LETTER **Communicated by DeLiang Wang**

Oscillatory Neural Networks with Self-Organized Segmentation of Overlapping Patterns

Thomas Burwick
Thomas.Burwick@neuroinformatik.rub.de
Institut für Neuroinformatik, Ruhr-Universität Bochum, 44306 Bochum, Germany

Temporal coding is considered with an oscillatory network model that generalizes the Cohen-Grossberg-Hopfield model. It is assumed that the frequency of oscillating units increases with stronger and more coherent input. We refer to this mechanism as acceleration. In the context of Hebbian memory, synchronization and acceleration take complementary roles, and their combined effect on the storage of patterns is profound. Acceleration implies the desynchronization that is needed for self-organized segmention of two overlapping patterns. The superposition problem is thereby solved even without including competition couplings. With respect to brain dynamics, we point to analogies with oscillation spindles in the gamma range and responses to perceptual rivalries.

1 Introduction

The segmentation of a pattern that is overlapping with other patterns, each of them memorized according to the Hebbian rule, is a classical challenge and a severe problem of neural network applications. The origin of the problem is easy to understand. Assume that the pattern has been retrieved and is activated. A sufficiently large overlap of this pattern with other patterns may then imply a large input to these patterns. As a consequence, the overlapping patterns may also get activated, and the unambigous segmentation of the first pattern becomes impossible unless some additional information is provided. In this letter, we consider this problem in the context of temporal coding. The additional information is then given by the temporal structure of the active neural units. This should permit distinguishing the separate patterns with respect to coherence. (For an introduction to theoretical aspects related to the memory of overlapping patterns, see Hertz, Krogh, & Palmer, 1991. For an introduction to temporal coding, see von der Malsburg, 1999. For reviews of its growing biological evidence, see Engel & Singer, 2001; Engel, Fries, & Singer, 2001. And for references to models of temporal coding based on oscillatory networks, see Burwick, 2005, 2006.)

In this letter, we present a new approach of dealing with this problem. The approach is a consequence of including an effect that we denote as acceleration. In contrast to the proposals described in Burwick (2005), the approach of this letter does not need competition couplings. Acceleration means that the frequency of oscillating neural units in a network increases with stronger and more coherent input. This is obviously an effect that may be in agreement with fundamental assumptions about neural behavior. It is interesting to see that it also arises naturally from a formal point of view. The consequences, however, may be surprising with respect to the storage of overlapping patterns. Synchronization and acceleration take complementary roles where acceleration introduces a desynchronization that implies self-organized segmentation of overlapping patterns. Given the problem that was described above, such a property may be an essential feature, and it is certainly worthwhile to study its consequences.

The relevance of temporal coding for information processing results from its relevance for the formation of assemblies. These are distributed representations of objects, where different parts of the network correspond to different features of the object (Hebb, 1949). Temporal coding proposes that the formation of these assemblies is based on temporal correlation of the corresponding parts of the network (von der Malsburg, 1981); see also (von der Malsburg, 1986). The challenge of linking the different features is also referred to as the binding problem. The presence of the binding problem is particularly obvious with some psychophysical phenomena where the brain fails to solve it. (In that respect, see the review on illusory conjunctions in Wolfe & Cave, 1999.) If perception is related to the storage and recall of patterns, as realized with Hebbian memory, then the problem of overlapping patterns will appear whenever different patterns share some features. A particular class of temporal coding applications is figure-ground segregation. As an example, see Shareef, Wang, and Yagel (1999) for an application to medical image segmention. Other applications are described in Wang, Freeman, Kozma, Lozowski, & Minai (2004).

We implement temporal coding with an oscillatory network model that is obtained as complex-valued generalization of the Cohen-Grossberg-Hopfield (CGH) model (Burwick, 2005). In section 2, the model is described, and acceleration is introduced by allowing for a complex phase of the couplings. In section 3, the couplings are chosen according to Hebbian memory, and the dynamics is related to coherences of the patterns. In section 4, the complementary role of synchronization and acceleration becomes obvious from studying the example of two overlapping patterns. In section 5, we comment on the possible inclusion of acceleration in the context of networks with nonsinusoidal couplings, in particular, networks of relaxation oscillators. In section 6, we briefly comment on fast implementations. Section 7 contains the summary.

2 Oscillatory Neural Networks with Synchronization and Acceleration

In section 2.1, the model that we use for the discussion in this letter is introduced. It may be fomulated as a complex-valued gradient system. In section 2.2, we discuss the role of complex-valued potential functions for oscillatory systems and their relation to minimization principles. In section 2.3, we give the potential function that implies the model. In section 2.4, this is then used to derive the exact forms of the couplings in terms of real coordinates. These will be given for all-order mode couplings. For the main arguments of this letter, however, it is sufficient to consider only the first-mode couplings that are specified in section 2.5.

2.1 The Model. The CGH model (Cohen & Grossberg, 1983; Hopfield, 1984) may be extended to an oscillatory network model that allows implementing features of temporal coding (Burwick, 2005); see also (Burwick, 2006). Given a network with N units, where each unit k is described in terms of the real-valued u_k and phase θ_k, $k = 1, \ldots, N$, the oscillatory system is of the form

$$\tau \frac{du_k}{dt} = \Lambda(u_k)\left(I_k - u_k + \frac{1}{N}\sum_{l=1}^{N} w_{kl}(u,\theta)V_l\right), \tag{2.1a}$$

$$\tau \frac{d\theta_k}{dt} = \tau\omega_k(u,\theta) + \underbrace{\frac{1}{N}\sum_{l=1}^{N} s_{kl}(u,\theta)V_l}_{\text{synchronization terms}}, \tag{2.1b}$$

where t is the time, τ is a timescale, I_k is an external input, and the amplitude V_k is related to u_k via the activation function g, chosen as

$$V_k = g(u_k) = \frac{1}{2}(1 + \tanh(u_k)). \tag{2.2}$$

In this letter, the frequency terms are

$$\omega_k(u,\theta) = \omega_{1,k} + \omega_{2,k}V_k + \underbrace{\frac{1}{N}\sum_{l=1}^{N} \Delta\omega_{kl}(u,\theta)V_l}_{\text{acceleration terms}}, \tag{2.3}$$

where the $\omega_{1,k}$ are the eigenfrequencies of the oscillators, and the $\omega_{2,k}$ parameterize the shear terms.

In the following sections, equation 2.1 with explicit expressions for the couplings w, s, and $\Delta\omega$ and the scaling factor Λ will be derived from a complex-valued gradient system. In section 2.3, we make an ansatz for a complex-valued potential function $\mathcal{L}$ and postulate that the dynamics takes the form

$$\tau \frac{dz_k}{dt} = -\frac{\partial \mathcal{L}}{\partial \bar{z}_k}, \tag{2.4}$$

where the complex coordinates are given by

$$x_k = z_k^2 = V_k \exp(i\theta_k), \qquad \bar{x}_k = \bar{z}_k^2 = V_k \exp(-i\theta_k). \tag{2.5}$$

Due to equation 2.2, the dynamics is restricted to the punctured unit disk,

$$0 < |x_k| = z_k \bar{z}_k = V_k = g(u_k) < 1. \tag{2.6}$$

(Given the quadratic dependencies in equation 2.5, using coordinates z_k, $\bar{z}_k$ corresponds to a formulation on the two-sheet Riemann surface.)

Equation 2.1 generalizes the classical CGH model by introducing phase and amplitude dependencies of the couplings in equation 2.1a and phase dynamics in equation 2.1b. The sum in equation 2.1b describes the synchronization terms, while the sum in equation 2.3 gives what we denote as acceleration terms. We consider only excitatory couplings. It will then be obvious from the explicit form of the s_{kl} that the synchronization terms have a synchronizing effect on the phases. The explicit form of the $\Delta\omega_{kl}$ will make obvious that the corresponding terms in equation 2.3 imply an increase in phase velocity of the units for stronger and more coherent input from the other units. We refer to this effect as acceleration.

The system of equation 2.1 that results from equation 2.4 will agree with the model presented in Burwick (2005) except for the presence of the acceleration terms (and the related modification of the couplings w). It turns out that the inclusion of acceleration terms is not a minor modification. Below, when formulating equation 2.1 as a complex-valued gradient system in sections 2.3 and 2.4, we find that the acceleration terms result from a complex phase of the coupling parameters. In the next section, we argue that including such complex-valued couplings is a natural choice in the context of oscillatory systems. In fact, complex-valued couplings are well-known ingredients of several physical systems (e.g., see Kuramoto, 1984). Given the reasonable interpretation in the context of neural systems, it may be oversimplifying to neglect the effects related to acceleration. Our examples will demonstrate that these effects are indeed substantial.

2.2 Complex-Valued Gradient Systems and Their Relation to Minimization Principles. In section 2.1, we mentioned that equation 2.1 may be derived as the gradient system of equation 2.4. A crucial aspect of this letter is that we allow for a complex-valued potential function $\mathcal{L}$,

$$\mathcal{L} = \mathcal{R} + i\,\Omega, \tag{2.7}$$

where $\mathcal{R}$ is the real and Ω the imaginary part. The acceleration terms result from an imaginary part of the coupling parameters, while the discussion in Burwick (2005) was focused on real-valued couplings.

