ELSEVIER

Dynamics of information and emergent computation in generic neural microcircuit models

Thomas Natschläger*, Wolfgang Maass

Institute for Theoretical Computer Science, Technische Universitaet Graz, A-8010 Graz, Austria

Abstract

Numerous methods have already been developed to estimate the information contained in single spike trains. In this article we explore efficient methods for estimating the information contained in the simultaneous firing activity of hundreds of neurons. Obviously such methods are needed to analyze data from multi-unit recordings. We test these methods on generic neural microcircuit models consisting of 800 neurons, and analyze the temporal dynamics of information about preceding spike inputs in such circuits. It turns out that information spreads with high speed in such generic neural microcircuit models, thereby supporting—without the postulation of any additional neural or synaptic mechanisms—the possibility of ultra-rapid computations on the first input spikes.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Neural microcircuit; Spiking neurons; Information theoretic methods; Neural coding; Computational power; Dynamic synapses; Linear regression; Bayesian classifier

1. Introduction

Common analytical tools of computational complexity theory cannot be applied to recurrent circuits with complex dynamic components, such as biologically realistic neuron models and dynamic synapses. In this article we explore the capability of information theoretic concepts to throw light on emergent computations in recurrent circuit of spiking neurons (we refer to p. 429 of Panzeri, Rolls, Battaglia, & Lavis, 2001 for a discussion of advantages in using information theoretic methods in this context). This approach is attractive since it may potentially provide a solid mathematical basis for understanding such computations. But it is methodologically difficult because of systematic errors caused by under-sampling problems that are ubiquitous even in extensive computer simulations of relatively small circuits. Previous work on these methodological problems had focused on estimating the information in spike trains, i.e. temporally extended protocols of the activity

of one or a few neurons. In contrast to that this paper addresses methods for estimating the information that is instantly available to a neuron that has synaptic connections to a large number of neurons. The proposed formalism to study simulated neural circuits has the advantage that it allows direct comparisons with experimental results on neural coding. In view of the very large existing literature on neural coding and relevant applications of information theory we cannot discuss here the preceding literature in detail. We refer to Borst and Theunissen (1999), deCharms and Zador (2000), Hertz (1999), Hertz and Panzeri (2003), Pola, Schultz, Petersen, and Panzeri (2003), and Rieke, Warland, van Stevenink, and Bialek (1997) for recent reviews. The dynamics of information in neural circuit models has previously been studied in Panzeri et al. (2001). In that study the speed of pattern completion was studied in a circuit model consisting of very realistic neuron models but static synapses. The network inputs consisted there of spatial patterns encoded by step currents, which represented fragments of more complete patterns from a fixed set of spatial patterns. Nevertheless, the results reported in that article about the speed of information processing are quite consistent with those reported in this article for the case where the network input consists of spike trains, and the fusion of information from several segments of these spike

inputs is examined (rather than the fusion of information between static input patterns and information stored in synaptic weights as in Panzeri et al., 2001).

We will define the specific circuit model used for our study in Section 2 (although the methods that we apply appear to be useful for to a much wider class of analog and digital recurrent circuits). The combination of information theoretic methods with methods from machine learning that we employ is discussed in Section 3. The results of applications of these methods to the analysis of the distribution and dynamics of information in a generic recurrent circuit of spiking neurons are presented in Section 4. Applications of these methods to the analysis of emergent computations are discussed in Section 5.

2. Our study case: a generic neural microcircuit model

As our study case for analyzing information in high-dimensional circuit states we used a randomly connected circuit with sparse, primarily local connectivity consisting of 800 leaky integrate-and-fire (I&F) neurons, 20% of which were randomly chosen to be inhibitory. Constants of neurons and synaptic parameters were chosen to reflect the diversity of parameters reported in experimental studies (see Destexhe & Marder, 2004 for a discussion).[1] The 800 neurons of the circuit were arranged on two 20×20 layers L1 and L2.[2] Circuit inputs consisting of five spike trains were injected into a randomly chosen subset of neurons in layer L1 (the connection probability was set to 0.25 for each of the five input channels and each neuron in layer L1). We modeled the (short term) dynamics of synapses according to the model proposed in Markram, Wang, and Tsodyks (1998), with the synaptic parameters U (use), D (time constant for depression), F (time constant for facilitation) randomly chosen from Gaussian distributions that model empirical data for such connections. Parameters of neurons and synapses were chosen as in Maass et al. (2002) to fit data from microcircuits in rat somatosensory cortex (based on Gupta, 2000; Markram et al., 1998).

---

[1] *Neuron parameters*: membrane time constant 30 ms, absolute refractory period 3 ms (excitatory neurons), 2 ms (inhibitory neurons), threshold 15 mV (for a resting membrane potential assumed to be 0), reset voltage 13.5 mV, constant nonspecific background current $I_b = 13.5$ nA, input resistance 1 MΩ.

