Pathological Effects of Cortical Architecture on Working Memory in Schizophrenia

C. D. Gore1,2, M. Bánya1,3, P. J. Gray1, V. Diwadkar4, P. Érdi1,3

Abstract

Neural connectivity of the prefrontal cortex is essential to working memory. Reduction of prefrontal connectivity and abnormal prefrontal dopamine modulation are common characteristics associated with schizophrenia.

Two experiments separately modeled the effects of exaggerated pruning and of synaptic depression to imitate schizophrenic performance in a prefrontal neural network. In the first model, effects of cortical pruning were simulated with a set of scale-free networks of neurons and compared with empirical results from the Sternberg working memory task. The second set of simulations were based on the synaptic theory of working memory. Simulations of this model measured memory duration in relation to synaptic facilitation and depression constants and in relation to the level of neural connectivity.

In the first set of simulations, modulating levels of cortical pruning resulted in a gain or loss in accuracy and speed of memory recollection. In the second set of simulations, increased facilitation time constants and decreased inhibitory time constants resulting in longer memory durations, and overly connected networks resulted in very low memory durations.

In the first model, the decline in memory performance can be attributed to the emergence of pathological memory behavior brought about by the warping of the basins of attraction. Collectively, the simulations demonstrate that a reduction of prefrontal cortical hubs can lead to schizophrenia like performance in neural networks, and may account for pathological working memory in the disorder.

Introduction

Working memory (WM) is a central cognitive function that involves the maintenance and online manipulation of information for brief periods of time [3, 13]. Numerous studies in animals (involving single unit recordings) and in humans (involving functional magnetic resonance imaging) suggest that the prefrontal cortex is essential to working memory and the maintenance of information [7, 14]. The anatomical and molecular bases of working memory in the prefrontal cortex have also been explored. WM has been linked to prefrontal dopaminergic function; the tuning functions of prefrontal neurons during WM tasks are impaired when dopamine receptors (D1) are hyper- or hypo-stimulated [30] indicating that an optimal range of dopamine function underlies intact working memory. These results are particularly relevant to the study of working memory function in schizophrenia.

Schizophrenia is characterized by altered dopaminergic function [5] that may mediate impairments in tasks of prefrontal function such as working memory [8, 19]. Working memory impairments in schizophrenia may also be related to altered structural integrity of the prefrontal cortex [21] and to altered network organization of the PFC and other regions of the brain. This is supported by in vivo imaging studies suggesting that the intact PFC exhibits hubs of neurons with significantly increased connectivity indices with other heteromodal cortical regions [16]. These general patterns of cortico-cortical connectivity from the prefrontal cortex to other regions may underlie the distributed underpinnings of working memory [6]. Further findings indicating disruption in fronto-cortical circuitry with reduced prefrontal hubs in schizophrenia [4, 19] may in part explain schizophrenia symptoms and disruptions in WM in the illness.
Dynamic systems approaches have begun to prove useful as approaches toward the formal characterization of schizophrenia. Pathological attractors may implement the dynamic generation of positive symptoms in the illness as these symptoms, including delusions and hallucinations can be activated in the absence of external cues. Related to the modifiability of the attractor-basin portrait, a model based on the NMDA receptor delayed maturation has also been suggested as a possible mechanism of the pathogenesis of schizophrenic psychotic symptoms [25].

Dynamical systems hypotheses are based on the assumption that pathological symptoms are related to changes in the geometry of the attractor basin portrait [20,24]. Network models of excitatory and inhibitory neurons built by leaky integrate-and-fire models have been used to design several simulation experiments to study the effects of changes in synaptic conductances on overall network performance. Reduction in synaptic conductances connected to glutamatergic NMDA receptors imply flatter attractor basins, and consequently less stable memory storage. Combined reduction of NMDA and GABA receptors imply such changes in the attractor structure, that may represent positive symptoms, as hallucinations and delusion.

Functional magnetic resonance imaging (fMRI) studies have demonstrated the PFC’s role in WM tasks. During clinical WM tasks, a stimulus is briefly presented to a subject often to be remembered during a delay period that takes place between the presentation of a stimulus and the execution of a task. An example of such a task is the Sternberg delay [29], where a subject is shown a string of capital, consonant letters (e.g., BGZXF) followed by a delay, after which time they are prompted with a letter and a position (i.e., Z=3) and asked whether or not that particular letter was in the given position in the string. Sternberg found that in doing this test, the response time for subjects is a linear function of the number of characters in the set. Neuroimaging studies using the Sternberg test [2] have shown that activation in the dorsolateral PFC increases with set size, as does response time and a decrement in accuracy. The results of the study are reproduced in Fig. 1.

