Maximal switchability of centralized networks

Sergei Vakulenko\textsuperscript{1,2}, Ivan Morozov\textsuperscript{3} and Ovidiu Radulescu\textsuperscript{4}

\textsuperscript{1} Institute for Mech. Engineering Problems Saint Petersburg, Russia
\textsuperscript{2} Saint Petersburg National Research University of Information Technologies, Mechanics and Optics. Saint Petersburg, Russia
\textsuperscript{3} University of Technology and Design, Saint-Petersburg, Russia
\textsuperscript{4} DIMNP—UMR 5235 CNRS/UM, Université de Montpellier, Montpellier, France

E-mail: vakulenfr@mail.ru and ovidiu.radulescu@umontpellier.fr

Received 23 March 2015, revised 26 May 2016
Accepted for publication 8 June 2016
Published 30 June 2016

Recommended by Professor James A Glazier

Abstract
We consider continuous time Hopfield-like recurrent networks as dynamical models for gene regulation and neural networks. We are interested in networks that contain \( n \) high-degree nodes preferably connected to a large number of \( N_s \) weakly connected satellites, a property that we call \( n/N_s \)-centrality. If the hub dynamics is slow, we obtain that the large time network dynamics is completely defined by the hub dynamics. Moreover, such networks are maximally flexible and switchable, in the sense that they can switch from a globally attractive rest state to any structurally stable dynamics when the response time of a special controller hub is changed. In particular, we show that a decrease of the controller hub response time can lead to a sharp variation in the network attractor structure: we can obtain a set of new local attractors, whose number can increase exponentially with \( N \), the total number of nodes of the network. These new attractors can be periodic or even chaotic. We provide an algorithm, which allows us to design networks with the desired switching properties, or to learn them from time series, by adjusting the interactions between hubs and satellites. Such switchable networks could be used as models for context dependent adaptation in functional genetics or as models for cognitive functions in neuroscience.

Keywords: networks, chaos, bifurcations
Mathematics Subject Classification numbers: 92B20, 37D45, 37D05
1. Introduction

Networks of dynamically coupled elements have imposed themselves as models of complex systems in physics, chemistry, biology and engineering [36]. The most studied propriety of networks is their topological structure. Structural features of networks are usually defined by the distribution of the number of direct connections a node has, or by various statistical properties of paths and circuits in the network [36, 3]. An important structure related property of networks is their scale-freeness [3, 7, 23, 24] often invoked as a paradigm of self-organization and spontaneous emergence of complex collective behaviour [10]. In scale-free networks the fraction \( P(k) \) of nodes in the network having \( k \) connections to other nodes (i.e. having degree \( k \)) can be estimated for large values of \( k \) as \( P(k) \sim k^{-\gamma} \), where \( \gamma \) is a parameter whose value is typically in the range \( 2 < \gamma < 3 \) [3]. In such networks, the degree is extremely heterogeneous. In particular, there are strongly connected nodes that can be named hubs, or centers. The hubs communicate to each other directly, or via a number of weakly connected nodes. The weakly connected nodes that interact mainly with hubs can be called satellites. Scale-free networks have also nodes of intermediate connectivity. Networks that have only two types of nodes, strongly connected hubs and weakly connected satellites are known as bimodal degree networks [51]. Because of the presence of a large number of hubs, scale-free or bimodal degree networks can be called centralized. Centralized connectivity has been found by functional imaging of brain activity in neuroscience [10], and also by large scale studies of the protein-protein interactions or of the metabolic networks in functional genetics [23, 24].

The centralized architecture was shown to be important for many emergent properties of networks. For instance, there has been a lot of interest in the resilience of networks with respect to attacks that remove some of their components [4]. It was shown that networks with bimodal degree connectivity are resilient to simultaneous targeted and random attacks [51], whereas scale-free networks are robust with respect to random attacks, but sensitive to targeted attacks that are directed against hubs [5, 11]. For this reason, the term ‘robust-yet-fragile’ was coined in relation to scale-free networks [8].

From a more dynamical perspective, a centralized architecture facilitates communication between hubs, stabilizes hubs by making them insensitive to noise [54, 55] and allows for hub synchronization even in the absence of satellite synchronization [41, 42, 48]. Another important question concerning networks is how to push their dynamics from one region of the phase space to another or from one type of behaviour to another, briefly how to control the network dynamics [13, 17, 25, 30, 35, 39, 40, 46, 49, 61]. Several authors used Kalman’s results for linear systems to understand how network structure influences network dynamics controllability, and in particular how to choose the control nodes [13, 30, 35]. As pointed out by [27, 34] several difficulties occur when one tries to apply these general results to real networks. Even for linear networks, the control of trajectories is nonlocal [49] and shortcuts are rarely allowed. As a result, even small changes of the network state may ask for control signals of large amplitude and energy [59]. The control of nonlinear networks is even more difficult and in this case we have no general results. Nonlinear networks can have several co-existing attractors and it is interesting to find out how to push the state of the network from one attractor basin to another. The ability of networks to change attractor under the effect of targeted perturbations can be called switchability. In relation to this, the paper [43] has introduced the terminology ‘stable yet switchable’ (SyS) meaning that the network remains stable given a context and is able to reach another stable state when a stimulus indicates a change of the context. It was shown, by numerical simulations, that centralized networks with bimodal degree distribution are more prone to SyS behavior than scale-free networks [43]. Switchability is important for practical reasons, for instance in drug design. In such applications, one uses pharmaceutical action on...
nodes to push a network that functions in a pathological attractor (such pathological attractors were discussed in relation to cancer [22] or neurological disorders [15, 47]) to a healthy functioning mode, characterized by a different attractor. Numerical methods to study switchability of linear [58] and nonlinear [12] networks were discussed in relation with drug design in cancer research. In theoretical biology, network switchability can be important for mathematical theories of genetic adaptation [37]. If one looks at organisms as complex systems and model them by networks, then adaptation to changes in the environment can be described as switching the network from one attractor to another one with a higher fitness [37]. An important question that is often asked with respect to tuning network dynamics is how many driver nodes are needed to control that dynamics. For linear networks, it was shown that this number is large if we aim to obtain a total control, which allows us to switch the network between any pair of states. This number can be as high as 80% for molecular regulatory networks [31]. This fact, as emphasized in [58], contradicts empirical results about cellular reprogramming and about adaptive evolution. Much less nodes are needed if instead of full controllability one wants switching between specific pairs of unexpected and desired states [58]. This concept, named ‘transittability’ in [58], is very similar to our switchability, but was studied only for linear systems.

In this paper, we study dynamical properties of large nonlinear networks with centralized architecture. We consider continuous time versions of the Hopfield model of recurrent neural networks [20] with a large number of neurons. The Hopfield model is based on the two-states McCulloch and Pitts formal neuron and uses symmetrical weight matrices to specify interactions between neurons. Like to the Hopfield version, we use a thresholding function to describe switching between the two neuron states, active and inactive. However, contrary to the original Hopfield version, we do not impose symmetrical interactions between neurons, in other words our weight matrix is not necessarily symmetric. This model has been successfully used to describe associative memories [20], neural computation [21, 32], disordered systems in statistical physics [50], neural activity [15, 29] and also to investigate space-time dynamics of gene networks in molecular biology [33, 57]. The choice of such type of dynamics is motivated by the existence of universal approximation results for multilayered perceptrons (see, for example, [6]). In particular, we have shown elsewhere that networks with Hopfield-type dynamics can approximate any structurally stable dynamics, including reaction-diffusion biochemical networks also largely used in biology [54].

Our aim is to study analytically the ability of a network with centralized architecture to be switchable. We employ a special notion of centrality. Many biological networks exhibit so-called dissortative mixing, i.e. high-degree nodes are preferably connected to low-degree nodes [1]. We will consider networks with strongly connected hubs. We also assume that each hub is under the action of at least weakly connected satellites, that on turn receive actions from all the hubs. For large networks, increases at least as fast as a power of , where and are constants and is the total number of nodes. We call this property -centrality. This network architecture ensures a large number of feed-back loops that produce complex dynamics. Furthermore, the dissortative connectivity implies functional heterogeneity of the hubs and satellites. The hubs play the role of controllers and the satellites sustain the feedback loops needed for attractor multiplicity. The large number of satellites guarantees a sufficient flexibility of the network dynamics and also buffer the perturbations transmitted to the hubs. This principle applies well to gene networks. The hubs in such networks can be the transcription factors, which are stabilized by numerous interactions with non-coding RNAs that represent the satellites [28]. In addition to structural conditions, we will consider a special correlation between time scales and connectivity of the nodes: the hubs have slow response, whereas the satellites respond rapidly. This condition is natural for
many real networks. The hubs have to cope with multiple tasks, therefore they must have more complex interaction than the satellites. Consequently, the hubs need more resources to be produced, decomposed, and react with other nodes, therefore their dynamics is slow. This property is obvious for gene networks, where transcription factors are complex proteins, much larger and more stable than the non-coding RNAs.

Our first result is valid without conditions on the structure and depends only on the condition on the timescales. We assume that there exist \( n \ll N \) slow nodes, whereas all the remaining ones are fast. Then, the dynamics of the network can be reduced to \( n \) variables. We prove the existence of an inertial manifold of dimension \( n \), which completely captures all network dynamics for large times. We recall that the fundamental concept of inertial manifold was introduced for infinite dimensional and multidimensional systems. The inertial manifolds are globally attracting invariant ones \([38]\). The large time dynamics of a system possessing an inertial manifold, is defined by a smooth vector field \( F \) of relatively small dimension, so-called inertial form. All attractors lie on inertial manifold \([38]\).

The second result holds under the structural assumption that the network is \( n/N_s \)-central. Under this condition, we show that the inertial forms \( F \) obtained from such networks are dense in the set of all smooth vector fields of dimension \( n \). This implies that given a certain combination of attractors defined by vector fields \( Q_i \) we can construct a centralized network that exhibits a combination of attractors that is topologically equivalent to the one given. Furthermore, we show that \( n/N_s \)-central networks can exhibit ‘maximal switchability’. By changing a control parameter \( \xi \), which determines the response time of a single network hub (‘controller’ hub), we can sharply change the network attractor. For instance we can switch from a situation when the network has a single rest point for \( \xi > \xi_0 \) to a situation when the network has a complicated global attractor for \( \xi < \xi_0 \), including a number of local attractors, which may be periodic or chaotic. The network state tends to the corresponding local attractor depending on the initial state of the control hub. This result shows in an analytical and rigorous way how nonlinear networks can be switched by only one control node. The possibility of switching nonlinear networks by a small number of nodes is crucial in theories of genetic adaptation. Indeed, phenomenological theories predict and empirical data confirm that the main part of the adaptive evolution process consists in only a few mutations producing large fitness changes \([37]\).

Our third result proves, in an analytical way, that the number of rest point local attractors (and therefore the network capacity) of \( n/N_s \)-central networks may be exponentially large in the number of nodes.

We also describe a constructive algorithm, which allows us to obtain a centralized network that performs a prescribed inertial dynamics and the desired switching properties of the network.