[2] *Connectivity structure*: We assumed that the neurons were located on the integer points of a three-dimensional grid in space, where $D(a, b)$ is the Euclidean distance between neurons a and b. The probability of a synaptic connection from neuron a to neuron b (as well as that of a synaptic connection from neuron b to neuron a) was defined as $C \exp(-D^2(a, b)/\lambda^2)$, where λ is a parameter which controls both the average number of connections and the average distance between neurons that are synaptically connected (we set $\lambda = 2$, see Maass et al., 2002 for details). Depending on whether a and b were excitatory (E) or inhibitory (I), the value of C was 0.3 (EE), 0.2 (EI), 0.4 (IE), 0.1 (II).

Since neural microcircuits in the nervous system often receive salient input in the form of spatio-temporal firing patterns (e.g. from arrays of sensory neurons, or from other brain areas), we have concentrated on circuit inputs of this type. Such firing pattern could for example represent visual information received during a saccade, or the neural representation of a phoneme or syllable in auditory cortex. Information dynamics and emergent computation in recurrent circuits of spiking neurons were investigated for input streams over 800 ms consisting of sequences of *noisy* versions of four of such firing patterns. We restricted our analysis to the case where in each of the four 200 ms segments one of two template patterns is possible, see Fig. 1. In the following, we write $s_i = 1$ ($s_i = 0$) if a noisy version of template 1 (0) is used in the ith time segment of the circuit input.

Fig. 2 shows the response of a circuit of spiking neurons (drawn from the distribution specified above) to the input stream exhibited in Fig. 1B. Each frame in Fig. 2 shows the current firing activity of one layer of the circuit at a particular point t in time. Since in such rather small circuit (compared for example with the estimated 10^5 neurons below a square millimeter of cortical surface) very few neurons fire at any given millisecond, we have replaced each spike by a pulse whose amplitude decays exponentially with a time constant of 30 ms. More precisely, the spike train from each presynaptic neuron was convolved with the kernel $e^{-\tau/30 \text{ ms}}$. This models the impact of a spike on the receptors and the membrane potential of a generic postsynaptic neuron. The resulting vector $\boldsymbol{r}(t) = \langle r_1(t), \ldots, r_{800}(t) \rangle$ consisting of 800 analog values from the 800 neurons in the circuit is exactly the 'liquid state' of the circuit at time t in the context of the abstract computational model introduced in Maass et al. (2002). In the subsequent sections, we will analyze the temporal dynamics of the information contained in these momentary circuit states $\boldsymbol{r}(t)$.[3]

3. Methods for analyzing the information contained in circuit states

The mutual information $MI(X, R)$ between two random variables X and R can be defined by $MI(X, R) = H(X) - H(X|R)$, where $H(X) = -\sum_{x \in \text{Range}(X)} p(x) \log p(x)$ is the entropy of X, and $H(X|R)$ is the expected value (with regard to R) of the conditional entropy of X given R, see e.g. Cover and Thomas (1991). It is well known that empirical estimates of the entropy tend to underestimate the true entropy of a random variable (see e.g. Panzeri & Treves,

---

[3] One should note that these circuit states do not reflect the complete current state of the underlying dynamical system, only those parts of the state of the dynamical system that are in principle 'visible' for neurons outside the circuit. The current values of the membrane potential of neurons in the circuit and the current values of internal variables of dynamic synapses of the circuit are not visible in this sense.

A   possible templates for spike train segments

B   an example for resulting input spike trains (with noise)
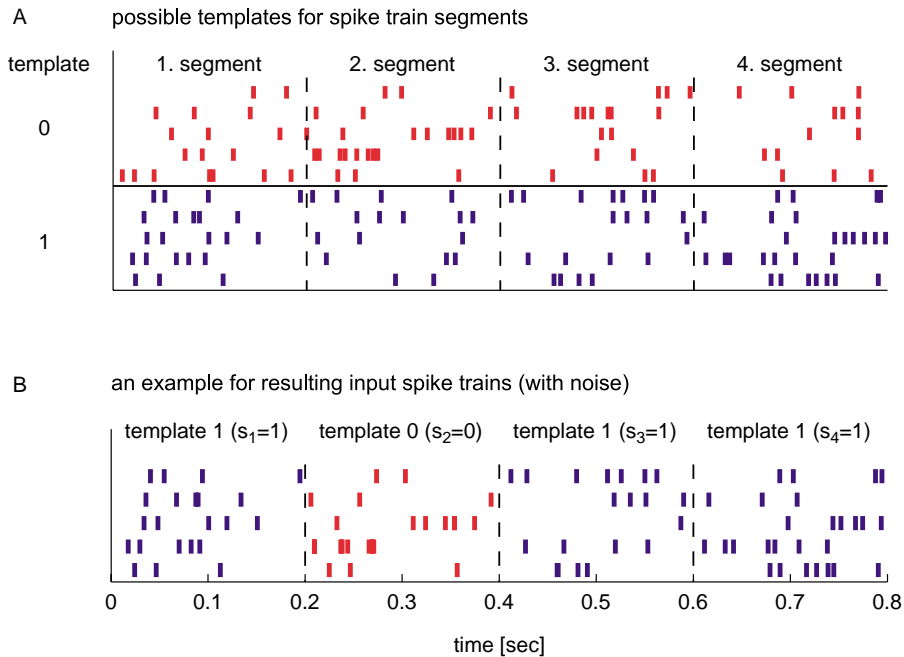

time [sec]

