Cross-Talk Induces Bifurcations in Nonlinear Models of Synaptic Plasticity

Terry Elliott
te@ecs.soton.ac.uk
Department of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.

Linear models of synaptic plasticity provide a useful starting-point for examining the dynamics of neuronal development and learning, but their inherent problems are well known. Models of synaptic plasticity that embrace the demands of biological realism are therefore typically nonlinear. Viewed from a more abstract perspective, nonlinear models of synaptic plasticity are a subset of nonlinear dynamical systems. As such, they may therefore exhibit bifurcations under the variation of control parameters, including noise and errors in synaptic updates. One source of noise or error is the cross-talk that occurs during otherwise Hebbian plasticity. Under cross-talk, stimulation of a set of synapses can induce or modify plasticity in adjacent, unstimulated synapses. Here, we analyze two nonlinear models of developmental synaptic plasticity and a model of independent component analysis in the presence of a simple model of cross-talk. We show that cross-talk does indeed induce bifurcations in these models, entirely destroying their ability to acquire either developmentally or learning-related patterns of fixed points. Importantly, the critical level of cross-talk required to induce bifurcations in these models is very sensitive to the statistics of the afferents' activities and the number of afferents synapsing on a postsynaptic cell. In particular, the critical level can be made arbitrarily small. Because bifurcations are inevitable in nonlinear models, our results likely apply to many nonlinear models of synaptic plasticity, although the precise details vary by model. Hence, many nonlinear models of synaptic plasticity are potentially fatally compromised by the toxic influence of cross-talk and other sources of noise and errors more generally. We conclude by arguing that biologically realistic models of synaptic plasticity must be robust against noise-induced bifurcations and that biological systems may have evolved strategies to circumvent their possible dangers.

1 Introduction

Synaptic plasticity is believed to underlie development, learning and memory in the nervous system (Bliss & Collingridge, 1993; Katz & Shatz, 2012).
Long-term potentiation (LTP; Bliss & Lømo, 1973) and long-term depression (LTD; Lynch, Dunwiddie, & Gribkoff, 1977) are currently the favored models for enduring changes in synaptic efficacy (Malenka & Bear, 2004). The Hebbian, pre- and postsynaptic activity requirements (Hebb, 1949) of LTP/D are typically taken to imply that LTP/D are exactly specific to the stimulated sets of synapses, but increasing evidence points to some degree of cross-talk between synapses in LTP/D. Cross-talk occurs when LTP/D in a stimulated set of synapses spills over into neighboring, unstimulated sets of synapses. It may arise during both early-phase LTP/D (e-LTP and e-LTD), either presynaptically (Bonhoeffer, Staiger, & Aertsen, 1989; Kossel, Bonhoeffer, & Bolz, 1990; Schuman & Madison, 1994) or postsynaptically (Engert & Bonhoeffer, 1997; Harvey & Svoboda, 2007), and late-phase LTP/D (l-LTP and l-LTD; Frey & Morris, 1997, 1998; Sajikumar & Frey, 2004). In e-LTP/D, the rapid diffusion of intracellular messengers such as calcium and Ras activity (Harvey & Svoboda, 2007; Harvey, Yasuda, Zhong, & Svoboda, 2008) or extracellular messengers such as neurotrophic (Lessmann, Gottmann, & Heumann, 1994; Korte et al., 1995; Levine, Dreyfus, Black, & Plummer, 1995) and gaseous factors (O’Dell, Hawkins, Kandel, & Arancio, 1991; Zhuo, Small, Kandel, & Hawkins, 1995), or the translocation of plasticity-related enzymes such as CaMKII from dendrites to spines (Rose, Jin, & Craig, 2009) may contribute to cross-talk. On longer timescales, the induction of l-LTP/D in one set of synapses can convert e-LTP/D induced by less persistent stimulation in a different set of neighboring synapses into l-LTP/D by the latter synapses’ capture of plasticity-related proteins; further, in so-called cross-tagging, l-LTP can stabilize e-LTD and l-LTD can stabilize e-LTP (Frey & Morris, 1998; Sajikumar & Frey, 2004).

Mathematical models of synaptic plasticity must exhibit the stability of developed or learned patterns of synaptic connectivity, although the stability versus plasticity dilemma is notorious (Abraham & Robins, 2005). One way to achieve stability is to ensure that models possess a fixed point structure, with some of the fixed points being stable (see Wiskott & Sejnowski, 1998; Elliott, 2003; Burkitt, Meffin, & Grayden, 2004) Fixed point structures can arise in both linear and nonlinear models of synaptic plasticity. Linear models can possess only a single fixed point, and this is usually unstable, with synaptic strengths growing unboundedly in linear Hebbian models. In order to admit multiple stable fixed points, linear Hebbian models are typically supplemented by constraints such as postsynaptic normalization (von der Malsburg, 1973). A more natural, flexible, and powerful way to incorporate multiple stable fixed points, however, is explicitly to formulate and construct nonlinear models of synaptic plasticity. Nonlinear models can exhibit extremely rich dynamics without the limitations characteristic of linear models. More generally, nonlinear models of synaptic plasticity can solve problems that linear models even in principle cannot solve. For example, the ability to extract independent components depends critically
on the presence of strength nonlinearities, while linear models can extract only principal components (Hyvarinen et al., 2001). Finally, while linear models of synaptic plasticity are of course a very useful starting point for examining the dynamics of neuronal development and learning, it is clear that the assumption of linearity constitutes a theoretician’s highly convenient but perhaps highly unrepresentative idealization. It is important to study models that are simple enough to embody the dynamics of interest but not so simple that they cease to be representative. Biologically realistic models of synaptic plasticity predicated on the need to understand the full and rich complexity of synaptic plasticity in real neurons will inevitably be nonlinear, and typically highly so. Understanding the nature of, and the possible problems associated with, nonlinear models of synaptic plasticity is therefore of paramount theoretical and experimental importance.

Unlike linear models, nonlinear models of synaptic plasticity are a subset of nonlinear dynamical systems. As a result, they are subject to bifurcation and other types of behavior in the presence of noise, errors, and perturbations more generally (Strogatz, 1994; Ermentrout & Terman, 2010). Nonlinear models operating with perfectly synapse-specific Hebbian plasticity may exhibit the fixed point structures demanded by the developmental or learning scenarios that they seek to embody. However, cross-talk between synapses during synaptic plasticity, leading to changes in strength of unstimulated synapses, may act as a form of noise or error in synaptic updates. Cross-talk may therefore behave as a control parameter and a possible source of bifurcations in nonlinear synaptic plasticity models. Such bifurcations could lead in principle to radical changes in the fixed point structures of models, destroying the capacity of nonlinear models to acquire developmentally or learning-related states.

Can cross-talk in synaptic plasticity in fact lead to a complete breakdown in the desired developmental or learning dynamics in otherwise perfectly synapse-specific nonlinear Hebbian models? A recent numerical study using independent component analysis (ICA) as the nonlinear model of synaptic plasticity has answered this question in the affirmative (Cox & Adams, 2009). Cox and Adams find that even for rather low levels of cross-talk, the fixed point structure of ICA can be entirely destroyed. In particular, above some critical level of cross-talk, the model stops operating as an independent component analyzer and ceases to extract independent components, with learning breaking down completely. These numerically seen bifurcations are likely not specific to ICA but should be generally observable in the presence of any nonlinearity in a synaptic plasticity rule. Since we have argued that all biologically plausible and biologically relevant models of synaptic plasticity will be nonlinear, the possibility that cross-talk between synapses can have such catastrophic consequences in nonlinear synaptic plasticity models is of the utmost import. It is therefore critical to begin to explore and assess the impact of noise and errors...
in general, and cross-talk in particular, on nonlinear models of synaptic plasticity.

In this letter, we examine in considerable detail this question of the impact of cross-talk on nonlinear models of synaptic plasticity, but analytically rather than numerically. The bifurcations and chaotic behaviors that nonlinear models can display are exquisitely sensitive to the details of the models, so a general analysis of nonlinear models of synaptic plasticity in the presence of cross-talk is likely impossible. We must therefore restrict the discussion here to certain specific but representative models, or classes of model, in order to illustrate the possibilities. Since we wish to obtain good analytical insight into the nature of the bifurcations induced by the presence of cross-talk, we must study models that are simple enough to analyze mathematically but complex enough to exhibit interesting dynamics. We focus on three particular, well-motivated models—two in the sphere of neuronal development and the third in the sphere of learning theory. Considering two such different spheres of operation of synaptic plasticity allows us to gauge to some degree the likely generality of our results. In all three cases, with models operating according to very different theoretical principles, we show that the level of cross-talk required to induce bifurcations, and therefore destroy the capacity of these models to acquire developmentally or learning-related patterns of fixed points, can be very small. In two of them, the bifurcation-inducing level of cross-talk can in fact be made arbitrarily small by changing input statistics or the number of afferents synapsing on a neuron. Such extreme sensitivity to even vanishingly small levels of cross-talk is profoundly disturbing. It suggests that the presence of cross-talk, and noise and errors more generally, in synaptic updates in real, biological synapses may impose severe limits and constraints on current theoretical models of synaptic plasticity. Such limits and constraints may force a change in the way in which we think about synaptic plasticity both theoretically and experimentally, and they may change the way in which we construct biologically relevant models of synaptic plasticity.

The remainder of the letter is organized as follows. After formulating the mathematical model of cross-talk in synaptic updates following Cox and Adams (2009) in section 2, we then analyze in detail the specific models mentioned above in the presence of cross-talk. In section 3, we examine nonlinear models of developmental processes. We first consider a very simple nonlinear model, called the minimal multiplicative model (MMM; Elliott, 2003), and derive its bifurcation properties as a function of the parameter that governs the level of cross-talk. We show that in general, it exhibits two distinct bifurcations with a multistable regime between the two bifurcations. At the first bifurcation point, developmental processes become compromised by the stabilization of a fixed point related to cross-talk dynamics; at the second bifurcation point, all developmentally relevant fixed points are destroyed in saddle node bifurcations. We then examine a broader class of developmental models, derived from the rate-based limit of our model...
of spike-timing-dependent plasticity (STDP; Appleby & Elliott, 2005), but under certain approximations to make a more general analysis tractable. With these approximations, we appear to obtain pure stability changes of fixed points, but we show that these changes must in general correspond to saddle node bifurcations. Although the bifurcation processes in the STDP model differ in detail from those in the MMM, there are still two sets of bifurcations with an intermediate, multistable regime. In section 4, we turn from developmental models to a learning model. In particular, we analyze ICA with whitened sources, considering particular choices of sources for which analysis is tractable. We find that ICA can exhibit an extremely rich bifurcation structure in the presence of cross-talk, with multiple cascades of bifurcations leading to the inability of ICA to acquire any sources at all, even in the presence of arbitrary small levels of cross-talk. Finally, in section 5, we discuss the extensive analysis performed here. We compare our results to those of Cox and Adams (2009), evaluate the generality of our results, and consider the possible implications of the impact of synaptic cross-talk on nonlinear synaptic plasticity models more generally.

2 Formulation

We consider models of synaptic plasticity that possess a fixed point structure. For simplicity, we examine a set of \( n \) presynaptic neurons synapsing on a single postsynaptic neuron with synaptic strengths \( v_i, i = 1, \ldots, n \). We suppose that these synaptic strengths, in the absence of cross-talk in synaptic updates, evolve according to an equation that can be written in the form

\[
\dot{v} = \Pi(a, v) - \Delta(a, v),
\]

where \( v = \{v_1, \ldots, v_n\}^T \), with \( T \) denoting the transpose, is the vector of synaptic strengths; \( \Pi(a, v) \) represents the change in strength due to potentiation processes, its components being nonnegative; \( \Delta(a, v) \) that due to depression processes, similarly nonnegative; the vector \( a = \{a_1, \ldots, a_n\}^T \) has as its components the instantaneous firing rates, \( a_i, i = 1, \ldots, n \), of the \( n \) afferents; a dot over a quantity denotes differentiation with respect to time; and any common “learning rate” has been absorbed into a rescaling of time. The fixed points are then given by the solutions of the equation \( \Pi(a, v) = \Delta(a, v) \) or, more usually, the solutions of this equation when averaged over the statistical ensemble of afferent activity patterns.

Many Hebbian models propose only a specific form for the potentiation process \( \Pi(a, v) \). Without compensating depression, such models are of course unstable, with the synaptic strengths \( v_i \) increasing without bound. Synaptic normalization is often used to stabilize these models. Under multiplicative synaptic normalization, synaptic strengths are multiplicatively rescaled in order to conserve some defined function of \( v \). Here we do
not consider the biologically rather implausible subtractive normalization (Goodhill & Barrow, 1994; Miller & MacKay, 1994). We have shown elsewhere that the need for subtractive normalization is particular only to linear (or quasi-linear) models of Hebbian synaptic plasticity and that large classes of nonlinear Hebbian models have acceptable fixed point structures under multiplicative normalization (Elliott, 2003; see also Elliott & Shadbolt, 2002). The two most popular choices of multiplicative normalization correspond to conserving \( n \cdot v \), where \( n = \{1, \ldots, 1\}^T \) is an \( n \)-component vector all of whose components are unity, and \( \cdot \) is the dot product, or to conserving \( v \cdot v = |v|^2 \). The first corresponds to the conservation of the total synaptic strength of the postsynaptic neuron, while the second corresponds to the conservation of the overall length of the vector of synaptic strengths. Following Miller and MacKay (1994), we refer to these, respectively, as M1 and M2 normalization. Without loss of generality, we may take either \( n \cdot v = 1 \) or \(|v|^2 = 1\). Under these two forms, we then have

\[
\Delta(a, v) = \begin{cases} 
[n \cdot \Pi(a, v)] v & \text{under M1} \\
[v \cdot \Pi(a, v)] v & \text{under M2}
\end{cases}
\] (2.2)

giving explicit models for the depression process. We can then write

\[
\dot{v} = \begin{cases} 
P_n \Pi(a, v) & \text{under M1} \\
P_v \Pi(a, v) & \text{under M2}
\end{cases}
\] (2.3)

where the two matrices \( P_n \equiv I - v n^T \) and \( P_v \equiv I - v v^T \), with \( I \) the \( n \times n \) identity matrix, are projection matrices that annihilate any instantaneous growth, respectively, in the \( n \) direction and so off the hyperplane \( n \cdot v = 1 \), or in the \( v \) direction and so off the hypersphere \( v \cdot v = 1 \). In such Hebbian, normalization-based models, the form for \( \Delta(a, v) \) is determined uniquely by \( \Pi(a, v) \) and the fixed point structure is induced by the interaction between the pure Hebbian term \( \Pi(a, v) \) and either M1 or M2. In other Hebbian models, however, \( \Delta(a, v) \) is not determined from \( \Pi(a, v) \) via synaptic normalization, but its form can be postulated simultaneously with that for \( \Delta(a, v) \). For example, under classical STDP, potentiation and depression processes arise as a result of differing responses to differently ordered pairs of pre- and postsynaptic spikes (for a review, see Caporale & Dan, 2008). A model that proposes forms for both \( \Pi(a, v) \) and \( \Delta(a, v) \) need not, however, be stable because solutions of the (averaged) equation \( \Pi(a, v) = \Delta(a, v) \) may not exist or may not be stable in relevant regions of parameter- or synaptic strength-space. Nevertheless, in addition to considering examples of models employing synaptic normalization, we will also consider the rate-based limit of a model of STDP that defines both \( \Pi(a, v) \) and \( \Delta(a, v) \) without synaptic normalization but is stable in all relevant regions.
The dynamics defined in equation 2.1 represent the evolution of synaptic strengths in a manner that is completely specific to the stimulated synapses. However, experimental data indicate, as discussed above, that there can be cross-talk in otherwise Hebbian synaptic updates, so that unstimulated or only weakly simulated synapses can undergo plasticity as a result of stimulation at other synapses. These processes can occur on different spatial scales and on different temporal scales and can involve different molecular and signaling pathways. Our purpose here is not to attempt to build a detailed model of such cross-talk, but rather to understand, analytically, the possible impact of cross-talk, very broadly construed, on the fidelity of Hebbian synaptic plasticity. In particular, we are interested in the possible impact of simple, analytically tractable models of cross-talk on the fixed point structures of Hebbian models.

Consider first cross-talk on only the potentiation term, \( \Pi(a, v) \). A very simple but mathematically elegant model can be built around a linear interpolation between two limits. For complete Hebbian specificity in synaptic updates, we require the synaptic update to be just \( \Pi(a, v) \). For a complete breakdown in Hebbian specificity in updates, we can imagine the average, otherwise specific Hebbian update across all synapses, \( n \cdot \Pi(a, v)/n \), being applied equally to all synapses. Thus, if we define, similarly to Cox and Adams (2009), the cross-talk matrix

\[
E = (1 - \epsilon)I + \epsilon \frac{1}{n} nn^T, \tag{2.4}
\]

then the parameter \( \epsilon \in [0, 1] \) smoothly and linearly interpolates between complete specificity at \( \epsilon = 0 \) and a complete breakdown in specificity at \( \epsilon = 1 \). Cox and Adams (2009) refer to such a matrix as an error matrix, but we prefer the more neutral term, cross-talk matrix. Our definition of the magnitude of cross-talk \( \epsilon \) differs from theirs by a scaling factor of \( 1/n \). We choose this convention so that \( \epsilon \in [0, 1] \) always, independent of \( n \), rather than in their case, \( \epsilon \in [0, 1/n] \), which depends on \( n \). Their diagonal entry, \( Q \equiv 1 - \epsilon(n - 1) \), then becomes our diagonal entry, \( 1 - \epsilon(n - 1)/n \). The cross-talk matrix \( E \) in equation 2.4 is the one whose impact on nonlinear, otherwise Hebbian synaptic plasticity we shall explore extensively below. Notice that \( n \cdot E \Pi(a, v) = n \cdot \Pi(a, v) \), since \( n \) is a left (and also right) eigenvector of \( E \) of unit eigenvalue. Hence, for any value of \( \epsilon \), there is a conservation of the total update across all synapses. With this choice of \( E \), we then write

\[
\dot{v} = E \Pi(a, v) - \Delta(a, v), \tag{2.5}
\]

where \( \Delta(a, v) \) is either postulated separately from \( \Pi(a, v) \) or induced by synaptic normalization, so that

\[
\dot{v} = PE \Pi(a, v). \tag{2.6}
\]
where $\mathbb{P}$ is the appropriate projection matrix for either $M_1$ or $M_2$. We will consider below synaptic plasticity rules in the presence of cross-talk of the form in equation 2.6.

Now consider cross-talk on only the depression term, $\Delta(a, v)$, and suppose that we want the depression term nevertheless to represent an exact normalization process. If we “break” the exact normalization, then we cannot be certain that any possible change in the fixed point structure is induced entirely by the cross-talk rather than by merely introducing instabilities due to the compromised normalization. We would then write, for a general cross-talk matrix $\mathbb{E}$,

$$
\dot{v} = \begin{cases}
\Pi(a, v) - \frac{n \cdot \Pi(a, v)}{n \cdot \mathbb{E} v} \mathbb{E} v & \text{under } M_1 \\
\Pi(a, v) - \frac{v \cdot \Pi(a, v)}{v \cdot \mathbb{E} v} \mathbb{E} v & \text{under } M_2
\end{cases}.
$$

These forms may be derived in one of two different ways. In the first derivation, we can suppose that the depression term is proportional to the update $\mathbb{E} v$ in the presence of cross-talk and then fix the constant of proportionality to ensure that the exact $M_1$ or $M_2$ normalization is preserved. However, such a derivation requires that the normalization dynamics have access to the pattern of cross-talk, which is biologically implausible. The second derivation supposes instead that the constant of proportionality is given by that which would have been induced had the update been exactly Hebbian, and so does not require access to the pattern of cross-talk. Thus, the system tries to normalize, but cross-talk in the update process leads to a failure to normalize exactly. If the system then iteratively renormalizes, but with each renormalization still necessarily subject to cross-talk, then the resulting process defines a geometrical series of iterative synaptic updates that converges to the same result obtained in the first, simpler derivation. Hence, the normalization dynamics need not in fact have access to the pattern of cross-talk. With the particular choice of $\mathbb{E}$ defined in equation 2.4, the dynamics in equation 2.7 become

$$
\dot{v} = \begin{cases}
\Pi(a, v) - [n \cdot \Pi(a, v)] \mathbb{E} v & \text{under } M_1 \\
\Pi(a, v) - [v \cdot \Pi(a, v)] \frac{\mathbb{E} v}{1 - \varepsilon + \frac{1}{\varepsilon} (n \cdot v)^2} & \text{under } M_2
\end{cases}.
$$

Under $M_1$, because $n$ is a left eigenvector of $\mathbb{E}$ of unit eigenvalue and since $n \cdot v = 1$ by assumption, the denominator term $n \cdot \mathbb{E} v$ collapses down to unity. However, under $M_2$, because the constraint is on $v \cdot v$ and not $n \cdot v$, the denominator term $v \cdot \mathbb{E} v$ does not so collapse. Hence, under $M_1$, the cross-talk parameter $\varepsilon$ appears linearly, while under $M_2$, it appears nonlinearly. This is highly unsatisfactory, since we want $\varepsilon$ to be introduced...
into different models in exactly the same functional fashion, at least for the abstract models of cross-talk considered here. We need to be certain that any possible changes in the fixed point structure arise as a function purely of the nonlinearities in the model of synaptic plasticity, and not in part as a function of the possible nonlinearities associated with the introduction of cross-talk itself. This difficulty arises because of the difference in the dynamics between M1 and M2 and the way in which they interact with the simple model of cross-talk in equation 2.4. To throw this problem into even sharper relief, we can consider a different model of cross-talk. Suppose that when $\varepsilon = 1$, the update is not applied equally across all synapses but rather in a strength-dependent manner. Such a proposal would fit neatly the requirements for a strength-dependent, multiplicative normalization process. We could then write,

$$E^* = (1 - \varepsilon)I + \varepsilon \frac{1}{n \cdot v} v n^T,$$

(2.9)

where we have fixed the constant of proportionality on the $v n^T$ term to ensure that $n$ is always a left eigenvector of $E^*$ with unit eigenvalue, so that the total synaptic update is conserved in the presence of cross-talk. We see that $v$ is a right eigenvector of $E^*$, also with unit eigenvalue, regardless of the form of synaptic normalization. With $E^*$ replacing $E$ in equation 2.7, the depression terms therefore collapse, with $E^*$ disappearing entirely from both M1 and M2 forms. Thus, depression processes that are governed by synaptic normalization containing cross-talk can interact in an unexpected and undesirable fashion with the precise form of the cross-talk matrix. With $E^*$, the cross-talk parameter $\varepsilon$ disappears under both M1 and M2, while with $E$, it appears linearly under M1 and nonlinearly under M2. These behaviors stand in contrast to cross-talk only on the potentiation term $\Pi(a, v)$, in which the cross-talk parameter $\varepsilon$ always appears linearly, interpolating smoothly between the two regimes of exact specificity and complete breakdown in specificity in potentiating synaptic updates.