Before going into the details of deriving equation 2.1 from equation 2.4, we add two comments on the imaginary part Ω of $\mathcal{L}$. In the first comment, we emphasize that such imaginary parts are natural for the most fundamental oscillator models. Consider a single oscillator with a complex-valued second-order potential function given by

$$\mathcal{L}_0 = -\frac{1}{2}\left(z\bar{z} - \frac{1}{2}(z\bar{z})^2 + i\tau\left(\omega_1 z\bar{z} + \frac{\omega_2}{2}(z\bar{z})^2\right)\right), \tag{2.8}$$

where ω_1, ω_2 are real parameters and τ is the timescale. Assuming a gradient dynamics like the one in equation 2.4, one obtains

$$\tau\frac{dz}{dt} = -\frac{\partial\mathcal{L}_0}{\partial\bar{z}} = \frac{z}{2}(1 - z\bar{z} + i\tau\,(\omega_1 + \omega_2 z\bar{z})). \tag{2.9}$$

With coordinates like the ones in equation 2.5 (without the index k), we may use

$$\frac{dz}{dt} = \frac{z}{2}\left(\frac{1}{V}\frac{dV}{dt} + i\,\frac{d\theta}{dt}\right) \tag{2.10}$$

to arrive at the real coordinate version of equation 2.9:

$$\tau\frac{dV}{dt} = V(1 - V), \tag{2.11a}$$

$$\frac{d\theta}{dt} = \omega_1 + \omega_2 V. \tag{2.11b}$$

This is then recognized as a basic oscillator model. The ω_1 is the eigenfrequency and $\omega_2 V$ is a shear term. The shear term implies a larger phase velocity for a larger amplitude. The oscillator amplitude approaches $V = 1$, and the frequency approaches $\omega = \omega_1 + \omega_2$. Thus, the frequency is determined by the imaginary part of $\mathcal{L}_0$. It should therefore be seen as a natural step to allow for an imaginary part Ω also in equation 2.7. Moreover, given

this simple example, it should not be surprising that the frequency dependency in equation 2.3 will result from such an imaginary part.

The second comment is on the relation of the imaginary part Ω to the minimization principle as it appears in the classical CGH model. Using

$$\frac{\partial \mathcal{L}}{\partial z_k} = \frac{\partial \overline{\mathcal{L}}}{\partial z_k} + 2i\frac{\partial \Omega}{\partial z_k} = -\tau \frac{d\bar{z}_k}{dt} + 2i\frac{\partial \Omega}{\partial z_k}, \tag{2.12}$$

one obtains

$$\begin{aligned}\frac{d\mathcal{L}}{dt} &= \sum_{k=1}^{N}\left(\frac{dz_k}{dt}\frac{\partial \mathcal{L}}{\partial z_k} + \frac{d\bar{z}_k}{dt}\frac{\partial \mathcal{L}}{\partial \bar{z}_k}\right) \\ &= -2\tau \sum_{k=1}^{N}\left(\left|\frac{dz_k}{dt}\right|^2 - \frac{i}{\tau}\frac{dz_k}{dt}\frac{\partial \Omega}{\partial z_k}\right).\end{aligned} \tag{2.13}$$

With $\Omega \equiv 0$, the system will have $d\mathcal{L}/dt \leq 0$, and local minima will be approached. With $\Omega \neq 0$, this may no longer be the case.

The examples of section 3 will indeed show that the real part of $\mathcal{L}$, that is, the analog of the CGH Lyapunov function, is not decreasing for all times if $\Omega \neq 0$. However, this should not be taken as a motivation to eliminate the imaginary part Ω. After all, such terms are natural ingredients of oscillatory systems. In fact, we will observe that this deviation from the classical picture of a monotonically decreasing Lyapunov function has a well-justified functional relevance. This will be obvious when considering the examples in section 4.

2.3 The Model as a Complex-Valued Gradient System. The dynamical system of equation 2.1 is obtained from equation 2.4 by choosing

$$\mathcal{L} = \mathcal{P} + \mathcal{W}, \tag{2.14}$$

where $\mathcal{P}$ describes the oscillator model of the single units,

$$\begin{aligned}\mathcal{P} = &\frac{1}{2}\sum_{k=1}^{N}\left(z_k\bar{z}_k \ln(z_k\bar{z}_k) + (1 - z_k\bar{z}_k)\ln(1 - z_k\bar{z}_k)\right) \\ &- \sum_{k=1}^{N}\left(J_{1,k}z_k\bar{z}_k + \frac{1}{2}J_{2,k}(z_k\bar{z}_k)^2\right),\end{aligned} \tag{2.15}$$

and $\mathcal{W}$ describes the couplings of the network,

$$\mathcal{W} = -\frac{1}{2N}\sum_{k,l=1}^{N} h_{kl}\sum_{\mu,\nu=1}^{H}\left(a_{\mu\nu}\,(z_k\bar{z}_k)^{\mu}\,(z_l\bar{z}_l)^{\nu} + \underbrace{\widetilde{b}_{\mu\nu}\,(z_k z_k)^{\mu}\,(\bar{z}_l\bar{z}_l)^{\nu}}_{\text{phase-couplings}}\right) \tag{2.16}$$

(Burwick, 2005). The $h_{kl} > 0$ are the weights of the network and assumed to be symmetric. In section 3, they will be specified according to Hebbian memory. The $a_{\mu\nu} > 0$ describe the coupling strengths of the different modes. For equation 2.16, it was assumed that these two dependencies factorize. For the coefficients in equation 2.15, we set

$$J_{1,k} = I_k + \frac{i}{2}\,\tau\omega_{1,k}, \quad J_{2,k} = \frac{i}{2}\,\tau\omega_{2,k}, \tag{2.17}$$

and for the phase-coupling coefficient in equation 2.16, we set

$$\widetilde{b}_{\mu\nu} = b_{\mu\nu} + ic_{\mu\nu}. \tag{2.18}$$

The coefficients $a_{\mu\nu}$, $b_{\mu\nu}$, and $c_{\mu\nu}$ are assumed to be real and symmetric with

$$h_{kl}a_{\mu\nu} = h_{lk}a_{\nu\mu},\ h_{kl}b_{\mu\nu} = h_{lk}b_{\nu\mu},\ \text{and}\ h_{kl}c_{\mu\nu} = h_{lk}c_{\nu\mu}. \tag{2.19}$$

The couplings $c_{\mu\nu}$ imply the acceleration terms, as may be seen in the next section.

2.4 The Dynamics in Real Coordinates. Using the function $\mathcal{L}$ of section 2.3 in equation 2.4, the resulting dynamics may be formulated in terms of the real coordinates given by equation 2.5. These coordinates imply

$$\frac{1}{z_k}\frac{dz_k}{dt} = \frac{1}{2}\left(\frac{1}{V_k}\frac{dV_k}{dt} + i\,\frac{d\theta_k}{dt}\right) \tag{2.20}$$

$$= (1 - g(u_k))\,\frac{du_k}{dt} + \frac{i}{2}\frac{d\theta_k}{dt}, \tag{2.21}$$

using

$$dg(u)/du = 2g(u)(1 - g(u)), \tag{2.22}$$

and the derivative is given by

$$\frac{1}{z_k}\frac{\partial}{\partial\bar{z}_k} = \frac{\partial}{\partial V_k} + \frac{i}{V_k}\frac{\partial}{\partial\theta_k}. \tag{2.23}$$

We also need to apply the inverse g^{-1} of the activation given in equation 2.2:

$$u_k = g^{-1}(V_k) = \frac{1}{2} \ln \frac{V_k}{1 - V_k}. \tag{2.24}$$

Equations 2.4 and 2.14 to 2.19 then imply equation 2.1, where the scaling factor is

$$\Lambda(u_k) = \frac{1}{1 - V_k} \tag{2.25}$$

(see Burwick, 2005, for a discussion of this factor), and the coupling terms are given by

$$w_{kl}(u, \theta) = h_{kl} \sum_{\mu,\nu=1}^{H} \mu \, W_{\mu\nu} \, (\nu\theta_l - \mu\theta_k) \; V_k^{\mu-1} V_l^{\nu-1}, \tag{2.26}$$

$$s_{kl}(u, \theta) = h_{kl} \sum_{\mu,\nu=1}^{H} 2\mu \, S_{\mu\nu} \, (\nu\theta_l - \mu\theta_k) \; V_k^{\mu-1} V_l^{\nu-1}, \tag{2.27}$$

$$\Delta\omega_{kl}(u, \theta) = h_{kl} \sum_{\mu,\nu=1}^{H} 2\mu \, C_{\mu\nu}(\nu\theta_l - \mu\theta_k) \, V_k^{\mu-1} V_l^{\nu-1}, \tag{2.28}$$

with

$$\begin{aligned} W_{\mu\nu}(\theta) &= a_{\mu\nu} + b_{\mu\nu} \cos(\theta) - c_{\mu\nu} \sin(\theta), \\ S_{\mu\nu}(\theta) &= b_{\mu\nu} \sin(\theta), \\ C_{\mu\nu}(\theta) &= c_{\mu\nu} \cos(\theta). \end{aligned}$$

For the sake of completeness, we have given this general dynamics. However, introducing acceleration and its effect does not require the inclusion of higher-mode couplings. In the next section, we therefore give the simpler form that describes the dynamics with first-mode couplings only.

2.5 The Dynamics with First-Mode Couplings Only. We set

$$a_{11} = a \quad \text{and} \quad \widetilde{b}_{11} = \frac{1}{2}(\sigma + i\tau\omega_3), \tag{2.29}$$

where $a > 0$, $\sigma > 0$, and $\omega_3 > 0$ are real and assume that all other higher-mode couplings vanish. In terms of real coordinates and using only the

first-mode couplings, the real and imaginary parts of $\mathcal{L} = \mathcal{R} + i\,\Omega$ are then given by

$$\mathcal{R} = \frac{1}{2}\sum_{k=1}^{N}\left(V_k \ln V_k + (1 - V_k)\ln(1 - V_k) - 2I_{1,k}V_k\right) - \frac{1}{2N}\sum_{k,l=1}^{N} h_{kl}\left(a + \frac{\sigma}{2}\cos(\theta_k - \theta_l)\right)V_k V_l, \tag{2.30}$$

$$\Omega = -\frac{\tau}{2}\sum_{k=1}^{N}\left(\omega_{1,k}V_k + \frac{1}{2}\omega_{2,k}V_k^2\right) - \frac{\tau\omega_3}{4N}\sum_{k,l=1}^{N} h_{kl}\cos(\theta_k - \theta_l)V_k V_l. \tag{2.31}$$

With the examples of section 4, we demonstrate how the dynamics of the real part relates to the retrieval of the patterns.

