*Physical Review A*, vol. 44 (6) (1991) 4017-4060
ingber@alumni.caltech.edu

**Statistical mechanics of neocortical interactions:
A scaling paradigm applied to electroencephalography**

Lester Ingber
*Science Transfer Corporation, P.O. Box 857, McLean, VA 22101*
(Received 10 April 1991)

A series of papers has developed a statistical mechanics of neocortical interactions (SMNI), deriving aggregate behavior of experimentally observed columns of neurons from statistical electrical-chemical properties of synaptic interactions. While not useful to yield insights at the single neuron level, SMNI has demonstrated its capability in describing large-scale properties of short-term memory and electroencephalographic (EEG) systematics. The necessity of including nonlinear and stochastic structures in this development has been stressed. In this paper, a more stringent test is placed on SMNI: The algebraic and numerical algorithms previously developed in this and similar systems are brought to bear to fit large sets of EEG and evoked potential data being collected to investigate genetic predispositions to alcoholism and to extract brain "signatures" of short-term memory. Using the numerical algorithm of Very Fast Simulated Re-Annealing, it is demonstrated that SMNI can indeed fit this data within experimentally observed ranges of its underlying neuronal-synaptic parameters, and use the quantitative modeling results to examine physical neocortical mechanisms to discriminate between high-risk and low-risk populations genetically predisposed to alcoholism. Since this first study is a control to span relatively long time epochs, similar to earlier attempts to establish such correlations, this discrimination is inconclusive because of other neuronal activity which can mask such effects. However, the SMNI model is shown to be consistent with EEG data during selective attention tasks and with neocortical mechanisms describing short-term memory previously published using this approach. This paper explicitly identifies similar nonlinear stochastic mechanisms of interaction at the microscopic-neuronal, mesoscopic-columnar and macroscopic-regional scales of neocortical interactions. These results give strong quantitative support for an accurate intuitive picture, portraying neocortical interactions as having common algebraic or physics mechanisms that scale across quite disparate spatial scales and functional or behavioral phenomena, i.e., describing interactions among neurons, columns of neurons, and regional masses of neurons.

PACS Nos.: 87.10.+e, 05.40.+j, 02.50.+s, 02.70.+d

# I. INTRODUCTION

## A. Methodology

In many complex systems, as spatial-temporal scales of observation are increased, new phenomena arise by virtue of synergistic interactions among smaller-scale entities—perhaps more properly labeled "quasientities"—which serve to explain much observed data in a parsimonious, usually mathematically aesthetic, fashion [1,2]. For example, in classical thermodynamics of equilibrium systems, it is possible to leap from microscopic molecular scales to macroscopic scales, to use the macroscopic concept of temperature to describe the average kinetic energy of microscopic molecular activity, or to use the macroscopic concept of pressure to describe the average rate of change of momentum per unit area of microscopic molecules bombarding the wall of a cavity.

However, many complex systems are in nonequilibrium, being driven by nonlinear and stochastic interactions of many external and internal degrees of freedom. For these systems, classical thermodynamics typically does not apply [3]. For example, the description of weather and ocean patterns, which attempt to include important features such as turbulence, rely on semiphenomenological mesoscopic models, those in agreement with molecular theories but not capable of being rigorously derived from them. Phase transitions in magnetic systems, and many systems similarly modeled [4-6], require careful treatment of a continuum of scales near critical points. In general, rather than having a general theory of nonequilibrium nonlinear process, there are several overlapping approaches, typically geared to classes of systems, usually expanding on nonlinear treatments of stochastic systems [1,2,7-9]. Many biological systems give rise to phenomena at overlapping spatial-temporal scales. For example, the coiling of DNA is reasonably approached by blending microscopic molecular-dynamics calculations with mesoscopic diffusion equations to study angular winding [10]. These approaches have been directed to study electroencephalography (EEG) [11], as well as other biological systems [12].

Therefore, it should not be surprising that the complex human brain supports many phenomena arising at different spatial-temporal scales. What is perhaps surprising is that it seems possible to study truly macroscopic neocortical phenomena such as EEG by appealing to a chain of arguments dealing with overlapping microscopic and mesoscopic scales. A series of papers has developed this statistical mechanics of neocortical interactions (SMNI) [13-25]. This approach permits us to find models of EEG whose variables and parameters are reasonably identified with ensembles of synaptic and neuronal interactions. This approach has only recently been made possible by developments in mathematical physics since the late 1970s, in the field of nonlinear nonequilibrium statistical mechanics. The origins of this theory are in quantum and gravitational field theory.

Other physical systems have varying degrees of theoretical support leading to Fokker-Planck partial differential equation descriptions. Here, new problems arise in nonlinear nonequilibrium systems, often requiring modeling with the introduction of a nonconstant coefficient of the second-derivative "diffusion" term. The spatial second-derivative term usually represents the kinetic energy, when the first derivative represents the momentum in the differential-equation description of a system. It was early noticed that a similar treatment of the gravitational equation [26] required a modification of the potential term of the corresponding Lagrangian. We now better understand the mathematical and physical similarities between classical stochastic processes described by Fokker-Planck equations and quantum processes described by the Schrödinger equation [27]. The Lagrangian, essentially equal to the kinetic energy minus the potential energy, to first order in an expansion about the most likely state of a quantum or stochastic system, gives a global formulation and generalization of the well-known relation, force equals mass times acceleration [28]. In the neocortex, the velocity corresponds to the rate of firing of a column of neurons, and a potential is derived which includes nearest-neighbor interactions between columns. The Lagrangian formulation also accounts for the influence of fluctuations about such most likely paths of the evolution of a system, by use of a variational principle associated with its development. The Lagrangian is therefore often more useful than the Hamiltonian, essentially equal to the kinetic energy plus the potential energy, related to the energy in many systems. As will be demonstrated, this is especially useful to obtain information about the system without solving the time-dependent Fokker-Planck equation; however, we also will describe neocortical phenomena requiring the full solution.

In its differential form, the momentum is proportional to the derivative operator. For classical systems, the coefficient of the square of the momentum is twice the diffusion, e.g., the second moment of a probability distribution describing some systems. The introduction of a nonconstant coefficient of even the first-derivative term requires careful treatment. Such a problem arises for a charged particle in an electromagnetic field [29], which originally was treated by physical arguments to enforce ‘‘minimal coupling,’’ whereby the momentum $p$ is simply replaced by $p - Ae/c$, where $A$ is the electromagnetic potential, $e$ is the electric charge, and $c$ is the speed of light. Minimal coupling excludes other $A^2$ terms from appearing in the equation of evolution, e.g., Schrödinger's equation. Such problems are related to the operator-ordering of the derivative operators with respect to their nonconstant coefficients. For classical systems, the analogous expression to $Ae/c$ is the drift, e.g., the first moment of a probability distribution describing some systems. In the neocortex, we derive nonlinear expressions for both the drift and diffusion. The detailed mathematical relationships to the physical content of these issues was only clarified in the late 1970s and early 1980s, and is relevant to the mathematics of the neocortex. The first real breakthrough was achieved by noting how these issues should be treated in the context of classical nonlinear nonequilibrium statistical mechanics [30,31].

While application has been made of these new mathematical physics techniques in quantum and classical statistical mechanics [32-34], we are not merely bringing over techniques to neuroscience from other disciplines out of curiousity. Indeed, the contention appears to be well supported that a mathematical investigation of the neocortex reasonably demands these mathematical techniques, to such an extent that it can be argued that, if the neocortex had been studied and sufficient data collected prior to mathematical developments in quantum or gravitational theory, then these mathematical techniques might have been developed in neuroscience first. The brain is sufficiently complex that it requires the same tools used for similar very complex physical systems. In many ways, we may consider the brain as the prototypical information processing system, and the mathematical techniques used here may be rigorously viewed as filters to describe the processing of this information.

This statistical-mechanics approach, at many junctions in its development, may be intuitively compared to the approach used in simple magnetic systems, a comparison made early in neuronal modeling [35]. While caution must be exercised to respect the integrity of the neocortical system, such comparisons can be useful pedagogically. The mathematical approach presented here also has been useful to describe phenomena in social systems, ranging from military command, control and communications [36-38], to political systems [39], to pricing bonds on financial markets [40-42]. In this context, it has been noted that the activity of neurons may resemble the activity of a throng of people, in which interactions take place at multiple hierarchical levels [43]. The numerical algorithms used in this paper were developed in part in the process of investigating these other systems.

### B. Top-down versus bottom-up

There are at least two ways to present this admittedly complex technical approach. First, we will take the top-down approach, essentially examining some macroscopic issues in EEG measurement. This will motivate us to then look to a bottom-up approach, starting with microscopic synaptic activity and neuronal interactions, then scaling up through mesocolumnar activity of columns of neurons, to finally achieve a reasonable macroscopic description of EEG activity. The confluence of these approaches is expected to yield a tractable approach to EEG analyses [17,24,44].

Section II presents the current dipole string model as a top-down approach to give a rationale for incorporating several scales of neuronal interactions in order to understand macroscopic EEG phenomena. We demonstrate how this dipole string model is the Euler-Lagrange variational limit of the derived bottom-up electric-potential EEG Lagrangian.

Section III presents a self-contained outline of the papers describing SMNI, highlighting the essential steps in aggregating from microscopic scales of synaptic-neuronal interactions, to mesoscopic scales of minicolumns and macrocolumns of neurons, to macroscopic scales of regional activity over centimeters of neocortex. Appendix A gives a brief derivation of the three equivalent mathematical representations used here, i.e., coupled Langevin stochastic rate equations, the Fokker-Planck partial-differential equation, and the path-integral equation defined by the Lagrangian.

Section IV presents a self-contained outline of previous verification of SMNI theory, applied to short-term memory (STM) at the mesoscopic scale, EEG systematics at the macroscopic scale, and statistical constraints on synaptic modification coding STM into long-term memories. Some discussion is given on the relevance of chaos in this system.

Section V presents an approach to directly fit SMNI theory to specific sets of EEG data, utilizing algorithms and physical insights gained in Secs. II-IV. It is demonstrated that SMNI can indeed fit these data within experimentally observed ranges of its underlying neuronal-synaptic parameters, and use the quantitative modeling results to examine physical neocortical mechanisms to discriminate high-risk and low-risk populations genetically predisposed to alcoholism. The SMNI model of STM is shown to be consistent with EEG data during selective attention tasks.

Section VI gives our conclusions, based on the success of the calculations presented in the previous sections. We emphasize the success in finding SMNI to be a reasonable model of neocortical interactions that can fit STM and EEG using only physically based parameters. We also emphasize the scaling of common similar nonlinear stochastic interactions at multiple levels, i.e., at microscopic-neuronal, mesoscopic-columnar, and macroscopic-regional scales.

## II. DIPOLE STRING MODEL

### A. Background

The human neocortex is a complex physical and biological system that processes information at multiple spatial and temporal scales. Connections between cortical neurons are of two major types: The short-range intracortical fibers (both excitatory and inhibitory) of average length less than 1 mm, and the long-range cortico-cortical fibers (exclusively excitatory) which form most of the white matter in humans and have an average length of several centimeters [45,46].

There are several noninvasive experimental or clinical methods of recording brain activity, e.g., EEG, magnetoencephalography (MEG), magnetic resonance imaging (MRI), positron-emission tomography (PET), single-photon-emission-computed tomography (SPECT). While MRI, PET, and SPECT offer better three-dimensional presentations of brain activity, EEG and MEG offer superior temporal resolutions on the order of neuronal relaxation times, i.e., milliseconds. Recently, it also has been shown that EEG and MEG offer comparable spatial resolutions on the order of several millimeters; a square millimeter is the approximate resolution of a macrocolumn representing the activity of approximately $10^5$ neurons.

Limiting cases of linear macroscopic theories of intracortical interaction predict local wave phenomena and obtain dispersion relations with typical wave numbers $k$= 10 to 100 cm$^{-1}$ and dominant frequencies in the general range of human spontaneous EEG (1−20 Hz) [17,47]. However, human scalp potentials are spatially filtered by both distance and tissue between cortical current sources and surface electrodes so that scalp EEG power is attenuated to about 1% of its cortical value at $k$= 1 or 2 cm$^{-1}$. This implies that spontaneous cortical activity is only measurable on the scalp if at least several hundred contiguous cortical macrocolumns, each macrocolumn being comprised of about $10^5$ neurons, produce coherent source activity. Many such columns of correlated sources cause substantial cortical power at low $k$, measurable on the scalp. The use of an average reference removes spectral power near $k = 0$, and skull and scalp tissue effects strongly attenuate power for $k > 0.5$ cm$^{-1}$. At $k$= 0.5 cm$^{-1}$, scalp power is typically below the noise level of the EEG machine. In one study, it was suggested that the $\alpha$ mode (9.5 Hz) is consistent with standing waves, whereas the $\mu$ mode (8.0 Hz) is more consistent with posterior-to-anterior traveling waves across the electrode array [24,46].

Vertically oriented neuronal columns generally receive input in the top two of six horizontal laminae and output their neuronal signals from the lower laminae. This justifies a dipole model of the neocortex, albeit care must be taken in this interpretation to account for spreading of electric potentials due to volume effects and skull and scalp mismatched conductivities [46,48-50]. Since coherent vertically oriented dipoles greatly enhance their contribution to the scalp electric potential, only a small fraction of columns approaching this idealized structure are required to justify this approximation [51].

It must be appreciated that this effective dipole model is just that, and is appropriate only at spatial scales of about a millimeter to a centimeter. Using a scalp EEG or MEG to infer sources at resolutions less than a millimeter is likely an inappropriate use of this tool as it underestimates the complex neurophysiology giving rise to electromagnetic recordings. Using scalp EEG or MEG on the order of tens of centimeters to infer a single source of regional activity also often is inappropriate, as in many situations there are likely to be many sources of activity giving rise to such recordings. The resolution of several millimeters to several centimeters appears to be quite appropriate to model statistically aggregated activity as an effective dipole.

These experimental considerations of scalp EEG have led to the development of a linear model of neocortical dynamics in which global delays are dominant and local delays are either neglected or treated as a perturbation [46]. Standing brain waves are predicted with resonant modes determined partly by cortical surface geometry. Both temporal and spatial properties predicted by this theoretical approach are in semiquantitative agreement with EEG states associated with minimal cognitive processing, i.e., awake resting with eyes closed ($\alpha$ rhythm), asleep, or under the influence of an anesthetic agent [48,49,52,53]. Recently, this approach has been extended to encompass theories based on linear interactions in which local delays (rise and decay times of postsynaptic potentials) are included [44].

There is ample experimental evidence of the existence of linear EEG phenomena in lower mammals [54,55] and humans [46,56] which occurs over limited ranges of experimental conditions (modulation depth of sinusoidal driving of the brain, for example). However, the use of linear and quasi-linear theories and the neglect of interactions across spatial scales are evidently crude approximations.

Complementing these studies, SMNI has shown that limiting cases of nonlinear treatments of the neocortex across microscopic-mesoscopic-macroscopic scales give similar numerical confirmation of EEG observables, thereby addressing both of the above approximations, i.e., nonlinearities and multiple scales of interaction. The SMNI approach has recently outlined a means to consistently include the long range cortico-cortical interactions that appear critical to the observed characteristics of scalp potentials [23].

### B. Modeling of observables

The mechanical string model of the neocortex assumes linear properties of connected local nonlinear oscillators [24]. Local cortical dynamics in dipole layers is considered analogous to the nonlinear mechanical oscillators that influence global modes. Macroscopic scalp potentials are analogous to the lower modes of string displacement. The linear global model of EEG waves is much more involved than the simple linear string. For example, multiple long-range fiber systems cause multiple branches of the dispersion relation, and distributed propagation velocities cause selective damping of macroscopic modes [44,46,53]. The application of boundary conditions to the two-dimensional cortical surface has predicted other EEG properties [44,46]. Many of these features can be included in our nonlinear generalization of the simple string described here.

The equation for the string displacement $\Phi$ is

$$\frac{\partial^2 \Phi}{\partial t^2} - c^2 \frac{\partial^2 \Phi}{\partial x^2} + [\omega_0^2 + f(\Phi)]\Phi = 0 \ , \tag{1}$$

for a linear array (length $l$) of sensors (electrodes) of size $s$. We can only observe wave numbers in the approximate range

$$\frac{\pi}{l} \le k \le \frac{\pi}{s} \ . \tag{2}$$

If the center-to-center spacing of sensors, considered to form a closed loop, is also $s$, $l = Ms$, where $M$= (number of sensors) − 1), $k$= $2n\pi/R$ for $n = \{ 1, 2, 3, \cdots \}$, and sensors span half the string, $l = R/2$, then

$$1 \le n \le M \tag{3}$$

for some maximum $M$, which is on the order of 3 to 7 in EEG studies using multielectrode recordings in two-dimensional arrays on the cortical surface. A typical circumference of the neocortex following a coordinate in and out of fissures and sulci is $R$ = 100 cm (about 50 cm along the scalp surface). If EEG power is mostly restricted to $k < 0.5$ cm$^{-1}$, only modes $n < 4$ are observed, independent of the number of electrodes.

Thus, the string displacement $\Phi$ (potential within the cortex) is given by

$$\Phi(x, t) = \sum_{n=1}^{\infty} G_n(t) \sin k_n x \ , \tag{4}$$

where we have assumed zero boundary conditions at the ends of the string for simplicity of presentation, but our observed $\Phi^\dagger$ is given by

$$\Phi^\dagger(x, t) = \sum_{n=1}^{M} G_n(t) \sin k_n x \ . \tag{5}$$

In the linear case, where $f(\Phi) = 0$ (equal linear oscillators to simulate local circuit effects in cortical columns), we have

$$\frac{\partial^2 \Phi}{\partial t^2} - c^2 \frac{\partial^2 \Phi}{\partial x^2} + \omega_0^2 \Phi = 0 \ ,$$

$$\Phi = \sum_{n=1}^{\infty} A_n \cos \omega_n t \sin k_n x \ ,$$

$$\omega_n^2 = \omega_0^2 + c^2 k_n^2 \ , \tag{6}$$

giving a dispersion relation $\omega_n(k_n)$. For the nonlinear case, $f(\Phi) \neq 0$, the restoring force of each spring is amplitude dependent [57].

Returning to the cortical medium, what can we say about

$$\Phi^\dagger(x, t) = \sum_{n=1}^{M} G_n(t) \sin k_n x \ , \tag{7}$$

the macroscopic observable displacement potential on the scalp or cortical surface? On the basis of previous studies of EEG dispersion relations [17,46], it would seem that we should be able to describe $\Phi^\dagger$ as a linear or quasilinear variable, but influenced by the local nonlinear behavior that crosses the hierarchical level from mesoscopic to macroscopic (columnar dipoles). How do we mathematically articulate this intuition, for the purposes of consistent description as well as lay the foundation for future detailed numerical calculations? We suggest answers to these questions in the ‘‘bottom-up’’ approach part of this paper.

We examine these issues by taking reasonable synaptic parameters, developing the statistical mechanics of neocortical interactions, and then determining whether they are consistent with observed EEG data. In fact, here we report fits of multichannel human scalp EEG data to these algebraic forms. A current project is investigating the response of the cortical system to given initial driving forces that match or mismatch firing patterns of columnar firings possessed by a given set of synaptic parameters, and under conditions of plastically changing synaptic parameters reflecting changes of these patterns. This should help in clinical diagnoses using the EEG tool.

### C. Outline of derivation of the nonlinear string model

We use a mechanical-analog model, the string model, derived explicitly for neocortical interactions using SMNI [24]. This defines a probability distribution of firing activity, which can be used to further investigate the existence of other nonlinear phenomena, e.g., bifurcations or chaotic behavior, in brain states.

Previous studies have detailed that the predominant physics of short-term memory and of (short-fiber contribution to) EEG phenomena takes place in a narrow ‘‘parabolic trough’’ in $M^G$ space, roughly along a diagonal line [16]. Here, $G$ represents $E$ or $I$, $M^E$ represents contributions to columnar firing from excitatory neurons, and $M^I$ represents contributions to columnar firing from inhibitory neurons. The object of interest within a short refractory time, $\tau$, approximately 5 to 10 msec, is the Lagrangian $\underline{L}$

for a mesocolumn, detailed further below. $\tau \underline{L}$ can vary by as much as a factor of $10^5$ from the highest peak to the lowest valley in $M^G$ space. Therefore, it is reasonable to assume that a single independent firing variable might offer a crude description of this physics. Furthermore, the scalp potential $\Phi$ can be considered to be a function of this firing variable. (Here, "potential" refers to the electric potential, not the potential term in the Lagrangian derived below.) In an abbreviated notation subscripting the time-dependence,

$$\Phi_t - \ll \Phi \gg = \Phi(M_t^E, M_t^I) \approx a(M_t^E - \ll M^E \gg) + b(M_t^I - \ll M^I \gg) , \tag{8}$$

where $a$ and $b$ are constants, and $\ll \Phi \gg$ and $\ll M^G \gg$ represent typical minima in the trough. In the context of fitting data to the dynamic variables, there are three effective constants, $\{ a, b, \phi \}$ ,

$$\Phi_t - \phi = a M_t^E + b M_t^I . \tag{9}$$

We scale and aggregate the mesoscopic probability distributions, $P$, over this columnar firing space to obtain the macroscopic probability distribution over the scalp-potential space:

$$P_\Phi[\Phi] = \int dM^E dM^I P[M^E, M^I] \delta[\Phi - \Phi'(M^E, M^I)] . \tag{10}$$

The parabolic trough described above justifies a form

$$P_\Phi = (2\pi\sigma^2)^{-1/2} \exp(-\frac{\Delta t}{2\sigma^2} \int dx\, L_\Phi) ,$$

$$L_\Phi = \frac{\alpha}{2} |\partial\Phi/\partial t|^2 + \frac{\beta}{2} |\partial\Phi/\partial x|^2 + \frac{\gamma}{2} |\Phi|^2 + F(\Phi) , \tag{11}$$

where $F(\Phi)$ contains nonlinearities away from the trough, $\sigma^2$ is on the order of $1/N$ given the derivation of $\underline{L}$ above, and the integral over $x$ is taken over the spatial region of interest. In general, there also will be terms linear in $\partial\Phi/\partial t$ and in $\partial\Phi/\partial x$.

Previous calculations of EEG phenomena, described below [17], show that the short-fiber contribution to the $\alpha$ frequency and the movement of attention across the visual field are consistent with the assumption that the EEG physics is derived from an average over the fluctuations of the system, e.g., represented by $\sigma$ in the above equation. I.e., this is described by the Euler-Lagrange equations derived from the variational principle possessed by $L_\Phi$ (essentially the counterpart to force equals mass times acceleration), more properly by the "midpoint-discretized" Feynman $L_\Phi$, with its Riemannian terms, as discussed below [14,15,23]. Hence, we can use the variational principle,

$$0 = \frac{\partial}{\partial t} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial t)} + \frac{\partial}{\partial x} \frac{\partial L_\Phi}{\partial(\partial\Phi/\partial x)} - \frac{\partial L_\Phi}{\partial\Phi} . \tag{12}$$

The result is

$$\alpha \frac{\partial^2\Phi}{\partial t^2} + \beta \frac{\partial^2\Phi}{\partial x^2} + \gamma\Phi - \frac{\partial F}{\partial\Phi} = 0 . \tag{13}$$

If there exist regions in neocortical parameter space such that we can identify $\beta/\alpha = -c^2$, $\gamma/\alpha = \omega_0^2$ (e.g., as explicitly calculated below),

$$\frac{1}{\alpha} \frac{\partial F}{\partial\Phi} = -\Phi f(\Phi) , \tag{14}$$

and we take $x$ to be one-dimensional, then we recover our nonlinear string, Eq. (1) above. Terms linear in $\partial\Phi/\partial t$ and in $\partial\Phi/\partial x$ in $L_\Phi$ in Eq. (11) can make other contributions, e.g., giving rise to damping terms.

The path-integral formulation has a utility beyond its deterministic Euler-Lagrange limit. We have utilized this to explicitly examine the long-time evolution of systems, to compare models to long-time correlations in data [36,37]. This use is being extended to other systems, in finance [41,42] and in EEG modeling as described here.

### D. Macroscopic coarse graining and renormalization

We now are in a position to address the issue posed originally of how to mathematically justify the intuitive coarse-graining of $\Phi$ to get $\Phi^\dagger$. In $L_\Phi$ above, consider terms of the form

$$\int \Phi^2 dx = \int dx \sum_n^\infty \sum_m^\infty G_n G_m \sin k_n x \sin k_m x$$

$$= \sum_n \sum_m G_n G_m \int dx \sin k_n x \sin k_m x$$

$$= (2\pi/R) \sum_n G_n^2 \ . \tag{15}$$

By similarly considering all terms in $L_\Phi$, we effectively define a short-time probability distribution for the change in node $n$, defined by

$$p_n[G_n(t + \Delta t) | G_m(t)] \ , \tag{16}$$

where we note that in general the $F(\Phi)$ term in $L_\Phi$ will require coupling between $G_n(t + \Delta t)$ and $G_m(t)$, $n \neq m$, likely including more than one $m$. Therefore, we can define

$$P_\Phi = p_1 p_2 \cdots p_\infty \ . \tag{17}$$

We now physically and mathematically can define a coarse-graining,

$$P_{\Phi^\dagger} = \int dk_{M+1} dk_{M+2} \cdots dk_\infty \ p_1 p_2 \cdots p_M p_{M+1} p_{M+2} \cdots p_\infty \ . \tag{18}$$

I.e., since we have a *bona fide* probability distribution $P_\Phi$, we can integrate over those fine-grained variables, which are not observed. This procedure is one contribution to algorithms used in ‘‘renormalization-group’’ theory [6], to account for multiple intermediate scales of interactions. While other criteria for use of this theory certainly are not present here, it is useful to recognize that this is a reasonable phenomenological approach to integrating over many scales of neocortical interactions.

The integration over the fine-grained wave numbers tends to smooth out the influence of the $k_n$’s for $n > M$, effectively ‘‘renormalizing’’

$$G_n \to G^\dagger{}_n \ ,$$

$$\Phi \to \Phi^\dagger \ ,$$

$$L_\Phi \to L^\dagger{}_{\Phi^\dagger} \ . \tag{19}$$

Eventually, laminar circuitry should be included in both the local and global models. Previous papers have detailed how this can be realized, but more numerical study is needed to determine the degree to which this can be accomplished. As reported here, the solutions are being tested by their goodness of fit to existing EEG data using methods of very fast simulated reannealing [58].

## III. SMNI MODEL

### A. Rationale

We begin our ‘‘bottom-up’’ approach by taking the viewpoint that, since there has been much progress made in mathematically describing the physics at finer spatial-temporal scales, we should use these descriptions to derive a development of the coarser EEG macroscopic scale described above. SMNI has reasonably detailed a consistent physics which at least parallels, with striking numerical specificity, short-term memory (STM) and EEG phenomena at the macrocolumnar scale of hundreds of thousands of neurons, in terms of aggregated physics transpiring at the single-neuronal level. The details of this SMNI development of STM and EEG will be used to support the ‘‘top-down’’ development described above.

A major contribution of this work is the analytic treatment of minicolumns [59]. Minicolumns are observed to span approximately $7 \times 10^2\ \mu\mathrm{m}^2$. Mesocolumnar domains are defined here as physiological (functional) units, with input converging from an anatomical (structural) macrocolumn extent of approximately 1000 minicolumns, and output diverging from an anatomical minicolumn out to a macrocolumnar extent. Calculations support observations of periodically alternating firings of columnar structures [60-64]. As pictured in Fig. 1, this microscopic scale is orders of magnitude larger than the molecular scale of membrane biophysics. Also note that ‘‘macrocolumns’’ spanning roughly $7 \times 10^5\ \mu\mathrm{m}^2$ have been defined as another physiological entity observed in the neocortex [59], but the macroscopic regions considered here are orders of magnitude larger than these. Mesocolumnar domains are sufficiently close to the scale of microscopic neurons to allow direct dependence of this theory on neuronal chemical and electrical properties. The proper stochastic treatment of their interaction permits their development into macroscopic regions responsible for global neocortical information processing. ‘‘Thermodynamic’’ entities corresponding to the ‘‘free-energy’’ potential, ‘‘temperature,’’ and order parameters of these macroscopic regions are derived by a statistical-mechanics paradigm [65].

Figure 1.

Relative to other biological entities, the intrinsic synaptic activity of the most highly evolved mammalian human neocortex functions via the most degenerate and the shortest-ranged neuronal interactions (on the order of micrometers). Here, ‘‘degenerate’’ reflects the mesoscopic state of approximate redundancy of connectivity among microscopic neurons, for purposes of describing this coarser scale. This suggests that many collective aspects of this system may be fruitfully studied similarly to other collective systems, e.g., including magnetic systems, lasers, and more general information-theoretic systems [8,66]. Collective effects, from clustering [67,68] or from statistical interactions [69], are proposed to be mechanisms of information processing, in addition to the ‘‘hard-wiring’’ mechanisms also possessed by other more ordered cortical entities [70,71].

Reasonable criteria for any physical approach to the neocortex should include the following three basic features. These also serve to illustrate the appropriate analogies between the neocortex and other collective physical systems.

i. Interactions. Short-ranged neuronal interactions over time periods of several milliseconds should be derived from even more microscopic synaptic activities [72]. [See Fig. 1(a).] Long-ranged spatial interactions from specific neuronal pathways, primarily composed of the relatively low population of long excitatory fibers from ipsilateral association, contralateral commissural, and thalamocortical processes must be consistently treated. These long-ranged interactions are also important for collective activity in the mammalian cortex [45], and they are included in this study. Longer-time, weaker and modulatory nonsynaptic influences arising from humoral and electrotonic interactions [73-75] are included, only as their averaged properties affect synaptic parameters.

ii. Statistics. Neurons separated by large distances, across $10^3$ to $10^8$ neurons, can be statistically coupled via their short-ranged interactions. [See Fig. 1(c).] Order parameters, the underlying independent variables at the appropriate scale of the theory, must be identified, and intrinsic fluctuations from the microscopic synaptic and neuronal systems, diffusion effects, must be included. There also are fluctuations of the mesoscopic system due to their aggregated neuronal interactions, derived here as gradient couplings between neighboring mesoscopic cells. These spatially ordered mesoscopic domains respect the observed anatomy and physiology of the neocortex [67,68], complementing earlier theories hypothesizing random neural networks [76,77].

iii. Dynamics. A viable formalism must be adopted to describe the statistical time evolution of the macroscopic nonequilibrium system over scales of $10^2$ to $10^4$ msec.

Although cooperativity between distant neurons is typically quite low [78], except perhaps in homologous regions of separate hemispheres, macroscopic regions reflect cooperative behavior, proposed here to best be understood as initiated at the mesoscopic level of interaction. The existence of collective spatial-temporal activity, embedded in a spontaneous noisy background, is supported by statistical analyses of electroencephalographic and magnetoencephalographic recordings [79,80]. As long as collective mechanisms arising in a physical system characterized by the above three features are considered to be viable sources of collective neocortical phenomena, then these features must be correctly formulated.

There is a large body of literature dealing with neuronal mechanisms that intuits phenomenological differential equations from rates of change of average excitatory and inhibitory neuronal firings, and then proceeds to search for collective behavior, limit cycles, and oscillatory behavior [81-86]. Mechanisms are sought to explain varied phenomena such as hysteresis in perception [84], perception and learning [87,88], and ontogenesis of columnar formation [88,89]. Comparisons with applications of these techniques to those used in other physical systems [1], illustrates that the pioneering application of these appropriate formalisms to the neocortical system still has much to offer. Much inspiration for these applications has come from work in nonequilibrium thermodynamics, which has been applied to specific systems, e.g., chemical reactions, lasers, magnetic systems, fluids, spin glasses, etc., as well as to the general formulation of complex nonlinear systems [1,2,9,69,90].

This study also distinguishes between neuronal mechanisms the neocortex uses to process information and the structures of information the neocortex processes. A Lagrangian is derived that operates on firings of the system. When integrated over a time period, this yields the nonequilibrium equivalent of a "thermodynamic potential." This Lagrangian is derived, not conveniently defined or hypothesized, from the short-time evolution of the probability distribution of columnar firing states. The exponential of minus the Lagrangian, essentially this short-time distribution up to an important normalization factor, operates as a weighting factor on all possible states, filtering or transforming (patterns of) input firings into output firings. "Information" is a concept well defined in terms of the probability eigenfunctions of electrical-chemical activity of this Lagrangian. The path-integral formulation presents an accurate intuitive picture of an initial probability distribution of patterns of firings being filtered by the (exponential of the) Lagrangian, resulting in a final probability distribution of patterns of firing.

