HIPPOCAMPUS 9:582–598 (1999)

Analysis of Information Transmission in the Schaffer Collaterals

Simon R. Schultz* and Edmund T. Rolls

Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

ABSTRACT: Hippocampal region CA1 seems from comparative studies to be particularly important in the primate brain, in addition to being crucial to memory function. Thus, it is an extremely appropriate place to begin a quantitative investigation of the information representation and transmission capabilities of cerebral neural networks. In this study, a mathematical model of the Schaffer collateral projection from CA3 to CA1 is described. From the model, the amount of information that can be conveyed by the Schaffer collaterals is calculated, i.e., the information that a pattern of firing in CA1 conveys about a pattern of firing in CA3, because of the connections between them. The calculation is performed analytically for an arbitrary probability distribution describing the pattern of CA3 firing and then solved numerically for particular input distributions. The effect of a number of issues on the information conveyed is examined. Consideration of the effect of the amount of analog resolution of firing rates in the patterns of activity in CA3 confirmed information transmission to be most efficient for binary codes, to a degree that depends on the sparseness of activity. For very sparse codes, a binary code allows more information to be received even in absolute terms, but for more distributed codes, slightly more information can be received by CA1 by making use of analog resolution. The pattern of convergence of connections from CA3 to CA1 is examined, i.e., the spatial distribution of the number of connections each CA1 neuron receives. It is found that the effect of the difference between a uniform convergence model and a proposed real convergence pattern (Bernard and Wheal, Hippocampus 1994;4:497–529) is minimal. The effect of the ratio of expansion between CA3 and CA1 due to the relative numbers of neurons in these two areas is studied. The Schaffer collaterals in all mammalian species reported in the literature seem to operate in a régime in which there is at least the scope for efficient transfer of information. In addition, the effect of topography (with respect to the transverse hippocampal axis) in the Schaffer collateral connectivity is examined. In the absence of spatial correlations, topography is found to have essentially no effect on information transmission. If spatial correlations in firing were present in CA3 (which, however, would be less efficient for memory storage in the recurrent collaterals), information transmission would be maximized by matching the topographic spread to the spatial scale of correlation. *Hippocampus 1999;9:582–598.*

KEY WORDS: hippocampus; CA1; information theory; neural network; spatial correlations

INTRODUCTION

The Schaffer collaterals unidirectionally link the CA3 and CA1 cytoarchitectonic fields of the hippocampus. For a number of reasons, the CA1 subfield is a particularly important region to study. In Alzheimer's disease, neuronal degeneration (Braak and Braak, 1991; Giannakopoulos et al., 1997; Simić et al., 1997) and loss (West et al., 1994) is seen more distinctively in CA1 than any other brain area; similarly, there is an abnormal distribution of neuropeptide-Y fibers in the hippocampus of Alzheimer-type brains, but it is particularly strong in area CA1 (Chan-Palay et al., 1986). Brain damage caused by ischemic/anoxic episodes (Zola-Morgan et al., 1986; Victor and Agamanolis, 1990) is particularly severe in CA1; CA1 neuronal loss in epileptic seizures is also evident (Sommer, 1880; Dam, 1980). Comparative neuroanatomy by using volumetric comparison (Stephan, 1983; West and Schwerdtfeger, 1985) and cell counts (Seress, 1988; West, 1990) shows a relative enlargement of the CA1 area inside the hippocampal formation in primates (compared with what would be expected from a "primitive" mammal of the same body weight), together with changes in cytoarchitecture. Anatomically, whereas hippocampal subregion CA3 seems to be the point of greatest convergence of information from various cortical areas, CA1 is uniquely situated as the first stage through which all output from CA3 back to the neocortex must pass. So, given the obvious importance of the CA1 region and its connections to brain function, it is important to gain a quantitative understanding of the transmission of information to CA1. By using mathematical techniques based on information theory, and drawing heavily from statistical mechanics, we now have the capability to do this, albeit with relatively simple models of the anatomy and physiology of the region.

The anatomy of the Schaffer collaterals has in fact been rather well explored. Autoradiographic techniques (Swanson et al., 1978), degeneration methods and retrograde transport (Lauberg, 1979), and anterograde transport (*Phaseolus vulgaris* leucoagglutinin, or PHA-L) (Ishizuka et al., 1990) have all been used to study the details of the connectivity between CA3 and CA1. In the Sprague-Dawley rat, around 304,000 CA3 pyramidal neurons project to 420,000 CA1 cells (Boss et al., 1987; Amaral et al., 1990). For comparison, in humans, there have been found to be approximately 2.7 million

*Correspondence to: Simon R. Schultz, Howard Hughes Medical Institute, New York University Center for Neural Science, 4 Washington Place, New York, NY 10003.
Accepted for publication 10 May 1999

neurons in CA3/2, and 16 million in CA1 (West, 1990; West and Gundersen, 1990; but see Seress, 1988). The projections from CA3 to CA1 terminate in stratum oriens and stratum radiatum of CA1. There exists a transverse topography in the connections, such that collaterals originating in CA3 cells close to the dentate gyrus tend to project into the superficial portion of stratum radiatum in the distal part of CA1 (near the subiculum), whereas CA3 cells closer to the border with CA1 tend to project to portions of stratum radiatum more proximal to the cell bodies, and even stratum oriens, in the closer parts of CA1 (Amaral and Witter, 1989). In contrast, the Schaffer collaterals possess very little topography along the longitudinal axis of the hippocampus (which goes by the name of the septotemporal axis in the rat, rostrocaudal in the primate)—a small injection of anterograde tracer in CA3 can spread to as much as 75% of the longitudinal extent of CA1 (Ishizuka et al., 1990; Amaral, 1993).

Another anatomical issue is *convergence*, i.e., on average, how many CA3 cells does each CA1 cell receive connections from? By a "connection" here we mean merely whether a CA1 cell receives one *or more* synapse(s) from a given CA3 cell, a connection may by means of more than one synaptic contact. In fact, recent studies of the mean number of quanta released during "minimal" stimulation suggest that Schaffer collateral fibers each connect with a CA1 cell by means of multiple release sites (Liao et al., 1992; Stricker et al., 1996; Larkman et al., 1997; but see Stevens and Wang, 1994). The number of connections received has been suggested, based on a connectivity model grounded in anterograde tracer data and Gaussian connectivity distributions, to vary slightly across the transverse axis of CA1 (Bernard and Wheal, 1994), such that mid-CA1 neurons receive more inputs (from CA3) than neurons proximal and distal to CA3.

The approach that we will take to analyze information transmission along the Schaffer collaterals is to consider them to be a kind of physiological transmission line (Shannon, 1948; Cover and Thomas, 1991; Treves, 1995). A simplified model of the physiology and anatomy of the Schaffer collaterals is established, together with the statistics of the firing rate distribution at the input to the model (the pattern of firing across the CA3 pyramidal cells), and the amount of information that could be transmitted to the CA1 pyramidal cells under these assumptions is computed. This is done by analytically obtaining a mathematical expression for the mutual information (which is independent of any particular input distribution) in the limit that the number of neurons in the field approaches infinity, and solving it numerically for a particular CA3 firing rate distribution. In some special cases, the mathematical expression can be evaluated directly without need for recourse to numerical solution.

The neural code used by the CA3 cells at the input to the system is specified by a probability distribution of symbols. In some parts of this paper, this distribution will be taken to be a binary distribution representing the presence or absence of a spike in some time period. In the section on analog resolution of coding, the distribution will more generally be taken to be a probability distribution of firing rates, which is to say that the information is represented by the number of spikes fired in a given time window. There is at least preliminary evidence that the majority (89%) of the information conveyed by hippocampal neurons about spatial stimuli is represented simply in the spike counts (Panzeri, 1996). There is also evidence that most of the information present in the spike count or rate is, in fact, present in a simple binary code, i.e., presence or absence of a spike (Panzeri, 1996; see also the discussion in the current paper).

This approach allows us to examine large rather than very small networks—infinity is a quite good approximation for the hundreds of thousands or millions of neurons in mammalian hippocampal subfields. However, it is only complementary to more detailed models of smaller networks, because of the level of physiological detail that can be incorporated into such an analysis. However, information transmission is fundamentally a network property, with information being encoded by populations of cells; therefore, our approach is well-suited to its analysis. Information as calculated from such a model is also related to quantities that can be measured from single unit and multiple single unit recordings (Treves et al., 1996) and, hence, is of neurophysiological interest in itself.

It should be noted that, in the current work, we are investigating a simple model of the physiology and anatomy of part of the hippocampus without reference to a particular theory of the function of the hippocampus. Although in part motivated by the theory of hippocampal function described in (Rolls and Treves, 1998), the assumptions for the calculation we describe here (which are spelled out in the next section) have been made on the grounds of simplicity and should be broadly compatible with a number of theories of hippocampal function (see e.g., Samsonovich and McNaughton, 1997; Murre, 1996; McClelland et al., 1995; O'Keefe, 1993; Hasselmo and Wyble, 1997). The aim of this type of exercise is to provide increasingly sophisticated methods of discriminating between theories on the basis of physiological and anatomic data. In comparison, an alternative modeling approach is to provide a precise functional specification of an idea of the role of the hippocampus (for example, "reinstatement of stored sparse random binary patterns into the entorhinal cortex"), to propose neural network models of hippocampal subregions that are capable of doing that, and to investigate parameters of the hippocampal model with regard to its ability to achieve this goal. The Schaffer collaterals were investigated by using this kind of approach in the study by Rolls (1995). The combination of both of these approaches, together with relevant experiments, should lead to a good understanding of the operation of this neural circuitry.

Analytical calculation of the information conveyed by the Schaffer collaterals allows us to investigate the effects of physiological, and particularly, anatomical parameters on information transmission. This provides an important guide for future, more detailed studies of the Schaffer collaterals, as we gain a better understanding of which parameters are likely to be important and which are not. Parameters such as the pattern of convergence onto CA1 cells, which we find to be of little informational import, are likely to have little influence on functional aspects of system behavior either. In contrast, parameters such as the expansion ratio (of the number of cells between CA3 and CA1) and the sparsity of coding are likely to be of critical functional importance,

because, as we shall show, they are of great importance to the informational characteristics of the system.

The calculation described here was motivated by that of Treves (1995). We have reproduced the results of that work, and in the present paper, we further explore the calculation, and by generalizing it in some respects, use it as a basis for examining in much greater detail the consequences of hippocampal anatomy (and, to a lesser extent, physiology) on information transmission in the Schaffer collaterals.

MATERIALS AND METHODS

A Model of the Schaffer Collateral Projection

The Schaffer collateral model describes, in a simplified form, the connections from the *N* CA3 pyramidal cells to the *M* CA1 pyramidal cells (see Fig. 1). Most Schaffer collateral axons project into the stratum radiatum of CA1, although CA3 neurons proximal to CA1 tend to project into the stratum oriens (Ishizuka et al., 1990); in the model, these projections are assumed to have the same effect on the recipient pyramidal cells. This can be considered to be a well-justified assumption, because the proximodistal position of a dendrite along a CA1 axon has been found not to affect the induced potential at the soma (Andersen et al., 1980). Inhibitory interneurons are considered to act only as regulators of pyramidal cell activity. The perforant path synapses to CA1 cells are at this stage ignored, as are the few CA1 recurrent collaterals. The system is considered for the purpose of analysis to operate in two distinct modes: storage and retrieval. During storage, the Schaffer collateral synaptic efficacies are modified by using a Hebbian rule reflecting the conjunction of pre- and postsynaptic activity. This modification has a slower time constant than that governing neuronal activity and, thus, does not affect the current CA1 output. During retrieval, the Schaffer collaterals relay a pattern of neural firing with synaptic efficacies that reflect all previous storage events. In the following, the superscript *S* is used to indicate the storage phase, and *R* is used to indicate the retrieval phase.

