The Diminishing Effect of Noise in BCM Neuron

A. Bazzani*,D. Remondini*, N. Intrator†*, G.C. Castellani‡

*Dept. of Physics and INFN, Univ. of Bologna, 40126 Bologna Italy,
†Inst. for Brain and Neural Systems, Brown Univ., 02912 Providence RI
‡Dept. of Physics and DIMORFIPA, Univ. of Bologna
E-mail: bazzani@bo.infn.it

Abstract

We study the statistical properties of a BCM neuron whose synapses are stimulated by inputs described as a stationary Markov process. We characterize the stable equilibria for the average equations of the synaptic weights as the maxima of a cubic function, associated to the third moments of the inputs distribution, constrained on an ellipsoid defined by the inputs covariance matrix. We do this approach to study the selectivity of a BCM neuron when the input signals are perturbed by a random noise as a function of the noise Root Mean Square (RMS) and the dimension of the input space. We prove the existence of critical values for the noise RMS which define the selectivity properties of the BCM neuron.

Introduction

The BCM theory of cortical plasticity has been introduced by Bienenstock, Cooper and Munro [Bienenstock et al., 1982] to account for the changes observed in cell response of visual cortex due to changes in visual environment. This learning model allows modeling and theoretical analysis of various visual deprivation experiments such as monocular deprivation (MD), binocular deprivation (BD) and reversed suture (RS) [Intrator and Cooper, 1992] and is in agreement with the many experimental results on visual cortical plasticity [Clothiaux et al., 1991, Shouval et al., 2000]. More recently it was shown that the so called ϕ function and its sliding threshold θ, a nonlinear function of the past neuronal history, support the framework of bidirectional synaptic plasticity. In particular the BCM theory [Bear et al., 1987] becomes consistent with two forms of neuronal plasticity observed in hippo-campus; the Long Term Depression (LTD) and the Long Term Potentiation (LTP) if the threshold is related to the calcium entry through the NMDA receptors. A laterally connected network of BCM neurons shows structural properties that are in good agreement with the columnar organization of visual cortex. A variant of this theory [Intrator and Cooper, 1992] supports an energy functional minimization and performs exploratory projection pursuit using a projection index that seeks for multi-modality in data distributions. More precisely, the average behavior of BCM neuron is characterized by the stationary points of an "energy" function that compares the third moments of the output of the neuron with the square of the second moments. This variant has been used in visual cortical

*Corresponding author:Phone (401)863-3857,Email nin@cns.brown.edu.

modeling [Shouval et al., 1996, Shouval et al., 1997, Blais et al., 1998] and in real-world high dimensional classification applications, e.g. : [Intrator and Gold, 1993] [Reisfeld and Yeshurun, 1992],[Huynh et al., 1998]. The analysis performed in [Intrator and Cooper, 1992] characterized the solution of n linearly independent vectors in an n dimensional space. It further showed that for sufficiently small Gaussian noise around each of the independent inputs, the solution of n clusters is close to the solution of the linearly independent case. Further analysis [Castellani et al., 1999] characterized the solution in the case of a laterally interacting network. This paper extends earlier analysis [Intrator and Cooper, 1992, Castellani et al., 1999] to the more general case of inputs distributions. In particular, it characterizes sufficient conditions on the noise so that the solution of a system of n linearly independent inputs with additional additive noise is "close" to the solution characterized by the non-noisy system. It further characterizes the noise properties as a function of the dimensionality of the system, and concludes that when the noise is bounded by $O(n^{1/4})$, the solution of the non-noisy system fully characterizes the solution of the clustered situation. In simple words, we show that due to the inherent sparsity of high-dimensional spaces, the case of n linearly independent input vectors, with additive Gaussian noise, becomes more realistic as the dimensionality of the systems becomes large and the linear independent case is a reasonable approximation. For this analysis, we develop a new method to compute the solutions and their stability. This method, in the case of a bimodal distribution, gives a closed relation for the two peaks discrimination depending on parameters like variance of each peak and peaks distance of the input distribution. The generalization to a multi-modal distribution follows in a natural way from the bidimensional case. We also discuss the average behavior of a network of connected BCM neurons.