## B. Microscopic neurons

### *1. General description*

Figure 1(a) illustrates the microscopic neuronal interaction scale, on the order of several micrometers. Neocortical neurons typically have many dendrites that receive quanta of chemical postsynaptic stimulation from many other neurons. The distribution of quanta transmitted across synapses takes place on the scale of $10^{-2}$ $\mu$m, as illustrated in the inset of Fig. 1(a*). Each quantum has thousands of molecules of chemical neurotransmitters that affect the chemically gated postsynaptic membrane. Chemical transmissions in the neocortex are believed to be either excitatory ($E$), such as glutamic acid, or inhibitory ($I$), such as $\gamma$ aminobutyric acid. There exist many transmitters as well as other chemicals that modulate their effects, but it is assumed that after millions of synapses between hundreds of neurons are averaged over, then it is reasonable to ascribe a distribution function $\Psi$ with a mean and variance for $E$ and $I$ interneuronal interactions.

While some neuroscientists do not accept the assumption that simple algebraic summation of excitatory depolarizations and inhibitory hyperpolarizations at the base of the inner axonal membrane determines the firing depolarization response of a neuron within its absolute and relative refractory periods [72], still many other neuroscientists agree that this assumption is reasonable when describing the activity of large ensembles of neocortical neurons, each one typically having many thousands of synaptic interactions.

This same averaging procedure makes it reasonable to ascribe a distribution function $\Gamma$ with a mean and variance for $E$ and $I$ intraneuronal interactions. A Gaussian $\Gamma$ is taken to describe the distribution of electrical polarizations caused by chemical quanta impinging on the postsynaptic membrane. These polarizations give a resultant polarization at the base of the neuron, the axon [extension in Fig. 1(a) cut by the double broken line]. The base of the axon of a large fiber may be myelinated. However, smaller neurons typically lack these distinguishing features. Experimental techniques are not yet sufficiently advanced to attempt the explicit averaging procedure necessary to establish the means and variances of $\Psi$ and $\Gamma$, and their parameters, *in vivo*. Differential attenuations of polarizations from synapses to the base of an axon are here only phenomenologically accounted for by including these geometric and physiological effects into $\Gamma$.

With a sufficient depolarization of approximately 10 to 20 mV at the soma, within an absolute and relative refractory period of approximately 5 msec, an action potential is pulsed down the axon and its many collaterals, affecting voltage-gated presynaptic membranes to release quanta of neurotransmitters. Not detailed here is the biophysics of membranes, of thickness $\approx 5 \times 10^{-3}$ $\mu$m, composed of biomolecular leaflets of phospholipid molecules [91-93]. At present, $\Psi$ and $\Gamma$ are taken to approximate this biophysics for use in macroscopic studies. The formalism adopted in this study is capable of using new microscopic functional dependences, gleaned from other experimental or theoretical investigations, and cranking them through to obtain similar macroscopic descriptions. Chemical independences of excitatory depolarizations and inhibitory hyperpolarizations are well established in the neocortex, and this independence is retained throughout this study.

It should be noted that experimental studies initially used to infer $\Psi$ and $\Gamma$ (e.g., at neuromuscular junctions) were made possible by deliberately reducing the number of quanta by lowering external calcium concentrations [94,95]. $\Psi$ was found to be Poissonian, but in that system, where hundreds of quanta are transmitted *in vivo*, $\Psi$ may well be otherwise; for example, Gaussian with independent mean and variance. Current research suggests a binomial distribution, having a Poisson limit [14,96,97]. Note that some investigators have shown a Bernoulli distribution to be more accurate in some cases [14,97,98], and that the very concept of quantal transmission, albeit that good fits to experimental data are achieved with this concept, is under review. In the neocortex, probably small numbers of quanta are transmitted at synapses, but other effects, such as nonuniformity and nonstationarity of presynaptic release sites, and nonlinear summation of postsynaptic potentials, may detract from a simple phenomenological Poisson description [72]. This short description serves to point out possible differences in $\Psi$ resulting from many sources. However, the derivation of synaptic interactions given here makes it plausible that for reasonable neuronal parameters, the statistical folding of $\Psi$ and $\Gamma$ is essentially independent of the functional form assumed for $\Psi$, just requiring specification of its numerical mean and variance.

The result of this analysis is to calculate the transition probability of the firing of neuron $j$, $p_{\sigma_j}$, given its interaction with its neighbors that also may fire or not fire. The result is given as the tabulated error function. Within the range where the total influences of excitatory and inhibitory firings match and exceed the average threshold potential of a given neuron, the probability of that neuron firing receives its major contribution to increase from 0 towards 1. A step function derived as $\tanh F^G$ is defined by the "threshold factor" $F_j$. That is, for $F^G \gg 1$, $\tanh F^G \rightarrow 1$, while for $F^G \ll 1$, $\tanh F^G \rightarrow -1$, and so the threshold region of $-1 < F^G < 1$ sensitively controls this important $\tanh F^G$ contribution to the drifts, the driving terms, in the Lagrangian. The mesoscopic development discussed below retains this sensitivity.

This is similar to the mathematical result obtained by others [99-101] who have modeled the neocortex after magnetic systems [35]. However, the following is derived more generally, and has the neural parameters more specifically denoted with different statistical significances given to $\Psi$ and $\Gamma$, as described above.

### *2. Conditional probability*

Consider $10^2 < N < 10^3$ neurons, labeled by $k$, interacting with a given neuron $j$. Each neuron may contribute many synaptic interactions to many other neurons. A neuron may have as many as $10^4 - 10^5$ synaptic interactions. Within time $\tau_n \approx 5$ msec, $\Psi$ is the distribution of $q$ quanta of chemical transmitter released from neuron $k$ to neuron $j$ ($k \neq j$) with mean $a_{jk}$, where

$$a_{jk} = A_{jk}(\sigma_k + 1)/2 + B_{jk} \ . \tag{20}$$

$A_{jk}$ is the conductivity weighting transmission of polarization, dependent on $k$ firing,

$$\sigma_k = \begin{cases} 1, & k \text{ fires,} \\ -1, & k \text{ does not fire} \end{cases} \tag{21}$$

and $B_{jk}$ is a background including some nonsynaptic and long-range activity. Of course, $A$ and $B$ are highly complicated functions of $kj$. This definition of $\sigma_k$ permits a decomposition of $a_{jk}$ into two different physical contributions. At this point there is a reasonable analogy to make with magnetic systems, where $\sigma_k$ might represent a unit spin. However, the details of the interactions between neurons differ from those between magnetic spins, and this greatly affects such comparisons.

Within the scope of the assumption that postsynaptic potential responses from numbers of presynaptic released quanta add algebraically, a Gaussian process is taken to represent this response for each quantum released. Application of the central limit theorem [102] then yields, for any $q$ quanta, a Gaussian process $\Gamma$ for imparting a potential $W_{jk}$ to neuron $j$:

$$\Gamma = (2\pi q \phi_{jk}^2)^{-\frac{1}{2}} \exp[-(W_{jk} - q v_{jk})^2]/(2q\phi_{jk}^2) ,$$

$$\lim_{q \to 0} \Gamma \equiv \delta(W_{jk}) , \tag{22}$$

where the polarization $v_{jk}$ can be positive (excitatory $E$) or negative (inhibitory $I$), and the Dirac $\delta$ function,

$$\delta(Z) = (2\pi)^{-1} \int_{-\infty}^{\infty} dQ \exp(iQZ) , \tag{23}$$

represents a well-behaved, strongly peaked distribution.

The probability $S_{jk}$ of developing $W_{jk}$ from $k$ is

$$S_{jk} = \sum_{q=0}^{\infty} \Gamma \Psi . \tag{24}$$

The probability $S_j$ of developing potential $W_j$ from all $N$ neurons is

$$S_j = \int \cdots \int dW_{j1} \cdots dW_{jN} \, S_{j1} \cdots S_{jN} \delta(W_j - \sum_k W_{jk}) . \tag{25}$$

The conditional probability $p_{+j}$ of neuron $j$ firing if $W_j > V_j$, the threshold of $j$, is

$$p_{+j} = \int_{V_j}^{\infty} dW_j S_j . \tag{26}$$

At this stage it is possible to include the probability of not firing by defining

$$p_{\sigma_j} = \begin{cases} p_{+j}, \ \sigma_j = +1 \\ p_{-j}, \ \sigma_j = -1, \end{cases} \tag{27}$$

by replacing

$$V_j \to \sigma_j V_j ,$$

$$Q \to \sigma_j Q . \tag{28}$$

*Poissonian* $\Psi$. For $a_{jk}$ small, take $\Psi$ as a Poisson with mean and variance $a_{jk}$:

$$\Psi = \exp(-a_{jk})(a_{jk})^q / q! . \tag{29}$$

The above yields

$$p_{\sigma_j} = \int_{V_j}^{\infty} dW_j (2\pi)^{-1} \int_{-\infty}^{\infty} dQ \exp(iQW_j)$$

$$\times \exp \left\{ -\sum_k a_{jk} [1 - \exp(-i\sigma_j Q v_{jk} - Q^2 \phi_{jk}^2 / 2)] \right\} . \tag{30}$$

An extremum approximation is now valid: The main contribution to $p_{\sigma_j}$ comes from nonoscillatory contributions from the second exponential above, where its argument has a minimum, rendering it Gaussian. Using

$$\phi_{jk} < |v_{jk}| \ll V_j \le W_j \;, \tag{31}$$

$p_{\sigma_j}$ may be calculated as

$$p_{\sigma_j} = \pi^{-\frac{1}{2}} \int\limits_{(\sigma_j F_j \sqrt{\pi}/2)}^{\infty} dz\, \exp(-z^2)$$

$$= \frac{1}{2}\,[1 - \mathrm{erf}(\sigma_j F_j \sqrt{\pi}/2)],$$

$$F_j = \frac{V_j - \sum_k a_{jk} v_{jk}}{(\pi \sum_{k'} a_{jk'} (v_{jk'}^2 + \phi_{jk'}^2))^{\frac{1}{2}}} \;. \tag{32}$$

"erf" is the tabulated error function, simply related to the "normal probability function." [102] $F_j^G$ is a "threshold factor," as $p_{\sigma_j}$ increases from 0 to 1 between $\infty > \sigma_j F_j > -\infty$ sharply within the range of $F_j \approx 0$.

*Gaussian* $\Psi$. The mean of a Poisson distribution of $q$ successes is

$$a_{jk} = \psi \bar{e} \;, \tag{33}$$

$\bar{e}$ is the large number of repetitions of an "experiment," likely correlated with the number of synaptic knobs [14,96], and $\psi$ is the small probability of success, the average probability of release of one quantum. For large $a_{jk}$, a Gaussian $\Psi'$ representing $\Psi$ is defined with mean $a_{jk}$ and variance $a_{jk}(1-\psi)$ [102],

$$\Psi' = \frac{[2a_{jk}(1-\psi)]^{-\frac{1}{2}} \exp[-(q - a_{jk})^2/(2a_{jk}(1-\psi))]}{\int\limits_{-[a_{jk}/(2(1-\psi))]^{\frac{1}{2}}}^{\infty} dz\, \exp(-z^2)} \;. \tag{34}$$

For a large number of closely spaced values of $q$, we can replace the sum by an integral,

$$\sum_q \rightarrow \int\limits_0^{\infty} dq \;. \tag{35}$$

Similar to the above derivation using the Poissonian $\Psi$, we derive

$$v_{jk'}^2 \rightarrow (1-\psi) v_{jk'}^2 \;. \tag{36}$$

*Arbitrary* $\Psi$. Examination of this derivation shows that we get the same result from the folding of $\Gamma$ and $\Psi$, for a wide range of reasonable $\Psi$ peaked near $q = a_{jk}$. If

$$|\sigma_j F_j| < 1, \tag{37}$$

then an asymptotic expression for $p_{\sigma_j}$ is

$$p_{\sigma_j} \approx \frac{\exp(-\sigma_j F_j)}{\exp F_j \; + \; \exp(-F_j)} \;. \tag{38}$$

This form of $p_{\sigma_j}$ exposes the linear dependence of the argument on $\sigma_j$ and $F_j$. Note that $\tanh(2z/\sqrt{\pi})$ is quite a good approximation to erf ($z$). Using the notation ($z$, tanh , erf), obtain: (0., 0. , 0.), (.01, .0113 , .0113), (.1, .1124 , .1125), (.5, .5111 , .5205), (1., .8105 , .8427), (2., .9783 , .9953), (2.5, .9929 , .9996), ( $\infty$, 1. , 1.). For small $z$ obtain:

$$\tanh(2z/\sqrt{\pi}) = (2/\sqrt{\pi})(z - z^3/2.36 + z^5/4.63 - z^7/8.98 \pm \cdots)$$

$$\mathrm{erf}(z) = (2/\sqrt{\pi})(z \; - z^3/3 \; + z^5/10 \; - z^7/42 \; \pm \cdots). \tag{39}$$

### C. Mesoscopic domains

#### *1. General description*

As is found for most nonequilibrium systems, e.g., for lasers, chemical systems, fluids, and ecological systems [1,8], a mesoscopic scale is required to formulate the statistical mechanics of the microscopic system, from which the macroscopic scale can be developed [1]. The neocortex is particularly interesting in this context in that a clear scale for the mesoscopic system exists, both anatomically (structurally) and physiologically (functionally). ``Minicolumns'' of about $N \approx 110$ neurons (about 220 in the visual cortex) comprise modular units vertically oriented relative to the warped and convoluted neocortical surface throughout most, if not all, regions of the neocortex [59,61,62,64,103,104]. Clusters of about 100 neurons have been deduced to be reasonable from other considerations as well [43]. Since the short-ranged interactions between neurons take place within ~1 mm, which is the extent of a ``macrocolumn'' comprising ~ $10^3$ minicolumns of $N^* \approx 10^5$ neurons, and since macrocolumns also exhibit rather specific information-processing features, this theory has retained the divergence-convergence of macrocolumn--minicolumn, efferent-afferent interactions by considering domains of minicolumns as having similar synaptic interactions within the extent of a macrocolumn. This macrocolumnar-averaged minicolumn is designated in this theory as a ``mesocolumn.''

This being the observed situation, it is interesting that $N \approx 10^2$ is just the right order of magnitude to permit a formal analysis using methods of mathematical physics just developed for statistical systems in the late 1970s [27,30]. $N$ is small enough to permit nearest-neighbor interactions to be formulated, such that interactions between mesocolumns are small enough to be considered gradient perturbations on otherwise independent mesocolumnar firing states. This is consistent with rather continuous spatial gradient interactions observed among columns [105], and with the basic hypothesis that nonrandom differentiation of properties among broadly tuned individual neurons coexists with functional columnar averages representing superpositions of patterned information [106]. This is a definite mathematical convenience; otherwise, a macrocolumn of ~ $10^3$ minicolumns would have to be described by a system of minicolumns with up to 16th-order next-nearest neighbors. (Consider 1000 minicolumns spread out in a two-dimensional grid about 33 by 33 minicolumns, and focus attention on the center minicolumn.) Also, $N$ is large enough to permit the derived binomial distribution of afferent minicolumnar firing states to be well approximated by a Gaussian distribution, a luxury not afforded an ``average'' neuron, even in this otherwise similar physical context. Finally, mesocolumnar interactions are observed to take place via one to several relays of neuronal interactions, so that their time scales are similarly $\tau \approx 5 - 10$ msec.

Even after statistically shaping the microscopic system, the parameters of the mesoscopic system are still macrocolumnar-averaged synaptic parameters, i.e., reflecting the statistics of millions of synapses with regard to their chemical and electrical properties. Explicit laminar circuitry, and more complicated synaptic interactions, e.g., dependent on all combinations of presynaptic and postsynaptic firings, can be included without loss of detailed analysis [15].

The mathematical development of mesocolumns establishes a mesoscopic Lagrangian $\underline{L}$, which may be considered as a ``cost function'' with variables $M^G$, $\dot{M}^G$, and $\nabla M^G$, and with parameters defined by the macrocolumnar-averaged chemical-electrical entities developed below. (See Fig. 2.)

The Einstein summation convention is used for compactness, whereby any index appearing more than once among factors in any term is assumed to be summed over, unless otherwise indicated by vertical bars, e.g., $|G|$. The mesoscopic probability distribution $P$ is given by the product of microscopic probability distributions $p_{\sigma_i}$, constrained such that the aggregate mesoscopic excitatory firings $M^E = \sum_{j \in E} \sigma_j$, and the aggregate mesoscopic inhibitory firings $M^I = \sum_{j \in I} \sigma_j$.

$$P = \prod_G P^G[M^G(r; t + \tau) | M^{\bar{G}}(r'; t)]$$

$$= \sum_{\sigma_j} \delta \left( \sum_{j \in E} \sigma_j - M^E(r; t + \tau) \right) \delta \left( \sum_{j \in I} \sigma_j - M^I(r; t + \tau) \right) \prod_j^N p_{\sigma_j}$$

$$\approx \prod_G (2\pi\tau g^{GG})^{-1/2} \exp(-N\tau \underline{L}^G) \ , \tag{40}$$

where the final form is derived using the fact that $N > 100$. $\bar{G}$ represents contributions from both $E$ and $I$ sources. This defines the Lagrangian, in terms of its first-moment drifts $g^G$, its second-moment diffusion matrix $g^{GG'}$, and its potential $\underline{V}'$, all of which depend sensitively on threshold factors $F^G$,

$$P \approx (2\pi\tau)^{-1/2} g^{1/2} \exp(-N\tau \underline{L}) \ ,$$

$$\underline{L} = (2N)^{-1}(\dot{M}^G - g^G) g_{GG'} (\dot{M}^{G'} - g^{G'}) + M^G J_G/(2N\tau) - \underline{V}' \ ,$$

$$\underline{V}' = \sum_G \underline{V}''^G_{G'} (\rho \nabla M^{G'})^2 \ ,$$

$$g^G = -\tau^{-1}(M^G + N^G \tanh F^G) \ ,$$

$$g^{GG'} = (g_{GG'})^{-1} = \delta^{G'}_G \tau^{-1} N^G \mathrm{sech}^2 F^G$$

$$g = \det(g_{GG'}) \ ,$$

$$F^G = \frac{(V^G - a^{|G|}_{G'} v^{|G|}_{G'} N^{G'} - \frac{1}{2} A^{|G|}_{G'} v^{|G|}_{G'} M^{G'})}{(\pi[(v^{|G|}_{G'})^2 + (\phi^{|G|}_{G'})^2](a^{|G|}_{G'} N^{G'} + \frac{1}{2} A^{|G|}_{G'} M^{G'}))^{1/2}} \ ,$$

$$a^G_{G'} = \frac{1}{2} A^G_{G'} + B^G_{G'} \ , \tag{41}$$

where $A^G_{G'}$ and $B^G_{G'}$ are macrocolumnar-averaged interneuronal synaptic efficacies, $v^G_{G'}$ and $\phi^G_{G'}$ are averaged means and variances of contributions to neuronal electric polarizations, and nearest-neighbor interactions $\underline{V}'$ are detailed in other SMNI papers [14,16]. $M^{G'}$ and $N^{G'}$ in $F^G$ are afferent macrocolumnar firings, scaled to efferent minicolumnar firings by $N/N^* \sim 10^{-3}$, where $N^*$ is the number of neurons in a macrocolumn. Similarly, $A^{G'}_G$ and $B^{G'}_G$ have been scaled by $N^*/N \sim 10^3$ to keep $F^G$ invariant. This scaling is for convenience only. This mathematical description defines the mesocolumn concept introduced previously.

---

Figure 2.

---

At this stage, severe approximation in modeling is typically required in order to proceed towards solutions. However, advantage can be taken of experimentally observed columnar structure to first attempt to analytically scale the neuronal system into mesoscopic domains that are still relatively microscopic compared to the macroscopic regions to be described [59-61,63,64,88]. For purposes of macroscopic description, the minicolumnar structure effectively spatially averages the neuronal interactions within one to several firing periods.

The following development is proposed, which: (1) reasonably includes and averages over millions of synaptic interactions that exist between groups of hundreds of neurons, (2) analytically establishes the integrity of columnar domains and specifies their interactions, and (3) prepares the formulation of (1) and (2) to foresee their analytic inclusion into studies of macroscopic regions.

For purposes of detailing anatomical or physiological properties of neurons, it is simply incorrect to postulate an "average" neuron. However, for the purpose of macroscopic brain function, when considering millions of neurons, we repeat that it is reasonable to at least respect the incredibly similar modular structure present in all regions of the neocortex [59,61,62,64,103,104], still allowing for the differentiation among the laminar structure of individual modules and among neurons active at different temporal and spatial scales.

The neocortex has about $5 \times 10^{10}$ neurons distributed rather uniformly over approximately $5 \times 10^8$ minicolumns. (The visual cortex has double this density.) Within these minicolumns, a "vertical" structure is defined perpendicular to six highly convoluted laminae of total thickness $\approx 2.5 \times 10^3$ $\mu$m, principally by the efferent pyramidal cells. They exhibit vertical apical bundling of their dendrites in the upper laminae, and some of their recurrent axonal collaterals also ascend to upper laminae. A number of other fusiform, Martinotti, and stellate cells (granule cells in the sensory cortex and basket cells in the motor cortex) also contribute to this vertical organization. In general, laminae I to IV are afferent and laminae V and VI are efferent [70].

However, "horizontal" dendritic basal arborizations (treelike structures) of the pyramidal cells, tangential to the laminae, horizontal axonal collaterals of the pyramidal cells, and horizontal processes of stellate, Martinotti, and neonatal horizontal cells, all impart horizontal stratification to columnar interactions. Therefore, although the columnar concept has anatomical and physiological support, the minicolumnar boundaries are not so clearly defined [107]. If this stratification and other long-ranged afferent input to groups of minicolumns are incorporated, then it is possible that future work may have to define a physiological unit that encompasses a mesocolumn consisting of one to perhaps several minicolumns outputting to a macrocolumar extent. This study formalizes these circumstances by defining a mesocolumn with extent greater than $10^2$ $\mu$m, as an intermediate integral physiological unit encompassing one minicolumn. [See Fig. 1(b).] Dynamic nearest-neighbor interactions between mesocolumns are analytically defined by their overlapping neuronal interactions, in accordance with the observed horizontal columnar stratifications outlined above. [See Fig. 1(b').] This approach permits future analytic modifications, as differences between inter- and intra-minicolumnar interactions and circuitries become experimentally clarified.

The resulting picture of columnar interactions is relatively simpler than a mass of interacting neurons, but not so simple to the point of uselessness. A collection of average excitatory and inhibitory neuronal firings, as depicted in Fig. 1(a'), now define a continuum of mesocolumnar firings. A zero-order binomial distribution is easily intuited: Let $G$ denote $E$ or $I$ firings. Using the magnetic analogy, consider $E$ as a spin-up and $I$ as a spin-down magnet. A column of $N^G$ neurons can have a total firing of $Nn^G$, within time $\tau$, where $n^G$ is the fraction firing, ranging by 2's between $-N^G \le Nn^G \le N^G$. (Count firing as +1, nonfiring as −1.) For convenience, assume $Nn^G > 0$, which arises from $Nn^G$ firings plus $\frac{1}{2}(N^G - Nn^G)$ cancelling pairs of firings and nonfirings. This gives a total of $\frac{1}{2}(N^G - Nn^G) + Nn^G = \frac{1}{2}(N^G + Nn^G)$ firings and $N^G - \frac{1}{2}(N^G + Nn^G)$ nonfirings. The degeneracy factor, as a function of the firing rate $Nn^G$, is the number of ways $N^G$ neurons can produce a given firing pattern, i.e., the binomial distribution. Note that the binomial coefficient is unity for states of all firing or all nonfiring, and peaks as $N^G!/[(N^G/2)!]^2 \approx 2^{N^G + \frac{1}{2}}(\pi N^G)^{-\frac{1}{2}}$ for $Nn^G = 0$. In the range $Nn^G \approx 0$, there is maximal degeneracy of information encoded by mesocolumnar firings. This argument analytically articulates the meaning of "neuronal degeneracy" and also of the ubiquitous, often ambiguous "average neuron." However, reasonable properties of mesocolumns, not of average neurons, are developed here for macroscopic study.

The properly calculated distribution contains nearest-neighbor mesocolumnar interactions expressed as spatial-derivative correction terms. This verifies that in macroscopic activity, where patterns of mesocolumnar firing vary smoothly over neighboring mesocolumns, it is consistent to approximate mesocolumnar interactions by including only second-order gradient correction terms. We calculate macroscopic states of mesocolumnar firings, which are subject to these constraints. Excitatory and inhibitory sensitivity to the neuronal parameters survives, similar to the sensitivity encountered by single neurons.

Nearest-neighbor interactions are "induced" between minicolumns in the following way. The bulk of short-ranged interactions engaging the neurons in a minicolumn do not take place within the minicolumn, but rather within a spatial extent the size of a macrocolumn, comprising roughly 1000 minicolumns. If we consider the area of influence of a minicolumn as extending out to a macrocolumn's reach, then the area of interactions engaged by nearest-neighbor minicolumns has an offset circle of influence [16]. (See Fig. 3.) Therefore, within one or two epochs spanning the refractory periods of a typical neuron, i.e., including the absolute refractory time after a firing during which no new spikes can be generated, and the relative refractory period during which spikes can be produced only at a decreased sensitivity [71], interactions engaged by a given minicolumn can be extended out to areas of influence engaged by their nearest neighbors. This is what physically is being calculated by a careful mathematical treatment of overlapping interactions. In this manner, microscopic degrees of freedom of many types of neurons (many of which are only crudely classified by the above definitions), synapses, neurotransmitters, cellular architecture, and circuitries, may be practically weighted and averaged for macroscopic considerations.

---

Figure 3.

---

In the steps described above, the mesocolumnar conditional probability that a given mesocolumn will fire is calculated, given its direct interactions with other mesocolumns just previously fired. Thus a transition rate from one state of mesocolumnar firing to another state closely following the first state is obtained. A string, or path of these conditional probabilities connects the mesocolumnar firings at one time to the firing at any time afterwards. Many paths may link the same initial and final state. In this way the long-time conditional probability of all possible mesocolumnar firings at any given time is obtained. A Lagrangian is thereby derived which explicitly describes the time evolution of the neocortical region in terms of its initial distribution of firings, and expressed in terms of its mesoscopic order parameters which retain a functional form derived from microscopic neuronal interactions. A major benefit derived from this formalism is a variational principle that permits extrema equations to be developed. This also makes it possible to draw analogies to the "orienting field" and "temperature" of other collective systems.

### *2. Mesocolumns*

Define a mesocolumn as a domain with $N$ neurons, with stochastic memory $\tau$ approximately equal to 1 to several $\tau_n$, the total refractory period of a typical neuron. Denote by indices $E$ and $I$ two chemically and anatomically independent firing fields; $G$ denotes either field. For this study we do not consider dynamic synaptic modifications, which typically take place as a result of one to many epochs of macroscopic temporal activity effecting such plastic changes [87]. Therefore, take as independent of spacetime,

$$|v_{jk}|,\ \phi_{jk} \rightarrow v^G_{G'},\ \phi^G_{G'}\ ,$$

$$V_j,\ A_{jk},\ B_{jk} \rightarrow V^G,\ A^G_{G'},\ B^G_{G'}\ ,$$

$$\tfrac{1}{2} A^G_{G'} + B^G_{G'} = a^G_{G'}\ . \tag{42}$$

The greater importance of $I$ synapses (circuitry, proximity to soma) increases effective inhibitory $v_{jk}$ and $a_{jk}$. Take

$$\sum_k v_{jk} A_{jk} \approx \sum_{G'} v^G_{G'} A^G_{G'} N^{G'}\ ,$$

$$v^G_E = -v^G_I > 0\ ,$$

$$N = N^E + N^I\ . \tag{43}$$

Define $Nn^G(r)$, a mesocolumn centered at the two-dimensional point $r$, as the mesocolumnar average of $\sigma_j$. [See Figs. 1(b) and 1(b').] Derive the conditional probability $P^G[M^G(r; t+\tau)|M^{\bar{G}}(r'; t)]$ for the firing transition to mesocolumn $M^G(r; t+\tau) \equiv Nn^G(r)$ at time $t+\tau$, from all contributing mesocolumns $M^G(r'; t) \equiv Nm^G(r')$ associated with neurons $\sigma_k$ at time $t$. $M^{\bar{G}}$ represents contributions from both $M^E$ and $M^I$ in $F^G_j$:

$$\prod_G P^G[M^G(r; t+\tau)|M^{\bar{G}}(r'; t)] = \sum_{\sigma_j=\pm 1} \delta(\sum_{j\in E} \sigma_j - Nn^E)\delta(\sum_{j\in I} \sigma_j - Nn^I) \prod_j^N p_{\sigma_j}\ . \tag{44}$$

For $r = r'$, the result of the mesoscopic averaging of independent $E$ and $I$ fields is easily intuited, as explained above: The contribution of $N^G$ averaging $\pm 1$ firings to establish a firing $Nn^G$ has degeneracy

$$\left[ \begin{matrix} N^G \\ \frac{1}{2}(N^G + n^G N) \end{matrix} \right] = \frac{N^G!}{(\frac{1}{2}(N^G + n^G N))!(\frac{1}{2}(N^G - n^G N))!}$$

$$-N^G \le Nn^G = M^G \le N^G \ . \tag{45}$$

A binomial distribution of $Nn^G$ is therefore anticipated, weighted by the averaged firing probability, with

$$F_j^G \to \text{typical } F^G \ . \tag{46}$$

However, the nearest-neighbor interactions must be calculated more precisely. Calculate

$$P^G[M^G(r; t+\tau)|M^{\bar{G}}(r'; t)] = (2\pi)^{-1} \int_{-\infty}^{\infty} dQ^G \exp(iNn^G Q^G) \prod_{j=1}^{N^G} C_j^G \ ,$$

$$C_j^G \ \{ \ F_j^G[M^{\bar{G}}(r+\varepsilon)] \ \} \ = \cosh(F_j^G + iQ^G)\mathrm{sech} F_j^G, \ \ r+\varepsilon = r' \ . \tag{47}$$

Expand $C_j^G$ about $Nn^G(r)$, using directional derivatives $D_{\hat{\varepsilon}}^{1,2}$,

$$C_j^G \approx [1 + |\varepsilon| D_{\hat{\varepsilon}}^1 + \tfrac{1}{2}|\varepsilon|^2 D_{\hat{\varepsilon}}^2] C^G(r),$$

$$\hat{\varepsilon} = \varepsilon/|\varepsilon| = (r'-r)/|r'-r| \ , \tag{48}$$

retaining only first- and second-order derivatives of $M^E$ and $M^I$, and obtain

$$P^G[M^G(r; t+\tau)|M^{\bar{G}}(r; t)] = (2\pi)^{-1} \int_{-\infty}^{\infty} dQ^G \exp(iNn^G Q^G)[\cosh(F^G + iQ^G)\mathrm{sech} F^G]^{N^G}$$

$$\times \prod_{j=1}^{N^G} (1 + dF_{j1}^G K_1 + dF_{j2}^G K_2),$$

$$K_1 = \sinh(iQ^G)\mathrm{sech}(F^G + iQ^G)\mathrm{sech} F^G \ ,$$

$$K_2 = -2K_1 \tanh F^G \ , \tag{49}$$

where $dF_{j1}^G$ and $dF_{j2}^G$ are more involved algebraic expressions given in the SMNI literature [14].

To zero order, neglect $K_1$ and $K_2$ terms, express $\cosh^{N^G}$ as a sum of exponentials, expand as a binomial, carry out the $\int dQ^G$, and obtain

$$P_0^G[M^G(t+\tau)|M^{\bar{G}}(t)] = (1 + f^G)^{-N^G} \binom{N^G}{\lambda^G} (f^G)^{\lambda^G},$$

$$f^G = \exp(-2F^G),$$

$$\lambda^G = [[\tfrac{1}{2}(Nn^G + N^G)]], \tag{50}$$

where $\lambda^G$ is the greatest integer in the double brackets on the right-hand side. This is the anticipated binomial distribution with mean

$$< \lambda^G >_0 = N^G \exp(-F^G)/[(\exp(F^G) + \exp(-F^G)],$$

$$< Nn^G >_0 = -N^G \tanh F^G \ , \tag{51}$$

and variance

$$< \lambda^G \lambda^{G'} >_0 - < \lambda^G >_0 < \lambda^{G'} >_0 = \frac{1}{4} \delta_{G'}^G N^G \mathrm{sech}^2 F^G \ . \tag{52}$$

Thus $n^G$ and $m^G$ are defined as mesocolumnar averaged neurons, with their anticipated zero-order statistical firing weights. This explicitly demonstrates how sensitive $< Nn^G >_0$ is to changes in sign of the threshold factor $F^G$ in the step function $\tanh F^G$: $\tanh(\pm\infty) = \pm 1$, $\tanh(\pm 1) = \pm 0.76$, $\tanh(0) = 0$. As expected, in the absence of interactions, setting $B^G_{G'} = 0$ and with $Nm^G = -N^G$, then $< Nn^G >_0 = -N^G$; no firing occurs. As $< Nn^G >_0 \geq 0$ maintains $< Nn^G >_0 > -N^G$ for all $Nm^G$, $B^G_{G'}$ may be taken to simulate nonsynaptic influences on $Nn^G$.