Finally, consider cross-talk on both the potentiation and the depression terms, with the depression term not governed by synaptic normalization so as to avoid the problems discussed above. Clearly, if we write

$$\dot{v} = E [\Pi(a, v) - \Delta(a, v)],$$

(2.10)

for some general but activity-independent form of the cross-talk matrix, then the fixed points in the absence of $E$ will remain fixed points in the presence of $E$. Under certain, reasonable assumptions on $E$, these fixed points will, moreover, be of identical characters and stabilities, so the cross-talk parameter $\varepsilon$ will exert no influence on the fixed point structure. (The fixed points may be additionally supplemented by null eigenvectors of $E$ when such solutions exist.) If $E$ is activity dependent and $\Pi(a, v)$ and
\( \Delta(\mathbf{a}, \mathbf{v}) \) depend on afferent activity in different ways, then a common cross-talk matrix may lead to different patterns of fixed points and characters. But for activity-independent forms, we need two separate, distinct cross-talk matrices characterizing the different forms of cross-talk that may attach to the biologically distinct and mechanistically different potentiation and depression processes.

Because of these two problems of the possible interaction of the cross-talk matrix with depression implemented via synaptic normalization and of the need in general for two distinct (activity-independent) cross-talk matrices when cross-talk is considered on both potentiation and depression terms, we will consider only cross-talk acting on the potentiation term \( \mathbf{\Pi}(\mathbf{a}, \mathbf{v}) \) as in equation 2.6, with the cross-talk matrix \( \mathbf{E} \) of the form defined in equation 2.4.

Before proceeding, we discuss all the eigenvectors and eigenvalues of the matrix \( \mathbf{E} \) in equation 2.4. It has a (unit norm) eigenvector \( \hat{\mathbf{n}} \equiv \mathbf{n}/\sqrt{n} \) with eigenvalue 1 and a further \( n - 1 \) repeated eigenvalues \( 1 - \epsilon \). We denote the degenerate (unit norm) eigenvectors corresponding to these repeated eigenvalues \( \mathbf{h}^1, \ldots, \mathbf{h}^{n-1} \). Because \( \mathbf{E} \) is symmetric, the set \( \{\hat{\mathbf{n}}, \mathbf{h}^1, \ldots, \mathbf{h}^{n-1}\} \) may be taken to form an orthonormal basis for \( \mathbb{R}^n \). We refer to the eigenvector \( \hat{\mathbf{n}} \) as the principal component (PC) of \( \mathbf{E} \) and the \( n - 1 \) degenerate eigenvectors \( \mathbf{h}^i, i = 1, \ldots, n - 1 \), as its subprincipal components (sub-PCs). Because the sub-PCs are degenerate and orthonormal, we may freely rotate them at will in the subspace orthogonal to \( \hat{\mathbf{n}} \), thereby picking any orthonormal basis for the sub-PCs of \( \mathbf{E} \) that we please. We may of course perform an eigenexpansion of \( \mathbf{E} \), writing

\[
\mathbf{E} = \hat{\mathbf{n}} \hat{\mathbf{n}}^T + (1 - \epsilon) \sum_{i=1}^{n-1} \mathbf{h}^i \mathbf{h}^i^T. \tag{2.11}
\]

Having characterized the eigenvectors of \( \mathbf{E} \), we now consider the impact of cross-talk on nonlinear models of synaptic plasticity.

3 Cross Talk in Developmental Models

Developmental models, such as models of the development of ocular dominance columns (ODCs) in the primary visual cortex, have been extensively studied (for reviews, see Swindale, 1996; van Ooyen, 2001). These models typically give rise, via activity-dependent competitive dynamics, to the emergence of segregated patterns of afferent connectivity in which all afferents’ synaptic strengths but one go to zero, leaving only a single afferent innervating the target cell. We have shown that the simplest nonlinear Hebbian model using multiplicative normalization that can segregate afferents whose activity patterns are positively correlated is the so-called minimal multiplicative model (MMM; Elliott, 2003). Because the Hebbian,
potentiation part of the MMM is nonlinear in synaptic strengths, the cross-talk matrix could potentially induce bifurcations in the model’s fixed point structure as the parameter $\varepsilon$ is varied. If such bifurcations occur, then development dynamics could be completely disrupted. We therefore now analyze the bifurcation structure of the MMM in the presence of cross-talk. We then analyze a second developmental model that does not employ synaptic normalization. This model is obtained as the rate-based limit of a stochastic model of STDP (Appleby & Elliott, 2005) and proposes specific and separate nonlinear forms for $\Pi(a, v)$ and $\Delta(a, v)$. Because the model is intrinsically stable in all regions of parameter space, no constraints on synaptic strength need be imposed, except for reflecting boundaries at zero synaptic strength (Elliott, 2008, 2010, 2011). This model therefore also possesses a fixed point structure corresponding to segregated states of afferent connectivity, but established by a dynamic regulation that is very different from that imposed by the device of multiplicative synaptic normalization.

### 3.1 The Minimal Multiplicative Model.

The MMM is defined by the potentiation term $\Pi(a, v)$ with components

$$\Pi_i = (a \cdot v) a_i v_i,$$

with $\Delta(a, v)$ obtained from $\Pi(a, v)$ by M1 normalization. The postsynaptic activity has been written in the standard, linear form $a \cdot v$, but the usual presynaptic term in the standard, linear Hebbian rule has been changed from $a_i$ to $a_i v_i$; this change can be interpreted anatomically (Elliott 2003). Averaging over the ensemble of afferent activity patterns, with $\langle aa^T \rangle = C$ being the correlation matrix with components $C_{ij} = \langle a_i a_j \rangle$, the angle brackets denoting the ensemble average, we obtain with the usual, standard assumptions regarding separation of timescales,

$$\dot{v}_i = (Cv)_i v_i - (v^T C v) v_i,$$

where angle brackets around the $v_i$ have been dropped for notational ease. Defining the $n$-dimensional set of vectors $\{e^1, \ldots, e^n\}$ to form a Cartesian basis, so that $(e^i)_j = \delta^i_j$, the Kronecker delta symbol, then under reasonable assumptions on $C$ in a slow, developmental context, equation 3.2 has $n$ stable, segregated fixed points $v = e^i$, $i = 1, \ldots, n$, and an unstable, unsegregated fixed point given by $nv = n$ (Elliott, 2003).\(^1\)

\(^1\)There are, of course, for $n > 2$ many other unstable fixed points, such as $(e^i + e^j)/2$ for $i \neq j$, but we are interested here principally in the segregated fixed points and the fully unsegregated fixed point $n/n$.  

---

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/NECO_a_00224 by guest on 16 May 2021
In the presence of the cross-talk matrix in equation 2.4 on the potentiating term $\Pi(a, v)$, equation 3.2 becomes

$$
\dot{v}_i = (1 - \varepsilon)(Cv)_i v_i + \frac{1}{n} \varepsilon(v^T Cv) - (v^T Cv)v_i,
$$

(3.3)

with no associated change induced in the depression term because $\sum_i a_i v_i \equiv a \cdot v$. To determine the impact of $\varepsilon \neq 0$ on the $\varepsilon = 0$ fixed point structure, we consider a form of $C$ corresponding to unbiased first- and second-order statistics in which all afferents are treated symmetrically. Such an assumption is standard in developmental models and characterizes normal rather than abnormal developmental dynamics. For example, the inputs from both eyes, via the lateral geniculate nucleus, to primary visual cortical cells are typically assumed to have equal means and variances, reflecting spatial and temporal homogeneity and isotropy during the development of ODCs. Similar assumptions would be made regarding motor neurons’ inputs to single muscle fibers during the developmental elimination of polyneuronal innervation at the vertebrate neuromuscular junction. Privileging any single afferent by giving it a higher mean activity or variance than other afferents would bias the competitive dynamics in favor of that afferent, so that it would inevitably win control of the postsynaptic cell, as in, for example, monocular deprivation paradigms. For more than two afferents, we also assume a common pairwise correlation coefficient between the activities of afferents. Such an assumption is not mandatory, but it simplifies our analysis. However, provided that presynaptic inputs to a postsynaptic cell have receptive fields that differ only slightly in their sampling of the relevant input space, the assumption of common pairwise correlation coefficients, while an idealization, is not unrealistic. Moreover, although the precise details of the bifurcation dynamics that we examine below will depend on the form of $C$, bifurcations will occur for any choice of $C$. Picking a form of $C$ corresponding to unbiased first and second order therefore allows us to demonstrate these bifurcations cleanly, without unnecessary complication.

Let the common mean activity of all afferents be $\mu$, the common variance in activity of all afferents be $\sigma^2$, and the common pairwise correlation coefficient between any pair of afferents’ activities be $\rho$. Then the correlation matrix $C$ has the form

$$
C = (\mu^2 + \rho \sigma^2) n n^T + (1 - \rho)\sigma^2 I,
$$

(3.4)

and the positivity of its eigenvalues implies that

$$
0 \leq \kappa \equiv \frac{(1 - \rho)\sigma^2}{\mu^2 + \sigma^2} < \frac{n}{n - 1},
$$

(3.5)
which defines the parameter $\kappa$. With this form of $C$, equation 3.3 becomes

\[
\dot{v}_i = (1 - \rho)\sigma^2 \left\{ (v_i - |v|^2)v_i - \varepsilon \left[ (\kappa^{-1} - 1) \left( v_i - \frac{1}{n} \right) + \left( v_i^2 - \frac{1}{n} |v|^2 \right) \right] \right\}.
\] (3.6)

We see that the unsegregated solution $nv = n$ remains with $\varepsilon \neq 0$, but the $\varepsilon = 0$, exactly segregated solutions $v = e^i$, $i = 1, \ldots, n$, are modified by $\varepsilon \neq 0$. We focus without loss of generality on the $v = e^1$ solution, and by the assumed symmetry between afferents and the form of the cross-talk matrix, we write the $\varepsilon \neq 0$ solution as $v = (1 - nw)e^1 + w n$ for some scalar $w$, to be determined. We then obtain

\[
\dot{w} = (1 - \rho)\sigma^2 (1 - nw) \left\{ n(n - 1)w^2 - [n + \varepsilon(n - 2)]w + \varepsilon \kappa^{-1} \right\}. \tag{3.7}
\]

The $w = 1/n$ solution corresponds to the $n v = n$, unsegregated fixed point. Its stability changes at the critical value

\[
\varepsilon^{(1)}_c = \frac{\kappa}{n - (n - 2)\kappa}, \tag{3.8}
\]

with $0 \leq \varepsilon^{(1)}_c < 1$ from equation 3.5, and is unstable for $\varepsilon < \varepsilon^{(1)}_c$ and stable for $\varepsilon > \varepsilon^{(1)}_c$. Thus, whereas with $\varepsilon = 0$ the unsegregated solution is always unstable, above the critical level of cross-talk $\varepsilon^{(1)}_c$, the unsegregated solution becomes stable. The unsegregated state $n v = n$ therefore undergoes a transcritical bifurcation at $\varepsilon = \varepsilon^{(1)}_c$. However, because the change in stability is governed by the quadratic factor on the right-hand side of equation 3.7, and not by a linear factor, this transcritical bifurcation is not associated with a change in stability of another fixed point except when $n = 2$, as we shall see below. For $n > 2$, this transcritical bifurcation occurs when a set of $n$ saddle node fixed points that we identify below pass through the segregated fixed point without any changes in the saddle nodes’ stabilities. Notice that as $\rho \to 1$ or $n \to \infty$, $\varepsilon^{(1)}_c \to 0$. By correlating the afferents strongly enough or considering enough of them, we can make this transcritical bifurcation occur for as small a value of $\varepsilon$ as we please.

The other two solutions for $w$ are the roots of the quadratic factor on the right-hand side of equation 3.7. For $\varepsilon = 0$, one solution is $w = 0$, or $v = e^1$, which we know to be stable. The other solution for $\varepsilon = 0$ is $w = 1/(n - 1)$, corresponding to $v = (n - e^1)/(n - 1)$. This other solution is easily shown to be an unstable saddle node for $n > 2$, while for $n = 2$, it corresponds to $n - e^1 \equiv e^2$, which is just the second, stable segregated solution, $v = e^2$. For $n > 2$, it is this fixed point and its $n - 1$ partner fixed points $(n - e^1)/(n - 1)$
for \( i > 1 \) at \( \varepsilon = 0 \) that move for \( \varepsilon > 0 \) and pass through the unsegregated fixed point at \( \varepsilon = \varepsilon_c^{(1)} \), changing the stability of the unsegregated fixed point. The discriminant of the quadratic factor is

\[
D(\varepsilon) = \varepsilon^2(n-2)^2 + 2n\varepsilon \left( (n-2) - 2(n-1)\kappa^{-1} \right) + n^2, \tag{3.9}
\]

and if there exists a real value of \( \varepsilon \), say \( \varepsilon_c^{(2)} \), for which \( D(\varepsilon_c^{(2)}) = 0 \), then the two roots of the quadratic equation become complex for \( \varepsilon > \varepsilon_c^{(2)} \) and a saddle node bifurcation occurs at \( \varepsilon = \varepsilon_c^{(2)} \) for \( n > 2 \) as these stable and unstable fixed points collide and annihilate. We obtain

\[
\varepsilon_c^{(2)} = \frac{n}{n-2} \left[ -1 + 2 \left( \frac{n-2}{n-1} \kappa \right)^{-1} \left( 1 - \sqrt{1 - \frac{n-2}{n-1} \kappa} \right) \right]. \tag{3.10}
\]

For \( n = 2 \), we can either take the limit \( n \to 2 \) in this equation or derive the critical value directly from the discriminant after explicitly setting \( n = 2 \). Either way, we obtain, for \( n = 2 \), \( \varepsilon_c^{(2)} = \kappa/2 \). The critical value \( \varepsilon_c^{(2)} \) is real if \( \kappa \leq \frac{n-1}{n-2} \), and since the upper bound on \( \kappa \) is \( \kappa < \frac{n}{n-1} < \frac{n-1}{n-2} \), the reality of \( \varepsilon_c^{(2)} \) is assured. We can verify that \( 0 \leq \varepsilon_c^{(2)} < 1 \). Furthermore, after some algebra, we can show that

\[
\varepsilon_c^{(2)} \geq \varepsilon_c^{(1)} , \tag{3.11}
\]

with equality only for the particular case \( n = 2 \). Thus, for \( n > 2 \), this saddle node bifurcation occurs after the transcritical bifurcation that occurs at \( \varepsilon = \varepsilon_c^{(1)} \). For \( n = 2 \), the saddle node bifurcation point and the transcritical bifurcation point coalesce, and since both solutions of the quadratic factor are stable prior to bifurcation, the dynamics for this specific \( n = 2 \) case are characterized by a supercritical pitchfork bifurcation occurring at \( \varepsilon = \varepsilon_c^{(1)} = \varepsilon_c^{(2)} \). In the limit \( \rho \to 1 \) or \( \kappa \to 0 \), we find that

\[
\varepsilon_c^{(2)} = \frac{n}{4(n-1)} \kappa + O(\kappa^2), \tag{3.12}
\]

so that \( \varepsilon_c^{(2)} \to 0 \) as \( \rho \to 1 \). However, in the limit \( n \to \infty \),

\[
\varepsilon_c^{(2)} \to -1 + 2\kappa^{-1} \left( 1 - \sqrt{1 - \kappa} \right), \tag{3.13}
\]

and this in general remains non-zero. Thus, while both \( \varepsilon_c^{(1)} \) and \( \varepsilon_c^{(2)} \) go to zero in the limit \( \rho \to 1 \), only \( \varepsilon_c^{(1)} \) goes to zero in the limit \( n \to \infty \), with \( \varepsilon_c^{(1)} \) in general remaining non-zero. Finally, computing the value of \( w \) from the root of the quadratic factor at the bifurcation point \( \varepsilon = \varepsilon_c^{(2)} \), we obtain
Bifurcations in Nonlinear Plasticity Models

\[
w(\varepsilon_c^{(2)}) = \frac{\kappa^{-1}}{n-2} \left( 1 - \sqrt{1 - \frac{n-2}{n-1}} \right), \tag{3.14}
\]

so that

\[
w(\varepsilon_c^{(2)}) \sim \begin{cases} 
\frac{1 - \kappa^{-1}}{n} \left( 1 - \sqrt{1 - \kappa} \right) & \text{for large } n \\
\frac{1}{2(n-1)} & \text{for small } \kappa
\end{cases} \tag{3.15}
\]

Thus, at this bifurcation point, \(w(\varepsilon_c^{(2)})\) is of order \(1/n\) for large \(n\). The overall factor controlling the deviation from \(1/n\) is \((1 - \sqrt{1 - \kappa})/\kappa \in [1/2, 1]\) for large \(n\), so \(w\) can jump from as low as \(1/(2n)\) to \(1/n\) at this second bifurcation point as it jumps to the unsegregated fixed point solution. A similarly sized jump can occur for small \(\kappa\), or \(\rho \sim 1\). For \(n = 2\), \(w(\varepsilon_c^{(2)}) = 1/2\), corresponding to exact coincidence with the unsegregated state \(n \mathbf{v} = \mathbf{n}\), reflecting the coalescing of the two bifurcation points into a supercritical pitchfork bifurcation.

We may now summarize these bifurcations dynamics. For \(n = 2\), there is a supercritical pitchfork bifurcation at \(\varepsilon_c \equiv \varepsilon_c^{(1)} = \varepsilon_c^{(2)}\). For \(\varepsilon < \varepsilon_c\), the unsegregated state, a fixed point for all \(\varepsilon\), is unstable, but for \(\varepsilon > \varepsilon_c\), it is stable. The two perfectly segregated fixed points at \(\varepsilon = 0\) are stable. As \(\varepsilon\) increases but remains below \(\varepsilon_c\), these two fixed points continue to be stable, but they drift away from being perfectly segregated to being only approximately segregated as the roots of the quadratic factor in equation 3.7 move away from \(w = 0\) and \(w = 1\). At \(\varepsilon = \varepsilon_c\), these real roots coalesce at \(w = 1/2\) and disappear, moving into the complex plane. For \(\varepsilon > \varepsilon_c\), only one fixed point remains, corresponding to the unsegregated state, which is now stable. The bifurcation diagram is sketched in Figure 1A.

For \(n > 2\), the dynamics are more complicated, with two distinct bifurcation processes. At \(\varepsilon = 0\), the \(n\) segregated fixed points are stable, and the unsegregated fixed point is unstable. For \(\varepsilon < \varepsilon_c^{(1)}\), the unsegregated fixed point remains unstable, but for \(\varepsilon > \varepsilon_c^{(1)}\), it is stable, a transcritical bifurcation occurring at \(\varepsilon = \varepsilon_c^{(1)}\) as \(n\) saddle node fixed points pass through the unsegregated fixed point. The \(n\) segregated fixed points drift as \(\varepsilon\) increases but remain stable as \(\varepsilon\) passes upward through \(\varepsilon_c^{(1)}\). However, as \(\varepsilon\) passes through \(\varepsilon_c^{(2)}\), all \(n\) (approximately) segregated fixed points disappear in saddle node bifurcations as they collide with the \(n\) saddle node fixed points that passed through the unsegregated fixed point at \(\varepsilon = \varepsilon_c^{(1)}\). Notice, therefore, that in the regime \(\varepsilon_c^{(1)} < \varepsilon < \varepsilon_c^{(2)}\), the (approximately) segregated fixed points and the unsegregated fixed point are all simultaneously stable, and, depending on the initial conditions, any of these \(n + 1\) stable fixed points could be found by the strength vector. This multistable regime can span a
Figure 1: Bifurcation diagrams for the minimal multiplicative model. Each graph shows the solutions \( w \) giving fixed points of the form \((1 - nw)\epsilon^1 + w n\) as a function of the cross-talk control parameter \( \epsilon \). Thick solid lines show stable solutions for \( w \), while thick dashed lines show unstable solutions. The thin dotted lines indicate critical values of \( \epsilon \) at which bifurcations occur, or values of \( w \) at such critical points. The bifurcation structure depends on whether \( n = 2 \) (A) or \( n > 2 \) (B).

large range of \( \epsilon \), since while \( \epsilon_c^{(1)} \) can be close to zero for large \( n \), \( \epsilon_c^{(2)} \) can remain non-zero in this limit. The bifurcation diagram for \( n > 2 \) is sketched in Figure 1B. In Figure 2 for \( n = 3 \) afferents, we illustrate the trajectories of all the fixed points on the positive part of the surface \( n \cdot v = 1 \) as \( \epsilon \) increases. The \( \epsilon = 0 \) stable fixed points at \( v = \epsilon_i \) and saddle node fixed points at \( v = (n - \epsilon_i)/(n - 1) \) move toward each other. As the latter pass through the unsegregated fixed point, it becomes stable. The approximately segregated fixed points and the saddle node fixed points subsequently collide and annihilate simultaneously.
Figure 2: Schematic illustration of the fixed points and their dependence on \( \varepsilon \) for \( n = 3 \) afferents in the minimal multiplicative model. The dynamics are constrained to exist on the plane \( \mathbf{n} \cdot \mathbf{v} = 1 \) in the positive hyperquadrant. This defines a triangular region, as shown. The tips of the triangle correspond to the \( \varepsilon = 0 \) segregated fixed points, shown as filled circles. The center of the triangle, shown as a gray circle, is the unsegregated fixed point \( \frac{1}{n} \mathbf{n} \), the location of which is independent of \( \varepsilon \). The three open circles at the centers of the lines connecting each pair of segregated fixed points correspond to saddle node fixed points. As \( \varepsilon \) increases, the fixed points move in the directions indicated by the arrows. The crosses denote the points at which the approximately segregated fixed points and the saddle node fixed points collide and annihilate in saddle node bifurcations.