Our method can be applied to study the selectivity property of a BCM neuron as a function of the second and third moments of the input distribution.

The paper is organized as follows:

1) in section 1 we develop a general method to compute the average equilibrium solutions of a BCM neuron and their stability;

2) in section 2 we discuss the explicit case of a bimodal distribution in the plane;

3) in section 3 we study the equilibrium solution of a BCM neural network.

1 BCM Neuron as Feature Detector

We consider the evolution of the synaptic weights $\vec{m} \in \mathbb{R}^n$ of a BCM neuron which is stimulated by the external signal $\vec{d}(t)$. We assume that $\vec{d}(t)$ is described by a stationary Markov process [J.L.Doob, 1953] whose covariance matrix and cubic moments are given by

$$E[d_i d_j] = C_{ij} \qquad E[d_i d_j d_k] = B_{ijk} \tag{1}$$

where $E[\quad]$ denotes the expectation value with respect to the probability distribution of the external signal. We consider the neuron in the linear approximation so that the activity x is defined $x = \vec{m} \cdot \vec{d}$. According to the BCM theory, the time evolution of the weights $\vec{m}$ are given by the equation

$$\dot{\vec{m}} = x(x-\theta)\vec{d} \qquad \vec{m}, \vec{d} \in \mathbb{R}^n \tag{2}$$

where θ is a threshold that depends on the past activity of the neuron; when the input signal is a stationary stochastic process, θ can be defined as $E[x^2]$ [Intrator, 1990]. We describe the mean evolution of the weights $\vec{m}$ by averaging the equations (2)

$$\dot{m}_i = E[x^2 d_i] - \theta E[x d_i] = B_{ijk} m_j m_k - C_{kl} m_k m_l C_{ij} m_j \tag{3}$$

This approximation is correct if the input signal $x(t)$ has a fast decaying correlation between the past history and the future [Khas'minskii, 1966]. The average BCM equations can be written in the covariant form $\dot{\vec{m}} = \partial\mathcal{E}/\partial\vec{m}$ if one introduces the "energy" [Intrator and Cooper, 1992]

$$\mathcal{E} = \frac{B_{ijk}m_i m_j m_k}{3} - \frac{(C_{ij}m_i m_j)^2}{4} \tag{4}$$

The study of the existence and the properties of critical points of the energy (4) could be not an easy task therefore we propose to simplify the problem by relating the existence of stable fixed point for the system (3) to the existence of local maxima of a cubic function constrained on an ellipsoid in the $\mathbb{R}^n$ space. The main idea is to write the system (3) in the base of the eigenvectors of the covariance matrix C and to show that the nontrivial fixed points correspond to the critical points of the cubic function $f(\vec{y}) = (B_{ijk}y_i y_j y_k)/3$, $\vec{y} \in \mathbb{R}^n$ constrained on the ellipsoid

$$\|\vec{y}\|^2 = \vec{y}\Lambda\vec{y} = 1 \tag{5}$$

where Λ is the diagonal form of the covariance matrix. Moreover the stable fixed points are related to the local maxima of the cubic function $f(\vec{y})$. The detail of the proof are reported in Appendix 1.

2 Analysis of single neuron behaviour

We shall apply the previous method to some explicit cases in order to point out the main properties and limits of the statistical analysis of the input space performed by a BCM neural network. To simplify the calculations, we do not consider the most general case but the results are completely generic. First we show how the selectivity of a BCM neuron is recovered by our approach. Let us consider a BCM neuron which is stimulated by n linearly independent vectors $\vec{v}_k \in \mathbb{R}^n$ $k = 1, .., n$; the input signal is defined

$$\vec{d} = \sum_{k=1}^{n} \xi_k \vec{v}_k \tag{6}$$

where $\vec{\xi}$ is a random vector which take the values on the canonical base $\hat{e}_k$ of $\mathbb{R}^n$ with equal probability. The covariance matrix reads

$$C = \sum_k \sum_h E\{\xi_k \xi_h \vec{v}_k \vec{v}_h\} = \frac{1}{n}\sum_{k=1}^{n} \vec{v}_k \vec{v}_k \tag{7}$$