Fig. 1. Input distribution used throughout the paper. Each input consists of five spike trains of length 800 ms generated from four segments of length 200 ms each. (A) For each segment two templates 0 and 1 were generated randomly (Poisson spike trains with a frequency of 20 Hz). (B) The actual input spike trains were generated by choosing randomly for each segment i, $i=1,...,4$, one of the two associated templates ($s_i=0$ or $s_i=1$), and then generating a noisy version by moving each spike by an amount drawn from a Gaussian distribution with mean 0 and SD 4 ms.

1996; Roulston, 1999). Hence, in situations where the true value of $H(X)$ is known (as is typically the case in neuroscience applications where X represents the stimulus, whose distribution is controlled by the experimentalist), the generic underestimate of $H(X|R)$ yield a generic overestimate of the mutual information $MI(X, R) = H(X) - H(X|R)$ for finite sample sizes. This undersampling effect has been addressed in a number of studies (see e.g. Hertz, 1999; Paninski, 2003; Pola et al., 2003; and the references therein), and has turned out to be a serious obstacle for a wide-spread application of information theoretic methods to the analysis of neural computation. The seriousness of this problem becomes obvious from results achieved for our study case of a generic neural microcircuit shown in Fig. 3A. The dashed line shows the dependence of 'raw' estimates MI_{raw} of the mutual information $MI(s_2, R)$ on the sample size[4] N, which ranges here from 10^3 to 2×10^5. The raw estimate of $MI(s_2, R)$ results from a direct application of the definition of MI to the observed occupancy frequencies for a discrete set of bins,[5]

where R consists here of just $d=5$ or $d=10$ components of the 800-dimensional circuit state $r(t)$ for $t=660$ ms, and s_2 is the bit encoded by the second input segment. For more components d of the current circuit state $r(t)$, e.g. for estimating the mutual information $MI(s_2, R)$ between the preceding circuit input s_2 and the current firing activity in a subcircuit consisting of $d=20$ or more neurons, even sample sizes beyond 10^6 are likely to severely overestimate this mutual information.

Several methods for correcting this bias towards overestimation of MI have been suggested in the literature. In Section 3.1 of Pola et al. (2003), it is proposed to subtract one of three possible bias correction terms B_{naive}, B_{full}, and B_{Bayes} from the raw estimate MI_{raw} of the mutual information. The effect of subtracting B_{naive} is shown for $d=5$ components of $r(t)$ in Fig. 3A. This correction is too optimistic for these applications, since the corrected estimate $MI_{naive} = MI_{raw} - B_{naive}$ at small sample sizes (e.g. 10^4) is still substantially larger than the raw estimate MI_{raw} at large sample sizes (e.g. 10^5). The subtraction of the second proposed term B_{full} is not applicable in our situation because it yields for $MI_{full} = MI_{raw} - B_{full}$ values lower than zero for all considered sample sizes. The reason is, that B_{full} is proportional to the quotient 'number of possible response bins'/ N and the number of possible response bins is in the order of 30^{10} in this example. It remains an open question how well the correction term B_{Bayes} can be made to perform in our setup.

Another way to correct MI_{raw} is proposed in Strong et al. (1998). This approach is based on a series expansion of MI

---

[4] In our case the sample size N refers to the number of computer simulations of the circuit response to new drawings of circuit inputs, with new drawings of temporal jitter in the input spike trains and initial conditions of the neurons in the circuit.

[5] For direct estimates of the MI the analog value of each component of the circuit state $r(t)$ has to be divided into discrete bins. We first linearly transformed each component of $r(t)$ such that it has zeros mean and variance $\sigma^2 = 1.0$. The transformed components are then binned with a resolution of $\epsilon = 0.5$. This means that there are four bins in the range $\pm \sigma$.
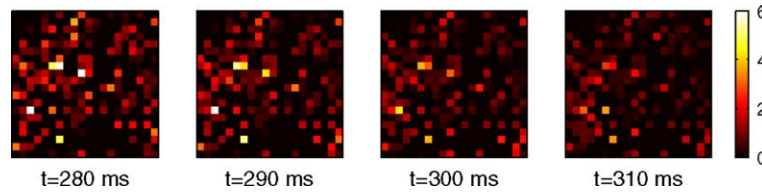
Fig. 2. Snapshots of the first 400 components of the circuit state $r(t)$ (corresponding to the neurons in the layer L1) at various times t for the input shown at the bottom of Fig. 1. Black denotes high activity, white no activity. A spike at time $t_s \leq t$ adds a value of $\exp(-(t-t_s)/(30 \text{ ms}))$ to the corresponding component of the state $r(t)$.

in $1/N$ (Panzeri & Treves, 1996) and is effectively a method to get an empirical estimate MI_{infinity} of the mutual information for infinite sample size ($N \to \infty$).[6] It can be seen in Fig. 3A that for moderate sample sizes MI_{infinity} also yields too optimistic estimates for MI.