3.2 The Rate-Based Limit of a Model of STDP. We now turn to an analysis of the rate-based limit of our stochastic model of STDP (Appleby & Elliott, 2005). Full details, derivation, and analysis of this model may be found elsewhere. In brief summary, the model proposes that the probability, but not amplitude, of synaptic strength change at single synapses depends on the timing of spikes. Thus, a spike pair will induce a fixed-amplitude change in strength or no change in strength, but the time window during which a change can be induced is stochastically determined by an internal, stochastic synaptic switch that is activated on the arrival of spikes. Averaged over multiple spike pairs or over multiple synapses, the classic, biphasic exponential response curve of STDP can be derived, but a single spike pair at a single synapse induces only a step-like change in strength. Such a model can be shown to define a non-Markovian random walk in synaptic strength, and on this view, it is possible to integrate out the switching and spiking dynamics of the neurons and determine a rate-based limit (Elliott, 2010).
When the potentiation and depression steps postulated by the model are of equal sizes, so that reflecting boundaries may be consistently erected in order to prevent synaptic strengths from becoming negative (Elliott, 2010, 2011), we can write the rate-based synaptic plasticity rule as

$$\dot{v}_i = P(a_i, a \cdot v) - D(a_i, a \cdot v),$$

(3.16)

where any overall factors have been absorbed into a rescaling of time. The functions $P$ and $D$ are the rates of potentiation and depression. For the purpose of generality, we do not initially specify the exact forms of these functions so that we may analyze a much broader class of models than just our own particular model of STDP. However, we will eventually need concrete forms for $P$ and $D$, and then we shall use those obtained from our model.

3.2.1 Preliminary General Analysis. The functions $P$ and $D$ will typically be complicated, nonlinear functions of their arguments, and averaging over the ensemble of afferent activity patterns will in general be analytically intractable. In order to proceed, we therefore make the simplifying assumption that the fluctuations in afferent activities around their common mean value, $\mu$, are sufficiently small that we can expand these functions up to $O(\sigma^2)$ only, discarding higher-order moments. Using the same form of correlation matrix as earlier, in equation 3.4, we may write

$$\langle P(a_i, a \cdot v) \rangle = P(\mu, \mu n \cdot v) + \frac{1}{2} \sigma^2 \left\{ \frac{\partial^2}{\partial x^2} + 2[(1 - \rho)v_i + \rho n \cdot v] \frac{\partial}{\partial x} \frac{\partial}{\partial y} \right\} P(\mu, \mu n \cdot v),$$

(3.17)

where $\partial_x$ and $\partial_y$ mean derivatives with respect to the first and second arguments, respectively, of $P$, and all derivatives are evaluated at $x = \mu$ and $y = \mu n \cdot v$. We may therefore take the ensemble-averaged function $\Pi(a, v)$ to be

$$\langle \Pi(a, v) \rangle = n P(\mu, \mu n \cdot v) + \frac{1}{2} \sigma^2 \left\{ n \frac{\partial^2}{\partial x^2} + 2[(1 - \rho)v + \rho n \cdot v] \frac{\partial}{\partial x} \frac{\partial}{\partial y} \right\} P(\mu, \mu n \cdot v).$$

(3.18)

An identical form follows for $\langle \Delta(a, v) \rangle$ with the replacement $P \rightarrow D$. Under the action of the cross-talk matrix on the potentiating term, we have

$$\mathbb{E}(\Pi(a, v)) = \langle \Pi(a, v) \rangle + (1 - \rho) \sigma^2 \left\{ \frac{\partial}{\partial x} \frac{\partial}{\partial y} P(\mu, \mu n \cdot v) \right\} (\mathbb{E} - n) v.$$

(3.19)
We may now determine the conditions on \( P \) and \( D \) that ensure that there exists a set of \( n \) stable, segregated fixed points and an unstable, unsegregated fixed point and how these conditions are modified by the cross-talk parameter \( \epsilon \). Because the depressing dynamics determined by \( D \) need not necessarily shut down depression when a synapse’s strength is zero, we must prevent a synapse’s strength from being driven negative by erecting reflecting boundaries at zero synaptic strength. Thus, at a segregated fixed point \( \mathbf{v} \) at which all components but, say, \( v_i \) are zero, the fixed point conditions become, in the presence of reflecting boundaries, \( \dot{v}_i = 0 \) and \( \dot{v}_j < 0 \) for \( j \neq i \).

At \( O(\sigma^0) \), so ignoring any fluctuations in activity around the mean firing rate, the fixed points are the solutions of the equation \( P(\mu, \mu \cdot \mathbf{v}) = D(\mu, \mu \cdot \mathbf{n} \cdot \mathbf{v}) \). This equation defines an entire hyperplane of solutions, \( \mathbf{n} \cdot \mathbf{v} = w_0 \), where \( w_0 \) determines the distance of this hyperplane from the origin. Clearly we require \( w_0 > 0 \), otherwise, nonnegative solutions for the strength vector are not available. Further, we require this hyperplane to be stable. It is easy to see that the condition for the stability of this hyperplane is

\[
\partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)] < 0, \tag{3.20}
\]

where the \( y \) derivative is now evaluated at \( y = \mu \cdot \mathbf{n} \cdot \mathbf{v} = \mu w_0 \), that is, on the fixed point hyperplane. All flow toward this hyperplane is parallel to the vector \( \mathbf{n} \), reflecting the fact that the stability matrix is proportional to \( -\mathbf{n} \mathbf{n}^T \), which has eigenvector \( \mathbf{n} \) with eigenvalue \( -n \), and \( n-1 \) repeated eigenvalues of 0, corresponding to the \( n-1 \) eigenvectors in the directions orthogonal to \( \mathbf{n} \). Because the cross-talk parameter \( \epsilon \) does not appear at \( O(\sigma^0) \), the conditions for the positivity and stability of this hyperplane are unaffected by \( \epsilon \).

The \( O(\sigma^2) \) terms split the degeneracy among all the points on the hyperplane \( \mathbf{n} \cdot \mathbf{v} = w_0 \), leaving only a small, finite subset of points as fixed points, the most pertinent of which are the \( n \) segregated states and the fully unsegregated state. We first examine the segregated states. Without loss of generality, we focus on the state \( w \mathbf{e}_1 \), where \( w \) is to be determined. Since we have expanded the dynamics to \( O(\sigma^2) \), we write \( w = w_0 + \sigma^2 w^*_2 \), where the \( w^*_2 \) term is the correction to the location of the \( O(\sigma^0) \) fixed point. To determine \( w^*_2 \), we solve \( \dot{v}_1 = 0 \) and find that, after some algebra,

\[
w^*_2 = -\left( \frac{1}{2\mu} \right) \frac{\partial_x + w_0 \partial_y} {\partial_y} \frac{\partial_x \partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]} {\partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]} \]

\[\epsilon(1-\rho) \left( \frac{n-1}{n\mu} \right) w_0 \frac{\partial_x \partial_y P(\mu, \mu w_0)} {\partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]}. \tag{3.21}\]

This gives the correction, and unlike \( w_0 \), it explicitly depends on \( \epsilon \). However, even in the presence of \( \epsilon \), the solution \( (w_0 + \sigma^2 w^*_2) \mathbf{e}_1 \) remains an exactly
segregated state. This is in contrast to the influence of cross-talk in the MMM, in which the fixed points are only approximately segregated for $\varepsilon \neq 0$.

The stability of the fixed point $w \ e^1$ in the $e^1$ direction is inherited from the stability of the $O(\sigma^0)$ fixed hyperplane, since $\mathbf{n} \cdot e^1 \neq 0$. There will be $O(\sigma^2)$ corrections to the stability in this direction, but the principal determinant of the character of this fixed point will be the stabilities in the other, $e^i$, $i \neq 1$, directions. Because the $O(\sigma^0)$ stability matrix has $n - 1$ repeated eigenvalues of 0, the stabilities in the directions orthogonal to $e^1$ will be governed by terms that are purely $O(\sigma^2)$, with vanishing $O(\sigma^0)$ contributions. Since a segregated fixed point has at least one component that is zero, the stability conditions in these orthogonal directions are, as discussed above, just $\dot{v}_i < 0$, $i \neq 1$. A little algebra shows that these conditions reduce to the requirement that

$$
(1 - \rho)\sigma^2 \partial_x \partial_y \left[(1 - \varepsilon)P(\mu, \mu w_0) - D(\mu, \mu w_0)\right] > 0 \quad (3.22)
$$

for stability of the segregated fixed point $w \ e^1$, and therefore for all $n$ segregated fixed points. For $\rho \neq 1$, we can immediately read off the critical level of cross-talk required to destabilize the segregated fixed points:

$$
\varepsilon_c = \frac{\partial_x \partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]}{\partial_x \partial_y P(\mu, \mu w_0)}.
$$

(3.23)

Assuming that the various positivity and stability conditions can be met for $\varepsilon = 0$, then the segregated fixed points are stable for $\varepsilon < \varepsilon_c$ and unstable for $\varepsilon > \varepsilon_c$. We discuss below the correct classification of this critical process.

We now briefly examine the unsegregated fixed point. We write the solution in the form $n^{-1} \ w \ \mathbf{n}$, with $w = w_0 + \sigma^2 w_2^u$. We require $\mathbf{v} = 0$, from which we obtain

$$
w_2^u = - \left[\frac{(n - 1)\rho + 1}{2n\mu}\right] \left(\frac{\partial_x + w_0 \partial_y}{\partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]}\right)^2
$$

$$
\quad - (1 - \rho)\left(\frac{n - 1}{2n\mu}\right) \frac{\partial^2_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]}{\partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)]}.
$$

(3.24)

The location of this fixed point does not depend on $\varepsilon$. As in the MMM, there is always an unsegregated fixed point, the location of which is unaffected by the presence of the cross-talk matrix $\mathbb{E}$ because $\mathbf{n}$ is an eigenvector of $\mathbb{E}$ of unit eigenvalue. The analysis of the stability of this unsegregated fixed point is slightly more involved than that for the segregated fixed points because it is necessary to determine the full stability matrix by expanding and linearizing about this fixed point in the standard manner. The stability
matrix can finally be written in the form

$$\begin{align*}
\{ \mu \partial_y [P(\mu, \mu w_0) - D(\mu, \mu w_0)] + O(\sigma^2) \} \mathbf{n} \mathbf{n}^T \\
+ \{ (1 - \rho)\sigma^2 \partial_x \partial_y [(1 - \varepsilon)P(\mu, \mu w_0) - D(\mu, \mu w_0)] \} \mathbb{I}.
\end{align*}$$

(3.25)

The $O(\sigma^2)$ term in the $\mathbf{n} \mathbf{n}^T$ part is independent of $\varepsilon$. We recognize the $O(\sigma^0)$ contribution as giving the stability condition for the hyperplane $\mathbf{n} \cdot \mathbf{v} = w_0$ in equation 3.20. In the $\mathbf{n}$ direction, this unsegregated fixed point therefore inherits its stability from the $O(\sigma^0)$ hyperplane, with $O(\sigma^2)$ corrections. The stabilities in the directions orthogonal to $\mathbf{n}$ are, however, governed by the term proportional to $\mathbb{I}$. This is purely $O(\sigma^2)$, and we recognize its form from equation 3.22. Thus, when the segregated fixed points are stable, the unsegregated fixed point is unstable, and vice versa. The segregated fixed points and the unsegregated fixed point therefore undergo coordinated but opposite stability changes at the same critical value of the cross-talk parameter, $\varepsilon = \varepsilon_c$. We discuss classification of this critical process below.

Little further progress can be made without an explicit specification of the forms for $P$ and $D$. In particular, we require these functions in order to characterize the nature of the critical processes in which fixed points’ stabilities change. We therefore now restrict the discussion to our stochastic model of STDP.

### 3.2.2 Particular Analysis for a Model of STDP

Having determined in general the stability criteria and their dependence on the cross-talk parameter $\varepsilon$, we now consider the particular case of our stochastic model of STDP (Appleby & Elliott, 2005). The average asymptotic rates of potentiation and depression in this model can be shown to be (Elliott, 2010)

$$\begin{align*}
P(x, y) &= \frac{x y (x + \lambda_{-})}{(x + y + \lambda_{+})(x + y + \lambda_{-}) - x y}, \\
D(x, y) &= \frac{x y (y + \lambda_{+})}{(x + y + \lambda_{+})(x + y + \lambda_{-}) - x y},
\end{align*}$$

(3.26, 3.27)

where $\lambda_{\pm}$ are inverse time constants that determine the average widths of the STDP time windows for potentiation in response to a presynaptic-before-postsynaptic ($\lambda_{+}$) spike pair or a postsynaptic-before-presynaptic ($\lambda_{-}$) spike pair. Typically we set $\lambda_{-} = 50$ Hz, or an average time window of 20 ms, and define a parameter $\gamma > 0$ via $\lambda_{+} = \lambda_{-}/\gamma$ that sets $\lambda_{+}$. These expressions are the asymptotic rates of potentiation and depression in this stochastic model, but asymptotic behavior is achieved swiftly (Elliott, 2010).

With $P$ and $D$ defined by equations 3.26 and 3.27, we can translate our general criteria into explicit requirements on the parameters defining the model. The $O(\sigma^0)$ location of the hyperplane of fixed points is
Figure 3: Contours of the critical value of the control parameter, $\varepsilon_c$, in the $\gamma-\hat{\mu}$ plane, where $\hat{\mu} = \mu/\lambda_-$, for the stochastic model of STDP. The thick line represents the boundary $\gamma = 1/(1 + \hat{\mu})$, or $\hat{\mu} = (1 - \gamma)/\gamma$, separating the parameter region in which $w_0 < 0$ to its left from the parameter region in which $w_0 > 0$ to its right. Within the $w_0 > 0$ region, contours of $\varepsilon_c$ are shown corresponding, moving left to right, to values of $\varepsilon_c = 2^{-1}, 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5},$ and $2^{-6}$.

Given by $w_0 = (\mu + \lambda_- - \lambda_+)/\mu$, and since we require $w_0 > 0$, we need $\gamma > \lambda_-/(\lambda_+ + \mu)$. In the absence of cross-talk, the condition for the stability of the segregated fixed points (and therefore for the instability of the unsegregated fixed point) reduces simply to $\gamma < 1$, or $\lambda_+ > \lambda_-$. The expression for the critical value $\varepsilon_c$ is rather messy, but we reproduce it here for the sake of completeness. Writing $\hat{\mu} = \mu/\lambda_-$, we obtain

$$
\varepsilon_c = \frac{1}{2}(1 - \gamma)(1 + 2\hat{\mu})[(1 - 2\gamma) - 2\gamma\hat{\mu}][(1 - 2\gamma) - 3\gamma\hat{\mu}]
$$

$$
\times [(1 - 2\gamma)(1 - 3\gamma + \gamma^2) + (2 - 13\gamma + 19\gamma^2 - 3\gamma^3)\hat{\mu}]
$$

$$
- \gamma(7 - 13\gamma - 6\gamma^2)\hat{\mu}^2 + \gamma^2(1 + 13\gamma)\hat{\mu}^3 + 6\gamma^3\hat{\mu}^4]^{-1}.
$$

(3.28)

Notice that $\varepsilon_c \to 0$ as $\gamma \to 1$ and as $\hat{\mu} \to \infty$. The critical value $\varepsilon_c$ depends on only on these two parameters, $\gamma$ and $\hat{\mu}$. In Figure 3 we show contours of $\varepsilon_c$ in that part of the $\gamma-\hat{\mu}$ plane for which $w_0 > 0$. The critical value $\varepsilon_c$ can be rather small in a large region of parameter space. For example, roughly speaking, for $\gamma > 3/4$, $\varepsilon_c < 0.1$.

In the MMM, the two critical values $\varepsilon_c^{(1)}$ and $\varepsilon_c^{(2)}$ both go to zero as $\rho \to 1$. Our expression for $\varepsilon_c$ in the stochastic STDP model, however, does not depend on $\rho$, since it is derived from a purely $O(\sigma^2)$ term with an overall
multiplier of $1 - \rho$. Thus, we must have $\rho < 1$, but we learn nothing about the behavior of $\varepsilon_c$ in the limit $\rho \to 1$ from this approximation. For $\rho \approx 1$, the purely $O(\sigma^2)$ term will be suppressed by the overall factor $1 - \rho$, and it is therefore unclear whether the expression for $\varepsilon_c$ remains a good approximation or whether it is dominated by higher-order corrections in the expansion. To address this, we compare our expression for $\varepsilon_c$ in equation 3.28 with the exact but numerical solutions of the fixed point equations, without the expansion in powers of $\sigma$. This requires a specification of parameter values and the form of the afferent activity ensemble. To this end, we consider just $n = 2$ afferents, set $\lambda_ - = 50$ Hz, and employ standard values used elsewhere, namely, $\mu = 50$ Hz, $\sigma = 12.5$ Hz, corresponding to 25% fluctuations around mean activity (Elliott, 2011). The ensemble is simply defined by the four activity patterns $a_i = \mu \pm \sigma$, for $i = 1$ and $i = 2$, with correlation coefficient $\rho$, so that the patterns $a^T_1 = (\mu - \sigma, \mu - \sigma)$ and $a^T_2 = (\mu + \sigma, \mu + \sigma)$ both have probability $(1 + \rho)/4$, while the patterns $a^T_3 = (\mu - \sigma, \mu + \sigma)$ and $a^T_4 = (\mu + \sigma, \mu - \sigma)$ both have probability $(1 - \rho)/4$. In Figure 4, we plot the ratio of the numerically obtained values of $\varepsilon_c$ to the approximate form for $\varepsilon_c$ in equation 3.28 as a function of the correlation coefficient $\rho$, for different choices of $\gamma$. We see that the analytical results are within at worst 10% of the numerical results and that, in particular, the numerical result for $\varepsilon_c$ does not go to zero at $\rho \to 1$, in contrast to the MMM. The form in equation 3.28 therefore remains a good approximation even for $\rho \approx 1$, and its nonvanishing behavior in the limit $\rho \to 1$ therefore reflects the exact behavior.
We saw in the MMM for \( n > 2 \) that there is a multistable regime \( \varepsilon_c^{(1)} < \varepsilon < \varepsilon_c^{(2)} \) in which the (approximately) segregated fixed points and the unsegregated fixed point are simultaneously stable. However, in the stochastic model of STDP, at \( O(\sigma^2) \), we have seen instead an as-yet-uncharacterized critical process in which the segregated fixed points and the unsegregated fixed point undergo simultaneous but opposite stability changes. In previous work unrelated to issues of cross-talk in Hebbian models of synaptic plasticity, we in fact found parameter regimes in which both sets of fixed points could be simultaneously stable in the stochastic model of STDP (Appleby & Elliott, 2006). This regime was beyond the small \( \sigma \) approximation necessary to derive analytical results in that work and here. It is therefore possible that by moving beyond this approximation, we would also find a cross-talk-induced multistable regime identical in character to that seen in the MMM.