Consistency of the above scheme requires a definition of this long-wavelength scale. The conditions placed on the above calculation are evaluated to be essentially

$$1 > (\rho \nabla M^G)^2/(24N), \tag{53}$$

where $\rho$ is the spatial extent of a minicolumn, which is consistent with this macroscopic development [16]. That is, we can consistently define a nearest-neighbor mesocolumnar potential to take into account interactions across thousands of minicolumns since the spatial changes in firings across minicolumns are effectively small enough to permit this perturbative procedure. Calculations confirm the consistency of this derivation of intercolumnar interactions with experimental observations.

The prototypical diffusion system describes Brownian motion, wherein the stochastic memory of the macroscopic system depends only on the immediate past history of the system at one specified unit of earlier time, and in a linearly functional manner [1,2]. Here, the $G$-space diffusion is expressed by a nonlinear dependence on this past firing state, and the stochastic memory must be carefully defined. Real-space diffusion is represented by the gradient couplings. These fluctuations are physically important for various excitations and possible critical behavior of second-order phase transitions between ordered and disordered states, e.g., as in magnetic systems [4,108].

Figures 1(c) and 1(c') illustrate how the mesocolumnar structure is a substrate for activity persisting for hundreds of milliseconds over a spatial region containing $\Lambda \approx 5 \times 10^5$ mesocolumns, spanning $\approx 10^{-2}$ of a total cortical area of $4 \times 10^{11}$ $\mu\text{m}^2$. Extended regional activity is possible, whereby conglomerates of 10 to 30 regions may interact [59].

### *3. Mesocolumnar short-time propagator*

To first order in $dF$, the distribution function $p$, of a given state of firing $Nn^G$ occurring given the earlier state $M^G$, can be defined in terms of variables that facilitate this development. For large $N^G$ and $N^G F^G$, this binomial distribution is asymptotically Gaussian [102]:

$$P^G[M^G(t+\tau)|M^{\bar{G}}(t)] = (2\pi\tau g'^{GG})^{\frac{1}{2}} \exp(-N\tau \underline{L}^G),$$

$$\underline{L}'^G = N^{-1}[(Nn^G - M^G)/\tau - g'^G]^2/(2g'^{GG}),$$

$$g'^G = -\tau^{-1}[M^G + N^G(1 + dF^G)\tanh F^G],$$

$$g'^{GG} = \tau^{-1} N^G (1 + dF^G) \text{sech}^2 F^G\ , \tag{54}$$

where no sum is implied over repeated indices, and we use $\underline{L}'$ to denote that we have not yet expanded the $dF$ terms.

Define time epochs $t$ in units of $\tau$, the total mesocolumnar refractory period, in terms of integer $s \geq 0$, from an initial time $t_0$:

$$t_s = s\tau + t_0\ . \tag{55}$$

For large time epochs to be considered, a continuum differential time scale is defined by $dt \leq \tau$. Within $\tau$, consistent with this long-time development and consistent with the previous mesocolumnar averaging of neuronal interactions, define

$$\underline{L}'^G = N^{-1}(\dot{M}^G - g'^G)^2/(2g'^{GG}),$$

$$\tau \dot{M}^G = \tau dM^G/dt = Nn^G - M^G\ . \tag{56}$$

Then the Markovian mesocolumnar short-time propagator, the conditional probability $p$, is developed for short times $\theta \approx \tau$ relative to any fixed time $t$:

$$p[M(t+\theta), t+\theta] = (2\pi\theta)^{-1} \int dM(t) g'(t)^{1/2} \exp \{ -N \underline{S}[M(t+\theta), M(t)] \} \; p[M(t), t],$$

$$dM = dM^E dM^I \; ,$$

$$g' = (\det g')^{-1} = (g'^{EE} g'^{II})^{-1} \; , \tag{57}$$

and $\underline{S}$ is defined by requiring that the system evolve by the principle of maximal probability:

$$\underline{S} = \min \int_t^{t+\theta} dt' \; \underline{L}'[\dot{M}(t'), M(t')],$$

$$\underline{L}' = \underline{L}'^E + \underline{L}'^I \; . \tag{58}$$

For small $\theta$, relative to the long times considered, with $N \gg 1$, contributions to $p$ at $t+\theta$ are heavily weighted within

$$|\Delta M^G| = |M^G(t+\theta) - M^G(t)| < \theta^{\frac{1}{2}} \; , \tag{59}$$

and therefore the quadratic $\dot{M}$ terms in $\underline{L}'$ must be carefully developed.

Finally, $\underline{L}'$ is expanded about $\underline{L}$, treating $dF^G$ as perturbations. This yields Eq. (41) in terms of $\underline{V}'$, the nearest-neighbor potential.

*4. Further development of mesocolumn model*

As pointed out in this derivation [14,15], this microscopic scale itself represents a high aggregation of submicroscopic scales, aggregating effects of tens of thousands of quanta of chemical transmitters as they influence the $5 \times 10^{-3}$ $\mu$m scale of biomolecular leaflets of phospholipid molecules. This microscopic scale has been aggregated up to the mesoscopic scale, again using the general property of probability distributions, that the aggregated distribution $P_q$ of variable $q$ is calculated from the joint distribution $P_{q_1 q_2}$ of underlying variables $q_1$ and $q_2$,

$$P_q(q) = \int dq_1 dq_2 P_{q_1 q_2}(q_1, q_2) \delta(q - (q_1 + q_2)) \; . \tag{60}$$

To summarize up to this point, the mathematical development of mesocolumns establishes a mesoscopic Lagrangian $\underline{L}$, defining the short-time probability distribution of firings in a minicolumn, composed of ~ $10^2$ neurons [59,61,62,64,103,104,109], given its just previous interactions with all other neurons in its macrocolumnar surround.

$$P = \prod_G P^G[M^G(r; t+\tau) | M^{\bar{G}}(r'; t)]$$

$$= \sum_{\sigma_j} \delta\left( \sum_{jE} \sigma_j - M^E(r; t+\tau) \right) \delta\left( \sum_{jI} \sigma_j - M^I(r; t+\tau) \right) \prod_j^N p_{\sigma_j}$$

$$\approx \prod_G (2\pi\tau g^{GG})^{-1/2} \exp(-N\tau \underline{L}^G) \; ,$$

$$P \approx (2\pi\tau)^{-1/2} g^{1/2} \exp(-N\tau \underline{L}) \; , \tag{61}$$

where $\underline{L}$ is defined in terms of its drift, diffusion, and potential, all of which depend sensitively on the threshold factor $F^G$,

$$F^G = \frac{V^G - a_{G'}^{|G|} v_{G'}^{|G|} N^{G'} - \frac{1}{2} A_{G'}^{|G|} v_{G'}^{|G|} M^{G'}}{\{\pi[(v_{G'}^{|G|})^2 + (\phi_{G'}^{|G|})^2](a_{G'}^{|G|} N^{G'} + \frac{1}{2} A_{G'}^{|G|} M^{G'})\}^{1/2}} \ . \tag{62}$$

In the first SMNI papers, long-ranged interactions were included in $\underline{L}$ by adding potential terms simulating these constraints, i.e., adding $J_G M^G$ to $\underline{L}$, where $J_G$ was numerically adjusted to account for these interactions.

In order to more properly include long-ranged fibers, so that interactions among macrocolumns may be included in the numerical studies, the $J_G$ terms are dropped, and more realistically replaced by a modified threshold factor $F^G$,

$$F^G = \frac{V^G - a_{G'}^{|G|} v_{G'}^{|G|} N^{G'} - \frac{1}{2} A_{G'}^{|G|} v_{G'}^{|G|} M^{G'} - a_{G'}^{\ddagger|G|} v_{G'}^{|G|} N^{\ddagger|G|'} - \frac{1}{2} A_{G'}^{\ddagger|G|} v_{G'}^{|G|} M^{\ddagger G'}}{(\pi[(v_{G'}^{|G|})^2 + (\phi_{G'}^{|G|})^2](a_{G'}^{|G|} N^{G'} + \frac{1}{2} A_{G'}^{|G|} M^{G'} + a_{G'}^{\ddagger|G|} N^{\ddagger G'} + \frac{1}{2} A_{G'}^{\ddagger|G|} M^{\ddagger G'}))^{1/2}} \ ,$$

$$A_E^{\ddagger I} = A_I^{\ddagger E} = A_I^{\ddagger I} = B_E^{\ddagger I} = B_I^{\ddagger E} = B_I^{\ddagger I} = 0 \ ,$$

$$a_E^{\ddagger E} = \frac{1}{2} A_E^{\ddagger E} + B_E^{\ddagger E} \ . \tag{63}$$

Here, afferent contributions from $N^{\ddagger E}$ long-ranged excitatory fibers, e.g., cortico-cortical neurons, have been added, where $N^{\ddagger E}$ might be on the order of 10% of $N^*$: Nearly every pyramidal cell has an axon branch that makes a cortico-cortical connection; i.e., the number of cortico-cortical fibers may be as high as $10^{10}$ [45].

At this point, attention is also drawn to the similar algebraic structure of the threshold factors in Eqs. (32) and (63), illustrating common forms of interactions between their entities, i.e., neurons and columns of neurons, respectively. The nonlinear threshold factors are defined in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges.

The net short-time probability distribution can be folded over and over (multiplied) in time increments $\Delta t$ to yield a path-integral algorithm for calculating the long-time probability distribution [27]. This result depends on the use of the Markov property of our distribution, wherein the short-time evolution of the system at time $t + \tau$ depends only on the state of the system at time $t$. For example, in a very compacted notation, labeling $u$ intermediate time epochs by $s$, i.e., $t_s = t_0 + s\Delta t$, in the limits $u \rightarrow \infty$ and $\Delta t \rightarrow 0$, and assuming $M_{t_0} = M(t_0)$ and $M_t = M(t \equiv t_{u+1})$ are fixed,

$$P[M_t|M_{t_0}] = \int \cdots \int dM_{t-\Delta t} dM_{t-2\Delta t} \cdots dM_{t_0+\Delta t}$$

$$\times P[M_t|M_{t-\Delta t}] P[M_{t-\Delta t}|M_{t-2\Delta t}] \cdots P[M_{t_0+\Delta t}|M_{t_0}] \ ,$$

$$P[M_t|M_{t_0}] = \int \cdots \int \underline{D}M \exp\left( - \sum_{s=0}^{u} \Delta t \underline{L}_s \right) ,$$

$$\underline{D}M = (2\pi \hat{g}_0^2 \Delta t)^{-1/2} \prod_{s=1}^{u} (2\pi \hat{g}_s^2 \Delta t)^{-1/2} dM_s \ . \tag{64}$$

Similarly, the short-time probability distribution $P$ can be folded over and over at each point $r$, to give a field-theoretic Lagrangian, $L(r, t)$. The above "prepoint-discretization" representation of $\underline{L}$ derived for the neocortex, e.g., $g_s^G = g^G[M^G(t_0 + s\Delta t)]$, disguises the Riemannian geometry induced by the nonconstant metric $g_{GG'}$, discussed further below.

**D. Macroscopic development**

*1. General description*

Inclusion of all the above microscopic and mesoscopic features of the neocortex permits a true non-phenomenological Gaussian-Markovian formal development for macroscopic regions encompassing ~ $5 \times 10^3$ macrocolumns of spatial extent ~ $5 \times 10^9$ $\mu\text{m}^2$, albeit one that is still highly nonlinear and nonequilibrium. The development of mesocolumnar domains presents conditional probability distributions for mesocolumnar firings with spatially coupled nearest-neighbor interactions. The macroscopic spatial folding of these mesoscopic domains and their macroscopic temporal folding of tens to hundreds of $\tau$, with a resolution of at least $\tau/N$ [16], yields a true path-integral formulation, in terms of a macroscopic Lagrangian possessing a variational principle for most-probable firing states. At this point in formal development, no continuous-time approximation has yet been made; this is done, with clear justification, only for some applications discussed below. Much of this algebra is greatly facilitated by, but does not require, the use of Riemannian geometry to develop the nonlinear means, variances, and "potential" contributions to the Lagrangian [27].

This formalism can also be recast in several other representations [27], perhaps more familiar to other investigators, and sometimes more useful for particular calculations. For example, a Hamiltonian formulation can be obtained, one that does not permit simple "energy"-type conservation approximations, but one that does permit the usual time-evolution picture. The time-dependent differential macroscopic probability distribution, or "propagator," is found to satisfy a true Fokker-Planck second-order partial-differential equation, expressing the rate of change of the distribution as the sum of contributions from nonlinear drifts and diffusion in the space of $E$ and $I$ firings. With respect to a reference stationary state, a well-defined information, or "entropy," can be formulated. Also, a set of Langevin rate equations for $E$ and $I$ firings can be obtained, expressing the rate of change of firings as the sum of drifts and multiplicative noise. The most-probable firing states derived variationally from the path-integral Lagrangian represent a reasonable average over the noise in the Langevin system; the noise cannot be indiscriminately neglected. Because of the presence of multiplicative noise, the Langevin system differs in its Itô (prepoint) and Stratonovich (midpoint) discretizations. Furthermore, there exists a midpoint-discretized covariant description, in terms of the Feynman Lagrangian $\underline{L}_F$, which is defined such that (arbitrary) fluctuations occur about solutions to the Euler-Lagrange variational equations. In contrast, the usual Itô and corresponding Stratonovich discretizations are defined such that the path integral reduces to the Fokker-Planck equation in the weak-noise limit. These points are discussed further below, and more explicitly derived in Appendix A.

*2. Regional short-time propagator*

Define the $\Lambda$-dimensional spatial vector $\tilde{M}_s$ at time $t_s$,

$$\tilde{M}_s = \{ M_s^\nu = M_s(r^\nu);\ \nu = 1, \cdots, \Lambda \} ,$$

$$M_s^\nu = \{ M_s^{G\nu};\ G = E, I \} . \tag{65}$$

For macroscopic space-time considerations, mesoscopic $\rho$ (spatial extent of a minicolumn) and $\tau$ scales are measured by $dr$ and $dt$. In the continuum limits of $r$ and $t$,

$$M_s^{G\nu} \rightarrow M^G(r, t),\ \dot{M}_s^{G\nu} \rightarrow dM^G/dt,$$

$$(M^{G,\nu+1} - M^{G\nu})/(r^{\nu+1} - r^\nu) \rightarrow \nabla_r M^G \ . \tag{66}$$

The previous development of mesocolumnar interactions via nearest-neighbor derivative couplings permits the regional short-time propagator $\tilde{P}$ to be developed in terms of the Lagrangian $\underline{L}$ [110]:

$$\tilde{P}(\tilde{M}) = (2\pi\theta)^{-\Lambda/2} \int d\tilde{M}\ g^{\Lambda/2} \exp[-N\tilde{S}(\tilde{M})]\tilde{P}(\tilde{M}),$$

$$\tilde{S} = \min \int_t^{t+\theta} dt' L[\dot{M}(t'), M(t')],$$

$$L = \Lambda\Omega^{-1} \int d^2 r \, \underline{L} \; , \tag{67}$$

where $\Omega$ is the area of the region considered, and

$$\Lambda\Omega^{-1} \int d^2 r = \Lambda\Omega^{-1} \int dx\, dy = \lim_{\Lambda \to \infty}^{\rho \to 0} \sum_{\nu=1}^{\Lambda} \; . \tag{68}$$

The Euler-Lagrange (EL) equations, giving the extrema $\ll M^G \gg$, are obtained from $\delta \underline{\tilde{S}} = 0$ [111]. The Einstein convention is used to designate summation over repeated indices, and the following notation for derivatives is used:

$$\begin{aligned}
(\cdots)_{:z} &= d(\cdots)/dz, \quad z = \{ \, x, y \, \} \, , \\
(\cdots)_{,G} &= \partial(\cdots)/\partial M^G, \\
(\cdots)_{,\dot{G}} &= \partial(\cdots)/\partial(dM^G/dt), \\
(\cdots)_{,G_{:z}} &= \partial(\cdots)/\partial(dM^G/dz), \\
(\cdots)_{,\nabla G} &= \hat{x}\partial(\cdots)/\partial(dM^G/dx) + \hat{y}\partial(\cdots)/\partial(dM^G/dy).
\end{aligned} \tag{69}$$

The EL equations are:

$$\begin{aligned}
\delta \underline{L} &= 0, \\
\delta_G \underline{L} &= \underline{L}_{,G} - \nabla \cdot \underline{L}_{,\nabla G} - \underline{L}_{,\dot{G}:t} = 0, \\
\nabla \cdot \underline{L}_{,\nabla G} &= \underline{L}_{,G_{:z}:z} \\
&= (\underline{L}_{,G_{:z}}{}_{,G'}) M^{G'}{}_{:z} + (\underline{L}_{,G_{:z}}{}_{,G'_{:z}}) M^{G'}{}_{:zz}, \\
\underline{L}_{,\dot{G}:t} &= (\underline{L}_{,\dot{G}}{}_{,G'}) \dot{M}^{G'} + (\underline{L}_{,\dot{G}}{}_{,\dot{G}'}) \ddot{M}^{G'}.
\end{aligned} \tag{70}$$

This exhibits the extremum condition as a set of 12 first-order differential equations in the 12 variables $\{ M^G, \dot{M}^G, \ddot{M}^G, M^G{}_{:z}, M^G{}_{:zz} \}$ in $r - t = (x, y, t)$ space, with coefficients nonlinear in $M^G$.

*3. Regional long-time propagator*

With $\tilde{P}$ properly defined by this space-time mesh, a path-integral formulation for the regional long-time propagator at time $t = (u + 1)\theta + t_0$ is developed:

$$\begin{aligned}
\tilde{P}[\tilde{M}(t)]\, d\tilde{M}(t) &= \int \cdots \int \underline{D}\tilde{M} \exp(-N \int_{t_0}^{t} dt' L), \\
\tilde{P}[\tilde{M}(t_0)] &= \delta(\tilde{M} - \tilde{M}_0), \\
\underline{D}\tilde{M} &= \prod_{s=1}^{u+1} \prod_{\nu=1}^{\Lambda} \prod_{G}^{E,I} (2\pi\theta)^{-\frac{1}{2}} (g_s^\nu)^{1/4} dM_s^{G\nu}.
\end{aligned} \tag{71}$$

Note that, even for $N\tau L \approx 1$, $N \int_{t_0}^{t} dt' \, L$ is very large for macroscopically large time $(t - t_0)$ and macroscopic size $\Lambda$, demonstrating how extrema of $L$ define peaked maximum probability states. This derivation can be viewed as containing the dynamics of macroscopic causal irreversibility, whereby $\tilde{P}$ is an unstable fixed point about which deviations from the extremum are greatly amplified [112].

*4. Riemannian geometry*

A series of papers has recognized that a few of the most popular Riemannian-geometric transformation properties possessed by physics systems might be advantageous for a theory of cortical interactions, i.e., most specifically in the cerebellum, and they have gone further to postulate this geometry as the essential component of their theory [113-115].

As developed most notably by Einstein [116], Riemannian geometry has been firmly established as a necessary component of the foundations of physics. There are still two viable camps of opinions, considering this geometry itself as a basic foundation [117], or considering the physical entities on which its transformations operate as the basic foundation [118]. However, there is unanimous agreement that Riemannian geometry is an essential theoretical construct to explain some observed physical phenomena. The existence of Riemannian geometry also is a natural mathematical consequence of properties possessed by quite general stochastic systems, including those models of neural systems assumed or endorsed by most investigators [119]. These properties have been stressed in the SMNI series of papers.

It is the purpose here to stress these general properties, and to make the short but important observation that there is indeed mathematical support on which to conjecture possible neural mechanisms that might exist as a result of invariance under Riemannian-geometric transformations. This observation then leads us back to the spirit, if not the essence, of other neuroscience investigators. However, whereas they *hypothesize* a Riemannian metric *between* cortical regions, SMNI *derives* a Riemannian metric *within* each cortical region, quite a physical distinction.

Corresponding to the differential-operator ordering problem in the Fokker-Plank equation is the discretization problem in the path integral and in the Langevin rate equations, both of which are equivalent mathematical representations of the Fokker-Plank equation [27,29-31,120,121]. An overview of these equations is required to at least note where the Riemannian geometry enters. Appendix A provides a brief derivation.

Consider a multivariate system with variance a general nonlinear function of the variables. The Einstein summation convention helps to compact the equations, whereby repeated indices in factors are to be summed over.

The Itô (prepoint) discretization for a system of stochastic differential equations is defined by

$$\bar{t}_s \in [t_s, t_s + \Delta t] \equiv [t_s, t_{s+1}] \ ,$$

$$M(\bar{t}_s) = M(t_s) \ ,$$

$$dM(\bar{t}_s)/dt = M(t_{s+1}) - M(t_s) \ . \tag{72}$$

The stochastic equations are then written as

$$dM^G/dt = f^G + \hat{g}_i^G \eta^i \ ,$$

$$i = 1, \cdots, \Xi \ ,$$

$$G = 1, \cdots, \Theta \ , \tag{73}$$

where $\Xi$ and $\Theta$ are the (arbitrary) ranges of $i$ and $G$, respectively. These equations might be used to describe the stochastic firing of mesocolumns, or if scaled by $N^{-1}$, they could be used to describe "average" neurons. Once parameters are fit to data, as we report here using the integral probability distribution representation below, these equations can be used to construct a brain simulation, e.g., to study the influence of changing contexts that cannot be so well incorporated into these equations.

The operator ordering (of the $\partial/\partial M^G$ operators) in the Fokker-Planck equation corresponding to this discretization is

$$\frac{\partial P}{\partial t} = VP + \frac{\partial(-g^G P)}{\partial M^G} + \frac{1}{2}\frac{\partial^2(g^{GG'}P)}{\partial M^G \partial M^{G'}} \ ,$$

$$g^G = f^G + \frac{1}{2}\,\hat{g}_i^{G'}\,\frac{\partial \hat{g}_i^G}{\partial M^{G'}}\;,$$

$$g^{GG'} = \hat{g}_i^G\,\hat{g}_i^{G'}\;. \tag{74}$$

where a ‘‘potential’’ $V$ is present in some systems, e.g., sometimes used to explicitly include boundary conditions. Above, a potential term was derived to describe nearest-neighbor interactions. As described below, this partial-differential representation is very useful in determining the stability and duration of STM.

The Feynman Lagrangian corresponding to this Fokker-Planck and set of Langevin equations may be written in the midpoint discretization, corresponding to

$$M(\bar{t}_s) = \frac{1}{2}\,[M(t_{s+1}) + M(t_s)]\;. \tag{75}$$

This discretization defines a covariant Lagrangian $\underline{L}_F$ that possesses a variational principle for arbitrary noise, and that explicitly portrays the underlying Riemannian geometry induced by the metric tensor $g_{GG'}$, calculated to be the inverse of the covariance matrix $g^{GG'}$:

$$P = \int \cdots \int \underline{D}M \exp\left(-\sum_{s=0}^{u} \Delta t \underline{L}_{Fs}\right),$$

$$\underline{D}M = g_{0_+}^{1/2}(2\pi\Delta t)^{-\Theta/2}\prod_{s=1}^{u} g_{s_+}^{1/2}\prod_{G=1}^{\Theta}(2\pi\Delta t)^{-1/2}dM_s^G\;,$$

$$\int dM_s^G \to \sum_{\iota=1}^{N^G}\Delta M_{\iota s}^G\,,\, M_0^G = M_{t_0}^G\,,\, M_{u+1}^G = M_t^G\;,$$

$$\underline{L}_F = \frac{1}{2}\,(dM^G/dt - h^G)g_{GG'}(dM^{G'}/dt - h^{G'}) + \frac{1}{2}\,h^G{}_{;G} + R/6 - V\;,$$

$$(\cdots)_{,G} = \frac{\partial(\cdots)}{\partial M^G}\;,$$

$$h^G = g^G - \frac{1}{2}\,g^{-1/2}(g^{1/2}g^{GG'})_{,G'}\;,$$

$$g_{GG'} = (g^{GG'})^{-1}\;,$$

$$g_s[M^G(\bar{t}_s), \bar{t}_s] = \det(g_{GG'})_s\,,\, g_{s_+} = g_s[M_{s+1}^G, \bar{t}_s]\;,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF}h^G = g^{-1/2}(g^{1/2}h^G)_{,G}\;,$$

$$\Gamma^F_{JK} \equiv g^{LF}[JK, L] = g^{LF}(g_{JL,K} + g_{KL,J} - g_{JK,L})\;,$$

$$R = g^{JL}R_{JL} = g^{JL}g^{JK}R_{FJKL}\;,$$

$$R_{FJKL} = \frac{1}{2}\,(g_{FK,JL} - g_{JK,FL} - g_{FL,JK} + g_{JL,FK}) + g_{MN}(\Gamma^M_{FK}\Gamma^N_{JL} - \Gamma^M_{FL}\Gamma^N_{JK})\;, \tag{76}$$

where $R$ is the Riemannian curvature, and we also have explicitly noted the discretization in the mesh of $M_{\iota s}^G$ by $\iota$. If $M$ is a field, e.g., also dependent on a spatial variable $x$ discretized by $\nu$, then the variables $M_s^G$ is increased to $M_s^{G\nu}$, e.g., as we have prescribed for the macroscopic neocortex. The term $R/6$ in $\underline{L}_F$ includes a contribution of $R/12$ from the WKB approximation to the same order of $(\Delta t)^{3/2}$ [27].

A prepoint discretization for the same probability distribution $P$ gives a much simpler algebraic form,

$$M(\bar{t}_s) = M(t_s) ,$$

$$\underline{L} = \frac{1}{2}(dM^G/dt - g^G)g_{GG'}(dM^{G'}/dt - g^{G'}) - V , \tag{77}$$

but the Lagrangian $\underline{L}$ so specified does not satisfy a variational principle useful for moderate to large noise; its associated variational principle only provides information useful in the weak-noise limit [122]. The neocortex presents us with a system of moderate noise. Still, this prepoint-discretized form has been quite useful in all systems examined thus far, simply requiring a somewhat finer numerical mesh. Note that although integrations are indicated over a huge number of independent variables, i.e., as denoted by $dM_s^{G\nu}$, the physical interpretation afforded by statistical mechanics makes these systems mathematically and physically manageable.

It must be emphasized that the output need not be confined to complex algebraic forms or tables of numbers. Because $\underline{L}_F$ possesses a variational principle, sets of contour graphs, at different long-time epochs of the path-integral of $P$, integrated over all its variables at all intermediate times, give a visually intuitive and accurate decision aid to view the dynamic evolution of the scenario. For example, this Lagrangian approach permits a quantitative assessment of concepts usually only loosely defined.

| Concept | Lagrangian equivalent |
|---|---|
| Momentum | $\Pi^G = \dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)}$ |
| Mass | $g_{GG'} = \dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)\partial(\partial M^{G'}/\partial t)}$ |
| Force | $\dfrac{\partial \underline{L}_F}{\partial M^G}$ |
| $F = ma$ | $\delta \underline{L}_F = 0 = \dfrac{\partial \underline{L}_F}{\partial M^G} - \dfrac{\partial}{\partial t}\dfrac{\partial \underline{L}_F}{\partial(\partial M^G/\partial t)}$ |

(78)

These physical entities provide another form of intuitive, but quantitatively precise, presentation of these analyses [38].

Using the Lagrangian formulation, a systematic numerical procedure has been developed for fitting parameters in such stochastic nonlinear systems to data using methods of very fast simulated re-annealing [58], and then integrating the path integral using a non-Monte Carlo technique especially suited for nonlinear systems [123-125]. This methodology has been applied with success to military modeling [36,126] and to financial markets [41,42], and we will be using it in this neocortical system to correlate EEG to behavioral states [24].

The key issue is that Riemannian geometry is not required to derive the mathematics of multiplicative Gaussian-Markovian systems. More interestingly, after this derivation, it can be demonstrated that the space of random variables actually induces a Riemannian geometry, obtained explicitly by simply (in hindsight) reorganizing terms in their defining equations. Then, the differential and path-integral representations can be rewritten only in terms of functions $f(M)$ of random variables $M$ that are tensor invariant under quite generally nonlinear point transformations, i.e., $M' = M'(M)$.

The derived probability distribution also is invariant under an equivalence class of discretizations or "look both ways" interpretations, or any shade in between. This is not the same as incorporating *bona fide* physical delays, e.g., those that can give rise to EEG wave propagation, in local circuits as emphasized in SMNI, or in long-ranged circuits [46].

The possibility of rewriting any theory or model of neural systems, which can be described by multiplicative Gaussian-Markovian dynamics, into an algebraic form invariant under Riemannian-geometric transformations, does not require that neural systems develop or elect mechanisms to take advantage of these transformations. However, the most obvious candidate for a physical consequence of invariance under such transformations is the information $\Upsilon$, developed in SMNI, sometimes loosely referred to as the "entropy" of the system. The invariance of $\Upsilon$ implies that, although different cortical regions may have different anatomical features and superficially appear to have quite different sets of firing states, they may indeed share, encode, or decode the same information using their own specific anatomy and physiology to develop their own sets of firing states, related to each other by (nonlinear) transformations.

This possibility is in the original spirit of other authors, who were attracted to the use of Riemannian geometry to explain how information in sensory regions might be transmitted to motor regions, albeit that their neural properties differ in many respects. SMNI develops columnar interactions, and here too it has been tempting to conjecture that local and global processing of columnar firing patterns is enhanced, if not primarily effected, by transmitting blocks of information that are invariant under nonlinear transformations of firing states.

Ultimately, these issues must be decided by experiment. There is presently no evidence, pro or con, to bear on the issue of the explicit Riemannian-geometric nature of information processing of neural firings. In principle, this could be accomplished by numerically fitting neuronal firing data to Lagrangians describing regions behaviorally proven to be processing similar information, similar to fits to data proposed for other artificial intelligence systems [127].

*5. Information, potential, and long-ranged interactions*

There have been attempts to use information as an index of EEG activity [128,129]. However, these attempts have focused on the concept of "mutual information" to find correlations of EEG activity under different electrodes. The SMNI approach at the outset recognizes that, for most brain states of late latency, at least a subset of regions being measured by several electrodes is indeed to be considered as one system, and their interactions are to be explicated by mathematical or physical modeling of the underlying neuronal processes. Then, it is not relevant to compare joint distributions over a set of electrodes with marginal distributions over individual electrodes. The concept of information, as expressed below, may yet prove to be a useful valid measure to compare different subjects within certain categories.

With reference to a steady state $\bar{P}(\tilde{M})$, when it exists, an analytic definition of the information gain $\hat{\Upsilon}$ in state $\tilde{P}(\tilde{M})$ is defined by [1,130]

$$\hat{\Upsilon}[\tilde{P}] = \int \cdots \int \underset{\sim}{D}\tilde{M}\ \tilde{P} \ln(\tilde{P}/\bar{P}), \tag{79}$$

where again a path integral is defined such that all intermediate-time values of $\tilde{M}$ appearing in the folded short-time distributions $\tilde{P}$ are integrated over. This is quite general for any system that can be described as Gaussian-Markovian [66], even if only in the short-time limit, e.g., the SMNI theory. (As time evolves, the distribution likely no longer behaves in a Gaussian manner, and the apparent simplicity of the short-time distribution must be supplanted by numerical calculations.) Although $\hat{\Upsilon}$ is well defined and useful for discussing macroscopic neocortical activity, it may not be as useful for all applications. Certainly many important local changes of information effected by the neocortical system are a function of the microscopic degrees of freedom already averaged over for the purposes of this study. However, it should also be noted that the path integral represents an enormous number of spatial-temporal degrees of freedom of the mesoscopic system. For example, even neglecting specific coding of presynaptic and postsynaptic membranes, detailed neuronal circuitry, and the dynamics of temporal evolution, in a hypothetical region of $10^9$ neurons with $10^{13}$ synapses: considering each synapse as only conducting or not conducting, there are $\approx \exp(7 \times 10^{12})$ possible synaptic combinations; considering only each neuron as firing or not firing, there are $\approx \exp(7 \times 10^8)$ neuronal combinations; considering only each mesocolumn as having integral firings between $-100$ and 100, there are $\approx \exp(5 \times 10^7)$ mesocolumnar combinations.