Using the parameter choices of equation 2.29, equation 2.1 reduces to the first-mode coupling part, where the couplings have no amplitude dependency:

$$\tau\frac{du_k}{dt} = \Lambda(u_k)\left(I_k - u_k + \frac{1}{N}\sum_{l=1}^{N} w_{kl}(\theta_l - \theta_k)V_l\right), \tag{2.32a}$$

$$\tau\frac{d\theta_k}{dt} = \tau\omega_k(u,\theta) + \underbrace{\frac{1}{N}\sum_{l=1}^{N} s_{kl}(\theta_l - \theta_k)V_l}_{\text{synchronization terms}}, \tag{2.32b}$$

with

$$\omega_k(u,\theta) = \omega_{1,k} + \omega_{2,k}V_k + \underbrace{\frac{1}{N}\sum_{l=1}^{N}\Delta\omega_{kl}(\theta_l - \theta_k)V_l}_{\text{acceleration terms}}, \tag{2.33}$$

and

$$w_{kl}(\theta) = h_{kl}\left(a + \frac{\sigma}{2}\cos(\theta) - \frac{\tau\omega_3}{2}\sin(\theta)\right), \tag{2.34}$$

$$s_{kl}(\theta) = h_{kl}\,\sigma\sin(\theta), \tag{2.35}$$

$$\Delta\omega_{kl}(\theta) = h_{kl}\,\omega_3\cos(\theta). \tag{2.36}$$

For introducing acceleration and its consequences, it is sufficient to consider only these first-mode couplings. Therefore, the simple form of equations 2.32 will be used for the remainder of this letter.

3 Hebbian Memory and Pattern Coherences

So far, we have not specified the couplings h_{kl}. In section 3.1, these are defined in terms of patterns that are stored according to Hebbian memory. In section 3.2, the coherence measures for patterns are introduced. These serve to describe the dynamics, in particular the ω_k of equations 2.33, in terms of activities, coherences, and phases of the patterns. In section 3.3, we begin to discuss the complementary role of synchronization and acceleration, which will become more apparent with the examples in the following section.

3.1 Hebbian memory. Consider the storage of P patterns ξ_k^p, with $p = 1, \ldots, P$ and $k = 1, \ldots, N$. In this letter, it will be sufficient to assume that $\xi_k^p \in \{0, 1\}$, where 1 (0) corresponds to an on-state (off-state). We refer to the units k with $\xi_k^p = 1$ as the units of pattern p. In equation 2.1, Hebbian memory may be used that is defined by

$$h_{kl} = \sum_{p=1}^{P} \lambda_p \, \xi_k^p \xi_l^p, \tag{3.1}$$

with $\lambda_p > 0$. The λ_p give the weights for patterns p. Let us also introduce

$$\Lambda_p = \rho_p \lambda_p, \tag{3.2}$$

where

$$\rho_p = \frac{N_p}{N}, \tag{3.3}$$

and N_p is the number of nonvanishing units of pattern p. The Λ_p may be interpreted as the weight of pattern p that also includes the pattern density ρ_p, that is, with equal λ_p, a larger pattern contributes more than a smaller pattern.

3.2 Dynamics in Terms of Activities, Coherences, and Phases of Patterns. The collective dynamics of the network may be described in terms of activity R, coherence C, and phase Ψ of each pattern, given by

$$R_p = \frac{1}{N_p} \sum_{k=1}^{N} \xi_k^p \, V_k, \tag{3.4}$$

$$Z_p = C_p \exp(i \Psi_p) = \frac{1}{N_p} \sum_{k=1}^{N} \xi_k^p V_k \exp(i \theta_k), \tag{3.5}$$

for $p = 1, \ldots, P$. Let us define

$$A_{pk}(u, \theta) = C_p \cos\left(\Psi_p - \theta_k\right), \tag{3.6}$$

$$S_{pk}(u, \theta) = C_p \sin\left(\Psi_p - \theta_k\right), \tag{3.7}$$

where $\left|A_{pk}\right| < 1$, $\left|S_{pk}\right| < 1$. For example, the synchronizing terms in equation 2.32b take the form

$$\begin{aligned}\frac{\sigma}{N}\sum_{l=1}^{N} h_{kl}\sin(\theta_l - \theta_k)V_l &= \frac{\sigma}{N}\mathrm{Im}\sum_{p=1}^{P}\sum_{l=1}^{N}\lambda_p\,\xi_k^p\xi_l^p\exp\left(i(\theta_l - \theta_k)\right)V_l \\ &= \sigma\,\mathrm{Im}\sum_{p=1}^{P}\xi_k^p\lambda_p\left\{\frac{\rho_p}{N_p}\sum_{l=1}^{N}\xi_l^p V_l\, e^{i\theta_l}\right\}e^{-i\theta_k}, \\ &= \sigma\sum_{p=1}^{P}\xi_k^p\,\Lambda_p\,S_{kp}(u, \theta).\end{aligned} \tag{3.8}$$

Steps analogous to that of equation 3.8 may be used to reformulate the other coupling terms of equations 2.32. This leads to an alternative form of equations 2.32 that relates the dynamics to pattern activities, coherences, and phases:

$$\tau(1 - V_k)\frac{du_k}{dt} = I_k - u_k + a\,Q_k(u) + \frac{\sigma}{2}A_k(u, \theta), \tag{3.9a}$$

$$\tau\frac{d\theta_k}{dt} = \tau\omega_k(u, \theta) + \sigma\,S_k(u, \theta), \tag{3.9b}$$

where

$$\omega_k(u, \theta) = \omega_{1,k} + \omega_{2,k}V_k + \omega_3 A_k(u, \theta), \tag{3.10}$$

and

$$Q_k(u) = \sum_{p=1}^{P}\xi_k^p\,\Lambda_p\,R_p(u), \tag{3.11}$$

$$A_k(u, \theta) = \sum_{p=1}^{P}\xi_k^p\,\Lambda_p\,A_{pk}(u, \theta), \tag{3.12}$$

$$S_k(u, \theta) = \sum_{p=1}^{P}\xi_k^p\,\Lambda_p\,S_{pk}(u, \theta). \tag{3.13}$$

Equation 3.11 gives the classical (i.e., phase-independent) contribution, while equations 3.12 and 3.13 describe, respectively, acceleration and synchronization terms. Due to the factor ξ_k^p in these definitions of Q_k, A_k, and S_k, every unit k couples only to patterns p that are active at unit k.

The tendency to synchronize two units k and l has to overcome any desynchronizing tendency that arises with different value of ω_k and ω_l. The synchronizing couplings in equation 3.9b are given by the σ-dependent terms. According to equations 3.13 and 3.7, these σ-dependent terms force each unit k to synchronize with the phases Ψ_p of patterns that are nonzero at this unit, that is, patterns p with $\xi_k^p = 1$. These patterns compete for the influence on unit k with a strength that is given by the product of the weight λ_p of the pattern, the pattern density ρ_p, and its coherence C_p. Thus, for example, with $\Lambda_p = \rho_p / P$ for every p, a less coherent or smaller pattern will have less influence on the unit, and a highly coherent and larger pattern will have a stronger influence. With $\Lambda_p = 1$, the influence of each pattern will depend only on its coherence C_p.

Notice also that the σ-dependent terms of equation 3.9a support the synchronization. With respect to equation 3.9a, the coupling contribution from pattern p is maximal for $\theta_k = \Psi^p$ and minimal for $\theta_k = \Psi^p + \pi$. This should be expected for couplings that favor synchronization. The a-dependent terms in equation 3.9a describe the classical CGH couplings that lead to on- and off-state amplitudes (together with the the phase-dependent modification of the couplings).

3.3 Avoiding Global Coherence: Complementarity of Synchronization and Acceleration. With the examples of section 4, we consider choices for the eigenfrequency and shear coefficients that obey

$$\omega_{1,k} = \omega_1, \tag{3.14}$$

$$\omega_{2,k} = \omega_2, \tag{3.15}$$

for some ω_1, ω_2. This is done to study the desynchronizing effect of acceleration without confusing it with desynchronizing effects that result from distributed eigenfrequency or shear coefficients. In this section, we begin the discussion of what dynamics should be expected.

In section 1, we described the superposition problems of overlapping patterns. This is not solved simply by introducing phases. Consider the situation without acceleration: $\omega_3 = 0$. Due to the synchronizing terms, every phase of a connected network component, given the choices of equations 3.14 and 3.15, will then approach the same value θ and arrive at a stable configuration of global coherence (this will be demonstrated with examples 1a and 2a in section 4). The superposition problem is present as for the classical case without phases.

The situation changes drastically, however, in the presence of acceleration: with $\omega_3 > 0$. Then global coherence may no longer be stable in case of overlapping patterns. A simple argument for this is the following. Consider the situation of global coherence. Using equations 3.14 and 3.15, and assuming that the units are on-state with a phase θ, one obtains

$$V_k \simeq 1, \quad \theta_k = \Psi \text{ for every } k : \frac{d\theta_k}{dt} \simeq \omega_1 + \omega_2 + \omega_3 \sum_p \xi_k^p \Lambda_p. \qquad (3.16)$$

With $\omega_3 > 0$, it follows immediately from equation 3.16 that global coherence may not be stable in case of overlapping patterns, since the $\sum \xi_k^p \Lambda_p$ will not be the same for different units k that participate in different sets of patterns (this will be demonstrated with examples 1b and 2b in section 4). Different parts of the patterns with overlap receive different inputs from the different patterns and thereby are subject to different accelerations. Thus, different parts of the pattern will have different phase velocities that destroy the coherence.

It is then obvious that synchronization and acceleration take complementary roles: synchronization tends to establish coherence of the network, but this will then imply acceleration effects that desynchronize overlapping patterns. The following examples of two overlapping patterns will show how this interplay of synchronization and acceleration may help to overcome the superposition problem that plagues Hebbian storage of overlapping patterns without acceleration.

