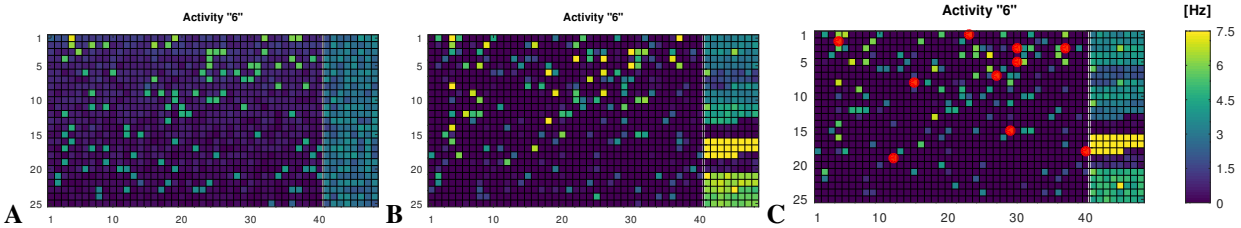

Figure 3: Neural rate representation for digit patterns. Excitatory neurons on the left side of each panel, inhibitory neurons on the right. (A) Activity after presentation with a pattern "6" before plasticity, (B) The same after application of plasticity rules. (C) Activity after stimulation of 10 high MI neurons (red). Patterns (A-C) were all classified correctly as representing digit "6". The high inhibitory activity contributes to specificity of representation.

With this classifier, we classify the representation on the network

1. before application of plasticity, when the pattern is presented to the naive network through input neurons P (Figure 3A),
2. after plasticity, when pattern is presented to the learned network through input neurons P (Figure 3B), and
3. after plasticity, when the representation is 'recalled' from stimulation of the 'concept' high-MI neurons alone (Figure 3C).

Before application of plasticity, we could obtain an accuracy of 0.98 through a supervised learning regime with only 10/500 patterns misclassified. The task could be learned by the network. After application of plasticity, patterns were applied in unsupervised fashion, and the accuracy was 0.874 for the 500 patterns (62/500 patterns were misclassified). The plasticity rules applied do not interfere much with the ability of the network to classify patterns derived from neural input. They do degrade the accuracy, but as we will see, there is a significant advantage to be gained from this form of plasticity.

3.2 Mutual Information Analysis of Representations

When we analyzed the neural representations for information content, by measuring mutual information (MI) between each neuron and the 10 digit input patterns, we could see that there is **no or weak** correlation between high MI and high baseline rate overall in the network (Figure 4). While the mutual information (for 500 patterns) in the network after plasticity is of course higher, the correlation between MI and baseline rate is $r = -0.09$ (Pearson), and -0.10 (Spearman) before plasticity, it is $r = 0.048$, 0.06 after plasticity.

Our aim was to identify high MI neurons for each digit which could be regarded as 'symbolic' abstractions of the whole pattern. I.e. we were looking for each digit pattern for the m neurons which are highest ranked by mutual information.

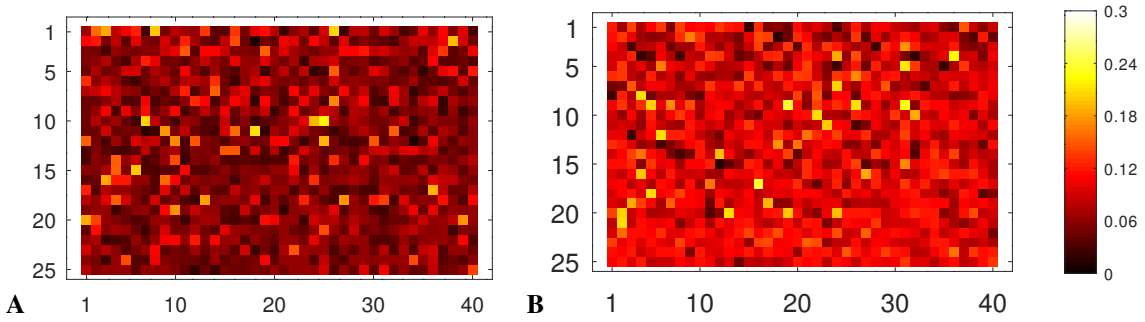

Figure 4: No correlation between baseline firing rate and MI. 1000 E neurons are shown on a 25x40 grid. The numbering is ordered by the baseline firing rate of each neuron (left to right and top to bottom). MI is calculated for all 500 patterns. A: before plasticity is applied, B: after plasticity is applied.