Since we can always perform the ortogonal change of variables that reduces the matrix (7) to a diagonal form on the initial vectors $\vec{v}_k$, without loss of generality we assume from the beginning that the covariance matrix has a diagonal form Λ and that the eigenvalues of C scale like $1/n$ as a function of the number of degrees of freedom. Then we compute the cubic defined by the third order moments

$$f(\vec{y}) = \frac{1}{3n}\sum_{k=1}^{n} (\vec{y}\cdot\vec{v}_k)^3 \qquad \vec{y} \in \mathbb{R}^n \tag{8}$$

According to the results of the previous section the existence of stable fixed points for the BCM average equation is related of the existence of local maxima of the cubic (8) constrained on the ellipsoid

$$\sum_{j=1}^{n} \lambda_j y_j^2 = 1 \tag{9}$$

where λ_j are the eigenvalues of the covariance matrix. This problem can be easily solved if we introduce the new variables $\vec{y}' = \sqrt{\Lambda}\vec{y}$ and the vectors $\vec{v}'_k = \sqrt{\Lambda^{-1}}\vec{v}_k$ so that the cubic (8) reads

$$f(\vec{y}') = \frac{1}{3n}\sum_{k=1}^{n}(\vec{y}' \cdot \vec{v}'_k)^3 \tag{10}$$

and $\vec{y}'$ are defined on the unit sphere. In Appendix 2 we show that the cubic (10) has always n local maxima on the unit sphere that are defined by the dual base of the vectors $\vec{v}'_k$. In such a case the BCM neuron has a complete selectivity for the space of initial inputs [Castellani et al., 1999].

Let us generalize the previuos model by adding a noise η to the input signal according to

$$\vec{d} = \sum_{k=1}^{n} \xi_k \vec{v}_k + \vec{\eta} \tag{11}$$

where $\vec{\eta}$ is a vector of independent gaussian variables with $E[\vec{\eta}] = 0$ and $E[\vec{\eta}^2] = \sigma^2/nI$ (I is the identity $n \times n$ matrix). We remark that the total noise variance σ^2 does not depend on the number n of the degrees of freedom; as a consequence the noise level on each components scales as $1/n$. In this way, the ratio between the norm of the signal and the norm of the noise (S/N ratio) is independent of the dimensionality. The eigenvalues of the correlation matrix have the same scaling law $1/n$ as a function of the number of degrees of freedoms. A straightforward calculation shows that the covariance matrix reads

$$C = \frac{1}{n}\sum_{k=1}^{n} \vec{v}_k \vec{v}_k + \frac{\sigma^2}{n} I \tag{12}$$

The matrix (12) has the same eigenvectors as the matrix (7) with eigenvalues $\lambda_k + \sigma^2/n$. The equilibrium points are the critical points of the cubic

$$f(\vec{y}') = \frac{1}{3n}\sum_{k=1}^{n}\left[(\vec{y}' \cdot \vec{v}'_k)^3 + 3\sigma^2(\vec{y}' \cdot \vec{v}'_k)\sum_{j=1}^{n}\frac{(\vec{y}'_j)^2}{n\lambda_j + \sigma^2}\right] \tag{13}$$

defined on the unit sphere. To simplify the calculations we assume that the eigenvalues of C are all equal to $(1+\sigma^2)/n$. Then the cubic (13) reads

$$f(\vec{y}') = \frac{1}{3n}\sum_{k=1}^{n}\left[((\vec{y}' \cdot \vec{v}'_k)^3 + \frac{3\sigma^2}{1+\sigma^2}(\vec{y}' \cdot \vec{v}'_k)\right] \tag{14}$$