Since cortical networks might have scale-free network features [28], we compared the performance of different scale-free networks in the prefrontal cortex in solving the Sternberg task. The question was whether a network architecture with smaller numbers of hubs showed reduced performance. The simulation results below indicate that such is indeed the case.

The synaptic theory of working memory [22] suggests that the duration and stability of working affected by calcium-dependent mechanisms. In our simulations, the duration of memory, that is, the time interval when a pattern can be recalled was defined as the performance measure. A simple recall task was simulated to assign a performance function to the two-dimensional parameter space of the rate constants. The effect of systematic cortical pruning was also studied by varying the overall connection probability from zero to one. The duration of memory was non-linearly affected by increases in connectivity, initially falling rapidly, then less precipitously. Below, possible bases of this behavior are explored.

**Comparative Studies on the Memory Performance of Different Scale-Free Networks**

**Generation and test of networks**

A traditional Hopfield neural network connects neurons in an all-to-all fashion [17], though this is known to be biologically implausible, as biological neural networks develop complex architectures that share common properties with scale-free (SF) networks [28]. SF networks are evolving networks, so their architectures are built over many iterations. The vertices in the resulting networks have degrees $k$ that are distributed according to a power law $P(k) \sim k^{-\gamma}$ where $P(k)$ is a function describing a node’s likelihood of receiving a new connection with each iteration. This results in a hierarchical architecture similar to real cortical architectures, with most of the neurons represented as sparse nodes, having very few connections, and a small portion of the nodes acting as hubs, with many connections. This is a more desirable template for network architecture for investigating the role of hubs in maintaining stable memory states.

**Implementation of the Sternberg task for scale-free neural networks**

The model employs the Barabási-Albert (BA) method [1] for building the network structure. A BA network has $N$ nodes and $n(N - n_0)$ edges. Here, $n_0$ is the initial number of nodes that seed the network with a fully connected graph, and $n$ (where $n \leq n_0$) is the number of connections added with each iteration of the algorithm. With the addition of each new node, previous nodes

![Fig. 1](sternberg_test.png) **Fig. 1** Sternberg Test with 6 s delay. Adapted from [2]. Error bars illustrate the standard deviation of the results. (a) Mean accuracy for task loads of 3, 5, and 8 letters. (b) Mean reaction times for task loads of 3, 5, and 8 letters.
in the network have a probability of receiving a new connection described by the linear preferential attachment algorithm:

\[ p_i = \frac{k_i}{\sum k_j} \]  

(1)

Sets of networks were generated by choosing values for \( n_0 \) and \( n \) in order to influence their respective degree distributions. A higher degree, in this case, represents a more dense neural hub. Profiles of sample degree distributions for four sets of \( n_0 \) and \( n \) are shown in Fig. 2.

Patterns are created for the Sternberg delay by generating sets of 3, 5, and 8 random consonants. Consonants are represented by their numerical equivalents 1 through 21 in binary form, using five binary digits per letter. Five digits are used because 21, the highest of the consonant values, is represented with five digits as 10101. Patterns are formed by scaling the sets of binary digits up to 600 by repeating digits as illustrated in Fig. 3. The number 600 is used because it is the least common multiple of the size of the consonants (5) and the set sizes 3, 5, and 8. This is simply to ensure that all networks have the same number of neurons.

The binary units are normalized to a range of 0 to 1 each, and a weight matrix is generated following the Hebbian learning rule

\[ \omega_{ij} = \sum_{k=1}^{n} k_i k_j \]  

(2)

assigning a synaptic strength \( w_{ij} \) to each network connection. Each weight matrix is mapped across a randomly generated BA graph, removing connections without corresponding edges. This weight matrix is undirected (admittedly not a realistic assumption), so each connection is bidirectional

\[ \omega_{ji} = \omega_{ij} \]  

(3)

Similar to the Sternberg delay, once a pattern is loaded to memory, a cue is offered with a letter and position. This is done by beginning the neurons with an incomplete form of the original pattern and running until convergence or divergence. For example, in Fig. 3, the cue for the second character \( M \) would be 00000 01101 00000 00000 00000.

Schizophrenia memory dynamics

It is generally accepted that schizophrenia is related to excessive pruning of cortical connections. Simple network studies have shown that cortical pruning may lead to the formation of “pathological attractors”, and unlearned pathological states may appear as a result of pruning connections from a network and from an imbalance of dopamine modulation.