The minimization of $\hat{\Upsilon}$ with respect to $\tilde{P}$, with $\tilde{M}$ constrained to its (possibly multivalued) mean trajectory $< \tilde{M} >$, is formalized by the use of Lagrange parameters $J_G$. This results in the Legendre transform of $\ln \tilde{P}$ [131], and is equivalent to the generating functional $\Upsilon$ defined in the presence of extrinsic sources $J_G$ [121,130]. In the neocortex, some of these $J_G$ sources were used to specify firing constraints imposed on a given region of mesocolumns from long-ranged extrinsic or inter-regional afferents, e.g. from ipsilateral association, contralateral commissural, and thalamocortical processes:

$$\hat{\Upsilon}\mathbf{[}\min \ \{ \ \tilde{P}[< \tilde{M} >_J] \ \} \ \mathbf{]} = \Upsilon[< \tilde{M} >_J]$$

$$= - \ln \tilde{P}(J) + \int dt \int d^2 r < \tilde{M}^G >_J \ J_G \ , \tag{80}$$

where $\tilde{P}(J)$ is calculated by replacing $\underline{L}$ by $\underline{L}^J$:

$$\underline{L}^J = \underline{L} + M^G J_G/(2N\tau). \tag{81}$$

If $J$ is distributed, then $\underline{D}J$ measures effects on $\underline{L}^J$. A Hamiltonian representation, including $J_G \dot{M}^G$ interactions, is also readily derived [132]. This may be useful for describing long-ranged constraints that directly affect rates of change of columnar firings. As noted above, more recent work proposes that the long-ranged fibers be incorporated directly into the $F^G$ threshold factors, by including them initially in the microscopic derivation [23,24].

$\Upsilon_{J=0}$ is a proper potential, i.e., it possesses free-energy-like properties, having a true minimum about $(< \tilde{M} >_J - \ll \tilde{M} \gg)$, where $\ll \tilde{M} \gg$ is the extremum of $\underline{L}$ obtained by maximizing $\tilde{P}(\tilde{M})$ [133]. Its lowest-order approximate $\Upsilon^{(0)}$ gives the mean-field approximation:

$$\Upsilon^{(0)} = \tilde{\underline{S}}(< \tilde{M} >). \tag{82}$$

$\underline{T}$, defined as the ‘‘kinetic-energy’’ $dF$-independent part of $\underline{L}$, is scale independent of $N$. Therefore, the small scale of the neocortical system, about which the system fluctuates, is derived to be $N^{-1}$, the inverse of the number of neurons in a mesocolumn. This is interpreted as the effective ‘‘temperature’’ or inherent noise of the system. Thus, STM defines a rather ‘‘hot’’ and volatile system, wherein the relevant activity takes place on the same order as $N^{-1}$. By contrast, some long-term-memory calculations described below [15] are consistent with the interpretation of transpiring at a much lower temperature taking place in a locally more stable environment.

### E. Relevance of chaos

Given the context of current studies in complex nonlinear systems [134], the question can be asked: What if EEG has chaotic mechanisms that overshadow the above stochastic considerations? The real issue is whether the scatter in data can be distinguished between being due to noise or chaos. In this regard, we note that several studies have been proposed with regard to comparing chaos to simple filtered (colored) noise [134,135]. Since we have previously derived the existence of multiplicative noise in neocortical interactions, then the previous references must be generalized, such that we must investigate whether EEG scatter can be distinguished from multiplicative noise.

A simple, coarse criteria used to determine whether chaos occurs in a dynamical system is to examine the nature of propagation of uncertainty, i.e., the variance. As an example of applying the above framework, in our analysis of military exercise data [36], we were able to fit the short-time attrition epochs (determined to be about 5 minutes from mesh considerations determined by the nature of the Lagrangian) with short-time nonlinear Gaussian-Markovian probability distributions with a resolution comparable to the spread in data. When we performed the long-time path-integral numerical calculations from some point (spread) at the beginning of the battle, we found that we could readily find a form of the Lagrangian that made physical sense and that also fit the multivariate variances as well as the means at each point in time of the rest of the exercise interval. I.e., there was not any degree of hypersensitivity to initial conditions that prevented us from ‘‘predicting’’ the long-time means and variances of the system. Of course, since the system is dissipative, there is a strong tendency for all moments to diminish in time, but in fact this exercise was of sufficiently modest duration (typically 1 to 2 h) that variances did increase somewhat during the middle of the battle.

In summary, this battalion-regiment scale of battle does not seem to possess chaos. Similar to serious work undertaken in several fields [135,136], here too, the impulse to identify ‘‘chaos’’ in a complex system has been premature. It is not supported by the facts, tentative as they are because of sparse data. Of course, some other combat conditions might show some elements of chaos in some spatial-temporal domain, and then the resolution of the analysis would determine the influence of that chaos in that scenario.

Similar caution should be exercised regarding neocortical interactions. A more purposeful project is to compare stochastic with deterministic models of data. Today much attention is turning to the use of deterministic chaotic models for short-time predictions of systems. For example, if only short-time predictions are required, and if a deterministic chaotic model could well describe stochastic data within these epochs, then this model might be more computationally efficient instead of a more "correct" stochastic model, which would be necessary for long-time predictions. The scales of time involved are of course system dependent, and the deterministic chaotic modeling of data is still in its infancy [137].

Similarly, the above SMNI-derived distributions can be used to help determine if chaos is a viable mechanism in EEG. While these studies are concerned with neocortical interactions, it is of interest to note a series of experimental and theoretical studies of nonlinear dynamics of the olfactory bulb in small mammals, in which distinctive EEG patterns on the bulb surface are shown to be associated with specific odors [138-140]. These studies demonstrating chaos are very much model dependent, and as such it is only fair to present the models as possessing chaos, not necessarily the actual physical system.

For example, it has been widely noted that the correlation dimension of data is difficult to calculate; perhaps it is often not even a well-founded concept, e.g., since the EEG of event-related potentials is likely nonstationary and very much context and subject dependent [141]. Its calculation, e.g., using the popular Grassberger-Procaccia algorithm [142], even when supplemented with finer statistical tests [143] and noise reduction techniques [144], may prove fruitful, but likely only as a sensitivity index relative to shifting contexts and complementary to other models of EEG data.

## IV. VERIFICATION OF SMNI

### A. Short-term memory

#### *1. General description*

The most detailed and dramatic application of the theory outlined here is to predict stochastic bounds for the phenomena of human STM capacity during focused selective attention [16,18,145-147], transpiring on the order of tenths of a second to seconds, limited to the retention of $7 \pm 2$ items [148]. This is true even for apparently exceptional memory performers who, while they may be capable of more efficient encoding and retrieval of STM, and while they may be more efficient in "chunking" larger patterns of information into single items, nevertheless they also are limited to a STM capacity of $7 \pm 2$ items [149]. This "rule" is verified for acoustical STM, but for visual or semantic STM, which typically require longer times for rehearsal in an hypothesized articulatory loop of individual items, STM capacity appears to be limited to two to four [150]. This STM capacity-limited chunking phenomenon also has been noted with items requiring varying depths and breadths of processing [145-147,151,152]. Another interesting phenomenon of STM capacity explained by this theory is the primacy versus recency effect in STM serial processing, wherein first-learned items are recalled most error-free, with last-learned items still more error-free than those in the middle [153].

The basic assumption being made is that a pattern of neuronal firing that persists for many $\tau$ cycles is a candidate to store the "memory" of activity that gave rise to this pattern. If several firing patterns can simultaneously exist, then there is the capability of storing several memories. The short-time probability distribution derived for the neocortex is the primary tool to seek such firing patterns. Since this distribution is exponentially sensitive to (minus) the Lagrangian function $\underline{L}$, it is more convenient to deal directly with $\underline{L}$, whereby its minima specify the most likely states that can be sustained at a given time. Then, several important features of these patterned states can be investigated, as is done for other physical systems [1], e.g., the evolution of these states, the "time of first passage" to jump from one state to another state, hysteresis between states that have different depths (values of the Lagrangian at these local minima), the stability of each state under external forces, etc.

We define the "stationary" (sometimes referred to as the "uniform") Lagrangian, $\bar{\underline{L}}$, by setting to zero all temporal and spatial derivatives of $M^G$, e.g., as appearing in Eq. (61). Contour plots of the stationary Lagrangian, $\bar{\underline{L}}$, for typical synaptic parameters balanced between predominately inhibitory and predominately excitatory firing states are examined at many scales when the background synaptic noise is only modestly shifted to cause both efferent and afferent mesocolumnar firing states to have a common most-probable firing, centered at $M^{*G} = M^G = 0$ [16]. Within the range of synaptic parameters considered, for values of $\tau\bar{\underline{L}} \sim 10^{-2}$, this "centering" mechanism causes the appearance of from 5 to 10 or 11 extrema for values of $\tau\bar{\underline{L}}$ on the order of $\sim 10^{-2}$. The centering mechanism is achieved by modestly shifting $B^G_{G'}$ to cause $(V^G - a^{|G|}_{G'} v^{|G|}_{G'} N^{G'})$ to go to zero, thereby driving the threshold factor $F^G$ to zero. (Note that at $F^G = 0$, the mesoscopic derivation of Gaussian $\underline{L}$ breaks down, so that we can only consider a finite region, heavily weighted by $N$, about this point.) In the absence of external constraints and this centering mechanism, no stable minima are found; i.e., the system either shuts down, with no firings, or it becomes epileptic, with maximal firings at the upper limits of excitatory or of excitatory and inhibitory firings. The appearance of these extrema due to the centering mechanism is clearly dependent on the non-linearities present in the derived Lagrangian, stressing competition and cooperation among excitatory and inhibitory interactions at columnar as well as at neuronal scales. (See Fig. 4.)

---

Figure 4.

---

It must be stressed that these numbers of minima are determined when the resolution of the contours is commensurate with the resolution of columnar firings, i.e., on the order of five to ten neuronal firings per columnar mesh point. If the resolution is forced to go down to one neuronal firing per columnar mesh point, then typically only about half these minima are found. The coarser resolution, in fact, is the one appropriate for numerical solution of the derived time-dependent path integral: Most important contributions to the probability distribution $P$ come from ranges of the time slice $\theta$ and the "action" $N\underline{L}$, such that $\theta N \underline{L} \le 1$. By considering the contributions to the first and second moments of $\Delta M^G$ for small time slices $\theta$, conditions on the time and variable meshes can be derived [123]. The time slice is determined by $\theta \le (N\bar{\underline{L}})^{-1}$ throughout the ranges of $M^G$ giving the most important contributions to the probability distribution $P$. The variable mesh, a function of $M^G$, is optimally chosen such that $\Delta M^G$ is measured by the covariance $g^{GG'}$ (diagonal in the neocortex due to the independence of $E$ and $I$ chemical interactions) or $\Delta M^G \sim (g^{GG}\theta)^{1/2}$ in the notation of the SMNI papers. For $N \sim 10^2$ and $\bar{\underline{L}} \sim 10^{-2}/\tau$, it is reasonable to pick $\theta \sim \tau$. Then it is calculated that optimal meshes are $\Delta M^E \sim 7$ and $\Delta M^I \sim 4$, essentially the resolutions used in the coarse contour plots.

Since the extrema of the Lagrangian appear to lie fairly well along a line in the two-dimensional $M^G$ space, and since coefficients of slowly varying $dM^G/dt$ terms in the nonstationary $\underline{L}$ are noted to be small perturbations on $\bar{\underline{L}}$ [15], a solution to the stationary probability distribution was hypothesized to be proportional to $\exp(-\Phi/D)$, where $\Phi = CN^2\bar{\underline{L}}$, the diffusion $D = N/\tau$, and $C$ is a constant. Surprisingly, at least until more recent research has shown the generality of such results [154], along the line of the extrema, for $C \approx 1$, this is determined to be an accurate solution to the full two-dimensional Fokker-Planck equation [18]. A weak-noise high-barrier regime defined by $\Delta\Phi/D > 1$, where $\Delta\Phi$ is the difference in $\Phi$ from minima to maxima, can be assumed for further analyses [155]. This is extremely useful, as a linear stability analysis shows that stability with respect to mesocolumnar fluctuations induced by several neurons changing their firings is determined by the second derivatives of $-\Phi$ [156]; here this just measures the parabolic curvature of $\bar{\underline{L}}$ at the extrema. Thus, all the extrema of the stationary Lagrangian are determined to be stable minima of the time-dependent dynamic system. Note however, that it is unlikely that a true potential exists over all $M^G$ space [122,157].

This stationary solution is also useful for calculating the time of first passage, $t_{vp}$, to fluctuate out of a valley in one minimum over a peak to another minimum [156]. It turns out that the values of $\tau\bar{\underline{L}} \sim 10^{-2}$ for which the minima exist are just right to give $t_{vp}$ on the order of tenths a second for about nine of the minima when the maximum of 10 to 11 are present. The other minima give $t_{vp}$ on the order of many seconds, which is large enough to cause hysteresis to dominate single jumps between other minima [16]. Thus, $7 \pm 2$ is the capacity of STM, for memories or new patterns that can be accessed in any order during tenths of a second, all as observed experimentally [149]. When the number of neurons per minicolumn is taken to be ~220, modeling the visual neocortex [16], then the minima become deeper and sharper, consistent with sharper depth of processing, but several minima become isolated from the main group. This effect might be responsible for the lowering of STM capacity for visual processing, as mentioned above. I.e., the statistical time of passage between clusters becomes many hours, longer than STM, while the time between minima within a cluster, now with only 2 to 4 minima per cluster, is on the order of tenths of a second, as observed. This effect also serves to illustrate that the "practical" number of emergent mesoscopic stable states does not necessarily increase with an increasing number of microscopic units. (See Fig. 5.)

Figure 5.

This estimate of the number of minima involves a very sensitive calculation. That is, if $N$ were a factor of 10 larger, or if $\tau \overline{\underline{L}} \sim 0.1$ at the minima, then $t_{vp}$ is on the order of hours instead of seconds, becoming unrealistic for STM durations. Alternatively, if $t_{vp}$ were much smaller, i.e., less than $\sim 5\tau$, this case would be inconsistent with observed time scales necessary for formation of any memory trace [158]. In this context, it is noted that the threshold factor of the probability distribution scales as $(N^* N)^{1/2}$, demanding that both the macrocolumnar divergence and minicolumnar convergence of mesocolumnar firings be tested by these calculations.

These results pose serious problems for other models, such as ‘‘mean-field’’ theories or reductionist doctrines. The mean-field approach essentially sets $N = 1$, and $N^*$ is effectively taken by some investigators to be $\sim 10^5$, the size of a macrocolumn, but others even consider it to be as large as $\sim 10^{10}$, the total number of neurons in the neocortex. The reductionist doctrine claims that only circuitries among a few to several neurons are responsible for a specific pattern of neocortical function, and this effectively sets $N \approx N^*$, on the order of a few neurons. It is hard to understand how both the capacity and duration of STM can be explained by these other models, even assuming they were or could be derived with realistic synaptic interactions and correct statistical dynamics.

The statistical nature of this storage and processing also explains the primacy versus recency effect: The deepest minima of the Lagrangian are more likely accessed than the others of this probability distribution, and these valleys are sharper than the others. I.e., they are more readily accessed and sustain their patterns against fluctuations more accurately than the relatively more shallow minima. The more recent memories or newer patterns may be presumed to be those having synaptic parameters more recently tuned and/or more actively rehearsed. Thus, both the nonlinearities and the statistical nature of this theory are tested by STM capacity. These insights have helped to correct the notions of some experimentalists who claimed they could not find this effect in the visual cortex: Their experimental paradigms were testing the visual cortex using rules of auditory capacity ( $7 \pm 2$), and therefore they were washing out this effect.

These calculations give experimental support to the derivation of the mesoscopic probability distribution, yielding similar algebraic structures of the threshold factors in Eqs. (32) and (63), illustrating common forms of interactions between their entities, i.e., neurons and columns of neurons, respectively. The nonlinear threshold factors are defined in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges.

*2. STM calculation*

Three cases of neuronal firings were considered [16]. Since STM duration is still long relative to $\tau$, stationary solutions of $\overline{\underline{L}}$, derived from $\underline{L}$ in Eq. (61), were investigated to determine how many stable minima, $\ll \bar{M}^G \gg$, may simultaneously exist within this duration. Also, individual mesocolumns were studied. I.e., take the uniform limit of $\dot{\bar{M}}^G = 0 = \nabla \bar{M}^G$. Although the $\dot{\bar{M}}^G = 0$ limit should only be taken for the midpoint-discretized Lagrangian $\underline{L}_F$, this is a small difference here [16].

A model of dominant inhibition describes how minicolumnar firings are suppressed by their neighboring minicolumns. For example, this could be effected by developing nearest-neighbor (NN) mesocolumnar interactions [15], but here the averaged effect is established by inhibitory mesocolumns (IC) by setting $A^I_E = A^E_I = 2A^E_E = 0.01 N^*/N$. Since there appears to be relatively little $I - I$ connectivity, set $A^I_I = 0.0001 N^*/N$. The background synaptic noise is taken to be $B^E_I = B^I_E = 2B^E_E = 10 B^I_I = 0.002 N^*/N$. As minicolumns are observed to have ~110 neurons (the visual cortex appears to have approximately twice this density) [59], and as there appear to be a predominance of $E$ over $I$ neurons [46], here take $N^E = 80$ and $N^I = 30$. Use $N^*/N = 10^3$, $J_G = 0$ (absence of long-ranged interactions), and $V^G$, $v^G_{G'}$, and $\phi^G_{G'}$ as estimated previously, i.e., $V^G = 10$ mV, $|v^G_{G'}| = 0.1$ mV, $\phi^G_{G'} = 0.1$ mV. The ‘‘threshold factors’’ $F^G_{\rm IC}$ for this IC model are then

$$F^E_{\rm IC} = \frac{0.5\bar{M}^I - 0.25\bar{M}^E + 3.0}{\pi^{1/2}(0.1\bar{M}^I + 0.05\bar{M}^E + 9.80)^{1/2}} \; ,$$

$$F^I_{\rm IC} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E - 45.8}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 11.2)^{1/2}} \ . \tag{83}$$

In the prepoint-discretized deterministic limit, the threshold factors determine when and how smoothly the "step functions" $\tanh F^G_{\rm IC}$ in $g^G(t)$ change $M^G(t)$ to $M^G(t+\theta)$. $F^I_{\rm IC}$ will cause afferent $\bar{M}^I$ to fire for most of its values, as $\bar{M}^I \sim -N^I \tanh F^I_{\rm IC}$ will be positive for most values of $\bar{M}^G$ in $F^I_{\rm IC}$, which is already weighted heavily with a term -45.8. Looking at $F^E_{\rm IC}$, it is seen that the relatively high positive values of efferent $\bar{M}^I$ require at least moderate values of positive efferent $\bar{M}^E$ to cause firings of afferent $\bar{M}^E$.

The calculations presented here support the contention that the neocortex functions at multiple hierarchies. While specific long-term memory (LTM) information is most likely coded at the microscopic neuronal level, the mesoscopic scale most likely provides the context for multiple most-probable firing patterns which process STM and which facilitate plastic synaptic encoding of LTM [15]. E.g., $\tau \underline{\bar{L}}$ can range from 0 to values greater than $10^3$ [14,15]. However, realistic constraints on STM duration dictate that only values of $\tau \underline{\bar{L}} \le 0.04$ are of interest here. Detailed mesoscalar calculations demonstrate that only this range exhibits sufficient nonlinear structure to support STM phenomena.

It is discovered that more minima of $\underline{\bar{L}}$ are created, or "restored," if the numerator of $F^G$ contains terms only in $\bar{M}^G$, tending to center $\underline{\bar{L}}$ about $\bar{M}^G = 0$. Of course, any mechanism producing more as well as deeper minima is statistically favored. However, this particular "centering" mechanism has plausible support: $M^G(t+\tau) = 0$ is the state of afferent firing with highest statistical weight. I.e., there are more combinations of neuronal firings, $\sigma_j = \pm 1$, yielding this state than any other $M^G(t+\tau)$; e.g., $\sim 2^{N^G+1/2}(\pi N^G)^{-1/2}$ relative to the states $M^G = \pm N^G$. Similarly, $M^{*G}(t)$ is the state of efferent firing with highest statistical weight. Therefore, it is natural to explore mechanisms that favor common highly weighted efferent and afferent firings in ranges consistent with favorable firing threshold factors $F^G \approx 0$. Another effect of this centering mechanism apparently is to shift minima of $\underline{\bar{L}}^G$ closer together, permitting them to often cooperate instead of compete.

Detailed calculations demonstrate that either $\underline{\bar{L}}^E$ or $\underline{\bar{L}}^I$ separately typically give rise to more multiple minima, $\approx 10$, than permitted by their sum $\underline{\bar{L}}$ at this resolution. This "loss" of minima apparently is an interesting consequence of $E - I$ competition at the mesoscopic scale. On one hand, since $\underline{\bar{L}}^G$ scales as $N^G/N$ for relatively large $\bar{M}^G$, $\underline{\bar{L}}^E$ dominates due to the larger $\bar{M}^E$ in its mean $g^E$. On the other hand, for relatively small $\bar{M}^G$, $g^G$ typically is small if there are several multiple minima in $\underline{\bar{L}}^G$, since most of the minima are found to cluster about the origin. Therefore, $\underline{\bar{L}}^G$ scales as $(N^G)^{-1}$ from the variances $(g^{GG})^{-1}$, and $\underline{\bar{L}}^I$ dominates for small $\bar{M}^G$.

The centering effect of the IC model of dominant inhibition, labeled here as the IC′ model, is quite easy for the neocortex to accommodate. For example, this can be accomplished simply by readjusting the synaptic background noise from $B^G_E$ to $B'^G_E$,

$$B'^G_E = \frac{V^G - (\frac{1}{2} A^G_I + B^G_I)v^G_I N^I - \frac{1}{2} A^G_E v^G_E N^E}{v^G_E N^G} \tag{84}$$

for both $G = E$ and $G = I$. This is modified straightforwardly when regional influences from $M^{\ddagger E}$ are included, as in in Eq. (63). In general, $B^G_E$ and $B^G_I$ (and possibly $A^G_E$ and $A^G_I$ due to actions of neuromodulators, and $J_G$ or $M^{\ddagger E}$ constraints from long-ranged fibers) are available to force the constant in the numerator to zero, giving an extra degree(s) of freedom to this mechanism. (If $B'^G_E$ would be negative, this leads to unphysical results in the square-root denominator of $F^G$. Here, in all examples where this occurs, it is possible to instead find positive $B'^G_I$ to appropriately shift the numerator of $F^G$.) In this context, it is experimentally observed that the synaptic sensitivity of neurons engaged in selective attention is altered, presumably by the influence of chemical neuromodulators on postsynaptic neurons [159].

By this centering mechanism, $B'^E_E = 1.38$ and $B'^I_I = 15.3$, and $F^G_{\rm IC}$ is transformed to $F^G_{\rm IC'}$,

$$F^E_{\mathrm{IC}'} = \frac{0.5\bar{M}^I - 0.25\bar{M}^E}{\pi^{1/2}(0.1\bar{M}^I + 0.05\bar{M}^E + 10.4)^{1/2}} ,$$

$$F^I_{\mathrm{IC}'} = \frac{0.005\bar{M}^I - 0.5\bar{M}^E}{\pi^{1/2}(0.001\bar{M}^I + 0.1\bar{M}^E + 20.4)^{1/2}} . \tag{85}$$

Note that, aside from the enforced vanishing of the constant terms in the numerators of $F^G_{\mathrm{IC}'}$, the only other change in $F^G_{\mathrm{IC}'}$ relative to $F^G_{\mathrm{IC}}$ is to moderately affect the constant terms in the denominators. This increases the number of minima of $\tau\underline{L}_{\mathrm{IC}'}$ to 4. The two minima clustered close to the origin are no longer discernible for $\tau\underline{L}_{\mathrm{IC}'} > 0.03$.

The other ''extreme'' of normal neocortical firings is a model of dominant excitation, effected by establishing excitatory mesocolumns (EC) by using the same parameters $\{ B^G_{G'}, v^G_{G'}, \phi^G_{G'}, A^I_I \}$ as in the IC model, but setting $A^E_E = 2A^I_E = 2A^E_I = 0.01N^*/N$. This yields

$$F^E_{\mathrm{EC}} = \frac{0.25\bar{M}^I - 0.5\bar{M}^E - 24.5}{\pi^{1/2}(0.05\bar{M}^I + 0.10\bar{M}^E + 12.3)^{1/2}} ,$$

$$F^I_{\mathrm{EC}} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.05\bar{M}^E + 7.24)^{1/2}} . \tag{86}$$

The negative constant in the numerator of $F^I_{\mathrm{EC}}$ inhibits afferent $\bar{M}^I$ firings. Although there is also a negative constant in the numerator of $F^E_{\mathrm{EC}}$, the increased coefficient of $\bar{M}^E$ (relative to its corresponding value in $F^E_{\mathrm{IC}}$), and the fact that $\bar{M}^E$ can range up to $N^E = 80$, readily permits excitatory firings throughout most of the range of $\bar{M}^E$. This permits three minima.

Applying the centering mechanism to EC, $B'^E_I = 10.2$ and $B'^I_I = 8.62$. The net effect in $F^G_{\mathrm{EC}'}$, in addition to removing the constant terms in the numerators of $F^G_{\mathrm{EC}}$, is to change the constant terms in the denominators: 12.3 in $F^E_{\mathrm{EC}}$ is changed to 17.2 in $F^E_{\mathrm{EC}'}$, and 7.24 in $F^I_{\mathrm{EC}}$ is changed to 12.4 in $F^I_{\mathrm{EC}'}$. Now six prominent minima are possible along a line through $\bar{M}^G = 0$, and two others are at $\bar{M}^G = \pm N^G$. Each pair of minima above and below the $\bar{M}^I = 0$ axis merge into single minima for $\tau\underline{L}_{\mathrm{EC}'} > 0.02$, and these lose resolution for $\tau\underline{L}_{\mathrm{EC}'} > 0.03$.

Now it is natural to examine a balanced case intermediate between IC and EC, labeled BC. This is accomplished by changing $A^E_E = A^I_E = A^E_I = 0.005N^*/N$. This yields

$$F^E_{\mathrm{BC}} = \frac{0.25\bar{M}^I - 0.25\bar{M}^E - 4.50}{\pi^{1/2}(0.050\bar{M}^E + 0.050\bar{M}^I + 8.30)^{1/2}} ,$$

$$F^I_{\mathrm{BC}} = \frac{0.005\bar{M}^I - 0.25\bar{M}^E - 25.8}{\pi^{1/2}(0.001\bar{M}^I + 0.050\bar{M}^E + 7.24)^{1/2}} . \tag{87}$$

Three minima are possible, on the boundaries of $\bar{M}^G$ space.

Applying the centering mechanism to BC, $B'^E_E = 0.438$ and $B'^I_I = 8.62$. The net effect in $F^G_{\mathrm{BC}'}$, in addition to removing the constant terms in the numerators of $F^G_{\mathrm{BC}}$, is to change the constant terms in the denominators: 8.30 in $F^E_{\mathrm{BC}}$ is changed to 7.40 in $F^E_{\mathrm{BC}'}$, and 7.24 in $F^I_{\mathrm{BC}}$ is changed to 12.4 in $F^I_{\mathrm{BC}'}$. Now ten minima are possible. The nine minima along the diagonal line lose resolution for $\tau\underline{L}_{\mathrm{BC}'} > 0.01$ above $\bar{M}^I = 0$ and for $\tau\underline{L}_{\mathrm{BC}'} > 0.02$ below $\bar{M}^I = 0$.

The effects of using the full Feynman Lagrangian $\underline{L}_F$ were considered, including all the Riemannian and other nonlinear corrections, discussed below. The net effect is to slightly raise the threshold at which minima dissipate, to about $\tau\underline{L}_{\mathrm{BC}'} \geq 0.03$, which is relevant for the duration of STM, discussed subsequently. However, note that the minima structure is essentially the same. (See Fig. 4.)

To demonstrate that multiple minima are not an effect of nonlinearities induced by the denominators of $F^G$, the net effect in $\underline{L}_{\mathrm{BC}'}$ by dropping the $\bar{M}^G$ terms in the denominators of $F^G_{\mathrm{BC}'}$ is such that the valleys of minima are only slightly increased. However, these denominators are still important contributions derived from synaptic interactions. E.g., even with the $\bar{M}^G$ terms dropped, the denominators contribute factors of ~1/5 to $F^G_{\mathrm{BC}'}$.

If $N^*$ is scaled larger or smaller, this effectively scales $A^G_{G'} = A^{*G}_{G'} N^*/N$ and $B^G_{G'} = B^{*G}_{G'} N^*/N$, disturbing the relatively sensitive balance that permits a few percent of efferent firings to affect their afferents. Then, the number of possible minima is typically reduced to one or two. If $N$ is scaled larger or smaller, the number of minima is altered and the duration of STM is affected, as discussed subsequently. However, for $N$ still in the range of a few hundred, the number of possible minima is not severely reduced. The case $N = 220$, e.g., the visual cortex was considered: For model BC′, the number of prominent minima found is 11, but they form clusters, with higher peaks between clusters than between minima within a cluster. Note that the larger $N$ sharpens the minima and therefore the resolution of visual information processing. (See Fig. 5.)

Note that the sharpness of the tanh $F^G$ step-function contribution to the mean firing is sensitive to a factor of $N^{\frac{1}{2}}$ in $F^G$. Additionally, the strength of coupling between mesocolumns scales as $N^{3/2}$. Thus the neuronal size of mesocolumns directly affects the breadth and depth of the information processing capability of the neocortex. It is interesting to note that the human visual cortex, which may be assumed to require the finest tuning in the neocortex, is unique in having twice the number of neurons per minicolumn than other regions of the neocortex [59].

These results are unchanged qualitatively for modest changes of any neocortical parameters. However, it is reasonable to conjecture that more drastic abnormal changes in the neocortical parameters might severely reduce the number of minima. This conjecture is based on calculations wherein $F^G$ do not possess the relatively sensitive balances allowing a few percent of efferent neurons to control firings in their afferent neurons. In calculations using these unrealistic or abnormal parameters only one or two minima survive.

*3. STM stability and duration*

The calculation of stability and time of duration in most likely states of firing starts by using the differential-equation Hamiltonian formulation of the path-integral Lagrangian, called the Fokker-Planck equation. For future reference, when EEG's are discussed below in the context of considering a given local minimum, note that the time-dependent differential macroscopic probability distribution $\tilde{P} = \prod_r P$, or "propagator," is found to satisfy a true Fokker-Planck equation, but one with nonlinear drifts and diffusions in the space of $E$ and $I$ firings. The Fokker-Planck equation for the region $\Omega$ is

$$\frac{\partial \tilde{P}}{\partial t} \approx \Omega^{-1} \int d^2 r [\frac{1}{2} (g^{GG'} \tilde{P})_{,GG'} - (g^G \tilde{P})_{,G} + N \underline{V}' \tilde{P}] ,$$

$$(\cdots)_{,G} \equiv \partial(\cdots)/\partial M^G . \qquad (88)$$

The true Fokker-Planck equation is actually more general, e.g., if long-ranged spatial structures are included, where the independent variables $M^G$ are fields which themselves may depend on space and time coordinates. The above equation is derived in the nearest-neighbor approximation from the general equation using functional derivatives [16],

$$\partial(\cdots)/\partial M^G \rightarrow \delta(\cdots)/\delta M^G ,$$

$$\delta(\cdots)/\delta M^G = (\cdots)_{,G} - \nabla_i (\cdots)_{,\nabla_i G} + \nabla_i^2 (\cdots)_{,\nabla_i^2 G} , \qquad (89)$$

where we have used the compacted notation introduced previously [16].

An estimate of a stationary solution $P_{\mathrm{stat}}$ to the Fokker-Planck differential equation for the probability distribution $P$ of $M^G$ firings for an uncoupled mesocolumn, i.e., $\underline{V}' = 0$, is given by the stationary limit of the short-time propagator,

$$P_{\mathrm{stat}} \approx N_{\mathrm{stat}} g^{1/2} \exp(-CN\tau \overline{\underline{L}}) ,$$

$$g = \det(g^{GG'})^{-1} \equiv \det(g_{GG'}) = g_{EE} g_{II} \ , \tag{90}$$

where $N_{\rm stat}$ and $C$ are constant factors. An estimate of the approximation made is estimated by seeking values of constants $C$, such that the stationary Fokker-Planck equation is satisfied exactly. Contour plots of $C$ versus $\bar{M}^G$ demonstrate that there exists real positive $C$ which may only range from ~ $10^{-1}$ to ~1, for which there exists unbroken contours of $C$ which pass through or at least border the line of minima [18]. At each point $\bar{M}^G$, this leaves a quadratic equation for $C$ to be solved. Dropping the $g^{1/2}$ factor results in $C$ not being real throughout the domain of $\bar{M}^G$.