It is necessary that any model of the Schaffer collaterals at the level of propagation of patterns of firing rates throughout the system includes a statistical description of noise, that is to say, of those components of the firing rate that would vary on different presentations of an identical pattern. This does not imply in any way that the noise is truly random; it may come from deterministic dynamics at some level of description, but must be described statistically at this level. Although it is assumed that the distributions of noise on the inputs to the CA1 cells is Gaussian, in the limit where an infinitely large number of neurons are considered, moments of the distributions higher than the second do not contribute to the calculation described below. Thus, any reasonably well-behaved distribution could be substituted for the Gaussian with no change in the result, as long as the mean and variance were kept the same. Any higher order moments due to, for instance, the distributions of the postsynaptic response

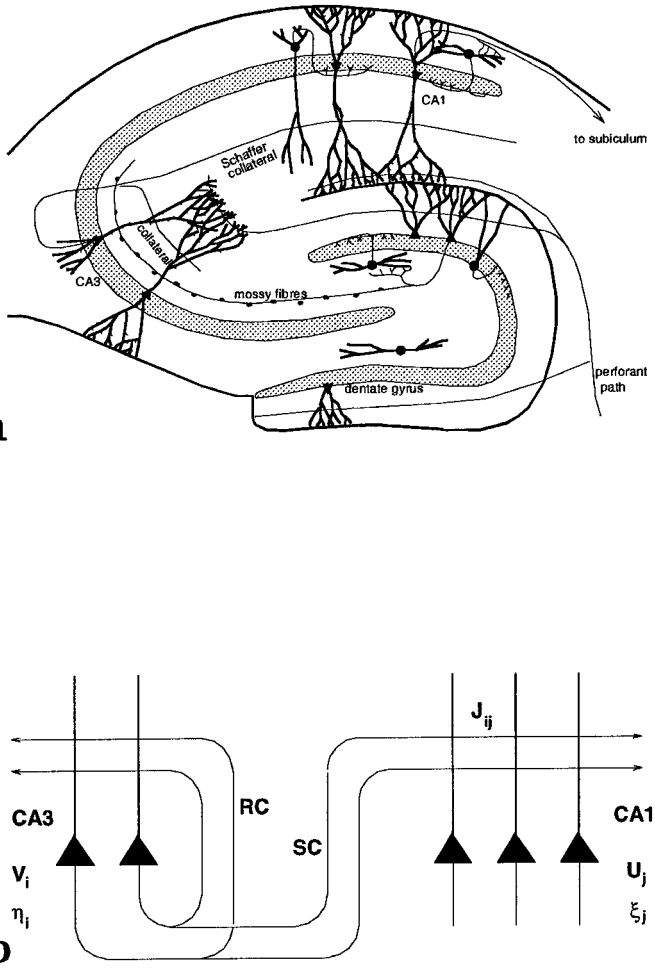

FIGURE 1. a: A schematic diagram that shows the Schaffer collaterals within the hippocampal formation. Information enters the hippocampus from layer 2 entorhinal cells by the perforant path, which projects into dentate gyrus and CA3 and CA1 areas. In addition to its perforant path inputs, CA3 receives a lesser number of mossy fiber synapses from the dentate granule cells. The axons of the CA3 pyramidal cells project commissurally, recurrently within CA3, and also forward to area CA1 by the Schaffer collateral pathway. Information leaves the hippocampus by means of backprojections to the entorhinal cortex from CA1 and the subiculum, and also by means of the fornix to the mammillary bodies and anterior nucleus of the thalamus. b: Schematic network architecture of the Schaffer collateral projection from CA3 to CA1. Please see text for explanation of terminology.

attributable to quantal synaptic release (Redman, 1990), could be expected not to affect the result as the relevant number of neurons becomes large.

The terminology we will adopt is as follows: i, CA3 cell index; j, CA1 cell index; N, Number of CA3 pyramidal cells; M, Number of CA1 pyramidal cells; η, CA3 storage firing rate vector; V, CA3 retrieval firing rate vector; ξ, CA1 storage firing rate vector; U, CA1 retrieval firing rate vector; ξ_0, CA1 storage effective threshold parameter; U_0, CA1 retrieval effective threshold parameter; c_{ij}, Connection presence matrix between CA3 and CA1; J_{ij}^S, Connection strength matrix between CA3 and CA1 during storage; J_{ij}^R,

Connection strength matrix between CA3 and CA1 during retrieval; J_{ij}^N, Modifications to connection strength matrix due to new patterns; σ_δ^2, CA3 retrieval noise variance; $\sigma_{\epsilon^S}^2$, CA1 storage noise variance; $\sigma_{\epsilon^R}^2$, CA1 retrieval noise variance; θ_μ, Retrieval lag parameter; γ, Degree of plasticity of connections; and σ_J^2, Variance of synaptic weights.

$\{\eta_i\}$ are the firing rates of each cell i of CA3 during storage. The probability density of finding a given firing pattern is taken to be:

$$P(\{\eta_i\}) = \prod_i P_\eta(\eta_i)\, d\eta_i \tag{1}$$

where η is the vector of the above firing rates. $P_\eta(\eta_i)$ is, thus, the probability density corresponding to the CA3 firing rate histogram. This assumption means that each cell in CA3 is taken to code for independent information, an idealized version of the idea that by this stage most of the redundancy present in earlier representations has been removed. Although not necessarily realistic, this testable assumption makes possible the analytical treatment of this problem. It may be possible to relax this assumption in future analyses, although there is some evidence that it may be a reasonable assumption (Treves et al., 1996). $P_\eta(\eta_i)$ is in general a continuous probability distribution (and in some theoretical analyses can be treated as such all the way through the calculation [Treves, 1990]); however, the numerics are carried out in the present work with discrete probability distributions.

$\{V_i\}$ are the firing rates of each cell i in CA3 during retrieval, and they are taken to reproduce the $\{\eta_i\}$ with some Gaussian distortion (noise), followed by rectification

$$V_i = [\eta_i + \delta_i]^+$$
$$\langle(\delta_i)^2\rangle = \sigma_\delta^2 \tag{2}$$

(the rectifying function $[x]^+ = x$ for $x > 0$, and 0 otherwise, ensures that a firing rate is a positive quantity. This results in the probability density of V_i having a point component at zero equal to the subzero contribution, in addition to the smooth component). This is to say, that the CA3 retrieved firing rates are simply the CA3 storage firing rates, with a noise distribution added, rectified so that they cannot go below zero. σ_δ can be related, e.g., to interference effects caused by the loading of other memory patterns in CA3 (see below and Treves and Rolls, 1991). This and the following noise terms are all taken to have zero means.

$\{\xi_j\}$ are the firing rates produced in each cell j of CA1, *during the storage* of the CA3 representation. They are determined for each CA1 cell by taking the sum over all connections from CA3 cells of the product of the CA3 neuron firing rate η_i with the synaptic weight J_{ij}. To this is added a Gaussian noise variable ϵ_j^S, and a threshold-linear transform (of threshold-gy_0) is applied.

$$\xi_j = \left[\xi_0 + \sum_i c_{ij} J_{ij}^S \eta_i + \epsilon_j^S\right]^+$$
$$\langle(\epsilon_j^S)^2\rangle = \sigma_{\epsilon^S}^2$$
$$\langle(J_{ij}^S)^2\rangle = \sigma_J^2. \tag{3}$$

The synaptic matrix is very sparse as each CA1 cell receives inputs from only C_j (of the order of 10^4) cells in CA3. The average of C_j across cells is denoted as C

$$c_{ij} = \{0, 1\}$$
$$\langle c_{ij}\rangle N = C_j \qquad (C \equiv \langle C_j\rangle). \tag{4}$$

$\{U_j\}$ are the firing rates produced in CA1 by the pattern $\{V_i\}$ retrieved in CA3

$$U_j = \left[U_0 + \sum_i c_{ij} J_{ij}^R V_i + \epsilon_j^R\right]^+$$
$$\langle(\epsilon_j^R)^2\rangle = \sigma_{\epsilon^R}^2$$
$$\langle\langle J_{ij}^R\rangle^2\rangle = \sigma_J^2. \tag{5}$$

Each weight of the synaptic matrix during retrieval of a specific pattern,

$$J_{ij}^R = \cos(\theta_\mu)\, J_{ij}^S + \gamma^{1/2}(\theta_\mu)\, H(\eta_i, \xi_j) + \sin(\theta_\mu)\, J_{ij}^N \tag{6}$$

consists of (1) the original weight during storage, J_{ij}^S, damped by a factor $\cos(\theta_\mu)$, where $0 < \theta_\mu < \pi/2$ parameterizes the time elapsed between the storage and retrieval of pattern μ (μ is a shorthand for the pattern quadruplet $\{\eta_i, V_i, \xi_j, U_j\}$); (2) the modification due to the storage of μ itself, represented by a Hebbian term $H(\eta_i, \xi_j)$, normalized so that

$$\langle\langle H(\eta, \xi)\rangle^2\rangle = \sigma_J^2; \tag{7}$$

γ measures the degree of *plasticity*, i.e., the mean square contribution of the modification induced by one pattern, over the overall variance, across time, of the synaptic weight; (3) the superimposed modifications J^N reflecting the successive storage of new intervening patterns, again normalized such that

$$\langle\langle J_{ij}^N\rangle^2\rangle = \sigma_J^2. \tag{8}$$

A factor of $\sin(\theta_\mu)$ is needed to ensure normalization of the stored and new synaptic strengths.

The mean value across all patterns of each synaptic weight is taken to be equal across synapses and, therefore, is taken into the threshold term. The synaptic weights J_{ij}^R and J_{ij}^S are, thus, of zero mean and variance σ_J^2 (all that affects the calculation is the first two moments of their distribution). Note that (obviously) since the mean synaptic strength is taken into the threshold term, this does not imply inhibitory synapses; the model is still one of excitatory synaptic transmission.

A plasticity model is used that corresponds to gradual decay of memory traces as new patterns are stored. Numbering memory patterns from 1, . . ., λ, . . ., ∞ backward, the model sets $\cos(\theta_\lambda) = \exp(-\lambda_{\gamma 0}/2)$ and $\gamma(\theta_\lambda) = \gamma_0 \exp(-\lambda\gamma_0)$. Thus, the strength of older memories fades exponentially with the number of intervening memories. The same forgetting model is assumed to apply to the CA3 network, and for this network, the maximum number of patterns can be stored when the plasticity $\gamma_0^{CA3} = 2/C$ (Treves, 1995).

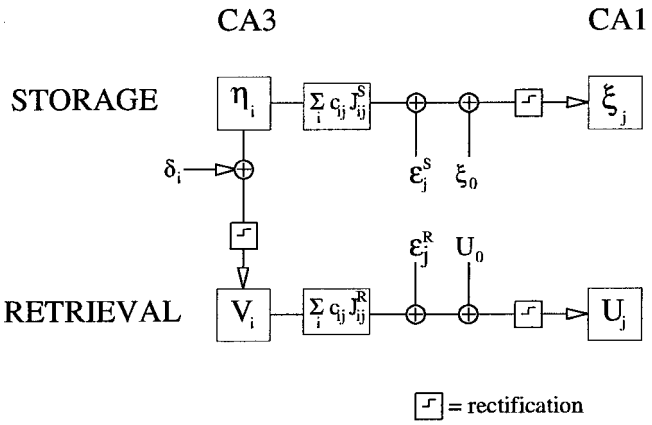

FIGURE 2. A block diagram illustrating the relationships between the variables present in the model. The input of the system could be considered to be the CA3 pattern during storage, η, and the output the CA1 pattern during retrieval, U. J_{ij}^R depends on J_{ij}^S, ξ_j, and η_i, as described in the text.

For the Hebbian term, the specific form

$$H(\eta_i, \xi_j) = \frac{h}{\sqrt{C}} (\xi_j - \xi_0) (\eta_i - \langle \eta_i \rangle) \tag{9}$$

is used, where h ensures the normalization given in Eq. 7.

The thresholds ξ_0 and U_0 are assumed to be of fixed value in the following analysis. This need not be the case, however, and as far as the model represents (in a simplified fashion) the real hippocampus, they might be considered to be tuned to constrain the sparseness of activity in CA1 in the storage and retrieval phases of operation, respectively, reflecting inhibitory control of neural activity. These threshold parameters have been set so that the neurons operate in a threshold-linear régime. Linear neurons can be simply recovered from the equations by taking the limit in which ξ_0 and U_0 approach positive infinity (such that the neurons never operate below threshold). A discussion of the effects of varying the threshold in this type of model can be found in (Schultz and Treves, 1998).

The block diagram shown in Figure 2 illustrates the relationships between the variables described in the preceding text.

