A New EEG Measure Using the 1-D Cluster Variation Method

Alianna J. Maren[1i] and Harold H. Szu[2]

[1]Northwestern University, Chicago IL, and Themasis Associates, Chicago, IL
[2]Catholic University of America, Washington, DC

ABSTRACT

A new information measure, drawing on the 1-D Cluster Variation Method (CVM), describes local pattern distributions (nearest-neighbor and next-nearest neighbor) in a binary 1-D vector in terms of a single interaction enthalpy parameter h for the specific case where the fractions of elements in each of two states are the same ($x_1=x_2=0.5$). An example application of this method would be for EEG interpretation in Brain-Computer Interfaces (BCIs), especially in the frontier of invariant biometrics based on distinctive and invariant individual responses to stimuli containing an image of a person with whom there is a strong affiliative response (e.g., to a person's grandmother). This measure is obtained by mapping EEG observed configuration variables (z_1, z_2, z_3 for next-nearest neighbor triplets) to h using the analytic function giving h in terms of these variables at equilibrium. This mapping results in a small phase space region of resulting h values, which characterizes local pattern distributions in the source data. The 1-D vector with equal fractions of units in each of the two states can be obtained using the method for transforming natural images into a binarized equi-probability ensemble (Saremi & Sejnowski, 2014; Stephens et al., 2013). An intrinsically 2-D data configuration can be mapped to 1-D using the 1-D Peano-Hilbert space-filling curve, which has demonstrated a 20 dB lower baseline using the method compared with other approaches (cf. SPIE ICA etc. by Hsu & Szu, 2014). This CVM-based method has multiple potential applications; one near-term one is optimizing classification of the EEG signals from a COTS 1-D BCI baseball hat. This can result in a convenient 3-D lab-tethered EEG, configured in a 1-D CVM equiprobable binary vector, and potentially useful for Smartphone wireless display. Longer-range applications include interpreting neural assembly activations via high-density implanted soft, cellular-scale electrodes.

KEYWORDS: EEG, Brain-Computer Interfaces, Cluster Variation Method, statistical thermodynamics, free energy minimization, pattern classification, neural networks, neural ensembles

1 BRAIN-COMPUTER INFORMATION INTERFACES VIA STATISTICAL THERMODYNAMICS

By harnessing the power of statistical thermodynamics, we create a foundation for interpreting the local pattern information available in active neural ensembles, as observable using current brain imaging and EEG methods, along with next-generation Brain-Computer Interfaces (BCIs). This paper presents, as a step in ongoing work, a new method for connecting unsupervised learning in biological neural networks with the global principle of free energy minimization.

One key result presented here is that by explicitly accounting for local pattern distributions in the entropy, we connect observable spatio-temporal patterns with specific enthalpy parameter values. This provides a basis for connecting biological learning and statistical thermodynamics.

Neural processes are part of an isothermal system consisting of a very large ($O(10^{11})$) units, with approximately 19% in the cortex ($O(10^{10})$ units).[1] These units can, very broadly speaking, be either active or non-active; we can treat them as "on" or "off." When we apply statistical thermodynamics, we invoke the notion that neural processes should tend towards a Minimal Free Energy (MFE) over time. Further, implicit in applying statistical thermodynamics, we have the understanding that neural processes – especially unsupervised learning – will not be time-invariant. That is, unlike traditional Newtonian physics, processes guided by free energy minimization will tend towards higher-entropy states.

Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII, edited by Harold H. Szu, Liyi Dai, Yufeng Zheng, Proc. of SPIE Vol. 9496, 949609 · © 2015 SPIE · CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2184651

At the same time, the biological energy put into a neural system leads to complex order, observable in both space and time. Thus, a consistent theme of this work is the interplay between two forces; one towards order (learning to associate distinct and reproducible responses to specific stimuli) and another towards greatest possible dispersion over possible states (increase in entropy).

The unique insight offered here is that by using a more complex entropy representation within the free energy equation, we explicitly address local pattern formation. We do not seek to characterize local patterns explicitly, but rather the degree to which different types of local patterns occur as a function of the interaction enthalpy term. This gives us a valuable potential connection between statistical thermodynamics and neural learning. The statistical thermodynamics observable result is in terms of configuration variables – variables representing nearest-neighbor, next-nearest-neighbor, and triplet configurations. The configuration variable values describe the different kinds of local patterns found in a free energy-minimized state associated with a specific h-value. This gives us a foundation, in future work, to associate h with neural activations and learning processes.

2 SPECIFIC NEURAL RESPONSES YIELD DISTINCT INTEGRATED SETS

Francis Crick, working with Christof Koch, wrote a plea on his deathbed to further our understanding of the claustrum – an area of the brain that integrates responses from many different cortical areas. "When holding a rose, you smell its fragrance and see its red petals while feeling its textured stem with your fingers," he wrote.[2] "A key property of conscious sensations is their integrated nature."

This ability to form very distinct yet highly reproducible responses to different stimuli is key to our using EEG – in both current and next-generation forms – for practical applications. One such application, suggested by Jenkins et al.[3], uses a person's distinctive EEG responses when they detect an image of a known and loved person (e.g., his or her grandmother), even when multiple other images are also present. Jenkins et al. suggested such imagery as a new biomarker; one which could not be readily foiled because it required the real-time response of a live person. *Figure 1* shows how Jenkins et al. suggested presenting an image that would induce an amygdala-based (emotionally-based recognition) response, in concert with other ("decoy") images.

Figure 1: A suite of images, including a person's grandmother (eliciting an emotionally-based EEG response) can provide stimulus in a new EEG-based biomarker recognition system. (Figure from Jenkins et al., 2014[3], used with permission.)