### 2. Problem statement and main assumptions

We consider the Hopfield-like networks \([20]\) described by the ordinary differential equations

\[
\frac{du_i}{dr} = \sigma \left( \sum_{j=1}^{N} W_{ij}u_j - h_i \right) - \lambda_i u_i, \tag{2.1}
\]

where \( u_i, h_i \) and \( \lambda_i > 0, i = 1, ..., N \) are node activities, activation thresholds and degradation coefficients, respectively. The matrix entry \( W_{ij} \) describes the action of the node \( j \) on the node \( i \), which is an activation if \( W_{ij} > 0 \) or a repression if \( W_{ij} < 0 \). Contrary to the original Hopfield
model, the interaction matrix $W$ is not necessarily symmetric. The function $\sigma$ is an increasing and smooth (at least twice differentiable) ‘sigmoidal’ function such that

$$\sigma(-\infty) = 0, \quad \sigma(+\infty) = 1, \quad \sigma'(z) > 0.$$  \hspace{1cm} (2.2)

Typical examples can be given by

$$\sigma = \frac{1}{1 + \exp(-h)}, \quad \sigma = \frac{1}{2} \left( \frac{h}{\sqrt{1 + h^2}} + 1 \right).$$  \hspace{1cm} (2.3)

The structure of interactions in the model is defined by a weighted digraph $(V, E, W)$ with the set $V$ of nodes, the edge set $E$ and weights $W_{ij}$. The nodes $v_i, j = 1 \ldots, N$ can be neurons or genes, depending on applications.

**Assumption 1.** Assume that if $W_{ij} \neq 0$, then $(i, j)$ is an edge of the graph, $(i, j) \in E$. This means that the $i$th node can act on the $j$th node only if it is prescribed by an edge of the digraph $(V, E, W)$. We also suppose that $(i, i) \notin E$, i.e. the nodes do not act on themselves.

Assume that the digraph $(V, E, W)$ satisfies a condition, which is a variant of the centrality property. This condition is a purely topological one and thus it is independent on the weights $W_{ij}$. To formulate this condition, we introduce a special notation.

Let us consider a node $v_j$. Let us denote by $S^*(j)$ the set of all nodes, which act on the neuron $j$:

$$S^*(j) = \{ v_i \in V : \text{edge } (i, j) \in E \}. \hspace{1cm} (2.4)$$

For each set of nodes $C \subset V$ we introduce the set $S(C)$ of the nodes, which are under action of all nodes from $C$ and which are not belonging to $C$:

$$S(C) = \{ v_i \in V : \text{for each } j \in C \text{ edge } (j, i) \in E \text{ and } v_j \notin C \}. \hspace{1cm} (2.5)$$

**$n/N_c$-Centrality assumption.** The graph $(V, E, W)$ is connected and there exists a set of nodes $C$ such that

i. $C$ consists of $n$ nodes;

ii. for each $j \in C$ the intersection $S^*(j) \cap S(C)$ contains at least $N_c$ nodes, where $N_c > c_0 N^\theta$ with constants $c_0 > 0, \theta \in (0, 1)$, which are independent of $j$ and $N$.

The nodes from $C$ can be interpreted as hubs (centers) and the nodes from $S(C)$ are the satellites. The condition ii implies that each center is under action of sufficiently many satellites. In turn, if we consider the union of these satellites, all the centers act on them (see figure 1). Such an intensive interaction leads, as we will see below, to a very complicated large time behaviour.

**3. Outline of main results**

Our results can be outlined as follows. The result on the inertial dynamics existence describes a situation, when the interaction topology is quite arbitrary. We assume that there exist $n$ slow nodes, say, $u_1, u_2, \ldots, u_n$ with $\lambda_i = O(1)$ whereas all the rest ones $u_{n+1}, \ldots, u_N$ are fast, i.e. the corresponding $\lambda_i$ have order $O(\kappa^{-1})$, where $\kappa$ is a small parameter. Then we show that there exists an inertial manifold of dimension $n$. We obtain, under general conditions, that for times $t \gg \kappa \log \kappa$ the dynamics of (2.1) is defined by the reduced equations
\[ \frac{du_i}{dt} = F_j(u_1, \ldots, u_n, W, h, \lambda), \]
\[ u_k = U_k(u_1, \ldots, u_n, W, h, \lambda), \quad k = n + 1, \ldots, N, \]
where \( F_j \) and \( U_k \) are some smooth functions of \( u_1, \ldots, u_n \), and \( h, \lambda \) denote the vector parameters \( (h_1, \ldots, h_N) \) and \( (\lambda_1, \ldots, \lambda_N) \), respectively. So, \( F \) gives us the inertial form on an inertial manifold.

The inertial form completely defines the dynamics for large times [38]. More interestingly, we can show that the vector field \( F \) is, in a sense, maximally flexible. Roughly speaking, by the number of nodes \( N \), the matrix \( W \) and \( h \) we can obtain all possible fields \( F \) (up to a small accuracy \( \varepsilon \), which can be done arbitrarily small as \( N \) goes to \( \infty \)), see section 5 for a formal statement of this flexibility property. For the networks this flexibility property holds under \( n/N_c \)-centrality assumption.

Let us introduce a special control parameter \( \xi \), which modulates the degradation coefficient \( \lambda_i \) for a hub: \( \lambda_i = \xi \lambda_i \) for some \( i \in C \). This hub is a ‘controller’. When we vary the coefficient \( \xi \), the interaction topology and the entries of the interaction matrix do not change, but the response time of the controller hub changes.

One can choose the network parameters \( N, W, \lambda \) in such a way that for \( \xi > \xi_0 \), the global attractor is trivial, it is a rest point, but for an open set of other values \( \xi \) the global attractor of (2.1) contains a number of local attractors.

This result can be interpreted as ‘maximal switchability’. A similar effect was found in [14] by numerical simulations for some models of neural networks. This effect describes a transition from neural resting states (NRS) to complicated global attractors, which occur as a reaction on learning tasks. Note that in [14] attractors consist of a number of steady states. In our case the global attractors can include many local attractors of all possible kinds including chaotic and periodic ones.

We end this section with a remark. Our method approximates vector fields by neural networks, but what can be said about the relationship between the trajectories of the simulated system and the ones corresponding to the neural network?

For chaotic and even for periodic attractors, direct comparison of trajectories is not a suitable test for the accuracy of the approximation. General mathematical arguments allow us
say only that these trajectories will be close for bounded times. For large times we can say nothing especially for general chaotic attractors. Consider the case when the attractor $A$ of the simulated system is transitive. This means the dynamics is ergodic and for smooth function $\phi$ the time averages

$$S_{F,\phi} = \lim_{T \to +\infty} T^{-1} \int_0^T \phi(v(t))dt$$

(3.3)

coincide with the averages $\int_A \phi(v) d\mu(v)$ over the attractor, where $\mu$ is an invariant measure on $A$.

Then, a suitable criterion of approximation is that the averages $S_{F,\phi}$ and the corresponding ones generated by the approximating centralized neural network, are close for smooth $\phi$:

$$|S_{F,\phi} - S_{G_{\text{an}},\phi}| = \text{Err}_{\text{approx}} < \delta(\epsilon, \phi)$$

(3.4)

where $G_{\text{an}}$ is the neural network approximation of $F$ and $\delta \to 0$ as $\epsilon \to 0$. This ‘stochastic stability’ property holds for hyperbolic (structurally stable) attractors [26, 56, 60].

4. Conditions on network parameters and attractor existence

Our first results do not use any assumptions on the network topology. However, we suppose that there are two types of network components that are distinguished by their time scales into slow nodes and fast nodes. To take into account the two types of the nodes, we use distinct variables $v_j$ for slow variables, $j = 1, \ldots, n$ and $w_i$ for the fast ones, $i = 1, \ldots, N - n = N_1$. The real matrix entry $A_{ji}$ defines the intensity of the action of the fast node $i$ on the slow node $j$. Similarly, the $n \times N_1$ matrix $B$, $N_1 \times N_1$ matrix $C$ and $n \times n$ matrix $D$ define the action of the slow nodes on the fast ones, the interactions between the fast nodes and the interactions between the slow nodes, respectively. We denote by $h_i$ and $\lambda_i$ the threshold and degradation parameters of the fast nodes and by $\tilde{h}_i$ and $\tilde{\lambda}_i$ the same parameters for the slow nodes, respectively. To simplify formulas, we use the notation

$$\sum_{j=1}^n D_{ji} v_j = D_i v, \quad \sum_{k=1}^N C_{jk} w_k = C_j w.$$

Then, equations (2.1) can be rewritten as follows:

$$\frac{dw_i}{dr} = \sigma(B_i v + C_i w - \tilde{h}_i) - \kappa^{-1} \tilde{\lambda}_i w_i,$$

(4.1)

$$\frac{dv_j}{dr} = \sigma(A_j w + D_j v - h_j) - \lambda_j v_j,$$

(4.2)

where $i = 1, \ldots, N_1$, $j = 1, \ldots, n$. Here unknown functions $w_i(t)$, $v_j(t)$ are defined for times $t \geq 0$. We assume that $\kappa$ is a positive parameter, therefore, the variables $w_i$ are fast.

We set the initial conditions

$$w_i(0) = \tilde{\phi}_i \geq 0, \quad v_j(0) = \phi_j \geq 0.$$

(4.3)

It is natural to assume that all concentrations are non-negative at the initial moment. It is clear that they stay non-negative for all times.
4.1. Global attractor exists

Let us prove that the network dynamics is correctly defined for all \( t \) and solutions are non-negative and bounded. For positive vectors \( r = (r_1, ..., r_n) \) and \( R = (R_1, ..., R_N) \), let us introduce the sets \( \mathcal{B} \) defined by

\[
\mathcal{B}(r, R) = \{ (w, v) : 0 \leq v_j \leq r_j, 0 \leq w_j \leq R_j, j = 1, ..., n, \ i = 1, ..., N \}.
\]

Note that

\[
\frac{dw_j}{dt} < 1 - \kappa^{-1} \tilde{\lambda} w_j.
\]

Thus, \( w_j(t) < X(t) \) for positive times \( t \), where

\[
\frac{dX}{dt} = 1 - \kappa^{-1} \tilde{\lambda} X, \quad X(0) = w_0(0).
\]

Therefore, resolving the last equation, and repeating the same estimates for \( v_i(t) \), one finds

\[
0 \leq w_j(x, t) \leq \bar{\phi}_j^+ \exp(-\tilde{\kappa}^{-1} \lambda_j t) + \kappa \bar{\lambda}_j^{-1} (1 - \exp(-\kappa^{-1} \tilde{\lambda} t)),
\]

\[
0 \leq v_j(x, t) \leq \phi_j^+ \exp(-\lambda_j t) + \tilde{\lambda}_j^{-1} (1 - \exp(-\lambda_j t)),
\]

(4.4)

Let us take arbitrary \( a > 1 \) and let \( r_j(a) = a \lambda_j^{-1} \) and \( R_j(a) = a \tilde{\lambda}_j^{-1} \). Estimates (4.4) show that solutions of (4.1) and (4.2) exist for all times \( t \) and they enter the set \( \mathcal{B}(r(a), R(a)) \) at a time moment \( t_0 \). The solutions stay in this set for all \( t > t_0 \), thus, this set is absorbing. This shows that system (4.1) and (4.2) defines a global dissipative semiflow \( \mathcal{S}_H \). Moreover, this semiflow has a global attractor contained in each \( \mathcal{B}(r(a), R(a)) \), where \( a > 1 \).