In Appendix 2, we show that the the approximation of the n dimensional noisy system to a system of n linearly independent neurons holds when the noise does not grow faster than

$$\sigma_c = \sqrt{\frac{n}{2\sqrt{n-1}}}. \tag{15}$$

This is due to the strong law of large numbers which implies that the effect of the noise is averaged over all dimensions. points; if $\sigma > \sigma_c$ the BCM neuron looses its selectivity since the fixed points move into the complex space. Consider now the effect of noise when the eigenvalues of the correlation matrix are all equal to $1/n + \sigma^2$ except for the last one $\lambda_n = (\lambda + \sigma^2)/n$. In such a case the cubic (14) reads

$$f(\vec{y}') = \frac{1}{3n}\sum_{k=1}^{n}\left((\vec{y}' \cdot \vec{v}'_k)^3 +\right.$$

$$3\frac{\sigma^2}{1+\sigma^2}(\vec{y}'\cdot\vec{v}'_k)\left(\sum_{j=1}^{n-1}(\vec{y}'_j)^2+\frac{1+\sigma^2}{\lambda+\sigma^2}(\vec{y}'_n)^2\right)\Bigg) \tag{16}$$

We discuss the existence of local maxima on the unit sphere in Appendix 2. For a fixed value of λ there is a critical value of σ for the existence of n local maxima of the cubic (16) on the unit sphere.

In the first case $\lambda = .5$ we see that for all n the critical value decreases with respect to σ_c and we loose the selecticity for the $\vec{v}_n$ vector; moreover the scaling law with n seems to be different and we do not improve the critical values σ_c increasing n as in the case (15). In the second case $\lambda = 2$ we still have a decreasing of the critical value with respect to σ except for $n \leq 4$ where an increase is observed; in such a case we loose the selectivity for the vectors $\vec{v}_k$ $k = 1, ..., n-1$. The initial behaviour can be explained because for $\lambda > 1$ we have a bigger signal and the selectivity of BCM neuron is enforced, but when n increases the presence of an eigenvector of the covariance matrix (12) with a big eigenvalue decreases the selectivity of the BCM neuron with respect to the other eigenvectors. The scaling law with n appears to be the same as in the first case.

3 Extension to BCM Neural Network

We consider a network of M N-dimensional neurons which are connected by synapses whose weights are given by a symmetric matrix L. The evolution of the weights $\vec{m}_r$ of the r-neuron is

$$\dot{\vec{m}}_r = x_r(x_r-\theta_r)\vec{d} \qquad r = 1, ..., N \tag{17}$$

where the output x_r of the r-neuron is defined $x_r = \vec{m}_r\cdot\vec{d} + L_{rs}x_s$, θ_r is the threshold of the r-neuron ($\theta_r = E[x_r^2]$). We assume that L is time-independent and that the matrix $I - L$ can be inverted. As a consequence the outputs x_r can be written in the form

$$x_s = \Delta_{sr}\vec{m}_r\cdot\vec{d} \tag{18}$$

where we have defined the matrix $\Delta = (I-L)^{-1}$; Δ turns out to be positive defined. If we introduce the modified weights $\vec{m}'_r = \Delta_{rs}\vec{m}_s$ and we average the equations (17) on the space of input signals, the evolution equations read

$$\dot{\vec{m}}'_r = \Delta_{rs}\left(B\vec{m}'^2_s - (\vec{m}'_s C\vec{m}'_s)C\vec{m}'_s\right) \tag{19}$$

where C and B are defined by eq. (1).
The equations (19) can be written in a covariant form $\dot{\vec{m}}'_r = \Delta_{rs}\partial\mathcal{E}/\partial\vec{m}'_s$ where the energy $\mathcal{E}$ is

$$\mathcal{E} = \sum_s \frac{B_{ijk}m'_{si}m'_{sj}m'_{sk}}{3} - \frac{(m'_{sl}C_{ln}m'_{sn})^2}{4} \tag{20}$$

Therefore we can compute the equilibrium solutions of the network by using the same procedure as for the single neuron. The equilibrium solutions correspond to the equilibria of a single BCM neuron and the whole network has $2^{NM}-1$ nontrivial equilibria.
It is not difficult to prove that the stable equilibria are the direct product of the stable solutions of the single neuron equations. Therefore a network of M N-dimensional neurons has N^M stable equilibria.
Finally we remark that the effect of the connections between the different neurons is contained in the matrix Δ that can change the relative stability between the different solutions. More precisely when the weights L are positive the solutions where all the neurons are in the same state turns out to be more stable (i.e. the corresponding eigenvalues of the linearized equations are bigger in absolute value) than the solutions where the neurons are in different states.