Although current EEG systems include commercial options with limited sensors and signal degradation due to both artifacts and noise, the progression over the past few decades has been towards EEG systems with larger numbers of contact sensors, and recent work on direct neural interfaces shows promise. In particular, there is high promise with the potential of carbon nanotubes[4] to provide interfaces which can provide two-way neural stimulation/signal reception together with high biocompatibility.

3 BCIs: MOVING TO STATISTICALLY-SIGNIFICANT SENSOR NUMBERS

Current EEG-based BCIs have sensor numbers between $O(10^1)$ and $O(10^2)$. Lalonde et al. recently reported efforts modeling *f*-EEG systems with $O(10^3)$ sensors, using a second order co-variance matrix defined as the electrode-pair fluctuation correlation function C(s~, s~') of independent thermodynamic source components[5]. Newer direct-neural Carbon NanoTube (CNT) sensors are much smaller and can thus be used with much greater sensor density. As CNTs are on the same order of size as individual neurons[6], it is reasonable that sensor density could become $O(10^5)$ and higher. At these numbers, statistical thermodynamics methods become very relevant, treating sensors as on/off units. Since (with CNTs) the sensor size approximates that of individual neurons, statistical thermodynamics methods would then be modeling local ensemble activation patterns within the brain.

This is a reasonable approach, since the brain operates as an isothermal reservoir. Further, recent work by Tkačik et al. provide evidence supporting local thermodynamic equilibrium in neural ensembles[7,8]. Nevertheless, there is a challenge with using traditional statistical thermodynamics formalisms, such as the classic Ising model, in which the entropy is expressed as a simple function of only the distribution of units among on/off states, where x_1 is the fraction of units found in the "on" state.

$$S = -[x_1 \ln x_1 + (1 - x_1) \ln(1 - x_1)] \tag{1}$$

The problem is that when this entropy formalism is included in a free energy expression that allows for both a per-unit activation energy (so the activation of "on" units is greater than those on the "off" state) and also an interaction energy, the phase map reveals limitations on model usefulness[9].

4 THE CLUSTER VARIATION METHOD: A STATISTICAL THERMODYNAMICS APPROACH FOR BRAIN-COMPUTER *INFORMATION* INTERFACES

Fortunately, there is a statistical thermodynamics method that can indeed usefully help interpret BCI data. This is the Cluster Variation Method (CVM), originally devised by Kikuchi in 1951[10], and since then largely applied to topics such as solid-state models[11,12]. Over the past decade, researchers have identified a stronger CVM role, including its relationship with other computational forms such as belief propagation and graph theory[13,14].

The essential core of the Cluster Variation Method is that, in addition to modeling the extent to which units are "on" or "off," we also model the distribution of local patterns: nearest-neighbors, next-nearest-neighbors, and triplets comprise a useful CVM. This can be applied to either 1-D or 2-D arrays. The configuration variables x_i (single unit), y_i (nearest neighbor), w_i (next-nearest neighbor), and z_i (triplet) are shown in the following *Figure 2*.

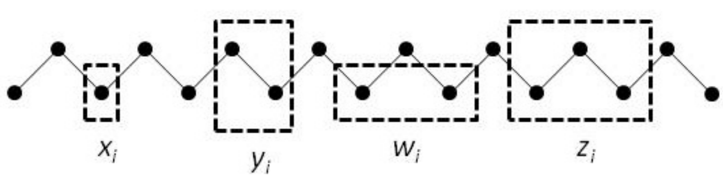

Figure 2: Configuration variables x_i, y_i, w_i, and z_i in a single zigzag chain for the 1-D Cluster Variation Method.

The distribution of the configuration variables z_i when all enthalpy terms are set to zero is shown in the following *Figure 3*.

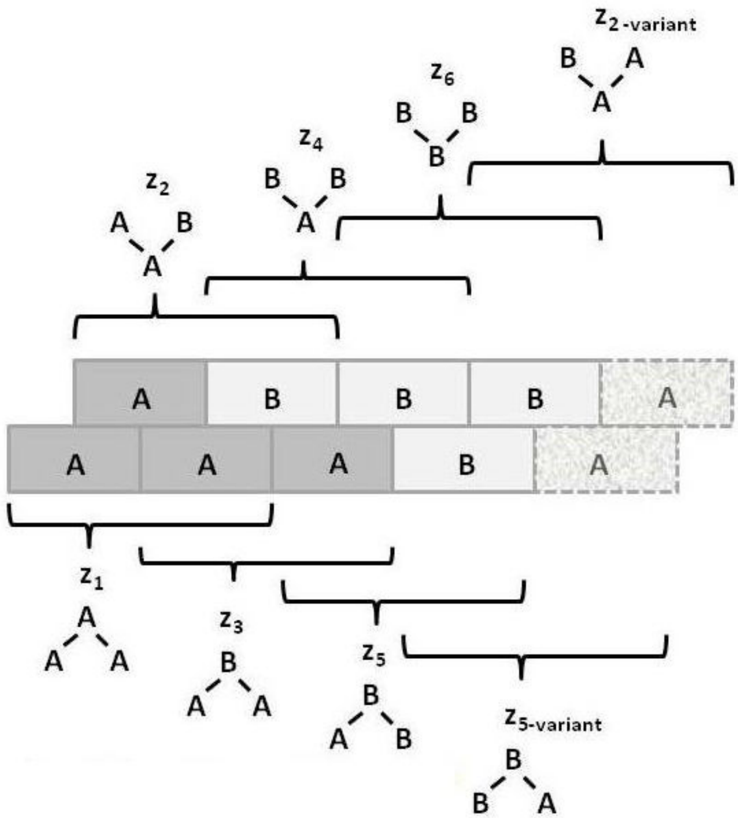

Figure 3: Micro-system for the 1-D Cluster Variation Method ensemble (single zigzag chain) with equilibrium-values for the configuration values z(i), where the interaction enthalpy is set to zero.