4.2. Assumptions for slow/fast networks

A simpler asymptotic description of system dynamics is possible under assumptions on network components timescales. We suppose here that the \( u \)-variables are fast and the \( v \)-ones are slow. We show then that the fast \( w \) variables are slaved, for large times, by the slow \( v \) modes. More precisely, one has \( w = \kappa U(v) + \tilde{w} \), where \( \kappa U(v) \) is a correction and \( \kappa > 0 \) is a small parameter. This means that, for large times, the fast nodes dynamics is completely controlled by the slow nodes.

To realize this approach, let us assume that the system parameters \( P = \{ A, B, C, D, h, \tilde{h}, \tilde{\lambda}, \lambda \} \) satisfy the following conditions:

\[
A = \kappa^{-1} \tilde{A},
\]

(4.5)

\[
|\tilde{A}|, |B|, |C|, |D| < c_0,
\]

(4.6)

\[
0 < c_1 < \tilde{\lambda}_i < c_2, \quad 0 < \lambda_i < c_3.
\]

(4.7)

Here all positive constants \( c_k \) are independent of \( \kappa \) for small \( \kappa \).

The scaling assumption on \( A \) is needed because, as we will prove later, \( w = O(\kappa) \) for small \( \kappa \). For the same reasons, \( C, w \) can be neglected with respect to \( B, v \) for small \( \kappa \), meaning that the action of centers on satellites is dominant with respect to satellites mutual interactions. In other words, these conditions describe a divide and rule control principle.
5. Realization of prescribed dynamics and maximally flexible systems

Our goal is to show that the network dynamics can realize, in a sense, arbitrary structurally stable dynamics of the centers. To precise this assertion, let us describe the method of realization of the vector fields for dissipative systems (proposed in [44]). More precisely, we are interested in systems enjoying the following properties:

A. These systems generate global semiflows \( S^\rho \) in an ambient Hilbert or Banach phase space \( H \). These semiflows depend on some parameters \( \rho \) (which could be elements of another Banach space \( B \)). They have global attractors and finite dimensional local attracting invariant \( C^1 \)-manifolds \( \mathcal{M} \), at least for some \( \rho \).

B. Dynamics of \( S^\rho \) reduced on these invariant manifolds can be, in a sense, almost completely tuned by variations of the parameter \( \rho \).

It can be described as follows. Assume the differential equations

\[
\frac{dq}{dt} = Q(q), \quad Q \in C^1(B^n)
\]

(5.1)

define a global semiflow in a unit ball \( B^n \subset \mathbb{R}^n \).

For any prescribed dynamics (5.1) and any \( \epsilon > 0 \), we can choose suitable parameters \( \rho = \rho(n, F, \epsilon) \) such that

B1. The semiflow \( S^\rho \) has a \( C^1 \)-smooth locally attracting invariant manifold \( \mathcal{M}_\rho \) diffeomorphic to \( B^n \);

B2. The reduced dynamics \( S^\rho|_{\mathcal{M}_\rho} \) is defined by equations

\[
\frac{dq}{dt} = \tilde{Q}(q, \rho), \quad \tilde{Q} \in C^1(B^n)
\]

(5.2)

where the estimate

\[
|Q - \tilde{Q}(q, \rho)| < \epsilon
\]

(5.3)

holds. In other words, one can say that, by \( \rho \), the reduced dynamics on the invariant manifold can be specified to within an arbitrarily small error.

Therefore, roughly speaking all robust dynamics (stable under small perturbations) can be generated by the systems, which satisfy above formulated properties. Such systems can be named maximally flexible. In order to show that maximal flexibility covers also the case of chaotic dynamics, let us recall some facts about chaos and hyperbolic sets.

Let us consider dynamical systems (global semiflows) \( S^t_1, ..., S^t_k \), \( t > 0 \), defined on the \( n \)-dimensional closed ball \( B^n \subset \mathbb{R}^n \) defined by finite dimensional vector fields \( F^{(i)} \in C^1(B^n) \) and having structurally stable attractors \( A_i \), \( i = 1, ..., k \). These attractors can have a complex form, since it is well known that structurally stable dynamics may be ‘chaotic’. There is a rather wide variation in different definitions of ‘chaos’. In principle, one can use here any concept of chaos, provided that this is stable under small \( C^1 \)-perturbations. To fix ideas, we shall use here, following [45], such a definition. We say that a finite dimensional dynamics is chaotic if it generates a compact invariant hyperbolic set \( \Gamma \), which is not a periodic cycle or a rest point (for a definition of hyperbolic sets see, for example, [45]). The hyperbolic sets give remarkable analytically tractable examples, where chaotic dynamics can be studied. For example, the Smale horseshoe is a hyperbolic set. If this set \( \Gamma \) is attracting we say that \( \Gamma \) is a chaotic (strange) attractor. In this paper, we use only the following basic property of hyperbolic sets,
so-called Persistence [45]. This means that the hyperbolic sets are, in a sense, stable (robust). This property can be described as follows. Let a system of differential equations be defined by a $C^1$-smooth vector field $Q$ on an open domain in $\mathbb{R}^n$ with a smooth boundary or on a smooth compact finite dimensional manifold. Assume this system defines a dynamics having a compact invariant hyperbolic set $\Gamma$. Let us consider $\epsilon$-perturbed the vector field $\tilde{Q} = Q + \epsilon \hat{Q}$, where $\hat{Q}$ is bounded in $C^1$-norm. Then, if $\epsilon > 0$ is sufficiently small, the perturbed field also generates dynamics with another compact invariant hyperbolic set $\tilde{\Gamma}$. The corresponding dynamics restricted to $\Gamma$ and $\tilde{\Gamma}$ respectively, are topologically orbitally equivalent (topological equivalence of two semiflows means that there exists a homeomorphism, which maps the trajectories of the first semiflows on the trajectories of the second one, see [45] for details).

We recall that chaotic structurally stable (persistent) attractors and invariant sets exist: this fact is well known from the theory of hyperbolic dynamics [45]. Thus, any kind of the chaotic hyperbolic sets can occur in the dynamics of the systems, for example, the Smale horseshoes, Anosov flows, and the Ruelle–Takens–Newhouse chaos, see [45]. Examples of systems satisfying these properties can be presented by some reaction-diffusion equations and systems [44, 52, 53], and neural network models [53].

6. Main results

For vectors $a = (a_1, ..., a_n)$ and $b = (b_1, ..., b_n)$ such that $a_i < b_i$ for each $i$ let us denote by

$$\Pi(a, b) = \{v \in \mathbb{R}^n : a_i \leq v_i \leq b_i\}$$

(6.1)
a $n$-dimensional box in $v$-space. Moreover, let us define $\Pi_0$ by $\Pi_0 = \Pi(0, \lambda^{-1})$, where the vector $\lambda^{-1}$ has components $(\lambda_1^{-1}, ..., \lambda_n^{-1})$.

**Theorem 6.1.** Under assumptions (2.2), (4.5)–(4.7) for sufficiently small $\kappa$ there exists a $n$-dimensional inertial manifold $M_n$ defined by

$$w_i = \kappa \hat{\lambda}_i^{-1} U_i(v, \kappa, P), \quad v \in \Pi_0$$

(6.2)

where $U_i \in C^{4+r}(\Pi_0)$, and $r \in (0, 1)$. The functions $U_i$ admit the estimate

$$|U_i(v, \kappa, P) - \sigma(B_i v - \tilde{h}_i)| \leq (4\kappa, \quad v \in \Pi_0)$$

(6.3)

The $v$ dynamics for large times takes the form

$$\frac{dv_i}{dt} = F_j(v, P) + \bar{F}_j(v, \kappa, P),$$

(6.4)

where $\bar{F}_j$ satisfy

$$|\bar{F}_j|_{C^1(\Pi_0)} < c_0 \kappa$$

(6.5)

with

$$F_j(v, P) = \sigma \left( \sum_{i=1}^{N-n} \hat{A}_j \hat{\lambda}_i^{-1} \sigma(B_i v - \tilde{h}_i) + D_j v - h_j \right) - \lambda_j v_j$$

(6.6)

Note that the matrix $C$ is not involved in relation (6.6), which defines the family of the vector fields $F$ (inertial forms). This property holds due to the property that
inter-satellite interactions are dominated by the satellite-center ones. The next assertion means that this principle allows us to create a network dynamics with prescribed dynamics (if the network satisfies \( n/N_s \)-centrality assumption and \( N_s \) is large enough). It is valid under the additional condition that the interaction graph \((V, E)\) verifies the centrality condition.

**Theorem 6.2.** Assume \( n/N_s \)-centrality assumption is satisfied. Then the family of the vector fields \( F \) defined by (6.6) is dense in the set of all \( C^1 \) vector fields \( Q \) defined on the unit ball \( B^n \subset \mathbb{R}^n \). In the other words, centralized Hopfield neural networks are maximally flexible.

Let us choose some \( i_C \) such that \( i_C \) belongs to \( C \). The corresponding node will be called a controller hub. We introduce the control parameter \( \xi \) by

\[
\lambda_{ic} = \xi \lambda, \quad (6.7)
\]

where we fix a positive \( \bar{\lambda} \).

Theorem 6.2 can be used to show the following

**Theorem 6.3 (Maximal switchability theorem).** Let us consider dynamical systems (global semiflows) \( S'_1, \ldots, S'_k, t \geq 0 \), defined on the n-dimensional closed ball \( B^n \subset \mathbb{R}^n \) defined by finite dimensional vector fields \( F^{(k)} \in C^1(B^n) \) and having structurally stable attractors \( A_l \), \( l = 1, \ldots, k \).

For sufficiently large \( N \) and any graph \((V, E)\) satisfying the \( n/N_s \)-centrality condition there exists a choice of interactions \( W_{ij} \) and thresholds \( h_i \) such that assumption 1 holds and

(i) there exist a \( \xi_0 \) such that for all \( \xi > \xi_0 \) the dynamics of network (2.1) has a rest point, which is a global attractor;

(ii) for an open interval of values \( \xi \) the global semiflow \( S'_H \) defined by (2.1) have local attractors \( B_l \) such that the restrictions of the semiflow \( S'_H \) to \( B_l \) are orbitally topological equivalent to the semiflows \( S_l \) restricted to \( A_l \).

Finally, let us give an estimate on the maximal number of equilibria \( N_{eq} \) of centralized networks. This number is a characteristics of the network capacity, flexibility and adaptivity. To proceed to these estimates, let us define a procedure, which can be named decomposition into ‘distar’ motifs. In the network interaction graph \((E, V)\) we choose some nodes \( v_1, \ldots, v_m \), which we conditionally consider as hubs. By ‘distar’ motif we understand a part of interaction graph consisting of the hub \( v_j \) and the subset \( S_j \) of the set \( S'_j \) (defined by (2.5)) consisting of the nodes connected in both directions to \( v_j \): \( S_j = \{ v_l \in V : (j, l) \text{ and } (l, j) \in E \} \). This distar motif becomes an usual star if directions of the edges are ignored. Consider the union \( U_n \) of all \( S_j \). Some nodes \( w \in U_n \) may belong to two different sets \( S_j \) and \( S_k \), where \( k \neq j \). We remove from the vertex set \( V \) all such nodes. After such removing we obtain a part of graph \( G_n = (V', E') \) of the initial graph \((E, V)\), which is a union of \( n \) disjoint distars \( S_1, \ldots, S_n \), where each \( S_j \) contains a single center \( \{ v_j \} \) and \( \mu(S_j) \) satellites connected with the center in both directions. Recall that the graph \((V', E')\) is a part of graph \((V, E)\) if \( V' \subset V \) and \( E' \subset E \). These numbers \( \mu(S_j) \) depend on the choice of hub nodes \( \{ v_1, \ldots, v_n \} \).