Mutual Information Analysis

The aim of the analysis is to calculate how much, on average, of the information present in the original pattern $\{\eta_i\}$ is still present in the effective output of the system, the pattern $\{U_j\}$, i.e., to average the mutual information

$$I(\{\eta_i\}, \{U_j\}) = \int \prod_i d\eta_i \int \prod_j dU_j P(\{\eta_i\}, \{U_j\}) \ln \frac{P(\{\eta_i\}, \{U_j\})}{P(\{\eta_i\})\, P(\{U_j\})} \tag{10}$$

over the variables c_{ij}, J_{ij}^S, J_{ij}^N. An outline of the calculation is given here, which should be sufficient for those with an appropriate background to reproduce the results. Those not familiar with replica calculations may refer to the final chapter of (Hertz et al., 1991), the appendices of (Rolls and Treves, 1998), and the book by Mezard et al. (1987) for background material. The full mathematical details of a slightly simpler model (see later) are presented in (Schultz and Treves, 1998); a preliminary exploration of the full model is described by Treves (1995).

$P(\{\eta_i\}, \{U_j\})$ is written (simplifying the notation) as

$$P(\eta, U) = P(U|\eta)P(\eta) = \int_V \int_\xi dV\, d\xi P(U|V, \xi, \eta)P(V|\eta)\, P(\xi|\eta)\, P(\eta) \tag{11}$$

where the probability densities implement the model defined above.

The average mutual information is evaluated using the replica trick, which amounts to

$$\log P = \lim_{n \to 0} \frac{(P^n - 1)}{n} \tag{12}$$

which involves a number of subtleties, for which (Mezard et al., 1987) can be consulted for a complete discussion. The important thing to note is that it is necessary to make an assumption about the form of symmetry of those variables to which are assigned two replica indices. The assumption made is that they are, in fact, symmetric. This assumption has been verified for a simple "transmission only" (no connection strength adaptation) version of the model to be valid down to noise variances sufficiently low that the mutual information is saturated at its ceiling (Schultz and Treves, 1998).

The expression for mutual information, thus, becomes

$$\langle I(\eta, U) \rangle_{c,J^S,J^N} = \lim_{n \to 0} \frac{1}{n} \left\langle \int d\eta\, dU P(\eta,U) \times \left\{ \left[\frac{P(\eta, U)}{P(\eta)} \right]^n - [P(U)]^n \right\} \right\rangle_{c,J^S,J^N} \tag{13}$$

where it is necessary to introduce $n + 1$ replicas of the variables δ_i, ϵ_j^S, ϵ_j^R, V_i, ξ_j, and, for the second term in curly brackets only, η_i.

The core of the calculation then is the calculation of the probability density $\langle P(\eta, U)^{n+1} \rangle$. The key to this is "self-consistent statistics" (Rolls and Treves, 1997; appendix 4): all possible values of each firing rate in the system are integrated, subject to a set of constraints that implement the model. The constraints are implemented using the integral form of the Dirac δ-function. Another set of Lagrange multipliers introduces macroscopic parameters

$$\begin{aligned} x^\alpha &= \frac{1}{N} \sum_i \frac{(\eta_i^\alpha - \langle \eta \rangle)}{\langle \eta \rangle} V_i^\alpha \theta(V_i^\alpha) \\ w^{\alpha\beta} &= \frac{1}{N} \sum_i \eta_i^\alpha V_i^\beta \theta(V_i^\beta) \\ y^{\alpha\beta} &= \frac{1}{N} \sum_i V_i^\alpha V_i^\beta \theta(V_i^\alpha)\, \theta(V_i^\beta) \\ z^{\alpha\beta} &= \frac{1}{N} \sum_i \eta_i^\alpha\, \eta_i^\beta \end{aligned} \tag{14}$$

where $\theta(x)$ is the Heaviside function, and α, β are replica indices.

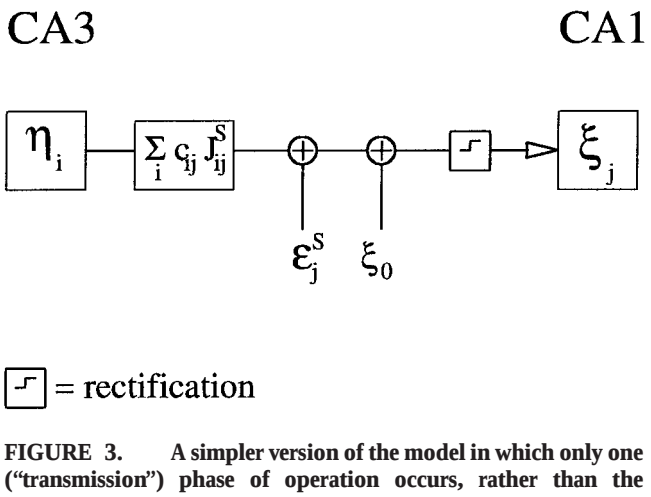

FIGURE 3. A simpler version of the model in which only one ("transmission") phase of operation occurs, rather than the separate storage and recall phases. In this simpler model, there is no Hebbian plasticity in the Schaffer collaterals. Please see text for explanation of terminology.

Making the assumption of replica symmetry, and performing the integrals over all microscopic parameters, with some algebra an integral expression is obtained for the average mutual information per CA3 cell. This integral over the macroscopic parameters and their respective Lagrange multipliers is evaluated by using the saddle-point approximation, which is exact in the limit of an infinite number of neurons (see, for example, Jeffreys and Jeffreys, 1972) to yield the expression given in Appendix A; the saddle-points of the expression must in general be found numerically. The parameters used for the numerical solutions are given in Appendix B.

Topographic connectivity is incorporated by taking the Schaffer collaterals to be composed of K elemental subregions. The connectivity C is generalized to $C_{\lambda\nu}$, which denotes the number of connections that a cell in subregion λ of CA1 receives from subregion ν in CA3. The Lagrange multipliers $w^{\alpha\beta}$, $y^{\alpha\beta}$, and $z^{\alpha\beta}$, thus, become $w^{\alpha\beta}_{(\nu)}$, $y^{\alpha\beta}_{(\nu)}$, and $z^{\alpha\beta}_{(\nu)}$. The information expression in Appendix A is stated in full generality, although in this study results for topography in the full Schaffer model were obtained by Monte Carlo simulation, and the analytical expression was used only in the special case discussed below, and for the full Schaffer collateral model without topography.

In a study by Schultz and Treves (1998), a special case of this calculation was considered, in which no plasticity was present in the Schaffer collaterals. In this situation, the retrieval phase of operation is dispensed with, and the storage phase becomes a single "transmission" phase of operation. This method simplifies the equations considerably. A block diagram illustrating this simpler model is shown in Figure 3. Further limiting cases were studied, and it was shown that in the limit of linear neurons, for low signal to noise ratio (i.e., very noisy activity), the information capacity of a simple Gaussian channel provides an upper bound on the transmitted information

$$I_{gauss} = \frac{r}{2} \ln \left[1 + \frac{\sigma_j^2 C(\langle \eta^2 \rangle_\eta - \langle \eta \rangle_\eta^2)}{\sigma_\epsilon^2} \right]. \qquad (15)$$

This upper bound is met for an optimal η (pattern) distribution and vanishing dependence on the same inputs of the output cells.

Monte Carlo Simulation

One issue was investigated by using a simulation of the model, rather than an explicit analytical calculation of the information. This issue was the effect of topography in the full Schaffer collateral model (as opposed to the single-phase model, in which the analytical calculation was used). To do this, information was calculated from simulated firing rates in the same way as that used to analyze single-unit neurophysiologic data (Heller et al., 1995; Rolls et al., 1997). This was a similar technique to that developed by Optican and Richmond (1987), except that an erroneous procedure for correcting the bias due to limited sampling in that study was replaced with an analytical evaluation of the bias (Miller, 1955; Panzeri and Treves, 1996) (a comparison of this approach and an alternative empirical procedure can be found in Golomb et al., 1997).

The actual simulation was very simple: patterns were created at random, and spatial correlations were generated by the following algorithm. The first cell took the value 1 with probability *P*, 0 with $1 - P$. The next cell took the value of the first cell with probability q, but with probability $1 - q$, its own random value was chosen (again 1 with probability *P*). This process was continued unidirectionally across the whole field of cells. The mean half-width in cells corresponding to a given q is thus $-\log 2/\log q$. Random connectivities were generated, and for each random connectivity, 50 trials of responses were generated according to the activity equations presented in the model description. The whole procedure was nested inside a Monte-Carlo loop, which executed 20 times and averaged the results.

The average information about all patterns s from the set S present in the set R of resonses r is given by:

$$I(S, R) = \sum_{s,r} P(s, r) \log_2 \frac{P(s, r)}{P(s)\, P(r)}, \qquad (16)$$

to which is added the leading term of the correction for the finite sampling bias:

$$\frac{1}{2N_T \ln 2} \left(\sum_s \tilde{R}_s - \tilde{R} - (S - 1) \right). \qquad (17)$$

$\tilde{R}_s$ is the number of response bins with nonzero occupancy probability $P(r = i|s)$, and $\tilde{R}$ is the number of response bins with nonzero occupancy probability $P(r = i)$.

It should be noted that this is an entirely different information measure than that computed analytically in the previous section, i.e., information about *which* pattern is present at the input of the system, rather than information about the firing rates that make up the pattern. This must be kept in mind. The main import of this, however, is that the two measures have different ceilings, i.e., the information about *which* pattern is bounded by the logarithm of the number of patterns that must be discriminated between and the true mutual information about the pattern by the actual entropy of the pattern. However, the two measures can be

expected to vary in a similar manner with changes in most parameters, and they are certainly conceptually similar quantities.

ANALOG RESOLUTION OF CODING

Investigation

Specification of the probability density $P(\eta)$ allows different distributions of firing rates in CA3 to be considered in the analysis. Clearly, the distribution of firing rates that should be considered in the analysis is that of the firing of CA3 pyramidal cells, computed over the time constant of storage (which we can assume to be the time constant of LTP), during only those periods in which biophysical conditions are appropriate for learning to occur. Unfortunately, this last caveat makes a simple fit of the firing-rate distribution from single-unit recordings fairly meaningless unless the correct assumptions regarding exactly what these conditions are in vivo can be made. Possibly the most useful thing we can do for the present is to assume that the distribution of firing rates during storage is graded, sparse, and exponentially tailed. This accords with the observations of neurophysiologists (Barnes et al., 1990; Rolls et al., 1998). The easiest way to introduce this to the current investigation is by means of a discrete approximation to the exponential distribution (a geometric distribution), with extra weight given to low firing rates. This method allows quantitative investigation of the effects of analog resolution on the information transmission capabilities of the Schaffer collateral model.

Although $P(\eta)$ is in general a continuous distribution, utilization of discrete distributions allows the issue of analog resolution to be addressed. The limiting cases are obviously the binary distribution and a continuous distribution such as an exponential distribution. The required discrete CA3 firing rate distributions were formed by the mixture of the unitary distribution and the discretised exponential, by using as mixture parameters the offset ϵ between their origins, and relative weightings. The distributions were constrained to have first and second moments $\langle\eta\rangle$, $\langle\eta^2\rangle$, and thus sparseness $\langle\eta\rangle^2/\langle\eta^2\rangle$, equal to a. In the cases considered here, a (CA3 sparseness) was allowed values of 0.05, 0.10, and 0.20 only. The threshold ξ_0 was set to a value (-0.4) for which the output (CA1) sparseness is approximately equal to the input sparseness. The width of the distribution examined was set to 3.0, and the number of discretized firing levels contained in this width parameterized as L. The binary distribution was completely specified by this; for distributions with a large number of levels, there was some degree of freedom, but its numerical effect on the resulting distributions was essentially negligible. Those distributions with a small number of levels ≥ 2 were non-unique and were chosen fairly arbitrarily for the following results, to be those that had entropies interpolating between the binary and large L situations. Some examples of the distributions used are shown in Figure 4a. For comparison Figure 4b shows the distribution of firing rates from a primate CA3 neuron.