The reduced free energy, using the CVM approach, for a one-dimensional system (single zigzag chain) of units is:

$$\overline{A_{1D}} = \frac{\beta A_{1D}}{N} = \beta\epsilon(z_2 + z_4 + z_3 + z_5) - 2\sum_{i=1}^{3} \beta_i \, Lf(y_i) + 2\sum_{i=1}^{6} \gamma_i Lf(z_i) + \mu\beta[1 - \sum_{i=1}^{6} \gamma_i z_i] + 4\lambda(z_3 + z_5 - z_2 - z_4), \quad (2)$$

where $\beta\epsilon$ is the reduced interaction energy between unlike units and μ and λ are Lagrange multipliers. Also, the unit enthalpy for "on" units is the same as for the "off" units, and is set to zero.

In this equation, $Lf(v) = vln(v) - v$, where v can respectively take on the values of x_i, y_i, and z_i. The full details of the equation and its subsequent analytic solution are available for both the 1-D case[15] and the 2-D CVM case[16]. Further, the equation makes use of the equivalence relation $2y_2 = z_2 + z_4 + z_3 + z_5$, where y_2 is the fraction of unlike (A-B) pairs.

Thus, a negative interaction enthalpy ϵ represents a stabilizing of unlike (A-B) pairs, and destabilizing of like (A-A and B-B) pairs. When we minimize the free energy, a negative ϵ should yield a system with relatively greater numbers of unlike pairs compared to a nominal system where the interaction enthalpy is zero. This means that there should be both more unlike pairs and also more triplets of unlike units (A-B-A and B-A-B); an anti-ferromagnetic-like ordering. Similarly, a positive value for ϵ should yield a greater preponderance of like pairs, and thus induce triplets and larger clusters of like-near-like; a ferromagnetic-like ordering.

We obtain the free energy minimum by computing the derivative of the free energy with respect to each of the configuration variables z_i, and setting each of these to zero.

For the specific case where $x_1 = x_2 = 0.5$, it is possible to compute an analytic solution for each of the configuration variables y_i and z_i in terms of h, where $h = e^{-\beta\epsilon/4}$.

As an example, of one of the solutions for free energy minimization is found for the configuration variable z_3:

$$z_3 = \frac{1}{2(h^2+1)^2} \tag{3}$$

5 1-D CVM ANALYTIC SOLUTION WHEN $x_1 = x_2 = 0.5$

Figure 4 presents the graphical results of the analytic solution for minimizing the free energy of Eqn. (2) when $x_1 = x_2 = 0.5$, for the three configuration variables y_2, z_1, and z_3, where y_2 is represents the fraction of unlike (A-B) nearest neighbors, z_1 represents the fraction of like (A-A-A) triplets, and z_3 represents the fraction of unlike-unlike (A-B-A) triplets. (When $x_1 = x_2 = 0.5$, values for z_3 are symmetric with z_4, values for z_1 are symmetric with z_6, and values for z_2 and z_5 are symmetric and can be computed from the previous triplet configuration variables.)

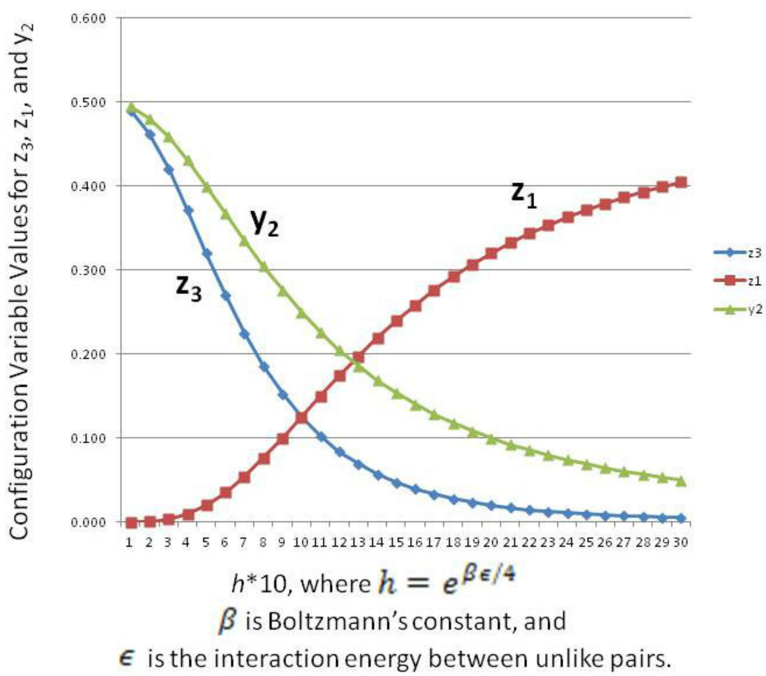

Figure 4: Results from the analytic solution for the configuration variables, z_i, in the 1-D zigzag chain, as a function of the reduced interaction energy term ϵ. The curves for y_2, z_1, and z_3 are shown. A detailed discussion of these results is available[15].

The left-most point on this graph is where *10h=1* (in *Figure 4*, where the point on the x-axis is $10h$). At this point, $h = e^{\beta\epsilon/4} = 0.1$, so that the interaction enthalpy ϵ between unlike (A-B) pairs is negative. We observe the expected results; as h decreases, $z_1 \rightarrow 0$, meaning that the fraction of like triplets (A-A-A) becomes vanishingly small. The fractions of unlike pairs (A-B) and unlike triplets (A-B-A), or $y_2$and z_3 respectively, approach their maximal possible values of 0.5. (Note that y_2 has a degeneracy of 2; the A-B pair is complemented by the B-A pair; thus y_2 can have a maximal value of 0.5. Further, z_3 is complemented by z_4 (B-A-B), so its maximal value is also 0.5.)