Delusions, or hallucinations, are common positive symptoms of schizophrenia. Here, delusional working memory is related to stable states in the neural network model. The energy function of the network

\[ E = -\frac{1}{2} \sum_{i,j} \omega_{ij} s_i s_j + \sum_i \vartheta_i s_i \]  

(4)

describes the memory landscape for a given state, where \( \omega \) is the weight, \( s \) is the state, and \( \vartheta \) is the unit threshold. Global minima of this function are stable states, learned from a pattern where the network is likely to converge. If there is no convergence, a network often oscillates between states. Delusions occur when the basin of attraction is too shallow as in Fig. 4. Here, the focus of attraction can easily slip and incorrectly stabilize in a local minimum state. This model focuses on the effects on the network’s performance and the shape of the energy function by varying the average degree \( k \) through various parameters to the BA method.

Results

Simulations of the Sternberg delay were run for sets of 3, 5, and 8 characters. One hundred iterations of the test were run for each set of parameters and set size and the results averaged. The behavior of the model’s results, shown in Fig. 5, depicts the relationship between nodal degree and memory performance. Accuracy is determined by the number of networks out of 100 that were able to successfully recall the stored pattern from one letter. Reaction time is an average of how many cycles a network requires to converge to a solution or diverge. It is shown that...
networks with stronger hubs remember patterns more accurately and more quickly.
For each of the four pairs of \( n \) and \( n_0 \), ten networks were generated where the energy levels for every possible state were measured according to Eq. 4. The lowest energy level, which is at the state of the stored memory, is recorded along with the next lowest energy level. ● Table 1 shows a comparison of values for \( n \), \( n_0 \), the resulting degree levels \( k \), and averages for the recorded energy levels. It is clear that the networks with stronger hubs had deeper attractor basins, while the networks with weaker hubs had basins that were more shallow. Correlating this with the data from ● Fig. 5, it is seen that the networks with more shallow basins are less accurate than the networks with deeper basins.

**Synaptic Theory of Working Memory: Some Further Studies**

\( \text{The model framework} \)

The synaptic theory of working memory was suggested by [22]. A simple model for the prefrontal cortex was specified, exploiting the general belief that in this brain region the excitatory synapses are facilitatory. Working memory is therefore generated and maintained by short-term synaptic facilitation.

A reduced short term plasticity model uses two variables, \( x \) is the available resource (released transmitter molecules) and \( u \) is the utilization variable (residual calcium level). The increase of \( u \) is called the facilitation, the decrease of \( x \) is the depression, and the product \( u \times x \) characterizes synaptic change. The process is controlled by two time constants; \( \tau_f \) and \( \tau_d \) denoting facilitatory and depressive time constants, respectively. The model is defined by Equations 5 and 6.

\[
\frac{dx}{dt} = \frac{1-x}{\tau_d} - u x (t-t_w)
\]

(5)

\[
\frac{du}{dt} = \frac{U - u}{\tau_f} + U (1-u) x (t-t_w)
\]

(6)

In [22] the time constants were fixed as \( \tau_D = 0.2s \) and \( \tau_F = 1.5s \) to express facilitation. The motivation for our experiment came from the ending of [22]: “…The model provides a possible target for a pharmacological interference with WM. In particular, manipulations that modify the facilitation/depression balance in the memory-related cortical areas … are predicted to have a strong effect on the stability and duration of memory”.

**Table 1** Energy levels for the different networks. It is clear that networks with a higher average degree have deeper energy profiles.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( n_0 )</th>
<th>( k )</th>
<th>Pattern energy</th>
<th>Next lowest energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>12</td>
<td>11.86</td>
<td>-7106.5</td>
<td>-6149.3</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>9.90</td>
<td>-5932.4</td>
<td>-5136.8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>7.93</td>
<td>-4753.0</td>
<td>-4126.2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>5.96</td>
<td>-3570.1</td>
<td>-3089.5</td>
</tr>
</tbody>
</table>

**Fig. 4** Comparison of a healthy attractor basin with one that is too shallow. In the second diagram, the focus of attraction slips from the memorized pattern to an incorrect local minimum. (a) Healthy attractor basin. (b) Shallow attractor basin.

**Fig. 5** Sternberg delay simulation. Results are shown for four levels of connectivity. Compare to Figure 1. (a) Accuracy comparison. (b) Reaction time comparison.
Definition of duration of working memory
Working memory was defined as the time between the end of the write-in signal (the population-specific increase in the background input that loads an item to the memory) and the last point in time when the object can be retrieved from the memory. Retrieval in the readout signal (a nonspecific increase in the background input), would produce a population spike (PS). This PS codes for the object loaded in the memory previously and refreshes the memory as well. The probability of observing a PS is mostly dependent on the actual level of synaptic efficacy. We can define a threshold value in efficacy that divides the two behaviors of the network (PS or not), so we can define the duration as the time between the endpoint of the write-in signal and the time-point when the efficacy falls under the threshold value.