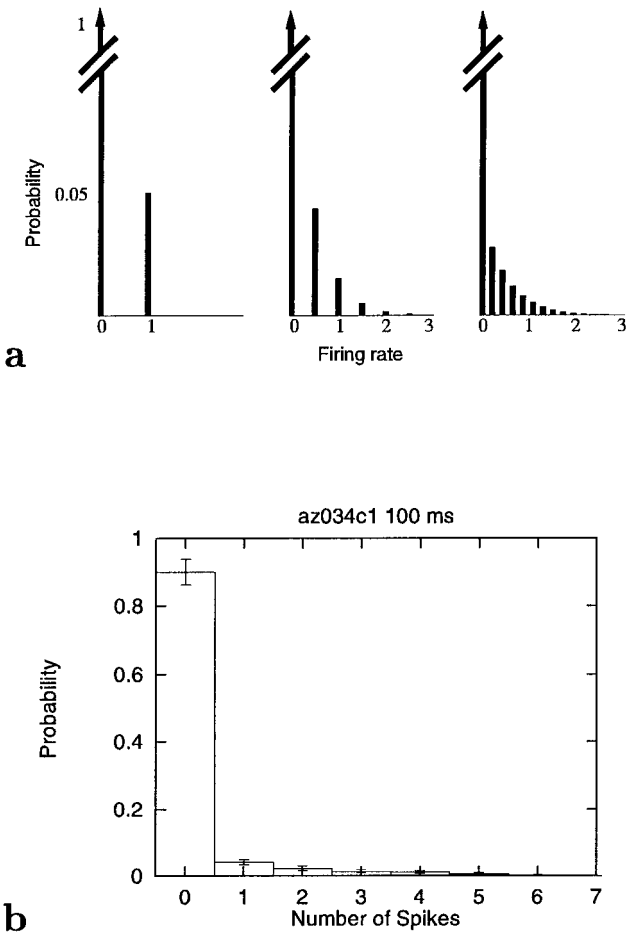

FIGURE 4. a: Some of the CA3 firing rate distributions used in the analysis. These are, in general, formed by the mixture of a unitary distribution and a discretized exponential. b: The distribution of the number of spikes emitted by a CA3 neuron in 100-ms periods. Taken from Rolls et al. (1998).

The total entropy per cell of the CA3 firing pattern, given a probability distribution characterised by L levels, is

$$h(\eta) = -\sum_{l=1}^{L} P_{\eta l}\langle\eta l\rangle \ln P_{\eta l}\langle\eta l\rangle. \quad (18)$$

RESULTS

The results are shown in Figure 5. The entropy present in the CA3 firing rate distributions is marked by asterisks. The mutual information conveyed by the retrieved pattern of CA1 firing rates, which must be strictly less than the CA3 entropy, is represented by circles. Note that for sparsity $a = 0.2$, it was not possible with the available computer time to solve the saddle-point equations for CA1 mutual information (circles) for 2 and 10 levels of analog resolution. However, in both of these cases, it is quite apparent where the points must lie (between the plus signs and asterisks

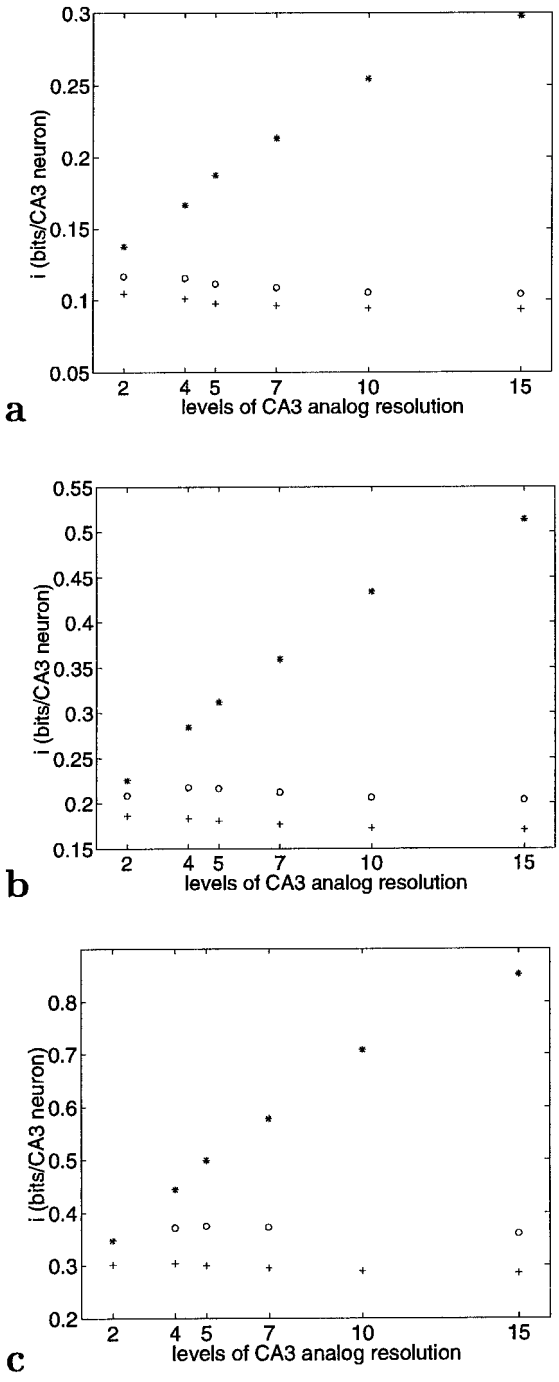

FIGURE 5. The mutual information between patterns of firing in CA1 and patterns of firing in CA3. Asterisks represent the entropy of the CA3 pattern distribution; circles represent the CA1 retrieved mutual information; and plus signs represent the CA1 information during the storage phase. The horizontal axis parameterizes the number of discrete levels in the input distribution: for codes with fine analog resolution, this is greater. (a) a = 0.05 (sparse); (b) a = 0.10; and (c) a = 0.20 (slightly more distributed).

because of physical constraints and near a line interpolated from the surrounding points). For more distributed codes than a = 0.2, we were also unable to solve the saddle-point equations numerically.

It is apparent that maximum information efficiency occurs in the binary limit for CA3 patterns. More remarkably, even in absolute terms, the information conveyed is maximal for low-resolution codes, at least for quite sparse codes. The results are qualitatively consistent over sparsenesses a ranging from 0.05 to 0.2; obviously with higher a (more distributed codes), entropies are greater. For more distributed codes (i.e., with signalling more evenly distributed over neuronal firing rates), it seems that there may be some small absolute increase in information with the use of analog signalling levels.

For comparison, the crosses in the figures show the information stored in CA1. This was computed using a simpler version of the calculation, in which the mutual information $i(\{\eta_i\}, \{\xi_j\})$ was calculated. Obviously, in this calculation, the CA3 and CA1 retrieval noises σ_δ and σ_ϵ^R are not present; on the other hand, neither is the Schaffer collateral memory term. Because the retrieved CA1 information is in every case higher than that stored, we can conclude that for the parameters considered, the additional Schaffer memory effect outweighs the deleterious effects of the retrieval noise distributions.

It follows from the forgetting model defined by Eq. 6 that information transmission is maximal when the plasticity (mean square contribution of the modification induced by one pattern) is matched in the CA3 recurrent collaterals and the Schaffer collaterals (Treves, 1995). It can be seen in Figure 6 that this effect is robust to the use of more distributed patterns.

ANATOMICAL CONVERGENCE

Investigation

It is assumed in a study by Treves (1995), that there is uniform convergence of connections from CA3 to CA1 across the extent of the CA1 subfield. In reality, of course, this finding may not be the case. Each CA1 pyramidal neuron may not receive the same

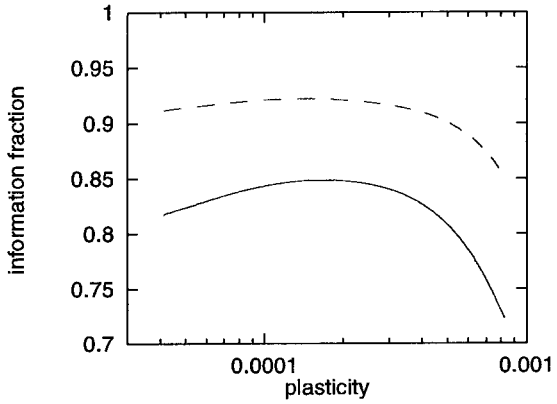

FIGURE 6. The dependence of information transmission on the degree of plasticity in the Schaffer collaterals, for $\alpha = 0.05$ (solid) and $\alpha = 0.10$ (dashed). A binary pattern distribution was used in this case.

number of connections from CA3: this quantity might vary across the transverse extent of CA1. Bernard and Wheal (1994) investigated this with a connectivity model constructed by simulating a *Phaseolus vulgaris* leucoagglutinin–labeling experiment, matched to the available anatomic data, with an assumption of a Gaussian distribution in the connectivity of any given neuron. Their conclusion was that mid-CA1 neurons receive more connections (8,000) than those in proximal and distal[1] CA1 (6,500). They assumed five synaptic contacts per fiber (a little more than the mean suggested by recent evidence; Larkman et al., 1997). The precise numbers are not important here; what *is* of interest is to consider the effect on information transmission of this spread in the convergence parameter C_j about its mean C, and specifically to see whether this relative amount of spread is significantly different to uniformity in information terms.

In this analysis, σ_j^2 is set to 1/C for all cells in the network, i.e., the neurons have unit gain. C is set by using the assumption of parabolic dependence of C_j on transverse extent, on the basis of Figure 5 of (Bernard and Wheal, 1994). To facilitate comparison with the results reported in Treves (1995), C is held at 10,000 for all results in this section. The model used (which we will refer to as the "realistic convergence" model), thus, is simply a scaled version of that due to Bernard and Wheal, with C_j = 7,143 at the proximal and distal edges of CA1, and C_j = 11,429 at the midpoint. Note that this refers to the number of CA3 cells contacting each CA1 cell; each may do so by means of more than one synapse. This convergence distribution was approximated in the numerical evaluation by 10 discrete convergence levels fitted to the above curve.

The saddle-point expression (Eq. 19) was evaluated numerically while varying the plasticity of the Schaffer connections, to give the relationships shown in Figure 7a between mutual information and γ_0^{CA1}. The information is expressed in the figure as a fraction of the information present when the pattern is stored in CA3 (Eq. 18).

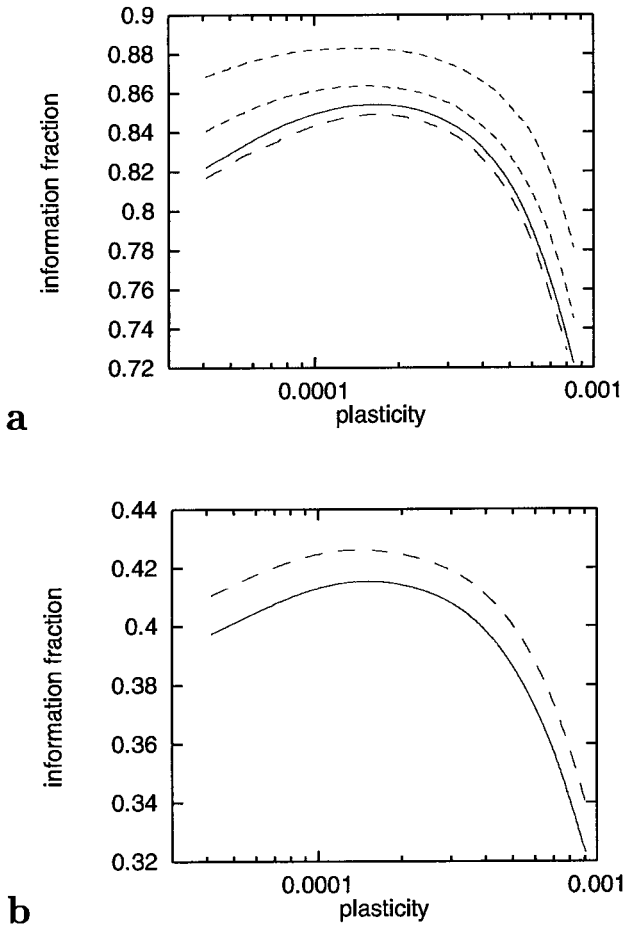

FIGURE 7. Information transmitted as a function of Schaffer collateral plasticity. a: Binary CA3 firing rate distributions. The solid line indicates the result for the realistic convergence model. The dashed lines indicate, in ascending order (1) uniform convergence, (2) two-tier convergence model with $C_j \in \{5000, 15000\}$, and (3) two-tier convergence model with $C_j \in \{2000, 18000\}$. b: With more realistic CA3 firing-rate distributions (the 10-level discrete exponential approximation from the previous section). The solid line indicates the result for uniform connectivity, and the dashed line represents the two-tier convergence model with $C_j \in \{5000, 15000\}$.