Restricting for the moment to \( n = 2 \) afferents, we confirm that this is indeed the case in Figure 5, in which we plot the two-dimensional flow for \( \dot{v} \) for the same parameters used in Figure 4, but using larger fluctuations \( \sigma/\mu = 3/4 \) in the four activity patterns, and setting \( \rho = -3/4 \). Increasing \( \varepsilon \), we observe a tristable regime in which the two segregated fixed points and the unsegregated fixed point are simultaneously stable. In this tristable regime, we clearly see an additional pair of unstable saddle node fixed points whose existence is necessary in order to render the global flow consistent. These dynamics demonstrate that the \( n = 2 \) unsegregated fixed point undergoes a subcritical pitchfork bifurcation, shedding a pair of saddle node fixed points and changing stability from unstable to stable at a critical value, \( \varepsilon = \varepsilon_c^{(1)} \). The segregated fixed points later, at \( \varepsilon = \varepsilon_c^{(2)} > \varepsilon_c^{(1)} \), undergo a change of stability when the saddle node fixed points created by the bifurcation at \( \varepsilon = \varepsilon_c^{(1)} \) reach them. The change of stability of the segregated fixed points is not strictly a critical process at all (a change in the character of the solutions of an equation at some critical parameter value), because the reflecting boundary conditions in part establish these segregated states as fixed points of the flow. For the segregated fixed point \( \mathbf{v} = w \mathbf{e}_1 \), say, the fixed point condition is simply \( \dot{v}_i = 0 \), determining \( w \), but the stability conditions are \( \dot{v}_i < 0 \), for \( i > 1 \). The reflecting boundaries halt the negative flow, creating the fixed point. However, as a saddle node created by the subcritical pitchfork bifurcation of the \( n = 2 \) unsegregated fixed point reaches a stable, segregated fixed point, the stability-inducing negative flow through the segregated fixed point changes. The saddle node actually passes straight through the segregated fixed point into a hyperquadrant with at least one negative component of \( \mathbf{v} \), and this reverses the direction of the flow through the segregated fixed point, violating the stability conditions and rendering the segregated fixed point unstable. For simplicity, we refer to this change of stability of the segregated fixed points in this model as a pseudocritical process.
Figure 5: The flow of the vector field $\dot{v}$ for $n = 2$ afferents in the stochastic model of STDP. Averaging over the four activity patterns $a_i = \mu \pm \sigma$, for $i = 1$ and $i = 2$, has been performed explicitly, avoiding the small $\sigma$ expansion. The common parameters used to generate these flows are $\mu = 50$ Hz, $\sigma = 37.5$ Hz so that $\sigma/\mu = 3/4$, $\lambda = 50$ Hz, $\gamma = 0.6$, and $\rho = -3/4$. The cross-talk control parameter $\varepsilon$ is set in each part as (A) 0.000, (B) 0.025, (C) 0.050, (D) 0.075, (E) 0.100, (F) 0.125.
Figure 6: The $v_1$ components of all the fixed points for $n = 2$ afferents in the stochastic model of STDP, as a function of the cross-talk control parameter $\varepsilon$. All other parameters and averaging are as in Figure 5. Solid lines show stable solutions and dashed lines unstable solutions. The line at constant value $v_1 \approx 0.32$ corresponds to the unsegregated fixed point, the location of which does not depend on $\varepsilon$. The two lines starting at $v_1 \approx 0.58$ and $v = 0$ and continuing beyond the collisions with the two saddle node fixed points at $\varepsilon = \varepsilon_c^{(2)} \approx 0.1197$ correspond to the two segregated fixed points. For $\varepsilon > \varepsilon_c^{(2)}$, the saddle nodes continue into the nonpositive hyperquadrants and are not shown.

To exhibit the subcritical pitchfork bifurcation of the $n = 2$ unsegregated fixed point explicitly, we show in Figure 6 the $v_1$ components of all three (non-zero) fixed points of the unapproximated dynamics as a function of $\varepsilon$, for the same parameters used to generate Figure 5. For these parameters, we observe the subcritical pitchfork bifurcation changing the stability of the unsegregated fixed point at $\varepsilon = \varepsilon_c^{(1)} \approx 0.0226$. The subsequent pseudocritical process occurs at $\varepsilon = \varepsilon_c^{(2)} \approx 0.1197$ when the saddle nodes pass through the segregated fixed points. To determine each of these critical values of $\varepsilon$, we must for this problem find the real root of a septic polynomial. A detailed analytical characterization of the dependence of these critical values on the five parameters $\mu, \sigma, \rho, \lambda, \gamma$ when $\sigma/\mu$ is not small is therefore intractable, even for $n = 2$ afferents.

The subcritical pitchfork bifurcation of the $n = 2$ unsegregated fixed point is not generic, however. For $n > 2$, a different critical process changes the stability of this fixed point. For concreteness, we now consider $n = 3$ afferents, but the results are generic for any $n \geq 3$. It is virtually impossible to visualize two-dimensional projections of the three-dimensional flow $\dot{v}$, so we consider a privileged two-dimensional slice through this...
flow in order to illustrate the bifurcation dynamics. We change variables to the new orthonormal system $u_1 = n \cdot v / \sqrt{3}$, $u_2 = (-1, -1, +2) v / \sqrt{6}$, and $u_3 = (-1, +1, 0) v / \sqrt{2}$. Up to a normalizing constant of proportionality, the coordinate $u_1$ corresponds to the total synaptic strength, and thus each value of $u_1 > 0$ defines a triangular surface in the nonnegative hyperquadrant of $\mathbb{R}^3$. The coordinates $u_2$ and $u_3$ then determine the location on this triangular surface, given $u_1$. We consider the particular slice of the three-dimensional space corresponding to $u_3 = 0$ or $v_1 = v_2$. The allowed range of $u_2$ is then $[-u_1 / \sqrt{2}, +u_1 \sqrt{2}]$. This slice contains the origin, the unsegregated fixed point, and the segregated fixed point $v = w e^3$, with $w$ being the strength of the third afferent in the segregated state. By the symmetry of the three afferents, in particular between $v_1$ and $v_2$, on this slice we have $\dot{u}_3 = 0$, so the flow exactly on the slice $u_3 = 0$ remains on it. Of course, we could choose the two other privileged slices containing the origin, the unsegregated fixed point, and either $v = w e^1$ or $v = w e^2$. Because any surface $u_1 = \text{const}$ is orthogonal to $n$, $\dot{u}_1$ does not depend on $\varepsilon$.

For $n = 3$ afferents, we use the same parameters as in Figures 5 and 6 except that for simplicity, we set $\rho = 0$ and also set the third-order correlation coefficient to 0, so that all three afferents’ activities are uncorrelated. In Figure 7, we plot the curves for $\dot{u}_1 = 0$ (independent of $\varepsilon$) and $\dot{u}_2 = 0$ in the allowed part of the $u_1$–$u_2$ plane for three different values of $\varepsilon$. The intersections of these curves define some of the fixed points in the slice $u_3 = 0$; the other fixed points are established in part by reflecting boundary conditions, as discussed. For $\varepsilon$ below a critical value, there are only two intersections, at $u_1 = 0$ and $u_2 = 0$, since the origin is always a fixed point (not shown), and at $u_1 \approx 0.3599$ and $u_2 = 0$. The $u_1 \neq 0$ fixed point is the unsegregated fixed point, and its $u_1$ location must be determined from the real root of a degree 11 polynomial. As $\varepsilon$ increases, the non-zero solution curve for $\dot{u}_1 = 0$ moves to the left, toward smaller values of $u_1$. At a critical value of $\varepsilon$, a pair of fixed points is created as the solution curve for $\dot{u}_1 = 0$ tangentially touches that for $\dot{u}_2 = 0$. Thus, we obtain a saddle node bifurcation for some critical value of $\varepsilon$, in which a pair of fixed points in this $u_3 = 0$ slice is created. In fact, these fixed points are saddle nodes. This critical touching occurs away from the $u_2 = 0$ axis, as demonstrated in the inset of Figure 7, so the unsegregated fixed point for $n = 3$ (and for $n > 3$) does not undergo a subcritical pitchfork bifurcation, as it does for $n = 2$. We find that this saddle node bifurcation occurs at the critical value $\varepsilon = \varepsilon_c^{(0)} \approx 0.0852$, for the parameter values used here. We see from Figure 7 that as $\varepsilon$ increases above $\varepsilon_c^{(0)}$, both saddle node fixed points must track along the $\dot{u}_1 = 0$ curve. Hence, one of this pair must pass straight through the unsegregated fixed point on the $u_2 = 0$ axis. As it does so, at critical value $\varepsilon = \varepsilon_c^{(1)} \approx 0.0866$, the unsegregated fixed point becomes stable in a transcritical bifurcation. Both fixed points then move toward the gray, bounding regions, passing into hyperquadrants with at least one negative $v_i$ component. The saddle node
Figure 7: Solution curves for $\dot{u}_1 = 0$ (dashed line) and $\dot{u}_2 = 0$ (solid lines) in the $u_1$–$u_2$ plane for $n = 3$ afferents in the stochastic model of STDP. With $u_3 = 0$, the variable $u_2$ takes the range $[-u_1/\sqrt{2}, +u_1\sqrt{2}]$, so areas outside this valid range of $u_2$ are shaded in gray. The curve for $\dot{u}_1 = 0$ does not depend on $\varepsilon$. For $\dot{u}_2 = 0$, we have used three different values of $\varepsilon$, moving right to left corresponding to $\varepsilon = 0.07$, $\varepsilon = 0.0852$, and $\varepsilon = 0.10$. The entire line $u_2 = 0$ is also a solution of $\dot{u}_2 = 0$. The inset magnifies the region around $u_1 = 0.36$, demonstrating that the solution curves for $\dot{u}_1 = 0$ and $\dot{u}_2 = 0$ first touch at some value $u_2 \neq 0$. The parameters used to generate these curves are otherwise identical to those used in Figure 5, except that we have set $\rho = 0$ and also set the third-order correlated coefficient to zero.

that passed through the unsegregated fixed point first reaches the lower boundary, $u_2 = -u_1/\sqrt{2}$, at $\varepsilon \approx 0.1399$. This boundary point corresponds to a vector $\mathbf{v} \propto (n - e^3)$ and is a saddle node separating the two segregated fixed points $\mathbf{v} = w e^1$ and $\mathbf{v} = w e^2$ (see Figure 2). As the saddle node created at $\varepsilon = \varepsilon_c(0)$ passes through this boundary saddle node, the character of the latter changes. The other saddle node created at $\varepsilon = \varepsilon_c(0)$ moves toward the other boundary, at $u_2 = +u_1\sqrt{2}$, corresponding to $\mathbf{v} = w e^3$, the third segregated fixed point. As this second saddle passes through the segregated fixed point, at $\varepsilon = \varepsilon_c(2) \approx 0.2090$, the segregated fixed point becomes unstable, as for the $n = 2$ dynamics. We can therefore interpret the intersections of the $\dot{u}_1 = 0$ curve with the two boundaries $u_2 = -u_1/\sqrt{2}$ and $u_2 = +u_1\sqrt{2}$ as defining the locations of the segregated fixed point $\mathbf{v} \propto e^3$ and its dual, saddle node fixed point $\mathbf{v} \propto (n - e^3)$ at these two fixed points’ corresponding (pseudo-)critical points.

Of course, two additional pairs of saddle node fixed pairs are also simultaneously created in the other two privileged slices, containing the other two segregated fixed points. For $n = 3$, we therefore obtain in total three
simultaneous saddle node bifurcations, giving rise to six saddle nodes that change the stabilities of the unsegregated fixed point and then the segregated fixed points. In Figure 8, we attempt to show these six fixed points by plotting the flow in the valid part of the $u_3-u_2$ plane for a fixed value of $u_1$, $u_1 = 0.35$, for $\epsilon = 0.11$. Of course, any plane $u_1 = \text{const.}$ is not actually fixed, that is, $\dot{u}_1 \neq 0$ in general, and we see from Figure 7 that $\dot{u}_1 = 0$ actually defines a curved triangular surface on which the fixed points live and move as $\epsilon$ changes. Thus, Figure 8 for fixed $u_1$ is somewhat misleading, in that the points advertised as fixed are not genuine fixed points of the three-dimensional flow. To see the six additional fixed points, we would need a projection with a broad depth of focus in the $u_1$ direction. Nevertheless, we clearly see six “fixed points” in this projection, albeit shadows of the real fixed points in different focal planes, which partition the triangular region into different basins of attraction, one of which contains the now-stable unsegregated fixed point and others contain the still-stable segregated fixed points.

Figure 8: The flow $(\dot{u}_3, \dot{u}_2)$ in the $u_3-u_2$ plane for the particular value $u_1 = 0.35$ for $n = 3$ afferents in the stochastic model of STDP. The parameters are identical to those used in Figure 7, except that we have set $\epsilon = 0.11$. Invalid ranges of the plane are shaded in gray. The points shown as small black disks are not strictly fixed points, because any slice $u_1 = \text{const.}$ is not fixed. In order to see all genuine fixed points for $n = 3$, we would require a projection of the full three-dimensional flow with a broad depth of focus. Nevertheless, the black disks are shadows of the real saddle node fixed points in the multistable regime in which the unsegregated and segregated fixed points are simultaneously stable. The flows in this particular projection illustrate the basins of attraction of these four stable fixed points.
For $n \geq 3$, we obtain $n$ saddle node bifurcations. These bifurcation
dynamics for $n \geq 3$ are strongly reminiscent of those in the MMM in section 3.1
but in reverse order. In the MMM, $2n$ fixed points, given at $\varepsilon = 0$ by $e^i$ and
$(n - e^i)/ (n - 1), i = 1, \ldots, n$, march toward each other as $\varepsilon$ increases, half
of them passing through the unsegregated fixed point, turning it stable,
and then all simultaneously annihilating in $n$ saddle node bifurcations (see
Figure 2). In the stochastic STDP model, instead $n$ saddle node bifurcations
create $2n$ saddle node fixed points. These fly apart, half of them passing
through the unsegregated fixed point, turning it stable, and then pass
through the segregated fixed points and their “duals,” changing their stabil-
ities. Thus, the two sets of dynamics are nearly mirror images of each other.
However, the stochastic STDP model must create saddle nodes ex nihilo
because the segregated fixed points, unlike in the MMM, remain exactly
segregated, so they cannot drift toward each other.

We saw none of these bifurcation dynamics, for either $n = 2$ or $n > 2$, in
the $O(\sigma^2)$ analysis performed in section 3.2.1. At $O(\sigma^2)$, we observe what
appears to be pure stability changes of the fixed points. The saddle nodes
would be visible in higher-order terms. At $O(\sigma^2)$, we therefore see only
a shadow of the full bifurcation processes. For the same parameters used
in Figures 5 and 6 for $n = 2$, but with $\sigma/\mu = 1/4$ rather than $3/4$, we can
compute the two critical values to be $\varepsilon_c^{(1)} \approx 0.2322$ and $\varepsilon_c^{(2)} \approx 0.2397$. Thus,
the regime of tristability for $n = 2$ afferents occurs over only a very narrow
range of the cross-talk parameter $\varepsilon$ when $\sigma/\mu$ is small. Our $O(\sigma^2)$ result
gives a single critical value of $\varepsilon_c = 1/4$, for comparison. Higher-order cor-
rections in the $\sigma$ expansion would therefore split this single, simultaneous,
apparently pure stability change into discrete, separate bifurcation pro-
cesses in general. The advantage of our approximate results above is that
they provide some analytical insight into these critical processes, providing
a reliable estimate of $\varepsilon_c \approx \varepsilon_c^{(0)} \approx \varepsilon_c^{(1)} \approx \varepsilon_c^{(2)}$ in the regime of validity of the ap-
proximation made—namely, small $\sigma/\mu$. Their disadvantage is that we see
only a shadow of the full (pseudo-)critical processes, but this disadvantage
is outweighed by both the generality of our analytical results for the general
functions $P$ and $D$ and the knowledge that even with a specification of $P$
and $D$ for any particular model, as here with our stochastic model of STDP,
exact analytical results will in general not be available.

4 Cross-Talk in a Learning Model

The nonlinear developmental models considered above have in some sense
a rather trivial fixed point structure, imposed by the experimental require-
ment to exhibit activity-dependent competitive dynamics, leading to the
emergence of segregated states of afferent connectivity from an initially
unsegregated state under normal rather than abnormal developmental
regimes (Purves & Lichtman, 1985). The more general analysis performed
in section 3.2 reveals that under a broad class of models, we expect to see a
cross-talk-induced stabilization of the unsegregated state and a destabi-
lation of the segregated states. The precise details of these (possibly pseudo-)
critical processes will inevitably depend on the particular nonlinearities of
the functions $P$ and $D$. And whether cross-talk will lead to the drifting of
segregated states away from perfect segregation before bifurcation (as in
the MMM), or whether these states remain perfectly segregated right up to
the (possibly pseudo-)critical process (as in the stochastic model of STDP),
will also depend on the exact details of the model. It is natural to wonder,
though, to what extent the very simple fixed point structure of these devel-
opmental models, and concomitantly the symmetry assumed in the choice
of $C$ in equation 3.4, contribute to the critical processes discussed above.

To this end, we now consider a popular nonlinear model of learning,
specifically ICA (Hyvarinen et al., 2001). In ICA, statistically independent
nongaussian sources (the independent components) are linearly mixed
through a mixing matrix. Samples of these mixtures form the inputs to
a neuronal network, and the synaptic strengths must be adjusted so that an
output neuron reports faithfully precisely one of the independent sources.
The learning task is thus for the synaptic strengths between the input neu-
rons and an output neuron to acquire a row of the inverse of the mixing
matrix. Our interest in ICA here is merely as a convenient, well-understood,
nonlinear model with a fixed point structure (the rows of the inverse mix-
ing matrix) that is nontrivial. Our use of ICA for this purpose implies no
commitment to whether the linear mixing of independent sources is a good
or even approximate characterization of the signals that (sensory) neurons
receive (see the discussion in Lyu & Simoncelli, 2009), or to whether bi-
ologically plausible models of synaptic plasticity can actually achieve an
implementation of ICA (see Triesch, 2007; Savin, Joshi, & Triesch, 2010). In
fact, we select ICA as an example principally because an existing numerical
study has already examined the impact of cross-talk on this model (Cox &
Adams, 2009).

Because we wish to continue analyzing a single postsynaptic neuron for
the purposes of analytical tractability, we focus on the single-unit ICA rule
(Hyvarinen & Oja, 1998). We select the cubic nonlinearity as an approxima-
tion to negentropy since this greatly facilitates analysis. With a whitened,
that is, orthogonal $n \times n$ mixing matrix $M$ with columns $\{m^1, \ldots, m^n\}$, the
synaptic plasticity rule, after explicitly evaluating the expectation value
over source probability distributions, becomes, in the presence of the cross-
talk matrix $E$,

$$\dot{v} = \gamma P_v \mathbb{E} \left[ \sum_{i=1}^{n} k_i (m^i \cdot v)^3 m^i + 3 v \right]. \tag{4.1}$$

Here $k_1, \ldots, k_n$ are the (excess) kurtoses of the $n$ sources, and $\gamma = \pm 1$ is the
standard factor necessary to take into account either supergaussian ($\gamma = +1$) or subgaussian ($\gamma = -1$) source distributions. The projection matrix $P_v,$
implements, as above, M2 normalization, preserving the magnitude of the strength vector \( \mathbf{v}, |\mathbf{v}|^2 = 1 \).

We must draw attention to an inconsistency in the formulation of these ICA dynamics in the presence of cross-talk. In section 2, we considered a careful partitioning of synaptic plasticity rules into their potentiation and depression parts and discussed the possible impact of cross-talk matrices acting on each part separately, arguing that these distinct forms of synaptic plasticity would be subject to different forms of cross-talk. However, the components of the terms in square brackets in equation 4.1 can be positive or negative, depending on the columns \( \mathbf{m}_i \) and also on the strength vector \( \mathbf{v} \), since in ICA, it can and in general must acquire some negative components. The term in square brackets therefore contains both potentiation and depression processes, but we have applied \( E \) to it as if it represented pure potentiation dynamics, and then allowed the normalizing projection matrix \( P_v \) to operate as if it implemented pure depression dynamics. Of course, this is not consistent with our formulation in section 2. Nevertheless, the inconsistency arises ultimately because models such as ICA violate Dale’s law, allowing excitatory synapses to become inhibitory, and vice versa. We could in principle employ hard nonlinearities to pull out only the positive parts and apply \( E \) to them, leaving any negative parts unchanged, but this would make interpretation difficult and analysis likely impossible, and it would fundamentally alter the dynamics of the ICA single-unit rule. We must therefore simply reconcile ourselves to this inconsistency, knowing that it arises not from our general formulation but from such models’ violation of Dale’s law.

When \( \epsilon = 0 \), it is straightforward to show that equation 4.1 has \( n \) stable fixed points at \( \mathbf{v} = +\mathbf{m}_i, i = 1, \ldots, n \), and a further \( n \) stable fixed points at the sign-reversed locations \( \mathbf{v} = -\mathbf{m}_i, i = 1, \ldots, n \) (Hyvarinen et al., 2001). It will also be necessary for what follows to characterize all the unstable fixed points that must exist for \( \epsilon = 0 \). Consider first the completely symmetric case in which all the sources’ kurtoses are equal. Any pair of sources \( \mathbf{m}_i \) and \( \mathbf{m}_j, i \neq j \), define a great circle on the hypersphere \( |\mathbf{v}|^2 = 1 \). On such a great circle, there are four stable fixed points at \( \pm \mathbf{m}_i \) and \( \pm \mathbf{m}_j \). Between these four stable fixed points, there must exist a further four unstable fixed points located at \( \pm \mathbf{m}_i \pm \mathbf{m}_j / \sqrt{2} \) (for all four combinations of signs). On the great sphere defined by any triplet of sources \( \mathbf{m}_i, \mathbf{m}_j, \text{and} \mathbf{m}_k \), for \( i, j, \text{and} k \) distinct, there must exist eight further unstable fixed points of the form \( \pm \mathbf{m}_i \pm \mathbf{m}_j \pm \mathbf{m}_k / \sqrt{3} \) (for all eight combinations of signs). In general, for any distinct set of \( r \) sources \( \mathbf{m}^{i_1}, \ldots, \mathbf{m}^{i_r} \) defining an \((r - 1)\)-sphere, there must be \( 2^r \) unstable fixed points of the form

\[
\frac{1}{\sqrt{r}} \left( \pm \mathbf{m}^{i_1} \pm \mathbf{m}^{i_2} \pm \cdots \pm \mathbf{m}^{i_r} \right).
\]

For \( 1 < r < n \), these unstable fixed points are saddle nodes, but for \( r = n \), they are repeller nodes. The total number of fixed points, stable and
unstable, is therefore $\sum_{r=1}^{n} C_{r} 2^r = 3^n - 1$. This number is the same as the total number of boundaries (e.g., vertices, edges, faces) of an $n$-cube. For $n = 2$, for example, the two sources and their sign-reversed analogs can be viewed as sitting at the centers of the four edges of a square and the four unstable (repeller) nodes at the four vertices. For $n = 3$, the three sources and their sign-reversed analogs can be viewed as sitting at the centers of the six faces of a cube; the 12 saddles at the centers of the 12 edges of the cube; the 8 repellors at the 8 vertices. We therefore deform the $(n - 1)$-sphere into an $n$-cube, identifying fixed points with the centers of the bounding $m$-cubes, $m = 0, \ldots, n - 1$. For general non-zero kurtoses $k_1, \ldots, k_n$, there must also be $3^n - 1$ fixed points, but the unstable fixed points are displaced from being equidistant between sets of adjacent sources on the hypersphere. In this case, the perfectly symmetric $n$-cube for equal kurtoses becomes distorted and buckled.