4 Conclusions

This paper continues earlier analysis of the BCM neuron and shows that a system with n linearly independent cluster centers and addiditive gaussian noise is governed by analysis of the (simple) case of n linearly independent input vectors. This holds for noise levels on the order of $n^{1/4}$ and implies that when the neuron dimensionality grows, the analysis of the linearly independent case become more appropriate for a larger class of clustered inputs. The mathematical tools which were introduced here, were able to characterize the dependency for a network of neurons and are not limited to Gaussian distributions only, but to general distributions which have the same covariance structure.

Aknowledgements

References

[Bear et al., 1987] Bear, M. F., Cooper, L. N., and Ebner, F. F. (1987). The physiological basis of a theory for synapse modification. *Science*, 237:42–48.

[Bienenstock et al., 1982] Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2:32–48.

[Blais et al., 1998] Blais, B. S., Intrator, N., Shouval, H., and Cooper, L. N. (1998). Receptive field formation in natural scene environments: comparison of single cell learning rules. *Neural Computation*, 10(7):1797–1813.

[Castellani et al., 1999] Castellani, G. C., Intrator, N., Shouval, H., and Cooper, L. N. (1999). Solutions of a BCM learning rule in a network of lateral interacting non-linear neurons. *Network:Comput.Neural Syst.*, 10:111.

[Clothiaux et al., 1991] Clothiaux, E. E., N Cooper, L., and Bear, M. F. (1991). Synaptic plasticity in visual cortex: Comparison of theory with experiment. *Journal of Neurophysiology*, 66:1785–1804.

[Huynh et al., 1998] Huynh, Q., Cooper, L. N., Intrator, N., and Shouval, H. (1998). Classification of underwater mammals using feature extraction based on time-frequency analysis and bcm theory. *IEEE-Signal Processing*, 46(5):1202–1207.

[Intrator, 1990] Intrator, N. (1990). An averaging result for random differential equations. Technical Report 54, Center For Neural Science, Brown University.

[Intrator and Cooper, 1992] Intrator, N. and Cooper, L. N. (1992). Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Networks*, 5:3–17.

[Intrator and Gold, 1993] Intrator, N. and Gold, J. I. (1993). Three-dimensional object recognition of gray level images: The usefulness of distinguishing features. *Neural Computation*, 5:61–74.

[J.L.Doob, 1953] J.L.Doob (1953). *Stochastic Processes*. J.Wiley, New York.

[J.Milnor, 1958] J.Milnor (1958). *Differential Topology*. Princeton.

[Khas'minskii, 1966] Khas'minskii, R. (1966). On stochastic processes defined by differential equations with a small parameter. *Theory Prob. Appl.*, 11:211–228.

[Reisfeld and Yeshurun, 1992] Reisfeld, D. and Yeshurun, Y. (1992). Robust detection of facial features by generalized symmetry. In *Proceedings of the 11th International Conference on Pattern Recognition*, The Hague, Netherlands.

[Shouval et al., 1996] Shouval, H., Intrator, N., Law, C. C., and N Cooper, L. (1996). Effect of binocular cortical misalignment on ocular dominance and orientation selectivity. *Neural Computation*, 8(5):1021–1040.

[Shouval et al., 1997] Shouval, H., Intrator, N., and N Cooper, L. (1997). BCM network develops orientation selectivity and ocular dominance in natural scene environment. *Vision Research*, 37(23):3339–3342.

[Shouval et al., 2000] Shouval, H. Z., Goldberg, D. H., Jones, J. P., Beckerman, M., and Cooper, L. N. (2000). Structured long-range connections can provide a scaffold for orientation maps. *Journal of Neuroscience*, 20(3).