Results

Two phenomena can be seen in the results. The first, as mentioned in the previous section (and discussed at more length in Treves, 1995), is that information transmission is maximal when the plasticity of the Schaffer collaterals is approximately matched with that of the preceding stage of information processing. The second phenomenon is the increase in information throughput with spread in the convergence about its mean. The increase in mutual information provided by those CA1 neurons with a greater number of connections than the mean more than compensates for the decrease in those with less than the mean. It must be remembered that what is being computed is the information provided by *all* CA1 cells about patterns of activity in CA3. In any case, the effect is rather small here: the realistic convergence model allows the transmission of only marginally more information than the uniform model. The uniform convergence approximation might be viewed as a reasonable one for future analyses, then.

Figure 7b shows that the situation for graded pattern distributions is almost identical. The numerical fraction of information transmitted is of course lower (but total transmitted information is similar; see previous section). The uniform and two-tier convergence models provide bounds between which the realistic case must lie.

EXPANSION RATIO

Another issue that this model allows us to address is the expansion ratio of the Schaffer collaterals, i.e., the ratio between

[1]Note that by proximal CA1, we refer to that part closest to CA3, and by distal, we refer to that part farther away from CA3, following the terminology of Amaral and Witter (1989).

the numbers of neurons in CA1 and CA3, M/N. It can be seen in Figure 8 that an expansion ratio of 2 (a "typical" biological value) is sufficient for CA1 to capture most of the information of CA3 and that, although the gains for increasing this are diminishing, there is a rapid drop-off in information transmission if it is reduced by any significant amount. The actual expansion ratio for different mammalian species reported in the literature is subject to some variation, with the method of nuclear cell counts giving ratios between 1.4 (Long-Evans rats) and 2.0 (humans) (Seress, 1988), whereas stereologic estimates range from 1.8 (Wistar rats) to 6.1 (humans) (West, 1990). It should be noted that in all these estimates, and particularly with larger brains, there is considerable error. However, in all cases the Schaffer collateral model seems to operate in a régime in which there is at least the scope for efficient transfer of information.

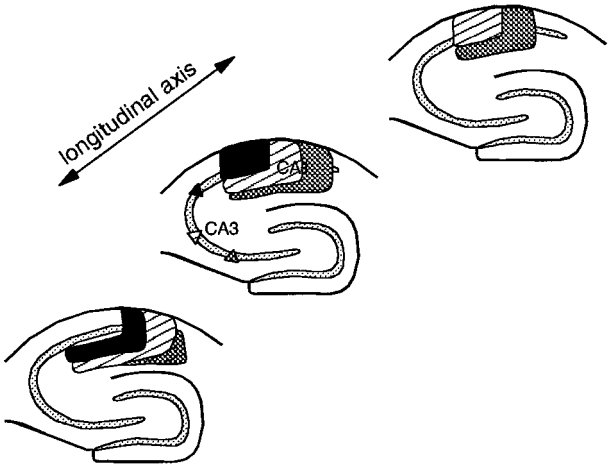

FIGURE 9. An illustration of the transverse topography of the Schaffer collaterals, as redrawn from (Amaral and Witter, 1995). CA3 cells closer to the border with CA1 tend to project just across into CA1, whereas CA3 cells close to the dentate gyrus tend to project to distal CA1, near the subiculum.

TOPOGRAPHIC CONNECTIVITY

Investigation

A noticeable anatomical feature of the Schaffer collaterals (as indeed of many other cortical projections), is the topography that exists in their connectivity. This is to say that, rather than each CA1 neuron receiving synapses from cells across the whole extent of CA3, there is a focal point in CA3 from which it receives the most synapses, and a (presumably Gaussian) spread such that the further a CA3 cell is from the focal point, the less likely the CA1 cell is to receive synapses from it. As summarized in the Introduction section, topography in the Schaffer collaterals exists to a far greater extent in the transverse axis than in the longitudinal axis of the hippocampus. It is this topography that will be investigated here, and it is illustrated in Figure 9.

Transverse topography was investigated in two ways. The first, valid only for CA3 firing rate patterns which were purely uncorrelated, was to numerically evaluate the mutual information expression in the topographic case. This was performed with $K = 50$ elemental topographic subregions. The connectivity matrix $C_{\lambda\nu}$ utilised was symmetric, with Gaussian spread σ, and with the total number of connections any neuron received normalised to $C = 10{,}000$. This was done only for the single phase model, rather than the full Schaffer collateral model. The reason for this is that it was desirable to at some point consider CA3 patterns which had some degree of spatial correlation amongst their firing rates. Although it may be possible to include spatially correlated patterns in such a mathematical analysis, it vastly complicates matters. Therefore, it was decided to investigate this particular issue in the full Schaffer collateral model by making use of a Monte Carlo simulation of the model, as described in the Materials and Methods section of the paper.

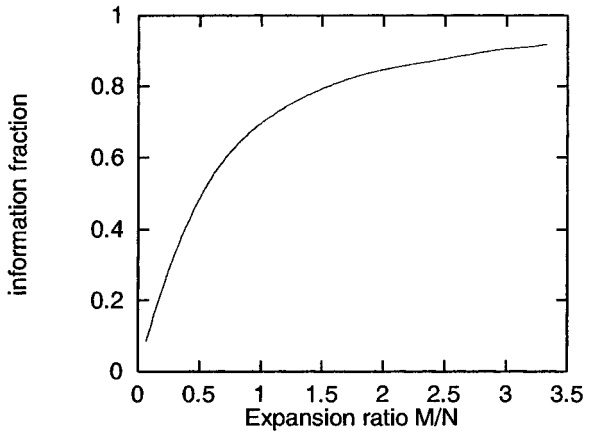

FIGURE 8. The dependence of the fraction of information transmitted on the expansion ratio (the number of CA1 cells divided by the number of CA3 cells).

Results

The resulting mutual information per CA3 cell for the single-phase model (Fig. 10) does not, to a very great extent at all, depend on the degree of spread of topography. For high topographic spread, connections are received from a wide area, and the information approaches that available with no topography at all. As the topography becomes more tightly constrained, the information drops somewhat, to a minimum at which the Gaussian spread is around $\sigma = 1/3$; but, below this, it increases again. The total magnitude of the deviation from the information available without any topography is very small. The results, thus, indicate that, in the absence of spatial correlations, topographic connectivity has very little effect on the amount of information that can be relayed by the Schaffer collaterals.

However, this relative constancy of the information conveyed is disrupted in the presence of spatially correlated patterns, at least as measured from the information about which pattern was present in the Monte Carlo simulation of the full Schaffer collateral model. The results of the simulation are shown in Figure 11. Note

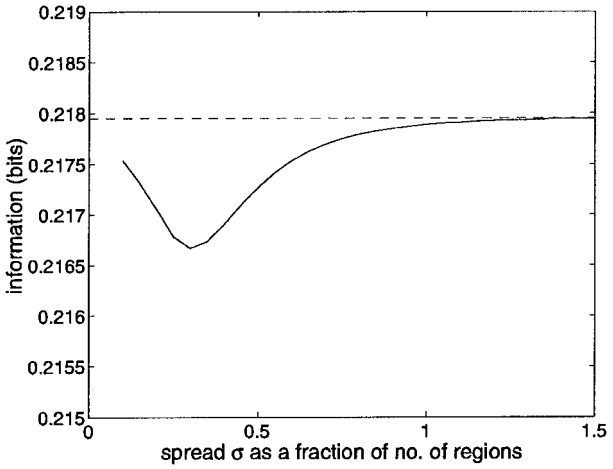

FIGURE 10. Information transmitted per CA3 cell, as a function of the topographic spread (the standard deviation of the spatial gaussian distribution of connections, expressed as a fraction of the total field width). The dashed line represents the information that would be available if no topography was present at all, i.e., the connectivity was uniform.

that this figure cannot be directly compared with Figure 10, because they refer to information about different things. In this figure, the ratio of transmitted to total information is plotted as a function of 1 − q. q, as mentioned in the Materials and Methods section, is the probability of a cell in the pattern sharing the value of the proceeding cell; the higher 1 − q, the more narrowly the correlations are confined in space. By plotting logarithmically versus 1 − q, we can see the dependence of information transmission on the spatial spread of the correlations. The first obvious point from the figure is that for higher amounts of spatial correlation, more information can be extracted. This may be a point specific to the information measure, i.e., with higher

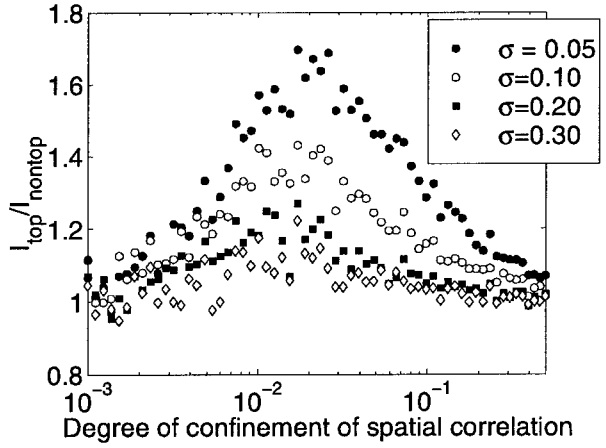

FIGURE 11. The ratio of transmitted information with and without topographic connectivity, as a function of the extent of spatial correlation internal to the patterns. Filled circles, topographic spread σ = 0.05; open circles, σ = 0.1; squares, σ = 0.2; diamonds, σ = 0.3. More widespread correlations are represented to the left of the plot, more narrowly confined to the right. As the spread σ becomes greater than 0.3, the information gain curve flattens out.

correlations, more information can be carried about which of the patterns is active at the input. It is not expected that this should necessarily carry over into the more general measure of information about the entire set of firing rates within the input pattern, because the entropy of the input pattern of course decreases with internal pattern correlations. The more important point that is apparent from the figure is that as the spread of the topography becomes higher (progressively lower curves on the figure), the maximum in the transmitted information moves to the left on this scale, i.e., for higher topographic spreads, the maximum is for patterns with wider spread in the spatial extent of the patterns. Transmission of information seems to be maximal when the spread of topography and the extent of spatial correlations are tuned to be matched. It is to be expected that this principle should carry over to the full information.

DISCUSSION

This study has examined quantitatively the effect of analog coding resolution on the total amount of information that can be transmitted in a model of the Schaffer collaterals. The tools used were analytical and numerical, and the focus was mostly on relatively sparse codes. What can these results tell us about the actual code used to signal information in the mammalian hippocampus? In themselves, of course, they can make no definite statement. It could be that there is a very clear maximum for information transmission in using binary codes for the Schaffer collaterals, and yet external constraints such as CA1 efferent processing might make it more optimal overall to use analog signalling. So, results from a single component study must be viewed with due caution. However, these results can provide a clear picture of the operating régime of the Schaffer collaterals, and that is after all a major aim of any analytical study.

The results from this paper bring out some interesting points. For instance, it is very clear from Figure 6 that, although nearly all of the information in the CA3 distribution can be transmitted by using a binary code, this information fraction drops off rapidly with analog level. The total amount of information transmitted is similar regardless of the amount of analog level to be signalled, but this is a well known and relatively general fact, and accords with common-sense intuition (Softky, 1995). However, the total amount of information that can be transmitted is only *roughly* constant. It seems, from this analysis, that although the total transmitted information drops off slightly with analog level for very sparse codes, the maximum moves in the direction of more analog levels for more evenly distributed codes. This finding provides some impetus for making more precise measurements of sparseness of coding in the hippocampus. The reason for improved efficiency with more analogue levels at higher sparsity of coding can be easily understood by closer inspection of Figure 5. The effect of plasticity is to provide an improvement to the amount of information that can be transmitted. This extra information is greater for more distributed coding. The extra

information provided by the plasticity effect is, all other things being equal, relatively constant across the number of levels of analog resolution. However, when the number of levels becomes very small, the pattern entropy that provides a ceiling on the transmitted information is quite low. This means that with more distributed coding, the added information due to plasticity pushes the transmitted information into the region of saturation due to the ceiling; thus, at sparsity $a = 0.2$, we see a situation for which more information can be transmitted at four or five levels of analogue resolution. This effect is strictly due to the presence of plasticity—without plasticity, binary coding would maximize transmitted information.