Thus we have defined a solution with potential $N^2 \underline{\bar{L}} = \int A \, dM$, drift $A$, and diffusion $N/\tau$. Stability of transient solutions, defined for $\underline{\delta} M^G$ about a stationary state by

$$\underline{\delta} \dot{M}^G \approx -A_{,G} \underline{\delta} M^G = -N^2 \underline{\bar{L}}_{,GG} \underline{\delta} M^G \ , \tag{91}$$

is therefore equivalent to $\ll \bar{M} \gg$ being a minimum of $\underline{\bar{L}}$.

Since the minima of the Lagrangian lie deep in a valley along a line, the time for first passage, $t_{vp}$, is estimated in analogy to a one-dimensional system as [156]

$$t_{vp} \approx \pi N^{-2} [|\underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_p)| \underline{\bar{L}}_{,GG'}(\ll \bar{M} \gg_v)]^{-1/2}$$

$$\times \exp \{ CN\tau [\underline{\bar{L}}(\ll \bar{M} \gg_p) - \underline{\bar{L}}(\ll \bar{M} \gg_v)] \} \ , \tag{92}$$

where $\ll \bar{M} \gg_v$ is the minimum at the valley of $\underline{\bar{L}}$ in question, and $\ll \bar{M} \gg_p$ is the maximum at a peak separating two minima. These equations are reasonable but crude estimates, and future numerical work must be done to detail the extent of their validity.

The exponential factor can be quite large in some instances, and quite small in others. As noted previously [15], differences in $\underline{\bar{L}}$ from valleys to peaks are still large relative to the Riemannian correction terms and relative to differential spatial-temporal contributions, thereby permitting this simpler analysis. However, values of $\tau \underline{\bar{L}}$ at maxima separating the far minima may be greater than 1, thereby yielding a very large $t_{vp}$, typical of many physical systems undergoing hysteresis [15]. Relaxation times $t_r$ about this stationary state are estimated by $|g^G_{,G}|^{-1}$ [156], and are on the order of $\tau$.

For changes $\Delta Z$ in synaptic parameters $Z = \{ A^*_{jk}, B^*_{jk}, V_j, v_{jk}, \phi_{jk}, N^{*G} \}$ that transpire within a $\Delta t$ of several tenths of a second to seconds, e.g., during typical attention spans, hysteresis is more probable than simple jumps between minima if the following inequalities are satisfied. These estimates necessarily require more details of the system in addition to $t_r$ and $t_{vp}$ [156]:

$$\left| \frac{t_r \, \Delta t \, N^2 \Delta \underline{\bar{L}}_{,G}}{\Delta Z} \right|^{-1} \frac{N}{2} \gg \frac{\Delta Z}{\Delta t} \gg \left| \frac{N \tau t_{vp} \, \Delta \underline{\bar{L}}}{\Delta Z} \right|^{-1} . \tag{93}$$

For $\Delta Z$ approximately corresponding to a significant increase in the synaptic efficacy of one neuron per minicolumn, this typically leads to

$$\frac{\Delta Z}{t_r} \gg \frac{\Delta Z}{\Delta t} \gg \frac{\Delta Z}{t_{vp}} \ , \tag{94}$$

where the last inequality may or may not hold, depending on the value of $t_{vp}$.

Therefore, it is possible for hysteresis to be highly more probable than simple jump behavior to another firing state. This provides a mechanism whereby an extended temporal firing pattern of information can be processed beyond the time scale of relaxation periods, e.g., reverberation among several local minima. It is to be expected that the effects of $J_G(r;t)$ on $\Delta Z(r;t)$ create more complex examples of spatial-temporal hysteresis. These sustaining mechanisms may serve to permit other biochemical processes to store information for longer time periods as stable synaptic modifications, e.g., LTM. As detailed previously [15], changes in synaptic parameters $\Delta Z$ may duplicate the effects of $J_G$, providing a mechanism whereby columnar firings encode long-range firing constraints. If this encoding of firing patterns can establish itself on short enough time scales, then columnar coding of long-range firings could be a precursor mechanism initiating the centering mechanism above, especially across large regions of the neocortex. Then, there would be a more uniform gradation of mechanism(s) establishing STM and LTM.

However, to address the issue of limited capacity of STM, it is reasonable to require that within time spans of tenths of a second to tens of seconds, simple jumps among minima are more probable than hysteresis. This permits all minima to be readily accessible during STM duration, in any ordering [149], at least more so than if hysteresis were more probable. In agreement with this empirical requirement, as detailed in the previous studies, it is found that $\tau[\underline{\bar{L}}(\ll \bar{M} \gg_p) - \underline{\bar{L}}(\ll \bar{M} \gg_v)]$~$0.01$—$0.03$ for these models using empirical values for synaptic parameters. Then for $|\tau \underline{\bar{L}}_{,GG'}|$~$10^{-3}$, $t_{vp}$~$10\tau$—$100\tau$, on the order of several tenths of a second to a second. Use of the full Feynman Lagrangian $\underline{\bar{L}}_F$ increases $t_{vp}$ slightly. For these relatively short $t_{vp}$ the second inequality above is violated, and simple jumps are more probable than hysteresis, as required for STM.

Under conditions of serial processing, the deeper valleys of $\underline{\bar{L}}$ representing the more likely firing states will be occupied first. In all cases considered here, some valleys are deeper than the others. This implies that the last several items in STM should be harder to encode (learn) and retain, with the possible exception of the last one or two items, which represent the most recent shifting of firing patterns $\bar{M}^G$ to these minima $\ll \bar{M} \gg_v$ of $\underline{\bar{L}}$. These conclusions are consistent with empirical observations, and are obtained independent of any other rehearsal mechanisms that may exist.

Calculations in these models establish that the prefactor most often is ~ $\tau$. However, points close to the corners $\bar{M}^G = \pm(N^E, N^I)$ have much more rapid variations. Therefore, minima at these corners, even when $\tau \underline{\bar{L}}(\ll \bar{M} \gg_p)$~$0.01$—$0.03$, because of their sharp peaks, typically have $t_{vp}$ on the order of tens of seconds to jump to minima clustered on the diagonal. This is within the range where hysteresis is more probable for these minima. Therefore, minima at the corners of $\bar{M}^G$ space most likely do not contribute to STM, bringing the number of available minima down to $7 \pm 2$ as empirically observed.

## B. EEG dispersion relations

### *1. General description*

Linear expansions about specific extrema, specified by the Euler-Lagrange variational equations, permit the development of stability analyses and dispersion relations in frequency-wave-number space [14,15,17]. Of course, such linear expansions are justified only after the nonlinear problem, e.g., such as that encountered for STM, is solved for locations of minima. It is noted in this regard that the corresponding wave propagation velocities pace interactions over several minicolumns, in order to be of magnitude sufficient to permit simultaneous information processing within ~ $10^{-1}$ sec with interactions mediated by long-ranged fibers possessing much greater propagation velocities ~600–900 cm/sec [17]. E.g., detailed auditory and visual processing can feed information to the association cortex where it can be processed simultaneously, possibly giving feedback to the primary sensory regions. The propagation velocities calculated by SMNI, ~1 cm/sec, also are consistent with observed movements of attention [160] and of hallucinations [161] across the visual field. This strongly suggests that nearest-neighbor mesocolumnar interactions as developed here are an important mechanism in these movements. These velocities scale strongly with the values of $\bar{M}^G$ minima, increasing with their distance from $\bar{M}^G$~0, the range of maximal firing combinations. This effect remains to be further investigated; the appropriate calculations should test the nearest-neighbor spatial dependence of the SMNI theory.

These mesoscopic dispersion relations also are consistent with global macroscopic dispersion relations derived and fitted to EEG data [46], yielding oscillatory solutions consistent with the $\alpha$ rhythm, i.e., $\omega \approx 10^2$ sec$^{-1}$, equivalent to $\nu = \omega/(2\pi) \approx 16$ cps (Hz). This suggests that these complementary local and global theories may be confluent, considered as a joint set of dispersion relations evolving from the deterministic limit of a joint Lagrangian, referred to as the "equations of motion," but linearly simplified in neighborhoods of minima of the stationary Lagrangian.

Other researchers have developed quite different approaches to investigating macroscopic neocortical activity, e.g., stressing that systematics of rhythmic EEG ($\alpha$ rhythm, sleep $\delta$, etc.) can be modeled by resonant modes of macroscopic dipole-layered firing patterns of the neocortex [46,48,49,52]. These resonances, in linearized coupled excitatory-inhibitory spatial-temporal integral equations describing dipole-layered sources, give rise to a macroscopic dispersion relation relating firing frequencies to spatial wave vectors, consistent with experimental observations. While many other investigators also accept dipole layers to model EEG activity, at least to the extent of recognizing activity perpendicular to laminae, they also demonstrate that there are respectable candidates for mechanisms that might fundamentally be responsible for macroscopic activity, other than those proposed here which detail synaptic dynamics of mesocolumnar interactions [11,162-167]. For example, given the present lack of experimental knowledge, it is possible to formulate macroscopic neocortical activity in terms of statistics of either membrane or synaptic microscopic neuronal activities, albeit that these two are obviously dependent on each other [168]. Therefore, the results of this statistical theory derived earlier [17] might be interpreted either as suggesting that mesocolumnar activity instigates macroscopic activity, or rather as suggesting that mesocolumnar activity strongly interacts with ongoing macroscopic activity that is instigated or sustained by other mechanisms.

The two approaches outlined above, i.e., local mesocolumnar versus global nonmesocolumnar, give rise to the important alternative conjectures suggested previously in this paper. Other studies also have proposed that EEG may be due to a combination of short-ranged and long-ranged interactions, which combine to form a single dispersion relation with multiple branches [44].

It is plausible that studies of the origin of rhythmic EEG will give direct insight into related mechanisms underlying evoked potentials. However, in contrast to the $\alpha$ rhythm and other gross EEG phenomena being gauges of general alertness to process information, the time-locked averaged evoked potentials appear to be a gauge of more selective attention to information being processed. Therefore, to derive a plausible picture of the nature of evoked potentials, it is more likely that more details of local interaction among columnar interactions must be included, such as those given below.

The first SMNI approach to scalp EEG assumed that the Euler-Lagrange variational limit of the stochastic Lagrangian was a suitable averaging procedure over masses of neurons contributing to this relatively coarse spatial phenomenon [17].

It should be noted that at this point in the development of our ‘bottom-up'' description we have overlapped with our initial ''top-down'' description as described in Sec. II, and therefore have provided a relatively first-principles approach to better understand these issues. We also show that most likely trajectories of the mesoscopic probability distribution, representing averages over columnar domains, give a description of the systematics of macroscopic EEG in accordance with experimental observations.

### *2. Euler-Lagrange variational equations*

This calculation begins by considering the Lagrangian $\underline{L}_F$, the Feynman midpoint-discretized Lagrangian. The Euler-Lagrange variational equation associated with $\underline{L}_F$ leads to a set of 12 coupled first-order differential equations, with coefficients nonlinear in $M^G$, in the 12 variables $\{ M^G, \dot{M}^G, \ddot{M}^G, \nabla M^G, \nabla^2 M^G \}$ in $(r;t)$ space. In the neighborhood of extrema $\ll \bar{M}^G \gg$, $\underline{L}_F$ can be expanded as a Ginzburg-Landau polynomial, i.e., in powers of $M^E$ and $M^I$. To investigate first-order linear oscillatory states, only powers up to 2 in each variable are kept, and from this the variational principle leads to a relatively simple set of coupled linear differential equations with constant coefficients:

$$
\begin{aligned}
0 &= \delta \underline{L}_F = \underline{L}_{F,\dot{G}:t} - \delta_G \underline{L}_F \\
&\approx - \underline{f}_{|G|} \ddot{\underline{M}}^{|G|} + \underline{f}^1_G \dot{\underline{M}}^{G\neg} - \underline{g}_{|G|} \nabla^2 \underline{M}^{|G|} + \underline{b}_{|G|} \underline{M}^{|G|} + \underline{b}\, \underline{M}^{G\neg} \ , \ G\neg \neq G \ , \\
&(\cdots)_{,\dot{G}:t} = (\cdots)_{,\dot{G}G'} \dot{M}^{G'} + (\cdots)_{,\dot{G}\dot{G}'} \ddot{M}^{G'} \ , \\
&\underline{M}^G = M^G - \ll \bar{M}^G \gg \ , \ \underline{f}^1_E = -\underline{f}^1_I \equiv \underline{f} \ .
\end{aligned} \tag{95}
$$

These equations are then Fourier transformed and the resulting dispersion relation is examined to determine for which values of the synaptic parameters and of the normalized wave-number $\underline{\xi}$, the conjugate variable to $r$, can oscillatory states, $\omega(\underline{\xi})$, persist. E.g., solutions are sought of the form

$$
\underline{M}^G = \mathrm{Re}\, \underline{M}^G_{\mathrm{osc}} \exp[-i(\underline{\xi} \cdot r - \omega t)] \ ,
$$

$$\underline{M}^G_{\rm osc}(r,t) = \int d^2\underline{\xi}\, d\omega\, \hat{\underline{M}}^G_{\rm osc}(\underline{\xi},\omega) \exp[i(\underline{\xi}\cdot r - \omega t)] \ . \tag{96}$$

For instance, a typical example is specified by extrinsic sources $J_E = -2.63$ and $J_I = 4.94$, $N^E = 125$, $N^I = 25$, $V^G = 10$ mV, $A^G_E = 1.75$, $A^G_I = 1.25$, $B^G_{G'} = 0.25$, and $v^G_{G'} = \phi^G_{G'} = 0.1$ mV. The synaptic parameters are within observed ranges [72], and the $J_G$'s are just those values required to solve the Euler-Lagrange equations at the selected values of $M^G$. The global minimum is at $\bar{M}^E = 25$ and $\bar{M}^I = 5$. This set of conditions yields (dispersive) dispersion relations

$$\omega\tau = \pm\{ -1.86 + 2.38(\xi\rho)^2; -1.25i + 1.51i(\xi\rho)^2 \} \ , \tag{97}$$

where $\xi = |\underline{\xi}|$. The propagation velocity defined by $d\omega/d\xi$ is ~1 cm/sec, taking typical wavenumbers $\xi$ to correspond to macrocolumnar distances ~ $30\rho$. Calculated frequencies $\omega$ are on the order of EEG frequencies ~ $10^2$ sec$^{-1}$, equivalent to $\nu = \omega/(2\pi)$= 16 cps (Hz). These mesoscopic propagation velocities permit processing over several minicolumns ~ $10^{-1}$ cm, simultaneous with the processing of mesoscopic interactions over tens of centimeters via association fibers with propagation velocities ~600–900 cm/sec. I.e., both intraregional and interregional information processing can occur within ~ $10^{-1}$ sec. Note that this propagation velocity is not ‘‘slow’’: Visual selective attention moves at ~8 msec/deg [160], which is ~ 1/2 mm/sec, if a macrocolumn of ~mm$^2$ is assumed to span 180 deg. This suggests that nearest-neighbor interactions play some part in disengaging and orienting selective attention.

### C. Calculating synaptic modifications

Perturbations of the synaptic and neuronal parameters used in SMNI present an approach to the discussion of plastic synaptic modifications. These changes are associated with new firing minima and their associated sets of eigenfunctions, related to learning new sets of information. This is especially true during development of synaptic formation, at a rate determined by successive small increments of these perturbations. Changes in the coefficients of gradient couplings also represent shifts in oscillatory states and in the degree of interaction between columnar firings.

To further clarify this methodology, an explicit calculation is given, demonstrating how a small increment of extrinsically imposed firing activity can be learned and stored as plastic synaptic modifications. Moderate changes in efficacies of even one neuron per mesocolumn give rise to moderate changes in macroscopic activity; therefore, macroscopic measurements can, in this theory, be sensitive to some microscopic details of neocortical interactions. Although most researchers believe that long-term potentiation (LTP) at synaptic sites is responsible for long-term memory (LTM) phenomena, they still are not as certain whether these changes take place predominantly at the presynaptic or postsynaptic sites [169]. Therefore, the SMNI study described below [15], looking at induced mesocolumnar effects from these alternative mechanisms, is still relevant.

Consider the change in probability of firing of neuron $j$, $p_{\sigma_j}$, associated with modifications of the neuronal parameters. Changes can occur in some parameters $Z$, where

$$Z \in \{ A^*_{jk}, B^*_{jk}, v_{jk}, \phi_{jk} \} \ , \tag{98}$$

which lead to

$$\Delta Z \in \{ \Delta A^*_{jk}, \Delta B^*_{jk}, \Delta v_{jk}, \Delta\phi_{jk} \} \ . \tag{99}$$

Now consider changes $\Delta Z$ in these parameters $Z$ that can be independent or proportional to the (repeated) firing of neuron(s) postsynaptically ($j$) or presynaptically ($k$).

$$\Delta Z = \Delta Z_1 + \sigma_k \Delta Z_2 + \sigma_k\sigma_j \Delta Z_3 + \sigma_j \Delta Z_4 \ . \tag{100}$$

All these $\Delta Z$ effects collect to modify the threshold factor $F_j$ derived for unit microscopic neuronal firing,

$$F_j \to F_j{}' = F_j + (\Delta F_{j1} + \Delta F_{j2}) + \sigma_j(\Delta F_{j3} + \Delta F_{j4}) \ ,$$

$$\Delta F_{j1,2,3,4} = (\partial F/\partial Z_{j1,2,3,4})\Delta Z_{j1,2,3,4} \ . \tag{101}$$

To order $\Delta F_{j1,2,3,4}$, so that the normalization of probability $p_+ + p_- = 1$ is preserved, $p_{\sigma_j}$ is modified as

$$p_{\sigma_j} \rightarrow p_{\sigma_j}{}' = \frac{\exp(\sigma_j F_j{}'')}{\exp F_j{}''' + \exp(-F_j{}''')} ,$$

$$F_j{}'' = F_j + \Delta F_{j1} + \Delta F_{j2} + (\Delta F_{j3} + \Delta F_{j4}) \tanh F_j ,$$

$$F_j{}''' = F_j + \Delta F_{j1} + \Delta F_{j2} + \Delta F_{j3} + \Delta F_{j4} . \quad (102)$$

(This form of $p_{\sigma_j}{}'$ corrects that given in a previous paper [15], changing the sign of the tanh term and removing this factor in $F_j{}'''$, thereby insuring that probability is conserved to order $\Delta Z$.)

Thus, the change in response of a single neuron associated with its synaptic modifications is a highly nonlinear function of the synaptic parameters $\{ Z, \Delta Z \}$. Nonlinearities persist even after mesocolumnar averaging, but then, because of the derived variational principle, explicit calculations can be performed to portray most-probable changes in patterns of columnar firings associated with changes in the Lagrangian:

$$\underline{L}_F \rightarrow \underline{L}_F + \Delta \underline{L}_F \approx \underline{L}_F + \sum_Z \frac{\partial \underline{L}_F}{\partial Z} \Delta Z . \quad (103)$$

To emphasize the point that linear-response models of neuronal activity should be scrutinized with respect to the underlying biophysics and mathematics, which they are assuming to be linear, the following equation is shown to represent the first-order change in $\underline{L}$ associated with modifications of only the columnar averaged efficacies $A^G_{G'}$, where sums are to taken only over $G'$ terms:

$$\Delta \underline{L} = \Delta F^G (2N^G N \tau)^{-1} \{ [(N^G)^2 + (\tau \dot{M}^G + M^G)^2] \sinh(2F^G) + 2N^G(\tau \dot{M}^G + M^G) \cosh(2F^G) \}$$

$$-\Delta A^G_{G'} (\partial \underline{V}''^G_{G'} / \partial A^G_{G'}) (\rho \nabla M^{G'})^2 ,$$

$$\Delta F^G = -(2F^G_d)^{-1} \Delta A^G_{G'} [v^G_{G'} (M^{G'} + N^{G'}) + \pi (v^G_{G'}{}^2 + \phi^G_{G'}{}^2)(M^{G'} + N^{G'}) F^G / (2F^G_d)] ,$$

$$F^G_d = [\pi (v^G_{G'}{}^2 + \phi^G_{G'}{}^2)(A^G_{G'} M^{G'}/2 + a^G_{G'} N^{G'})]^{1/2} ,$$

$$\Delta A^G_{G'} = \Delta A^G_{G'1} + \Delta A^G_{G'2} + (\Delta A^G_{G'3} + \Delta A^G_{G'4}) \tanh F^G . \quad (104)$$

Examining $\Delta A^G_{G'}$, it is clear that even after mesocolumnar averaging, groups of synaptic modifications dependent on postsynaptic firings can be discerned from groups of modifications independent of this activity, by the additional $\tanh F^G$ factor. However, since

$$\sigma_k (1 + \sigma_k) = 1 + \sigma_k , \quad (105)$$

mesocolumnar averaging washes out discrimination of $\Delta A^G_{G'1,3}$ from $\Delta A^G_{G'2,4}$ unless these possess additional distinguishing functional features. Similar calculations are proposed to further investigate phenomena as encountered in habituation [170].

For instance, a system considered was $N^E = 125$, $N^I = 25$, $V^G = 10$ mV, $A^G_E = 1.25$, $A^G_I = 1.75$, $B^G_{G'} = 0.25$, $v^G_{G'} = \phi^G_{G'} = 0.1$ mV, and $J_G = 0$ (no long-ranged connectivity) [15]. This system was synaptically modified about its most probable firing state by $\Delta A^{E^G}_{E3} = -0.01 \tanh F^G$, e.g., numerically equivalent to a substantial change in $A_{jk}$ of one $E$ neuron per mesocolumn in a region. Then, the change in the uniform Lagrangian was

$$\tau \Delta \overline{\underline{L}} \approx -4.87 \times 10^{-4} + 3.99 \times 10^{-6} (\bar{M}^E - \ll \bar{M}^E \gg)$$

$$-9.80 \times 10^{-5} (\bar{M}^I - \ll \bar{M}^I \gg) . \quad (106)$$

The shifts in the most-probable firing state $\ll \bar{M}^G \gg$ associated with this synaptic modification are observed to be algebraically equivalent, within a constant increment to $\underline{L}$, to those that could also have been caused by extrinsic stimulations measured by $J_E/(2\tau N) = 3.99 \times 10^{-6}$ and $J_I/(2\tau N) = -9.80 \times 10^{-5}$. This shifts $\ll \bar{M}^G \gg$ and $\overline{\underline{L}}$, $(\ll \bar{M}^E \gg, \ll M^I \gg, \overline{\underline{L}})$,

$$(89.02, 23.14; 1.59 \times 10^{-3}) \to (89.20, 23.19; -3.25 \times 10^{-4}) \; . \tag{107}$$

Changes in the coefficients of spatial derivatives $[(\nabla M^E)^2 \| (\nabla M^I)^2]$ are

$$[-5.33 \times 10^{-5} \| 9.65 \times 10^{-4}] \to [-5.72 \times 10^{-5} \| 9.65 \times 10^{-4}] \; . \tag{108}$$

Changes in the coefficients of temporal derivatives $[(\dot{M}^E)^2 : (\dot{M}^I)^2 : : \dot{M}^E : \dot{M}^I]$ are

$$[4.81 \times 10^{-5} : 1.09 \times 10^{-3} : : 5.38 \times 10^{-4} : -6.14 \times 10^{-4}]$$

$$\to [4.82 \times 10^{-5} : 1.10 \times 10^{-3} : : 5.43 \times 10^{-4} : -5.28 \times 10^{-4}] \; . \tag{109}$$

From the Euler-Lagrange equations, $\omega\tau$ is shifted,

$$\pm \{ \, 0.392i - 1.68i(\xi\rho)^2 ; 1.01i - 0.541i(\xi\rho)^2 \, \}$$

$$\to \pm \; \{ \, 0.396i - 1.79i(\xi\rho)^2 ; 1.01i - 0.550i(\xi\rho)^2 \, \} \; . \tag{110}$$

These numerical estimates suggest that the sensitivity of mesocolumnar statistics to microscopic dynamics barely might be within the present range of experimental determination, e.g., to changes induced by small clusters of neurons. This calculation also represents an explicit demonstration of how extrinsic constraints on firing patterns can be learned and coded by plastic synaptic modifications. In general, there exist (a set of) synaptic modifications $\Delta Z(r; t')$ that reproduce the most probable firing states $\ll \tilde{M}(t') \gg$ induced by $J_G(r; t')$.

Given the above, a precise scenario of neocortical information processing is thereby detailed, from coding of long-ranged firings from stimuli external to a macrocolumn by short-ranged mesocolumnar firings, to STM storage via hysteresis, and to LTM storage via plastic deformation [15]. This scenario enables SMNI to place some statistical bounds on such mechanisms. For example, as calculated in detail, some STM items can be held for long enough epochs via hysteresis [15,16], from which plastic synaptic changes can be used to store these in LTM [15]. In contrast to the appearance of multiple minima in the interior of $M^G$ space under conditions of realistic but sensitive adjustment of synaptic interactions, which are candidates for multiple STM [16], typically one or at most a few minima appear at the corners of $M^G$ space. This typically occurs when the system likely has minima outside the physical boundary. These boundary states correspond to all $G$ neurons collectively firing or not firing [15]. When these corner minima are present, they are typically much deeper than those found for the interior minima, corresponding to longer-lived states with properties of hysteresis rather than simple jumps. These corner minima are therefore candidates for LTM phenomena. Similar properties of corner minima in simpler models of the neocortex have been shown to satisfy properties desirable for multistable perception and for collective computational properties [171].

## V. DIRECT FIT OF SMNI TO EEG

### A. Data collection

The project uses the collection of EEG spontaneous and averaged evoked potential (AEP) data from a multi-electrode array under a variety of conditions. We are fitting data being collected at several centers in the United States, sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) project [172,173]. Another paper to appear in the neuroscience literature will address issues of neuroscientific and clinical import. These experiments, performed on carefully selected sets of subjects, suggest a genetic predisposition to alcoholism that is strongly correlated to EEG AEP responses to patterned targets.

For the purposes of this paper, it suffices to explain that we fit data obtained from 19 electrode sites on each of 49 subjects, of which 25 are considered to be high risk with respect to a propensity to alcoholism, and 24 are considered to be low risk. Each subject participated in EEG-monitored pattern-matching tasks. The time epoch during which the P300 EP exists was extracted (the P300 EP is named for its appearance over 300 msec after an appropriate stimulus), yielding 191 time epochs of 5.2 msec for each of the above circumstances. Each set of 192 pieces of data is obtained by having the subject perform similar pattern-matching tasks, e.g., about 100 such tasks, time-locking the EEG responses to the initiation of the task, and averaging over the set of tasks for each time epoch.

### B. Algebraic development

We take Eq. (10) as the basic probability distribution to fit this data. This can be developed straightforwardly using $\delta(Z)$, given in Eq. (23), and $P[M^G(t+\tau)|M^G(t)]$, given in Eq. (41).

Alternately, advantage can be taken of the prepoint discretization. We also take advantage of and extend the results gained for the STM analysis discussed previously. Accordingly, we assume a linear relationship (about minima to be fit to data) between the $M^G$ firing states and the measured scalp potential $\Phi_\nu$, at a given electrode site $\nu$ representing a macroscopic region of neuronal activity:

$$\Phi_\nu - \phi = aM^E + bM^I \ , \tag{111}$$

where $\{ \phi, a, b \}$ are constants determined for each electrode site. In the prepoint discretization, the postpoint $M^G(t+\Delta t)$ moments are given by

$$m \equiv \langle \Phi_\nu - \phi \rangle = a \langle M^E \rangle + b \langle M^I \rangle$$

$$= ag^E + bg^I \ ,$$

$$\sigma^2 \equiv \langle (\Phi_\nu - \phi)^2 \rangle - \langle \Phi_\nu - \phi \rangle^2 = a^2 g^{EE} + b^2 g^{II} \ , \tag{112}$$

where the $M^G$-space drifts $g^G$, and diffusions $g^{GG'}$, have been derived above. Note that the macroscopic drifts and diffusions of the $\Phi$'s are simply linearly related to the mesoscopic drifts and diffusions of the $M^G$'s. For the prepoint $M^G(t)$ firings, we assume the same linear relationship in terms of $\{ \phi, a, b \}$.

The data we are fitting are consistent with invoking the "centering" mechanism discussed above. Therefore, for the prepoint $M^E(t)$ firings, we also take advantage of the parabolic trough derived for the STM Lagrangian, and take

$$M^I(t) = cM^E(t) \ , \tag{113}$$

where the slope $c$ is determined for each electrode site. This permits a complete transformation from $M^G$ variables to $\Phi$ variables.

Similarly, as appearing in the modified threshold factor $F^G$ given in Eq. (63), each regional influence from electrode site $\mu$ acting at electrode site $\nu$, given by afferent firings $M^{\ddagger E}$, is taken as

$$M^{\ddagger E}_{\mu\to\nu} = d_\nu M^E_\mu (t - T_{\mu\to\nu}) \ , \tag{114}$$

where $d_\nu$ are constants to be fitted at each electrode site, and $T_{\mu\to\nu}$ is the delay time estimated for inter-electrode signal propagation, based on current anatomical knowledge of the neocortex and of velocities of propagation of action potentials of long-ranged fibers, typically on the order of one to several multiples of $\tau$= 5 msec. Some terms in which $d$ directly affects the shifts of synaptic parameters $B^G_{G'}$ when calculating the centering mechanism also contain long-ranged efficacies (inverse conductivities) $B^{*E}_{E'}$. Therefore, the latter were kept fixed with the other electrical-chemical synaptic parameters during these fits. In future fits, we will experiment taking the $T$'s as parameters.

This defines the conditional probability distribution for the measured scalp potential $\Phi_\nu$,

$$P_\nu[\Phi_\nu(t+\Delta t)|\Phi_\nu(t)] = \frac{1}{(2\pi\sigma^2\Delta t)^{1/2}} \exp(-L_\nu \Delta t) \ ,$$

$$L_\nu = \frac{1}{2\sigma^2} (\dot{\Phi}_\nu - m)^2 \ , \tag{115}$$

where $m$ and $\sigma$ have been derived just above. As discussed above in defining macroscopic regions, the probability distribution for all electrodes is taken to be the product of all these distributions:

$$P = \prod_\nu P_\nu \ ,$$

$$L = \sum_{\nu} L_{\nu} \ . \tag{116}$$

Note that we are also strongly invoking the current belief in the dipole or nonlinear-string model. The model SMNI, derived for $P[M^G(t + \Delta t)|M^{\bar{G}}(t)]$, is for a macrocolumnar-averaged minicolumn; hence we expect it to be a reasonable approximation to represent a macrocolumn, scaled to its contribution to $\Phi_{\nu}$. Hence we use $L$ to represent this macroscopic regional Lagrangian, scaled from its mesoscopic mesocolumnar counterpart $\underline{L}$. However, the above expression for $P_{\nu}$ uses the dipole assumption to also use this expression to represent several to many macrocolumns present in a region under an electrode: A macrocolumn has a spatial extent of about a millimeter. A scalp electrode has been shown just recently, under extremely favorable circumstances, to have a resolution as small as several millimeters, directly competing with the spatial resolution attributed to magnetoencephalography; often most data represents a resolution more on the order of up to several centimeters, many macrocolumns. Still, it is often argued that typically only several macrocolumns firing coherently account for the electric potentials measured by one scalp electrode [51]. Then, we are testing this model to see if the potential will scale to a representative macrocolumn. The results presented here seem to confirm that this approximation is in fact quite reasonable.

As noted in a previous SMNI paper [16], the structure of STM survives an approximation setting $M^G = 0$ in the denominator of $F^G$, after applying the "centering" mechanism. To speed up the fitting of data in this first study, this approximation was used here as well.

The resolution of this model is certainly consistent with the resolution of the data. For example, for the nonvisual neocortex, taking the extreme of permitting only unit changes in $M^G$ firings, it seems reasonable to always be able to map the observed electric potential values $\Phi$ from a given electrode onto a mesh a fraction of $4N^E N^I \approx 10^4$.