4 Examples: Self-Organized Segmentation of Two Overlapping Patterns

In this section, we demonstrate the combined effect of synchronization and acceleration terms by considering numerical examples for the case of two overlapping patterns. In section 4.1, we describe the network architectures, some parameters, the numerical approach, and initial conditions that are used. Then we compare the dynamics with and without the acceleration terms. In section 4.2, this is done for the case that one pattern is dominating, $\Lambda_1 > \Lambda_2$. In section 4.3, neither of the patterns is dominating: $\Lambda_1 = \Lambda_2$. In section 4.4, we comment on the corresponding behavior of the potential functions (Lyapunov functions in case of vanishing acceleration). In sections 4.5 and 4.6, network behavior is traced back to the dynamics of overlapping and nonoverlapping parts ("patches") of the patterns. The observed dynamics corresponds to oscillation spindles. This is illustrated in section 4.7, where the effect of changing the shear parameter ω_2 is demonstrated.

4.1 Network Architectures and Numerical Setup. We consider $P = 2$ patterns for networks with $N = 36$ (examples 1 and 2) and $N = 30$ units

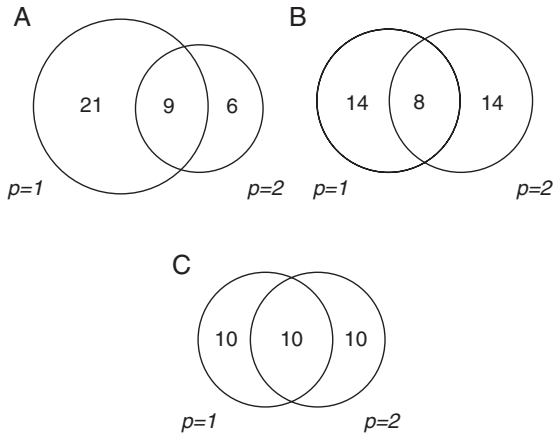

Figure 1: The three architectures of the examples in section 4. (A) Example 1. The first of the two patterns has $N_1 = 30$ nonvanishing units and the second $N_2 = 15$. The patterns overlap at 9 units. (B, C) The analog architectures for examples 2 and 3. The example networks have (A, B) $N = 36$ and (C) $N = 30$ units.

(example 3) (see Figure 1). The patterns in the Hebbian couplings of equation 3.1 are equally weighted with $\lambda_p = 1/2$ for $p = 1, 2$. Thus, $\Lambda_p = \rho_p/2$. The coupling parameters a, σ, and ω_3 will be specified for each example. Let us use equation 2.2 to define a shifted input I'_k through

$$I_k + \frac{a}{N}\sum_{l=1}^{N} h_{kl} g(u_l) = I'_k + \frac{a}{2N}\sum_{l=1}^{N} h_{kl} \tanh(u_l). \tag{4.1}$$

The input is then chosen so that $I'_k = 0$. This implies that $u_k = 0$: $V_k = 1/2$ is a threshold for (decoherent) on- and off-states of unit k (see Burwick, 2005).

We use identical eigenfrequencies and shear parameters as specified in equations 3.14 and 3.15. This ensures that any desynchronizing effect arises from the acceleration terms. With examples 1 and 2, we choose $\omega_1 = \omega_2 = 0$. The phase velocity will then be generated only by acceleration and synchronization. Only example 3 will use $\omega_2 > 0$. This example will then serve to demonstrate the effect of nonvanishing shear terms and the relation to oscillation spindles.

The numerical study is performed with a real parameter $0 < \varepsilon^2 \ll 1$, and modifying the effective timescale $\tau(1 - V_k)$ of equation 2.1a with equation 2.25 by using this regularization parameter for replacement:

$$\tau(1 - V_k) \to \tau(1 - V_k + \varepsilon^2) \tag{4.2}$$

(see Burwick, 2005). The system of equation 2.1 is then discretized with the Euler approximation, where the time step is given by $dt = \varepsilon^2\tau$. The following examples are based on $\varepsilon^2 = 0.001$. The initial values are chosen so that initial amplitudes are small, while phases are equally distributed. Examples 1 and 2 use the same initial conditions (example 3 has different N).

4.2 Examples 1: With a Dominating Pattern. We study the network with the architecture displayed in Figure 1A. Due to our choice of $\Lambda_p = \rho_p/2$ (see section 4.1), a comparison gives

$$\Lambda_1 = \rho_1/2 = 10/24 > \Lambda_2 = \rho_2/2 = 5/24. \tag{4.3}$$

In consequence, one may expect that the first pattern is dominating the second. Coupling parameters are chosen as $a/N = 2$, $\sigma = a$. Since we are interested in the acceleration effect, we compare two examples. Example 1a uses $\omega_3 = 0$, and example 1b uses $\omega_3 = \pi\sigma/(2\tau) > 0$. For the resulting dynamics in terms of the pattern coherences, see the top diagrams of Figure 2.

4.2.1 Example 1a. In case of $\omega_3 = 0$, we find that both patterns become coherent, and due to the overlap, we may conclude that global coherence is reached (see the top diagram of Figure 2A). The individual patterns may no longer be distinguished by activity or coherence. This is the superposition problem for overlapping patterns.

4.2.2 Example 1b. In case of $\omega_3 = \pi\sigma/(2\tau) > 0$, the situation changes decisively. Now, only pattern 1 becomes coherent, while pattern 2 alternates between moments of coherence and periods of decoherence (see the top diagram of Figure 2B). The dominating pattern 1 has been segmented from the nonoverlapping part of pattern 2. No superposition problem arises.

4.3 Examples 2: Without a Dominating Pattern—Pattern Switching. Next, we consider the architecture of Figure 1B: $N_1 = N_2 = 22$ with 8 overlapping units. The pattern weights are now equal:

$$\Lambda_1 = \Lambda_2 = \rho_1/2 = 11/36. \tag{4.4}$$

Neither of the patterns may be expected to dominate the other. Coupling parameters are $a/N = 2$, $\sigma = a/2$. Again, we compare the cases with and without acceleration (see Figure 3).

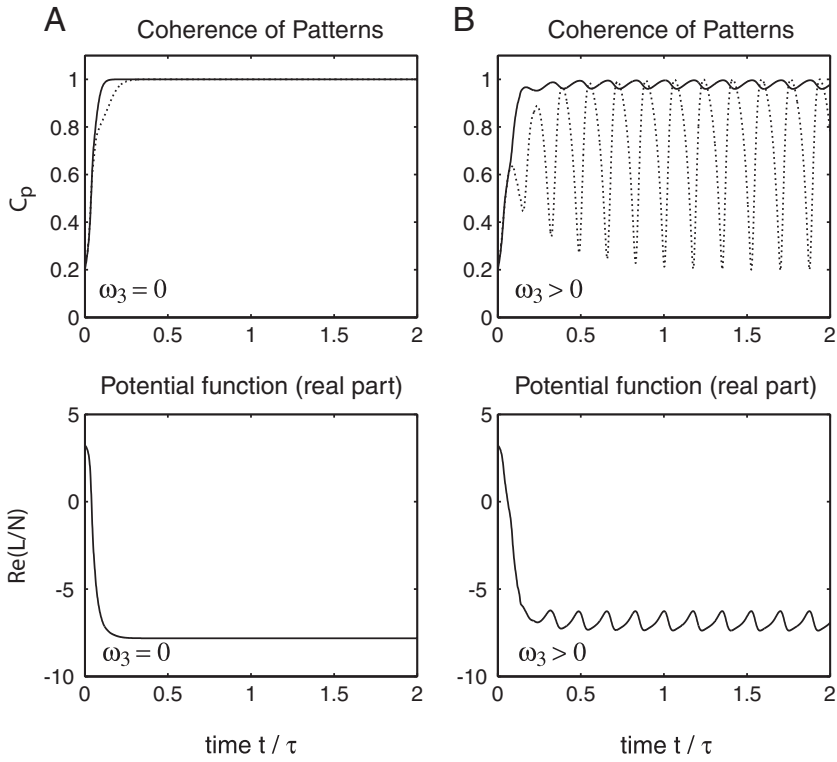

Figure 2: Examples 1a and 1b, with the network architecture of Figure 1A. (A) Example 1a: Without acceleration, $\omega_3 = 0$. (B) Example 1b: with acceleration, $\omega_3 > 0$. The top diagrams show the coherences for patterns $p = 1$ (solid line) and $p = 2$ (dotted line). The bottom diagrams give the real part of the potential function (i.e., Lyapunov function in case of $\omega_3 = 0$). See the discussion in sections 4.2 and 4.4.

4.3.1 Example 2a. In case of $\omega_3 = 0$, we find the superposition problem for overlapping patterns as with examples 1a (see the top diagram of Figure 3A).