We will prove the following theorem:

**Theorem 6.4.** The maximal possible number \( N_{eq}(E, N) \) of equilibria of a network with a given interaction graph \((E, V)\), where \( V \) consists of \( N \) nodes, satisfies

\[
N_{eq} \geq \sup \mu(S_j) \mu(S_2) \cdots \mu(S_n), \quad (6.8)
\]
where the supremum is taken over all integers $n > 0$ and all graphs $G_n$, which are parts of interaction graph $(V, E)$ and consist of $n$ disjoint distars. Here $\mu(S_l)$ is the number of the nodes in the distar $S_l$.

Consider now graphs, which are unions of identical distars. The degree of the center of each distar is $\mu(S_l) = \lambda^{-1}_l - n$. Then, the maximal possible number $N_{eq}$ of equilibria in such a centralized network (2.1) with $N$ nodes and $n$ centers satisfies $N_{eq} \geq [(N - n)/n]^n$, where $[x]$ denotes the floor of a real number $x$. Note that for a fixed $N$ the maximum of $(N/n)^n$ over $n = 1, 2, ...$ is attained at $n = \lfloor N/5 \rfloor$, when the distars contain 5 satellites each. Therefore we obtain the estimate $N_{eq} \geq 4^{\lfloor N/5 \rfloor}$.

### 7. Proof of theorem 6.1

Let us start by proving a lemma

**Lemma 7.1.** Under assumptions (4.5)–(4.7) for sufficiently small positive $\kappa < \kappa_0$ solutions $(u, v)$ of (4.1)–(4.3) satisfy

$$w_i(t) = \kappa U_i(v(t), B, \hat{h}) + \tilde{w}_i(t),$$

(7.1)

where $U = (U_1, ..., U_n)$ is defined by

$$U_i(v, B, \hat{h}) = \lambda^{-1}_i \sigma(B_i v(t) - \hat{h}_i).$$

(7.2)

Then, for some $T_0$ function $\tilde{w}$ satisfies the estimates

$$|\tilde{w}(t)| < c_1 \kappa^2, \quad t > T_0$$

(7.3)

where $c_1$ does not depend on $t$ and $\kappa$. The time moment $T_0$ depends on initial data and the network parameters.

**Proof.** Let us introduce a new variables $\tilde{w}_i$ by (7.1). They satisfy the equations

$$\frac{d\tilde{w}_i}{dt} = H_i(v, \tilde{w}) - \kappa^{-1}\lambda_i \tilde{w}_i,$$

(7.4)

where

$$H_i(v, \tilde{w}) = \kappa Z_i(v) + W_i(v, \tilde{w}),$$

and

$$Z_i(v) = \sum_{j=1}^n \frac{\partial U_i(v)}{\partial v_j} (\sigma(\lambda_i U_i + D_j v - h_j) - \xi \lambda_j v_j),$$

and

$$W_i(v, \tilde{w}) = \sigma(B_i v + C_j w - \hat{h}_j) - \sigma(B_i v - \hat{h}_j).$$

Let us estimate $H_i(v, \tilde{w})$ for sufficiently large $t$. According to (4.4), for such times we can use that $(v, w) \in B(r(a), R(a))$, where $a > 1$. In this domain $B(r(a), R(a))$ one has $|Z_i| < c_2$ and sup $|W_i| < c_3 \kappa$, where $c_2, c_3$ are independent of $\kappa$. Therefore,

$$H_i(v(t), \tilde{w}(t)) < c_0 \kappa, \quad t > T_0(\kappa, \textbf{P}).$$

Now, as above in section 4.1, equation (7.4) entails estimate (7.3). The assertion is proved. □
Proof of theorem 6.1. The rest part of the proof of theorem 6.1 uses the well known tech-
nique of invariant manifold theory, see, for example, [19, 38, 45]. Let us consider the domain
\( D_k = \{ w : |w| < c \nu_k^2 \} \). Theorem 6.1.7 [19] shows that for \( d \in (0, 1) \) there is a locally attractive
\( C^{1+d} \)-smooth invariant manifold \( \mathcal{M}_d \). Relation (6.3) follows from (7.3). The global attractivity
of this manifold also follows from (7.3). The theorem is proved.

8. Proof of theorems 6.2–6.4

8.1. Proof of theorem 6.2

The main idea of the subsequent statement is to study the dependence of the fields \( F_j \) defined
by equation (6.6) on the parameters \( P \). To this end, we apply a special method stated in the
next section.

Let us formulate a lemma, that gives us a key tool and which implies theorem 6.2.

Lemma 8.1. Assume
\[ a_i > \delta \lambda_i, \quad b_i < (1 - \delta) \lambda_i \quad i = 1, \ldots, n. \]  
(8.1)

Let \( Q = (Q'(v), \ldots, Q_n(v)) \) be a \( C^1 \) smooth vector field on \( \Pi(a, b) \) and \( \delta > 0 \) verify
\[ -\delta < Q(v) < \delta, \quad v \in \Pi(a, b), \quad i = 1, \ldots, n. \]  
(8.2)

Then there are parameters \( P \) of the neural network such that the field \( F \) defined by (6.6)
satisfies the estimates
\[ \sup_{v \in \Pi(a, b)} |F(v, P) - Q(v)| < \epsilon, \]  
(8.3)

\[ \sup_{v \in \Pi(a, b)} |\nabla F(v, P) - \nabla Q(v)| < \epsilon. \]  
(8.4)

In other words, the fields \( F \) are dense in the vector space of all \( C^1 \) smooth vector fields satisfy-
ing to (8.2).

Proof. The proof uses the standard results of the multilayered network theory.

Step 1. The first preliminary step is as follows. Let us solve the system of equations
\[ \sigma(R_j) = Q_j(v) + \lambda_j v_j, \quad v \in \Pi(a, b) \]  
(8.5)

with unknown \( R_j \). Here \( R_j \) are the regulatory inputs of the sigmoidal functions. These equations
have a unique solution due to conditions (2.2), (8.1) and (8.2): the right hand sides
\( V_j + \lambda_j v_j \) range in \((0, 1)\). The solutions \( R_j(v) \) are \( C^1 \)-smooth vector fields.

Step 2. Consider relation (6.6). We choose entries \( A_j \) and \( B_j \) in a special way. First, let us
set \( A_j = 0 \) if \( i \notin S^j(j) \), where the set \( S^j(j) \) is defined in the \( n/V_j \)-centrality assumption, see
condition ii. Recall that \( S^j(j) \) is the set of the satellites acting on the center \( j \). Note that then
sum (6.6) can be rewritten as
\[ F_j(v, P) = \sigma \left( \sum_{i \in S^j(j)} \tilde{A}_j \tilde{A}_j^{-1} \sigma(B_j v - \tilde{h}_i) + D_j v - \tilde{h}_j \right) - \lambda_j v_j. \]  
(8.6)
Using the result of step 1 and this relation, we see that our problem is reduced to the following: to approximate $R_j(v)$ in $C^1$ norm with a small accuracy $O(\varepsilon)$ by

$$H_j(v, P) = \sum_{i \in S^*(j)} \tilde{A}_{ji} \lambda_{i-1} \sigma(B_i v - \tilde{h}_i) + D_i v - h_j. \quad (8.7)$$

Note that, according to the centrality assumption, the set $S^*(j)$ contains $N > CN^\theta$ elements. Moreover, due to this assumption, the sum $B_i = \sum_k B_{kji}$ involves all $k$, $k = 1, \ldots, n$. Therefore, since $n$ is fixed and $N$ can be taken arbitrarily large, the theorem on the universal approximation by multilayered perceptrons (see, for example, [6]) implies that the fields $H = (H_1, \ldots, H_n)$ are dense in the Banach space of all the vector fields on $\Pi(a, b)$ (with $C^1$-norm). Therefore, $H_j$ approximate $R_j$ with $O(\varepsilon)$-accuracy in $C^1$-norm. This finishes the proof. \hfill \Box

8.2. Proof of theorem 6.3

Ideas behind proof. Before stating a formal proof, we present a brief outline, which describes main ideas of the proof and the architecture of the switchable network. The network consists of two modules. The first module is a generating one and it is a centralized neural network with $n$ centers $v_1, \ldots, v_n$ and satellites $w_1, \ldots, w_n$. The second module consists of a center $v_{n+1} = z$ and $m$ satellites $\tilde{w}_1, \ldots, \tilde{w}_m$. The satellites from this module interact only with the module center $z$, i.e. in this module the interactions can be described by a distar graph. Only the center of the second module interacts with the neurons of the first (generating) module. We refer to the second module as a switching one. This architecture is shown on figure 2.

For the switching module the corresponding equations have the following form. Let us consider a distar interaction motif, where a node $z$ is connected in both directions with $m$ nodes $\tilde{w}_1, \ldots, \tilde{w}_m$. We set $n = 1$ and $N_1 = m$, $\lambda_1 = 1$, $D = 0$, $C = 0$, $\lambda_1 = 1$, and $A_{ij} = \kappa^{-1} \tilde{a}_{ij}$ in equations (4.1) and (4.2). By such notation the equations for the switching module can be rewritten in the form

$$\frac{d\tilde{w}_i}{dr} = \sigma(\tilde{b}_i z - \tilde{h}_i) - \kappa^{-1} \tilde{w}_i, \quad (8.8)$$

$$\frac{dz}{dr} = \sigma(\kappa^{-1} \sum_{j=1}^m \tilde{a}_{ij} \tilde{w}_j - h) - \xi \tilde{z}, \quad (8.9)$$

where $i = 1, \ldots, m$ and $\tilde{b}_i, \tilde{a}_{ij}, \tilde{\lambda}_i > 0$.

Under above assumptions on the network interactions, equations for generating module can be represented as follows:

$$\frac{dw_i}{dr} = \sigma(B_i v + C_i w - d_i z - \tilde{h}_i) - \kappa^{-1} \tilde{\lambda}_i w_i, \quad (8.10)$$

$$\frac{dv}{dr} = \sigma(A_{ij} w + D_{ij} v - d_j z - h_j) - \lambda_{ij} v_j, \quad (8.11)$$

where $i = 1, \ldots, N$, $j = 1, \ldots, m$ and $d_i, \tilde{a}_i$ are coefficients.