### C. Numerical methodology

Coarse statistical fits first identify major correlated electrode sites within each class of behavioral activity. Then, the macrocolumnar-averaged synaptic parameters in the nonlinear string model, including long-ranged interactions between these electrode locations, are fit to the EEG data within in each class using methods of simulated annealing [58]. Recently, two major computer codes have been developed, which are key tools for the use of this approach.

The first code, very fast simulated reannealing (VFSR) [58], fits short-time probability distributions to observed data, using a maximum likelihood technique on the "effective" Lagrangian (including the exponential prefactor). This algorithm has been developed to fit observed data to a large class of theoretical cost function over a $D$-dimensional parameter space [58], adapting for varying sensitivities of parameters during the fit. The annealing schedule for the "temperatures" (artificial fluctuation parameters) $T_i$ decrease exponentially in "time" (cycle number of iterative process) $k$, i.e., $T_i = T_{i0} \exp(-c_i k^{1/D})$.

Heuristic arguments have been developed to demonstrate that this algorithm is faster than the fast Cauchy annealing [174], $T_i = T_0/k$, and much faster than Boltzmann annealing [175], $T_i = T_0/\ln k$. To be more specific, the $k$ th estimate of parameter $\alpha^i$,

$$\alpha^i_k \in [A_i, B_i] \ , \tag{117}$$

is used with the random variable $x^i$ to get the $(k+1)$th estimate,

$$\alpha^i_{k+1} = \alpha^i_k + x^i(B_i - A_i) \ ,$$

$$x^i \in [-1, 1] \ . \tag{118}$$

The generating function is defined as

$$g_T(x) = \prod_{i=1}^{D} \frac{1}{2\ln(1 + 1/T_i)(|x^i| + T_i)} \equiv \prod_{i=1}^{D} g^i_T(x^i) \ ,$$

$$T_i = T_{i0} \exp(-c_i k^{1/D}) \ . \tag{119}$$

The cost function $C$ used here is defined by

$$C = Ldt + \frac{1}{2}\ln(2\pi dt) - \ln(g) \ , \tag{120}$$

where the Lagrangian $L$ and the determinant of the metric $g$ have been defined previously.

The second code develops the long-time probability distribution from the Lagrangian fit by the first code. A robust and accurate histogram-based (non-Monte Carlo) path-integral algorithm to calculate the long-time probability distribution has been developed to handle nonlinear Lagrangians [123-125], which was extended to two-dimensional problems [36].

The histogram procedure recognizes that the distribution can be numerically approximated to a high degree of accuracy as sum of rectangles at points $M_i$ of height $P_i$ and width $\Delta M_i$. For convenience, just consider a one-dimensional system. The above path-integral representation can be rewritten, for each of its intermediate integrals, as

$$\begin{aligned} P(M; t + \Delta t) &= \int dM' [g_s^{1/2} (2\pi\Delta t)^{-1/2} \exp(-L_s \Delta t)] P(M'; t) \\ &= \int dM' G(M, M'; \Delta t) P(M'; t) \ , \end{aligned}$$

$$P(M; t) = \sum_{i=1}^{N} \pi(M - M_i) P_i(t) \ ,$$

$$\pi(M - M_i) = \begin{cases} 1 \ , & (M_i - \frac{1}{2}\Delta M_{i-1}) \le M \le (M_i + \frac{1}{2}\Delta M_i) \ , \\ 0 \ , & \text{otherwise} \ . \end{cases} \tag{121}$$

This yields

$$P_i(t + \Delta t) = T_{ij}(\Delta t) P_j(t) \ ,$$

$$T_{ij}(\Delta t) = \frac{2}{\Delta M_{i-1} + \Delta M_i} \int_{M_i - \Delta M_{i-1}/2}^{M_i + \Delta M_i/2} dM \int_{M_j - \Delta M_{j-1}/2}^{M_j + \Delta M_j/2} dM' G(M, M'; \Delta t) \ . \tag{122}$$

$T_{ij}$ is a banded matrix representing the Gaussian nature of the short-time probability centered about the (varying) drift.

This histogram procedure has been extended to two dimensions, i.e., using a matrix $T_{ijkl}$ [36]. Explicit dependence of $L$ on time $t$ also can be included without complications. We have recently extended it to an arbitrary number of dimensions, but care must be used in developing the mesh in $\Delta M^G$, which is dependent on the diffusion matrix. Presently, this constrains the dependence of the covariance of each variable to be a (nonlinear) function of that variable, in order to present a straightforward rectangular underlying mesh.

Since integration is inherently a smoothing process [42], fitting data with the short-time probability distribution, effectively using an integral over this epoch, permits the use of coarser meshes than the corresponding stochastic differential equation. For example, the coarser resolution is appropriate, typically required, for numerical solution of the time-dependent path integral. Similar to the discussion above for STM, by considering the contributions to the first and second moments of $\Delta\Phi$ for small time slices $\theta$, conditions on the time and variable meshes can be derived [123]. The time slice essentially is determined by $\theta \le \bar{L}_\nu^{-1}$, where $\bar{L}\nu$ is the uniform electric-potential Lagrangian, with $d\Phi/dt = 0$, throughout the ranges of $\Phi$ giving the most important contributions to the probability distribution $P$. Thus, $\theta$ is roughly measured by the diffusion divided by the square of the drift. Here, an upper bound can be calculated by looking at the averages over all time epochs of the drifts and diffusions under each electrode, yielding a mean ratio typically on the order of 1 to 10. The 5.2-msec sampling rate of our data is certainly less than this $\theta$. This might seem to be too fine a temporal mesh for purposes of doing the path integral, but that calculation would have to weigh heavily the most likely trajectories, not a mean over averaged EP observed data.

The boundary conditions for the mesocolumnar probability distribution are reflecting at the four walls confining $-N^G \le M^G \le N^G$. This methodology readily permits such inclusion in its numerical implementation and we have performed such calculations in other systems using the method of images [12], which has proven to be about as accurate as boundary element methods for the systems investigated [36]. These also can be incorporated into the cost function being fit, but since we have invoked the centering mechanism, favoring interior minima, and to increase the speed of our runs, we did not put the boundary conditions into the fits reported below.

### D. Present results

For this first study, we used some current knowledge of the P300 EP phenomena to limit ourselves to just five electrodes per subject, corresponding to hypothesized fast and slow components of P300. The first component appears to be generated along the brain midline, from frontal (Fz) to central (Cz) to parietal (Pz) areas; a delay time of one 5.2-msec epoch was assumed for each relay. The slow component appears to be generated from Pz, branching out to lateral areas P3 and P4; a delay time of two 5.2-msec epochs was assumed for each branch. Since P300 has such a quite broad rise, peak, and decay over a large fraction of a second, regional delays are not expected to be very important here. Data currently being collected on more stringent time-locked STM tasks are expected to provide stronger tests of the importance of such delays. Furthermore, the relative lack of sensitivity of fits to such delays here suggests that volume conductance effects are large in these data, and Laplacian techniques to localize EEG activities are required to get more electrode-specific sensitivity to such delays. However, the main emphasis here is to determine whether SMNI is consistent with EEG data collected under conditions of selective attention, and these results appear to be quite strong.

The P300 EP, so named because of its appearance over 300 msec after an appropriate stimulus, has been demonstrated to be negatively correlated (reduction in amplitude and delay) with a number of psychiatric diseases, e.g., schizophrenia and depression, and typically is most active at sites Pz, P3 and P4 [176]. Here, the suggestion is that there also is some correlation with some precursor activity at Fz and Cz.

Thus, in this paper we are reporting fits to 46,550 pieces of data. As described above in the section deriving $P[\Phi(t+\Delta t)|\Phi(t)]$, we have: four parameters at site Fz, corresponding to coefficients $\{ \phi, a, b, c \}$; five parameters at Cz, $\{ \phi, a, b, c, d_{\mathrm{Fz}\to\mathrm{Cz}} \}$; five parameters at Pz, $\{ \phi, a, b, c, d_{\mathrm{Cz}\to\mathrm{Pz}} \}$; five parameters at P3, $\{ \phi, a, b, c, d_{\mathrm{Pz}\to\mathrm{P3}} \}$; and five parameters at P4, $\{ \phi, a, b, c, d_{\mathrm{Pz}\to\mathrm{P4}} \}$. This represents a 24-parameter fit for 950 points of data (each electrode offset by two points to account for delays) for each of 49 subjects.

The VFSR runs took several CPU hours each on a personal Sun SPARCstation 2 (28.5 MIPS, 21 SPECmarks) running under GNU g++, a C++ compiler developed under the GNU project at the Massachusetts Institute of Technology, which proved to yield faster runs than using Sun's bundled non-ANSI C, depending on how efficiently the simulated annealing run could sense its way out of local minima. Runs were executed for inclusion of delays between electrodes, as discussed above. All runs reported here were completed in approximately 400 CPU hours. Typically, at least one to three significant-figure consistencies between finer resolution runs per parameter were obtained by exiting the global simulated annealing runs after either two sets of 100 acceptances or 20,000 trials led to the same best estimate of the global minima. Each trial typically represented a factor of 3 to 5 other generated sets of randomly selected parameters, which did not satisfy the physical constraints on the electrode sets of $\{ M^G \}$, $\{ M^{*E} \}$ and the centering mechanism (which required calculation of new synaptic parameters $\{ B^G_{G'} \}$ for each new set of regional connectivity parameters $\{ d \}$). Some efficiency was gained by using the means and extremes of the observed electric potentials as a guide for the ranges of the sets of intercept parameters $\{ \phi \}$.

Then, several more significant-figure accuracy was obtained by shunting the code to a local fitting procedure, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [177], where it either exited naturally or was forcefully exited, saving the lowest cost function to date, after exceeding a limit of 1000 function calls. The local BFGS runs enforced the above physical constraints by adding penalties to the cost functions calculated with trial parameters, proportional to the distance out of range.

These sets of EEG data were obtained from subjects while they were reacting to pattern-matching tasks requiring varying states of selective attention taxing their short-term memory. To test the assumptions made in the model, after each subject's data set was fitted to its probability distribution, the data were again filtered through the fitted Lagrangian, and the mean and mean-square values of $M^G$ were recorded as they were calculated from $\Phi$ above. The average values from the 49 subjects, each in turn representing averages over 190 points of collected EEG data and the standard deviations were calculated simply from $\{[< (M^G)^2 > - < M^G >^2] n/(n-1)\}^{1/2}$, where $n = 49$. Table I gives the means and standard deviations of the effective firings, $M^G$, aggregated from all subjects under each electrode, and the weight $d$ of the regional (time-delayed) inputs $M^{*E}$. Also given for the firings under each electrode are the means and standard deviations of the standard deviations aggregated from all time epochs from each subject; these may give some insights into correlations of different states of selective attention with respect to coherency and stability of columnar firings. The weights $d$, as with the synaptic parameters affected by the centering mechanism, are the same across time epochs for a given subject, and therefore do not have any dispersion as reported for the firings. Indeed, although $M^G$ were permitted to roam throughout their physical ranges of $\pm N^E = \pm 80$ and $\pm N^I = \pm 30$ (in the nonvisual neocortex as is the case for all these regions), their observed effective (regional- and macrocolumnar-averaged) minicolumnar firing states were observed to obey the centering mechanism. I.e., this numerical result is consistent with the assumption that $M^G \approx 0 \approx M^{*E}$ in $F^G$.

Table I.

Based on psychiatric and family-history evaluations, these 49 subjects were classified into two groups, as possibly having high-risk and low-risk genetic propensities to alcoholism. The high-risk group was composed of 25 subjects, and the low-risk group was composed of the remaining 24 subjects. Tables II and III give the same statistics as in Table I, for each subgroup.

While neuroscientists are very hopeful that such discrimination is possible on the basis of EEG, not all research laboratories have been able to produce such evidence. It seems that the long time epochs spanning P300, e.g., the 1-sec epoch fitted in Table I, are just not selective enough of features occurring within about 1/2 second. When the smaller epoch is examined, occurring at a different temporal location within the 1-sec epoch for each subject, some distinguishing P300 patterns appear to emerge between low-risk and high-risk groups [178]. The SMNI approach is seen as a viable approach to interpret these correlations into physical variables reflecting neocortical activity.

Figure 6 gives one of the 49 sets of data fit here. This P300 is particularly well synchronized and well defined with respect to the five electrodes. The first two dips are known as N100 and N200; the first broad peak afterwards is the P300. It is believed that the P300 shape may be an indication of the genetic marker for alcoholism. Since all genes are not dominant, assuming such genetic markers for alcoholism exist, even properly classified low-risk or high-risk subjects need not possess the discriminating EEG trace. In fact, quite a few of these 49 subjects have such similar patterns of data that they could well be classified in either group if the only criterion were their EEG. A larger subject sample is required to discern if P300 can be a correct identifier of such behavioral traits. Then, it still remains to be proved that such statistical calculations as presented here can establish quantitative norms for these two groups.

Figure 6.

The results in Tables II and III reflect the lack of correlation observed between EEG recordings and behavioral states when the temporal epoch used is too wide, as was generally done in earlier EEG studies. Here, this calculation is offered as a control. We are performing fits on the 1/2-sec epochs containing the selected P300 features that may give tighter correlations, and these will be reported at a later date.

Based on the plots and numerical calculations performed in this study, it is the opinion of this author, not necessarily shared by other neuroscientists, that the high degree of synchrony of almost equal wave-forms of AEP, especially of N100 and N200, at these several electrode sights is actually due to volume conductance effects [51] that are swamping the underlying local activity at each electrode. I.e., each electrode is effectively recording an average global response, a known circumstance which can be caused by several effects. If true, then if the recording of more localized activity is desired, Laplacian techniques, e.g., derived from clusters of five closely-spaced electrodes, are required.

Independent of the circumstances of this suggestion, the conclusions drawn here, regarding the SMNI description of EEG under conditions of selective attention, are not radically altered since we still are fitting the SMNI model to valid EEG AEP data. The approach here has been used to investigate the consistency of SMNI with previous STM studies using this theory. The 1-sec epoch does signify a period of relatively intense selective attention, and it is of interest to see if this is consistent with the SMNI centering mechanism. This has been accomplished independent of the success of finding EEG correlates of genetic predispositions to alcoholism.

Table II.

Table III.

In the context of current paradigms of neuroscience, and in the context of appreciating that SMNI offers statistical bounds on neocortical interactions, i.e., not necessarily offering specific neuronal mechanisms correlated to specific brain states, this information addresses “connectivity” and “coherency.” Connectivity is simply addressed by examining the bottom third set of rows of numbers in each table, where the coefficients $d$ of regional influences are given. Coherency is addressed in an approximate way, by examining the last two rows of numbers in each table, where a measure of the variance of the mean firings and connectivity and the variances among these means can be obtained. I.e., it might be assumed that the degree of selective attention is related to the degree of connectivity and to the degree of focus of firing states.

As just discussed, comparison of the data in Tables II and III shows no statistical differences between these groups in terms of the physical SMNI model. However, even within these wide statistical margins, it is clear that the values of the columnar firings, permitted to roam throughout their physical ranges during the fitting procedure, do in fact cluster according to the centering mechanism derived above in the context of STM. While it is premature to speculate on the relevance of these differences between these two subgroups, these data are offered here to demonstrate the kind of information this approach can offer.

Again, it must be stressed that such nonlinear stochastic modeling may be essential in such noisy systems, to first fit the functional nature of both the drifts and the diffusions, in order to then extricate some distinguishing signatures of subsystems. Further work is in progress with variations of SMNI and other models, to give some insights into just what nonlinear stochastic features are essential to fit such EEG data.

This particular five-electrode circuitry may yet be amenable to our path-integral algorithm. Each variable typically requires a mesh size that is a product of a postpoint mesh, $\Delta$, determined by the square root of the (non-constant) diffusion, times $2\Delta' + 1$, where $\Delta'$ is the size of the off-diagonal prepoint taken about the (non-constant) drift; $\Delta'$ is typically 3−7 to get good numerical accuracy. For five variables requiring a product of all five $\Delta$ and $\Delta'$ values, this does not seem a reasonable calculation on today's standard computers. However, here we really have five distributions of only one or two variables each of which must be simultaneously propagated: $P(\mathrm{Fz})$, $P(\mathrm{Cz}|\mathrm{Fz})$, $P(\mathrm{Pz}|\mathrm{Cz})$, $P(\mathrm{P2}|\mathrm{Pz})$, and $P(\mathrm{P2}|\mathrm{Pz})$. This work is in progress.

These results, in addition to their importance in reasonably modeling EEG with SMNI, also have a deeper theoretical importance with respect to the scaling of neocortical mechanisms of interaction across disparate spatial scales and behavioral phenomena: As has been pointed out previously, SMNI has given experimental support to the derivation of the mesoscopic probability distribution, yielding similar algebraic structures of the threshold factors in Eqs. (32) (63), illustrating common forms of interactions between their entities, i.e., neurons and columns of neurons, respectively. The nonlinear threshold factors are defined in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges. It also was noted that the most likely trajectories of the mesoscopic probability distribution, representing averages over columnar domains, give a description of the systematics of macroscopic EEG in accordance with experimental observations. In this section, we have demonstrated that we can derive the macroscopic regional probability distribution, Eq. (115), as having the same functional form as the mesoscopic Eq. (41), where the macroscopic drifts and diffusions of the potentials described by the $\Phi$'s are simply linearly related to the (nonlinear) mesoscopic drifts and diffusions of the columnar firing states given by the $M^G$'s. Then, this macroscopic probability distribution gives a reasonable description of experimentally observed EEG.

## VI. CONCLUSION

We have outlined in some detail a reasonable approach to extract more ‘‘signal’’ out of the ‘‘noise’’ in EEG data, in terms of physical dynamical variables, than by merely performing regression statistical analyses on collateral variables. To learn more about complex systems, we inevitably must form functional models to represent huge sets of data. Indeed, modeling phenomena is as much a cornerstone of 20th century science as is collection of empirical data [179].

We have been able to fit these sets of EEG data taken during selective attention tasks, using parameters either set to experimentally observed values, or being fitted within experimentally observed values. The ranges of columnar firings are consistent with a centering mechanism derived for STM in earlier papers.

The ability to fit data to these particular SMNI functional forms goes beyond nonlinear statistical modeling of data. The plausibility of the SMNI model, as emphasized in this and previous SMNI papers, as spanning several important neuroscientific phenomena, suggests that the fitted functional forms may yet help to explicate some underlying biophysical mechanisms responsible for the normal and abnormal behavioral states being investigated, e.g., excitatory and/or inhibitory influences, background electromagnetic influences from nearby firing states (by using SMNI synaptic conductivity parameters in the fits).

There is much more work to be done. We have not yet addressed the “renormalization” issues raised, which are based on the nature of EEG data collection and which are amenable to this framework. While the fitting of these distributions certainly compacts the experimental data into a reasonable algebraic model, a prime task of most physical theory, in order to be useful to clinicians (and therefore to give more feedback to theory) even more data reduction must be performed. We are experimenting with path-integral calculations and some methods of “scientific visualization” to determine what minimal, or at least small, set of “signatures” might suffice, which would be faithful to the data yet useful to clinicians. We also are examining the gains that might be made by putting these codes onto a parallel processor, which might enable real-time diagnoses based on noninvasive EEG recordings.

In order to detail such a model of EEG phenomena we found it useful to seek guidance from ‘‘top-down’’ models; e.g., the nonlinear string model representing nonlinear dipoles of neuronal columnar activity. In order to construct a more detailed ‘‘bottom-up’’ model that could give us reasonable algebraic functions with physical parameters to be fit by data, we then needed to bring together a wealth of empirical data and modern techniques of mathematical physics across multiple scales of neocortical activity. At each of these scales, we had to derive and establish reasonable procedures and submodels for climbing from scale to scale. Each of these submodels could then be tested against some experimental data to see if we were on the right track. For example, at the mesoscopic scale we checked the consistency of SMNI with known aspects of visual and auditory short-term memory; at the macroscopic scale we checked consistency with known aspects of EEG and the propagation of information across the neocortex. Here, we have demonstrated that the currently accepted dipole EEG model can be derived as the Euler-Lagrange equations of an electric-potential Lagrangian.

The theoretical and experimental importance of specific scaling of interactions in the neocortex has been quantitatively demonstrated: We have shown that the explicit algebraic form of the probability distribution for mesoscopic columnar interactions is driven by a nonlinear threshold factor of the same form taken to describe microscopic neuronal interactions, in terms of electrical-chemical synaptic and neuronal parameters all lying within their experimentally observed ranges; these threshold factors largely determine the nature of the drifts and diffusions of the system. This mesoscopic probability distribution has successfully described STM phenomena and, when used as a basis to derive the most likely trajectories using the Euler-Lagrange variational equations, it also has described the systematics of EEG phenomena. In this paper, we have taken the mesoscopic form of the full probability distribution more seriously for macroscopic interactions, deriving macroscopic drifts and diffusions linearly related to sums of their (nonlinear) mesoscopic counterparts, scaling its variables to describe interactions among regional interactions correlated with observed electrical activities measured by electrode recordings of scalp EEG, with apparent success. These results give strong quantitative support for an accurate intuitive picture, portraying neocortical interactions as having common algebraic or physics mechanisms that scale across quite disparate spatial scales and functional or behavioral phenomena, i.e., describing interactions among neurons, columns of neurons, and regional masses of neurons.

It seems reasonable to speculate on the evolutionary desirability of developing Gaussian-Markovian statistics at the mesoscopic columnar scale from microscopic neuronal interactions, and maintaining this type of system up to the macroscopic regional scale. I.e., this permits maximal processing of information [130]. There is much work to be done, but we believe that modern methods of statistical mechanics have helped to point the way to promising approaches.

## ACKNOWLEDGMENTS

Henri Begleiter, of the Neurodynamics Laboratory at the State University of New York Health Center at Brooklyn, suggested the electrode circuitry used here as well as providing his EEG data. William Brock, at the University of Wisconsin, provided helpful suggestions on the statistical development of the results. Paul Nunez, of the Brain Physics Group at Tulane University, has provided much information about current work in the EEG community as well as providing his EEG data during the early stage of code development. The University of Maryland Institute for Advanced Computer Studies graciously provided their computer facilities until I purchased my own Sun SPARCstation 2 for production runs.

## APPENDIX A: DERIVATION OF PATH INTEGRAL

This appendix outlines the derivation of the path-integral representation of the nonlinear Langevin equations, via the Fokker-Planck representation. This serves to point out the importance of properly treating nonlinearities, and to emphasize the deceptive simplicity of the Langevin and Fokker-Planck representations of stochastic systems. There are a few derivations in the literature, but the following blend seems to be the most concise. All details may be found in the references given here [27,180-182].

The Stratonovich (midpoint discretized) Langevin equations can be analyzed in terms of the Wiener process $dW^i$, which can be rewritten in terms of Gaussian noise $\eta^i = dW^i/dt$ if care is taken in the limit [27].

$$dM^G = f^G(t, M(t))dt + \hat{g}_i^G(t, M(t))dW^i \ ,$$

$$\dot{M}^G(t) = f^G(t, M(t)) + \hat{g}_i^G(t, M(t))\eta^i(t) \ ,$$

$$dW^i \to \eta^i dt \ ,$$

$$M = \{ \ M^G; G = 1, \cdots, \Lambda \ \} \ ,$$

$$\eta = \{ \ \eta^i; i = 1, \cdots, N \ \} \ .$$

$$\dot{M}^G = dM^G/dt \ ,$$

$$< \eta^j(t) >_\eta = 0 \ ,$$

$$< \eta^j(t), \eta^{j'}(t') >_\eta = \delta^{jj'}\delta(t - t') \ , \qquad \text{(A1)}$$

$\eta^i$ represents Gaussian white noise, and moments of an arbitrary function $F(\eta)$ over this stochastic space are defined by a path-type integral over $\eta^i$,

$$< F(\eta) >_\eta = \bar{N}^{-1} \int D\eta F(\eta) \exp\left( -\frac{1}{2} \int_{t_0}^{\infty} dt \eta^i \eta^i \right),$$

$$\bar{N} = \int D\eta \exp\left( -\frac{1}{2} \int_{t_0}^{\infty} dt \eta^i \eta^i \right),$$

$$D\eta = \lim_{v \to \infty} \prod_{\alpha=0}^{v+1} \prod_{j=1}^{N} (2\pi\theta)^{-1/2} dW_\alpha^j \ ,$$

$$t_\alpha = t_0 + \alpha\theta \ ,$$

$$\frac{1}{2} \int dt \eta^i \eta^i = \frac{1}{2\theta} \sum_\beta \sum_i (W_\beta^i - W_{\beta-1}^i)^2 \ ,$$

$$< \eta^i >_\eta = 0 \ ,$$

$$< \eta^i(t)\eta^j(t') >_\eta = \delta^{ij}\delta(t - t') \ . \qquad \text{(A2)}$$

Non-Markovian sources $\hat{\eta}$, and their influence throughout this development, can be formally treated by expansions about the Markovian process by defining

$$< F(\hat{\eta}) >_\eta = \bar{N}_\xi^{-1} \int D\hat{\eta} F \ \exp\left( -\frac{1}{2} \int \int dt dt' \hat{\eta}(t) \Delta_\xi^{-1}(t - t') \hat{\eta}(t') \right),$$

$$\int dt\ \Delta_\xi^{-1}(t-t')\Delta_\xi(t'-t'') = \delta(t-t'')\ , \tag{A3}$$

with $\xi$ defined as an interval centered about the argument of $\Delta_\xi$. Letting $\xi \to 0$ is an unambiguous procedure to define the Stratonovich prescription used below.

In terms of a specific stochastic path $\eta$, a solution to Eq. (A1), $M_\eta^G(t; M_0, t_0)$ with $M_\eta^G(t_0; M_0, t_0) \equiv M_0$, the initial condition on the probability distribution of $M_\eta$ is

$$P_\eta[M, t|M_0, t_0] = \delta(M - M_\eta(t; M_0, t_0))\ . \tag{A4}$$

Using the conservation of probability condition,

$$P_{\eta,t} + (\dot{M}^G P_\eta)_{,G} = 0\ ,$$

$$(\cdots)_{,G} = \partial(\cdots)/\partial M^G\ ,$$

$$(\cdots)_{,t} = \partial(\cdots)/\partial t\ , \tag{A5}$$

the evolution of $P_\eta$ is written as

$$P_{\eta,t}[M, t|M_0, t_0] = \{\ [-f^G(t, M) - \hat{g}(t, M)\eta^i]P_\eta\ \}_{,G}\ . \tag{A6}$$

To perform the stochastic average of Eq. (A6), the "functional integration by parts lemma" [29] is used on an arbitrary function $Z(\eta)$ [181],

$$\int D\eta\ \frac{\hat{\delta} Z(\eta)}{\hat{\delta}\eta^i} = 0\ . \tag{A7}$$

Applied to $Z = Z' \exp(-\frac{1}{2}\int_{t_0}^{\infty} dt \eta^i \eta^i)$, this yields

$$< \eta^i Z' >_\eta = < \delta Z'/\delta\eta^i >_\eta\ . \tag{A8}$$

Applying this to $\hat{F}(M_\eta) = \int dM\ P_\eta F(M)$,

$$\int dM\ \frac{\hat{\delta} P_\eta}{\hat{\delta}\eta^i}\ F(M) = \frac{\partial \hat{F}(M_\eta)}{\partial M_\eta^G}\ \frac{\hat{\delta} M_\eta^G}{\hat{\delta}\eta^i}$$

$$= -\frac{1}{2}\int dM\ F(M)(\hat{g}_j^G \delta^{ij} P^\eta)_{,G}\ , \tag{A9}$$

where $\hat{\delta}$ designates functional differentiation. The last equation has used the Stratonovich prescription,

$$M_\eta^G(t) = M_0^G + \int dt' \hat{H}(t-t')\hat{H}(t-t_0)(f^G + \hat{g}_i^G \eta^i)\ ,$$

$$\lim_{t' \to t-0} \frac{\hat{\delta} M_\eta^G(t)}{\hat{\delta}\eta^i(t')} = \frac{1}{2}\hat{g}_j^G[t, M_\eta(t)]\delta_{ij}\ ,$$

$$\hat{H}(z) = \begin{cases} 1, & z \ge 0 \\ 0, & z < 0\ . \end{cases} \tag{A10}$$

Taking the averages $< P_{\eta,t} >_\eta$ and $< \eta^i P_\eta >_\eta$, the Fokker-Planck is obtained from Eq. (A9). If some boundary conditions are added as Lagrange multipliers, these enter as a "potential" $V$, creating a Schrödinger-type equation:

$$P_{,t} = \frac{1}{2}(g^{GG'}P)_{,GG'} - (g^G P)_{,G} + VP\ ,$$

$$P = < P_\eta >_\eta\ ,$$

$$g^G = f^G + \frac{1}{2}\hat{g}_i^{G'}\hat{g}_{i,G'}^G ,$$

$$g^{GG'} = \hat{g}_i^G \hat{g}_i^{G'} ,$$

$$(\cdots)_{,G} = \partial(\cdots)/\partial M^G . \qquad \text{(A11)}$$

Note that $g^G$ replaces $f^G$ in Eq. (A1) if the Itô (prepoint discretized) calculus is used to define that equation.

To derive the path-integral representation of Eq. (A11), define operators $\hat{M}^G$, $\hat{p}_G$, and $\hat{H}$,

$$[\hat{M}^G, \hat{p}_{G'}] \equiv \hat{M}^G \hat{p}_{G'} - \hat{p}_{G'}\hat{M}^G = i\delta_{G'}^G ,$$

$$[\hat{M}^G, \hat{M}^{G'}] = 0 = [\hat{p}_G, \hat{p}_{G'}] ,$$

$$P_{,t} = -i\hat{H}P ,$$

$$\hat{H} = -\frac{i}{2}\hat{p}_G\hat{p}_{G'}g^{GG'} + \hat{p}_G g^G + iV , \qquad \text{(A12)}$$

and define the evolution operator $U(t, t')$ in terms of "bra" and "ket" probability states of $M$,

$$\hat{M}^G|M^G> = M^G|M^G> ,$$

$$\hat{p}_G|M^G> = -i\partial/\partial M^G|M^G> ,$$

$$<M'|M> = \delta(M' - M) ,$$

$$<M|p> = (2\pi)^{-1}\exp(ip \cdot M) ,$$

$$P[M, t|M_0, t_0] = <M|U(t, t_0)|M_0> ,$$

$$\hat{H}(t')U(t', t) = iU(t', t)_{,t'} ,$$

$$U(t, t) = 1 ,$$

$$U(t_\rho, t_{\rho-1}) \approx 1 - i\theta\hat{H}(t_{\rho-1}) , \qquad \text{(A13)}$$

where $\rho$ indexes units of $\theta$ measuring the time evolution. This is formally integrated to give the path integral in the phase space $(p, M)$,

$$P[M_t|M_0] = \int_{M(t_0)=M_0}^{M(t)=M_t} DM\ Dp\ \exp[\int_{t_0}^{t} dt'(ip_G M^G - \frac{1}{2}p_G p_{G'}g^{GG'} - ip_G g^G + V)] ,$$

$$DM = \lim_{u\to\infty}\prod_G\prod_{\rho=1}^{u} dM_\rho^G ,$$

$$Dp = \lim_{u\to\infty}\prod_G\prod_{\rho=1}^{u+1}(2\pi)^{-1}dp_{G\rho} ,$$

$$t_\rho = t_0 + \rho\theta . \qquad \text{(A14)}$$

The integral over each $dp_{G\rho}$ is a Gaussian and simply calculated. This gives the path integral in coordinate space $M$, in terms of the prepoint discretized Lagrangian,

$$P[M_t|M_0] = \int DM \prod_{\rho=0}^{u} (2\pi\theta)^{-\Lambda/2} g(M_\rho, t_\rho)^{1/2}$$

$$\times \exp \left\{ -\frac{1}{2} \theta g_{GG'}(M_\rho, t_\rho)[\Delta^G_\rho/\theta - g^G(M_\rho, t_\rho)] \right.$$

$$\left. \times[\Delta^{G'}_\rho/\theta - g^{G'}(M_\rho, t_\rho)] + \theta V(M_\rho, t_\rho) \right\} ,$$

$$L_I(\dot{M}^G, M^G, t) = \frac{1}{2} (\dot{M}^G - g^G) g_{GG'} (\dot{M}^{G'} - g^{G'}) - V ,$$

$$g = \det(g_{GG'}) ,$$

$$g_{GG'} = (g^{GG'})^{-1} ,$$

$$\Delta^G_\rho = M^G_{\rho+1} - M^G_\rho . \qquad \text{(A15)}$$

This can be transformed to the Stratonovich representation, in terms of the Feynman Lagrangian $L$ possessing a covariant variational principle,

$$P[M_t|M_0] = \int DM \prod_{\rho=0}^{u} (2\pi\theta)^{-\Lambda/2} g(M_\rho + \Delta_\rho, t_\rho + \theta/2)^{1/2}$$

$$\times \exp\left( -\min \int_{t_\rho}^{t_\rho+\theta} dt' L(M(t'), \dot{M}(t'), t') \right), \qquad \text{(A16)}$$

where "min" specifies that Eq. (A11) is obtained by constraining $L$ to be expanded about that $M(t)$ which makes the action $S = \int dt' L$ stationary for $M(t_\rho) = M_\rho$ and $M(t_\rho + \theta) = M_{\rho+1}$.