Our results show that these neurons exist with a low degree of overlap (Figure 5). While with larger m more shared high MI neurons appear, there are mostly unique high MI neurons which are spread over the whole network. Of course this is also a reflection of the data (digits), which are fairly well separated in information-theoretical terms.

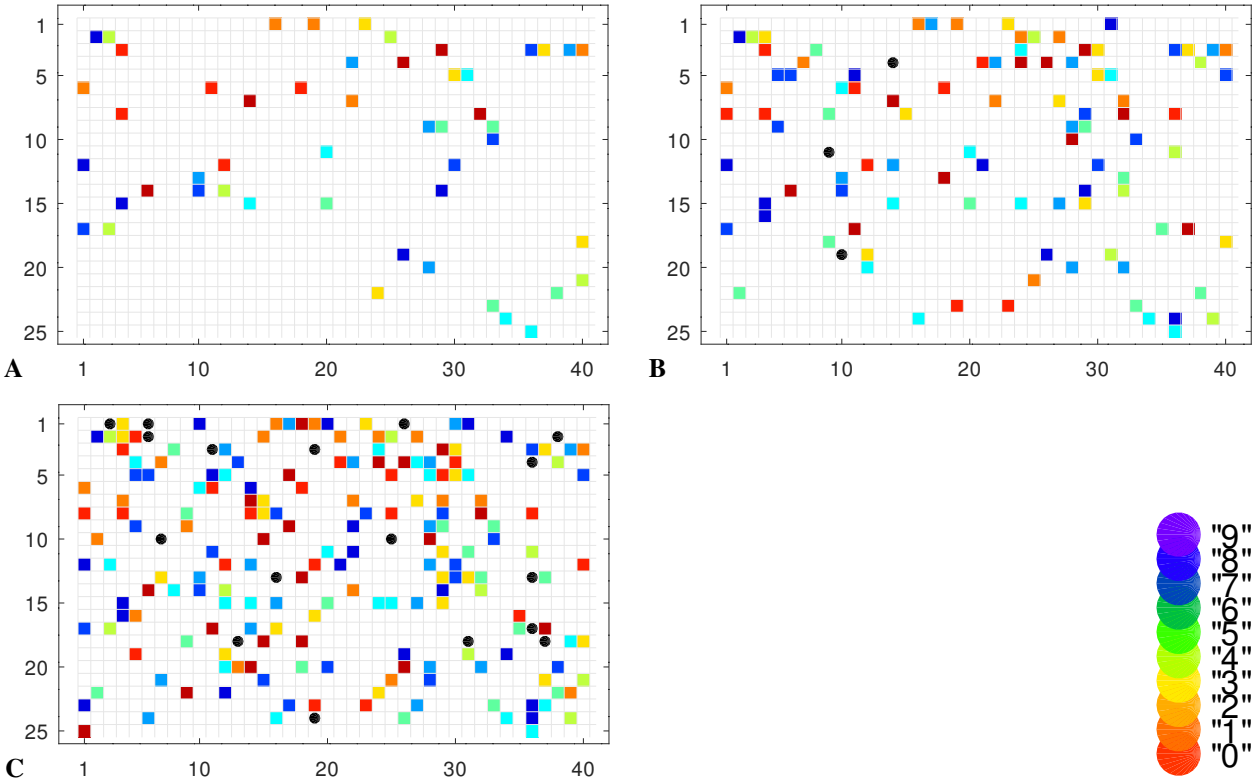

Figure 5: Ranked high MI neurons (m) for a naive network representation, unique for each digit (0-9), marked by color. (A) $m = 5$ neurons per pattern, (B) $m = 10$ neurons, and (C) $m = 20$ neurons. Shared high MI neurons are marked as round and black. They are sparse even at higher m.