4.3.2 Example 2b. In case of $\omega_3 = \pi\sigma/\tau > 0$, the effect of the acceleration terms is again profound (see the top diagram of Figure 3B). The dynamics changed qualitatively and is now different from example 1b. Again, each pattern is retrieved coherently but not with a steady state. Instead, an alternating series of states is generated, where each state corresponds to the coherent retrieval of one individual pattern, while the other pattern approaches decoherence. We refer to the kind of behavior that is illustrated with Figure 3C as pattern switching. The superposition problem of

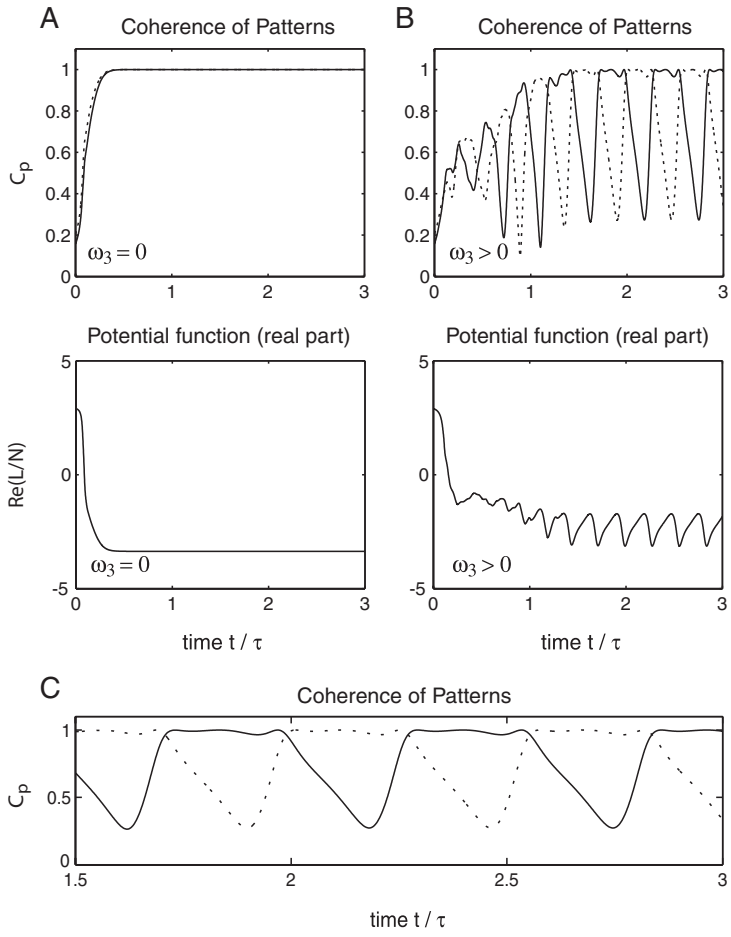

Figure 3: Examples 2a and 2b, with the network architecture of Figure 1B. (A) Example 2a: without acceleration, $\omega_3 = 0$. (B) Example 2b: with acceleration, $\omega_3 > 0$. The top diagrams show the coherences for patterns $p = 1$ (solid line) and $p = 2$ (dotted line). The bottom diagrams of A and B give the real part of the potential function (i.e., Lyapunov function in case of $\omega_3 = 0$). (C) A magnified view on the coherences in the top diagram of B, with time window from $t = 1.5\tau$ to $t = 3.0\tau$. The dynamics shows pattern-switching behavior. See the discussion in sections 4.3 and 4.4.

example 2a is no longer present. With pattern switching, there is no simultaneous coherent retrieval of the two patterns.

4.4 Nonclassical Gradient Systems. Considering examples 1 and 2, we may also compare the dynamics of the potential functions. We concentrate

this discussion on $\mathcal{R}$, the real part of $\mathcal{L}$. This allows us to compare the acceleration effect with the classical situation, where the dynamics minimizes $\mathcal{R}$. Notice that for computing $\mathcal{R}$, we use equation 2.30 with

$$(1 - V_k) \ln (1 - V_k) \to \left(1 - V_k + \varepsilon^2\right) \ln \left(1 - V_k + \varepsilon^2\right) \tag{4.5}$$

in order to avoid numerical problems with the case $V_k \simeq 1$. The parameter $0 < \varepsilon^2 \ll 1$ was introduced in section 4.1.

4.4.1 Examples 1a and 2a. Without acceleration, $\omega_3 = 0$, the potential function $\mathcal{L}$ is a real-valued Lyapunov function, $\mathcal{L} = \mathcal{R}$. The corresponding dynamics of $\mathcal{R}$ is displayed with the bottom diagrams of Figures 2A and 3A. We find that $\mathcal{R}$ is minimized as expected. Although the model describes a phase dynamics, the behavior corresponds to the classical picture of the CGH model.

4.4.2 Examples 1b and 2b. With acceleration, $\omega_3 > 0$, the potential function has a nonvanishing imaginary part, $\mathcal{L} = \mathcal{R} + i\,\Omega$ (see the discussion in section 2.2). The resulting dynamics of $\mathcal{R}$ has been displayed with the bottom diagrams of Figures 2B and 3B. In contrast to the cases of vanishing acceleration, the real part of $\mathcal{L}$ is not decreasing for all time. Instead, it reaches a minimal level, and then its values oscillate around this level. With example 2b, this frequency is related to the switching between the patterns (see Figure 3B).

This confirms that a deviation from the classical picture of a monotonically decreasing Lyapunov function may have functional relevance. With examples 1b and 2b, it is related to segmentation of overlapping patterns.

4.5 Decomposing the Networks into Patches. In the next section, examples 1b and 2b are discussed based on the concept of patches (Burwick, 2005). In this section, we review this concept and introduce the quantities needed for the discussion.

The networks of this section consist of three patches, referred to as 10, 11, and 01. These notations correspond to the binary values $\alpha = \xi_k^1 \xi_k^2$. Thus, patch $\alpha = 10$ collects the units k with $\xi_k^1 = 1$, $\xi_k^2 = 0$, and so forth. The nonoverlapping part of pattern $p = 1$ is patch $\alpha = 10$, the nonoverlapping part of pattern $p = 2$ is patch $\alpha = 01$, and the overlap is given by patch $\alpha = 11$. Patch $\alpha = 00$ (the "background") is empty since the patterns cover the complete network.

The patch coherences $\widetilde{C}_\alpha$ and phases $\widetilde{\Psi}_\alpha$ are defined in analogy to the pattern coherences of equation 3.5. This leads to

$$\widetilde{Z}_{10} = \widetilde{C}_{10} \exp(i\,\widetilde{\Psi}_{10}) = \frac{1}{\widetilde{N}_{10}} \sum_{k \in K_{10}} x_k, \tag{4.6}$$

$$\widetilde{Z}_{11} = \widetilde{C}_{11} \exp(i\widetilde{\Psi}_{11}) = \frac{1}{\widetilde{N}_{11}} \sum_{k \in K_{11}} x_k, \tag{4.7}$$

$$\widetilde{Z}_{01} = \widetilde{C}_{01} \exp(i\widetilde{\Psi}_{01}) = \frac{1}{\widetilde{N}_{01}} \sum_{k \in K_{01}} x_k, \tag{4.8}$$

where $x_k = V_k \exp(i\theta_k)$, K_α is the set of units k that constitute patch α, and $\widetilde{N}_\alpha$ is the number of units that belong to this patch. These expressions are related to the pattern coherences and phases via

$$Z_1 = C_1 \exp(i\Psi_1) = \rho_1^{10} \widetilde{Z}_{10} + \rho_1^{11} \widetilde{Z}_{11}, \tag{4.9a}$$

$$Z_2 = C_2 \exp(i\Psi_2) = \rho_2^{11} \widetilde{Z}_{11} + \rho_2^{01} \widetilde{Z}_{01}, \tag{4.9b}$$

where $\rho_1^{10} = \widetilde{N}_{10}/N_1$, and so forth. The analogous definitions may be given for the patch activities R_{10}, R_{11}, and R_{01}.

4.6 Dynamics in Terms of Patches. In this section, we use the concept of patches to arrive at some qualitative understanding of the dynamics that is observed with examples 1b and 2b. For these examples, we assumed equations 3.14 and 3.15 for every k. This was done to restrict the origin of desynchronization to the acceleration terms in order to get a better view on their effect. Units may then have different ω_k only due to different accelerations (see equation 2.33). In section 3.3, we argued that global coherence of a pattern may not be stable if the pattern is overlapping with other patterns. The argument, however, does not exclude the coherence of units that participate in the same set of patterns, that is, units that belong to the same patch (see section 4.5). Indeed, in the following, we observe that the dynamics of examples 1b and 2b generate coherent patches. The dynamics of the complete patterns may then be understood with respect to phase differences between the patches. The observed decoherences of the patterns are related to mutual decoherence of coherent patches.

4.6.1 Example 1b. As illustrated in Figure 2B, the dominating pattern $p = 1$ becomes coherent with small oscillations close to $C_1 \simeq 1$, while pattern $p = 2$ is alternating between states of coherence and decoherence. Despite the observed decoherence, it turns out that all three patches of the network reach coherent states with $\widetilde{C}_\alpha \simeq 1$ (see Figure 4). This implies that the observed decoherences in Figure 2B result from phase differences between the patch phases $\widetilde{\Psi}_\alpha$. This is indeed the case (see Figure 5).

Figure 5B illustrates that the two patches $\alpha = 10$ and $\alpha = 11$ of the dominating pattern $p = 1$ are close to synchronization. A careful inspection of the figure allows understanding of the origin of the small oscillation at $C_1 \simeq 1$. The acceleration terms in equations 3.9b and 3.10, together with equations

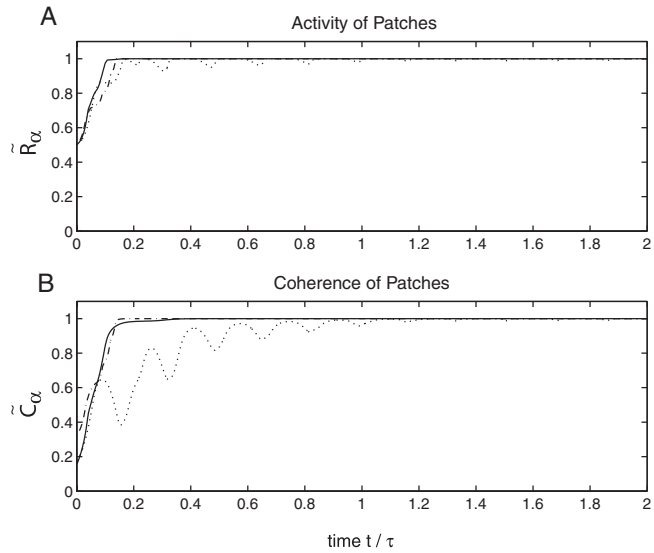

Figure 4: Example 1b. (A) Activities $\widetilde{R}_\alpha$ and (B) coherences $\widetilde{C}_\alpha$ of the patches, with $\alpha = 10$ (solid line), $\alpha = 01$ (dotted line), and $\alpha = 11$ (dash-dotted line). See the discussion in section 4.6.