When the interaction enthalpy ϵ is zero ($\beta\epsilon \cong 0$), we have $h = e^{\beta\epsilon/4} = 1$, seen on Figure 4 at the *10h = 10* on the x-axis. At this point and in the immediate neighborhood, we would expect the most "disordered" state. The cluster variables should all achieve their nominal distributions; $z_1 = z_3 = 0.125$, and $y_2 = 0.25$. This is precisely what we observe. (Nominal distributions for the configuration variables when the interaction enthalpy is zero are given in Maren[15].)

As we move to the RHS of *Figure 4*, where $\epsilon > 0$ and $h = e^{\beta\epsilon/4} > 1$, we have the case of a positive interaction energy between unlike units (the A-B pairwise combination). We would expect that as ϵ increases as a positive value, that we would minimize y_2, and also see smaller values for those triplets that involve non-similar pair combinations. That is, the A-B-A triplet, or z_3, would approach zero. We observe this on the RHS of the above graph. This is the case where as $h = e^{\beta\epsilon/4}$ moves into the positive range (0-3, or $10h \rightarrow 30$), we see that y_2 and z_3 fall towards zero. In particular, z_3 becomes very small. Correspondingly, this is also the situation in which $z_1 = z_6$ becomes large; we see z_1 taking on values > 0.4 when h > 2.9.

This is the realm of creating a highly structured system where large "domains" of like units mass together. These large domains (comprised of overlapping A-A-A and B-B-B triplets) stagger against each other, with relatively few instances of "islands" (e.g., the A-B-A and B-A-B triplets.)

As applied to neural systems, we would expect that areas of activated neural ensembles would be more cohesive rather than being disjoint.

6 CONNECTING THE FREE ENERGY TO UNSUPERVISED LEARNING

We begin with a multispectral data set $\vec{X}$, which may be found per pixel $\vec{x}$ in modeling biological optical systems, or as activations of neurons or neural ensembles, and which is also a function of time t. We then describe $\vec{X}$ as a linear mixture, being the product of a mixing matrix $[\boldsymbol{A}]$ operating on a set of time-varying sensor data $\vec{S}(\vec{x}, t)$:

$$\vec{X}(\vec{x}, t) = [\boldsymbol{A}]\vec{S}\,(\vec{x}, t), \tag{4}$$

without knowing the impulse response mixing matrix function $[\boldsymbol{A}]$ of the imaging or neural response system. In this case, the vector quantity $\vec{S}(\vec{x}, t)$ represents inputs into the system. This may be sensor data impinging on a biological optical system, or may constitute sensor-driven stimulus or stimulus from afferent neurons in a neuron or neural ensemble.

We are also, at the same time, parameterizing the entropy *S* as a set of measurements of sensor data or inputs into a system, with the note that entropy truly represents a measurement of the degree of uniformity of data distribution.

Eqn. (4) presents us with a need to perform blind de-convolution to obtain the inverse of $[\boldsymbol{A}]$, which characterizes the Blind Source Separation (BSS) problem

$$\vec{S}\,(\vec{x}, t) \approx [\boldsymbol{W}]\vec{X}(\vec{x}, t), \tag{5}$$

where the inverse solution of source stimulus is based on estimating the neural network learning weight matrix $[\boldsymbol{W}]$. This suggests that we would be able to identify the entropy *S* in parameterized form $\vec{S}$ as a function of the observable neural responses $\vec{X}$.

We now address system entropy, drawing on the notion of the tendency towards entropic increase and noting the following relations between the Reservoir (RV) and a sub-system black body (bb):

$$\Delta S > 0\,; \Delta S_{RV}T_o + \Delta S_{bb}T_o > 0\,; -E_{bb} + S_{bb}T_o \equiv -H_{bb} > 0 \tag{6}$$

where use is made of the conservation of total energy $\Delta S_{RV}T_o + \Delta E_{bb} = 0$.

We expand the internal energy *H* using a Taylor series in order to estimate the change vector of entropy sources $\Delta\vec{S} \equiv \vec{S} - \vec{S}_o$.

$$\Delta H_{bb} = (\frac{\Delta H_{bb}}{\Delta\vec{S}})\,\Delta\vec{S} = \vec{\mu}\cdot\{[\boldsymbol{W}]\vec{X} - \vec{S}_{bbo}\} \rightarrow 0 \tag{7}$$

$$\vec{\mu} \equiv (\frac{\Delta H}{\Delta\vec{S}}); \tag{8}$$

where $\vec{S}_{bbo}$ is the subsystem ("black body") entropy at some initial time 0, $[\boldsymbol{W}]\vec{X}$ is the entropy at some later time *t* (see Eq. 5), and $\vec{\mu}$ is known as Lagrange energy slope vector parameter. Such a blind de-convolution formulism is known as *Lagrange Constraint Neural Network (LCNN)* for *Blind Sources Separation (BSS) based on Helmholtz MFE*, which we express as

$$\frac{\boldsymbol{\delta}[\boldsymbol{w_{i,j}}]}{\boldsymbol{\delta t}} = \frac{\delta H}{\boldsymbol{\delta}[\boldsymbol{w_{i,j}}]} = \vec{X}_i\vec{\mu}_j, \tag{9}$$

where $[w_{i,j}]$, the vector elements giving the dependence of entropy on observable activations, need to be determined.