We may now examine the fixed point structure of equation 4.1 for $\varepsilon > 0$ for those cases for which the complexity of the bifurcation structure of the initial $3^n - 1$ fixed points for $\varepsilon = 0$ can be handled. First, we examine the fixed point structure for maximal cross-talk, at $\varepsilon = 1$. Then we consider the case in which all the source vectors remain fixed points in the presence of any level of cross-talk, so that only pure stability changes can occur. Next, we study the case in which essentially exactly two sources are present, although the analysis is slightly more general. Finally, we consider the toy but informative case of a single non-gaussian source, with all other sources being Gaussian.

**4.1 Fixed Point Structure for $\varepsilon = 1$.** When $\varepsilon = 1$, the cross-talk matrix reduces to $E = \hat{n} \hat{n}^T$. We can then read off by inspection the fixed point solutions $v = \pm \hat{n}$ since manifestly $P_\hat{n} \hat{n} = 0$. We determine the stability of this pair of fixed points by writing $v = \hat{n} + y$ for a small perturbation $y$. We must have $\hat{n} \cdot y = 0$ to preserve the norm of $v$ at $O(|y|)$. Writing the single-unit ICA rule with a cubic nonlinearity in the form with an expectation value over the source vector $s$,

$$\dot{v} = \gamma P_v E \langle (v^T M s)^3 M s \rangle_s,$$

we obtain

$$\dot{y} = -\gamma y \langle (\hat{n}^T M s)^4 \rangle_s.$$  (4.2)

Hence, for $\gamma = +1$ or supergaussian sources, the PC of $E$ is a stable fixed point at $\varepsilon = 1$, while for $\gamma = -1$ or subgaussian sources, it is an unstable fixed point at $\varepsilon = 1$. Whether stable or unstable, this pair of fixed points implies the existence of at least a second pair of fixed points, of opposite stability, since the symmetric pair of fixed points at $\pm \hat{n}$ in isolation cannot induce consistent flows on the hypersphere $|v|^2 = 1$. 

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/NECO_a_00224 by guest on 16 May 2021
Unfortunately, the locations of these additional fixed points for $\varepsilon = 1$ cannot in general be determined by inspection. Because $E = \hat{n} \hat{n}^T$, the action of $E$ on the terms in square brackets in equation 4.1 leaves a vector proportional to $\hat{n}$. Writing

$$\Omega(v) = \sum_{i=1}^{n} k_i (m^i \cdot v)^3 m^i + 3v,$$

(4.3)

Equation 4.1 becomes, simply,

$$\dot{v} = \gamma [\hat{n} \cdot \Omega(v)] P \hat{n}.$$  (4.4)

The additional fixed points are therefore the unit norm solutions of $\hat{n} \cdot \Omega(v) = 0$. At such solutions, the vector $\Omega(v)$ must lie in the $(n-1)$-dimensional space orthogonal to the PC of $E$, which is the space spanned by the $n-1$ degenerate sub-PCs of $E$. However, because all the columns of $M$ appearing in $\Omega(v)$ span the whole space $\mathbb{R}^n$, these additional solutions in general do not lie purely within the span of the sub-PCs of $E$ (except when one of the columns of $M$ coincides with $\hat{n}$). The unit norm solutions of $\hat{n} \cdot \Omega(v) = 0$ form an $(n-2)$-dimensional space, so solutions that are purely within the span of the sub-PCs of $E$ form in general an $(n-3)$-dimensional subspace (except when one of the columns of $M$ coincides with $\hat{n}$, and then the two spaces exactly coincide).

Notice that equation 4.4 means that for $\varepsilon = 1$, the flow $\dot{v}$ is always in a direction parallel to $\hat{n}$ but projected onto the hypersphere $|v|^2 = 1$. Let $\theta$ be the angle between the two vectors $v$ and $\hat{n}$, and let the $n-2$ angles $\Phi = \{\phi_1, \ldots, \phi_{n-2}\}$ parametrize the hypersphere in the directions orthogonal to $\hat{n}$. We will refer to this hypersphere in the directions orthogonal to $\hat{n}$ by the symbol $S_{\hat{n}}^\perp$. Then $v = v(\theta; \Phi)$ and the flow for $\varepsilon = 1$ satisfies

$$\dot{\theta} = -\gamma [\hat{n} \cdot \Omega(v(\theta; \Phi))] \sin \theta,$$  (4.5)

with fixed points in $\theta$ at $\theta = 0$ (or $+\hat{n}$), $\theta = \pi$ (or $-\hat{n}$) and at $\theta$ satisfying $\hat{n} \cdot \Omega(v(\theta; \Phi)) = 0$. There is no flow in the $\Phi$ angular directions, that is, $\phi_i \equiv 0$, $i = 1, \ldots, n-2$. Hence, for any choice of these other angles $\Phi$, there must exist a $\theta$ fixed point solution of $\hat{n} \cdot \Omega(v(\theta; \Phi)) = 0$ with $\theta \in (0, \pi)$.

In general, then, we have the following geometric picture for the fixed points on the hypersphere $|v|^2 = 1$ when $\varepsilon = 1$. For any $n$, we have a pair of fixed points at $\pm \hat{n}$, stable for supergaussian sources and unstable for subgaussian sources. There is an $(n-2)$-dimensional submanifold of additional fixed points on this $(n-1)$-dimensional surface $|v|^2 = 1$. These additional fixed points must be unstable for supergaussian sources and stable for subgaussian sources. For $n = 2$, there are precisely two such additional fixed points (sign paired), in general not coinciding with the sign-paired vectors.
orthogonal to \( \hat{n} \), or \( \pm h^1 \). For \( n = 3 \), there is a continuous, closed curve of such solutions extending around the 2-sphere. This closed curve of fixed points in general intersects the great circle \( S^2_n \) in only two sign-paired, diametrically opposite locations, by the intermediate value theorem. For \( n = 4 \), the additional fixed points form a closed surface of solutions on a 3-sphere, intersecting with the great sphere \( S^2_n \) in a closed curve of solutions. Similar pictures hold for higher \( n \). Except when a column of \( M \) coincides precisely with \( \hat{n} \), the location of this \((n-2)\)-dimensional submanifold of additional stable or unstable fixed points must be computed explicitly by solving \( \hat{n} \cdot \Omega(v(\theta; \Phi)) = 0 \) for \( \theta \) given a specification of the other angles \( \Phi \).

In Figure 9 we show for \( n = 3 \) two examples of the flows \( \dot{v} \) and the two corresponding one-dimensional subspaces of fixed point solutions when \( \varepsilon = 1 \). We write

\[
v(\theta, \phi) = \hat{n} \cos \theta + (h^1 \cos \phi + h^2 \sin \phi) \sin \theta,
\]

where the two vectors \( h^1 = (1, 0, -1)^T / \sqrt{2} \) and \( h^2 = (-1, 2, -1)^T / \sqrt{6} \) define a choice of basis for the space orthogonal to \( \hat{n} \) and the set \( \{\hat{n}, h^1, h^2\} \) thus forms an orthonormal basis for \( \mathbb{R}^3 \) coinciding with the eigenvectors of \( E \). For the choices of sources and kurtoses described in the caption of Figure 9, to obtain the locations of the fixed points other than \( \pm \hat{n} \), we solve \( \hat{n} \cdot \Omega(v(\theta, \phi)) = 0 \) for \( \theta \), given the angle \( \phi \). We see, depending on the kurtoses of the sources, the flows moving either toward (supergaussian) or away from (subgaussian) \( \pm \hat{n} \), in directions parallel to \( \hat{n} \) but projected onto the 2-sphere. In each case, the continuous, closed curve of fixed points insects the great circle \( S^2_n \) in two opposite locations. For the subgaussian choice of kurtoses used, we see that the maximum deviation of the closed curve of fixed points away from \( S^2_n \) is greater than that for the supergaussian choice.

### 4.2 Pure Stability Changes of Sources Under Cross-Talk

Consider now the case when one of the sources \( m^i \), for some \( i \), coincides with the PC of \( E \). Without loss of generality, we may assume that \( m^1 = \hat{n} \). We may also rotate the degenerate sub-PCs of \( E \) into alignment with the \( m^i \), \( i > 1 \), performing any sign changes as necessary. Thus, when one of the columns of \( M \) coincides with \( \hat{n} \), we may always choose the eigenvectors of \( E \) so that the columns of \( M \) exactly coincide with them. In this case, it is clear that the source vectors \( \pm m^i \), \( i = 1, \ldots, n \), remain fixed points of equation 4.1 for any choice of \( \varepsilon > 0 \), although their stabilities may change as \( \varepsilon \) increases.

To examine the stability of these source fixed points as a function of \( \varepsilon \), we write \( E m^i = \lambda_i m^i \), where the eigenvalues are \( \lambda_1 = 1 \) and \( \lambda_i = 1 - \varepsilon \) for \( i > 1 \). Equation 4.1 becomes
Figure 9: The flow and fixed point locations for $\varepsilon = 1$ in single-unit cubic ICA for $n = 3$ sources. (A, B) Supergaussian sources with $k_1 = 1, k_2 = 1/2,$ and $k_3 = 3/2$ are used. (C, D) Subgaussian sources with $k_1 = -1, k_2 = -1/2,$ and $k_3 = -3/2$ are used. (A, C) The flows $\dot{v}$ on the 2-sphere. The thick line in each case shows the continuous, closed curve of fixed point locations other than at $\pm \hat{n}$; the large arrow indicates the direction of the vector $\hat{n}$. We see that the flows lie parallel or antiparallel to $\hat{n}$, but projected onto the 2-sphere. (B, D) $\cos \theta$ as a function of the angle $\phi$ for the two closed curves of fixed points in A and C. The angle $\theta$ is the angle of a vector with respect to $\hat{n}$, and $\phi$ is the rotation of this vector in the plane orthogonal to $\hat{n}$, defined in equation 4.6. Each curve intersects the great circle $S_{\hat{n}}^1$ defined by $\theta = \pi/2$ at precisely two, diametrically opposite locations. In all cases, we have taken the columns of the mixing matrix $M$ to be $m_1 = (1, 1, 0)^T / \sqrt{2}, m_2 = (1, -1, 0)^T / \sqrt{2}$ and $m_3 = (0, 0, 1)^T$.

Consider first the stability of $m_1 = \hat{n}$, or the PC direction of $E$. Writing $v = \hat{n} + y$ with $\hat{n} \cdot y = 0$ for a small perturbation $y$, we readily obtain

$$\dot{v} = \gamma P_v \left[ \sum_{i=1}^{n} k_i \lambda_i (m^i \cdot v)^3 m^i + 3 E v \right].$$

(4.7)
\[ \dot{y} = -\gamma (3 \varepsilon + k_1) y. \] (4.8)

For \( k_1 > 0 \) and thus \( \gamma = +1 \), the fixed point \( \hat{n} \) is stable for any value of \( \varepsilon \in [0, 1] \). However, for \( k_1 < 0 \) and thus \( \gamma = -1 \), there is some critical process at \( \varepsilon_c = -k_1 / 3 \in (0, \frac{2}{3}] \), with the fixed point \( \hat{n} \) being stable for \( \varepsilon < \varepsilon_c \) and unstable for \( \varepsilon > \varepsilon_c \). Now consider the stability of \( \mathbf{m}_i \), some \( i > 1 \). Linearly perturbing as usual, with \( \mathbf{m}_i \cdot y = 0 \), we obtain

\[ \dot{y} = \gamma \left[ - (1 - \varepsilon) k_i y + 3 \varepsilon (\hat{n} \cdot y) \hat{n} \right]. \] (4.9)

In directions orthogonal to \( \hat{n} \) and \( \mathbf{m}_i \), the dynamics are linearly stable for any \( \varepsilon \in [0, 1] \). However, in the \( \hat{n} \) direction, we have

\[ \hat{n} \cdot \dot{y} = \gamma \left[ - (1 - \varepsilon) k_i + 3 \varepsilon (\hat{n} \cdot y) \right] (\hat{n} \cdot y). \] (4.10)

For \( k_i < 0 \) and thus \( \gamma = -1 \), the fixed point \( \mathbf{m}_i \), \( i > 1 \), is stable for any value of \( \varepsilon \in [0, 1] \). However, for \( k_i > 0 \) and thus \( \gamma = +1 \), we have some critical process at \( \varepsilon_c^{(i)} = k_i / (3 + k_i) \in (0, 1) \). For \( \varepsilon < \varepsilon_c^{(i)} \), the fixed point \( \mathbf{m}_i \), \( i > 1 \), is stable, but for \( \varepsilon > \varepsilon_c^{(i)} \), it is unstable.

Summarizing these changes of stability of the sources, for the subgaussian case \( \gamma = -1 \), the PC of \( E \) source \( \mathbf{m}_1 = \hat{n} \) becomes unstable above the critical value \( \varepsilon_c = -k_1 / 3 \), but all the sub-PC sources, \( \mathbf{m}_i \) for \( i > 1 \), remain stable for all \( \varepsilon \). Thus, a single source is lost. However, for the supergaussian case \( \gamma = +1 \), we obtain a richer structure in which each sub-PC source \( \mathbf{m}_i \), \( i > 1 \), becomes unstable above its own critical value \( \varepsilon_c^{(i)} \); the PC source \( \mathbf{m}_1 = \hat{n} \) remains stable. In this supergaussian case, we therefore obtain in general a cascade of \( n - 1 \) changes of stability as each sub-PC source is lost.

Since the sources in this case are by assumption always fixed points for any \( \varepsilon \in [0, 1] \), the loss of the stability of any given source must represent a subcritical pitchfork bifurcation. To demonstrate this, we must identify a set of unstable nodes and track their locations as a function of \( \varepsilon \). Above, we saw that there are in general saddle nodes between all adjacent pairs of source fixed points. We therefore consider the great circle on the unit hypersphere that passes through any pair of distinct source vectors, say, \( \mathbf{m}_i \) and \( \mathbf{m}_j \), \( i \neq j \). Thus, we consider a strength vector of the restricted form

\[ \mathbf{v} = \mathbf{m}_i \cos \psi + \mathbf{m}_j \sin \psi, \] (4.11)

where the angle \( \psi \) parametrizes the great circle through these two sources. We then obtain the dynamics in \( \psi \) as

\[ \dot{\psi} = \frac{1}{2} \gamma \sin 2\psi \left[ \lambda_j (3 + k_j \sin^2 \psi) - \lambda_i (3 + k_i \cos^2 \psi) \right]. \] (4.12)

We have fixed points at \( \psi = 0, \pi/2, \pi, \) and \( 3\pi/2 \), corresponding to \( \pm \mathbf{m}_i \) and \( \pm \mathbf{m}_j \), and of course a linear stability analysis around these points must
reproduce the results above. We have further fixed points when the term in square brackets in equation 4.12 is zero. When $\varepsilon = 0$, the locations of these further fixed points are given by the distinct solutions of $\cos^2 \psi = k_j / (k_i + k_j)$. These represent the four solutions interdigitated between the four fixed points $\pm m^1$ and $\pm m^j$ discussed above. If $k_i = k_j$, then they are symmetrically located at $\psi = \pi / 4, 3\pi / 4, 5\pi / 4$, and $7\pi / 4$, exactly midway between the sources. If both sources $m^i$ and $m^j$ correspond to sub-PCs of $\mathbf{E}$, then an overall factor of $1 - \varepsilon$ pulls out of equation 4.12, so that the locations and stabilities of the saddle node fixed points are not affected by $\varepsilon < 1$, at least by dynamics occurring on this great circle. We therefore need to consider only the case where one of the sources corresponds to the PC of $\mathbf{E}$, so that, say, $i = 1$. At $\varepsilon = 1$, $\lambda_1 = 1$, and $\lambda_j = 0$ and the term in square brackets in equation 4.12 collapses to $-(3 + k_1 \cos^2 \psi) < 0$ for any valid choice of $k_1 \geq -2$, so these four saddle node solutions are absent. There is therefore a bifurcation in these saddle node solutions for some value of $\varepsilon$. The bifurcations in these solutions occur when

$$ (1 - \varepsilon)(3 + 0 k_j) = (3 + 1 k_1) \Rightarrow \varepsilon = -k_1 / 3, \quad (4.13) $$

$$ (1 - \varepsilon)(3 + 1 k_j) = (3 + 0 k_1) \Rightarrow \varepsilon = +k_j / (3 + k_j), \quad (4.14) $$

that is, at exactly the same critical values of $\varepsilon$ at which the relevant source fixed points change stability. These critical processes are therefore, as advertised, subcritical pitchfork bifurcations. For the subgaussian case, in which just the PC of $\mathbf{E}$ sources $m^1 = \pm \hat{n}$ turn unstable at $\varepsilon_c = -k_1 / 3$, we have a subcritical pitchfork bifurcation in which for $n > 2$, not two but (at least) $2(n - 1)$ saddle nodes simultaneously coalesce with the source fixed points. But for the supergaussian case, we have a cascade, in general, of $n - 1$ successive, distinct subcritical pitchfork bifurcations as each sub-PC source in turn becomes unstable.

In Figures 10 and 11 we illustrate these dynamics for $n = 3$ sources. For supergaussian sources (see Figure 10), we see the two pairs of sub-PC sources undergoing subcritical pitchfork bifurcations as four pairs of saddle nodes coalesce with the sub-PC source fixed points, changing the sub-PC source fixed points’ stabilities. Each sub-PC source turns unstable at a critical value of $\varepsilon$ specific to that source, as discussed. We also see that the two pairs of saddle nodes between the sub-PC sources undergo bifurcations as the eight repeller nodes collide on $S_{\hat{n}}^+$. These pitchfork bifurcations involving only the saddle and repeller nodes occur for smaller values of $\varepsilon$ than the subcritical pitchfork bifurcations involving the sub-PC sources. For these pitchfork bifurcations, however, all fixed points are unstable before and after the bifurcation, so no changes of stability occur. For subgaussian sources (see Figure 11), we see the pair of PC of $\mathbf{E}$ sources undergoing subcritical pitchfork bifurcations. In addition to the saddle nodes between the PC sources and the sub-PC sources colliding with the PC sources, we also
Figure 10: The dependence of fixed points on the cross-talk parameter $\varepsilon$ in single-unit cubic ICA for $n = 3$ supergaussian sources corresponding to a choice of eigenvectors of $E$. (A) Fixed points on a Mercator projection of the 2-sphere, with $\theta$ and $\phi$ being the standard angles in spherical polar coordinates. At $\varepsilon = 0$, the locations of the 6 fixed points corresponding to the sources are indicated by circles; the 12 fixed points corresponding to the saddle nodes are indicated by squares; the 8 fixed points corresponding to the repeller nodes are indicated by triangles. The paths traversed by the fixed points (here, only some of the unstable fixed points) as $\varepsilon$ increases are shown by solid lines for saddle nodes and long-dashed lines for repeller nodes. The thin dotted line shows the great circle $S_{\hat{n}}$. (B) A subregion of the Mercator projection showing a subset of the fixed points that undergo pitchfork bifurcations is shown explicitly against $\varepsilon$, in contrast to the suppression of the $\varepsilon$ coordinate in A. Line styles are as in A, but supplemented by a short-dashed line representing source fixed points. To generate this figure, we have selected sources corresponding to $m_1 = \hat{n}$, $m_2 = \frac{1}{6}(3 - \sqrt{3}, 2\sqrt{3}, -3 - \sqrt{3})^T$ and $m_3 = \frac{1}{6}(3 + \sqrt{3}, -2\sqrt{3}, -3 + \sqrt{3})^T$, with kurtoses $k_1 = 3/2$, $k_2 = 1$ and $k_3 = 1/2$.

see that the eight repeller nodes simultaneously collide with the PC sources as the saddle nodes collide with them, so that eight unstable nodes (four saddles and four repellers) in total collide with each PC source.

For completeness, we examine the bifurcations of these unstable nodes on the great spheres defined by triplets of sources, including the PC of $E$ source. For $n = 3$, as in Figures 10 and 11, these unstable nodes are repellers,
while for \( n > 3 \), they are saddles. With the two sub-PC sources being \( \mathbf{m}^i \) and \( \mathbf{m}^j, \ i \neq j \), we must consider a strength vector of the form

\[
\mathbf{v} = (\mathbf{m}^i \cos \psi + \mathbf{m}^j \sin \psi) \cos \chi + \mathbf{m}^1 \sin \chi.
\]  

(4.15)

We then derive

\[
\dot{\psi} = \frac{1}{2} \gamma (1 - \varepsilon) \cos^2 \chi \sin 2\psi (k_j \sin^2 \psi - k_i \cos^2 \psi),
\]  

(4.16)

\[
\dot{\chi} = \frac{1}{2} \gamma \sin 2\chi [3 \varepsilon + k_1 \sin^2 \chi - (1 - \varepsilon) (k_i \cos^4 \psi + k_j \cos^4 \psi) \cos^2 \chi].
\]  

(4.17)

We recognize the nontrivial solutions \( \cos^2 \psi = k_j / (k_i + k_j) \) of the \( \psi \) equation as the nontrivial solutions of equation 4.12 for two sub-PC sources, and we then find the nontrivial solutions for \( \chi \) as

\[
\cos^2 \chi = \frac{(3 \varepsilon + k_1)(k_i + k_j)}{k_1(k_i + k_j) + (1 - \varepsilon)k_i k_j},
\]  

(4.18)
when real solutions exist. The reality conditions give the locations of the bifurcations of these triplet solutions:

\[
\cos^2 \chi = 0 \Rightarrow \epsilon = -k_1/3, \quad (4.19)
\]

\[
\cos^2 \chi = 1 \Rightarrow \epsilon = k_i k_j / \left[ 3(k_i + k_j) + k_i k_j \right]. \quad (4.20)
\]

Hence, the first condition confirms that for subgaussian sources, these triplet solutions vanish as the PC source pair turns unstable. The second condition gives the bifurcations of these triplet solutions for supergaussian sources. We may verify that 

\[
k_i k_j / \left[ 3(k_i + k_j) + k_i k_j \right] < k_i / (3 + k_i) \text{ or } k_j / (3 + k_j),
\]

so these triplet solutions vanish before the corresponding sub-PC sources turn unstable, again confirming the observation in Figure 10.