A further point worth discussing is the functional role of topography, as opposed to its quantitative import with regard to information transmission. Given that it seems from our theoretical study here that, in the absence of spatially correlated patterns in CA3, the presence of topography does not greatly affect the amount of information that can be conveyed by the Schaffer collaterals, any explanation for the gradients of connectivity (in which we will include topography and convergence patterns) based on the mere quantity of information transmitted can be dismissed. Topography would seem, thus, to be merely a minimum wiring phenomenon: cortical regions need only certain types of information, and they get it only from certain places.

The implications of this can be clarified slightly with regard to the efferent microanatomy of the hippocampus. The hippocampus has been proposed as the point of integration of two processing streams representing visual and spatial information (Smith and Milner, 1981; Ungerleider and Mishkin, 1982). In particular it has been suggested that the CA3 subregion of the hippocampus, by virtue of the extensive, associative (i.e., recurrent) connectivity of the pyramidal cells within it, could be the only site where object and place information are entirely integrated, or almost so (Rolls, 1989; Treves and Rolls, 1994; Rolls and Treves, 1998). Although the connectivity within CA3 is such that there is essentially a single CA3 recurrent network, however, it should not be thought that it is entirely uniform, and this idea bears on the current discussion. In actual fact, cells proximal in CA3 are more limited in their projections, tending to project mostly within their own subfield, although they receive connections from all over CA3. In contrast, mid and distal CA3 cells project throughout the transverse axis and much more extensively longitudinally (Hjorth-Simonsen, 1973; Swanson et al., 1978; Ishizuka et al., 1990). However, although there are connections throughout CA3, it now seems that most of the connections a CA3 neuron receives are from its own subregion (CA3a,b,c, where c is the proximal and a the distal subregion) (Li et al., 1994). This could not really be called topography, because there is too much spread, but it is certainly the beginnings of one, which is reflected in a more full-blown topography in the CA3-CA1 projection. Most of the outputs of CA1 are to the subiculum, and there is in turn substantial topography in this projection (Tamamaki et al., 1987; Tamamaki and Nojyo, 1990, 1995). Within the subiculum, different populations of cells project to different target structures or to different areas within the target structures (Swanson and Cowan, 1977; Groenewegen et al., 1982, 1987; Van Groen et al., 1986; Witter et al., 1990). These populations of cells in some cases overlap (e.g., projections to the entorhinal cortex and septal region originate in a restricted population of cells, which overlap on the proximodistal axis) and in some cases do not (e.g., projections to retrosplenial cortex being completely separated). Thus, there is certainly scope for particular parts of the CA3 network to be selected preferentially for output to individual cortical and subcortical structures on the basis of the type of information represented in them, the degree of topographic segregation of the information increasing through CA3, CA1, subiculum, and so on. Topography at the level of the Schaffer collaterals thus, is merely a reflection of gross neuroanatomy. This does suggest some degree of inhomogeneity in the way in which information is represented in the CA3 network. Note, however, that inhomogeneity in the representation does not in any way suggest correlation in neuronal firing in any one time period.

The possibility should be considered that, as a result of active, inhibitory channelling, there should be more precise *functional* topography than that suggested by the anatomic studies. However, in the case of the longitudinal axis of the hippocampus electrophysiology supports the anterograde tracer studies (Finnerty and Jefferys, 1993), and there is no reason at this stage to believe that the situation should be different across the transverse axis.

The study of the expansion ratio (the number of CA3 cells divided by the number of CA1 cells) indicated that there is little advantage in information terms for increasing the expansion ratio beyond about 2. If we take the expansion ratio measurement of 6 in humans to be correct, then the Schaffer collaterals as a single network would be operating well into the region of diminishing information return. Now, we are not suggesting that the Schaffer collaterals operate only as an information relay, but rather, turning the argument the other way around, that whatever other functional role they may have, it is likely that (given the "bottleneck" nature of the anatomy described above) they must operate in a régime in which efficient information transfer is possible. An expansion ratio as high as 6, thus, would lead us to speculate—but it must remain pure speculation at this stage—that if the human CA1 were to operate as an efficient recoder of information to the neocortex from the hippocampus, and the expansion ratio was as high as 6, then there should be approximately three largely topographically segregrated networks in human CA1, corresponding to three broad "streams" of output.

In the Schaffer collateral model described in this study, CA1 pyramidal neurons have been represented only with respect to the input they receive from CA3 by means of the Schaffer collaterals. In actual fact, there are perforant path projections from the entorhinal cortex directly to CA1 (Steward, 1976; Wyss, 1981; Witter et al., 1988; Desmond et al., 1994), which will surely play a role in the information processing performed by CA1 neurons (Soltesz and Jones, 1995; Levy et al., 1995; Buzsaki et al., 1995; Hasselmo and Schnell, 1994). The perforant path to CA1 seems to be in nature excitatory, but normally "covered" by dominant feedforward inhibitory connections (Empson and Heinemann, 1995a,b). However, the question of what its functional role is still seems to be completely open, despite advances in the understanding of its physiology. The model of the Schaffer collaterals we have

presented here is an attempt to describe only the CA3 input to CA1 and, thus, makes no statement about the effect of perforant path inputs to CA1. It is possible to describe an extended model, along the same lines as the one presented here, which takes as its input the pattern of activity in the entorhinal cortex, that uses a simple functional description of the processing done by the dentate gyrus and CA3 and computes the CA1 activity attributable to both perforant path and Schaffer collateral inputs. Such a model is presented in (Fulvi-Mari et al., 1998). However, a model of this complexity is laborious to solve numerically, and further efforts in this direction will be contingent on future computational resources.

The assumption was made in the analytical calculation that there were no correlations between neurons in CA3. This assumption was of course made to perform the analytical calculation, which have been significantly more complicated otherwise. A question which can be asked is, how would spatial correlations in CA3 affect the role of plasticity in the Schaffer collateral connections. It is reasonable to think that there may be some nontrivial effect, because synaptic weight statistics, at least under criteria of optimality, reflect correlations in the input patterns (Nadal and Parga, 1993). Therefore, this is an interesting question for future study, possibly by further computer simulation.

There are a number of respects in which the analytical calculation itself could be extended without increasing too much the difficulty of the calculation.[2] One possibility would be to consider different forms of the plasticity term given by Eq. 9. The form used in the model we have described was a simple equation motivated by the presence of associative long-term potentiation in the Schaffer collaterals (Schwartzkroin and Wester, 1975; Kelso et al., 1986; Bliss and Collingridge, 1993; Nicoll and Malenka, 1995). The calculation essentially could be carried through in an identical manner for different forms of plasticity specified by a generalized form of this equation, and the effect on information transmission of details of the physiology of long-term potentiation could be examined.

The same techniques used here, i.e., the replica trick together with the "self-consistent statistics" approach to writing down probability densities of the states of the system, could also be used to calculate the mutual information that could be transmitted by or retrieved from a recurrent network with diluted connectivity such as CA3. It might also be possible to analyze in this way cascaded stages of feedforward and/or recurrent networks (Coolen and Viana, 1996) to gain an understanding of some general information-theoretic features of cortical connectivity.

Of course, there are many details of the physiology of this region that have either not been taken into account at all or that have been crudely assumed to be set so as to provide the most simplistic function possible. This model was never intended to be anything more than a first step in the direction of a precisely quantitative account of network function. We believe that our rather liberal application of Occam's Razor is justified by the need for a first-order model with which other, more detailed studies can be compared. This should hopefully increase substantially the insight gained from such more detailed studies, in which the simple picture can often be obscured by the multiplicity of parameters.

Advances in neuroanatomy are yielding extremely detailed information about the fine structure of the hippocampal formation (see, e.g., Freund and Buzsáki, 1996). A number of authors (Bernard and Wheal, 1994; Patton and McNaughton, 1995) have presented detailed constructions of hippocampal connectivity, and proposed that they be used as the basis for computer simulations of great complexity. Although the possible utility of such a tool cannot be denied, we urge caution in its application: simulation of physiological details will in itself contribute little or nothing to understanding of the working of the system. However, we hope that by complementing such detailed analysis with an approach in which simplifying principles are used whenever possible, advances will be made toward understanding the organization of the hippocampus.

Clearly, it is essential to further constrain the model by fitting the parameters as sufficient neurophysiologic data become available. As more parameters assume biologically measured values, the sensible ranges of values that as-yet unmeasured parameters can take will become clearer. It will then be possible to address further issues such as the quantitative importance of the constraint of dendritic length (i.e., the number of synapses per neuron) on information processing. It may also, in conjunction with studies of the effects of neuronal degeneration on single-cell physiology, offer the prospect of providing a quantitative account of information loss in neurodegenerative diseases such as Alzheimer's.

One thing this analysis does is to provide us with information about which are the important parameters of the system with regard to information transmission and which are not. This of course has impact on which quantities are of the most use to obtain experimentally. On the anatomic side, for instance, it could be said that there is little point in measuring the distribution of convergence onto CA1 neurons; in stark contrast, divergence (expansion) to CA1 from CA3 neurons is important with regard to information transmission. On the physiologic side, it seems that a very important parameter is the sparsity (a mean invariant measure of the spread) of the activity distributions (a fact pointed to by converging lines of evidence; see Treves and Rolls [1994] and Schultz and Treves [1998]), rather than the specific details of the firing distributions themselves, which do not matter beyond the second moment. Another issue is whether or not there are correlations between cells in region CA3? This finding is crucial for the issue of whether topography affects information transmission at all: if there are not, it affects it not at all; if there are, then it becomes important to determine their spatial scale, which should match up with the spatial spread of topographic connectivity in the respective axis.

In summary, we have used techniques for the analysis of neural networks to quantitatively investigate the effect of a number of biological issues on information transmission by the Schaffer collaterals. We envisage that these techniques, developed further and applied in a wider context to networks in the medial temporal

[2] It is important to avoid, if at all possible, adding an extra level of integration in the solution outlined in Eq. 19.

lobe, will yield considerable insight into the organization of the mammalian hippocampal formation.

Acknowledgments

We thank S. Panzeri for many useful discussions that were both the motivation for and invaluable assistance throughout the course of this work. S.S. has been in receipt for the duration of the work described here of a studentship from the Oxford McDonnell-Pew Centre for Cognitive Neuroscience.

REFERENCES

Amaral DG. 1993. Emerging principles of intrinsic hippocampal organization. Curr Opin Neurobiol 3:225–229.

Amaral DG, Ishizuka N, Claiborne B. 1990. Neurons, numbers and the hippocampal network. In: Storm-Mathisen J, Zimmer J, Ottersen OP, editors. Understanding the brain through the hippocampus, Vol. 83. Progress in brain research. Amsterdam: Elsevier Science.

Amaral DG, Witter MP. 1989. The three-dimensional organization of the hippocampal formation: a review of anatomical data. Neuroscience 31:571–591.

Amaral DG, Witter MP. 1995. Hippocampal formation. In: Paxinos G, editor. The rat nervous system. San Diego: Academic Press. 711–755.

Andersen P, Silfvenius H, Sunderberg SH, Sveen O. 1980. A comparison of distal and proximal dendritic synapses on CA1 pyramids in hippocampal slices in vitro. J Physiol (Lond) 307:273–299.

Barnes CA, McNaughton BL, Mizumori SJ, Leonard BW, Lin LH. 1990. Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. Prog Brain Res 83:287–300.

Bernard C, Wheal HV. 1994. Model of local connectivity patterns in CA3 and CA1 areas of the hippocampus. Hippocampus 4:497–529.

Bliss TVP, Collingridge GL. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31–39.

Boss BD, Turlejski K, Stanfield BB, Cowan WM. 1987. On the numbers of neurons in fields CA1 and CA3 of the hippocampus of Sprague-Dawley and Wistar rats. Brain Res 406:280–287.