Eqn. (9) proposes a means for connecting Hebbian neural learning $\frac{\delta[w_{i,j}]}{\delta t}$ with the local-in-time neural activation vector $\vec{X}_i$ mediated by the correlated change in system free energy with the change in entropy.

Biologically speaking, the Lagrange constraint $\vec{\mu}_j$ could potentially be correlated with the house-keeping glial cells, one per brain neuron, 10 billion glia cells and 10 billion neurons, driving the brain to isothermal equilibrium at 37°C[17, 18].

More specifically, we envision a potential connection between (Hebbian) neural learning and the observable local patterns (degree to which activated neurons or neural ensembles are clustered together). This avenue is part of ongoing inquiry.

7 APPLYING THE 1-D CVM METHOD TO OBSERVABLE DATA SETS

The analytic CVM solution is possible only when the unit distribution is equiprobable. This is not as onerous a demand as it might seem. A two-step process can transform a map of 2-D data into a 1-D array, and from there, generate an equiprobable distribution data set for which the CVM is a useful analytic tool.

Hsu and Szu[19] have described how the Peano-Hilbert method can generate a 1-D space-filling curve from a 2-D data map, such as is found in EEG and other topographic-mapping or imaging tasks. They demonstrated that this method resulted in a 20 dB lower baseline than the prior frequency spectrum.

To apply the 1-D CVM, is necessary to have an equiprobable data set. Stephens et al. developed an elegant method for this purpose.[20] Saremi and Sejnowski applied this method to 1-D data slices taken from natural images, in order to assess whether the images could be considered at a critical point.[21] While they did not find such evidence (applying a straightforward Ising model to a 1-D system), their use of the Stephens et al. method sets the stage for analysis with the CVM approach.

One possible application of this approach would be to EEG data, which is the most common means for non-invasive inputs to a Brain-Computer Interface (BCI). To do this, the Peano-Hilbert method identified by Hsu and Szu, followed by generating an equiprobable dataset using the method of Stephens et al., could then be described by an *h*-value. The choice of granularity in identifying data units, along with the resulting *h*-values, would potentially assist in characterizing neural responses to different stimuli. We note here that applying the Peano-Hilbert method of space-filling curves to analyzing EEG data becomes natural when the actual configuration of the EEG sensors follows such a layout.[22]

8 BIOLOGICAL PLAUSIBILITY FOR FREE ENERGY MINIMIZATION IN NEURAL SYSTEMS

As we approach modeling both neural dynamics and resultant activations within dense electrode arrays, we need to give attention to how neural ensembles interact with each other. Early work by Donald Hebb gave rise to the notion that organized behavior,[23] so that certain neural assemblies become jointly activated. Other crucial early work includes ideas about cooperative dynamics originally espoused by luminaries such as Edelman,[24] followed by striking observations of long-range synaptic correlations by Singer and colleagues.[25] More recent studies emphasize the role of learned synchronization in neural ensembles.[26,27]

In particular, we are interested in statistical thermodynamics-based approaches that can potentially yield phase transition models and other useful insights for neural systems.[7,28,29] Barton and Cocco have developed a selective cluster expansion (SCE) algorithm that yields an approximate solution to the inverse Ising inference problem, which they have applied to experimental data for multielectrode recordings.[30] The advantage of this approach is that they can disentangle networks of direct interactions from correlations, and then extract structural information on selective connections.

A method such as Barton and Cocco's is useful for understanding correlated neural activations at the detail-level which will influence next-generation BCIs. However, there is advantage to an approach that correlates observable local patterns with the single, simple h-value, as was described in Section 5.

In order to take maximal advantage of this approach for BCI, though, we need to map the local response pattern characteristics to a representation of the applied BCI stimulus. This is particularly true since we do not expect a local pattern consistency across the entire cortical response region. Rather, areas of activation should be correlated with both overall and specific nature of the stimulus.

A useful means for doing this is to create a hybrid neural network; one which maps an overall activation pattern characterization (locally-based *h-values*) to a response-characterization layer. This kind of hybrid neural network [involving a base-layer CVM was first developed using a Hebbian learning rule for pattern classification.[31,32]

9 USING THE CVM WITHIN A NEURAL NETWORK ARCHITECTURE

The preceding sections showed how the 1-D (or potentially, the 2-D) CVM could be used as a means for modeling neural processes approaching a local-in-time free energy minimum. The key differentiator in the CVM as compared with most Ising-like free energy equations was the explicit inclusion of nearest-neighbor and triplet configuration variables in the entropy term. Under the condition of equiprobable distribution of units into on/off states, we have an analytic solution for these configuration variables in terms of a single parameter, h, which is a function of the interaction enthalpy between neighboring units.

Section 7 showed how the equiprobable distribution ($x_1 = x_2 = 0.5$) requirement could be addressed via a two-step mapping process. This would allow the 1-D CVM to be applied as a pattern-characterizing tool, and the resultant h-value would indicate the nearest-neighbor interaction enthalpy associated with the observed configuration values.

A complementary use for the CVM (either 1-D or 2-D) is as a hidden layer component in a complex neural network.

There is a distinct and unique advantage of using such a neural network-based system for pattern classification based on inputs from a complex biological system. This is that we build on understanding that neural activation processes, modeled as a collection of interacting ensembles, will seek some form of local-in-time free energy minimization. This should hold, even as neural activations are dominantly driven by neural stimulus inputs. However, as neural ensembles learn to form stable patterns in response to known stimuli, there are thus two processes at work: activation due to input stimulus and modification of which neural units become active in order to achieve free energy minimization.

A complex, hybrid neural network incorporating these two processes includes a CVM layer which receives sustained activations from an input layer, and is allowed to come to free energy equilibrium. The resulting free energy-minimized layer then becomes input to a classification system, with the usual options for training weights from the input-to-hidden and hidden-to-output nodes.