For \( n = 3 \), the above is a complete characterization of all the bifurcation processes that occur as \( \epsilon \) increases, the triplet solutions being the eight repeller fixed points. In principle, we could extend this analysis to \( n > 3 \) sources and consider all the higher-order unstable fixed points and their bifurcations. For subgaussian sources, however, it is clear that since the PC source fixed points become unstable at \( \epsilon = \epsilon_c = -k_1/3 \) and all the sub-PC sources are always stable, only one set of bifurcations occurs, when all the \( 2(3^n - 1) \) unstable fixed points not on \( S_{\hat{n}}^\perp \) simultaneously collide with \( \pm \hat{n} \) at \( \epsilon = \epsilon_c \). All the fixed points on \( S_{\hat{n}}^\perp \), including the \( 2^{n-1} \) stable sub-PC source fixed points and the \( 3^{n-1} - 2^{n-1} - 1 \) unstable saddle node fixed points, do not change their stabilities for \( \epsilon < 1 \), but at \( \epsilon = 1 \), this entire hypersphere \( S_{\hat{n}}^\perp \) spanned by the sub-PCs of \( E \) becomes stable. For supergaussian sources, we have multiple cascades of pitchfork bifurcations. There are subcritical pitchfork bifurcations involving the sub-PC source fixed points and the saddle nodes between these sub-PC fixed points and the PC source fixed points at the critical values \( \epsilon = \epsilon_c^{(i)} \). There are also earlier pitchfork bifurcations in which higher-order unstable fixed points collide on \( S_{\hat{n}}^\perp \) and vanish. These are bifurcations involving purely unstable fixed points, without any change of stability. From the point of view of learning, however, the principal dynamics of interest to us are the bifurcations involving the sources’ fixed points, so bifurcations involving purely higher-order unstable fixed points need not be analyzed in detail.

### 4.3 Exact Analysis for \( n = 2 \) Sources.

The pitchfork bifurcations we saw in section 4.2 arose because the columns of \( M \) coincided with the eigenvectors of \( E \), so that the sources remained fixed points of equation 4.1 for any choice of the cross-talk parameter \( \epsilon \). The preservation of a fixed point through a bifurcation process reflects such a symmetry or invariance in an equation. By breaking the symmetry, pitchfork bifurcations break open into saddle node bifurcations, with the pair of solutions that collide and annihilate now consisting of the previously central prong of the pitchfork and one of its outer prongs. The other outer prong continues through the
Figure 12: The dependence of fixed points on the cross-talk parameter $\varepsilon$ in single-unit cubic ICA for $n = 3$ subgaussian sources rotated slightly away from a choice of eigenvectors of $E$. The subgaussian sources used in Figure 11 have been rotated anticlockwise by $\pi/180$ radians about the $x$-axis so that the sources are no longer perfectly aligned with the eigenvectors of $E$. The format of this figure is otherwise identical to Figure 11.

bifurcation, being deflected during the bifurcation to a degree that depends strongly on the extent of the symmetry breaking. We would therefore expect that in rotating the sources away from being exactly eigenvectors of $E$, all the pitchfork bifurcations we saw and analyzed above would turn into saddle node bifurcations. We verify this in Figure 12 for the subgaussian sources used in Figure 11, after having applied an anticlockwise rotation of just $\pi/180$ radians or 1 degree about the $x$-axis to the mixing matrix $M$ and therefore to the three source vectors $m^i$, $i = 1, 2,$ and 3. We clearly see that the four simultaneous pitchfork bifurcations that previously occurred around the paired PC of $E$ source fixed points, turning them unstable, now break open into four saddle node bifurcations. Whereas before the fixed point $\hat{n}$, say, survived the simultaneous pitchfork bifurcations, merely turning unstable but tracking up from $\varepsilon = 0$ to $\varepsilon = 1$, now it is annihilated by a saddle node fixed point, and the unstable fixed point that ends up at $\hat{n}$ for $\varepsilon = 1$ started out as one of repeller node fixed points at $\varepsilon = 0$.

For a general rotation of the sources, or equivalently a general orthogonal $M$, an analysis of the bifurcation structure of equation 4.1, would entail
solving $n - 1$ coupled nonlinear equations to determine the fixed points of the flow $\dot{v}$ on the unit hypersphere $|v|^2 = 1$ as a function of $\varepsilon$. For small rotations of the sources away from the eigenvectors of $E$, we could perform a perturbative analysis, determining the changes in the bifurcation points as a function of the small angles. Such a small-angle analysis would not, however, add significantly to the qualitative understanding gained in section 4.2. An expansion of the sources and the strength vector $v$ in the eigenvectors of $E$ is also an obvious strategy; our analysis represents the deployment of this approach for one particular “privileged” choice of sources exactly coinciding with the eigenvectors of $E$. In general, though, the components in such an eigenexpansion would interact nonlinearly through equation 4.1, emasculating the usual power of an eigenexpansion in a linear setting. Inevitably, in the presence of such nonlinearity and the complexity that can arise from the bifurcation structure of fully $3^n - 1$ initial, $\varepsilon = 0$ fixed points, not to mention the possibility of chaotic dynamics in the presence of $n \geq 4$ sources (chaotic dynamics are not possible for $n = 2$ and $n = 3$ sources by the Poincaré-Bendixson theorem), a general analysis is likely intractable.

We therefore now restrict to considering a rotation of the first source, $\mathbf{m}_1$, away from being exactly $\hat{n}$. Because the sources are by assumption orthogonal or have been made so by whitening, at least a second source must in the process acquire an $\hat{n}$ component. We therefore consider a rotation of, say, $\hat{n}$ and $\mathbf{h}_1$ to form the two sources $\mathbf{m}_1$ and $\mathbf{m}_2$, but all other sources remain in $S_{\hat{n}}^\bot$ for the purposes of analytical tractability. Hence, we write

$$\begin{align*}
\mathbf{m}_1 &= \hat{n}\cos \alpha - \mathbf{h}_1\sin \alpha, \\
\mathbf{m}_2 &= \hat{n}\sin \alpha + \mathbf{h}_1\cos \alpha, \\
\mathbf{m}_i &= \mathbf{h}_{i-1} \quad \text{for } i > 2,
\end{align*}$$

(4.21)

where $\alpha$ is the rotation applied in the $\hat{n}$–$\mathbf{h}_1$ plane to generate the sources $\mathbf{m}_1$ and $\mathbf{m}_2$ or, equivalently, the angle between the source $\mathbf{m}_1$ and the (positive) PC of $E$. We may view this choice of sources as an analysis of either exactly $n = 2$ sources in terms of an eigenexpansion involving the eigenvectors of $E$, or $n > 2$ sources for a very specific selection of sources.

Manifestly, the sources $\mathbf{m}_i$ for $i > 2$ remain fixed points for any $\varepsilon$, as in section 4.2. In particular, for the subgaussian case, these sources are always stable, while for the supergaussian case, they undergo subcritical pitchfork bifurcations at $\varepsilon = \varepsilon^{(i)} = k_i/(3 + k_i)$, $i = 3, \ldots, n$, becoming unstable.

At $\varepsilon = 0$, the source fixed points $\pm \mathbf{m}_1$ and $\pm \mathbf{m}_2$ must be stable, while at $\varepsilon = 1$, we already know that $\pm \hat{n}$ are either stable or unstable, depending on the sign of $\gamma$. Thus, we must determine how the fixed points $\pm \mathbf{m}_1$ and $\pm \mathbf{m}_2$ move as $\varepsilon$ changes. It suffices to restrict to the plane spanned by $\hat{n}$ and $\mathbf{h}_1$, so we write the strength vector in the form

$$v = \mathbf{m}_1\cos \psi + \mathbf{m}_2\sin \psi = \hat{n}\cos(\psi - \alpha) + \mathbf{h}_1\sin(\psi - \alpha).$$

(4.22)
We then find that the angle $\psi$ evolves according to

$$\dot{\psi} = \frac{1}{2} \gamma \sin 2\psi \left( k_2 \sin^2 \psi - k_1 \cos^2 \psi \right)$$

$$+ \gamma \varepsilon \left[ k_1 \cos^3 \psi \sin \alpha - k_2 \sin^3 \psi \cos \alpha - 3 \sin(\psi - \alpha) \right] \cos(\psi - \alpha).$$

(4.23)

The $\varepsilon = 0$ fixed points are, as expected, at $\sin 2\psi = 0$ (the four fixed points corresponding to the two sources) and at $\cos^2 \psi = k_2/(k_1 + k_2)$ (the four saddle node fixed points between the four source fixed points). In order to determine the fixed points for $\varepsilon > 0$, we write $z = \tan \psi$. In doing so, we essentially lose information about half the circle, but this is unimportant since this merely corresponds to a loss of information about the sign of $v$, and equation 4.1 is invariant under a change of signs. Writing

$$f(z; \alpha) = \frac{1}{2} \left\{ - (3 + k_1) \sin 2\alpha + z \left[ (6 + k_1) \cos 2\alpha - k_1 \right] ight.$$ 

$$+ z^3 \left[ (6 + k_2) \cos 2\alpha + k_2 \right] + z^4 (3 + k_2) \sin 2\alpha \right\},$$

(4.24)

we find that $z$ evolves according to

$$(1 + z^2) \dot{z} = \gamma \left[ z(k_2z^2 - k_1) - \varepsilon f(z; \alpha) \right].$$

(4.25)

The fixed points are solutions of either $\dot{z} = 0$ or $\dot{\zeta} = 0$, where $\zeta = z^{-1} = \cot \psi$. For $\varepsilon = 0$, the fixed points are at $z = 0$, $z = \pm \sqrt{k_1/k_2}$, and $\zeta = 0$. However, for $\varepsilon > 0$ and $\sin 2\alpha \neq 0$, the $z^0$ and $z^4$ terms in $f(z; \alpha)$ destroy the $z = 0$ and $\zeta = 0$ solutions. For $\varepsilon = 1$, we have a fixed point at $z = \tan \alpha$ or $\psi = \alpha$, corresponding to the known fixed point $v = +\hat{n}$ at $\varepsilon = 1$, but the three other fixed points must be determined as the roots of a cubic. For general $\varepsilon$, the four fixed points, when they are real, become the solutions $z = z(\varepsilon)$ of the quartic equation $z(k_2z^2 - k_1) = \varepsilon f(z; \alpha)$. While we can write down these solutions explicitly, they are an uninformative mess, and their bifurcation structure is completely opaque. It is much simpler to view $\varepsilon$ as the dependent variable rather than $z$, write $\varepsilon = \varepsilon(z)$, and immediately obtain

$$\varepsilon(z) = \frac{z(k_2z^2 - k_1)}{f(z; \alpha)},$$

(4.26)

for $z \in (-\infty, +\infty)$ supplemented by $\zeta = 0$. Of course, this just amounts to viewing the graph of the solutions $z(\varepsilon)$ as a function of $\varepsilon$ on its side, so that we see the single-valued solution $\varepsilon(z)$ as a function of $z$. Bifurcations in the solutions $z(\varepsilon)$ therefore occur at maxima or minima of $\varepsilon(z)$ as a function of $z$. 

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/NECO_a_00224 by guest on 16 May 2021
We illustrate these solutions for a particular choice of the two supergaussian kurtoses $k_1 = 3$ and $k_2 = 3/10$ in Figure 13, for different values of the rotation angle $\alpha$. We define an “approximate source” to be a stable fixed point for $\varepsilon \neq 0$ that is connected to a stable source fixed point for $\varepsilon = 0$ via a continuous path of stable fixed points as $\varepsilon$ decreases. Similarly, we define an “approximate PC of $E$” to be a stable fixed point for $\varepsilon \neq 0$ that is connected to a stable PC of $E$ fixed point for $\varepsilon = 1$ via a continuous path of stable fixed points as $\varepsilon$ increases. For $2\alpha/\pi = 0.5$ (see Figure 13A), we see just one bifurcation, in which the approximate source $m_2$ is lost, annihilated by a saddle node fixed point (a repeller node for $n = 2$). The other source, $m_1$, with the larger kurtosis, slowly transforms as $\varepsilon$ increases into the stable fixed point at $+\hat{n}$, corresponding to the solution $\psi = \pi/4$ since at $\varepsilon = 1$, the fixed point solution $v = \hat{n}$ maps onto $\psi = \alpha$. Increasing $\alpha$ further (see Figures 13B and 13C), the source $m_1$ continues to transform continuously into $+\hat{n}$, but as $\alpha$ approaches a critical value, the transformation becomes sharper and sharper, the curve acquiring a vertical component at the critical angle $2\alpha_c/\pi \approx 0.7369$ for this choice of $k_1$ and $k_2$. Above the critical angle (see Figures 13D and 13E), we have three bifurcations. The first bifurcation corresponds to the appearance of the approximate PC of $E$ fixed point. This bifurcation corresponds to a saddle node bifurcation in which a stable and an unstable fixed point appear at a critical value of $\varepsilon$. The third bifurcation corresponds to the loss of the approximate source $m_1$ as it is annihilated by the (in general) saddle node fixed point created in the second bifurcation. Thus, there is an intermediate regime, between the second and third bifurcations, in which the fixed point smoothly connected to the source $m_1$ at $\varepsilon = 0$ and to the PC of $E$ fixed point at $\varepsilon = 1$ are simultaneously stable. These two stable fixed points do not, however, smoothly transform into each other via a path of stable fixed points, as they do below the critical angle. At the third bifurcation point, a strength vector at the approximate source vector $m_1$ will suddenly jump to the approximate PC of $E$ vector as the approximate source fixed point vanishes in a saddle node bifurcation. For $\alpha$ above this critical angle, both approximate source fixed points are annihilated in two separate saddle node bifurcations. Increasing $\alpha$ still further, there is a second critical angle $2\alpha''_c/\pi \approx 0.8050$ for this choice of $k_1$ and $k_2$ at which we move from three bifurcations back to just one bifurcation (see Figure 13F). Above this second critical angle, the first approximate source $m_1$ is annihilated in a saddle node bifurcation, while the second source, with the smaller kurtosis, moves smoothly into the PC of $E$ stable fixed point at $\varepsilon = 1$.

We now show that this bifurcation structure is in fact completely generic for $k_1 \neq k_2$, for both supergaussian and subgaussian sources. In the case of subgaussian sources, the stable fixed points in the $\hat{n} - h_1$ plane at $\varepsilon = 1$ must be computed explicitly by solving $z(k_2z^2 - k_1) = f(z; \alpha)$, similarly to solving $\hat{n} \cdot \Omega(v) = 0$ for subgaussian sources in section 4.1. In analogy
Figure 13: The dependence of fixed points on the cross-talk parameter $\varepsilon$ in single-unit cubic ICA for supergaussian sources, two of which have been rotated by an angle $\alpha$ away from $\hat{n}$ and $h^1$ (in the $\hat{n}$--$h^1$ plane) so that these two sources no longer coincide with the eigenvectors of $\mathbb{E}$. Shown are all the fixed points associated with the two rotated sources, from equation 4.22. In all cases, the kurtoses of the two rotated sources, $m^1$ and $m^2$, are set to $k_1 = 3$ and $k_2 = 3/10$. As the rotation angle relative to the PC of $\mathbb{E}$ is increased, the bifurcation structure changes. (A) $2\alpha/\pi = 0.5$. (B) $2\alpha/\pi = 0.7$. (C) $2\alpha/\pi = 0.7369$. (D) $2\alpha/\pi = 0.8$. (E) $2\alpha/\pi = 0.804$. (F) $2\alpha/\pi = 0.81$. Stable fixed points are shown with solid lines, unstable fixed point with dashed lines.
with the approximate PC of $E$ fixed points, we will refer to the stable fixed point that ends up for $\varepsilon = 1$ at the relevant solution of $z (k_2 z^2 - k_1) = f(z; \alpha)$ as the approximate $\Omega$ fixed point. For subgaussian sources, as for the supergaussian sources in Figure 13, there is still a range of angles $\alpha$ in which the dynamics exhibit three saddle node bifurcations, so that there is an intermediate regime in which an approximate source fixed point and the approximate $\Omega_1$ fixed point are simultaneously stable but are not connected to each other by a continuous path of stable fixed points.

As noted in relation to equation 4.26, bifurcations occur at maxima or minima of $\varepsilon(z)$ as a function of $z$. Computing the derivative, we find

\[
\frac{d \varepsilon(z)}{dz} \propto -k_2 (3 + k_2) z^6 \sin 2\alpha + 3k_1 (3 + k_2) z^4 \sin 2\alpha \\
+ 4 \left[ 3(k_1 + k_2) + k_1 k_2 \right] z^3 \cos 2\alpha \\
- 3k_2 (3 + k_1) z^2 \sin 2\alpha + k_1 (3 + k_1) \sin 2\alpha.
\] (4.27)

The qualitative characteristics of the extrema of $d\varepsilon(z)/dz$ change when the number of real zeros of this sextic polynomial in $z$ changes. Thus, qualitative changes in the bifurcation structure occur at the zeros of the discriminant of this polynomial. Because the right-hand side of equation 4.25 is a quartic polynomial in $z$, there can never be more than two stable fixed points (four, counting the lost signs) in the $\hat{n}$–$h_1$ plane, and because $d\varepsilon(z)/dz$ is proportional to a sextic polynomial and its maxima and minima must be created or destroyed in pairs, there cannot be more than three bifurcations associated with the fixed points in this plane. When there are three bifurcations, the first one cannot correspond to the creation of the approximate PC of $E$ or $\Omega$ fixed point and its unstable partner, because we would then have three (six, counting the sign multiplicity) stable fixed points and three (six) unstable fixed points as solutions of a quartic, which is clearly impossible. In principle, it is possible that the final bifurcation could lead to the creation of the approximate PC of $E$ or $\Omega$ fixed point. In such a scenario, the two (four) stable approximate source fixed points would have been lost in the first two bifurcations, leaving a range of $\varepsilon$ between the second and third bifurcations in which no fixed points exist at all in the $\hat{n}$–$h_1$ plane. For exactly $n = 2$ sources, this would lead to periodic solutions in which the strength vector would rotate indefinitely around the unit circle. However, the structure of equation 4.25 also rules out this possibility. When $\sin 2\alpha \neq 0$, the $z^0$ and $z^4$ terms on the right-hand side ensure that $\dot{z} < 0$ at $z = 0$ and $\dot{z} > 0$ at $\zeta = 0$ for $\alpha \in (0, \pi/2)$, with the inequalities reversed for $\alpha \in (\pi/2, \pi)$. Thus, by the intermediate value theorem, for $\sin 2\alpha \neq 0$, there must always be at least two (four, with signs) fixed points. When $\sin 2\alpha = 0$ or when $\varepsilon = 0$, we explicitly have two fixed points at $z = 0$ and $\zeta = 0$. Hence, oscillatory solutions
cannot arise in this case. The only possibility when three bifurcations occur is that the approximate PC of $E$ or $\Omega$ fixed point is created at the second bifurcation, with one source being lost before and the other after this second bifurcation.

The discriminant will be a function of $\alpha$, $k_1$, and $k_2$, so we can regard its zeros as determining the critical angles as a function of the kurtoses, $\alpha(k_1, k_2)$. Defining

$$q_1(k_1, k_2) = (3 + k_1)(3 + k_2)(k_1^2 + k_2^2),$$  
$$(4.28)$$

$$q_2(k_1, k_2) = [(3 + k_1)k_2^2 + (3 + k_2)k_1^2]\sqrt{(3 + k_1)(3 + k_2)},$$  
$$(4.29)$$

$$q_3(k_1, k_2) = (3 + k_1)k_2^2 + (3 + k_2)k_1^3 + k_1k_2[3(k_1 + k_2) + k_1k_2],$$  
$$(4.30)$$

the discriminant vanishes when

$$\cos 2\alpha = \pm \frac{k_1 - k_2}{k_1 + k_2} \quad \text{or} \quad \cos 2\alpha = \pm \sqrt{\frac{q_1(k_1, k_2) \pm q_2(k_1, k_2)}{q_3(k_1, k_2)}},$$

for all four combinations of signs in the second set of solutions. Some of these critical values, however, do not correspond to bifurcations with $\epsilon \in [0, 1]$ or do not give real critical angles $\alpha$. The required critical angles are given by

$$\cos 2\alpha'(k_1, k_2) = \begin{cases} -\sqrt{\frac{q_1(k_1, k_2) - q_2(k_1, k_2)}{q_3(k_1, k_2)}} & \text{for } k_1 \geq k_2 \\ +\sqrt{\frac{q_1(k_1, k_2) - q_2(k_1, k_2)}{q_3(k_1, k_2)}} & \text{for } k_1 < k_2 \end{cases},$$  
$$(4.31)$$

$$\cos 2\alpha''(k_1, k_2) = \begin{cases} -\frac{1}{k_1 + k_2} & \text{for } k_1 \geq k_2 \\ +\frac{1}{k_1 + k_2} & \text{for } k_1 < k_2 \end{cases}. \quad (4.32)$$

For $k_1 \neq k_2$, we obtain an interval $(\alpha'_c, \alpha''_c) \subset [0, \pi/2]$ and a symmetric interval above $\pi/2$, $(\pi - \alpha'_c, \pi - \alpha''_c) \subset [\pi/2, \pi]$, in which three bifurcations occur for any sign of $\gamma$. Between the second and third bifurcations, one of the approximate source fixed points and either the approximate PC of $E$ $(\gamma = +1)$ or the approximate $\Omega$ $(\gamma = -1)$ fixed point are simultaneously stable. Outside these intervals, only one bifurcation occurs, with one of

\footnote{However, if we consider ICA without whitening, then we do find that for $n = 2$ sources, there are indeed parameter regimes in which oscillatory solutions arise. We do not pursue this here.}
Figure 14: The regions in the $k_2$–$\alpha$ plane in which triple bifurcations occur. (A) Supergaussian sources with $k_1 = +1$. (B) Subgaussian sources with $k_1 = -1/2$. The two solid lines in each case show the two critical angles $\alpha_c'(k_1, k_2)$ and $\alpha_c''(k_1, k_2)$, with the shaded areas between being the regions in parameter space in which an approximate source fixed point and either the approximate PC of $E$ or $\Omega$ fixed points are simultaneously stable in some interval of the cross-talk parameter $\varepsilon$. Outside these regions, only single bifurcations occur in which one of the approximate source fixed points is lost in a single saddle node bifurcation.

the approximate source fixed points being annihilated in a saddle node bifurcation and the other source fixed point moving continuously into the PC of $E$ fixed point or the $\Omega$ fixed point along a continuous path of stable fixed points. We show in Figure 14 these triple bifurcation regions in the $k_2$–$\alpha$ plane for two fixed choices of $k_1$, corresponding to supergaussian sources ($k_1 = +1$; see Figure 14A) and subgaussian sources ($k_1 = -1/2$; see
Figure 14B). The two critical angle curves intersect at \( k_1 = k_2 \), so changing the fixed value of \( k_1 \) in these figures essentially has the effect of translating the curves to the left or to the right, but with some distortion because the limits \( k_2 \to 0 \) and \( k_2 \to \infty \) are privileged.