A further method for dealing with generic overestimates of MI has been proposed in Strong et al. (1998). This method is based on the equation $MI(X, R) = H(R) - H(R|X)$ and compares the raw estimates of $H(R)$ and $H(R|X)$ with the so-called Ma-bounds (Ma, 1981), and suggests to judge raw estimates of $H(R)$ and $H(R|X)$, and hence raw estimates of $MI(X, R) = H(R) - H(R|X)$, as being trustworthy as soon as the sample size is so large that the corresponding Ma-bounds (which are conjectured to be less affected by undersampling) assume values below the raw estimates of $H(R)$ and $H(R|X)$.[7] According to this criterion a sample size of 9×10^3 would be sufficient in the case of 5-neuron subcircuits (i.e. $d = 5$ components of $r(t)$), c.f. Fig. 3B.[8] However, Fig. 3A shows that the raw estimate MI_{raw} is still too high for $N = 9 \times 10^3$, since MI_{raw} assumes a substantially smaller value at $N = 2 \times 10^5$.

In view of this unreliability of—even corrected—estimates for the mutual information we have employed standard methods from machine learning in order to derive *lower*

bounds for the MI (see for example Hertz, 1999; Paninski, 2003 for references to preceding related work). This method is computationally feasible and yields with not too large sample sizes reliable lower bounds for the MI even for large numbers of components of the circuit state. In fact, we will apply it in Sections 4 and 5 even to the full 800-component circuit state $r(t)$. This method is quite simple. According to the data processing inequality (Cover & Thomas, 1991), one has $MI(X, R) \geq MI(X, h(R))$ for *any* function h. Obviously $MI(X, h(R))$ is easier to estimate than $MI(X, R)$ if the dimension of $h(R)$ is substantially lower than that of R, especially if $h(R)$ assumes just a few discrete values. Furthermore, the difference between $MI(X, R)$ and $MI(X, h(R))$ is minimal if $h(R)$ throws away only that information in R that is not relevant for predicting the value of X. Hence, it makes sense to use as h a predictor or classifier that has been trained to predict the current value of X. Similar approaches for estimating a lower bound were motivated by the idea of predicting the stimulus (X) given the neural response (R) (see Hertz, 1999; Paninski, 2003 and the references therein). To get an unbiased estimate for $MI(X, h(R))$ one has to make sure that $MI(X, h(R))$ is estimated on data which have not been used for the training of h. To make the best use of the data one can alternatively use cross-validation or even leave-one-out (see Duda, Hart, & Storck, 2001) to estimate $MI(X, h(R))$. Fig. 3C–F shows for three different predictors h how the resulting lower bounds for the MI depend on the sample size N.

It is noteworthy that the lower bounds $MI(X, h(R))$ derived with the empirical Bayes classifier[9] increase significantly with the sample size[10] and converge quite well to the upper bounds $MI_{\text{raw}}(X, R)$. This reflects the fact that the estimated joint probability density between X and R gets more and more accurate. Furthermore, the computationally less demanding[11] use of linear classifiers h also

---

[6] For a given sample size N the raw estimate for the mutual information is estimated using several subsample sizes $n1 < n2 < \cdots < nk < N$. Let $MI_{\text{raw}}i$ be the estimate using the subsample size n_i. MI_{infinity} is estimated as the intercept of the least squares fit of the model $MI_{\text{raw}}i = MI_{\text{infinity}} + \frac{1}{ni}C1 + \frac{1}{ni^2}C2$ to the observed data points $\langle ni, MI_{\text{raw}}i \rangle$.

[7] Due to (Ma, 1981) a lower bound for the entropy $H(R)$ (and analogous for the entropy $H(R|X)$) can be estimated by counting coincidences. These Ma-bounds are given by (Strong et al., 1998)

$$H(R)_{\text{Ma}} = -\sum_{R_i} P(R_i) \cdot \log\left(P(R_i) \frac{2n_c(R_i)}{N(R_i)(N(R_i)-1)} \right)$$

where the regions R_i are disjoint subspaces of the response space, $P(R_i)$ is the probability that any response belongs to region R_i, $N(R_i)$ is the number of observed responses in region R_i, and $n_c(R_i)$ is the number of observed coincidences in region R_i. According to Strong et al. (1998) the lower bound is tightest if the responses are distributed closely to uniform in the individual regions R_i. Hence, we have chosen the regions R_i such that the frequency counts of the individual responses in any given region R_i differ at most by 1.