One way of proceeding is to expand Eq. (A15) and compare to Eq. (A16), but it is somewhat easier to expand Eq. (A16) and compare to Eq. (A15) [182]. It can be shown that expansions to order $\theta$ suffice, and that $\Delta^2 = O(\theta)$.

Write $L$ in the general form

$$L = \frac{1}{2} g_{GG'} \dot{M}^G \dot{M}^{G'} - h_G \dot{M}^G + b$$

$$= L^0 + \Delta L ,$$

$$L^0 = \frac{1}{2} g_{GG'}(M(t), t) \dot{M}^G \dot{M}^{G'} ,$$

$$g_{GG'}(M(t), t) = g_{GG'}(M(t), t') + g_{GG',t'}(M(t), t')(t - t') + O((t - t')^2) , \qquad \text{(A17)}$$

where $h_G$ and $b$ must be determined by comparing expansions of Eq. (A15) and Eq. (A16). Only the $L^0$ term is dependent on the actual $M(t)$ trajectory, and so

$$\int_{t_\rho}^{t_\rho+\theta} dt\ \Delta L = \left(\frac{1}{4} g_{GG',t} \Delta^G \Delta^{G'} - h_G \Delta^G - \frac{1}{2} h_{G,G'} \Delta^G \Delta^{G'} + \theta b\right)\Big|_{(M,t)} , \qquad \text{(A18)}$$

where "$|_{(M,t)}$" implies evaluation at $(M, t)$.

The determinant $g$ is expanded as

$$g(M + \Delta, t + \theta/2)^{1/2} \approx g^{1/2}(M, t) \exp\left(\frac{\theta}{4g} g_{,t} + \frac{1}{2g} \Delta^G g_{,G}\right.$$

$$+ \frac{1}{4g} \Delta^G \Delta^{G'} (g_{,GG'} + g^{-1} g_{,G} g_{,G'}))|_{(M,t)} \ . \tag{A19}$$

The remaining integral over $L^0$ must be performed. This is accomplished using the variational principle applied to $\int L^0$ [180],

$$g_{GH} \ddot{M}^H = -\frac{1}{2} (g_{GH,K} + g_{GK,H} - g_{KH,G}) \dot{M}^K \dot{M}^H \ ,$$

$$\ddot{M}^F = -\Gamma^F_{JK} \dot{M}^J \dot{M}^K \ ,$$

$$\Gamma^F_{JK} \equiv g^{LF} [JK, L] = g^{LF} (g_{JL,K} + g_{KL,J} - g_{JK,L}) \ ,$$

$$(\frac{1}{2} g_{GH} \dot{M}^G \dot{M}^H)_{,t} = 0 \ ,$$

$$\int_t^{t+\theta} L^0 dt \approx \frac{\theta}{2} g_{GH} \dot{M}^G \dot{M}^H |_{(M,t+\theta)} \ . \tag{A20}$$

Differentiating the second equation in Eq. (A20) to obtain $\dddot{M}$, and expanding $\dot{M}(t+\theta)$ to third order in $\theta$,

$$\dot{M}(t+\theta) = (\frac{1}{\theta} \Delta^G - \frac{1}{2\theta} \Gamma^G_{KL} \Delta^K \Delta^L + \frac{1}{6\theta} (\Gamma^G_{KL,N} + \Gamma^G_{AN} \Gamma^A_{KL}) \Delta^G \Delta^L \Delta^N)|_{(M,t)} \ . \tag{A21}$$

Now Eq. (A16) can be expanded as

$$P[M_t | M_0] dM(t) = \int \underline{D}M \prod_{\rho=0}^{u} \exp\left(-\frac{1}{2\theta} g_{GG'}(M,t) \Delta^G \Delta^{G'} + B\right) ,$$

$$\underline{D}M = \prod_{\rho=1}^{u+1} g_\rho^{1/2} \prod_G (2\pi\theta)^{-1/2} dM^G_\rho \ . \tag{A22}$$

Expanding $\exp B$ to $O(\theta)$ requires keeping terms of order $\Delta$, $\Delta^2$, $\Delta^3/\theta$, $\Delta^4/\theta$, and $\Delta^6/\theta^2$. Under the path integral, evaluated at $(M,t)$, and using "$\doteq$" to designate the order of terms obtained from $\int d\Delta \, \Delta^n \exp(-\frac{1}{2\theta} \Delta^2)$,

$$[\Delta^G \Delta^H]_{\text{order}} = \theta g^{GH} \ ,$$

$$[\Delta^G \Delta^H \Delta^K]_{\text{order}} = \theta (\Delta^G g^{HK} + \Delta^H g^{GH} + \Delta^K g^{GH}) \ ,$$

$$[\Delta^G \Delta^H \Delta^A \Delta^B]_{\text{order}} = \theta^2 (g^{GH} g^{AB} + g^{GA} g^{HB} + g^{GB} g^{HA}) \ ,$$

$$[\Delta^A \Delta^B \Delta^C \Delta^D \Delta^E \Delta^F]_{\text{order}} = \theta^3 (g^{AB} g^{CD} g^{EF} + 14 \text{ permutations}) \ . \tag{A23}$$

This expansion of $\exp B$ is to be compared to Eq. (A15), expanded as

$$P[M_t | M_0] dM(t) \approx \int \underline{D}M \prod_{\rho=0}^{u} \exp\left(-\frac{1}{2\theta} g_{GG'} \Delta^G \Delta^{G'}\right)$$

$$\times [1 + g_{GG'} g^G \Delta^{G'} + \theta V + O(\theta^{3/2})] \ , \tag{A24}$$

yielding identification of $h_G$ and $b$ in Eq. (A17),

$$h^G = g^{GG'} h_{G'} = g^G - \frac{1}{2} g^{-1/2} (g^{1/2} g^{GG'})_{,G'} \ ,$$

$$b = \frac{1}{2} h^G h_G + \frac{1}{2} h^G{}_{;G} + R/6 - V \ ,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF} h^G = g^{-1/2} (g^{1/2} h^G)_{,G} \ ,$$

$$R = g^{JL} R_{JL} = g^{JL} g^{FK} R_{FJKL} \ . \tag{A25}$$

The result is

$$P[M_t | M_{t_0}] dM(t) = \int \cdots \int \underset{\sim}{D} M \exp\left( -\min \int_{t_0}^{t} dt' L \right) \delta\big(M(t_0) = M_0\big) \delta\big(M(t) = M_t\big) \ ,$$

$$\underset{\sim}{D} M = \lim_{u \to \infty} \prod_{\rho=1}^{u+1} g^{1/2} \prod_G (2\pi\theta)^{-1/2} dM^G_\rho \ ,$$

$$L(\dot{M}^G, M^G, t) = \frac{1}{2} (\dot{M}^G - h^G) g_{GG'} (\dot{M}^{G'} - h^{G'}) + \frac{1}{2} h^G{}_{;G} + R/6 - V \ ,$$

$$h^G = g^G - \frac{1}{2} g^{-1/2} (g^{1/2} g^{GG'})_{,G'} \ ,$$

$$g_{GG'} = (g^{GG'})^{-1} \ ,$$

$$g = \det(g_{GG'}) \ ,$$

$$h^G{}_{;G} = h^G_{,G} + \Gamma^F_{GF} h^G = g^{-1/2} (g^{1/2} h^G)_{,G} \ ,$$

$$\Gamma^F_{JK} \equiv g^{LF} [JK, L] = g^{LF} (g_{JL,K} + g_{KL,J} - g_{JK,L}) \ ,$$

$$R = g^{JL} R_{JL} = g^{JL} g^{JK} R_{FJKL} \ ,$$

$$R_{FJKL} = \frac{1}{2} (g_{FK,JL} - g_{JK,FL} - g_{FL,JK} + g_{JL,FK}) + g_{MN} (\Gamma^M_{FK} \Gamma^N_{JL} - \Gamma^M_{FL} \Gamma^N_{JK}) \ . \tag{A26}$$

In summary, because of the presence of multiplicative noise, the Langevin system differs in its Itô (prepoint) and Stratonovich (midpoint) discretizations. The midpoint-discretized covariant description, in terms of the Feynman Lagrangian, is defined such that (arbitrary) fluctuations occur about solutions to the Euler-Lagrange variational equations. In contrast, the usual Itô and corresponding Stratonovich discretizations are defined such that the path integral reduces to the Fokker-Planck equation in the weak-noise limit. The term $R/6$ in the Feynman Lagrangian includes a contribution of $R/12$ from the WKB approximation to the same order of $(\Delta t)^{3/2}$ [27].

**REFERENCES**

[1] H. Haken, *Synergetics*, 3rd ed. (Springer, New York, 1983).

[2] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, NY, 1973).

[3] S.-K. Ma, *Statistical Mechanics* (World Scientific, Philadelphia, 1985).

[4] K.G. Wilson and J. Kogurt, "The renormalization group and the $\epsilon$ expansion," *Phys. Rep. C* **12**, 75-200 (1974).

[5] K.G. Wilson, "Problems in physics with many scales of length," *Sci. Am.* **241**, 158-179 (1979).

[6] S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin/Cummings, Reading, MA, 1976).

[7] C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer-Verlag, Berlin, Germany, 1983).

[8] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[9] R. Kubo, K. Matsuo, and K. Kitahara, "Fluctuation and relaxation of macrovariables," *J. Stat. Phys.* **9**, 51-96 (1973).

[10] B. Cartling, "From short-time molecular dynamics to long-time stochastic dynamics of proteins," *J. Chem. Phys.* **91**, 427-438 (1989).

[11] E. Basar, *EEG-Brain Dynamics* (Elsevier, Amsterdam, 1980).

[12] N.S. Goel and N. Richter-Dyn, *Stochastic Models in Biology* (Academic Press, New York, NY, 1974).

[13] L. Ingber, "Towards a unified brain theory," *J. Social Biol. Struct.* **4**, 211-224 (1981).

[14] L. Ingber, "Statistical mechanics of neocortical interactions. I. Basic formulation," *Physica D* **5**, 83-107 (1982).

[15] L. Ingber, "Statistical mechanics of neocortical interactions. Dynamics of synaptic modification," *Phys. Rev. A* **28**, 395-416 (1983).

[16] L. Ingber, "Statistical mechanics of neocortical interactions. Derivation of short-term-memory capacity," *Phys. Rev. A* **29**, 3346-3358 (1984).

[17] L. Ingber, "Statistical mechanics of neocortical interactions. EEG dispersion relations," *IEEE Trans. Biomed. Eng.* **32**, 91-94 (1985).

[18] L. Ingber, "Statistical mechanics of neocortical interactions: Stability and duration of the 7±2 rule of short-term-memory capacity," *Phys. Rev. A* **31**, 1183-1186 (1985).

[19] L. Ingber, "Towards clinical applications of statistical mechanics of neocortical interactions," *Innov. Tech. Biol. Med.* **6**, 753-758 (1985).

[20] L. Ingber, "Statistical mechanics of neocortical interactions," *Bull. Am. Phys. Soc.* **31**, 868 (1986).

[21] L. Ingber, "Applications of biological intelligence to Command, Control and Communications," in *Computer Simulation in Brain Science: Proceedings, University of Copenhagen, 20-22 August 1986*, ed. by R. Cotterill (Cambridge University Press, London, 1988), p. 513-533.

[22] L. Ingber, "Statistical mechanics of mesoscales in neocortex and in command, control and communications ($C^3$): Proceedings, Sixth International Conference, St. Louis, MO, 4-7 August 1987," *Mathl. Comput. Modelling* **11**, 457-463 (1988).

[23] L. Ingber, "Mesoscales in neocortex and in command, control and communications ($C^3$) systems," in *Systems with Learning and Memory Abilities: Proceedings, University of Paris 15-19 June 1987*, ed. by J. Delacour and J.C.S. Levy (Elsevier, Amsterdam, 1988), p. 387-409.

[24] L. Ingber and P.L. Nunez, "Multiple scales of statistical physics of neocortex: Application to electroencephalography," *Mathl. Comput. Modelling* **13**, 83-95 (1990).

[25] L. Ingber, "Statistical mechanics of multiple scales of neocortical interactions," in *Neocortical Dynamics and Human EEG Rhythms*, ed. by P.L. Nunez (Oxford University Press, New York, NY,

1995), p. 628-681.

[26] B.S. DeWitt, "Dynamical theory in curved spaces. I. A review of the classical and quantum action principles," *Rev. Mod. Phys.* **29**, 377-397 (1957).

[27] F. Langouche, D. Roekaerts, and E. Tirapegui, *Functional Integration and Semiclassical Expansions* (Reidel, Dordrecht, The Netherlands, 1982).

[28] R.P. Feynman, R.B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison Wesley, Reading, MA, 1963).

[29] L.S. Schulman, *Techniques and Applications of Path Integration* (J. Wiley & Sons, New York, 1981).

[30] R. Graham, "Covariant formulation of non-equilibrium statistical thermodynamics," *Z. Physik* **B26**, 397-405 (1977).

[31] R. Graham, "Lagrangian for diffusion in curved phase space," *Phys. Rev. Lett.* **38**, 51-53 (1977).

[32] L. Ingber, "Riemannian corrections to velocity-dependent nuclear forces," *Phys. Rev. C* **28**, 2536-2539 (1983).

[33] L. Ingber, "Path-integral Riemannian contributions to nuclear Schrödinger equation," *Phys. Rev. D* **29**, 1171-1174 (1984).

[34] L. Ingber, "Riemannian contributions to short-ranged velocity-dependent nucleon-nucleon interactions," *Phys. Rev. D* **33**, 3781-3784 (1986).

[35] B.G. Cragg and H.N.V. Temperley, "The organization of neurons: A cooperative analogy," *Electroencephal. clin. Neurophysiol.* **6**, 85-92 (1954).

[36] L. Ingber, H. Fujio, and M.F. Wehner, "Mathematical comparison of combat computer models to exercise data," *Mathl. Comput. Modelling* **15**, 65-90 (1991).

[37] L. Ingber, "Statistical mechanics of combat and extensions," in *Toward a Science of Command, Control, and Communications*, ed. by C. Jones (American Institute of Aeronautics and Astronautics, Washington, D.C., 1993), p. 117-149.

[38] L. Ingber, "Statistical mechanical measures of performance of combat," in *Proceedings of the 1991 Summer Computer Simulation Conference 22-24 July 1991, Baltimore, MD*, ed. by D. Pace (Society for Computer Simulation, San Diego, CA, 1991), p. 940-945.

[39] W. Weidlich and G. Haag, *Concepts and Models of a Quantitative Sociology* (Springer, Berlin, 1983).

[40] L. Ingber, "Statistical mechanics of nonlinear nonequilibrium financial markets," *Math. Modelling* **5**, 343-361 (1984).

[41] L. Ingber, M.F. Wehner, G.M. Jabbour, and T.M. Barnhill, "Application of statistical mechanics methodology to term-structure bond-pricing models," *Mathl. Comput. Modelling* **15**, 77-98 (1991).

[42] L. Ingber, "Statistical mechanical aids to calculating term structure models," *Phys. Rev. A* **42**, 7057-7064 (1990).

[43] T.H. Bullock, "Reassessment of neural connectivity and its specification," in *Information Processing in the Nervous System*, ed. by H.M. Pinsker and W.D. Willis, Jr. (Raven Press, New York, NY, 1980).

[44] P.L. Nunez, "Generation of human EEG rhythms by a combination of long and short-range neocortical interactions," *Brain Topography* **1**, 199-215 (1989).

[45] V. Braitenberg, *Cortical Architectonics: General and Areal* (Raven, New York, 1978).

[46] P.L. Nunez, *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford University Press, London, 1981).

[47] A. Van Rotterdam, F.H. Lopes da Silva, J. van der Ende, M.A. Viergever, and A.J. Hermans, "A model of the spatial-temporal characteristics of the alpha rhythm," *Bull. Math. Biol.* **44**, 283-305 (1982).

[48] P.L. Nunez, "A study of origins of the time dependencies of scalp EEG: I—Theoretical basis," *IEEE Trans. Biomed. Eng.* **28**, 271-280 (1981).

[49] P.L. Nunez, "A study of origins of the time dependencies of scalp EEG: II—Experimental support of theory," *IEEE Trans. Biomed. Eng.* **28**, 281-288 (1981).

[50] S.J. Williamson and L. Kaufman, "Evolution of neuromagnetic topographic mapping," *Brain Topography* **3**, 113-127 (1990).

[51] P.L. Nunez, "Localization of brain activity with Electroencephalography," in *Advances in Neurology, Vol. 54: Magnetoencephalography*, ed. by S. Sato (Raven Press, New York, NY, 1990), p. 39-65.

[52] P.L. Nunez, "The brain wave equation: A model for the EEG," *Math. Biosci.* **21**, 279-297 (1974).

[53] P.L. Nunez, "Global contributions to cortical dynamics: theoretical and experimental evidence for standing wave phenomena," in *Dynamics of Sensory and Cognitive Processing by the Brain*, ed. by E. Basar (Springer-Verlag, New York, 1988), p. 173-182.

[54] F.H. Lopes da Silva, A. Van Rotterdam, W. Storm van Leeuwen, and A.M. Tieler, "Dynamic characteristics of visual evoked potentials in the dog. II. Beta frequency selectivity in evoked potentials and background activity," *Electroencephal. clin. Neurophysiol.* **29**, 260-268 (1970).

[55] J.J. Wright, R.R. Kydd, and G.J. Lees, "Gross electrocortical activity as a linear wave phenomena, with variable temporal damping," in *Dynamics of Sensory and Cognitive Processing by the Brain*, ed. by E. Basar (Springer-Verlag, New York, 1988), p. 183-195.

[56] L.H. Van der Tweel and H.F.E. Lunel, "Human visual responses to sinusoidally modulated light," *Electroencephal. clin. Neurophysiol.* **18**, 587-598 (1965).

[57] N.B. Tufillaro, "Nonlinear and chaotic string vibrations," *Am. J. Phys.* **57**, 408-414 (1989).

[58] L. Ingber, "Very fast simulated re-annealing," *Mathl. Comput. Modelling* **12**, 967-973 (1989).

[59] V.B. Mountcastle, "An organizing principle for cerebral function: The unit module and the distributed system," in *The Mindful Brain*, ed. by G.M. Edelman and V.B. Mountcastle (Massachusetts Institute of Technology, Cambridge, 1978), p. 7-50.

[60] K.A. Fitzpatrick and T.J. Imig, "Auditory cortico-cortical connections in the owl monkey," *J. Comp. Neurol.* **192**, 589-601 (1980).

[61] P.S. Goldman and W.J.H. Nauta, "Columnar distribution of cortico-cortical fibers in the frontal association, limbic, and motor cortex of the developing rhesus monkey," *Brain Res.* **122**, 393-413 (1977).

[62] D.H. Hubel and T.N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.* **160**, 106-154 (1962).

[63] D.H. Hubel and T.N. Wiesel, "Functional architecture of Macaque monkey visual cortex," *Proc. R. Soc. London Ser. B* **198**, 1-59 (1977).

[64] E.G. Jones, J.D. Coulter, and S.H.C. Hendry, "Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys," *J. Comp. Neurol.* **181**, 291-348 (1978).

[65] L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskii, *Statistical Physics* (Pergamon, New York, NY, 1980).

[66] H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems* (Springer, Berlin, 1988).

[67] J. Szentágothai, "The 'module-concept' in cerebral cortex architecture," *Brain Res.* **95**, 475-496 (1975).

[68] J. Szentágothai and M.A. Arbib, "Conceptual models of neural organization," *Neurosci. Res. Bull.* **12**, 307-510 (1974).

[69] A.K. Katchalsky, V. Rowland, and R. Blumenthal, "Dynamic patterns of brain assemblies," *Neurosci. Res. Prog. Bull.* **12**, 1-187 (1974).

[70] A.K. Afifi and R.A. Bergman, *Basic Neuroscience* (Urban and Schwarzenberg, Baltimore, MD, 1978).

[71] G. Sommerhoff, *Logic of the Living Brain* (Wiley, New York, NY, 1974).

[72] G.M. Shepherd, *The Synaptic Organization of the Brain*, 2nd ed. (Oxford University Press, New York, NY, 1979).

[73] R.K. Dismukes, "New concepts of molecular communication among neurons," *Behav. Brain Sci.* **2**, 409-448 (1979).

[74] P.C. McGeer, J.C. Eccles, and E.G. McGeer, *Molecular Neurobiology of the Mammalian Brain* (Plenum, New York, NY, 1978).

[75] F.O. Schmitt, P. Dev, and B.H. Smith, "Electrotonic processing of information by brain cells," *Science* **193**, 114-120 (1976).

[76] J.S. Griffith, *Mathematical Neurobiology* (Academic Press, New York, NY, 1971).

[77] W.S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.* **12**, 207-219 (1943).

[78] W.R. Adey, "Long-range electromagnetic field interactions at brain cell surfaces," in *Magnetic Field Effect on Biological Systems*, (Plenum, New York, NY, 1978).

[79] E.R. John, "Switchboard versus statistical theories of learning and memory," *Science* **177**, 850-864 (1972).

[80] S.J. Williamson, L. Kaufman, and D. Brenner, "Evoked neuromagnetic fields of the human brain," *J. Appl. Phys.* **50**, 2418-2421 (1979).

[81] A. Babloyantz and L.K. Kaczmarek, "Self-organization in biological systems with multiple cellular contacts," *Bull. Math. Biol.* **41**, 193-201 (1979).

[82] W.J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, NY, 1975).

[83] L.K. Kaczmarek and A. Babloyantz, "Spatiotemporal patterns in epileptic seizures," *Biol. Cybernetics* **26**, 199-208 (1977).

[84] H.R. Wilson, "Hysteresis in binocular grating perception: Contrast effects," *Vision Res.* **17**, 843-851 (1977).

[85] H.R. Wilson and J.D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophys. J.* **12**, 1-23 (1972).

[86] H.R. Wilson and J.D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetik* **13**, 55-80 (1973).

[87] D.O. Hebb, *The Organisation of Behavior* (Wiley, New York, NY, 1949).

[88] A. Takeuchi and S. Amari, "Formation of topographic maps and columnar microstructures in nerve fields," *Biol. Cybernetics* **35**, 63-72 (1979).

[89] Ch. von der Marlsburg, "Development of ocularity domains and growth behavior of axon terminals," *Biol. Cybernetics* **32**, 49-62 (1979).

[90] N.G. van Kampen, "Fluctuations in closed and open non-linear systems," in *Statistical Physics*, ed. by L. Pál and P. Szépfalusy (North-Holland, Amsterdam, The Netherlands, 1976), p. 29-48.

[91] A. Caillé, D. Pink, F. de Verteuil, and M.J. Zuckermann, "Theoretical models for quasi-two-dimensional mesomorphic monolayers and membrane bilayers," *Can. J. Phys.* **58**, 1723-1726 (1980).

[92] A.C. Scott, "The electrophysics of a nerve fiber," *Rev. Mod. Phys.* **47**, 487-533 (1975).

[93] I. von der Heydt, N. von der Heydt, and G.M. Obermair, "Statistical model of current-coupled ion channels in nerve membranes," *Z. Physik* **41**, 153-164 (1981).

[94] I.A. Boyd and A.R. Martin, "The end-plate potential in mammalian muscle," *J. Physiol. (London)* **132**, 74-91 (1956).

[95] B. Katz, *Nerve, Muscle, and Synapse* (McGraw-Hill, new York, NY, 1966).

[96] H. Korn, A. Mallet, and D.S. Faber, ‘‘Fluctuating responses at a central synapse: *n* of binomial fit predicts number of stained presynaptic boutons,’’ *Science* **213**, 898-900 (1981).

[97] D.H. Perkel and M.W. Feldman, ‘‘Neurotransmitter release statistics: Moment estimates for inhomogeneous Bernoulli trials,’’ *J. Math. Biol.* **7**, 31-40 (1979).

[98] H. Korn and A. Mallet, ‘‘Transformation of binomial input by the postsynaptic membrane at a central synapse,’’ *Science* **225**, 1157-1159 (1984).

[99] W.A. Little, ‘‘The existence of persistent states in the brain,’’ *Math. Biosci.* **19**, 101-120 (1974).

[100] W.A. Little and G.L. Shaw, ‘‘Analytic study of the memory storage capacity of a neural network,’’ *Math. Biosci.* **39**, 281-290 (1978).

[101] G.L. Shaw and R. Vasudevan, ‘‘Persistent states of neural networks and the random nature of synaptic transmission,’’ *Math. Biosci.* **21**, 207-218 (1974).

[102] J. Mathews and R.L. Walker, *Mathematical Methods of Physics, 2nd ed.* (Benjamin, New York, NY, 1970).

[103] C.D. Gilbert and T.N. Wiesel, ‘‘Functional organization of the visual cortex,’’ *Prog. Brain Res.* **58**, 209-218 (1983).

[104] T.J. Imig and R.A. Reale, ‘‘Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex,’’ *J. Comp. Neurol.* **192**, 293-332 (1980).

[105] R.W. Dykes, ‘‘Parallel processing of somatosensory information: A theory,’’ *Brain Res. Rev.* **6**, 47-115 (1983).

[106] R.P Erickson, ‘‘The across-fiber pattern theory: An organizing principle for molar neural function,’’ *Sensory Physiol.* **6**, 79-110 (1982).

[107] R.W. Dykes, ‘‘The anatomy and physiology of the somatic sensory cortical regions,’’ *Prog. Neurobiol.* **10**, 33-88 (1978).

[108] P.C. Hohenberg and B.I. Halperin, ‘‘Theory of dynamic critical phenomena,’’ *Rev. Mod. Phys.* **49**, 435-479 (1977).

[109] J. Szentágothai, ‘‘The neural network of the cerebral cortex: A functional interpretation,’’ *Proc. Roy. Soc. Lon.* **201**, 219-248 (1978).

[110] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[111] N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields* (Wiley, New York, NY, 1980).

[112] U. Deininghaus and R. Graham, ‘‘Nonlinear point transformations and covariant interpretation of path integrals,’’ *Z. Physik B* **34**, 211-219 (1979).

[113] A. Pellionisz and R. Llinás, ‘‘Brain modeling by tensor network theory and computer simulation.

The cerebellum: distributed processor for predictive coordination,'' *Neuroscience* **4**, 323-348 (1979).

[114] A. Pellionisz and R. Llinás, ''Tensorial approach to the geometry of brain function: cerebellar coordination via a metric tensor,'' *Neuroscience* **5**, 1125-1136 (1980).

[115] A.J. Pellionisz, ''Coordination: a vector-matrix description of transformations of overcomplete CNS coordinates and a tensorial solution using the Moore-Penrose generalized inverse,'' *J. Theor. Biol.* **110**, 353-375 (1984).

[116] A. Einstein, A. Lorentz, H. Minkowski, and H. Weyl, *The Principle of Relativity* (Dover, New York, 1923).

[117] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[118] S. Weinberg, *Gravitational Cosmology* (Wiley, New York, 1972).

[119] M.A. Arbib and S.-I. Amari, ''Sensori-motor transformations in the brain (with a critique of the tensor theory of cerebellum),'' *J. Theor. Biol.* **112**, 123-155 (1985).

[120] H. Dekker, ''Quantization in curved spaces,'' in *Functional Integration: Theory and Applications*, ed. by J.P. Antoine and E. Tirapegui (Plenum, New York, 1980), p. 207-224.

[121] H. Grabert and M.S. Green, ''Fluctuations and nonlinear irreversible processes,'' *Phys. Rev. A* **19**, 1747-1756 (1979).

[122] R. Graham, D. Roekaerts, and T. Tél, ''Integrability of Hamiltonians associated with Fokker-Planck equations,'' *Phys. Rev. A* **31**, 3364-3375 (1985).

[123] M.F. Wehner and W.G. Wolfer, ''Numerical evaluation of path-integral solutions to Fokker-Planck equations. I.,'' *Phys. Rev. A* **27**, 2663-2670 (1983).

[124] M.F. Wehner and W.G. Wolfer, ''Numerical evaluation of path-integral solutions to Fokker-Planck equations. II. Restricted stochastic processes,'' *Phys. Rev. A* **28**, 3003-3011 (1983).

[125] M.F. Wehner and W.G. Wolfer, ''Numerical evaluation of path integral solutions to Fokker-Planck equations. III. Time and functionally dependent coefficients,'' *Phys. Rev. A* **35**, 1795-1801 (1987).

[126] L. Ingber and D.D. Sworder, ''Statistical mechanics of combat with human factors,'' *Mathl. Comput. Modelling* **15**, 99-127 (1991).

[127] L. Ingber, ''Statistical mechanics algorithm for response to targets (SMART),'' in *Workshop on Uncertainty and Probability in Artificial Intelligence: UC Los Angeles, 14-16 August 1985*, ed. by P. Cheeseman (American Association for Artificial Intelligence, Menlo Park, CA, 1985), p. 258-264.

[128] W. Gersch, ''Non-stationary multichannel time series analysis,'' in *Methods of Brain Electrical and Magnetic Signals. EEG Handbook*, ed. by A.S. Gevins and A. Remond (Elsevier, New York, NY, 1987), p. 261-296.

[129]
N.J.I. Mars and F.H. Lopes da Silva, ‘‘EEG analysis methods based on information theory,’’ in *Methods of Brain Electrical and Magnetic Signals. EEG Handbook*, ed. by A.S. Gevins and A. Remond (Elsevier, New York, NY, 1987), p. 297-307.

[130]
R. Graham, ‘‘Path-integral methods on nonequilibrium thermodynamics and statistics,’’ in *Stochastic Processes in Nonequilibrium Systems*, ed. by L. Garrido, P. Seglar and P.J. Shepherd (Springer, New York, NY, 1978), p. 82-138.

[131]
H. Goldstein, *Classical Mechanics, 2nd ed.* (Addison Wesley, Reading, MA, 1980).

[132]
E.S. Abers and B.W. Lee, ‘‘Gauge theories,’’ *Phys. Rep. C* **9**, 1-141 (1973).

[133]
D.J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena* (McGraw-Hill, New York, NY, 1978).

[134]
R. Pool, ‘‘Is it chaos, or is it just noise?,’’ *Science* **243**, 25-28 (1989).

[135]
P. Grassberger, ‘‘Do climatic attractors exist?,’’ *Nature* **323**, 609-612 (1986).

[136]
W.A. Brock, ‘‘Distinguishing random and deterministic systems: Abridged version,’’ *J. Econ. Theory* **40**, 168-195 (1986).

[137]
H.D.I. Abarbanel, R. Brown, and J.B. Kadtke, ‘‘Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra,’’ *Phys. Rev. A* **41**, 1782-1807 (1990).

[138]
W.J. Freeman and C.A. Skarda, ‘‘Spatial EEG patterns, nonlinear dynamics and perception: The neo-Sherringtonian view,’’ *Brain Res. Rev.* **10**, 147-175 (1985).

[139]
B. Baird, ‘‘Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb,’’ *Physica D* **22**, 150-175 (1986).

[140]
W.J. Freeman, ‘‘Simulation of chaotic EEG patterns with a dynamic model of the olfactory system,’’ *Biol. Cybern.* **55**, 139-150 (1987).

[141]
P.E. Rapp, T.R. Bashore, J.M. Marinerie, A.M. Albano, I.D. Zimmerman, and A.I. Mees, ‘‘Dynamics of brain electrical activity,’’ *Brain Topography* **2**, 99-118 (1989).

[142]
P. Grassberger and I. Procaccia, ‘‘Measuring the strangeness of strange attractors,’’ *Physica D* **9**, 189-208 (1983).

[143]
W.A. Brock, W.D. Dechert, J.A. Scheinkman, and B. LeBaron, ‘‘A test for the independence based on the correlation dimension,’’ SSRI Working Paper No. 8702, Revised, U. Wisconsin, Madison, WI, 1988.

[144]
D.S. Broomhead and G.P. King, ‘‘Extracting qualitative dynamics from experimental data,’’ *Physica D* **20**, 217-236 (1986).