The value of this approach is that the hidden layer thus undergoes two distinct learning processes; one involving the usual weight training,[33] and another – internally-stabilized via learned lateral connections – incorporating free energy minimization, and thus ensuring that the activation patterns produced within the hidden layers are responsive to the patterns of the input layer, which are taken directly from neural activations.

One important reason for investigating this kind of hybrid neural system is that the hidden layer activations can be designed to have temporal persistence, even when the initial input stimulus fades. The activation persistence is driven by maintaining a free energy-minimized state, where some fluctuation of unit activations is allowable. This provides a new means of modeling time-varying pattern evolution.

There is a strong tradition within the neural network community of having certain neural learning methods based on free energy minimization.[34-39] What is unique about using the CVM as the free energy minimization basis is that we provide an architecture by which the hidden layer can behave in a neuromorphic sense; the formation of local pattern configurations can mimic the extent to which such are observed in an actual neural system. This provides a means for modeling neural activations within the statistical thermodynamic ensemble of on/off units.

This approach was the basis for early development of the first CORTECON neural networks, in which a hidden layer first received activation from an input layer, and then came to free energy equilibrium using the 1-D CVM. Persistent inputs allowed certain active units within the hidden (CVM) layer to develop intra-layer connections that stabilized certain local clusters in response to given stimuli. Then, the stabilized hidden nodes generated signals to a pattern classification output layer. In the first stage, the CORTECON had a single major free-energy/output-weight-learning cycle; in future developments, this could be iterated to yield a selection of hidden layer responses that both achieved free energy minimization and effective pattern classifying output signals. [40-42] Ideally, the free energy minimized hidden layer would then also have local pattern configurations whose h-values were representative of the local patterns found in the input stimulus; preferably coming from activated neural ensembles.

10 DIRECTIONS FOR FUTURE WORK

EEG-based COTS systems are finding increased use in multiple out-of-the-lab applications, where the dry-electrode interface can be contained within a baseball cap.[43,44] Wireless EEG systems, such as those connected to a smartphone, are also now available.[45] Thus, even with current technologies, the opportunities to use EEG in practical daily applications are increasing. Further, there is a movement towards soft electrodes which withstand a subject's better than more traditional stiff-wire electrodes **[zz4]**. A recent EEG system contains both connectors and electrodes which are arranged in a spatially-varying design based on Peano curves; this provides stretchability along longitudinal axes.

This paper envisions a future scenario in which the density of device contact points increases substantially. This is likely within a realistic timeframe, as wireless BCIs become available.[46]

One of the most immediate applications of the 1-D CVM for modeling pattern formation is in interpreting EEG signals, where the *h*-value provides an additional input into a classic pattern classification scheme. Given that EEG systems are finding greater use as COTS BCI components, this can become a useful addition. In particular, the CVM approach can assist characterization of sustained responses to known signals. These are neural responses which have learned, over time, to achieve free energy minimization as well as forming a distinct and characteristic response to specific stimuli. The CVM value lies in emulating the free energy minimization inherent within the neural processes.

This approach has widespread applications beyond pattern characterization in neural systems. One obvious application is in providing a useful measure describing local patterns within natural images, building on work established by Stephens et al.[20] Another potential application is to text data mining, particularly in entity co-reference determination,[47] where the CVM method provides an alternative to Bayesian prior probabilities for determining the likely entity-type for a given entity in the text sequence. It would be necessary to determine two sets of entities, where the distribution of each would be near equiprobability. (It may be necessary to do perturbation analyses, or to determine regions on the z_i as functions of h to which various entity sets match.) Nevertheless, it is conceivable that the CVM approach could complement Kullback-Leibler or other information measures.[48]

11 CONCLUSIONS

The Cluster Variation Method (CVM) has been used for decades to model processes in solid state lattices. Now, it has potentially surprising and powerful applications in the realm of Brain-Computer Interfaces (BCIs). Its value lies in its ability to create free energy-minimized local pattern configurations that can be trained to emulate those found in neural activation patterns. Recent experimental and theoretical work evidences the strong likelihood of free energy minimization processes within neural systems. As the density of neural interface direct connections increases, and the connection granularity approaches the neural-size level, it becomes more important than ever to model the statistical thermodynamics as well as the activation-induced learning in neural ensembles. An analytic solution to the 1-D CVM can provide a simple and expedient connection between local pattern configurations and the interaction enthalpy, leading to characterization of neural activation patterns in terms of a single thermodynamic-based parameter. This provides value in characterizing ensembles and in improving pattern classification.

REFERENCES

[1] Azevedo, F.A., Carvalho, L.R., Grinberg, L.T., Farfel, J.M., Ferretti, R.E., Leite, R.E., Jacob Filho, W., Lent, R., Herculano-Houzel, S., "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain," *J. Comp. Neurol.* **513**(5), 532-41 (2009 Apr 10). DOI: 10.1002/cne.21974

[2] Crick, F. C., Koch, C., "What is the function of the claustrum?" *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360** (1458): 1271–1279 (2005 Jun 29).

[3] Jenkins, J., Sweet, C., Sweet, J., Noel, S., Szu, H. "Authentication, privacy, security can exploit brainwave by biomarker," *Proc. SPIE STA (Sensing Technology + Applications) Conference: Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII* (H. H. Szu and L. Dai, Eds.), **9118**, 91180U (2014 June).