Braak H, Braak E. 1991. Neuropathological staging of Alzheimer-related changes. Acta Neuropathol 82:239–259.

Buzsaki G, Penttonen M, Bragin A, Nádasdy Z, Chrobak JJ. 1995. Possible physiological role of the perforant path-CA1 projection. Hippocampus 5:141–146.

Chan-Palay V, Lang W, Haesler V, Strohler C, Yasarpil G. 1986. Distribution of altered hippocampal neurons and axons immunoreactive with antisera against neuropeptide Y in Alzheimer's type dementia. J Comp Neurol 248:376–394.

Coolen ACC, Viana L. 1996. Feed-forward chains of recurrent attractor neural networks near saturation. J Physiol A 29:7855–7866.

Cover TM, Thomas JA. 1991. Elements of information theory, New York: John Wiley.

Dam AM. 1980. Epilepsy and neuron loss in the hippocampus. Epilepsia 21:617–629.

Desmond NL, Scott CA, Jane JA Jr, Levy WB. 1994. Ultrastructural identification of entorhinal cortical synapses in CA1 stratum lacunosum-moleculare of the rat. Hippocampus 4:594–600.

Empson RM, Heinemann U. 1995a. Perforant path connections to area CA1 are predominantly inhibitory in the rat hippocampal-entorhinal cortex combined slice preparation. Hippocampus 5:104–107.

Empson RM, Heinemann U. 1995b. The perforant path projection to hippocampal area CA1 in the rat hippocampal-entorhinal cortex combined slice. J Physiol 484:707–720.

Finnerty GT, Jefferys JGR. 1993. Functional connectivity from CA3 to the ipsilateral and contralateral CA1 in the rat dorsal hippocampus. Neuroscience 56:101–108.

Freund TF, Buzsáki G. 1996. Interneurons of the hippocampus. Hippocampus 6:347–470.

Fulvi-Mari C, Panzeri S, Rolls ET, Treves A. 1998. A quantitative model of information processing in CA1. In: Baddeley R, Földiák P, Hancock P, editors. Information theory and the brain. Cambridge, UK: Cambridge University Press.

Giannakopoulos P, Hof PR, Micehl J-P, Guimon J, Bouras C. 1997. Cerebral cortex pathology in aging and Alzheimer's disease: a quantitative survey of large hospital-based geriatric and psychiatric cohorts. Brain Res Rev 25:217–245.

Golomb D, Hertz J, Panzeri S, Treves A, Richmond B. 1997. How well can we estimate the information carried in neuronal responses from limited samples? Neural Comput 9:649–655.

Groenewegen HJ, Room P, Witter MP, Lohman AHM. 1982. Cortical afferents of the nucleus accumbens in the cat studied with anterograde and retrograde transport techniques. Neuroscience 7:977–995.

Groenewegen HJ, Vermeulen-van der Zee E, te Kortschot A, Witter MP. 1987. Organization of the projections from the subiculum to the ventral striatum in the rat. Neuroscience 23:103–120.

Hasselmo ME, Schnell E. 1994. Laminar selectivity of the cholinergic supression of synaptic transmission in rat hippocampal region CA1: computational modeling and brain slice physiology. J Neurosci 14:3898–3914.

Hasselmo ME, Wyble BP. 1997. Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. Behav Brain Res 89:1–34.

Heller J, Hertz JA, Kjaer TW, Richmond BJ. 1995. Information flow and temporal coding in primate pattern vision. J Comp Neurosci 2:175–193.

Hertz JA, Krogh A, Palmer RG. 1991. Introduction to the theory of neural computation. Wokingham, UK: Addison-Wesley.

Hjorth-Simonsen A. 1973. Some intrinsic connections of the hippocampus in the rat: an experimental analysis. J Comp Neurol 147:145–162.

Ishizuka N, Weber J, Amaral DG. 1990. Organization of intrahippocampal projections originating from CA3 pyramidal cells in the rat. J Comp Neurol 295:580–623.

Jeffreys H, Jeffreys BS. 1972. Methods of mathematical physics, 3rd ed. Cambridge, UK: Cambridge University Press.

Kelso SR, Ganong AH, Brown TH. 1986. Hebbian synapses in hippocampus. Proc Natl Acad Sci USA 83:5326–5330.

Larkman AU, Jack JJB, Stratford KJ. 1997. Assessment of the reliability of amplitude histograms from excitatory synapses in rat hippocampal CA1 in vitro. J Physiol (Lond) 505:443–456.

Lauberg S. 1979. Commissural and intrinsic connections of the rat hippocampus. J Comp Neurol 184:685–708.

Levy WB, Colbert CM, Desmond NL. 1995. Another network model bites the dust: entorhinal inputs anre no more than weakly excitatory in the hippocampal CA1 region. Hippocampus 5:137–140.

Li X-G, Somogyi P, Ylinen A, Buzsáki G. 1994. The hippocampal CA3 network: an in vivo intracellular labelling study. J Comp Neurol 339:181–208.

Liao D, Jones A, Malinow R. 1992. Direct measurement of quantal changes underlying long-term potentiation in CA1 hippocampus. Neuron 9:1089–1097.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419–457.

Mezard M, Parisi G, Virasoro M. 1987. Spin glass theory and beyond. Singapore: World Scientific. Singapore.

Miller GA. 1955. On the bias of information estimates. In: Quasler H, editor. Information theory in psychology: problems and methods. Glencoe, Ill: The Free Press. p 95–100.

Murre JMJ. 1996. Tracelink: a model of amnesia and consolidation of memory. Hippocampus 6:675–684.

Nadal J-P, Parga N. 1993. Information processing by a perceptron in an unsupervised learning task. Network 4:295–312.
Nicoll RA, Malenka RC. 1995. Contrasting properties of two forms of long-term potentiation in the hippocampus. Nature 377:115–118.
O'Keefe J. 1993. Hippocampus, theta, and spatial memory. Curr Biol 3:917–924.
Optican LM, Richmond BJ. 1987. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex: III. Information theoretic analysis. J Neurophysiol 57:162–178.
Panzeri S. 1996. Quantitative methods for analyzing information processing in the mammalian cortex. PhD thesis, International School for Advanced Studies, via Beirut 2-4, 34013 Trieste, Italy. fttp://axp0.cns.ox.ac.uk/pub/users/panzeri/thesis.ps.gz.
Panzeri S, Treves A. 1996. Analytical estimates of limited sampling biases in different information measures. Network 7:87–107.
Patton PE, McNaughton B. 1995. Connection matrix of the hippocampal formation: I. The dentate gyrus. Hippocampus 5:245–286.
Redman SJ. 1990. Quantal analysis of synaptic potentials in neurons of the central nervous system. Physiol Rev 70:165–198.
Rolls ET. 1989. Functions of neuronal networks in the hippocampus and neocortex in memory. In Byrne JH, Berry WO, editors. Neural models of plasticity: experimental and theoretical approaches. San Diego: Academic Press. p 240–265.
Rolls ET. 1995. A model of the operation of the hippocampus and entorhinal cortex in memory. Int J Neural Syst 6(Suppl):51–70.
Rolls ET, Treves A. 1998. Neural networks and brain function. Oxford, UK: Oxford University Press.
Rolls ET, Treves A, Robertson RG, Georges-François P, Panzeri S. 1998. Information about spatial view in an ensemble of primate hippocampal cells. J Neurophysiol 79:1797–1813.
Rolls ET, Treves A, Tovee MJ, Panzeri S. 1997. Information in the neuronal representations of individual stimuli in the primate temporal visual cortex. J Comput Neurosci 4:309–333.
Samsonovich A, McNaughton BL. 1997. Path integration and cognitive mapping in a continuous attractor neural network model. J Neurosci 17:5900–5920.
Schultz S, Treves A. 1998. Stability of the replica-symmetric solution for the information conveyed by a neural network. Phys Rev E 57:3302–3310.
Schwartzkroin P, Wester K. 1975. Long-lasting facilitation of a synaptic potential following tetanization in the in vitro hippocampal slice. Brain Res 89:107–119.
Seress L. 1988. Interspecies comparison of the hippocampal formation shows increased emphasis on the Regio superior in the Ammon's Horn of the human brain. J Hirnforsch 3:335–340.
Shannon CE. 1948. A mathematical theory of communication. AT&T Bell Labs Tech J 27:379–423.
Simić G, Kostović I, Winblad B, Bogdanović N. 1997. Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease. J Comp Neurol 379:482–494.
Smith ML, Milner B. 1981. The role of the right hippocampus in the recall of spatial location. Neuropsychologia 19:781–793.
Softky WR. 1995. Fine analog coding minimizes information transmission. Neural Networks 9:15–24.
Soltesz I, Jones RSG 1995. The direct perforant path input to CA1: excitatory or inhibitory? Hippocampus 5:101–103.
Sommer W. 1880. Erkrankung des Ammonshorns als aetiologisches Moment der Epilepsie. Arch Psychiatr Nervenkr 10:631–675.
Stephan H. 1983. Evolutionary trends in limbic structures. Neurosci Biobehav Rev 7:367–374.
Stevens CF, Wang Y. 1994. Changes in reliability of synaptic function as a mechanism for plasticity. Nature 371:704–707.
Steward O. 1976. Topographic organization of the projections from the entorhinal area to the hippocampal formation of the rat. J Comp Neurol 167:347–370.
Stricker C, Field AC, Redman SJ. 1996. Statistical analysis of amplitude fluctuations in EPSCs evoked in rat CA1 pyramidal neurons in vitro. J Physiol (Lond) 490:419–441.
Swanson LW, Cowan WM. 1977. An autoradiographic study of the organization of the efferent connections of the hippocampal formation in the rat. J Comp Neurol 172:49–84.
Swanson LW, Wyss JM, Cowan WM. 1978. An autoradiographic study of the organization of intrahippocampal association pathways in the rat. J Comp Neurol 181:681–716.
Tamamaki N, Abe K, Nojyo Y. 1987. Columnar organization in the subiculum formed by axon branches originating from single CA1 pyramidal neurons in the rat hippocampus. Brain Res 412:156–160.
Tamamaki N, Nojyo Y. 1990. Disposition of slab-like modules formed by axon branches originating from single CA1 pyramidal neurons in the rat hippocampus. J Comp Neurol 291:509–519.
Tamamaki N, Nojyo Y. 1995. Preservation of topography in the connections between the subiculum, field CA1, and the entorhinal cortex in rats. J Comp Neurol 353:379–390.
Treves A. 1990. Graded-response neurons and information encodings in autoassociative memories. Physiol Rev A 42:2418–2430.
Treves A. 1995. Quantitative estimate of the information relayed by the Schaffer collaterals. J Comput Neurosci 2:259–272.
Treves A, Barnes CA, Rolls ET. 1996. Quantitative analysis of network models and of hippocampal data. In: Ono T, McNaughton BL, Molotchnikoff S, Rolls ET, Nishijo H, editors. Perception, memory and emotion: frontier in neuroscience. Amsterdam: Elsevier. p 567–579.
Treves A, Rolls ET. 1991. What determines the capacity of autoassociative memories in the brain. Network 2:371–397.
Treves A, Rolls ET. 1994. A computational analysis of the role of the hippocampus in learning and memory. Hippocampus 4:373–391.
Treves A, Skaggs WE, Barnes CA. 1996. How much of the hippocampus can be explained by functional constraints? Hippocampus 6:666–674.
Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In Ingle DJ, Goodale MA, Wansfield RJ, editors. Analysis of visual behavior. Cambridge, MA: MIT Press. p 549–586.
Van Groen T, Van Haren FJ, Witter MP, Groenewegen HJ. 1986. The organization of the reciprocal connections between the subiculum and the entorhinal cortex in the cat: I. A neuroanatomical tracing study. J Comp Neurol 250:485–497.
Victor M, Agamanolis D. 1990. Amnesia due to lesions confined to the hippocampus: a clinico-pathologic study. J Cog Neurosci 2:246–257.
West MJ. 1990. Stereological studies of the hippocampus: a comparison of the hippocampal subdivisions of diverse species including hedgehogs, laboratory rodents, wild mice and men. In: Storm-Mathisen J, Zimmer J, Ottersen OP, editors. Progress in Brain Research, Vol. 83. Amsterdam: Elsevier Science. p 13–36.
West MJ, Gundersen HJG. 1990. Unbiased stereological estimation of the number of neurons in the human hippocampus. J Comp Neurol 296:1–22.
West MJ, Schwerdtfeger WK. 1985. An allometric study of hippocampal components. A comparative study of the brains of the European hedgehod (*Erinaceus europaeus*), the tree shrew (*Tupaia glis*), and the marmoset (*Callithrix jacchus*). Brain Behav Evol 27:93–105.
West WJ, Coleman PD, Flood DG, Troncoso JC. 1994. Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. Lancet 344:769–772.
Witter MP, Griffioen AW, Jorritsma-Byham B, Krijnen JLM. 1988. Entorhinal projections to the hippocampal CA1 region in the rat: an underestimated pathway. Neurosci Lett 85:193–198.
Witter MP, Ostendorf RH, Groenewegen HJ. 1990. Heterogeneity in the dorsal subiculum of the rat. Distinct neuronal zones project to different cortical and subcortical targets. Eur J Neurosci 2:718–725.
Wyss JM. 1981. An autoradiographic study of the efferent connections of the entorhinal cortex in the rat. J Comp Neurol 199:495–512.
Zola-Morgan S, Squire LR, Amaral DG. 1986. Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. J Neurosci 6:2950–2967.