3.12 and 3.6, are key to understanding this oscillation. Every patch is part of a set of patterns and receives acceleration from each of these patterns that is proportional to the pattern coherence (the C_p in equation 3.6) and increases with closer phase differences (the $\cos\left(\Psi_p - \theta_k\right)$ in equation 3.6). Thus, consider first a situation where the patch phases $\widetilde{\Psi}_{10}$, $\widetilde{\Psi}_{11}$ are close to each other, that is, close to the pattern phase Ψ_1, while patch phase $\widetilde{\Psi}_{01}$ is close to $\Psi_1 + \pi$. This implies minimal coherence C_2. In consequence, the overlap will receive only minimal acceleration from pattern $p = 2$. Thus, both patches will be closest in phase velocity, and this makes maximal coherence C_1 possible. This explains why maximal coherence C_1 and minimal coherence C_2 are correlated in Figure 5A. Correspondingly, we find maximal synchronization in Figure 5B and mininal synchronization in Figure 5C.

Next, we study how such a state of maximal C_1 and minimal C_2 evolves. With our example, we use $\omega_1 = \omega_2 = 0$. Therefore, the phase velocities are generated only by acceleration and synchronization. Since patches $\alpha = 10$, 11 are driven by pattern $p = 1$ with maximal C_1, while patch $\alpha = 01$ is driven only by the minimal C_2, the former two patches will have higher phase velocity than the latter patch. In consequence, the assumed situation, where $\widetilde{\Psi}_{01}$ is close to $\Psi_1 + \pi$ changes, and the phase of the overlap $\alpha = 11$ will approach the phase of the nonoverlapping part of pattern $p = 2$, patch $\alpha = 01$, and therefore C_2 increases. In consequence, the overlap is more strongly

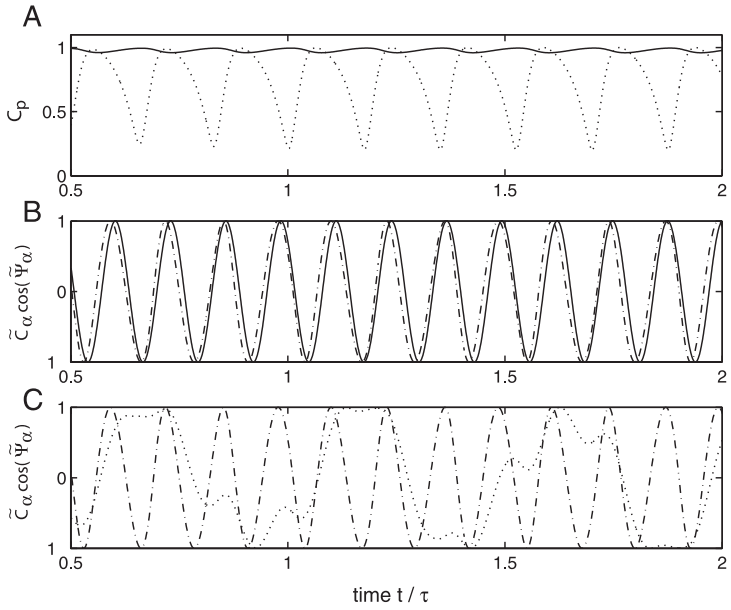

Figure 5: Example 1b. (A) The pattern coherences C_p, with time window from $t = 0.5\tau$ to $t = 2\tau$, for $p = 1$ (solid line) and $p = 2$ (dotted line). The dominating pattern $p = 1$ gets coherent. (B) Patch coherences and phases in terms of $\widetilde{C}_\alpha \cos(\widetilde{\Psi}_\alpha)$. The solid line gives $\alpha = 10$ and the dash-dotted line $\alpha = 11$. (C) The corresponding diagram for $\alpha = 11$ and $\alpha = 01$, where the latter is given by the dotted line. See the discussion in section 4.6.

accelerated by pattern $p = 2$, an acceleration that the nonoverlapping part of pattern $p = 1$ does not feel. Therefore, the phase velocities of the patches $\alpha = 10$ and $\alpha = 11$ become different, and the coherence C_1 decreases. Of course, the described effect becomes maximal when the phase of the overlap $\widetilde{\Psi}_{11}$ reaches the phase $\widetilde{\Psi}_{01}$ (i.e., when C_2 is maximal). This explains why maximal C_2 is also correlated with minimal C_1. Again, these considerations are confirmed by the correlation of minimal synchronization in Figure 5B and maximal synchronization in Figure 5C.

4.6.2 Example 2b. With example 2b, the patches become active and coherent as with example 1b (see Figure 6). Therefore, the observed decoherences of example 2b, in particular the phenomenon of pattern switching, should be traced back to differences between phases of the patches (see Figure 7). We describe example 2b with reference to the foregoing discussion of example 1b.

Consider again the situation where pattern $p = 1$ is coherent and the phases of patches $\alpha = 10, 11$ move toward the phase of patch $\alpha = 01$ with

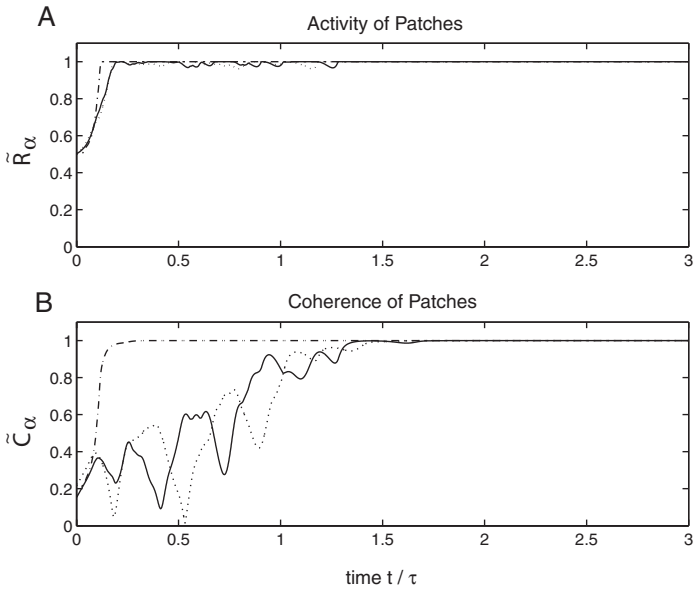

Figure 6: Example 2b. (A) Activities $\widetilde{R}_\alpha$ and (B) coherences $\widetilde{C}_\alpha$ of the patches, with $\alpha = 10$ (solid line), $\alpha = 01$ (dotted line), and $\alpha = 11$ (dash-dotted line). See the discussion in section 4.6.

higher velocity. We have already described that this leads to patch $\alpha = 11$'s receiving stronger acceleration than patch $\alpha = 10$. This may be interpreted as patch $\alpha = 01$ pulling phase $\widetilde{\Psi}_{11}$ away from phase $\widetilde{\Psi}_{10}$ and toward phase $\widetilde{\Psi}_{01}$. In case this effect is strong enough, one may even expect that the pulling leads to a transfer of synchronization away from the synchronization of $\widetilde{\Psi}_{10}$ and $\widetilde{\Psi}_{11}$ toward synchronization $\widetilde{\Psi}_{11}$ and $\widetilde{\Psi}_{01}$. This is not the case with example 1b, where pattern $p = 1$ dominated pattern $p = 2$. However, this is exactly what happens with example 2b, where neither pattern dominates the other (see Figure 7). The same mechanism that caused the dynamics of example 1b, resulting only in a weak oscillation of C_1 close to $C_1 \simeq 1$, and the alternation of C_2 between coherence and decoherence, leads to the transfer of synchronization that was denoted as pattern switching.

In summary, the breakdown of the coherent retrieval of one of the patterns and the switch to coherently retrieving the other pattern may be understood as an effect of sufficiently strong pulling away the overlap phase $\widetilde{\Psi}_{11}$ from the nonoverlapping phase $\widetilde{\Psi}_{10}$ ($\widetilde{\Psi}_{01}$) of pattern $p = 1$ ($p = 2$), due to its approaching $\widetilde{\Psi}_{01}$ ($\widetilde{\Psi}_{10}$) so that synchronization of the overlap is transferred to the other patch phase $\widetilde{\Psi}_{01}$ ($\widetilde{\Psi}_{10}$), making the other pattern $p = 2$ ($p = 1$) coherent. In consequence, phases $\widetilde{\Psi}_{11}$ and $\widetilde{\Psi}_{01}$ ($\widetilde{\Psi}_{10}$) begin to move

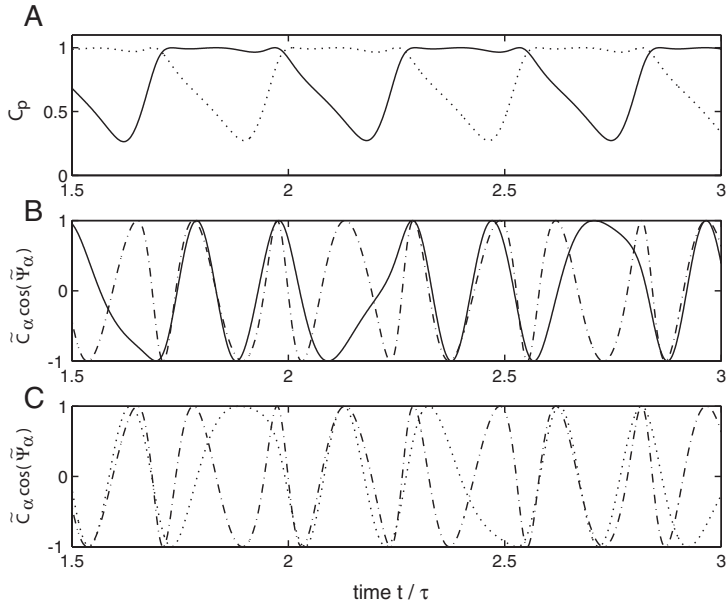

Figure 7: Example 2b. (A) Pattern coherences C_p, with a time window from $t = 1.5\tau$ to $t = 3\tau$, with $p = 1$ (solid line) and $p = 2$ (dotted). The diagram illustrates the pattern switching. (B) Patch coherences and phases in terms of $\widetilde{C}_\alpha \cos(\widetilde{\Psi}_\alpha)$. The solid line gives $\alpha = 10$ and the dash-dotted line $\alpha = 11$. (C) The corresponding diagram for $\alpha = 11$ and $\alpha = 01$, where the latter is given by the dotted line. See the discussion in section 4.6.

faster than $\widetilde{\Psi}_{10}$ ($\widetilde{\Psi}_{01}$), and the situation repeats, now transferring the synchronization from pattern $p = 2$ ($p = 1$) back to $p = 1$ ($p = 2$), and so on.