Results
Fig. 6 shows explorations of the two-dimensional parameter space. These intuitive results indicate that if facilitation relaxes slower and depression relaxes faster, memory duration will increase. One could define a regime to be considered normal, and so there would be two regimes: one for too short and one for too long memory fading time. Second, the connectivity was changed by setting the overall connection probability from 0 to 0.8. Somewhat counter-intuitively, the duration was reduced by increasing the connectivity, as one can see in Fig. 7.

However, if we look at the model setting, the cause for this behavior is obvious. We apply an external input on all the cells, which is modeled by a Gaussian noise with a large mean and small deviation. The mean is actually above the firing threshold of the cells, so if there were no other dynamics, they would fire permanently with a frequency defined by the refractory period. The principal effect of the cells on each other is the inhibition, allowing them to follow different firing patterns.

Discussion
We have previously presented computational studies of the reduced cognitive abilities of schizophrenia patients in fronto-hippocampal tasks [9,11,12]. In this paper, we provide further examples regarding details of the structure-function relationship behind schizophrenia.

A loss of neural hubs in the PFC was observed in schizophrenia patients [4], and the Sternberg delay [29] proved to be a good working memory exercise, utilizing the PFC [2]. Modulating network generation parameters in order to vary the average nodal degree produced the predicted pathological attractor basins. Along with warping the energy function, these pathological attractors proved to have a strong impact on the resulting networks’ accuracy and time to recall a pattern.

Our second experiment shows pathological memory behavior when the network is interconnected. These experiments, along with previous in vitro studies, imply that, whether connected too strongly or too weakly, abnormal connectivity has a measurable effect on memory performance.

How might the time constants be regulated?
Calcium binding proteins (e.g., neuronal calcium sensor NCS-1) modify short-term plasticity (at least hippocampal cell cultures) by switching pair-pulsed depression to facilitation [27]. As facilitation in our models appears to signify normal performance but not schizophrenic performance, we might expect that NCS-1 concentration would be large in the normal brain, but low in the schizophrenia brain. This prediction is inconsistent with empirical evidence: NCS-1 is up-regulated in the PFC of schizophrenia patients [18], suggesting that facilitatory synapses should be normal and not schizophrenic. However, the molecular machinery might be much more complicated for the following reasons:
(i) NCS-1 might be double localized pre- and postsynaptically [23], (ii) NCS-1 is a part of a network of proteins.

As we are far removed from being able to give a realistic detailed mechanism for changing the balance between facilitation and depression, we studied the dynamic properties of the system in the two-dimensional parameter space of the time constants. To evaluate the performance of the memory system we had to define the duration of the memory. The hypothesis was that the shift in balance between facilitation and depression might modify the duration of the memory. Phenomenologically two types of pathology, “too short” and “too long” could emerge. The question is whether the duration of memory depends on the two time constants and reduced connectivity, respectively.
Acknowledgements

PE thanks the Henry Luce Foundation for general support. CDG thanks Robert Trenary for answering questions related to neural networks and Madeline McAuley for general help. VD acknowledges the MH68680 grant from NIMH.

Disclosure

The authors of this paper did not receive financial support by pharmaceutical companies for this work.

Affiliations
1 Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, Michigan, USA
2 Department of Computer Science, Western Michigan University, Kalamazoo, Michigan, USA
3 Department of Biophysics, KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, Budapest, Hungary
4 Department of Psychiatry and Behavioral Neuroscience, Wayne State Univ. Medical School, Detroit, Michigan, USA

References
7 Cohen JD, Ferman SD, Braver TS et al. Activation of prefrontal cortex in a non-spatial working memory task with functional MRI. Human brain mapping 1994; 1: 293–304
10 Durstewitz D, Seamans JK. The computational role of dopamine D1 receptors in working memory. Neural Netw 2002; 15 (4-6): 561–572
19 Lewis DA, Gonzalez-Burgos G. Neuroplasticity of neocortical circuits in schizophrenia. Neuropsychopharmacology 2008; 33: 141–165
23 Negyessy I, Goldman-Rakic PS. Subcellular localization of the dopamine D2 receptor and coexistence with the calcium-binding protein neuronal calcium sensor-1 in the primate prefrontal cortex. J Comp Neurol 2005; 488: 464–475
25 Ruppin E. NMDA receptor delayed maturation and schizophrenia. Med Hypotheses 2000; 54: 693–697
29 Sternberg S. High-Speed Scanning in Human Memory. Science 1966; 153 (3736): 652–654