[145]
L. Ingber, ‘‘Editorial: Learning to learn,’’ *Explore* **7**, 5-8 (1972).

[146]
L. Ingber, *Karate: Kinematics and Dynamics* (Unique, Hollywood, CA, 1981).

[147]
L. Ingber, *Elements of Advanced Karate* (Ohara, Burbank, CA, 1985).

[148]
G.A. Miller, "The magical number seven, plus or minus two," *Psychol. Rev.* **63**, 81-97 (1956).

[149]
K.A. Ericsson and W.G. Chase, "Exceptional memory," *Am. Sci.* **70**, 607-615 (1982).

[150]
G. Zhang and H.A. Simon, "STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses," *Memory & Cognition* **13**, 193-201 (1985).

[151]
L. Ingber, *The Karate Instructor's Handbook* (Physical Studies Institute-Institute for the Study of Attention, Solana Beach, CA, 1976).

[152]
L. Ingber, "Attention, physics and teaching," *J. Social Biol. Struct.* **4**, 225-235 (1981).

[153]
B.B. Murdock, Jr., "A distributed memory model for serial-order information," *Psychol. Rev.* **90**, 316-338 (1983).

[154]
P. Colet, H.S. Wio, and M. San Miguel, "Colored noise: A perspective from a path-integral formalism," *Phys. Rev. A* **39**, 6094-6097 (1989).

[155]
S.R. Shenoy and G.S. Agarwal, "First-passage times and hysteresis in multivariable stochastic processes: The two-mode ring laser," *Phys. Rev. A* **29**, 1315-1325 (1984).

[156]
G.S. Agarwal and S.R. Shenoy, "Observability of hysteresis in first-order equilibrium and nonequilibrium phase transitions," *Phys. Rev. A* **23**, 2719-2723 (1981).

[157]
R. Graham and T. Tél, "Existence of a potential for dissipative dynamical systems," *Phys. Rev. Lett.* **52**, 9-12 (1984).

[158]
B. Libet, "Brain stimulation in the study of neuronal functions for conscious sensory experience," *Human Neurobiol.* **1**, 235-242 (1982).

[159]
V.B. Mountcastle, R.A. Andersen, and B.C. Motter, "The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex," *J. Neurosci.* **1**, 1218-1235 (1981).

[160]
Y. Tsal, "Movements of attention across the visual field," *J. Exp. Psychol.* **9**, 523-530 (1983).

[161]
J.D. Cowan, "Spontaneous symmetry breaking in large scale nervous activity," *Int. J. Quant. Chem.* **22**, 1059-1082 (1982).

[162]
S.-I. Amari, "Field theory of self-organizing neural nets," *IEEE Trans. Syst. Man Cybern.* **13**, 741-748 (1983).

[163]
P. Anninos, S. Zenone, and R. Elul, "Artificial neural nets: Dependence of the EEG amplitude's probability distribution on statistical parameters," *J. Theor. Biol.* **103**, 339-348 (1983).

[164]
W.J. Freeman, "Spatial properties of an EEG event in the olfactory bulb and cortex," *Electroencephal. clin. Neurophysiol.* **44**, 586-605 (1978).

[165]
W.R. Klemm and C.J. Sherry, "Serial ordering in spike trains: What's it 'trying to tell us'?," *Int. J. Neurosci.* **14**, 15-33 (1981).

[166]
M. Steriade, "Mechanisms underlying cortical activation: Neuronal organization and properties of the midbrain reticular core and intralaminar thalamic nuclei," in *Brain Mechanisms and Perceptual Awareness*, ed. by O. Pompeiano and C.A. Marsan (Raven, New York, 1981), p. 327-377.

[167]
R.D. Traub and R. Llinás, "Hippocampal pyramidal cells: Significance of dendritic ionic conductances for neuronal function and epileptogenesis," *J. Neurophysiol.* **42**, 476-496 (1979).

[168]
F.E. Dudek, G.I. Hatton, and B.A. Macvicar, "Intracellular recordings from the paraventricular nucleus in slices of rat hypothalamus," *J. Physiol.* **301**, 101-114 (1980).

[169]
M. Barinaga, "The tide of memory, turning," *Science* **248**, 1603-1605 (1990).

[170]
E.R. Kandel and J.H. Schwartz, *Principles of Neural Science* (Elsevier, New York, NY, 1981).

[171]
A.H. Kawamoto and J.A. Anderson, "A neural model of multistable perception," *Acta Psychologica* **59**, 35-65 (1985).

[172]
C. Holden, "Probing the complex genetics of alcoholism," *Science* **251**, 163-164 (1991).

[173]
B. Porjesz and H. Begleiter, "Event-related potentials individuals at risk for alcoholism," *Alcohol* **7**, 465-469 (1990).

[174]
H. Szu and R. Hartley, "Fast simulated annealing," *Phys. Lett. A* **122**, 157-162 (1987).

[175]
S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, "Optimization by simulated annealing," *Science* **220**, 671-680 (1983).

[176]
K. Maurer, T. Dierks, W.K. Strik, and L. Frölich, "P3 topography in psychiatry and psychopharmacology," *Brain Topography* **3**, 79-84 (1990).

[177]
D.F. Shanno and K.H. Phua, "Minimization of unconstrained multivariate functions," *ACM Trans. Mathl. Software* **2**, 87-94 (1976).

[178]
B. Porjesz and H. Begleiter, "Individuals at risk for alcoholism: Neurophysiologic processes," in *Banbury Report 33: Genetics and Biology of Alcoholism*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, MA, 1990), p. 137-157.

[179]
M. Jammer, *The Philosophy of Quantum Mechanics* (Wiley & Sons, New York, NY, 1974).

[180]
K.S. Cheng, "Quantization of a general dynamical system by Feynman's path integration formulation," *J. Math. Phys.* **13**, 1723-1726 (1972).

[181]
F. Langouche, D. Roekaerts, and E. Tirapegui, ‘‘Discretization problems of functional integrals in phase space,’’ *Phys. Rev. D* **20**, 419-432 (1979).

[182]
F. Langouche, D. Roekaerts, and E. Tirapegui, ‘‘Short derivation of Feynman Lagrangian for general diffusion process,’’ *J. Phys. A* **113**, 449-452 (1980).

**TABLE CAPTIONS**

TABLE I. Means and standard deviations of averages over EEG recordings from 49 subjects, representing 190 points of data per subject, are consistent with the centering mechanism during selective attention tasks. Under each electrode the means and standard deviations of $M^G$ are given. Also given for each electrode are the means and standard deviations of the individual-subject standard deviations, here labeled as $\sigma$, aggregated from each subject. The physical bounds for all $M^E$ under these nonvisual regions are $\pm N^E = \pm 80$. Also given are the weights $d$ of the regional time-delayed contributions $dM^{*E}$. The physical bounds for all $M^E$ and $M^{*E}$ under these nonvisual regions are $\pm N^E = \pm N^{*E} = \pm 80$; the physical bounds for all $M^I$ are $\pm N^I = \pm 30$.

TABLE II. The same kind of information is presented as given in Table I; however, here the statistics is developed from a subgroup of 25 subjects who were previously classified as being of high risk with respect to having a genetic propensity to alcoholism.

TABLE III. The same kind of information is presented as given in Table I; however, here the statistics is developed from a subgroup of 24 subjects who were previously classified as being of low risk with respect to having a genetic propensity to alcoholism.

## FIGURE CAPTIONS

FIG. 1. *Scales illustrated*. Illustrated are three biophysical scales of neocortical interactions: (a), ($a^*$), (a') microscopic neurons; (b) and (b') mesocolumnar domains; (c) and (c') macroscopic regions. In ($a^*$) synaptic interneuronal interactions, averaged over by mesocolumns, are phenomenologically described by the mean and variance of a distribution $\Psi$. Similarly, in (a) intraneuronal transmissions are phenomenologically described by the mean and variance of $\Gamma$. Mesocolumnar averaged excitatory ( $E$) and inhibitory ( $I$) neuronal firings are represented in (a'). In (b) the vertical organization of minicolumns is sketched together with their horizontal stratification, yielding a physiological entity, the mesocolumn. In (b') the overlap of interacting mesocolumns is sketched. In (c) macroscopic regions of the neocortex are depicted as arising from many mesocolumnar domains. These are the regions designated for study here. (c') sketches how regions may be coupled by long-ranged interactions.

FIG. 2. *Minima structure of nonlinear Lagrangian*. Examination of the minima structure of the spatially averaged and temporally averaged Lagrangian provides some quick intuitive details about the most likely states of the system. This is supported by further analysis detailing the actual spatial-temporal minima structure. Illustrated is the surface of the static (time-independent) mesoscopic neocortical Lagrangian $\bar{\underline{L}}$ over the excitatory-inhibitory firing plane ($\bar{M}^E - \bar{M}^I$), for a specific set of synaptic parameters. All points on the surface higher than $5 \times 10^{-3}/\tau$ have been deleted to expose this fine structure.

FIG. 3. *Nearest neighbors*. Nearest-neighbor interactions between mesocolumns are illustrated. Afferent minicolumns of $\sim 10^2$ neurons are represented by the inner circles, and efferent macrocolumns of $\sim 10^5$ neurons by the outer circles. Illustrated are the nearest-neighbor interactions between a mesocolumn, represented by the thick circles, and its nearest neighbors, represented by thin circles. The area outside the outer thick circle represents the effective number of efferent macrocolumnar nearest-neighbor neurons. I.e., these neurons outside the macrocolumnar area of influence of the central minicolumn are contacted through interactions with neurons in the central macrocolumn.

FIG. 4. *Valleys of STM*. Contours of the Lagrangian illustrate "valleys" that trap firing-states of mesocolumns. ( $\tau \bar{\underline{L}}_{\rm BC}$ can be as large as $10^3$.) These valleys are candidates for short-term memories. Detailed calculations support the identification of the inner valleys with stable short-term-memory states having durations on the order of tenths of a second. (a) Contours for values less than 0.04 are drawn for $\tau \bar{\underline{L}}_{\rm BC}$, where BC designates the balanced case of firing states being at a moderate level of excitatory and inhibitory firings. The $\bar{M}^E$ axis increases to the right, from $-N^E = -80$ to $N^E = 80$. The $\bar{M}^I$ axis increases to the right, from $-N^I = -30$ to $N^I = 30$. In each cluster, the smaller values are closer to the center. Note the absence of any closed contours in the interior space. (b) Contours for values less than 0.04 are drawn for $\tau \bar{\underline{L}}_{\rm BC'}$, where BC′ designates that the "centering mechanism" has been turned on. A right brace } signifies enclosure of other nested closed contours above and to the left of this brace.

FIG. 5. *Modeling the visual cortex STM*. When $N = 220$, modeling the number of neurons per minicolumn in the visual neocortex, then only clusters containing 2–4 up to 5–6 minima are found, consistent with visual STM. These minima are narrower, consistent with the sharpness required to store visual patterns.

FIG. 6. *EEG traces*. Illustrated is 1 sec of EEG data, time-locked to a particular stimulus pattern and averaged over about 100 such events, from 5 electrodes, from a subject previously classified as high-risk with respect to genetic propensity to alcoholism. The abscissa labels the index of each epoch; there are 192 points of data for each electrode, and each time epoch represets 5.2 msec. The ordinate is the amplitude of the "averaged epoch potential" in mV. The first dip, here at about epoch 18, is classified as N100, occurring at about 100 msec after a stimulus pattern. The second dip, here at about epoch 30, is classified as N200. The broad peak approximately centered about epoch 80 is the P300.

**Table I**

| Electrode | Mean $M^E$ | Standard Deviation $M^E$ | Mean $\sigma(M^{E)}$ | Standard Deviation $\sigma(M^E)$ |
|---|---|---|---|---|
| Fz | 1.13661 | 3.8754 | 6.09339 | 5.30891 |
| Cz | -0.533493 | 4.83208 | 6.31146 | 5.59003 |
| Pz | -0.3158 | 3.87471 | 5.44242 | 5.50453 |
| P3 | -0.121703 | 10.1069 | 8.152 | 7.08701 |
| P4 | -0.0208276 | 7.47837 | 11.0526 | 7.04522 |

| Electrode | Mean $M^I$ | Standard Deviation $M^I$ | Mean $\sigma(M^{I)}$ | Standard Deviation $\sigma(M^I)$ |
|---|---|---|---|---|
| Fz | 1.83249 | 7.11368 | 11.4131 | 2.71679 |
| Cz | 0.229446 | 5.89307 | 11.5578 | 2.68969 |
| Pz | -0.255393 | 6.37452 | 12.4699 | 2.86198 |
| P3 | -0.0234756 | 7.39736 | 10.5579 | 3.2693 |
| P4 | -0.0271411 | 6.25121 | 12.0525 | 2.52846 |

| Electrode | Mean $d$ | Standard Deviation $d$ |
|---|---|---|
| Fz $\rightarrow$ Cz | 0.389722 | 0.291677 |
| Cz $\rightarrow$ Pz | 0.377305 | 0.266958 |
| Pz $\rightarrow$ P3 | 0.536313 | 0.288519 |
| Pz $\rightarrow$ P4 | 0.485525 | 0.294742 |

**Table II**

| Electrode | Mean $M^E$ | Standard Deviation $M^E$ | Mean $\sigma(M^{E)}$ | Standard Deviation $\sigma(M^E)$ |
|---|---|---|---|---|
| Fz | 0.340677 | 4.44151 | 5.87727 | 4.73137 |
| Cz | -0.0332101 | 4.4386 | 5.79371 | 4.79331 |
| Pz | 0.487081 | 4.02893 | 5.02483 | 5.99425 |
| P3 | -0.882276 | 4.32405 | 5.81738 | 5.59617 |
| P4 | -0.940474 | 8.48643 | 10.6133 | 6.98527 |

| Electrode | Mean $M^I$ | Standard Deviation $M^I$ | Mean $\sigma(M^{I)}$ | Standard Deviation $\sigma(M^I)$ |
|---|---|---|---|---|
| Fz | 0.474634 | 7.11011 | 11.8372 | 2.27895 |
| Cz | 1.23924 | 5.45593 | 11.6565 | 2.96615 |
| Pz | 0.217116 | 6.45096 | 11.5355 | 2.966 |
| P3 | -1.61882 | 7.68062 | 10.8889 | 3.16314 |
| P4 | -0.334122 | 6.39578 | 12.2809 | 2.11318 |

| Electrode | Mean $d$ | Standard Deviation $d$ |
|---|---|---|
| Fz $\rightarrow$ Cz | 0.391295 | 0.322489 |
| Cz $\rightarrow$ Pz | 0.395463 | 0.254323 |
| Pz $\rightarrow$ P3 | 0.526201 | 0.286918 |
| Pz $\rightarrow$ P4 | 0.543161 | 0.282187 |

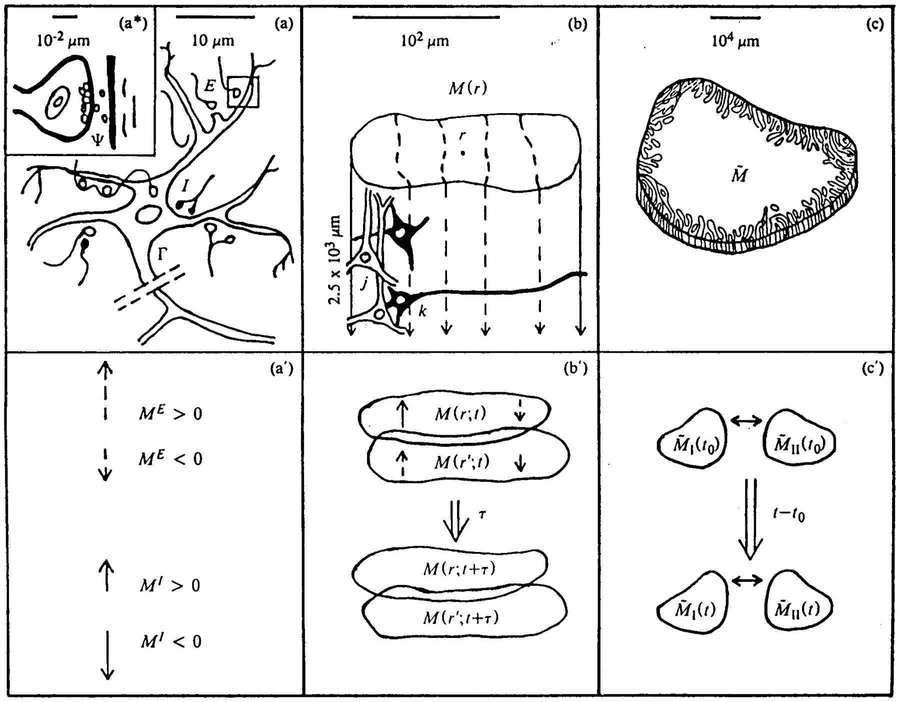
**Table III**

| Electrode | Mean $M^E$ | Standard Deviation $M^E$ | Mean $\sigma(M^{E)}$ | Standard Deviation $\sigma(M^E)$ |
|---|---|---|---|---|
| Fz | 1.9657 | 3.05825 | 6.31853 | 5.94619 |
| Cz | -1.05462 | 5.25502 | 6.85079 | 6.37529 |
| Pz | -1.15213 | 3.60052 | 5.87741 | 5.03529 |
| P3 | 0.670561 | 13.8703 | 10.5839 | 7.74884 |
| P4 | 0.937138 | 6.29979 | 11.5102 | 7.2279 |

| Electrode | Mean $M^I$ | Standard Deviation $M^I$ | Mean $\sigma(M^{I)}$ | Standard Deviation $\sigma(M^I)$ |
|---|---|---|---|---|
| Fz | 3.24693 | 6.98325 | 10.9714 | 3.096 |
| Cz | -0.822421 | 6.25717 | 11.4549 | 2.42812 |
| Pz | -0.747589 | 6.39397 | 13.4433 | 2.44476 |
| P3 | 1.63834 | 6.85532 | 10.2132 | 3.40917 |
| P4 | 0.292631 | 6.2177 | 11.8146 | 2.92689 |

| Electrode | Mean $d$ | Standard Deviation $d$ |
|---|---|---|
| Fz $\rightarrow$ Cz | 0.388083 | 0.26272 |
| Cz $\rightarrow$ Pz | 0.35839 | 0.283736 |
| Pz $\rightarrow$ P3 | 0.546846 | 0.295968 |
| Pz $\rightarrow$ P4 | 0.425487 | 0.301391 |

**Figure 1**



(a*)
$10^{-2}$ μm
Ψ
(a)
10 μm
E
I
Γ
(b)
$10^2$ μm
M(r)
r
2.5 x $10^3$ μm
j
k
(c)
$10^4$ μm
$\bar{M}$
(a′)
$M^E > 0$
$M^E < 0$
$M^I > 0$
$M^I < 0$
(b′)
M(r;t)
M(r′;t)
τ
M(r;t+τ)
M(r′;t+τ)
(c′)
$\bar{M}_I(t_0)$
$\bar{M}_{II}(t_0)$
$t-t_0$
$\bar{M}_I(t)$
$\bar{M}_{II}(t)$

**Figure 2**

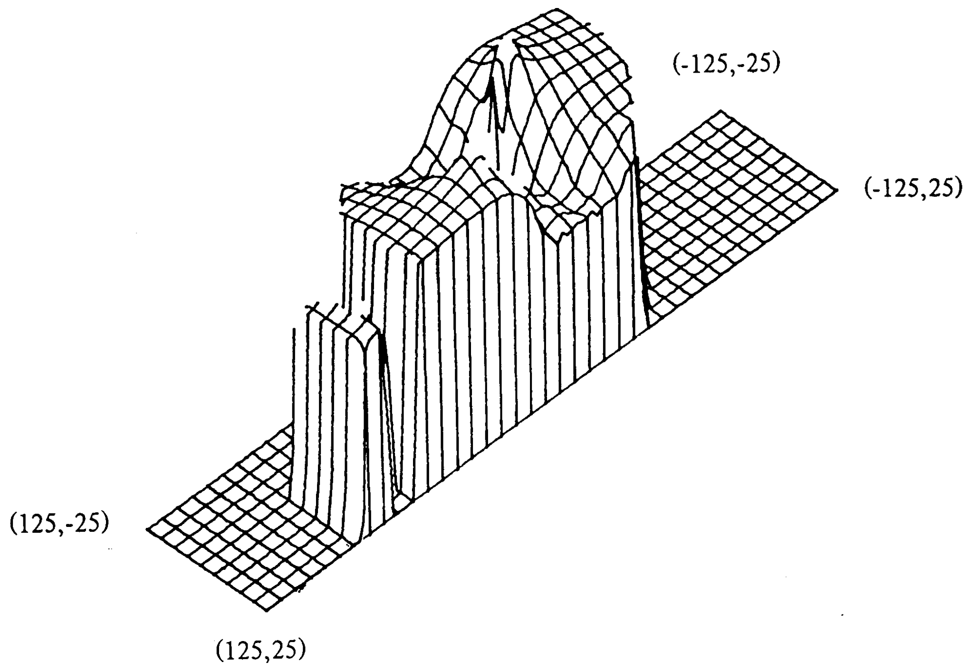


**Figure 3**

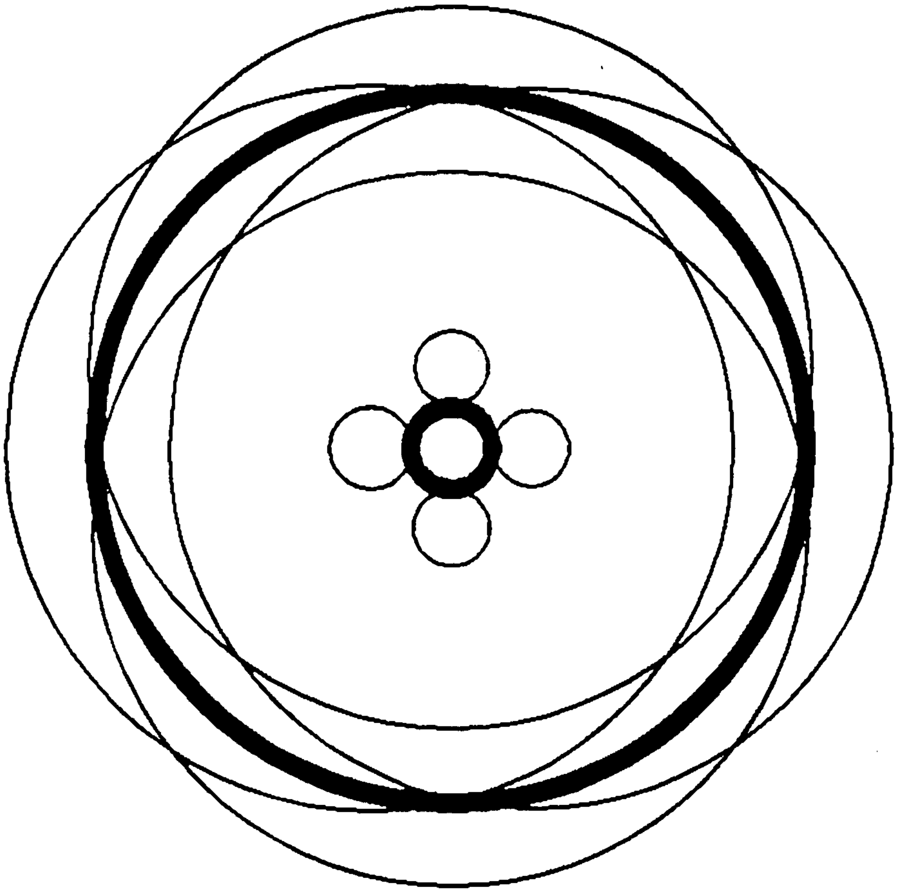


**Figure 4**

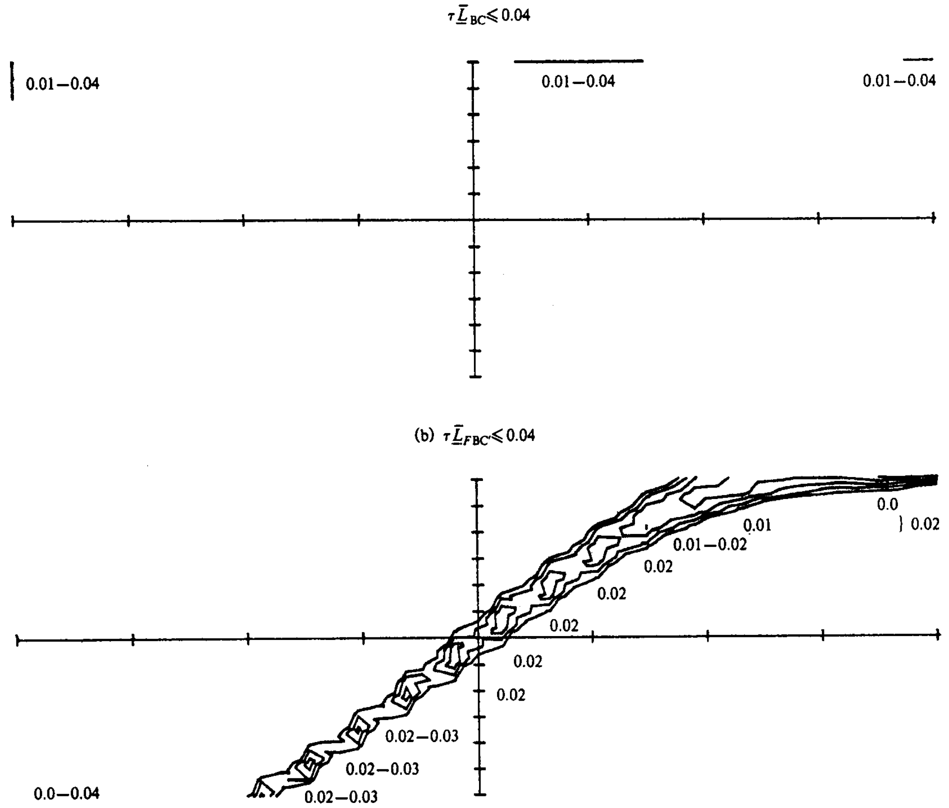

$\tau \underline{\bar{L}}_{BC} \leqslant 0.04$
0.01−0.04
0.01−0.04
0.01−0.04
(b) $\tau \underline{\bar{L}}_{FBC} \leqslant 0.04$
0.0
0.01
} 0.02
0.01−0.02
0.02
0.02
0.02
0.02
0.02
0.02−0.03
0.02−0.03
0.02−0.03
0.0−0.04

**Figure 5**

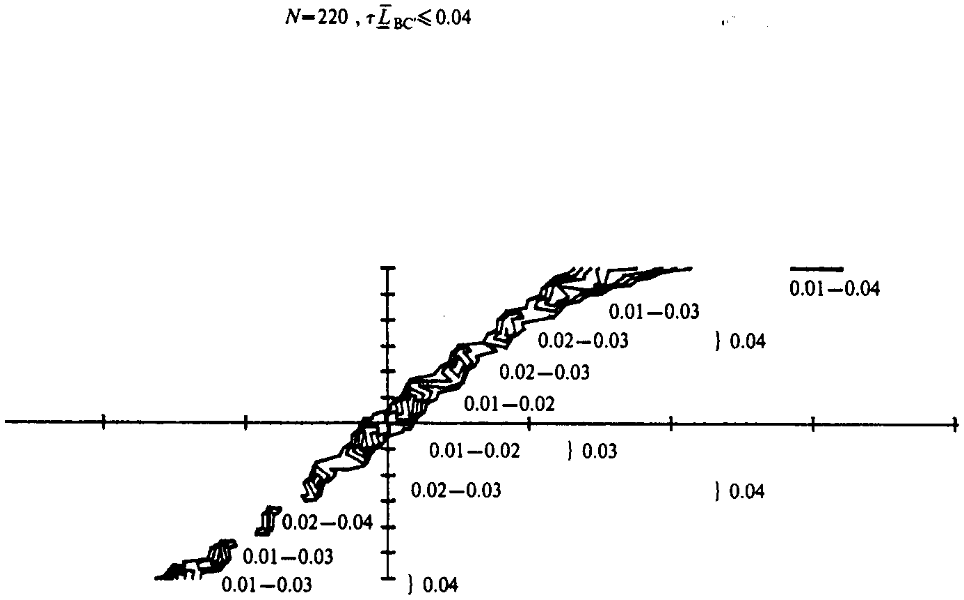
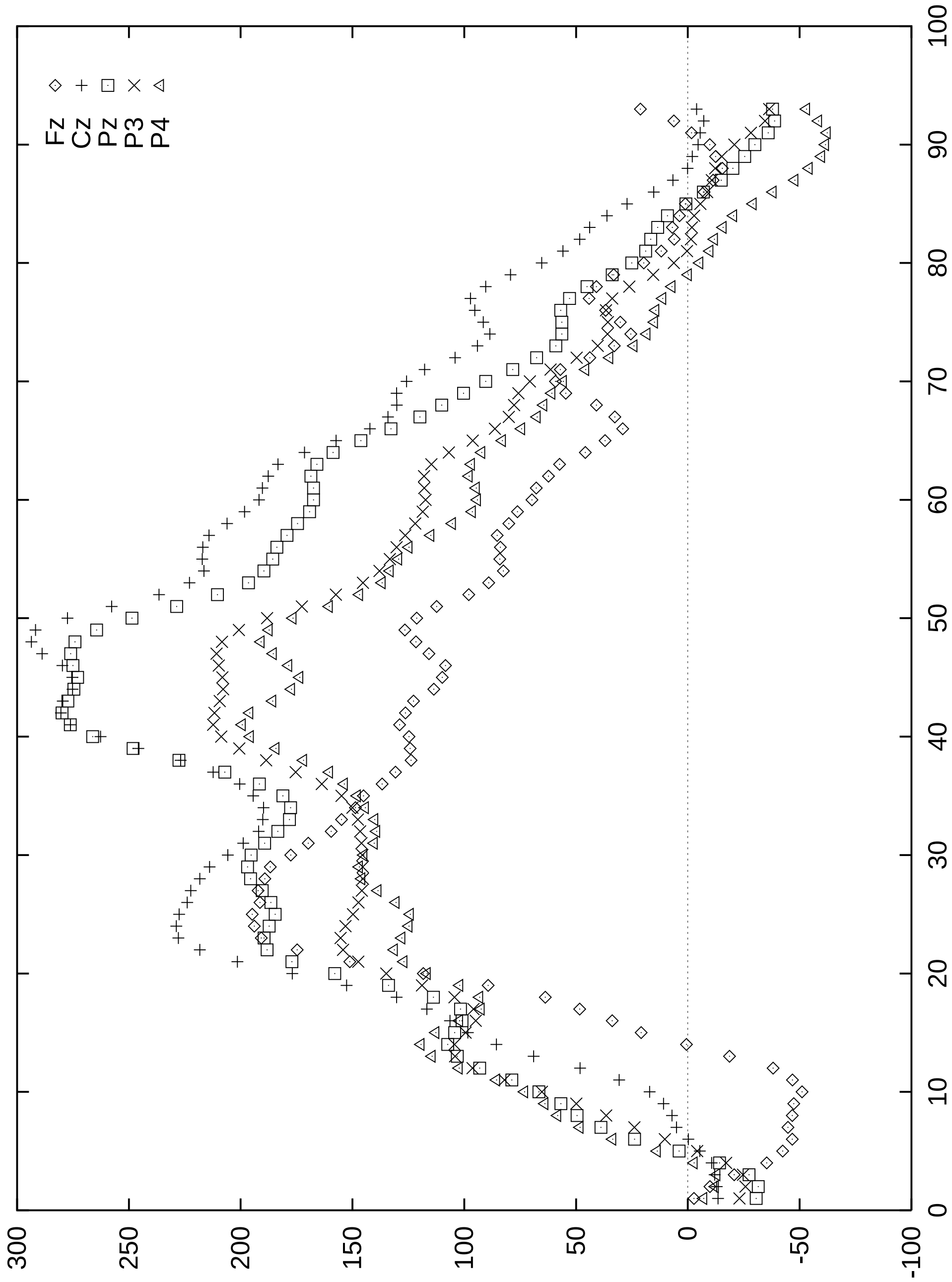

N=220 , τL̲BC' ≤ 0.04
0.01−0.04
0.01−0.03
0.02−0.03
} 0.04
0.02−0.03
0.01−0.02
0.01−0.02
} 0.03
0.02−0.03
} 0.04
0.02−0.04
0.01−0.03
0.01−0.03
} 0.04

**Figure 6**


Fz
Cz
Pz
P3
P4
0
10
20
30
40
50
60
70
80
90
100
300
250
200
150
100
50
0
-50
-100