APPENDIX A

The replica symmetric solution for the mutual information is much the same as that derived by Treves (1995), except that it has been generalized to allow for topographic connectivity. In the limit of a single topographic subdivision, the equations reduce to those of Treves (1995).

$$\begin{aligned}\langle I\rangle = \mathrm{extr}_{y_A^{(\lambda)},\,\tilde{y}_A^{(\lambda)}} \Bigg\{ & \sum_{j=1}^{M}\prod_{\lambda=1}^{K}\Gamma(y_A^{(\lambda)}, w^0, z^0, C_j = C_{\lambda\nu,\gamma}) \\ & - \frac{N}{2}\sum_{\lambda=1}^{K} y_A^{(\lambda)}\,\tilde{y}_A^{(\lambda)} \\ & + N\int D\tilde{s}_1 \left\langle \prod_{\lambda=1}^{K} F(\tilde{s}_1, 0, \eta, \tilde{y}_A^{(\lambda)}, 0, 0) \right. \\ & \left. \times \ln \prod_{\lambda=1}^{K} F(\tilde{s}_1, 0, \eta, \tilde{y}_A^{(\lambda)}, 0, 0) \right\rangle_\eta \Bigg\} \\ & - \mathrm{extr}_{y_B^{(\lambda)},\,\tilde{y}_B^{(\lambda)},\,w_B^{(\lambda)},\,\tilde{w}_B^{(\lambda)},\,z_B^{(\lambda)},\tilde{z}_B(\lambda)} \\ & \cdot \Bigg\{ \sum_{j=1}^{M}\prod_{\lambda=1}^{K}\Gamma(y_B^{(\lambda)}, w_B^{(\lambda)}, z_B^{(\lambda)}, C_j = C_{\lambda\nu}, \gamma\rangle \\ & - \frac{N}{2}\sum_{\lambda=1}^{K}(y_B^{(\lambda)}\,\tilde{y}_B^{(\lambda)} + 2w_B^{(\lambda)}\,\tilde{w}_B^{(\lambda)} + z_B^{(\lambda)}\,\tilde{z}_B^{(\lambda)}) \\ & + N\int D\tilde{s}_1 D\tilde{s}_2 \prod_{\lambda=1}^{K}\langle F(\tilde{s}_1, \tilde{s}_2, \eta, \tilde{y}_B^{(\lambda)}, \tilde{w}_B^{(\lambda)}, \tilde{z}_B^{(\lambda)})\rangle_\eta \\ & \times \ln \prod_{\lambda=1}^{K}\langle F(\tilde{s}_1, \tilde{s}_2, \eta, \tilde{y}_B^{(\lambda)}, \tilde{w}_B^{(\lambda)}, \tilde{z}_B^{(\lambda)})\rangle_\eta \Bigg\} \end{aligned} \tag{19}$$

where taking the extremum means evaluating each of the two terms, separately, at a saddle-point over the variables indicated. The notation is as follows. N is the number of CA3 cells, whereas the sum over j is over M CA1 cells. λ is the elemental topographic subregion, numbered from 1 to K. F is given by

$$\begin{aligned} F(\tilde{s}_1, \tilde{s}_2, \eta, \tilde{y}, \tilde{w}, \tilde{z}) = & \left\{ \phi\left[\frac{\eta + \sigma_\delta^2(\tilde{s}_+ - \tilde{w}\eta)}{\sigma_\delta\sqrt{1 + \sigma_\delta^2\tilde{y}}}\right]\frac{1}{\sqrt{1 + \sigma_\delta^2\tilde{y}}} \right. \\ & \times \exp\frac{[\eta + \sigma_\delta^2(\tilde{s}_+ - \tilde{w}\eta)]^2}{2\sigma_\delta^2(1 + \sigma_\delta^2\tilde{y})} \\ & \left. + \phi\left[\frac{-\eta}{\sigma_\delta}\right]\exp\frac{\eta^2}{2\sigma_\delta^2}\right\} \\ & \times \exp\left[\eta\tilde{s}_- - \frac{\eta^2}{2\sigma_\delta^2}(1 + \sigma_\delta^2\tilde{z})\right] \end{aligned} \tag{20}$$

and has to be averaged over P_η and over the Gaussian variables of zero mean and unit variance $\tilde{s}_1$, $\tilde{s}_2$.

$$Ds \equiv \frac{ds}{\sqrt{2\pi}}\exp -s^2/2 \qquad \phi(x) \equiv \int_{-\infty}^{x} Ds. \tag{21}$$

$\tilde{y}$, $\tilde{w}$, and $\tilde{z}$ are saddle-point parameters. $\tilde{s}_+$ and $\tilde{s}_-$ are linear combinations of $\tilde{s}_1$, $\tilde{s}_2$:

$$\begin{aligned} \tilde{s}_\pm = & \sum_{k=1}^{2}(\mp 1)^{(k-1)} \\ & \times \sqrt{\frac{[\sqrt{(\tilde{y} - \tilde{z})^2 + 4\tilde{w}^2} \mp (-1)^k(\tilde{y} - \tilde{z})]\,(\tilde{y}\tilde{z} - \tilde{w}^2)}{[\tilde{y} + \tilde{z} + (-1)^k\sqrt{(\tilde{y} - \tilde{z})^2 + 4\tilde{w}^2}]\;\sqrt{(\tilde{y} - \tilde{z})^2 + 4\tilde{w}^2}}}\,\tilde{s}_k \end{aligned} \tag{22}$$

in the last two lines of Eq. 19, but in the second line of Eq. 19, one has $\tilde{s}_+ = \tilde{s}_1\sqrt{\tilde{y}_A}$, $\tilde{s}_- = 0$.

Γ is effectively an entropy term for the CA1 activity distribution, given by

$$\begin{aligned} \Gamma(y^{(\lambda)}, w^{(\lambda)}, z^{(\lambda)}, C_{\lambda\nu,\gamma}) = & \int \frac{ds_1 ds_2}{2\pi\sqrt{\det T'_j}}\exp -(s_1\;s_2) \\ & \cdot \frac{(T'_j)^{-1}}{2}\begin{pmatrix} s1 \\ s2 \end{pmatrix} \times \left[\int_{-\infty}^{0} dU G(U)\right. \\ & \cdot \ln\int_{-\infty}^{0} dU' G(U') \\ & \left. + \int_0^\infty dU G(U)\ln G(U)\right] \end{aligned} \tag{23}$$

where

$$\begin{aligned} G(U) = G(U; s_1, s_2, y^{(\lambda)}, w^{(\lambda)}, z^{(\lambda)}, C_{\lambda\nu,\gamma}) \\ = \phi\left[\frac{(\xi_0 - s_2)(T_{yj} + 2g_jT_{wj} + g_j^2T_{zj}) + (U - U_0 + s_1\,g_js_2)\,(T_{wj} + g_jT_{zj})}{\sqrt{(T_{yj}T_{z_j} - T_{wj}^2)(T_{yj} + 2g_jT_{wj} + g_j^2T_{zj})}}\right] \\ \times \frac{1}{\sqrt{2\pi(T_{yj} + 2g_jT_{wj} + g_j^2T_{zj})}} \\ \cdot \exp -\frac{(U - U_0 + s_1 + g_js_2)^2}{2(T_{yj} + 2g_jT_{wj} + g_j^2T_{zj})} \\ + \phi\left[\frac{-(\xi_0 - s_2)\,T_{yj} - (U - U_0 + s_1 + g_j\xi_0)\,T_{wj}}{\sqrt{(T_{yj}T_{z_j} - T_{wj}^2)\,T_{yj}}}\right] \\ \times \frac{1}{\sqrt{2\pi T_{yj}}}\exp -\frac{(U - U_0 + s_1 + g_j\xi_0)^2}{2T_{yj}} \end{aligned} \tag{24}$$

and

$$\begin{aligned} T_{yj} &= \sigma_{\epsilon^R}^2 + \sigma_J^2\sum_{\nu=1}^{K} C_{\lambda\nu}(y^0 - y^{(\nu)}) \\ T_{wj} &= \sigma_J^2\sum_{\nu=1}^{K} C_{\lambda\nu}(w^0 - w^{(\nu)})\cos(\theta) \\ T_{zj} &= \sigma_{\epsilon^S}^2 + \sigma_J^2\sum_{\nu=1}^{K} C_{\lambda\nu}(z^0 - z^{(\nu)}) \\ T'_j &= \sigma_J^2\sum_{\nu=1}^{K} C_{\lambda\nu}\begin{pmatrix} y^{(\nu)} & w^{(\nu)}\cos(\theta) \\ w^{(\nu)}\cos(\theta) & z^{(\nu)} \end{pmatrix} \end{aligned} \tag{25}$$

are effective noise terms.

$$g_j = h\,\frac{\sum_{\nu=1}^{K} C_{\lambda\nu}}{C}\,x^0\langle\eta\rangle_\eta\sqrt{C\gamma(\theta)} \tag{26}$$

y, w, z are saddle-point parameters (conjugated to $\tilde{y}$, $\tilde{w}$ and $\tilde{z}$), and x^0, y^0, w^0, z^0 are corresponding single-replica parameters fixed as

$$x^0 = \frac{1}{N}\sum_i \left\langle \frac{(\eta_i - \langle\eta\rangle_\eta)}{\langle\eta\rangle_\eta} V_i \right\rangle = \left\langle \frac{(\eta - \langle\eta\rangle_\eta)}{\langle\eta\rangle_\eta} \times \left[\eta\phi\left(\frac{\eta}{\sigma_\delta}\right) + \frac{\sigma_\delta}{\sqrt{2\pi}}\exp - \frac{1}{2}\left(\frac{\eta}{\sigma_\delta}\right)^2\right]\right\rangle_\eta$$

$$y^0 = \frac{1}{N}\sum_i \langle V_i^2\rangle = \left\langle [\sigma_\delta^2 + \eta^2]\,\phi\left(\frac{\eta}{\sigma_\delta}\right) + \frac{\eta\sigma_\delta}{\sqrt{2\pi}}\exp - \frac{1}{2}\left(\frac{\eta}{\sigma_\delta}\right)^2\right\rangle_\eta$$

$$w^0 = \frac{1}{N}\sum_i \langle \eta_i V_i\rangle = \left\langle \eta\left[\eta\phi\left(\frac{\eta}{\sigma_\delta}\right) + \frac{\sigma_\delta}{\sqrt{2\pi}}\exp - \frac{1}{2}\left(\frac{\eta}{\sigma_\delta}\right)^2\right]\right\rangle_\eta$$

$$z^0 = \frac{1}{N}\sum_i \eta_i^2 = \langle\eta^2\rangle_\eta. \qquad (27)$$

APPENDIX B

Parameters used were, except where otherwise indicated in the text:

σ_δ	0.30
σ_ϵ^S	0.20
σ_ϵ^R	0.20
C	10000
σ_J^2	0.0001
ζ_0	−0.4
U_0	−0.4
M / N	2.0