These equations involve $z$ as a parameter. This fact can be used in such a way. Consider the system of the differential equations

$$\frac{dv}{dr} = Q(v, z), \quad v = (v_1, \ldots, v_n) \quad (8.12)$$
where $z$ is a real control parameter. Let $z_1, ..., z_{m+1}$ be some values of this parameter. We find a vector field $Q$ such that for $z = z_l$, where $l = 1, ..., m$, the dynamics defined by (8.12) has the prescribed structurally stable invariant sets $\Gamma_l$. Furthermore, according to theorem 6.2, for each positive $\epsilon$ we can choose the parameters $\tilde{\lambda}, \tilde{\nu}, \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta}, \tilde{\epsilon}, \tilde{\zeta}, \tilde{\xi}$ of the system (8.10) and (8.11) such that the dynamics of this system will have structurally stable invariant sets $\tilde{\Gamma}_l$ topologically equivalent to $\Gamma_l$.

For the switching module we adjust the center-satellite interactions and the center response time parameter $\xi$ in such a way that for a set of values $\xi$ the switching module has the dynamics of system (8.8) and (8.9) with $m$ different stable hyperbolic equilibria $z = z_1, z_2, ..., z_{m+1}$ and for sufficiently large $\xi$ system (8.8) and (8.9) has a single equilibrium close to $z_1 = 0$.

Existence of such a choice will be shown in coming lemma 8.2. Then the both modules form a network having need dynamical properties formulated in the assertion of theorem 6.3.

**Proof.** Let us formulate some auxiliary assertions. First we consider the switching module.

**Lemma 8.2.** Let $m$ be a positive integer and $\beta \in (0, 1)$. For sufficiently small $\kappa > 0$ there exist $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta}, \tilde{\epsilon}, \tilde{\zeta}, \tilde{\xi}$ such that

i. for an open interval of values $\xi$ system (8.8) and (8.9) has $m$ stable hyperbolic rest points $z_j \in (j - 1 + \beta, j + \beta)$, where $j = 1, ..., m$;

ii. for $\xi > \xi_0 > 0$ system (8.8) and (8.9) has a single stable hyperbolic rest point.
Proof. Let $h = 0$. To find equilibria $z$, we set $d\tilde{w}/dt = 0$, and express $\tilde{w}$ via $z$. Then we obtain the following equation for the rest points $z$:

$$
\xi_z = \sigma \left( \sum_{j=1}^{m} a_{j} \sigma (\tilde{b} z - \tilde{b}_j) \right).
$$

(8.13)

For especially adjusted parameters equation (8.13) has at least $m$ solutions, which give stable equilibria of system (8.8) and (8.9). To show it, we assume that $0 < \kappa \ll 1$, $\tilde{b}_j = \tilde{b} = \kappa^{-1/2}$ and $\tilde{b}_j = \kappa_0$. We obtain then

$$
V(\xi_z) = \sum_{j=1}^{m} \sigma (\tilde{b}_j z - \mu_j) + O(\kappa) = F_m(z, \beta, \kappa),
$$

(8.14)

where $V(z)$ is a function inverse to $\sigma(z)$ defined on $(0, 1)$. For sufficiently large $\kappa$, the plot of the function $F_m$ is close to a stairway (see figure 3). Let

$$
\xi = 1, \quad \tilde{a}_1 = V(\mu_1) + \kappa, \quad \tilde{a}_j = V(\mu_j) - V(\mu_{j-1}), j = 2, ..., m.
$$

The intersections of the curve $V(z)$ with the almost horizontal pieces of the plot of $F_m$ give us $m$ stable equilibria of system (8.8) and (8.9). These equilibria $z_j$ lie in the corresponding intervals $(j-1 + \beta, j + \beta)$. For sufficiently large $\xi$ we have a single rest stable point $z$ at 0. The lemma is proved.

□

Consider compact invariant hyperbolic sets $\Gamma_1, ..., \Gamma_m$ of semiflows defined by arbitrarily chosen $C^1$ smooth vector fields $Q^{(l)}$ on the unit ball $B^n \subset \mathbb{R^n}$, where $l = 1, ..., m$.

Lemma 8.3. Let $\Pi(a, b)$ be a box in $\mathbb{R}^n$ and $m > 1$ be a positive integer. There is a $C^1$-smooth vector field $Q$ on $\Pi(a, b) \times [0, m + 1]$ such that equation (5.1) defines a semiflow having hyperbolic sets $\Gamma_1, ..., \Gamma_m$ and the restriction of this field on $\Pi(a, b) \times [0, 1]$ has an attractor consisting of a single hyperbolic rest point.

Proof. The proof uses the following idea. For $k \in \{2, ..., m + 1\}$ let $Q^{(k)}(v)$ be a vector field on $\Pi(a, b)$ having $\Gamma_{k-1}$ as an invariant compact hyperbolic set. Moreover, suppose that $Q^{(1)}$ has a single globally attracting rest point in $\Pi(a, b)$, $z_j \in (j-1 + \beta, j + \beta)$, where $j = 1, ..., m$ and $\beta \in (0, 1)$. Let $\chi_l(z)$ be smooth functions of $z \in \mathbb{R}$ such that

$$
\chi_l(z_i) = \delta_{lk}, \quad l \in \{1, ..., m\}, \quad k = 1, ..., m
$$

where $\delta_{lk}$ stands for the Kronecker delta. Let $Q(v, z)$ be the vector field on $\Pi(a, b) \times [0, m + 1]$ defined by

$$
Q(v, z) = \sum_{k=1}^{m} Q^{(k)}_l \chi_l(z), \quad l \in \{1, ..., n\},
$$

(8.15)

for first $n$ components and $n + 1$ th component of this field (denoted by $z$) is defined by

$$
Q_{n+1}(v, z) = F_m(z, \beta, \kappa),
$$

(8.16)

where $F_m$ is defined by (8.14). For $\beta \in (0, 1)$ the function $F_m$ has stable roots at the points $z = 1, 2, ..., m$. We observe that the equation for $z$-component $dz/dt = F_m(z, \beta, \kappa)$ does not involve $v$. By applying lemma 8.2 we note that solutions $z(t, z(0))$ of the Cauchy problem for this differential equation verify $|z(t) - z_i| < \exp(-\epsilon t)$, if $z(0)$ lies in an open neighbourhood of $z_i$. To conclude the proof, we consider the system
The right hand sides of this system define the field \( Q \) of dimension \( n + 1 \) from the assertion of lemma 8.3. To check this fact, we apply lemmas 8.1 and 8.2 that completes the proof.

Next, to finish the proof of theorem 6.3, let us take a box \( \Pi(a, b) \), where \( 0 < a_i < b_i \). The semiflows defined by differential equations \( \frac{dv_i}{dt} = Q(v, z), \ i = 1, ..., n \), \( \frac{dz}{dt} = F_m(z, \beta, \kappa) - \xi \bar{\lambda} z = Q_{n+1}(z) \).

The right hand sides of this system define the field \( Q \) of dimension \( n + 1 \) from the assertion of lemma 8.3. To check this fact, we apply lemmas 8.1 and 8.2 that completes the proof.

Next, to finish the proof of theorem 6.3, let us take a box \( \Pi(a, b) \), where \( 0 < a_i < b_i \). The semiflows defined by differential equations \( \frac{dv_i}{dt} = \delta Q(v) \) are orbitally topologically equivalent for all \( \delta > 0 \). We approximate the first \( n \) components of the field \( Q \) by our neural network using lemma 8.3. We multiply here \( Q \) on an appropriate positive \( \delta \) to have a field with components bounded by sufficiently small number in order to apply lemma 8.1. Namely, we take \( \delta \) such that \( a_i > \delta/(\xi_i \bar{\lambda}_i) \) and \( b_i < (1-\delta)/(\xi_i \bar{\lambda}_i) \) and apply lemma 8.1. Note that this approximation does not involve the control parameter \( \xi \). Indeed, this parameter is involved only in the approximation of \( Q_{n+1} \), which can be done independently, see the distar graph lemma 8.2. This concludes the proof of theorem 6.3.

**Remark 8.4.** In theorem 6.3, we assume that the vector field \( Q(v) \) is given. However, by centralized networks we can solve the problem of identification of dynamical systems supposing that the trajectories \( v(t) \) are given on a sufficiently large time interval whereas \( Q \) is unknown or we know this field only up to unknown parameters. An example, where we consider an identification construction for a modified noisy Lorenz system, can be found in section 9.
8.3. Proof of theorem 6.4

Let us refer to the distar centers as hubs and to periphery nodes as satellites. We suppose that satellites do not interact each with others and a satellite interacts only with the corresponding hub. Therefore the interaction graph resulting from the ‘hub disconnecting’ construction consists of \( n \) disconnected distar motifs.

**Step 1.** Let \( n = 1 \). We apply lemma 8.2 to the distar graphs, see the proof of the previous theorem. Then we have \( m_1 \) stable equilibria, where \( m_1 \) is the number of satellites in the distar motif.

**Step 2.** In the case \( n > 1 \) we consider the disconnected interaction graph consisting of \( n \) distar motifs, where the \( j \)th distar motif contains \( m_j \) nodes. One has \( m_1 + m_2 + \ldots + m_n = N - n \) and totally the graph consists of \( N \) nodes. For each distar we adjust the parameters as above (see step 1). We obtain thus \( m_1m_2\ldots m_n \) of equilibria and the theorem is proven.

9. Algorithm of construction of switchable network with prescribed dynamics

The proof of theorem 6.3 can be used to construct practically feasible algorithms, which solve the problem of construction of a switchable network with prescribed dynamical properties. As a matter of fact, we can address two different, but related problems. The first problem is the synthesis of a neural network with prescribed attractors and switchability properties. The second problem is the identification of a neural network from time series. First we state the solution of the first problem and after we describe how to resolve the second one by analogous methods.

The prescribed network properties for the synthesis problem are stated in theorem 6.3. We describe here a step by step algorithm, allowing to construct a network with these properties.

Consider structurally stable dynamical systems defined by the equations

\[
\frac{dv}{dt} = Q^{(l)}(v) \quad v = (v_1, \ldots, v_n) \in \Pi(a, b) \subset \mathbb{R}^n,
\]

where \( l = 1, \ldots, m \) and \( \Pi(a, b) \) is a defined by (6.1). We suppose that the fields \( Q^{(l)}(v) \) are sufficiently smooth, for example, \( Q^{(l)} \in C^\infty(\Pi(a, b)) \). Without any loss of generality we can assume that

\[
1 < a_i < b_i,
\]

(9.2)

(otherwise we can shift variables \( v_i \) setting \( v_i = \tilde{v}_i - c_i \)).

**Step 1.** Find a sufficiently small \( \varepsilon \) such that perturbations of vector fields \( Q(v)^{(l)} \), which are \( \varepsilon \) small in \( C^1 \) norm, do not change topologies of semiflows defined by 9.1. Actually, it is hard to compute such a value of \( \varepsilon \), so, in practice we simply choose a small \( \varepsilon \) by the trial and error method.

**Step 2.** We find a vector field \( Q(v,z) \) with \( n + 1 \) components, where \( z = v_{n+1} \in [a_{n+1}, b_{n+1}] \subset \mathbb{R} \) such that the first \( n \) components of \( Q(v,z) \) are defined by relations (8.15) and the \( n + 1 \) component is defined by (8.16). Let \( D = \Pi(a, b) \times [a_{n+1}, b_{n+1}] \).