[8] These kind of results depend on a division of the space of circuit states into subspaces, which is required for the calculation of the Ma-bound. In our case we have chosen the subspaces such that the frequency counts of any two circuit states in the same subspace differ by at most 1.

[9] The empirical Bayes classifier operates as follows: given observed (and discretized) d components $r^{(d)}(t)$ of the state $r(t)$ it predicts the input which was observed most frequently for the given state components $r^{(d)}(t)$ (maximum a posterior classification, see e.g. Duda et al., 2001). If $r^{(d)}(t)$ was not observed so far a random guess about the input is made.

[10] In fact, in the limit $N \to \infty$ the Bayes classifier is the optimal classifier for the discretized data in the sense that it would yield the lowest classification error—and hence the highest lower bound on mutual information—over all possible classifiers.

[11] In contrast to the Bayes classifier the linear classifiers (both for analog and discrete data) yield already for relatively small sample sizes N good results which do not improve much with increasing N.
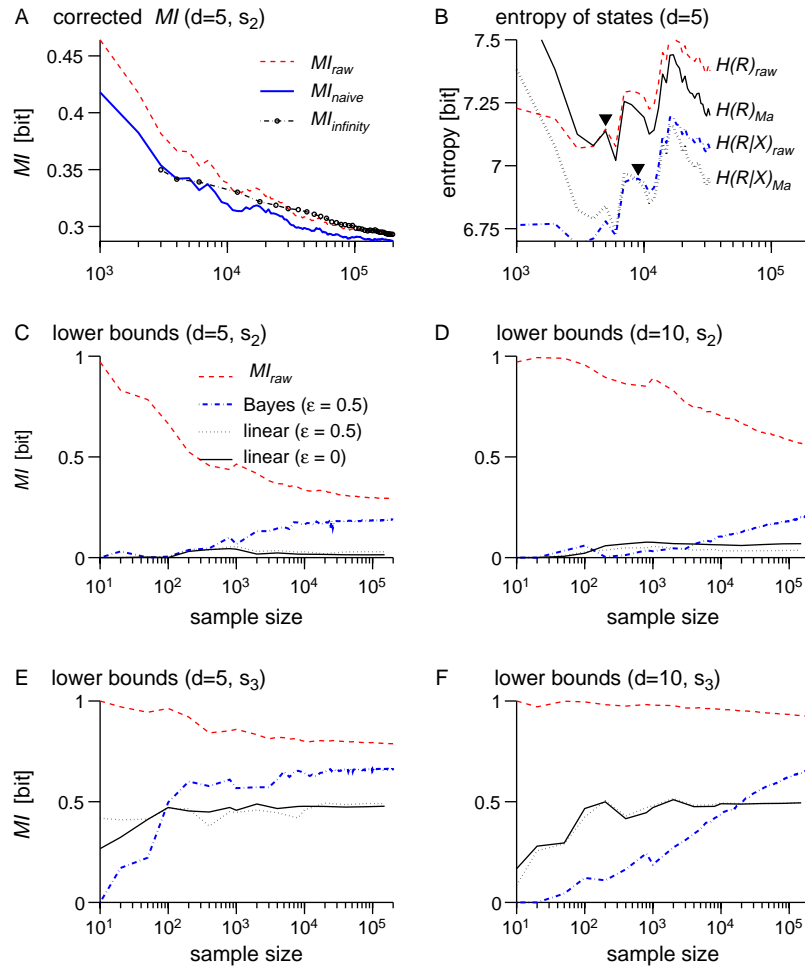
Fig. 3. Estimated mutual information depends on sample size. In all panels *d* denotes the number of components of the circuit state *r*(*t*) at time *t*=660 ms (or equivalently the number of neurons considered). (A) Dependence of the 'raw' estimate MI_{raw} and two corrected estimates MI_{naive} and $MI_{infinity}$ of the mutual information $MI(s_2, R)$ (see text). (B) Estimates of the entropies $H(R)$ and $H(R|X)$. The 'raw' estimates are compared with the corresponding Ma-bounds (see text). The filled triangle marks the sample size from which on the Ma-bound is below the raw estimate. (C) Lower bounds $MI(s_2, h(R))$ for the mutual information obtained via classifiers *h* which are trained to predict the actual value of s_2 given the circuit state *r*(*t*). Results are shown for (a) an empirical Bayes classifier (discretization $\epsilon = 0.5$, see footnotes 5 and 9), (b) a linear classifier trained on the discrete ($\epsilon = 0.5$) data and (c) for a linear classifier trained on the analog data ($\epsilon = 0$). In the case of the Bayes classifier $MI(s_2, h(R))$ was estimated by employing a leave-one-out procedure (which is computationally efficient for a Bayes classifier), whereas for the linear classifiers a test set of size 5×10^4 was used (hence no results beyond a sample size of 1.5×10^5). (D) Same as (C) but for *d*=10. (E) Same as (C) but for $MI(s_3, h(R))$. (F) Same as (E) but for *d*=10.