3.3 Localist Plasticity and Recall

In order to store the representations of pattern input that appear on the naive, randomly initialised network, we apply plasticity rules. Our goal is to compress the information into selected neurons and their surrounding ensembles, and recall patterns by stimulating those high-level, "concept" neurons. There are experimental indications for constructions like this, and they also have enormous advantages in intelligence applications.

We have shown that MI is distributed such that a few neurons have high MI. We want to use these neurons like symbolic abstractions. In a number of experimental publications [2, 9], a hub-spoke representation for each pattern, with few, high-order, "concept" neurons and a larger amount of attached "feature" neurons, has been sometimes suggested. We believe that such structures imply great advantages for recall. It should be sufficient to target only the high-order neurons for activation, which then activate their feature structures, and in this way reconstruct the whole pattern.

Accordingly, we use a simple, one-shot plasticity rule, focused on high MI neurons. We may label this type of plasticity rule "localist", because it only affects high-order neurons and their synaptic environment. The basic idea is to form a trace of the input pattern activation as recorded in the intrinsic excitability and synaptic connections of selected neurons (cf.2.5) . The localist restriction of parameter updates is the decisive difference to a distributed network update rule, and ensures its specific recall properties. It is also computationally very efficient.

Foremost, we can show that the mutual information values for each pattern as the sum of the MI values over the high-ranking neurons for $m = 10$ increases as a result of plasticity (Fig. 6).

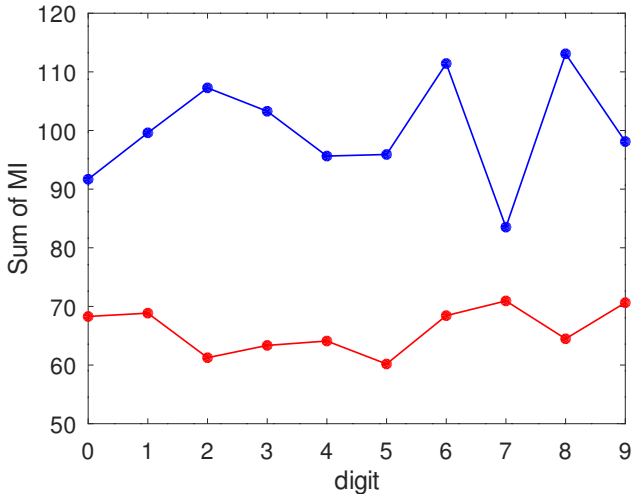

Figure 6: Sum of MI per digit before applying plasticity (red) and after application (blue)

The goal of localist plasticity is to allow restoration of representations with high accuracy from activation of high MI neurons alone. With this type of representation, pattern recall by localist stimulation of such 'concept' neurons