To describe the next steps, first let us introduce the functions

\[
G_j(v, P) = \sum_{i=1}^{N} \tilde{A}_{ij} \sigma(B_i v - h_i),
\]

(9.3)

where the parameter \( P = \{N, \tilde{A}_j, B_k, h_{ij}, j = 1, \ldots, n + 1, i, k = 1, \ldots, N\} \) and \( v = (v_1, \ldots, v_z, z) \).

Let us observe that dynamical systems \( dq/dt = Q(q) \) and \( dq/dt = \gamma Q(q) \) with \( \gamma > 0 \) have the same trajectories, invariant sets and attractors, therefore, instead of \( Q \) we can use \( \gamma Q \). We choose a \( \gamma > 0 \) and a small positive \( \delta < 1 \) such that
\( -\delta < \gamma^Q(\vec{v}) < \delta, \quad \vec{v} \in D_i, \quad i = 1, \ldots, n + 1 \) \hspace{1cm} (9.4)

and

\( a_i > \delta/\lambda_i, \quad b_i < (1 - \delta)/\lambda_i \quad i = 1, \ldots, n + 1 \) \hspace{1cm} (9.5)

for \( \lambda_i > 1 \).

Then (9.4) and (9.5) imply that

\[ 0 < \gamma^Q(\vec{v}) + \lambda_j \vec{v}_j < 1, \quad \vec{v} \in D, \quad j = 1, \ldots, n + 1. \] \hspace{1cm} (9.6)

Let \( \sigma^{-1} \) be the function inverse to \( \sigma \). Due to (9.6) the functions

\[ R_j(\vec{v}) = \sigma^{-1}(\gamma^Q(\vec{v}) + \lambda_j \vec{v}_j) \] \hspace{1cm} (9.7)

are correctly defined and smooth on \( D \).

Now we solve the following approximation problem.

To find the number \( N \), the matrices \( \bar{A}, \bar{B} \), and vector \( h \) such that

\[ |R_j(\vec{v}) - G_j(\vec{v}, \mathbf{P})| + |D_{\epsilon}(R_j(\vec{v}) - G_j(\vec{v}, \mathbf{P}))| \leq \epsilon/2, \quad j = 1, \ldots, n + 1. \] \hspace{1cm} (9.8)

This problem can be resolved by standard algorithms, which perform approximations of functions by multilayered perceptrons [6]. Note that these standard methods are based on iteration procedures, which can use a large running time.

We describe here a new variant of the algorithm for this approximation problem, which uses a wavelet-like approach. This approach does not exploit any iteration procedures or linear system solving. All the procedure reduces to a computation of the Fourier and wavelet coefficients. However, this algorithm is numerically effective only for sufficiently smooth \( R_j \) with fast decreasing Fourier coefficients and for not too large dimensions \( n \).

The solution of the approximation problem (9.8) proceeds in the two steps.

**Step 3.** We reduce the \( n + 1 \)-dimensional problem (9.8) to a set of one-dimensional ones as follows. Let us approximate the functions \( R_j \) by the Fourier expansion:

\[ \sup_{\vec{v} \in D} |R_j(\vec{v}) - \hat{R}_j(\vec{v})| + |\nabla_\vec{v}(R_j(\vec{v}) - \hat{R}_j(\vec{v}))| < \epsilon/4, \] \hspace{1cm} (9.9)

where

\[ \hat{R}_j(\vec{v}) = \sum_{k \in K_0} \hat{R}_j(k) \exp(ik(\vec{v}, \vec{v})), \] \hspace{1cm} (9.10)

\( (k, \vec{v}) = k_1 \vec{v}_1 + k_2 \vec{v}_2 + \ldots + k_{n+1} \vec{v}_{n+1} \) and the set \( K_0 \) of vectors \( k \) is a finite subset of the \( (n + 1) \)-dimensional lattice \( L_0 \)

\[ K_0 \subset L_0 = \{ k = (k_1, \ldots, k_{n+1}) : k_i = (a_i - b_i)^{-1} \pi m_i, \text{ for some } m_i \in \mathbb{Z} \}. \] \hspace{1cm} (9.11)

The Fourier coefficients \( \hat{R}_j(k) \) can be computed by

\[ \hat{R}_j(k) = (\text{volume}(D))^{-1} \int_D R_j(\vec{v}) \exp(-ik(\vec{v}, \vec{v}))d\vec{v}. \]

In order to satisfy (9.9), we take a sequence of extending sets \( K_0 \). For some \( K_0 \) relation (9.9) will be satisfied because the Fourier coefficients \( \hat{R}_j(k) \) fastly decrease in \( |k| \).
Step 4. We exploit the fact that the problem (9.8) is linear with respect to the coefficients \( \bar{A}_{ij} \). For each \( k \in K_D \) we resolve the following one-dimensional problem. Let

\[
g(q, M, a, \beta, \tilde{h}) = \sum_{i=1}^{M} a_i \sigma(\beta_i(q - \tilde{h}_i)). \tag{9.12}
\]

We are seeking for integer \( M > 0 \) and the vectors \( a = (a_1, ..., a_M) \), \( \beta = (\beta_1, ..., \beta_M) \) and \( \tilde{h} = (\tilde{h}_1, ..., \tilde{h}_M) \) such that

\[
\sup_{q \in k} |W_{j,k}(q) - g(q, M, a, \beta, \tilde{h})| < \epsilon (10|K_D|)^{-1}, \tag{9.13}
\]

\[
\sup_{q \in k} |dW_{j,k}(q) dq - g'(q, M, a, \beta, \tilde{h})| < \epsilon_1 \leq \epsilon (10|K_D|)^{-1}, \tag{9.14}
\]

where \( |K_D| \) is the number of the elements \( k \) in the set \( K_D \).

To resolve these one-dimensional approximation problems, we apply a method based on the wavelet theory. Notice that this method is numerically effective. First we observe that if (9.14) is fulfilled with a sufficiently small \( \epsilon_1 \), then, to satisfy (9.13), it is sufficient to add a constant term of the form \( \sum a_i \sigma(\beta_i) \) with \( b_{M+1} = 0 \) to the sum in the right hand side of (9.12).

Let us define the function \( \psi \) by

\[
\psi(q) = \sigma'(q) - \sigma'(q - 1). \tag{9.16}
\]

We observe that

\[
\int_{-\infty}^{\infty} \psi(q) dq = 0 \tag{9.17}
\]

and \( \psi(q) \to 0 \) as \( |q| \to \infty \), therefore, \( \psi \) is a wavelet-like function.

Let us introduce the following family of functions indexed by the real parameters \( r, h \):

\[
\psi_{r,h}(q) = |r|^{-1/2} \psi(r^{-1}(q - \xi)). \tag{9.18}
\]

For any \( f \in L_2(\mathbb{R}) \) we define the wavelet coefficients \( T_f(r, \xi) \) of the function \( f \) by

\[
T_f(r, \xi) = \langle f, \psi_{r,h} \rangle = \int_{-\infty}^{\infty} df(q) \psi_{r,h}(q). \tag{9.19}
\]

For any smooth function \( f \) with a finite support \( I_R = (-R, R) \) one has the following fundamental relation:

\[
f = \sum_{r=0}^{\infty} \int_{-\infty}^{\infty} r^{-2} dr \psi_{r,h} T_f(r, \xi) \psi_{r,h} = f_{\text{wav}}. \tag{9.20}
\]
for some constant $c_\psi$. This equality holds in a weak sense: the left hand side and the right hand side define the same linear functionals on $L_2(\mathbb{R})$, i.e. for each smooth, well localized $g$ one has

$$\langle f, g \rangle = \langle f_{\text{wav}}, g \rangle.$$  

Let $\delta(\epsilon) \ll \epsilon$ be a small positive number. According to (9.20) we can find positive integers $p_1, p_2$, points $r_1, \ldots, r_{p_2}$, $\xi_1, \ldots, \xi_{p_2}$, and a constant $\bar{c}_\psi$ such that the integral in the right hand side of (9.20) can be approximated by a finite sum:

$$\sup | f(q) - \tilde{f}_{\text{wav}}(q) | < \delta,$$  

where

$$\tilde{f}_{\text{wav}} = \bar{c}_\psi \sum_{l=1}^{p_1} \sum_{l=1}^{p_2} r_l^{-2} T_l(r_l, \xi_l) \psi_{l} \xi_l.$$  

In our case for each $(j, k)$ we set $f = W_j, k(q)$ for $q \in I_k$ and $f = 0$ for $q \notin I_k$. We can take $r_k = r_l l / p_l$, where $r_l$ is large enough, and $\xi_k = q_{\min} + (q_{\max} - q_{\min}) l / p_2$, where $q_{\min} < q_{\max}$. We can renumerate the points $(r_l, \xi_l)$ by a single index $l = 1, \ldots, p$, where $p = p_1 p_2$, that gives us $r_l, \xi_l$ and the wavelet coefficients $T_l = \bar{c}_\psi T_l(r_l, \xi_l),$.

Having $p$, $r_l$, $\xi_l$ and the wavelet coefficients $T_l$, we obtain the following solution of the approximation problem (9.12):

$$M(j, k) = p, \quad \beta_{2j-1}(j, k) = r_l^{-1} \xi_l, \quad \hat{h}_{2j}(j, k) = r_l^{-1}(\xi_l + 1),$$  

$$\beta_{2j-1}(j, k) = \beta_{2j}(j, k) = r_l^{-1}, \quad \alpha_{2j-1}(j, k) = -\alpha_{2j}(j, k) = T_l,$$

where we have introduced the index $(j, k)$ in notation for the solution $(M, a, \beta, \tilde{h})$ to emphasize that problem (9.12) depends on this index.

Finally, in the end of this step we obtain the coefficients

$$M(j, k), a_l(j, k), \ldots, a_{M(j,k)}(j, k), \beta_l(j, k), \ldots, \beta_{M(j,k)}(j, k), \hat{h}(j, k), \ldots, \hat{h}_{M(j,k)}(j, k).$$  

(9.22)

**Step 5.** We construct a network with $n + 1$ centers $v_1, \ldots, v_{n+1}$ and $N$ satellites as follows. Let $C = 0$ and $D = 0$, i.e. we assume that the satellites don’t interact among themselves and there are no direct interactions between the centers. The number of satellites is defined by

$$N = \sum_{j=1}^{n+1} \sum_{k \in K_0} M(j, k).$$

Each satellite can be equipped with a triple index $(i, j, k)$, where $j = 1, \ldots, n + 1$, $k \in K_0$ and $i \in \{1, \ldots, M(j, k)\}$. We set that all $h_j = 0$, $\lambda_y = 1$, and $\lambda_j$ are chosen as above. The threshold $h_{i, j, k}$ for the satellite with the index $(i, j, k)$ is defined by

$$h_{i, j, k} = \hat{h}(i, j, k),$$

where $\hat{h}(i, j, k)$ are obtained at the step 4 (see (9.22)).

Furthermore, we define the matrices $\tilde{A}$ and $B$ as follows. One has

$$B_{\beta_{i,j,k}}(i, j, k) = \beta(j, k) k_1,$$

(this relation describes an action of the $i$th center on the satellite with index $(i, j, k)$) and

$$\tilde{A}_{\beta_{i,j,k}}(i, j, k) = a(j, k)$$
(this relation describes an action of the \( l \)th center on the satellite with index \((i, j, k)\)). Here \( i \in \{1, ..., M(j, k)\}, j, l = 1, ..., n + 1 \) and \( k \in K_D \).