The only region in parameter space in which triple bifurcations with the associated bistability of approximate source and PC of \( E \) or \( \Omega_1 \) fixed points do not occur for some range of \( \alpha \) corresponds to \( k_1 = k_2 \). When \( k_1 = k = k_2 \), \( \alpha'(k, k) = \alpha''(k, k) = \pi/4 \), so these critical angles coincide. For \( \alpha \neq \pm \pi/4 \) in this case, we obtain only one saddle node bifurcation. Which of the two approximate sources \( m_1 \) and \( m_2 \) is lost in a saddle node bifurcation then depends on \( \alpha \) relative to \( \pm \pi/4 \) and on the sign of \( \gamma \).

For \( \alpha = \pm \pi/4 \), however, equation 4.25 reduces to

\[
(1 + z^2) \dot{z} = \gamma(z^2 - 1) \left\{ kz \mp \frac{1}{2} \epsilon \left[ (3 + k)(z^2 + 1) \pm kz \right] \right\},
\]

so that the fixed points \( z = \pm 1 \), corresponding to \( \psi = \pm \pi/4 \) or \( v = \pm \hat{n} \) and \( v = \pm h^1 \), exist for any \( \epsilon \). In this exactly symmetrical case, therefore, the two stable approximate source fixed points \( m_1 \) and \( m_2 \) undergo a supercritical pitchfork bifurcation, with one of the pair of (now exact) fixed points \( \pm \hat{n} \) or \( \pm h^1 \) turning stable. Regardless of the value of \( \alpha \), we can explicitly compute the location of the bifurcation when \( k_1 = k = k_2 \). We find that

\[
\epsilon = \frac{2k}{k + \gamma \left\{ [(6 + k) \cos 2\alpha]^{2/3} + [(6 + 2k) \sin 2\alpha]^{2/3} \right\}^{3/2}}.
\]

For the particular values \( \alpha = \pm \pi/4 \), the supercritical pitchfork bifurcations therefore occur when \( \epsilon = 2k/(k + \gamma(6 + 2k)) \).

In Figure 15 we plot the critical values of the cross-talk parameter \( \epsilon \) at which bifurcations occur, \( \epsilon_c \), as a function of the angle \( \alpha \). In Figure 15A for \( k_1 > k_2 \) and \( \gamma = +1 \), we sketch the generic structure, with the single and triple bifurcation regions. In Figures 15B, 15C, and 15D, we show explicit examples for specific choices of \( k_1 \) and \( k_2 \) with \( \gamma = +1 \). In Figure 15B for \( k_1 > k_2 \), the choice \( k_2 = 0 \), although corresponding to a Gaussian source \( m_2 \), produces a badge-shaped region situated symmetrically in the interval \( \alpha \in (\alpha'_c, \pi - \alpha'_c) \) since for \( k_2 = 0 \), \( \alpha''_c = \pi/2 \). For \( k_2 \) non-zero, the bottom of the badge breaks open, the two free ends of the badge being pulled to \( \alpha = 0 \) and \( \alpha = \pi \). This breaking and deformation process decreases the sizes of the two intervals in which triple bifurcations occur and increases the critical value \( \epsilon_c \) at which the second source \( m_2 \) is lost in a saddle node bifurcation. The interval \( (\alpha''_c, \pi - \alpha''_c) \) corresponds to the broken bottom of the badge, and in this region, only one bifurcation occurs, in which the first source \( m_1 \) is lost. The critical value \( \epsilon_c \) for \( \alpha \in (0, \alpha'_c) \) is lower than that for \( \alpha \in (\alpha''_c, \pi/2) \) because \( k_2 < k_1 \), so a lower value of the cross-talk parameter is required to break the weaker of the two sources, \( m^2 \). For \( k_1 = k_2 \), in Figure 15C, the two
Figure 15: Critical values of the cross-talk parameter, \( \varepsilon_c \), at which bifurcations occur as a function of the angle \( \alpha \) defining the sources \( m^1 \) and \( m^2 \) in terms of \( \hat{n} \) and \( h^1 \). (A) The generic diagram for \( k_1 > k_2 \) and \( \gamma = +1 \). For \( \alpha \in (\alpha'_c, \alpha''_c) \) and \( \alpha \in (\pi - \alpha''_c, \pi - \alpha'_c) \), three bifurcations occur at the three distinct values of \( \varepsilon_c \) corresponding to each value of \( \alpha \). For \( \alpha \in [0, \alpha'_c) \) and \( \alpha \in (\pi - \alpha'_c, \pi] \), a single bifurcation occurs, with the approximate source \( m^2 \) with \( k_2 < k_1 \) being lost at the corresponding value of \( \varepsilon_c \). For \( \alpha \in (\alpha''_c, \pi - \alpha''_c) \), only a single bifurcation still occurs, but in this range, the approximate source \( m^1 \) with \( k_1 > k_2 \) is lost. (B) The diagram for the specific choices of kurtoses \( k_1 = 3 \) and \( k_2 = 3/100 \) (solid line) and \( k_1 = 3 \) and \( k_2 = 3/10 \) (dashed line). Also shown for reference is the case in which \( k_1 = 3 \) and \( k_2 = 0 \) (thin dotted line). (C) The diagram for the specific choice \( k_1 = 3 \) and \( k_2 = 3 \). Here there is no triple bifurcation range in \( \alpha \). (D) The diagram for the specific choices of kurtoses \( k_1 = 3/100 \) and \( k_2 = 3 \) (solid line) and \( k_1 = 3/10 \) and \( k_2 = 3 \) (dashed line). Also shown for reference is the case in which \( k_1 = 0 \) and \( k_2 = 3 \) (thin dotted line). Compared to A, the sources have effectively been relabeled, introducing a shift of \( \pi/2 \) into the angle \( \alpha \).
sources are identically distributed, and here there is no triple bifurcation region. Instead, which of the two identical sources is lost depends on the angle $\alpha$ in relation to the symmetric points $\alpha = \pm \pi/4$ (or $\alpha = \pi/4$ and $\alpha = 3\pi/4$) and the sign of $\gamma$. Even for $k_1 \approx k_2$, the triple bifurcation region is so small as to be irrelevant in practice. The case $k_1 < k_2$, Figure 15D, is identical to the case $k_1 > k_2$, Figure 15B, except for a relabeling of the two sources $m^1$ and $m^2$ and thus a translation of $\alpha$ by $+\pi/2$ or $-\pi/2$. For $\gamma = -1$, we obtain identical diagrams, except that, say, $k_2 < k_1$ now means that the source $m^2$ has the more significant excess kurtosis (a flatter distribution), so we obtain a diagram like that in Figure 15D rather than that in Figure 15B.

The critical angle $\alpha_c'$ corresponds to a so-called cusp catastrophe (Strogatz, 1994). We denote the values of $\epsilon$ at the three bifurcation points in the triple bifurcation region by $\epsilon_c', \epsilon_c''$, and $\epsilon_c'''$. For $\alpha < \alpha_c'$, the surviving approximate source fixed point for $\epsilon > \epsilon_c'$ smoothly and continuously moves as $\epsilon$ increases into the PC of $E$ fixed point at $\epsilon = 1$, considering $\gamma = +1$ for concreteness. However, for $\alpha \in (\alpha_c', \alpha_c'')$, this continuous process does not occur. Increasing $\epsilon$ above $\epsilon_c'$, the surviving approximate source fixed point changes smoothly until we reach $\epsilon = \epsilon_c''$. Pushing $\epsilon$ slightly above $\epsilon_c''$, this approximate source fixed point undergoes a catastrophic change, disappearing entirely. Any strength vector near this fixed point will then jump to the approximate PC of $E$ fixed point that appeared at $\epsilon = \epsilon''$. Conversely, decreasing $\epsilon$ from unity, the approximate PC of $E$ fixed point tracks down to $\epsilon = \epsilon''$. Reducing $\epsilon$ slightly, any strength vector near the approximate PC of $E$ fixed point will then jump to the approximate source fixed point that exists for $\epsilon > \epsilon_c'$. These dynamics are akin to phase transitions with the critical angle $\alpha_c'$ playing the role of a critical point in a physical system, corresponding to the beginning or end of a phase boundary. For $\alpha < \alpha_c'$, the approximate source for $\epsilon > \epsilon_c'$ and the approximate PC of $E$ can be moved smoothly and continuously into each other by varying $\epsilon$. There is thus no real distinction between these two fixed points in this regime. However, for $\alpha > \alpha_c'$, the two fixed points are quite distinct, with no continuous process existing that can transform one into the other, and we have, in effect, two different, distinct phases. At $\alpha = \alpha_c'$, the phase boundary between these two phases starts or ends. Notice that the second critical angle $\alpha_c''$ is rather different and does not correspond to a cusp catastrophe. The cusp catastrophe at $\alpha = \alpha_c'$ arises at the point at which a stable part of the solution $\psi$ develops a vertical component, while the critical process at $\alpha = \alpha_c''$ arises at the point at which $\epsilon_c' = \epsilon_c''$, where two parts of the solution touch and “pinch off.” The presence of a cusp at $\alpha = \alpha_c'$ and its absence at $\alpha = \alpha_c''$ is clear in the diagrams in Figure 15. In Figure 16, we show the solutions for $\psi$ as a function of $\epsilon$ and $\alpha$, with $k_1$ set to unity and $k_2$ set to zero for clarity so that the first bifurcation at $\epsilon = \epsilon_c'$ corresponding to the loss of the approximate source $m^2$ is absent. The cusp catastrophe is indicated with an arrow and corresponds to the critical angle $\alpha_c'$ at which the solution $\psi$ starts to fold back on itself, as in Figure 13C.
Figure 16: Solutions $\psi$ for the strength vector in the $\mathbf{n} - \mathbf{h}$ plane plotted against the cross-talk parameter $\varepsilon$ and the rotation angle $\alpha$ defining the two sources $\mathbf{m}_1$ and $\mathbf{m}_2$. The kurtosis $k_1$ has been set to unity, and for clarity, the kurtosis $k_2$ has been set to zero so that, in effect, the second source undergoes a bifurcation at $\varepsilon = 0$, and thus it does not appear on the plot. Solutions in the negative region $\psi \in (-\pi/2, 0)$ are not shown, again for clarity. The head of the arrow indicates the location of the cusp catastrophe or critical angle $\alpha'_c$ (with $k_2 = 0$ the second critical angle is $\alpha''_c = \pi/2$), or the critical point corresponding to the termination of the phase boundary between the approximate source $\mathbf{m}_1$ and the approximate PC of $\mathbb{E}$.

4.4 Single Nongaussian Source. The analysis of $n$ sources, with just two of them rotated in the $\mathbf{n} - \mathbf{h}$ plane, leads to a very rich bifurcation structure, and in general to regimes of distinct fixed point phases. Extending the analysis to a fully general orthogonal $\mathbb{M}$ would presumably lead to even richer and very much more complicated dynamics. The possibility of limit cycles for $n \geq 3$ (white) sources and chaotic dynamics for $n \geq 4$ sources is immensely enticing but would lead to a largely numerical jungle. We defer exploring this jungle and instead take the reverse step. We strip ICA right down to a toy version of just a single nongaussian source in an $n$-dimensional space. In doing so, we will quickly see that the analysis performed in section 4.3 remains valid with only minimal changes, but without the distraction of multiple sources.

Let the single nongaussian source be $\mathbf{m}$ with kurtosis $k$. We can project out any $\mathbf{n}$ component of this source. Writing $\cos \alpha = \mathbf{n} \cdot \mathbf{m}$, we define the unit norm vector $\mathbf{h}$ via $\mathbb{P}_\mathbf{n} \mathbf{m} = (I - \mathbf{n} \mathbf{n}^T) \mathbf{m} \equiv -\mathbf{h} \sin \alpha$, so that we may write $\mathbf{m} = \mathbf{n} \cos \alpha - \mathbf{h} \sin \alpha$, with $\mathbf{n} \cdot \mathbf{h} = 0$. Because $\mathbf{h}$ is orthogonal to $\mathbf{n}$, we have $\mathbb{E} \mathbf{h} = (1 - \varepsilon)\mathbf{h}$. We know that at $\varepsilon = 0$, the vectors $\pm \mathbf{m}$ are stable fixed points and that at $\varepsilon = 1$, the vectors $\pm \mathbf{n}$ are fixed points, either stable or unstable depending on the sign of $\gamma$. The PC of $\mathbb{E}$ and the source vector $\mathbf{m}$ therefore pick out two privileged directions on the hypersphere $|v|^2 = 1$, and hence we look for fixed points on the great circle through $\mathbf{n}$ and $\mathbf{m} \neq \pm \mathbf{n}$;
if $m = \pm \hat{n}$, the analysis in section 4.2 immediately applies. We parametrize this great circle by defining the vector $m^\perp = \hat{n} \sin \alpha + h \cos \alpha$ with $m \cdot m^\perp = 0$ and then writing

$$v = m \cos \psi + m^\perp \sin \psi = \hat{n} \cos(\psi - \alpha) + h \sin(\psi - \alpha), \quad (4.35)$$

similarly to equation 4.22. Of course, this definition of $m^\perp$ does not imply that $m^\perp$ corresponds to one of the $n - 1$ sources whose kurtoses have been set to zero, but just provides a convenient parametrization of the great circle through $\hat{n}$ and $m$.

The angle $\psi$ evolves according to

$$\dot{\psi} = -\frac{1}{2} \gamma k \sin 2\psi \cos^2 \psi + \gamma \varepsilon [k \cos^3 \psi \sin \alpha - 3 \sin(\psi - \alpha)] \cos(\psi - \alpha). \quad (4.36)$$

This equation is by construction nothing but equation 4.23 with $k_1 = k$ and $k_2 = 0$, so the above analysis directly carries over to this toy case of a single nongaussian source. We therefore obtain an essentially identical bifurcation structure except that any bifurcations associated with the sources whose kurtoses have been set to zero are absent. There is a single critical angle, $\alpha_c = \alpha'_c(k, 0)$, given by

$$\cos 2\alpha_c = -\gamma \sqrt{\frac{3 + k - \sqrt{3} \sqrt{3 + k}}{k}}. \quad (4.37)$$

For $\alpha \in [0, \alpha_c)$ or $\alpha \in (\pi - \alpha_c, \pi]$, no bifurcations occur, and the source $m$ moves smoothly into the PC of $E$ or $\Omega$ fixed point. For $\alpha \in (\alpha_c, \pi - \alpha_c)$, there are two bifurcations. This reduction in the number of bifurcations from three to two is confirmed by equation 4.27, the right-hand side of which reduces from a sextic to a quartic polynomial when $k_1 = k$ and $k_2 = 0$. The first is the creation of the approximate PC of $E$ or $\Omega$ fixed point, the second the annihilation of the approximate source $m$ fixed point, with a bistable regime between these two bifurcations. For $\varepsilon \in [0, 1)$, the fixed point solutions given by equation 4.36 characterize all possible stable fixed points in the presence of a single nongaussian source since the stable fixed point dynamics are restricted to the subspace spanned by the source vector $m$ and the PC of $E$, $\hat{n}$. At $\varepsilon = 1$, we additionally obtain the manifold of fixed points given by the solutions of $\hat{n} \cdot \Omega(v) = 0$, which are stable for a single subgaussian source.

---

3This observation explains the absence of the $z^1$ and $z^5$ terms from the right-hand side of equation 4.27.
5 Discussion

Viewing nonlinear models of synaptic plasticity as a subset of nonlinear dynamical systems affords a much deeper understanding of the richness and diversity of behaviors that can be exhibited by them, but also of the concomitant dangers of the sudden onset of instabilities, bifurcations, and even chaotic dynamics driven by small parameter variations, perturbations, noise, and other influences. This broader view is common and longstanding in many areas of biology (see, for example, May, 1981; Murray, 1989), including neuroscience, in which the nonlinear dynamical systems’ approach to action potential generation in particular and collective neuronal firing dynamics in general is widely adopted (Ermentrout & Terman, 2010). By comparison, many models of synaptic plasticity remain rooted in assumptions of linearity or based on optimization principles that guarantee the existence of gradient systems. Yet integrable systems are the exception, not the rule. And even if evolution singles out integrable systems as furnishing the underlying principles of neuronal development, learning, and memory, instantiation of such principles in a biological substrate will inevitably introduce nonintegrability and fluctuations. The nonlinear dynamical systems’ view of biologically relevant models of synaptic plasticity therefore appears inevitable and inescapable.

In building biologically relevant models of synaptic plasticity, we must take seriously the constraints imposed by biology. One of those constraints must be the fidelity with which changes in synaptic strength can be implemented. Experimental evidence indicates that changes in strength are not confined to the pre- or postsynaptic aspects of synapses, but can spill over into neighboring synapses, through the intra- or extracellular diffusion of plasticity-related molecules. In consequence, this cross-talk between synapses sets fundamental limits on the accuracy with which synapses can change their strengths and provides an irreducible source of error or noise in synaptic updates. We have shown that cross-talk therefore also provides an irreducible source of bifurcations in nonlinear models of synaptic plasticity.

5.1 Bifurcations in Nonlinear Models of Synaptic Plasticity Under Cross-Talk. We first considered developmental models. In the MMM, we found that in general (i.e., for \( n > 2 \) afferents) the unsegregated fixed point undergoes a transcritical bifurcation at \( \varepsilon = \varepsilon_c^{(1)} \), while the (approximate) segregated fixed points all undergo saddle node bifurcations at \( \varepsilon = \varepsilon_c^{(2)} > \varepsilon_c^{(1)} \). Thus, there is a multistable regime in which the (approximate) segregated fixed points and the unsegregated fixed point are simultaneously stable. As \( \rho \to 1 \) or \( n \to \infty \), \( \varepsilon_c^{(1)} \to 0 \), while \( \varepsilon_c^{(2)} \to 0 \) only for \( \rho \to 1 \). The level of cross-talk required to stabilize the unsegregated fixed point can therefore be arbitrarily small. We then considered a class of developmental models including our own stochastic model of STDP. At the level of approximation
considered, we found coordinated pure stability changes at $\varepsilon = \varepsilon_c$ of the unsegregated fixed point and all the segregated fixed points. Examining this critical process in more detail, we found for our model of STDP that in general, $n$ simultaneous saddle node bifurcations occur at $\varepsilon = \varepsilon_c^{(0)}$. Half the saddle nodes pass straight through the unsegregated fixed point, turning it stable in a transcritical bifurcation at $\varepsilon = \varepsilon_c^{(1)} > \varepsilon_c^{(0)}$. The other half pass straight through the $n$ segregated fixed points at $\varepsilon = \varepsilon_c^{(2)} > \varepsilon_c^{(1)}$, rendering them unstable in a pseudocritical process. At the level of approximation at which we appear to observe pure stability changes, we have that $\varepsilon_c \approx \varepsilon_c^{(0)} \approx \varepsilon_c^{(1)} \approx \varepsilon_c^{(2)}$. Like the MMM, there is a multistable regime in which the unsegregated and segregated fixed points are simultaneously stable. Unlike the MMM, $\varepsilon_c$ cannot be made arbitrarily close to zero by taking $\rho \to 1$ or $n \to \infty$. However, explicit calculations showed that $\varepsilon_c$ can be very small. In broad outline, therefore, these two developmental models exhibit overall rather similar cross-talk-induced dynamics. In particular, the bifurcation processes are nearly exact mirror images of each other, with $n$ pairs of fixed points annihilating in the MMM, but with $n$ pairs being created in the stochastic STDP model. The reason for this difference is that in the MMM, the segregated fixed points drift as a function of $\varepsilon$, whereas in the stochastic model of STDP, the segregated fixed points remain perfectly segregated.