yields significant lower bounds for *MI*(*X*, *R*), especially if the true value of *MI*(*X*, *R*) is not too small. In our application this does not even require high numerical precision, since a coarse binning (see footnote 3) of the analog components of *r*(*t*) suffices, see Fig. 3C–F. All estimates of *MI*(*X*, *R*) in the subsequent sections are lower bounds *MI*(*X*, *h*(*R*)) computed via linear classifiers *h*. This is meaningful since we will focus there on cases where even these simple lower bounds for *MI* assume high values. These types of lower bounds for *MI*(*X*, *R*) are of particular interest from the point of view of neural computation, since a linear classifier can in principle be approximated by a neuron that is trained (for example by a suitable variation of the perceptron learning rule) to extract information about *X* from the current circuit state *R*. Hence, a high value of a lower bound *MI*(*X*, *h*(*R*)) for such *h* shows not only that information about *X* is present in

the current circuit state *R*, but also that this information is in principle accessible for other neurons.

4. Distribution and dynamics of information in circuit states

We have applied the method of estimating lower bounds for mutual information via linear classifiers described in Section 3 to analyze the spatial distribution and temporal dynamics of information for our study case described in Section 2. Fig. 4 shows the temporal dynamics of information (estimated every 20 ms as described in Section 3) about input bits s_i (encoded as described in Section 2) for different components of the circuit state *r*(*t*) corresponding to different randomly drawn subsets of
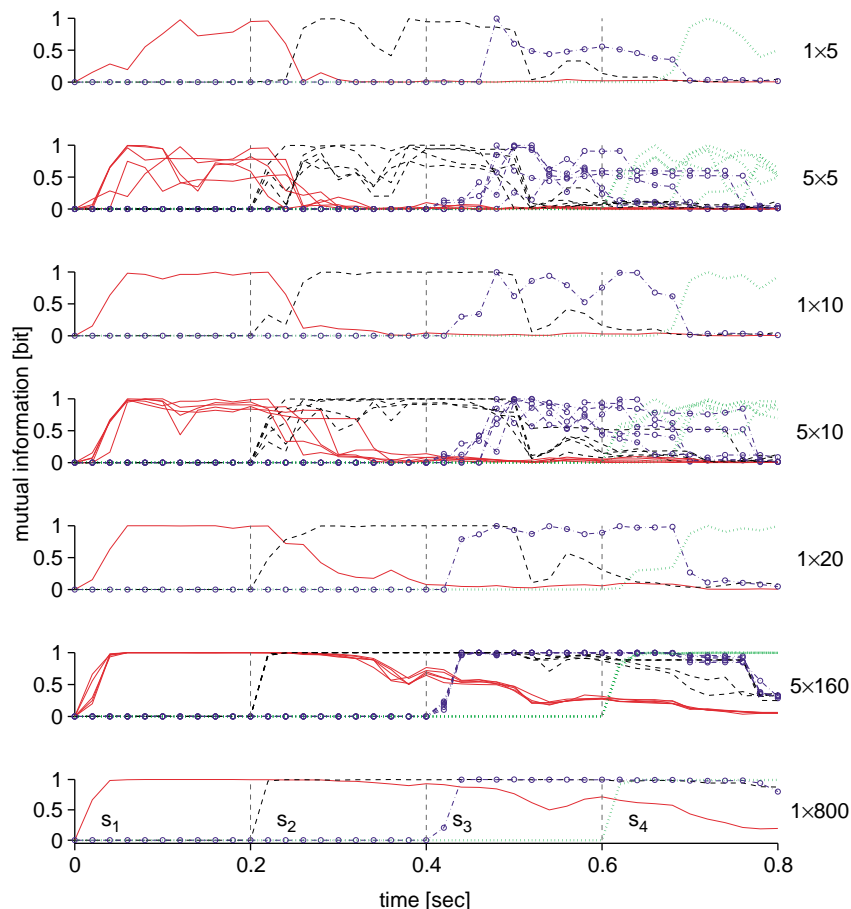
Fig. 4. Information in subset of neurons. Shown are lower bounds for mutual information $MI(s_i, h(R))$ obtained with a linear classifier h operating on d components of the circuit state $r(t)$. The numbers $a \times d$ to the right of each panel specify the number of components d used by the linear classifier and for how many different choices a of such subsets of size d the results are plotted in that panel.

neurons in the circuit. One sees that even subsets of just five neurons absorb substantial information about the input bits s_i; however, with a rather slow onset of the information uptake at the beginning of a segment and little memory retention when this information is overwritten by the next input segment. By merging the information from different subsets of neurons the uptake of new information gets faster and the memory retention grows. Note that for large sets of neurons (160 and 800) the information about each input bit s_i jumps up to its maximal value right at the *beginning* of the corresponding ith segment of the input trains.