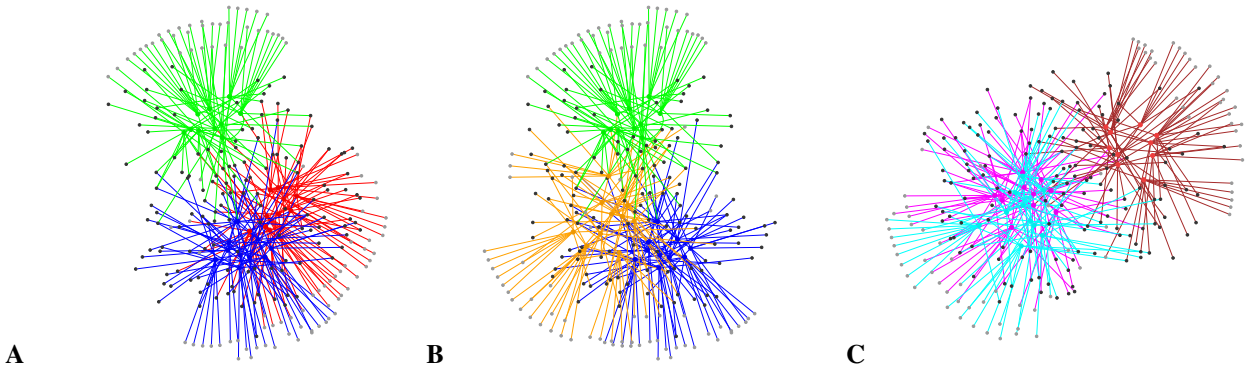

Figure 7: Network connection graph after plasticity for three sets of 3 digits. Target neurons for the strongest 100 connections from high-MI neurons per digit and their connections are shown. Grey indicates that a target neuron is digit-specific, black that there are overlaps. High-MI neurons ($m = 10$) for digits are shown as colored larger dots. A: digits "0" (red), "2" (blue), "6" (green). B: digits "4" (orange), "2" (blue), "6" (green). C: digits "1" (magenta), "5" (brown), "7" (light blue).

can be investigated. To understand the process we look at the neural representation after plasticity has been applied. The feature representations for each concept may be overlapping, and the same features can be re-used in multiple representations (cf. Fig.7 for examples).

3.4 Classification of recalled patterns

The classifier has been trained with the neural representations of the 50 variants for each digit, presented to the network before plasticity is applied. The activity of all ($n = 1000$) E neurons is recorded over 300ms. Each representation is labeled by its corresponding digit. Training validation with 100 patterns yields a classification accuracy of about 98%, using AutoGluon [1], Fig. 8).

This trained classifier is now subjected to recalled representations on the network, after plasticity has been applied (Figure 8).The representations unfold from selected stimulation of high-MI neurons. Again, the classifier returns, for each representation, the class probabilities. A representation is classified correctly if the class probability is highest for the corresponding digit, and above 50%. For $m = 5$ a recall of 100% could be achieved, for $m = 10$ two digits ("0", "5") and for $m = 20$ one digit ("0") was not recognized. (cf. Figure 8). We used the training/testing paradigm from machine learning to ascertain the similarity of the recalled representation with the original naive representation in a simple way. This result already shows, without further analysis, that recall of a complete representation from a few 'concept' neurons is possible and that the recalled representations are similar to the original, naive representations which resulted from input to ($k = 400$) neurons.

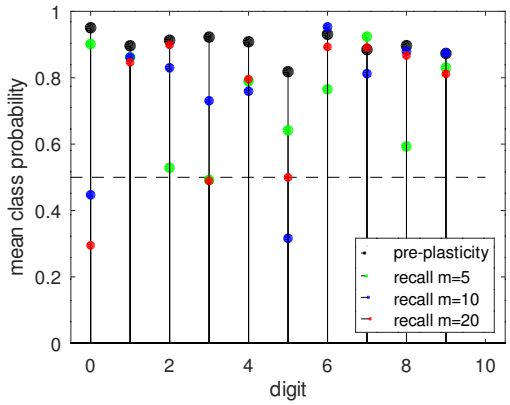

Figure 8: Pattern recall: classification for each digit pattern (A) by supervised training (black) and (B) from recalled representations with $m = 5$ (green), $m = 10$ (blue), and $m = 20$ (red) in unsupervised fashion (testing)

By stimulating the 'concept' neurons alone we can retrieve a version of the original input representation, sufficient to be recognized by the downstream classifier. Thus there is local storage at a small set of neurons for large patterns of hundreds of neurons, such that stimulation of these neurons allows to 'unfold' and spread activations to retrieve large representations. For each pattern, only a small subset of neurons and synapses are affected by plasticity. As we can see from Fig. 9, this result is partially achieved through reduction of activation on all neurons not affected by an input pattern as in Fig. 9A (pre-plasticity) vs. B and C (post-plasticity). These percentages are very stable across patterns, and they are also similar with the A,B,C response classification as in section 2.4.