4.7 Examples 3: Shear and Oscillation Spindles. So far, we have used vanishing shear, $\omega_{2,k} = \omega_2 = 0$. Now we want to study the effect of nonvanishing shear. We demonstrate that nonvanishing shear generates an oscillation spindle structure. Consider a network with $N = 30$, and $P = 2$ patterns of $N_1 = N_2 = 20$ nonvanishing units. We assume that the patterns overlap at 10 units (see Figure 1C). The parameter choices are $a/N = 2$, $\sigma = a/2$, $\omega_3 = \pi\sigma/\tau$, and $\omega_{1,k} = \omega_1 = 0$ for every k. Example 3a is defined by $\omega_2 = 0$ and example 3b by $\omega_2 = \omega_3$.

4.7.1 Example 3a. First, we consider the case of vanishing shear terms, $\omega_2 = 0$ in equation 2.33. The example shows pattern switching as for the case of example 2b (see Figure 8). Figures 8C and 8D display the characteristic alternating coherent retrieval of the patterns and the oscillating real part of

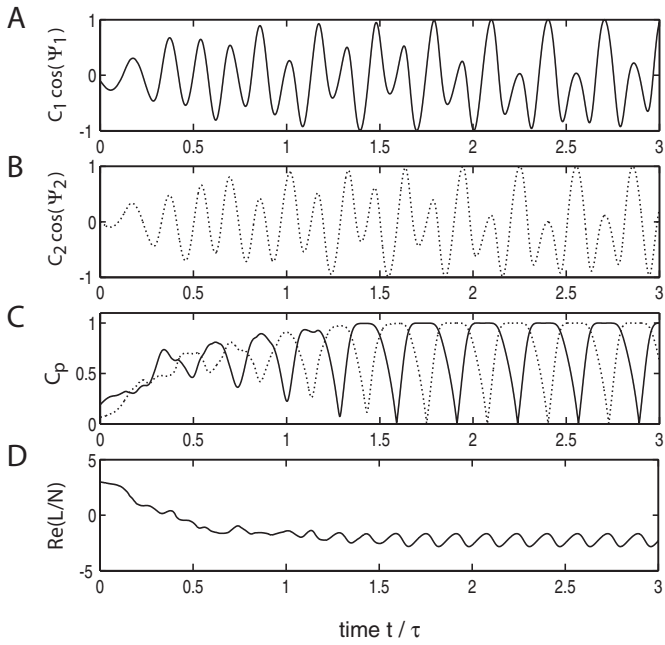

Figure 8: Example 3a, with vanishing shear terms, $\omega_2 = 0$. (A) Coherence and phase of pattern $p = 1$ in terms of $C_1 \cos(\Psi_1)$ and (B) pattern $p = 2$ in terms of $C_2 \cos(\Psi_2)$. (C) Coherences C_1 (solid line) and C_2 (dotted line). (D) The corresponding values of $\mathcal{R}/N$, the real part of the potential function $\mathcal{L}/N$. See the discussion in section 4.7.

$\mathcal{L}$ that we oberserved and discussed in sections 4.3 and 4.4. Here, we are interested in the oscillation behavior of the single patterns $p = 1$ and $p = 2$, as displayed in the Figures 8A and 8B, respectively.

4.7.2 Example 3b. Second, we allow for nonvanishing shear terms and choose $\omega_2 = \omega_3$. The resulting dynamics is displayed in Figure 9. The occurrence of pattern switching (see Figures 9C and 9D), is rather unchanged, as are the frequency of pattern switching and oscillations of $\mathcal{L}$. However, as is to be expected due to larger ω_2, the patterns now oscillate with higher frequency. The combination of low-frequency pattern switching and high-frequency oscillation of the single oscillators results in oscillation spindles of the single patterns (see Figures 9A and 9B). This is the main statement that we want to make with this example: the frequency of pattern switching and oscillations of the single oscillators may be different, allowing for oscillation spindles. Notice that we also included the dynamics of the real part of $\mathcal{L}$ in Figure 9D to illustrate that its oscillation

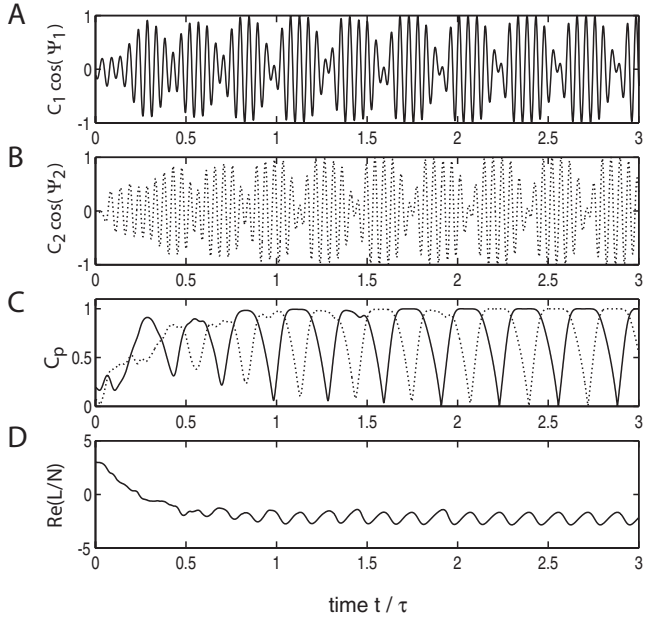

Figure 9: Example 3b, where the shear terms are nonvanishing, $\omega_2 = \omega_3$. (A) Coherence and phase of pattern $p = 1$ in terms of $C_1 \cos(\Psi_1)$ and (B) pattern $p = 2$ in terms of $C_2 \cos(\Psi_2)$. (C) Coherences C_1 (solid line) and C_2 (dotted line). (D) The corresponding values of $\mathcal{R}/N$, the real part of the potential function $\mathcal{L}/N$. See the discussion in section 4.7.

is related to pattern switching frequency and not the frequency of single oscillators.

5 Higher-Mode Couplings and Relaxation Oscillators

In this letter, we studied phase models for the oscillators with sinusoidal couplings. In section 2.4, we also gave the form of the higher-mode couplings, but these were not used for the examples in section 4. The relevance of higher-mode couplings for the model with inhibition and without acceleration has been demonstrated in Burwick (2005). In the following, we comment on their effect, with a comparison with Wilson-Cowan or relaxation oscillators.

Pattern switching phenomena have been observed with other models. For example, in Wang, Buhmann, and von der Malsburg (1990), the oscillators, have been described as extended Wilson-Cowan oscillators, and different patterns were stored according to Hebbian memory. There, pattern switching was observed as a consequence of delayed self-inhibition

in the presence of competition couplings (noise or modulation of external input were mentioned as alternative mechanisms to cause the switching). In succeeding work, networks of coupled relaxation oscillators were considered by including a second timescale (Terman & Wang, 1995; Wang & Terman, 1995, 1997). Coupled relaxation oscillators have been compared to phase oscillators with sinusoidal couplings, finding differences in synchronization time and sensitivity to the distribution of eigenfrequencies (Somers & Kopell, 1993, 1995). Further clarification of this comparison has been obtained by giving a phase model formulation for relaxation oscillator networks. These were formulated (in the limit of weak couplings) as phase oscillators with discontinuous connection functions (Izhikevich, 2000).

Given the phase formulation of the relaxation oscillator networks, it may be of interest to study whether acceleration can also be implemented with these relaxation networks. This could be achieved by using their phase model formulation and the higher-order terms that were given in section 2.4. Even the noted discontinuities may be implemented, at least in principle, by using Fourier decomposition similar to that of the Heaviside (step) function. As a future research topic, these higher-mode couplings should be considered. The studies in Somers and Kopell (1993, 1995) and Izhikevich (2000) showed that the synchronization behavior of relaxation oscillator networks is less sensitive to the distribution of eigenfrequencies in comparison with sinusoidally coupled networks. To study the corresponding feature in the presence of acceleration terms would be of particular interest, given the possible interpretation of acceleration terms as generalized eigenfrequency terms with dynamical structure (see equations 2.1, 2.3, and 2.28).

6 A Note on Fast Implementation Models

We want to add one remark on a possible modification of the system of equation 2.1. The proposed modification may be of use for faster implementation and should be of particular value when realizing more advanced applications. The u_k dynamics of the considered model comes with an inverse timescale given by

$$\begin{aligned}\frac{\Lambda(u_k)}{\tau} &= \frac{1}{\tau}\frac{1}{1-V_k}\\ &= \frac{1}{\tau}\lim_{\kappa\to\infty}\left(1+V_k+V_k^2+\ldots+V_k^\kappa\right)\\ &= \frac{1}{\tau}\lim_{\kappa\to\infty}\Lambda_\kappa(u_k). \qquad (6.1)\end{aligned}$$

The limit in equation 6.1 is finite, since $0 < V_k = g(u_k) < 1$.

The presence of Λ gives the model the simplest form in terms of complex coordinates, allowing for a straightforward derivation of the model as a

simple gradient system. In effect, it implies shorter timescales for on-states. Therefore, in the context of finite implementations, for example, with a Euler approximation, using the Λ then requires choosing short discrete time steps dt. With the regularization of section 4.1, these were restricted to $dt = \varepsilon^2\tau$. Otherwise, overshooting occurs in the case of on-states.