Appendix 1

We study the existence of stable equilibrium for the average BCM equation (3). It is convenient to write the equation (3) in the base of the eigenvectors of the matrix C. Let O be the orthogonal matrix which diagonalizes the correlation matrix according to $\Lambda = OCO^T$, where Λ is a diagonal matrix. We define the new synaptic weights $\vec{w} = O\vec{m}$. The energy (4) takes the form

$$\mathcal{E} = \frac{\beta\vec{w}^3}{3} - \frac{(\vec{w}\Lambda\vec{w})^2}{4} \tag{21}$$

and the averaged equations read

$$\dot{\vec{w}} = \beta\vec{w}^2 - \Lambda\vec{w}\|\vec{w}\|^2 \tag{22}$$

where β are the cubic moments of the new inputs $O\vec{d}$ and $\|\vec{w}\|^2 = \vec{w}\Lambda\vec{w}$ is the metric defined by the correlation matrix. We observe that if an eigenvalue of the correlation matrix is zero the corresponding weight is constant since the r.h.s. of eq. (22) is zero too. Then we can reduce the dimensionality of the system (22) and without loss of generality we assume $\|\vec{w}\| = 0$ if and only if $\vec{w} = 0$.
The equilibrium solutions are defined by the equations

$$\beta\vec{w}^2 - \Lambda\vec{w}\|\vec{w}\|^2 = 0 \tag{23}$$

Since we are interested in the non-trivial solutions of system (23), it is convenient to use the following **Lemma** I:
A vector $\vec{w}^*$ is a non-trivial solution of the system (23) if and only if $\vec{y}^* = \frac{\vec{w}^*}{\|\vec{w}^*\|^2}$ where $\vec{y}^*$ is a non-trivial solution of the system

$$\beta\vec{y}^2 = \Lambda\vec{y} \tag{24}$$

We observe that by definition $\|\vec{y}^*\| = 1/\|n^*\|$ and that the equation (24) has a covariant form $\partial\mathcal{E}/\partial\vec{y} = 0$ if we introduce the reduced energy

$$\mathcal{E} = \frac{\beta\vec{y}^3}{3} - \frac{\|\vec{y}\|^2}{2} \tag{25}$$

Then we prove the **Proposition** I:

let us consider the homogeneous cubic function $f(\vec{y}) = \beta\vec{y}^3/3$ defined on a n-dimensional ellipsoid $\|\vec{y}\|^2 = r$, then each stationary point $\vec{y}^*$ of $f(\vec{y})$ corresponds to a non-trivial equilibrium solution of the system (23) $\vec{w}^* = 3f(\vec{y}^*)/r^4\,\vec{y}^*$

Proof

By using the method of the Lagrange multipliers, we introduce the new function

$$F(\vec{y}) = f(\vec{y}) - \frac{\gamma}{2}(\|\vec{y}\|^2 - r^2) \tag{26}$$

where γ is a real parameter. Then the stationary points of $f(\vec{y})$ are given by the equation

$$\frac{\partial F}{\partial \vec{y}} = \beta \vec{y}^2 - \gamma \Lambda \vec{y} = 0 \qquad \frac{\partial F}{\partial \gamma} = -\frac{1}{2}(\|\vec{y}\|^2 - r^2) = 0 \tag{27}$$

If $\vec{y}^*$ is a solution of eqs. (27), then $\gamma r^2 = 3f(\vec{y}^*)$ and $\vec{w}^* = \gamma/r^2\, \vec{y}^*$ turns out to be a solution of eq. (23). In particular if we choose $r = 1$ and scale the variables according to $\hat{\vec{y}} = \sqrt{\Lambda^{-1}}\vec{y}$, the existence of fixed directions for the BCM equation (3) is related to the stationary points of a homogeneous cubic function $f(\hat{\vec{y}})$ on the unit sphere.