[4] Vitale, F., Summerson, S. R., Aazhang, B., Kemere, C., Pasquali, M., "Neural stimulation and recording with bidirectional, soft carbon nanotube fiber microelectrodes," *ACS Nano* 150324172405006 (2015). DOI: 10.1021/acsnano.5b01060. Online: http://pubs.acs.org/doi/pdfplus/10.1021/acsnano.5b01060

[5] Lalonde, F., Gogtayand, N., Giedd, J., Vydelingum, N., Brown, D., Tran, B., Hsu, C., Hsu, M.-K., Cha, J., Jenkins, J., Ma, L., Willey, J., Wu, J., Oh, K., Landa, J., Lin, C. T., Jung, T. P., Makeig, S., Morabito, C., Moon, Q., Yamakawa, T., Lee, S.-Y., Lee, J.-H., Szu, H. H., Kaur, B., Byrd, K., Dang, K., Krzywicki, A., Familoni, B. O., Larson, L., Harkrider, S., Krapels, K. A., and Dai, L., "Brain order disorder 2nd group report of f-EEG," *Proc. SPIE STA (Sensing Technology + Applications) Conference: Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XII* (H. H. Szu and L. Dai, Eds.), **9118** (June 2014). DOI: 10.1117/12.2051706

[6] Cogan, S. F., "Neural stimulation and recording electrodes," *Annu. Rev. Biomed. Eng*. **10**, 275-309 (2008). DOI: 10.1146/annurev.bioeng.10.061807.160518.

[7] Tkačik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., and Berry II, M., "Searching for collective behavior in a large network of sensory neurons," *PLoS Comput. Biol.* **10**, e1003408 (March 2014). DOI: 10.1371/journal.pcbi.1003408.
[8] Tkačik, G., Marre, O., Mora, T., Amodei, D., Berry II, M., and Bialek, W., "The simplest maximum entropy model for collective behavior in a neural network," *J. Stat. Mech.*, **2013**, P03011 (2013). DOI:10.1088/1742-5468/2013/03/P03011.
[9] Maren, A.J., "Statistical thermodynamics: introduction to phase space and metastable states," *Tech. Rep. THM TR2014-001 (ajm)*, Themasis (2014).
[10] Kikuchi, R., "A theory of cooperative phenomena," *Phys. Rev.* **81**, 988-1003 (1951).
[11] Mohri, T. "Short range ordering and local displacement of alloys studied by CVM," *Int'l. J. Computational Materials Science and Engineering*, **1**(2), 1250018 (12 p) (2012). DOI: 10.1142/S2047684112500182
[12] Maren, A., Lin, S. H., Langley, R. H., Eyring, L., "A theoretical model of solid state phase reactions: hysteresis and kinetics," *J. Solid State Chemistry* **53**(3), 329-343 (1984).
[13] Pelizzola, A., "Cluster Variation Method in statistical physics and probabilistic graphical models," *J. Phys. A: Math. Gen.*, **38**, R309 (2005). DOI 10.1088/0305-4470/38/33/R01.
[14] Yedidia, J., Freeman, W., and Weiss, Y., "Understanding belief propagation and its generalizations," *Tech. Rep. MERL TR-2001-22*, Mitsubishi Electric Research Laboratories, January 2002. www.merl.com.
[15] Maren, A.J., The Cluster Variation Method I: 1-D single zigzag chain: basic theory, analytic solution and free energy variable distributions at midpoint (x1 = x2 = 0.5). *Tech. Rep. THM TR2014-002 (ajm)*, Themasis (2014). http://www.aliannajmaren.com/Downloads/Cluster_Variation_Method_I_Basic_Theory_2014-06-25.pdf. DOI: 10.13140/2.1.4415.6485.
[16] Maren, A.J., The Cluster Variation Method II: 2-D grid of zigzag chains: basic theory, analytic solution and free energy variable distributions at midpoint (x1 = x2 = 0.5). *Tech. Rep. THM TR2014-003 (ajm)*, Themasis (2014). http://www.aliannajmaren.com/Downloads/Cluster_Variation_Method_II_Basic_Theory_2-D_2014-07-07.pdf DOI: 10.13140/2.1.4112.5446.
[17] Szu, H., Miao, L., and Qi, H., "Unsupervised learning with mini free energy," *Proc. SPIE Conference: Independent Component Analyses, Wavelets, Unsupervised Nano-Biomimetic Sensors, and Neural Networks V* (H. H. Szu and J. Agee, Eds.), **6576** (2007 Apr 9). DOI: 10.1117/12.725198.
[18] Szu, H. and Kopriva, I., "Unsupervised learning with stochastic gradient," *Neurocomputing* **68**, 130160 (2005).
[19] Hsu, C. and Szu, H. H., "Low discrepancy sampling of parameter surface using adaptive Space-Hilbert-Curves (SFC)," *Proc. SPIE STA (Sensing Technology + Applications) Conference: Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XII* (H. H. Szu and L. Dai, eds.), **9118**, 91180P1-91180P8 (2014 June). DOI: 10.1117/12.2053306.
[20] Stephens, G. J., Mora, T., Tkačik, G., and Bialek, W., "Statistical thermodynamics of natural images," *Phys. Rev. Lett.* **110** (2013).
[21] Saremi, S. and Sejnowski, T. J., "On criticality in high-dimensional data," *Neural Computation*, 1-11 (2014).
[22] Norton, J. J. S., et al., "Soft, curved electrode systems capable of integration on the auricle as a persistent brain–computer interface," *Proc. Natl. Acad. Sci.* 112(13), 3920–3925 (2015). DOI: 10.1073/pnas.1424875112
[23] Hebb, D.O., [The Organization of Behavior: A Neuropsychological Theory] Psychology Press, Florence, KY (2002, orig. pub. 1949).
[24] Edelman, G. M., and Mountcastle, V. B., [The Mindful Brain: Cortical Organization and the Group-selective Theory of Higher Brain Function] MIT Press, Cambridge, MA (1978).
[25] Singer, W., "Synchronization of cortical activity and its putative role in information processing and learning," *Ann. Rev. Physiology* **55**, 349 – 374 (1993).
[26] Engel, A. K., Gerloff, C., Hilgetag, C. C., and Nolte, G., "Intrinsic coupling modes: multiscale interactions in ongoing brain activity," *Neuron* **9**, 38 ff. (2013).
[27] Moran, R., Pinotsis, D. A., and Friston, K., "Neural masses and fields in dynamic causal modeling," *Front. Comp. Neurosci.* **7**, 57 ff. (2013). DOI: 10.3389/fncom.2013.00057
[28] Cowan, J. D., Neuman, J., and van Drongelen, W., "Self-organized criticality in a network of interacting neurons," *J. Stat. Mech.* **04**, P04030 (2013). DOI:10.1088/1742-5468/2013/04/P04030
[29] Advani, M., Lahiri, S., and Ganguli, S., "Statistical mechanics of complex neural systems and high dimensional data," *J. Stat. Mech.* **04**, P03014 (2013). DOI:10.1088/1742-5468/2013/03/P03014.
[30] Barton, J., and Cocco, S., "Ising models for neural activity inferred via selective cluster expansion: structural and coding properties," *J. Stat. Mech.* **04**, P03002 (2013). DOI:10.1088/1742-5468/2013/03/P03002.