**Remark.** This algorithm can be simplified if instead networks (4.1) and (4.2) we use analogous networks where satellites act on centers in a linear way:

\[
\frac{dw_i}{dt} = \sigma(B_iw + C_iw - h_i) - \kappa^{-1}w_i, \tag{9.23}
\]

\[
\frac{dv_j}{dt} = (A_jw - h_j) - \lambda_jv_j, \tag{9.24}
\]

where \( i = 1, ..., N_i, j = 1, ..., n, \) and the fields \( Q^{(l)} \) are defined by polynomials (note that Jackson’s theorems [2] guarantee that any \( Q \) can be approximated by a polynomial field on \( \Pi(a, b) \) in \( C^{1}\)-norm). Then we can simplify Step 3 and Step 4 of the algorithm as follows. We observe that we can set \( \gamma = 1 \) and in this case the functions \( R_j \) have the form

\[
R_j(\tilde{v}) = Q_j(\tilde{v}) + \lambda_j\tilde{v}_j. \tag{9.25}
\]

On step 3 for polynomial functions \( R_k(v) \) we can also use simple algebraic transformations, instead of the Fourier decomposition, to reduce the multidimensional approximation problem to one dimensional ones. On step 4 the function \( \psi \) defined by (9.16) is well localized and therefore alternatively step 4 can be realized by standard programs using radial basic functions and the method of least squares (see an example on the Lorenz system below).

Let us turn now to the problem of identification of a neural network from time series produced by a dynamical system \( dv/dt = Q(v, P) \), \( v \in \mathbb{R}^d \) with unknown parameters \( P \). Assume that we observe a time series \( v(t_1), v(t_2), ..., v(t_k) \) and the time interval between observations is small: \( t_{i+1} - t_i = \Delta t \ll 1 \). We want to construct a network with \( n \) centers, which produces, in a sense, analogous time series. According to (3.4), a suitable criterion of trajectory similarity is as follows. We can approximate the averages \( S_{Q, \varphi} \) from (3.3) by the time series

\[
S_{Q, P, \varphi} \approx K^{-1}\Delta T \sum_{k=1}^{K} \phi(v(t_k)) = S_{Q, P, \varphi}^{(K)}. \tag{9.26}
\]

Then, if the network identification is correct, the averages defined by time series and the corresponding ones generated by the approximating centralized neural network, should be close for smooth weight functions \( \phi \):

\[
|S_{Q, P, \varphi}^{(K)} - S_{Q, \varphi}^{(K)}| = Err_{\text{approx}} < \delta(\phi) \ll 1, \tag{9.27}
\]

where \( G_{\text{approx}} \) is the approximation of \( Q \) by the neural network.

As a first step, we can approximate the unknown field \( Q(v) \) by finite differences, for example, using the relation

\[
Q(\tilde{v}, P) = (v(t_{i+1}) - v(t_i))\Delta t^{-1}, \quad \tilde{v}_i = (v(t_{i+1}) + v(t_i))/2. \tag{9.28}
\]

For other values \( v \) the field \( Q \) can be reconstructed, for example, by a linear interpolation. The neural network approximation of \( Q \) can be obtained by applying the steps 2–5 of the synthesis algorithm described above.

We end this section with an illustration of the simplified variant of the identification and synthesis algorithm, see the preceding remark.

As an example, we describe a solution of the following identification problem. Consider time series generated by the Lorenz system perturbed by noise. The Lorenz system involves a controller parameter. Adjusting the values of this parameter, we can obtain chaotic dynamics,
time periodic one or dynamics with convergent trajectories. We are going to find a centralized network, which also has a controller parameter and can generate all this rich variety of trajectories. For chaotic and periodic trajectories this neural approximation should exhibit dynamics with analogous ergodic properties (in the sense of (9.27)).

Recall that the Lorenz system has the form

\[
\begin{align*}
\dot{x} &= \alpha(y - x), \\
\dot{y} &= x(\rho - z) - y, \\
\dot{z} &= xy - \beta z.
\end{align*}
\]  

(9.29)

This system shows a chaotic behaviour for \(\alpha = 10, \beta = 8/3\) and \(\rho = 28\) and \(\rho \in (0, 1)\) this system has a globally attracting rest point.

We introduce new variables \(v_1 = x, v_2 = y, v_3 = z\) and \(v_4 = \rho\) and consider a more complicated modified Lorenz system with a controller parameter: (compare with the proof of theorem 6.3):

\[
\begin{align*}
\dot{v}_1 &= \alpha(v_2 - v_1) - f_1, \\
\dot{v}_2 &= \rho v_3(v_2 - v_3) - r_2 v_2 - f_2, \\
\dot{v}_3 &= r_1 v_1 v_2 - \beta z = f_3, \\
\dot{v}_4 &= \varpi(v_4, b_0, h_0) - \xi v_4 = f_4,
\end{align*}
\]

(9.30)

where \(\varpi\) is a regularized step function defined by \(H(w) = (1 + \exp(-b_0(w - h_0)))^{-1}\) with \(b_0 \gg 1\) and \(h_0 = 1\). We set \(\xi = 0.5, r_1 = 14, r_2 = 1, r_3 = 1\). The initial data for the fourth component \(v_4 = v(0)\) is a controller parameter. For large \(b_0\) the differential equation for \(v_4\) has two stable equilibria: \(v_4 \approx 0\) and \(v_4 \approx 2\). Therefore, for \(v_0 \in (0, 1)\) system (9.30) and (9.31) has a globally attracting rest point and for \(v_0 > 1\) the attractor of this system is chaotic Lorenz one. The parameters of this system are \(P = (\alpha, \beta, r_1, r_2, r_3)\).

Suppose we observe trajectories \(v(t), t \in [0, T]\) of system (9.30) at some time moments \(t_0 = 0, h = dt, ..., t_p = p\Delta t\). In order to simulate experimental errors we have perturbed the system with additive noise. We are going to find a centralized network, which has an attractor with, in a sense, similar statistical characteristics. More precisely, we aim to minimize \(\text{Err}_{\text{approx}}\) from relation (9.27). For identification procedure we use a centralized network with 4 centers \(v_1, v_2, v_3\) and \(v_4\). In this case steps 3, 4 can be simplified if we use this specific form of the modified Lorenz system. The last center \(v_4\) serves as a controller.

We state the algorithm for the modified Lorenz system, however, the method is general and feasible for identification by trajectories generated by all low-dimensional dynamical systems defined by polynomial vector fields.

First we set

\[C = D = 0.\]  

(9.32)

This means that only satellites act on centers and vice versa. To find the matrices \(A, B\) and the thresholds \(h_i\), we solve the following approximation problems:

\[
R(A, B, h) \rightarrow \min, \quad R = \sum_{i=1}^{N} \sum_{j=1}^{p} (Q_i(t_j) - S_i(v(t_j, A, B, h)))^2
\]

(9.33)

where

\[
Q_i(t_j) = (v(t_j + \Delta t) - v(t_j))/\Delta t, \quad S_i(v, A, B, h) = \sum_{k=1}^{N} A_{ik} \sigma\left(\sum_{j=1}^{p} B_{kj} v_j - h_k\right).
\]

(9.34)

This approximation problem is nonlinear with respect to \(B\) and \(h\). We can simplify this problem by the following heuristic method. Each function \(f_j(v)\) defined on a open bounded domain can be represented as a linear combination of functions \(g_f(v \cdot k_i)\), where vectors \(k_i\).
belong to a finite set of vectors $K_i$. For example, for system (9.30) and (9.31) the components $f_j$ for $j = 1, 2, 3$ can be represented as linear combinations of monomials:

$$f_j(v) = g_j(v) - \lambda_j v_j, \quad g_j(v) = \sum_{l=1}^{11} C(j,l) T_l(v)$$  \hspace{1cm} (9.35)

where

$$T_l = v_l, \quad l = 1, 2, 3, 4$$

$$T_{2l+1} = (v_1 + v_l)^2, \quad T_{2l+2} = (v_1 - v_l)^2, \quad l = 2, 3, 4, \quad T_{14} = 1,$$

and $\lambda_1 = \alpha, \quad \lambda_2 = 1, \quad \lambda_3 = \beta$. Therefore, $K_1 = \{k_{11} = (1, 0, 0, 0)\}, \quad K_2 = \{k_{12} = (1, 0, 1, 0), \quad k_{22} = (1, 0, -1, 0), \quad k_{32} = (1, 0, 0, 1), \quad k_{42} = (1, 0, 0, -1)\}, \quad K_3 = \{k_{13} = (1, 1, 0, 0), \quad k_{33} = (1, -1, 0, 0)\}, \quad K_4 = \{k_{14} = (1, 0, 0, 0)\}$. Let $n_i$ be the number of the vectors contained in the set $K_i$, $n_1 = 1, n_2 = 4, \quad n_3 = 2$ and $n_4 = 1$. In this case of the modified Lorenz system, the set $K_D$ from (9.11) is the union of sets $K_i, i = 1, ... , 4$.

We take a sufficiently large $N_L$, a large $b_0$ and define the auxiliary thresholds $\bar{h}_{k_0,j}$, where $j = 1, ..., N_L$, by

$$\bar{h}_{k_0,j} = \min_{x = 1, ..., p, \in K_i} v(t_x) \cdot k_b + j \left(\max_{x = 1, ..., p, \in K_i} v(t_x) \cdot k_b - \min_{x = 1, ..., p, \in K_i} v(t_x) \cdot k_b\right)/N_L.$$

We seek coefficients $\bar{A}_{i,k_b}$ and $C_i$, which minimize $R_i(\bar{A}, C_i)$ for $i = 1, 2, 3, 4$:

$$R_i(\bar{A}, C_i) \rightarrow \min, \quad R_i = \sum_{j=1}^{p} (Q(t_j) - \bar{S}_i(v(t_j), \bar{A}, C_i))^2$$  \hspace{1cm} (9.36)

where

$$\bar{S}_i(v, \bar{A}, C) = C_i + \sum_{l=1}^{n_i} \sum_{j=1}^{N_L} \bar{A}_{i,l} \sigma(b_0(v_l) \cdot v - \bar{h}_{k_0,j})).$$  \hspace{1cm} (9.37)

Note that since $\bar{S}_i$ are linear functions of $\bar{A}_{i,k_b}$ and $C_i$, problems (9.36) can be solved by the least square method. The important advantage of this approach is that approximations can be done independently for different components $i$.

This approximation produces a centralized network involving 4 centers and $N = 8 N_L + 8$ satellites. Indeed, each vector $k_b$ associated with a quadratic term $T_l$, gives us $N_L$ satellites to approximate this term. Moreover, we use 4 satellites for approximations of the linear terms and 4 satellites are necessary for constants $C_i$ in the right hand sides of (9.37).