In a non-degenerate situation the system (24) has $2^n - 1$ non-trivial solutions, which corresponds to the maximum number of non-trivial equilibria of a BCM neuron with n synapses. In this case we observe that the cubic function $f(\vec{y})$ has $2(2^n - 1)$ critical points on a sphere due to the ambiguity in the choice of the sign of the stationary points.
The stable equilibrium solutions are the maxima of the energy function (4). To study the stability of the fixed points we linearize the eqs. (23) around an equilibrium solution $\vec{w}^*$ by introducing the variables $\delta\vec{w} = \vec{w} - \vec{w}^*$. The leading terms of eq. read

$$\delta\dot{\vec{w}} = 2\beta\vec{w}^*\delta\vec{w} - \Lambda\delta\vec{w}\|\vec{w}^*\|^2 - 2\Lambda\vec{w}^*(\vec{w}^* \cdot \delta\vec{w}) + \beta\delta\vec{w}^2 + O(\|\delta\vec{w}\|^3) \tag{28}$$

We study the stability of the fixed point $\vec{w}^*$ by considering the time derivative of $\sum_k \delta w_k^2$ which corresponds to a Ljapounov function for the system (28).
We distinguish the cases $\vec{w}^* = 0$ and $\vec{w}^* \neq 0$. In the first case eq. (28) reduces

$$\delta\dot{\vec{w}} = \beta\delta\vec{w}\delta\vec{w} \tag{29}$$

so that the trivial solution is always unstable along the directions $\delta\vec{w}$ which satisfy $f(\delta\vec{w}) > 0$. In the second case we introduce the vector $\vec{y}^* = \vec{w}^*/\|\vec{w}^*\|^2$ and the stability depends on the eigenvalues of the symmetric matrix

$$S = 2\beta\vec{y}^* - \Lambda - 2\|\vec{w}^*\|^2(\Lambda\vec{y}^*)^2 \tag{30}$$

The following **Lemma** II holds:
The equilibrium solution $\vec{w}^*$ is stable if and only if the matrix S is negative definite in the linear space tangent to the ellipsoid

$$\|\vec{y}\|^2 = 1/\|\vec{w}^*\| \tag{31}$$

at the point $\vec{y}^*$.
Proof
Using eq. (24), we get by a direct calculation $S\vec{y}^* = -\Lambda\vec{y}^*$.We decompose any vector $\delta\vec{w}$ according to $\delta\vec{w} = \delta\vec{w}^* + \delta\vec{v}$ where $\delta\vec{w}^*$ is parallel to $\vec{y}^*$ and $\delta\vec{v}$ belongs to the tangent space of the ellipsoid (31) at the point $\vec{y}^*$ (i.e. $\delta\vec{v}\Lambda\vec{y}^* = 0$). The stability of the equilibrium solution $\vec{w}^*$ is a consequence of the inequality

$$\delta\vec{w}S\delta\vec{w} = -\|\delta\vec{w}^*\|^2 + 2\delta\vec{w}^*S\delta\vec{v} + \delta\vec{v}S\delta\vec{v} < 0 \qquad \forall\, \delta\vec{w} \tag{32}$$

According to our hypothesis we have

$$\delta\vec{w}^*S\delta\vec{v} \propto 2\vec{y}^*\beta\vec{y}^*\delta v - 3\vec{y}^*\Lambda\delta\vec{v} = 0 \tag{33}$$

Therefore the inequality $\delta\vec{v}S\delta\vec{v} < 0\ \forall\ \delta\vec{v}$ is a necessary and sufficient condition for the stability.
As a consequence we have **Proposition** II:
Let $\vec{w}^*$ a nontrivial equilibrium position of the system (23), if the eigenvalues of the reduced symmetric matrix

$$\hat{S} = 2\beta\vec{y}^* - \Lambda \tag{34}$$

defined on the tangent space of the ellipsoid (31) at the point $\vec{y}^* \propto \vec{w}^*$, are all negative then $\vec{w}^*$ is a stable equilibrium.

By a straightforward calculation one can see that $\vec{y}^* \hat{S} \vec{y}^* = \|\vec{y}^*\|^2$, so that the matrix $\hat{S}$ at the stable equilibrium solutions $\vec{y}^*$ without any restriction has $n-1$ negative eigenvalues and 1 positive eigenvalue. Then due to topological reasons [J.Milnor, 1958] the number of stable solutions among the $2^n - 1$ equilibria are at most n, that corresponds to the case of maximal selectivity of the BCM neuron. We observe that the matrix $\hat{S}$ is the Hessian matrix of the reduced energy function (25). By using the previous arguments as a **Corollary** to the Proposition I we get:

the equilibrium solution $\vec{y}^*$ of the system (28) defines a stable direction if and only if it is a local maximum of the cubic $f(\vec{y}) = \beta \vec{y}^3/3$ on the ellipsoid $\|\vec{y}\| = 1$.