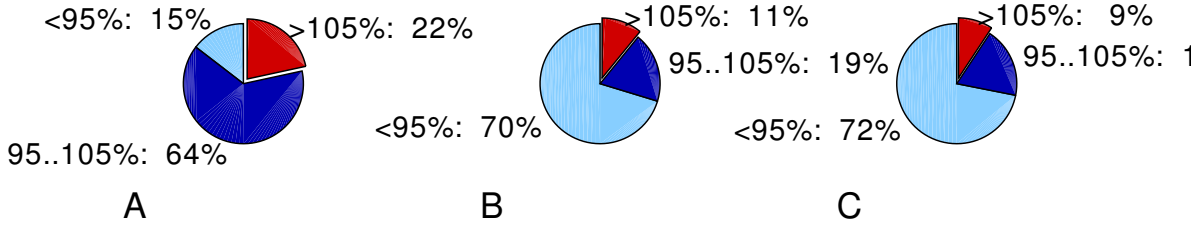

Figure 9: Response of the 1000 E neurons to stimulation: percentage of E neurons with a response of less than 95% of baseline (light blue), between 95% and 105% of baseline (dark blue), and above 105% (red). A. patterns prior to plasticity (mean of 50 variations for "1"), B. after plasticity (mean of 50 variations for "1") and C. after targeted stimulation with ($m = 10$ per digit)(mean over 10 digits).

This means that learning of patterns minimally affects the response of the network to other patterns. The network shows considerable independence in its response to input. This is based on a simple learning rule, which could be further developed. It is remarkable in this model that all pattern memories are stored in parallel, and stimulation of pattern-specific neurons will reproduce the associated distributed pattern with considerable accuracy (numbers?).

There are implications for experimental and technical results.

- experimental: We show that a cortical model is able to form high-level 'concept' neurons where individual pattern information is stored in a localized manner. A version of the pattern information can be retrieved from these neurons, forming a short-term ensemble as in Fig.7. These results fit with recent experimental work [5]. It has been empirically shown that up-regulation of intrinsic excitability occurs for those neurons which are part of a pattern (engram) [10].
- technical: The patterns which are entered into the network alter the impact of high information 'concept' neurons and their connections in a single trial. Over time, a number of different patterns can be stored in this way. In this case we were only interested in pattern storage and separation, but an application to overlap and similarity (generalization, abstraction) is the obvious next step.

3.5 Parameter Distributions

Plasticity changes the initial Gaussian distribution, which can be regarded as a narrow lognormal distribution to a wider lognormal distribution with heavy-tail characteristics (cf. Figure 10). The overall distribution of intrinsic gain parameters a and b after plasticity approaches a lognormal distribution, where the high-MI neurons are substantially strengthened (Figure 10 A). It is interesting that plasticity changes the intrinsic strength distribution towards a power tail distribution, because this will also alter network characteristics.

The picture is more complex for synaptic weights, because only few weights are changed by plasticity in our example. For $m = 10$ high MI (=control) neurons, only ≈ 6400 AMPA connections (out of $\approx 10^6$, that is 0.5%) are adapted. This leads to an uneven distribution of synaptic weights (Figure 10 B). Extrapolating for larger numbers of weight updates, a lognormal distribution would be reached.