[31] Maren, A. J., “Hybrid and complex networks,” [Handbook of Neural Computing Applications] (A.J. Maren, R. Pap, C. Harston; Academic, New York, 203-218 (1990).
[32] Maren, A. J., “A logical topology of neural networks,” *Proc. Second Workshop on Neural Networks – WINN/AIND 91 (Workshop Neural Networks: Academia, Industry, NASA, and Defense; Auburn, GA, Feb. 14-16)* (1991).
[33] Werbos, P. J., “Backpropagation through time: what it does and how to do it,” *Proc. IEEE* **78**(10), 1550-1560 (1990 Oct),
[34] Hopfield, J., “Neural networks and physical systems with emergent collective computational abilities,” *Proc. Natl. Acad. Sci.USA* 79, 25542558 (April 1982).
[35] Kirkpatrick, S., Gelatt Jr., C. D., and Vecchi, M. P., “Optimization by simulated annealing,” *Science* **220** (4598), 671-680 (1983).
[36] Hinton, G. E., and Sejnowski, T. J., “Learning and relearning in Boltzmann Machines,” [Parallel Distributed Processing: Explorations in the Microstructure of Cognition] vol. 1, MIT Press, Cambridge, MA, 282-317 (1986).
[37] Szu, H. “NI can do better compressive sensing,” *Natural Intelligence: the INNS Magazine* **1**(2), 5-22 (2012 Winter).
[38] Szu, H., “Thermodynamics energy for both supervised and unsupervised learning neural nets at a constant temperature*,” Int. J. Neur. Syst.* **09**, 175 (1999). DOI: 10.1142/S0129065799000162
[39] Szu, H. “Thermodynamics energy for both supervised and unsupervised learning neural nets at a constant temperature,” *Int. J. Neural Syst.* **9**(3), 175-186 (1999).
[40] Schwartz, E. and Maren, A. J., “Domains of interacting neurons: a statistical thermodynamics model,” *World Congress on Neural Networks (WCNN’93 – Portland)*, (Portland, OR: July 11-15, 1993), I-577 – I-580 (1993).
[41] Maren, A. J., Schwartz, E., and Seyfried, J., “Configurational entropy stabilizes pattern formation in a hetero-associative neural network,” *1992 IEEE Intl. Conf. Systems, Man, & Cybernetics* (Chicago, Il; Oct. 18-21), 89-93, (1992). DOI: 10.1109/ICSMC.1992.271796
[42] Maren, A.J., “Free energy as driving function in neural networks,” Presented at the *1993 Symposium on Nonlinear Theory and Its Applications* (Hawaii: Dec. 5-10), (1993). DOI 10.13140/2.1.1621.1529
[43] Lee, S.-C., Shin, Y.-H., Woo, S., Kim, K., and Lee, H.-N., “Review of wireless brain-computer interface systems” [Chapter 11: Brain-Computer Interface Systems – Recent Progress and Future Prospects] (Ed. R. Fazel-Rezai) (June 5, 2013). DOI: 10.5772/56436
[44] Sullivan, T. J., Deiss, S. R., Jung, T.-P., and Cauwenberghs, G., “A brain-machine interface using dry-contact, low-noise EEG sensors,” *Proc. IEEE Intl. Symp. Circuits and Systems - ISCAS*, 1986-1989 (2008). DOI: 10.1109/ISCAS.2008.4541835
[45] Stopczynski, A., Stahlhut, C. Larsen, J. E., Petersen, M. K., Hansen, L. K., “The Smartphone Brain Scanner: a portable real-time neuroimaging system,” *PLOS* (2014 Feb. 5). DOI: 10.1371/journal.pone.0086733
[46] Regalado, A., “A brain-computer interface that works wirelessly,” MIT Technology Review (January 14, 2015). Online at: http://www.technologyreview.com/news/534206/a-brain-computer-interface-that-works-wirelessly/, accessed 05/21/2015.
[47] Sleeman, J, and Finin, T., “Taming wild big data,” *AAAI Fall Symposium on Natural Language Access to Big Data*, AAAI Press (2014 Nov.).
[48] Kullback, S., and Leibler, R., “On information and sufficiency,” *Ann Math Statist* **22** (1), 79-86 (1951).

[i] Author to whom correspondence should be addressed: Alianna J. Maren, Northwestern University School of Professional Studies (SPS), Master of Science in Predictive Analytics (MSPA) program; alianna.maren@northwestern.edu.