Appendix 2

We compute the local maxima of the cubic function

$$f(\vec{y}') = \frac{1}{3n} \sum_{k=1}^{n} (\vec{y}' \cdot \vec{v}'_k)^3 \tag{35}$$

where $\vec{y}'$ is defined on the unit sphere of $\mathbb{R}^n$. We introduce the new variables $c_k = \vec{y}' \cdot \vec{v}'_k$ that are related to the output of the BCM neuron. We remark that there exists an invertible linear relation between the c and the $\vec{y}$ variables since we have assumed that the initial vectors $\vec{v}_k$ are independent. Moreover it is straigthforward to check that $\sum_k c_k^2 = n$. In the c variables the cubic (10) has the simple form

$$f(c) = \frac{1}{3n} \sum_{k=1}^{n} c_k^3 \tag{36}$$

constrained on the unit sphere. Let us suppose that $c_n \neq 0$, then by differentiating eq. (36) we get the system

$$c_k^2 + c_n^2 \frac{\partial c_n}{\partial c_k} = 0 \qquad k = 1, ..., n-1 \qquad \frac{\partial c_n}{\partial c_k} = -\frac{c_k}{c_n} \tag{37}$$

Therefore the critical points with $c_n \neq 0$ are computed by the system

$$c_k(c_k - c_n) = 0 \qquad k = 1, ..., n-1 \tag{38}$$

It is easy to check that equation (38) defines an unique local maximum $c_k = 0$ for $k = 1, ..., n-1$ and $c_n = 1$ which corresponds to the choice of synaptic weights that select the n-th input vector $\vec{v}_n$. The neuron output c_j for all the other input vectors is zero. The other local maxima can be computed in the same way.

According to eq. (14), the effect of a gaussian noise added to the input vectors under the hypothesis that the correlation matrix is proportional to the identity reduces to a problem of studying the critical points

$$f(\vec{y}') = \frac{1}{3n} \sum_{k=1}^{n} \left[(\vec{y}' \cdot \vec{v}'_k)^3 + 3\hat{\sigma}^2 (\vec{y}' \cdot \vec{v}'_k) \right] \tag{39}$$

constrained on the unit sphere where

$$\hat{\sigma}^2 = \frac{\sigma^2}{1 + \sigma^2} \tag{40}$$

Introducing the variables $c_k = \vec{y}' \cdot \vec{v}'_k$ the equation (39) reads

$$f(u) = \frac{1}{3n} \sum_{k=1}^{n} c_k \left(c_k^2 + 3\hat{\sigma}^2 \right) \tag{41}$$

with the constraint $\sum_k c_k^2 = n/(1 + n\sigma^2)$ We again assume $c_n \neq 0$ and consider the effect of noise on the selectivity property of the n-th input vector. If we differentiate eq. (41) we get the system

$$(c_n c_k - \hat{\sigma}^2)(c_k - c_n) = 0 \qquad k = 1, ..., n-1 \tag{42}$$

If $c_k < c_n$ we have a local maximum at $c_k = \hat{\sigma}^2/c_n$ where

$$c_n = \frac{n}{(1+\sigma^2)} \left(\sqrt{\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{(n-1)}{n^2}\sigma^4}} \right) \tag{43}$$

In the limit $\sigma \to 0$ this solution converges towards a stable solution $c_k = 0$ for $k = 1, ..., n-1$ and $c_n = n$. It follows that when the noise σ reaches the critical value

$$\sigma_c = \sqrt{\frac{n}{2\sqrt{n-1}}} \tag{44}$$

the solution moves into the complex space. Beyond the critical value (44) the BCM neuron is no more able to develop selectivity for the n-th input vector.