We then turned to ICA as an example of a nonlinear learning model with a potentially much richer fixed point structure than the somewhat trivial fixed point structure required of developmental models. The richer fixed point structure means that a general analysis for $n \geq 3$ sources is likely intractable. For choices of the sources corresponding to the eigenvectors of the cross-talk matrix, so that the sources are always fixed points of the dynamics regardless of the level of cross-talk, we found for supergaussian sources in general a cascade of $n - 1$ successive subcritical pitchfork bifurcations involving the sources in the sub-PC of $\mathbb{E}$ directions. For subgaussian sources, just the PC of $\mathbb{E}$ direction destabilizes in a subcritical pitchfork bifurcation involving all the saddle node fixed points away from $S_n^\perp$. The subcritical pitchfork bifurcations are not generic, however. In rotating the sources away from alignment with the eigenvectors of $\mathbb{E}$, pitchfork bifurcations break open into saddle node bifurcations. For an arbitrary rotation, analysis for $n \geq 3$ again becomes difficult. Restricting to a rotation of just two sources with PC of $\mathbb{E}$ components, which also constitutes a general analysis of the $n = 2$ source case, we found that the dynamics in the plane spanned by the two rotated sources can exhibit either one or three bifurcations, depending on the angle between the sources and the PC of $\mathbb{E}$. Outside the critical angular interval, one source is lost in a saddle node bifurcation. The other source then smoothly transforms into the PC of $\mathbb{E}$ in the supergaussian case, or the $\Omega_1$ fixed point in the subgaussian case. Which of the two sources is lost is determined by whether the angle is below or above the critical angular interval. Inside the critical angular interval, three bifurcations occur. First,
one source is lost (always the one with the smaller kurtosis). Then the approximate PC of $E$ or $\Omega$ fixed point is created in a second bifurcation. Finally, the second source is lost in the third bifurcation. Inside this critical angular interval, there is thus always a multistable regime in which a source and the approximate PC of $E$ or $\Omega$ fixed points are simultaneously stable for some range of values of $\varepsilon$.

5.2 Comparison Between Developmental and Learning Dynamics.

There are two principal differences between the results for the developmental models and the ICA model. The first is that in the ICA model, we can observe cascades of bifurcations in which sources are lost at different critical values of $\varepsilon$. In the developmental models, however, all (perhaps approximately) segregated fixed points undergo coordinated, simultaneous bifurcations. This difference reflects the fact that the afferents’ activity patterns are typically assumed to be unbiased in developmental models. Since developmental models should be sensitive to slow processes occurring over hours or days (Purves & Lichtman, 1985), temporal and spatial averaging, coupled with isotropy and homogeneity, should ensure sufficiently unbiased distributions for all practical purposes. In the ICA model, we have explicitly allowed different kurtoses, and thus although the sources have been assumed to be white (or made so) and therefore have identical first- and second-order cumulants, their higher-order cumulants differ. Even when we set the kurtoses equal in the analysis of two rotated sources, one or another source is lost, depending on the rotation angle relative to the PC of $E$. Only when the rotation angle is $\pm \pi/4$, so that both sources are equidistant from the PC of $E$, do we obtain a supercritical pitchfork bifurcation in which both sources are lost simultaneously. In this case, the PC of $E$ is always a fixed point, and it becomes stable during the supercritical pitchfork bifurcation for supergaussian sources.

The second difference, relating to the first, is the role played by the PC of $E$. In the developmental models, the unsegregated state is always a fixed point for any value of $\varepsilon$. That the unsegregated state is a fixed point at $\varepsilon = 0$ is a consequence of treating the afferents in a completely symmetric manner. That it is a fixed point for any value of $\varepsilon$ is a consequence of the rather simple but highly convenient form of $E$ used in equation 2.4. Of course, the unsegregated state, proportional to $n$, is nothing but the (unnormalized) PC of $E$. Thus, in the developmental models, the PC of $E$ is always a fixed point of the dynamics by a combination of symmetry and accident. However, in the ICA model, the PC of $E$ can become a fixed point of the dynamics only in the trivial limit $\varepsilon \to 1$ unless there is a symmetry that enforces this fixed point, although we have seen that the approximate PC of $E$ fixed point can be created in a bifurcation for $\varepsilon < 1$ and can be extremely close to the exact PC of $E$ even for $\varepsilon$ very different from unity. Thus, in the developmental models, the PC of $E$ is always a fixed point,
while in the ICA model, the PC of $\mathbf{E}$ as a fixed point must typically be induced, continuously or discontinuously, by the dynamics.

The differences are therefore not fundamental. In particular, both sets of models would exhibit similar bifurcation dynamics in different, limiting regimes. We can obtain in the ICA model the coordinated bifurcations seen in the developmental models by setting all the kurtoses equal and rotating the sources so that they are all equidistant from the PC of $\mathbf{E}$. However, to observe cascades of bifurcations in developmental models, it would be necessary to violate the symmetrical treatment of afferents. The purest way of doing this, other than introducing disparities into their mean activities (reproducing abnormal developmental regimes), would be to allow their second-order moments to differ, so that the correlation matrix $\mathbf{C}$ no longer takes the simple form in equation 3.4. We would then expect that the degeneracy of the $n$ saddle node bifurcations that we saw in these models would be split and that we would instead observe in general a cascade of $n$ bifurcations as each segregated fixed point is either annihilated (in the MMM) or turned unstable (in the stochastic STDP model).

5.3 Comparison to Results of Cox and Adams (2009). Our analysis of the ICA model largely corroborates the qualitative findings of Cox and Adams (2009) in their numerical study of ICA in the presence of cross-talk. They focused mainly on the Bell-Sejnowski model of ICA (Bell & Sejnowski, 1995), which acquires all $n$ sources simultaneously by implementing antiredundancy dynamics across $n$ postsynaptic neurons. However, they also briefly examined single-unit ICA with a hyperbolic tangent nonlinearity (Hyvarinen & Oja, 1998). Single-unit ICA provides a dynamically much simpler model, without contamination by antiredundancy terms. It also facilitates a greater level of mathematical analysis with the choice of cubic nonlinearity, as here.

Cox and Adams (2009) argue, implicitly in places, that the presence of (activity) cumulants greater than second order is important in driving the cross-talk-induced breakdown of ICA. As evidence in favor of this view, they cite their study of a neuronal implementation of principal component analysis, which is sensitive only to second-order activity cumulants (Râdulescu, Cox, & Adams, 2009). In that study, they found a continuous, gradual breakdown in the acquisition of principal components, rather than the discontinuous, sudden collapse in the acquisition of independent components that can occur in ICA.

It is critical, however, to distinguish carefully between nonlinearities in strengths in a synaptic plasticity rule and nonlinearities in activities in a synaptic plasticity rule. The bifurcations in the fixed points of stable strength vectors studied here and by Cox and Adams (2009) can arise only in the presence of nonlinearities in strengths in synaptic plasticity rules, essentially by definition. In particular, a synaptic plasticity rule linear in strengths but nonlinear in activities can never exhibit any bifurcations in strength dynamics;
conversely, a synaptic plasticity rule nonlinear in strengths but linear in activities can exhibit bifurcations in strength dynamics. Of course, simple neuronal models often employ a linear model of postsynaptic activity (or, rather, “activation”) of the form $v \cdot a$, so any nonlinearity in strengths introduced by a nonlinear dependence on $v \cdot a$ necessarily introduces nonlinearities in activities. Nevertheless, the very possibility of the presence of bifurcations in the fixed points of stable strength vectors in nonlinear models of synaptic plasticity arises because of, and only because of, strength nonlinearities. Activity nonlinearities may indeed modify these bifurcations, but the bifurcations could not exist without the underlying strength nonlinearities. Hebbian models, whether linear or nonlinear in strengths, are at least quadratic in activities, since they postulate a dependence on both pre- and postsynaptic activity (Hebb, 1949), giving rise to the usual dependence on at least the correlation matrix. The MMM studied above is, however, nonlinear in strengths but sensitive only to first- and second-order moments. This model exhibits bifurcations in its stable fixed points and therefore provides an explicit counterexample to the suggestion of Cox and Adams (2009) that higher-order cumulants are important in driving bifurcations in nonlinear models of synaptic plasticity.

5.4 Generality of Results: The Dangers of Even Low Levels of Cross-Talk. A completely general analysis of nonlinear models of synaptic plasticity under the influence of cross-talk or noise and errors more generally is almost certainly impossible. The pattern of bifurcations induced by the variation of control parameters in nonlinear models of any type, biological or otherwise, is typically highly idiosyncratic, although universality may provide a wider perspective in certain limits. It is therefore inevitable that we must be content with a largely case-by-case analysis of representative examples of models, and preferably models that are sufficiently simple to admit a fair degree of analysis but not so simple that they are biologically irrelevant.

By considering two developmental models and a learning model, we have sought to provide a flavor of the rich and varied bifurcation dynamics that can be induced by the presence of cross-talk in nonlinear models of synaptic plasticity. The developmental dynamics are simpler, reflecting the choice of the unbiased correlation matrix typical in developmental studies, but the ICA dynamics expand this simplicity into enormous complexity, even for just two sources. Extending to three or four afferents or sources (depending on whether normalization dynamics reduce the effective space by one dimension) for general input statistics, in developmental or learning models, would lead to wild, untamable complexity, with limit cycles, strange attractors, and chaotic dynamics.

Despite these caveats, it is clear from the construction of the cross-talk matrix in equation 2.4 that in the trivial limit $\epsilon \to 1$, cross-talk will completely destroy all developmental and learning dynamics because synaptic
updates cease to have any Hebbian specificity at all. Such a limit, however, is completely irrelevant from a biological point of view. The central and paramount question, then, is how this limit is in fact approached in models of synaptic plasticity. The purpose of our detailed analysis has essentially been to attempt to explore this question in the context of specific, concrete, and well-motivated models operating in different spheres of synaptic plasticity, either development or learning.

In all three models, we found the existence of multistable regimes in which the desired stable fixed points (either developmental or learning related) simultaneously coexist with a stable fixed point that is very close to, or identical to, the PC of $E$. This multistable regime can exist over a broad range of the cross-talk parameter $\varepsilon$. Furthermore, in the MMM and under ICA, the lower range of this interval, corresponding to the creation or stabilization of this PC of $E$-related fixed point at $\varepsilon = \varepsilon_c^{(1)}$, can be made arbitrarily small, depending on the input statistics and the number of inputs. Even in the STDP model, $\varepsilon_c^{(1)}$ can be very small. Thus, although the desired fixed points may not necessarily have been lost to bifurcations, when this further stable fixed point appears, the dynamics are no longer guaranteed to find the desired fixed points, and the outcome of learning and development then becomes vulnerable to a dependence on initial conditions. Moreover, the desired stable fixed points themselves can be lost at critical levels of cross-talk that can be very small, depending on the input statistics but not the number of inputs. This is even true in the stripped-down, toy version of ICA with a single nongaussian source. This case might perhaps be regarded as the simplest, cleanest, and thus most robust learning scenario imaginable, since there is only one source to acquire. Surely a single nongaussian source can be robustly acquired from the surrounding Gaussian noise? But even here, the critical cross-talk level can depend on that source’s kurtosis.

Thus, the trivial limit $\varepsilon \to 1$ can in fact exert a powerful and profoundly toxic influence on the dynamics right down to $\varepsilon$ arbitrarily close to 0, depending in a detailed, model-dependent manner on the input statistics and the number of afferents. Even very modest levels of cross-talk can therefore entirely compromise the acquisition of the desired, stable fixed points and destroy the capacity for development or learning. To what extent this result applies generally in very broad classes of nonlinear models of synaptic plasticity is unclear. However, we have analyzed in detail three very different models of synaptic plasticity, operating according to very different, well-motivated theoretical principles, precisely in order to attempt to gauge an answer to this question. All three models exhibit multistable regimes, and in all three models, the critical levels of cross-talk can be very small—in two of them, arbitrarily small. While far from conclusive, these very similar results in very different models at least suggest that the disastrous impact of even low levels of cross-talk on the integrity and viability of nonlinear models of synaptic plasticity should be a source of serious and grave concern for both theoreticians and experimentalists. If extant models of synaptic
Bifurcations in Nonlinear Plasticity Models

plasticity are not robust against cross-talk, and noise and errors generally, then we have simply not understood the basis on which real biological systems operate in the face of these unavoidable potential sources of destabilizing bifurcations.

If we have uncovered some commonality in the impact of cross-talk on very different nonlinear models of synaptic plasticity, then perhaps that commonality reflects the common cross-talk matrix $E$ itself? An obvious criticism of our analysis of the impact of cross-talk on nonlinear models, which may be thought to circumvent all the concerns we have raised, is our restriction to the particular, idealized cross-talk matrix in equation 2.4. Specifically, such a matrix is what Cox and Adams (2009) refer to as an “error-onto-all” matrix. Clearly, the error-onto-all assumption is an idealization, at least at the level of single synaptic updates. Cross-talk during single synaptic updates would be expected to be at least localized to the immediate neighborhood of the stimulated synapse, either pre- or postsynaptically. Nevertheless, rather like a diffusion process implemented by a one-step stochastic matrix involving transitions only between adjacent sites, we would expect the impact of even localized cross-talk to accumulate and spread, in a plasticity step-by-plasticity step manner (Adams & Cox, 2002). Thus, although it is an idealization, the precise form of the cross-talk matrix in equation 2.4 is not expected to be unrepresentative. In particular, we would still expect similar results in models of synaptic plasticity that explicitly consider the axonal and dendritic structure of pre- and postsynaptic neurons, with cross-talk more restricted within their arbors.

5.5 Consequences of Cross-Talk for Our Understanding of Synaptic Plasticity. If synaptic cross-talk in particular, and noise and errors more generally, provide an irreducible, potential source of bifurcations in the dynamics of synaptic plasticity, destroying the capacity of nervous systems to acquire developmental-, memory-, and learning-related patterns of fixed points, then what are the consequences experimentally and theoretically?

Tantalizingly, experimental evidence does exist that may be interpreted as demonstrating the capacity for bifurcations in stable fixed point structures during pharmacological and molecular-genetic manipulations. As discussed, synaptic plasticity appears to be subject to cross-talk on different time and length scales, depending on the actions of intra- and extracellular messengers. Focusing on just one of these, manipulations of retrograde neurotrophic factors do alter patterns of synaptic connectivity in a manner entirely consistent with the above analysis. Mice lacking brain-derived neurotrophic factor (BDNF) exhibit impaired hippocampal CA1 LTP (Korte et al., 1995, 1996), and hippocampal CA1 LTP is believed to underlie, for example, spatial learning (Morris, Anderson, Lynch, & Baudry, 1986; McHugh, Blum, Tsien, Tonegawa, & Wilson, 1996). During visual cortical development, the neurotrophic factors BDNF and NT-4/5 are implicated in ODC development. Visual cortical infusion of BDNF and NT-4/5 (Cabelli,
Hohn, & Shatz, 1995) and blockade of the endogenous ligands of the trkB receptor (Cabelli, Shelton, Segal, & Shatz, 1997) prevent ODC formation, while NT-4/5 rescues lateral geniculate cells from the effects of monocular deprivation (Riddle, Lo, & Katz, 1995). These data suggest that learning and developmental dynamics can be shifted pharmacologically from one regime with one set of fixed points (learning- or developmental-related patterns of connectivity) to another regime with a set of fixed points in which normal learning and developmental dynamics have been abolished. While these manipulations are typically very coarse, it is possible that by carefully controlling the concentrations of exogenous or endogenous neurotrophic factors in the visual cortex, for example, developmental dynamics could be moved to the multistable regime discussed above. We would then expect to observe neither perfect ODC formation nor the complete abolition of ODC formation, but rather a mixture of cortical cells, some of which have segregated inputs, and others of which do not have segregated inputs, depending on the initial balance of their inputs. Our analysis would even lead us to predict that the critical levels of neurotrophic factors required to induce this shift could be under environment control, a function of the correlation coefficient between the two eyes’ geniculate inputs to cortical cells. Perhaps such experiments are better suited to in vitro preparations, in which there is much greater scope for controlled pharmacological manipulations and the control of firing patterns (see, for example, Campenot, 1982a, 1982b; Kimpinski, Campenot, & Mearow, 1997).

Notwithstanding these pharmacological and molecular-genetic manipulations, cross-talk under normal circumstances does not in fact appear to have the potentially toxic impact that our analysis suggests. After all, learning, memory, and development do occur, and occur robustly and routinely in real, normal nervous systems. We have presented the nonlinear dynamical systems view as both unavoidable and inescapable, but biology quite self-evidently avoids and escapes the potential pitfalls of cross-talk. How do we reconcile these apparent antinomies into a coherent synthesis? We do so by arguing that our understanding of synaptic plasticity may be incomplete, needing perhaps minor, but perhaps even major, revision.

Biological systems are intrinsically noisy, cross-talk being such an example. But models of synaptic plasticity are typically built as theoretical abstractions, with little concern for the noisy and dirty reality of real biological systems. Issues of noise and errors are often regarded as “second-order” problems and of not much theoretical interest. But if noise and errors are fundamental and irreducible, and if they can radically undermine the viability of models, then we have no choice but to tackle the issues face on. To do so is to engage with the real, biological problems of developing and learning in the real, noisy world rather than in the abstract and pure space of the mathematician’s imagination. Theoretical approaches to the understanding of synaptic plasticity must of course continue to be guided by principles based on information processing and so on, but these approaches
must also take serious account of the way in which biology implements, in its noisy and dirty substrate, these principles. If these principles cannot operate in the presence of even low levels of noise and error, then they cannot be biologically relevant, although they may nevertheless have value in artificial domains in which noise and errors can be engineered or designed out. Thus, we must always ask a two-sided question: Will it work in principle, and will it work in practice? For biology, both sides are equally important, and neither can be relegated to being merely “second order.” In the context of our analysis of the impact of cross-talk here, this means that while nonlinear models cannot avoid bifurcations under the influence of cross-talk, the critical cross-talk levels should be as high as possible. In particular, if the critical levels can be made arbitrarily small, then these models offer little theoretical insight into real, biological plasticity. Interestingly, while both the MMM and the cubic single-unit ICA model analyzed above fail on this score, the model of STDP passes the test: its critical level of cross-talk cannot be made arbitrarily small.

What the minimum critical level of cross-talk should be in models depends on the actual level of cross-talk between real synapses. This will depend on the density of synapses on dendrites. Cox and Adams (2009) have argued that the nervous system, and the cerebral cortex in particular, must actively employ strategies to mitigate the deleterious impact of cross-talk between synapses in order to safeguard the viability of learning. They draw an analogy with error correction in DNA replication. In DNA proofreading, errors in base-pair copying are detected, excised, and corrected by DNA polymerase and exonuclease, vastly reducing the overall error rate to around $10^{-9}$ per base-pair per replication cycle. They argue that a similar “Hebbian proofreading” mechanism must exist in order to detect, correct, and reduce cross-talk-induced errors in synaptic plasticity (Adams & Cox, 2002; Cox & Adams, 2009). They propose that the structure of the neocortex, with its apparently redundant thalamic input to different cortical layers, has evolved to provide two separate copies of afferent input precisely to facilitate the checking and correction required during synaptic plasticity (Adams & Cox, 2002). While we agree that the potential dangers of cross-talk-induced bifurcations are so great that biology may have evolved strategies to reduce or mitigate the impact of cross-talk, it seems to us extravagant to suggest that the principal evolutionary function of the neocortex is essentially to correct synaptic errors induced by cross-talk. Nevertheless, if the actual level of cross-talk between synapses is too high for most realistic models to accommodate comfortably in terms of their operating distance from critical bifurcation behavior, then indeed active strategies against cross-talk, and noise and errors in general, may be necessary. Biology is, after all, extravagant: error correction in DNA replication is in principle a simple and, in hindsight, obvious requirement, but its molecular implementation is undoubtedly extravagant (even if beautiful), the extravagance of its molecular implementation being far outweighed by the
utter necessity to minimize replication errors. Simple theoretical principles can often have, in biology, fantastically complicated, baroque, and elaborate implementations, and thus the charge of “extravagance” demonstrates a fundamental misunderstanding of the implementational complexity of biology. This implementational complexity is a reality that theoreticians must accept and embrace if theoretical neuroscience is to have the same impact on experimental neuroscience that theoretical genetics had on the emergence of molecular genetics (Judson, 1996).

We have focused exclusively on one particular aspect of the nonlinear dynamical systems view, bifurcation dynamics, and have not even broached other classical features of nonlinear dynamics. Extreme sensitivity to initial conditions and its attendant chaotic dynamics on a fractal geometry is perhaps the prototypical example (Strogatz, 1994). For example, although the backpropagation algorithm may not be biologically relevant, it is known that even this standard weapon in the arsenal of the connectionist community exhibits extreme sensitivity to initial conditions and that a fractal structure may be induced in synaptic strength space in the presence of this nonlinear synaptic plasticity algorithm (Kolen & Pollack, 1990; see also French, 1999). On reflection, such possibilities should not be unexpected. After all, the incremental, step-by-step updating of synaptic strengths postulated by our current views of synaptic plasticity merely implements an iterated map. If the iterated map is nonlinear, then chaotic regimes are ubiquitous, indeed universal (May, 1976). The problems associated with cross-talk-induced bifurcations may therefore be only the tip of a large, nonlinear iceberg. The root cause of these problems is repetitive synaptic updates. These wider considerations may therefore hint at the need for a more radical revision of our understanding of synaptic plasticity than even Hebbian proofreading, and that perhaps the prevailing incrementalist, gradualist account of synaptic plasticity is wanting.

Acknowledgments

I am deeply indebted to Paul Adams for encouraging me to examine the consequences of synaptic cross-talk and for numerous, extensive, and ongoing discussions. We radically disagree, however, over the interpretation and implications of cross-talk in models of synaptic plasticity and brain function more generally.

References


Bifurcations in Nonlinear Plasticity Models


Received February 4, 2011; accepted July 29, 2011.