The numerical simulations give the following results. The trajectories to identify are produced by the Euler method applied to the system (9.30) and (9.31) perturbed by noise, where the time step 0.005 on the interval [0, 50], the noise is simulated by $\epsilon(t)\omega(t)$, where $\omega(t)$ is the standard white noise and $\epsilon = 0.05$. As a result of minimization procedure, we have obtained the errors $R_i$ of the order 0.01–0.1. The trajectories of the system (9.30) and (9.31) perturbed by noise and the corresponding neural networks are not close but they have a similar form and statistical characteristics that is confirmed by the value $Err_{approx}$ (defined by (9.27)), which is 0.008, where the test function $\phi$ is $\phi(v) = v_1^2 + v_2^2/2 - 2v_3$. These results are illustrated by figure 4.
10. Conclusion and discussion

In this paper, we have proposed a complete analytic theory of maximally flexible and switchable Hopfield networks. We shown that dynamics of a network with $n$ slow components $v_1, \ldots, v_n$ can be reduced to a system of $n$ differential equations defined by a smooth $n$ dimensional vector field $F(v)$. If these slow components are hubs, i.e. they are connected with a number of other weakly connected nodes (satellites) and center-satellite interactions dominate inter-satellite forces, then the network becomes maximally flexible. Namely, by adjusting only center-satellite interactions we can obtain smooth $F$ of arbitrary forms.

These networks are also maximally switchable. We describe networks of a special architecture, which contains a controller hub. By changing the state of this hub and the hub response time parameter $\xi$ one can completely change the network dynamics from an unique global attractive steady state to any combination of periodic or chaotic attractors.

Our results provide a rigorous framework for the idea that centralized networks are flexible. We also propose mechanisms for switching between attractors of these networks with controller hubs. In functional genomics there are numerous examples when transitions between attractors of gene regulatory networks can be triggered by controller proteins having multiple states sometimes resulting from interactions with micro-RNA satellites [9]. Similarly, neurons having multiple internal states can trigger phase transitions of brain networks suggesting that single neuron activation could be used for neural network control [16].

The proofs of our results are constructive and are based on an algorithm allowing the network reconstruction. This algorithm has several potential applications in biology. Identified
networks can be used to study emergent network properties such as robustness, controllability and switchability. Gene networks with the desired switchability properties could be build by synthetic biology tools for various applications in biotechnology. Furthermore, maximal switchable network models can be used in neuroscience to relate structure and function in the brain activity, or in genetics to explain how a minimal number of mutations can induce large phenotypic changes from one type of adaptive behavior to another one.

Acknowledgments

SV was financially supported by Government of Russian Federation, Grant 074-U01, also supported in part by grant RO1 OD010936 (formerly RR07801) from the US NIH and by grant 16-01-0048 of Russian Fund of Basic Research. OR was supported by the Labex EPIGENMED (ANR-10-LABX-12-01). The authors are grateful to the anonymous referees for their useful remarks, that helped improve the text.

References

[1] Menche J, Valleriani A and Lipowsky R 2010 Dynamical processes on dissortative scale-free networks Europhys. Lett. 89 18002
[2] Achieser N I 2013 Theory of Approximation (New York: Dover)
[3] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks Rev. Mod. Phys. 74 47
[4] Albert R, Jeong H and Barabási S-L 2000 Error and attack tolerance of complex networks Nature 406 378–82
[5] Bar-Yam Y and Epstein I R 2004 Response of complex networks to stimuli Proc. Natl Acad. Sci. USA 101 4341–5
[6] Barron A R 1993 Universal approximation bounds for superpositions of a sigmoidal function IEEE Trans. Inf. Theory 39 930–45
[7] Bascompte J 2007 Networks in ecology Basic Appl. Ecol. 8 485–90
[8] Carlson J M and Doyle J 2002 Complexity and robustness Proc. Natl Acad. Sci. 99 2538–45
[9] Carthew R W 2006 Gene regulation by micrornas Curr. Opin. Gen. Dev. 16 203–8
[10] Chialvo D R 2010 Emergent complex neural dynamics Phys. Rev. Lett. 86 3682
[11] Cohen R, Erez K, Ben-Avraham D and Havlin S 2001 Breakdown of the internet under intentional attack Phys. Rev. Lett. 86 3682
[12] Cornelius S P, Kath W L and Motter A E 2013 Realistic control of network dynamics Nat. Commun. 4 1942–6
[13] Cowan N J, Chastain E J, Vilhena D A, Freudenberg J S and Bergstrom C T 2012 Nodal dynamics, not degree distributions, determine the structural controllability of complex networks PloS One 7 e38398
[14] Deco G and Jirsa V K 2012 Ongoing cortical activity at rest: criticality, multistability, and ghost attractors J. Neurosci. 32 3366–75
[15] Edwards R, Beuter A and Glass L 1999 Parkinsonian tremor and simplification in network dynamics Bull. Math. Biol. 61 157–77
[16] Fujisawa S, Matsuki N and Iegaya Y 2006 Single neurons can induce phase transitions of cortical recurrent networks with multiple internal states Cerebral Cortex 16 639–54
[17] Gao J, Liu Y-Y, D’Souza R M and Barabási A-L 2014 Target control of complex networks Nat. Commun. 5 5415–9
[18] Hale J K 2010 Asymptotic Behavior of Dissipative Systems vol 25 (Providence, RI: American Mathematical Society)
[19] Henry D 1981 Geometric Theory of Semilinear Parabolic Equations vol 840 (Berlin: Springer)
[20] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities Proc. Natl Acad. Sci. 79 2554–8
[21] Hopfield J J et al 1986 Computing with neural circuits—a model Science 233 625–33
[22] Huang S, Embregt I and Kaufman S 2009 Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective Semin. Cell Biol. 20 869–76
[23] Jeong H, Mason S P, Barabási A-L and Oltvai Z N 2001 Lethality and centrality in protein networks Nature 411 41–2
[24] Jeong H, Tombor B, Albert R, Oltvai Z N and Barabási A-L 2000 The large-scale organization of metabolic networks Nature 407 651–4
[25] Jia T and Barabási A-L 2013 Control capacity and a random sampling method in exploring controllability of complex networks Sci. Rep. 3 2354–8
[26] Kifer Y 1986 General random perturbations of hyperbolic and expanding transformations J. Anal. Math. 47 111–50
[27] Lai Y C 2014 Controlling complex, non-linear dynamical networks Natl Sci. Rev. 1 339–41
[28] Li X, Cassidy J J, Reinke C A, Fischboeck S and Carthew R W 2009 A microRNA imparts robustness against environmental fluctuation during development Cell 137 273–82
[29] Li Z and Hopfield J J 1989 Modeling the olfactory bulb and its neural oscillatory processings Biol. Cybern. 61 379–92
[30] Lin C T 1974 Structural controllability IEEE Trans. Autom. Control 19 201–8
[31] Liu Y Y, Slotine J-J and Barabási A-L 2011 Controllability of complex networks Nature 473 167–73
[32] Maass W, Schnitger G and Sontag E D 1991 On the computational power of sigmoid versus Boolean threshold circuits IEEE Proc. 32nd Annual Symp. on Foundations of Computer Science pp 767–76
[33] Mjolsness E, Sharp D H and Reinitz J 1991 A connectionist model of development J. Theor. Biol. 152 429–53
[34] Motter A E 2015 Networkcontrology Chaos: Interdiscip. J. Nonlinear Sci. 25 097621
[35] Nepusz T and Vicsek T 2012 Controlling edge dynamics in complex networks Phys. Rev. E 86 046106
[36] Newman M E J 2003 The structure and function of complex networks SIAM Rev. 45 167–256
[37] Allen Orr H 2005 The genetic theory of adaptation: a brief history Nat. Rev. Gen. 6 119–27
[38] Nicolaenko B, Constantin P, Foias C and Temam R 1989 Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations (Applies Mathematical Sciences Series vol 70) (New York: Springer)
[39] Pan Y and Li X 2014 Structural controllability and controlling centrality of temporal networks PloS One 9 e94998
[40] Pasqualetti F, Zampieri S and Bullo F 2014 Controllability metrics, limitations and algorithms for complex networks IEEE Trans. Control Netw. Syst. 1 40–52
[41] Pereira T 2010 Hub synchronization in scale-free networks Phys. Rev. E 82 036201
[42] Pereira T, Eroglu D, Baris Bagci G, Tirnakli U and Jensen H J 2013 Connectivity-driven coherence in complex networks Phys. Rev. Lett. 110 234103
[43] Perumal S and Minai A A 2009 Stable-yet-switchable (sys) attractor networks Neural Networks (Berlin: Springer) pp 2509–16
[44] Polačik P 1991 Complicated dynamics in scalar semilinear parabolic equations in higher space dimension J. Differ. Equ. 89 244–71
[45] Ruelle D 2014 Elements of Differentiable Dynamics and Bifurcation Theory (Amsterdam: Elsevier)
[46] Ruths D and Ruths J 2014 Control profiles of complex networks Science 343 1373–6
[47] Stam C J, Jelles B, Achtereekte H A M, Rombouts S A R B, Slaets J P J and Keunen R W M 1995 Investigation of EEG non-linearity in dementia and Parkinson’s disease Electroencephalogr. Clin. Neurophysiol. 95 309–17
[48] Stroud J, Barahona M and Pereira T 2015 Dynamics of cluster synchronisation in modular networks: implications for structural, functional networks Applications of Chaos, Nonlinear Dynamics in Science and Engineering, Vol 4 (New York: Springer) pp 107–30
[49] Sun J and Motter A E 2013 Controllability transition and nonlocality in network control Phys. Rev. Lett. 110 208701
[50] Talagrand M 1998 Rigorous results for the hopfield model with many patterns Probab. Theory Relat. Fields 110 177–275
[51] Tanizawa T, Paul G, Cohen R, Havlin S and Eugene Stanley H 2005 Optimization of network robustness to waves of targeted and random attacks Phys. Rev. E 71 047101
[52] Vakulenko S A 1994 A system of coupled oscillators can have arbitrary prescribed attractors J. Phys. A: Math. Gen. 27 2335
[53] Vakulenko S A 2000 Dissipative systems generating any structurally stable chaos Adv. Differ. Equ. 5 1139–78
[54] Vakulenko S and Radulescu O 2012 Flexible and robust patterning by centralized gene networks *Fundam. Inf.* **118** 345–69
[55] Vakulenko S A and Radulescu O 2012 Flexible and robust networks *J. Bioinform. Comput. Biol.* **10** 1241011
[56] Viana M 1998 Dynamics: a probabilistic and geometric perspective *Doc. Math. J. I* 557
[57] Vohradský J 2001 Neural network model of gene expression *FASEB J.* **15** 846–54
[58] Wu F-X, Wu L, Wang J, Liu J and Chen L 2014 Transittability of complex networks and its applications to regulatory biomolecular networks *Sci. Rep.* **4** 4819–23
[59] Yan G, Ren J, Lai Y-C, Lai C-H and Li B 2012 Controlling complex networks: how much energy is needed? *Phys. Rev. Lett.* **108** 218703
[60] Young L-S 1986 Stochastic stability of hyperbolic attractors *Ergod. Theor. Dynam. Syst.* **6** 311–9
[61] Yuan Z, Zhao C, Di Z, Wang W-X and Lai Y-C 2013 Exact controllability of complex networks *Nat. Commun.* **4** 2447–51