5. Emergent computation in recurrent circuits of spiking neurons

In this section, we apply the same method to analyze not only how much information a current circuit state contains about preceding inputs, but also its information about various potential target outputs. This provides an interesting new method for analyzing neural computation, rather than just neural communication and coding. Computations require in general that some information from the input is

surpressed (e.g. for position-invariant object recognition, but even for just computing the XOR of 2 bits), hence an optimal communication channel is not necessarily an optimal module for computation tasks. There exist 16 different Boolean functions $f(s_1, s_2)$ that depend just on the first two of the 4 bits $s_1,\ldots,s_4$. Fig. 5B,C shows that all these Boolean functions f are autonomously computed by the circuit, in the sense that the current circuit state contains high mutual information with the target output $f(s_1, s_2)$ of this function f. Furthermore, the information about the result $f(s_1, s_2)$ of this computation can be extracted *linearly* from the current circuit state $r(t)$ (in spite of the fact that the computation of $f(s_1, s_2)$ from the spike patterns in the input requires highly nonlinear computational operations). This is shown in Fig. 5B,C for those five Boolean functions of two variables that are nontrivial in the sense that their output really depends on both input variables. There exist five other Boolean functions which are nontrivial in this sense, which are just the negations of the five Boolean functions shown (and for which the mutual information analysis therefore yields exactly the same result). In Fig. 5D, corresponding results are shown for parity functions that depend on three of the 4 bits s_1, s_2, s_3, s_4. These Boolean functions are the most
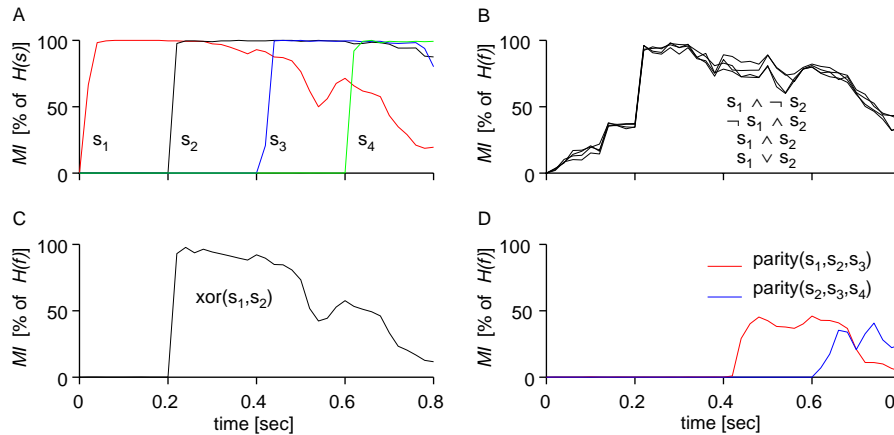
Fig. 5. Emergent computations. (A) Dynamics of information about input bits as in the bottom row of Fig. 4. $H(s)$ denotes the entropy of a segment s_i (which is 1 bit for $i=1, 2, 3, 4$). (B–D) Lower bounds for the mutual information $MI(f, h(R))$ for various Boolean functions $f(s_1,...,s_4)$ obtained with a linear classifier h operating on the full 800-component circuit state $R = r(t)$. $H(f)$ denotes the entropy of a Boolean function $f(s_1,...,s_4)$ if the s_i are independently uniformly drawn from $\{0, 1\}$.

difficult ones to compute in the sense that knowledge of just 1 or 2 of their input bits does not give any advantage in guessing the output bit.

One noteworthy feature in all these emergent computations is that information about the result of the computation is already present in the current circuit state long before the complete spatio-temporal input patterns that encode the relevant input bits have been received by the circuit. In fact, the computation of $f(s_1, s_2)$ automatically just uses the temporal order of the first spikes in the pattern encoding s_2, and merges information contained in the order of these spikes with the 'context' defined by the preceding input pattern. In this way the circuit supports a potential ultra-rapid computation within just 20 ms of the beginning of the second pattern s_2. The existence of such ultra-rapid neural computations has previously already been inferred (Thorpe et al., 1996), but models that would support the possibility of such ultra-rapid computations on the basis of generic models for recurrent neural microcircuits have been missing. Fig. 5 demonstrates that just a suitable trained linear readout neuron is needed in order to produce from the current state of a generic neural microcircuit model the result of an ultra-rapid computation on complex spike inputs.