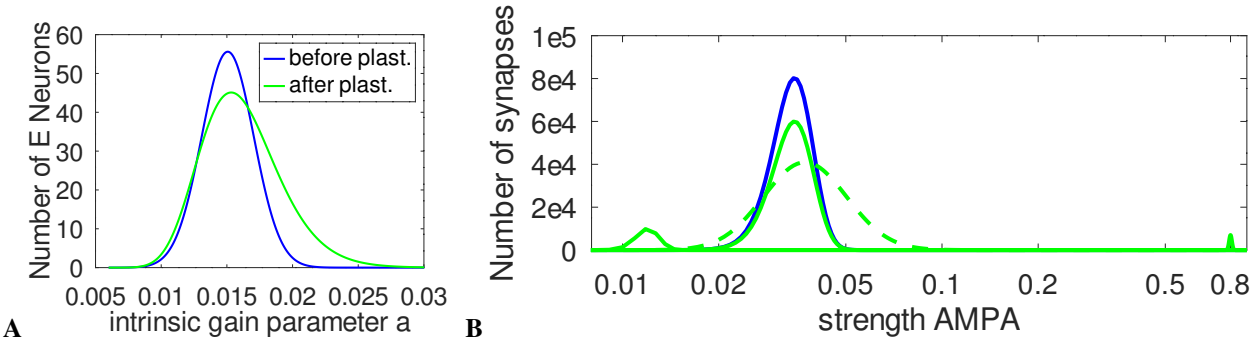

Figure 10: A: Distribution of the intrinsic gain parameters a and b before applying plasticity (blue), and after plasticity (green). The Gaussian distribution is streteched to a lognormal distribution. B: Distribution of strength of E-E AMPA connections prior to applying plasticity (blue), fitted to a Gaussian. After plasticity, the distribution consists of three different groups (green): strong synapses, which have been strengthened during plasticity, a group of weakened synapses, and a majority of unchanged synapses. The results can be approximated by a lognormal distribution (dashed) under the assumption that more synapses are updated.

4 Discussion

We outlay a mechanism to store patterns in cortical-like networks using an ensemble, more precisely a concept-feature ensemble, idea. The plasticity rule stores patterns on the network by localist adaptation of high MI neurons and their synaptic connections, creating a hub-spoke structure. By stimulating single-pattern high MI 'concept' neurons (recall), the resulting representation unfolds on the network and is similar enough to be recognized by a trained classifier. When these neurons are being activated by direct stimulation, they activate a whole set of related neurons to recreate the original pattern. In this network model individual neurons are recruited as pattern storage elements. We thus achieve a localist memory with a distributed component.

Such concept-feature ensemble need not be just passive storage of pattern information. The 'concept' neurons may also act as 'control' neurons, since they form a set of high-level neurons interacting with each other and having access to their feature neurons as needed.

We have not discussed the biological mechanisms which could underlie such selective plasticity, but there are possibilities in the cell-internal memory which filters the information by a number of indicators, such as small molecules and proteins [14]. This way of representing pattern information in a network allows for parallel storage of related patterns. We have not yet analysed the possibilities for blends and interference between patterns which serve to create high-order memory.

References

[1] N. Erickson, et al.: AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. arXiv preprint arXiv:2003.06505 (2020).

[2] E. Gal, O. Amsalem, et.al.: The Role of Hub Neurons in Modulating Cortical Dynamics. Front Neural Circuits. **15**:718270. (2021)

[3] E.M. Izhikevich: Which model to use for cortical spiking neurons? IEEE Trans Neural Netw. **15**(5):1063–1070 (2004)

[4] H. Jaeger: Adaptive nonlinear system identification with echo state networks. In Advances in Neural Information Processing Systems (NIPS) 15, MIT Press, pp. 593-600 (2003)

[5] S.A. Josselyn, S. Tonegawa: Memory engrams: Recalling the past and imagining the future. Science. **367**(6473) (2020)

[6] J. Kim, W. Leahy, E. Shlizerman. Neural Interactome: Interactive Simulation of a Neuronal System. Frontiers in Computational Neuroscience. **13**: 8 (2019)

[7] Y. LeCun and C. Cortes: MNIST handwritten digit database. `http://yann.lecun.com/exdb/mnist/` (2010)

[8] W. Maass, T. Natschlaeger, H. Markram: Real-time computing without stable states: A new framework for neural computation based on perturbations. Neural Computation, **14**(11):2531-2560 (2002)

[9] M. Agetsuma, I. Sato, et.al.: Activity-dependent organization of prefrontal hub-networks for associative learning and signal transformation. Nat Commun. **14**(1):5996. (2023)

[10] M. Pignatelli, T.J. Ryan, D.S. Roy, C. Lovett, L.M. Smith, S. Muralidhar, and S. Tonegawa: Engram Cell Excitability State Determines the Efficacy of Memory Retrieval. Neuron, **101**(2):274-284 (2019)

[11] G. Scheler: Extreme pattern compression in lognormal networks. F1000Research **6**:2177(poster) (2016)

[12] G. Scheler: Logarithmic distributions prove that intrinsic learning is Hebbian. F1000Research **6**:1222 (2017)

[13] G. Scheler: Neuromodulation influences synchronization and intrinsic readout. F1000Research, **7**:1277 (2018)

[14] G. Scheler: Sketch of a novel approach to a neural model. arXiv q-bio.NC. 2209.06865 (2023)