Consider the following modification. For the sake of fast implementation, it may be advisable to choose Λ_κ with some finite κ instead of Λ. Using some Λ_κ instead of Λ will then destroy the simple form of the system in terms of complex coordinates, but may allow fast implementations while keeping the essential features studied in this letter. Another possibility would be using fourth-order Runge-Kutta methods. A detailed treatment of these proposals is beyond the scope of this letter. For the purpose of this letter, the described discretization allowed sufficiently fast implementations.

7 Summary

In this letter, we studied an oscillatory network model that is a generalization of the classical Cohen-Grossberg-Hopfield (CGH) system. In Burwick (2005), this generalization was obtained by going from the real-valued to a complex-valued system. Here, we took another step in this direction by allowing for complex-valued couplings. It turns out that the resulting couplings have a natural interpretation in terms of neural properties. They describe an increase in phase velocity of the neural units for stronger and more coherent input. This mechanism is what we refer to as acceleration. It embeds into the framework of temporal coding a core feature that is usually identified with rate coding. Nevertheless, the new perspectives opened by including acceleration may come as a surprise.

We concentrated our discussion on the superposition problem of overlapping patterns. The combined effect of synchronization and acceleration was demonstrated for two overlapping patterns with a number of examples. The patterns were stored according to Hebbian memory, implying the usual excitatory couplings. With the parameters of our examples, the patterns became active, and without including acceleration, global coherence was reached. The patterns therefore could not be segmented with respect to coherence. The situation changed in the presence of acceleration. We distinguished two cases. First, one of the patterns dominates the other. In that case, the dominating pattern becomes coherent, thereby segmenting itself from the nonoverlapping part of the other pattern that shows a different phase dynamics. Second, neither of the patterns dominated the other. In that case, temporal segmenting of the two patterns occurred, realized with an alternating chain of transient coherent retrievals of the individual patterns. We referred to this behavior as pattern switching. For both cases, the superposition problem was solved with respect to the pattern coherences.

The approach of this letter does not need inhibitory couplings. In particular, we did not include the mutually inhibitory couplings that constitute

competition. In Burwick (2005), four mechanisms were described that may solve the superposition problem for overlapping patterns. Each of these mechanisms was built on competition. Therefore, it may come as a surprise that acceleration allows segmenting overlapping patterns without explicitly including such competition. Notice also that the resulting segmentation is self-organized in the sense that it arises not in response to some external input but as a consequence of the internal (excitatory) couplings.

We mentioned that even the simplest oscillatory systems have a natural formulation in terms of complex-valued gradient systems, where complex value refers not only to taking derivatives with respect to complex coordinates but also to taking derivatives of a potential function $\mathcal{L}$ with imaginary part. This gradient structure is reminiscent of formulating the CGH model as a real-valued gradient system that minimizes some real-valued Lyapunov function. However, there is an essential difference: nontrivial phase dynamics follows from the imaginary part of $\mathcal{L}$, and these are not compatible with the classical picture of minimizing the Lyapunov function (at least not with respect to the phase dynamics). There is a choice to be made between a classical system that minimizes the Lyapunov function or a system that has a more standard oscillatory character. Several earlier approaches toward oscillatory neural networks based on phase models used the more classical approach. Consider the treatment in Hoppensteadt and Izhikevich (1997, section 10.4), where the different eigenfrequencies and shear terms were eliminated from the general Ginzburg-Landau phase model in order to arrive at a gradient system that is complex valued (i.e., the derivative is with respect to complex coordinates) but with a real-valued potential function (Lyapunov function).

Here, we took another approach by keeping the network dynamics truly oscillatory. From a formal point of view, this amounted to allowing for imaginary parts of the potential function $\mathcal{L}$. In fact, we allowed for the imaginary parts not only with regard to the oscillator models, allowing for different eigenfrequencies and shear, but also with regard to the couplings, leading to acceleration. This implied the favorable features that were mentioned before. In consequence, the real part of $\mathcal{L}$ (corresponding to the Lyapunov function in the classical and earlier phase model approaches) decreased toward a minimal level, where the decrease is mainly caused by the amplitude dynamics. Then, however, it oscillates around this level. For example, in case of pattern switching, we found that the frequency of the pattern switching is related to the frequency of this oscillating $\mathcal{L}$.

Future work should study network behavior with more than two overlapping patterns, as well as the presence of asymmetric connections. The latter aspect may contribute to understanding how the mechanism that was presented in this letter may become part of some hierarchical architecture. Also, a stability analysis may provide a more systematic approach to the phenomena described here. We also mentioned that the applicability of the acceleration mechanism to relaxation oscillator networks should

be considered. Moreover, it would be of interest to study the combined effect of synchronization and acceleration in the context of more advanced applications. This could be based on the fast implementation models described in section 6.

In this letter, a new mechanism is introduced and described in the context of simple examples. More challenging real-world applications will be the decisive test for the relevance of the described mechanism. Discussing these is beyond the scope of this letter. Notice, however, that synchronization and acceleration may be part of the solution, but the combination with additional mechanisms may be needed. Advanced applications may also involve additional problems like occlusion and may require implementations with hierarchical architecture. In analogy to brain dynamics, it could be argued that a combination of temporal and labeled line coding might be necessary (see Singer, 2003, for an introduction to the differences of these coding schemes). Although we argued that inhibition is not necessary to generate desynchronization, competition based on inhibition may play an essential role for the complete picture of the hierarchical architecture, supporting also the matching of bottom-up and top-down information flow. Given the biological example, only such a complete machinery should be expected to realize advanced and truly intelligent applications. In that respect, we consider the specification of the hierarchical processing to be of utmost importance.

Finally, let us add some remarks on biological analogies. It is quite common to identify temporal coding with synchronization. Here, we argued that it may be worth while also considering the combination of synchronization with acceleration. The examples of this letter demonstrated that the impact of including acceleration may be fundamental. We emphasized that the observed dynamics corresponds to oscillation spindles, a fundamental dynamics that is observed also in brain dynamics. Moreover, the pattern switching that we described seems to be analogous to Gestalt switching in response to perceptual rivalries.

Nevertheless, whether the combination of synchronization and acceleration will be of relevance for understanding information processing in the brain is an open question. Its answer will depend on whether a meaningful interpretation of the network properties may be found. It should be obvious that any interpretation in terms of biological quantities needs to be accompanied with some averaging, temporal or spatial, or both. Consider, for example, an interpretation of network units as time-averaged groups of biological neurons that oscillate with frequencies in the gamma range. Despite the participation in the high-frequency wave, the underlying biological neurons may not be at saturation with respect to their individual firing rates. Thus, there may indeed be some mechanism where stronger and more coherent input implies a higher frequency of the collective firing, even without leaving the gamma range. Such a property would correspond to acceleration.

Acknowledgments

It is a pleasure to thank Christoph von der Malsburg for valuable discussions.

References

Burwick, T. (2005). *Pattern recognition based on coherent activity in complex-valued neural networks*. Manuscript submitted for publication.

Burwick, T. (2006). Oscillatory networks: Pattern recognition without a superposition problem. *Neural Computation, 18*, 356–380.

Cohen, M. A., & Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transaction on Systems, Man, and Cybernetics, SMC-13*, 815–826.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamics predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience, 2*, 704–716.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences, 5*, 16–25.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Reading, MA: Addison-Wesley.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceeding of the National Academy of Sciences of the U.S.A., 81*, 3088–3092.

Hoppensteadt, F. C., & Izhikevich, E. M. (1997). *Weakly connected neural networks*. New York: Springer-Verlag.

Izhikevich, E. M. (2000). Phase equations for relaxation oscillators. *SIAM J. Appl. Math., 60*, 1789–1805.

Kuramoto, Y. (1984). *Chemical oscillations, waves, and turbulence*. New York: Springer-Verlag.

Shareef, N., Wang, D., & Yagel, R. (1999). Segmentation of medical images using Legion. *IEEE Transactions on Medical Imaging, 18*, 74–91.

Singer, W. (2003). Synchronization, binding and expectancy. In M. Arbib (Ed.), *Brain theory and neural networks* (2nd ed., pp. 1136–1143). Cambridge, MA: MIT Press.

Somers, D., & Kopell, N. (1993). Rapid synchronization through fast threshold modulation. *Biological Cybernetics, 68*, 393–407.

Somers, D., & Kopell, N. (1995). Waves and synchrony in networks of oscillators of relaxation and non-relaxation type. *Physica D, 88*, 1–14.

Terman, D., & Wang, D. (1995). Global competition and local cooperation in a network of neural oscillators. *Physica D, 81*, 148–176.

von der Malsburg, C. (1981). *The correlation theory of brain function* (Internal Rep. 81-2). Göttingen: Max-Planck Institute for Biophysical Chemistry.

von der Malsburg, C. (1986). Am I thinking assemblies? In G. Palm & A. Aertsen (Eds.), *Brain theory: Proceedings of First Trieste Meeting on Brain Theory, October 1–4, 1984* (pp. 161–176). New York: Springer-Verlag.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron, 24*, 95–104.

Wang, D., Buhmann, J., & von der Malsburg, C. (1990). Pattern segmentation in associative memory. *Neural Computation, 2*, 94–106.

Wang, D., Freeman, W., Kozma, R., Lozowski, A., & Minai, A. (Eds.). (2004). *IEEE Transactions on Neural Networks, 15*(5). [Special Issue].

Wang, D., & Terman, D. (1995). Locally excitatory globally inhibitory oscillator networks. *IEEE Transaction on Neural Networks, 6*, 283–286.

Wang, D., & Terman, D. (1997). Image segmentation based on oscillatory correlation. *Neural Computation, 9*, 805–836.

Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron, 24*, 11–17.

Received April 3, 2006; accepted August 14, 2006.