6. Discussion

We have analyzed the dynamics of information in high-dimensional circuit states of a generic neural microcircuit model. We have focused on that information which can be extracted by a linear classifier from the current state of a neural circuit (a linear classifier may be viewed as a coarse model for the classification capability of a biological neuron). It had previously been shown that quite similar linear decoding methods can be used to reconstruct moving visual stimuli from the current state of ensembles of neurons in the LGN of cats (Stanley, 1999), and hand movements from the current state of ensembles of cortical neurons in primates (Wessberg et al., 2000). In the context of our circuit simulations such linear decoding has the additional advantage that significant lower bounds for the information content of high-dimensional circuit states can already be achieved for relatively small sample sizes. Our results show that information about current and preceding circuit inputs is quickly spread throughout the circuit in a rather uniform manner. Furthermore, our results show that a generic neural microcircuit model has inherent capabilities to process new input in the context of other information that arrived several hundred ms ago, and that information about the outputs of numerous potentially interesting target functions automatically accumulates in the current circuit state. While this article has focused on the spread of information in generic cortical microcircuit models, it has not yet addressed the open problem which computational operations on the accumulated information are especially supported by the structure of these circuits. A first step in this direction is made (for a more detailed microcircuit model) in Häusler and Maass (2005). Such emergent computation in generic models for circuits of spiking neurons is extremely fast, and therefore provides an interesting alternative to models based on special-purpose constructions for explaining empirically observed (Thorpe et al., 1996) ultra-rapid computations in neural systems. A closely related computational model for computations on time series has been proposed independently in Jäger (2002) and Jäger and Haas (2004) in the context of artificial neural network models.

Further research will have to clarify a number of interesting questions which are left open. For example, it would be fruitful to analyze which coding of information in the input (and which number of input channels) to a neural microcircuit model is most effective in injecting new information rapidly into the circuit, or alternatively, is most effective in making sure that this information decays

less quickly in the circuit. Simultaneously other information contained in input spike trains should perhaps decay particularly fast since it just represents noise (from the perspective of specific types of computations). In addition it will be interesting to see how such 'optimal input coding' from the perspective of a specific cortical area depends on various anatomical and physiological parameters of the microcircuits involved. Possibly such investigation would provide functional explanations for subtle differences in the anatomy and physiology of neural microcircuits in primary sensory areas for different sensory modalities (vision, auditory, somatosensory, etc.)

The method for analyzing information contained in high-dimensional circuit states that we have explored in this article for a generic neural microcircuit model should also be applicable to biological data from multi-unit recordings, *fMRI*, etc. since significant lower bounds for mutual information were achieved in our study case already for sample sizes in the range of a few hundred (see Fig. 3). This may provide further insight into the dynamics of information and emergent computations in biological neural systems.

Acknowledgements

References

Borst, A., & Theunissen, F. F. (1999). Information theory and neural coding. *Nature Neuroscience*, 2, 947–957.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.

deCharms, R. C., & Zador, A. (2000). Neural representation and the cortical code. *Annu Rev Neurosci*, 23, 613–647.

Destexhe, A., & Marder, E. (2004). Plasticity in single neuron and circuit computations. *Nature*, 431, 789–795.

Duda, R. O., Hart, P. E., & Storck, D. G. (2001). *Pattern Classification* (2nd ed.). New York: John Wiley.

Gupta, A., Wang, Y., & Markram, H. (2000). Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science*, 287, 273–278.

Häusler and Maass, 2005] Häusler, S. and Maass, W. (2005). Emergent computational properties of lamina-specific cortical microcircuit models. *submitted for publication*.

Hertz, J. (1999). Reading the information in the outcome of neural computation. In *Frontiere della Vita*, Vol. 3 (pp. 631–649). Sistermi Intelligenti, Institute della Enciclopedia Italiana. online available via http://www.nordita.dk/~hertz/papers/infit.ps.gz.

Hertz, J., & Panzeri, S. (2003). *Sensory coding and information transmission The handbook of brain theory and neural networks* (2ndv ed.). MIT Press.

Jäger, H. (2002). Short term memory in echo state networks. GMD Report 152, German National Research Center for Information Technology.

Jäger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication.. *Science*, 304, 78–80.

Ma, S.-K. (1981). Calculation of entropy from data. *Journal of Statistical Physics*, 26(2).

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

Markram, .H., Wang, Y., & Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *PNAS*, 95, 5323–5328.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Computation*, 15, 1191–1253.

Panzeri, S., Rolls, E. T., Battaglia, F., & Lavis, R. (2001). Speed of feedforward and recurrent processing in multilayer networks of integrate-and-fire neurons. *Network*, 12(4), 423–440.

Panzeri, S., & Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures. *Network: Computation in Neural Systems*, 7, 87–107.

Pola, G., Schultz, S. R., Petersen, R. S., & Panzeri, S. (2003). In R. Kötter (Ed.), *A Practical Guide to Information Analysis of Spike Trains* (pp. 139–153). Boston: Kluwer Academic Publishers.

Rieke, F., Warland, D., van Steveninck, R. R. D., & Bialek, W. (1997). *SPIKES: Exploring the neural code*. Cambridge, MA: MIT Press.

Roulston, M. S. (1999). Estimating the errors on measured entropy and mutual information. *Physica D*, 125, 285–294.

Stanley, G. B., Li, F. F., & Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *The Journal of Neuroscience*, 19(18), 8036–8042.

Strong, S., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, E. (1998). Entropy and information in neural spike trains. *Physical Review Letters*, 80(1), 197–200.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.

Wessberg, J., Stambaugh, C. R., Kralik, J. D., Beck, P. D., Laubach, M., Chapin, J. K., et al. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(6810), 361–365.